



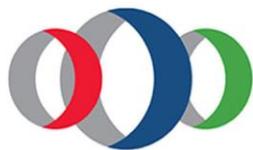
Vol. 1, September 2017
ISSN : 1907 - 2562

PROSIDING

SEMINAR NASIONAL MATEMATIKA 2017

Departemen Matematika FMIPA UI
Depok, 11 Februari 2017

Disponsori oleh:



IndonesiaRe

Didukung oleh:



TOKIOMARINE
INSURANCE GROUP



THE REAL LIFE
COMPANY



BPJS
Ketenagakerjaan



PERSATUAN AKTUARIS INDONESIA
(THE SOCIETY OF ACTUARIES OF INDONESIA)

SNM 2017



KATA SAMBUTAN

SEMINAR NASIONAL MATEMATIKA 2017

Dekan FMIPA Universitas Indonesia

Assalamu'alaikum Warahmatullahi Wabarakatuh

Salam sejahtera untuk kita semua.



Atas nama Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Indonesia, dengan bangga saya mengucapkan selamat kepada semua peserta pada Seminar Nasional Matematika 2017 yang diselenggarakan pada tanggal 11 Februari 2017 di Universitas Indonesia, Depok. Ucapan terima kasih saya sampaikan kepada pihak IndoMS Pusat dan IndoMS Wilayah JABAR, Banten, dan DKI Jakarta atas kepercayaannya kepada Universitas Indonesia dalam hal ini Departemen Matematika FMIPA sebagai tuan rumah kegiatan sarasehan dan sosialisasi program kerja IndoMS Pusat dan IndoMS Wilayah JABAR, Banten, dan DKI Jakarta.

Seminar Nasional ini merupakan seminar yang telah dilaksanakan secara bergantian oleh Universitas Indonesia dan Universitas Padjadjaran sejak 20 tahun yang lalu. Pihak Universitas Indonesia sebagai salah satu perguruan tinggi yang menjadi pelopor perkembangan peran ilmu pengetahuan di Indonesia tidak henti-hentinya mendorong segenap civitas akademika, termasuk di FMIPA UI untuk menghirilkan penelitiannya agar dapat memberikan dampak nyata pada kemajuan bangsa dan tanah air.

Saya ucapkan terima kasih kepada para pembicara utama, peserta dan tentunya kepada panitia pelaksana SNM 2017 ini. Semoga kegiatan ini dapat memberikan manfaat yang besar kepada kita semua dan bangsa Indonesia.

Salam hangat,

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Dekan FMIPA Universitas Indonesia

Dr. rer. nat. Abdul Haris

Gubernur IndoMS JABAR, Banten, dan DKI Jakarta



Assalamu'alaikum Warahmatullahi Wabarakatuh

Salam sejahtera untuk kita semua.

Atas nama *Indonesian Mathematical Society (IndoMS)*, sebuah kebanggaan yang besar bagi saya untuk menyampaikan selamat kepada semua peserta Seminar Nasional Matematika (SNM) 2017 yang diadakan pada tanggal 11 Februari 2017 di Departemen Matematika FMIPA UI, Depok.

IndoMS pada tahun ini bekerjasama dengan pihak penyelenggara lokal, mengadakan cukup banyak aktivitas temu ilmiah di berbagai daerah di Indonesia, termasuk salah satunya pada tahun ini yaitu SNM 2017 yang dirangkaikan dengan Sarasehan IndoMS Wilayah JABAR, Banten, dan DKI Jakarta serta sosialisasi program kerja IndoMS Pusat. Penyelenggaraan SNM 2017 tidak hanya merupakan program berkelanjutan dari pihak IndoMS, Universitas Indonesia dan Universitas Padjadjaran, namun juga merupakan sebuah kegiatan yang akan membawa peluang besar kepada seluruh pihak yang terlibat untuk menyeminarkan dan mendiskusikan hasil penelitian di berbagai bidang matematika.

Kami mengucapkan terima kasih kepada para pembicara utama, peserta dari berbagai daerah di Indonesia, dan panitia SNM 2017. Ucapan terima kasih khususnya kami sampaikan kepada Departemen Matematika, FMIPA Universitas Indonesia yang bersedia menjadi tuan rumah. Saya berharap agar SNM 2017 ini dapat memberikan manfaat yang besar kepada kita semua.

Salam hangat,

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Gubernur IndoMS JABAR, Banten dan DKI Jakarta.

Alhadi Bustamam, Ph.D.

Ketua Panitia Seminar Nasional Matematika 2017



Salam sejahtera bagi kita semua.

Matematika sebagai salah satu bidang ilmu yang penerapannya banyak digunakan di berbagai bidang, telah diterapkan pula pada berbagai kajian dan penelitian di masalah lingkungan. Pentingnya masalah pelestarian dan bagaimana mengatasi perubahan-perubahan fenomena lingkungan tersebut menjadi dasar dalam penentuan tema utama pada Seminar Nasional Matematika (SNM) 2017 ini, yakni “Peranan Matematika dalam Memahami Fenomena Lingkungan”.

Seminar Nasional Matematika merupakan perkembangan dari Seminar Matematika Bersama UI-UNPAD yang telah dilaksanakan sejak lebih dari 20 tahun yang lalu. SNM merupakan salah satu forum nasional bagi para matematikawan, peminat atau pemerhati Matematika dan para pengguna Matematika untuk saling berbagi pengetahuan dan pengalaman terhadap hasil penelitian dan penerapan matematika di berbagai hal. Melalui SNM 2017 diharapkan peserta yang berasal dari berbagai perguruan tinggi dan institusi di Indonesia dapat berpartisipasi dan berkontribusi sesuai dengan kepakaran bidang masing-masing di dalam mengatasi dan menyelesaikan masalah lingkungan beserta berbagai fenomenanya. Makalah yang masuk ke pihak penyelenggara meliputi berbagai bidang, seperti Analisis dan Geometri, Aljabar, Statistika dan aplikasinya, Matematika Keuangan dan Aktuaria, Kombinatorika, Komputasi, Pendidikan Matematika, Optimisasi, Pemodelan Matematika dan bidang terapan lainnya.

Penyelenggara SNM 2017 memberikan apresiasi yang setinggi-tingginya kepada berbagai pihak, antara lain Himpunan Matematika Indonesia wilayah Jabar, DKI Jakarta, dan Banten, Program Studi Matematika Universitas Padjadjaran, serta FMIPA UI yang telah memberikan dukungan dan bantuan dalam penyelenggaraan seminar nasional ini. Tidak lupa kami ucapkan terima kasih yang sebesar-besarnya kepada para sponsor yang telah berkontribusi dan kepada panitia SNM 2017 sehingga SNM 2017 dapat terselenggara.

Hormat kami,
Ketua Panitia SNM 2017

Bevina D. Handari Ph.D

UCAPAN TERIMA KASIH

Panitia Seminar Nasional Matematika 2017 menyampaikan ucapan terima kasih dan penghargaan kepada Pimpinan Universitas, Pimpinan Fakultas, Pimpinan Departemen, dan para sponsor, atas dukungannya dalam bentuk dana, fasilitas, dan lain-lain, untuk terselenggaranya seminar ini.

Secara khusus Panitia Seminar Nasional Matematika 2017 menyampaikan ucapan terima kasih kepada:

1. Rektor Universitas Indonesia
2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam
3. Ketua Departemen Matematika FMIPA Universitas Indonesia
4. Ketua Jurusan Matematika FMIPA Universitas Padjadjaran
5. Direktur Utama PT Reasuransi Indonesia Utama
6. Rektor Universitas Gunadarma
7. Direktur Utama PT Tokio Marine Life Insurance Indonesia
8. Direktur Utama PT AIA Financial Indonesia
9. Direktur Utama PT BNI Life Insurance
10. Direktur Utama BPJS Ketenagakerjaan
11. Ketua Persatuan Aktuaris Indonesia (PAI)
12. Direktur Utama PT Asuransi Cigna

Panitia Seminar Nasional Matematika 2017 juga mengucapkan terima kasih kepada pembicara utama Prof. Dr. Jatna Supriatna, M.Sc (Ketua RCCC Universitas Indonesia), Dr. Sri Purwani (Dosen Departemen Matematika FMIPA Universitas Padjadjaran), Dr. Ardhasena Sopaheluwakan (Kepala Bidang Litbang Klimatologi dan Kualitas Udara BMKG), para pemakalah pada sesi paralel, setiap tamu undangan, dan seluruh peserta Seminar Nasional Matematika 2017.

DAFTAR PANITIA SNM 2017

PELINDUNG

1. Prof. Dr. Ir. Muhammad Anis, M.Met. (Rektor Universitas Indonesia)
2. Dr. rer. nat. Abdul Haris (Dekan FMIPA Universitas Indonesia)

KOMISI PENGARAH

1. Alhadi Bustamam, Ph.D. (Gubernur IndoMS JABAR, DKI Jakarta, dan Banten, sekaligus sebagai Ketua Departemen Matematika, FMIPA Universitas Indonesia)
2. Prof. Dr. A.K. Supriatna (Ketua Jurusan Matematika, FMIPA Universitas Padjadjaran)

PANITIA PELAKSANA

1. Ketua : Bevina D. Handari, Ph.D.
2. Sekretaris : Dr. Dipo Aldila
3. Bendahara : Dra. Siti Aminah, M.Kom.
4. Pendanaan : Mila Novita, S.Si., M.Si.
Dr. Titin Siswantining, DEA.
5. Acara : Nora Hariadi, S.Si., M.Si.
Dra. Ida Fithriani, M.Si.
6. Makalah dan Prosiding : Dra. Siti Nurrohmah, M.Si.
Dr. rer. nat. Hendri Murfi
7. Perlengkapan : Maulana Malik, S.Si., M.Si.
Dr. Saskya Mary Soemartojo, M.Si.
Suci Fratama Sari, S.Si., M.Si.
Gianinna Ardanewari, S.Si., M.Si.

DAFTAR ISI

KATA SAMBUTAN	ii
Dekan FMIPA Universitas Indonesia	iii
Gubernur IndoMS JABAR, Banten, dan DKI Jakarta.....	iv
Ketua Panitia Seminar Nasional Matematika 2017	v
UCAPAN TERIMA KASIH	vi
DAFTAR PANITIA SNM 2017	vii
DAFTAR ISI.....	viii
PEMBICARA UTAMA.....	xiv
PERANAN MATEMATIKA DALAM MEMAHAMI FENOMENA LINGKUNGAN.....	xv
Prof. Dr. Jatna Supriatna, M.Sc	xv
UNDERSTANDING INDONESIAN ENVIRONMENTAL PHENOMENA, AND IMPROVING HUMAN LIVES	xvii
Dr. Sri Purwani	xvii
PERSPEKTIF SINGKAT IKLIM DI INDONESIA: PEMODELAN DAN STATUS PERUBAHAN IKLIM.	xviii
Dr. Ardhasena Sopaheluwakan	xviii
SESI PARALEL.....	1
STATISTIKA DAN APLIKASINYA.....	1
PENERAPAN GSTAT-R UNTUK PERHITUNGAN PREDIKSI POLUTAN TIMBAL MENGGUNAKAN METODE ORDINARY KRIGING	2
ANNISA NUR FALAH ¹ , KANKAN PARMIKANTI ² , BUDI NURANI RUCHJANA ³ , dan ATJE SETIAWAN ABDULLAH ⁴	2
PERBANDINGAN MODEL STAR DENGAN VAR DALAM MEMPREDIKSI DEBIT AIR SUNGAI CITARUM	13
KANKAN PARMIKANTI ¹ DAN E. RUSYAMAN ²	13

PEMODELAN INDEKS HARGA KONSUMEN BEBERAPA KOTA DI PROVINSI SUMATERA UTARA MENGGUNAKAN MODEL GSTARI – ARCH	20
HOT BONAR ¹ , BUDI NURANI RUCHJANA ² , DAN GUMGUM DARMAWAN ³	20
ANALISIS HARGA DAGING AYAM DI BANDUNG DENGAN MENGGUNAKAN VECTOR ERROR CORRECTION MODEL	27
RIBUT NURUL TRI WAHYUNI.....	27
MODEL SPATIO TEMPORAL DALAM PEMETAAN PENYAKIT DIARE DI KOTA BANDUNG	36
I GEDE NYOMAN MINDRA JAYA ¹ , BUDI NURANI RUCHJANA ²	36
PENGARUH INFRASTRUKTUR SOSIAL DAN EKONOMI TERHADAP PEREKONOMIAN DI INDONESIA: APLIKASI SARMA FIXED EFFECT MODEL.....	44
PRIANGGA A.W. ¹ , RIBUT NURUL T. W. ²	44
DETERMINAN KEMISKINAN DARI ASPEK HETEROGENITAS WILAYAH DENGAN GEOGRAPHICAL WEIGHTED REGRESSION (GWR).....	60
SUGIARTO	60
METODE RIDGE REGRESSION UNTUK MENGATASI MULTIKOLINERITAS PADA VARIABEL-VARIABEL YANG MEMENGARUHI PENDAPATAN ASLI DAERAH KABUPATEN KAPUAS HULU.....	70
KHOTIMAH ¹ , WINIH BUDIARTI ²	70
PENGARUH ENERGI TERMODINAMIKA DAN MOMEN DIPOL TERHADAP KESTABILAN KOMPLEKS GADOLINIUM DENGAN LIGAN TURUNAN DBDTP MENGGUNAKAN REGRESI BERGANDA	84
NURDENI.....	84
MODEL REGRESI ZERO-INFLATED JOINT BIVARIATE POISSON DAN APLIKASINYA UNTUK ANALISIS DATA KEMISKINAN DI BOGOR.....	90
FEVI NOVKANIZA ¹ , KHAIRIL ANWAR NOTODIPUTRO ² , KUSMAN SADIK ³	90
DAMPAK PENGETAHUAN, PENYULUHAN, DAN KEPEDULIAN LINGKUNGAN TERHADAP PERILAKU PENGELOLAAN SAMPAH DENGAN MENGGUNAKAN REGRESI LOGISTIK BINER.....	101
SUGIARTO	101
POLA PEROKOK USIA 10-24 DI INDONESIA TAHUN 2015.....	113

DIN NURIKA AGUSTINA.....	113
DETERMINAN PENGETAHUAN PRIA KAWIN TERHADAP TANDA BAHAYA PADA MASA KEHAMILAN DI INDONESIA TAHUN 2012.	122
ANITA DESMARINI ¹ , ITA WULANDARI ²	122
ANALISIS KECENDERUNGAN USIA WANITA MELAKUKAN KOHABITASI BERDASARKAN TINGKAT PENDIDIKAN MENGGUNAKAN REGRESI LOGISTIK BINER	133
ROBERT KURNIAWAN ¹ , DESNACITA HARLY PUTRI ²	133
PENANGANAN MASALAH PELANGGARAN ASUMSI KENORMALAN SISAAN DENGAN REGRESI KUANTIL	139
FERRA YANUAR ¹ , HAZMIRA YOZZA ² , SITY KHAIRIYAH ³	139
METODE LOGISTIC REGRESSION UNTUK TEKNOLOGI CREDIT SCORING.....	149
ULFA FAUZIAH ¹ , ZUHERMAN RUSTAM ²	149
MODEL PELUANG PUNAHNYA KUMAN MYCOBACTERIUM TUBERCULOSA YANG RESISTEN TERHADAP RIFAMPISIN DAN TUBERKULOSIS-MULTI DRUG RESISTANT (TB -MDR)	158
JONNER NAINGGOLAN	158
PENGARUH FOREIGN DIRECT INVESTMENT (FDI) TERHADAP KETIMPANGAN PENDAPATAN DI INDONESIA: APLIKASI IV-2SLS MODEL	164
RIBUT NURUL TRI WAHYUNI ¹	164
POLUSI DAN PERTUMBUHAN EKONOMI: SEBUAH ANALISIS IMPULSE RESPONSE FUNCTION.....	170
ANUGERAH KARTA MONIKA ¹	170
DAMPAK EKSPOR KARET ALAM TERHADAP PEREKONOMIAN INDONESIA DAN PERAMALANNYA: ANALISIS TABEL I-O DAN ARIMA	179
RAMADHAN I. P. ¹ , RIBUT NURUL T. W. ²	179
PEMODELAN ERROR CORRECTION MECHANISM DALAM ANALISIS ALIRAN MODAL DARI DALAM DAN LUAR NEGERI TERHADAP PEREKONOMIAN INDONESIA.....	188
AGATHA HERDIANI BRIA ¹ , KRISMANTI TRI WAHYUNI ²	188
PERAMALAN PRODUKSI DAN EKSPOR MINYAK KELAPA SAWIT INDONESIA TRIWULAN I 2015 – TRIWULAN III 2016 (PERBANDINGAN METODE EXPONENTIAL SMOOTHING DENGAN ARIMA)	199

RIO KUNCORO JATIKUSUMA ¹ , DEWI PURWANTI ²	199
PENGARUH TINGKAT KEMANDIRIAN FISKAL, INVESTASI, DAN TENAGA KERJA TERHADAP PERTUMBUHAN EKONOMI KABUPATEN/KOTA DI JAWA TENGAH TAHUN 2010-2014	210
ANDHIKA RAHMADANI ¹ , DEWI PURWANTI ²	210
ANALISIS DISPARITAS PEMBANGUNAN DAERAH TERTINGGAL, TERDEPAN, DAN TERLUAR DI INDONESIA TAHUN 2008-2013.....	223
LINDA ANNISA ¹ , ITA WULANDARI ²	223
PENERAPAN REGRESI DATA PANEL UNTUK MELIHAT PENGARUH PEMBANGUNAN INFRASTRUKTUR TERHADAP PERTUMBUHAN EKONOMI DI PULAU SUMATERA.....	233
ULFA MARDIYAH ¹ , WINIH BUDIARTI ²	233
KONVERGENSI EKONOMI ANTARKABUPATEN/KOTA DI PROVINSI PAPUA BARAT.....	249
I MADE WAHYU DWI SEPTIKA ¹ , DEWI PURWANTI ²	249
GENERALIZED LINIER MODELS (GLM) KEJADIAN KECELAKAAN LALU LINTAS DI KOTA MERAUKE.....	262
HARIANI FITRIANTI ¹ , MARKUS PALOBO ²	262
APLIKASI STOCHASTIC PRODUCTION FRONTIER UNTUK MENGUKUR EFISIENSI TEKNIS USAHA TANI PADI DI PROVINSI KEPULAUAN BANGKA BELITUNG.....	272
FITRI KARTIASIH ¹ DAN ADI SETIAWAN ²	272
MODEL PERSAMAAN STRUKTURAL UNTUK MENGETAHUI FAKTOR-FAKTOR YANG MEMPENGARUHI PRESTASI BELAJAR MATEMATIKA SISWA SMAN KECAMATAN PASAR REBO BERDASARKAN TINGKAT KEDISIPLINAN.....	283
ISMAH.....	283
PENELUSURAN RUMUS UJI KUBINGER SEBAGAI PADANAN ANOVA DUA ARAH (STUDI KASUS: PENGARUH FAKTOR METODE PREPARASI DAN FAKTOR MEREK TERHADAP KADAR TIMBAL PADA JAMU CINA)	296
FITRI CATUR LESTARI.....	296
EFEKTIVITAS ORGANISASI TATA KELOLA MINYAK DAN GAS BUMI DENGAN PEMODELAN STATISTIK “STRUCTURAL EQUATION MODELLING (SEM)” (STUDI KASUS DI KABUPATEN KUTAI KARTANEGARA)	309
RADITA ARINDYA ¹	309

PENERAPAN METODE ORDERING DUALY IN TRIANGLES (ORDIT) PADA MASALAH TINDAK KEJAHATAN DI INDONESIA .	320
TARUGA RUNADI ¹ , YEKTI WIDYANINGSIH ²	320
APLIKASI ANALISIS KONJOIN DESAIN FULLPROFILE RESPON RATING PADA PREFERENSI MAHASISWA TERHADAP KUALITAS DOSEN SEKOLAH TINGGI ILMU STATISTIK	328
FITRI CATUR LESTARI.....	328
PENDEKATAN DISTRIBUSI GUMBEL DALAM BAGAN KENDALI RANGE – EXPONENTIALLY WEIGHTED MOVING AVERAGE (R- EWMA) PADA PROSES PRODUKSI COKLAT DI INDONESIA PADA TAHUN 2009-2013.....	335
ERNA TRI HERDIANI ¹ , SITI AISYAH ALDILLAH ² , SRI ASTUTI THAMRIN ³	335
PENGELOMPOKAN KABUPATEN/KOTA DI JAWA BARAT MENGUNAKAN HYBRID CLUSTERING BERDASARKAN INDIKATOR PEMBANGUNAN MANUSIA BERBASIS GENDER	345
MARIA ULPAH ¹ , ZULHANIF ² , TITI PURWANDARI ²	345
METODE BOOTSTRAP BAGI MODEL POISSON DENGAN PENCILAN DAN OVERDISPERSI	366
KUTHA ARDANA ¹ , KHAIRIL ANWAR NOTODIPUTRO ² , KUSMAN SADIK ³	366
KELUARGA PENAKSIR MEAN POPULASI YANG MEMANFAATKAN AUXILIARY ATTRIBUTE BESERTA PARAMETERNYA	376
DJULIO RAHMAWAN ¹ • IDA FITHRIANI ² • SITI NURROHMAH ³	376
PENGGUNAAN METODE JACKKNIFE EMPIRICAL LIKELIHOOD UNTUK MEMBENTUK INTERVAL KEPERCAYAAN DENGAN PARAMETER MEAN ABSOLUTE DEVIATION.....	386
REVITA WIDASARI ¹ , IDA FITHRIANI ² , DAN SASKYA MARY SOEMARTOJO ³	386
PROPORTIONAL REDUCTION IN UNCERTAINTY (PRU) DALAM SELEKSI VARIABEL CALON DEBITURKREDIT MIKRO KOMERSIAL.....	395
ANNA CHADIDJAH ¹ , ACHMAD ZANBAR SOLEH ²	395
DETERMINAN BAYI DENGAN BERAT BADAN LAHIR RENDAH DI NUSA TENGGARA TIMUR DENGAN METODE GWNBR	401
LIZA KURNIA SARI ¹	401
ANALISIS CLUSTER MENGGUNAKAN METODE FUZZY C-MEANS BERDASARKAN SIMPLE RANDOM SAMPLING.....	414

ESTI RAMADITIA MULATSIH ¹ , IDA FITHRIANI ² , SITI NURROHMAH ³	414
DISTRIBUSI KUMARSWAMY-GEOMETRIK	420
DEANA ROSAN ¹ , SITI NURROHMAH ² , IDA FITHRIANI ³	420

PEMBICARA UTAMA

SEMINAR NASIONAL MATEMATIKA 2017

PERANAN MATEMATIKA DALAM MEMAHAMI FENOMENA LINGKUNGAN

Prof. Dr. Jatna Supriatna, M.Sc

Ketua RCCC Universitas Indonesia

Abstrak: Pembangunan berkelanjutan (*SDG-Sustainable Development Goal*) yang dicanangkan PBB untuk menggantikan *Millenium Development Goal* (MDG) sudah dimulai sejak awal 2016 dan akan berakhir 2030. Dari 17 goal dari SDG, 10 goal adalah *traditional development*, satu goal adalah kerjasama antar pemangku kepentingan (SDG 17) dan 6 goal adalah *emerging issues* dalam permasalahan lingkungan yaitu Energi terbarukan (SDG 7), Pembangunan kota dan masyarakat (SDG 11), Konsumsi bertanggung jawab (12), Perubahan iklim (SDG 13), Laut dan kehidupan bawah air (SDG 14), dan Kehidupan Flora dan Fauna di darat (SDG 15). Ke enam permasalahan lingkungan dalam pembangunan berkelanjutan yang baru ini tidak ada dalam target pembangunan MDG, sehingga banyak sekali diperlukan riset untuk dapat membuat berbagai kebijakan yang berdasarkan *evidence based decision*, mengadaptasikan rencana sesuai dengan kesiapan dan ketersediaan, pembuatan berbagai *computer and mathematical model* pengembangan SDG sampai 2030, mengarusutamakan SDG ke dalam rencana pembangunan RPJM/RPJP pemerintah pusat dan daerah dan bagaimana membuat MRV (*Measuring, Reporting, Verification*) dari setiap goal yang baru. Peranan pakar matematika sangat besar dalam membantu pelaksanaan pembangunan berkelanjutan. Sebagai contoh adalah masalah perubahan iklim. Masalah perubahan iklim adalah masalah terbesar dunia saat ini. Hasil survey Asahi Glass Foundation (2013) tampak bahwa masalah dunia terbesar saat ini adalah perubahan iklim (20%) dibanding dengan masalah lingkungan lainnya yang berkisar antara 10% (polusi) , keanekaragaman hayati (6%) dan yang lainnya. Model-model matematika dan komputer diperlukan untuk mengetahui dampak perubahan iklim terhadap kenaikan permukaan laut, cuaca ekstrim, kesehatan, ekonomi, pertanian, flora dan fauna, ketersediaan pakan, air dan lainnya dalam bentuk *time series*. Untuk MRV, diperlukan pedoman Pelaksanaan Pengukuran, Pelaporan, dan Verifikasi Aksi Mitigasi dan adaptasi dari setiap program di setiap

sektor pemerintah, swasta dan juga termasuk masyarakat. Capaian Aksi Mitigasi dan adaptasi Perubahan Iklim yang akurat, transparan, dan dapat dipertanggungjawabkan hanya dapat dilakukan apabila dilakukan oleh berbagai pakar terintegrasi termasuk pakar matematika dan statistik. Pemerintah harus mengatur (i) tatacara Pengukuran Aksi Mitigasi adaptasi dan Perubahan Iklim, (ii) tatacara pelaporan aksi mitigasi dan adaptasi perubahan iklim (iii) tatacara verifikasi capaian aksi mitigasi dan adaptasi perubahan iklim (iv) tatacara penilaian. Semua pengaturan tersebut memerlukan perhitungan yang pasti dan mendalam karena dampak dari perubahan iklim dapat menghancurkan perekonomian, membahayakan keberadaan ekosistem manusia, dalam jangka panjang dapat mempengaruhi peradaban dunia.

UNDERSTANDING INDONESIAN ENVIRONMENTAL PHENOMENA, AND IMPROVING HUMAN LIVES

Dr. Sri Purwani

Departemen Matematika, FMIPA Universitas Padjadjaran

Abstract: The universe and the environment around us were created perfectly by Allah. However, we find a lot of damage and disaster everywhere (Ar-Rum 30:41). This case, afflicting the environment and people of Indonesia, of course was through a long process. Indonesia, the country with the largest ocean border in the world, has experienced prosperity, well-being and peace in society. Understanding what the cause and how the process of occurrence, can provide answers for future improvements.

Human beings as part of the environment face the same thing. Various disease emerges, afflicts human survival. Imaging Sciences as a branch of knowledge is widely used in medical images analysis, range from disease detection, such as Alzheimer's, asthma, cancer and so on, up to image-guided surgery. This field involves many disciplines, hence providing opportunities for mathematicians to conduct research collaboration with scientists from various disciplines.

Registration and Segmentation, two important processes in the analysis of medical images, aims to find correspondence between two or more images, and attempts to extract structures/tissues within images, respectively. Previously, both processes are done separately. However, information from one process can be used to assist the other, and vice versa. Therefore, we tried to combine both processes implemented on database of MR brain images.

One of Petrovic et al. paper shows that adding structural information in their registration stage improved the result significantly, compared to registration using intensity alone. However, they only used little structural information. We attempted to include more structural information/segmentation in our new methods, and implemented groupwise registration to sets of images, consisting of tissue fraction images, intensity image and images with other structural information. The results of the registration were evaluated by using ground-truth annotation. It was found that ensemble registration using structural information can give a consistent improvement over registration using intensity alone of 25%-35%.

PERSPEKTIF SINGKAT IKLIM DI INDONESIA: PEMODELAN DAN STATUS PERUBAHAN IKLIM.

Dr. Ardhasena Sopaheluwakan

Kepala Bidang Litbang Klimatologi dan Kualitas Udara

Pusat Penelitian dan Pengembangan Badan Meteorologi Klimatologi dan Geofisika
(BMKG)

Abstrak: Iklim memiliki peranan penting dalam mendukung perikehidupan di bumi ini. Memiliki pengetahuan mengenai evolusi iklim (lampau dan kini) akan memberikan pemahaman untuk penggunaannya pada sektor yang penting, semisal pertanian dan ketahanan pangan. Sedangkan memiliki kemampuan untuk prediksi iklim yang akan datang, akan memberikan keunggulan untuk perencanaan strategis pembangunan bangsa-bangsa agar perikehidupannya dapat berkelanjutan (*sustainable development*).

Untuk mendapatkan deskripsi yang lengkap atas dinamika iklim di atmosfer, melibatkan pemodelan dengan rentang skala ruang yang sangat besar, melibatkan ukuran dari micrometer (butiran awan) hingga ribuan kilometer (*planetary scale*), yang melingkupi rentang ukuran ruang hingga 10^{14} meter. Pada saat ini pemodelan yang tersedia baru memenuhi sebagian dari skala rentang yang besar tersebut, sehingga tantangan untuk melengkapinya masih terbuka lebar. Presentasi ini akan memberikan beberapa *highlight* mengenai pemodelan iklim, karakter iklim di Indonesia, dan perubahan iklim yang sedang terjadi.

SESI PARALEL
STATISTIKA DAN
APLIKASINYA

SEMINAR NASIONAL MATEMATIKA 2017

PENERAPAN GSTAT-R UNTUK PERHITUNGAN PREDIKSI POLUTAN TIMBAL MENGGUNAKAN METODE ORDINARY KRIGING

ANNISA NUR FALAH¹, KANKAN PARMIKANTI², BUDI
NURANI RUCHJANA³, dan ATJE SETIAWAN ABDULLAH⁴

1 Mahasiswa Program Studi S-2 Matematika Departemen Matematika FMIPA Universitas
Padjadjaran, annisanurfalah02@gmail.com

2 Departemen Matematika FMIPA Universitas Padjadjaran, parmikanti@yahoo.com

3 Departemen Matematika FMIPA Universitas Padjadjaran, bnurani@gmail.com

4 Departemen Ilmu Komputer FMIPA Universitas Padjadjaran, atjesetiawan@gmail.com

Abstrak. Data spasial merupakan data yang dikumpulkan berdasarkan lokasi dalam ruang tertentu, sebagai contoh dataran banjir yang tercemar. Pada dataran banjir sungai Meuse terdapat salah satu logam yang mencemari yaitu Timbal (Pb), sehingga perlu diprediksi di lokasi mana saja yang terdapat polutan timbal. *Kriging* adalah salah satu metode untuk memprediksi indeks polutan pada dataran banjir di lokasi yang tidak tersampel dengan input semivariogram teoritis terbaik. Salah satu Metode Kriging yang dapat digunakan adalah Metode *Ordinary Kriging* yang merupakan salah satu metode kriging yang paling sederhana ketika rata-rata populasi tidak diketahui yang biasanya diaplikasikan pada data spasial. Perhitungan prediksi polutan timbal tersebut dapat menggunakan GStat-R untuk mendapatkan hasil yang akurat. Pada perhitungan prediksi dengan metode *Ordinary Kriging* diperlukan *library gstat* dan *library sp* serta algoritma dalam GStat-R, untuk diterapkan pada data dataran banjir sungai Meuse agar diperoleh prediksi indeks polutan di lokasi yang tidak tersampel.

Algoritma metode *ordinary kriging* diterapkan pada *GStat-R* untuk memprediksi indeks polutan timbal di lokasi yang tidak tersampel. Perhitungan prediksi indeks polutan timbal menggunakan GStat-R mudah, cepat dan akurat karena rata-rata variansi *kriging* yang dihasilkan minimum. Pada GStat-R juga dapat menampilkan kontur yang menunjukkan lokasi polutan timbal berada serta tingkat kandungan polutan timbal dengan interval tertentu di setiap lokasi.

Kata kunci: *Ordinary Kriging*, GStat-R, Polutan Timbal.

1. Pendahuluan

Geostatistika merupakan perpaduan dari ilmu pertambangan, geologi, matematika, dan statistika [3]. Data yang digunakan dalam ilmu geostatistika merupakan data observasi berdasarkan pengamatan lokasi. Dalam geostatistika terdapat proses prediksi yang biasa digunakan untuk memprediksi cadangan mineral, prediksi ini dinamakan *Kriging* [5]. *Kriging* dinamakan berdasarkan nama

D.G. Krige, seorang insinyur pertambangan dari Afrika Selatan, yang pertama kali mengembangkan teknik *moving averages* untuk memperkirakan kadar emas untuk menghilangkan efek regresi. Matheron mengembangkan metode *Kriging* dan aplikasinya.

Kriging adalah metode prediksi di lokasi tidak tersampel yang menggunakan data spasial. Proses perhitungan *Kriging* dapat dibedakan beberapa macam, yaitu: *Ordinary Kriging*, *Simple Kriging*, dan *Universal Kriging*. Metode *Ordinary Kriging* merupakan salah satu metode *Kriging* yang memiliki asumsi rata-rata (mean) tidak diketahui. Misalnya, data yang digunakan dalam penelitian ini adalah data kandungan polutan timbal di Dataran Banjir Sungai Meuse dari software R yang diasumsikan rata-ratanya tidak diketahui.

Program GStat-R adalah program untuk model geostatistika, dapat digunakan untuk prediksi dan simulasi dalam satu, dua, atau tiga dimensi dari sampel perhitungan variogram. Pada awalnya program ini berdiri sendiri, namun pada tahun 2004 dibuat paket dalam perangkat lunak R sebagai salah satu paket berisi fungsi untuk pengolahan data spasial atau aplikasi geostatistika [2]. Program GStat-R mudah digunakan, mempercepat perhitungan dan memberikan hasil yang akurat termasuk rata-rata variansi *Kriging*. Dalam penelitian ini digunakan program GStat-R dengan metode *Ordinary Kriging* untuk perhitungan prediksi kandungan polutan timbal di lokasi tidak tersampel.

2. Metode Penelitian

Kriging merupakan salah satu teknik atau metode analisis data yang sering digunakan dalam Geostatistika. Secara umum, kriging digunakan untuk menginterpolasikan suatu nilai kandungan mineral berdasarkan nilai-nilai yang diketahui. *Kriging* adalah metode estimasi yang memberikan sifat *the best unbiased linear estimator* (BLUE) pada nilai titik atau rata-rata blok. Maksud dari “best” disini adalah variansi minimum [4].

Semivariogram eksperimental dan teoritis Semivariogram adalah fungsi yang menggambarkan tingkat korelasi spasial dari variabel acak spasial $Z(x)$. Misalnya di bidang pertambangan emas, semivariogram memberikan gambaran untuk dua sampel yang diambil dari daerah pertambangan akan berfluktuasi dalam persentase emas tergantung pada jarak antara sampel. Untuk memperoleh struktur korelasi spasial dapat digunakan asumsi *intrinsic stationarity* [1]. Ini berarti nilai rata-rata dan variansi dari selisih $[Z(x+h) - Z(x)]$ harus independen dari lokasi x , sebagai berikut:

$$E[Z(x+h) - Z(x)] = 0 \quad (1)$$

$$Var[Z(x+h) - Z(x)] = 2\gamma(h) \quad (2)$$

Fungsi $\gamma(h)$ disebut sebagai semivariogram. Berdasarkan persamaan (1) dan (2), semivariogram eksperimental dapat didefinisikan sebagai:

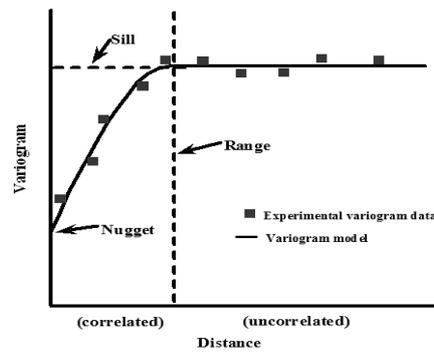
$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i + h) - Z(x_i)]^2 \quad (3)$$

dengan:

- $\hat{\gamma}(h)$: nilai semivariogram eksperimental dengan jarak h
- $Z(x_i)$: nilai observasi di lokasi x_i

$Z(x_i + h)$: nilai observasi di lokasi $x_i + h$
 $N(h)$: banyaknya pasangan titik yang berjarak h

Umumnya semivariogram eksperimental tidak memiliki properti isotropi. Isotropi berarti keseragaman dalam semua arah. Fungsi isotropik terbaik-dikenal sebagai model semivariogram teoritis spherical, gaussian dan eksponensial [6]. Model-model teoritis harus dicocokkan dengan semivariogram eksperimental melalui tiga parameter: sill (c), range (a) dan jarak (h) seperti digambarkan pada Gambar 1 di bawah ini:



Gambar 1. Grafik Semivariogram Eksperimental

Semua model semivariogram yang telah memenuhi semua sifat yang diperlukan dijelaskan sebagai berikut:

Model Spherical

$$\gamma(h) = \begin{cases} c \left[1,5 \left(\frac{h}{a} \right) - 0,5 \left(\frac{h}{a} \right)^3 \right], & h < a \\ c, & h \geq a \end{cases} \quad (4)$$

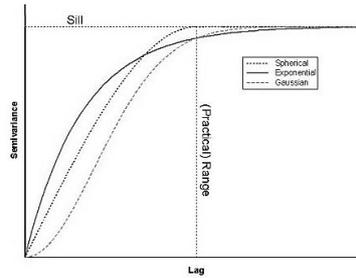
Model Eksponensial

$$\gamma(h) = \begin{cases} c \left[1 - \exp\left(\frac{-3h}{a} \right) \right], & h < a \\ c, & h \geq a \end{cases} \quad (5)$$

Model Gaussian

$$\gamma(h) = \begin{cases} c \left[1 - \exp\left(\frac{-3h^2}{a^2} \right) \right], & h < a \\ c, & h \geq a \end{cases} \quad (6)$$

Keterangan: c : sill, a : range, h : jarak



Gambar 2. Model Semivariogram Spherical, Eksponensial dan Gaussian

Metode *Ordinary Kriging*

Metode *Ordinary Kriging* merupakan salah satu metode kriging yang mempunyai asumsi rata-rata (*mean*) tidak diketahui. Rata-rata (*mean*) di setiap lokasi bernilai konstan. Jika variabel random $Z(x)$ diasumsikan stasioner dengan mean konstan m diketahui, kemudian $E[Z(x)] = m = E[Z(x_i)]$, dan estimator kriging bersifat linear, maka persamaannya sebagai berikut [7]:

$$\hat{Z}(x) = \sum_{i=1}^n (\lambda_i Z(x_i)) \quad (7)$$

dengan λ adalah bobot kriging. Mean dari kesalahan estimasi adalah:

$$E \left[\sum_{i=1}^n \left\{ \lambda_i \left[Z(x_i) - \hat{Z}(x) \right] \right\} \right] = m = \left(\sum_{i=1}^n \lambda_i - 1 \right)$$

Karena rata-rata m diasumsikan tidak diketahui maka didapat estimator *ordinary*

$$\text{kriging } \hat{Z}(x) = \sum_{i=1}^n \left\{ \lambda_i Z(x_i) \right\} \text{ dengan syarat } \sum_{i=1}^n \lambda_i = 1$$

Bobot kriging λ kemudian dihitung dengan menurunkan variansi minimum dan mensubstitusi fungsi semivariogram sebagai input, dinyatakan:

$$\lambda = CD^{-1}$$

dengan C dan D merupakan matriks dengan elemen sebagai berikut:

$$\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ p \end{bmatrix} = \begin{bmatrix} Cz_{11} & \dots & Cz_{1n} & 1 \\ \vdots & \vdots & \vdots & 1 \\ Cz_{n1} & \vdots & Cz_{nn} & \vdots \\ 1 & \dots & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} Cz_{10} \\ \vdots \\ Cz_{n0} \\ 1 \end{bmatrix}$$

3. Hasil dan Pembahasan

Data Penelitian

Data penelitian yang digunakan merupakan Data Set Meuse yang terdiri dari empat logam berat di dataran banjir di sepanjang sungai Meuse. Proses yang mengatur distribusi logam berat berupa sedimen tercemar yang dibawa oleh air sungai, dan sebagian besar disimpan dekat dengan tepi sungai, dan daerah dengan

elevasi rendah. Sampel dikumpulkan dalam dataran banjir sungai Meuse. Data set terdiri dari 164 titik sampel tanah atas konsentrasi logam berat (ppm).



Sumber: Program R 3.2.5

Gambar 3. Kandungan Timbal pada 164 lokasi di Dataran Banjir Sungai Meuse

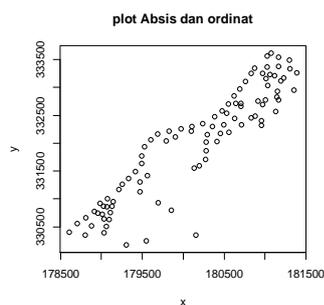
Berikut ini disajikan Dataset Meuse yang terdiri dari 164 lokasi, dengan koordinat x dan y dalam satuan meter dan kandungan polutan timbal dalam satuan ppm .

Tabel 1. Data Kandungan Timbal di 164 Lokasi

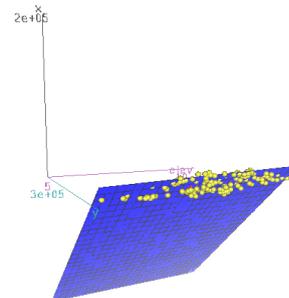
Lokasi	$x(m)$	$y(m)$	Timbal (ppm)
1	181072	333611	299
2	181025	333558	277
...
164	180627	330190	124

Sumber: Dataset dari Program 3.2.5

Berdasarkan Tabel 1 untuk melihat kestasioneran dari data dapat dilihat dengan plot 2D dan 3D seperti terlihat pada Gambar 4 dan Gambar 5 di bawah ini:



Gambar 4. Plot 2D



Gambar 5. Plot 3D

Grafik pada Gambar 4 dan Gambar 5 menunjukkan tidak banyak outlier, sehingga dengan kata lain 164 sampel dalam dataset memenuhi asumsi stasioneritas. Hal ini menunjukkan bahwa kandungan timbal stasioner dan dapat dianalisis lebih lanjut menggunakan ordinary point kriging.

Statistika deskriptif dari data kandungan timbal didapat dengan menggunakan perintah `summary(data)` pada program R maka dapat dilihat sebagai berikut:

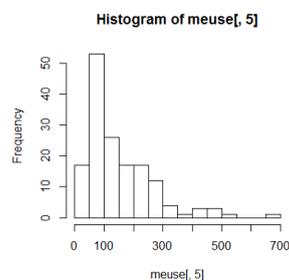
Tabel 2. Statistika Deskriptif Data Kandungan Timbal

	$x(m)$	$y(m)$	Timbal (ppm)
Min.	178605	330179	37.0
1st Qu.	179442	331065	72.5
Median	180283	332213	123.0
Mean	180156	332022	153.4
3rd Qu.	180935	332778	207.0
Max.	181390	333611	654.0

Tabel 2 memperlihatkan ringkasan data berdasarkan titik koordinat x yang memiliki nilai minimum 178605 m dan maksimum 181390 m , koordinat y memiliki nilai minimum 330179 m dan nilai maksimum 333611 m , sedangkan untuk kandungan timbal memiliki nilai minimum 37.0 ppm dan maksimum 654.0 ppm .

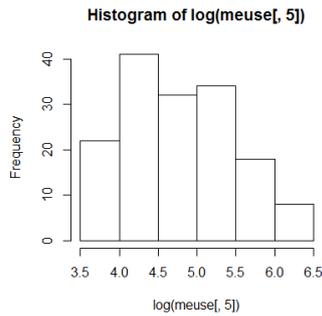
Normalitas

Normalitas dataset merupakan salah satu persyaratan yang digunakan dalam menggunakan metode *ordinary point kriging*. Untuk melihat histogram dari data kandungan timbal dapat menggunakan perintah `hist(meuse[, 5])` dan dapat dilihat pada Gambar 6 di bawah ini:



Gambar 6. Histogram Data Kandungan Timbal

Berdasarkan Gambar 6 dapat dilihat bahwa tidak berdistribusi normal. Oleh karena itu, salah satu cara untuk mendapatkan data tersebut berdistribusi normal dilakukan transformasi log, dan plot data yang ditransformasikan dapat dilihat pada Gambar 7 di bawah ini:



Gambar 7. Histogram Transformasi Log Data Kandungan Timbal

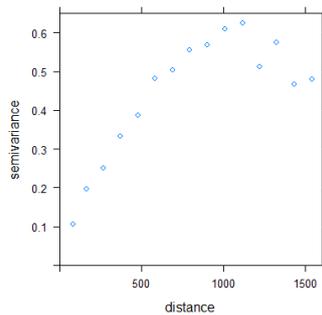
Semivariogram Eksperimental dan Teoritis

Untuk mendapatkan model semivariogram teoritis yang terbaik, pertama secara manual kita dapat menghitung nilai semivariogram eksperimental menggunakan persamaan (3), tetapi untuk menyederhanakan proses perhitungan kita menggunakan program R yang dapat mencari nilai semivariogram eksperimental dapat dilihat pada Tabel 3.

Tabel 3. Semivariogram Eksperimental

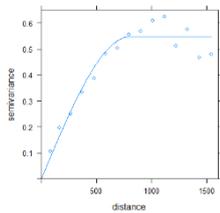
No	Banyaknya pasangan	Jarak (<i>m</i>)	Semivariogram Eksperimental
1	57	79.29244	0.1046520
2	299	163.97367	0.1965929
3	419	267.36483	0.2507668
4	457	372.73542	0.3330690
5	547	478.47670	0.3875716
6	533	585.34058	0.4817750
7	574	693.14526	0.5031432
8	564	796.18365	0.5545787
9	589	903.14650	0.5693882
10	543	1011.29177	0.6098806
11	500	1117.86235	0.6253271
12	477	1221.32810	0.5126165
13	452	1329.16407	0.5755737
14	457	1437.25620	0.4676728
15	415	1543.20248	0.4804887

Tabel 3 menunjukkan nilai rata-rata kandungan polutan timbal. Hal ini menunjukkan bahwa hasil dari pengolahan semivariogram eksperimental, diperoleh nilai rata-rata dari beberapa sampel di setiap kelas dan jarak masing-masing pasangan dari sampel dengan nilai semivariogram yang berbeda. Dapat dilihat dari plot nilai semivariogram eksperimental pada Gambar 8 sebagai berikut:

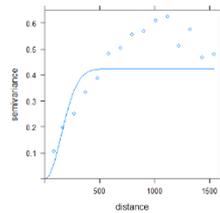


Gambar 8. Plot Semivariogram Eksperimental

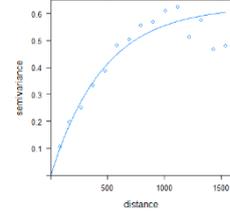
Untuk memperoleh model semivariogram teoritis terbaik, berdasarkan plot semivariogram eksperimental dipilih nilai sill = 1,5005 , range = 1000 dan nugget = 0. Berdasarkan tiga parameter tersebut, maka kita memilih yang terbaik model semivariogram teoritis dengan mencocokkan model semivariogram teoritis ke semivariogram eksperimental. Dalam pencocokkan model semivariogram teoritis, digunakan tiga model yaitu spherical, eksponensial dan gaussian. Hasil pencocokan ketiga model tersebut dapat dilihat pada Gambar 9, 10 dan 11. Selanjutnya model terbaik dapat dipilih dengan menghitung jumlah error minimum.



Gambar 9. Spherical



Gambar 10. Gaussian



Gambar 11. Eksponensial

Tabel 4. MSE Model Semivariogram Teoritis

MSE Model Semivariogram Teoritis		
Spherical	Gaussian	Eksponensial
3.175989e-05	0.0001764284	2.089753e-05

Berdasarkan hasil pencocokkan ketiga model semivariogram teoritis dengan model semivariogram eksperimental, diperoleh model semivariogram teoritis yang terbaik adalah model eksponensial karena memiliki mean square error yang paling minimum.

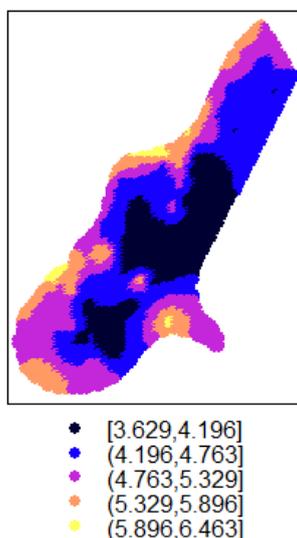
Hasil Prediksi Metode *Ordinary Kriging*

Setelah semivariogram teoritis telah dipilih, kita dapat menghitung prediksi kandungan timbal di lokasi tidak tersampel dengan menggunakan bantuan paket GStat-R yang dapat menghasilkan prediksi 3103 titik lokasi.

Tabel 5. Hasil Prediksi Kandungan Timbal

Lokasi	x	y	Prediksi	Variansi Error
1	181180	333740	5.378727	0.300037753
2	181140	333700	5.470063	0.223575378
3	181180	333700	5.388589	0.249639247
4	181220	333700	5.307014	0.275494696
5	181100	333660	5.583244	0.129930745
...
3099	179060	329620	5.308642	0.224043481
3100	179100	329620	5.282084	0.225899268
3101	179140	329620	5.259009	0.223019914
3102	179180	329620	5.240887	0.212364749
3103	179220	329620	5.225671	0.201056549

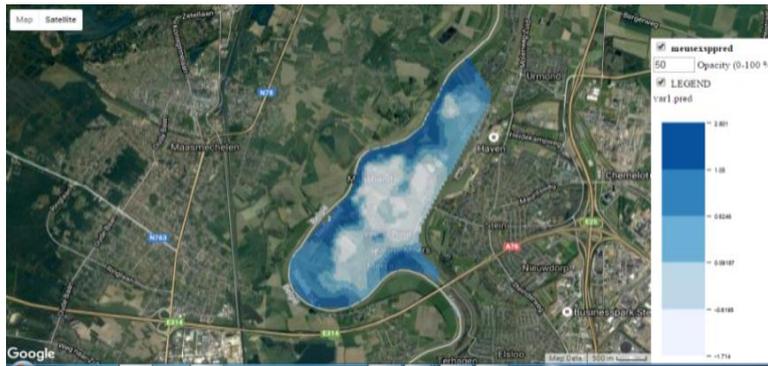
Dengan menggunakan fungsi dalam bahasa pemrograman R, kontur hasil prediksi kandungan timbal dapat digambarkan seperti terlihat pada Gambar 12 di bawah ini:



Gambar 12. Kontur Hasil Prediksi Kandungan Timbal

Berdasarkan Gambar 12 dapat disimpulkan bahwa warna hitam terdapat kandungan timbal paling sedikit sekitar 3.629 – 4.196 ppm. Warna biru terdapat kandungan timbal sekitar 4.196 – 4.763 ppm. Warna ungu terdapat kandungan timbal berada diantara rata-rata sekitar 4.763 – 5.329 ppm. Warna orange terdapat kandungan timbal sekitar 5.329 – 5.896 ppm. Warna kuning terdapat kandungan timbal paling banyak sekitar 5.896 – 6.463 ppm.

Selanjutnya, akurasi prediksi metode *ordinary kriging* dapat diverifikasi secara visual dengan memproyeksikan hasil untuk peta seperti yang ditunjukkan pada Gambar 13.



Gambar 13. Proyeksi Google Map

4. Kesimpulan

Algoritma metode *ordinary kriging*, penentuan model semivariogram teoritis terbaik sebagai input untuk metode *ordinary kriging* dan menampilkan kontur hasil prediksi melalui *script* dalam *software* R yaitu dimulai dari aktivasi *library* *gstat* dan *library* *sp*, pemanggilan data, melihat *summary* dari data yang telah dipanggil, menghitung semivariogram eksperimental, menentukan model semivariogram teoritis terbaik, menentukan prediksi di lokasi yang tidak tersampel menggunakan metode *ordinary kriging*, dan terakhir menampilkan kontur hasil prediksi.

Penerapan algoritma metode *ordinary kriging* dengan *software* R untuk memprediksi indeks polutan timbal yang tidak tersampel dan menampilkan kontur hasil prediksi di dataran banjir sungai Meuse yaitu dengan memanggil data set *meuse* dan data set *meuse.grid*, lalu dari proses selanjutnya diperoleh model eksponensial sebagai model semivariogram teoritis terbaik dan hasil prediksi yang didapat cukup akurat yaitu dengan memperoleh rata-rata variansi kriging yang cukup kecil. Pada kontur yang dihasilkan dalam R menunjukkan bahwa untuk yang berwarna gelap memiliki kadar polutan yang rendah, sedangkan untuk warna yang terang memiliki kadar polutan yang tinggi.

Pernyataan Terima Kasih. Paper ini didanai Hibah oleh Penelitian Kerjasama Luar Negeri (KLN) dan Publikasi Internasional 2017 Nomor Kontrak: 718/UN.6.3.1/PL/2017 serta *Academic Leadership Grant* Universitas Padjadjaran Nomor Kontrak: 872/UN.6.3.1/LT/2017.

Referensi

- [1] Armstrong, M., Basic Linear Geostatistics, Springer, 1998.
- [2] Bivand, R., Pebesma, E. and Rubio, V. Applied Spatial Data Analysis with R, 2 ed., New York: Springer, 2013.
- [3] Choudhury, N. H., Rahman, A. and Ferdousi, S. "Kriging Infill of Missing Data and Temporal Analysis of Rainfall in North," *Climatology & Weather Forecasting*, vol. 3, no. 3, pp. 1-5, 2015.

- [4] Cressie, N., *Statistics for Spatial Data*, 2 ed., New Jersey: Wiley-Interscience, 2015, p. 928.
- [5] Middelkoop, H., "Heavy-metal pollution of the river Rhine and Meuse floodplains in the Netherlands," *Netherlands Journal of Geosciences*, vol. 79, no. 4, pp. 411-428, 2000.
- [6] Montero, J.M., G. Fernández-Avilés and J. Mateu, *Spatial and Spatio-Temporal Geostatistical Modeling and Kriging*, John Wiley & Sons, 2015.
- [7] Pebesma, E., "Multivariable geostatistics in S: the gstat package," *Computers & Geosciences*, vol. 30, no. 7, pp. 683-691, 2004.

PERBANDINGAN MODEL STAR DENGAN VAR DALAM MEMPREDIKSI DEBIT AIR SUNGAI CITARUM

KANKAN PARMIKANTI¹ DAN E. RUSYAMAN²

¹Departemen Matematika FMIPA Unpad, rusyaman@unpad.ac.id
²Departemen Matematika FMIPA Unpad, parmikanti@unpad.ac.id

Abstrak. Untuk meminimalisasi dampak banjir bagi warga yang tinggal di dekat aliran sungai, perlu alat atau metode yang dapat memprediksi dan mendeteksi debit air sungai. Untuk kebutuhan tersebut, dalam makalah ini akan dibahas mengenai pembuatan dua buah model dan membandingkannya untuk mengetahui model mana yang lebih baik. Kedua model tersebut adalah *Space Time* Autoregressi (STAR) dan Vektor autoregresi (VAR), sedangkan untuk memilih yang terbaik dilihat dari nilai *error*-nya. Dengan mengambil data debit air sungai Citarum di tiga daerah yang berbeda, hasil penelitian ini menunjukkan bahwa model STAR dengan Metode Kuadrat Terkecil (MKT) ternyata lebih baik daripada model STAR.

Kata kunci : Banjir, STAR, VAR, Debit air, Citarum

1. Pendahuluan

Tingkat curah hujan yang tinggi akhir-akhir ini banyak menimbulkan banjir akibat meluapnya air sungai. Dengan demikian perlu adanya suatu tindakanantisipasi untuk meminimalkan akibat banjir tersebut. Salah satu antisipasi tersebut adalah dengan mendeteksi secara lebih dini debit air yang akan datang mengalir. Untuk keperluan tersebut perlu dibuat model yang dapat memperkirakan dan meramalkan debit air sungai. Secara teoritis terdapat banyak model yang dapat digunakan untuk mengolah data debit air sungai hasil observasi dari waktu ke waktu, yaitu yang berkaitan dengan deret waktu (*time series*). Data ini dapat diolah menjadi suatu model yang bertujuan untuk memprediksi kejadian di masa yang akan datang tanpa harus melakukan observasi lagi. Model yang paling sederhana adalah model yang disebut dengan model autoregresi orde1 atau disingkat dengan AR(1), yaitu model yang menunjukkan bahwa pengamatan periode sekarang akan tergantung pada pengamatan satu periode sebelumnya [3]. Karena kesederhanaannya, model ini lebih dikenal dan lebih sering digunakan hingga saat ini.

Apabila sumber yang diamati lebih dari satu lokasi, maka model AR ini dapat dikembangkan menjadi model multivariat Vektor Auto Regresi atau disingkat dengan VAR dan model *Space Time* Autoregresi atau STAR, di mana pengamatan periode sekarang di suatu lokasi, tidak saja dipengaruhi oleh

pengamatan satu periode sebelumnya di lokasi tersebut, tapi juga dipengaruhi oleh pengamatan di lokasi lain disekitarnya.[1].

. Dengan demikian penerapan model deret waktu merupakan masalah yang menarik untuk dikaji, baik secara teori maupun aplikasi. Dalam makalah ini akan dikaji bagaimana cara membuat model VAR(1) dengan menggunakan Metode Kuadrat Terkecil dan menggunakan Metode Yule-Walker. Selain itu akan dibahas pula bagaimana membuat model STAR(1,1), yaitu model STAR dengan lag waktu 1 (satu periode sebelumnya) dan lag spasial 1 (satu kelompok penelitian) di beberapa lokasi dengan menggunakan Metode Kuadrat Terkecil.

Selanjutnya, akan dibandingkan model mana yang lebih baik dengan melihat nilai *Mean Square Error*.

2. Hasil – Hasil Utama

.Ada dua jenis model peramalan yang utama, yaitu model regresi dan model deret waktu (*time series*). Jika dalam model regresi peramalan masa depan tidak harus didasarkan pada masa lalu, misalnya hubungan antara tinggi badan dan berat badan ideal, maka dalam model deret waktu peramalan masa depan dilakukan berdasarkan nilai masa lalu. Tujuan metode peramalan deret waktu adalah menemukan pola atau model dalam deret data historis dan mengekstrapolasikan pola tersebut ke masa depan. Tujuan dari analisis deret waktu secara umum meliputi dua hal, yaitu memahami atau memodelkan proses stokastik yang timbul pada rangkaian observasi dan memperkirakan nilai masa depan berdasarkan data deret waktu yang telah terjadi. [6]

Dalam bagian ini, penulis akan menguraikan beberapa kajian teori tentang model VAR dan STAR yang selanjutnya akan diaplikasikan untuk data debit air sungai Citarum Jawa Barat

2.1 Model Vektor Autoregresi Orde-p: VAR(p)

Dalam banyak kesempatan, data *time series* sering kali merupakan hasil observasi dari beberapa lokasi yang berbeda, misalnya N lokasi. Dalam hal ini model yang cocok digunakan bukan lagi AR, tetapi model Vektor Auto Regresi yang disingkat VAR dengan N buah variabel (multivariat). Model VAR berorde p ditulis $VAR(p)$ artinya ketergantungan terhadap p waktu sebelumnya.

Model vektor autoregresi berorde-1 yang dinotasikan dengan VAR(1) merupakan bentuk khusus dari model vektor autoregresi berorde- p , di mana model VAR(1) untuk bivariat adalah:

$$\mathbf{Z}_t = \Phi \mathbf{Z}_{t-1} + \vec{a}_t \quad \dots(1)$$

dengan

$$\mathbf{Z}_t = \begin{pmatrix} z_{1t} \\ z_{2t} \\ z_{3t} \end{pmatrix}, \quad \Phi = \begin{pmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{pmatrix}, \quad \text{dan} \quad \vec{a}_t \sim iid(\mathbf{0}, \sigma^2 \mathbf{I}_3)$$

Untuk mendapatkan nilai parameter, dengan menggunakan metode Yule – Walker diperoleh

$$\begin{pmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{22} \\ \phi_{22} & \phi_{22} & \phi_{22} \end{pmatrix}' = \begin{pmatrix} \gamma_{11}(0) & \gamma_{12}(0) & \gamma_{13}(0) \\ \gamma_{21}(0) & \gamma_{22}(0) & \gamma_{23}(0) \\ \gamma_{31}(0) & \gamma_{32}(0) & \gamma_{33}(0) \end{pmatrix}^{-1}$$

$$\begin{pmatrix} \gamma_{11}(1) & \gamma_{12}(1) & \gamma_{13}(1) \\ \gamma_{21}(1) & \gamma_{22}(1) & \gamma_{23}(1) \\ \gamma_{31}(1) & \gamma_{32}(1) & \gamma_{33}(1) \end{pmatrix}$$

... (2)

di mana

- $\gamma_{ij}(k)$ adalah fungsi cross-kovarian antara komponen Z_i dan Z_j , yaitu

$$= E(Z_{i,t} - \mu_i)(Z_{j,t+k} - \mu_j) = \Sigma(Z_{i,t} - \mu_i)(Z_{j,t+k} - \mu_j)$$

Selanjutnya untuk mendapatkan nilai parameter dengan menggunakan MKT, dari (1) diperoleh

$$\Phi'_{(3 \times 3)} = [\mathbf{Z}'(1)_{(3 \times T)} \quad \mathbf{Z}(1)_{(T \times 3)}]^{-1} \cdot [\mathbf{Z}'(1)_{(3 \times T)} \quad \mathbf{Z}_{(T \times 3)}] \quad \dots (3)$$

di mana

$$[\mathbf{Z}'(1)_{(3 \times T)} \quad \mathbf{Z}(1)_{(T \times 3)}]^{-1} =$$

$$\left[\begin{pmatrix} z_1(1) & z_1(2) & \dots & z_1(T) \\ z_2(1) & z_2(2) & \dots & z_2(T) \\ z_3(2) & z_3(2) & \dots & z_3(T) \end{pmatrix} \cdot \begin{pmatrix} z_1(1) & z_2(1) & z_3(1) \\ z_1(2) & z_2(2) & z_3(1) \\ \vdots & \vdots & \vdots \\ z_1(T) & z_2(T) & z_3(T) \end{pmatrix} \right]^{-1}$$

$$[\mathbf{Z}'(1)_{(3 \times T)} \quad \mathbf{Z}_{(T \times 3)}] =$$

$$\begin{pmatrix} z_1(1) & z_1(2) & \dots & z_1(T) \\ z_2(1) & z_2(2) & \dots & z_2(T) \\ z_3(2) & z_3(2) & \dots & z_3(T) \end{pmatrix} \cdot \begin{pmatrix} z_1(2) & z_2(2) & z_3(2) \\ z_1(3) & z_2(3) & z_3(3) \\ \vdots & \vdots & \vdots \\ z_1(T+1) & z_2(T+1) & z_3(T+1) \end{pmatrix}$$

Dengan demikian model VAR(1) yang diperoleh adalah:

$$\begin{cases} Z_1(t) = \hat{\phi}_{11} Z_1(t-1) + \hat{\phi}_{12} Z_2(t-1) + \hat{\phi}_{13} Z_3(t-1) \\ Z_2(t) = \hat{\phi}_{21} Z_1(t-1) + \hat{\phi}_{22} Z_2(t-1) + \hat{\phi}_{23} Z_3(t-1) \\ Z_3(t) = \hat{\phi}_{31} Z_1(t-1) + \hat{\phi}_{32} Z_2(t-1) + \hat{\phi}_{33} Z_3(t-1) \end{cases}$$

2.2 Model STAR (1;1)

Model STAR dapat digunakan untuk memodelkan masalah-masalah dalam berbagai bidang ilmu. Misalnya Kyriakidis dan Journel (1999) telah menggunakannya dalam bidang geologi, Epperson (2000) menggunakan model STAR pada masalah genetika berdasarkan waktu dan lokasi, Giacomini dan Granger (2004) memanfaatkan Model STAR secara luas di bidang ekonomi, serta Ruchjana dalam [4] menggunakan pengembangan model ini yaitu GSTAR untuk menyelesaikan masalah produksi minyak bumi. Demikian pula Hesti dalam [2] mengembangkan STAR jadi STARI untuk data yang tidak stasioner, serta Suhartono dan Dhoriva telah membandingkan model VAR dengan STAR dalam memodelkan produksi perkebunan teh di Jawa Barat. Untuk kesederhanaan model, kajian

difokuskan pada lag waktu 1 (satu periode sebelumnya) dan lag spasial 1 (satu kelompok penelitian) untuk model STAR(1,1) di beberapa lokasi.

Model STAR(1,1) dinyatakan:

$$(4) \quad \mathbf{z}(t) = \phi_{01} \mathbf{z}(t-1) + \phi_{11} \mathbf{W} \mathbf{z}(t-1) + \mathbf{e}(t) \quad \dots$$

dengan $\mathbf{z}(t)$: vektor pengamatan ($p \times 1$) dari p lokasi pada waktu t .

\mathbf{W} : matriks bobot ($p \times p$) pada lag spasial 1.

t : waktu pengamatan ($t=1, 2, 3, 4, \dots, n$).

ϕ_{01} : parameter model pada lag spasial 0 dan lag waktu 1.

ϕ_{11} : parameter model pada lag spasial 1 dan lag waktu 1.

dan vektor *error*: $\mathbf{e}(t) \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_N)$.

iid = *independent identically distributed*

Persamaan model STAR(1,1) untuk 3 lokasi dapat disajikan dalam bentuk matriks sebagai berikut :

$$\begin{pmatrix} z_1(t) \\ z_2(t) \\ z_3(t) \\ e_1(t) \\ e_2(t) \\ e_3(t) \end{pmatrix} = \phi_{01} \begin{pmatrix} z_1(t-1) \\ z_2(t-1) \\ z_3(t-1) \end{pmatrix} + \phi_{11} \begin{pmatrix} 0 & w_{12} & w_{13} \\ w_{21} & 0 & w_{23} \\ w_{31} & w_{32} & 0 \end{pmatrix} \begin{pmatrix} z_1(t-1) \\ z_2(t-1) \\ z_3(t-1) \end{pmatrix} + \dots(5)$$

Secara umum persamaan (5) dapat ditulis sebagai:

$$\begin{cases} z_1(t) = \phi_{01} z_1(t-1) + \phi_{11} w_{12} z_2(t-1) + \phi_{11} w_{13} z_3(t-1) + e_1(t) \\ z_2(t) = \phi_{01} z_2(t-1) + \phi_{11} w_{21} z_1(t-1) + \phi_{11} w_{23} z_3(t-1) + e_2(t) \\ z_3(t) = \phi_{01} z_3(t-1) + \phi_{11} w_{31} z_1(t-1) + \phi_{11} w_{32} z_2(t-1) + e_3(t). \end{cases} \dots(6)$$

atau

$$\begin{pmatrix} z_1(t) \\ z_2(t) \\ z_3(t) \end{pmatrix} = \begin{pmatrix} z_1(t-1) & w_{12} z_2(t-1) + w_{13} z_3(t-1) \\ z_2(t-1) & w_{21} z_1(t-1) + w_{23} z_3(t-1) \\ z_3(t-1) & w_{31} z_1(t-1) + w_{32} z_2(t-1) \end{pmatrix} \begin{pmatrix} \phi_{01} \\ \phi_{11} \end{pmatrix} + \begin{pmatrix} e_1(t) \\ e_2(t) \\ e_3(t) \end{pmatrix}$$

Dengan demikian, penaksiran parameter pada model STAR(1;1) dapat dilakukan dengan menggunakan Metode Kuadrat Terkecil (MKT), karena model STAR(1;1) di atas dapat dinyatakan sebagai model linear:

$$\mathbf{Y} = \mathbf{X} \boldsymbol{\beta} + \mathbf{e}(t), \quad \mathbf{e}(t) \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$$

di mana

$$\mathbf{Y} = \mathbf{z}(t) = \begin{pmatrix} z_1(t) \\ z_2(t) \\ z_3(t) \end{pmatrix},$$

$$\mathbf{X} = [z(t-1) \quad \mathbf{W} z(t-1)] = \begin{pmatrix} z_1(t-1) & w_{12} z_2(t-1) + w_{13} z_3(t-1) \\ z_2(t-1) & w_{21} z_1(t-1) + w_{23} z_3(t-1) \\ z_3(t-1) & w_{31} z_1(t-1) + w_{32} z_2(t-1) \end{pmatrix}$$

,

$$\beta = \begin{pmatrix} \phi_{01} \\ \phi_{11} \end{pmatrix}.$$

Dengan metode kuadrat terkecil, yaitu dengan meminimumkan kuadrat *error*, maka diperoleh rumus penaksir parameter

$$\beta = \begin{pmatrix} \phi_{01} \\ \phi_{11} \end{pmatrix} = (X^T \cdot X)^{-1} X^T Y \quad \dots$$

(7)

3. Hasil dan Pembahasan

Data yang diperoleh merupakan data debit sungai Citarum hulu dari 3 lokasi daerah aliran sungai Citarum sebagai berikut:

Z1: Daerah aliran sungai Cidurian-Sukapada

Z2: Daerah aliran sungai Cigulung-Maribaya

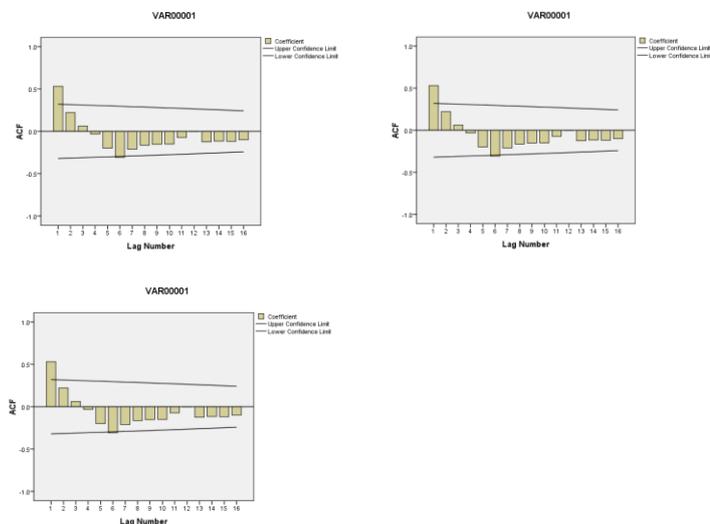
Z3: Daerah aliran sungai Citarum-Majalaya

Data diperoleh dari PUSAIR Bandung yang merupakan data debit sungai Citarum hulu tahun 2009 hingga tahun 2012. Data yang diteliti berupa data rekap harian debit sungai Citarum.

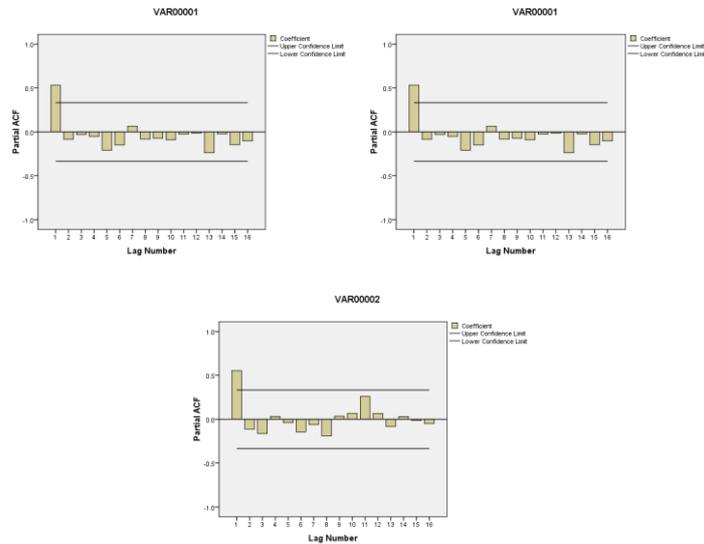
Tabel 1. Sebaran data debit air sungai hasil 32 pengamatan

Lokasi	Minimum	Rata-rata	STD DEV	Makasimum
Z1	17,1	24,96	5,65	35,95
Z2	24,58	46,76	18,74	61,93
Z3	57,95	195,11	209,06	438

Berikut adalah plot ACF dan PACF



Gambar.1 ACF Z_1, Z_2, Z_3



Gambar.1 PACF Z_1, Z_2, Z_3

Dari ACF dan PACF terlihat bahwa data stasioner terpotong di lag-1, artinya berorde $p = 1$.

Model VAR(1) yang ditunjukkan dalam (6) menggunakan metode Yule-Walker adalah:

$$\begin{cases} z_1(t) = 0,614 z_1(t-1) - 0,0695 z_2(t-1) + 0,003 z_3(t-1) \\ z_2(t) = -0,0797 z_1(t-1) + 0,6272 z_2(t-1) + 0,0084 z_3(t-1) \\ z_3(t) = -3,558 z_1(t-1) + 1,5809 z_2(t-1) + 0,5578 z_3(t-1) \end{cases}$$

Model VAR(1) pada (6) menggunakan metode MKT diperoleh adalah:

$$\begin{cases} z_1(t) = 0,95397 z_1(t-1) + 0,03230 z_2(t-1) - 0,00441 z_3(t-1) \\ z_2(t) = 0,39548 z_1(t-1) + 0,77102 z_2(t-1) - 0,00156 z_3(t-1) \\ z_3(t) = -1,00554 z_1(t-1) + 2,36798 z_2(t-1) + 0,50545 z_3(t-1) \end{cases}$$

Model STAR(1,1) pada (6) menggunakan metode MKT dengan dengan mengambil matriks bobot seragam

$$W = \begin{bmatrix} 0 & w_{12} & w_{13} \\ w_{21} & 0 & w_{23} \\ w_{31} & w_{32} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0,5 & 0,5 \\ 0,5 & 0 & 0,5 \\ 0,5 & 0,5 & 0 \end{bmatrix}$$

diperoleh nilai parameter: $\beta = \begin{bmatrix} \phi_{01} \\ \phi_{11} \end{bmatrix} = \begin{bmatrix} 0,877 \\ 0,039 \end{bmatrix}$.

Sehingga modelnya adalah

$$\begin{cases} z_1(t) = 0,877 z_1(t-1) + 0,0195 z_2(t-1) + 0,0195 z_3(t-1) \\ z_2(t) = 0,877 z_2(t-1) + 0,0195 z_1(t-1) + 0,0195 z_3(t-1) \\ z_3(t) = 0,877 z_3(t-1) + 0,0195 z_1(t-1) + 0,0195 z_2(t-1) \end{cases}$$

Untuk menentukan model mana yang lebih baik, dihitung *Mean Square Error* (MSE) dari setiap model. Nilai MSE untuk VAR dengan metode Yule-Walker dan STAR diperoleh hasil seperti terlihat pada tabel berikut.

Tabel-1: Nilai MSE

Lokasi	MSE VAR Y-W	MSE VAR MKT	MSE STAR
Cidurian-Sukapada	162,14	16,02	19,18
Cigulung-Maribaya	368,44	79,35	91,73
Citarum-Majalaya	13873,16	5371,98	6557,78

4. Simpulan dan Saran

Dengan menggunakan MSE sebagai patokan perbandingan *error*, diperoleh hasil bahwa model VAR dengan MKT lebih baik daripada model STAR maupun model VAR dengan metode Yule-Walker. Saran yang dapat penulis sampaikan adalah mengembangkan model STAR atau VAR dengan memberikan faktor eksogen yang mungkin berpengaruh pada model supaya menjadi lebih baik.

Daftar Pustaka

- [1] Borovkova, S.A, Lopuhaa, H.P, dan Ruchjana, B.N(2002), Generalized, STAR Model with Experimental Weights, In M. Staionopolousand G.Toulomi (Eds), Proceeding of the 17th International Workshop on Statistical Modelling, Chania, hal 139-147.
- [2] Hesti Anita Retnaningsih, 2014, *Pemodelan Generalized Space Time Autoregrssive Integrated dengan Differncing Musiman* pada Data Nonstasioner, Jurusan Matematika Universitas Brawijaya.
- [3] Ruchjana, B. N. 2002. The Stationary of The Space Time Autoregressive Model. *Majalah Ilmiah Himpunan Matematika Indonesia (MIHMI)*, Vol. 8 No. 2, ISSN: 0854-1380, hal. 151-159.
- [4] Ruchjana, B. N, 2002. Suatu Model Generalisasi Space Time Autoregresi (GSTAR) Orde1 dan Aplikasinya pada Data Produksi Minyak Bumi. Disertasi Program S3 Matematika ITB. Indah Dipublikasikan. Bandung; ITB.
- [5] Suhartono dan Dhoriva Urwatul Wutsqa, 2007, Perbandingan Model VAR dan STAR pada Peramalan Produksi Teh di Jawa Barat
- [6] Wei, W.W.S., (1989), *Time Series Analysis and Multivariate Methods*, Addison Wesley Publishing, Company, Inc. New York.

PEMODELAN INDEKS HARGA KONSUMEN BEBERAPA KOTA DI PROVINSI SUMATERA UTARA MENGUNAKAN MODEL GSTARI – ARCH

HOT BONAR¹, BUDI NURANI RUCHJANA², DAN GUMGUM
DARMAWAN³

¹BPS Provinsi Sumatera Utara dan Mahasiswa Magister Statistika Terapan Universitas
Padjadjaran, Bandung, bonarsitumorang@bps.go.id

²Departemen Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas
Padjadjaran, Bandung, budi.nurani@unpad.ac.id

³Departemen Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas
Padjadjaran, Bandung, gumstat@gmail.com

Abstrak. Inflasi merupakan salah satu indikator makro ekonomi yang dihitung berdasarkan persentase perubahan Indeks Harga Konsumen (IHK). IHK dapat diamati sebagai pengamatan yang melibatkan unsur lokasi dan waktu secara simultan, sehingga dapat dikategorikan sebagai data *space time*. Pemodelan dan peramalan data *space time* seperti IHK dapat dilakukan menggunakan pendekatan model *space time* berupa pengembangan model *Generalized Space Time Autoregressive* (GSTAR). Data deret waktu finansial seperti IHK, tingkat suku bunga, dan kurs mata uang sering menunjukkan pola data yang tidak stasioner dan memiliki volatilitas yang tinggi, sehingga berimplikasi pada varians *error* yang tidak konstan. Penelitian ini mempelajari dan mengembangkan model GSTAR untuk data yang tidak stasioner dan memiliki varians *error* yang tidak konstan berupa model *Generalized Space Time Autoregressive Integrated with ARCH Error* (GSTARI-ARCH) pada data IHK empat kota di Provinsi Sumatera Utara. Identifikasi orde optimum berdasarkan plot *Space Time Partial Autocorrelation Function* (STPACF) dan nilai minimum AIC diperoleh model GSTARI(1,1,1)-ARCH(1). Penaksiran parameter model GSTARI(1,1,1)-ARCH(1) dilakukan dengan metode *Generalized Least Square* (GLS) menggunakan matriks bobot antar lokasi berupa bobot invers jarak. Hasil peramalan data IHK empat kota di Provinsi Sumatera Utara memberikan nilai *Mean Absolute Percentage Error* MAPE pada data *out sample* sebesar 0,65 persen yang menunjukkan bahwa model GSTARI-ARCH dapat direkomendasikan sebagai model alternatif dalam pengambilan kebijakan sektor ekonomi terkait inflasi di Provinsi Sumatera Utara.

Kata kunci: *Space Time, GSTARI – ARCH, GLS, IHK, MAPE.*

1. Pendahuluan

Perkembangan inflasi di Provinsi Sumatera Utara dihitung berdasarkan perubahan IHK pada empat kota, yaitu Kota Medan, Kota Pematangsiantar, Kota Sibolga, dan Kota Padangsidimpuan. Data IHK empat kota di Provinsi Sumatera Utara ini dihitung setiap bulan, sehingga dapat dikategorikan sebagai data deret waktu. Perkembangan IHK selain dipengaruhi oleh jumlah permintaan komoditas

barang dan jasa yang dikonsumsi oleh masyarakat di kota tersebut juga dipengaruhi oleh tingkat harga dan ketersediaan barang dan jasa di kota lainnya. Kedekatan geografis antar satu kota dengan kota lainnya juga turut mempengaruhi perilaku perubahan harga barang dan jasa yang dikonsumsi masyarakat, terutama terhadap barang dan jasa pembentuk IHK.

Fenomena ini menunjukkan bahwa IHK empat kota di Provinsi Sumatera Utara dapat diamati sebagai pengamatan yang melibatkan unsur lokasi dan waktu secara simultan, sehingga dapat dikategorikan sebagai data *space time*. Pemodelan data *space time* seperti IHK yang melibatkan unsur lokasi dan waktu secara simultan dapat dilakukan menggunakan pengembangan model *Generalized Space Time Autoregressive* (GSTAR) yang diperkenalkan oleh Ruchjana, Borovkova, dan Lopuhaa [1, 7, 8].

Data deret waktu finansial seperti IHK, tingkat suku bunga, dan kurs mata uang sering menunjukkan pola data yang tidak stasioner dan memiliki volatilitas yang tinggi, sehingga berimplikasi pada varians *error* yang tidak konstan. Nainggolan [4, 5] mengembangkan model *Generalized Space Time Autoregressive with ARCH Error* (GSTAR-ARCH) untuk memodelkan data *space time* yang memiliki varians *error* yang tidak konstan. Pemodelan keheterogenan varians ini pertama kali diperkenalkan oleh Engle [2]. Model GSTAR dan GSTAR-ARCH memiliki kelemahan jika digunakan untuk memodelkan data *space time* yang tidak stasioner, untuk itu dalam penelitian ini dikembangkan model *Generalized Space Time Autoregressive Integrated with ARCH Error* (GSTARI-ARCH) untuk data yang tidak stasioner dan memiliki varians *error* yang tidak konstan pada data IHK empat kota di Provinsi Sumatera Utara.

Matriks bobot yang digunakan dalam penelitian ini dibatasi hanya menggunakan bobot invers jarak dengan lokasi penelitian pada empat kota inflasi di Provinsi Sumatera Utara. Nilai bobot invers jarak diperoleh berdasarkan perhitungan jarak sebenarnya antar lokasi. Penaksiran parameter mengikuti prosedur yang telah disampaikan oleh Nainggolan [4, 5], meliputi: (1) penaksiran parameter model GSTARI dengan metode *Ordinary Least Square* (OLS); (2) penaksiran parameter varians *error* dengan metode *Maximum Likelihood* (ML); dan (3) penaksiran parameter GSTARI-ARCH dengan metode *Generalized Least Square* (GLS). Ketepatan model dalam penelitian ini dilihat berdasarkan nilai *Mean Absolute Percentage Error* (MAPE) pada data *out sample* yang diperoleh.

2. Tinjauan Pustaka

2.1 Analisis Data Time Series

Analisis data *time series* merupakan suatu metode yang mempelajari data deret waktu berdasarkan teori yang mendasarinya serta bagaimana melakukan peramalan atau prediksi. Wei [9] menjelaskan bahwa model *time series* berdasarkan jumlah variabel yang diamati dapat dibagi menjadi 2, yakni analisis data deret waktu univariat (mengacu kepada data deret waktu yang terdiri dari satu observasi atau pengamatan yang diukur selama kurun waktu tertentu) dan analisis deret waktu multivariat (melibatkan beberapa variabel runtun waktu untuk memodelkan dan menjelaskan interaksi serta pergerakan diantara sejumlah variabel deret waktu).

2.2 Model Space Time Autoregressive (STAR)

Diperkenalkan oleh Pfeifer-Deutch [6], model STAR merupakan model *time series* yang tidak hanya dipengaruhi oleh waktu, tetapi juga dipengaruhi oleh faktor lokasi. Asumsi dalam model STAR adalah bahwa karakteristik untuk semua lokasi yang diamati bersifat homogen. Persamaan model STAR ($p; \lambda_1, \lambda_2, \dots, \lambda_k$) dituliskan sebagai berikut:

$$\mathbf{Z}(t) = \sum_{k=1}^p \sum_{l=0}^{\lambda_k} [\phi_{kl} \mathbf{W}^{(l)} \mathbf{Z}(t-k)] + \mathbf{e}(t). \quad (1)$$

2.3 Model Generalized Space Time Autoregressive (GSTAR)

Model STAR mengasumsikan karakteristik untuk semua lokasi yang diamati bersifat homogen, sedangkan kenyataan di lapangan sering kali ditemukan keheterogenan yang tinggi antar lokasi pengamatan. Ruchjana [7] merevisi kelemahan model STAR dengan mengembangkan sebuah model yang dikenal dengan model GSTAR. Model ini menghasilkan model *space time* dengan parameter-parameter yang tidak harus sama untuk dependensi waktu maupun dependensi lokasi. Dalam notasi matriks, model GSTAR ($p; \lambda_1, \lambda_2, \dots, \lambda_k$) dapat ditulis sebagai berikut:

$$\mathbf{Z}(t) = \sum_{k=1}^p \sum_{l=0}^{\lambda_k} [\Phi_{kl} \mathbf{W}^{(l)} \mathbf{Z}(t-k)] + \mathbf{e}(t). \quad (2)$$

2.4 Model Generalized Space Time Autoregressive with ARCH Error (GSTAR-ARCH)

Pada data ekonomi seperti IHK, nilai tukar, indeks harga saham, dan variabel ekonomi lain sering terjadi pola data dengan fluktuasi yang tinggi, sehingga tidak lagi memenuhi asumsi bahwa varians *error* konstan setiap waktu (homoskedastis). Nainggolan [4, 5] memperkenalkan model GSTAR-ARCH untuk mengatasi fenomena varians *error* yang tidak konstan tersebut. Model GSTAR-ARCH orde (1) untuk lag spasial, waktu, dan ARCH *error* dapat dituliskan sebagai GSTAR(1,1)-ARCH(1) dengan bentuk umum sebagai berikut:

$$\mathbf{Z}(t) = \Phi_{10} \mathbf{Z}(t-1) + \Phi_{11} \mathbf{W}^{(1)} \mathbf{Z}(t-1) + \mathbf{e}(t), \quad (3)$$

dengan:

$$\begin{aligned} \mathbf{e}(t) &= \mathbf{D}_t \boldsymbol{\eta}_t, \\ (e_t | F_{t-1}) &\sim N(\mathbf{0}, \boldsymbol{\Sigma}_t). \end{aligned}$$

2.5 Model Generalized Space Time Autoregressive Integrated with ARCH Error (GSTARI-ARCH)

Mulyaningsih [3] mengemukakan bahwa model STAR dan GSTAR memiliki kelemahan apabila digunakan untuk data yang tidak stasioner. Data deret waktu yang tidak stasioner dengan varians *error* yang tidak konstan dan berubah setiap waktu dapat dimodelkan melalui pengembangan model GSTARI-ARCH. Bentuk umum model GSTARI(p,d,l)-ARCH(p) dapat dituliskan sebagai berikut:

$$\mathbf{Y}(t) = \sum_{k=1}^p \sum_{l=0}^{\lambda_k} [\Phi_{kl} \mathbf{W}^{(l)} \mathbf{Y}(t-k)] + \mathbf{e}(t), \quad (4)$$

dengan:

$$\begin{aligned} \mathbf{Y}(t) &= \mathbf{Z}(t) - \mathbf{Z}(t-1); \mathbf{Y}(t-k) = \mathbf{Z}(t-k) - \mathbf{Z}(t-k-1), \\ \mathbf{e}(t) &= \mathbf{D}_t \boldsymbol{\eta}_t, \\ (e_t | F_{t-1}) &\sim N(\mathbf{0}, \boldsymbol{\Sigma}_t). \end{aligned}$$

$\mathbf{Z}(t)$: vektor variabel berukuran ($n \times 1$) pada waktu t

- $\mathbf{Z}(t-1)$: vektor variabel berukuran $(n \times 1)$ pada waktu $(t-1)$
 $\mathbf{Z}(t-k)$: vektor variabel berukuran $(n \times 1)$ pada waktu $(t-k)$
 Φ_{kl} : $\text{diag}(\phi_{kl}^{(1)}, \dots, \phi_{kl}^{(n)})$, yaitu matriks diagonal parameter autoregressive pada lag waktu k dan lag spasial l berukuran $(n \times n)$
 $\mathbf{W}^{(l)}$: matriks bobot berukuran $(n \times n)$ pada lag spasial l (dimana $l = 0, 1, \dots$), dan pembobot tersebut dipilih untuk $w_{ii} = 0$ dan $\sum_{i \neq j} w_{ij} = 1$
 $\mathbf{e}(t)$: vektor residual berukuran $(n \times 1)$ pada waktu t .

3. Hasil dan Pembahasan

3.1 Identifikasi Model

Data yang digunakan dalam penelitian ini adalah data IHK empat kota di Provinsi Sumatera Utara, yakni Kota Medan, Kota Pematangsiantar, Kota Sibolga, dan Kota Padangsidempuan periode Januari 2008 sampai Desember 2016, yang bersumber dari BPS Provinsi Sumatera Utara. Pengolahan data dalam penelitian ini menggunakan bantuan *software R* [10], analisis data *time series* dilakukan dengan terlebih dahulu membagi data menjadi dua bagian yaitu data *in sample* (yakni data IHK masing-masing kota periode Januari 2008 – Desember 2015) dan data *out sample* (yakni data IHK masing-masing kota periode Januari – Desember 2016). Analisis dan pemodelan selanjutnya dilakukan dengan menggunakan data *in sample* sedangkan pengujian ketepatan prediksi akan dilakukan dengan menggunakan data *out sample*. Identifikasi model diawali dengan pemeriksaan stasioneritas dengan uji *Augmented Dickey Fuller*. Adapun hipotesis yang digunakan adalah sebagai berikut:

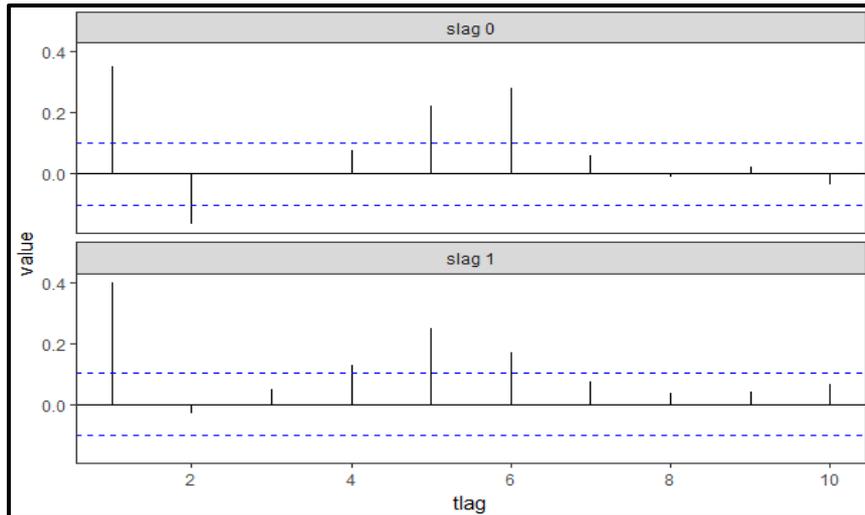
- $H_0: \gamma = 0$, terdapat *unit root* atau data tidak stasioner
 $H_1: \gamma < 0$, tidak terdapat *unit root* atau data stasioner

Tolak H_0 jika p-value lebih kecil dari nilai kekeliruan (α), yang artinya tidak terdapat *unit root* atau data sudah stasioner. Hasil uji berdasarkan Tabel 1. menunjukkan bahwa data IHK empat kota di Provinsi Sumatera Utara stasioner setelah melalui proses *differencing* satu kali.

Tabel 1. Hasil Uji Stasioneritas *Augmented Dickey Fuller* (ADF)

Lokasi	Kondisi Data	ADF Test	
		p-value	Kesimpulan
Medan	Asli (level)	0,8774	Tidak stasioner
	<i>First difference</i>	< 0,0001	Stasioner
Pematangsiantar	Asli (level)	0,7765	Tidak stasioner
	<i>First difference</i>	< 0,0001	Stasioner
Sibolga	Asli (level)	0,5728	Tidak stasioner
	<i>First difference</i>	< 0,0001	Stasioner
Padangsidempuan	Asli (level)	0,4614	Tidak stasioner
	<i>First difference</i>	< 0,0001	Stasioner

Setelah pengujian stasioneritas dilakukan, tahapan analisis selanjutnya adalah identifikasi orde lag waktu dan lag spasial model yang akan digunakan. Plot *Space Time Partial Autocorrelation Function* (STPACF) data IHK empat kota di Provinsi Sumatera Utara yang sudah stasioner dengan bobot invers jarak ditunjukkan melalui Gambar 1. berikut:



Gambar 1. Grafik Plot STPACF dengan Bobot Invers Jarak

Orde lag spasial model GSTARI-ARCH yang dipilih dalam penelitian ini adalah satu ($l=1$), hal ini dikarenakan empat lokasi dalam penelitian ini berada dalam satu provinsi. Dengan melihat plot STPACF pada lag spasial 1, maka kemungkinan orde lag waktu untuk model GSTARI-ARCH adalah orde lag waktu 1, 5, dan 6 (ditunjukkan oleh garis terpotong pada gambar di atas).

Penentuan orde lag waktu yang optimal dari berbagai kemungkinan yang diberikan oleh plot STPACF dapat dilakukan berdasarkan *Akaike's Information Criterion* (AIC) yang paling minimum. Kriteria ini umumnya dapat digunakan dalam pemilihan orde waktu model *time series* multivariat, dimana kriteria minimum AIC dapat memberikan nilai AIC untuk masing-masing orde sehingga dapat membandingkan dan menentukan lag waktu yang optimal terhadap model GSTARI-ARCH yang akan dibangun. Berdasarkan hasil olahan data IHK in sample yang digunakan, diperoleh nilai AIC terkecil (sebesar -3,9323) pada orde AR(1) dan MA(0), sehingga dalam penelitian ini selanjutnya pemodelan data IHK empat kota di Provinsi Sumatera Utara menggunakan model GSTARI(1,1,1)-ARCH(1).

3.2 Penaksiran Parameter dan *Diagnostic Test*

Penaksiran parameter model GSTARI-ARCH dalam penelitian ini mengikuti prosedur penaksiran parameter yang disampaikan oleh Nainggolan [5], secara umum dapat dibagi dalam 3 (tiga) tahapan, diawali dengan penaksiran parameter model GSTARI dengan metode *Ordinary Least Square* (OLS), dilanjutkan dengan menaksir varians bersyarat *error* model GSTARI dengan metode *Maximum Likelihood* (ML), dan terakhir penaksiran parameter GSTARI-ARCH dengan metode *Generalized Least Square* (GLS). Sebelum melakukan penaksiran parameter, terlebih dahulu dihitung matriks bobot yang akan digunakan, dimana dalam penelitian ini digunakan matriks bobot invers jarak sebagai berikut:

$$W_{ij} = [w_{ij}] = \begin{bmatrix} 0 & 0,6170 & 0,2067 & 0,1763 \\ 0,5252 & 0 & 0,2647 & 0,2101 \\ 0,1461 & 0,2197 & 0 & 0,6342 \\ 0,1335 & 0,1869 & 0,6796 & 0 \end{bmatrix}$$

Berdasarkan hasil taksiran parameter yang diperoleh, model GSATRI(1,1,1)-ARCH(1) untuk masing-masing lokasi dapat dituliskan sebagai berikut:

$$\hat{Z}_1(t) = 1,3169 Z_1(t-1) - 0,3169 Z_1(t-2) + 0,1705 Z_2(t-1) - 0,1705 Z_2(t-2) + 0,0571 Z_3(t-1) - 0,0571 Z_3(t-2) + 0,0487 Z_4(t-1) - 0,0487 Z_4(t-2)$$

$$\hat{Z}_2(t) = 0,2823 Z_1(t-1) - 0,2823 Z_1(t-2) + 0,9978 Z_2(t-1) + 0,0022 Z_2(t-2) + 0,1423 Z_3(t-1) - 0,1423 Z_3(t-2) + 0,1129 Z_4(t-1) - 0,1129 Z_4(t-2)$$

$$\hat{Z}_3(t) = 0,0812 Z_1(t-1) - 0,0812 Z_1(t-2) + 0,1220 Z_2(t-1) - 0,1220 Z_2(t-2) + 0,9002 Z_3(t-1) + 0,0998 Z_3(t-2) + 0,3523 Z_4(t-1) - 0,3523 Z_4(t-2)$$

$$\hat{Z}_4(t) = 0,0507 Z_1(t-1) - 0,0507 Z_1(t-2) + 0,0710 Z_2(t-1) - 0,0710 Z_2(t-2) + 0,2582 Z_3(t-1) - 0,2582 Z_3(t-2) + 0,9956 Z_4(t-1) + 0,0044 Z_4(t-2)$$

Uji *diagnostic* dilakukan untuk memeriksa apakah asumsi model yang terbentuk sudah terpenuhi. Uji *diagnostic* dilakukan terhadap *error* model yang terbentuk, masing-masing mencakup uji *white noise*, normalitas, dan homoskedastisitas. Hasil pengujian *multivariate white noise* dengan *Portmanteu test* menunjukkan nilai *p-value* lebih besar dari nilai $\alpha = 0,05$ untuk semua lag waktu, sehingga dapat dikatakan bahwa *error* sudah memenuhi asumsi *multivariate white noise*. Pengujian *multivariate normal* dengan *Royston's test* menunjukkan nilai *p-value* lebih besar dari nilai $\alpha = 0,05$ yang berarti *error* sudah memenuhi asumsi multivariat normal. Demikian juga dengan hasil pengujian asumsi homoskedastisitas *error* dengan *Multivariate ARCH-LM test*. Hasil uji menunjukkan nilai *p-value* lebih besar dari nilai $\alpha = 0,05$ yang berarti bahwa *error* model sudah tidak mengandung efek ARCH atau *error* sudah homoskedastis.

3.3 Peramalan

Ketepatan nilai prediksi model sebagaimana dijelaskan pada bagian sebelumnya dilihat dari nilai *Mean Absolute Percentage Error* (MAPE) pada data *out sample*. Nilai MAPE yang diperoleh sebesar 0,65 persen menunjukkan model GSTARI(1,1,1)-ARCH(1) dengan bobot invers jarak dapat digunakan sebagai model alternatif untuk menghitung dan memprediksi IHK empat kota di Provinsi Sumatera Utara. Peramalan IHK dengan model GSTARI(1,1,1)-ARCH(1) selama tiga bulan ke depan disajikan pada tabel berikut:

Tabel 2. Hasil Peramalan IHK Empat Kota di Provinsi Sumatera Utara

Bulan	Medan	Pematang siantar	Sibolga	Padang sidempuan
Januari 2017	133,65	132,12	132,81	125,91
Februari 2017	134,16	132,51	133,14	126,17
Maret 2017	134,94	133,10	133,62	126,55

Hasil peramalan IHK menggunakan model GSTARI(1,1,1)-ARCH(1) dengan bobot lokasi invers jarak menunjukkan bahwa IHK di Kota Medan, Pematangsiantar, Sibolga, dan Padangsidempuan cenderung akan mengalami peningkatan hingga bulan Maret 2017.

4. Kesimpulan

Model GSTARI-ARCH dibangun berdasarkan fenomena data IHK empat kota di Provinsi Sumatera Utara yang tidak stasioner sehingga dilakukan proses *differencing* untuk memperoleh data yang stasioner. Pengujian terhadap varians

error model menunjukkan adanya efek ARCH *error*, yakni kondisi yang menunjukkan varians *error* yang tidak konstan, sehingga pemodelan data IHK empat kota di Provinsi Sumatera Utara dimodelkan dengan model GSTARI-ARCH.

Berdasarkan hasil identifikasi orde model melalui plot STPACF dan nilai AIC terkecil, pemodelan dan peramalan data IHK empat kota di Provinsi Sumatera Utara dalam penelitian ini menggunakan model GSTARI(1,1,1)-ARCH(1) dengan bobot invers jarak. Berdasarkan nilai MAPE yang diperoleh yakni sebesar 0,65 persen menunjukkan bahwa model GSTARI-ARCH dengan bobot invers jarak dapat digunakan sebagai model alternatif untuk menghitung dan meramalkan IHK empat kota di Provinsi Sumatera Utara. Model ini secara umum dapat pula direkomendasikan sebagai model alternatif dalam pengambilan kebijakan sektor ekonomi terkait inflasi di Provinsi Sumatera Utara.

Pernyataan Terima Kasih. Paper ini didanai oleh Hibah Penelitian Kerjasama Luar Negeri (KLN) dan Publikasi Internasional 2017 Nomor Kontrak: 718/UN6.3.1/PL/2017 serta *Academic Leadership Grant* Universitas Padjadjaran Tahun 2017 Nomor Kontrak: 872/UN6.3.1/LT/2017

Referensi

- [1] Borovkova, S. A., Lopuhaa, H. P., and Ruchjana, B. N., Consistency and Asymptotic Normality of Least Squares Estimators of Generalized STAR Models, *Statistica Neerlandica* (2008), 482-508.
- [2] Engle, R. F., Autoregressive Conditional Heteroscedasticity with Estimates of The Variance of United Kingdom Inflation, *Econometrica* 50, (1982), 987 – 1008.
- [3] Mulyaningsih, T., *Model Generalized Space Time Autoregressive Integrated Untuk Peramalan Indeks Harga Konsumen Beberapa Kota di Jawa Tengah*, Tesis: Tidak dipublikasikan, (Bandung: Universitas Padjadjaran, 2015).
- [4] Nainggolan, N., Ruchjana, B. N., Darwis, S. and Siregar, R. E., GSTAR Models with ARCH Errors and the Simulation, *Proceeding of the Third International Conference on Mathematics and Natural Sciences (ICMNS)*, (2010), 1075-1084.
- [5] Nainggolan, N., *Pengembangan Model GSTAR dengan Galat ARCH dan Penerapannya pada Inflasi*, Disertasi: Tidak dipublikasikan, (Bandung: Program Pascasarjana Universitas Padjadjaran, 2011).
- [6] Pfeifer, P. E., and Deutsch, S. J., A Three Stage Iterative Procedure for Space Time Modeling, *Technometrics* 22, (1980), 35-47.
- [7] Ruchjana, B. N., *Suatu Model Generalisasi Space Time Autoregresi dan Penerapannya pada Produksi Minyak Bumi*, Disertasi: Tidak dipublikasikan, (Bandung: Program Doktor Institut Teknologi Bandung, 2002).
- [8] Ruchjana, B. N., Borovkova, S. A., and Lopuhaa, H. P., Least Squares Estimation of Generalized Space Time AutoRegressive (GSTAR) Model and Its Properties, *AIP Conf. Proc.*, (2012), 61-64.
- [9] Wei, W.W.S., *Time Series Analysis Univariate and Multivariate Methods, Second Edition* (Pearson Addison Wesley, Boston, 2006).
- [10] Tsay, R. S., *Multivariate Time Series Analysis with R and Financial Application*, (John Wiley and Sons, Inc., USA, 2014).

ANALISIS HARGA DAGING AYAM DI BANDUNG DENGAN MENGGUNAKAN *VECTOR ERROR CORRECTION MODEL*

RIBUT NURUL TRI WAHYUNI

Sekolah Tinggi Ilmu Statistik, rnurult@stis.ac.id

Abstrak. Kebutuhan konsumsi daging ayam penduduk Indonesia cenderung terus meningkat sejalan dengan meningkatnya jumlah penduduk dan kesadaran masyarakat akan pentingnya protein hewani. Namun beberapa tahun terakhir, harga daging ayam cenderung berfluktuasi seiring meningkatnya harga barang substitusi (harga daging sapi) dan harga pakan ternak yang mayoritas merupakan komoditas impor. Berdasarkan data Badan Pusat Statistik (BPS) selama lima tahun terakhir (2011-2015), Provinsi Jawa Barat, Jawa Timur, dan Jawa Tengah merupakan sentra produksi daging ayam yang memasok kebutuhan nasional dengan kontribusi masing-masing 47,12 persen, 12,37 persen, 6,94 persen. Fluktuasi harga daging ayam di tiga provinsi tersebut tentunya akan sangat berpengaruh terhadap harga daging ayam Indonesia. Pergerakan harga di tiga provinsi tersebut juga cenderung menunjukkan pola yang sama sehingga ada kemungkinan perubahan harga di salah satu provinsi dipengaruhi oleh provinsi lainnya. Tujuan dari penelitian ini adalah melihat pengaruh harga daging ayam di daerah sekitarnya terhadap harga daging ayam di Jawa Barat, diwakili Kota Bandung, sebagai sentra produksi daging ayam terbesar di Indonesia dengan *Vector Error Correction Model* (VECM). VECM merupakan bentuk *Vector Autoregression* (VAR) yang terestriksi. Data yang digunakan merupakan data harga eceran daging ayam per kilogram periode mingguan dari BPS selama 2011-2015. Hasilnya menunjukkan bahwa harga daging ayam di Bandung dipengaruhi oleh harga daging ayam di Bandung, di Semarang, dan di Surabaya.

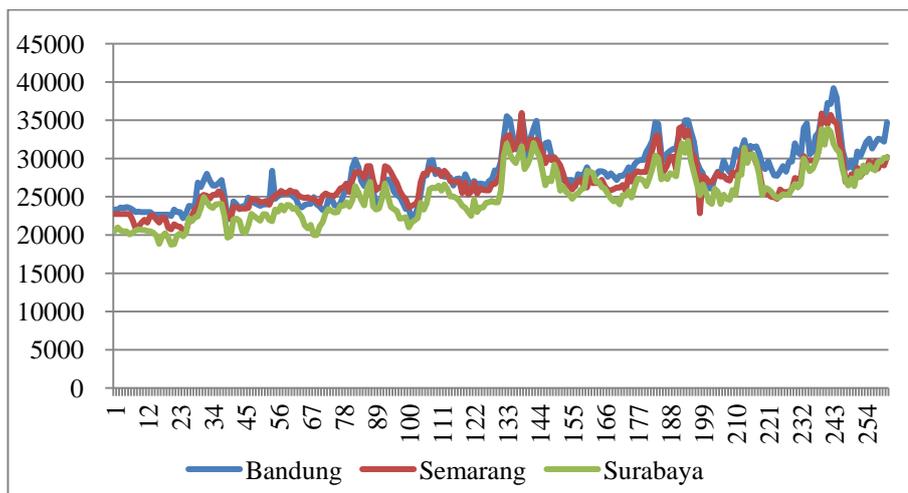
Kata kunci : harga daging ayam, VECM, VAR.

1. Pendahuluan

Kebutuhan konsumsi daging ayam penduduk Indonesia cenderung terus meningkat sejalan dengan meningkatnya jumlah penduduk Indonesia dan kesadaran masyarakat akan pentingnya protein hewani. Daging ayam lebih banyak dikonsumsi dibandingkan dengan daging sapi karena harga daging ayam lebih terjangkau dibandingkan daging sapi, khususnya daging ayam ras. Indonesia mulai tahun 2010 sudah swasembada daging ayam. Dengan kata lain, kebutuhan daging ayam dapat dicukupi dari produksi dalam negeri. Neraca produksi dan konsumsi menunjukkan nilai positif, artinya produksi dalam negeri masih mencukupi untuk kebutuhan konsumsi dalam negeri (Kementerian Pertanian [1]).

Meskipun sudah swasembada daging ayam, harga daging ayam beberapa tahun terakhir cenderung berfluktuasi seiring meningkatnya harga barang substitusi (harga daging sapi) dan harga pakan ternak yang mayoritas

merupakan komoditas impor. Berdasarkan rata-rata populasi ayam selama lima tahun terakhir (2011-2015) per provinsi, Provinsi Jawa Barat, Jawa Timur, dan Jawa Tengah merupakan sentra produksi daging ayam yang memasok kebutuhan nasional dengan besaran kontribusi masing-masing 47,12 persen, 12,37 persen, 6,94 persen¹. Fluktuasi harga daging ayam di tiga provinsi tersebut tentunya akan sangat berpengaruh terhadap harga daging ayam Indonesia. Berdasarkan data BPS (BPS [2][3][4][5][6][7][8][9][10][11]) yang bisa dilihat di Gambar 1, pergerakan harga daging ayam di tiga provinsi tersebut cenderung menunjukkan pola yang sama sehingga ada kemungkinan perubahan harga di salah satu provinsi dipengaruhi oleh provinsi lainnya. Berdasarkan latar belakang di atas, penelitian ini bertujuan untuk melihat pengaruh harga daging ayam di ibukota Provinsi Jawa Tengah (Semarang) dan Jawa Timur (Surabaya) terhadap harga daging ayam di ibukota Provinsi Jawa Barat (Bandung).



Gambar 1. Harga daging ayam mingguan di Ibukota Provinsi Jawa Barat, Jawa Tengah, dan Jawa Timur tahun 2011-2015 (rupiah)

¹http://www.pertanian.go.id/ASEM2015-NAK/Prod_DagingAyamRasPedaging_Prop_2015.pdf

2. Hasil – Hasil Utama

Metode Penelitian

Data yang digunakan dalam penelitian ini berasal dari publikasi “Perkembangan Mingguan Harga Eceran Beberapa Jenis Bahan Pokok di Ibukota Provinsi Seluruh Indonesia” tahun 2011-2015 yang diterbitkan oleh BPS. Metode yang digunakan adalah *Vector Error Correction Model* (VECM) karena datanya tidak stasioner dan ada kointegrasi. Spesifikasi VECM merestriksi hubungan jangka panjang variabel endogen agar konvergen ke dalam hubungan kointegrasinya, namun tetap membiarkan keberadaan dinamisasi jangka pendek. VECM merupakan bentuk *Vector Autoregression* (VAR) yang terestriksi (Juanda dan Junaidi [12]).

Konsep VAR diperkenalkan oleh Sims yang membuat persamaan simultan dan antar variabel diberlakukan sama sehingga tidak ada lagi variabel endogen dan eksogen. Model VAR tidak banyak tergantung pada teori, melainkan hanya perlu menentukan variabel dan jumlah lag dalam sistem. Kelebihan analisis VAR adalah metodenya sederhana tanpa membedakan variabel endogen dan eksogen, metode *Ordinary Least Square* (OLS) dapat diterapkan pada setiap persamaan secara terpisah, dan mayoritas hasil *forecast* lebih bagus dibanding model persamaan simultan. Meskipun memiliki banyak kelebihan, ternyata VAR juga memiliki banyak kelemahan, antara lain: modelnya tidak struktural (tidak menggunakan teori), model kurang sesuai untuk analisis kebijakan, lag yang banyak dapat mengakibatkan parameter yang diestimasi menjadi banyak sehingga membutuhkan data dengan *series* yang panjang, data harus stasioner, dan koefisien dalam estimasi VAR sulit diinterpretasikan (Hidayati [13]). Model VAR dengan n variabel endogen dapat ditulis sebagai berikut:

$$Y_{1t} = \beta_{01} + \sum_{i=1}^p \beta_{i1} Y_{1t-i} + \sum_{i=1}^p \alpha_{i1} Y_{2t-i} + \dots + \sum_{i=1}^p \gamma_{i1} Y_{nt-i} + e_{1t}$$

.....

$$Y_{nt} = \beta_{0n} + \sum_{i=1}^p \beta_{in} Y_{1t-i} + \sum_{i=1}^p \alpha_{in} Y_{2t-i} + \dots + \sum_{i=1}^p \gamma_{in} Y_{nt-i} + e_{nt}$$

Langkah pembentukan model VAR adalah:

1. Melakukan uji stasioneritas dengan *Augmented Dicky Fuller (ADF) test*. *ADF test* digunakan untuk melihat ada atau tidaknya *unit root* dalam suatu data. Jika data memiliki *unit root*, maka data tersebut tidak stasioner sehingga tidak dapat digunakan dalam peramalan data. Persamaan dari uji ADF adalah sebagai berikut:

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \alpha \sum_{i=1}^m \Delta Y_{t-1} + \varepsilon_t$$

Y_t = data pada periode ke- t

β_1 = *intercept*

β_2 = koefisien tren

t = tren

ε_t = *error term*

m = panjangnya lag

Berdasarkan model tersebut kita dapat memilih model yang akan digunakan untuk melakukan Uji ADF yaitu:

1. Model dengan *intercept* dan tren, seperti pada model diatas.
2. Model dengan *intercept* dan tanpa tren, yaitu:

$$\Delta Y_t = \beta_1 + \delta Y_{t-1} + \alpha \sum_{i=1}^m \Delta Y_{t-1} + \varepsilon_t$$

3. Model tanpa *intercept* dan tren, yaitu:

$$\Delta Y_t = \delta Y_{t-1} + \alpha \sum_{i=1}^m \Delta Y_{t-1} + \varepsilon_t$$

Dalam pemilihan model uji ADF yang digunakan, kita terlebih dahulu melihat ada atau tidaknya *intercept* dan tren pada data yang kita gunakan. Ketika ada tren dan *intercept* pada data, maka digunakan model dengan

intercept dan tren. Ketika tidak ada tren, maka digunakan model dengan *intercept* saja. Hipotesisnya adalah

H_0 : data tidak stasioner

H_1 : data stasioner

Jika nilai absolut statistik ADF lebih besar dari nilai kritisnya, maka tolak H_0 dan dapat disimpulkan data stasioner. Jika data tidak stasioner, maka perlu dilakukan transformasi agar data menjadi stasioner.

2. Melakukan uji kausalitas Granger untuk melihat apakah antar variabel, misalnya X dan Y, memiliki hubungan dua arah atau satu arah saja. Uji ini dimaksudkan untuk mendeteksi arah hubungan antara dua variabel dengan lag-lag yang terdistribusi dalam suatu sistem VAR. Hipotesisnya adalah

H_0 : variabel X tidak *Granger cause* variabel Y

H_1 : variabel X *Granger cause* variabel Y

Jika *p-value* kurang dari 0,05; maka keputusannya tolak H_0 dan dapat disimpulkan bahwa variabel X memiliki hubungan dengan variabel Y.

3. Menentukan panjang lag optimum, yaitu panjang lag dengan nilai *Final Prediction Error (FPE)*, *Akaike Information Criterion (AIC)*, *Schwarz Information Criterion (SIC)*, dan *Hannan-Quinn Information Criterion (HQ)* terkecil. Selain itu, digunakan juga kriteria tambahan, yaitu nilai *adjusted R²*-nya tertinggi. Kriteria ini dibutuhkan untuk mencari model terbaik.
4. Melakukan uji kointegrasi. Kombinasi linier dari dua atau lebih variabel *time series* yang non stasioner kemungkinan dapat menghasilkan residual yang stasioner dan variabel-variabel tersebut dikatakan terkointegrasi. Artinya, antar variabel dalam sistem VAR memiliki hubungan jangka panjang. Jika variabel stasioner dan tidak ada kointegrasi, maka digunakan VAR. Jika variabel tidak stasioner dalam level dan stasioner dalam *difference* namun ada kointegrasi, maka digunakan VECM. Uji kointegrasi yang digunakan adalah uji Johansen dengan kesimpulan ada kointegrasi jika *p-value* kurang dari 0,05.
5. Menganalisis *Impulse Response Function (IRF)*. Analisis IRF digunakan untuk mengetahui pengaruh dari suatu variabel terhadap *impulse* pada variabel lain dalam sistem yang melibatkan sejumlah variabel di dalamnya.
6. Menganalisis *Forecast Error Variance Decomposition (FEVD)*. Analisis FEVD digunakan untuk melihat kontribusi atau proporsi atau peranan dari suatu variabel terhadap *shock* pada variabel lain dalam sistem.

Hasil dan Pembahasan

Langkah pertama untuk mendapatkan model adalah melakukan uji stasioneritas. Tujuan dilakukannya uji stasioneritas adalah menghindari terjadinya regresi palsu (*spurious regression*), yaitu regresi yang menggambarkan hubungan antara dua variabel atau lebih yang nampak signifikan secara statistik tapi sebenarnya tidak sehingga dapat mengakibatkan *misleading* dalam penelitian. Apabila semua variabel stasioner, maka koefisien dalam model akan menjadi valid. VECM dapat digunakan apabila variabel dalam model tidak stasioner pada tingkat level atau $I(0)$. Dalam penelitian ini, uji stasioneritas yang digunakan adalah *ADF test*. Apabila nilai absolut t-statistik pada *ADF test* kurang dari nilai kritis tabel ADF dengan tingkat kepercayaan tertentu, maka dapat disimpulkan datanya tidak stasioner. Hasil uji stasioneritas dapat dilihat di Tabel 1. Berdasarkan hasil tersebut,

dapat dilihat bahwa seluruh variabel dalam penelitian ini telah stasioner pada derajat yang sama, yaitu pada derajat integrasi satu (I(1)).

Tabel 1. Hasil uji stasioneritas

Variabel	Derajat integrasi	t-statistik pada <i>ADF test</i>	nilai kritis tabel ADF dengan $\alpha = 5$ persen	Kesimpulan
Harga Bandung	Level	-2,68	-2,87	Tidak stasioner
Harga Surabaya	Level	-2,80	-2,87	Tidak stasioner
Harga Semarang	Level	-3,15	-2,87	Stasioner
Harga Bandung	<i>First difference</i>	-12,84	-2,87	Stasioner
Harga Surabaya	<i>First difference</i>	-16,11	-2,87	Stasioner
Harga Semarang	<i>First difference</i>	-17,07	-2,87	Stasioner

Tabel 2. Hasil uji kausalitas Granger

Hipotesis null	F-Statistik	<i>p-value</i>
SEMARANG does not Granger Cause BANDUNG	4,56331	0,0014
BANDUNG does not Granger Cause SEMARANG	4,62569	0,0013
SURABAYA does not Granger Cause BANDUNG	6,47688	0,0001
BANDUNG does not Granger Cause SURABAYA	3,90401	0,0043
SURABAYA does not Granger Cause SEMARANG	5,82378	0,0002
SEMARANG does not Granger Cause SURABAYA	2,10589	0,0806

Langkah selanjutnya setelah melakukan uji stasioneritas adalah melakukan uji kausalitas Granger. Hasilnya bisa dilihat di Tabel 2. Hasilnya ada hubungan dua arah antara harga Semarang-Bandung, Surabaya-Bandung, dan Surabaya-Semarang. Langkah ketiga adalah menentukan panjang lag optimum untuk seluruh variabel yang digunakan dalam model, yaitu spesifikasi lag yang menghasilkan nilai modulus kurang dari satu. Hasilnya bisa dilihat di Tabel 3. Dari tabel *AR roots*, ternyata spesifikasi lag dengan nilai modulus kurang dari satu berada pada lag 34. Berdasarkan kriteria nilai *Final Prediction Error (FPE)*, *Akaike Information Criterion (AIC)*, *Schwarz Information Criterion (SIC)*, dan *Hannan-Quinn Information Criterion (HQ)*, panjang lag optimum yang terpilih adalah 1 dan 4.

Tabel 3. Nilai FPE, AIC, SIC, dan HQ

Lag	FPE	AIC	SIC	HQ
0	1.71e+19	52.79726	52.84253	52.81553
1	8.99e+17	49.85384	50.03490*	49.92690*
2	8.55e+17	49.80381	50.12066	49.93166
3	8.37e+17	49.78160	50.23424	49.96425
4	8.29e+17*	49.77277*	50.36120	50.01021
5	8.55e+17	49.80255	50.52677	50.09478

Keterangan: *) panjang lag optimum

Keberadaan variabel yang stasioner pada derajat yang sama dapat

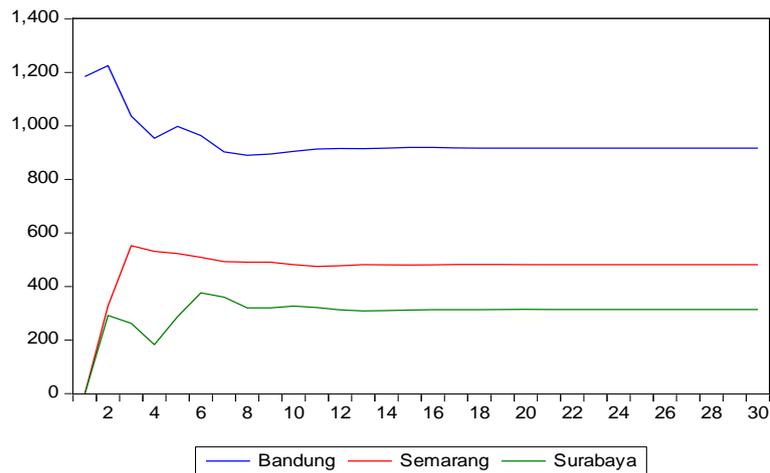
mengindikasikan adanya hubungan linier jangka panjang antar variabel tersebut. Pernyataan tersebut dapat dibuktikan dengan menggunakan uji kointegrasi (*Johansen cointegration*). Hasil dari uji kointegrasi adalah persamaan kointegrasi memiliki *intercept* tanpa tren dengan satu kointegrasi pada *trace test* ($p\text{-value} = 0,0151$) dan hasilnya bisa dilihat di Lampiran 1. Dengan demikian, model yang digunakan dalam penelitian ini adalah VECM. Berdasarkan nilai *adjusted R²* terbesar dan AIC terkecil pada VECM, panjang lag optimum yang terpilih adalah lag 4. Persamaan yang dihasilkan dari VECM dapat dilihat di Tabel 4.

Tabel 4. Hasil estimasi VECM

Variabel	Koefisien	t-statistik	Kesimpulan
Jangka panjang			
Harga Bandung (-1)	1,00		
Harga Surabaya (-1)	-1,34	-10,14	Signifikan
Harga Semarang (-1)	0,28	2,04	Signifikan
C	-1691,75		
Jangka pendek			
D(harga bandung (-1))	-0,02	-0,2	Tidak signifikan
D(harga bandung (-2))	-0,23	-2,29	Signifikan
D(harga bandung (-3))	-0,04	-0,42	Tidak signifikan
D(harga bandung (-4))	-0,01	-0,12	Tidak signifikan
D(harga surabaya (-1))	0,05	0,41	Tidak signifikan
D(harga surabaya (-2))	-0,24	-2,06	Signifikan
D(harga surabaya (-3))	-0,29	-2,56	Signifikan
D(harga surabaya (-4))	-0,03	-0,24	Tidak signifikan
D(harga semarang (-1))	0,21	2,09	Signifikan
D(harga semarang (-2))	0,42	3,78	Signifikan
D(harga semarang (-3))	0,30	2,65	Signifikan
D(harga semarang (-4))	0,09	0,88	Tidak signifikan
<i>Error correction</i>	-0,26	-3,06	Signifikan

Pada persamaan jangka pendek di Tabel 4, harga ayam di Bandung pada lag 2 secara signifikan berpengaruh negatif terhadap harga ayam di Bandung pada periode berjalan dengan tingkat kepercayaan 95 persen. Kenaikan harga di Bandung kemungkinan direspon dengan penambahan *supply* ayam sehingga terjadi penurunan harga pada periode berikutnya. Harga ayam di Surabaya pada lag 2 dan lag 3 secara signifikan berpengaruh negatif terhadap harga ayam di Bandung pada periode berjalan dengan tingkat kepercayaan 95 persen. Sedangkan harga ayam di Semarang pada lag 1, lag 2, dan lag 3 secara signifikan berpengaruh positif terhadap harga ayam di Bandung pada periode berjalan dengan tingkat kepercayaan 95 persen. Hal terpenting dari persamaan jangka pendek ini adalah nilai dari *error correction term* (*speed of adjustment*) yang signifikan dan negatif yang berarti model memiliki spesifikasi yang valid sehingga hasil VECM dapat digunakan untuk melihat pengaruh jangka panjang. *Error correction term* sebesar -0,26 (baris paling bawah di Tabel 4) mengindikasikan disequilibrium harga ayam di minggu sebelumnya akan dikoreksi di minggu sekarang sebesar 26 persen dan dibutuhkan 4 minggu (pembulatan dari 3,8) untuk dapat kembali ke kondisi keseimbangan.

Pada persamaan jangka panjang di Tabel 4, harga ayam di Surabaya pada lag 1 secara signifikan berpengaruh negatif terhadap harga ayam di Bandung pada periode berjalan dengan tingkat kepercayaan 95 persen. Sebaliknya, harga ayam di Semarang pada lag 1 secara signifikan berpengaruh positif terhadap harga ayam di Bandung pada periode berjalan dengan tingkat kepercayaan 95 persen.



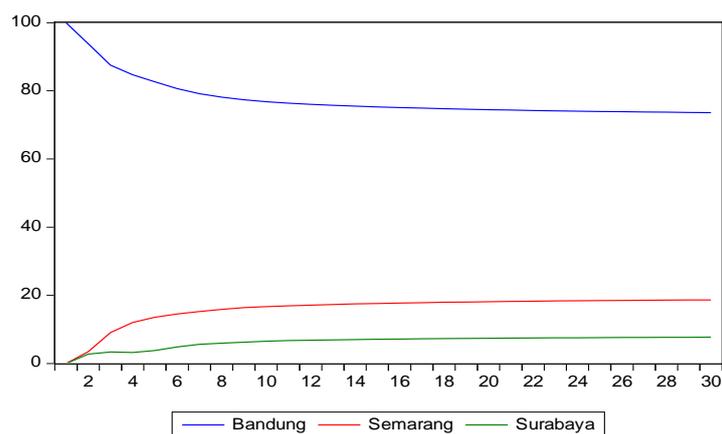
Gambar 2. Hasil IRF

Respon variabel harga ayam di Bandung akibat adanya *shock* pada harga ayam di Semarang dan Surabaya bisa dilihat di Gambar 2 dan secara kuantitatif bisa dilihat di Tabel 5. Gambar 2 dan Tabel 5 merupakan hasil dari IRF. Adanya *shock* pada harga ayam di Semarang awalnya direspon positif oleh harga ayam di Bandung pada periode kedua sebesar 327,4 rupiah dan meningkat lagi pada periode ketiga sebesar 552,2 rupiah dan masih positif cenderung menurun pada periode setelahnya sampai relatif stabil pada kisaran 481-482 rupiah (minggu ke-13). Sedangkan *shock* pada harga ayam di Surabaya awalnya direspon positif oleh harga ayam di Bandung pada periode kedua sebesar 291,8 rupiah, kemudian menurun pada periode ketiga dan keempat. Setelah periode keempat, responnya masih positif dan cenderung meningkat sampai relatif stabil pada kisaran 313-314 rupiah (minggu ke-15).

Tabel 5. Respon harga ayam di Bandung akibat *shock* pada harga ayam di Semarang dan Surabaya

Periode	Bandung	Semarang	Surabaya
1	1184,266	0,000000	0,000000
2	1225,499	327,3983	291,7699
3	1037,099	552,2423	262,1614
4	954,1265	531,1671	182,5319
5	998,2900	523,2789	286,8973
6	964,2906	508,8480	376,4767
7	903,2427	493,1395	360,4668
8	890,1666	491,0644	320,3103
9	894,9943	490,8434	320,2139
10	905,0584	481,2783	326,5391
11	913,5140	474,8125	321,2241
12	915,8242	477,8329	312,5946
13	915,2459	481,5459	308,8413
14	917,4491	481,1708	309,6771
15	920,1050	480,3132	312,1948

Hasil FEVD dapat dilihat di Gambar 3. Analisis FEVD dapat menggambarkan besarnya peranan dari setiap variabel di dalam sistem VAR karena adanya *shock*. Kontribusi terbesar pada variabel harga ayam di Bandung adalah varians pada harga ayam di Bandung itu sendiri. Kontribusi tersebut semakin menurun pada periode-periode selanjutnya. Sedangkan pengaruh terbesar kedua dalam memengaruhi variabel harga ayam di Bandung adalah varians pada harga ayam di Semarang yang cenderung meningkat pada periode-periode selanjutnya. Dengan demikian, harga ayam di Semarang memiliki peran yang lebih besar dalam memengaruhi harga ayam di Bandung dibandingkan dengan peran harga ayam di Surabaya. Ini mungkin dipengaruhi oleh kedekatan lokasi karena jarak Bandung-Semarang lebih dekat dibanding jarak Bandung-Surabaya.



Gambar 3. Hasil FEVD

3. Kesimpulan

Dengan menggunakan VECM yang hasilnya bisa dilihat di Tabel 4, dapat dilihat bahwa harga daging ayam di ibukota Provinsi Jawa Tengah (Semarang) dan Jawa Timur (Surabaya) berpengaruh terhadap harga daging ayam di ibukota Provinsi Jawa Barat. Ketika terjadi *shock* pada harga ayam di Semarang, harga ayam di Bandung membutuhkan waktu 13 minggu untuk mencapai kondisi keseimbangannya. Sedangkan *shock* pada harga ayam di Surabaya, harga ayam di Bandung membutuhkan waktu 15 minggu untuk mencapai kondisi keseimbangannya. Kontribusi terbesar pada variabel harga ayam di Bandung adalah varians pada harga ayam di Bandung itu sendiri, kemudian harga daging ayam di Semarang.

Referensi

- [1] Kementerian Pertanian, 2015, *Outlook Komoditas Pertanian Sub Sektor Peternakan Daging Ayam*, Kementerian Pertanian.
- [2] BPS, 2011a, *Perkembangan Mingguan Harga Eceran Beberapa Jenis Bahan Pokok di Ibukota Provinsi Seluruh Indonesia Januari-Juni 2011*, BPS.
- [3] BPS, 2011b, *Perkembangan Mingguan Harga Eceran Beberapa Jenis Bahan Pokok di Ibukota Provinsi Seluruh Indonesia Juli-Desember 2011*, BPS.
- [4] BPS, 2012a, *Perkembangan Mingguan Harga Eceran Beberapa Jenis Bahan*

- Pokok di Ibukota Provinsi Seluruh Indonesia Januari-Juni 2012*, BPS.
- [5] BPS, 2012b, *Perkembangan Mingguan Harga Eceran Beberapa Jenis Bahan Pokok di Ibukota Provinsi Seluruh Indonesia Juli-Desember 2012*, BPS.
- [6] BPS, 2013a, *Perkembangan Mingguan Harga Eceran Beberapa Jenis Bahan Pokok di Ibukota Provinsi Seluruh Indonesia Januari-Juni 2013*, BPS.
- [7] BPS, 2013b, *Perkembangan Mingguan Harga Eceran Beberapa Jenis Bahan Pokok di Ibukota Provinsi Seluruh Indonesia Juli-Desember 2013*, BPS.
- [8] BPS, 2014a, *Perkembangan Mingguan Harga Eceran Beberapa Jenis Bahan Pokok di Ibukota Provinsi Seluruh Indonesia Januari-Juni 2014*, BPS.
- [9] BPS, 2014b, *Perkembangan Mingguan Harga Eceran Beberapa Jenis Bahan Pokok di Ibukota Provinsi Seluruh Indonesia Juli-Desember 2014*, BPS.
- [10] BPS, 2015a, *Perkembangan Mingguan Harga Eceran Beberapa Jenis Bahan Pokok di Ibukota Provinsi Seluruh Indonesia Januari-Juni 2015*, BPS.
- [11] BPS, 2015b, *Perkembangan Mingguan Harga Eceran Beberapa Jenis Bahan Pokok di Ibukota Provinsi Seluruh Indonesia Juli-Desember 2015*, BPS.
- [12] Juanda, B., Junaidi, 2012, *Ekonometrika Deret Waktu Teori & Aplikasi*, IPB Press.
- [13] Hidayati, S., 2009, Analisis Hubungan Kinerja Sistem Keuangan (Perbankan dan Pasar Modal) Terhadap Pertumbuhan Ekonomi Indonesia Periode 1999-2008, *Universitas Indonesia [Tesis]*.

Lampiran

Lampiran 1. Hasil uji kointegrasi

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.090424	39.72666	35.19275	0.0151
At most 1	0.055843	18.30702	20.26184	0.0908
At most 2	0.023266	5.320359	9.164546	0.2501

Trace test indicates 1 cointegrating eqn(s) at the 0.05 level

* denotes rejection of the hypothesis at the 0.05 level

**MacKinnon-Haug-Michelis (1999) p-values

MODEL *SPATIO TEMPORAL* DALAM PEMETAAN PENYAKIT DIARE DI KOTA BANDUNG

I GEDE NYOMAN MINDRA JAYA¹, BUDI NURANI
RUCHJANA²

¹Departmen Statistika FMIPA UNPAD, mindra@unpad.ac.id
²Departmen Matematika FMIPA UNPAD, bnurani@gmail.com

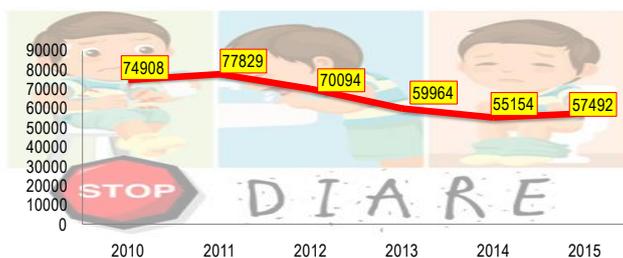
Abstrak. Pemodelan pemetaan penyakit khususnya penyakit menular menjadi suatu keharusan bagi upaya menekan penyebarluasan penyakit [1]. Peta penyebaran penyakit menular akan sangat membantu pengambil kebijakan dalam merumuskan tindakan apa yang perlu diambil dalam rangka menekan dan mengurangi dampak dari penyakit [2]. Salah satu kasus penyakit menular yang sampai saat ini masih menjadi masalah adalah Diare. Angka prevalensi diare di Kota Bandung mendekati angka prevelansi diare Jawa Barat. Diperlukan upaya serius untuk menanggulangi kejadian diare di Kota Bandung. Salah satu upaya yang dapat dilakukan adalah melakukan deteksi dini tentang pola penyebaran dan faktor-faktor yang mempengaruhi kejadian diare di Kota Bandung. Pemodelan pemetaan diare umumnya dilakukan hanya mengakomodasi variansi spasial tanpa memperhatikan faktor waktu. Penelitian ini bertujuan membuat pemodelan *spatio temporal* untuk dapat mengidentifikasi pola penyebaran diare di Kota Bandung menurut ruang dan waktu serta mengkaji berbagai faktor lingkungan yang berpengaruh terhadap angka kasus diare di kota Bandung. Hasil penelitian menemukan bahwa perilaku hidup sehat dan kepadatan penduduk adalah dua faktor yang berpengaruh pada tingginya angka kasus diare di Kota Bandung.

Kata kunci: Diare, Pemetaan Penyakit, Spati Temporal

1. Pendahuluan

Penyakit Diare adalah salah satu jenis penyakit menular yang menjadi masalah kesehatan masyarakat di dunia [3]. Diare adalah keadaan dimana seseorang kehilangan cairan tubuh dan elektrolit yang berlebihan melalui feses, ditandai dengan bertambahnya frekuensi defekasi lebih dari tiga kali dalam sehari disertai perubahan konsistensi tinja menjadi cair, dengan/tanpa darah dan lendir [4].

Kota Bandung sebagai salah satu kota dengan penduduk terpadat di Indonesia juga tidak lepas dari masalah penyakit Diare. Menurut data Dinas Kesehatan Kota Bandung, terjadi trend penurunan angka kejadian diare dari tahun 2010 sampai tahun 2015. Namun demikian angka kasus ini masih dipandang tinggi dengan rata-rata mencapai 57.536 kasus setiap tahunnya (Gambar 1).



Sumber: Diolah dari data Dinas Kesehatan Kota Bandung

Gambar 1. Angka Kejadian Diare di Kota Bandung Periode 2010-2015

Untuk menekan angka kasus dan penyebaran penyakit diare diperlu upaya preventif dari pemerintah dan juga masyarakat sehingga angka kasus diare dapat diturunkan secara berkala. Studi epidemiology membahas bagaimana suatu penyakit terjadi, bagaimana pola penyebaran dan faktor apa saja yang menjadi penyebab semakin tingginya angka kasus satu penyakit [5]

Angka prevalensi diare di Kota Bandung mendekati angka prevelansi diare Jawa Barat. Diperlukan upaya serius untuk menanggulangi kejadian diare di Kota Bandung. Salah satu upaya yang dapat dilakukan adalah melakukan deteksi dini tentang pola penyebaran dan faktor-faktor yang mempengaruhi kejadian diare di Kota Bandung.

Pemodelan dan pemetaan penyakit diare menjadi suatu keharusan bagi upaya menekan penyebarluasan penyakit [1]. Peta penyebaran penyakit diare akan sangat membantu pengambil kebijakan dalam merumuskan tindakan apa yang perlu diambil dalam rangka menekan dan mengurangi dampak dari penyakit diare [2]. Pemodelan dan pemetaan diare umumnya dilakukan hanya mengakomodasi variansi spasial tanpa memperhatikan faktor waktu. Penelitian ini bertujuan membuat pemodelan *spatio temporal* untuk dapat mengidentifikasi pola penyebaran diare di Kota Bandung menurut ruang dan waktu serta mengkaji berbagai faktor lingkungan yang berpengaruh terhadap angka kasus diare di kota Bandung.

2. Metode Penelitian

Penyakit Menular

Penyakit menular menjadi ancaman serius bagi kesehatan masyarakat [2]. Penyakit ini disebabkan oleh patogen dan ditularkan melalui berbagai cara diantaranya dari orang ke orang melalui kontak baik langsung maupun tidak langsung, melalui agen seperti hewan. Berdasarkan kejadiannya, terdapat empat jenis penyakit menular. Penyakit yang terjadi kadang-kadang pada populasi diklasifikasikan sebagai sporadis; jika terjadi terus-menerus dikatakan endemik; jika sejumlah besar korban terinfeksi dalam waktu singkat dikatakan epidemi; dan jika itu terjadi di seluruh dunia dalam waktu singkat dikatakan sebagai pandemi.

Diare merupakan salah satu jenis penyakit yang sangat menular baik disebabkan oleh bakteri ataupun virus. Perilaku hidup tidak sehat akan sangat mempercepat penularan penyakit ini pada orang lain. Membuang air besar tidak secara sehat, tidak mencuci bersih tangan saat makan, dan memakan makanan yang

terjangkit lalat menjadi pemicu penularan penyakit diare.

Model Dasar Pemetaan Penyakit

Standardized Morbidity / Mortality Ratio (SMR)

SMR adalah model dasar dari pemetaan penyakit. Model ini merupakan sebuah ukuran statistik yang sering digunakan dalam studi epidemiologi untuk mengidentifikasi lokasi dengan tingkat resiko tinggi. Misalkan angka kasus pada lokasi ke- i dinyatakan dengan y_i dan angka harapan kasus pada lokasi yang sama dinyatakan dengan e_i ($i = 1,2,3, \dots, N$), SMR pada lokasi ke- i didefinisikan sebagai berikut [6]:

$$SMR_i = \frac{y_i}{e_i}, \quad (1)$$

dengan e_i didefinisikan sebagai:

$$e_i = N_i \times \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n N_i}, \quad (2)$$

dengan N_i menyatakan ukuran populasi berisiko pada lokasi ke- i . Lokasi dengan nilai SMR lebih besar dari 1 dapat diidentifikasi sebagai lokasi dengan resiko relative tinggi [7]. SMR sendiri merupakan penaksir maksimum likelihood (ML) parameter relative risk yang diperoleh sebagai berikut:

$$y_i | e_i \theta_i \sim \text{Poisson}(e_i \theta_i). \quad (3)$$

$$\hat{\theta}_i^{ML} = \frac{y_i}{e_i}, \quad (4)$$

merupakan penaksir tak bias dengan varians sebagai berikut:

$$V(\hat{\theta}_i^{ML}) = \frac{\hat{\theta}_i^{ML}}{e_i}. \quad (5)$$

Penaksir resiko relative SMR memiliki beberapa kelemahan 1) penaksir ini tidak memasukkan informasi ketergantungan spasial dan temporal, 2) memiliki sensitifitas yang tinggi terhadap perbedaan ukuran populasi, yang berakibat pada ketidakhandalan sebagai penaksir resiko relative dan 3) tidak tepat digunakan untuk kasus-kasus dengan angka kejadian yang kecil, dan 4) tidak memasukkan informasi lain faktor-faktor yang mungkin berpengaruh pada variasi angka kasus [6].

Untuk mengatasi tersebut mengusulkan pendekatan metode Bayesian untuk mengatasi ketidakhandalan dari SMR [6], dengan prinsip meminjam informasi lokasi tetangga untuk meningkatkan kualitas penaksiran resiko relative melalui pemanfaatan distribusi prior.

Ukuran SMR dapat dikembangkan untuk data *spatio temporal* dengan menambahkan indeks waktu pada elemen penyusunnya menjadi SMR_{it} . SMR_{it} menyatakan resiko relatif SMR pada lokasi ke- i dan periode waktu ke- t yang dirumuskan sebagai berikut :

$$SMR_{it} = \frac{y_{it}}{e_{it}}, \quad (6)$$

dengan angka harapan kasus pada lokasi ke- i dan periode waktu ke- t , e_{it} didefinisikan sebagai:

$$e_{it} = N_{it} \times \frac{\sum_{i=1}^n \sum_{t=1}^T y_{ii}}{\sum_{i=1}^n \sum_{t=1}^T N_{it}}, \quad (7)$$

dengan N_{it} menyatakan populasi penduduk yang berisiko pada lokasi ke- I dan

periode waktu ke- t

Model Spatio Temporal

Resiko relatif secara empiris tidak pernah sama untuk setiap periode waktu hal ini dikarenakan adanya berbagai variasi faktor-faktor yang berpengaruh tidak hanya secara spasial namun juga menurut waktu.

Model dasar yang digunakan untuk mengakomodasi adanya variasi spasial dan temporal sebagai penyempurnaan dari metode SMR adalah model Poisson Log-Normal. *Poisson log-normal* merupakan campuran dari dua distribusi yaitu distribusi *poisson* dan distribusi *log-normal*.

Dengan data pengamatan angka kasus diare pada wilayah i ($i = 1, \dots, n$) dan waktu ke t ($t = 1, 2, \dots, T$), y_{it} , berdistribusi *Poisson* dengan parameter ($E_{it}\theta_{it}$), dan selanjutnya model log-normal didefinisikan sebagai berikut [7]:

$$y_{it} \sim \text{Poisson}(E_{it}\theta_{it}) \quad (6)$$

dengan

$$\begin{aligned} E(y_{it}) &= E_{it}\theta_{it} \\ \theta_{it} &= \exp(\alpha + \mathbf{x}_{it}^T \boldsymbol{\beta} + u_{it}) \\ \eta_{it} &= \alpha + \mathbf{x}_{it}^T \boldsymbol{\beta} + u_{it} \end{aligned} \quad (7)$$

dan $\ln(\theta_{it}) \sim i.i.d \text{ Normal}(\mu, \sigma^2)$

θ_{it} : parameter risiko relatif

α : koefisien intersep yang menyatakan rata-rata risiko relatif

$\boldsymbol{\beta}$: vektor koefisien regresi untuk variabel independen \mathbf{X}

u_{it} : komponen acak yang menjelaskan ketergantungan spasial dan temporal.

Komponen u_{it} dapat didefinisikan dalam berbagai model. Salah satu model untuk mendefinisikan u_{it} yang paling banyak digunakan adalah model linear yang diberikan sebagai berikut [8] :

$$u_{it} = (\beta_0 + \phi_i) + (\mu + \delta_i) \left(\frac{t - \bar{t}}{T} \right) \quad (8)$$

dengan distribusi prior untuk masing-masing parameter dinyatakan sebagai berikut:

$$\begin{aligned} \phi_i | \phi_{-i}, \mathbf{W} &\sim \mathbf{N} \left(\frac{\rho_\phi \sum_{j=1}^K w_{ij} \phi_j}{\rho_\phi \sum_{j=1}^K w_{ij} + 1 - \rho_\phi}, \frac{\tau_\phi^2}{\rho_\phi \sum_{j=1}^K w_{ij} + 1 - \rho_\phi} \right) \\ \delta_i | \delta_{-i}, \mathbf{W} &\sim \mathbf{N} \left(\frac{\rho_\delta \sum_{j=1}^K w_{ij} \delta_j}{\rho_\delta \sum_{j=1}^K w_{ij} + 1 - \rho_\delta}, \frac{\tau_\delta^2}{\rho_\delta \sum_{j=1}^K w_{ij} + 1 - \rho_\delta} \right) \\ \tau_\phi^2, \tau_\delta^2 &\sim \text{Inverse - Gamma}(a, b) \\ \rho_\phi, \rho_\delta &\sim \text{Uniform}(0, 1) \\ \beta_0 &\sim \mathbf{N}(\mu_{\beta_0}, \sigma_{\beta_0}^2) \\ \gamma &\sim \mathbf{N}(\mu_\gamma, \sigma_\gamma^2) \end{aligned}$$

dan $\bar{t} = \left(\frac{1}{T}\right) \sum_{t=1}^T t$ menyatakan rata-rata indeks waktu, dan modifikasi linear temporal trend dinyatakan dalam $t^* = \frac{t - \bar{t}}{T}$. Setiap lokasi memiliki rata-rata risiko relatif yang berbeda dengan besaran $(\beta_0 + \phi_i)$ dan memiliki temporal trend

yang juga berbeda dan dinyatakan oleh $(\mu + \delta_i)$. Setiap elemen acak $\phi = (\phi_1, \dots, \phi_n)$ dan $\delta = (\delta_1, \dots, \delta_n)$ dimodelkan dengan *Conditional Autoregressive* (CAR) prior untuk mengakomodasi autokorelasi spasial. Parameter ρ_ϕ dan ρ_δ menyatakan parameter ketergantungan spasial. Parameter ini diasumsikan mengikuti distribusi prior uniform. Sedangkan varians τ_ϕ^2 dan τ_δ^2 diasumsikan mengikuti distribusi inverse gamma dengan parameter (a, b) . Nilai yang ditetapkan untuk parameter prior inverse gamma adalah $a = 0.001$ dan $b = 0.001$. Nilai (a, b) yang ditetapkan merujuk pada definisi *vague prior* yaitu prior yang ditetapkan sedemikian sehingga distribusi prior tidak memberikan pengaruh yang besar pada distribusi posterior dengan kata lain *vague prior* dipilih karena peneliti mengoptimalkan informasi dari data dibandingkan subjektivitas peneliti. Kemudian parameter intersep α diasumsikan mengikuti distribusi normal dengan $\mu_\alpha = 0$ dan $\sigma_\alpha^2 = 1000$. Nilai yang besar untuk σ_α^2 dikenal dengan istilah *vague prior*. Semakin besar nilai σ_α^2 maka semakin kecil peranan distribusi prior dalam membangun distribusi posteriornya.

Untuk mengestimasi parameter *spatio temporal* trend digunakan metode MCMC yang sudah terpasilitas dalam package CARBayesST [8]

3. Hasil Analisis

Angka kasus diare di Kota Bandung tidak mengalami perubahan yang signifikan, dengan angka kasus pada kisaran 57537 kasus setiap tahunnya. Kejadian Diare tidak terlepas dari kepadatan penduduk dan PHBS masyarakatnya. Hasil pengumpulan data menunjukkan secara rata-rata angka kepadatan penduduk sebesar 15960/km² dan PHBS secara rata-rata sebesar 18.12%, (Tabel 1).

Tabel 1. Statistik Variabel Penelitian

Tahun	Statistik		
	Kasus	Kepadatan	PHBS
2013	59964	15790.3	21.87
2014	55154	15713.2	14.53
2015	57492	16376.6	17.95
Rata-Rata	57537	15960.0	18.12

Untuk mengetahui model pengaruh dari variabel kepadatan penduduk dan PHBS terhadap angka kasus diare di Bandung dilakukan pemodelan *spatio temporal* dengan hasil sebagai berikut (Tabel 2):

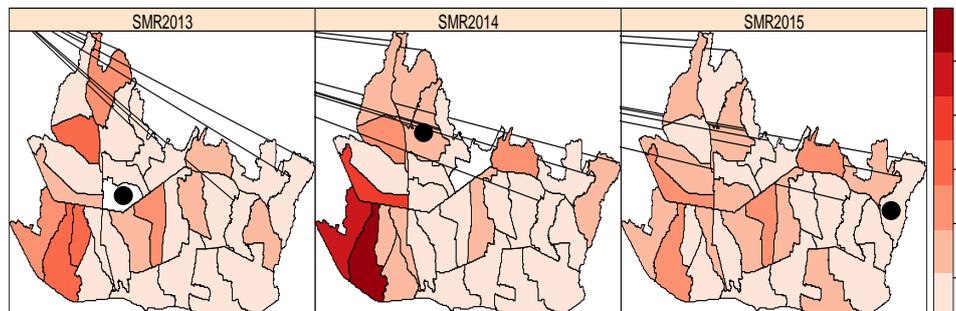
Tabel 2. Pemodelan *Spatio Temporal*

	Median	2.50%	97.5%
Intersep	5.9601	5.6487	6.1552
Kepadatan Penduduk (X1)	0.0001	0.0001	0.0001
PHBS (X2)	-0.0015	-0.0023	-0.0008
ρ_ϕ	0.1491	0.0076	0.6373
ρ_δ	0.0711	0.0025	0.4353

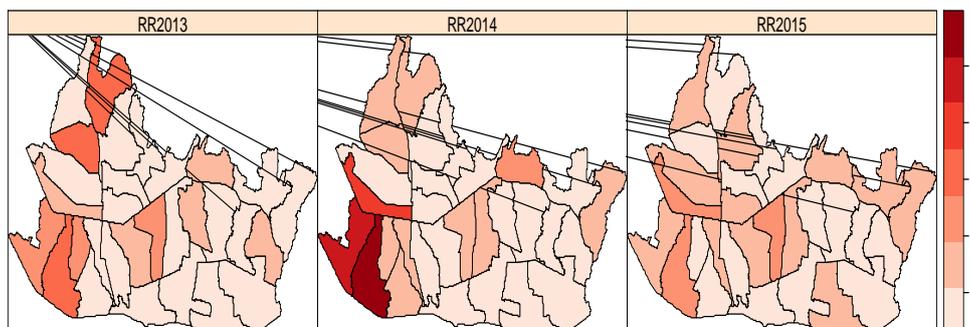
	Median	2.50%	97.5%
DIC=857.7943			

Hasil perhitungan *Spatio Temporal Model* (STMODEL) pada (Tabel 2) menemukan bahwa variabel kepadatan penduduk dan PHBS menjadi faktor penyebab utama tingginya angka kasus diare di Kota Bandung. Koefisien regresi kepadatan penduduk sebesar 0.0001 menjelaskan bahwa peningkatan kepadatan penduduk 10 orang/km² maka angka kasus akan meningkat sebesar 0.1%. Sedangkan koefisien regresi untuk PHBS sebesar -0.0015 menyatakan bahwa peningkatan perilaku hidup bersih dan sehat (PHBS) akan mampu menurunkan angka kasus mencapai 0.15%. Koefisien spasial autoregressive untuk intersep dan slop masing-masing sebesar 0.1491 dan 0.0711 menyatakan adanya ketergantungan spasial untuk angka kasus diare di Kota Bandung yang artinya bahwa tingginya angka kasus di satu kecamatan dapat menyebabkan tingginya angka kasus pada kecamatan yang berdekatan.

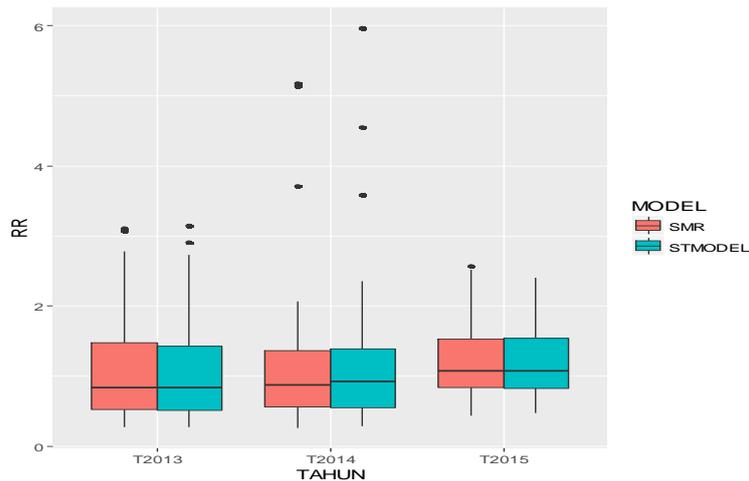
Selanjutnya dilakukan perbandingan model SMR dengan STMODEL untuk menunjukkan bahwa perhitungan dengan SMR akan memberikan hasil yang berbeda dibandingkan dengan STMODEL.



(a). MAP SMR



(b). MAP SMR



(c). BOXPLOT

Gambar 3. Perbandingan Model SMR dan STMODEL

Secara visual dapat dilihat dari Gambar 3 (a-c) bahwa pemetaan dengan menggunakan model SMR dengan STMODEL memberikan hasil yang berbeda khususnya pada Gambar 3(a) yang ditandai dengan tanda bulat hitam. Hasil taksiran resiko relatif terlihat jelas berbeda dari gambar (c) Boxplot. Berdasarkan Gambar 3 (c) ditemukan bahwa kecenderungan nilai resiko relatif berdasarkan STMODEL lebih homogen dibandingkan dengan metode SMR.

Berdasarkan ST Model diketahui bahwa pada periode Tahun 2013-2015 Kecamatan Babakan Ciparai yang ditandai dengan warna gelap pada Peta (b) memiliki angka resiko relatif paling tinggi dibandingkan dengan kecamatan lainnya di Kota Bandung.

4. Kesimpulan

Berdasarkan hasil pengolahan data yang dilakukan diperoleh simpulan bahwa terdapat perbedaan hasil estimasi parameter resiko relatif kasus diare antara metode SMR dan STMODEL.

Hasil analisis menunjukkan bahwa kepadatan penduduk dan perilaku hidup bersih dan sehat berpengaruh pada angka kasus diare

Pernyataan terima kasih. Terimakasih Kepada Rektor Universitas Padjadjaran atas dana HIBAH KLN.

Referensi

- [1] Lawson, A. B. (2006). Statistical Methods Methods in Spasial Epidemiology. England: John Wiley & Sons.
- [2] Chen, D., Moulin, B., & Wu, J. (2015). Introduction to Analyzing and Modeling Spasial and Temporal Dynamics of Infectious Diseases . In D.

- Chen, B. Moulin, & J. Wu, *Analyzing and Modeling Spasial and Temporal Dynamics of Infectious Diseases* (pp. 3-17). New Jersey: John Wiley & Sons.
- [3] Kementrian Kesehatan RI. (2011). *Situasi Diare di Indonesia*.
 - [4] Suraatmaja, Sudaryat. 2007. *Kapita Selekta Gastroenterologi*. Sagung Seto, Jakarta
 - [5] Elliott, P., & Wartenberg, D. (2004). *Spasial Epidemiology: Current Approaches and Future Challenges*. *Environmental Health Perspectives* , 112 (9), 998-1006.
 - [6] Clayton, D., & Kaldor, J. (1987). *Empirical Bayes Estimates of Age-Standardized Relative Risks for Use in Disease Mapping*. *Biometrics* , 671-681.
 - [7] Tango, T. (2010). *Statistical Methods for Disease Clustering*. Japan: Springer.
 - [8] Lee, D., Rushworth, A., dan Napier, G. (2015) *CARBayesST version 2.2: An R Package for Spatio-temporal Areal Unit Modelling with Conditional Autoregressive Priors*. CRAN-R

PENGARUH INFRASTRUKTUR SOSIAL DAN EKONOMI TERHADAP PEREKONOMIAN DI INDONESIA: APLIKASI SARMA *FIXED EFFECT MODEL*

PRIANGGA A.W.¹, RIBUT NURUL T. W.²

1 Badan Pusat Statistik, 12.7310@stis.ac.id
2 Sekolah Tinggi Ilmu Statistik, rnurult@stis.ac.id

Abstrak. Berdasarkan *World Economic Forum Report* tahun 2014, *Global Competitive Index* (GCI) Indonesia berada di peringkat ke-34 dunia. Peringkat tersebut masih dibawah Singapura (2), Malaysia (20), dan Thailand (31). Salah satu komponen pembentuk GCI adalah keberadaan infrastruktur. Untuk meningkatkan infrastruktur di Indonesia, pemerintahan Jokowi gencar melakukan pembangunan infrastruktur, meliputi: infrastruktur sosial (pendidikan dan kesehatan) serta infrastruktur ekonomi (jalan dan listrik). Pembangunan infrastruktur tersebut diharapkan bisa meningkatkan perekonomian di Indonesia. Selain infrastruktur, interaksi antarwilayah juga kemungkinan berpengaruh terhadap perekonomian. Penelitian ini ingin melihat ada tidaknya pengaruh spasial terhadap perekonomian di Indonesia dan mengidentifikasi variabel infrastruktur yang berpengaruh terhadap perekonomian di Indonesia. Model yang digunakan adalah *Spatial Autoregressive and Moving Average with Fixed Effect* (SARMA-FE). Hasil penelitian menunjukkan adanya pengaruh spasial terhadap perekonomian di Indonesia. Panjang jalan, jumlah listrik yang terjual, jumlah fasilitas kesehatan, dan jumlah fasilitas pendidikan ternyata juga berpengaruh signifikan terhadap perekonomian di Indonesia.

Kata kunci : infrastruktur, perekonomian, SARMA-FE.

1. Pendahuluan

Berdasarkan *World Economic Forum Report* tahun 2014, *Global Competitive Index* (GCI) Indonesia berada di peringkat ke-34 dunia. Peringkat tersebut masih dibawah Singapura (2), Malaysia (20), dan Thailand (31). Salah satu komponen pembentuk GCI adalah keberadaan infrastruktur. Pada tahun 2014, infrastruktur Indonesia berada di peringkat ke-56 dunia, dibawah Singapura, Malaysia, dan Thailand. Untuk meningkatkan infrastruktur di Indonesia, pemerintahan Jokowi gencar melakukan pembangunan infrastruktur baik di Jawa maupun luar Jawa, meliputi: pembangunan jalan, penyediaan listrik, fasilitas kesehatan, fasilitas pendidikan, dan prasarana infrastruktur lainnya. Pembangunan fasilitas kesehatan dan pendidikan merupakan infrastruktur sosial. Sedangkan pembangunan jalan dan penyediaan listrik merupakan infrastruktur ekonomi. Pembangunan infrastruktur tersebut diharapkan bisa meningkatkan perekonomian di Indonesia. Selain infrastruktur, interaksi antarwilayah juga kemungkinan berpengaruh terhadap perekonomian. Penelitian ini ingin melihat ada tidaknya pengaruh spasial terhadap perekonomian di Indonesia

dan mengidentifikasi variabel infrastruktur sosial dan ekonomi yang berpengaruh terhadap perekonomian di Indonesia.

2. Hasil – Hasil Utama

Metode Penelitian

Menurut Anselin [1], ekonometrika spasial merupakan metodologi yang berkaitan dengan efek spasial, seperti autokorelasi spasial dan heterogenitas spasial. Model regresi yang digunakan untuk menggambarkan hubungan antarvariabel dengan memperhatikan efek spasial yang dimilikinya dinamakan model regresi spasial. Model regresi spasial bisa juga diterapkan pada data panel sehingga model yang digunakan disebut model regresi panel spasial. Menurut Elhorst [2], terdapat beberapa model regresi panel spasial, antara lain:

1. *Spatial Autoregressive with Fixed Effect (SAR-FE)*

$$y_{it} = \lambda \sum_{j=1}^N w_{ij} y_{jt} + x_{it} \beta + \mu_i + \varepsilon_{it}$$

dengan y_{it} = variabel dependen di wilayah i pada waktu t , x_{it} = variabel independen di wilayah i pada waktu t , λ = koefisien spasial lag, w_{ij} = matriks penimbang spasial, dan μ_i = efek spesifik spasial. Pada model ini, terdapat parameter λ yang menggambarkan koefisien spasial lag.

2. *Spatial Autoregressive with Random Effect (SAR-RE)*

$$y_{it} = \lambda \sum_{j=1}^N w_{ij} y_{jt} + x_{it} \beta + \varepsilon_{it}$$

$$\varepsilon_{it} = \mu_i + v_{it}$$

dengan v_{it} = vektor error. Pada model ini, terdapat parameter λ yang menggambarkan koefisien spasial lag.

3. *Spatial Error Model with Fixed Effect (SEM-FE)*

$$y_{it} = x_{it} \beta + \mu_i + \varepsilon_{it}$$

$$\varepsilon_{it} = \rho \sum_{j=1}^N w_{ij} \varepsilon_{jt} + v_{it}$$

Pada model ini, terdapat parameter ρ yang menggambarkan koefisien spasial error.

4. *Spatial Error Model with Random Effect (SEM-RE)*

$$y_{it} = x_{it} \beta + \varepsilon_{it}$$

$$\varepsilon_{it} = \mu_i + \rho \sum_{j=1}^N w_{ij} \varepsilon_{jt} + v_{it}$$

Pada model ini, terdapat parameter ρ yang menggambarkan koefisien spasial error.

5. *Spatial Lag and Error Model with Fixed Effect (SARMA-FE)*

$$y_{it} = \lambda \sum_{j=1}^N w_{ij} y_{jt} + x_{it} \beta + \mu_i + \varepsilon_{it}$$

$$\varepsilon_{it} = \rho \sum_{j=1}^N w_{ij} \varepsilon_{jt} + v_{it}$$

Pada model ini, terdapat parameter λ yang menggambarkan koefisien spasial lag dan parameter ρ yang menggambarkan koefisien spasial error.

6. *Spatial Lag and Error Model with Random Effect (SARMA-RE)*

$$y_{it} = \lambda \sum_{j=1}^N w_{ij} y_{jt} + x_{it} \beta + \varepsilon_{it}$$

$$\varepsilon_{it} = \mu_i + \rho \sum_{j=1}^N w_{ij} \varepsilon_{jt} + v_{it}$$

Pada model ini, terdapat parameter λ yang menggambarkan koefisien spasial lag dan parameter ρ yang menggambarkan koefisien spasial error.

Langkah pertama adalah melakukan estimasi parameter model dengan menggunakan metode *maximum likelihood*. Matriks penimbang yang digunakan adalah *k-nearest neighbor*. k wilayah yang memiliki jarak terdekat dari suatu wilayah adalah tetangga bagi wilayah tersebut, sehingga nilai penimbangnya (w_{ij}) = 1. Penelitian BPS [3] menyatakan bahwa secara umum Indonesia memiliki 3 tetangga (berbatasan langsung dengan 3 wilayah). Matriks penimbang spasial (\mathbf{W}) dapat digambarkan sebagai berikut:

$$\mathbf{W} = \begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ \vdots & & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{pmatrix}$$

dengan w_{ij} adalah penimbang wilayah i dengan wilayah j . Nilai $w_{ij} = w_{ji}$ dan $w_{ii} = 0$ untuk $i = 1, 2, \dots, n$.

Setelah diperoleh estimasi parameter model, langkah kedua adalah melakukan uji *Lagrange Multiplier (LM)* dan atau uji *LM robust* pada model *SAR* dan model *SEM*. Uji *LM* pada model *SAR* digunakan untuk memilih model yang lebih baik antara model *pooled* dan *SAR*. Hipotesisnya adalah

$$H_0 : \lambda = 0 \text{ dan } \rho = 0$$

$$H_1 : \lambda \neq 0 \text{ dan } \rho = 0$$

Uji *LM* pada model *SEM* digunakan untuk memilih model yang lebih baik antara model *pooled* dan *SEM*. Hipotesisnya adalah

$$H_0 : \lambda = 0 \text{ dan } \rho = 0$$

$$H_1 : \lambda = 0 \text{ dan } \rho \neq 0$$

Jika kedua uji di atas memberi kesimpulan yang sama, bahwa model *SAR* dan model *SEM* adalah model yang lebih baik digunakan dibanding model *pooled*, maka perlu dilakukan uji lanjutan, yaitu uji *LM robust*. Uji ini memiliki kegunaan yang sama dengan uji sebelumnya, sehingga hipotesis yang digunakan juga tetap sama. Jika uji *LM robust* memberikan hasil yang sama, yaitu model *SAR* dan model *SEM* merupakan model yang lebih baik digunakan dibanding model *pooled*, maka peneliti dapat menggunakan model dengan *p-value* yang lebih kecil, menggunakan model gabungan, atau memilih salah satu di antara keduanya.

Langkah ketiga melakukan uji Hausman (Baltagi [4]). Uji tersebut untuk melihat efek yang akan digunakan dalam model, apakah efek tetap atau efek random. Hipotesisnya adalah

$$H_0 : m = 0$$

$$H_1 : m \neq 0$$

Statistik uji yang digunakan adalah

$$m = \hat{\mathbf{q}}' [\mathbf{Var}(\hat{\mathbf{q}})]^{-1} \hat{\mathbf{q}}$$

$$\hat{\mathbf{q}} = \hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{RE}$$

$$\mathbf{Var}(\hat{\mathbf{q}}) = \mathbf{Var}(\hat{\boldsymbol{\beta}}_{FE}) - \mathbf{Var}(\hat{\boldsymbol{\beta}}_{RE})$$

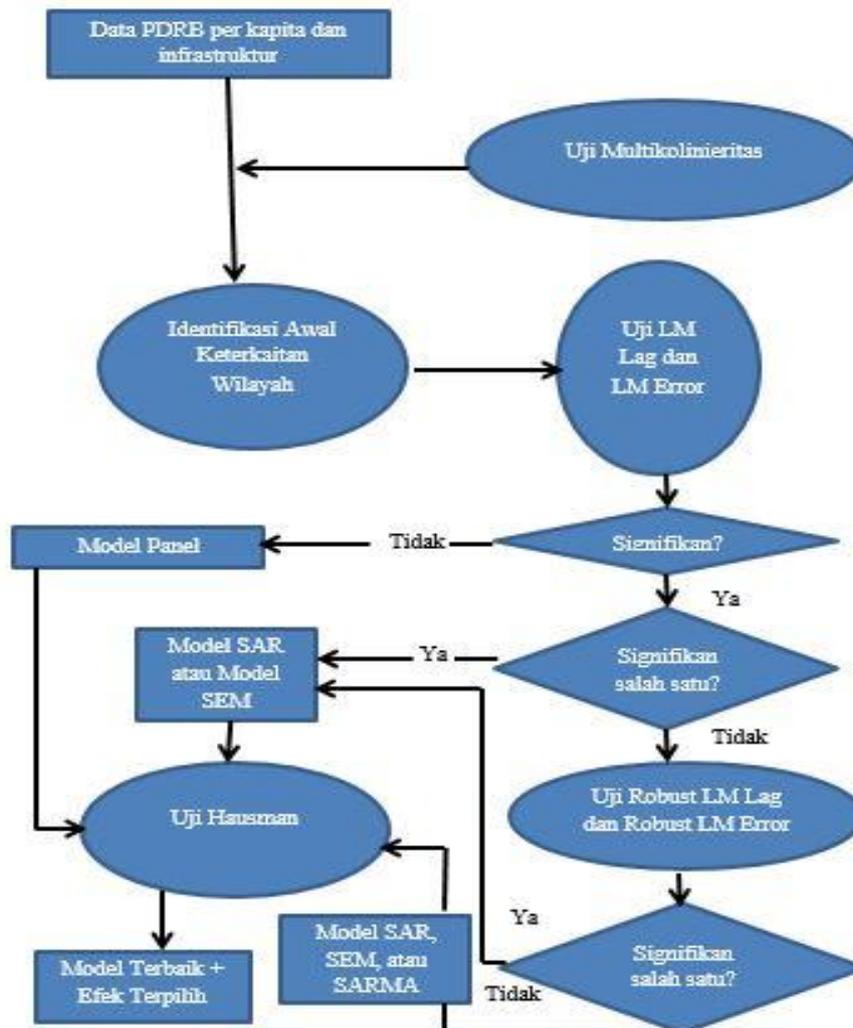
Statistik m mendekati distribusi χ_k^2 , dengan k adalah dimensi dari vektor β .

Langkah keempat melakukan pemilihan model terbaik dengan menggunakan koefisien determinasi (R^2) dan *Akaike Information Criterion*¹ (AIC). Model yang terbaik adalah model yang memiliki nilai R^2 terbesar. atau memiliki nilai AIC terkecil. Langkah selanjutnya adalah melakukan uji asumsi. Uji

¹ AIC merupakan ukuran kualitas relatif model statistik untuk suatu set data.

normalitas dilakukan mengingat *Maximum Likelihood Estimator* (MLE) mengasumsikan residual berdistribusi normal (Bain dan Engelhart [5]). Hipotesis dalam pengujian kenormalan menggunakan uji Kolmogorov-Smirnov satu arah (Daniel [6]). Hipotesis nol dalam pengujian ini adalah error berdistribusi normal. Selain itu, tidak adanya multikolinearitas antar variabel independen juga harus terpenuhi. Apabila nilai VIF lebih besar 10, maka diindikasikan adanya gejala multikolinearitas (Wooldridge [7]).

Penelitian ini menggunakan data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) berupa data PDRB perkapita (Y), panjang jalan per luas wilayah ($JALAN$), jumlah fasilitas kesehatan yang meliputi puskesmas dan rumah sakit per kapita ($FASKES$), jumlah fasilitas pendidikan meliputi SD/MI, SMP/MTs, dan SMA/SMK/MA per jumlah murid ($FASDIK$), dan jumlah listrik yang terjual per kapita ($LISTRİK$). Panjang jalan per luas wilayah dan jumlah listrik yang terjual perkapita merupakan infrastruktur ekonomi. Sedangkan jumlah fasilitas kesehatan yang meliputi puskesmas dan rumah sakit per kapita dan jumlah fasilitas pendidikan meliputi SD/MI, SMP/MTs, dan SMA/SMK/MA per jumlah murid merupakan infrastruktur sosial. Data yang digunakan merupakan data 33 provinsi dari tahun 2010 hingga 2014. Proses pembentukan model regresi panel spasial bisa dilihat di Gambar 1.



Gambar 1. Skematik diagram pemilihan model regresi panel spasial

Hasil dan Pembahasan

Langkah awal untuk melakukan pembentukan model regresi panel spasial adalah dengan melakukan uji asumsi non-multikolinearitas. Hasil pengujian non-multikolinearitas menunjukkan bahwa variabel yang dipakai telah memenuhi uji asumsi non-multikolinearitas. Oleh karena itu, pemodelan panel spasial dapat dilakukan ke tahap selanjutnya. Berdasarkan Tabel 1, hasil uji *LM* menunjukkan hasil yang signifikan, baik *LM error* maupun *LM lag*. Hasil uji *LM robust* menunjukkan bahwa model spasial yang dapat digunakan dalam menggambarkan pengaruh infrastruktur terhadap perekonomian adalah model SEM dan SARMA.

Uji Hausman dilakukan untuk menentukan efek yang terbaik bagi model terpilih, baik efek tetap maupun efek random. Berdasarkan hasil pengujian pada Tabel 2 diperoleh kesimpulan bahwa efek terbaik untuk model SEM adalah efek random, sedangkan efek terbaik untuk model SARMA adalah efek tetap.

• Tabel 1. Hasil uji *LM* dan uji *LM robust*

Tahapan Pemilihan	<i>p-value</i>	Kesimpulan
-------------------	----------------	------------

Model		
<i>LM error</i>	2,2e-16*	Model SEM dapat digunakan
<i>LM lag</i>	5,551e-16*	Model SAR dapat digunakan
<i>LM error robust</i>	0,000162*	Model SEM dapat digunakan
<i>LM lag robust</i>	0,1604	Model SAR tidak dapat digunakan
SARMA	2,2e-16*	Model SARMA dapat digunakan

Keterangan: * sig. 0,05.

• Tabel 2. Hasil uji Hausman

Model	<i>p-value</i>	Kesimpulan
SEM	0,222	<i>Random</i>
SARMA	0,0192*	<i>Fixed</i>

Keterangan: * sig. 0,05.

Untuk menentukan model terbaik, maka dilakukan uji pemeriksaan terhadap model dengan efek terpilih. Pemeriksaan model terbaik dilakukan dengan cara membandingkan nilai R^2 dan AIC. Model yang terbaik adalah model yang memiliki nilai R^2 terbesar dan nilai AIC terkecil.

• Tabel 3. Hasil pengujian *goodness of fit*

Kriteria	SEM-RE	SARMA-FE
(1)	(2)	(3)
R^2	0,3594409	0,9975809
AIC	-494,4524	-1412,982

Berdasarkan hasil pengujian *Goodness of Fit* pada Tabel 3, nilai R^2 pada model SARMA-*Fixed Effect* (SARMA-FE) yaitu sebesar 0,9975809 lebih besar dari SEM-*Random Effect* (SEM-RE) yang hanya sebesar 0,3594409. Hal ini menunjukkan bahwa model SARMA-FE merupakan model yang terbaik. Hasil pengujian nilai AIC juga menunjukkan hasil yang sama. Nilai AIC SARMA-FE sebesar -1412,982 lebih kecil dibandingkan nilai AIC SEM-RE yang hanya sebesar -494,4524. Oleh karena itu, model SARMA-FE merupakan model terbaik dalam penelitian ini.

Setelah mendapatkan model terbaik, maka dilakukan uji normalitas. Hasil pengujian normalitas menggunakan uji *Normal Probability Plot* dan Kolmogorov-Smirnov menunjukkan bahwa data berdistribusi normal. Oleh karena itu, model SARMA-FE merupakan model yang terbaik untuk memodelkan pengaruh infrastruktur sosial dan ekonomi terhadap perekonomian di Indonesia. Persamaan yang terbentuk adalah sebagai berikut:

$$\widehat{Y}_{it} = \mu_i + 0,37419W_{ij}\widehat{Y}_{jt}^* + 0,20594LISTRİK_{it}^* + 0,06297JALAN_{it} - 1546,2FASKES_{it}^* + 2,2846FASDIK_{it}^{**} - 0,41603W_{ij}\varepsilon_{jt}^* \quad (1)$$

Keterangan: * sig. 0,05; ** sig. 0,1.

Nilai *adjusted R squared* pada penelitian ini adalah 0,9975049. Hal ini menunjukkan bahwa variabel-variabel independen yang terdapat dalam model dapat menjelaskan variasi PDRB perkapita sebesar 99,75 persen. Selain itu, jika dilihat dari nilai p-value masing-masing variabel independen, variabel panjang jalan, listrik yang terjual, jumlah fasilitas kesehatan, dan jumlah fasilitas pendidikan signifikan pada tingkat kepercayaan 90 persen. Hal ini menunjukkan bahwa perekonomian bergantung pada panjang jalan, jumlah listrik yang terjual, jumlah fasilitas kesehatan, dan jumlah fasilitas pendidikan.

Berdasarkan persamaan yang terbentuk, diperoleh nilai μ_i untuk masing-masing provinsi yang berbeda (Lampiran 1). μ_i merupakan efek spesifik spasial yang terbentuk karena penggunaan *fixed effect*. Efek spesifik spasial merupakan gambaran heterogenitas setiap provinsi. Efek spesifik spasial mencerminkan adanya variabel lain yang dimiliki oleh suatu provinsi tetapi tidak dimiliki oleh provinsi lainnya. Dengan kata lain, setiap provinsi memiliki karakteristik masing-masing yang tercermin dalam variabel lain di luar model. Apabila diasumsikan variabel independen tidak berubah maka PDRB perkapita hanya dipengaruhi oleh efek spesifik spasial. Dengan demikian model *fixed effect* dapat menjelaskan adanya perbedaan perilaku antar provinsi di Indonesia.

Jumlah listrik yang terjual perkapita memiliki hubungan yang positif terhadap perekonomian di suatu provinsi. Dari persamaan terlihat bahwa setiap kenaikan rasio jumlah listrik yang terjual perkapita (Mwh perkapita) sebesar 1 satuan akan meningkatkan PDRB perkapita sebesar 20,59 juta rupiah dengan asumsi variabel lainnya konstan. Lee *et. al.* [8] dalam penelitiannya menyimpulkan bahwa kekurangan kapasitas listrik menjadi hambatan besar pada perkembangan perusahaan-perusahaan besar di Nigeria. Oleh karena itu, banyak perusahaan besar yang menyediakan listrik dengan cara memproduksi listrik secara mandiri. Hal tersebut menyebabkan pengeluaran perusahaan semakin besar sehingga keuntungan yang dihasilkan akan berkurang karena produktivitas yang tidak optimal. Sibarani [9] dalam penelitian di Indonesia menyebutkan bahwa listrik berpengaruh positif dan signifikan terhadap pertumbuhan ekonomi di Indonesia. Ketersediaan pasokan listrik yang cukup akan mampu meningkatkan pertumbuhan ekonomi di Indonesia.

Panjang jalan per luas wilayah memiliki hubungan yang positif terhadap perekonomian di suatu provinsi. Setiap kenaikan rasio panjang jalan per luas wilayah sebesar 1 satuan akan meningkatkan PDRB perkapita sebesar 6,3 juta rupiah dengan asumsi variabel lain bernilai konstan. Njoh [10] menunjukkan bahwa jalan memiliki hubungan yang positif terhadap pertumbuhan ekonomi. Selain itu, banyak sektor perekonomian lain yang sangat bergantung pada keberadaan transportasi, seperti industri, pertambangan, dan pertanian. Njoh merekomendasikan sarana dan prasarana transportasi harus diprioritaskan sebagai sasaran investasi.

Jumlah fasilitas pendidikan per jumlah murid bertanda positif dan signifikan. Hal ini menunjukkan setiap kenaikan rasio jumlah fasilitas pendidikan per jumlah murid sebesar 1 satuan akan meningkatkan PDRB perkapita sebesar 228,46 juta rupiah dengan asumsi variabel lain bernilai konstan. Sabates *et. al.* [11] menyebutkan bahwa keberadaan sekolah menjadi daya tarik bagi masyarakat untuk sekolah. Semakin sedikit jumlah sekolah di suatu wilayah akan meningkatkan angka putus sekolah, sedangkan semakin banyak jumlah sekolah akan meningkatkan jumlah masyarakat untuk bersekolah sehingga jumlah murid

bertambah. Ketersediaan fasilitas pendidikan berfungsi untuk menarik minat masyarakat untuk bersekolah dan menimba ilmu. Semakin tinggi ilmu dan pengetahuan yang dikuasai akan meningkatkan kualitas dan kemampuan seseorang hingga terbentuk tenaga kerja yang terampil dan berpengalaman. Tenaga kerja yang terampil tersebut akan meningkatkan produktivitas dan perekonomian negara.

Nilai variabel rasio jumlah fasilitas kesehatan perkapita dalam penelitian ini bernilai negatif. Setiap kenaikan rasio jumlah kesehatan perkapita akan menurunkan PDRB perkapita sebesar 154,62 milyar rupiah dengan asumsi variabel lain bernilai konstan. Hasil penelitian ini bertentangan dengan penelitian Bloom, Canning, dan Sevilla [12] yang menyebutkan bahwa kesehatan berpengaruh terhadap pertumbuhan ekonomi. Peningkatan dalam bidang kesehatan dapat meningkatkan output tidak hanya melalui produktivitas tenaga kerja, melainkan juga akumulasi modal. Faktor-faktor yang memengaruhi tetapi tidak dilibatkan dalam penelitian ini adalah jumlah dokter dan tenaga medis lainnya, jumlah kamar rawat inap, dan jumlah peralatan penunjang kesehatan. Tingginya rasio fasilitas kesehatan perkapita di Indonesia Timur tidak sebanding dengan rasio dokter per puskesmas. Menurut data dari Kemenkes RI, rasio dokter per puskesmas tertinggi terkonsentrasi di Indonesia Barat. Hal ini mengindikasikan adanya ketidakmerataan distribusi tenaga medis di Indonesia. Devarajan, Swaroop, dan Zou [13] menunjukkan bahwa pengeluaran yang biasanya dianggap produktif bisa menjadi tidak produktif jika terjadi pembiayaan yang berlebihan yang umumnya terjadi pada negara berkembang. Hal inilah yang menyebabkan variabel fasilitas kesehatan bertanda negatif.

Faktor spasial terbukti berpengaruh secara signifikan terhadap model yang terbentuk baik pada variabel dependen (λ) maupun *error* (ρ). Koefisien λ yang bernilai positif menunjukkan bahwa PDRB perkapita di suatu provinsi dipengaruhi oleh PDRB perkapita provinsi tetangga. Setiap kenaikan PDRB perkapita di provinsi-provinsi tetangga sebesar 1 satuan akan meningkatkan kenaikan PDRB perkapita di suatu provinsi sebesar 37,42 juta rupiah dengan asumsi variabel lain bernilai konstan. Hal ini sesuai dengan penelitian Ramirez dan Loboguerrero [14] yang menunjukkan bahwa pertumbuhan ekonomi di suatu negara dipengaruhi oleh pertumbuhan ekonomi negara tetangga.

Koefisien ρ merupakan pengaruh spasial *error*. Dalam penelitian ini, ρ bernilai negatif. Persamaan yang terbentuk menunjukkan bahwa setiap kenaikan nilai variabel lain diluar model dari provinsi tetangga sebesar 1 satuan akan menyebabkan penurunan PDRB perkapita di suatu provinsi sebesar 41.6 juta rupiah dengan asumsi variabel lainnya konstan. Resende *et. al.* [15] menyebutkan bahwa koefisien spasial bisa bernilai negatif yang disebabkan oleh variabel lain di luar model dan kesalahan menggunakan matriks penimbang. Penelitian Garretsen dan Peeters [16] juga memperoleh nilai ρ negatif. Faktor-faktor yang mungkin memberi pengaruh negatif dalam penelitian ini namun tidak dimasukkan dalam model adalah bencana alam dan faktor keamanan.

2. Kesimpulan

Faktor yang signifikan berpengaruh terhadap pertumbuhan ekonomi dalam tingkat signifikansi 10 persen adalah koefisien spasial lag, koefisien spasial error, panjang jalan, jumlah listrik yang terjual, jumlah fasilitas pendidikan, dan jumlah

fasilitas kesehatan. Koefisien spasial lag, panjang jalan, jumlah listrik yang terjual, dan jumlah fasilitas pendidikan signifikan dan bertanda positif. Hal ini menunjukkan bahwa ketiga variabel tersebut memberi efek positif terhadap perekonomian di Indonesia. Koefisien spasial error dan jumlah fasilitas kesehatan signifikan dan bertanda negatif. Hal ini menunjukkan bahwa kedua variabel tersebut memberi efek yang negatif bagi perekonomian.

Referensi

- [1] Anselin, L., 1988, *Spatial Econometrics: Methods and Models*, Kluwer Academic Publisher.
- [2] Elhorst, J. P., 2010, *Handbook of Spatial Analysis*, Springer.
- [3] BPS, 2011, *Analisis Dampak Spasial pada Peramalan Perekonomian dan Ketenagakerjaan Provinsi*, BPS.
- [4] Baltagi, B. H., 2005, *Econometric Analysis of Panel Data*, John Wiley & Sons. Ltd.
- [5] Bain, Lee J. dan Engelhart, Max, 1992, *Introduction to Probability and Mathematical Statistics*, PWS-KENT Publishing Company.
- [6] Daniel, W. W., 1990, *Applied Nonparametrics Statistics Second Edition*, Cengage Learning.
- [7] Wooldridge, J. M., 2013, *Introductory Econometrics a Modern Approach 5th Edition*, South-Western Cengage Learning.
- [8] Lee, Anas, Verma, dan Murray, 1996, *Why Manufacturing Firms Produce Some Electricity Internally*, The World Bank.
- [9] Sibarani, M. H., 2002, *Kontribusi Infrastruktur terhadap Pertumbuhan Ekonomi Indonesia (26 Provinsi di Indonesia Tahun 1983-1997) [Tesis]*, Universitas Indonesia.
- [10] Njoh, A. J., 2012, Impact of Transportation Infrastructure on Development in East Africa and the Indian Ocean Region, *Journal of Urban Planning and Development*, 1-9.
- [11] Sabates, Akyeampong, Westbrook, dan Hunt, 2010, School Drop out: Patterns, Causes, Changes and Policies, *Education for All Global Monitoring Report*.
- [12] Bloom, Canning, dan Sevilla, 2004, The Effects of Health on Economic Growth : A Production Function Approach, *World Development*, 1-13.
- [13] Devarajan, Swaroop, dan Zou, 1996, The Composition of Public Expenditure and Economic Growth, *Journal of Monetary Economics*, 313-344.
- [14] Ramirez, Maria Teresa dan Loboguerrero, Ana Maria, 2002, Spatial Dependence and Economic Growth: Evidence from a Panel of Countries, *JEL*.
- [15] Resende, Carvalho, Sakowski, dan Cravo, 2015, *Evaluating Multiple Spatial Dimension of Economic Growth in Brazil Using Spatial Panel Data Models (1970-2000)*, Instituto for Applied Economic Research - IPEA.
- [16] Garretsen, H dan Peeters. J., 2007, FDI and the Relevance of Spatial Linkages: Do Third Country Effects Matter for Dutch FDI? *DNB Working Paper No. 162*.

Lampiran

Lampiran 1. Efek spesifik spasial

Provinsi	μ_i	Provinsi	μ_i
Aceh	-0.15951	Nusa Tenggara Barat	0.016529
Sumatera Utara	-0.11072	Nusa Tenggara Timur	0.058387
Sumatera Barat	-0.10493	Kalimantan Barat	0.036597
Riau	-0.2882	Kalimantan Tengah	-0.1048
Jambi	-0.35984	Kalimantan Selatan	-0.07082
Sumatera Selatan	-0.02448	Kalimantan Timur	0.325513
Bengkulu	-0.00901	Sulawesi Utara	0.431632
Lampung	-0.21495	Sulawesi Tengah	0.435726
Kep. Babel	-0.24832	Sulawesi Selatan	-0.07229
Kepulauan Riau	-0.08356	Sulawesi Tenggara	-0.08952
DKI Jakarta	-0.07954	Gorontalo	-0.0767
Jawa Barat	-0.05168	Sulawesi Barat	0.070652
Jawa Tengah	-0.04633	Maluku	-0.03828
DI Yogyakarta	0.968154	Maluku Utara	-0.08527
Jawa Timur	-0.08358	Papua Barat	0.004829
Banten	0.271565	Papua	-0.02872
Bali	-0.18854		

•

• Lampiran 2. Uji LM

Lagrange multiplier diagnostics for spatial dependence

```
data:
model: lm(formula = fm, data = compile)
weights: Mat_panel_listw
```

LMerr = 77.95, df = 1, p-value < 2.2e-16

Lagrange multiplier diagnostics for spatial dependence

```
data:
model: lm(formula = fm, data = compile)
weights: Mat_panel_listw
```

LMLag = 65.693, df = 1, p-value = 5.551e-16

Lagrange multiplier diagnostics for spatial dependence

```
data:
model: lm(formula = fm, data = compile)
weights: Mat_panel_listw
```

RLMerr = 14.227, df = 1, p-value = 0.000162

Lagrange multiplier diagnostics for spatial dependence

```
data:
model: lm(formula = fm, data = compile)
weights: Mat_panel_listw
RLMLag = 1.9703, df = 1, p-value = 0.1604
```

Lagrange multiplier diagnostics for spatial dependence

```
data:
model: lm(formula = fm, data = compile)
weights: Mat_panel_listw

SARMA = 79.921, df = 2, p-value < 2.2e-16
```

Lampiran 3. Uji Hausman

```
> hausman (semfe, semre)
Hausman test for spatial models
```

```
data:
chisq = 5.7087, df = 4, p-value = 0.222
alternative hypothesis: one model is inconsistent
```

```
> hausman (sarmafe, sarmare)
Hausman test for spatial models
```

```
data:
chisq = 13.4865, df = 5, p-value = 0.0192
alternative hypothesis: one model is inconsistent
```

• Lampiran 41. Uji goodness of fit

Nilai R^2

```
> Rsqr2(semre$residuals, compile$y)
[1] 0.3594409
> Rsqr2(sarmafe$residuals, compile$y)
[1] 0.9975809
```

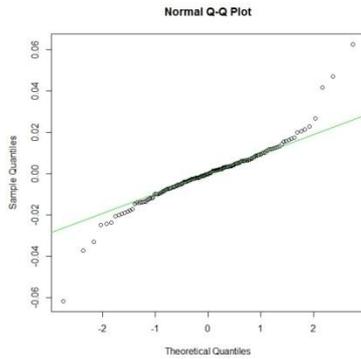
Nilai AIC

```
> godfit(semre)
[1] -494.4524
```

```
> godfit(sarmafe)
[1] -1412.982
```

• Lampiran 5. Uji asumsi

```
> vif(f)
LISTRIK JALAN FASDIK FASKES
9.376805 8.618449 1.815449 1.550836
```



One-sample Kolmogorov-Smirnov test
 data: sarmafe\$residuals
 D = 0.092333, p-value = 0.12
 alternative hypothesis: two-sided

•

• **Lampiran 6. Matriks penimbang terstandarisasi**

	1	1	1	1	1	1	1	1	1	2	3	3	3	3	3	3	5	5	5	6	6	6	7	7	7	7	7	7	8	8	9	9							
	1	2	3	4	5	6	7	8	9	1	1	2	3	4	5	6	1	2	3	1	2	3	4	1	2	3	4	5	6	1	2	1	4						
1	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0							
1	2	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0						
1	3	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0						
1	4	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0					
1	5	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0					
1	6	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
1	7	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
1	8	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
1	9	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
2	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
3	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
3	2	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
3	3	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
3	4	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
3	5	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
3	6	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
5	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
5	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0			
5	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0			
6	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0			
6	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0			
6	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0			
6	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0			
7	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0			
7	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	
7	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	0

74	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	0		
75	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0
76	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	
81	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0		
82	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0		
91	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0			
94	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0			

Sumber : Output R (diolah)

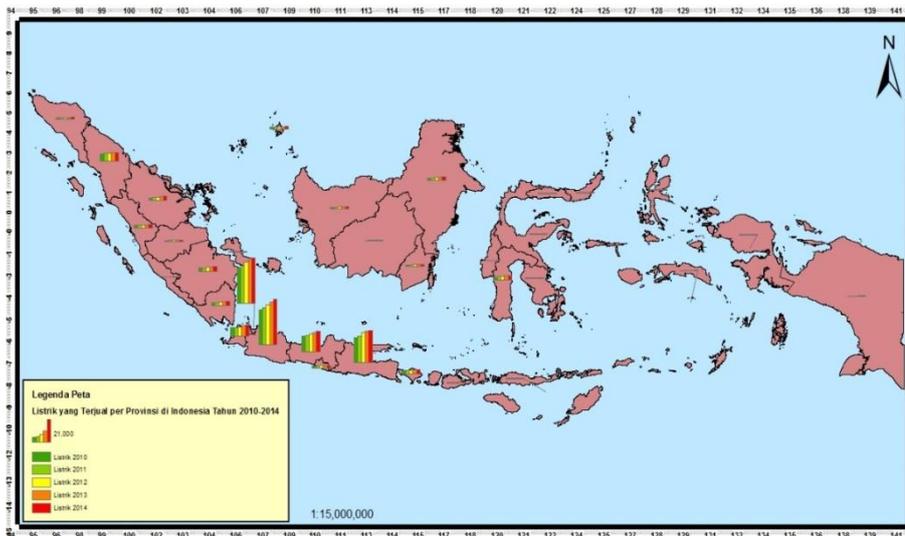
Lampiran 7. Tabel tiga tetangga terdekat

Provinsi	Tetangga 1	Tetangga 2	Tetangga 3
Aceh	Sumatera Utara	Sumatera Barat	Riau
Sumatera Utara	Aceh	Sumatera Barat	Riau
Sumatera Barat	Sumatera Utara	Riau	Jambi
Riau	Sumatera Utara	Sumatera Barat	Jambi
Jambi	Sumatera Barat	Sumatera Selatan	Bengkulu
Sumatera Selatan	Bengkulu	Lampung	Kep. Bangka Belitung
Bengkulu	Jambi	Sumatera Selatan	Lampung
Lampung	Sumatera Selatan	DKI Jakarta	Banten
Kep. Bangka Belitung	Jambi	Sumatera Selatan	Lampung
Kep. Riau	Kalimantan Barat	Kep. Bangka Belitung	Riau
DKI Jakarta	Lampung	Jawa Barat	Banten
Jawa Barat	DKI Jakarta	Jawa Tengah	Banten
Jawa Tengah	Jawa Barat	D.I. Yogyakarta	Jawa Timur
D.I. Yogyakarta	Jawa Barat	Jawa Tengah	Jawa Timur
Jawa Timur	Jawa Tengah	D.I. Yogyakarta	Bali
Banten	Lampung	DKI Jakarta	Jawa Barat
Bali	D.I. Yogyakarta	Jawa Timur	Nusa Tenggara Barat
Nusa Tenggara Barat	Jawa Timur	Bali	Sulawesi Selatan
Nusa Tenggara Timur	Nusa Tenggara Barat	Sulawesi Selatan	Sulawesi Tenggara
Kalimantan Barat	Kep. Riau	Kalimantan Tengah	Kalimantan Selatan
Kalimantan Tengah	Kalimantan Barat	Kalimantan Selatan	Kalimantan Timur
Kalimantan Selatan	Kalimantan Tengah	Kalimantan Timur	Sulawesi Barat
Kalimantan Timur	Kalimantan Tengah	Kalimantan Selatan	Sulawesi Barat
Sulawesi Utara	Sulawesi Tengah	Gorontalo	Maluku Utara
Sulawesi Tengah	Sulawesi Selatan	Gorontalo	Sulawesi Barat
Sulawesi Selatan	Sulawesi Tengah	Sulawesi	Sulawesi Barat

		Tenggara	
Sulawesi Tenggara	Sulawesi Tengah	Sulawesi Selatan	Sulawesi Barat
Gorontalo	Sulawesi Utara	Sulawesi Tengah	Sulawesi Barat
Sulawesi Barat	Sulawesi Tengah	Sulawesi Selatan	Sulawesi Tenggara
Maluku	Sulawesi Utara	Maluku Utara	Papua Barat
Maluku Utara	Sulawesi Utara	Gorontalo	Maluku
Papua Barat	Maluku	Maluku Utara	Papua
Papua	Maluku	Maluku Utara	Papua Barat

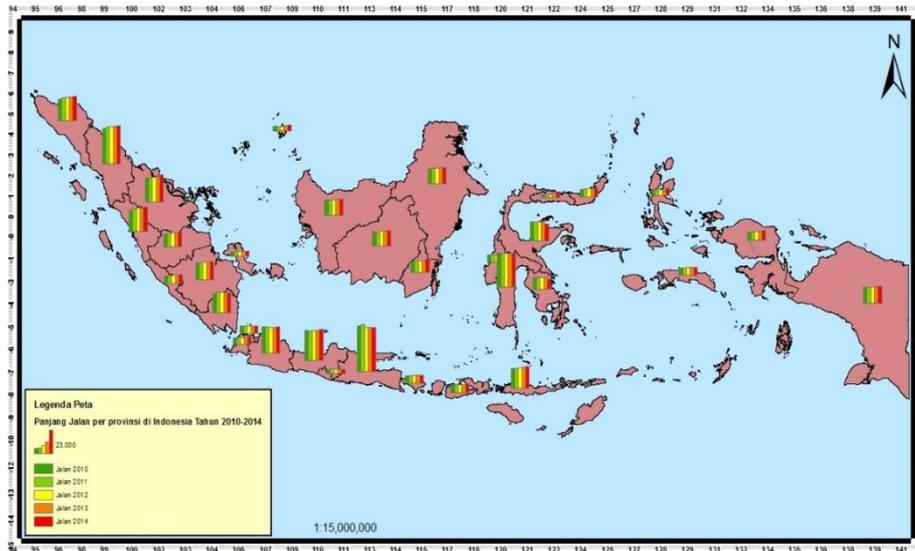
Sumber : Output R (diolah)

- **Lampiran 8. Peta jumlah listrik yang terjual per provinsi di Indonesia**
 - tahun 2010 - 2014



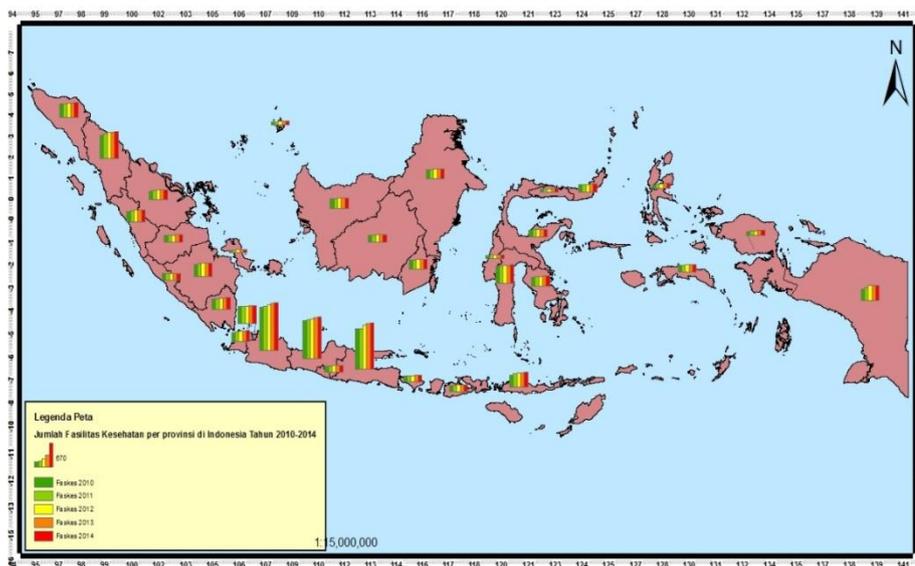
Sumber : PLN (diolah)

- **Lampiran 9. Peta panjang jalan per provinsi di Indonesia tahun 2010 - 2014**



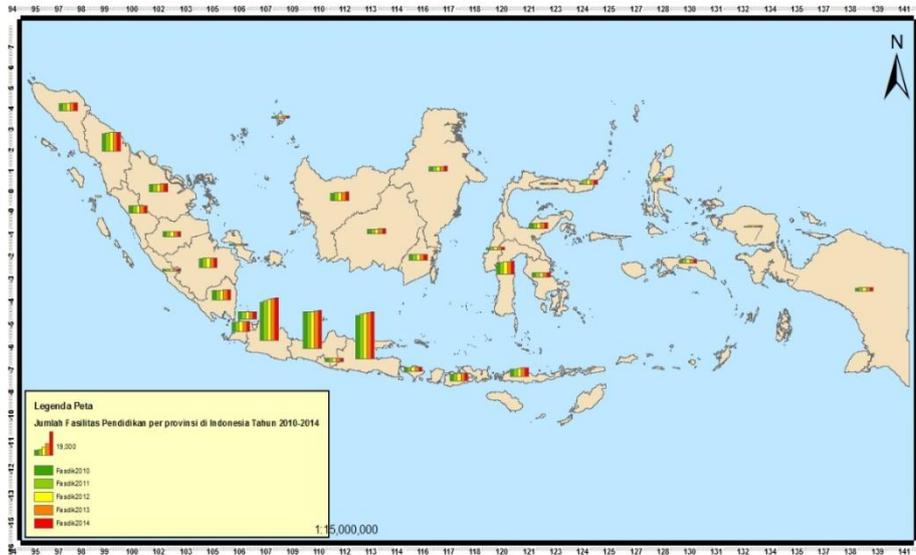
Sumber : BPS (diolah)

- **Lampiran 10. Peta jumlah fasilitas kesehatan per provinsi di Indonesia tahun 2010 - 2014**



Sumber : BPS (diolah)

- **Lampiran 11. Peta jumlah fasilitas pendidikan per provinsi di Indonesia tahun 2010 - 2014**



Sumber : BPS (diolah)

DETERMINAN KEMISKINAN DARI ASPEK HETEROGENITAS WILAYAH DENGAN GEOGRAPHICAL WEIGHTED REGRESSION (GWR)

SUGIARTO

Jurusan Statistika, Sekolah Tinggi Ilmu Statistik (STIS)- Jakarta, soegie@stis.ac.id

Abstrak. Penelitian mengenai determinan kemiskinan wilayah tidak ada habisnya, namun kebanyakan penelitian sebelumnya melihat kemiskinan wilayah dari aspek global, tidak melihat dari aspek lokal wilayah. Analisis *geographical weighted regression (GWR)* dapat menganalisis tidak hanya pengaruh variabel secara global namun juga pengaruh ke masing-masing lokal wilayah. Penelitian ini bertujuan untuk melihat determinan kemiskinan wilayah dari aspek heterogenitas wilayah/propinsi di Indonesia. Data yang digunakan dalam penelitian ini berasal dari berbagai publikasi BPS tahun 2015. Pengolahan dilakukan menggunakan metode *geographical weighted regression (GWR)*, yang menjadi variabel responnya persentase penduduk miskin dan variabel bebasnya adalah Indeks Pembangunan Manusia (IPM), PDRB perkapita, dan Luas Lantai Rumah Tangga. Hasil penelitian menunjukkan bahwa variabel yang mempengaruhi kemiskinan wilayah tiap propinsi di Indonesia berbeda-beda.

Kata kunci : Kemiskinan wilayah, *geographical weighted regression*.

1. Pendahuluan

Masalah kemiskinan adalah masalah yang dihadapi oleh hampir seluruh negara didunia, tidak hanya negara-negara miskin, atau berkembang saja tetapi juga dihadapi oleh negara maju. Menurut data BPS [1][2] selama kurun waktu hampir 10 tahun Indonesia dihadapkan oleh penduduk miskin yang cukup tinggi yaitu lebih dari 10 persen dari jumlah penduduk, walaupun secara persentase menunjukkan arah yang menurun. Namun bila dilihat dari koefisien gini, nilainya menunjukkan angka semakin besar. Koefisien gini menunjukkan ada tidaknya ketimpangan pendapatan antar penduduk (nilainya 0 s.d 1), semakin besar koefisien gini semakin besar ketimpangannya.

Tabel 1. Statistik Kemiskinan Indonesia tahun 2007-2015

	2007	2008	2009	2010	2011	2012	2013	2014	2015
Kemiskinan Relatif (% dari populasi)	16.6	15.4	14.2	13.3	12.5	11.7	11.5	11.0	11.1
Kemiskinan Absolut (dalam jutaan)	37	35	33	31	30	29	29	28	29
Koefisien Gini/ Rasio Gini	0.35	0.35	0.37	0.38	0.41	0.41	0.41	0.41	0.41

Sumber data: BPS

Berbagai program pengentasan kemiskinan telah dilaksanakan pemerintah, mulai dari awal reformasi sampai saat ini, namun begitu belum ada program yang efektif menurunkan kemiskinan di Indonesia secara signifikan. Padahal jumlah pendanaan penanggulangan kemiskinan setiap tahun tidak sedikit bahkan jumlahnya terus meningkat. Namun karena tersebar di berbagai departemen/instansi, menyebabkan pelaksanaannya seringkali keterkaitan antar program penanggulangan kemiskinan di pusat maupun di daerah belum optimal [3].

Menurut Tim Nasional Percepatan Penanggulangan Kemiskinan (TNP2K)[4], adanya variasi yang besar dalam kemiskinan antar daerah menjadi salah satu penyebab kurang suksesnya program pengentasan kemiskinan. Angka kemiskinan di Jawa dan Bali yang relatif maju bervariasi antara 3,8 hingga sekitar 17 persen. Sementara di wilayah Papua, angka kemiskinan mencapai 37 persen. Tapi karena populasi penduduk di Jawa dan Bali lebih besar, dimana 60 persen dari seluruh penduduk miskin Indonesia ada di daerah ini, dan hanya 3 persen yang tinggal di Papua, ini menyebabkan adanya 'dilema' dalam strategi kebijakan kemiskinan. Apakah kebijakan sebaiknya diarahkan pada daerah dengan tingkat kemiskinan tinggi di Kawasan Timur Indonesia, atau di wilayah dengan jumlah absolute penduduk miskin terbanyak (Jawa dan Bali).

Ada tiga ciri yang menonjol dari kemiskinan di Indonesia. Pertama, banyak rumah tangga yang berada di sekitar garis kemiskinan nasional. Kedua, ukuran kemiskinan didasarkan pada pendapatan, sehingga tidak menggambarkan batas kemiskinan yang sebenarnya. Ketiga, mengingat sangat luas dan beragamnya wilayah Indonesia, perbedaan antar daerah merupakan ciri mendasar dari kemiskinan di Indonesia, dimana masing-masing wilayah memiliki karakteristik dan penyebab yang berbeda-beda. Sehingga penanganan pengentasan kemiskinan di setiap wilayah juga disesuaikan dengan faktor yang dominan yang menjadi penyebab di wilayah tersebut.

Penelitian terhadap determinan kemiskinan suatu wilayah telah banyak dilakukan namun sifatnya masih global. Padahal seperti telah dijelaskan sebelumnya, kebijakan pengentasan kemiskinan sebaiknya dilakukan berdasarkan faktor dominan yang menjadi penyebab kemiskinan di wilayah tersebut. Analisis *Geographically Weighted Regression* (GWR) merupakan metode analisis yang memfasilitasi fenomena yang menjadi penyebab dominan terhadap kemiskinan di suatu wilayah. Model GWR merupakan pengembangan dari analisis regresi yang dikenal dengan regresi global. Namun, berbeda dengan regresi global yang nilai parameter modelnya konstan, parameter model GWR berbeda-beda pada setiap lokasi penelitian. Parameter pada model GWR dihitung pada setiap lokasi pengamatan, sehingga setiap lokasi pengamatan mempunyai nilai parameter regresi yang berbeda-beda. Penelitian ini bertujuan untuk melihat determinan kemiskinan dari aspek heterogenitas wilayah sehingga nantinya dapat diperoleh informasi tentang faktor-faktor yang dominan mempengaruhi kemiskinan di suatu wilayah. Dari penelitian-penelitian sebelumnya, banyak variabel yang bisa menjadi determinan kemiskinan suatu wilayah, namun dalam penelitian ini peneliti hanya membatasi pada tiga variabel yang mempengaruhi kemiskinan yaitu dari nilai Indeks Pembangunan Manusia (IPM) (yang mencerminkan kualitas manusia),

PDRB perkapita (yang mencerminkan pendapatan per kapita) dan luas lantai rumah tangga (yang mencerminkan salah satu indikator rumah tidak layak huni).

2. Hasil-Hasil Utama

2.1. Sumber Data dan Variabel Penelitian

Penelitian ini menggunakan informasi yang ada dalam publikasi BPS yaitu buku Statistik Indonesia [1,2]. Selain itu beberapa data diambil dari *website* BPS. Informasi yang dikumpulkan dalam penelitian ini mengacu pada kejadian tahun 2015. Variabel-variabel yang digunakan dalam penelitian ini diambil dari kajian teori dari penelitian sebelumnya. Adapun variabel yang digunakan antara lain:

KEMISKINAN (Y) = Persentase penduduk miskin di provinsi-i
 IPM (X₁) = Indeks pembangunan manusia provinsi ke-i
 PDRBPERKPT(X₂) = PDRB perkapita provinsi ke-i
 LUASLANTAI (X₃) = Persentase Rumah Tangga dengan Luas Lantai $\leq 7,2m^2$ provinsi ke-i

Karena tiap variabel berbeda satuannya, peneliti melakukan standarisasi pada masing-masing variable, sehingga pengaruh satuannya hilang. Adapun proses pengolahan data pada penelitian ini menggunakan software Geoda dan GWR4.

2.2. Metode Analisis

Dalam penelitian ini untuk mencari determinan kemiskinan peneliti menggunakan metode analisis GWR. Seperti dijelaskan pada latar belakang, analisis GWR merupakan pengembangan dari analisis regresi. Namun, berbeda dengan regresi global yang nilai parameter modelnya konstan, parameter model GWR berbeda-beda pada setiap lokasi penelitian. Sehingga akan terdapat sebanyak $n(p+1)$ parameter yang diestimasi pada model GWR, dimana n adalah jumlah lokasi pengamatan dan $(p+1)$ adalah jumlah parameter. Pada model GWR, masing-masing lokasi pengamatan bergeoreferensi, yaitu memiliki koordinat spasial. Koordinat spasial pada lokasi pengamatan ke-i dilambangkan dengan (u_i, v_i) . Adapun model GWR dapat dituliskan sebagai berikut [5]:

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^p \beta_k(u_i, v_i) x_{ik} + \varepsilon_i; i = 1, 2, \dots, n \quad (1)$$

y_i : Nilai observasi variabel respon pada pengamatan ke-i
 $\beta_0(u_i, v_i)$: Nilai konstanta/*intercept* pada pengamatan ke-i
 $\beta_k(u_i, v_i)$: Koefisien regresi variabel prediktor ke-k pada pengamatan ke-i
 x_{ik} : Nilai observasi variabel prediktor ke-k pada pengamatan ke-i
 ε_i : *Error* ke-i yang diasumsikan identik, independen, dan berdistribusi normal dengan rata-rata nol dan varian konstan σ^2
 u_i, v_i : Menyatakan titik koordinat (*longitude, latitude*) lokasi pengamatan ke-i

Berdasarkan Fotheringham [5], ada dua asumsi yang digunakan dalam model GWR:

- Error term* adalah independen, identik, dan mengikuti distribusi normal dengan rata-rata nol dan varians konstan [$\varepsilon_i \sim \text{IIDN}(0, \sigma^2)$]

b. $E(\hat{y}_i) = E(y_i)$ untuk semua i

Dalam melakukan analisis GWR peneliti melakukan beberapa tahapan sebagai berikut:

1. Melakukan pemilihan peubah penjelas dan peubah respon yang digunakan dalam analisis berdasarkan studi literatur sebelumnya dan menduga parameter model regresi klasik menggunakan analisis regresi klasik berganda.
2. Melakukan uji asumsi klasik pada model regresi yaitu asumsi kenormalan menggunakan Uji *Jarque-Bera*, asumsi multikolinieritas dengan melihat nilai *Variance Inflation Factor* (VIF), asumsi autokolerasi dengan Uji *Durbin-Watson*, serta uji kehomogenan data dengan Uji *Breusch-Pagan*. Jika asumsi klasik terpenuhi maka regresi global yang terbaik, sedangkan jika asumsi tidak dapat dipenuhi, maka dapat dilanjutkan ke langkah berikutnya.
3. Menguji efek keragaman spasial dengan menggunakan uji *Breusch-Pagan* dengan hipotesis $H_0 : \sigma^2(u_i, v_i) = \dots = \sigma^2(u_n, v_n) = \sigma^2$ (Data homogen) dan H_1 : minimal ada satu $\sigma^2(u_i, v_i) \neq \sigma^2(u_j, v_j)$ untuk $i \neq j$, dengan $i, j = 1, 2, \dots, n$ atau (Data Heterogen). Statistik uji Breusch-Pagan sebagai berikut:

$$BP = \frac{1}{2} h' Z (Z' Z)^{-1} Z' h \text{ dimana } h = \left(\frac{e_i^2}{\sigma^2} - 1 \right) \quad (2)$$

Z adalah vektor amatan peubah respon y yang berukuran $(n \times 1)$ dan sudah dibakukan untuk setiap pengamatan dengan $i = 1, 2, \dots, n$, e_i^2 adalah kuadrat galat untuk pengamatan ke- i dan σ^2 merupakan ragam dari e_i . Jika H_0 gagal tolak maka digunakan model regresi, jika H_0 ditolak digunakan model GWR dan lanjut ke langkah berikutnya

4. Menentukan nilai lebar jendela optimum dengan melihat Cross Validation (CV) yang minimum. Formula yang digunakan untuk mengitung CV adalah sebagai berikut:

$$CV = \sum_{i=1}^n [y_i - \hat{y}_{\neq i}(b)]^2 \quad (3)$$

Dimana $\hat{y}_{\neq i}(b)$ merupakan nilai prediksi y_i dengan pengamatan ke- i dikeluarkan dari proses estimasi.

5. Menghitung matriks jarak dan matriks pembobot $W(u_i, v_i)$ dengan fungsi kernel normal. Adapun formulanya sebagai berikut:

$$W(u_i, v_i) = \exp \left(1 - \left(\frac{d_{ij}^2}{h_i(q)} \right)^2 \right) \quad (4)$$

Dengan $h_i(q)$ adalah lebar jendela adaptif yang menetapkan q sebagai jarak tetangga terdekat dari wilayah i .

6. Menguji kebaikan model dengan melakukan uji F.

$$Fh = \frac{(RSS_{OLS} - RSS_{GWR})/V}{RSS_{GWR}/\delta 1} = \frac{GWR_{IMP}/V}{RSS_{GWR}/\delta 1} \quad (5)$$

Dimana $v = \text{tr}(R_0 - R_1)$ dan $\delta 1 = \text{tr}(R_1)$.

Hipotesis yang diuji adalah H_0 : model GWR dan OLS sama dan H_1 : model GWR lebih baik dibanding model OLS. Jika hasil uji F tidak signifikan maka tidak terdapat perbedaan antara model GWR, sedangkan jika H_0 ditolak maka lanjut ke langkah berikutnya.

7. Melihat kebaikan model dengan melihat nilai koefisien determinasi (R^2), jumlah kuadrat galat (RSS), BIC dan AIC dari model regresi global dan GWR.
8. Jika model GWR lebih baik, maka dilanjutkan dengan pengujian pendugaan parameter secara parsial di setiap Provinsi di Indonesia.

Bentuk hipotesisnya adalah sebagai berikut:

$$\begin{aligned}
 H_0 : \beta_k (u_i, v_i) &= 0 \\
 H_0 : \beta_k (u_i, v_i) &\neq 0 \\
 \text{dengan } i &= 1, 2, \dots, n ; k = 1, 2, \dots, p \\
 T_{hit} &= \frac{\hat{\beta}(u_i, v_i)}{\sigma \sqrt{C_{kk}}} \quad (6)
 \end{aligned}$$

$\hat{\beta}(u_i, v_i)$ adalah estimasi parameter (koefisien GWR) sedangkan C_{kk} adalah elemen diagonal ke-k dari matriks pembobotnya. Hipotesis nol akan ditolak jika $|T_{hit}| > t_{\alpha/2, df}$ dimana $df = \frac{\delta_1^2}{\delta_2^2}$.

2.3 Hasil dan Pembahasan

2.3.1 Hasil Penelitian

Sesuai dengan tahapan penelitian, pada tahap awal dilakukan pemodelan regresi klasik berganda berdasarkan variabel-variabel yang diperoleh dari kajian teori penelitian sebelumnya. Dalam penelitian ini pengolahan data menggunakan SPSS 19 dan diperoleh hasil sebagai berikut:

Tabel 2. Nilai Koefisien dan *p-value* Regresi Global

Variabel	Koefisien	t-stat	p-value
Intersept	0.00	0.00	1.00
IPM	-0.4388404	-2.464772	0.01966
PDRBPERKPT	-0.1934261	-1.158954	0.25562
LUASLANTAIRT	0.3509695	2.205165	0.03525
F-Statistik	= 11.3075		
P-Value (F-Stat)	= 0.000		
R-Square	= 0.5307		

Berdasarkan output diatas diperoleh hasil uji simultan (Uji F-Stat) model global yang dibentuk sudah cukup baik (dengan $p\text{-value } 0,0000 < 0,10$). Adapun model regresi global yang terbentuk sebagai berikut:

$$\text{KEMISKINAN} = -\beta_1 \text{IPM} - \beta_2 \text{PDRBPERKPT} + \beta_3 \text{LUASLANTAIRT} + \varepsilon_i$$

Selanjutnya dilakukan uji asumsi klasik yaitu memenuhi asumsi kenormalan, tidak ada multikolinearitas, tidak ada autokorelasi dan harus homogen. Dari output diperoleh hasil sebagai berikut:

Tabel 3. *Test on Normality of Error*

TEST	df	VALUE	P-VALUE	KEPUTUSAN
Jarque-Bera	2	0,7166	0,69886	Data Berdistribusi Normal
<i>MULTICOLLINEARITY CONDITION NUMBER 2.4639</i>				

Tabel 4. Nilai VIF Variabel Prediktor

Variabel	VIF	KEPUTUSAN
IPM	2.026	

PDRBPERKPT	1.781	Nilai VIF < 10 berarti tidak ada multikolinearitas
LUASLANTAIRT	1.619	
(Keputusan Data tidak ada multikolinearitas)		
Test Durbin-Watson = 2.034 (Keputusan Data tidak ada autokolerasi)		

Tabel 5. *Diagnostics For Heteroskedasticity*

TEST	df	VALUE	P-VALUE	KEPUTUSAN
Breusch-Pagan	3	10,3537	0,01579	HETEROGEN

Berdasarkan uji asumsi klasik terhadap *error* diperoleh hasil bahwa ada satu asumsi yang terlanggar yaitu *error* tidak homogen (bersifat heterogen). Sehingga penggunaan model GWR dapat dilanjutkan. Langkah selanjutnya adalah mencari lebar jendela optimum dengan melihat nilai CV yang terkecil. Dalam penelitian ini dipakai software GWR4 untuk mencari nilai CV yang minimum. Fungsi Kernel yang digunakan adalah metode *adaptif Gaussian*. Alasan dipakainya metode ini, karena memberikan *bandwith* yang optimum tidak terlalu besar dan tidak terlalu kecil selain itu metode *adaptif gaussian* merupakan fungsi kernel yang baik karena memberikan nilai pembobot mengikuti kurva normal. Dari metode ini diperoleh nilai CV dan *bandwidth* terbaik sebagai berikut:

Tabel 6. Nilai CV dan *Bandwidth* Terbaik

Indikator	Adaptive Gaussian
Best Bandwidth Size	19
Minimum CV	0.576

Berdasarkan output diatas diperoleh *bandwidth* terbaik dengan metode penimbang kernel *Adaptive Gaussian* adalah 19. Artinya, terdapat 19 tetangga (provinsi) terdekat yang signifikan mempengaruhi suatu provinsi. Hasil matriks jarak dan matriks pembobot $W(u_i, v_i)$ sebagai berikut:

Tabel 7. Nilai Matriks pembobot Variabel Prediktor

Variabel	Min	Q1	Med	Q3	Max	OLS
(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>Intercept</i>	-0.230	-0.175	-0.117	-0.047	0.003	0.000
IPM	-0.488	-0.372	-0.209	-0.137	-0.112	-0.439
PDRBPERKPT	-0.328	-0.315	-0.288	-0.195	-0.129	-0.193
LUASLANTAIRT	0.113	0.160	0.236	0.355	0.383	0.351

Model GWR menghasilkan estimasi parameter yang berbeda-beda untuk setiap lokasi. Sedangkan model OLS menghasilkan nilai estimasi parameter yang sama untuk semua lokasi penelitian yaitu yang ditunjukkan oleh kolom (7). Sementara kolom (2) sampai (6) merupakan statistik deskriptif dari estimasi koefisien model GWR yang meliputi nilai minimum, kuartil pertama, median, kuartil ketiga, dan nilai maksimum. Model GWR memungkinkan suatu variabel penjelas yang sama memiliki hubungan dengan variabel respon yang positif di suatu lokasi dan nilai negatif di lokasi lainnya.

Langkah berikutnya adalah menguji model GWR yang dibentuk. Untuk melihat apakah model GWR lebih baik dibanding model regresi global, maka dilakukan *goodness of fit test* Pengujian ini bertujuan untuk mengidentifikasi model GWR yang terbentuk mampu secara signifikan menjelaskan hubungan variabel respon

dan variabel penjelasnya.

Hipotesis:

H_0 : model GWR dan OLS sama

H_1 : model GWR lebih baik dibanding model OLS

Tabel 8. Nilai ANOVA Model GWR

<i>Source</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F-hit</i>	<i>p-value</i>
Global Residual	15.487	30.000			
GWR Improvement	3.320	3.105	1.069		
GWR Residual	12.168	26.895	0.452	2.363	0.000

Nilai *p-value* pada Tabel 8. menunjukkan angka yang sangat signifikan pada tingkat signifikansi 10 persen dan dapat diambil kesimpulan bahwa model GWR secara statistik lebih baik dibandingkan model regresi global dalam menjelaskan kemiskinan dan faktor-faktor yang mempengaruhinya. Keunggulan dari model GWR dibandingkan dengan model RLB adalah model GWR mampu melakukan prediksi pada data yang mengandung heterogenitas spasial, jika nilai *difference of criterion* yang dihasilkan negatif. Sebaliknya, nilai positif pada *difference of criterion* menunjukkan bahwa variabel bebas tersebut bersifat global. Secara lengkap, nilai *difference of criterion* dari masing-masing variabel bebas disajikan pada tabel berikut:

Tabel 9. Nilai *DiFF of Criterion* Variabel

Variabel	F	DOF for F test		DIFF of Criterion
Intercept	4,034486	0,437	28,042	-0,019232
IPM	5,250727	0,476	28,042	-0,027507
PDRBPERKPT	1,431272	0,503	28,042	0,027380
LUASLANTAIRT	2,705807	0,492	28,042	-0,033366

Terlihat bahwa variabel IPM dan luas lantai rumah tangga memiliki nilai *difference of criterion* yang negatif. Dengan demikian, variabel bebas tersebut secara signifikan memiliki heterogenitas spasial atau bersifat lokal pada masing-masing provinsi di Indonesia.

Setelah didapatkan bahwa model GWR adalah model yang lebih baik daripada model regresi global namun untuk lebih memperkuat kita bandingkan beberapa indikator yang dapat digunakan untuk membandingkan model GWR dan model regresi global diantaranya nilai R^2 , AIC, *residual sum of square*, BIC/MDL.

Tabel 10. Nilai R-Square, AIC, RSS, dan BIC Model GWR dan Global

Pembanding	GWR	Regresi Global
R-Square	0.6313	0.5307
AIC	75.4668	79.7527
RSS	12.1677	15.4875
BIC/MDL	86.0876	87.3845

Dari beberapa ukuran indikator terlihat bahwa model GWR lebih baik dari model regresi global. Selanjutnya masing-masing variabel akan di uji secara parsial. Dari

pengujian parsial dengan menggunakan tingkat signifikansi 10% diperoleh variabel yang signifikan dimasing-masing provinsi sebagai berikut:

Tabel 11. Variabel independen yang signifikan memengaruhi kemiskinan provinsi di Indonesia

Kelompok	Provinsi	Variabel yang signifikan
(1)	(2)	(3)
1	DI Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Kepulauan Bangka Belitung, Kepulauan Riau, DKI Jakarta, Jawa Barat, Jawa Tengah, DI Yogyakarta, Jawa Timur, Banten, Bali, Kalimantan Barat	PDRB perkapita
2	Nusa Tenggara Barat	PDRB perkapita dan luas lantai rumah tangga
3	Nusa Tenggara Timur, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, Papua	IPM dan luas lantai rumah tangga
4	Kalimantan Tengah, Kalimantan Selatan	Tidak ada variabel yang signifikan

2.3.2 Pembahasan

Berdasarkan Tabel 11, terbentuk 4 kelompok provinsi yang memiliki kesamaan variabel yang signifikan berpengaruh terhadap kemiskinan di Indonesia Tahun 2015. Ada 2 kelompok besar yaitu kelompok 1 dan kelompok 3. Dimana kelompok 1 adalah provinsi yang terletak dikawasan barat Indonesia sedangkan kelompok 3 adalah provinsi yang sebagian besar terletak dikawasan timur Indonesia. Provinsi-provinsi yang berada kelompok 1, kemiskinan wilayahnya lebih dipengaruhi oleh variabel PDRB perkapita yang mencerminkan pendapatan perkapita suatu daerah. Sementara itu provinsi-provinsi yang berada kelompok 3, kemiskinan wilayahnya lebih dipengaruhi oleh variabel IPM dan luas lantai rumah tangga dalam hal ini adalah persentase rumah tangga yang memiliki luas lantai $\leq 7,2m^2$. Sedangkan Provinsi Nusa Tenggara Barat lebih dipengaruhi oleh variabel PDRB perkapita dan luas lantai rumah tangga. Dua provinsi yaitu Provinsi Kalimantan Tengah dan Kalimantan Selatan tidak dipengaruhi oleh variabel dependen. Dua pengelompokan besar yaitu 1 dan 3 hampir sama dengan yang disampaikan oleh TNP2K pada latar belakang sebelumnya.

Dari hasil uji parsial didapat 34 persamaan yang berbeda untuk masing-masing provinsi. Untuk koefisien model masing-masing provinsi bisa dilihat pada Tabel 12. berikut ini:

Tabel 12. Koefisien Model GWR

No	Provinsi	β_0	β_1	β_2	β_3	LocalR ²
1	Aceh	-0.0820	-0.2502	-0.2934	0.2729	0.4687
2	Sumatera Utara	-0.1107	-0.2048	-0.3104	0.2424	0.4343
3	Sumatera Barat	-0.1440	-0.1596	-0.3259	0.1980	0.3991

4	Riau	-0.1540	-0.1486	-0.3265	0.1933	0.3903
5	Jambi	-0.1950	-0.1123	-0.3275	0.1477	0.3545
6	Sumatera Selatan	-0.1733	-0.1436	-0.3186	0.1659	0.3827
7	Bengkulu	-0.1719	-0.1338	-0.3282	0.1608	0.3750
8	Lampung	-0.1803	-0.1382	-0.3189	0.1478	0.3774
9	Bangka Belitung	-0.1871	-0.1351	-0.3146	0.1578	0.3733
10	Kepulauan Riau	-0.1840	-0.1257	-0.3224	0.1719	0.3686
11	DKI Jakarta	-0.2075	-0.1185	-0.3132	0.1125	0.3543
12	Jawa Barat	-0.1962	-0.1300	-0.3127	0.1210	0.3684
13	Jawa Tengah	-0.1718	-0.1586	-0.3080	0.1510	0.4041
14	DI Yogyakarta	-0.1563	-0.1713	-0.3102	0.1653	0.4200
15	Jawa Timur	-0.1240	-0.2110	-0.2996	0.2127	0.4631
16	Banten	-0.1954	-0.1267	-0.3165	0.1271	0.3646
17	Bali	-0.0775	-0.2680	-0.2832	0.2684	0.5190
18	Nusa Tenggara Barat	-0.0670	-0.2850	-0.2745	0.2847	0.5330
19	Nusa Tenggara Timur	-0.0023	-0.4248	-0.1842	0.3623	0.6321
20	Kalimantan Barat	-0.2303	-0.1145	-0.2973	0.1391	0.3422
21	Kalimantan Tengah	-0.1614	-0.1915	-0.2802	0.2152	0.4335
22	Kalimantan Selatan	-0.1422	-0.2085	-0.2800	0.2303	0.4549
23	Kalimantan Timur	-0.0897	-0.3066	-0.2305	0.3237	0.5258
24	Kalimantan Utara	-0.1042	-0.2971	-0.2292	0.3186	0.5096
25	Sulawesi Utara	-0.0283	-0.4506	-0.1431	0.3789	0.6331
26	Sulawesi Tengah	-0.0638	-0.3587	-0.1983	0.3683	0.5660
27	Sulawesi Selatan	-0.0504	-0.3420	-0.2246	0.3477	0.5675
28	Sulawesi Tenggara	-0.0283	-0.4110	-0.1735	0.3833	0.6106
29	Gorontalo	-0.0363	-0.4295	-0.1561	0.3816	0.6164
30	Sulawesi Barat	-0.0798	-0.3058	-0.2291	0.3486	0.5376
31	Maluku	-0.0066	-0.4730	-0.1349	0.3698	0.6573
32	Maluku Utara	-0.0176	-0.4696	-0.1323	0.3740	0.6502
33	Papua Barat	-0.0022	-0.4866	-0.1286	0.3596	0.6698
34	Papua	0.0035	-0.4882	-0.1323	0.3539	0.6744

Sebagai gambaran bentuk salah satu model persamaan untuk Provinsi Aceh sebagai berikut:

$$\text{KEMISKINAN}_{\text{ACEH}} = -0,0820 - 0,2502\text{IPM}_{\text{ACEH}} - 0,2934\text{PENDAPATAN}_{\text{ACEH}} + 0,2729\text{LUASLANTAIRT}_{\text{ACEH}}$$

dengan $R^2 = 0.4687$ * Signifikan pada $\alpha = 10$ persen

Dari model persamaan Provinsi Aceh diatas, variabel IPM berpengaruh negatif terhadap kemiskinan di Provinsi Aceh namun tidak signifikan dengan koefisien sebesar 0.2502, artinya bila IPM meningkat satu satuan maka kemiskinan akan menurun sebesar 0.2502 satuan. Variabel PDRB perkapita berpengaruh negatif terhadap kemiskinan di Provinsi Aceh dan signifikan dengan koefisien sebesar 0.2934 artinya bila pendapatan meningkat satu satuan maka kemiskinan akan menurun sebesar 0.2934 satuan. Sedangkan variabel luas lantai rumah tangga berpengaruh positif terhadap kemiskinan di Provinsi Aceh namun tidak signifikan dengan koefisien sebesar 0.2729, artinya bila semakin banyak rumah tangga dengan luas lantai $\leq 7,2\text{m}^2$ maka kemiskinan akan meningkat sebesar 0.2729 satuan. Nilai R^2 sebesar 0.4687 mempunyai arti bahwa ketiga variabel yaitu IPM, PDRB perkapita, dan luas lantai rumah tangga mempengaruhi kemiskinan di Provinsi Aceh sebesar 46,87 persen sedangkan 53,13 persennya dipengaruhi oleh faktor lain. Untuk koefisien masing-masing provinsi lainnya bisa dilihat pada lampiran.

3. Kesimpulan

Berdasarkan hasil dan pembahasan diatas dapat disimpulkan bahwa variabel dependen memiliki pengaruh yang berbeda-beda pada setiap wilayah, tergantung dari seberapa dominannya variabel tersebut mempengaruhi provinsi tersebut. Ada 2 kelompok besar yaitu kelompok 1 dan kelompok 3. Dimana kelompok 1 adalah provinsi yang sebagian besar terletak dikawasan barat Indonesia dimana variabel yang signifikan mempengaruhi adalah PDRB perkapita sedangkan kelompok 3 adalah provinsi yang sebagian besar terletak dikawasan timur Indonesia dimana variabel yang signifikan mempengaruhi adalah IPM dan luas lantai rumah tangga. Berdasarkan determinan kemiskinan wilayah yang berbeda-beda tersebut seharusnya pemerintah dalam pengentasan kemiskinan bisa lebih fokus menjalankan program-program yang sesuai dengan variabel yang paling mempengaruhi di provinsi tersebut. Sehingga program pengentasan kemiskinan bisa lebih efektif.

Referensi

- [1] Badan Pusat Statistik, Statistik Indonesia, Jakarta, 2015
- [2] Badan Pusat Statistik, Statistik Indonesia, Jakarta, 2016
- [3] Royat, Sujana, *Prosiding Seminar Nasional Meningkatkan Peran Sektor Pertanian dalam Penanggulangan Kemiskinan*, Pusat Analisis Sosial Ekonomi dan Kebijakan Pertanian Departemen Pertanian, 2008
- [4] Tim Nasional Percepatan Penanggulangan Kemiskinan (TNP2K), *Penanggulangan Kemiskinan: Situasi Terkini, Target Pemerintah, dan Program Percepatan*, 2010
- [5] Fotheringham, A.S., Brunson, C., & Charlton, M. *Geographically Weighted Regression*, John Wiley & Sons, Chichester, UK, 2002.
- [6] Anselin, L., *Exploratory Spatial Data Analysis and Geographic Information Systems*, National Center for Geographic Information and Analysis of California Santa Barbara, 1993.

METODE RIDGE REGRESSION UNTUK MENGATASI MULTIKOLINIERITAS PADA VARIABEL-VARIABEL YANG MEMENGARUHI PENDAPATAN ASLI DAERAH KABUPATEN KAPUAS HULU

KHOTIMAH¹, WINIH BUDIARTI²

¹Badan Pusat Statistik Kabupaten Pontianak, khotimah@bps.go.id

²Sekolah Tinggi Ilmu Statistik, winih@stis.ac.id

Abstrak. Analisis regresi linear berganda merupakan salah satu teknik analisis data statistika yang sering digunakan untuk mengkaji hubungan atau pengaruh variabel-variabel bebas terhadap variabel tak bebas. Penelitian ini bertujuan untuk mengetahui pengaruh variabel jumlah penduduk, pengeluaran pemerintah, dan PDRB sektor unggulan terhadap variabel tak bebas Pendapatan Asli Daerah (PAD) Kabupaten Kapuas Hulu. Pendapatan Asli Daerah (PAD) merupakan tolok ukur yang penting untuk menentukan tingkat kemampuan daerah dalam melaksanakan otonomi daerah secara nyata dan bertanggungjawab. Untuk memperoleh variabel sektor unggulan dari sisi Pendapatan Domestik regional Bruto (PDRB) digunakan metode *Location Quotient* yang menghasilkan sektor pertanian dan sektor konstruksi sebagai sektor unggulan dari Kabupaten Kapuas Hulu. Hasil analisis regresi linear berganda diperoleh bahwa semua variabel bebas mempengaruhi variabel tak bebas namun terdapat pelanggaran asumsi nonmultikolinieritas. Menurut Neter [11] untuk mengatasi terjadinya multikolinieritas digunakan metode *ridge regression*. *Ridge regression* terbukti mampu mengatasi multikolinieritas yang terjadi ditandai dengan nilai *Variance Inflation Factor* (VIF) tidak lebih dari 10. Berdasarkan hasil *ridge regression* variabel yang memengaruhi PAD Kabupaten Kapuas Hulu adalah jumlah penduduk, pengeluaran pemerintah, PDRB sektor pertanian, dan PDRB sektor konstruksi.

Kata kunci : *ridge regression*, multikolinieritas, pendapatan asli daerah, *location quotient*, regresi linear berganda

1. Pendahuluan

Kabupaten Kapuas Hulu merupakan kabupaten dengan luas wilayah terbesar kedua di Kalimantan Barat setelah Kabupaten Ketapang dengan luas wilayah sebesar 20.33 persen terhadap luas wilayah Kalimantan Barat. Potensi Sumber Daya Alam (SDA) yang melimpah baik yang berada didalam tanah berupa bahan tambang, diatas tanah berupa kayu didalam hutan tropis yang memiliki nilai sangat tinggi. Akan tetapi, potensi dan besarnya wilayah tersebut, belum menjamin

keberhasilan pembangunan ekonominya. Hal tersebut terlihat dari pertumbuhan ekonominya dimana pertumbuhan ekonomi Kabupaten Kapuas Hulu yaitu 4,56 persen pada tahun 2011, 5,61 persen pada tahun 2012, dan 6,01 persen pada tahun 2013 masih di bawah Kabupaten Sintang yang memiliki wilayah terbesar ketiga. Jika dibandingkan dengan kabupaten Sambas yang sama-sama berbatasan dengan Malaysia, pertumbuhan ekonomi Kabupaten Kapuas Hulu masih di bawahnya. Bahkan dibandingkan dengan daerah pemekaran seperti Kabupaten Kayong Utara dan Kabupaten Kubu Raya, pertumbuhan ekonomi Kabupaten Kapuas Hulu masih lebih rendah.

Kabupaten Kapuas Hulu merupakan salah satu daerah yang diberikan hak otonomi daerah. Menurut UU No. 32 Tahun 2004, Otonomi daerah adalah hak, wewenang, dan kewajiban daerah otonom untuk mengatur dan mengurus sendiri urusan pemerintahan dan kepentingan masyarakat setempat sesuai dengan peraturan perundang-undangan. Otonomi daerah memberikan hak kepada daerah otonom untuk mengatur daerahnya sendiri, sehingga mereka lebih leluasa untuk meningkatkan pelayanan kepada masyarakat, selain itu otonomi daerah juga mampu mempermudah pemerintah khususnya pemerintah daerah otonom untuk mengerti kebutuhan masyarakatnya.

Diberikannya hak otonomi daerah diharapkan pemerintah daerah Kabupaten Kapuas Hulu mampu menggali sumber-sumber keuangan lokal, khususnya melalui Pendapatan Asli Daerah (PAD). Pendapatan Asli Daerah (PAD) merupakan semua penerimaan daerah yang berasal dari sumber ekonomi asli daerah. Jika PAD meningkat maka dana yang dimiliki oleh pemerintah daerah akan lebih besar. Hal tersebut akan meningkatkan kemandirian daerah, sehingga pemerintah daerah akan berinisiatif untuk lebih menggali potensi-potensi daerah yang dapat meningkatkan pertumbuhan ekonomi [14]. PAD Kabupaten Kapuas Hulu tergolong rendah dibandingkan kabupaten/kota lainnya di Provinsi Kalimantan Barat.

Ada beberapa hal yang perlu diperhatikan untuk meningkatkan PAD, yaitu: (1) Sektor unggulan yang berada pada daerah tersebut. Dengan memperhatikan sektor yang unggul, kita dapat mengetahui potensi yang ada di daerah tersebut yang mempunyai daya saing dalam beberapa periode tahun terakhir dan kemungkinan prospek sektor ekonomi di masa mendatang; (2) Pengeluaran pemerintah yang digunakan untuk administrasi pembangunan dan sebagian lain untuk kegiatan pembangunan di berbagai jenis infrastruktur yang penting. Perbelanjaan perbelanjaan tersebut akan meningkatkan pengeluaran agregat dan mempertinggi tingkat kegiatan ekonomi; (3) Jumlah penduduk yang dapat memengaruhi suatu perekonomian di suatu daerah. Jika penduduk di suatu daerah banyak menandakan banyaknya SDM yang terdapat di daerah tersebut sehingga dapat menyerap tenaga kerja secara produktif.

Berdasarkan uraian di atas, peneliti ingin mengkaji bagaimana pengaruh sektor unggulan, pengeluaran pemerintah, dan jumlah penduduk terhadap PAD Kabupaten Kapuas Hulu. Waktu yang digunakan pada penelitian ini tahun 1991 sampai dengan tahun 2013. Peneliti ingin menganalisis PAD yang berada di Kabupaten Kapuas Hulu saat terjadinya krisis ekonomi tahun 1998, kasus antara suku Dayak dan Madura tahun 1999, kenaikan BBM tahun 2004 dan krisis ekonomi global tahun 2008.

Untuk melihat pengaruh beberapa variabel independen terhadap variabel dependen digunakan analisis regresi berganda. Analisis regresi berganda merupakan salah satu teknik analisis data statistika yang sering kali digunakan

untuk mengkaji hubungan antara beberapa variabel dan meramal suatu variabel. Dalam penelitian ini yang ingin dikaji adalah hubungan atau pengaruh dua atau lebih variabel bebas terhadap variabel tidak bebas, maka model regresi yang sesuai adalah model regresi linear berganda (*multiple linear regression model*). Untuk mendapatkan model regresi linear berganda dapat diperoleh dengan melakukan estimasi terhadap parameter-parameternya menggunakan metode tertentu. Adapun metode yang dapat digunakan untuk mengestimasi parameter model regresi linear sederhana maupun regresi linear berganda adalah dengan metode kuadrat terkecil atau *ordinary least square* dan metode kemungkinan maksimum atau *maximum likelihood estimation* [11]. Suatu model regresi dengan metode kuadrat terkecil (OLS) dapat dikatakan sebagai model regresi terbaik apabila memenuhi asumsi-asumsi linear. Dalam Supranto [15], beberapa asumsi tersebut adalah: a. Normalitas yaitu $\epsilon_i \sim N(0, \sigma^2)$; b. Homoskedastisitas yaitu $\text{Var}(\epsilon_i) = E(\epsilon_i^2) = \sigma^2$; c. Nonautokorelasi $\text{cov}(\epsilon_i, \epsilon_j) = 0$; $i \neq j$; dan d. Nonmultikolinearitas. Asumsi-asumsi ini digunakan agar persamaan regresi layak digunakan. Oleh karena itu, sebelum melakukan analisis terhadap hasil output regresi linear berganda, terlebih dahulu dilakukan pengujian terhadap asumsi-asumsi tersebut. Apabila terjadi pelanggaran asumsi maka diperlukan perlakuan khusus atau metode-metode tertentu agar model yang dihasilkan dapat digunakan.

Apabila ketiga asumsi pertama terpenuhi namun asumsi ke-empat terlanggar yaitu asumsi nonmultikolinieritas maka akan diterapkan metode Regresi Ridge untuk menstabilkan nilai koefisien regresi karena adanya multikolinieritas. Multikolinieritas merupakan adanya hubungan linear sempurna ataupun tidak sempurna antar variabel bebas. Multikolinieritas ini dapat mengakibatkan Variansi besar (dari taksiran OLS), Interval kepercayaan lebar, dan Uji-t (t rasio) tidak signifikan [12]. Regresi Ridge merupakan metode estimasi koefisien regresi yang diperoleh melalui penambahan konstanta bias c pada diagonal $X^T X$. Sehingga diperoleh persamaan regresi linier yang baru dan tidak mengandung multikolinieritas.

Adapun tujuan yang ingin dicapai dalam penelitian ini adalah untuk melihat pengaruh jumlah penduduk, pengeluaran pemerintah, dan sektor unggulan terhadap PAD Kabupaten Kapuas Hulu tahun 1991-2013 dengan menggunakan Regresi Ridge untuk mengatasi masalah Multikolinieritas dalam variabel-variabel independen sehingga diperoleh persamaan regresi linier berganda yang lebih baik.

2. Hasil – Hasil Utama

Metodologi

Ruang lingkup lokasi penelitian ini dilakukan di Kabupaten Kapuas Hulu yang merupakan sebuah kabupaten yang terletak paling ujung bagian barat Provinsi Kalimantan Barat dan berbatasan langsung dengan Serawak (Malaysia Timur). Referensi waktu objek penelitian yang digunakan adalah selama dua puluh tiga tahun, yaitu dari tahun 1991 sampai dengan tahun 2013. Pengambilan referensi waktu selama 23 (dua puluh tiga) tahun ini dilandasi oleh keterbatasan data Pendapatan Asli Daerah dan pengeluaran pemerintah di Kabupaten Kapuas Hulu.

Data yang digunakan dalam penelitian ini adalah data sekunder yang meliputi: a). Data Produk Domestik Regional Bruto (PDRB) tahunan Kabupaten Kapuas Hulu tahun 1991-2013 berdasarkan atas dasar harga konstan tahun 1983,

tahun 1993, dan tahun 2000. Kemudian data PDRB Provinsi Kalimantan Barat atas dasar harga konstan tahun 2000. Semua data PDRB berdasarkan atas dasar harga konstan tahun 1983, tahun 1993, dan tahun 2000 telah disamakan tahun dasarnya berdasarkan harga konstan tahun 2000. b). Data Pendapatan Asli Daerah (PAD) dengan periode tahun 1991 sampai dengan tahun 2013 yang bersumber dari publikasi Dirjen Perimbangan dan Keuangan dan Badan Pusat Statistik. c). Data pengeluaran pemerintah dengan periode tahun 1991 sampai dengan tahun 2013 yang bersumber dari yang bersumber dari publikasi Dirjen Perimbangan dan Keuangan. d). Data jumlah penduduk periode tahun 1991 sampai dengan tahun 2013 yang bersumber dari yang bersumber dari publikasi Badan Pusat Statistik.

Beberapa definisi yang ada dalam penelitian ini dijelaskan sebagai berikut:

1. Pendapatan Asli Daerah adalah pendapatan yang diperoleh daerah yang dipungut berdasarkan peraturan daerah sesuai dengan peraturan perundang-undangan.
2. Penduduk adalah orang yang biasa (sehari-hari) tinggal di wilayah itu.
3. Pengeluaran pemerintah adalah seluruh pembelian atau pembayaran barang dan jasa untuk kepentingan nasional, seperti pembelian persenjataan dan alat-alat kantor pemerintah, pembangunan jalan dan bendungan, gaji pegawai negeri, angkatan bersenjata, dan lainnya.
4. PDRB adalah jumlah nilai tambah bruto (*gross value added*) yang timbul dari seluruh sektor perekonomian di suatu wilayah. Nilai tambah adalah nilai yang ditambahkan dari kombinasi faktor produksi dan bahan baku dalam proses produksi.
5. Sektor unggulan adalah sektor yang keberadaannya pada saat ini telah berperan besar kepada perkembangan perekonomian suatu wilayah, karena mempunyai keunggulan-keunggulan. Untuk mengidentifikasi sektor ekonomi potensial atau sektor unggulan dari sisi PDRB dengan melalui alat analisis *Location Quotient* [5].

Untuk menjawab tujuan penelitian ini digunakan analisis regresi linear berganda. Metode yang digunakan adalah *Ordinary Least Square* (OLS) dengan tingkat kepercayaan 95 persen. Pemilihan model ini dilakukan oleh peneliti karena tujuan penelitian hanya ingin melihat pengaruh dari variabel bebas terhadap variabel terikat. Dalam penelitian ini analisis regresi linear berganda digunakan untuk melihat pengaruh sektor unggulan, jumlah penduduk, dan pengeluaran pemerintah terhadap Pendapatan Asli Daerah (PAD) Kabupaten Kapuas Hulu dari tahun 1991 sampai dengan tahun 2013.

Pembentukan Model Regresi

Analisis regresi linier berganda yang akan dilakukan dalam penelitian ini menggunakan model sebagai berikut:

$$PAD = \beta_0 + \beta_1 JP + \beta_2 PP + \beta_{3k} PDRBSk + \varepsilon_i$$

dengan

PAD	= Pendapatan Asli Daerah
JP	= jumlah penduduk
PP	= pengeluaran pemerintah
PDRBSk	= PDRB sektor unggulan k
k	= sektor ekonomi

Model di atas digunakan untuk mengetahui pengaruh jumlah penduduk, pengeluaran pemerintah, dan PDRB sektor unggulan k terhadap pendapatan asli daerah di Kabupaten Kapuas Hulu.

Asumsi Klasik

a). Uji kenormalan

$H_0 : \varepsilon_i \sim N(0, \sigma^2)$, $i = 1, 2, \dots, 23$ (residual berdistribusi normal, rata-rata 0 dan varian σ^2)

$H_1 : \varepsilon_i \not\sim N(0, \sigma^2)$, $i = 1, 2, \dots, 23$ (residual tidak berdistribusi normal)

Statistik uji:

$$JB = n \left[\frac{S^2}{6} + \frac{(K - 3)^2}{24} \right]$$

Keterangan:

n : jumlah observasi

S : *skewness* (kemencengan)

K : *kurtosis* (keruncingan)

Apabila nilai *Jarque-Berra* nya lebih kecil dari nilai *Chi-Square* (χ^2) dengan derajat bebas dua maka gagal tolak H_0 dan asumsi normal terpenuhi.

b). Uji Heteroskedastisitas

Pengujian asumsi homoskedastisitas pada penelitian ini dilakukan dengan menggunakan uji *White* tanpa *cross term*. Adapun hipotesis yang digunakan adalah:

$H_0 = \text{Var}(\varepsilon_i) = \sigma^2$, untuk setiap t (homoskedastisitas)

$H_1 = \text{Var}(\varepsilon_i) \neq \sigma^2$, untuk setiap t (heteroskedastisitas)

Kriteria yang digunakan dalam uji *White* yaitu:

$Obs * R^2 > \chi^2$ dengan derajat bebas 4 atau probabilitas $Obs * R^2 < 0.05$, maka tolak H_0 yang artinya ada heteroskedastisitas, tetapi apabila probabilitas $Obs * R^2 > 0.05$, maka gagal tolak H_0 yang artinya tidak adanya heteroskedastisitas.

c). Uji Autokorelasi

$H_0: \text{cov}(e_i, e_j) = 0$ (tidak adanya autokorelasi antar error)

$H_1: \text{cov}(e_i, e_j) \neq 0$ (adanya autokorelasi)

Statistik uji:

$$d = \frac{\sum_{t=2}^N (e_t - e_{t-1})^2}{\sum_{t=1}^N e_t^2} \quad t = 1, 2, \dots, 23$$

Bandungkan nilai d yang dihitung dengan nilai dL dan dU dari tabel *Durbin Watson* dengan aturan sebagai berikut:

- Bila $d < dL$ ada korelasi positif
- Bila $dL \leq d \leq dU$ kita tidak dapat mengambil kesimpulan apa-apa
- Bila $dL < d < 4 - dU$ tidak ada korelasi
- Bila $4 - dU \leq d \leq 3 - dL$ kita tidak dapat mengambil kesimpulan apa-apa
- Bila $d > 4 - dL$ ada korelasi negatif

d). Uji Multikolinearitas

Adapun hipotesis pada uji multikolinearitas ini adalah:

$H_0: \rho_{ij} = 0$ (tidak terdapat korelasi antara variabel bebas i dan j)

$H_1: \rho_{ij} \neq 0$ (terdapat korelasi antara variabel bebas i dan j)

Hipotesis nol akan ditolak jika nilai *Variance Inflation Factor* (VIF) lebih besar dari 10 berarti variabel bebas mengalami multikolinearitas [16]. VIF merupakan elemen diagonal utama dari invers matriks korelasi. Apabila nilai VIF atau Faktor variansi inflasi kecil maka multikolinieritas lebih sederhana, tetapi jika Faktor inflasi melebihi 10 maka dikatakan terdapat multikolinieritas.

Pengujian Parameter Model

a). Uji Keseluruhan Parameter (*Overall F-test*)

$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ (tidak ada variabel bebas yang berpengaruh secara signifikan terhadap variabel terikat secara bersamaan)

H_1 : minimal ada $\beta_i \neq 0$, dengan $i = 1, 2, 3, 4$ (minimal ada 1 variabel bebas yang berpengaruh secara signifikan terhadap variabel terikat)

Statistik uji: Uji F

$$F = \frac{MSR}{MSE}$$

Keputusan: tolak H_0 jika F hitung $> F_{0.05}(4,18)$

b). Uji Individual Parameter (t-test)

$H_0: \beta_k = 0$; $k = 1, 2, 3, 4$ (variabel bebas tidak berpengaruh signifikan terhadap variabel terikat)

$H_1: \beta_k \neq 0$; $k = 1, 2, 3, 4$ (variabel bebas mempunyai pengaruh signifikan terhadap variabel terikat)

Uji-t didefinisikan sebagai berikut:

$$t = \frac{b_k}{s(b_k)}$$

Keputusan: tolak H_0 jika $|t| > t_{\alpha/2}$ atau p-value $< \alpha$

Ridge Regression

Regresi Ridge dapat digunakan untuk mengatasi korelasi yang tinggi antara beberapa variabel bebas [11]. Multikolinieritas yang terdapat dalam regresi linier berganda yang mengakibatkan matriks $X^T X$ -nya hampir singular yang pada gilirannya menghasilkan nilai estimasi parameter yang tidak stabil. Regresi Ridge merupakan metode estimasi koefisien regresi yang diperoleh melalui penambahan konstanta bias c pada diagonal $X^T X$. Nilai c untuk koefisien regresi ridge diantara 0 hingga 1.

Dalam bentuknya yang sederhana adalah sebagai berikut: $\beta(c) = (X^T X + cI)^{-1} X^T Y$. Umumnya sifat dari penafsiran ridge ini memiliki variansi yang minimum sehingga diperoleh nilai VIF-nya yang merupakan diagonal utama dari matriks $(X^T X + cI)^{-1} X^T X (X^T X + cI)^{-1}$. Pada dasarnya Regresi Ridge merupakan metode kuadrat terkecil. Perbedaannya adalah bahwa pada metode regresi ridge, nilai variabel bebasnya ditransformasikan dahulu melalui prosedur centering dan rescaling.

Ridge Regression merupakan salah satu metode untuk mengatasi multikolinieritas [11]. *Ridge Regression* merupakan modifikasi dari metode kuadrat terkecil (OLS) dengan memasukkan estimator bias pada koefisien regresi. *Ridge Regression* lebih baik digunakan karena peluang untuk mendekati nilai parameter yang sesungguhnya lebih besar.

Dalam OLS secara umum penghitungan estimator b diberikan dengan rumus:

$$(X'X) = X'Y$$

Sehingga menghasilkan persamaan regresi:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi} + \varepsilon_i$$

Ketika semua variabel ditransformasikan dengan transformasi korelasi maka diperoleh persamaan dalam bentuk matriks sebagai berikut:

$$r_{xx} = \begin{bmatrix} 1 & r_{12} & \cdots & r_{1,p-1} \\ r_{21} & 1 & \cdots & r_{2,p-1} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p-1,1} & r_{p-1,2} & \cdots & 1 \end{bmatrix} \quad r_{yx} = \begin{bmatrix} r_{y1} \\ r_{y2} \\ \vdots \\ r_{y,p-1} \end{bmatrix}$$

Dimana: r_{xx} = matriks korelasi antar variabel X
 r_{xy} = matriks korelasi antar variabel X dan variabel Y sehingga dari transformasi tersebut dihasilkan persamaan regresi *Standardized Regression Models* sebagai berikut:

$$Y_i^* = \beta_1^* X_{1i}^* + \beta_2^* X_{2i}^* + \cdots + \beta_k^* X_{ki}^* + \varepsilon_i^*$$

Dimana nilai $Y_i^* = \frac{1}{\sqrt{n-1}} \left(\frac{Y_i - \bar{Y}}{s_Y} \right)$ dan $X_{ki}^* = \frac{1}{\sqrt{n-1}} \left(\frac{X_{ki} - \bar{X}_k}{s_{X_k}} \right)$

Nilai estimator *Ridge Regression* yang distandarkan diperoleh dengan memasukkan sebuah nilai konstan bias (c) yang nilainya dalam interval $0 \leq c \leq 1$ ke dalam persamaan OLS (Neter, 1989). Nilai estimator *Standardized Ridge Regression* dapat dihitung dengan formula:

$$b^R = (r_{xx} + cI)^{-1} r_{xy}$$

Dimana:

I = matriks Identitas

c = konstan bias

b_R = estimator *Standardized Ridge Regression*

Sehingga setelah ditransformasikan terbentuklah persamaan regresi baru sebagai berikut:

$$\hat{Y}_i' = b_1^R X_1 + b_2^R X_2 + \cdots + b_{p-1}^R X_{p-1}$$

Nilai c yang dimasukkan ke dalam persamaan OLS akan menghasilkan estimator yang bias namun cenderung menghasilkan nilai koefisien yang lebih stabil dibandingkan dengan nilai koefisien regresi dalam OLS. Semakin besar nilai c yang dimasukkan maka akan menghasilkan komponen bias dari total MSE yang semakin mengecil dan nilai estimator *Ridge Regression* b_R yang semakin stabil. Hal ini menunjukkan bahwa nilai c pada estimator *Ridge Regression* b_R memiliki total MSE yang lebih kecil dibandingkan dengan nilai MSE pada OLS estimator b. Namun, dalam *Ridge Regression* terdapat kendala yaitu bahwa nilai optimum dari variasi nilai kecil c terkecil dimana koefisien *Ridge Regression* sudah stabil dan nilai *Variance Inflation Factor* (VIF) mempunyai nilai yang cukup kecil dengan nilai koefisien Determinasi (R^2) yang relatif besar. Nilai VIF adalah elemen diagonal dari matriks yang dapat diperoleh dengan formula sebagai berikut:

$$(r_{xx} + cI)^{-1} r_{xx} (r_{xx} + cI)^{-1}$$

Adapun koefisien determinasi (R^2) dalam *Ridge Regression* adalah:

$$R_R^2 = 1 - SSE_R \quad \text{dimana} \quad SSE_R = \sum (Y_i' - \hat{Y}_i')^2$$

Tahap selanjutnya adalah mentransformasi kembali model *Ridge Regression* yang diperoleh menjadi model regresi semula. Nilai estimator diperoleh dengan formula:

$$b_i = \left(\frac{S_Y}{S_X} \right) b_i^R$$

Dimana:

bi = koefisien regresi biasa

SY = standar deviasi variabel Y

SX = standar deviasi variabel X

Sedangkan nilai b0 diperoleh dari formula:

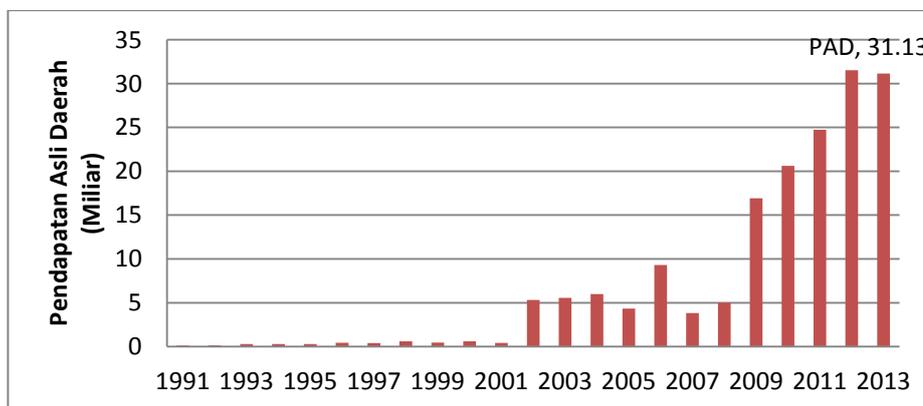
$$b_0 = \bar{Y} - b_1\bar{X}_1 - b_2\bar{X}_2 - \dots - b_{p-1}\bar{X}_{p-1}$$

Sehingga akan terbentuk persamaan *Ridge Regression*:

$$\hat{Y}'_i = b_0 + b_1X_1 + b_2X_2 + \dots + b_{p-1}X_{p-1} + e$$

Hasil dan Pembahasan

Pendapatan Asli Daerah Kabupaten Kapuas Hulu tahun 1991-2013 secara keseluruhan mengalami peningkatan seperti terlihat pada Gambar 1. Pada tahun 1997 PAD Kabupaten Kapuas Hulu menjadi Rp 398 juta dari 434 juta di tahun 1996. Di tahun 2001 PAD Kabupaten Kapuas Hulu mengalami penurunan lagi menjadi Rp 419 juta dari Rp 606 juta di tahun 2000. Hal ini disebabkan adanya krisis ekonomi yang terjadi di Indonesia. Kemudian pada 2005 mengalami penurunan kembali menjadi Rp 4.3 Miliar dari Rp 6 miliar. Ini terjadi karena adanya kenaikan BBM yang berdampak pada PAD. Pada tahun 2006 sempat mengalami kenaikan menjadi Rp 9 miliar. Kemudian di tahun 2007 PAD Kabupaten Kapuas Hulu mengalami penurunan kembali. Pada tahun 2008 hingga tahun 2012 PAD Kabupaten Kapuas Hulu meningkat menjadi Rp 3,15 Miliar di tahun 2012, akan tetapi mengalami penurunan kembali tahun 2013 sebesar Rp 3,11 Miliar.

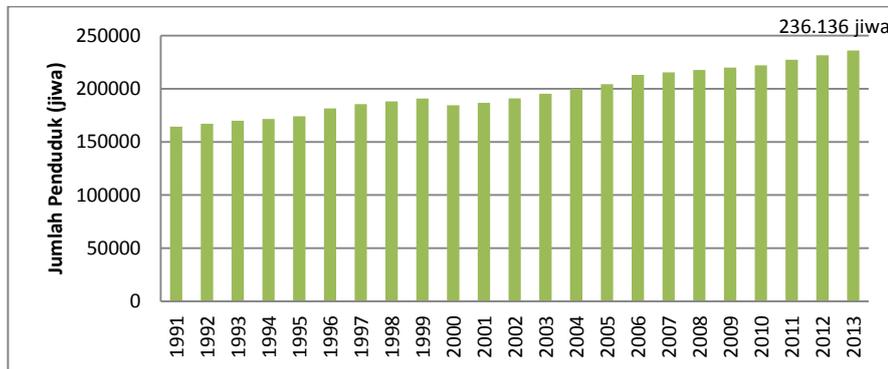


Sumber : BPS dan Dinas pendapatan, Pengelolaan Keuangan dan Aset Daerah

Gambar 1. Pendapatan Asli Daerah Kabupaten Kapuas Hulu Tahun 1991-2013

Demikian pula dengan jumlah penduduk Kabupaten Kapuas Hulu pada umumnya mengalami perkembangan (Gambar 2). Pada tahun 1991 jumlah penduduk Kabupaten Kapuas Hulu sebanyak 164.301 jiwa dan pada tahun 2013 meningkat sebanyak 236.136 jiwa. Namun pada tahun 2000, Penduduk Kabupaten Kapuas Hulu berkurang hingga 6.314 jiwa. Hal ini disebabkan terjadinya konflik antara suku Dayak dan Madura pada tahun 1999 yang berdampak pada tahun 2000.

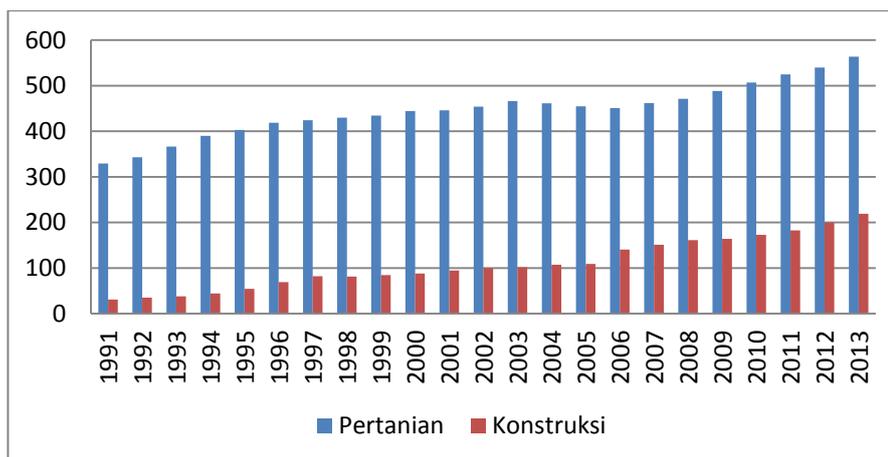
Banyak penduduk yang pindah dari daerah Kabupaten Kapuas Hulu. Kemudian sejak tahun 2001 hingga tahun 2013 penduduk Kabupaten Kapuas Hulu semakin bertambah hingga 236.136 jiwa.



Sumber : BPS Kabupaten Kapuas Hulu Tahun 1991-2013 (diolah)

Gambar 2. Jumlah Penduduk Kabupaten Kapuas Hulu Tahun 1991-2013

Sektor unggulan ini diperoleh dari hasil metode LQ. Jika $LQ > 1$, maka dapat dikatakan sektor tersebut sebagai sektor unggulan. Berdasarkan penghitungan LQ sektor unggulan di Kabupaten Kapuas Hulu adalah sektor pertanian dan sektor bangunan. Dalam sektor pertanian ini, subsektor yang memberikan sumbangan terbesar yaitu subsektor tanaman bahan makanan, subsektor kehutanan dan subsektor perikanan. Untuk sektor bangunan, penyumbang terbesar pada sektor tersebut adalah pembangunan jalan, jembatan, dan bangunan lainnya selain bangunan tempat tinggal.



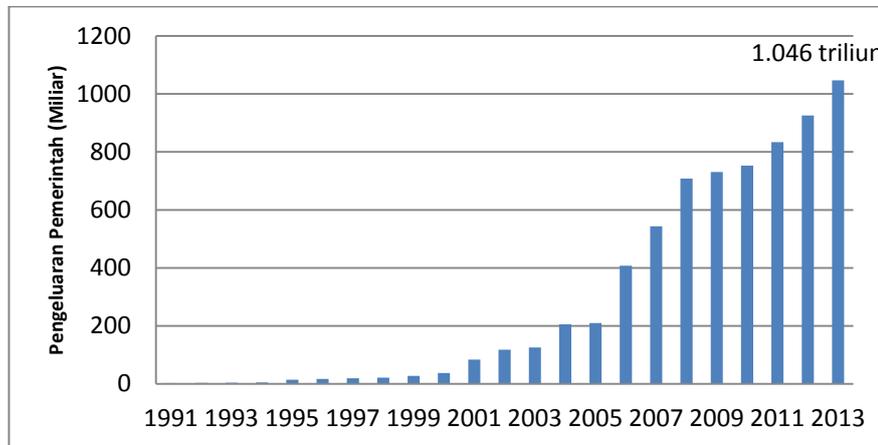
Sumber :

BPS Kabupaten Kapuas Hulu (diolah)

Gambar 3. PDRB Sektor Unggulan Kabupaten Kapuas Hulu Tahun 1991-2013

Terlihat pada Gambar 3, PDRB sektor pertanian dari tahun 1991-2013 secara umum mengalami peningkatan. Pada tahun 2003 sektor pertaniannya tercatat sebesar Rp 466 Miliar dan pada tahun 2004 turun menjadi Rp 461 Miliar, hingga pada tahun 2006 menjadi Rp 450 miliar. Kemudian dari tahun 2007 hingga tahun 2013 meningkat sebesar Rp 563 miliar. Sektor konstruksi hampir sama halnya seperti sektor pertanian. Secara umum, PDRB sektor konstruksi ini

mengalami peningkatan dari tahun 1991-2013. Pada tahun 1998, PDRB pada sektor bangunan ini mengalami penurunan sebesar Rp 549 juta. Hal ini disebabkan karena adanya krisis ekonomi pada tahun 1998.



Sumber : BPS Kabupaten Kapuas Hulu (diolah)

Gambar 4. Pengeluaran Pemerintah Kabupaten Kapuas Hulu Tahun 1991-2013

Berdasarkan Gambar 4, pengeluaran pemerintah Kabupaten Kapuas Hulu dari tahun 1991 hingga tahun 2013 mengalami peningkatan. Pada tahun 1991 pengeluaran pemerintah Kabupaten Kapuas Hulu sebesar Rp 2.4 miliar dan pada tahun 2013 meningkat sebesar Rp 1.04 triliun. Pengeluaran pemerintah ini bertujuan untuk melindungi dan meningkatkan kualitas kehidupan masyarakat Kabupaten Kapuas Hulu dalam upaya memenuhi kewajiban daerah yang diwujudkan dalam bentuk peningkatan pelayanan dasar, pendidikan, kesehatan, fasilitas sosial dan fasilitas umum yang layak serta mengembangkan sistem jaminan sosial. Peningkatan kualitas kehidupan masyarakat diwujudkan melalui prestasi kerja dalam pencapaian standar pelayanan.

Seperti yang telah dijelaskan sebelumnya, analisis regresi linear berganda harus memenuhi asumsi-asumsi klasik yang terdiri dari normalitas, homoskedastisitas, non-autokorelasi, dan non multikolinearitas. Asumsi-asumsi ini digunakan agar persamaan regresi layak digunakan. Oleh karena itu, sebelum melakukan analisis terhadap hasil output regresi linear berganda, terlebih dahulu dilakukan pengujian terhadap asumsi-asumsi tersebut.

Pengujian Asumsi Model Regresi linear Berganda

- a) Asumsi kenormalan
Uji Kenormalan dapat dideteksi dengan uji *Jarque-Berra*. Hasil pengolahan menunjukkan $p\text{-value}$ (0,957) > tingkat signifikansi (0,05) atau nilai *Jarque-Berra* (0,08) < tabel *Chi-Square* (35,17), yang artinya residual data berdistribusi normal. Dengan demikian asumsi kenormalan terpenuhi.
- b) Asumsi homoskedastisitas
Uji heteroskedastisitas dapat dideteksi dengan uji *White*. Hasil yang ditunjukkan bahwa tidak ada hubungan yang signifikan antara variabel bebas dengan residualnya dimana nilai $\text{obs} \cdot R^2$ (15,38) < tabel *Chi-Square* (35,17). Hal ini berarti tidak terdapat heteroskedastisitas di dalam persamaan regresi yang terbentuk sehingga asumsi homoskedastisitas terpenuhi.

- c) Asumsi non-autokorelasi
 Uji asumsi nonautokorelasi dilakukan untuk menguji apakah terdapat korelasi antar error. Data time series sering menunjukkan adanya korelasi serial, namun dalam regresi klasik agar dapat menggunakan estimasi OLS diperlukan asumsi nonautokorelasi yaitu tidak adanya korelasi antar error bukan antar datanya. Dengan uji *Durbin-Watson* diperoleh nilai $d = 2,359$. Berdasarkan tabel *Durbin-Watson*, didapat nilai $dL = 1,17$ dan nilai $dU = 1,54$. Hal ini menandakan tidak terjadi korelasi antar error karena nilai d berada di antara dL dan $4-dU$.
- d) Asumsi non multikolinearitas
 Uji multikolinearitas dapat dilihat dengan nilai VIF. Jika nilai $VIF > 10$ maka terjadi multikolinearitas. Pada semua variabel bebas terjadi multikolinearitas yang ditunjukkan dengan nilai $VIF > 10$. Mengingat asumsi multikolinearitas sangat penting untuk dipenuhi karena mempengaruhi interpretasi koefisien dalam model, maka untuk mengatasi adanya multikolinearitas tersebut digunakan *Ridge Regression* [11].

Model Regresi Linear Berganda

Berdasarkan hasil pengolahan data diperoleh persamaan Regresi Linear Berganda sebagai berikut:

$$\ln PAD = -160.713 + 14.289 \ln JP + 1.039 \ln PP + 9.023 \ln PDRB \text{Pertanian} - 5.375 \ln PDRB \text{Konstruksi}$$

dengan

- $\ln PAD$ = ln Pendapatan Asli Daerah
 $\ln JP$ = ln jumlah penduduk
 $\ln PP$ = ln pengeluaran pemerintah
 $\ln PDRB \text{Pertanian}$ = ln PDRB sektor pertanian
 $\ln PDRB \text{Konstruksi}$ = ln PDRB sektor konstruksi

Tabel 1. Ringkasan hasil regresi linear berganda

Variabel (1)	Koefisien (2)	SE (3)	p-value (4)	VIF (5)	Adj-R ² (6)
intercept	-160.713	50.56	0.0052		0.9318
ln JP	14.289	4.545	0.0017	23.529	
ln PP	1.039	0.283	0.0056	30.983	
ln PDRB Pertanian	9.023	3.222	0.002	16.973	
ln PDRB Konstruksi	-5.375	1.485	0.0119	66.971	

Sumber: Hasil pengolahan dengan e-views 6

Pengujian Keberartian Model

- a). Uji Keseluruhan Parameter

Untuk menguji apakah model layak digunakan maka dapat dilakukan pengujian terhadap hubungan antara variabel tak bebas dengan variabel bebas. Dari tabel *Anova* dapat diketahui bahwa hasil pengujian Dilihat dari *p-value* $0,00 < \text{tingkat signifikansi} (0,05)$. Dengan demikian dapat diputuskan bahwa tolak H_0 dengan tingkat signifikansi 5 persen. Artinya bahwa setidaknya minimal ada satu β_i yang tidak bernilai nol. Hal ini menyimpulkan bahwa secara bersama-sama keempat variabel bebas yaitu jumlah penduduk,

pengeluaran pemerintah, PDRB sektor pertanian, dan PDRB sektor konstruksi secara statistik memiliki pengaruh terhadap Pendapatan Asli Daerah (PAD) Kabupaten Kapuas Hulu. Oleh karena itu, dapat diputuskan bahwa dengan tingkat kepercayaan 95 persen model yang diperoleh layak digunakan dalam analisis.

b). Uji Parameter Individual

Setelah pengujian parameter secara simultan selanjutnya setiap parameter akan diuji satu persatu dengan menggunakan uji-t. Dari hasil pengolahan (Tabel 1) dapat dilihat bahwa keempat variabel bebas yaitu jumlah penduduk, pengeluaran pemerintah, PDRB sektor pertanian, dan PDRB sektor konstruksi secara statistik memiliki pengaruh terhadap Pendapatan Asli Daerah Kabupaten Kapuas Hulu dengan semua nilai *p-value* pada variabel bebas < tingkat signifikansi (0,05). Namun, pada persamaan model RLB, terjadi multikolinearitas yang menandakan nilai VIF pada semua variabel bebas > 10. Adanya multikolinearitas dapat diatasi menggunakan *Ridge Regression*.

Model Regresi Ridge

Adanya multikolinearitas dalam model dapat menyebabkan standar error semakin besar sehingga selang keyakinan estimator semakin lebar. Tentunya akan mengurangi presisi dari nilai estimasi. Untuk mengatasi hal tersebut digunakan metode regresi ridge. Estimasi model regresi ridge yang digunakan adalah *Ordinary Ridge Regression (ORR)* yang distandarkan dengan tahapan sebagai berikut:

1. Mentransformasi semua variabel dengan transformasi korelasi.
2. Memilih konstanta bias c.
3. Menentukan persamaan regresi ridge.
4. Transformasi persamaan regresi ridge ke bentuk awal.

Dengan menggunakan software Herodes 1.4 untuk mengolah data dengan menggunakan metode *ORR* diperoleh konstanta bias c yang optimum adalah sebesar 0.091. Setelah memilih konstanta c, membuat persamaan regresi ridge, dan mentransformasi kembali ke bentuk awal, diperoleh persamaan regresi sebagai berikut

$$\ln PAD = -73,324 + 6,258 \ln JP + 0,335 \ln PP + 2,369 \ln \text{Pertanian} + 0,084 \ln \text{Konstruksi}$$

dengan

- $\ln PAD$ = ln Pendapatan Asli Daerah
- $\ln JP$ = ln jumlah penduduk
- $\ln PP$ = ln pengeluaran pemerintah
- $\ln PDRB \text{Pertanian}$ = ln PDRB sektor pertanian
- $\ln PDRB \text{Konstruksi}$ = ln PDRB sektor konstruksi

Tabel 2. Ringkasan hasil pengolahan *ridge regression*

Variabel (1)	Koefisien (2)	SE (3)	VIF (5)	Adj-R ² (6)
intercept	-73.324			0.843
ln JP	6.258	1.836	1.666	
ln PP	0.335	0.096	1.555	
ln PDRB Pertanian	2.369	1.666	1.968	

ln PDRB Konstruksi	0.084	0.257	0.87	
--------------------	-------	-------	------	--

Sumber: Hasil pengolahan dengan Herodes 1.4

Berdasarkan Tabel 2, dapat disimpulkan bahwa *Ridge Regression* dapat digunakan untuk menghilangkan multikolinearitas dalam persamaan regresi tanpa banyak menghilangkan informasi penting dalam model terlihat dari nilai VIF yang lebih kecil yaitu di bawah 10 dibandingkan dengan menggunakan metode OLS biasa. Selain itu terlihat dari adj-R^2 sebelum dan sesudah menggunakan *Ridge Regression* juga tidak berbeda jauh yaitu 93,18 persen dan 84,3 persen.

Interpretasi Model

Berdasarkan hasil di atas dapat disimpulkan bahwa variabel-variabel yang memengaruhi Pendapatan Asli Daerah (PAD) di Kabupaten Kapuas Hulu adalah jumlah penduduk, pengeluaran pemerintah, PDRB sektor pertanian, dan PDRB sektor konstruksi dengan tingkat signifikansi 5 persen.

Koefisien regresi pada variabel jumlah penduduk (b_1) menunjukkan angka yang bernilai positif yaitu 6,258. Angka ini secara signifikan berpengaruh terhadap Pendapatan Asli Daerah (PAD) dan dapat diartikan bahwa setiap kenaikan satu persen jumlah penduduk dapat mengakibatkan kenaikan Pendapatan Asli Daerah (PAD) sebesar 6,258 persen dengan asumsi variabel pengeluaran pemerintah, sektor pertanian, dan sektor konstruksi dalam keadaan konstan. Koefisien regresi pada variabel pengeluaran pemerintah (b_2) menunjukkan angka yang bernilai positif yaitu 0,335. Angka ini secara signifikan berpengaruh terhadap Pendapatan Asli Daerah (PAD) dan dapat diartikan bahwa setiap kenaikan satu persen pengeluaran pemerintah dapat mengakibatkan kenaikan Pendapatan Asli Daerah (PAD) sebesar 0,335 persen dengan asumsi variabel jumlah penduduk, sektor pertanian, dan sektor konstruksi dalam keadaan konstan.

Koefisien regresi pada variabel sektor pertanian (b_{31}) menunjukkan angka yang bernilai positif yaitu 2,369. Angka ini secara signifikan berpengaruh terhadap Pendapatan Asli Daerah (PAD) dan dapat diartikan bahwa setiap kenaikan satu persen sektor pertanian dapat mengakibatkan kenaikan Pendapatan Asli Daerah (PAD) sebesar 2,369 persen dengan asumsi variabel jumlah penduduk, pengeluaran pemerintah, dan sektor konstruksi dalam keadaan konstan. Koefisien regresi pada variabel sektor konstruksi (b_{35}) menunjukkan angka yang bernilai positif yaitu 0,084. Angka ini secara signifikan berpengaruh terhadap Pendapatan Asli Daerah (PAD) dan dapat diartikan bahwa setiap kenaikan satu persen sektor konstruksi dapat mengakibatkan kenaikan Pendapatan Asli Daerah (PAD) sebesar 0,084 persen dengan asumsi variabel jumlah penduduk, pengeluaran pemerintah, dan sektor pertanian, dalam keadaan konstan.

3. Kesimpulan

Pendapatan Asli Daerah Kabupaten Kapuas Hulu mengalami peningkatan dari tahun 1991-2013 demikian juga dengan variabel-variabel yang diduga memengaruhinya yaitu jumlah penduduk, pengeluaran pemerintah, dan PDRB sektor unggulan. Sektor unggulan Kabupaten Kapuas Hulu adalah sektor pertanian dan sektor konstruksi.

Berdasarkan hasil Regresi Linear Berganda, semua variabel berpengaruh secara signifikan terhadap Pendapatan Asli Daerah Kabupaten Kapuas Hulu yaitu

jumlah penduduk, pengeluaran pemerintah, PDRB sektor pertanian dan PDRB sektor konstruksi. Namun, terjadi multikolinearitas yang diatasi dengan menggunakan *ridge regression*. Selain mampu mengatasi multikolinieritas yang ditandai dengan nilai VIF menjadi < 10 , model regresi ridge juga menghasilkan hasil yang sama yaitu semua variabel berpengaruh secara signifikan terhadap Pendapatan Asli Daerah Kabupaten Kapuas Hulu.

Referensi

- [1] Badan Pusat Statistik. Berbagai Edisi. *Kabupaten Kapuas Hulu dalam Angka*. Jakarta: Badan Pusat Statistik.
- [2] Badan Pusat Statistik. Berbagai Edisi. *Kalimantan Barat dalam Angka*. Jakarta: Badan Pusat Statistik.
- [3] Badan Pusat Statistik (2001). *Indikator-indikator Ekonomi*. Jakarta: Badan Pusat Statistik
- [4] Badan Pusat Statistik. (2011). *Publikasi Pertumbuhan dan Persebaran Penduduk Indonesia*, Hasil Sensus 2010. Jakarta: Badan Pusat Statistik.
- [5] Darmawan. (2012). *Analisis Struktur Ekonomi dengan Metode Shift share dan Loction Quotient serta Dampaknya terhadap Pengembangan Wilayah Kota Medan*. [Tesis]. Medan: Universitas Sumatera Utara.
- [6] Datu, I. R. (2012). *Analisis Faktor yang Mempengaruhi PAD di Kota Makassar Tahun 1999-2009*. Makassar: Universitas Hasanuddin.
- [7] Dirjen Perimbangan dan Keuangan. Melalui www.djpk.go.id. Di akses tanggal 25 Agustus 2015.
- [8] Dornbush, Riduger et al. (2004). *Makroekonomi*. (Penerjemah: Yusuf Wibisono & Indra Mirazudin). Jakarta: PT. Media Global Edukasi.
- [9] Gujarati, Damodar N. (2003). *Dasar-dasar Ekonometrika*, 3th ed. (Penerjemah: Julius A & Mulyadi). Jakarta: Erlangga.
- [10] Johnson, R. A & Wichern, D.W. (2002). *Applied Multivariate Statistical Analysis*, 5th ed. Pearson Educational International.
- [11] Neter, John et al. (1989). *Applied Linear Regression Models*, 2th ed. Boston
- [12] Nachrowi, N. D & Usman, H. (2002). *Penggunaan Teknik Ekonometri*. Jakarta: PT. Raja Grafindo Persada.
- [13] Santoso, P.B & Rahayu, R.P. (2005). *Analisis Pendapatan Asli Daerah (PAD) dan Faktor-faktor yang Mempengaruhi Dalam Upaya Pelaksanaan Otonomi Daerah di Kabupaten Kediri*. Jurnal Dinamika Pembangunan Vol.2 No. 1 / Juli 2005.
- [14] Sidik, M et al. 2002. *Dana Alokasi Umum – Konsep, Hambatan, dan Prospek di Era Otonomi Daerah*. Jakarta: Penerbit Buku Kompas
- [15] Supranto, J. (2004). *Ekonometri, buku kedua*. Bogor: PT. Ghalia Indonesia.
- [16] Webster, A. L. (1992). *Applied Statistics for Business and Economics*. USA: McGraw-Hill Book

PENGARUH ENERGI TERMODINAMIKA DAN MOMEN DIPOL TERHADAP KESTABILAN KOMPLEKS GADOLINIUM DENGAN LIGAN TURUNAN DBDTP MENGUNAKAN REGRESI BERGANDA

NURDENI

Universitas Padjadjaran
nurdeni@unindra.ac.id

Abstrak

Metode statistika dan matematika dapat digunakan dalam berbagai bidang, dalam penelitian ini diterapkan pada bidang kimia. Penelitian ini bertujuan untuk mengetahui pengaruh energi dan momen dipol terhadap kestabilan kompleks gadolinium dengan ligan turunan DBDTP. Metode penelitian yang digunakan adalah regresi linier berganda, dengan asumsi *error* model berdistribusi normal dengan mean nol dan variansi konstan, dan terdapat korelasi antara energi dengan momen dipol terhadap kestabilan, serta parameter model bersifat linear dengan tanda koefisien negatif. Hasil penelitian ini secara simultan menunjukkan bahwa energi dan momen dipol berdasarkan uji F berpengaruh terhadap kestabilan. Hasil secara parsial dengan uji t menunjukkan bahwa kontribusi variabel energi berpengaruh sangat signifikan sebesar 94,77%, sedangkan kontribusi variabel momen dipol berpengaruh tidak signifikan sebesar 0,264%. Koefisien determinasi yang dihasilkan adalah sebesar 0,946 atau 94,6%. Nilai ini menunjukkan bahwa kontribusi energi dan momen dipol terhadap kestabilan yaitu sebesar 94,6%, sedangkan sisanya sebesar 5,4% disebabkan oleh faktor lain diluar model yang diteliti.

Kata Kunci: Energi Termodinamika, Gadolinium, Ligan DBDTP, Regresi Berganda.

1. Pendahuluan

Gadolinium merupakan salah satu unsur dari Unsur Tanah Jarang (UTJ) yang cukup populer saat ini, karena penggunaannya sebagai senyawa *contrast agent* untuk keperluan pencitraan medis *Magnetic Resonance Imaging* (MRI) yang dapat memperjelas visualisasi jaringan tubuh [1]. Gadolinium digunakan sebagai pengontras dikarenakan memiliki sifat paramagnetik yang tinggi, hal ini disebabkan oleh adanya 7 elektron tidak berpasangan dalam konfigurasi atomnya.

Pemisahan UTJ sangat sulit dilakukan karena mempunyai sifat yang sangat mirip antara satu unsur dengan unsur yang lainnya. Metode-metode pemisahan

yang telah banyak digunakan untuk memperoleh UTJ adalah metode kristalisasi bertingkat, pengendapan bertingkat, ekstraksi pelarut, dan metode kromatografi kolom pertukaran ion [2]. Ekstraksi pelarut adalah suatu teknik yang handal untuk memisahkan UTJ. Salah satu inovasi dari ekstraksi adalah pengembangan ekstraktan yang membentuk kompleks dengan unsur transisi maupun tanah jarang seperti Tris (Dibutylidithiophosphato) Gadolinium (III) [3].

Untuk mengekstraksi UTJ, sejumlah kajian telah dilakukan dengan menggunakan pelarut organik yang mengandung ligan-ligan pembentuk kompleks. Salah satu ligan pengompleks yang telah digunakan adalah ligan dibutilditiofosfat (DBDTP) dengan beberapa pelarut organik untuk mengekstraksi Gd dan Sm [4]. Seiring dengan kemajuan teknologi, pemisahan UTJ telah dilakukan dengan pendekatan kimia komputasi, dengan cara ini sangat meminimalkan biaya dan mengefektifkan waktu analisis.

Kemampuan gadolinium untuk membentuk senyawa kompleks dengan beberapa ligan telah dibuktikan dalam beberapa penelitian, diantaranya dengan menggunakan *di-butyl-dithiocarbamate* (DBDTC) dan *di-butyl-dithiophosphat* (DBDTP). Seperti yang telah dilakukan

[5] ekstraksi logam gadolinium telah dilakukan dengan menggunakan ligan DBDTP dengan membentuk senyawa kompleks dengan hasil *recovery* diatas 90%. Selain itu, telah dilakukan desain turunan ligan DBDTP untuk ekstraksi logam gadolinium dengan beberapa deskriptor yang mempengaruhi kestabilan senyawa kompleks [5].

Kestabilan senyawa kompleks diperoleh dengan menggunakan pendekatan termodinamika, dimana dengan pendekatan ini suatu senyawa kimia dapat dikatakan stabil atau tidak stabil. Adapun faktor-faktor yang mempengaruhi kestabilan senyawa kompleks diantaranya faktor energi termodinamika dan momen dipol, sehingga tujuan dari penelitian adalah untuk mengetahui pengaruh energi termodinamika dan momen dipol terhadap kestabilan kompleks Gadolinium dengan ligan turunan DBDTP.

2. Hasil - Hasil Utama

Nilai konstanta stabilitas menunjukkan kestabilan dari suatu kompleks.

Semakin besar konstanta stabilitasnya semakin stabil senyawa kompleksnya, faktor yang mempengaruhi kestabilan diantaranya adalah energi termodinamika dan momen dipol. Untuk mengetahui kontribusi pengaruh dari kedua faktor itu maka digunakan regresi berganda. Bahan yang digunakan dalam penelitian ini adalah 20 senyawa kompleks gadolinium dengan ligan turunan DBDTP merupakan data sekunder dari hasil penelitian [5] seperti yang disajikan pada Tabel 1.

Tabel 1
Data Sekunder

Ligan	Y	X ₁	X ₂
L01	71.576	-8315.438	4.390
L02	71.555	-8313.316	19.440
L03	71.299	-8309.689	27.080
L04	82.887	-9182.365	15.795
L05	82.422	-9159.264	8.758
L06	109.516	-11246.766	16.298
L07	96.078	-10203.499	16.537
L08	96.105	-10206.804	15.542
L09	75.126	-8595.564	21.410
L10	83.198	-9211.217	10.402
L11	94.244	-10070.714	3.646
L12	94.880	-10113.161	9.375
L13	98.267	-10360.196	9.866
L14	94.267	-10069.479	3.640
L15	132.393	-13016.74	16.221
L16	131.100	-12903.069	5.479
L17	119.423	-12005.75	14.570
L18	-19.833	-5047.704	8.595
L19	-19.879	-5044.772	15.215
L20	-21.130	-5047.335	13.393

Dimana :

- Y = Kestabilan kompleks
- X₁ = Energi termodinamika (*Joule*)
- X₂ = Momen Dipol (*Deybe/D*)

Hasil perhitungan melalui program paket *software* SPSS 20.00 dari data sekunder diatas adalah sebagai berikut :

Table 2. Hasil Perhitungan Pengujian Koefisien Korelasi Ganda Variabel X_1 dan X_2 terhadap Y

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.972 ^a	.946	.939	11.23651
a. Predictors: (Constant), X2, X1				

Dari Tabel 2 dapat dijelaskan bahwa secara bersama-sama variabel energi termodinamika dan momen dipol memberikan kontribusi sebesar 94,6 % terhadap variable kestabilan kompleks gadolinium dengan ligan turunan DBDTP.

Table 3. Hasil Perhitungan Pengujian Signifikansi Koefisien Regresi Variabel X_1 dan X_2 terhadap Y

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	37391.085	2	18695.543	148.073	.000 ^b
	Residual	2146.404	17	126.259		
	Total	39537.489	19			
a. Dependent Variable: Y						
b. Predictors: (Constant), X2, X1						

Berdasarkan tabel 3 dapat dinyatakan bahwa terdapat pengaruh yang sangat signifikan energi termodinamika dan momen dipol secara bersama-sama terhadap kestabilan kompleks gadolinium dengan ligan turunan DBDTP. Hal ini dibuktikan dengan perolehan nilai Sig. $0,000 < 0,05$ dan $F_h = 148,073$.

Table 4. Hasil Perhitungan Persamaan Regresi Ganda Variabel X_1 dan X_2 terhadap Y

Coefficients ^a						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-107.690	12.542		-8.587	.000
	X1	-.019	.001	-.976	-17.176	.000
	X2	.325	.416	.044	.780	.446
a. Dependent Variable: Y						

Hal ini memiliki pengertian bahwa ketika variable energy termodinamika

naik 1 satuan, maka variable kestabilan kompleks gadolinium dengan turunan ligan DBDTP akan turun sebesar 0,019 satuan, ketika momen dipol konstan. Sedangkan ketika variable momen dipol naik 1 satuan, maka variable kestabilan kompleks gadolinium dengan turunan ligan DBDTP akan naik sebesar 0,325 satuan ketika energy termodinamika konstan.

Sementara itu, persamaan garis regresi ganda dapat dinyatakan dengan $\hat{Y} = -107,690 - 0,019 X_1 + 0,325 X_2$. Hal ini memiliki pengertian bahwa variabel energi termodinamika memberikan kontribusi sebesar -0,019 ketika variable X_2 konstan dan variable momen dipol memberikan kontribusi sebesar 0,325 ketika variable X_1 konstan terhadap variable kestabilan kompleks gadolinium dengan ligan turunan DBDTP.

Dari Table 4 dapat dinyatakan bahwa terdapat pengaruh yang sangat signifikan energi termodinamika terhadap kestabilan kompleks gadolinium dengan ligan turunan DBDTP. Hal ini dibuktikan dengan perolehan nilai Sig. $0,000 < 0,05$ dan $t_h = -17,176$.

Adapun kontribusi variabel energi termodinamika terhadap kestabilan kompleks gadolinium dengan ligan turunan DBDTP adalah :

$$KD = \text{Nilai } \beta_{x_1y} \times \text{Nilai Korelasi Parsialnya } (r_{x_1y}) \times 100 \% \quad [6]$$

$$KD = (-0,976) \times (-0,971) \times 100 \% = 94,77 \%$$

Dari hasil perhitungan di atas dapat dinyatakan bahwa kontribusi energi termodinamika dalam meningkatkan kestabilan kompleks gadolinium dengan ligan turunan DBDTP sebesar 94,77 %.

Dari Table 4 dapat dinyatakan bahwa terdapat pengaruh yang tidak signifikan momen dipol terhadap kestabilan kompleks gadolinium dengan ligan turunan DBDTP. Hal ini dibuktikan dengan perolehan nilai Sig. $0,446 > 0,05$ dan $t_h = 0,780$.

Adapun kontribusi variabel momen dipol terhadap kestabilan kompleks gadolinium dengan ligan turunan DBDTP adalah :

$$KD = \text{Nilai } \beta_{x_2y} \times \text{Nilai Korelasi Parsialnya } (r_{x_2y}) \times 100 \%$$

$$KD = 0,044 \times (-0,060) \times 100 \% = (-0,26) \%$$

Dari hasil perhitungan di atas dapat dinyatakan bahwa kontribusi momen dipol dalam meningkatkan kestabilan kompleks gadolinium dengan ligan turunan DBDTP sebesar (-0,26) %.

3. Kesimpulan

Terdapat pengaruh yang sangat signifikan energi termodinamika dan momen dipol secara bersama-sama terhadap kestabilan kompleks gadolinium dengan ligan turunan DBDTP. Hal ini dibuktikan dengan kontribusi yang diberikan sebesar 94,6% terhadap kestabilan kompleks gadolinium dengan ligan turunan DBDTP, sisanya 5,4 % dipengaruhi oleh faktor lain yang terkandung dalam unsur *error* atau unsur lain.

Dengan melihat pengaruh yang sangat signifikan maka model ini dapat digunakan untuk prediksi kestabilan kompleks gadolinium dengan ligan turunan DBDTP.

Referensi

- [1] Gupta, C. K & Krishnamurthy N. 2005. *Extractive metallurgy of rare earth*. CRC Press, Boca Raton London.
- [2] Amin, A. 2009. Pemisahan unsur Sm dan Yttrium dari mineral tanah jarang dari teknik membrane cair berpendukung (*Supported Liquid Membrane*). *Jurnal Rekayasa Kimia & Lingkungan*. 7 (1), 15-23.
- [3] Pratomo *et al.*, Synthesis, Characterization, and Molecular Modelling of Bis(Aquo)Tris(Dibutyldithiophosphato) Gadolinium (III). *Procedia Chemistry* 17 (2015) 207 – 215. Elsevier.
- [4] Sofyatin, Nunik Nurlina, Anggraeni, A. & Bahti, H.H. (2016). Penentuan Koefesien Distribusi, Efisiensi Ekstraksi dan Faktor Pemisahan pada Ekstraksi Gadolinium dan Samarium dengan Ligan Dibutilditiofosfat. *Chimica et Natura*. 47-51.
- [5] Novian, 2016. *Disain Turunan Ligan Dbdtp Untuk Ekstraksi Logam Gadolinium, Dari Kelompok Lantanida Menengah (Terbium Dan Disprosium) Menggunakan Metode Semi-Empirik Pm7*. Universitas Padjadjaran. Bandung.
- [6] Suparman I.A. *Aplikasi Komputer Dalam Penyusunan Karya Ilmiah*. Pustaka Mandiri, 2013. 116.

MODEL REGRESI *ZERO-INFLATED JOINT BIVARIATE POISSON* DAN APLIKASINYA UNTUK ANALISIS DATA KEMISKINAN DI BOGOR

FEVI NOVKANIZA¹, KHAIRIL ANWAR NOTODIPUTRO²,
KUSMAN SADIK³

1 Departemen Matematika FMIPA UI, fevi.novkaniza@sci.ui.ac.id
2 Departemen Statistika FMIPA IPB, khairilnotodiputro@gmail.com
3 Departemen Statistika FMIPA IPB, kusmansadik@gmail.com

Abstrak. Model regresi untuk dua atau lebih data *count* banyak diterapkan dalam berbagai bidang ilmu. Salah satu model regresi untuk dua data *count* adalah model regresi *Jointly Bivariate Poisson* (JBP). Dalam model regresi JBP, variabel respon bivariat berkorelasi dan dapat tergantung pada satu atau lebih kovariat. Namun seringkali diperoleh bentuk data *count* bivariat dengan kasus (0,0) yang banyak (*excess zero* atau *zero-inflated*) sehingga terjadi overdispersi. Untuk memodelkan data *count* berkorelasi dengan *excess zero*, banyak model yang digunakan seperti model regresi *Zero-Inflated Poisson* (ZIP), *Zero-Inflated Negative Binomial* (ZINB), *Zero-Inflated Beta Binomial* (ZIBB) dan lainnya. Ide dasar dari model regresi tersebut adalah distribusi variabel respon berasal dari campuran beberapa distribusi, seperti distribusi *degenerate*, Bernoulli, Poisson, Binomial, Binomial Negatif, Beta dan lain-lain. Dalam tulisan ini dijelaskan tentang model regresi *Zero-Inflated Jointly Bivariate Poisson* (ZIJBP) dan estimasi parameter model menggunakan metode maksimum *likelihood*. Sebagai aplikasi model regresi ZIJBP diterapkan untuk analisis data kemiskinan di Kabupaten dan Kota Bogor menggunakan data Susenas tahun 2013 dan Podes tahun 2014. Hasil model regresi ZIJBP menunjukkan bahwa jumlah rumah tangga miskin di desa Bogor dengan indeks kedalaman kemiskinan diatas indeks kedalaman kemiskinan Kabupaten dan Kota Bogor tidak dipengaruhi oleh sumber penghasilan pertanian dan sumber air yang digunakan.

Kata kunci: bivariat, Poisson, *zero-inflate*

1. Pendahuluan

Dalam berbagai bidang ilmu seperti epidemiologi, kesehatan masyarakat, biologi, sosiologi, psikologi, engineering, ekologi, dan pertanian, seringkali diperoleh data *count* dengan banyak 0 (*excess zero* atau *zero-inflated*). Untuk memodelkan data *count* dengan *excess zero*, sudah banyak penelitian yang dilakukan seperti model regresi *Zero-Inflated Poisson* (ZIP), *Zero-Inflated Negative Binomial* (ZINB), *Zero-Inflated Beta Binomial* (ZIBB) dan lainnya. Ide dasar dari model regresi tersebut adalah distribusi variabel respon berasal dari campuran beberapa distribusi, seperti distribusi *degenerate*, Bernoulli, Poisson, Binomial, Binomial Negatif, Beta dan lain-lain. Seringkali model regresi untuk data *count* melibatkan tidak hanya satu variabel respon, tetapi dua variabel respon

yang memiliki distribusi bivariat. Salah satunya adalah dua variabel respon yang memiliki distribusi *bivariate* Poisson. Distribusi *bivariate* Poisson sering digunakan untuk data *count bivariate*. Sebagai contoh dua variabel respon dengan distribusi *bivariate* Poisson adalah jumlah darurat dan non darurat kunjungan seseorang ke rumah sakit, jumlah klaim asuransi dengan dan tanpa cedera tubuh, dan lainnya. Distribusi *bivariate* Poisson pertama kali diusulkan oleh Holgate pada tahun 1964 dan dalam perkembangannya terdapat beberapa definisi untuk distribusi Poisson bivariat. Menurut Johnson [4], salah satu metode untuk mengkonstruksi distribusi *bivariate* Poisson adalah melalui metode reduksi *trivariate* yaitu :

Distribusi *Jointly Bivariate Poisson*

Menurut Karlis dan Ntzoufra [4], jika X_1, X_2 dan X_3 masing-masing menotasikan peubah acak Poisson dengan parameter λ_1, λ_2 dan $\lambda_3 > 0$ dan mendefinisikan $Y_1 = X_1 + X_3$ dan $Y_2 = X_2 + X_3$, maka fungsi kepadatan peluang bersama untuk peubah acak Y_1 dan Y_2 adalah :

$$f_{JBP}(y_1, y_2; \lambda_1, \lambda_2, \lambda_3) = e^{-(\lambda_1 + \lambda_2 + \lambda_3)} \frac{\lambda_1^{y_1}}{y_1!} \frac{\lambda_2^{y_2}}{y_2!} \sum_{r=0}^{\min(y_1, y_2)} \binom{y_1}{r} \binom{y_2}{r} r! \left(\frac{\lambda_3}{\lambda_1 \lambda_2} \right)^r \quad (1)$$

Distribusi marginal Y_1 adalah distribusi Poisson dengan mean $\lambda_1 + \lambda_3$ dan distribusi marginal Y_2 juga distribusi Poisson dengan mean $\lambda_2 + \lambda_3$. Kovariansi antara peubah acak Y_1 dan Y_2 adalah λ_3 (λ_3 merupakan ukuran kebergantungan antara peubah acak Y_1 dan Y_2) sehingga korelasi antara Y_1 dan Y_2 :

$$\rho_{Y_1, Y_2} = \frac{\lambda_3}{\sqrt{(\lambda_1 + \lambda_3)(\lambda_2 + \lambda_3)}}$$

Jika nilai $\lambda_3 = 0$, maka Y_1 dan Y_2 independen dan distribusi *bivariate* Poisson menjadi distribusi *Double Poisson*, yaitu distribusi yang diperoleh dari perkalian dua distribusi Poisson univariat.

Model Regresi *Jointly Bivariate Poisson*

Model regresi yang memiliki variabel respon bivariat berdistribusi Poisson bivariat disebut model regresi *Jointly Bivariate Poisson* (JBP). Dalam model regresi ini, parameter distribusi variabel respon bivariat yaitu λ_1, λ_2 dan λ_3 bisa tergantung pada satu atau lebih kovariat yang dihubungkan melalui fungsi penghubung kanonik yaitu fungsi logaritma. Model JBP memiliki bentuk :

$$(Y_{1i}, Y_{2i}) \sim JBP(\lambda_{1i}, \lambda_{2i}, \lambda_{3i})$$

$$\log \lambda_{ki} = \mathbf{x}_i^T \boldsymbol{\beta}_k, \quad k = 1, 2, 3 \quad \text{dan } i = 1, 2, \dots, n$$

dimana :

\mathbf{x}_i = vektor kovariat dengan panjang m untuk observasi ke-i yang berhubungan dengan parameter ke-k

$\boldsymbol{\beta}_k$ = vektor koefisien regresi

Dalam melakukan estimasi parameter model, digunakan metode maksimum *likelihood*. Misalkan observasi independen (y_{1i}, y_{2i}) dengan vektor ke- i memiliki distribusi Poisson bivariat sebagai berikut :

$$f_{JBP}(y_{1i}, y_{2i}; \lambda_{1i}, \lambda_{2i}, \lambda_{3i}) = e^{-(\lambda_{1i} + \lambda_{2i} + \lambda_{3i})} \varphi(y_{1i}, y_{2i}) \quad (2)$$

dimana :

$$\varphi(y_{1i}, y_{2i}) = \sum_{r=0}^{\min(y_1, y_2)} \frac{\lambda_{3i}^r \lambda_{1i}^{(y_{1i}-r)} \lambda_{2i}^{(y_{2i}-r)}}{r! (y_{1i}-r)! (y_{2i}-r)!}, \quad i = 1, 2, \dots, n$$

Fungsi skor yang bersesuaian adalah :

$$U_{1j} = \frac{\partial \log L}{\partial \beta_{1j}} = \sum_{i=1}^n \lambda_{1i} x_{ji} \left[\frac{\varphi(y_{1i}-1, y_{2i})}{\varphi(y_{1i}, y_{2i})} - 1 \right]$$

$$U_{2j} = \frac{\partial \log L}{\partial \beta_{2j}} = \sum_{i=1}^n \lambda_{2i} x_{ji} \left[\frac{\varphi(y_{1i}, y_{2i}-1)}{\varphi(y_{1i}, y_{2i})} - 1 \right]$$

$$U_{3j} = \frac{\partial \log L}{\partial \beta_{3j}} = \sum_{i=1}^n \lambda_{3i} x_{ji} \left[\frac{\varphi(y_{1i}-1, y_{2i}-1)}{\varphi(y_{1i}, y_{2i})} - 1 \right]$$

dimana $j = 1, 2, \dots, m$

Model Regresi Zero-Inflated Bivariate Poisson

Untuk memodelkan data *count* dengan *excess zero*, sudah banyak penelitian yang dilakukan seperti model regresi *Zero-Inflated Poisson* (ZIP), *Zero-Inflated Generalized Poisson* (ZIGP), *Zero-Inflated Negative Binomial* (ZINB), *Zero-Inflated Beta Binomial* (ZIBB) dan lainnya. Untuk memodelkan data *count* bivariat, perlu dikembangkan model regresi ZIP bivariat untuk 2 variabel respon yang berkorelasi dan memiliki *excess zero*. Model ini pertama kali diperkenalkan oleh Karlis dan Ntzoufras (2005) dan disebut dengan model regresi *Zero-Inflated Jointly Bivariate Poisson* (ZIJBP). Model regresi ZIJBP dapat memperbaiki masalah overdispersi dan masalah korelasi yang ditemui dalam variabel respon bivariat. Menurut AlMuhayfith dkk [3], dalam model regresi ZIJBP vektor respon bivariat memiliki fungsi kepadatan peluang:

$$f_{ZIJBP}(y_{1i}, y_{2i}) = \begin{cases} \phi + (1-\phi)f_{JBP}(0,0; \lambda_{1i}, \lambda_{2i}, \lambda_{3i}), & y_{1i} = y_{2i} = 0 \\ (1-\phi)f_{JBP}(y_{1i}, y_{2i}; \lambda_{1i}, \lambda_{2i}, \lambda_{3i}), & y_{1i} \text{ atau } y_{2i} = 0 \end{cases}$$

dimana $f_{JBP}(y_{1i}, y_{2i}; \lambda_{1i}, \lambda_{2i}, \lambda_{3i})$ adalah fungsi kepadatan peluang bersama pada persamaan (2) dan $0 < \phi < 1$. Bentuk momen dari distribusi ZIJBP adalah:

$$\begin{aligned} E(Y_k) &= (1-\phi)(\lambda_k + \lambda_3) \\ Var(Y_k) &= E(Y_k)[1 + \phi(\lambda_k + \lambda_3)], \quad k = 1, 2 \\ E(Y_1 Y_2) &= (1-\phi)[(\lambda_1 + \lambda_3)(\lambda_2 + \lambda_3) + \lambda_3] \\ Cov(Y_1, Y_2) &= (1-\phi)[\lambda_3 + \phi(\lambda_1 + \lambda_3)(\lambda_2 + \lambda_3)] \end{aligned} \quad (3)$$

Distribusi marjinal dari variabel respon univariat pada model regresi ZIJBP bukanlah distribusi Poisson saja tetapi campuran dari distribusi dimana tingkat kebergantungan antara dua variabel respon dipertimbangkan. Distribusi marjinal dari variabel respon Y_1 atau Y_2 adalah distribusi *Zero-Inflated Poisson* (ZIP).

Berikut ini akan dijelaskan tentang distribusi ZIP untuk variabel respon Y_1 . Dalam distribusi ZIP, vektor respon $\mathbf{Y}_1 = (Y_{11}, Y_{12}, \dots, Y_{1n})'$ independen dan

$$Y_1 \sim \begin{matrix} 0 & , \text{dengan probabilitas } \phi \\ \sim \text{Poisson}(\lambda_1) & , \text{dengan probabilitas } 1 - \phi \end{matrix}$$

sehingga Y_1 adalah:

$$Y_{1i} = \begin{cases} 0, & \text{dengan probabilitas } \phi_i + (1 - \phi_i)e^{-\lambda_i} \\ k, & \text{dengan probabilitas } \frac{(1 - \phi_i)e^{-\lambda_i}\lambda_{1i}^k}{k!}, k = 1, 2, \dots \end{cases}$$

Mean dan variansi distribusi ZIP adalah:

$$E(Y_1) = (1 - \phi)\lambda_1 \\ V(Y_1) = (1 - \phi)(\lambda_1 + \lambda_1^2) - ((1 - \phi)\lambda_1)^2$$

Distribusi ini akan menghampiri distribusi Poisson (λ) saat $\phi \rightarrow 0$.

Bukti:

Misalkan $U \sim \text{Bernoulli}(1 - \phi)$ dan $Z \sim \text{Poisson}(\lambda_1)$. Definisikan: $Y_1 = UZ$

$$P(Y_1 = 0) = P(U = 0) + P(U = 1, Z = 0) = \phi + (1 - \phi)\lambda_1 \quad (\text{sifat independen})$$

$$P(Y_1 = 1) = P(U = 1, Z = k), k \neq 0$$

Karena X dan Z independen, maka:

$$\mu = E(Y_1) = E(UZ) = E(U)E(Z) = (1 - \phi)\lambda_1 \\ \text{Var}(Y_1) = E(Y_1^2) - \mu^2 = E(U^2)E(Z^2) - \mu^2 = (1 - \phi)(\lambda_1 + \lambda_1^2) - [(1 - \phi)\lambda_1]^2 \square$$

Untuk melakukan estimasi parameter model regresi ZIJBP digunakan prosedur maksimum *likelihood* sebagai berikut:

1. Misalkan (y_{1i}, y_{2i}) adalah sebuah sampel random observasi dari model regresi ZIJBP
2. Fungsi likelihood untuk sampel random observasi bivariat tersebut adalah:

$$L(\phi_i, \lambda_{1i}, \lambda_{2i}, \lambda_{3i}; y_{1i}, y_{2i}) = \prod_{i=1}^n \left\{ (\phi_i + (1 - \phi_i)f_{JBP}(0,0; \lambda_{1i}, \lambda_{2i}, \lambda_{3i}))^{1-a_i} \left((1 - \phi_i)f_{JBP}(y_{1i}, y_{2i}; \lambda_{1i}, \lambda_{2i}, \lambda_{3i}) \right)^{a_i} \right\}$$

$$\text{dimana: } a_i = \begin{cases} 1, & \text{jika } (y_{1i}, y_{2i}) \neq (0,0) \\ 0, & \text{jika } (y_{1i}, y_{2i}) \text{ lainnya} \end{cases}$$

Seperti halnya dalam model regresi JBP, model regresi ZIJBP juga menggunakan fungsi penghubung kanonik log untuk parameter λ_{ki} , $k = 1, 2, 3$ dan fungsi penghubung logit untuk proporsi ϕ_i yaitu:

$$\log \lambda_{ki} = \mathbf{x}_i^T \boldsymbol{\beta}_k, k = 1, 2, 3 \\ \text{logit}(\phi_i) = \gamma_0 + \sum_{j=1}^m \gamma_j z_{ji}, \quad i = 1, 2, \dots, n,$$

Untuk mencari estimasi parameter model regresi ZIJBP, dapat digunakan metode maksimum *likelihood* dengan menyelesaikan sistem persamaan berikut ini:

$$\frac{\partial \log L}{\partial \beta_{1j}} = 0, \quad \frac{\partial \log L}{\partial \beta_{2j}} = 0, \quad \frac{\partial \log L}{\partial \beta_{3j}} = 0 \quad \text{dan} \quad \frac{\partial \log L}{\partial \gamma_j} = 0, \quad j = 1, 2, \dots, m$$

Karena bentuk sistem persamaan diatas sulit diselesaikan secara analitis, dibutuhkan pendekatan numerik dengan menggunakan algoritma Newton-Raphson. Sebagai contoh, misalkan parameter Poisson untuk respon bivariat pada model regresi ZIJBP merupakan fungsi logaritma dari kovariat x_1, x_2, x_3 dengan koefisien $\beta_1, \beta_2, \beta_3$ dan parameter *zero-inflated* ϕ berhubungan dengan 1 kovariat z_1 yaitu:

$$\lambda_1 = \exp(x_1^t \beta_1), \quad \lambda_2 = \exp(x_2^t \beta_2), \quad \lambda_3 = \exp(x_3^t \beta_3), \quad \phi = \frac{e^{z_1^t \gamma_1}}{1 + e^{z_1^t \gamma_1}}$$

Definisikan $\theta = (\beta_1, \beta_2, \beta_3, \gamma_1)^t := (\theta_1, \theta_2, \theta_3, \theta_4)^t$. Misalkan diperoleh data $(\mathbf{X}_1, \mathbf{Y}_1, \mathbf{Z}_1), \dots, (\mathbf{X}_n, \mathbf{Y}_n, \mathbf{Z}_n)$ dimana $i = 1, 2, \dots, n$ dan $\mathbf{X}_i = (X_{1i}, X_{2i}, X_{3i})$, $\mathbf{Y}_i = (Y_{1i}, Y_{2i})$ dan $\mathbf{Z}_i = (Z_{1i})$. Fungsi likelihood untuk data tersebut dapat dinyatakan sebagai:

$$L(\theta | \mathbf{X}, \mathbf{Y}, \mathbf{Z}) = \prod_{i=1}^n \left\{ G_{1i}^{1(Y_{1i}=0, Y_{2i}=0)} G_{2i}^{1(Y_{1i}=y_{1i}, Y_{2i}=0)} G_{3i}^{1(Y_{1i}=0, Y_{2i}=y_{2i})} G_{4i}^{1(Y_{1i}=y_{1i}, Y_{2i}=y_{2i})} \right\},$$

dimana:

$$G_{1i} = Pr(Y_{1i} = 0, Y_{2i} = 0), \quad G_{2i} = Pr(Y_{1i} = y_{1i}(> 0), Y_{2i} = 0), \\ G_{3i} = Pr(Y_{1i} = 0, Y_{2i} = y_{2i}(> 0)), \quad G_{4i} = Pr(Y_{1i} = y_{1i}(> 0), Y_{2i} = y_{2i}(> 0))$$

Fungsi *loglikelihood* adalah:

$$l = l(\theta | \mathbf{X}, \mathbf{Y}, \mathbf{Z}) = \sum_{i=1}^n \left\{ \mathbf{1}(Y_{1i} = 0, Y_{2i} = 0) \ln G_{1i} + \mathbf{1}(Y_{1i} = y_{1i}, Y_{2i} = 0) \ln G_{2i} \right. \\ \left. + \mathbf{1}(Y_{1i} = 0, Y_{2i} = y_{2i}) \ln G_{3i} + \mathbf{1}(Y_{1i} = y_{1i}, Y_{2i} = y_{2i}) \ln G_{4i} \right\}$$

Taksiran maksimum *likelihood* untuk θ adalah:

$$\hat{\theta} = \arg \max_{\theta} l(\theta | \mathbf{X}, \mathbf{Y}, \mathbf{Z})$$

(4)

Untuk menyelesaikan persamaan (4) digunakan algoritma Newton-Raphson dengan nilai awal $\theta^{(0)}$ dan dihitung:

$$\theta^{(t+1)} = \theta^{(t)} - \{V(\theta^{(t)})\}^{-1} S(\theta^{(t)})$$

dimana $S(\theta) = \frac{\partial l}{\partial \theta}$, $V(\theta) = \frac{\partial^2 l}{\partial \theta \partial \theta^t}$ untuk nilai t berturutan sampai konvergen.

2. Hasil – Hasil Utama

Untuk mengetahui bagaimana model regresi ZIJBP dapat menangani kasus overdispersi dan 2 variabel respon berkorelasi, berikut ini dilakukan simulasi dengan menggunakan program R untuk pembangkitan data. Data simulasi sebanyak 100 dari 1 variabel kovariat X yang dibangkitkan dari distribusi Uniform (0,1) dan 2 variabel respon Y_1 dan Y_2 dengan parameter distribusi Poisson bivariat $(\lambda_1, \lambda_2, \lambda_3)$. Korelasi antara Y_1 dan Y_2 sebesar 0.942 dan terdapat 50 % kasus *zero-inflated*. Setting parameter model yang digunakan adalah:

$$\lambda_{1i} = \exp(0.1 + 0.2 * x_i)$$

$$\lambda_{2i} = \exp(0.3 + 0.4 * x_i)$$

$$\lambda_{3i} = \exp(0.5 + 0.6 * x_i), i=1,2, \dots, 100$$

Dengan menggunakan software SAS 9.4, pada dua tabel berikut ini disajikan perbandingan hasil estimasi parameter beserta galat bakunya untuk model regresi *Jointly Bivariate Poisson* (JBP) dan Model Regresi *Zero-Inflated Jointly Bivariate Poisson* (ZIJBP).

Tabel 1: Hasil Estimasi Parameter dan Galat Baku Model Regresi JBP

Parameter	Estimasi	Galat Baku
$\beta_{01} = 0.1, \beta_{11} = 0.2$	$\widehat{\beta}_{01} = -1.1264, \widehat{\beta}_{11} = -0.00058$	$se \widehat{\beta}_{01} = 0.3967, se \widehat{\beta}_{11} = 0.6475$
$\beta_{02} = 0.3, \beta_{21} = 0.4$	$\widehat{\beta}_{02} = -1.3058, \widehat{\beta}_{21} = 0.02148$	$se \widehat{\beta}_{02} = 0.4396, se \widehat{\beta}_{21} = 0.7153$
$\beta_{03} = 0.5, \beta_{31} = 0.6$	$\widehat{\beta}_{03} = 0.2815, \widehat{\beta}_{31} = 0.2695$	$se \widehat{\beta}_{03} = 0.1765, se \widehat{\beta}_{31} = 0.2784$

* -2 Log Likelihood = 633.5, AIC = 645.5, BIC = 661.2

Tabel 2: Hasil Estimasi Parameter dan Galat Baku Model Regresi ZIJBP

Parameter	Estimasi	Galat Baku
$\beta_{01} = 0.1, \beta_{11} = 0.2$	$\widehat{\beta}_{01} = -0.2283, \widehat{\beta}_{11} = 0.03756$	$se \widehat{\beta}_{01} = 0.4560, se \widehat{\beta}_{11} = 0.7556$
$\beta_{02} = 0.3, \beta_{21} = 0.4$	$\widehat{\beta}_{02} = -0.3649, \widehat{\beta}_{21} = 0.05588$	$se \widehat{\beta}_{02} = 0.5035, se \widehat{\beta}_{21} = 0.8326$
$\beta_{03} = 0.5, \beta_{31} = 0.6$	$\widehat{\beta}_{03} = 0.8565, \widehat{\beta}_{31} = 0.3061$	$se \widehat{\beta}_{03} = 0.2063, se \widehat{\beta}_{31} = 0.3256$
$\bar{\lambda}_1 = 1.232397$	$\widehat{\lambda}_1 = 0.81212$	
$\bar{\lambda}_2 = 1.684459$	$\widehat{\lambda}_2 = 0.715472$	
$\bar{\lambda}_3 = 2.310371$	$\widehat{\lambda}_3 = 2.786251$	
	$\widehat{\phi} = 0.4828$	$se \widehat{\phi} = 0.05075$

* -2 Log Likelihood = 478.5, AIC = 492.5, BIC = 510.8

Dari Tabel 1 dan 2 terlihat bahwa AIC model regresi ZIJBP lebih kecil dibandingkan model regresi JBP.

2.1 Aplikasi Model Regresi ZIJBP

Masalah kemiskinan merupakan salah satu persoalan mendasar yang menjadi pusat perhatian pemerintah di negara manapun. Salah satu aspek penting untuk mendukung Strategi Penanggulangan Kemiskinan adalah tersedianya data kemiskinan yang akurat dan tepat sasaran. Pengukuran kemiskinan yang dapat dipercaya dapat menjadi instrumen tangguh bagi pengambil kebijakan dalam memfokuskan perhatian pada kondisi hidup orang miskin. Data kemiskinan yang baik dapat digunakan untuk mengevaluasi kebijakan pemerintah terhadap kemiskinan, membandingkan kemiskinan antar waktu dan daerah, serta menentukan target penduduk miskin dengan tujuan untuk memperbaiki kondisi mereka.

Kemiskinan merupakan ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan maupun bukan makanan yang diperoleh dari data pengeluaran perkapita. Kemiskinan berdasarkan penyebabnya terbagi menjadi dua yaitu kemiskinan kultural dan kemiskinan struktural. Kemiskinan kultural disebabkan faktor adat dan budaya, sedangkan Kemiskinan struktural yaitu sekelompok orang yang miskin karena mereka miskin. Kemiskinan secara konseptual juga terbagi dua yaitu kemiskinan relatif dan kemiskinan absolut.

Kemiskinan relatif ditetapkan secara subjektif oleh masyarakat setempat, sedangkan kemiskinan absolut berdasarkan kemampuan memenuhi kebutuhan makanan maupun bukan makanan.

Badan Pusat Statistik (BPS) pada tahun 1984 melakukan perhitungan jumlah dan persentase penduduk miskin untuk pertama kalinya. Dalam mengukur kemiskinan BPS menggunakan konsep kemampuan memenuhi kebutuhan dasar, sehingga kemiskinan dapat diartikan sebagai ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan maupun bukan makanan yang diperoleh dari data pengeluaran perkapita. Penduduk miskin adalah penduduk yang memiliki rata-rata pengeluaran perkapita per bulan dibawah garis kemiskinan (GK). Garis kemiskinan merupakan penjumlahan dari garis kemiskinan makanan (GKM) dan garis kemiskinan non makanan (GKNM). Garis kemiskinan makanan merupakan pengeluaran minimum makanan yang disetarakan dengan 2100 kilokalori perkapita per hari, sedangkan garis kemiskinan non makanan merupakan kebutuhan minimum untuk perumahan, sandang, pendidikan dan kesehatan.

BPS menggunakan konsep kemampuan memenuhi kebutuhan dasar dalam mengukur kemiskinan. Dengan pendekatan ini, kemiskinan dipandang sebagai ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur dari sisi pengeluaran. Berdasarkan hal tersebut, BPS menggunakan tiga indikator kemiskinan, yaitu:

1. *Head count index* (HCI- P_0), merupakan presentase penduduk dibawah garis kemiskinan (GK)
2. Indeks kedalaman kemiskian (*Poverty gap index*- P_1), merupakan ukuran rata-rata kesenjangan pengeluaran masing-masing penduduk miskin terhadap garis kemiskinan. Semakin tinggi nilai indeks, semakin jauh rata-rata pengeluaran penduduk dari garis kemiskinan.
3. Indeks keparahan kemiskinan (*Poverty severity index*- P_2), memberikan gambaran mengenai penyebaran pengeluaran diantara penduduk miskin

P_{aij} didefinisikan sebagai:

$$P_{aij} = \left(\frac{z - E_{ij}}{z} \right)^\alpha I(E_{ij} < z), \quad j = 1, \dots, N_i, \quad \alpha = 0,1,2$$

dengan :

$I(E_{ij} < z) = 1$ jika $E_{ij} < z$ (penduduk miskin) dan

$I(E_{ij} < z) = 0$ jika $E_{ij} \geq z$ (penduduk tidak miskin).

Untuk $\alpha = 0$ diperoleh proporsi penduduk di bawah garis kemiskian (GK) yang disebut *head count index* (HCI - P_0). Untuk $\alpha = 1$ merupakan indeks kedalaman kemiskinan (*poverty gap index* - P_1). Untuk $\alpha = 2$ merupakan indeks keparahan kemiskinan (*poverty severity index* - P_2).

Sumber data yang digunakan dalam aplikasi model regresi ZIJB ini berdasarkan Survei Sosial Ekonomi Nasional tahun 2013 (Pengeluaran per kapita per bulan) dan Potensi Desa (**Podes**) tahun 2014 di Kabupaten dan Kota Bogor. Jumlah desa tersampel adalah sebanyak 147 desa.

Variabel respon yang digunakan adalah:

Y_{1i} = Jumlah rumah tangga (ruta) di desa ke-i dengan pengeluaran per kapita per bulan berada dibawah garis kemiskinan (GK) Kabupaten dan Kota Bogor. Berdasarkan publikasi BPS [1] yaitu "DATA DAN INFORMASI

KEMISKINAN KABUPATEN/KOTA” tahun 2013, Garis Kemiskinan Kabupaten Bogor adalah Rp. 271.970,- dan Kota Bogor adalah Rp. 360.518,-

Y_{2i} = Jumlah rumah tangga (ruta) di desa ke-i yang memiliki indeks kedalaman kemiskinan diatas Indeks kedalaman kemiskinan Kabupaten dan Kota Bogor. Berdasarkan publikasi BPS tahun 2013, Indeks Kedalaman Kemiskinan (P_1) Kabupaten Bogor adalah 0.0139 sedangkan Kota Bogor adalah 0.0113.

Kovariat yang digunakan adalah:

X_{1i} = Proporsi Ruta di desa ke-i dengan sumber penghasilan utama bidang pertanian

X_{2i} = Sumber air untuk minum sebagian besar keluarga di desa ke-i

$$X_{2i} = \begin{cases} 1, & \text{air kemasan/ledeng} \\ 2, & \text{sumur/sumur bor atau pompa} \\ 3, & \text{mata air/sungai/danau/air hujan / lainnya} \end{cases}$$

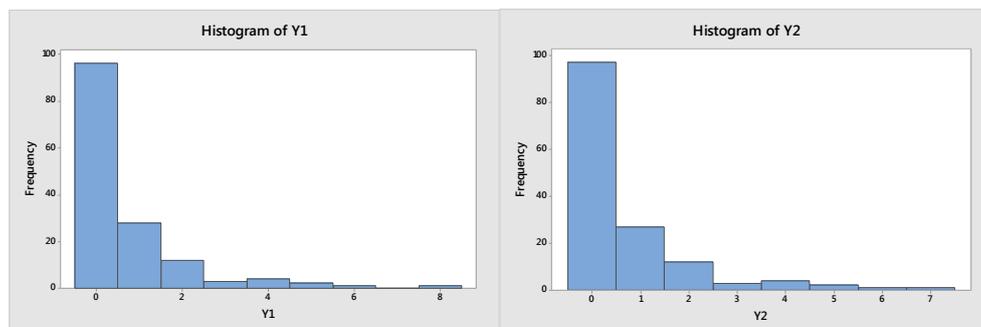
X_{3i} = Sumber air untuk mandi/cuci sebagian besar keluarga di desa ke-i

$$X_{3i} = \begin{cases} 1, & \text{ledeng dengan meteran/ledeng tanpameteran} \\ 2, & \text{sumur/sumur bor atau pompa} \\ 3, & \text{mata air/sungai/danau/air hujan/lain nya} \end{cases}$$

Berikut disajikan statistik deskriptif dan histogram untuk setiap variabel respon dan kovariat yang digunakan dalam model:

Tabel 3: Statistik Deskriptif

Variabel	N	Mean	SE		Minimum	Q1	Median	Q3	Maximum
			Mean	StDev					
Y_1	147	0.687	0.107	1.302	0	0	0	1	8
Y_2	147	0.673	0.104	1.267	0	0	0	1	7
X_1	147	0.01386	0.00187	0.02264	0	0	0	0.031	0.0986
X_2	147	1.8639	0.0489	0.5926	1	1	2	2	3
X_3	147	1.9388	0.0504	0.6107	1	2	2	2	3



Gambar 1: Histogram Y_1 dan Y_2

Dari gambar 1, terdapat pasangan nilai variabel respon (0,0) sebanyak 97 desa dari total 147 desa tersampel di Kabupaten dan Kota Bogor, artinya sebanyak 65.9864

% adalah kasus *excess zero*. Karena korelasi antara Y_1 dan Y_2 besar yaitu 0.996 dan terjadinya *excess zero*, maka akan digunakan model regresi *Zero-Inflated Jointly Bivariate Poisson* (ZIJBP). Bentuk model ZIJBP adalah:

$(y_{1i}, y_{2i}) \sim$ Distribusi ZIJBP dengan fungsi kepadatan peluang adalah:

$$f_{ZIJBP}(y_{1i}, y_{2i}) = \begin{cases} \phi + (1-\phi)f_{JBP}(0,0; \lambda_{1i}, \lambda_{2i}, \lambda_{3i}), & y_{1i} = y_{2i} = 0 \\ (1-\phi)f_{JBP}(y_{1i}, y_{2i}; \lambda_{1i}, \lambda_{2i}, \lambda_{3i}), & y_{1i} \text{ atau } y_{2i} = 0 \end{cases}$$

Untuk $i=1,2,\dots,147$

$$\lambda_{1i} = \exp(\beta_{01} + \beta_{11}x_1 + \beta_{12}x_2 + \beta_{13}x_3)$$

$$\lambda_{2i} = \exp(\beta_{02} + \beta_{21}x_1 + \beta_{22}x_2 + \beta_{23}x_3)$$

$$\lambda_{3i} = \exp(\beta_{03} + \beta_{31}x_1 + \beta_{32}x_2 + \beta_{33}x_3)$$

$$\text{logit } \phi = \gamma \text{ atau } \phi = \frac{\exp(\gamma)}{1 + \exp(\gamma)}$$

Dengan menggunakan program SAS 9.4 diperoleh hasil estimasi parameter model regresi ZIJBP sebagai berikut:

Tabel 4: Hasil Estimasi Maksimum Likelihood Parameter Model Regresi ZIJBP Dengan 3 Kovariat

Parameter	Estimasi	SE	DF	t value	Pr> t
β_{01}	-5.6581	3.1441	147	-1.80	0.0740
β_{11}	36.5960	30.2697	147	1.21	0.2286
β_{12}	2.5771	4.4690	147	0.58	0.5650
β_{13}	-2.2019	4.4521	147	-0.49	0.6216
β_{02}	-3.3455	2066.07	147	0	0.9987
β_{21}	-0.04002	110.69	147	0	0.9997
β_{22}	-4.7568	2181.08	147	0	0.9983
β_{23}	-4.8386	4238.51	147	0	0.9991
β_{03}	-1.2285	0.4984	147	-2.46	0.0149
β_{31}	0.04330	5.4246	147	0.01	0.9936
β_{32}	-0.3358	0.3151	147	-1.07	0.2884
β_{33}	1.0137	0.3168	147	3.20	0.0017
γ	-0.09325	0.3060	147	-0.3	0.7610
ϕ	0.4767	0.07634	147	6.24	<0.0001

* -2Log Likelihood = 336.9, AIC = 362.9, BIC= 401.7

Dengan menggunakan $\alpha = 0.05$, dari tabel diatas terlihat bahwa hanya parameter β_{03} dan β_{33} serta ϕ yang signifikan. Berdasarkan nilai estimasi maksimum likelihood untuk setiap parameter model, diperoleh nilai rata-rata untuk setiap estimasi parameter adalah $\bar{\lambda}_1 = 0.019778$, $\bar{\lambda}_2 = 5.06665E-07$, $\bar{\lambda}_3 = 1.241284$. Karena X_1 memiliki korelasi yang signifikan dengan Y_1 dan Y_2 , akan digunakan hanya 2 kovariat yaitu X_1 dan X_2 dengan hasil estimasi parameter model sebagai berikut:

Tabel 5: Hasil Estimasi Maksimum Likelihood Parameter Model Regresi ZIJBP Dengan 2 Kovariat

Parameter	Estimasi	SE	DF	t value	Pr> t
β_{01}	-6.1845	3.0891	147	-2.00	0.0471
β_{11}	31.0646	28.7568	147	1.08	0.2818
β_{12}	0.6771	1.4398	147	0.47	0.6389
β_{02}	-9.4418	2568.48	147	0.00	0.0071
β_{21}	-0.07577	50.7838	147	0.00	0.9988
β_{22}	-10.5470	2574.46	147	0.00	0.9967
β_{03}	-0.6432	0.4407	147	-1.46	0.1465
β_{31}	3.7002	5.4644	147	0.68	0.4994
β_{32}	0.4243	0.2222	147	1.95	0.0537
γ	0.01159	0.2897	147	0.04	0.9681
ϕ	0.5029	0.07241	147	6.95	<0.0001

* -2Log Likelihood = 346.1, AIC = 366.1, BIC = 396.0

Dengan menggunakan $\alpha = 0.05$, dari tabel diatas terlihat bahwa β_{01}, β_{02} yang signifikan. Berdasarkan nilai estimasi maksimum likelihood untuk setiap parameter model, diperoleh nilai rata-rata untuk setiap estimasi parameter adalah $\widehat{\lambda}_1 = 0.020638$, $\widehat{\lambda}_2 = 5.24528E-10$, $\widehat{\lambda}_3 = 1.299387$. Dengan menggunakan moment distribusi ZIJBP pada persamaan (3), dapat dihitung rata-rata Y_1 adalah taksiran korelasi antara Y_1 dan Y_2 sehingga rata-rata taksiran korelasi antara Y_1 dan Y_2 adalah 0.996596. Hasil ini sangat mirip dengan nilai korelasi sampel antara Y_1 dan Y_2 pada tabel 4. Namun demikian, tidak terdapat satupun variabel kovariat signifikan yang mempengaruhi jumlah ruta miskin di desa Kabupaten dan Kota Bogor. Untuk itu perlu dicari lagi kovariat yang berpengaruh terhadap jumlah rumah tangga miskin di desa Kabupaten dan Kota Bogor sebagai indikator kemiskinan.

3. Kesimpulan

Model regresi *Zero-Inflated Jointly Bivariate Poisson* (ZIJBP) merupakan model regresi yang dapat digunakan untuk respon berupa data count bivariat berkorelasi dan terjadi over dispersi karena banyaknya kasus (0,0) atau terjadi *excess zero*. Variabel respon bivariat memiliki distribusi ZIJBP dengan 3 parameter dan setiap parameter bisa tergantung pada satu atau lebih kovariat melalui fungsi penghubung kanonik log. Untuk mencari estimasi parameter model regresi ZIBP dapat digunakan metode maksimum likelihood dan melalui bantuan algoritma Newton-Raphson. Dalam aplikasi model regresi ZIJBP pada data Podes 2014 untuk rumah tangga di Kabupaten dan Kota Bogor, hasil analisa model menunjukkan bahwa sumber penghasilan pertanian dan air yang digunakan bukanlah kovariat signifikan yang mempengaruhi banyaknya ruta miskin di desa Kabupaten dan Kota Bogor. Untuk itu perlu dicari lagi kovariat yang berpengaruh

terhadap jumlah rumah tangga miskin di desa Kabupaten dan Kota Bogor sebagai indikator kemiskinan.

Referensi

- [1] Badan Pusat Statistik. 2013. Data dan informasi kemiskinan kabupaten/kota tahun 2013. Jakarta (ID): BPS.
- [2] AlMuhayfith, E. Fatimah, Abdulhamid A. Alzaid, Maha A. Omair, 2016. *On bivariate Poisson regression models*, Journal of King Saud University – Science (2016) 28, 178–189
- [3] Johnson, N.L., Kotz, S., Balakrishnan, N., 1997. *Discrete Multivariate Distributions*. John Wiley & Sons, New York
- [4] Karlis, D., Ntzoufras, I., 2005. *Bivariate Poisson and diagonal inflated bivariate Poisson regression models in R*. J. Stat. Softw. 14, 1–

DAMPAK PENGETAHUAN, PENYULUHAN, DAN KEPEDULIAN LINGKUNGAN TERHADAP PERILAKU PENGELOLAAN SAMPAH DENGAN MENGGUNAKAN REGRESI LOGISTIK BINER

SUGIARTO

Jurusan Statistika, Sekolah Tinggi Ilmu Statistik (STIS)- Jakarta, soegie@stis.ac.id

Abstrak. Penelitian mengenai perilaku masyarakat dalam pengelolaan sampah masih jarang ditemui. Padahal masalah sampah merupakan masalah bersama. Perlakuan yang kurang baik terhadap sampah akan menimbulkan kerusakan lingkungan di sekitar kita. Program penyuluhan/pelatihan pengelolaan sampah sedang digalakkan pemerintah namun dampaknya belum terlihat hal ini ditunjukkan dengan volume sampah yang semakin meningkat setiap harinya. Selain itu tingkat kepedulian terhadap lingkungan di masyarakat sejauh ini masih kurang. Sehingga penelitian mengenai bagaimana perilaku pengelolaan sampah di masyarakat dan dampak pengetahuan, penyuluhan/pelatihan dan kepedulian terhadap lingkungan terhadap pengelolaan sampah rumah tangga sangat menarik untuk diteliti. Data yang digunakan dalam penelitian ini berasal dari Survei Perilaku Peduli Lingkungan Hidup (SPPLH) yang diselenggarakan oleh BPS tahun 2013. Metode analisis yang digunakan adalah metode regresi logistik biner, hal tersebut sesuai dengan karakteristik variabel tidak bebas yaitu pengelolaan sampah dengan 2 kategori mengelola sampah dan tidak mengelola sampah. Sedangkan variabel bebas yang diteliti keterkaitannya dengan perilaku pengelolaan sampah adalah variabel pengetahuan, keikutsertaan dalam pelatihan/penyuluhan lingkungan dan kepedulian terhadap pengelolaan lingkungan. Hasil penelitian menunjukkan bahwa pengetahuan, keikutsertaan penyuluhan/pelatihan pengelolaan sampah dan kepedulian terhadap lingkungan, berpengaruh terhadap pengelolaan sampah rumah tangga.

Kata kunci : Pengelolaan Sampah, Regresi Logistik Biner.

1. Pendahuluan

Masalah sampah selalu dihadapi oleh manusia karena manusia menghasilkan sampah setiap hari. Bertambahnya jumlah penduduk, perubahan pola konsumsi, dan meningkatnya aktivitas kegiatan ekonomi meningkatkan jumlah sampah yang dihasilkan setiap harinya. Sehingga apabila masalah sampah tidak segera diatasi, volumenya akan terus bertambah setiap harinya. Selain semakin besarnya volume sampah, saat ini permasalahan sampah menjadi semakin rumit karena manajemen pengelolaan sampah yang tidak baik. Mulai dari permasalahan tempat penampungan sementara, pengumpulan, pengangkutan sampai pada tempat pembuangan akhir dan pengolahan. Permasalahan sampah yang dihadapi dapat dikatakan sebagai masalah kultural, tidak hanya masalah sosial maupun ekonomi karena dampaknya terkena pada berbagai sisi kehidupan [1]. Meningkatnya

sampah dapat juga menimbulkan banyak masalah lain yang akan muncul misalnya masalah pencemaran lingkungan, masalah penyakit, dan masalah sosial lainnya. Berdasarkan data dari Dinas Kebersihan seluruh Indonesia yang dirangkum dalam publikasi Statistik Lingkungan Hidup Indonesia 2013 dan 2014 tentang produksi dan volume sampah yang terangkut di kota-kota besar, menunjukkan bahwa sampah yang dihasilkan oleh rumah tangga belum bisa terangkut sepenuhnya oleh petugas kebersihan [2,3]. Untuk itu peningkatan jumlah sampah harus disertai dengan pengelolaan sampah yang baik dan kontinu yang dimulai dari skala rumah tangga, dimana rumah tangga merupakan sumber penghasil sampah yang utama.

Di negara maju sosialisasi mengenai pentingnya pengelolaan sampah mulai dilakukan kepada penduduknya, terutama pada rumah tangga, untuk mengelola sampah yang dihasilkan mulai dari melakukan pemilahan sampah sampai dengan melibatkan penduduk untuk melakukan daur ulang sampah, misalnya menjadi pakan ternak atau pupuk kompos. Di Indonesia program-program pengelolaan sampah juga sudah mulai digalakan oleh pemerintah, hal ini terbukti dengan diwujudkannya Undang-Undang Nomor 18 Tahun 2008 tentang pengelolaan sampah, namun demikian masih belum terlihat dampaknya. Hal ini terlihat dari data yang menunjukkan persentase rumah tangga yang tidak pernah melakukan pengelolaan sampah cukup besar yaitu 73,16 persen [2,3].

Tabel 1. Persentase Rumah Tangga dengan Perlakuan yang Paling Sering Dilakukan Terhadap Sampah

Perlakuan Terhadap Sampah	2013	2014
(1)	(2)	(3)
Didaur ulang	0.25	0.19
Dibuat kompos	0.89	0.53
Diangkut petugas/ dibuang ke TPS/TPA	25.61	25.35
Dijual ke pengumpul barang bekas	0.46	0.5
Dijadikan makanan ternak	0.18	0.26
Ditimbun/ dikubur	4.61	3.58
Dibakar	53.06	54.65
Dibuang ke laut/ sungai/ got	7.82	5.54
Dibuang sembarangan	7.12	9.39

Sumber: BPS, SPPLH

Tabel 1. memberikan informasi bahwa, terjadi penurunan rumah tangga yang melakukan daur ulang sampah, pengelolaan sampah menjadi kompos dan terjadi peningkatan rumah tangga yang membakar sampah dan membuang sampah sembarangan. Bila dilihat dari data-data ini terlihat bahwa peran serta rumah tangga dalam pengelolaan sampah masih kurang. Untuk itu diperlukan peran serta dari seluruh pihak, tidak hanya pemerintah namun juga keterlibatan masyarakat dalam mengatasi masalah sampah dengan cara menggiatkan program pengelolaan sampah pada masyarakat.

Dalam upaya meningkatkan kesadaran pengelolaan sampah, pemerintah telah melakukan berbagai program penyuluhan/pelatihan dalam pengelolaan sampah, serta memberikan peningkatan wawasan pengetahuan mengenai kepedulian lingkungan melalui iklan layanan masyarakat serta broket dan brosur, sehingga diharapkan muncul rasa kepedulian terhadap lingkungan. Namun sampai saat ini masih sedikit penelitian yang mengukur apakah pengetahuan masyarakat terhadap lingkungan, program penyuluhan/pelatihan dan kepedulian masyarakat terhadap lingkungan berpengaruh terhadap cara pengelolaan sampah pada

masyarakat atau tidak. Berdasarkan penjelasan diatas, penelitian merasa tertarik mengenai bagaimana perilaku pengelolaan sampah di masyarakat dan bagaimana dampak pengetahuan, penyuluhan/pelatihan dan kepedulian terhadap lingkungan terhadap pengelolaan sampah rumah tangga .

2. Hasil-Hasil Utama

2.1. Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang dikumpulkan dari Survei Perilaku Terhadap Lingkungan Hidup yang diselenggarakan oleh BPS tahun 2013. Sampel dalam penelitian ini adalah rumah tangga (ruta) sebanyak 70.406 yang tersebar di seluruh wilayah Indonesia. Untuk mengukur pengetahuan perilaku terhadap lingkungan, keikutsertaan penyuluhan/pelatihan, dan kepedulian lingkungan peneliti melakukan pendekatan rumah tangga, dimana rumah tangga diwakilkan oleh Kepala Rumah Tangga (KRT). Pendekatan KRT dengan alasan karena KRT dianggap merupakan orang yang bertanggung jawab dan mempunyai peran penting untuk memutuskan segala sesuatu di rumah tangga.

2.2. Variabel Penelitian

Variabel tidak bebas dalam penelitian ini bersifat kategorik yaitu melakukan pengelolaan sampah (kode 1) dan tidak melakukan pengelolaan sampah (kode 0). Sedangkan untuk variabel bebas meliputi variabel pengetahuan, keikutsertaan penyuluhan dan kepedulian terhadap lingkungan. Adapun kategori variabel bebas sebagai berikut:

Tabel 2. Variabel-variabel bebas dalam penelitian

No	Nama Variabel	Kategori	Variabel Dummy	Dummy
(1)	(2)	(3)	(4)	(5)
1	Pengetahuan	Skala rasio	X_1	
2	Penyuluhan	1. Tidak Ikut Penyuluhan/Pelatihan Pengelolaan lingkungan/sampah	D_2	0
		2. Ikut Penyuluhan/Pelatihan Pengelolaan lingkungan/sampah		1
3	Kepedulian	1. Tidak Peduli Terhadap Lingkungan	D_3	0
		2. Peduli Terhadap Lingkungan		1

Definisi variabel:

1. **Pengelolaan Sampah** adalah mencakup antara lain perlakuan KRT terhadap sampah rumah tangga, seperti 1. didaur ulang, 2. dibuat kompos, 3. diangkut dibuang ke TPA/TPS, 4. di jual ke pengumpul barang bekas, 5. dijadikan makanan ternak, 6. Ditimbun/dikubur, 7. Dibakar, 8. dibuang kelaut/sungai/got, 9. dibuang sembarangan. Poin 1-5 oleh peneliti dianggap rumah tangga melakukan pengelolaan, sedangkan poin 6-9 tidak melakukan pengelolaan.

2. **Pengetahuan** adalah pengetahuan perilaku peduli lingkungan KRT, mencakup antara lain pengetahuan responden akan pencemaran udara, pemanasan suhu bumi, pemborosan air, penyediaan area resapan air, penghematan listrik dan bahan bakar, perlakuan terhadap sampah dan sampah yang mengandung bahan kimia, dan sumber informasi pengetahuan yang dimiliki responden.
3. **Penyuluhan** adalah keikutsertaan KRT dalam penyuluhan/pelatihan pengelolaan sampah atau lingkungan
4. **Kepedulian** adalah mencakup antara lain keikutsertaan KRT pada kegiatan kerja bakti di lingkungan rumah, keberadaan pencemaran dan jenisnya, dan upaya menanggulangi pencemaran.

2.3. Metode Analisis

Metode analisis yang digunakan dalam penelitian ini adalah analisis deskriptif untuk melihat gambaran pengelolaan sampah pada masyarakat dan inferensia dengan menggunakan Regresi Logistik Biner untuk mengetahui dampak pengetahuan, keikutsertaan penyuluhan dan kepedulian lingkungan terhadap perilaku pengelolaan sampah rumah tangga.

Regresi Logistik Biner

Metode regresi merupakan suatu cara untuk menggambarkan hubungan antara satu variabel tidak bebas/respon dengan satu atau lebih variabel bebas/penjelas. Dalam beberapa kasus sering dijumpai variabel tidak bebas diskrit dengan dua nilai atau variabel biner. Dalam kasus seperti ini, metode analisis yang dapat digunakan adalah regresi logistik. Perbedaan pokok antara logistik dengan regresi linier adalah pada variabel tidak bebasnya yang bersifat biner [4].

Dalam regresi terdapat suatu ukuran rata-rata nilai dari dependen variabel, dengan syarat diketahui nilai dari variabel independen. Ukuran ini disebut *conditional mean* dengan notasi “ $E(Y|x)$ ”, dimana Y menyatakan variabel tidak bebas dan x menyatakan nilai spesifik dari variabel bebas. Pada data dikotomi nilai $E(Y|x)$ akan bernilai lebih besar dari atau sama dengan nol dan kurang dari atau sama dengan satu ($0 \leq E(Y|x) \leq 1$). pada distribusi logistik, $\pi(x) = E(Y|x)$ digunakan untuk merepresentasikan *conditional mean* dari Y dengan syarat x. Bentuk spesifik dari model regresi logistik adalah sebagai berikut :

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p)}{1 + \exp(\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p)} \quad (1)$$

dengan menerapkan transformasi logit, maka dapat dinyatakan sebagai berikut:

$$\begin{aligned} g(x) &= \ln \pi(x) - \ln[1 - \pi(x)] \\ &= \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p \end{aligned} \quad (2)$$

yang merupakan fungsi linier dalam parameter-parameternya.

Pada regresi logistik, peubah respon biasanya dinotasikan sebagai $y = E(Y/x) + \varepsilon$, dimana ε merupakan salah satu dari dua kemungkinan nilai, yaitu:

$$\varepsilon = \begin{cases} 1 - \pi(x), & \text{jika } y=1 \text{ dengan peluang } \pi(x) \\ - \pi(x) & , \text{ jika } y=0 \text{ dengan peluang } 1 - \pi(x) \end{cases}$$

dimana ε mengikuti distribusi binomial dengan rata-rata nol dalam ragam $\pi(x)[1 - \pi(x)]$ (Hosmer [4]).

Jika beberapa dari variabel bebas bersifat diskrit dan berskala nominal seperti tingkat pendidikan, jenis kelamin dan lainnya, maka variabel ini sebaiknya dimasukkan dalam model sebagai *dummy*. Apabila terdapat p peubah bebas dan peubah ke- j merupakan peubah kategori dengan k nilai, maka diperlukan variabel *dummy* sebanyak $k-1$. Misalkan peubah bebas ke- j , yaitu x_j mempunyai k_j tingkatan, maka variabel *dummy* k_{j-1} dinotasikan D_{ju} dengan koefisien β_{ju} dimana $u = 1, 2, \dots, k_{j-1}$. Maka model transformasi logitnya menjadi :

$$g(x) = \beta_0 + \beta_1 X_1 + \dots + \sum_{u=1}^{k_{j-1}} \beta_{ju} D_{ju} + \dots + \beta_p X_p \quad (3)$$

a. Pendugaan Parameter

Dalam model regresi logistik, nilai harapan peubah respon tidak linier terhadap parameter β , maka penduga β diperoleh melalui metode *maximum Likelihood*.

b. Pengujian Parameter

Uji Simultan

Untuk menguji signifikansi model maka digunakan uji nisbah kemungkinan (Uji Simultan), dimana :

H_0 : $\beta_1 = \beta_2 = \dots = \beta_p = 0$

H_1 : minimal ada satu $\beta_j \neq 0$; (minimal ada satu variabel bebas yang memengaruhi pengelolaan sampah rumah tangga dimana $j = 1, 2, \dots, p$)

Statistik ujinya adalah :

$$G = -2 \ln [L_0 - L_k] \quad (4)$$

Dimana $L_0 =$ Likelihood tanpa peubah penjelas

$L_k =$ Likelihood dengan peubah penjelas

Statistik G ini mengikuti sebaran *Chi-Square* dengan derajat bebas p , dimana H_0 akan ditolak jika nilai $G > \chi^2_{(\alpha;p)}$ atau *p-value* < tingkat signifikan (α). Jika H_0 ditolak berarti minimal terdapat satu variabel bebas yang memengaruhi variabel tidak bebas.

Uji Parsial

Untuk melihat β_j mana yang signifikan dapat digunakan Uji Parsial dengan menggunakan Uji Wald. Adapun Hipotesisnya sebagai berikut:

H_0 : $\beta_j = 0$ (tidak ada pengaruh variabel bebas ke- j terhadap pengelolaan sampah rumah tangga); $j = 1, 2, \dots, p$

H_0 : $\beta_j \neq 0$ (ada pengaruh variabel bebas ke- j terhadap pengelolaan sampah rumah tangga); $j = 1, 2, \dots, p$

Statistik ujinya adalah :

$$W = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)} \quad (5)$$

Dimana : $\hat{\beta}_j$ = penduga parameter β_j ; $se(\hat{\beta}_j)$ = penduga galat baku dari β_j

W diasumsikan mengikuti sebaran normal baku, H_0 akan ditolak jika $W < Z_{\alpha/2}$ atau $W > Z_{1-\alpha/2}$.

Nilai statistik uji G maupun uji W ini umumnya diberikan dalam output paket program komputer lengkap dengan *standar error*nya. Pada uji W jika H_0 ditolak, berarti variabel bebas (X) mempengaruhi variabel tidak bebas (Y). Tetapi hal ini juga dapat dilihat dari nilai signifikansinya (*p-value*). Jika *p-value* < tingkat signifikan (α), maka variabel bebas mempengaruhi variabel tidak bebas .

Uji Goodness of fit

Untuk mengetahui apakah model yang dihasilkan telah sesuai dalam menjelaskan variabel tidak bebas digunakan Uji *Goodness of fit* [4]. Pengujian ini dapat dilakukan dengan menggunakan statistik uji *Hosmer and Lemeshow test*.

Hipotesis yang digunakan yaitu:

H_0 : Model sesuai (tidak terdapat perbedaan antara hasil observasi dan hasil prediksi dari model).

H_1 : Model tidak sesuai (terdapat perbedaan antara hasil observasi dan hasil prediksi dari model).

Statistik uji yang digunakan adalah sebagai berikut:

$$\hat{C} = \sum_{k=1}^g \frac{(o_k - n_k \bar{\pi}_k)^2}{n_k \bar{\pi}_k (1 - \bar{\pi}_k)} \quad (6)$$

Keterangan:

n_k = jumlah subjek pada grup ke-k

g = jumlah grup

$o_k = \sum_{j=1}^{c_k} y_j$ = jumlah nilai variabel tidak bebas; c_k = banyaknya kombinasi variabel bebas pada grup ke-k

$\bar{\pi}_k = \sum_{j=1}^{c_k} \frac{m_j \hat{\pi}_j}{n_k}$ = rata-rata dari estimasi probabilitas; m_j = jumlah subjek dengan c_k kombinasi variabel bebas.

Statistik \hat{C} mengikuti distribusi Khi-Kuadrat (χ^2) dengan derajat bebas $g-2$. Hipotesis nol akan ditolak $\hat{C} > \chi^2_{(\alpha, g-2)}$ atau *p-value* < α yang artinya model tidak sesuai atau tidak dapat menjelaskan variabel tidak bebas.

c. Odds Ratio (Rasio Kecenderungan)

Odds Ratio adalah suatu ukuran yang menunjukkan rasio untuk mengalami suatu kejadian tertentu antara suatu bagian populasi dengan ciri tertentu dan bagian populasi yang lain yang tidak memiliki ciri tertentu tersebut [4].

Perbandingan tingkat resiko antara dua nilai peubah penjas, misal $x = a$ dan $x = b$ berdasarkan regresi logistik adalah :

$$\begin{aligned} \hat{\theta} &= \frac{\hat{\pi}_b (1 - \hat{\pi}_a)}{(1 - \hat{\pi}_b) \hat{\pi}_a} \\ \ln \hat{\theta} &= \log it(\hat{\pi}_b) - \log it(\hat{\pi}_a) \\ &= \hat{\beta}_j (b - a) \end{aligned}$$

Jika $b-a=1$ maka pendugaan *odds ratio* dapat diperoleh dari nilai koefisien β untuk setiap kenaikan satu satuan X_j dengan asumsi X yang lain tetap, sehingga :

$$\hat{\theta}_j = \exp(\hat{\beta}_j) \quad (7)$$

Dengan selang kepercayaan sebesar 100 (1- α) % untuk θ adalah

$$\exp[\hat{\beta}_j \pm Z_{1-\alpha/2} .se(\hat{\beta}_j)] \quad (8)$$

2.4 Hasil dan Pembahasan

2.4.1 Gambaran Pengelolaan Sampah di Indonesia

Data BPS dari hasil SPPLH 2013 menunjukkan bahwa sebagian besar rumah tangga mengelola sampahnya dengan beberapa cara pembuangan sampah (lebih dari satu jenis perlakuan) yaitu sebesar 61,41 persen dan sisanya mengelola sampah hanya dengan satu jenis perlakuan saja. Cara pembuangan sampah yang banyak dilakukan oleh rumah tangga adalah dengan cara dibakar (69,92 persen) dan dibuang ke TPA (27,98 persen). Di daerah pedesaan lebih parah lagi persentase rumah tangga yang membakar sampah mencapai 87,36 persen. Angka ini menunjukkan bahwa perilaku masyarakat Indonesia dalam mengelola sampah masih kurang ramah terhadap lingkungan.

Jika dilihat pengelolaan sampah menurut provinsi, maka persentase tertinggi rumah tangga yang membakar sampah terdapat di Provinsi Aceh yaitu sebesar 86,07 persen, menyusul Provinsi Gorontalo (86,07 persen), Provinsi Nusa Tenggara Timur (84,04 persen). Sementara provinsi dengan persentase terkecil rumah tangga yang membakar sampah tercatat di Provinsi DKI Jakarta yaitu sebesar 12,65 persen. Hal ini menunjukkan di DKI Jakarta masyarakatnya sudah cenderung peduli terhadap lingkungan. Selain itu sistem pembuangan melalui TPA juga cukup baik dibandingkan provinsi lainnya.

Perlakuan sampah yang juga kurang baik terhadap lingkungan lainnya adalah menimbun/mengubur sampah, membuang sampah ke got/sungai/laut, dan membuang sampah sembarangan. Perlakuan menimbun/mengubur sampah ternyata masih cukup banyak dilakukan oleh rumah tangga yaitu mencapai sekitar 22,29 persen. Sedangkan rumah tangga yang membuang sampah ke saluran got/sungai/laut tercatat sebesar 14,16 persen dan rumah tangga yang membuang sampah secara sembarangan sebesar 19,88 persen. Perlakuan terhadap sampah seperti ini bisa mengakibatkan kondisi lingkungan di sekitar tempat tinggal yang semakin buruk, karena bisa menimbulkan polusi udara seperti bau, juga dapat menimbulkan banjir akibat tertutupnya saluran got atau aliran sungai.

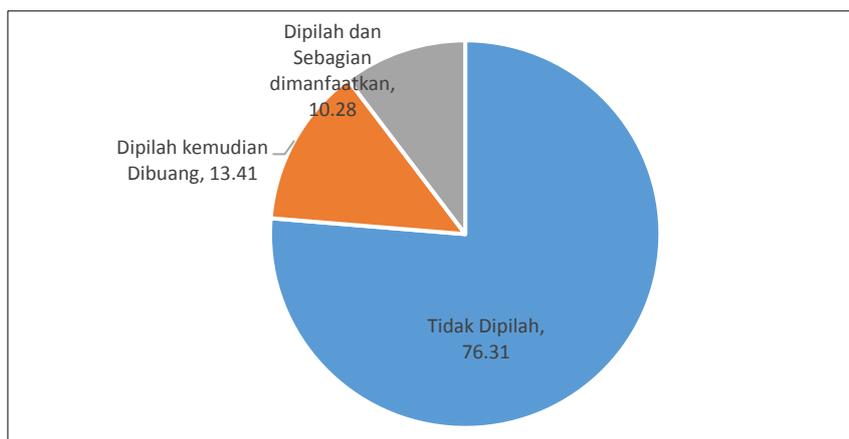
Data survei SPPLH 2013 menunjukkan bahwa kesadaran masyarakat untuk memanfaatkan sampah masih kecil hal ini terlihat dari persentase rumah tangga yang melakukan daur ulang sampah dan membuat pupuk atau kompos dari sampah masih sedikit yaitu berturut-turut sebesar 1,76 persen dan 5,56 persen. Sementara rumah tangga yang memanfaatkan kembali sampah yang dihasilkan sebagai makanan ternak tercatat ada sekitar 9,78 persen.

Tabel 3. Persentase Rumah Tangga Menurut Perlakuan Yang Pernah Dilakukan dan Perlakuan Utama Terhadap Sampah

Perlakuan Terhadap Sampah	Pernah Dilakukan	Utama
(1)	(2)	(3)
Didaur ulang	1.76	0.25
Dibuat kompos	5.56	0.89
Diangkut petugas/ dibuang ke TPS/TPA	27.98	25.61
Dijual ke pengumpul barang bekas	26.10	0.46
Dijadikan makanan ternak	22.29	0.18
Ditimbun/ dikubur	69.92	4.61
Dibakar	14.16	53.06
Dibuang ke laut/ sungai/ got	19.88	7.82
Dibuang sembarangan	9.78	7.12

Sumber: BPS, SPPLH

Jika dilihat menurut perlakuan utama terhadap sampah, maka tercatat bahwa membuang sampah dengan cara dibakar merupakan perlakuan utama terhadap sampah yang paling banyak dilakukan oleh rumah tangga yaitu mencapai 53,06 persen, diikuti cara membuang sampah dengan cara diangkut petugas/dibuang ke TPS/TPA (25,61 persen), dan dibuang ke got/sungai/laut sebesar 7,82 persen. Sebagian besar provinsi di Indonesia terlihat bahwa persentase terbesar cara membuang sampah yang utama oleh rumah tangga adalah dengan cara dibakar, kecuali untuk Provinsi Kepulauan Riau, DKI Jakarta, Bali, dan Kalimantan Timur, persentase terbesar cara membuang sampah adalah dengan cara diangkut petugas/dibuang ke TPS/TPA, dan untuk Provinsi Maluku dan Maluku Utara, persentase terbesar cara membuang sampah adalah dengan cara dibuang ke got/sungai/laut.



Sumber: BPS, SPPLH 2013

Grafik 1. Persentase Rumah Tangga Menurut Pemilahan Sampah

Pemilahan sampah antara sampah organik dan non organik sangat bermanfaat, karena hal ini akan memudahkan proses pengolahan sampah pada tahapan berikutnya. Data hasil SPPLH 2013 menunjukkan bahwa kesadaran masyarakat untuk memilah sampah rumah tangga masih sangat rendah, hal ini

ditunjukkan dengan tingginya persentase rumah tangga yang menyatakan tidak memilah sampah yaitu sebesar 76,31 persen.

Dilihat menurut alasan sebagian besar rumah tangga tidak melakukan pemilahan dikarenakan malas sebagai alasan utama sebesar 40,56 persen, hal ini dimungkinkan karena rumah tangga merasa capek atau tidak ada waktu untuk melakukan pemilahan sampah. Selanjutnya sebesar 33,33 persen rumah tangga menyatakan tidak melakukan pemilahan sampah karena tidak mengetahui bahwa sampah harus dipilah sebelum dibuang. Sementara rumah tangga yang tidak melakukan pemilahan sampah karena alasan tidak ada fasilitas, tidak menguntungkan, dan tidak ada peraturan berturut-turut sebesar 11,79 persen; 10,40 persen; dan 3,92 persen. Masih kurangnya kesadaran masyarakat untuk melakukan pemilahan sampah diharapkan menjadi tantangan bagi pemerintah untuk terus menggalakkan program pemisahan sampah, sehingga partisipasi masyarakat menjadi lebih meningkat. Selain itu pemberian edukasi kepada petugas sampah juga perlu ditingkatkan agar sampah yang sudah dipilah di rumah tangga tidak tercampur lagi ketika berada di dalam TPA.

2.4.2 Dampak/Pengaruh variabel Penyuluhan, Pengetahuan Perilaku Peduli lingkungan, dan Kepedulian Lingkungan terhadap Pengelolaan Sampah Rumah Tangga

Untuk mengetahui apakah variabel pengetahuan perilaku peduli lingkungan peduli lingkungan, keikutsertaan penyuluhan, dan kepedulian lingkungan mempengaruhi pengelolaan sampah rumah tangga digunakan analisis regresi logistik biner. Untuk mendapatkan model dan variabel apa saja yang signifikan maka akan dilakukan pengujian parameter secara simultan, uji parsial, pengujian kecocokan model (*goodness of fit*). Proses pengolahan data untuk mendapatkan model regresi logistik menggunakan software SPSS 21 dengan metode Enter. Berdasarkan hasil uji simultan dengan menggunakan uji Rasio *Likelihood*, diperoleh hasil sebagai berikut.

Tabel 4. Rasio *Likelihood* dan *omnibus test of model coefficients* untuk uji simultan

	<i>-2 Log likelihood</i>	<i>Chi-Square</i>	<i>df</i>	<i>p-value</i>
(1)	(2)	(3)	(4)	(5)
Model	94180,138	3361,490	3	0,000

Berdasarkan hasil uji simultan dengan tingkat signifikan 5% menggunakan uji Rasio *Likelihood* dan *omnibus test of model coefficients* diperoleh nilai statistik G sebesar 3361.490 dengan *p-value* sebesar 0,000 dan hasilnya menolak H_0 . Dengan demikian hasil uji ini menunjukkan bahwa dengan tingkat kepercayaan 95%, dapat dikatakan minimal terdapat satu variabel bebas yang berpengaruh terhadap pengelolaan sampah rumah tangga. Setelah pengujian variabel secara simultan menyatakan bahwa minimal ada satu variabel yang signifikan maka dilakukan pengujian variabel secara parsial untuk mengetahui variabel bebas mana saja yang ternyata memengaruhi pengelolaan sampah rumah tangga. Berdasarkan hasil uji parsial dengan tingkat signifikan 5% menggunakan uji *Wald*, diperoleh hasil sebagai berikut.

Tabel 5. Penduga parameter, statistik uji *wald* dan *odds ratio* variabel tidak bebas

Variabel	$\hat{\beta}$	Se	Wald	df	p-value	Exp($\hat{\beta}$)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
PENGETAHUAN	0,141	0,003	1838,401	1	0,000*	1,151
PENYULUHAN	0,525	0,044	142,306	1	0,000*	1,690
KEPEDULIAN	0,469	0,023	430,820	1	0,000*	1,598
Constant	-1,399	0,028	2434,389	1	0,000*	0,247

Keterangan: * Signifikan di $\alpha = 5\%$,

Berdasarkan Tabel 5, terlihat bahwa dari tiga variabel bebas ternyata semua signifikan memengaruhi pengelolaan sampah rumah tangga pada taraf signifikansi 5 persen. Hal ini dapat dilihat dari nilai *p-value* yang lebih kecil dari tingkat signifikansi (α). Adapun persamaan peluang regresi logistik yang terbentuk sebagai berikut:

$$\hat{\pi}(D) = \frac{\exp(-1,399 + 0,141X_1 + 0,525D_2 + 0,469D_3)}{1 + \exp(-1,399 + 0,141X_1 + 0,525D_2 + 0,469D_3)}$$

Persamaan peluang tersebut adalah persamaan nonlinier, agar diperoleh persamaan linier maka dilakukan transformasi logit. Bentuk persamaan transformasi logit yang diperoleh sebagai berikut:

$$\hat{g}(D) = -1,399 + 0,141X_1^* + 0,525D_2^* + 0,469D_3^*$$

Keterangan:

*) signifikan pada $\alpha = 0,05$

Nilai intersep sebesar -1,399 menunjukkan bahwa KRT yang mempunyai karakteristik pengetahuan terhadap lingkungan yang rendah, tidak pernah ikut serta dalam penyuluhan/pelatihan, dan tidak peduli terhadap lingkungan mempunyai peluang untuk melakukan pengelolaan sampah sebesar 0,197. Hal ini menunjukkan bahwa KRT yang mempunyai karakteristik pengetahuan terhadap lingkungan yang rendah, tidak pernah ikut serta dalam penyuluhan/pelatihan, dan tidak peduli terhadap lingkungan memiliki peluang yang sangat kecil untuk melakukan pengelolaan sampah rumah tangga.

Tahap selanjutnya dilakukan pengujian *goodness of fit* untuk mengetahui apakah model yang terbentuk telah sesuai dalam menjelaskan pengelolaan sampah rumah tangga. Hasil pengujian dapat dilihat pada tabel *Hosmer and lemeshow test* berikut ini:

Tabel 6. *Hosmer and lemeshow test*

Nilai Khi-Kuadrat	df	p-value
(1)	(2)	(3)
220,536	7	0,000

Berdasarkan Tabel 6, dapat diketahui bahwa nilai statistik uji *Hosmer and lemeshow* sebesar 220,536 dengan *p-value* sebesar 0,000. Karena nilai *p-value* lebih kecil dari 0,05, maka tolak H_0 . Dengan demikian, dapat disimpulkan bahwa dengan tingkat kepercayaan sebesar 95 persen, model yang digunakan tidak sesuai dalam menjelaskan pengelolaan sampah pada rumah tangga. Walaupun dalam uji

goodness of fit ini tidak sesuai namun menurut agresti bahwa uji ini bisa diabaikan asal uji model secara simultan dan parsial telah memenuhi [5]. Sebagai penguat bahwa model telah cukup baik untuk melakukan prediksi bisa digunakan tabel klasifikasi model. Dimana bila ketepatan pengklasifikasian model lebih dari 50 persen sudah dianggap cukup baik. Berdasarkan tabel klasifikasi model terlihat bahwa nilai *overall* persentase sebesar 59,7. Angka ini menunjukkan bahwa secara keseluruhan model mempunyai tingkat kesesuaian sebesar 59,7 persen dalam memprediksi perilaku pengelolaan sampah rumah tangga. Artinya model sudah cukup baik.

Tabel 7. Tabel Klasifikasi Model

<i>Observed</i>		<i>Predicted</i>			
		<i>Pengelolaan Sampah</i>		<i>Percentage</i>	
		<i>Tidak Dikelola</i>	<i>Dikelola</i>	<i>Correct</i>	
(1)		(2)	(3)	(4)	
<i>Step 1</i>	Pengelolaan	Tidak Dikelola	15259	18901	44,7
	Sampah	Dikelola	9463	26783	73,9
	<i>Overall Percentage</i>				59,7

a. *The cut value is ,500*

2.4.3 Diskusi dan Pembahasan

Berdasarkan persamaan regresi logistik yang terbentuk dapat diketahui seberapa besar dampak/pengaruh dan kecenderungan dari variabel bebas yang berpengaruh signifikan terhadap pengelolaan sampah rumah tangga. Nilai kecenderungan dapat dilihat dari nilai $\exp(\beta)$ pada Tabel 5. Semua variabel bebas signifikan memengaruhi variabel respon.

Persamaan regresi logistik menunjukkan bahwa variabel pengetahuan, penyuluhan, dan kepedulian berpengaruh secara positif terhadap perilaku pengelolaan sampah pada rumah tangga. Dalam keadaan semua variabel bebas lainnya tetap, variabel pengetahuan memiliki pengaruh sebesar 0,141 atau KRT yang memiliki pengetahuan lebih tinggi memiliki kecenderungan 1,15 kali lebih besar dibandingkan KRT yang memiliki pengetahuan lebih rendah. Kemudian KRT yang pernah mengikuti penyuluhan/pelatihan pengelolaan sampah atau lingkungan memiliki kecenderungan 1,69 kali lebih besar untuk mengelola sampah dibandingkan KRT yang tidak pernah mengikuti penyuluhan/pelatihan pengelolaan sampah atau lingkungan. Sedangkan KRT yang memiliki rasa kepedulian terhadap lingkungan memiliki kecenderungan 1,59 kali lebih besar untuk mengelola sampah dibandingkan KRT yang tidak memiliki rasa kepedulian terhadap lingkungan. Dari hasil pembahasan diatas terbukti bahwa keikutsertaan KRT mengikuti penyuluhan/pelatihan pengolahan sampah atau lingkungan mempunyai dampak yang cukup besar yaitu hampir dua kali lipat dibandingkan KRT yang tidak pernah mengikuti penyuluhan/pelatihan untuk melakukan pengolahan sampah rumah tangga. Begitu juga dengan kepedulian KRT terhadap lingkungan mempunyai dampak yang cukup besar yaitu hampir satu setengah kali lipat dibandingkan KRT yang tidak peduli terhadap lingkungan untuk melakukan pengolahan sampah rumah tangga. Sedangkan pengetahuan KRT terhadap lingkungan karena nilainya hanya sebesar 1,1 kali atau secara statistik dianggap dampaknya kurang begitu signifikan.

3. Kesimpulan

Pengelolaan sampah rumah tangga di Indonesia masih sangat kurang hal ini terlihat dari masih banyaknya masyarakat yang melakukan pembakaran sampah rumah tangga dalam pengelolaan sampahnya dan masih sedikitnya rumah tangga yang melakukan daur ulang sampah, pemanfaatan sampah menjadi pupuk kompos atau pakan ternak.

Keikutsertaan KRT mengikuti penyuluhan/pelatihan pengolahan sampah atau lingkungan dan kepedulian KRT terhadap lingkungan memiliki dampak yang cukup signifikan untuk rumah tangga melakukan pengelolaan sampah rumah tangga dimana kecenderungannya hampir dua kali lipat dibandingkan kondisi sebaliknya. Dari hasil ini diharapkan bahwa program-program penyuluhan dan pelatihan yang sudah dilakukan selama ini hendaknya terus dilanjutkan dan ditingkatkan lagi.

Referensi

- [1] Azwar, Azrul. (1995). *Pengantar Ilmu Kesehatan Lingkungan*. Jakarta: Mutiara Sumber Widya.
- [2] Badan Pusat Statistik, *Indikator Perilaku Peduli Lingkungan Hidup*, Jakarta, 2013
- [3] Badan Pusat Statistik, *Indikator Perilaku Peduli Lingkungan Hidup*, Jakarta, 2014
- [4] Hosmer, D.W. dan Lemeshow, S. (2000). *Applied Logistic Regression Second Edition*. New York: John Wiley and Sons, Inc.
- [5] Agresti, A., & Kateri, M. (2011). *Categorical data analysis*. Springer Berlin Heidelberg.

POLA PEROKOK USIA 10-24 DI INDONESIA TAHUN 2015

DIN NURIKA AGUSTINA

Badan Pusat Statistik, Jl. dr. Sutomo N0.6-8, Jakarta 10710
din_nurika@bps.go.id

Abstrak. Hampir semua orang mengetahui bahwa perilaku merokok memiliki pengaruh negartif, utamanya pada kondisi ekonomi dan kesehatan. Namun demikian jumlah perokok di Indonesia masih tergolong tinggi. Menurut WHO, lebih dari sepertiga (36,1 persen) penduduk usia lebih dari 15 tahun di Indonesia adalah perokok, dengan rata-rata pengeluaran per bulan untuk rokok kretek mereka adalah sekitar 369.948 rupiah, fakta lainnya adalah hampir tiga dari lima orang pelajar terpapar asap rokok di rumah mereka. Pentingnya berhenti merokok di usia tua dimaksudkan menghentikan pengaruh lebih parah dari racun rokok yang telah terakumulasi di dalam tubuh dan sangat penting untuk menghindari resiko menjadi perokok pasif bagi orang di sekitar perokok aktif yang bisa jadi adalah ibu hamil dan anak-anak. Hal yang tak kalah penting adalah mencegah dan berhenti merokok di usia muda, agar racun rokok tidak mengancam bonus demografi Indonesia di tahun 2020-2030, dengan rusaknya derajat kesehatan pada usia produktif di masa bonus demografi tersebut. Penelitian ini bertujuan mengeksplorasi pola perokok usia 10-24 tahun ditinjau dari berbagai faktor internal responden yaitu jenis kelamin, umur, pendidikan, status perkawinan, dan status ekonomi serta faktor eksternal responden yaitu tipe daerah tempat tinggal dan karakteristik kepala rumah tangga responden (status merokok, pendidikan, dan perkawinan). Data yang digunakan pada penelitian ini adalah data Survei Sosial Ekonomi Nasional (Susenas) tahun 2015. Meskipun Susenas bukan merupakan survei yang didesain khusus sebagai penelitian terhadap perilaku perokok, namun besarnya ukuran contoh, adanya pertanyaan tentang perilaku merokok dan banyaknya peubah yang dapat digali untuk mengeksplorasi pola perokok, membuat penulis menentukan Susenas sebagai data penelitian. Selanjutnya metode analisis yang digunakan pada penelitian ini adalah regresi logistik biner, karena tersedianya peubah respon yang berkategori biner dan berbagai kemudahan yang dimiliki oleh metode analisis ini guna melihat pola perokok usia 10-24 tahun di Indonesia tahun 2015. Hasil dari penelitian ini adalah diperolehnya pola berbagai peubah internal maupun eksternal pada perilaku perokok. Informasi tentang pola tersebut diharapkan dapat menjadi pertimbangan bagi pihak-pihak yang bermaksud melakukan berbagai intervensi terhadap perokok usia 10-24 tahun, sehingga intervensi tersebut dapat dilakukan secara lebih efisien dan tepat sasaran untuk menghentikan perilaku merokok.

Kata kunci : perokok, usia 10-24 tahun, regresi logistik biner, Susenas 2015.

1. Pendahuluan

Berbagai survey menunjukkan bahwa jumlah perokok di Indonesia masih cukup tinggi. *Global Adults Tobacco Survey* (GATS) [7] melaporkan bahwa lebih dari sepertiga (36,1 persen) penduduk usia lebih dari 15 tahun di Indonesia adalah perokok, dengan rata-rata pengeluaran per bulan untuk rokok kretek mereka sekitar 369.948 rupiah. Selanjutnya menurut hasil *Global Youth Tobacco Survey* (GYTS) [6] sekitar 20,3 persen pelajar usia 13-15 tahun di Indonesia menggunakan tembakau baik berupa rokok maupun bentuk lain, dan hampir tiga dari lima pelajar (57,3 persen) terpapar asap rokok di rumahnya.

Hampir semua orang mengetahui bahwa perilaku merokok mengakibatkan kerugian utamanya ekonomi dan kesehatan. Bahkan peraturan pemerintah no. 109 tahun 2012 [5] yang salah satunya mengatur kewajiban dicantumkan gambar dan tulisan peringatan konsekuensi buruk bagi kesehatan akibat perilaku merokok pada kemasan rokok, semakin mensosialisasikan dampak buruk rokok pada kesehatan terutama bagi perokok aktif. Namun demikian masih tingginya jumlah perokok di Indonesia mungkin dikarenakan dampak kerugian kesehatan akibat merokok tidak dialami secara langsung sesaat setelah merokok namun cenderung merupakan akumulasi dari kandungan zat beracun dalam tubuh perokok. Ditambah adanya zat adiktif nikotin membuat kecanduan merokok, sehingga konsekuensi buruknya kesehatan tidak menjadi pertimbangan.

Berhenti merokok di usia tua dimaksudkan untuk menghentikan pengaruh lebih parah dari racun rokok yang telah terakumulasi di tubuh dan sangat penting untuk menghindari resiko menjadi perokok pasif bagi orang di sekitar perokok aktif yang bisa jadi adalah wanita hamil dan anak-anak. Penggunaan tembakau tersebut oleh perokok aktif maupun oleh perokok pasif telah terbukti menimbulkan dampak buruk terkait berbagai masalah kesehatan yang serius sebagaimana hasil penelitian *U.S. Department of Health and Human Services* [4]. Diperlukan berbagai intervensi terhadap perokok utamanya perokok usia muda untuk menghentikan perilaku merokok tersebut.

Hal yang tak kalah penting adalah mencegah dan berhenti merokok di usia muda, agar racun rokok tidak mengancam bonus demografi Indonesia di tahun 2020-2030. Bonus demografi adalah suatu fenomena dimana struktur penduduk sangat menguntungkan dari sisi pembangunan karena jumlah penduduk usia produktif sangat besar, sedangkan usia non produktif (usia muda dan usia lanjut) tidak banyak. Berhenti merokok di usia muda tentanya sangat dibutuhkan mengingat kualitas usia produktif dari sisi kesehatan dapat dirugikan oleh racun rokok.

Tulisan ini bertujuan mengeksplorasi pola berbagai faktor internal dan eksternal pada perokok usia muda yaitu 10-24 tahun. Faktor internal responden yang dieksplorasi pada penelitian ini adalah jenis kelamin, umur, pendidikan, status perkawinan, status ekonomi dan faktor eksternal responden adalah tipe daerah tempat tinggal dan status kepala rumah tangga responden (status merokok, pendidikan, dan perkawinan). Informasi pola tersebut diharapkan dapat

menjadi pertimbangan sehingga berbagai intervensi terhadap perokok usia 10-24 tahun dapat dilakukan secara efisien dan tepat sasaran.

Sistematika tulisan ini terbagi atas tiga bagian. Bagian pertama berisi pendahuluan meliputi latar belakang, tujuan, dan sistematika penulisan. Bagian ke dua adalah metodologi yang terdiri dari data, peubah dan metode analisis. Bagian ke tiga adalah hasil dan pembahasan yang meliputi signifikansi model, interpretasi model serta kesimpulan.

2. Metodologi

2.1. Data

Data yang digunakan pada penelitian ini adalah data Survei Sosial Ekonomi Nasional (Susenas) yang dilakukan oleh Badan Pusat Statistik (BPS) tahun 2015 di bulan Maret menggunakan kuesioner Kor (VSEN2015.K). Ukuran contoh Susenas tersebut adalah 300.000 rumah tangga menghasilkan data representatif estimasi tingkat kabupaten/kota. Data Susenas dipilih data individu dengan batasan usia 10-24 tahun, dan dilakukan penggabungan data kepala rumah tangga untuk mendapatkan karakteristik kepala rumah tangga individu tersebut. Meskipun Susenas tidak didesain khusus sebagai alat penelitian perilaku perokok, namun besarnya ukuran contoh, adanya pertanyaan tentang perilaku merokok dan banyaknya peubah yang dapat dieksplorasi, membuat penulis memilih Susenas sebagai data penelitian.

2.2. Peubah

Menurut penelitian Strauss J et.al. [3] beberapa faktor yang memiliki pengaruh terhadap partisipasi merokok adalah jenis kelamin, umur, pengeluaran per kapita dan pendidikan orang tua. Berdasarkan penelitian lintas negara yang dilakukan oleh Agaku et.al. [1] disimpulkan bahwa wilayah, jenis kelamin, penghasilan dan umur memberikan perbedaan yang signifikan dalam prevalensi merokok. Mempertimbangkan hasil penelitian-penelitian tersebut dan ketersediaan peubah pada Susenas 2015 maka ditentukan faktor-faktor yang dieksplorasi yaitu faktor internal responden: jenis kelamin, umur, pendidikan, status perkawinan, status ekonomi dan faktor eksternal responden: tipe daerah tempat tinggal dan status kepala rumah tangga responden (status merokok, pendidikan, dan perkawinan). Faktor-faktor yang diteliti pada tulisan ini merupakan peubah-peubah dengan karakteristik sebagaimana dijelaskan pada Tabel.1.

Tabel 1. Peubah yang Dieksplorasi Pada Pola Perokok Usia 10-24 Tahun di Indonesia, 2015.

Jenis Peubah	Nama Peubah	Tipe Data	Keterangan
(1)	(2)	(3)	(4)
Tak bebas	Status merokok	Kategorik	1: Merokok 0: Tidak merokok
Bebas	Tipe daerah	Kategorik	1: Perdesaan 0: Perkotaan
	Jenis kelamin	Kategorik	1: Laki-laki 0: Perempuan

Kelompok umur	Kategorik	3: 20-24 1: 14-16	2: 17-19 0: 10-13
Status sekolah	Kategorik	1: Tidak sekolah 0: Masih sekolah	
Ijazah	Kategorik	0: <= SMP 1: >= SMU	
Status perkawinan	Kategorik	1: Pernah kawin 0: Belum kawin	
Status bekerja	Kategorik	1: Bekerja 0: Tidak bekerja	

Tabel 1. Lanjutan.

Jenis Peubah	Nama Peubah	Tipe Data	Keterangan
(1)	(2)	(3)	(4)
Bebas	Pengeluaran per kapita	Kategorik	2: 20% pengeluaran per kapita tertinggi 1: 40% persen pengeluaran per kapita menengah 0: 40% persen pengeluaran per kapita terendah
	Status perkawinan kepala rumah tangga (KRT)	Kategorik	1: Kawin 0: Tidak kawin
	Jenis kelamin KRT	Kategorik	1: Laki-laki 0: Perempuan
	Ijazah KRT	Kategorik	0: <= SMP 1: >= SMU
	Status merokok KRT	Kategorik	1: Merokok 0: Tidak merokok

2.3. Metode Analisis

Metode analisis yang digunakan pada penelitian ini adalah analisis deskriptif dan analisis regresi logistik biner. Analisis deskriptif digunakan untuk melihat gambaran karakteristik status merokok usia 10-24 tahun. Sedangkan regresi logistik biner dipilih untuk memberikan informasi terkait pola peubah bebas yang dieksplorasi pada penelitian ini terhadap peubah tak bebasnya sebagaimana telah diuraikan pada Tabel 1.

Menurut Agresti A.[2] model yang paling populer untuk data biner adalah regresi logistik. Model regresi dengan bentuk linier untuk transformasi logit peluang bagi peubah tak bebas ($\hat{\pi}(x)$) adalah :

$$\text{logit} [\hat{\pi}(x)] = \ln \left(\frac{\pi(x)}{1 - \pi(x)} \right) = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p$$

Jika beberapa dari variabel penjelas bersifat kategorik, maka variabel tersebut masuk dalam model sebagai peubah boneka (*dummy*). Sebagai contoh apabila terdapat sejumlah p peubah bebas, kemudian peubah bebas ke-j (yaitu X_j)

merupakan peubah kategorik yang memiliki nilai sejumlah k, maka diperlukan peubah boneka sebanyak k-1 untuk peubah bebas ke-j tersebut. Selanjutnya peubah boneka dinotasikan dengan D_{ju} dengan parameter model β_{ju} dimana $u = 1, 2, \dots, k-1$. Model regresi dengan bentuk linear untuk transformasi logit menjadi :

$$g(x) = \beta_0 + \beta_1 X_1 + \dots + \sum_{u=1}^{k_j-1} \beta_{ju} D_{ju} + \dots + \beta_p X_p$$

Selanjutnya dalam model regresi logistik tersebut, penduga parameter model (β_j) diperoleh melalui metode Maximum Likelihood.

Untuk menguji signifikansi model maka digunakan uji nisbah kemungkinan (Uji Simultan), dengan hipotesis sebagai berikut:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0; \quad H_1 : \text{Minimal ada satu } \beta_j \neq 0; \\ j = 1, 2, \dots, p.$$

Statistik ujinya adalah :

$$G = -2 \text{Ln} [L_0 - L_k] \quad \text{dimana}$$

L_0 = Likelihood tanpa peubah penjelas; L_k = Likelihood dengan peubah penjelas

Statistik G ini mengikuti sebaran *Chi-Square* dengan derajat bebas p, dimana tolak H_0 jika nilai $G > \chi^2_{(p; \alpha)}$. Jika H_0 ditolak berarti minimal ada satu $\beta_j \neq 0$.

Untuk melihat β_j mana yang signifikan dapat digunakan uji parsial, salah satunya adalah Uji Wald. Adapun Hipotesisnya sebagai berikut:

$$H_0: \beta_j = 0 \quad (\text{tidak ada pengaruh peubah bebas ke-j terhadap peubah tak bebas}); \\ H_0: \beta_j \neq 0 \quad (\text{ada pengaruh peubah bebas ke-j terhadap peubah tak bebas}); \\ j = 1, 2, \dots, p.$$

Statistik ujinya adalah :

$$W = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)}$$

dimana: $\hat{\beta}_j$ = penduga parameter β_j ; $se(\hat{\beta}_j)$ = galat baku dari penduga β_j

W diasumsikan mengikuti sebaran normal baku, H_0 ditolak jika $W < Z_{\alpha/2}$ atau $W > Z_{1-\alpha/2}$ atau $p\text{-value} < \text{tingkat signifikan } (\alpha)$, maka peubah bebas ke-j memiliki pengaruh terhadap peubah tak bebas.

3. Hasil dan Pembahasan

3.1. Signifikansi Model

Untuk mengetahui signifikansi persamaan peluang yang terbentuk, dilakukan uji simultan menggunakan uji Rasio *Likelihood* dengan hasil sebagaimana Tabel 2.:

Tabel 2. Hasil *Omnibus Test Of Model Coefficients* untuk Uji Simultan.

Uji Simultan	Khi-Kuadrat	<i>p-value</i>
(1)	(2)	(3)

Model	25.989.520,92	0,000
-------	---------------	-------

Hasil uji simultan dengan tingkat signifikansi 5% menggunakan uji Rasio *Likelihood* pada *omnibus test of model coefficients* diperoleh nilai statistik G sebesar 25.989.520,92 dengan *p-value* sebesar 0,000. Dengan demikian dapat disimpulkan bahwa H0 ditolak. Dengan kata lain pada tingkat kepercayaan 95%, dapat dibuktikan bahwa minimal terdapat satu peubah bebas yang berpengaruh terhadap status merokok pada usia 10-24 tahun.

Hasil uji simultan meyakinkan untuk pelaksanaan uji lanjutan yaitu uji parsial. Hasil uji parsial dengan tingkat signifikansi 5% menggunakan uji *Wald*, sebagaimana dapat dilihat pada Tabel 3. menyatakan bahwa seluruh peubah bebas baik yang merupakan peubah internal maupun peubah eksternal responden yang telah ditentukan sebagai peubah yang akan dieksplorasi pada tulisan ini (seperti yang telah dijabarkan pada Tabel 1.) seluruhnya memiliki pengaruh yang signifikan terhadap status merokok pada usia 10-24 tahun. Signifikansi pengaruh tersebut dapat dilihat dari besaran *p-value* yang lebih kecil dari tingkat signifikansi model yang telah ditentukan sebelumnya yaitu 5%.

Persamaan peluang regresi logistik dibentuk sesuai perhitungan hasil penduga parameter. Dikarenakan keseluruhan peubah yang digunakan pada penelitian ini merupakan peubah dengan tipe data kategorik, maka seluruh peubah bebas berupa peubah boneka (*dummy*). Berdasarkan Tabel 3. persamaan peluang regresi logistik yang terbentuk adalah:

$$\hat{\pi}(D) = \frac{\exp\left(\begin{matrix} -10,207-0,005D_1+5,092D_2+2,547D_{31}+3,786D_{32}+4,354D_{33}-0,39D_4+1,184D_5 \\ -0,53D_6+0,841D_7+0,441D_{81}+0,608D_{82}-0,539D_9-0,416D_{10}-0,238D_{11}+1,018D_{12} \end{matrix}\right)}{1 + \exp\left(\begin{matrix} -10,207-0,005D_1+5,092D_2+2,547D_{31}+3,786D_{32}+4,354D_{33}-0,39D_4+1,184D_5 \\ -0,53D_6+0,841D_7+0,441D_{81}+0,608D_{82}-0,539D_9-0,416D_{10}-0,238D_{11}+1,018D_{12} \end{matrix}\right)} \dots\dots\dots(1)$$

Persamaan peluang pada persamaan (1) merupakan persamaan non linier, Transformasi logit dilakukan sehingga diperoleh persamaan linier. Persamaan transformasi logit yang memiliki bentuk linear adalah sebagai berikut:

$$\hat{g}(D) = -10,207 - 0,005D_1 + 5,092D_2 + 2,547D_{31} + 3,786D_{32} + 4,354D_{33} - 0,39D_4 + 1,184D_5 - 0,53D_6 + 0,841D_7 + 0,441D_{81} + 0,608D_{82} - 0,539D_9 - 0,416D_{10} - 0,238D_{11} + 1,018D_{12} \dots\dots\dots(2)$$

Tabel 3. Penduga Parameter, Galat Baku, Statistik Uji *Wald*, Signifikansi dan *Odds Ratio* Variabel Tak Bebas.

	Peubah Bebas	$\hat{\beta}$	S.E.	Wald	Sig.	Exp($\hat{\beta}$)
	(1)	(2)	(3)	(4)	(6)	(7)
D ₁	Tipe daerah (perkotaan)	-0,005	0,001	22,57	0,00	0,99
D ₂	Jenis kelamin (laki-laki)	5,092	0,003	2.749.059,92	0,00	162,65
D ₃	Kelompok umur			1.243.639,42	0,00	
D ₃₁	Kelompok umur (14-16)	2,547	0,005	229.044,83	0,00	12,77

D ₃₂	Kelompok umur (17-19)	3,786	0,005	522.885,13	0,00	44,07
D ₃₃	Kelompok umur (20-24)	4,354	0,005	671.551,99	0,00	77,81
D ₄	Ijazah (>=SMU)	-0,390	0,001	91.955,01	0,00	0,68
D ₅	Status sekolah (tidak sekolah)	1,184	0,001	672.308,93	0,00	3,27
D ₆	Status perkawinan (kawin)	-0,530	0,002	94.367,99	0,00	0,59
D ₇	Status bekerja (bekerja)	0,841	0,001	461.265,78	0,00	2,32
D ₈₁	Pengeluaran per kapita			185.560,61	0,00	
D ₈₁	Pengeluaran per kapita (40 persen pengeluaran menengah)	0,441	0,001	132.458,32	0,00	1,55
D ₈₂	Pengeluaran per kapita (20 persen pengeluaran tertinggi)	0,608	0,002	142.232,29	0,00	1,84
D ₉	Status perkawinan KRT (kawin)	-0,539	0,002	75.175,26	0,00	0,58
D ₁₀	Jenis kelamin KRT (laki-laki)	-0,416	0,002	31.081,74	0,00	0,66
D ₁₁	Ijazah KRT (>=SMU)	-0,238	0,001	38.229,74	0,00	0,79
D ₁₂	Status merokok KRT (merokok)	1,018	0,001	64.7036,66	0,00	2,77
	Constant	-10,207	0,006	2.724.969,28	0,00	0,00

Ukuran kebaikan model pada persamaan yang telah terbentuk, dapat ditinjau dari kemampuan prediksi model tersebut terhadap data observasi. Sebagaimana dapat dilihat pada Tabel 4, dari persamaan peluang (2) yang dibentuk, dengan menggunakan titik potong 0,300, secara keseluruhan ketepatan prediksi yang dimiliki oleh model ini adalah sebesar 89,94 persen. Jika dilakukan pengamatan lebih lanjut berdasarkan status merokok pada data observasi dan prediksi status merokok yang dihasilkan persamaan tersebut, maka ketepatan prediksi yang dimiliki oleh persamaan yang terbentuk adalah sebesar 91,80 persen untuk memprediksi status tidak merokok dan sekitar 77,92 persen untuk memprediksi status merokok.

Tabel 4. Data Observasi dan Prediksi Status Merokok.

Status Merokok (data observasi)	Prediksi Status Merokok		Persentase Ketepatan Prediksi
	Tidak Merokok	Merokok	
(1)	(2)	(3)	(4)
Tidak merokok	52.615.901	4.453.487	92.2

Merokok	1.457.977	6.719.793	82.2
Persentase Ketepatan Keseluruhan			90.9

Keterangan: Menggunakan titik potong 0,300.

3.2. Interpretasi Model

Arah pengaruh dari peubah bebas terhadap peubah tak bebas dapat ditinjau dari nilai positif atau negatif penduga parameter peubah bebas. Peubah bebas yang memiliki nilai penduga parameter positif dapat diinterpretasikan memberikan andil dalam meningkatkan peluang peubah tak bebas, demikian sebaliknya dengan penduga parameter yang bernilai negatif.

Terdapat beberapa peubah bebas yang memiliki nilai penduga parameter positif yang berarti berpengaruh terhadap meningkatnya peluang peubah tak bebas yaitu perilaku merokok pada usia 10-24 tahun. Peubah bebas tersebut adalah peubah jenis kelamin (laki-laki), kelompok umur (14-16), kelompok umur (17-19), kelompok umur (20-24), status sekolah (tidak sekolah), status bekerja (bekerja), pengeluaran per kapita (40 persen pengeluaran menengah), pengeluaran per kapita (20 persen pengeluaran tertinggi), dan status merokok KRT (merokok).

Peubah bebas yang memiliki nilai penduga parameter negatif berarti berpengaruh terhadap menurunnya peluang peubah tak bebas perilaku merokok pada usia 10-24 tahun. Peubah bebas tersebut adalah tipe daerah (perkotaan), ijazah (\geq SMU), status perkawinan (kawin), status perkawinan KRT (kawin), jenis kelamin KRT (laki-laki), dan ijazah KRT (\geq SMU).

Asumsi yang digunakan untuk menginterpretasikan pengaruh suatu peubah bebas terhadap peubah tak bebas adalah nilai peubah bebas lain tidak mengalami perubahan atau dengan kata lain peubah bebas lainnya tetap. Salah satu contoh interpretasi pada peubah bebas dengan nilai penduga parameter positif adalah peubah status sekolah (tidak sekolah). Seseorang dengan status tidak sekolah akan meningkatkan peluang untuk memiliki perilaku merokok 3 (tiga) kali lipat dibandingkan seseorang dengan status sekolah. Contoh interpretasi peubah bebas dengan penduga parameter bernilai negatif adalah peubah ijazah KRT (\geq SMU). Seseorang anak yang memiliki kepala rumah tangga berijazah \geq SMU akan menurunkan peluang memiliki perilaku merokok 0,788 kali lipat dibandingkan seorang anak dengan kepala rumah tangga berijazah \leq SMP.

3.3. Kesimpulan

Berbagai informasi tentang perokok usia 10-24 tahun dapat diperoleh dari tulisan ini. Salah satunya adalah rekomendasi terhadap prioritas intervensi yang bertujuan mencegah dan menghentikan perilaku merokok. Berdasarkan hasil penelitian ini, intervensi penting diberikan secara lebih intensif pada laki-laki khususnya kelompok usia 20-24 tahun, dengan status sekolah yaitu tidak sekolah, memiliki pekerjaan, tinggal di kelompok rumah tangga dengan pengeluaran per kapita 40 persen pengeluaran menengah dan pengeluaran per kapita 20 persen pengeluaran tertinggi, tingkat pendidikan rendah ($<$ SMU), dan berstatus belum kawin.

Karakteristik eksternal yang perlu diperhatikan untuk suksesnya intervensi guna mencegah dan menghentikan perilaku merokok adalah terkait dengan karakteristik kepala rumah tangga. Intervensi penting diberikan secara lebih intensif pada usia 10-24 tahun dengan karakteristik kepala rumah tangga (KRT) sebagai perokok aktif, status perkawinan KRT belum kawin atau pernah kawin (cerai hidup, cerai mati), jenis kelamin KRT perempuan, dan ijazah KRT rendah (<SMU).

Referensi

- [1] Agaku I.T., Yusuf O.A., Vardavas C.I., Connolly G. 2013. *Predictors and Patterns of Cigarette and Smokeless Tobacco Use Among Adolescents in 32 Countries, 2007-2011*. Diakses melalui www.jahonline.org pada bulan Februari 2017.
- [2] Agresti A. 2007. *An Introduction to Categorical Data Analysis*. New Jersey: JohnWiley & Sons, Inc.
- [3] Strauss J, Witoelar F, Rukumnuaykit P. 2006. *Smoking Behavior among Youth in a Developing Country: The Case of Indonesia*. California (LA): Westin Bonaventure.
- [4] U.S. Department of Health and Human Services. 2012. *Preventing Tobacco Use Among Youth and Young Adults: A Report of the Surgeon General*. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.
- [5] Peraturan Pemerintah Republik Indonesia Nomor 109 tahun 2012. Pengamanan Bahan yang Mengandung Zat Adiktif Berupa Produk Tembakau Bagi Kesehatan. 24 Desember 2012. Lembaran Negara Republik Indonesia Tahun 2012 Nomor 278. Jakarta.
- [6] World Health Organization, Regional Office for South-East Asia. 2015. *Global Youth Tobacco Survey (GYTS): Indonesia report, 2014*. New Delhi: WHO-SEARO.
- [7] World Health Organization Regional Office for South-East Asia, Ministry of Health, Badan Pusat Statistik (BPS-Statistics Indonesia), National Institute of Health Research and Development, Centers for Disease Control and Prevention CDC Foundation, John Hopkins Bloomberg School of Public Health, RTI International. 2015. *Global Adult Tobacco Survey: Indonesia Report 2011*. Indonesia: WHO-SEARO.

DETERMINAN PENGETAHUAN PRIA KAWIN TERHADAP TANDA BAHAYA PADA MASA KEHAMILAN DI INDONESIA TAHUN 2012

ANITA DESMARINI¹, ITA WULANDARI²

1 Badan Pusat Statistik, 12.7027@stis.ac.id

2 Sekolah Tinggi Ilmu Statistik, ita.wulandari@stis.ac.id

Abstrak. Pengetahuan tanda bahaya pada masa kehamilan yang dimiliki oleh pria kawin merupakan hal yang paling mendasar dan diperlukan dalam mengupayakan penurunan kematian maternal. Penelitian ini bertujuan untuk mengetahui karakteristik pria kawin dalam kaitannya dengan pengetahuan terhadap tanda bahaya pada masa kehamilan dan untuk mengetahui variabel-variabel yang signifikan beserta kecenderungannya dalam mempengaruhi pengetahuan pria kawin. Data yang digunakan adalah *raw* data SDKI 2012 dengan unit analisis pria kawin yang memiliki anak terakhir berusia nol hingga dua tahun. Metode analisis yang digunakan adalah metode analisis regresi logistik biner. Hasil menunjukkan bahwa status ekonomi, pendidikan, banyaknya anak lahir hisap (ALH), dan konsultasi ke petugas mempengaruhi pengetahuan pria kawin. Semakin tinggi status ekonomi maupun tingkat pendidikan, semakin tinggi pula kecenderungan pria kawin untuk memiliki pengetahuan mengenai tanda bahaya pada masa kehamilan. Pria kawin yang memiliki lebih dari dua ALH maupun yang melakukan konsultasi ke petugas kesehatan, cenderung memiliki pengetahuan dibandingkan dengan pria kawin yang memiliki maksimal dua ALH, maupun pria kawin yang tidak melakukan konsultasi.

Kata kunci : regresi logistik biner, obstetri, kehamilan, pengetahuan, tanda bahaya, pria kawin

1. Pendahuluan

Kematian maternal merupakan salah satu gambaran mengenai kualitas kesehatan suatu negara. Perhatian terhadap kematian maternal sangat penting. Hal ini sesuai dengan *International Covenant on Economic, Social, and Cultur Rights* tahun 1966 pasal 12.1 yang menyebutkan bahwa kesehatan, termasuk kesehatan reproduksi dan seksualitas, merupakan hal yang penting dalam pengembangan potensi manusia serta pembangunan dan diakui sebagai hak asasi yang wajib untuk dipenuhi.[1] Ukuran kematian maternal yang seringkali digunakan untuk membandingkan besarnya kematian maternal antar negara ialah rasio kematian maternal. Kematian maternal selalu menjadi perhatian karena kematian maternal merupakan suatu indikator kesehatan ibu dan anak. Menurut data dari WHO, Indonesia merupakan salah satu negara penyumbang kematian maternal tertinggi di dunia.[2] Dibandingkan dengan negara-negara lain yang berada di kawasan Asia, Indonesia menduduki urutan keempat tertinggi setelah Timor Leste, Bangladesh

dan Kamboja di tahun 2010. Sedangkan pada tahun 2015, WHO menyebutkan bahwa di kawasan Asia tenggara, Indonesia menduduki urutan keempat sebagai negara dengan rasio kematian maternal tertinggi setelah Laos, Myanmar dan Kamboja.[3]

Rasio kematian maternal di Indonesia telah menunjukkan penurunan sejak tahun 1997 hingga tahun 2007. Namun pada tahun 2012, rasio kematian maternal di Indonesia menunjukkan peningkatan menjadi 359 kematian maternal per 100.000 kelahiran hidup. Rasio tersebut masih jauh dari target rasio maternal pada MDGs (*Millenium Development Goals*) yang harus dicapai pada tahun 2015. Berakhirnya MDGs pada tahun 2015, dilanjutkan dengan SDGs (*Sustainable Development Goals*) yang merupakan kelanjutan dari program MDGs dengan ditambahkan beberapa target yang belum tercakup dalam program sebelumnya. SDGs mencantumkan target kematian maternal yang harus dicapai Indonesia, yaitu 70 kematian maternal per 100.000 kelahiran hidup. Hal ini menunjukkan bahwa diperlukan kebijakan yang lebih mendalam untuk menurunkan rasio kematian maternal di Indonesia.

Salah satu penyumbang kematian maternal adalah komplikasi yang terjadi pada masa kehamilan. Menurut Kementerian Kesehatan RI, penyumbang kematian maternal di Indonesia banyak disebabkan oleh perdarahan, infeksi, dan keracunan kehamilan (eklampsia). Komplikasi pada masa kehamilan yang paling banyak menyebabkan kematian maternal merupakan akibat dari kurangnya pengetahuan mengenai tanda bahaya tersebut. Padahal kebanyakan kematian maternal yang terjadi masih dapat dicegah.[4] Pengetahuan ini penting untuk dimiliki bukan hanya oleh wanita saja, melainkan juga oleh pria kawin sebagai pasangan yang berperan dalam mengambil keputusan yang tepat ketika tanda bahaya komplikasi masa kehamilan tersebut muncul. Ketika pengetahuan mengenai tanda bahaya masa kehamilan dapat diketahui, maka tindakan yang tepat dapat dilakukan dengan segera. Pengetahuan tersebut penting supaya keluarga dapat segera membawa ibu untuk mendapatkan pertolongan di fasilitas kesehatan terdekat secara tepat waktu. *Women Research Institute* menyebutkan bahwa keterlambatan dalam mengenali tanda-tanda bahaya dan membuat keputusan merupakan salah satu penyebab kematian ibu.

Terdapat “tiga terlambat” yang menyebabkan tingginya kematian maternal, salah satunya ialah terlambat mengenali tanda bahaya dan mengambil keputusan. Padahal kematian maternal disebabkan oleh berbagai hal yang sebenarnya bisa dicegah. Pengenalan tanda bahaya sebagai upaya pencegahan kematian maternal sangat perlu dilakukan oleh keluarga, terutama pria kawin sebagai pasangan yang berperan dalam pengambilan keputusan.[5]

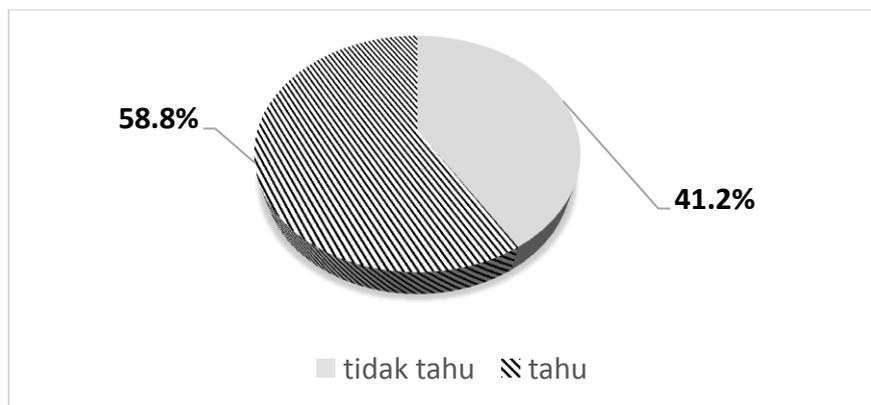
Pengetahuan pria kawin terhadap tanda bahaya pada masa kehamilan dapat dipengaruhi oleh berbagai hal, seperti umur pria, pendidikan, daerah tempat tinggal, status ekonomi, banyaknya anak, status pekerjaan, keterpaparan terhadap media informasi, dan kepemilikan alat transportasi. Peninjauan hal-hal tersebut dilakukan agar upaya dalam meningkatkan pengetahuan pria kawin terhadap tanda bahaya pada masa kehamilan dapat dilakukan secara maksimal dan menyeluruh, mengingat bahwa masalah kematian maternal perlu mendapatkan perhatian karena memiliki kaitan dengan status sosial, budaya, ekonomi, pendidikan, pengetahuan, sarana kesehatan dan kompetensi tenaga kesehatan.

2. Hasil – Hasil Utama

Gambaran Umum Pengetahuan Pria Kawin terhadap Tanda Bahaya pada Masa Kehamilan di Indonesia Tahun 2012

Indonesia masih memiliki rasio kematian maternal yang tinggi. Walaupun terus mengalami penurunan, namun pencapaian penurunan tersebut masih belum mencapai target MDGs di tahun 2015. Target tersebut ialah 102, sedangkan rasio kematian maternal yang dicapai pada tahun 2015 masih mencapai 126 kematian maternal per 100.000 kelahiran hidup. Kematian maternal seharusnya dapat dicegah dengan mengenali tanda bahaya sedini mungkin agar tindakan yang tepat dapat segera diberikan demi menyelamatkan jiwa ibu.

Penelitian ini mengkaji salah satu upaya pendeteksian dini pada hal yang dapat menyebabkan kematian maternal, yaitu melalui pengetahuan pria kawin terhadap tanda bahaya pada masa kehamilan. Dari 9.306 sampel pria kawin yang memenuhi syarat sebagai sampel dan berhasil diwawancari, terdapat 1.273 pria kawin yang termasuk sebagai *eligible sample*, yaitu pria kawin yang memiliki anak terakhir berusia 0 sampai 2 tahun. Berdasarkan hasil penelitian, berikut akan dipaparkan hasil mengenai gambaran umum karakteristik pria kawin berdasarkan pengetahuan terhadap tanda bahaya pada masa kehamilan di Indonesia. Karakteristik tersebut antara lain tempat tinggal, pendidikan, status ekonomi, banyaknya ALH, konsultasi dengan petugas kesehatan, kehadiran ketika istri/pasangan memeriksakan kehamilan, dan tempat istri/pasangan melahirkan.



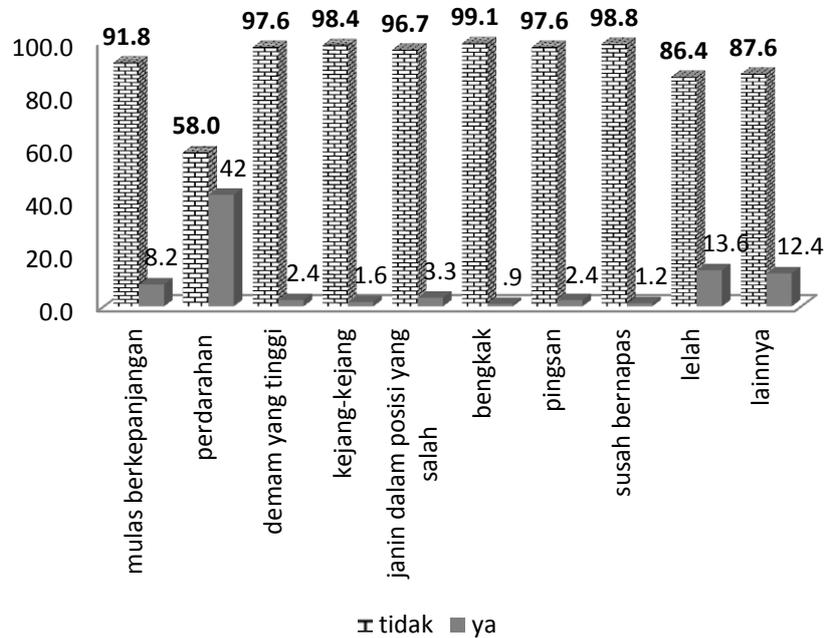
Sumber : SDKI 2012

Gambar 1. Persentase pria kawin berdasarkan pengetahuan tanda bahaya pada masa kehamilan di Indonesia tahun 2012

Gambar 1 menunjukkan bahwa sebanyak 58,8 persen pria kawin mampu menyebutkan minimal satu tanda bahaya pada masa kehamilan, namun sebanyak 41,2 persen pria kawin sama sekali tidak dapat menyebutkan tanda bahaya pada masa kehamilan, padahal pengetahuan mengenai hal yang membahayakan pada masa kehamilan perlu diketahui oleh pria kawin sebagai pasangan dan pengambil keputusan dalam rumah tangga. Pengetahuan yang dimiliki mengenai tanda bahaya tersebut merupakan langkah awal antisipasi agar tindakan yang cepat dan

tepat dapat dilakukan kepada ibu hamil yang menunjukkan tanda bahaya tersebut.

Berdasarkan hasil mengenai gambaran pengetahuan tersebut, selanjutnya akan dijelaskan secara rinci mengenai gambaran umum karakteristik pria kawin berdasarkan pengetahuan mereka terhadap tanda bahaya pada masa kehamilan.



Sumber : SDKI 2012

Gambar 2. Pengetahuan Pria Kawin Terhadap Sepuluh Tanda Bahaya Pada Masa Kehamilan

Data yang disajikan dalam gambar 2 merupakan sepuluh tanda bahaya pada masa kehamilan yang tercantum pada kuesioner SDKI 2012 dan ditanyakan oleh responden pria kawin. Dari kesepuluh tanda bahaya tersebut, perdarahan merupakan tanda bahaya yang paling banyak disebutkan. Padahal perdarahan merupakan tanda bahaya yang paling sering terjadi, bahkan sering menimbulkan keadaan darurat yang dapat menyebabkan kematian ibu.

Pengaruh Variabel-Variabel Independen terhadap Pengetahuan Pria Kawin terhadap Tanda Bahaya pada Masa Kehamilan

Uji Simultan

Uji simultan dilakukan untuk mengetahui pengaruh seluruh variabel independen yang ada di dalam model secara simultan atau bersama-sama. Hasil penelitian menunjukkan bahwa setidaknya terdapat satu variabel x yang dalam model yang memengaruhi variabel y.

Variabel yang memengaruhi pengetahuan pria kawin terhadap tanda bahaya pada masa kehamilan dianalisis menggunakan metode analisis regresi logistik biner. Pengolahan dilakukan dengan tingkat kepercayaan 90 persen.

Pengujian ini dilakukan dengan membandingkan statistik uji G dengan χ^2

tabel atau dapat dilakukan dengan membandingkan *p-value* dengan taraf nyata 0.1. Hipotesis yang digunakan dalam pengujian ini ialah sebagai berikut:

- H_0 : $\beta_1 = \beta_2 = \dots = \beta_7 = 0$ (tidak terdapat pengaruh antara sekumpulan variabel independen terhadap pengetahuan pria kawin terhadap tanda bahaya pada masa kehamilan)
- H_1 : terdapat minimal satu $\beta_j \neq 0$ (terdapat minimal satu variabel independen yang memengaruhi variabel pengetahuan pria kawin terhadap tanda bahaya pada masa kehamilan untuk semua $j = 1, 2, \dots, 7$)

Pada pengujian ini, diharapkan bahwas *p-value* lebih kecil dari α , yaitu 0,1. Hasil pengolahan menunjukkan bahwa *p-value* yang diperoleh sebesar 0,000. Hal ini berarti dapat dikatakan bahwa terdapat minimal satu variabel independen yang memengaruhi pengetahuan pria kawin terhadap tanda bahaya pada masa kehamilan di Indonesia tahun 2012.

Uji Parsial

Setelah dilakukannya pengujian secara simultan, maka selanjutnya ialah melihat keberartian setiap variabel independen dalam model secara parsial. Hipotesis yang digunakan dalam pengujian ini adalah:

- H_0 : $\beta_j = 0$ (tidak terdapat pengaruh antara variabel independen ke- j dengan pengetahuan pria kawin terhadap tanda bahaya pada masa kehamilan)
- H_1 : $\beta_j \neq 0$ (terdapat pengaruh antara variabel independen ke- j dengan variabel pengetahuan pria kawin terhadap tanda bahaya pada masa kehamilan untuk semua $j = 1, 2, \dots, 7$)

Jika hasil menunjukkan tolak H_0 , maka dapat dikatakan bahwa variabel independen ke- j memengaruhi pengetahuan pria kawin terhadap tanda bahaya pada masa kehamilan. Hal ini berlaku ketika nilai statistik *Wald* lebih besar dari χ^2 tabel atau ketika *p-value* lebih kecil dari $\alpha = 0,1$.

Berdasarkan hasil pengolahan dengan tingkat kepercayaan sebesar 90 persen, didapatkan bahwa terdapat empat variabel yang secara parsial signifikan memengaruhi pengetahuan pria kawin terhadap tanda bahaya pada masa kehamilan. Keempat variabel tersebut antara lain status ekonomi, pendidikan, banyaknya ALH, dan konsultasi dengan petugas kesehatan. Sisanya merupakan variabel-variabel yang secara parsial tidak signifikan memengaruhi pengetahuan pria kawin terhadap tanda bahaya pada masa kehamilan, yaitu tempat tinggal, kehadiran saat istri/pasangan memeriksakan kehamilan, dan tempat istri/pasangan melahirkan.

Berdasarkan nilai koefisien $\hat{\beta}$ dari masing-masing variabel, model peluang regresi logistik adalah:

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7)}$$

$$= \frac{\exp(-1,422 + 0,176D_1 + 0,459D_{21} + 0,428D_{22} + 0,648D_{23} + 0,720D_{24} + 0,372D_{31} + 0,773D_{32} + 1,205D_{33} + 0,228D_4 + 0,300D_5 + 0,182D_6 + 0,088D_7)}{1 + \exp(-1,422 + 0,176D_1 + 0,459D_{21} + 0,428D_{22} + 0,648D_{23} + 0,720D_{24} + 0,372D_{31} + 0,773D_{32} + 1,205D_{33} + 0,228D_4 + 0,300D_5 + 0,182D_6 + 0,088D_7)}$$

Tabel 1. Nilai penduga parameter, statistik uji *Wald*, derajat bebas, dan tingkat signifikansi berdasarkan variabel independen

Variabel Independen	B	Wald	Df	Sig.
(1)	(2)	(3)	(4)	(5)
Tempat Tinggal	0,176	1,651	1	0,199
Status Ekonomi		11,399	4	0,022
Status Ekonomi (<i>Second</i>)	0,459	5,428	1	0,020
Status Ekonomi (<i>Middle</i>)	0,428	4,378	1	0,036
Status Ekonomi (<i>Fourth</i>)	0,648	8,715	1	0,003
Status Ekonomi (<i>Highest</i>)	0,720	9,631	1	0,002
Pendidikan		25,863	3	0,000
Pendidikan (Primer)	0,372	2,019	1	0,155
Pendidikan (Sekunder)	0,773	8,186	1	0,004
Pendidikan (Diatas Sekunder)	1,205	16,491	1	0,000
Banyaknya ALH	0,228	2,856	1	0,091
Konsultasi ke Petugas Kesehatan	0,300	3,274	1	0,070
Kehadiran pada Pemeriksaan Kehamilan	0,182	0,579	1	0,447
Tempat Istri/Pasangan Melahirkan	0,088	0,372	1	0,542
<i>Constant</i>	-1,422	16,302	1	0,000

Sehingga persamaan yang terbentuk dari transformasi logitnya adalah:

$$\begin{aligned}
 g(x) &= \ln \left[\frac{\hat{\pi}(x)}{1 - \hat{\pi}(x)} \right] \\
 &= -1,422 + 0,176D_1 + \mathbf{0,459}D_{21} * + \mathbf{0,428}D_{22} * + \mathbf{0,648}D_{23} * + \\
 &\quad \mathbf{0,720}D_{24} * + 0,372D_{31} + \mathbf{0,773}D_{32} * + \mathbf{1,205}D_{33} * + \mathbf{0,228}D_4 * \\
 &\quad + \mathbf{0,3}D_5 * + 0,182D_6 + 0,088D_7 \\
 &= -1,422 + 0,176 \text{ tempat tinggal} + \mathbf{0,459} \text{ status ekonomi (second)} * + \mathbf{0,428} \\
 &\quad \text{status ekonomi (middle)} * + \mathbf{0,648} \text{ status ekonomi (fourth)} * + \mathbf{0,720} \\
 &\quad \text{status ekonomi (highest)} * + 0,372 \text{ pendidikan (primer)} + \mathbf{0,773} \\
 &\quad \text{pendidikan (sekunder)} * + \mathbf{1,205} \text{ pendidikan (diatas sekunder)} * + \\
 &\quad \mathbf{0,2284} \text{ banyaknya ALH} * + \mathbf{0,3} \text{ Konsultasi ke Petugas Kesehatan} * + \\
 &\quad 0,182 \text{ Kehadiran pada Pemeriksaan Kehamilan} + 0,088 \text{ Tempat} \\
 &\quad \text{Istri/Pasangan Melahirkan}
 \end{aligned}$$

Keterangan:

*Signifikan pada tingkat kepercayaan 90 persen (tingkat signifikansi 5 persen)

Interpretasi nilai $\hat{\beta}$ ialah sebagai berikut:

1. Nilai *intercept* dapat dilihat dari koefisien $\hat{\beta}_0 = -1,422$. Hal ini berarti bahwa apabila semua variabel nilainya nol, maka besarnya peluang pria kawin tidak memiliki pengetahuan terhadap tanda bahaya pada masa kehamilan ($y = 0$) adalah 0,8056 (didapatkan dari penghitungan $1 - [\exp(\beta) / 1 + \exp(\beta)] = 1 - [\exp(-1,422) / 1 + \exp(-1,422)] = 0,8056$)
2. Nilai *slope* untuk variabel banyaknya ALH dan konsultasi ke petugas kesehatan masing-masing ialah 0,228 dan 0,3. Artinya, terdapat pengaruh positif antara variabel-variabel tersebut dengan pengetahuan pria kawin terhadap tanda

bahaya pada masa kehamilan. Proporsi pria kawin untuk memiliki pengetahuan tersebut ialah lebih besar pada pria kawin dengan anak lebih dari dua dan pria kawin yang melakukan konsultasi ke petugas kesehatan.

3. Variabel status ekonomi pada masing-masing dummy memiliki nilai *slope* 0,459; 0,428; 0,648; dan 0,720. Sedangkan untuk variabel pendidikan pada masing-masing dummy adalah 0,372; 0,773; 1,205. Artinya, terdapat pengaruh positif antara status ekonomi maupun pendidikan terhadap pengetahuan pria kawin terhadap tanda bahaya pada masa kehamilan. Proporsi pria kawin dengan status ekonomi *highest* untuk memiliki pengetahuan mengenai tanda bahaya pada masa kehamilan lebih besar daripada pria kawin dengan status ekonomi yang berada dibawahnya (*fourth, middle, second, dan lowest*). Proporsi pria kawin dengan pendidikan diatas sekunder untuk memiliki pengetahuan mengenai tanda bahaya pada masa kehamilan lebih besar daripada pria kawin dengan pendidikan yang berada dibawahnya (sekunder, primer, dan tidak sekolah).

Pengujian *Goodness of Fit*

Pengujian *goodness of fit* digunakan untuk menentukan kesesuaian model regresi logistik yang dibentuk. Uji ini menggunakan hipotesis sebagai berikut:

H₀: model *fit*/sesuai (tidak ada perbedaan antara hasil observasi dan hasil prediksi dari model)

H₁: model tidak *fit*/tidak sesuai (terdapat perbedaan antara hasil observasi dan hasil prediksi dari model)

Nilai χ^2 yang dihasilkan ialah sebesar 5,044 dengan χ^2 tabel dengan derajat bebas 7 dan $\alpha = 10$ persen adalah 12,017. Selain itu, kita dapat mengambil keputusan terhadap hipotesis dengan menggunakan *p-value*, yaitu 0,753. Nilai ini lebih besar jika dibandingkan dengan $\alpha = 0,1$; sehingga hasilnya adalah tolak H₀. Hal ini menunjukkan bahwa pada tingkat kepercayaan sebesar 90 persen, dapat dikatakan bahwa model regresi logistik yang digunakan sudah *fit* atau cukup baik dalam menjelaskan data.

Odds Ratio

Odds ratio dapat digunakan untuk melihat kecenderungan pria kawin untuk memiliki pengetahuan terhadap tanda bahaya pada masa kehamilan dengan melihat nilai $\exp(\hat{\beta}_1)$ variabel yang signifikan.

Tabel 2. Nilai *odds ratio* regresi logistic

Variabel Independen (1)	<i>Odds Ratio</i> (2)
Tempat Tinggal	1,193
Status Ekonomi (<i>Second</i>)	1,582
Status Ekonomi (<i>Middle</i>)	1,535
Status Ekonomi (<i>Fourth</i>)	1,911
Status Ekonomi (<i>Highest</i>)	2,054
Pendidikan (Primer)	1,450
Pendidikan (Sekunder)	2,167
Pendidikan (Diatas Sekunder)	3,337

Banyaknya ALH	1,256
Konsultasi ke Petugas Kesehatan	1,350
Kehadiran pada Pemeriksaan Kehamilan	1,200
Tempat Istri/Pasangan Melahirkan	1,092

Berdasarkan tabel 2, nilai *odds ratio* diinterpretasikan sebagai berikut:

1. Pria kawin dengan status ekonomi *second*, *middle*, *fourth*, dan *highest* masing-masing memiliki kecenderungan 1,582; 1,535, 1,911; dan 2,054 kali lebih besar untuk memiliki pengetahuan dibandingkan dengan pria kawin dengan status ekonomi *lowest*. Sedangkan pria kawin dengan pendidikan primer, sekunder, dan diatas sekunder masing-masing memiliki kecenderungan 1,450; 2,167; dan 3,337 kali lebih besar untuk memiliki pengetahuan dibandingkan dengan pria kawin yang tidak sekolah.
2. Banyaknya ALH lebih dari dua dan melakukan konsultasi dengan petugas kesehatan memberikan kecenderungan 1,256 dan 1,35 kali lebih besar bagi pria kawin untuk memiliki pengetahuan dibandingkan dengan pria kawin yang memiliki maksimal dua ALH dan tidak melakukan konsultasi dengan petugas kesehatan.
3. Kedua poin diatas memiliki asumsi bahwa variabel lainnya konstan.

3. Kesimpulan

1. Pria kawin di Indonesia yang memiliki pengetahuan mengenai tanda bahaya pada masa kehamilan ialah sebanyak 58,8 persen. Adapun tanda bahaya yang paling banyak disebutkan oleh pria kawin ialah perdarahan, yaitu sebanyak 42 persen.
2. Variabel yang signifikan dalam memengaruhi pengetahuan pria kawin terhadap tanda bahaya pada masa kehamilan adalah pendidikan, status ekonomi, banyaknya ALH, dan konsultasi dengan petugas kesehatan. Semua variabel tersebut berpengaruh secara positif terhadap pengetahuan pria kawin. Sedangkan variabel yang tidak signifikan dalam memengaruhi pengetahuan pria kawin antara lain tempat tinggal, kehadiran pada saat istri/pasangan memeriksakan kehamilan, dan tempat istri/pasangan melahirkan.
3. Kecenderungan pria kawin untuk memiliki pengetahuan akan meningkat seiring dengan meningkatnya status ekonomi dan pendidikan. Pria kawin yang memiliki lebih dari dua ALH dan melakukan konsultasi dengan petugas kesehatan memiliki kecenderungan yang lebih besar untuk memiliki pengetahuan dibandingkan dengan pria kawin yang memiliki ALH sebanyak satu atau dua orang dan tidak melakukan konsultasi.

Referensi

- [1] Borges, Danielle da Costa Leite., 2016, *EU Health Systems and Distributive Justice*. Newyork Routledge.
- [2] World Health Organization (WHO)., 2013, *World Health Statistics 2013*. WHO Library Cataloguing-in-Publication Data
- [3] World Health Organization (WHO)., 2015, *Maternal Mortality Data by Country*. WHO Global Health Observatory Data Repository.
- [4] Kululanga, Lucy I; Johanne Sundby, Address Malata; dan Ellen Chirwa., 2012, Male

- Involvement in Maternity Health Care in Malawi. *African Journal of Reproductive Health*, Maret 2012, 16(1): 145-158.
- [5] Maryunani, Anik., 2015, *Buku Saku Kegawatdaruratan Maternal dan Neonatal Terpadu Pengenalan Praktis Program Kesehatan Terkini: Program Penyelamatan Ibu dan Bayi di Indonesia (Program EMAS)*. Jakarta Timur: Trans Info Media
- [6] Astutik, Reni Yuli., 2015, *Asuhan Kebidanan Masa Nifas dan Menyusui*. Jakarta Timur: CV. Trans Info Media
- [7] Seers, D., 1969, The Meaning of Development. *Institute of Development Studies Communication Series*, 44.
- [8] Aarnio, Pauliina; Pia Olsson; A Chimbiri; dan Teija Kulmala., 2009, Male Involvement in Antenatal HIV Counseling and Testing: Exploring Men's Perceptions in Rural Malawi. *AIDS Care, Taylor & Francis (Routledge)*.
- [9] Amin, Sajeda dan Ashish Bajracharya., 2011, Marriage and First Birth Intervals in Early and Late Marrying Societies: An Exploration of Determinants. *Paper prepared for the 2011 Annual Meetings of the Population Association of America Washington, DC March 31-April 2, 2011*.
- [10] August Furaha; Andrea B Pembe; Rose Mpembeni; Pia Axemo; dan Elisabeth Darj., 2015, Men's Knowledge of Obstetric Danger Signs, Birth Preparedness and Complication Readiness in Rural Tanzania. *Jurnal: Plos One*.
- [11] Badan Pusat Statistik., 2012, *Survei Demografi dan Kesehatan 2012*. Jakarta : Badan Pusat Statistik
- [12] Bakhari, Ayman A; Ahmad MS Almrstani; Osama S Bajouh; Abdullah F Alahmari; Ziyad T Alzahrani; dan Rayan M Alghamdi., 2015, Awareness of Obstetric Danger Signs among Husbands in Outpatient Department of Obstetrics and Gynaecology at King Abdulaziz University Hospital (KAUH), Jeddah, Saudi Arabia in September 2015 (Cross-Sectional Study). Artikel penelitian: *E Cronicon Gynaecology*, volume 2, terbitan 3 Desember 2015, halaman 192-197.
- [13] Bogale, Daniel dan Desalegn Markos., 2015, Knowledge of Obstetric Danger Signs Among Child Bearing Age Women in Goba District, Ethiopia: A Cross-Sectional Study. Artikel: *BioMed Central*.
- [14] Debelew, Gurmesa Tura; Mesganaw Fantahun Afework; dan Alemayehu Worku Yalew., 2014, Factors Affecting Birth Preparedness and Complication Readiness in Jimma Zone, Southwest Ethiopia: a Multilevel Analysis. *Open access journal: Pan African Medical Journal*.
- [15] Debiso, Alemu Tamiso; Behailu Merdekios Gello; dan Marelign Tilahun Malaju., 2015, Factors Associated with Men's Awareness of Danger Signs of Obstetric Complications and Its Effect on Men's Involvement in Birth Preparedness Practice in Suthern Ethiopia, 2014. *Hindawi Publishing Corporation Advances in Public Health*, volume 2015, article ID 386084.
- [16] Dile, Mulugeta; Daniel Taddesse, Molla Gedefaw, dan Tarekegn Asmama., 2015, Knowledge of Obstetric Danger Signs and its Associated Factors in Debaytilatgin District, Ethiopia: A Community Based Cross Sectional Study. *Gynecology and Obstetrics Journal*, volume 5, terbitan 9.
- [17] Djuhaeni, Henni., 2009, *Kebijakan Departemen Kesehatan Tentang PKMRS pada*

Penyuluhan Kelompok bagi Rumah Sakit Swanta Se Jawa Barat.

- [18] Dunn, Alice; Sayeed Haque; dan Michael Innes., 2011, Rural Kenyan Men's Awareness of Danger Signs of Obstetric Complications. Open access journal: *The Pan African Medical Journal*, 2011; 10 : 39.
- [19] Hailu, Desta dan Hailemariam Berhe., 2014, Knowledge about Obstetric Danger Signs and Associated Factors among Mothers in Tsegedie District, Tigray Region, Ethiopia 2013: Community Based Cross-Sectional Study. *Jurnal Plos One*, volume 9, terbitan 2, e83459.
- [20] Hosmer, David W dan Stanley Lemeshow., 2000, *Applied Logistic Regression [Second Edition]*. New York: John Wiley & Son
- [21] Hurlock, E. B., 1991, *Psikologi Perkembangan: Suatu Pendekatan Sepanjang Rentang Kehidupan*. Surabaya: Erlangga
- [22] Imron, Riyanti; Yusari Asih; dan Nelly Indrasari., 2016, *Buku Ajar Asuhan Kebidanan Patologi dalam Kehamilan, Persalinan, Nifas dan Gangguan Reproduksi*. Jakarta Timur: CV Trans Info Media
- [23] JHPIEGO., 2004, *Monitoring Birth Preparedness and Complication Readiness (Tools and Indicators for Maternal and Newborn Health)*. Maryland : JHPIEGO
- [24] Nigatu, Solomon Gedlu; Abebaw Gebyehu Worku; dan Abel Fekadu Dadi., 2015, Level of Mother's Knowledge About Neonatal Danger Signs and Associated Factors in North West of Ethiopia: A Community Based Study. *Bio Med Central (BMC) Research Notes*, 8: 309.
- [25] Notoatmodjo, Soekidjo., 1997, *Ilmu Kesehatan Masyarakat*. Jakarta: Rineka Cipta
- [26] Notoatmodjo, Soekidjo., 2012, *Metodologi Penelitian Kesehatan*. Jakarta: Rineka Cipta.
- [27] Pembe, Andrea B; David P Urassa; Anders Carlstedt; Gunill Lindmark; Lennarth Nystrom; dan Elisabeth Darj., 2009, Rural Tanzanian Women's Awareness of Danger Signs of Obstetric Complications *Bio Med Central (BMC)*, 9: 12.
- [28] Purwoastuti, Endang dan Elisabeth Siwi Walyani., 2015, *Ilmu Kesehatan Masyarakat dalam Kebidanan: Konsep, Teori dan Aplikasi*. Yogyakarta: Pustaka Baru Press
- [29] Rismalinda., 2015, *Buku Ajar Asuhan Kebidanan Kehamilan*. Jakarta: CV Trans Info Media
- [30] Say Iale; Doris Chou; Alison Gemmill; Ozge Tuncalp; Ann-Beth Moller; Jane Daniels; A Metin Gulmezoglu; Marleen Temmerman, dan Leontine Alkema., 2014, Global Causes of Maternal Death: A WHO Systematic Analysis. *Artikel: Lancet Glob Health 2014*.
- [31] Sekoni dan Owoaje., 2014, Male Knowledge of Danger Signs of Obstetric Complications in An Urban City in South West Nigerial. *Jurnal: Annals of Ibadan Postgraduate Medicine*, Volume 12, Nomor 2, Desember 2014.
- [32] Solomon, Abayneh Akililu; Negash Wkgari Amanta; Endeshaw Admasu Chirkose; dan Marta Berta Badi., 2015, Knowledge About Danger Signs of Pregnancy and Associated Factors Among Pregnant Women in Debra Birhan Town, Central Ethiopia. *Science Publishing Group*, 3(2): 269-273.
- [33] Teng, See Poh et al., 2015, Knowledge of Pregnancy Danger Signs and Associated Factors Among Malaysian Mothers. *British Journal of Midwifery; Nov 2015*, Vol. 23

Issue 11, p800.

- [34] Walpole, Ronald E., 1982, *Pengantar Statistika Edisi ke-3* [Alih bahasa: Ir. Bambang Sumantri]. Jakarta: PT Gramedia Pustaka Utama
- [35] Women Research Institute (WRI)., 2012, Mengurangi Angka Kematian Ibu.
- [36] World Health Organization (WHO)., 2016, *World Health Statistics 2016: Monitoring Health for The SDGs.* WHO Library Cataloguing-in-Publication Data.

ANALISIS KECENDERUNGAN USIA WANITA MELAKUKAN KOHABITASI BERDASARKAN TINGKAT PENDIDIKAN MENGGUNAKAN REGRESI LOGISTIK BINER

ROBERT KURNIAWAN¹, DESNACITA HARLY PUTRI²

1,2 Jurusan Komputasi Statistik, Sekolah Tinggi Ilmu Statistik (STIS) – Jakarta,
robertk@stis.ac.id

Abstrak. Penelitian terkait dengan kohabitasi atau lebih sering disebut dengan istilah “kumpul kebo” masih jarang dilakukan di Indonesia. Oleh karena itu, penelitian ini bertujuan untuk menginvestigasi bagaimana analisis kecenderungan usia wanita melakukan kohabitasi dengan tingkat pendidikannya. Data yang digunakan adalah data Survei Demografi Kesehatan Indonesia (SDKI) tahun 2012, dan dipilih Provinsi Jawa Barat sebagai lokus penelitian. Pengolahan data dilakukan dengan metode regresi logistik biner, dengan variabel respon adalah usia melakukan kohabitasi pada wanita yang dikategorikan menjadi 2, yaitu anak-anak dan remaja. Sedangkan untuk variabel bebasnya yaitu tingkat pendidikan dan daerah tempat tinggal.

Kata kunci: Kohabitasi, Kumpul kebo, SDKI, Regresi Logistik Biner.

1. Pendahuluan

Kohabitasi menurut Kamus Besar Bahasa Indonesia (KBBI) adalah perihal tinggal serumah tanpa ikatan perkawinan. Biasa disebut dalam pergaulan sehari-hari dengan istilah “kumpul kebo”. Kohabitasi merupakan hal yang tabu untuk dibicarakan di Indonesia, karena untuk membicarakannya masih terganjal dengan sopan santun dan adat istiadat ketimuran, tetapi perilaku tersebut ternyata sangat berkembang di masyarakat.

Akan tetapi perilaku kohabitasi di dunia barat, perilaku kohabitasi lebih banyak ditemui dan tidak menjadikan perilaku kohabitasi ini sesuatu hal yang tabu, justru diatur dalam hukum Negara, seperti di Belanda. Berdasarkan Kolkman dkk [1] dalam bukunya menjelaskan bahwa kohabitasi di Belanda, jika laki-laki dan perempuan yang tinggal bersama tanpa menikah, maka dalam dirinya sendiri tidak memiliki konsekuensi hukum apapun bagi hubungan antara dua individu tersebut. Jika memiliki anak dari hubungan tersebut, dan jika masa kohabitasi tersebut berakhir, maka keduanya tidak memiliki klaim terhadap satu sama lain untuk hak pemeliharaan anaknya.

Nelson, dkk [2] dalam hasil penelitiannya menyatakan bahwa anak atau remaja untuk melakukan kohabitasi karena pengaruh dari keluarganya yang juga

melakukan kohabitasi. Dapat dikatakan remaja akan mengikuti apa yang dilihat dan dilakukan oleh orang tuanya yang *single parent* yang tinggal serumah dengan pacar dari ibunya atau bapaknya. Olausson, dkk [3] meneliti terkait dengan remaja yang hamil diluar nikah di Swedia berasal dari keluarga yang sosial ekonominya rendah. Hal ini menyebabkan banyaknya kematian bayi yang disebabkan *free sex* yang dilakukan oleh remaja di Swedia. Hal tersebut sejalan dengan yang terjadi di Amerika, disebutkan bahwa sepertiga dari remaja wanita yang melakukan kohabitasi mempunyai anak, dan banyak yang menjadi *single mother*, Manning dan Cohen [4].

Terjadinya perilaku kohabitasi di beberapa negara barat dan Asia, termasuk Indonesia tak lepas dari beberapa faktor yang mempengaruhinya. Salah satunya adalah faktor pendidikan. Hal ini sejalan dengan penelitian Yu dan Xie [5] di China, bahwa kohabitasi salah satunya dipengaruhi oleh lokasi tempat tinggalnya (dalam hal ini desa/kota) dan pendidikannya. Rendahnya pendidikan pada wanita dari golongan masyarakat miskin di Colombia, juga memicu banyak terjadinya perilaku kohabitasi di kalangan masyarakat, Esteve dkk [6].

Hal diatas senada dengan penelitian yang dilakukan oleh Candraningrum, Dhewy dan Pratiwi [7] di Sukabumi yang menjelaskan bahwa fakta yang terkumpul di lapangan mengkonfirmasi bahwa faktor kehamilan diluar nikah mendominasi sebagai penyebab pernikahan anak. Hal ini dipicu karena rendahnya pendidikan anak perempuan dan kemiskinan keluarga. Sehingga satu-satunya solusi untuk mengatasi “kecelakaan” tersebut yaitu dengan menikahkan anaknya untuk menutupi rasa malu keluarga.

Berdasarkan hal tersebut, maka penelitian ini bertujuan untuk menginvestigasi hubungan asosiasi antara usia wanita dalam melakukan perilaku kohabitasi dan pendidikan dengan membandingkan antara yang bertempat tinggal di desa atau kota. Penelitian ini dilakukan pada lokus Provinsi Jawa Barat, pada tahun 2012.

2. Metodologi

2.1 Data

Data yang digunakan dalam penelitian ini adalah data Survei Demografi dan Kesehatan Indonesia (SDKI) 2012. Observasi yang digunakan adalah jumlah wanita dengan usia pertama kali melakukan kohabitasi. Dalam penelitian ini, variabel tak bebasnya dibagi menjadi dua kategori, yaitu kategori usia anak-anak, yaitu mereka yang melakukan kohabitasi pertama kali kurang dari 15 tahun, sedangkan kategori kedua yaitu kategori usia remaja, yaitu mereka yang melakukan kohabitasi pertama kali pada usia lebih dari sama dengan 15 tahun.

Variabel bebas yang digunakan adalah klasifikasi daerah tempat tinggal, yang dibagi menjadi 2 kategori, yaitu perdesaan dan perkotaan. Dan variabel pendidikan tinggi yang ditamatkan, dibagi menjadi 2 kategori, yaitu SD ke bawah dan SMP ke atas.

2.2 Regresi Logistik

Regresi logistik merupakan suatu model matematika yang dapat digunakan untuk menggambarkan hubungan beberapa variabel X dengan variabel respon yang dikotomis/biner [8]. Penelitian ini menggunakan metode regresi *binary* logistik multivariat karena variabel respon merupakan variabel dikotomi dan banyak variabel bebas yang menjadi variabel prediktornya multivariat. Penelitian ini hanya menggunakan dua variabel bebas. Nilai observasi untuk model regresi dengan variabel respon merupakan variabel kategorik memenuhi hubungan sebagai berikut:

$$Y = f(\mathbf{x}) + \varepsilon \quad (1)$$

Sehingga fungsi logistik yang digunakan:

$$f(\mathbf{x}) = \frac{1}{1 + e^{-(\beta_0 + \sum_i \beta_i x_i)}} \quad (2)$$

dimana $f(\mathbf{x})$ akan bergerak mulai dari 0 untuk $\beta_0 + \sum_i \beta_i x_i$ mendekati $-\infty$ dan 1 untuk $\beta_0 + \sum_i \beta_i x_i$ mendekati ∞ . Fungsi logistik merupakan fungsi peluang sehingga fungsi regresi logistik dapat dikatakan sebagai $P(Y=1|X_1, X_2)$ yang berarti bahwa peluang $Y=1$ untuk nilai tertentu dari X_1 sampai dengan X_2 . Dalam perhitungan estimasi, fungsi $f(\mathbf{x})$ tersebut harus diubah menjadi bentuk linear dalam bentuk *odds* yaitu peluang sebuah kejadian dibandingkan dengan peluang kejadian tersebut tidak terjadi. Sehingga notasi persamaan dapat ditulis menjadi:

$$\text{logit}(f(\mathbf{x})) = \text{logit}\left(\frac{f(\mathbf{x})}{1-f(\mathbf{x})}\right) = \beta_0 + \sum_{i=1}^n \beta_i x_i + e_i \quad (3)$$

dimana n adalah total jumlah variabel bebasnya. Pengujian signifikansi variabel sangat perlu dilakukan untuk mengetahui apakah variabel yang berada dalam model secara signifikan berpengaruh terhadap model. Dalam regresi logistik uji signifikansi variabel dapat ditunjukkan melalui uji *Wald Statistic*, di mana *Wald Statistic* adalah rasio dari kuadrat koefisien regresi dengan standar error kuadrat koefisien regresi, dan statistik Wald secara asimtotik berdistribusi *Chi-Square*.

$$W_j = \frac{\beta_j^2}{SE^2(\beta_j)} \quad (4)$$

Hasil pengujian statistik Wald dari setiap koefisien dalam models menyatakan bahwa variabel tersebut secara signifikan berpengaruh terhadap variabel respon.

Nilai *R-Square* merupakan statistik yang menunjukkan seberapa besar variabel dalam model dapat menjelaskan perubahan model. Berbeda dengan regresi linear sederhana, pada regresi logistik, statistik *R-Square* yang digunakan adalah statistik *Pseudo R-Square Nagelkerke* dan *Cox & Snell*, Hosmer dan Lemeshow [9]. Statistik *Cox & Snell* menghitung proporsi varian yang tidak dapat dijelaskan dan yang dihilangkan dengan penambahan variabel dalam model. Serupa dengan statistik *Cox & Snell*, statistik *Nagelkerke* juga mencerminkan proporsi variansi yang tidak dapat dijelaskan yang dihilangkan dengan penambahan variabel dalam model, tetapi nilai statistik ini berkisar antara 0 hingga 1. Statistik uji *Nagelkerke* dapat diterangkan menurut persamaan berikut :

$$R^2 = \frac{1 - \frac{(-2LL_{null})^{2/n}}{(-2LL_k)^{2/n}}}{1 - (-2LL_{null})^{2/n}} \quad (5)$$

Uji untuk mengetahui apakah model yang digunakan sudah sesuai dalam menjelaskan variabel respon adalah uji Hosmer dan Lemeshow [9], dengan hipotesis bahwa tidak terdapat perbedaan antara hasil observasi dengan hasil prediksi dari modelnya. Statistik uji yang digunakan sebagai berikut:

$$\hat{C} = \sum_{k=1}^g \frac{(O_{1k} - n'_k \bar{\pi}_k)^2}{n'_k \bar{\pi}_k (1 - \bar{\pi}_k)} \quad (6)$$

Dimana n'_k adalah jumlah subjek pada kelompok ke k, g adalah jumlah total kelompok, O_{1k} adalah sebagai jumlah variabel bebas yang tidak teridentifikasi pada kelompok ke k, $\bar{\pi}_k$ adalah rata-rata dari estimasi peluang pada kelompok ke k.

3. Hasil dan Pembahasan

3.1 Analisis Deskriptif

Berdasarkan tabel 1 terlihat bahwa rata-rata wanita melakukan kohabitasi pertama pada umur 17 tahun untuk daerah perdesaan, sedangkan di perkotaan rata-rata pada umur 19 tahun. Hal ini mengindikasikan bahwa rata-rata usia melakukan kohabitasi baik di daerah perdesaan dan perkotaan untuk wanita adalah pada usia remaja, yaitu pada usia 16 – 19 tahun.

Hal serupa juga bisa dilihat dari pendidikan yang ditamatkan. Rata-rata usia melakukan kohabitasi pada pendidikan SD ke bawah, pada usia sekitar 17 tahun. Sedangkan untuk pendidikan SMP ke atas rata-rata wanita melakukan kohabitasi pada usia 19 tahun. Hal ini mengindikasikan bahwa untuk semua pendidikan baik SD ke bawah atau SMP ke atas, rata-rata melakukan kohabitasi pertama kali adalah pada usia remaja, yaitu sekitar usia 16 – 19 tahun.

Berdasarkan hasil diatas, dapat di tarik kesimpulan bahwa usia melakukan kohabitasi pertama kali pada wanita pernah kawin di Jawa Barat terjadi pada usia remaja. Hal ini dipengaruhi oleh rasa ingin tahu dari remaja yang sangat besar terhadap seksual. Sebaliknya untuk pengetahuan tentang seksual pada semua jenjang pendidikan dirasa masih kurang.

Tabel 1. Rata-rata Usia Pertama Melakukan Kohabitasi dirinci Menurut Variabel Bebas Daerah Tempat Tinggal dan Pendidikan Tertinggi yang Ditamatkan

Variabel	Rata-rata usia pertama melakukan kohabitasi
Daerah Tempat Tinggal	
- Desa	16,99
- Kota	18,95
Pendidikan tertinggi yang di tamatkan	
- SD Kebawah	16,75
- SMP Keatas	19,61

Sumber: Hasil Pengolahan Data SDKI.

3.2 Regresi Logistik

Dari pengolahan data diperoleh nilai *Chi-Square* dari Uji Hosmer-Lemeshow sebesar 0,442 dengan nilai *p-value* sebesar 0.802 dengan tingkat

signifikansi 5 persen. Ini berarti gagal tolak H_0 , yang artinya tidak terdapat perbedaan antara hasil observasi dengan hasil prediksi dari modelnya. Berdasarkan hasil yang diperoleh seperti pada tabel 2 dibawah ini, bahwa semua variabel yang di peroleh signifikan terhadap model, hal ini dapat dilihat dari semua nilai *p-value* yang kurang dari tingkat signifikansi yang telah ditetapkan, yaitu 0,05.

Sehingga model matematis dari usia pertama melakukan kohabitasi pada wanita pernah kawin di Jawa Barat adalah sebagai berikut:

$$\text{logit}(f(x)) = 1.100 + 0.572x_{11} + 2.075x_{21}$$

Tabel 2. Koefisien, Statistik Uji, *P-value*, dan *Odd Ratio* Berdasarkan Variabel Bebas

Variabel	Koefisien	Statistik Wald	<i>P-value</i>	<i>Odd Ratio</i>
Kontanta	1.100	109.075	0.000	3.006
Daerah tempat tinggal (Perkotaan)	0.572	11.445	0.001	1.771
Pendidikan (Lebih dari SMP)	2.075	73.057	0.000	7.964

Sumber: Hasil Pengolahan Data SDKI.

Hal yang perlu diperhatikan dalam analisis regresi logistik ini adalah asosiasi kecenderungan yang dapat dilihat dari nilai *Odd Ratio*. Nilai *Odd Ratio* sebesar 1.771 untuk variabel daerah tempat tinggal berarti bahwa usia remaja pada wanita yang pernah kawin yang tinggal di perkotaan untuk melakukan kohabitasi pertama lebih besar 1,771 kali dibandingkan usia remaja di perdesaan.

Sedangkan nilai *Odd Ratio* pada variabel pendidikan tertinggi lebih dari SMP sebesar 7,964, berarti bahwa pada wanita pernah kawin usia remaja yang berpendidikan lebih dari sama dengan SMP memiliki kecenderungan melakukan kohabitasi sebesar 7,964 kali lebih besar dibandingkan usia remaja yang pendidikannya SD kebawah.

Berdasarkan dua nilai kecenderungan tersebut dapat diartikan bahwa usia remaja beresiko lebih besar melakukan kohabitasi dibandingkan dengan usia anak-anak. Hubungan asosiasi kecenderungan tersebut sejalan dengan analisis deskriptif sebelumnya yang menyatakan bahwa usia remaja berpeluang besar untuk melakukan kohabitasi dibandingkan pada usia anak-anak.

Hal ini membuktikan bahwa penelitian ini sejalan dengan penelitian yang dilakukan Manning dan Cohen [4], Olausson [3], yang menyatakan bahwa remaja yang melakukan kohabitasi mempunyai hubungan yang sangat kuat untuk mempunyai anak dan tidak sedikit juga yang mengalami kematian saat melahirkan anak pertama.

4. Kesimpulan

Penelitian ini menyimpulkan bahwa usia pertama kali melakukan kohabitasi pada wanita pernah kawin di Jawa Barat adalah pada usia remaja. Pada usia remaja di perkotaan beresiko sangat tinggi melakukan kohabitasi pertama kali

dibandingkan dengan di pedesaan, dan usia remaja dengan pendidikan SMP ke atas beresiko lebih besar melakukan kohabitasi pertamanya.

Hal ini mengindikasikan bahwa remaja lebih rentan melakukan *free sex* dan mungkin dikarenakan pengetahuan terkait dengan reproduksi sangat kurang, termasuk kurang mengetahui efek di masa depannya. Sehingga kedepannya dimungkinkan akan lebih banyak pernikahan di usia muda, yang berakibat perceraian semakin meningkat. Oleh karena itu, perlu adanya perhatian dari pemerintah, khususnya dinas pendidikan yang bekerjasama dengan LSM terkait untuk memberikan penyuluhan sejak dini kepada remaja pada jenjang pendidikan SMP ke atas, yaitu antara usia 17 tahun sampai 19 tahun. Peran orang tua juga seyogyanya ditingkatkan dalam mengawasi anak remajanya, khususnya wanita.

Referensi

- [1] Kolkman, Wilbert D., et.al., 2012, (eds), Hukum Tentang Orang, Hukum Keluarga Dan Hukum Waris Di Belanda Dan Indonesia, (Denpasar: Pustaka Larasan; Jakarta: Universitas Indonesia, Universitas Leiden, Universitas Gronigen)
- [2] Nelson, S. (2001). Beyond the two-parent family: How teenagers fare in cohabitating couple and blended families.
- [3] Olausson, P. M. O., Cnattingius, S., & Goldenberg, R. L. (1997). Determinants of poor pregnancy outcomes among teenagers in Sweden. *Obstetrics & Gynecology*, 89(3), 451-457.
- [4] Manning, W. D., & Cohen, J. A. (2015). Teenage cohabitation, marriage, and childbearing. *Population research and policy review*, 34(2), 161-177.
- [5] Yu, J., & Xie, Y. (2015). Cohabitation in China: Trends and determinants. *Population and Development Review*, 41(4), 607-628.
- [6] Esteve, A., Saavedra, A. C., López-Colás, J., López-Gay, A., & Lesthaeghe, R. J. (2016). The Boom of Cohabitation in Colombia and in the Andean Region: Social and Spatial Patterns. In *Cohabitation and Marriage in the Americas: Geo-historical Legacies and New Trends* (pp. 187-215). Springer International Publishing.
- [7] Candraningrum, D., Dhewy, A., & Pratiwi, A. M. (2016). Fear of Zina, Poor Education, and Poverty: Status of Girls in Child-Marriage in Sukabumi West Java. *Jurnal Perempuan*, 21(1), 77-95.
- [8] Kleinbaum, D. G. & Klein, M. (2002). *“Logistic Regression A Self-Learning Text, 2nd edition”*. New York: Spinger-verlag, Inc.
- [9] Hosmer, D. W., and S. Lemeshow. (2000). *Applied Logistic Regression, 2nd ed.* New York: John Wiley and Sons.

PENANGANAN MASALAH PELANGGARAN ASUMSI KENORMALAN SISAAN DENGAN REGRESI KUANTIL

FERRA YANUAR¹, HAZMIRA YOZZA², SITY KHAIRIYAH³

1 Jurusan Matematika FMIPA Unand, ferrayanuar@yahoo.co.id,
2 Jurusan Matematika FMIPA Unand, hazmirayozza@gmail.com
3 Jurusan Matematika FMIPA Unand, sitykhairiyah09@gmail.com

Abstrak. Analisis regresi melibatkan pemodelan antara dua buah peubah yaitu peubah respon dan peubah prediktor. Penggunaan analisis regresi berdasarkan pada beberapa asumsi yaitu kenormalan sisaan, multikolinearitas, kehomogenan ragam sisaan dan non-autokorelasi. Semua asumsi harus dipenuhi supaya diperoleh penduga parameter yang bersifat BLUE (Best Linier Unbiased Estimator). Penelitian ini bertujuan menerapkan metode regresi kuantil dalam mengatasi masalah pelanggaran asumsi kenormalan sisaan pada data bangkitan dimana galat berdistribusi eksponensial. Penduga parameter terbaik dievaluasi dengan menghitung nilai *mean square error* (MSE) terkecil. Kajian ini membuktikan bahwa metode kuantil menghasilkan nilai MSE yang lebih kecil dari MKT. Konsistensi hasil dugaan parameter dengan metode kuantil juga dievaluasi dengan melakukan studi simulasi. Penelitian ini menghasilkan bahwa kedua kelompok data bangkitan menghasilkan nilai dugaan parameter yang hampir sama. Dengan demikian metode regresi kuantil terbukti mampu mengatasi terlanggarnya asumsi kenormalan sisaan.

Kata kunci : BLUE, regresi kuantil, kenormalan sisaan, MSE, studi simulasi.

1. Pendahuluan

Analisis regresi adalah metode statistika yang bertujuan untuk memodelkan hubungan kausal antara dua jenis variabel yaitu variabel tak bebas (Y) dengan satu atau lebih variabel bebas (X). Dengan menggunakan analisis regresi dapat dilakukan pendugaan terhadap parameter model. Metode pendugaan parameter biasanya menggunakan Metode Kuadrat Terkecil (MKT). Metode MKT ini diterapkan jika beberapa asumsi terpenuhi, seperti asumsi kenormalan, non-multikolinieritas, kehomogenan ragam sisaan dan non autokorelasi. Semua asumsi harus terpenuhi supaya didapatkan penduga parameter yang bersifat BLUE (*Best Linier Unbiased Estimator*). Tetapi MKT ini sensitif terhadap penyimpangan asumsi tersebut, misalnya data tidak memenuhi asumsi kenormalan, variansi data tak homogen (heterokedastisitas), terdapat masalah multikolinearitas, autokorelasi

dan sebagainya. Pada data yang memiliki masalah tersebut tidak tepat jika digunakan metode MKT.

Metode regresi Median kemudian muncul untuk mengatasi kelemahan MKT. Metode regresi Median dilakukan dengan mengganti pendekatan rata-rata (*mean*) pada MKT menjadi median. Hal ini dilakukan dengan mempertimbangkan, apabila data berbentuk lonceng atau data tidak simetris. Pada kenyatannya, pendekatan regresi median juga dianggap kurang tepat karena regresi ini hanya melihat pada dua kelompok data. Padahal ada kemungkinan data bisa terbagi menjadi lebih dari dua kelompok. Dengan pertimbangan istilah ini berkembanglah metode regresi kuantil. Metode ini menggunakan pendekatan pendugaan parameter dengan memisahkan atau membagi data menjadi kuantil-kuantil, dengan menduga fungsi kuantil bersyarat pada suatu sebaran data tersebut dan meminimumkan sisaan mutlak berbobot yang tidak simetris. Beberapa peneliti sebelumnya telah membuktikan bahwa metode regresi kuantil ini mampu mengatasi asumsi-asumsi yang tidak terpenuhi, diantaranya adalah asumsi kenormalan, non-multikolinieritas, kehomogenan ragam serta tidak adanya korelasi antara sisaan (Yanuar, 2013 & Davino et al., 2014).

Penelitian ini bertujuan menerapkan metode regresi kuantil untuk mengatasi pelanggaran asumsi kenormalan sisaan dengan menggunakan data simulasi, yaitu 25 dan 50 kali simulasi.

2. Metode Regresi Kuantil

Pada bagian diuraikan beberapa teori terkait analisis regresi kuantil, uji signifikansi model dan uji kebaikan model.

2.1. Analisis Regresi Kuantil

Kuantil adalah salah satu cara dengan pembagian sekelompok data kepada beberapa bagian yang sama setelah data diurutkan dari yang paling kecil atau paling besar.

Fungsi kuantil dilambangkan dengan $Q_\theta, 0 \leq \theta \leq 1$. Misalkan Y variabel acak dengan fungsi distribusi kumulatif

$$F_y = P(Y \leq y)$$

Kuantil ke- θ dari Y adalah

$$\begin{aligned} Q_y(\theta) &= F_y^{-1}(\theta) \\ &= \text{inv} \{ y : F_y(y) \geq \theta \} \end{aligned}$$

Model persamaan regresi klasik,

$$y_i = \mathbf{x}_i' \boldsymbol{\beta} + e_i, \quad i = 1, \dots, n \quad (2.1)$$

Dengan y_i adalah variabel tak bebas untuk setiap data i , \mathbf{x}_i' adalah vektor $1 \times p$ yang terdiri dari p buah variabel bebas, dimana $E(y_i) = \mathbf{x}_i' \boldsymbol{\beta}$, $\boldsymbol{\beta}$ adalah vektor $1 \times p$ dari parameter dan e_i adalah galat untuk setiap data observasi i . Pendugaan parameter dengan regresi klasik, dengan meminimumkan jumlah kuadrat galat, dituliskan sebagai

$$\min \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.2)$$

yang mana \hat{y}_i merupakan nilai dugaan bagi setiap data observasi i . Pendugaan berdasarkan median, yaitu dengan meminimumkan jumlah galat mutlak dapat ditulis dengan menggunakan persamaan

$$\min \sum_{i=1}^n |y_i - \hat{y}_i| \quad (2.3)$$

Seterusnya, persamaan linear bagi kuantil ke- τ dapat dituliskan dengan persamaan

$$y_i = x_i' \beta_\tau + e_i \quad i = 1, \dots, n \quad (2.4)$$

Nilai τ berada dalam selang (0,1). Berdasarkan konsep median, dugaan untuk β dari regresi kuantil ke- τ diperoleh dengan meminimumkan jumlah nilai mutlak dari galat dengan pembobot τ untuk galat positif dan pembobot $(1-\tau)$ untuk galat negatif. Konsep ini dijadikan sebagai suatu penyelesaian untuk meminimumkan masalah dengan persamaan

$$\min_{a \in \mathbb{R}} \left\{ \sum_{i=1}^n \tau |y_i - a| + \sum_{i=1}^n (1-\tau) |y_i - a| \right\} \quad 0 < \tau < 1 \quad (2.5)$$

Pendugaan umum bagi kuantil ke- τ berdasarkan persamaan (2.5), dapat dituliskan sebagai berikut :

$$\hat{\beta}(\tau) = \min_{\beta \in \mathbb{R}} \left\{ \sum_{i=1}^n \tau |y_i - x_i' \beta_\tau| + \sum_{i=1}^n (1-\tau) |y_i - x_i' \beta_\tau| \right\} \quad (2.6)$$

Atau dapat juga dituliskan dengan

$$\hat{\beta}(\tau) = \min_{\beta \in \mathbb{R}} \sum_{i=1}^n \rho_\tau (y_i - x_i' \beta_\tau) \quad 0 < \tau < 1 \quad (2.7)$$

Makna yang dapat ditambahkan untuk menjelaskan persamaan (2.5) dan (2.6), yaitu bahwa semua observasi yang lebih besar dari nilai kuantil, dikalikan dengan pembobot τ dan observasi yang nilainya kurang dari kuantil dikalikan dengan pembobot $(1-\tau)$. Penggunaan pembobot ini berfungsi untuk mengoptimalkan nilai LAD observasi (Yanuar, F., 2013). Adapun pemilihan pembobot yang akan digunakan adalah berdasarkan subjektif peneliti.

2.2 Uji Signifikansi Parameter Model

Statistik uji t pada regresi kuantil bertujuan untuk menguji pengaruh koefisien regresi β pada model secara terpisah

Pengujian hipotesisnya adalah :

$$H_0 : \beta_j = 0 \text{ (pengaruh variabel ke-}j\text{ tidak signifikan)}$$

$$H_1 : \beta_j \neq 0 \text{ (pengaruh variabel ke-}j\text{ signifikan)}$$

untuk $j = 1, 2, \dots, p$. Rumus umum untuk uji t adalah :

$$t_j = \left(\frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \right) \quad (2.8)$$

dengan $\hat{\beta}_j$ merupakan koefisien regresi ke- j , $SE(\hat{\beta}_j)$ merupakan galat baku koefisien regresi ke- j . Daerah penolakannya adalah jika $|t_{hit}| > t_{\alpha/2, v}$, dimana α adalah tingkat signifikansi dan v adalah derajat bebas

2.3 Uji Normalitas (Shapiro – Wilk)

Uji Shapiro wilk digunakan untuk mengetahui apakah data berdistribusi normal atau tidak. Hipotesis dari pengujian ini sebagai berikut :

H_0 : X_1, X_2, \dots, X_n berdistribusi normal

H_1 : X_1, X_2, \dots, X_n tidak berdistribusi normal

Statistik uji yang digunakan adalah :

$$W = \frac{1}{\sum_{i=1}^n (X_i - \bar{X})^2} \left[\sum_{i=1}^k a_i (X_{n-i+1} - X_i) \right]^2 \quad (2.9)$$

dimana a_i merupakan koefisien uji Shapiro wilk (diperoleh dari tabel), X_{n-i+1} adalah data ke $n - i + 1$, X_i menyatakan data ke i dan \bar{X} adalah rata – rata data. Nilai statistik uji (W) yang diperoleh kemudian dibandingkan dengan nilai tabel Shapiro Wilk. Jika nilai $W <$ nilai tabel maka H_0 ditolak dan sebaliknya. Uji Shapiro wilk dianggap lebih akurat digunakan jika data kurang dari 50.

2.4 Indikator Keباikan Model

2.4.1 *PseudoR*²

Model regresi kuantil sederhana dengan n variabel bebas dapat dibentuk sebagai berikut :

$$Q_\tau(\hat{y}|\mathbf{x}) = \hat{\beta}_0(\tau) + \hat{\beta}_1(\tau)\mathbf{x} + \dots + \hat{\beta}_n(\tau)\mathbf{x} \quad (2.10)$$

Indikator kebaikan model tersebut dapat diduga dengan *pseudoR*² yang dirumuskan sebagai berikut (Davino et al., 2014) :

$$pseudoR_\tau^2 = 1 - \frac{RAWS_\tau}{TASW_\tau} \quad (2.11)$$

dimana :

$$RAWS_\tau = \sum_{y_i \geq \hat{\beta}_0(\tau) + \hat{\beta}_1(\tau)x_i + \dots + \hat{\beta}_n(\tau)x} \tau |y_i - \hat{\beta}_0(\tau) - \hat{\beta}_1(\tau)x_i - \dots - \hat{\beta}_n(\tau)x_i| \\ + \sum_{y_i < \hat{\beta}_0(\tau) + \hat{\beta}_1(\tau)x_i + \dots + \hat{\beta}_n(\tau)x} (1 - \tau) |y_i - \hat{\beta}_0(\tau) - \hat{\beta}_1(\tau)x_i - \hat{\beta}_n(\tau)x_i|$$

dan

$$TASW_\tau = \sum_{y_i \geq \hat{\tau}} \tau |y_i - \hat{\tau}| + \sum_{y_i < \hat{\tau}} (1 - \tau) |y_i - \hat{\tau}|$$

Nilai *RAWS* _{τ} selalu lebih kurang dari nilai *TASW* _{τ} , sehingga nilai *pseudoR* _{τ} ² akan berada pada rentang 0 sampai 1. Semakin dekat nilai *pseudoR* _{τ} ² ke satu maka model akan semakin baik. Namun nilai kebaikan *pseudoR* _{τ} ² tidak dapat digunakan untuk menguji kebaikan model secara keseluruhan, hanya dapat digunakan untuk menguji kebaikan model perkuantil yang dipilih.

2.4.2 *Mean Square Error (MSE)*

Suatu pendugaan parameter yang diperoleh dikatakan baik apabila memiliki nilai bias dan ragam yang kecil. Oleh karena itu, untuk melihat kebaikan pendugaan parameter berdasarkan nilai bias dan ragam secara bersamaan maka diperlukan *Mean Square Error (MSE)* (Schulze, 2004) :

$$MSE(\hat{\beta}_j(\tau_q)) = Var \left((\hat{\beta}_j(\tau_q)) \right) + Bias \left(\hat{\beta}_j(\tau_q) \right)^2 \quad (2.12)$$

dimana *MSE*($\hat{\beta}_j(\tau_q)$) adalah nilai MSE dengan $j = 1, 2, \dots, p$, τ_q menyatakan kuantil ke- q , $q=1, 2, \dots, k$. *Var*($\hat{\beta}_j(\tau_q)$) adalah nilai variansi dari kuantil yang terpilih dan *Bias*($\hat{\beta}_j(\tau_q)$) adalah nilai bias dari kuantil yang diperoleh yaitu selisih nilai dugaan dengan nilai tetapan dengan rumus :

$$Bias \left(\hat{\beta}_j(\tau_q) \right) = \frac{1}{n} \sum_{i=1}^n \left((\hat{\beta}_j(\tau_q)) - (\beta_j(\tau_q)) \right) \quad (2.13)$$

3. Penerapan Metode Regresi Kuantil pada Data Bangkitan

Data yang digunakan pada penelitian ini adalah data bangkitan berukuran 30. Data untuk variabel X_1 dibangkitkan dari sebaran $N(0,1)$ dan variabel X_2 dibangkitkan dari sebaran $t_{(1)}$. Variabel tak bebas Y ditetapkan dengan nilai $y = 2x_1 + 0.5x_2 + \varepsilon_1$ dimana $\varepsilon_1 = \varepsilon - 1$ dengan ε dibangkitkan dari sebaran EXP (1). Variabel sisaan yang dibangkitkan diatur sedemikian hingga menyebabkan variabel respon tidak menyebar normal. Data tersebut kemudian dimodelkan dengan menggunakan metode regresi kuantil.

3.1 Hasil Estimasi Parameter Model Dengan Metode Regresi Kuantil dan Metode Kuadrat Terkecil

3.1.1 Hasil Estimasi Parameter Model Dengan Metode Regresi Kuantil

Berikut ini adalah hasil pemodelan data dengan menggunakan metode kuantil. Nilai kuantil yang dipilih adalah kuantil ke- $\tau = 0.10, 0.25, 0.50, 0.75$ dan 0.90 . Model hipotesis yang digunakan pada kasus ini adalah $y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \varepsilon_1$. Tabel 3.1 berikut menampilkan rata-rata hasil pendugaan parameter model untuk β_0, β_1 , dan β_2 untuk 50 kali simulasi dan nilai standar error untuk setiap model.

Tabel 3.1 Nilai Parameter Dugaan Regresi Kuantil

Kuantil ke- τ	Parameter	Nilai dugaan	Standar Error
0.1	b0	0.096095*	0.03563
	b1	1.908114*	0.03929
	b2	0.492669*	0.01064
0.25	b0	0.281717*	0.10791
	b1	1.943742*	0.04788
	b2	0.506277*	0.02488
0.5	b0	0.754225*	0.05933
	b1	1.874276*	0.06018
	b2	0.504950*	0.02898
0.75	b0	1.458130*	0.10922
	b1	1.932415*	0.11904
	b2	0.524259*	0.04173
0.9	b0	2.488928*	0.29725
	b1	1.759540*	0.26250
	b2	0.530678*	0.05694

(*setiap nilai dugaan signifikan pada $\alpha = 0,05$).

Tabel 3.1 memberikan informasi bahwa nilai dugaan mutlak koefisien model ($\hat{\beta}_1$ dan $\hat{\beta}_2$) pada kuantil ke-0,1 masing-masing adalah 1.908114 dan

0.49267. Nilai tersebut sudah cukup dekat dengan nilai awal yang ditetapkan yaitu $b_1 = 2$ dan $b_2 = 0.5$. Hasil yang tidak jauh berbeda juga diperoleh pada kuantil ke-25 yaitu $\hat{\beta}_1 = 1.94374$ dan $\hat{\beta}_2 = 0.506277$. Kuantil ke-0,5 nilai koefisien modelnya yaitu $\hat{\beta}_1 = 1.87427599$ dan $\hat{\beta}_2 = 0.50494997$. Kuantil ke-0,75 nilai koefisien modelnya adalah $\hat{\beta}_1 = 1.932415$ dan $\hat{\beta}_2 = 0.524259$. Kuantil ke-0,90 nilai koefisien modelnya $\hat{\beta}_1 = 1.759540$ dan $\hat{\beta}_2 = 0.530678$.

Selanjutnya adalah melakukan uji kebaikan model dengan menghitung nilai $PseudoR^2$. Tabel 3.2 menampilkan nilai $PseudoR^2$ pada masing-masing kuantil Analisis ini menghasilkan nilai $PseudoR^2$ yang cukup beragam pada masing-masing kuantil namun tetap diatas 50%, dengan nilai $PseudoR^2$ terbesar adalah pada model dugaan yang berada pada kuantil ke 0.1 yaitu sebesar 0.85476

Tabel 3.2 Nilai $PseudoR^2$

Kuantil ke- τ	$PseudoR^2$
0.10	0.85476
0.25	0.806906
0.50	0.751773
0.75	0.706754
0.90	0.676100

Sebagaimana diketahui bahwa nilai $PseudoR^2$ ini juga menyatakan besarnya kebaikan suatu model dugaan. Dengan demikian semua model dugaan yang dihasilkan dari metode kuantil ini sudah memenuhi kriteria kebaikan model. Dapat terlihat bahwa ada kecenderungan penurunan nilai $PseudoR^2$ dengan bertambahnya nilai kuantil.

3.1.2 Hasil Estimasi Parameter Model Dengan Metode Kuadrat Terkecil

Tabel 3.3 di bawah ini menampilkan rata-rata hasil pendugaan parameter model untuk β_0 , β_1 , dan β_2 dengan menggunakan Metode Kuadrat Terkecil dan nilai standar error untuk setiap model.

Tabel 3.3 Nilai Parameter Dugaan Metode Kuadrat Terkecil

Parameter	Nilai dugaan	Standar Error	R^2
b0	0.92869	0.04552	0.91875
b1	1.87760	0.05001	
b2	0.48589	0.03268	

Tabel 3.3 memberikan informasi bahwa nilai dugaan mutlak koefisien model ($\hat{\beta}_1$ dan $\hat{\beta}_2$) masing-masing adalah 1.87760 dan 0.48589. Nilai tersebut sudah cukup dekat dengan nilai awal yang ditetapkan yaitu $b_1 = 2$ dan $b_2 = 0.5$. Nilai R^2 yang merupakan indikator kebaikan model adalah 0.91875.

3.2 Perbandingan Hasil Estimasi Parameter Model dengan Metode Regresi Kuantil dan Metode Kuadrat Terkecil

Kebaikan pendugaan parameter regresi dapat dilihat berdasarkan nilai *mean square error* (MSE). Perbandingan nilai MSE dari Metode Regresi Kuantil dengan Metode Kuadrat Terkecil disajikan pada Tabel 3.4 berikut.

Tabel 3.4 Nilai MSE Regresi kuantil dan Metode Kuadrat Terkecil

Parameter	MSE Regresi Kuantil					MSE MKT
	0.1	0.25	0.5	0.75	0.9	
b0	0.02384	0.11114	0.53243	0.52511	0.88926	0.86453
b1	0.00204	0.00282	0.01406	0.01447	0.09856	0.01748
b2	0.00013	0.00066	0.00084	0.00189	0.00362	0.00127

Berdasarkan tabel di atas dapat diketahui bahwa nilai MSE Regresi Kuantil lebih kecil dari MSE MKT untuk tiap parameternya. Dengan demikian dapat dikatakan bahwa penduga parameter dengan metode regresi kuantil adalah lebih baik daripada MKT terutamanya pada data sisaan yang melanggar asumsi kenormalan.

3.3 Uji Normalitas Nilai \hat{Y}

Selanjutnya dilakukan uji normalitas terhadap nilai \hat{Y} untuk setiap kuantil yang diamati. Uji normalitas bertujuan untuk memastikan bahwa sebaran nilai \hat{Y} telah menyebar normal.

Tabel 3.5. Uji Kenormalan dengan Uji Shapiro –Wilk

Kuantil ke- τ	Shapiro-Wilk		
	Nilai W	df	Sig.
0,10	0.985	30	0.938
0,25	0,965	30	0,911
0,50	0,922	30	0,914
0,75	0,911	30	0,904
0,90	0,901	30	0,917

Berdasarkan Tabel 3.5 diketahui bahwa pada kuantil ke 0,10 nilai \hat{Y} mempunyai nilai *sig* = 0.938. Jika dibandingkan dengan nilai $\alpha = 0.05$, nilai *sig* = 0.938 > $\alpha = 0.05$ maka dapat disimpulkan bahwa \hat{Y} telah berdistribusi normal. Hal yang sama juga diperoleh untuk kuantil yang lain. Dengan demikian setelah menggunakan pendugaan parameter dengan metode regresi kuantil, sebaran

untuk nilai \hat{Y} telah berdistribusi normal.

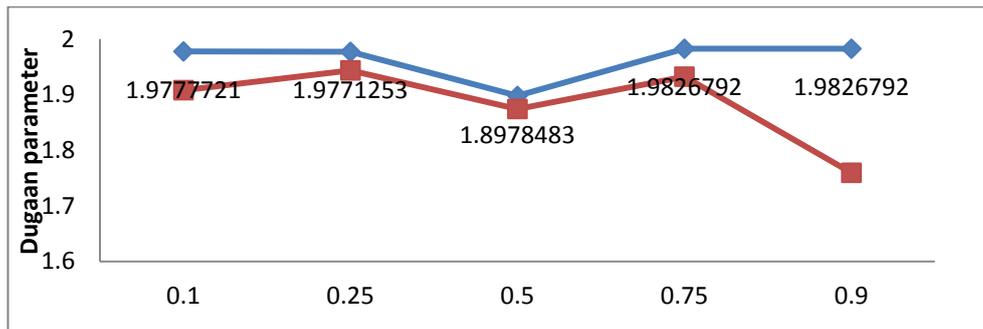
3.4 Studi Simulasi Untuk Mendeteksi Konsistensi Hasil Metode Regresi Kuantil

Berikut ini dilakukan kajian simulasi (25 dan 50 kali simulasi) untuk mengetahui konsistensi dari hasil metode regresi kuantil. Studi simulasi dilakukan dengan membangkitkan kelompok data sebanyak 25 kali dan 50 kali untuk setiap X_1 , X_2 dan Y dengan mengikuti aturan yang sama dengan data sebelum, kemudian dilakukan pendugaan parameter model dengan metode regresi kuantil. Selanjutnya dicari rata-rata nilai setiap parameter model yang diperoleh. Hasil perbandingan rata-rata nilai β_0 , β_1 , dan β_2 untuk 25 kali dan 50 kali simulasi disajikan pada tabel berikut.

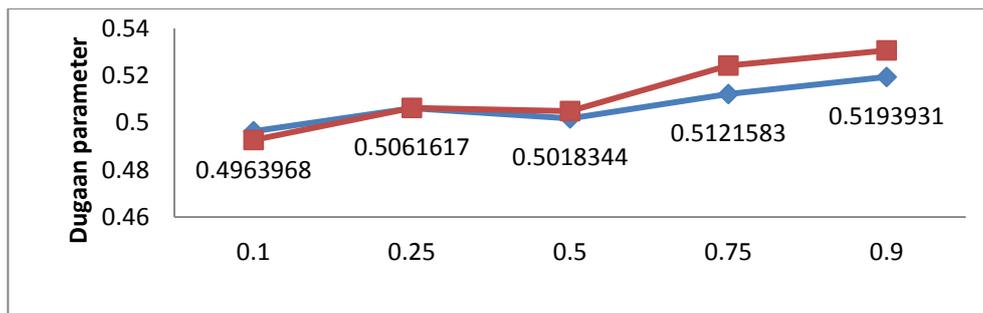
Tabel 3.6 Perbandingan Rata-rata Nilai Parameter Model untuk 25 dan 50 Kali Simulasi

Kuantil ke- τ	Parameter	Nilai Dugaan (Galat Baku)			
		25 kali simulasi		50 kali Simulasi	
0.1	b0	0.15023	(0.031063)	0.09610	(0.03563)
	b1	1.97777	(0.030508)	1.90812	(0.03929)
	b2	0.49640	(0.013070)	0.49267	(0.01064)
0.25	b0	0.31542	(0.051960)	0.28172	(0.10791)
	b1	1.97712	(0.050485)	1.94374	(0.04787)
	b2	0.50617	(0.033065)	0.50628	(0.02488)
0.5	b0	0.72726	(0.055006)	0.75423	(0.05933)
	b1	1.89785	(0.058084)	1.87428	(0.06018)
	b2	0.50183	(0.030231)	0.50495	(0.02898)
0.75	b0	1.41886	(0.138413)	1.45813	(0.10921)
	b1	1.98268	(0.146818)	1.93242	(0.11904)
	b2	0.51216	(0.052303)	0.52426	(0.04173)
0.9	b0	2.40856	(0.387163)	2.48893	(0.29725)
	b1	1.82781	(0.223863)	1.75954	(0.26251)
	b2	0.51939	(0.073960)	0.53068	(0.05694)

Berikut dapat dilihat trend perubahan nilai koefisien regresi ($\hat{\beta}_1$ dan $\hat{\beta}_2$) pada tiap bangkitan untuk setiap kuantil terpilih yang diilustrasikan pada Gambar 3.1 berikut.



(a)



(b)

Gambar 3.1 Nilai Rata-rata Parameter untuk 25 dan 50 Simulasi Untuk Setiap Kuantil, (a) Koefisien β_1 dan (b) Koefisien β_2

Keterangan gambar : warna merah untuk simulasi 25 kali
warna biru untuk simulasi 50 kali

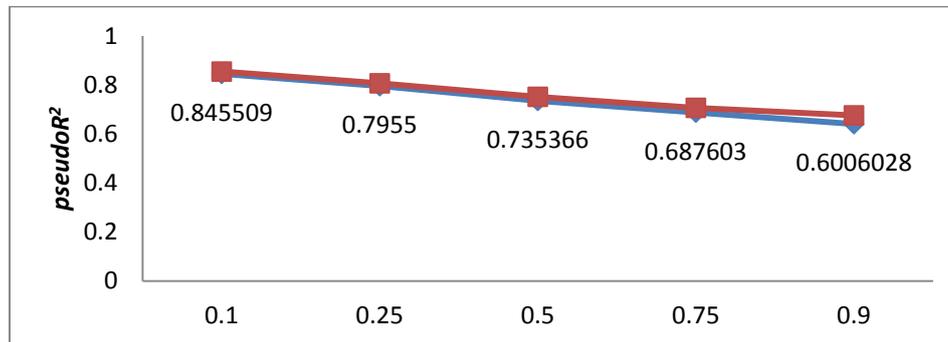
Berdasarkan tabel dan gambar di atas dapat dilihat bahwa nilai dugaan parameter untuk 25 dan 50 kali bangkitan menghasilkan nilai yang cukup dekat.

Kemudian dilakukan uji kebaikan model dengan menghitung nilai $pseudoR^2$. Nilai $PseudoR^2$ untuk setiap model disajikan pada tabel di bawah ini.

Tabel 3.7 Nilai $pseudoR^2$ Untuk 25 dan 50 kali Simulasi

Kuantil ke- τ	Nilai $pseudoR^2$ 25 Simulasi	Nilai $pseudoR^2$ 50 simulasi
0.10	0.84551	0.85476
0.25	0.79550	0.80690
0.50	0.73536	0.75177
0.75	0.68760	0.70675
0.90	0.64071	0.67610

Gambar 3.2 berikut ini mengilustrasikan perubahan nilai $PseudoR^2$ untuk setiap nilai kuantil yang diamati pada 25 dan 50 kali simulasi. Dapat terlihat bahwa ada kecenderungan penurunan nilai $PseudoR^2$ dengan bertambahnya nilai kuantil.



Gambar 3.2 Nilai $PseudoR^2$ Untuk Setiap Kuantil Pada Setiap Simulasi

Keterangan gambar : warna merah untuk simulasi 25 kali
warna biru untuk simulasi 50 kali

3. Kesimpulan

Penelitian ini bertujuan untuk membandingkan hasil dugaan parameter antara MKT dan metode kuantil untuk data dengan galatnya tidak berdistribusi normal. Kajian ini memilih nilai kuantil adalah kuantil ke- $\tau = 0.10, 0.25, 0.50, 0.75$ dan 0.90 . Pendugaan parameter pada kuantil yang berbeda-beda ini menghasilkan nilai dugaan yang hampir sama untuk setiap parameter yang bersesuaian. Kajian ini juga membuktikan bahwa penduga parameter yang terbaik adalah penduga dengan metode kuantil, karena penduga tersebut menghasilkan nilai MSE yang lebih kecil dari MKT. Konsistensi dari penduga kuantil diuji dengan melakukan studi simulasi sebanyak 25 dan 50 kali bangkitan. Didapati bahwa hasil dugaan parameter untuk kedua bangkitan menghasilkan nilai yang hampir sama.

Referensi

- [1] Davino, Christina, dkk. 2014. *Quantile Regression Theory and Applications*. Pondicherry. India
- [2] Schulze, N. 2004. *Applied Quantile Regression: Microeconomic, Financial, and Environmental Analyses*. Tubingen: Universitat Tubingen
- [3] Yanuar, F. 2013. Quantile Regression Approach to Determine the Indicator of Health Status. *Scientific Research Journal*, 1 : 17 - 23
- [4] Yu K. 2003. Quantile regression: Applications and Current Research Area. *The Statistician*, 52 (3): 331-350

METODE LOGISTIC REGRESSION UNTUK TEKNOLOGI CREDIT SCORING

ULFA FAUZIAH¹, ZUHERMAN RUSTAM²

1 Departemen Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas
Indonesia, ulfa.fauziah@sci.ui.ac.id

2 Departemen Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas
Indonesia, rustamzr@yahoo.com

Abstrak. Meningkatnya kebutuhan manusia sebanding dengan meningkatnya pengeluaran akan uang. Banyak orang yang memiliki permasalahan dalam memenuhi kebutuhan tersebut, yaitu pendapatan yang diperoleh tidak sebanding dengan besarnya pengeluaran yang dibutuhkan. Pinjaman Bank dapat dimanfaatkan untuk mengatasi masalah ini. Namun pihak Bank sebagai pihak pemberi pinjaman tidak akan begitu saja memberikan pinjaman, model Teknologi *Credit Scoring* dapat dimanfaatkan untuk menyaring pinjaman dikarenakan terdapat risiko yang besar bagi Bank jika penyaringan ini tidak dilakukan. Model *Logistic Regression* dapat digunakan untuk menghubungkan probabilitas kegagalan pinjaman (kredit macet) dengan beberapa atribut peminjam yang diperlukan seperti besar pendapatan, besar pinjaman, usia peminjam dan lain-lain. Sehingga diharapkan metode *Logistic Regression* dapat digunakan untuk mencari probabilitas kredit macet dari peminjam dimana probabilitas tersebut digunakan untuk menentukan kategori kredit yaitu kredit macet atau kredit lancar.

Kata kunci: Teknologi Credit Scoring, Logistic Regression.

1. Pendahuluan

Banyaknya jenis kredit yang ditawarkan oleh perusahaan-perusahaan finansial membuat penggunaan kredit sudah menjadi hal yang lazim bagi masyarakat Indonesia. Jenis kredit yang ditawarkan mulai dari barang-barang yang termasuk dalam kebutuhan sandang, pangan dan papan. Namun, dari banyaknya jenis kredit yang ditawarkan tersebut, kredit uang merupakan salah satu bentuk kredit yang paling diminati, hal ini karena uang merupakan hal yang tidak dapat dipisahkan dari kehidupan manusia. Kredit uang yang biasanya diajukan tersebut digunakan untuk alasan yang berbeda, sebagian besar ada yang menggunakan untuk investasi, dan ada juga untuk keperluan konsumtif.

Permintaan kredit yang tinggi oleh masyarakat tidak akan membuat pihak bank (kreditor) memberikan begitu saja pinjaman, dikarenakan terdapat resiko yang besar. Resiko yang dimaksud disini adalah ketidakmampuan masyarakat untuk membayar kembali pinjaman ke pihak bank karena suatu keadaan yang berbeda atau biasa disebut dengan istilah *default*. Akibat *default* tersebut, pihak kreditor akan mengalami banyak kerugian apabila menarik kembali kredit yang

bermasalah (kredit macet) yang dialami pihak debitor. Walaupun dengan jaminan yang sudah diberikan, pihak bank juga tetap akan mengalami kerugian seperti rugi waktu, rugi tenaga dan terutama kerugian dana operasional yang akan mempengaruhi kestabilan bank tersebut.

Sebelum memberikan pinjaman, pihak bank (kreditor) akan melakukan analisis kredit. Analisis kredit disini dengan melakukan penilaian (*scoring*) terhadap berbagai data yang dimiliki calon debitor. *Scoring* ini akan menghasilkan dua kategori, yaitu “kredit lancar” dan “kredit macet”. Apabila calon debitor masuk kedalam kategori “kredit lancar” maka pengajuannya akan dilanjutkan (*accept*), tetapi jika masuk ke dalam kategori “kredit macet” maka pengajuan pinjamannya akan dipertimbangkan terlebih dahulu atau bahkan ditolak.

Terdapat banyak metode yang sudah digunakan dalam *credit scoring*, antara lain *Neural Network*, *Genetic Algorithm*, *Fuzzy Logic*, dan lain-lain. Dalam penelitian ini digunakan metode *Logistic Regression* untuk mendapatkan *output* berupa probabilitas dimana probabilitas tersebut digunakan untuk menentukan resiko kredit (kredit macet atau kredit lancar) [1].

2. Metode dan Bahan yang Digunakan

2.1 Analisis Kredit

Seperti yang sudah dijelaskan sebelumnya, sebelum pihak bank memberikan pinjaman, akan dilakukan analisis kredit. Analisis yang biasanya dilakukan oleh pihak bank dikenal dengan istilah analisis 5C, yaitu [2] :

1. *Character*

Character menggambarkan sifat atau watak seseorang. Untuk membaca watak seseorang dapat dilihat dari latar belakang calon debitor, seperti latar belakang pekerjaan, gaya hidup dan lain-lain.

2. *Capacity*

Capacity merupakan kemampuan calon debitor untuk membayar kembali pinjamannya. Kemampuan keuangan calon debitor ini sangat penting karena akan menentukan kualitas dari pembayaran kembali pinjaman. Kemampuan keuangan ini dapat dilihat dari latar belakang pendidikan dan bagaimana cara calon debitor mengelola usahanya.

3. *Capital*

Analisis kredit terhadap objek kredit berbentuk modal harus dilakukan secara detail. Semakin besar modal yang dimiliki, maka akan semakin baik dalam memberikan keyakinan pihak bank dalam memberikan kredit.

4. *Collateral*

Collateral merupakan jaminan atau agunan yang diberikan pihak debitor untuk mengajukan kredit. Jaminan ini merupakan sumber pembayaran yang kedua, apabila suatu saat terjadi kredit yang bermasalah maka pihak bank dapat mengeksekusi jaminan yang diberikan, pihak bank dapat menjual kembali jaminan tersebut dan hasil penjualan dapat digunakan untuk membayar kembali pinjaman debitor tersebut.

5. *Condition*

Dalam memberikan kredit, pihak bank juga harus melihat kondisi ekonomi, politik dan sosial masa sekarang dan memprediksi keadaan di masa yang akan datang. Prospek usaha yang dimiliki calon debitor juga harus dilihat

dan dilakukan penilaian. Hal ini juga akan mempengaruhi pengembalian pinjaman debitur.

2.2 Data Penelitian

Dalam penelitian ini digunakan data dari salah satu perusahaan finansial cabang Depok. Data tersebut merupakan data selama 4 bulan, yaitu bulan Desember 2015, Januari 2016, Februari 2016 dan Maret 2016 dengan jumlah 302 debitur. Atribut yang digunakan meliputi besar pendapatan perbulan, besar pinjaman, usia, klasifikasi pekerjaan, status tempat tinggal serta kepemilikan jaminan. Untuk besar pendapatan perbulan, besar pinjaman, dan usia merupakan variabel numerik, sedangkan klasifikasi pekerjaan, status tempat tinggal dan kepemilikan jaminan merupakan variabel kategorik.

Pada penelitian ini, klasifikasi pekerjaan dibedakan menjadi 2 kategori yaitu karyawan swasta yang diberikan nilai 1, dan wiraswasta diberikan nilai 2. Status tempat tinggal dibedakan menjadi kategori mengontrak yang diberikan nilai 1, dan rumah (tempat tinggal) sendiri diberikan nilai 2, dan terakhir kepemilikan jaminan juga dibedakan menjadi kategori jaminan atas nama orang lain yang diberikan nilai 1, dan jaminan atas nama sendiri diberikan nilai 2. Dalam menganalisis kredit, digunakan analisis 5C, yaitu *Character* menggunakan atribut usia dan klasifikasi pekerjaan, *Capacity* menggunakan atribut besar pendapatan perbulan, *Capital* menggunakan atribut besar pinjaman yang diajukan, *Collateral* menggunakan atribut kepemilikan jaminan, dan *Condition* menggunakan atribut status tempat tinggal. Tabel 1 menjelaskan mengenai variabel-variabel yang digunakan.

Tabel 1. Variabel yang Digunakan

No	Variabel	Keterangan
1.	x_{1i}	Besar pendapatan perbulan data ke i
2.	x_{2i}	Besar pinjaman data ke i
3.	x_{3i}	Usia pada data ke i
4.	x_{4i}	Klasifikasi pekerjaan pada data ke i
5.	x_{5i}	Status tempat tinggal pada data ke i
6.	x_{6i}	Kepemilikan jaminan data ke i
7.	$\mu(x_{1i})$	Nilai <i>fuzzy</i> besar pendapatan perbulan data ke i
8.	$\mu(x_{2i})$	Nilai <i>fuzzy</i> besar pinjaman data ke i
9.	$\mu(x_{3i})$	Nilai <i>fuzzy</i> usia data ke i
10.	$\mu(x_{4i})$	Nilai <i>fuzzy</i> klasifikasi pekerjaan data ke i
11.	$\mu(x_{5i})$	Nilai <i>fuzzy</i> status tempat tinggal pada data ke i
12.	$\mu(x_{6i})$	Nilai <i>fuzzy</i> kepemilikan jaminan pada data ke i
13.	β_0, \dots, β_6	Model parameter
14.	\hat{Y}	Nilai yang akan diobservasi
15.	$\hat{\pi}$	Probabilitas kredit macet seorang peminjam

2.3 Metode Logistic Regression

Metode *Logistic Regression* merupakan salah satu metode statistik yang biasa digunakan untuk menghubungkan antara variabel *dependent* dengan variabel *independent*. Variabel *independent*-nya dapat berbentuk kontinu, diskrit atau kategorik, maupun gabungan keduanya.

Sedangkan untuk variabel *dependent*-nya mengikuti probabilitas dari distribusi Bernoulli. Probabilitas akan bernilai 1 jika sukses dan dilambangkan dengan π , serta probabilitas bernilai 0 jika gagal ($1 - \pi$). Pendekatan model persamaan *logistic regression* digunakan karena dapat menjelaskan hubungan antara variabel *independent* dan peluangnya yang bersifat tidak linear, ketidaknormalan sebaran dari variabel *dependent*, serta keragaman respon yang tidak konstan dan tidak dapat dijelaskan oleh model regresi linear biasa [3]. Pada *logistic regression*, hubungan antara nilai yang akan diprediksi dengan variabel respon tidak linear, dikenal dengan transformasi logit yang mempunyai persamaan [5].

$$\pi = \frac{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k}}{1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k}} \quad (1)$$

Sehingga

$$Y = \ln\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k \quad (2)$$

dimana $\frac{\pi}{1-\pi}$ merupakan *probability odds* dan β_0, \dots, β_k merupakan parameter model.

Dalam mengestimasi parameter pada model *logistic regression*, terdapat beberapa metode antara lain:

1. Maximum Likelihood Estimation (MLE)

Mempunyai fungsi likelihood [4]

$$L(\beta) = \ln[l(\beta)] = \sum_{i=1}^n \{y_i \ln[\pi(x_i)] + (1 - y_i) \ln[1 - \pi(x_i)]\} \quad (3)$$

Dengan memaksimumkan persamaan (3) dan mendiferensialkannya terhadap β serta menyamakannya dengan 0 maka parameter β dapat diperoleh.

2. Least Square Method

Pada metode ini parameter β dipilih dengan meminimumkan jumlah eror kuadrat dari variabel input x dan nilai y yang diamati berdasarkan pada model. Persamaan (2) dapat dituliskan dalam notasi matriks menjadi [1]

$$y = X\beta \quad (3)$$

dengan

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix}, \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \quad (4)$$

Berdasarkan persamaan (4) dapat diperoleh model estimasi

$$\hat{y} = X\hat{\beta} \quad (5)$$

Sehingga penduga (estimator) adalah $\hat{\beta}$ sebagai berikut

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (6)$$

3. Hasil Percobaan

Langkah awal dalam melakukan penelitian ini adalah membagi data menjadi data *train* dan data *test*. Dalam penelitian ini digunakan beberapa pembagian data *train* dan data *test*. Selanjutnya dicari parameter $\hat{\beta}$ dengan menggunakan persamaan (6) dimana variabel X merupakan data *train* dan variabel Y merupakan aktual data. Tabel berikut merupakan hasil parameter $\hat{\beta}$ yang telah diperoleh

Tabel 2. Parameter $\hat{\beta}$

Data train					
	90%	80%	70%	60%	50%
$\hat{\beta}_0$	0,0494	0,0395	0,0639	0,0545	0,0612
$\hat{\beta}_1$	0,0724	0,1464	-0,3993	-11,461	-13,956
$\hat{\beta}_2$	0,1261	0,1373	0,1059	0,0861	0,1053
$\hat{\beta}_3$	0,0861	0,1066	0,1063	0,1426	0,1593
$\hat{\beta}_4$	-0,0185	-0,0316	-0,0399	-0,0379	-0,0274
$\hat{\beta}_5$	-0,0753	-0,0757	-0,0641	-0,0344	-0,0461
$\hat{\beta}_6$	-0,0994	-0,0974	-0,0871	-0,0675	-0,0926

Dari parameter $\hat{\beta}$ yang telah diperoleh dengan menggunakan beberapa data *train* maka persamaan (2) menjadi:

- Jika menggunakan 90% data *train*

$$Y = 0,0494 + 0,0724x_1 + 0,1261x_2 + 0,0861x_3 + (-0,0185)x_4 + (-0,0753)x_5 + (-0,0994)x_6 \quad (7)$$
- Jika menggunakan 80% data *train*

$$Y = 0,0395 + 0,1464x_1 + 0,1373x_2 + 0,1066x_3 + (-0,0316)x_4 + (-0,0757)x_5 + (-0,0974)x_6 \quad (8)$$
- Jika menggunakan 70% data *train*

$$Y = 0,0639 + (-0,3993)x_1 + 0,1059x_2 + 0,1063x_3 + (-0,0399)x_4 + (-0,0641)x_5 + (-0,0871)x_6 \quad (9)$$
- Jika menggunakan 60% data *train*

$$Y = 0,0545 + (-11,461)x_1 + 0,0861x_2 + 1426x_3 + (-0,0379)x_4 + (-0,0344)x_5 + (-0,0675)x_6 \quad (10)$$
- Jika menggunakan 50% data *train*

$$Y = 0,0612 + (-13,956)x_1 + 0,1053x_2 + 0,1593x_3 + (-0,0274)x_4 + (-0,0461)x_5 + (-0,0926)x_6 \quad (11)$$

Setelah parameter $\hat{\beta}$ ditemukan selanjutnya adalah mencari masing-masing hasil percobaan $\hat{\pi}$ menggunakan persamaan (1) dengan variabel x_i adalah data *test* (untuk data *train* 90% maka data *test* yang digunakan adalah 10%, untuk data *train* 80% maka data *test* yang digunakan adalah 20%, dst) dan parameter $\hat{\beta}$ seperti pada tabel 2. Dari semua hasil percobaan dan membandingkannya dengan aktual data, debitor yang mengalami kredit macet mempunyai nilai $\hat{\pi} \geq 0,5500$. Tabel berikut merupakan hasil percobaan.

Tabel 4. Hasil percobaan 10% data *test*

Aktual Data	Hasil Percobaan	
	Kredit Lancar	Kredit Macet
Kredit Lancar	25	0
Kredit Macet	5	0

Tabel 5. Hasil percobaan 20% data *test*

Aktual Data	Hasil Percobaan	
	Kredit Lancar	Kredit Macet
Kredit Lancar	53	1
Kredit Macet	6	0

Tabel 6. Hasil percobaan 10% data *test*

Aktual Data	Hasil Percobaan	
	Kredit Lancar	Kredit Macet
Kredit Lancar	82	0
Kredit Macet	8	0

Tabel 7. Hasil percobaan 10% data *test*

Aktual Data	Hasil Percobaan	
	Kredit Lancar	Kredit Macet
Kredit Lancar	110	0
Kredit Macet	11	0

Tabel 8. Hasil percobaan 10% data *test*

Aktual Data	Hasil Percobaan	
	Kredit Lancar	Kredit Macet
Kredit Lancar	135	4
Kredit Macet	11	1

Setelah mendapatkan hasil, selanjutnya adalah menentukan tingkat akurasi dengan menggunakan tabel matriks *confusion* berikut.

Tabel 9. Matriks *Confusion*

Forecast	Observed		
	Positive	Negative	Total
Positive	True Positive (TP)	False Positive (FP)	TP+FP
Negative	False Negative (FN)	True Negative (TN)	FN+TN

Total	TP + FN	FP+TN	TP+FP+FN+TN
--------------	---------	-------	-------------

Keterangan:

- *True positive* adalah jumlah data yang mengalami kredit lancar, dan hasil percobaan juga mengalami kredit lancar.
- *False positive* adalah jumlah data yang mengalami kredit macet, namun hasil percobaan dinyatakan mengalami kredit lancar.
- *True negative* adalah jumlah data yang mengalami kredit macet, dan hasil percobaan juga mengalami kredit macet.
- *False negative* adalah jumlah data yang sebenarnya mengalami kredit lancar namun hasil percobaan menyatakan mengalami kredit macet.

Dari tabel tersebut, untuk menentukan tingkat akurasi, digunakan rumus

$$Akurasi = \frac{TP+TN}{TP+FP+FN+TN} \times 100\% \quad (12)$$

Sehingga didapatkan tingkat akurasi seperti pada tabel berikut`

Tabel 10. Tingkat Akurasi

Data Test	Tingkat Akurasi
10%	83,33333%
20%	88,33333%
30%	91,11111%
40%	90,90909%
50%	90,06623%

4. Kesimpulan

Dengan melihat dan membandingkan antara hasil yang diperoleh dengan aktual data, maka dapat disimpulkan pada penelitian ini kredit yang bermasalah (kredit macet) mempunyai probabilitas lebih besar sama dengan 0.5500. Hasil yang diperoleh dengan menggunakan metode *logistic regression* pada penelitian ini mempunyai tingkat akurasi yang berbeda jika menggunakan data *train* dan data *test* yang berbeda seperti pada tabel 10, tingkat akurasi tertinggi diperoleh jika menggunakan 70% data *train* yaitu sebesar 91,11111%

Referensi

- [1] Sohn, S.Y., Kim, D.H., Yoon, J.H., 2016, *Technology credit scoring model with fuzzy logistic regression*. Applied soft computing 43. 150-158.
- [2] Sinungan., Muchdarsyah., 1989, *Uang dan Bank*. Rineka Cipta.
- [3] Agresti, Alan., 1990, *Categorical Data Analysis*. John Wiley & Sons, Inc.
- [4] Hosmer, David W., Lemeshow, Stanley, 2000, *Applied Logistic Regression* 2nd Edition. John Wiley & Sons, Inc.
- [5] *Logistic Regression* 2nd Edition. John Wiley & Sons, Inc.

Lampiran

Tabel Hasil percobaan 10% data *test*

No	Aktual Data	Y	II	Keterangan Hasil Percobaan
1	Kredit Lancar	0	0,5247	Kredit Lancar
2	Kredit Lancar	0	0,5199	Kredit Lancar
3	Kredit Lancar	0	0,5182	Kredit Lancar
4	Kredit Lancar	0	0,5389	Kredit Lancar
5	Kredit Macet	1	0,5428	Kredit Lancar
6	Kredit Lancar	0	0,5274	Kredit Lancar
7	Kredit Macet	1	0,5046	Kredit Lancar
8	Kredit Macet	1	0,5049	Kredit Lancar
9	Kredit Lancar	0	0,5288	Kredit Lancar
10	Kredit Lancar	0	0,4986	Kredit Lancar
⋮	⋮	⋮	⋮	⋮
30	Kredit Lancar	0	0.5149	Kredit Lancar

Tabel Hasil percobaan 20% data *test*

No	Aktual Data	Y	II	Keterangan Hasil Percobaan
1	Kredit Lancar	0	0,5169	Kredit Lancar
2	Kredit Lancar	0	0,5312	Kredit Lancar
3	Kredit Lancar	0	0,5006	Kredit Lancar
4	Kredit Lancar	0	0,5157	Kredit Lancar
5	Kredit Lancar	0	0,5202	Kredit Lancar
6	Kredit Lancar	0	0,5330	Kredit Lancar
7	Kredit Lancar	0	0,5040	Kredit Lancar
8	Kredit Lancar	0	0,5395	Kredit Lancar
9	Kredit Lancar	0	0,5273	Kredit Lancar
10	Kredit Lancar	0	0,5538	Kredit Macet
⋮	⋮	⋮	⋮	⋮
60	Kredit Lancar	0	0,5121	Kredit Lancar

Tabel Hasil percobaan 30% data *test*

No	Aktual Data	Y	II	Keterangan Hasil Percobaan
1	Kredit Lancar	0	0,4948	Kredit Lancar
2	Kredit Lancar	0	0,5139	Kredit Lancar
3	Kredit Lancar	0	0,5282	Kredit Lancar
4	Kredit Lancar	0	0,5419	Kredit Lancar
5	Kredit Lancar	0	0,5238	Kredit Lancar
6	Kredit Lancar	0	0,5053	Kredit Lancar
7	Kredit Macet	1	0,5143	Kredit Lancar
8	Kredit Lancar	0	0,5222	Kredit Lancar
9	Kredit Lancar	0	0,5423	Kredit Lancar
10	Kredit Lancar	0	0,5122	Kredit Lancar
⋮	⋮	⋮	⋮	⋮

90	Kredit Lancar	0	0,5079	Kredit Lancar
----	---------------	---	--------	---------------

Tabel Hasil percobaan 40% data *test*

No	Aktual Data	Y	II	Keterangan Hasil Percobaan
1	Kredit Lancar	0	0,5066	Kredit Lancar
2	Kredit Lancar	0	0,5213	Kredit Lancar
3	Kredit Lancar	0	0,5091	Kredit Lancar
4	Kredit Lancar	0	0,5206	Kredit Lancar
5	Kredit Lancar	0	0,5232	Kredit Lancar
6	Kredit Lancar	0	0,5145	Kredit Lancar
7	Kredit Lancar	0	0,5179	Kredit Lancar
8	Kredit Lancar	0	0,5203	Kredit Lancar
9	Kredit Lancar	0	0,5410	Kredit Lancar
10	Kredit Lancar	0	0,5459	Kredit Lancar
∴	∴	∴	∴	∴
121	Kredit Lancar	0	0,5032	Kredit Lancar

Tabel Hasil percobaan 50% data *test*

No	Aktual Data	Y	II	Keterangan Hasil Percobaan
1	Kredit Lancar	0	0,5153	Kredit Lancar
2	Kredit Lancar	0	0,5068	Kredit Lancar
3	Kredit Lancar	0	0,5420	Kredit Lancar
4	Kredit Lancar	0	0,5366	Kredit Lancar
5	Kredit Lancar	0	0,5097	Kredit Lancar
6	Kredit Lancar	0	0,4956	Kredit Lancar
7	Kredit Lancar	0	0,5193	Kredit Lancar
8	Kredit Lancar	0	0,5045	Kredit Lancar
9	Kredit Lancar	0	0,5406	Kredit Lancar
10	Kredit Lancar	0	0,4884	Kredit Lancar
∴	∴	∴	∴	∴
151	Kredit Lancar	0	0,5062	Kredit Lancar

MODEL PELUANG PUNAHNYA KUMAN MYCOBACTERIUM TUBERCULOSA YANG RESISTEN TERHADAP RIFAMPISIN DAN TUBERKULOSIS-MULTI DRUG RESISTANT (TB -MDR)

JONNER NAINGGOLAN

Jurusan Matematika FMIPA Universitas Cenderawasih Jayapura, jonn_cesil@yahoo.co.id

Abstrak. Pada tulisan ini mengkaji model peluang punahnya penyakit tuberkulosis (TB) yang berbasiskan pencabangan Galton-Watson dan pembangkit distribusi Binomial Negatif dengan pendekatan peluang berbobot. Respon kuman *Mycobacterium tuberculosis* terhadap rifampisin dan TB-MDR, dapat dilihat dengan tes *in-vivo*. Berdasarkan hasil test dapat dideteksi adanya *Mycobacterium tuberculosis* yang sensitif menunjukkan *Mycobacterium* negatif (gagal), dan resisten menunjukkan *Mycobacterium* positif (sukses). Kajian model difokuskan untuk memprediksi besarnya transmisi sukses dan probabilitas punahnya kuman *Mycobacterium tuberculosis* yang resisten terhadap *rifampisin* dan TB-MDR di suatu daerah populasi.

Kata Kunci : *Mycobacterium tuberculosis*, model peluang, transmisi sukses, *rifampisin* dan TB-MDR.

1. Pendahuluan

Penyakit tuberkulosis (TB) adalah penyakit infeksi paru-paru (*pulmones*) yang disebabkan oleh *Mycobacterium tuberculosis* (Varaine et al. [7]). Secara global penyakit tuberkulosis merupakan penyakit pembunuh nomor dua setelah penyakit jantung koroner (*World Health Organization* (WHO) [9]). Akibat penyakit tuberkulosis setiap hari sekitar 5.000 jiwa meninggal dunia (Gerberding et al. [2]). Berdasarkan data WHO pada tahun 2010, terdapat 8,8 juta kasus dan 1,35 juta meninggal akibat penyakit tuberkulosis (WHO [9]). Pencegahan individu yang masih sehat agar tidak terinfeksi penyakit tuberkulosis dapat dilakukan dengan cara memberikan vaksinasi *Bacillus Calmette Guerin* (BCG) kepada bayi yang berumur 2-3 bulan (WHO [9]). Individu yang divaksinasi BCG dapat memproteksi 70-80% infeksi penyakit tuberkulosis (Crofton et al. [1]). Pada tahun 1990 WHO telah membuat dan merealisasikan strategi penanggulangan penyakit tuberkulosis global yang dikenal sebagai strategi *Directly Observed Treatment Short-course* (DOTS) dan terbukti sebagai strategi penanggulangan yang secara ekonomis sangat efektif (Gerberding et al. [2]).

Di Indonesia diperkirakan telah terjadi 5100 kasus tuberkulosis *multidrug resistant* (TB-MDR) dari kasus TB baru yang tercatat pada tahun 2010. Hal ini membuat Indonesia masuk dalam daftar negara dengan masalah MDR-TB yang serius (WHO [9]). Lebih dari 90% isolat yang resisten *rifampisin* juga resisten terhadap *isoniazid*, sehingga resistensi terhadap *rifampisin* merupakan pertanda yang mewakili terjadinya MDR (Syarifudin dkk. [5]). Hasil penelitian yang diperoleh juga menunjukkan bahwa lebih dari 96% resistensi terhadap *rifampisin* terjadi akibat mutasi pada segmen 81-bp gen *rpoB* dari kodon 507-533 yang disebut sebagai *Rifampicin Resistance Determining Region* (Syarifudin dkk. [5]).

Diagnosis TB-MDR dengan uji sensitivitas obat memberikan hasil yang sangat sensitif dan spesifik tetapi membutuhkan waktu yang lama (6-9 minggu) mengingat laju pertumbuhan *Mycobacterium tuberculosis* yang sangat lambat (Syaifudin dkk. [5]). Saat ini, pemanfaatan berbagai metode molekular berbasis *Polymerase Chain Reaction* telah dirancang untuk mempercepat deteksi mutasi terkait masalah resistensi agen antimikroba (Syaifudin dkk. [5]).

Fokus utama DOTS adalah penemuan dan penyembuhan pasien, prioritas diberikan kepada pasien penyakit tuberkulosis aktif (WHO [9]). Masalah penanggulangan penyakit tuberkulosis dunia saat ini adalah timbulnya resistensi terhadap Obat Anti Tuberkulosis (OAT). Resistensi terhadap OAT dikategorikan: Monoresisten, poliresistensi, dan *extensively drug resistant* (XDR). Monoresisten yaitu resisten terhadap salah satu obat *isoniazid* atau *rifampisin*, poliresistensi atau resisten ganda *multidrug resistant* tuberkulosis yaitu resisten terhadap lebih dari satu OAT, dan *extensively drug resistant* yaitu resistensi terhadap MDR ditambah salah satu obat kapreomisin, kanamisin, dan amikasin (Gerberding et al. [2]). Pada tahun 1994, hasil survey WHO pada 135 negara dari 194 negara di seluruh dunia telah mengalami resistensi terhadap OAT, sedangkan laporan WHO tahun 2007 di seluruh dunia telah terjadi monoresistensi 10,3% dan poliresistensi 17,0% (WHO [9]).

Masalah resistensi terhadap OAT merupakan masalah besar pada penanggulangan penyakit tuberkulosis di Indonesia. Resistensi terhadap OAT ada yang primer dan sekunder. Resistensi primer yaitu seseorang yang baru pertama kali terinfeksi dari orang yang terinfeksi penyakit tuberkulosis yang resisten terhadap OAT, sedangkan resistensi sekunder yaitu diperoleh pada saat proses pengobatan (Gerberding et al. [2]). Resistensi terhadap OAT di Indonesia pada tahun 1995-1997 telah terjadi resistensi primer 4,6%-5,8% dan resistensi sekunder 22,95%-26,07% (Munir dkk. [4]). Ancaman resistensi terhadap OAT memunculkan wacana perlunya regulasi OAT serta menekankan urgensi ketersediaan obat lini kedua (Gerberding et al. [2]).

Masalah yang dibahas dalam penelitian ini yaitu bagaimana model untuk memprediksi besarnya probabilitas punahnya kuman *Mycobacterium tuberculosis* yang resisten terhadap *rifampisin* dan TB-MDR di suatu daerah populasi. Tujuan penelitian ini adalah menentukan besarnya probabilitas punahnya kuman *Mycobacterium tuberculosis* yang resisten terhadap *rifampisin* dan TB-MDR di suatu daerah populasi.

2. Model Probabilitik Punahnya Kuman *Mycobacterium Tuberculosis*

2.1 Obat Antituberkulosis Lini Pertama dan Resistensi

Penanggulangan penyakit TB sampai saat ini obat utama yang disarankan WHO adalah *isoniazid*, *rifampisin*, *pirazinamid*, *etambutol* dan *streptomisin*. Obat antituberkulosis dengan model yang dikaji dalam tulisan ini adalah *rifampisin* dan TB-MDR. Kendala paling utama dalam usaha pemberantasan TB adalah resistensi *Mycobacterium tuberculosis* terhadap obat antituberkulosis. Resistensi adalah kemampuan kuman untuk berkembang biak walaupun telah diberi obat standar (Iloni [3]). *Rifampisin* adalah salah satu obat antituberkulosis utama kelompok antimikobakterial dan digunakan untuk pengobatan berbagai jenis infeksi. Sering digunakan dalam bentuk kombinasi dengan antibakterial lainnya untuk menghindari resistensi dan untuk pengobatan tuberkulosis. Penggunaan pada

pasien dewasa secara umumnya adalah 600 mg per hari melalui mulut pada keadaan lambung kosong. Sedangkan pada pasien anak-anak diberikan dosis 10 mg/kg hingga 20 mg/kg per hari dengan batas maksimum 600 mg per hari.

Munculnya mutasi pada gen-gen yang menjadi sasaran obat sehingga obatnya tidak efektif. Teknik yang akurat untuk mendeteksi keberadaan bakteri yang resisten terhadap obat dibutuhkan saat ini karena tingkat akurasi teknik konvensional yang ada adalah rendah dan kurang spesifik sehingga perlu penerapan teknik biologi molekuler.

Untuk mengetahui mekanisme molekuler resistensi *Mycobacterium tuberculosis* terhadap monoresisten rifampin dan multiresisten TB dengan menggunakan tes *in-vivo*, yaitu dengan pemeriksaan sampel TB berupa sputum dan pus yang diperiksa pada Laboratorium Mikrobiologi. Sputum yaitu bahan yang dikeluarkan dari bronkus dan paru yang dalam hal ini kemungkinan mengandung Batang Tahan Asam (BTA). Pus yaitu cairan hasil proses peradangan yang terbentuk dari sel darah putih (leukosit) dan cairan encer yang sering disebut nanah. Salah satu usaha untuk memberikan informasi pada bidang kesehatan tentang model peluang punahnya kuman terhadap obat antituberkulosis diuraikan beberapa model matematik berikut.

2.2 Resistensi Kuman *Mycobacterium Tuberculosis*

Pengamatan penderita tuberkulosis dilakukan dengan cara pemeriksaan dan pengobatan. Untuk mengetahui seseorang penderita penyakit tuberkulosis masih sensitif terhadap rifampisin digunakan dengan pemeriksaan sputum dan pus pada individu yang terinfeksi tuberkulosis. Pembacaan dilakukan setelah masa inkubasi 4-6 minggu, dimana kontrol 10^{-3} lebih dari 500 koloni, separuh media tertutup oleh koloni yang terisolir, 200-500 koloni (Ilone [3]). Jika jumlah koloni pada media berisi obat sama atau lebih dibandingkan dengan jumlah koloni pada kontrol 10^{-5} maka dinyatakan resisten. Jika jumlah koloni pada media berisi obat tidak ada atau kurang dibandingkan dengan jumlah koloni pada kontrol 10^{-5} maka dinyatakan sensitif (S), sensitif menunjukkan kuman negatif (gagal). Resisten (R_I , R_{II} , R_{III}) menunjukkan kuman positif (sukses). Resistensi pertama yaitu *Mycobacterium tuberculosis* yang monoresisten terhadap obat anti tuberkulosis rifampisin (R_I), dan resistensi terhadap *multidrug resistant tuberculosis* (TB-MDR): isoniazid + rifampisin, isoniazid + etambutol + rifampisin, isoniazid + streptomisin + rifampisin (R_{II}), TB-MDR lain: etambutol + rifampisin, streptomisin + etambutol + rifampisin (R_{III}) (Ilone [3]).

2.3 Kombinasi Permodelan Binomial Negatif dengan Teorema Galton-Watson

Untuk memprediksi peluang punahnya kuman *Mycobacterium tuberculosis* dan penyebaran populasi digunakan perpaduan antara permodelan Binomial Negatif dengan teorema Galton-Watson. Besarnya probabilitas Binomial Negatif $b^*(x; k; p)$ (Taneyhill, et al., 1999) ditentukan oleh rumus:

$$b^*(x; k; p) = \binom{x-1}{k-1} p^k q^{x-k}, x = k, k+1, k+2, \dots$$

Pada teorema Galton-Watson, jika s adalah peluang bahwa proses akan punah, maka harus memenuhi persamaan $G(s) = s$, dengan kata lain: peluang kepunahan adalah titik keseimbangan $G(s)$. Sebuah proses pencabangan akan punah secara

pasti jika 1 adalah satu-satunya titik keseimbangan $G(s)$ atau akan berakhir dengan peluang $s < 1$, jika $G(s)$ mempunyai 2 titik keseimbangan yang mungkin muncul dalam interval $0 \leq s \leq 1$. Fungsi pembangkit peluang $G(s)$, dengan $0 \leq s \leq 1$ dinyatakan sebagai berikut (Taneyhill et al. [6]):

$$G(s) = \sum_{g=0}^{\infty} p_g s^g = p_0 + p_1 s + p_2 s^2 + p_3 s^3 + \dots, \quad (1)$$

dimana s adalah peubah acak punahnya partikel. Agar deret ini konvergen sebagai akibat adanya suku kuadrat, pangkat tiga dan seterusnya, maka $0 \leq s \leq 1$. $G(s)$ dapat dinyatakan dalam bentuk yang lebih sederhana, yaitu: $G(s) = (1 - q + qs)^N$ sebagai fungsi pembangkit peluang dari distribusi Binomial

$$p_g = \binom{N}{g} q^g (1 - q)^{N-g},$$

dimana g adalah hasil yang diharapkan terjadi, N adalah ukuran sampel, q adalah peluang bahwa sebuah unsur (item) dalam sampel merupakan hasil yang diharapkan.

2.4 Transmisi Sukses

Transmisi sukses dipengaruhi oleh faktor variabilitas pada proses yang terakhir, obat antituberkulosis dengan *Mycobacterium tuberculosis*. Transmisi yang terjadi pada hospes terbagi atas 2 jenis yaitu: transmisi sukses dengan peluang p dan transmisi gagal dengan peluang $1 - p = q$. Banyaknya transmisi sukses dinyatakan dengan k , merupakan transmisi yang menyebabkan individu terinfeksi, akibat antituberkulosis tidak mampu melawan kuman *Mycobacterium tuberculosis*.

Jumlah transmisi sukses oleh satu *Mycobacterium tuberculosis* adalah ekspektasi bersyarat yang dinyatakan dalam R^* , dan diperoleh melalui langkah-langkah berikut:

$$E[R = y] = \sum_{y=0}^{\infty} y \frac{P(R^*, R)}{P(R=y)} = \frac{1}{1-p^k} \sum_{y=1}^{\infty} y p^k q^y \binom{y+k-1}{y}.$$

Misalkan $z = y - 1$. Maka

$$E[R^* | R = y] = \frac{1}{1-p^k} \sum_{z=0}^{\infty} p^k q^{z+1} \frac{k+z}{z!(k-1)!} = \left(\frac{1}{1-p^k} \right) \left(\frac{k(1-p)}{p} \right) = \frac{\bar{R}}{1-p^k} \quad (2)$$

dimana:

y = frekuensi transmisi; $\bar{R} = \frac{k(1-p)}{p}$ adalah laju transmisi dalam satuan orang/tahun; p = peluang sukses kuman masih mampu bertahan hidup; k = banyaknya kuman yang sukses. Misalkan nilai taksiran peluang p adalah \hat{p} dan taksiran k adalah k^* , maka model transmisi sukses menjadi $R^* = \frac{\bar{R}}{1-\hat{p}^{k^*}}$.

2.5 Model Probabilitas Punahnya Parasit

Model probabilitas punahnya kuman *Mycobacterium tuberculosis* diperoleh dengan menggabungkan fungsi pembangkit probabilitas distribusi Binomial Negatif dengan titik punah proses Galton-Watson yang dirumuskan sebagai berikut:

$$\varphi_{NB} = \sum_{x=0}^{\infty} f(x) s^x = \left(\frac{\hat{p}}{1-(1-\hat{p})s} \right)^{k^*}, \quad (3)$$

dan variansinya adalah:

$$\sigma_{NB}^2 = \frac{k^*(1-\hat{p})}{\hat{p}^2},$$

dimana: $f(x)$ = fungsi pembangkit probabilitas Binomial Negatif; s^x = fungsi pembangkit probabilitas punahnya kuman.

2.6 Pengujian Kelayakan Model (*Goodness of fit test*)

Data yang digunakan dalam model ini adalah distribusi Binomial Negatif. Data di lapangan yaitu data resistensi kuman *Mycobacterium tuberculosis* yang resisten terhadap obat antituberkulosis di suatu daerah populasi. Hipotesis testing digunakan untuk uji kelayakan model (*Goodness of fit test*). Uji kelayakan model dilakukan untuk menguji kesesuaian data dengan asumsi dasar dalam permodelan. Langkah kerja *goodness fit test* antara frekwensi observasi dan ekspektasi didasarkan pada:

$$\chi^2 - \text{hitung} = \sum_{i=1}^r \frac{(O_i - E_i)^2}{E_i},$$

(4)

dengan χ^2 : Nilai variabel random yang berdistribusi samplingsnya sangat mendekati distribusi kai-kuadrat dengan derajat kebebasan $\nu = r - 1$, O_i = frekuensi observasi yang ke- i , E_i = frekuensi ekspektasi yang ke- i . Dalam hal ini hipotesis ditolak jika χ^2 -hitung > χ^2 -tabel = $\chi^2_{\alpha, r-1}$, untuk taraf signifikansi α (Walpole and Myers [8]). Sebagai aplikasinya diberikan data sebagian wilayah Jakarta.

3. Hasil Dan Pembahasan

Berdasarkan sampel TB berupa sputum dan pus yang diperiksa pada Laboratorium Mikrobiologi Klinik FKUI yang berasal dari wilayah Jakarta dan sekitarnya pada periode September 2005 hingga Desember 2007, dan sesuai dengan kriteria inklusi yang ada diperoleh jumlah data yang digunakan sebagai sampel penelitian sejumlah 676 individu (Ilone [3]) seperti pada Tabel 1.

Tabel 1. Frekuensi Observasi dan Ekspektasi Tingkat Resistensi di Suatu Populasi

Test Resistensi	Frekuensi Observasi	Probabilitas	Frekuensi Ekspektasi
<i>In-vivo</i>	O_i	\hat{p}	$E_i = n\hat{p}$
R_I	49	1/3	154
R_{II}	59	1/3	154
R_{III}	54	1/3	154

Dari data pada Tabel 1 di atas, dapat ditentukan titik punahnya kuman *Mycobacterium tuberculosis* yang resisten terhadap *rifampisin*, *TB-MDR* dan *TB-MDR lainnya* dengan menggunakan fungsi pembangkit momen Teorema Galton-Watson dari persamaan (1), $G(s) = p_0 + p_1s + p_2s^2 + p_3s^3$. Nilai-nilai p_0 , p_1 , p_2 dan p_3 merupakan peluang terjadinya sensitif (S) dan resistensi derajat R_I , R_{II} dan R_{III} . Sehingga diperoleh s yang memenuhi persamaan ini adalah $s = 0.996$.

(1) Dengan menggunakan persamaan (4) untuk taraf signifikansi $\alpha = 0,05$ diperoleh χ^2 -hitung =

$0,93 < \chi^2$ -tabel = 5,99, sehingga data tersebut di atas berdistribusi Binomial Negatif.

(2) Transmisi sukses dengan menggunakan persamaan (2) diperoleh $R^* = 27$ orang per tahun.

(3) Peluang punahnya kuman *Mycobacterium tuberculosis* yang resisten terhadap obat antituberkulosis dengan menggunakan persamaan (3) diperoleh $\phi_{NB} = 0.8$.

4. Kesimpulan

Model probabilistik punahnya kuman *Mycobacterium tuberculosis* yang resisten terhadap obat antituberkulosis di suatu daerah populasi diperoleh rumusan permodelan sebagai berikut:

(1) Model transmisi sukses pada penderita di suatu daerah populasi rumusannya berbentuk:

$$R^* = \frac{\bar{R}}{1-\hat{p}k^*}.$$

Berdasarkan perhitungan diperoleh transmisi sukses kuman *Mycobacterium tuberculosis* adalah 27 orang per tahun.

(2) Model probabilitas punahnya kuman *Mycobacterium tuberculosis* yang resisten terhadap suatu obat antituberkulosis di suatu daerah populasi rumusannya berbentuk:

$$\phi_{NB} = \left(\frac{\hat{p}}{1-(1-\hat{p})s} \right)^{k^*}.$$

Berdasarkan perhitungan dari data yang diperoleh peluang punahnya kuman *Mycobacterium tuberculosis* berdasarkan data yang diperoleh di laboratorium FKUI pada bulan September 2005 sampai Desember 2007 di wilayah Jakarta (Ilone [3]) adalah 0,8, hal ini berarti cukup besar peluang punahnya kuman *Mycobacterium tuberculosis*. Namun peluang punahnya kuman *Mycobacterium tuberculosis* tersebut belum tentu sama dengan di daerah yang lain.

Referensi

- [1] Crofton, S.J., Horne, M., and Miller, F., 2002, *Clinical Tuberculosis*, MacMillan Education Ltd, London.
- [2] Gerberding, J. L. et al., 2003, Treatment of Tuberculosis, *American Thoracic Society, CDC, and Infectious Diseases Society of America, MMWR*, 52, (11), 1-32.
- [3] Ilone, S., 2008, Deskripsi Monoresisten Rifampisin, Multi-Drug, Serta Multiresisten *Mycobacterium Tuberculosis* pada Sampel yang Diteliti Departemen Mikrobiologi FKUI Periode September 2005-Desember 2007. *Jurnal FKUI*, 4, 34-40.
- [4] Munir, S.M., Nawas, A., dan Soetoyo, D. K., 2010, Pengamatan Pasien Tuberculosis Paru dengan Multidrug Resistant (TB-MDR) di Poliklinik Paru RSUP Persahabatan, *J. Respir Indo*, 30 (2), 90-104.
- [5] Syaifudin, M., Rosilawati, M.L., Irawan, H., dan Bela, B., 2007, Identifikasi *Mycobacterium Tuberculosis* dan Analisis Mutasi Gen RpoB dan KatG Penyebab Resistensi Ganda dengan Teknik Molekuler. Jakarta: *Laporan Penelitian Balitbang BATAN*.
- [6] Taneyhill D.E., Dunn, A.M., and Hatcher, M.J., 1999, The Galton-Watson Branching Process a Quatitative Tool in Parasitology. *Parasitology Today*, 15, 159-165.
- [7] Varaine, F. et al., 2013, *Tuberculosis*, Practical Guide for Clinicians, Nurses, Laboratory Technicians and Medical Auxiliaries, *Midicians Sans Frontieres*.
- [8] Walpole, R.E. and Myers, R. H., 1986, *Probability and Statistic for Engineers and Scientists*, Macmillan Publishing Company, New York.
- [9] WHO, 2012, Global Tuberculosis Control 2012, *WHO Library Cataloguing-in-Publication Data (2011)*.

PENGARUH *FOREIGN DIRECT INVESTMENT (FDI)* TERHADAP KETIMPANGAN PENDAPATAN DI INDONESIA: APLIKASI *IV-2SLS MODEL*

RIBUT NURUL TRI WAHYUNI¹

1 Sekolah Tinggi Ilmu Statistik, rnurult@stis.ac.id

Abstrak. *Foreign Direct Investment (FDI)* merupakan kegiatan menanam modal untuk melakukan usaha di wilayah negara Republik Indonesia yang dilakukan oleh penanam modal asing dalam bentuk pembelian aset-aset perusahaan (barang modal). Sedangkan gini rasio digunakan sebagai ukuran ketimpangan pendapatan di suatu negara. Selama periode 2007-2014, FDI dan gini rasio di Indonesia menunjukkan pergerakan yang sama, yaitu cenderung meningkat, padahal tujuan utama dari FDI adalah mengurangi ketimpangan pendapatan. Kontroversi tentang bagaimana pengaruh FDI terhadap ketimpangan pendapatan terjadi di kalangan para peneliti. Penelitian ini bertujuan untuk mengetahui pengaruh FDI terhadap ketimpangan pendapatan di Indonesia pada tahun 2014 dengan menggunakan *Instrument Variable-Two Stage Least Square (IV-2SLS) model*. Hasil analisis inferensia menunjukkan bahwa FDI berpengaruh signifikan positif terhadap ketimpangan pendapatan. Dari hasil penelitian ini disarankan kepada pemerintah perlu mengevaluasi kembali kebijakan perdagangan internasional khususnya FDI agar pemerataan pendapatan penduduk dapat tercapai.

Kata kunci : FDI, ketimpangan pendapatan, gini rasio, *IV-2SLS model*.

1. Pendahuluan

Di negara berkembang kegiatan perekonomian dapat berkembang lebih cepat apabila tersedia lebih banyak modal dan kemampuan untuk menggunakan tambahan modal dengan lebih efektif. Hal ini dapat diatasi dengan adanya modal dari luar negeri, terutama dari negara maju. Menurut Undang-Undang Republik Indonesia Nomor 25 Tahun 2007 tentang Penanaman Modal, yang dimaksud dengan Penanaman Modal Asing (PMA) adalah kegiatan menanam modal untuk melakukan usaha di wilayah negara Republik Indonesia yang dilakukan oleh penanam modal asing, baik yang menggunakan modal asing sepenuhnya maupun yang berpatungan dengan penanam modal dalam negeri. Tujuan penanaman modal menurut Undang-Undang Republik Indonesia Nomor 25 Tahun 2007 tentang Penanaman Modal adalah meningkatkan pertumbuhan ekonomi nasional, menciptakan lapangan kerja, meningkatkan pembangunan ekonomi berkelanjutan, meningkatkan kemampuan daya saing dunia usaha nasional, meningkatkan kapasitas dan kemampuan teknologi nasional, mendorong pengembangan ekonomi kerakyatan, mengolah ekonomi potensial menjadi kekuatan ekonomi riil dengan

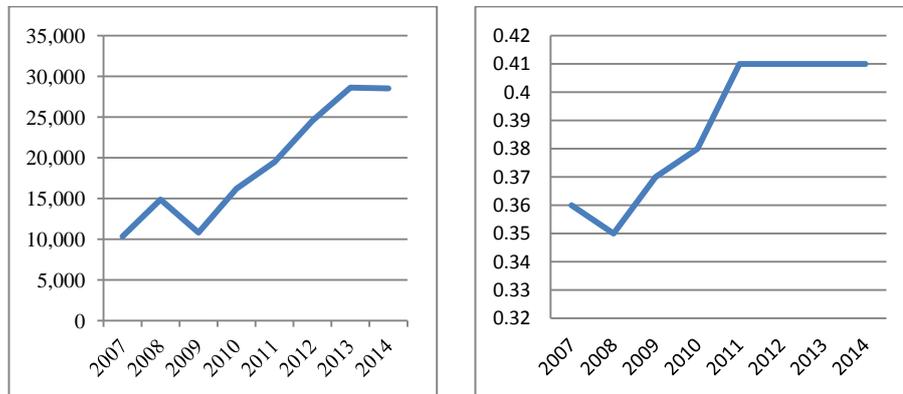
menggunakan dana yang berasal, baik dari dalam negeri maupun luar negeri, dan meningkatkan kesejahteraan masyarakat.

PMA dapat dilakukan dalam dua bentuk, yaitu: investasi portofolio dan *Foreign Direct Investment* (FDI) (Salvatore [1]). Investasi portofolio adalah investasi yang hanya melalui aset finansial seperti obligasi dan saham dalam mata uang domestik. Investasi ini biasanya dilakukan melalui lembaga keuangan seperti bank dan perusahaan investasi. Sedangkan FDI merupakan investasi dalam bentuk aset-aset perusahaan seperti barang modal. FDI dapat dilakukan dengan membeli sebagian atau seluruh saham. Menurut Todaro dan Smith [2], FDI merupakan investasi luar negeri yang dilakukan oleh perusahaan multinasional swasta, yang mana dana-dana investasinya langsung digunakan untuk menjalankan kegiatan bisnis seperti mendatangkan mesin-mesin dan membeli bahan baku produksi.

Motivasi yang melatarbelakangi investor untuk menanamkan modalnya ke luar negeri ada empat. Pertama, FDI ditujukan untuk memperluas pasar dengan mencari pasar baru atau mempertahankan pasar lama (*market seeking*). Biaya transportasi dan kebijakan pemerintah merupakan alasan utama motivasi ini. Kedua, FDI dilakukan untuk mendapatkan faktor-faktor produksi yang berbentuk sumber daya alam atau tenaga kerja yang lebih murah di negara lain (*resource seeking*). FDI jenis ini biasanya dilakukan oleh perusahaan-perusahaan yang berbasis di sektor primer. Ketiga, FDI dilakukan untuk meningkatkan efisiensi karena bisa meningkatkan skala ekonomi (*efficiency seeking*). Keempat, perusahaan multinasional menggunakan FDI untuk mendapatkan aset strategis, baik berwujud atau tidak berwujud, yang penting untuk strategi jangka panjang mereka, tetapi tidak tersedia di negara asal (*strategic asset seeking*). Hymer dalam Jensen dan Rosas [3] berpendapat bahwa FDI terjadi karena pasar yang tidak sempurna. Perusahaan asing lebih memilih melakukan investasi di negara lain dalam bentuk FDI dibanding memproduksi barang tersebut di negaranya sendiri karena membuat produk multinasional dianggap lebih menguntungkan. Berdasarkan alasan tersebut, maka dapat disimpulkan bahwa FDI dipengaruhi oleh variabel lain, misalnya jumlah penduduk, sumber daya alam, upah tenaga kerja, pendapatan penduduk, dan variabel lainnya di negara tujuan FDI.

Perusahaan multinasional berbeda jika dibandingkan dengan perusahaan dalam negeri. Perusahaan multinasional biasanya lebih banyak membutuhkan tenaga kerja *high skill* dan upah yang diterima tenaga kerja mereka cenderung lebih tinggi dibanding upah yang diterima tenaga kerja di perusahaan dalam negeri. Selain itu, perusahaan multinasional juga menggunakan teknologi yang baru. Penelitian sebelumnya menunjukkan bahwa ada hubungan antara FDI dari perusahaan multinasional dengan ketimpangan pendapatan di negara tujuan FDI. Pertama, perusahaan multinasional akan membawa modal ke negara tujuan sehingga menurunkan *returns to capital* di negara tujuan dan meningkatkan upah tenaga kerja. Modal asing akan bersaing dengan modal dalam negeri sehingga akan meningkatkan upah tenaga kerja dan menurunkan *returns to capital* perusahaan dalam negeri. Dampak ini akan mempercepat konvergensi dari upah tenaga kerja relatif terhadap *returns to capital* sehingga menurunkan ketimpangan pendapatan. Kedua, perusahaan multinasional akan membayar tenaga kerja di atas perusahaan dalam negeri. Jika perusahaan multinasional meningkatkan upah tenaga kerja *skill*, maka kondisi ini akan meningkatkan perbedaan pendapatan antara tenaga kerja *skill* dan *unskill* serta menurunkan perbedaan pendapatan antara tenaga kerja *skill* dengan pemilik modal. Sebaliknya, jika perusahaan multinasional membutuhkan tenaga kerja *unskill* dan membayar mereka di atas upah yang dibayarkan

perusahaan dalam negeri, maka FDI akan menurunkan ketimpangan pendapatan karena peningkatan pendapatan tenaga kerja umumnya terjadi di distribusi pendapatan paling bawah.



(a) FDI (ribu dollar AS)

(b) Gini rasio

Gambar 1. FDI dan gini rasio Indonesia tahun 2007-2014

Sejak pemberlakuan Undang-Undang Republik Indonesia Nomor 25 Tahun 2007, perkembangan FDI dan ketimpangan pendapatan menunjukkan tren yang sama, yaitu cenderung meningkat (Gambar 1). Ini mengindikasikan kemungkinan adanya hubungan antara FDI dan ketimpangan pendapatan. Data FDI yang digunakan merupakan data *flow*, bukan *stock*, sesuai dengan penelitian Jensen dan Rosas [3] serta Blanc-Brude *et. al.* [4]. Ketimpangan pendapatan menggunakan indikator gini rasio. Penelitian-penelitian sebelumnya memberikan hasil yang berbeda mengenai dampak FDI terhadap ketimpangan pendapatan (positif (Wood; Reuveny dan Li dalam Jensen dan Rosas [3]) dan negatif (Jensen dan Rosas [3])). Di negara maju, FDI bisa meningkatkan ketimpangan pendapatan. Sedangkan di negara berkembang, FDI bisa menurunkan ketimpangan pendapatan (Wood dalam Jensen dan Rosas [3]). Berdasarkan permasalahan tersebut, tujuan dalam penelitian ini yang ingin dicapai adalah mengetahui bagaimana pengaruh FDI terhadap ketimpangan pendapatan di Indonesia.

2. Hasil – Hasil Utama

Metode Penelitian

Data yang digunakan dalam penelitian ini berupa data *cross section* di 33 provinsi di Indonesia tahun 2014 yang berasal dari data Badan Pusat Statistik (BPS). Sebagaimana dijelaskan sebelumnya, FDI dipengaruhi oleh variabel lain sehingga FDI tidak bisa diberlakukan sebagai variabel eksogen. Jika tetap diberlakukan sebagai variabel eksogen dan estimasi parameternya menggunakan model OLS, maka hasil estimasinya menjadi tidak efisien, bias, dan tidak konsisten. Oleh karena itu, diperlukan *instrument variable* untuk mendapatkan model FDI. Untuk menjelaskan hubungan antara FDI dan gini rasio, langkah pertama adalah mengestimasi model FDI (persamaan 2), kemudian hasil estimasi FDI digunakan untuk membuat model gini rasio (persamaan 1). Dengan menggunakan metode tersebut, kita bisa mencari model regresi *instrument variable two-stage least squares* (IV-2SLS) dengan persamaan sebagai berikut:

$$gini_{2014} = \beta_0 + \beta_1 \ln FDI_{2014}^* + \beta_2 tpt_{2014} + \beta_3 po_{2014} \quad (1)$$

$$\ln FDI_{2014}^* = \gamma_0 + \gamma_1 \ln pdrbk_{2014} + \gamma_2 \ln ump_{2014} + \gamma_3 tpt_{2014} + \gamma_4 \ln FDI_{2013} \quad (2)$$

Gini rasio di Indonesia pada tahun 2014 ($gini_{2014}$) merupakan fungsi dari instrumen FDI tahun 2014 ($\ln FDI_{2014}^*$), tingkat pengangguran terbuka tahun 2014 (tpt_{2014}), dan persentase penduduk miskin tahun 2014 (po_{2014}). Instrumen FDI tahun 2014 ($\ln FDI_{2014}^*$) diperoleh dari regresi antara FDI tahun 2014 dalam juta dollar AS ($\ln FDI_{2014}$) dengan variabel instrumen PDRB atas dasar harga konstan tahun 2014 (2010=100) dalam milyar rupiah ($\ln pdrbk_{2014}$), upah minimum provinsi perbulan tahun 2014 dalam rupiah ($\ln ump_{2014}$), tingkat pengangguran terbuka tahun 2014 (tpt_{2014}), dan FDI tahun 2013 dalam juta dollar AS ($\ln FDI_{2013}$).

Untuk melihat apakah variabel instrumen di atas bisa digunakan sebagai instrumen FDI, maka dilakukan *overidentifying restriction test*. Syarat model *overidentified* penting untuk menjamin bahwa asumsi tidak ada korelasi antara variabel instrumen dan residual terpenuhi. Wooldridge [5] menjelaskan tentang langkah-langkah melakukan *overidentifying restriction test*. Pertama, estimasi persamaan struktural dengan IV-2SLS (persamaan 1) dan diperoleh residualnya, misalkan e_i . Kedua, regresikan e_i terhadap seluruh variabel instrumen dengan metode OLS sehingga diperoleh R^2 . Ketiga, dengan hipotesis null bahwa seluruh variabel instrumen tidak berkorelasi dengan e_i , maka $nR^2 \sim \chi_{df}^2$ dengan df adalah jumlah variabel instrumen dikurangi jumlah variabel endogen. Jika nilai statistik uji melebihi nilai kritis pada $\alpha = 5$ persen, maka hipotesis null ditolak dan dapat disimpulkan bahwa paling tidak ada satu variabel instrumen yang tidak eksogen. Setelah mendapatkan model regresi IV-2SLS, langkah selanjutnya adalah melakukan uji asumsi klasik (non multikolinearitas antar variabel, homoskedastisitas residual, dan kenormalan residual) pada persamaan 1 dan persamaan 2.

Hasil dan Pembahasan

Hasil dari regresi IV-2SLS dengan menggunakan variabel instrumen untuk FDI bisa dilihat di Tabel 1 dan hasilnya sudah sesuai teori. Semua variabel instrumen berkorelasi dengan FDI sehingga instrumen yang digunakan sudah valid dengan $R^2 = 0,9007$ (Tabel 1). Sedangkan regresi antara e_i dengan variabel instrumen dapat dilihat pada Tabel 2. *Overidentifying restriction test* diperoleh $nR^2 = 0$ dengan $\chi_{3;0,05}^2 = 7,815$ dan dapat disimpulkan bahwa semua variabel instrumen adalah variabel eksogen (model *overidentified*).

Tabel 1. Hasil regresi OLS variabel FDI dan variabel instrumen

Source	SS	df	MS			
Model	98.2755155	4	24.5688789	Number of obs = 33		
Residual	10.834143	28	.386933678	F(4, 28) = 63.50		
Total	109.109658	32	3.40967683	Prob > F = 0.0000		
				R-squared = 0.9007		
				Adj R-squared = 0.8865		
				Root MSE = .62204		

	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lnfdi2014						
lnpdrbk2014	.4028979	.1306202	3.08	0.005	.1353346	.6704612
lnump2014	.8929984	.5030526	1.78	0.087	-.1374581	1.923455
tpt2014	-.1880604	.0578396	-3.25	0.003	-.3065395	-.0695813
lnfdi2013	.7993869	.0836014	9.56	0.000	.6281373	.9706366
_cons	-15.22009	7.348799	-2.07	0.048	-30.27343	-.1667615

Tabel 2. Hasil regresi OLS e_i dan variabel instrumen

Source	SS	df	MS			
Model	1.7764e-15	4	4.4409e-16	Number of obs = 33		
Residual	10.8341432	28	.386933686	F(4, 28) = 0.00		
Total	10.8341432	32	.338566975	Prob > F = 1.0000		
				R-squared = 0.0000		
				Adj R-squared = -0.1429		
				Root MSE = .62204		

e	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lnpdrbk2014	-1.30e-09	.1306202	-0.00	1.000	-.2675633	.2675633
lnump2014	-9.53e-09	.5030526	-0.00	1.000	-1.030457	1.030456
tpt2014	-1.43e-09	.0578396	-0.00	1.000	-.1184791	.1184791
lnfdi2013	1.22e-09	.0836014	0.00	1.000	-.1712497	.1712497
_cons	1.54e-07	7.3488	0.00	1.000	-15.05333	15.05333

Tabel 3. Hasil regresi IV-2SLS

Source	SS	df	MS			
Model	.008814999	3	.002938333	Number of obs = 33		
Residual	.036809244	29	.001269284	F(3, 29) = 2.31		
Total	.045624242	32	.001425758	Prob > F = 0.0966		
				R-squared = 0.1932		
				Adj R-squared = 0.1097		
				Root MSE = .03563		

	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
gini2014						
estlnfdi2014	.0090038	.0038774	2.32	0.027	.0010736	.016934
tpt2014	-.0019007	.0030756	-0.62	0.541	-.008191	.0043896
po2014	.0020716	.0011189	1.85	0.074	-.0002168	.00436
_cons	.3113299	.0323076	9.64	0.000	.2452535	.3774063

Tabel 4. Hasil uji asumsi homoskedastisitas dan kenormalan residual

Persamaan	<i>p-value</i> <i>Breusch-Pagan test</i>	<i>p-value</i> <i>Shapiro-Wilk W test</i>
1	0,2400	0,53147
2	0,1617	0,07253

Setelah diperoleh regresi IV-SLS, dilakukan uji asumsi pada persamaan gini rasio dan instrumen FDI. Kesimpulannya tidak ada pelanggaran asumsi residual dengan tingkat kepercayaan 95 persen (residual berdistribusi normal dan homoskedastis) dan antar variabel tidak ada multikolinearitas (korelasi antar variabel kurang dari 0,8 (Gujarati [6]).

Dampak FDI terhadap gini rasio adalah positif, artinya liberalisasi FDI akan mendorong peningkatan gini rasio. Ketika FDI meningkat sebesar 1 persen, maka gini rasio akan meningkat sebesar 0,0090038 dengan tingkat kepercayaan 95 persen, asumsi variabel lain konstan. Hasil ini sesuai dengan penelitian Reuveny dan Li dalam Jensen dan Rosas [3], *flow* FDI bisa meningkatkan ketimpangan pendapatan. Sedangkan Alderson dan Nielsen dalam Jensen dan Rosas [3],

menemukan hubungan yang positif antara *stock* FDI dan ketimpangan pendapatan. FDI di Indonesia cenderung membutuhkan tenaga kerja *skill* sehingga upah tenaga kerja *skill* akan meningkat, sedangkan upah tenaga kerja *unskill* cenderung tetap atau bahkan turun. Akibatnya, perbedaan pendapatan antara tenaga kerja *skill* dan *unskill* akan cenderung meningkat dan perbedaan pendapatan antara tenaga kerja *skill* dan pemilik modal akan cenderung menurun. Secara total, FDI akan meningkatkan ketimpangan pendapatan (gini rasio).

4. Kesimpulan

Dengan menggunakan *IV-2SLS model* dapat dilihat bahwa FDI signifikan positif memengaruhi ketimpangan pendapatan di Indonesia tahun 2014. Kemungkinan FDI mengakibatkan permintaan tenaga kerja *skill* meningkat sehingga upahnya meningkat. Sebaliknya, peningkatan FDI tidak banyak menyerap tenaga kerja *unskill* sehingga upahnya cenderung tetap atau bahkan turun. Akibatnya, ketimpangan pendapatan di Indonesia semakin meningkat.

Referensi

- [1] Salvatore, D., 2012, *International Economics*, Wiley Global Education.
- [2] Todaro, M., Smith, S., 2003, *Development Economics*, Pearson Education.
- [3] Jensen, N. M., Rosas, G., 2007, Foreign Direct Investment and Income Inequality in Mexico, 1990-2000, *International Organization Vol. 61, No. 3 Summer, 2007*, 467-487.
- [4] Blanc-Brude, F., Cookson, G., Piesse, J., Strange, R., 2014, The FDI Location Decision: Distance and The Effects of Spatial Dependence. *International Business Review*, 23(4), 797-810.
- [5] Wooldridge, J. M., 2009, *Introductory Econometrics: A Modern Approach Fourth Edition*, South-Western Cengage Learning.
- [6] Gujarati, D. N., 2004, *Basic Econometrics Fourth Edition*, Mc Graw-Hill.

POLUSI DAN PERTUMBUHAN EKONOMI: SEBUAH ANALISIS *IMPULSE RESPONSE FUNCTION*

ANUGERAH KARTA MONIKA¹

1Sekolah Tinggi Ilmu Statistik, ak.monika@stis.ac.id

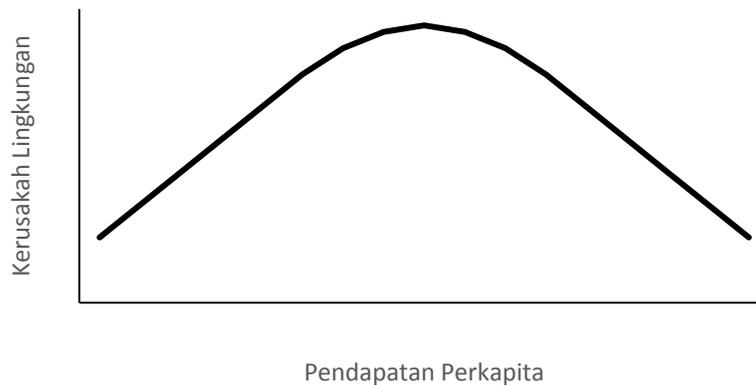
Abstrak. Hipotesis *Environmental Kuznets Curve* (EKC) menyatakan bahwa seiring dengan kenaikan pendapatan suatu negara, emisi yang timbul juga mengalami peningkatan sampai pada level ambang batas tertentu, kemudian emisi mengalami penurunan. Di sisi lain antara pendapatan dan emisi juga menunjukkan adanya hubungan kausalitas dimana masing-masing variabel saling mempengaruhi. Penelitian ini merupakan pengembangan dari hipotesis EKC dengan menjelaskan bagaimana hubungan kausalitas antara pendapatan ekonomi dan emisi CO₂ di Indonesia menggunakan *vector error correction*. Hubungan ini akan menunjukkan bagaimana respon dari pendapatan ekonomi dan emisi CO₂ bila terjadi *shock* pada salah satu variabel. Hasilnya kenaikan pertumbuhan ekonomi direspon positif oleh kenaikan emisi CO₂ dan stabil pada tahun kelima. Sedangkan kenaikan pertumbuhan emisi CO₂ direspon negatif oleh kenaikan pertumbuhan ekonomi dan stabil pada periode ketujuh.

Kata kunci : Environmental Kuznets Curve, Vector Error Correction Mechanism, Emisi CO₂

1. Pendahuluan

Studi lingkungan hidup sangat erat kaitannya dengan pembangunan berkelanjutan (*sustainable development*) Untuk memahami pembangunan berkelanjutan di masa yang akan datang, diperlukan pemahaman yang akurat mengenai kondisi dan lingkungan di masa lalu sehingga dapat diketahui hubungan antara pembangunan ekonomi dan pembangunan berkelanjutan.

Di bidang ekonomi lingkungan, hipotesis EKC digunakan untuk mempertimbangkan hubungan antara pertumbuhan ekonomi dan lingkungan. Hipotesis ini diperoleh melalui observasi data. Ketika salah satu plot pendapatan per-kapita di sepanjang sumbu horisontal dan indeks per-kapita degradasi lingkungan pada sumbu vertikal untuk negara tertentu, dia akan umumnya menemukan hubungan yang mengambil bentuk kurva berbentuk U terbalik. Dengan kata lain, degradasi lingkungan mengalami peningkatan pada tahap awal pembangunan ekonomi, dan setelah pendapatan per kapita mengalami peningkatan pada level tertentu degradasi lingkungan akan mengalami penurunan seiring dengan meningkatnya pendapatan.



Gambar 1. Hipotesis *Environmental Kuznets Curve* (EKC)
 Sumber : Uchiyama,

Gambar 1 memperlihatkan hubungan antara pendapatan perkapita dengan kerusakan lingkungan. Seiring meningkatnya pendapatan perkapita maka terjadi peningkatan kerusakan lingkungan. Hingga pada satu titik, ketika pendapatan perkapita semakin bertambah, kerusakan lingkungan mengalami penurunan.

Simon Kuznets menemukan hubungan antara pendapatan nasional per kapita dan ketimpangan pendapatan, kemudian beliau menyajikan hipotesis hubungan yang berkembang pada tahap awal pembangunan ekonomi dan kemudian mengalami penciutan yang dikenal dengan Kuznets Curve Hypothesis. Ketika pemerintah mempertimbangkan pembangunan berkelanjutan, memahami pengaruh pertumbuhan ekonomi terhadap lingkungan menjadi lebih penting dibanding hanya memperhatikan pertumbuhan ekonomi itu sendiri, sehingga konsep EKC juga menjadi menarik ketika suatu wilayah ingin menerapkan pembangunan berkelanjutan [5].

Model statis sederhana dan langsung untuk menjelaskan EKC di mana kurva tersebut tergantung pada kenaikan hubungan teknologi antara konsumsi barang yang diinginkan dan pengurangan produk sampingan yang tidak diinginkan. Hasilnya, kurva tidak tergantung pada dinamika pertumbuhan, lembaga-lembaga politik, atau bahkan eksternalitas, dan konsisten dengan kegagalan pasar atau efisiensi. Sehingga disimpulkan dengan menghadirkan dukungan empiris untuk meningkatkan pengembalian polutan udara secara umum [2].

de Bruyn and Heintz [3] menyajikan lima faktor yang mempengaruhi EKC, yaitu (1) perubahan perilaku dan preferensi, (2) perubahan institusi, (3) perubahan teknologi dan organisasi, (4) perubahan struktur dan (5) realokasi internasional.

Secara umum, studi empiris telah berusaha untuk menjelaskan tingkat degradasi lingkungan per-kapita dengan cara persamaan polinomial pendapatan per kapita. Pembuktian secara empiris terhadap EKC yang dilakukan oleh peneliti-peneliti tersebut menggunakan data panel dengan individu yang dipilih adalah negara. Model regresi yang standar digunakan untuk menjelaskan EKC adalah sebagai berikut:

$$E_{it} = \beta_0 + \beta_1 y_{it} + \beta_2 y_{it}^2 + \beta_3 y_{it}^3 + \beta_4 x_{it} + u_{it} \quad (1)$$

Dimana E adalah indeks polutan lingkungan, y adalah pendapatan perkapita dan x adalah variabel kontrol lainnya. Notasi i dan t menunjukkan negara dan periode

waktu yang digunakan. Jika $\beta_1 > 0$, $\beta_2 < 0$ dan $\beta_3 = 0$ dan signifikan secara statistik, dan titik balik dimana level polutan mulai turun, pada saat itulah terjadi EKC.

Sementara itu, penelitian ini bertujuan untuk memperlihatkan bagaimana hubungan antara polutan dan pendapatan suatu negara yang ditunjukkan oleh pertumbuhan polutan dan pertumbuhan ekonomi (dihitung dari pertumbuhan Produk Domestik Bruto) terjadi.

Indonesia, sebagai negara berkembang yang selalu berpaku pada nilai pertumbuhan ekonomi sebagai indikator utama bagi keberhasilan pembangunan ekonominya perlu memperhatikan hubungan antara pertumbuhan polutan dan pertumbuhan ekonominya. Harapannya tentu agar tercipta pembangunan ekonomi yang berkelanjutan, sehingga pembangunan yang dilakukan pada saat ini dapat dinikmati oleh generasi selanjutnya tanpa mengorbankan lingkungan.

2. Metodologi

Untuk memperlihatkan hubungan tersebut, digunakan data deret waktu emisi CO₂ dan Produk Domestik Bruto (PDB) dari tahun 1971 sampai 2013. Karena yang ingin dilihat adalah pertumbuhannya, maka kedua data tersebut ditransformasi dalam bentuk log. Selain itu data dalam bentuk log ini juga stasioner di level. Asumsi stasioneritas ini diperlukan bila kita ingin menggunakan model deret waktu, agar estimasi parameter yang dilakukan memenuhi asumsi Gauss Markov.

Metode yang digunakan adalah Vector Autoregressive (VAR). Model ini merupakan system persamaan simultan dimana jumlah persamaan yang dibentuk adalah sejumlah variabel yang digunakan. Masing-masing persamaan memiliki keterkaitan baik secara matematis maupun teoritis.

VAR digunakan ketika kita akan melakukan analisis terhadap beberapa variabel, tetapi kita tidak bisa membedakan mana yang menjadi variabel independen dan mana yang menjadi variabel dependen. Dengan karakteristik persamaan yang demikian, maka dapat dikatakan VAR adalah metode statistik yang menjelaskan hubungan antar variabel secara simultan. VAR akan membentuk sistem persamaan yang terdiri dari banyaknya variabel yang digunakan dalam sistem. Masing-masing persamaan merupakan fungsi dari variabel independen beserta lag-nya serta lag dari variabel dependennya. Jumlah lag ini menunjukkan berapa ordo dari VAR.

Berikut adalah penurunan formula VAR

$$Y_t = b_{10} - b_{12}Z_t + \gamma_{11}Y_{t-1} + \gamma_{12}Z_{t-1} + \varepsilon_{Yt} \quad (2)$$

$$Z_t = b_{20} - b_{21}Y_t + \gamma_{21}Y_{t-1} + \gamma_{22}Z_{t-1} + \varepsilon_{Zt} \quad (3)$$

Kemudian dibentuk menjadi

$$Y_t + b_{12}Z_t = b_{10} + \gamma_{11}Y_{t-1} + \gamma_{12}Z_{t-1} + \varepsilon_{Yt} \quad (4)$$

$$Z_t + b_{21}Y_t = b_{20} + \gamma_{21}Y_{t-1} + \gamma_{22}Z_{t-1} + \varepsilon_{Zt} \quad (5)$$

Jika ditulis dalam bentuk matriks

$$\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} Y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} Y_{t-1} \\ Z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{Yt} \\ \varepsilon_{Zt} \end{bmatrix}$$

$$B X_t = \Gamma_0 + \Gamma_1 X_{t-1} + \varepsilon_t$$

Kemudian dikali dengan B^{-1}

$$B^{-1}B X_t = B^{-1}\Gamma_0 + B^{-1}\Gamma_1 X_{t-1} + B^{-1}\varepsilon_t \quad (6)$$

$$X_t = A_0 + A_1 X_{t-1} + e_t \quad (7)$$

Sehingga

$$Y_t = a_{10} + a_{11}Y_{t-1} + a_{12}Z_{t-1} + e_{Yt} \quad (8)$$

$$Z_t = a_{20} + a_{21}Y_{t-1} + a_{22}Z_{t-1} + e_{Zt} \quad (10)$$

Dalam bentuk Lag Operator

$$Y_t = a_{10} + a_{11}LY_t + a_{12}LZ_t + e_{Yt} \quad (11)$$

$$Z_t = a_{20} + a_{21}LY_t + a_{22}LZ_t + e_{Zt} \quad (12)$$

Bila dibuat dalam bentuk Autoregressive² ordo 1 atau AR(1) maka bentuknya

$$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t \quad (13)$$

Sedangkan bila disusun dalam bentuk Vektor, akan membentuk Vector Autoregressive ordo 1 atau VAR(1), penulisannya menjadi seperti berikut ini beserta penurunannya

$$X_t = A_0 + A_1 X_{t-1} + e_t \quad (14)$$

$$= A_0 + A_1(A_0 + A_1 X_{t-2} + e_{t-1}) + e_t$$

$$= (I + A_1)A_0 + A_1^2 X_{t-2} + A_1 e_{t-1} + e_t$$

$$= (I + A_1)A_0 + A_1^2(A_0 + A_1 X_{t-3} + e_{t-2}) + A_1 e_{t-1} + e_t$$

$$= (I + A_1 + A_1^2)A_0 + A_1^3 X_{t-3} + A_1^2 e_{t-2} + A_1 e_{t-1} + e_t$$

$$= (I + A_1 + A_1^2)A_0 + A_1^3(A_0 + A_1 X_{t-4} + e_{t-3}) + A_1^2 e_{t-2} + A_1 e_{t-1} + e_t$$

$$X_t = (I + A_1 + A_1^2 + A_1^3)A_0 + A_1^4 X_{t-4} + A_1^3 e_{t-3} + A_1^2 e_{t-2} + A_1 e_{t-1} + e_t \quad (15)$$

I = matriks identitas 2 x 2

Bila dilakukan hingga n kali, maka bentuk persamaan (15) menjadi

$$X_t = (I + A_1 + A_1^2 + \dots + A_1^n)A_0 + \sum_{i=0}^n A_1^i e_{t-1} + A_1^{n+1} X_{t-(n-1)} \quad (16)$$

Dimana:

$$\mu = \frac{\bar{Y}}{\bar{Z}}$$

$$\bar{Y} = \frac{[a_{10}(I - a_{22}) + a_{12}a_{20}]}{\Delta}$$

$$\bar{Z} = \frac{[a_{20}(I - a_{11}) + a_{21}a_{10}]}{\Delta}$$

$$\Delta = (I - a_{11})(I - a_{22}) - a_{12}a_{21}$$

² Autoregressive (AR) merupakan suatu observasi pada waktu t dinyatakan sebagai fungsi linier terhadap p waktu sebelumnya ditambah dengan sebuah residual acak at yang white noise yaitu independen dan berdistribusi normal dengan rata-rata 0 dan varian konstan σ^2 , ditulis at $\sim N(0, \sigma^2)$.

Impulse Response Function

Pada persamaan VAR, bila koefisien regresinya diinterpretasikan seperti menginterpretasikan regresi linear berganda akan terasa janggal karena variabel endogennya adalah variabel eksogen dari periode sebelumnya yang tidak bisa diintervensi karena sudah terjadi. Selain itu, variabel eksogen di persamaan yang satu merupakan variabel endogen di persamaan yang lain.

Oleh karena itu, untuk menginterpretasikan persamaan VAR, diperlukan teknik lain dengan cara memberikan guncangan (*shock*) terhadap suatu variabel kemudian melihat reaksi dari variabel lainnya. Cara seperti ini disebut dengan *Impulse Response Function* (IRF).

IRF bermanfaat untuk menunjukkan bagaimana respon suatu variabel dari sebuah shock dalam variabel itu sendiri dan variabel endogen lainnya. Formula IRF dapat diturunkan sebagai berikut:

Misalkan dari persamaan VAR(1)

$$X_t = A_0 + A_1 X_{t-1} + e_t$$

Setelah hingga ke-n seperti pada persamaan (16), maka

$$\begin{aligned} X_t &= (I + A_1 + A_1^2 + \dots + A_1^n)A_0 + \sum_{i=0}^n A_1^i e_{t-1} + A_1^{n+1} X_{t-(n-1)} \\ &= \mu + \sum_{i=0}^{\infty} A_1^i e_{t-1} \\ &= \begin{bmatrix} \bar{Y} \\ \bar{Z} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \begin{bmatrix} e_{1t-1} \\ e_{2t-1} \end{bmatrix} \\ &= \begin{bmatrix} \bar{Y} \\ \bar{Z} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i B^{-1} \varepsilon_t \\ &= \begin{bmatrix} \bar{Y} \\ \bar{Z} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \frac{1}{1-b_{12}b_{22}} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} e_{Yt-1} \\ e_{Zt-1} \end{bmatrix} \\ X_t &= \begin{bmatrix} \bar{Y} \\ \bar{Z} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} \phi_{11}(i) & \phi_{12}(i) \\ \phi_{21}(i) & \phi_{22}(i) \end{bmatrix} \begin{bmatrix} \varepsilon_{Yt-1} \\ \varepsilon_{Zt-1} \end{bmatrix} \end{aligned} \quad (17)$$

$$X_t = \begin{bmatrix} \bar{Y} \\ \bar{Z} \end{bmatrix} + \sum_{i=0}^{\infty} \phi \varepsilon \quad (18)$$

Dimana:

$\phi_{11}(i)$ = respon i periode pada variabel Y akibat perubahan ε_{Yt-1} yang disebabkan shock pada Y

$\phi_{12}(i)$ = respon i periode pada variabel Y akibat perubahan ε_{Zt-1} yang disebabkan shock pada Z

$\phi_{21}(i)$ = respon i periode pada variabel Z akibat perubahan ε_{Yt-1} yang disebabkan shock pada Y

$\phi_{22}(i)$ = respon i periode pada variabel Z akibat perubahan ε_{Zt-1} yang disebabkan shock pada Z

Nilai IRF dihitung antar waktu forecast dan direpresentasikan pada persamaan (2) pada matriks ϕ .

Data yang digunakan dalam penelitian ini adalah emisi karbondioksida (CO₂) untuk menggambarkan kerusakan lingkungan dan GDP (*Gross Domestic*

Product atau PDB) untuk menggambarkan pendapatan. Emisi karbondioksida yang digunakan adalah yang berasal dari pembakaran bahan bakar fosil dan pembuatan semen termasuk karbon dioksida yang dihasilkan selama konsumsi bahan bakar padat, cair, dan gas dan pembakaran gas. Data GDP yang digunakan adalah GDP atas dasar harga berlaku. Semua data diunduh dari <http://data.worldbank.org/country> dengan series tahunan (1971-2013). Data ditransformasi dalam bentuk logaritma untuk menggambarkan pertumbuhan dari variabelnya.

3. Hasil dan Pembahasan

Dengan data log emisi CO₂ dan log PDB dihasilkan persamaan VAR dengan ordo 2 atau ditulis VAR(2). Hasil AR Root menunjukkan bahwa data stabil digunakan dalam system VAR karena tidak terdapat root. Sistem persamaan VAR(2) tersebut ditulis sebagai berikut:

$$\log(\text{CO}_2) = 0.73 \log(\text{CO}_2)_{t-1} - 0.21 \log(\text{CO}_2)_{t-2} + 0.29 \log(\text{GDP}_{t-1}) + 0.18 \log(\text{GDP}_{t-2}) - 11.28 \quad (3)$$

$$\log(\text{GDP}) = -0.0075 \log(\text{CO}_2)_{t-1} + 0.0327 \log(\text{CO}_2)_{t-2} + 1.25 \log(\text{GDP}_{t-1}) - 0.29 \log(\text{GDP}_{t-2}) + 1.15 \quad (4)$$

Selanjutnya dilakukan uji kausalitas Granger untuk melihat apakah terjadi hubungan timbal balik antara pertumbuhan emisi CO₂ dengan pertumbuhan ekonomi [4]. Hipotesis nolnya adalah pertumbuhan ekonomi tidak mempengaruhi pertumbuhan emisi CO₂. Hipotesis nol ditolak bila nilai p-value yang dihasilkan lebih kecil dari tingkat signifikansi yang digunakan. Dari hasil olahan data, nilai p-value yang dihasilkan adalah 0,0493. Nilai ini lebih kecil dibandingkan tingkat signifikansi yang digunakan (α) 5%, maka hipotesis nol tersebut ditolak, sehingga dapat disimpulkan bahwa pertumbuhan ekonomi mempengaruhi pertumbuhan emisi CO₂. Sebaliknya dengan hipotesis nol pertumbuhan emisi CO₂ tidak mempengaruhi pertumbuhan ekonomi, uji kausalitas granger menunjukkan bahwa Ho gagal ditolak. Sehingga pertumbuhan emisi CO₂ tidak mempengaruhi pertumbuhan ekonomi. Uji Kausalitas Granger ini ditunjukkan pada Tabel 1 berikut ini.

Tabel 1. VAR Granger Causality

Dependent variable: LOG(CO2)			
Excluded	Chi-sq	df	Prob.
LOG(GDP)	6.020274	2	0.0493
All	6.020274	2	0.0493

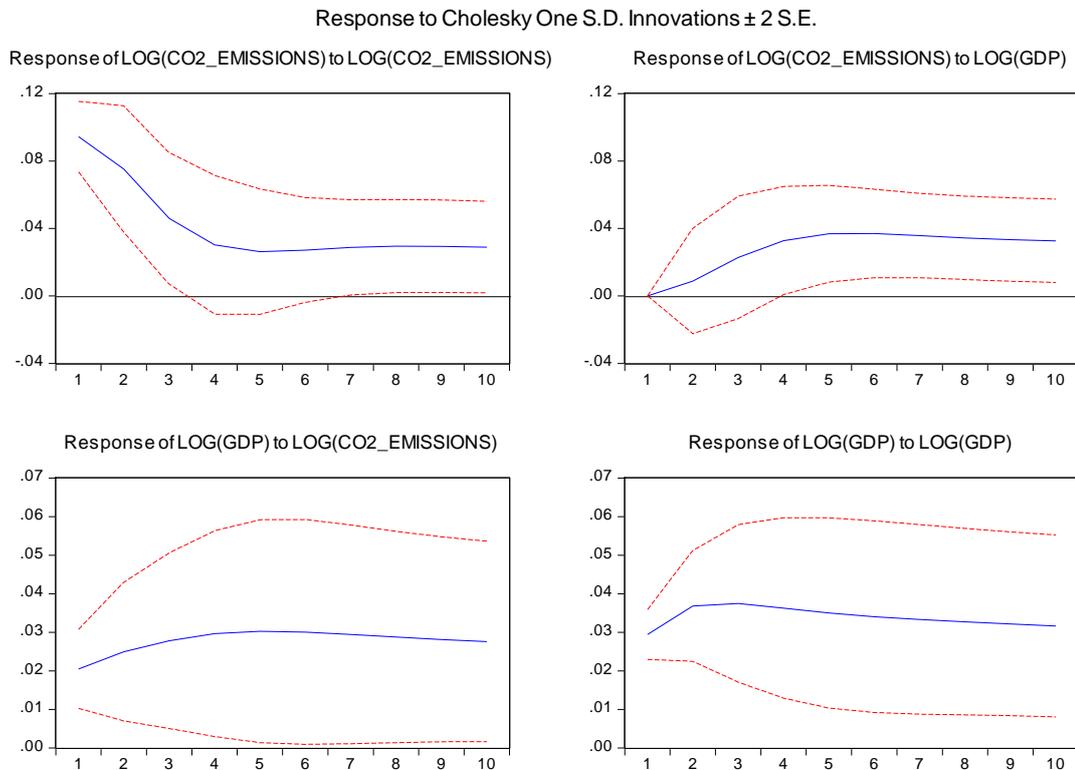
Dependent variable: LOG(GDP)

Excluded	Chi-sq	df	Prob.
LOG(CO2)	0.19.7240	2	0.9061
All	0.197240	2	0.9061

Sumber: Hasil pengolahan data dengan eviews 8

Untuk mengidentifikasi respon pertumbuhan emisi CO₂ pada IRF dalam model VAR digunakan standar *Cholesky Decomposition*³ yang bertujuan untuk *generate impulse response* yang tergantung secara kursial pada urutan (*ordering*) variable dalam sistem. Pada penelitian ini periode waktu yang digunakan alam menganalisis respon pertumbuhan ekonomi terhadap CO₂ diproyeksikan dalam tahun ke depan. Melalui IRF dapat diketahui bagaimana respon pertumbuhan emisi CO₂ terhadap guncangan dari variabel pertumbuhan ekonomi. Respon tersebut disajikan pada gambar 2.

Gambar 2. Respon Pertumbuhan Emisi CO₂ dan Pertumbuhan Ekonomi Terhadap Guncangan Pada Masing-Masing Variabel.



³ Dalam aljabar linear, Cholesky decomposition atau Cholesky factorization adalah dekomposisi matriks hermitian positif definit In linear algebra menjadi produk matriks segitiga bawah dan transpose konjugasinya. Ini sangat berguna untuk solusi numeric yang efisien dan simulasi monte carlo.

Pada gambar 2, sumbu x adalah periode ramalan dari data yang digunakan. Angka 1 menggambarkan 1 periode berikutnya dari data terakhir yaitu tahun 2014. Sedangkan sumbu y menggambarkan respon pada variabel yang terkena dampak jika terjadi guncangan (*shock*) pada variabel pasangannya. Penjelasan detail untuk gambar 2 adalah sebagai berikut. Panel 1 hingga 4 mengikuti arah jarum jam.

1. Respon pertumbuhan emisi CO₂ terhadap guncangan pertumbuhan emisi CO₂
Guncangan pertumbuhan emisi CO₂ sebesar satu standar deviasi pada tahun pertama akan menyebabkan peningkatan pada pertumbuhan emisi CO₂ sebesar 9,4 persen. Respon ini mengalami penurunan di tahun kedua dan terus turun hingga tahun keenam responnya relative stabil.
2. Respon pertumbuhan emisi CO₂ terhadap guncangan pertumbuhan ekonomi
Guncangan pertumbuhan ekonomi sebesar satu standar deviasi pada tahun pertama tidak menimbulkan respon pada pertumbuhan emisi CO₂. Tetapi pada tahun kedua pertumbuhan emisi merespon 0,8 persen atas guncangan pada pertumbuhan ekonomi. Respon positif ini terus mengalami kenaikan di tahun ketiga dan terus naik hingga tahun kelima responnya relative stabil.
3. Respon pertumbuhan ekonomi terhadap guncangan pertumbuhan emisi CO₂
Guncangan pertumbuhan emisi CO₂ sebesar satu standar deviasi pada tahun pertama menimbulkan respon positif sebesar 2 persen pada pertumbuhan ekonomi. Respon ini terus naik, hingga tahun kelima dan cenderung stabil.
4. Respon pertumbuhan ekonomi terhadap guncangan pertumbuhan emisi CO₂
Guncangan pertumbuhan emisi CO₂ sebesar satu standar deviasi pada tahun pertama menimbulkan respon positif pada pertumbuhan ekonomi sebesar 2,9 persen dan pada tahun ketiga responnya naik sebesar 3,9 persen. Respon ini cenderung ini stabil pada tahun ketiga dan seterusnya.

Tabel 2 berikut ini merupakan representasi dari grafik 2 dalam bentuk nilai responnya. Kolom 2 dan 3 adalah respon dari variabel emisi CO₂ bila terjadi guncangan pada variabel log CO₂ dan log GDP. Kolom 4 dan 5 adalah respon dari variabel log GDP bila terjadi guncangan pada variabel log CO₂ dan log GDP

Tabel 2. Nilai IRF untuk pertumbuhan ekonomi dan pertumbuhan emisi CO₂

Period	Response of LOG(CO ₂):		Response of LOG(GDP):	
	LOG(CO ₂)	LOG(GDP)	LOG(CO ₂)	LOG(GDP)
1	0.094434	0.000000	0.020491	0.029405
2	0.075265	0.008757	0.024947	0.036822
3	0.045945	0.022820	0.027781	0.037460
4	0.030280	0.032789	0.029621	0.036273
5	0.026213	0.036817	0.030257	0.034985
6	0.027148	0.037011	0.030035	0.034016
7	0.028763	0.035756	0.029430	0.033308

8	0.029522	0.034405	0.028757	0.032720
9	0.029417	0.033380	0.028137	0.032160
10	0.028867	0.032646	0.027583	0.031593

Cholesky Ordering: LOG(CO2) LOG(GDP)

4. Kesimpulan

Dari hasil analisis IRF di atas dapat disimpulkan bahwa pertumbuhan emisi CO₂ tidak langsung merespon terhadap guncangan pada pertumbuhan ekonomi pada tahun pertama. Artinya ketika pertumbuhan ekonomi mengalami perubahan sebesar satu standar deviasi, pertumbuhan emisi CO₂ baru merespon di tahun kedua dan seterusnya. Sehingga dapat disimpulkan diperlukan waktu untuk terjadi polusi bila terjadi pembangunan ekonomi yang ditunjukkan dengan pertumbuhan ekonomi. Lamanya waktu sampai terjadi polusi akibat pertumbuhan ekonomi itu adalah satu tahun, karena pada tahun kedua terjadi respon positif atas guncangan tersebut. Sejalan dengan itu, tahap awal *Environmental Kuznets Curve* memang terjadi di Indonesia. Untuk tahap berikutnya diperlukan metode regresi lain yang menggambarkan hipotesis ini.

Referensi

- [1] Andreoni, James & Arik Levinson, 2004, *The Simple Analytic of the Environmental Kuznets Curve*, *Journal of Public Economics* (2004), 80, 260-286.
- [2] Arouri, Mohamed El Hedi et al., 2012, *Energy Consumption, Economic Growth and Co2 Emissions in Middle East and North African Countries*, *IZA Discussion Paper* (2012),6412.
- [3] de Bruyn, S. M., & Heintz, R. J., 2002, *The environmental Kuznets curve hypothesis*. In J. C.J.M.vandenBergh(Ed.), *Handbook of environmental and resource economics* pp.656–677. Oxford: Edward Elgar.
- [4] Enders, Walter, 2014, *Applied Econometric Time Series (4th edition)*, Wiley.
- [5] K. Uchiyama, 2016, *Environmental Kuznets Curve Hypothesis and Carbon Dioxide Emissions*, *Development Bank of Japan Research Series*, p.11-29
- [6] Stern, David I., 2004, *The Rise and Fall of the Environmental Kuznets Curve*. *World Development*, 32(8), 1419-1439
- [7] <http://data.worldbank.org/country> diunduh 30 Januari 2017.

DAMPAK EKSPOR KARET ALAM TERHADAP PEREKONOMIAN INDONESIA DAN PERAMALANNYA: ANALISIS TABEL I-O DAN ARIMA

RAMADHAN I. P.¹, RIBUT NURUL T. W.²

1 Badan Pusat Statistik, 12.7328@stis.ac.id

2 Sekolah Tinggi Ilmu Statistik, rnurult@stis.ac.id

Abstrak. Karet alam merupakan salah satu komoditas ekspor penting bagi Indonesia. *Share* karet alam terhadap total ekspor Indonesia berada di peringkat ke-5. Penurunan harga karet alam di pasar internasional mengakibatkan nilai ekspor karet alam semakin menurun dari tahun ke tahun. Hal ini akan berdampak bagi perekonomian Indonesia, terutama sektor karet sendiri karena mayoritas karet alam di Indonesia diekspor ke luar negeri dan hanya sedikit yang diserap di dalam negeri. Dari latar belakang tersebut, peneliti ingin melihat dampak penurunan nilai ekspor karet alam terhadap perekonomian Indonesia. Selain itu, peneliti juga melakukan peramalan nilai ekspor karet pada tahun 2016 dan dampaknya terhadap perekonomian di tahun tersebut. Metode analisis yang digunakan adalah analisis tabel Input-Output (I-O) dan metode ARIMA. Hasil analisis menunjukkan bahwa sektor yang paling terpengaruh akibat penurunan nilai ekspor karet alam adalah sektor karet, jasa keuangan perbankan dan jasa pertanian, kehutanan, dan perikanan. Total potensi output, pendapatan masyarakat, dan lapangan kerja yang hilang pada periode 2011-2015 adalah masing-masing sebesar 61,7 trilyun rupiah, 18,35 trilyun rupiah, dan 1,54 juta tenaga kerja. Model peramalan terbaik yang didapatkan adalah ARIMA (1,2,1). Peramalan untuk tahun 2016 adalah terjadi penurunan nilai ekspor sebesar 3,7 trilyun rupiah. Perkiraan dampak penurunan nilai ekspor karet terhadap output, pendapatan, dan tenaga kerja masing-masing adalah masing-masing sebesar 4,5 trilyun rupiah, 1,3 trilyun rupiah, dan 114 ribu tenaga kerja.

Kata kunci : karet, ekspor, analisis tabel Input-Output, ARIMA.

1. Pendahuluan

Sektor pertanian merupakan sektor yang penting dalam perekonomian. Hal ini dilihat dari kontribusinya terhadap Produk Domestik Bruto (PDB) yang besar, yaitu sebesar 13,52 persen pada tahun 2015 atau berada di peringkat kedua setelah sektor industri. Sektor ini juga kuat dalam menghadapi guncangan ekonomi pada waktu krisis ekonomi (Mubyarto dalam Utomo [1]). Salah satu sub sektor yang berperan penting adalah sub sektor perkebunan. Pada tahun 2015, sub sektor perkebunan merupakan penyumbang PDB terbesar pada sektor pertanian yaitu sebesar 3,57 persen. Selain itu sub sektor perkebunan juga merupakan penyedia bahan baku bagi sektor industri, penyerap tenaga kerja dan penghasil devisa (BPS

[2]).

Karet alam merupakan salah satu komoditas hasil perkebunan yang cukup penting dalam kegiatan perekonomian di Indonesia. Karet juga salah satu komoditas ekspor Indonesia yang cukup penting sebagai penghasil devisa negara di luar minyak dan gas. Sekitar 83,46 persen produksi karet alam Indonesia diekspor ke luar negeri dan hanya sebagian kecil yang dikonsumsi di dalam negeri (BPS [2]). Selain itu, Indonesia juga merupakan salah satu pengeksport karet alam terbesar di dunia, berada di peringkat kedua setelah Thailand.

Pada tahun 2012, terjadi penurunan harga yang drastis pada komoditas karet, terutama karet alam (*natural rubber*). Hal ini tentu memengaruhi ekspor karet Indonesia karena hampir sebagian besar karet yang diekspor merupakan karet alam. Tren turunnya harga karet alam ini masih berlanjut hingga tahun 2015. Penurunan harga karet alam mengakibatkan banyak kerugian. Di beberapa daerah seperti di Riau dan Bengkulu, terjadi pemutusan hubungan kerja (PHK) tenaga kerja di perkebunan dan pabrik karet yang diakibatkan oleh rendahnya harga karet di luar negeri. Selain berdampak kepada perusahaan besar, turunnya harga karet juga berdampak kepada perkebunan rakyat. Berdasarkan data Kementerian Perdagangan, di tingkat petani harga karet pada tahun 2014 hanya berkisar 5 ribu rupiah per kilogram, padahal biasanya berada di atas 10 ribu rupiah dan bahkan pernah mencapai 18 hingga 22 ribu rupiah per kilogramnya. Turunnya harga ini juga secara langsung akan mengurangi pendapatan petani sehingga bisa jadi petani karet akan beralih ke pekerjaan lain yang lebih menguntungkan.

Berdasarkan latar belakang di atas, tujuan dilakukannya penelitian ini adalah melihat dampak perubahan ekspor karet alam terhadap perekonomian dan melakukan proyeksi ekspor karet alam Indonesia di masa yang akan datang.

2. Hasil – Hasil Utama

Metode Penelitian

Nazara [3] menyatakan bahwa dalam keseimbangan umum, seluruh sektor perekonomian adalah satu kesatuan sistem dengan keseimbangan (atau tidak keseimbangan) di satu sektor akan berpengaruh terhadap keseimbangan (atau ketidakseimbangan) di sektor-sektor lainnya. Dengan kata lain, ketika suatu sektor mengalami perubahan dalam permintaan akhirnya, maka sektor-sektor lainnya akan terpengaruh dengan adanya perubahan tersebut. Untuk melihat dampak ekspor karet terhadap perekonomian, digunakan analisis dampak menggunakan analisis tabel Input-Output. Analisis dampak yang digunakan adalah dampak output, pendapatan, dan tenaga kerja. Rumus yang digunakan adalah sebagai berikut:

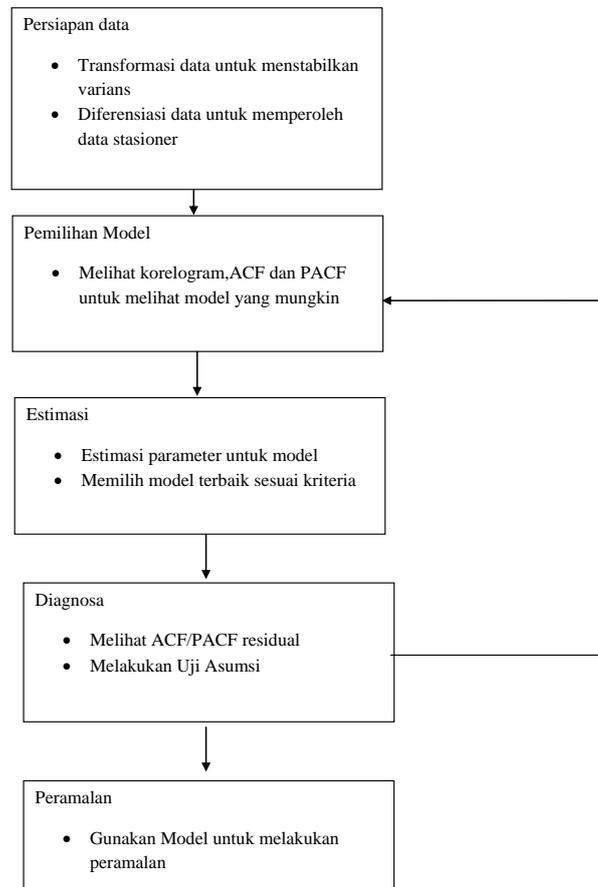
$$\Delta X_i = (I - A^d)^{-1} \Delta E \quad (1)$$

$$\Delta G_i = g_i (I - A^d)^{-1} \Delta E \quad (2)$$

$$\Delta L_i = l_i (I - A^d)^{-1} \Delta E \quad (3)$$

dengan A^d adalah matriks koefisien input, ΔX_i adalah matriks perubahan output yang diciptakan akibat perubahan ekspor karet, ΔE adalah matriks perubahan ekspor karet, ΔG_i adalah matriks perubahan upah/gaji akibat perubahan ekspor karet, g_i adalah matriks diagonal koefisien upah/gaji (rasio antar upah/gaji dengan output yang dihasilkan), ΔL_i adalah matriks perubahan tenaga kerja akibat

perubahan ekspor karet, dan I_i adalah matriks diagonal koefisien tenaga kerja (rasio antara tenaga kerja dengan output yang dihasilkan) (Sutomo [4]).



Gambar 1. Tahapan pembentukan model ARIMA

Untuk melakukan peramalan ekspor, metode yang digunakan adalah *Autoregressive Integrated Moving Average* (ARIMA). Tahapan dalam membentuk model ARIMA ada pada Gambar 1. Pemilihan model terbaik yang dihasilkan berdasarkan nilai Akaike Information Criterion (AIC) dan Schwarz Criterion (SC) (Gujarati [5]). Model yang terpilih adalah model yang menghasilkan nilai AIC dan SC yang terkecil. Setelah diperoleh persamaan tersebut, maka dilakukan estimasi untuk nilai ekspor Indonesia pada tahun 2016. Dari nilai tersebut dihitung nilai perubahannya terhadap periode sebelumnya dan selanjutnya dapat dilakukan analisis dampak perubahan ekspor pada tahun 2016.

Data yang digunakan dalam penelitian ini adalah data yang bersumber dari Badan Pusat Statistik (BPS) berupa Tabel Input-Output 185 dan 9 Sektor tahun 2010, data tenaga kerja 9 sektor tahun 2010, dan data nilai ekspor karet alam bulanan pada periode 2011-2015.

Analisis I-O

Total turunnya output perekonomian Indonesia akibat perubahan ekspor karet alam pada periode 2011-2015 sebesar -61,714 trilyun rupiah. Sektor yang mengalami penurunan output terbesar akibat turunnya ekspor karet alam adalah

sektor karet, yaitu mengalami penurunan output sebesar 52,18 trilyun rupiah. Hal ini terjadi karena sektor karet adalah sektor yang berhubungan langsung dengan ekspor karet alam sehingga ketika karet alam turun maka otomatis output karet alam akan ikut turun juga. Sektor lainnya yang terkena dampak penurunan output akibat penurunan ekspor karet adalah jasa keuangan dan perbankan serta jasa pertanian, kehutanan, dan perikanan.

Jasa keuangan dan perbankan menjadi sektor kedua terbesar yang terkena dampak akibat penurunan karet ekspor karena sebagian besar petani karet membutuhkan pinjaman, terutama dari lembaga keuangan dalam melakukan kegiatan produksinya. Riantini [6] menyatakan bahwa petani karet hanya memiliki modal sekitar 52 persen untuk melakukan peremajaan terhadap tanaman karetinya setiap tahun. Kekurangan dana tersebut diperoleh dari sumber lain, antara lain pinjaman ke lembaga keuangan. Ketika output karet berkurang, otomatis pinjaman terhadap lembaga keuangan juga berkurang dikarenakan petani karet takut tidak bisa mengembalikan sehingga output jasa keuangan dan perbankan juga ikut turun.

Tabel 1. Dampak perubahan ekspor karet alam terhadap output Indonesia Tahun 2011-2015

Sektor	Besar Perubahan Output (Milyar Rupiah)					
	2011	2012	2013	2014	2015	Total
Karet	41.600,152	-39.013,628	-12.032,289	-27.860,152	-14.874,519	-52.180,435
Jasa keuangan dan perbankan	789,508	-740,420	-228,355	-528,744	-282,296	-990,306
Jasa pertanian, kehutanan, dan perikanan	705,811	-661,926	-204,146	-472,690	-252,369	-885,322
Bangunan dan instalasi listrik, gas, air minum, dan komunikasi	635,456	-595,946	-183,797	-425,573	-227,213	-797,073
Pupuk	622,486	-583,782	-180,046	-416,887	-222,576	-780,805
Perdagangan selain mobil dan sepeda motor	517,464	-485,290	-149,670	-346,552	-185,024	-649,072
Barang-barang hasil kilang minyak dan gas bumi	442,909	-415,371	-128,106	-296,622	-158,366	-555,556
Prasarana pertanian	345,062	-323,608	-99,805	-231,093	-123,380	-432,823
Minyak bumi	252,631	-236,924	-73,070	-169,190	-90,331	-316,884
Jasa persewaan dan jasa penunjang usaha	235,213	-220,589	-68,032	-157,525	-84,103	-295,036
Lainnya	3.054,022	-2.864,135	-883,335	-2.045,317	-1.091,994	-3.830,759
Total	49.200,714	-46.14.,619	-14.230,650	-32.950,345	-17.592,170	-61.714,069

Dampak pendapatan digunakan untuk melihat pengaruh penurunan ekspor karet alam terhadap upah/gaji yang diterima rumah tangga di suatu negara/wilayah. Besarnya dampak pendapatan akibat penurunan ekspor karet alam adalah sebesar -18,355 trilyun rupiah. Lima sektor yang mengalami penurunan pendapatan terbesar adalah sektor karet sebesar 16,807 trilyun rupiah; sektor jasa keuangan dan perbankan sebesar 248,217 milyar rupiah; jasa pertanian, kehutanan, dan perikanan sebesar 187,269 milyar rupiah; perdagangan selain mobil dan motor sebesar 122,622 milyar rupiah; dan sektor pupuk sebesar 109,874 milyar rupiah. Penurunan pendapatan searah dengan penurunan output karena ada hubungan yang proporsional. Penurunan output yang dihasilkan otomatis mengakibatkan

penurunan pendapatan.

Tabel 2. Dampak perubahan ekspor karet alam terhadap upah dan gaji Indonesia tahun 2011-2015.

Sektor	Besar Perubahan Pendapatan Rumah Tangga (Juta Rupiah)					
	2011	2012	2013	2014	2015	Total
Karet	13.399.679	-12.566.543	-3.875.678	-8.973.936	-4.791.179	-16.807.657
Jasa keuangan perbankan	197.888	-185.584	-57.236	-132.528	-70.757	-248.217
Jasa pertanian, kehutanan, dan perikanan	149.297	-140.015	-43.182	-99.986	-53.383	-187.269
Perdagangan selain mobil dan motor	97.791	-91.711	-28.285	-65.492	-34.966	-122.662
Pupuk	87.595	-82.149	-25.336	-58.664	-31.321	-109.874
Bangunan dan instalasi, listrik, gas, air minum, dan komunikasi	82.909	-77.754	-23.980	-55.525	-29.645	-103.995
Jasa persewaan dan jasa penunjang usaha	56.835	-53.301	-16.439	-38.063	-20.322	-71.290
Barang-barang hasil kilang minyak dan gas bumi	53.706	-50.367	-15.534	-35.968	-19.203	-67.365
Prasarana pertanian	48.423	-45.412	-14.006	-32.430	-17.314	-60.739
Jasa pemerintah umum	47.693	-44.727	-13.794	-31.940	-17.053	-59.822
Lainnya	412.090	-386.468	-119.191	-275.982	-147.346	-516.897
Total	14.633.906	-13.724.031	-4.232.662	-9.800.513	-5.232.488	-18.355.788

Koefisien tenaga kerja merupakan perbandingan antara tenaga kerja dalam suatu sektor produksi dengan output yang dihasilkan sektor tersebut. Semakin kecil nilai koefisien tenaga kerja menunjukkan sektor tersebut merupakan sektor padat modal dan semakin besar nilai koefisien tenaga kerja menunjukkan sektor tersebut merupakan sektor padat karya. Dampak tenaga kerja digunakan untuk melihat perubahan penyerapan tenaga kerja pada suatu sektor akibat penurunan ekspor karet alam . Ketika ekspor karet alam menurun, maka diperkirakan akan terjadi penurunan jumlah tenaga kerja atau terjadi pemutusan hubungan kerja (PHK) terutama pada perkebunan nasional dan swasta.

Besarnya penurunan tenaga kerja akibat perubahan ekspor karet alami adalah sebesar 1,54 juta orang. Lima sektor yang mengalami penurunan terbesar adalah sektor pertanian, peternakan, kehutanan dan perikanan sebesar 1,42 juta orang; industri pengolahan sebesar 50,64 ribu orang; perdagangan, hotel, dan restoran sebesar 32,49 ribu orang; listrik, gas, dan air bersih sebesar 32,49 ribu orang; dan bangunan sebesar 32,49 ribu orang. Turunnya penyerapan tenaga kerja berbanding lurus terhadap penurunan output. Karena ketika output menurun, otomatis pengeluaran dari perusahaan harus dikurangi sehingga salah satu jalan yang ditempuh oleh perusahaan adalah dengan pemutusan hubungan kerja karyawannya.

Tabel 3. Dampak perubahan ekspor karet alam terhadap tenaga kerja Indonesia tahun 2011-2015

Sektor	Besar Perubahan Tenaga Kerja					
	2011	2012	2013	2014	2015	Total
Pertanian, peternakan, kehutanan, dan perikanan	1.139.328	-1.068.490	-329.535	-763.023	-407.377	-1.429.097
Industri pengolahan	40.373.47	-37.863.2	-11.677.5	-27.038.6	-14.435.9	-50.641.8
Perdagangan, hotel dan Restoran	25.909.35	-24.298.4	-7.493.93	-17.351.8	-9.264.13	-32.499
Listrik, gas, dan air bersih	7.545.591	-7.076.44	-2.182.46	-5.053.38	-2.698	-9.464.68
Bangunan	7.478.371	-7.013.4	-2.163.02	-5.008.36	-2.673.96	-9.380.37
Keuangan, real estate, dan jasa perusahaan	3.825.485	-3.587.63	-1.106.47	-2.561.98	-1.367.84	-4.798.43
Pengangkutan dan komunikasi	3.136.869	-2.941.83	-907.298	-2.100.8	-1.121.62	-3.934.68
Jasa-jasa	2.432.523	-2.281.28	-703.575	-1.629.09	-869.771	-3.051.19
Pertambangan dan penggalian	954.783	-895.419	-276.158	-639.43	-341.391	-1197.62
Total	1.230.985	-1.154.447	-356.046	-824.406	-440.150	-1.544.064

ARIMA

Null Hypothesis: D(LNNILAI,2) has a unit root
 Exogenous: Constant, Linear Trend
 Lag Length: 0 (Automatic - based on SIC, maxlag=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-14.78857	0.0000
Test critical values:		
1% level	-4.127338	
5% level	-3.490662	
10% level	-3.173943	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(LNNILAI,3)
 Method: Least Squares
 Date: 08/31/16 Time: 02:23
 Sample (adjusted): 2011M04 2015M12
 Included observations: 57 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(LNNILAI(-1),2)	-1.617787	0.109394	-14.78857	0.0000
C	-0.019979	0.044499	-0.448964	0.6553
@TREND(2011M01)	0.000516	0.001268	0.407161	0.6855
R-squared	0.802191	Mean dependent var		0.002990
Adjusted R-squared	0.794864	S.D. dependent var		0.347727
S.E. of regression	0.157492	Akaike info criterion		-0.807685
Sum squared resid	1.339406	Schwarz criterion		-0.700156
Log likelihood	26.01903	Hannan-Quinn criter.		-0.765896
F-statistic	109.4951	Durbin-Watson stat		2.304775
Prob(F-statistic)	0.000000			

Gambar 2. Hasil uji ADF pada data difeferensiasi ke-2

Sebelum dilakukan pengujian stasioneritas, data dilihat apakah membentuk tren atau tidak. Hasil plot menunjukkan adanya tren pada data ekspor karet, sehingga pada pengujian stasioneritas dimasukkan efek tren. Berdasarkan hasil uji *Augmented Dickey Fueller* (ADF), nilai dari statistik uji t lebih kecil dari nilai $t_{0,05}$ yang berarti data stasioner pada level dengan tingkat kepercayaan 95 persen. Namun karena masih mengandung tren, maka data pada level tersebut tetap dilakukan diferensiasi sehingga tidak ada lagi efek tren pada data. Setelah dilakukan diferensiasi ke-2 dapat dilihat bahwa data yang dihasilkan sudah stasioner dan tidak lagi mengandung tren, sehingga peramalan menggunakan metode ARIMA dapat dilakukan. Langkah selanjutnya adalah penentuan lag maksimal yang akan digunakan untuk AR dan MA. Penentuan lag maksimal ditentukan berdasarkan output korelogram (Gambar 2).

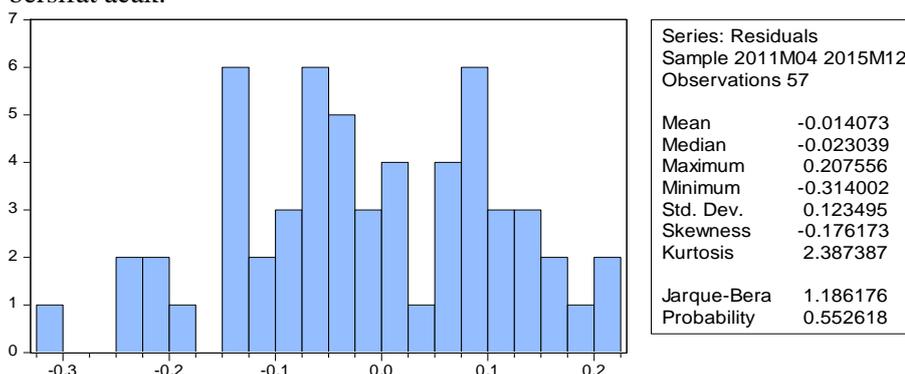
ACF lag yang melewati garis bartlett adalah lag ke-1. Sedangkan PACF lag yang melewati garis bartlett adalah lag ke-5. Oleh karena itu, lag maksimal untuk AR adalah 5 dan lag maksimal untuk MA adalah 1. Selain itu, korelogram menunjukkan bahwa tidak ada pola musiman pada data sehingga tidak perlu dilakukan diferensiasi musiman.

Setelah pembentukan model dilakukan, maka dilihat nilai AIC dan SC dari masing-masing model. Pemilihan model terbaik dilihat berdasarkan nilai AIC dan SC yang terkecil. Berdasarkan Tabel 4, dapat dilihat bahwa model ARIMA(1,2,1) merupakan model terbaik karena memiliki nilai AIC dan SC yang terkecil.

Tabel 4. Nilai AIC dan SC model ARIMA yang terbentuk

Model ARIMA	AIC	SC
ARIMA (1,2,1)	-1,244	-1,137
ARIMA (2,2,1)	-1,210	-1,065
ARIMA (3,2,1)	-1,179	-0,997
ARIMA (4,2,1)	-1,143	-0,922
ARIMA (5,2,1)	-1,222	-0,962

Setelah model diperoleh, maka dilakukan pengujian asumsi. Asumsi model yang diuji adalah *white noise*, normalitas, non-autokorelasi, dan homoskedastisitas. Asumsi *white noise* diuji menggunakan *korelogram Q-statistics*, normalitas menggunakan uji Jacque-Berra, non-autokorelasi menggunakan *LM-test* dan heteroskedastisitas menggunakan uji *White*. Berdasarkan korelogram *Q-statistics*, model yang dihasilkan memenuhi asumsi *white noise*. Asumsi tersebut terpenuhi karena *p-value* pada semua lag bernilai lebih besar dari taraf uji sehingga dapat dikatakan *error* pada model yang dihasilkan bersifat *random walk* atau *error* bersifat acak.



Gambar 3. Uji normalitas ARIMA(1,2,1)

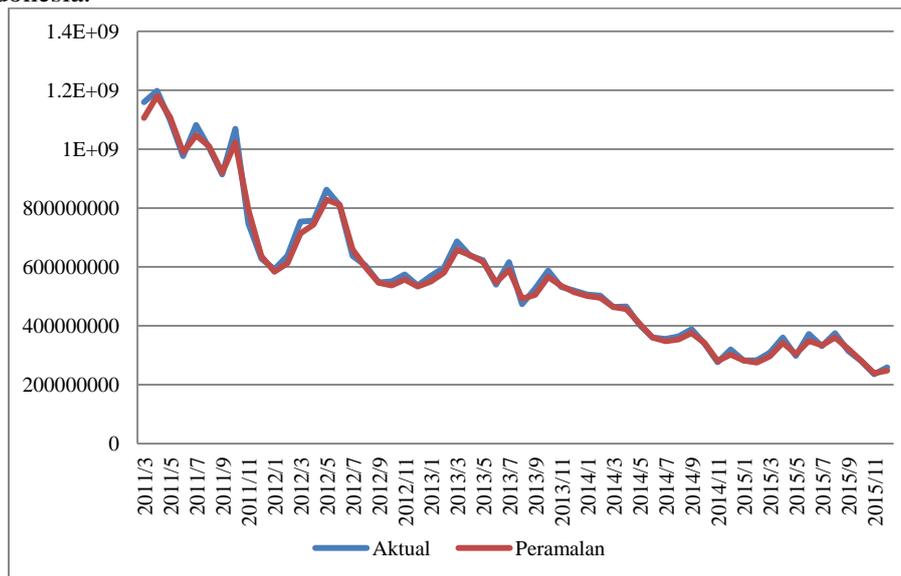
Berdasarkan uji Jacque-Berra pada Gambar 3, dapat dikatakan bahwa model yang dihasilkan memenuhi asumsi normalitas. Asumsi tersebut terpenuhi karena *p-value* bernilai lebih besar dari taraf uji sehingga dapat dikatakan *error* pada model yang dihasilkan mengikuti distribusi normal. Berdasarkan *LM test*, model yang dihasilkan memenuhi asumsi non-autokorelasi. Sedangkan hasil uji White menunjukkan bahwa model yang dihasilkan memenuhi asumsi homoskedastisitas. Asumsi tersebut terpenuhi karena *p-value* bernilai lebih besar dari taraf uji sehingga dapat dikatakan varians *error* pada model yang dihasilkan konstan.

Berdasarkan tahapan di atas, model yang dihasilkan adalah:

$$\hat{Y}_t = 100.000,196 + 5,931Y_{t-1} + 0,5708Y_{t-2} + 0,8027Y_{t-3} - 0,971\varepsilon_{t-1} \quad (4)$$

Dari persamaan diatas, dapat dilihat bahwa nilai ekspor (*Y*) sekarang dipengaruhi oleh nilai ekspor bulan sebelumnya, dua bulan sebelumnya, dan tiga bulan sebelumnya. Hal ini sejalan dengan hasil penelitian Mladenovic, Lepojevic dan Milic [7] yang menyatakan bahwa nilai ekspor negara Serbia dipengaruhi oleh nilai ekspor tiga bulan sebelumnya.

Untuk melihat apakah model peramalannya sudah cukup baik meramalkan nilai ekspor karet alam atau belum, maka perlu dibandingkan antara peramalannya dengan nilai sesungguhnya. Berdasarkan Gambar 4, dapat dilihat bahwa model peramalan yang dihasilkan sudah cukup baik dalam meramalkan nilai ekspor karet Indonesia.



Gambar 4. Nilai aktual dan peramalan ekspor karet alam Indonesia tahun 2011-2015

Dari model yang dihasilkan, maka dilakukan peramalan untuk tahun 2016. Hasil peramalan nilai ekspor karet Indonesia tahun 2016 adalah 7,529 miliar rupiah dan turun 3,754 miliar rupiah dibanding tahun sebelumnya. Setelah dilakukan simulasi dengan ramalan nilai ekspor karet tahun 2016, maka diperoleh besarnya penurunan output akibat penurunan ekspor adalah sebesar 4,569 triliun rupiah, besarnya penurunan pendapatan akibat penurunan ekspor adalah sebesar

1,358 trilyun rupiah, dan banyak orang yang terancam terkena phk sebanyak 114.317 orang.

5. Kesimpulan

Nilai ekspor karet alam Indonesia dipengaruhi oleh nilai ekspor karet alam pada periode sebelumnya. Ketika periode-periode sebelumnya mengalami penurunan, maka nilai ekspor pada periode selanjutnya akan mengalami tren menurun juga. Dampak ekspor karet alam Indonesia terhadap output, pendapatan dan tenaga kerja searah dengan perubahan nilai ekspornya. Besar penurunan output, pendapatan dan tenaga kerja juga proporsional terhadap penurunan nilai ekspor karet. Semakin besar penurunan nilai ekspor maka penurunan output, pendapatan, dan tenaga kerja akan semakin besar pula.

Referensi

- [1] Utomo, Y. K., 2015, *Dampak Investasi Sektor Pertanian Terhadap Perekonomian di Provinsi Jawa Timur (Pendekatan Analisis Input-Output)*, Universitas Negeri Jember.
- [2] Badan Pusat Statistik, 2015, *Statistik Karet Indonesia*, Badan Pusat Statistik.
- [3] Nazara, S., 2005, *Analisis Input-Output*, Lembaga Penerbit Fakultas Ekonomi Universitas Indonesia.
- [4] Sutomo, S., 2015, *Sistem Data dan Perangkat Analisis Ekonomi Makro*, Corleone Books.
- [5] Gujarati, D. N., 2004, *Basic Econometrics Fourth Edition*, Mc Graw-Hill.
- [6] Riantini, M., 2011, Potensi Modal Petani dalam Melakukan Peremajaan Karet di Kabupaten Musi Rawas Sumatera Selatan, *Agrisep Volume 10 Nomor 1 Maret 2011*, 19-27.
- [7] Mladenovic, J., Lepojevic, V., dan Milic, V. J., 2016, Modelling and Prognosis Of the Export of the Republic Serbia by Using Seasonal Jolt-Winters and ARIM, *Method.Ec0j0k9c Themes 2016*, 233-260.

PEMODELAN ERROR CORRECTION MECHANISM DALAM ANALISIS ALIRAN MODAL DARI DALAM DAN LUAR NEGERI TERHADAP PEREKONOMIAN INDONESIA

AGATHA HERDIANI BRIA¹, KRISMANTI TRI WAHYUNI²

¹ Badan Pusat Statistik, agatha.bria@bps.go.id
² Sekolah Tinggi Ilmu Statistik, krismanti@stis.ac.id

Abstrak. Proses pembangunan yang berkelanjutan membutuhkan dana yang besar sebagai modal. Modal pembangunan di Indonesia yang berasal dari dalam negeri dapat berupa tabungan domestik dan hasil ekspor, sedangkan luar negeri antara lain modal asing dan utang luar negeri. Utang luar negeri Indonesia digunakan untuk menutupi defisit anggaran pemerintah. Sejak diberlakukannya desentralisasi fiskal tahun 2011, anggaran pemerintah tidak pernah mengalami surplus, dengan pertumbuhan utang mencapai 50,50 persen sampai dengan kuartal kedua tahun 2013. Penelitian ini menggunakan analisis time series dengan pemodelan *Error Correction Mechanism* (ECM) dan membuktikan bahwa utang pemerintah tersebut tidak berpengaruh nyata terhadap PDB Indonesia, baik dalam jangka pendek maupun jangka panjang. PDB Indonesia dipengaruhi secara signifikan oleh modal asing dan tabungan domestik pada jangka pendek. Sedangkan pada jangka panjang, PDB juga dipengaruhi oleh ekspor selain kedua faktor tersebut.

Kata kunci: utang luar negeri, modal asing, ekspor, tabungan domestik, *Error Correction Mechanism*.

1. Pendahuluan

Indonesia memanfaatkan sumber modal yang berasal dari dalam negeri maupun luar negeri untuk mendukung berjalannya kegiatan ekonomi. Salah satu modal dari luar negeri yang dimanfaatkan sebagai sumber investasi yaitu utang luar negeri pemerintah. Daryanto [1] mengatakan bahwa utang luar negeri pemerintah dibutuhkan untuk membiayai defisit anggaran. Ketika kebutuhan dana kegiatan ekonomi yang berasal dari dalam negeri tidak cukup tersedia, maka kekurangannya dipenuhi dari luar negeri. Untuk menutupi kekurangan modal tersebut, pemerintah meminta bantuan dari negara lain dalam bentuk utang atau yang biasa disebut utang luar negeri pemerintah.

Anggaran pemerintah cenderung mengalami defisit, dan sejak tahun 2011 anggaran pemerintah tidak pernah mengalami surplus. Hingga kuartal kedua tahun 2013, defisit anggaran pemerintah Indonesia sebesar 9,8 triliun rupiah. Defisit anggaran ini menyebabkan pemerintah menarik utang luar negeri untuk memenuhi

kebutuhan anggaran. Utang luar negeri pemerintah Indonesia cenderung meningkat, mengalami pertumbuhan sebesar 50,5 persen dari tahun 2001 hingga kuartal kedua tahun 2013. Peningkatan utang luar negeri pemerintah ini menunjukkan ketergantungan pemerintah akan utang luar negeri dalam memenuhi kebutuhan anggaran yang kurang untuk pelaksanaan kegiatan ekonomi Indonesia.

Peranan utang luar negeri pemerintah dalam perekonomian dapat dijelaskan oleh beberapa pandangan para ahli yang melihatnya dari sisi yang berbeda-beda. Pandangan Neo Klasik mengindikasikan bahwa kenaikan utang luar negeri untuk membiayai pengeluaran pemerintah hanya menaikkan pertumbuhan ekonomi dalam jangka pendek, namun dalam jangka panjang akan menurunkan Produk Domestik Bruto (Barsky et al [2]). Paham Keynesian berpendapat bahwa kebijakan menutup defisit anggaran dengan utang luar negeri dalam jangka pendek akan menguntungkan perekonomian dengan adanya pertumbuhan ekonomi menurut Eisner [3]. Pandangan Ricardian yang dijelaskan oleh Barro [4] dan [5] serta Evans [6] menunjukkan bahwa kebijakan utang luar negeri untuk membiayai defisit anggaran belanja pemerintah tidak akan memengaruhi pertumbuhan ekonomi.

Dalam hal ini, utang luar negeri dilihat sebagai modal dalam penciptaan pendapatan nasional yang selanjutnya digunakan untuk mengukur pertumbuhan ekonomi. Pertumbuhan ekonomi merupakan sebuah proses kenaikan output per kapita dalam jangka panjang dan terlihat adanya aspek dinamis dalam suatu perekonomian, yaitu terlihat bagaimana perekonomian suatu negara yang berkembang atau berubah dari waktu ke waktu. Suatu perekonomian dikatakan tumbuh apabila dalam jangka waktu yang cukup panjang mengalami kenaikan output per kapita berdasarkan Boediono [7]. Indikator yang digunakan untuk mengukur pertumbuhan ekonomi adalah Produk Domestik Bruto (PDB). Menurut Todaro [8], pertumbuhan ekonomi yang tinggi dan prosesnya yang berkelanjutan merupakan kondisi utama bagi kelangsungan pembangunan ekonomi. Karena penduduk dan kebutuhan ekonomi semakin bertambah, maka penambahan pendapatan sangat dibutuhkan setiap tahunnya. Hal ini hanya bisa didapat lewat peningkatan output agregat (barang dan jasa) atau Produk Domestik Bruto (PDB) setiap tahun.

Mankiw [9] menyatakan bahwa Produk Domestik Bruto merupakan ukuran terbaik dari kinerja perekonomian. Tujuan PDB adalah meringkas aktivitas ekonomi dalam nilai uang tunggal selama periode waktu tertentu. PDB juga menjadi indikator yang menyatakan berapa banyak uang yang mengalir mengelilingi aliran sirkuler perekonomian per unit waktu atau dapat dihitung sebagai pendapatan total dari setiap orang dalam perekonomian dan pengeluaran total atas output barang dan jasa perekonomian. Oleh karena itu PDB juga merupakan cerminan dari kinerja ekonomi karena mengukur total pendapatan masyarakat suatu perekonomian dan output barang dan jasa yang dapat memenuhi permintaan rumah tangga, perusahaan dan pemerintah.

Dalam penelitian ini, PDB yang digunakan yaitu PDB atas dasar harga konstan dengan tahun dasar 2000, yang dihitung dengan pendekatan produksi. Sebagai nilai pasar semua barang dan jasa akhir yang diproduksi dalam perekonomian selama kurun waktu tertentu, faktor produksi yang meningkatkan PDB terdiri dari kapital dan tenaga kerja, sesuai dengan fungsi produksi Cobb

Douglas (Mankiw [9]). Utang luar negeri merupakan salah satu modal yang digunakan dalam pembangunan, selain Penanaman Modal Asing (PMA), ekspor dan tabungan domestik yang akan dibahas dalam penelitian ini.

Utang luar negeri pemerintah merupakan penerimaan negara baik dalam bentuk devisa dan/atau devisa yang dirupiahkan, rupiah, maupun dalam bentuk barang dan/atau jasa yang diperoleh dari pemberi pinjaman luar negeri yang harus dibayar kembali dengan persyaratan tertentu, tidak termasuk didalamnya utang luar negeri yang dipinjam oleh masyarakat Indonesia secara pribadi maupun secara kolektif sebagai suatu perusahaan atau kegiatan usaha (Bank Indonesia [10]). Menurut Todaro [8], utang luar negeri adalah seluruh pinjaman serta konsensual baik secara resmi dalam bentuk uang tunai maupun bentuk-bentuk aktiva yang lainnya secara umum ditujukan untuk mengalihkan sejumlah sumber daya negara-negara maju ke negara berkembang untuk kepentingan pembangunan atau mempunyai maksud sebagai distribusi pendapatan. Pinjaman atau utang luar negeri merupakan salah satu alternatif pembiayaan yang diperlukan dalam pembangunan dan dapat digunakan untuk meningkatkan investasi guna menunjang pertumbuhan ekonomi (Basri [11]).

Investasi yang berasal dari luar negeri lainnya adalah Penanaman Modal Asing (PMA), yaitu total investasi yang dilakukan oleh warga negara asing baik perorangan maupun badan usaha ke suatu negara tertentu. Data PMA yang digunakan dalam penelitian ini yaitu PMA yang sudah direalisasikan. Modal asing dalam industrialisasi pembangunan ekonomi dapat menciptakan kesempatan kerja. Modal asing juga membantu memodernisasi masyarakat dan memperkuat sektor negara maupun sektor swasta. Penggunaan modal asing yang demikian penting untuk mempercepat pembangunan ekonomi negara-negara terbelakang (Jhingan [12]).

Pemanfaatan modal asing seyogyanya hanya merupakan tambahan ketika modal dari dalam negeri tidak mencukupi. Oleh karena itu penelitian ini menggunakan modal asing dan dalam negeri sebagai faktor produksi yang sangat penting dalam meningkatkan pendapatan nasional. Modal dalam negeri yang digunakan adalah tabungan domestik dan ekspor. Tabungan domestik (*domestic saving*) dapat didefinisikan sebagai pendapatan total dalam perekonomian yang tersisa setelah dipakai untuk pengeluaran pemerintah dan konsumsi. Model Solow mengemukakan bahwa tingkat tabungan adalah determinan penting dari persediaan modal. Jika tingkat tabungan tinggi, perekonomian akan mempunyai persediaan modal yang besar dan tingkat output yang tinggi. (Mankiw [9]). Sedangkan ekspor adalah suatu kegiatan pengiriman barang dan jasa yang dijual oleh penduduk suatu negara kepada penduduk negara lain untuk mendapatkan mata uang asing dari negara pembeli. Ekspor terjadi karena adanya kelebihan penawaran domestik suatu negara. Selain itu, suatu negara akan melakukan ekspor apabila negara tersebut mempunyai keunggulan dalam komoditi tertentu, dilihat dari melimpahnya sumber daya yang dimiliki, maupun dari keefisienan memproduksi komoditi tertentu dibandingkan dengan negara lain. Dengan demikian ekspor memberikan pemasukan devisa bagi negara yang bersangkutan yang kemudian akan digunakan untuk membiayai kebutuhan impor maupun pembiayaan program pembangunan di dalam negeri.

Peranan modal dalam meningkatkan pertumbuhan ekonomi di Indonesia

merupakan hal menarik yang perlu diteliti karena beberapa pandangan ekonom yang berbeda. Secara khusus, munculnya perbedaan pandangan mengenai utang luar negeri terhadap pertumbuhan ekonomi memotivasi peneliti untuk menguji secara statistik apakah utang luar negeri pemerintah menjadi modal yang memberikan dampak positif atau negatif terhadap pertumbuhan ekonomi di Indonesia, sebagaimana diringkas dalam kerangka pikir yang digunakan (Gambar 1). Guna menganalisis pertumbuhan ekonomi Indonesia dalam jangka pendek dan jangka panjang, secara khusus penelitian ini bertujuan untuk: 1) menganalisis pertumbuhan ekonomi dan aliran modal, baik yang berasal dari dalam negeri maupun luar negeri; 2) menganalisis pengaruh variabel modal terhadap PDB dalam jangka panjang dan jangka pendek. Data penelitian yang digunakan bersumber dari Badan Pusat Statistik (BPS), Statistik Utang Luar Negeri Indonesia (SULNI) dan Neraca Arus Dana (NAD) yang diterbitkan oleh Bank Indonesia (BI), serta Badan Koordinasi Penanaman Modal (BKPM). Data yang digunakan berupa data kuartalan (dalam bentuk time series) periode 2001-2013.

Data deret waktu dikumpulkan secara berkala pada interval tertentu, misalnya kuartalan seperti dalam penelitian ini. *Error Correction Model* (ECM) adalah model yang dapat merepresentasikan data series waktu non stasioner terkointegrasi yang memiliki hubungan jangka panjang (adanya kointegrasi). Prosedur analisisnya adalah sebagai berikut:

1) Melakukan uji akar unit untuk menguji stasioneritas pada masing-masing variabel yang digunakan dan berada pada *difference* yang sama. Teknik analisis yang menggunakan data deret waktu berlandaskan pada asumsi bahwa data bersifat konstan dan bebas dari waktu ke waktu sehingga data yang digunakan dapat memberikan hasil yang terhindar dari kemungkinan adanya kesalahan atau bias terhadap estimasi (Gujarati [13]). Data yang tidak stasioner akan menyebabkan terjadinya autokorelasi atau pengaruh variabel pada periode sebelumnya terhadap periode saat ini, sehingga berisiko menghasilkan estimasi yang bersifat palsu (*spurious regression*). Uji yang digunakan adalah uji *Augmented Dickey Fuller* atau biasa disebut sebagai *unit root test* dengan hipotesis H_0 terdapat sebuah unit root atau data tidak stasioner.

2) Jika seluruh variabel stasioner pada *difference* yang sama maka dilakukan estimasi persamaan jangka panjang, yang dapat terjadi karena adanya kointegrasi antar variabel yang nonstasioner di tingkat *level*. Kombinasi linier yang dilakukan pada variabel-variabel nonstasioner akan menunjukkan keadaan yang terkointegrasi apabila hasil dari kombinasi linier tersebut bersifat stasioner (Enders [14]). ECM dilakukan ketika kointegrasi masih belum mampu untuk menyesuaikan atau mengembalikan keseimbangan pada jangka pendek, sehingga ECM digunakan untuk mengoreksi ketidakseimbangan yang terjadi pada jangka pendek (Gujarati [13]). Persamaan jangka panjang dalam penelitian ini adalah:

$$LN_PDB = c_0 + c_1LN_ULNP + c_2LN_PMA + c_3LN_EKS + c_4LN_TD \quad (1)$$

dengan LN_PDB pertumbuhan PDB, LN_UNLP pertumbuhan utang luar negeri pemerintah, LN_PMA pertumbuhan PMA, LN_EKS adalah pertumbuhan ekspor, LN_TD pertumbuhan tabungan domestik dan c_1 , c_2 , c_3 dan c_4 adalah intersep.

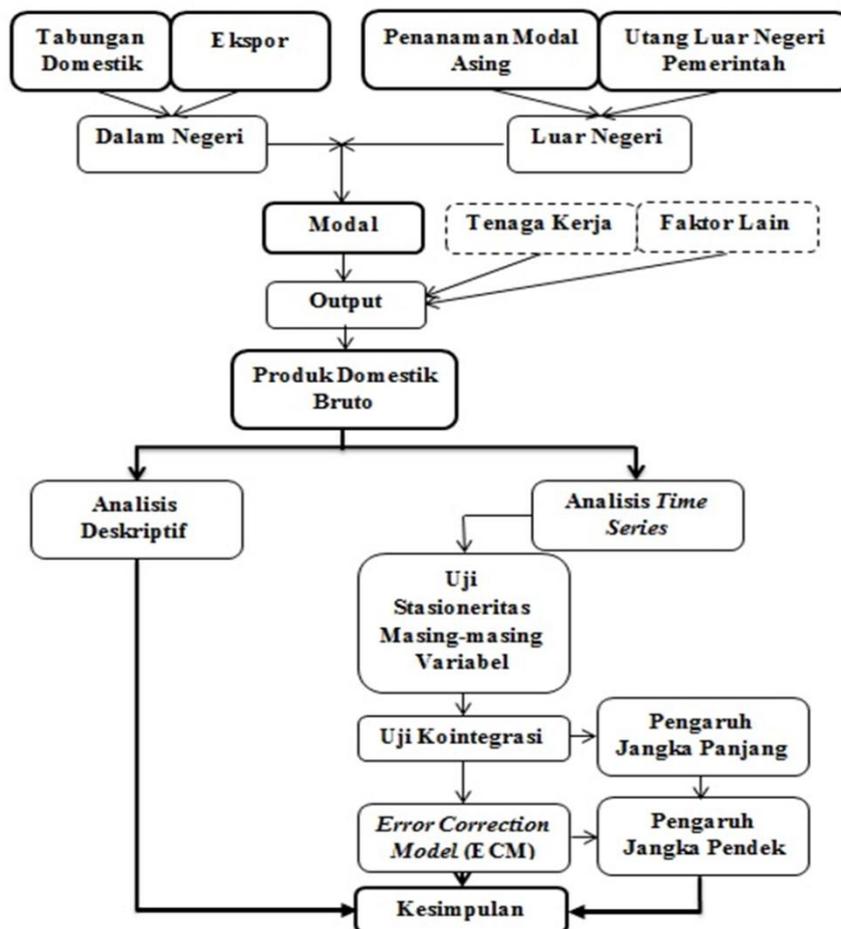
3) Membentuk persamaan jangka pendek dengan menggunakan *Error Correction Term* (ECT), dengan cara meregresikan variabel-variabel yang stasioner pada order yang sama dan lag pertama residual dari hasil persamaan jangka

panjangnya. Persamaannya adalah:

$$LN_PDB = c_0 + c_1LN_ULNP + c_2LN_PMA + c_3LN_EKS + c_4LN_TD + c_5ECT \quad (2)$$

Koefisien regresi dari ECT (c_5) diharapkan bernilai negatif dan kurang dari satu, yang menunjukkan adanya λ (*speed of adjustment*) atau akan terjadi keseimbangan pada jangka panjang. Adanya ECT dapat melihat ketidakseimbangan perubahan variabel dependen akibat variabel-variabel independen dalam model.

- 4) Menilai kesesuaian model dengan melihat keberartian model dengan menggunakan uji F untuk mengetahui pengaruh variabel-variabel independen secara bersama-sama terhadap variabel dependen dan uji t untuk pengujian regresi secara parsial masing-masing variabel independennya berpengaruh terhadap variabel dependen. Selanjutnya adalah pemenuhan asumsi klasik yang terdiri dari uji Durbin Watson untuk pengujian non autokorelasi, uji Jarque-Bera untuk normalitas, uji White untuk homoskedastisitas dan *Variance Inflation Factor/VIF* untuk non multikolinieritas.

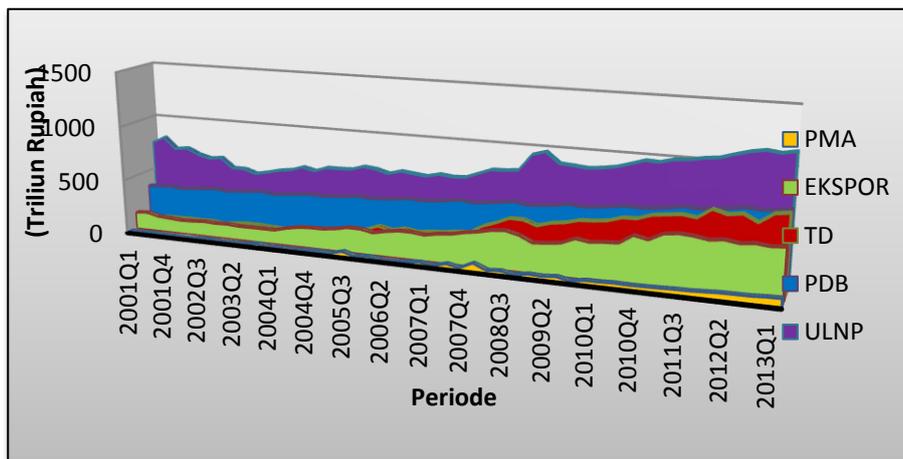


Gambar 1 Kerangka Pikir Penelitian

2. Hasil – Hasil Utama

Perkembangan PDB dan Aliran Modal di Indonesia

Produk Domestik Bruto Indonesia mengalami pertumbuhan yang cenderung meningkat tiap periodenya. Penurunan PDB yang terbesar terjadi pada kuartal IV tahun 2002 dengan 3,86 persen. Hal ini diakibatkan oleh pada tahun 2002 terjadi pengalihan kekuasaan pemerintahan Indonesia. Namun, pertumbuhan PDB Indonesia cenderung positif dengan rata-rata pertumbuhan sebesar 1,35 persen dengan peningkatan terbesar terjadi pada kuartal III tahun 2009 sebesar 3,87 persen. Hal tersebut menunjukkan bahwa perekonomian Indonesia tetap stabil walaupun pada tahun 2008 terjadi krisis keuangan global. Pada akhir periode, pertumbuhan PDB mencapai 2,57 persen. Perkembangan PDB menunjukkan pertumbuhan ekonomi karena yang digunakan adalah PDB riil (Gambar 2).



Gambar 2 Perkembangan PDB, Utang Luar Negeri Pemerintah, Penanaman Modal Asing (PMA), Ekspor dan Tabungan Domestik Indonesia Tahun 2001 – 2013 (Trilyun Rupiah)

Secara rata-rata, utang luar negeri pemerintah cenderung mengalami pertumbuhan yang positif sebesar 0,95 persen walaupun mengalami penurunan dan peningkatan secara bergantian. Utang luar negeri pemerintah yang terkecil terjadi pada kuartal II 2003 sebesar 536,9 triliun rupiah. Nilai utang luar negeri pemerintah mengalami peningkatan yang tinggi pada periode IV 2008 sebanyak 18,69 persen dari periode sebelumnya menjadi sebesar 932,2 triliun rupiah. Hal ini disebabkan oleh terjadinya krisis keuangan global pada tahun 2008 yang menyebabkan pemerintah Indonesia membutuhkan modal yang lebih besar untuk mempertahankan kegiatan produksi dalam negeri agar tetap stabil. Namun, setelah periode krisis, utang luar negeri pemerintah mulai kembali stabil dengan pertumbuhan rata-rata sebesar 1,2 persen. Pengendalian ekonomi yang baik berdampak pada menurunnya nilai utang yang terbesar terjadi pada kuartal II 2009 sebanyak 9,5 persen dari nilai periode sebelumnya menjadi 874,2 triliun rupiah. Hingga akhir periode, utang luar negeri pemerintah mengalami pertumbuhan sebesar 2 persen dengan mencapai nilai 1.132 triliun rupiah.

Perkembangan penanaman modal asing di Indonesia mengalami fluktuasi

yang cukup tinggi. Pada awal periode, realisasi penanaman modal asing di Indonesia sebesar 8,6 triliun rupiah dan terus berfluktuasi dengan rata-rata pertumbuhan sebesar 31,78 persen hingga mencapai nilai 71,2 triliun rupiah pada akhir periode. Pertumbuhan penanaman modal asing di Indonesia melampaui seratus persen terjadi pada delapan periode dengan peningkatan terbesar terjadi pada kuartal I 2008 sebesar 328,06 persen dari periode sebelumnya menjadi sebesar 72,8 triliun rupiah. Walau mengalami pertumbuhan melampaui seratus persen pada beberapa periode, penanaman modal asing di Indonesia juga banyak mengalami pertumbuhan yang negatif. Pertumbuhan negatif terbesar terjadi pada kuartal IV 2005 dimana penanaman modal asing menurun 71,63 persen dari periode sebelumnya.

Nilai ekspor yang diperoleh Indonesia melalui perdagangan luar negerinya cenderung mengalami trend positif, tetapi juga berfluktuasi. Dengan pertumbuhan rata-rata sebesar 1,09 persen, nilai ekspor Indonesia mengalami pertumbuhan negatif terbesar pada kuartal I 2009 yaitu turun 17,06 persen dari periode sebelumnya. Hal ini disebabkan oleh pasar internasional yang mengalami gangguan akibat krisis global yang terjadi pada tahun 2008 sehingga kemampuan pasar impor internasional mengalami penurunan dan berdampak pada hasil ekspor Indonesia. Namun, nilai ekspor Indonesia juga mengalami peningkatan yang terbesar pada kuartal IV 2010 yang mencapai 421,3 triliun rupiah, meningkat 22,97 persen dari periode sebelumnya. Nilai ekspor Indonesia mencapai 453,3 triliun rupiah pada akhir periode penelitian.

Keadaan tabungan domestik Indonesia cenderung memiliki trend positif dan mengalami fluktuasi yang tidak cukup berarti. Dengan nilai terkecil pada awal periode sebesar 80,7 triliun rupiah, tabungan domestik Indonesia terus mengalami peningkatan dengan rata-rata pertumbuhan sebesar 2,2 persen hingga mencapai nilai tertingginya pada akhir periode sebesar 692,9 triliun rupiah. Walau demikian, tabungan domestik Indonesia pernah mengalami penurunan terbesarnya pada kuartal IV 2002 hingga mencapai 94,7 triliun rupiah yang turun 12,15 persen dari periode sebelumnya. Namun, tabungan domestik Indonesia juga mengalami peningkatan terbesarnya pada kuartal I 2008 yang mencapai 25,53 persen hingga mencapai nilai 328,9 triliun rupiah.

Pengaruh Aliran Modal terhadap PDB di Indonesia

PDB merupakan ukuran yang digunakan untuk menghitung pertumbuhan ekonomi dan ukuran ini tidak dapat dilepaskan dari pengaruh variabel-variabel ekonomi lainnya, terutama modal. Modal yang digunakan dalam proses produksi dibagi menjadi modal yang berasal dari dalam negeri dan modal yang berasal dari luar negeri. Modal yang berasal dari dalam negeri berupa nilai ekspor dan tabungan domestik, sedangkan modal dari luar negeri berupa utang luar negeri pemerintah dan Penanaman Modal Asing (PMA). Variabel-variabel ini yang akan dilihat pengaruhnya terhadap Produk Domestik Bruto baik dalam jangka panjang maupun jangka pendek, dengan menggunakan model analisis *Error Correction Mechanism* (ECM) atau model koreksi kesalahan untuk menganalisis pengaruh antara variabel bebas terhadap variabel tak bebas. Semua variabel yang digunakan dalam penelitian ini sudah dibentuk dalam logaritma natural (distandardisasi) untuk menyamakan satuan tiap variabel sehingga analisis yang dilakukan menjadi lebih mudah. Pada tahap pertama, dilakukan uji stasioneritas dan diperoleh bahwa semua variabel penelitian tidak stasioner pada level dan stasioner pada *difference* pertama. Selanjutnya dilakukan uji kointegrasi, yang dilakukan untuk mendeteksi stabilitas

hubungan jangka panjang antara dua variabel atau lebih. Jika hasil regresi antar variabel X dan Y menghasilkan residual yang stasioner maka variabel X dan Y dikatakan terkointegrasi, yang berarti mempunyai hubungan jangka panjang. Menurut Gujarati [13], pengujian ini hanya valid jika dilakukan pada data asli yang nonstasioner. Enders [14] memberikan catatan penting tentang definisi kointegrasi, diantaranya adalah kointegrasi merupakan kombinasi linear dari variabel-variabel yang *seriesnya* nonstasioner dan semua variabel yang diuji harus terintegrasi (stasioner) pada order yang sama. Hasil uji kointegrasi menunjukkan bahwa terdapat kointegrasi antara variabel dalam penelitian, sehingga metode ECM lebih efektif digunakan dalam menganalisis pengaruh modal terhadap pertumbuhan ekonomi di Indonesia.

ECM merupakan kelanjutan dari pengujian kointegrasi yang bertujuan untuk mengoreksi penyimpangan terhadap keseimbangan jangka panjang. Dalam ECM, hubungan dinamis jangka pendek antarvariabel dalam model dipengaruhi oleh deviasi/penyimpangan dari keseimbangan jangka panjang (Enders [14]).

Hasil penaksiran terhadap persamaan jangka pendek pada penelitian ini telah memenuhi asumsi normalitas, homoskedastisitas, nonautokorelasi dan nonmultikolinearitas sehingga menunjukkan hasil estimasi OLS yang dihasilkan bersifat BLUE dan layak untuk diinterpretasikan. Dalam jangka panjang maupun jangka pendek, secara bersama-sama utang luar negeri pemerintah, PMA, ekspor dan tabungan domestik memberikan pengaruh yang signifikan terhadap pertumbuhan ekonomi Indonesia dengan tingkat kepercayaan 95 persen. Dalam jangka panjang, perubahan pertumbuhan ekonomi Indonesia dipengaruhi oleh PMA, ekspor dan tabungan domestik secara signifikan dan positif. Tabungan domestik memiliki pengaruh yang lebih besar terhadap pertumbuhan ekonomi Indonesia dibandingkan dengan PMA dan ekspor, yang ditunjukkan dengan nilai koefisien yang paling besar yaitu 0,2124. Artinya, jika terjadi peningkatan tabungan domestik sebesar 1 persen, maka PDB akan meningkat sebesar 0,21 persen, dan sebaliknya. Kenaikan/penurunan 1 persen pada PMA akan meningkatkan/menurunkan produk domestik bruto sebesar 0,02 persen. Sedangkan PDB akan meningkat/menurun sebesar 0,07 persen jika ekspor meningkat/menurun sebesar 1 persen. Utang luar negeri pemerintah dalam jangka panjang tidak memberikan pengaruh yang signifikan terhadap pertumbuhan ekonomi Indonesia.

Tabel 1 Output Eviews yang Menunjukkan Persamaan Jangka Panjang

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	14.01282	0.705130	19.87269	0.0000
LN_ULNP	0.004700	0.045567	0.103137	0.9183
LN_EKS	0.073636	0.041617	1.769372	0.0836
LN_TD	0.212413	0.025915	8.196451	0.0000
LN_PMA	0.021507	0.009622	2.235222	0.0304
R-squared	0.976515	Mean dependent var		20.00076
Adjusted R-squared	0.974427	S.D. dependent var		0.200628
S.E. of regression	0.032083	Akaike info criterion		-3.946321
Sum squared resid	0.046320	Schwarz criterion		-3.755118
Log likelihood	103.6580	Hannan-Quinn criter.		-3.873510
F-statistic	467.7756	Durbin-Watson stat		0.710423
Prob(F-statistic)	0.000000			

Keterangan variabel:

LN_UNLP : Pertumbuhan Utang Luar Negeri Pemerintah
 LN_EKS : Pertumbuhan Ekspor
 LN_TD : Pertumbuhan Tabungan Domestik
 LN_PMA : Pertumbuhan Penanaman Modal Asing

Dari persamaan jangka pendek, diperoleh bahwa PMA dan tabungan domestik signifikan berpengaruh terhadap pertumbuhan ekonomi Indonesia secara positif. Sama seperti pada jangka panjang, tabungan domestik memberikan pengaruh yang dominan terhadap PDB Indonesia dengan memberikan perubahan peningkatan/penurunan sebesar 0,09 persen pada tiap 1 persen perubahan peningkatan/penurunannya pada jangka pendek. Sedangkan PMA yang mengalami perubahan meningkat/menurunnya sebesar 1 persen akan menyebabkan peningkatan/penurunan PDB sebesar 0,01 persen. Utang luar negeri pemerintah dan ekspor tidak memberikan pengaruh yang signifikan terhadap produk domestik bruto Indonesia dalam jangka pendek.

Koefisien determinasi jangka panjang sebesar 0,9744 menunjukkan bahwa dalam jangka panjang, 97,44 persen dari variasi PDB dapat dijelaskan oleh variasi utang luar negeri pemerintah, PMA, ekspor dan tabungan domestik, sedangkan 2,56 persen lainnya dipengaruhi oleh variabel-variabel lain diluar model. Dalam jangka pendek, hanya 16,77 persen dari variasi PDB dapat dijelaskan oleh variasi utang luar negeri pemerintah, PMA, ekspor dan tabungan domestik. Selain itu, koefisien kointegrasi sebesar -0,2508 menjelaskan bahwa kecepatan *error correction* untuk mengoreksi perilaku modal dalam jangka pendek untuk menuju keseimbangan jangka panjang hanya sebesar 25,08 persen sehingga membutuhkan waktu yang cukup lama untuk mencapai posisi keseimbangan jangka panjang, karena pada kuartal pertama hanya terkoreksi sebesar 25,08 persen.

Tabel 2 Output Eviews yang Menunjukkan Persamaan Jangka Pendek

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.007719	0.003583	2.154401	0.0369
DLN_ULNP	-0.058091	0.060856	-0.954560	0.3451
DLN_PMA	0.009488	0.004752	1.996814	0.0522
DLN_EKS	0.071040	0.042808	1.659489	0.1043
DLN_TD	0.090202	0.037370	2.413740	0.0201
ECT_1	-0.250856	0.115033	-2.180730	0.0347
R-squared	0.254466	Mean dependent var		0.013465
Adjusted R-squared	0.167776	S.D. dependent var		0.023173
S.E. of regression	0.021140	Akaike info criterion		-4.761068
Sum squared resid	0.019216	Schwarz criterion		-4.529417
Log likelihood	122.6462	Hannan-Quinn criter.		-4.673180
F-statistic	2.935356	Durbin-Watson stat		2.344844
Prob(F-statistic)	0.022905			

Keterangan variabel:

DLN_UNLP : Perubahan pertumbuhan Utang Luar Negeri Pemerintah
 DLN_EKS : Perubahan pertumbuhan Ekspor
 DLN_TD : Perubahan pertumbuhan Tabungan Domestik
 DLN_PMA : Perubahan pertumbuhan Penanaman Modal Asing
 ECT_1 : *Error Correction Term* (koefisien kointegrasi)

Penelitian ini menunjukkan bahwa dalam jangka pendek maupun jangka panjang, utang luar negeri pemerintah tidak berpengaruh signifikan terhadap PDB. Hasil analisis ini tidak sesuai dengan teori dari Rostow yang menyatakan bahwa mobilisasi dana tabungan (dalam mata uang domestik maupun asing) yang salah satunya adalah dalam bentuk utang luar negeri akan memengaruhi peningkatan pertumbuhan ekonomi di negara-negara berkembang. Hal ini disebabkan oleh pengalokasian utang luar negeri pemerintah yang tidak hanya digunakan sebagai modal investasi dalam kegiatan produksi, tetapi dialokasikan juga untuk kebijakan pemberantasan pengangguran dan kemiskinan di dalam negeri, menurut Kementerian Keuangan yang diungkapkan oleh Firmanzah [15]. Selain itu, dalam jangka pendek utang luar negeri pemerintah yang baru ditarik digunakan juga untuk membayar cicilan utang sebelumnya sehingga justru menyebabkan penurunan pertumbuhan PDB dalam jangka pendek. Penelitian ini menunjukkan posisi keadaan utang luar negeri Indonesia sesuai dengan pandangan Ricardian yang menyatakan bahwa kebijakan utang luar negeri untuk membiayai defisit anggaran belanja pemerintah tidak akan memengaruhi pertumbuhan ekonomi baik dalam jangka panjang maupun jangka pendek.

Hasil analisis ini menguatkan ketidaksetujuan para pakar ekonomi mengenai peranan positif eksistensi utang luar negeri pemerintah dalam perekonomian di negara-negara berkembang. Hasil penelitian ini sejalan dengan Syaparudin dan Hermawan [16] serta Susilo dan Seyastuti [17]. Utang luar negeri pemerintah justru menjadi bumerang bagi negara penerima. Beberapa alasan yang menyebabkan kegagalan dalam menggunakan dana pinjaman untuk pembangunan ekonomi negara berkembang antara lain adalah (1) ketidakmampuan pemerintah negara penerima dalam memanfaatkan utang secara efektif, (2) utang luar negeri pemerintah lebih bermotifkan politik dibandingkan ekonomi, (3) utang yang diterima dikorupsi oleh pejabat negara berkembang, dan (4) tidak bekerjanya mekanisme pasar akibat kegagalan pasar (*market failure*) seperti monopoli dan oligopoli.

3. Kesimpulan

Secara umum, perkembangan PDB, utang luar negeri pemerintah, ekspor dan tabungan domestik tidak begitu berfluktuasi dan memiliki trend yang cenderung positif selama tahun 2001 hingga 2013. Sedangkan perkembangan PMA sangat berfluktuasi setiap periodenya, tetapi juga memiliki trend yang cenderung positif dalam kurun waktu yang sama. Dalam jangka panjang, PDB Indonesia dipengaruhi secara nyata dengan arah yang positif oleh PMA, ekspor dan tabungan domestik. Tabungan domestik memberikan pengaruh yang paling besar terhadap PDB dibandingkan dengan variabel lainnya. Utang luar negeri pemerintah tidak memberikan pengaruh yang nyata secara statistik terhadap PDB Indonesia selama periode penelitian. Dalam jangka pendek, PMA dan tabungan domestik secara nyata memengaruhi PDB dengan arah yang positif. Tabungan domestik juga memberikan pengaruh yang paling besar dalam jangka pendek ini. Ekspor dan utang luar negeri pemerintah memberikan pengaruh yang nyata secara statistik terhadap PDB Indonesia selama periode penelitian. Hasil *Error Correction Term* menunjukkan bahwa dibutuhkan waktu yang relatif lama bagi PDB untuk mencapai posisi keseimbangan jangka panjang dari ketidakseimbangan jangka pendeknya.

Hasil analisis pengaruh utang luar negeri pemerintah terhadap PDB Indonesia menunjukkan pengelolaan utang luar negeri pemerintah yang kurang efektif dalam mendukung peningkatan PDB. Hal ini membuktikan pandangan bahwa utang luar negeri dapat menjadi beban pembangunan ekonomi, bukan sebagai sumber pembiayaan pembangunan yang efektif. Kondisi ini menempatkan Indonesia sesuai dengan pandangan Ricardian, bahwa utang luar negeri pemerintah tidak memberikan pengaruh yang nyata terhadap pertumbuhan ekonomi Indonesia selama tahun 2001 hingga 2013.

Referensi

- [1] Daryanto, 2001, *Hutang Luar Negeri Indonesia: Masalah dan Alternatif Solusinya*, Agrimedia.
- [2] Barsky et al, 1986, *Ricardian Consumers with Keynesian Propensities*, American Economic Review.
- [3] Eisner, Robert, 1989, Budget Deficit: Rhetoric and Reality, *The Journal of Economic Perspectives*.
- [4] Barro, J. Robert, 1974, Are Government Bond Net Wealth?, *Journal of Political Economic Perspectives*.
- [5] Barro, J. Robert, 1989, The Ricardian Approach to Budget Deficits, *Journal of Political Economic Perspectives*.
- [6] Evans, Paul, 1988, Are Consumers Ricardian? Evidence for The United States, *Journal of Political Economy*.
- [7] Boediono, 1991, *Mikroekonomi*, Yogyakarta: BPFE.
- [8] Todaro, 2000, *Pembangunan Ekonomi Dunia Ketiga Edisi Ketujuh*, Jakarta: Erlangga.
- [9] Mankiw, Gregory N., 2007, *Macroeconomics Sixth Edition*, Worth Publishers.
- [10] Bank Indonesia, 2010, *Statistik Utang Luar Negeri Indonesia*, Jakarta: Bank Indonesia.
- [11] Basri, Faisal, 2004, *Kebijakan Pinjaman Luar Negeri*, Jakarta: Bank Indonesia.
- [12] Jhingan, 2007, *Ekonomi Pembangunan dan Perencanaan*. Jakarta: PT. Raja Grafindo Persada.
- [13] Gujarati, Damodar, 2004, *Basic Econometric Fourth Edition*, New York: McGraw-Hill.
- [14] Enders, Walter, 2004, *Applied Econometric Time Series 2nd Edition*, United States of America: John Wiley & Sons, University of Alabama.
- [15] Firmanzah, 22 Oktober 2012, Sekretariat Kabinet Republik Indonesia, *Pemanfaatan Utang Negara*. Diakses pada tanggal 30 Mei 2014 melalui <http://www.setkab.go.id/artikel-6115-pemanfaatan-utang-negara.html>.
- [16] Syaparudin dan Heri Hermawan, 2005, Hutang Luar Negeri Pemerintah: Kajian dari Sisi Permintaan dan Pengaruhnya terhadap Produk Domestik Bruto Indonesia Periode 1980-2002, *Simposium Riset Ekonomi*.
- [17] Susilo, Y.S dan Seyastuti, 2013, Pengaruh Utang Luar Negeri dan Ekspor terhadap Pertumbuhan Ekonomi Indonesia Periode 2000.1-2008.4., *Journal Akuntansi dan Manajemen*.

PERAMALAN PRODUKSI DAN EKSPOR MINYAK KELAPA SAWIT INDONESIA TRIWULAN I 2015 – TRIWULAN III 2016 (PERBANDINGAN METODE *EXPONENTIAL SMOOTHING* DENGAN ARIMA)

RIO KUNCORO JATIKUSUMA¹, DEWI PURWANTI²

¹ Badan Pusat Statistik, 12.7348@stis.ac.id

² Sekolah Tinggi Ilmu Statistik, dewip@stis.ac.id

Abstrak. Indonesia merupakan negara produsen utama minyak kelapa sawit (CPO) dunia dengan rata-rata pertumbuhan produksi mencapai 7,9% tiap tahunnya. Industri minyak kelapa sawit merupakan salah satu industri yang prospek ekspornya pun cukup cerah. Namun demikian, kegiatan produksi dan ekspor tetap harus dipantau agar tidak terjadi kelebihan produksi maupun kelangkaan domestik akibat ekspor yang berlebih. Oleh karena itu, pemerintah maupun swasta harus membuat kebijakan-kebijakan yang tepat dengan berdasar pada data statistik. Sedangkan, pada kenyataannya data komoditas kelapa sawit yang komprehensif kurang *up-to-date*, sehingga perlu dilakukan peramalan agar kebijakan dapat dibuat lebih dini. Penelitian ini akan meramalkan volume produksi dan ekspor minyak kelapa sawit Indonesia secara terpisah dengan menggunakan model peramalan terbaik. Minyak kelapa sawit yang dimaksud di sini adalah produk olahan dalam bentuk *Crude Palm Oil* (CPO) dan berkode HS 151110000. Metode yang akan dibandingkan adalah *Autoregressive Integrated Moving Average* dan *Exponential Smoothing*. Hasil perbandingan menunjukkan bahwa SARIMA lebih baik untuk peramalan variabel produksi, sedangkan *Double Exponential Smoothing* lebih baik untuk peramalan variabel ekspor. Dari hasil peramalan diketahui bahwa produksi untuk periode 2016 menunjukkan kenaikan, begitu juga halnya dengan ekspornya.

Kata kunci : *crude palm oil*, ARIMA, *Seasonal ARIMA*, *Double Exponential Smoothing* *Triple Exponential Smoothing*

1. Pendahuluan

Hingga saat ini, sebagian besar penduduk Indonesia masih berkecimpung dalam dunia pertanian. Pada Agustus 2015, dari jumlah penduduk usia 15 tahun ke atas yang bekerja, tercatat sekitar 37.748.228 jiwa atau 32,87% bekerja di sektor pertanian. Jumlah pekerja di sektor ini merupakan yang paling besar dibanding sektor lain. Sektor pertanian juga berkontribusi besar terhadap total PDB Indonesia yaitu sebesar 1.560.399,3 miliar rupiah atau sekitar 13,52% pada tahun 2015. Jumlah tersebut memberikan sumbangan terbesar kedua setelah sektor Industri Pengolahan yang menyumbang 20,84% dari total PDB.

Badan Pusat Statistik (BPS) menyatakan bahwa salah satu subsektor pertanian yang cukup besar potensinya adalah subsektor perkebunan. Pada tahun

2015, subsektor ini merupakan penyumbang terbesar terhadap nilai tambah sektor pertanian yaitu sebesar 26,39% atau menyumbang 6,16% terhadap total PDB. Subsektor ini merupakan penyedia bahan baku untuk sektor industri, penyerap tenaga kerja, dan penghasil devisa. Oleh karena itu untuk negara sedang berkembang seperti Indonesia alangkah baiknya apabila sektor ini benar-benar diperhatikan.

Minyak kelapa sawit adalah salah satu komoditas primadona yang berasal dari subsektor perkebunan. Dalam kurun waktu 2012 hingga 2014 tercatat bahwa volume produksi dan ekspor minyak kelapa sawit selalu meningkat. Kementerian Perindustrian mencatat juga bahwa produk kelapa sawit dan olahannya merupakan kelompok hasil industri nomor satu dari sepuluh kelompok hasil industri dengan nilai ekspor terbesar tahun 2015 dan nilainya mencapai sekitar 20,7 milyar US\$. Kontribusi produksi minyak sawit mentah Indonesia adalah terbesar dunia dan hanya memiliki pesaing negara Malaysia dan Thailand, sehingga diharapkan Indonesia dapat memanfaatkan hal tersebut sebagai target tambahan pengembangan dan penjualan CPO pada tahun-tahun berikutnya.

Pembuatan kebijakan pemerintah dan keputusan-keputusan perusahaan yang berkaitan mesti terarah. Salah satu instrumen yang dapat digunakan sebagai bahan pertimbangan adalah ketersediaan data komoditas yang bersangkutan. Namun pada kenyataannya data yang berkaitan dengan kelapa sawit masih belum lengkap atau lambat dirilis. Adanya *time lag* antara pengumpulan data hingga data dipublikasikan ini menyebabkan perlunya dilakukan suatu peramalan agar kebijakan dapat dibuat lebih dini.

Masalah *time lag* tersebut dapat diatasi dengan memproduksi angka taksiran sendiri yang dapat dihasilkan dengan melakukan peramalan. Dalam dunia statistika dan matematika terdapat banyak metode peramalan. Namun diantara yang paling umum digunakan adalah metode ARIMA dan *Exponential Smoothing*. Kedua metode tersebut lebih sering digunakan karena mudah dilakukan dan hasilnya lebih baik dibanding dengan metode lainnya. Dengan demikian, dari dua metode yang umum digunakan tersebut akan dibandingkan mana yang lebih baik.

2. Hasil-Hasil Utama

Minyak kelapa sawit atau lebih dikenal dengan sebutan *Crude Palm Oil* (CPO) berasal dari buah segar kelapa sawit yang didapatkan dengan cara mengekstrak buah sawit tersebut. Produksi kelapa sawit yang ada dalam penelitian ini adalah dalam bentuk produk olahan berupa minyak kelapa sawit (*Crude Palm Oil*). Data produksi dalam penelitian ini merupakan penjumlahan antara produksi pada perusahaan perkebunan dan produksi pada perkebunan rakyat dalam satuan ton. Data volume ekspor minyak sawit Indonesia dalam penelitian ini dalam satuan ton dengan kode HS 151110000.

Penelitian ini menggunakan analisis *time series* (runtun waktu) yang akan diterapkan pada data volume produksi dan ekspor minyak kelapa sawit. Metode analisis *time series* hanya menggunakan series histori dari suatu data variabel yang akan diramalkan untuk membuat suatu model yang dapat digunakan untuk meramalkan nilai-nilai yang akan datang. Dalam melakukan peramalan penelitian ini menggunakan metode ARIMA dan *Exponential Smoothing*.

ARIMA (*Auto Regressive Integrated Moving Average*) merupakan suatu metode analisis *time series*. Metode ini biasa disebut juga “Metode Box-Jenkins” yang mengacu pada dua orang ahli statistika yang telah berhasil mengembangkan

dan menyempurnakan metode tersebut. Notasi yang diusulkan Box dan Jenkins adalah ARIMA (p,d,q) . Kemudian untuk mengatasi masalah data dengan pola musiman digunakan *seasonal* ARIMA dengan notasi $(p,d,q)(P,D,Q)_s$.

Metode *exponential smoothing* merupakan metode peramalan dengan mengadakan penghalusan terhadap masa lalu, yaitu dengan mengambil rata-rata dari nilai beberapa tahun untuk menaksir nilai pada beberapa tahun ke depan. Metode *single exponential* digunakan untuk data tanpa tren dan musiman. *Double exponential* digunakan untuk data dengan tren tanpa musiman. *Triple exponential* digunakan ketika data memiliki pola tren dan juga musiman.

Pemilihan model terbaik untuk peramalan dilakukan dengan memperhatikan nilai *Root Mean Square Error* (RMSE). Nilai RMSE yang mendekati nol berarti semakin baik pula model tersebut.

Model Autoregressive (AR)

Model AR menunjukkan nilai prediksi variabel dependen Y_t hanya merupakan fungsi linier dari sejumlah Y_t sebelumnya (regresi diri). Model *autoregressive* orde p , atau AR(p) mempunyai bentuk sebagai berikut:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t$$

dimana,

Y_t	=	variabel respon pada waktu t
$Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$	=	variabel respon pada masing-masing selang waktu t-1, t-2, ... , t-p.
$\phi_0, \phi_1, \phi_2, \dots, \phi_p$	=	koefisien regresi
ε_t	=	residual
p	=	ordo AR
t	=	periode waktu

Model Moving Average (MA)

Model MA menyatakan bahwa nilai variabel dependen Y_t hanya dipengaruhi oleh nilai residual periode sebelumnya. Model *moving average* orde q , atau MA(q) mempunyai bentuk persamaan sebagai berikut:

$$Y_t = \mu + \varepsilon_t - \omega_1 \varepsilon_{t-1} - \omega_2 \varepsilon_{t-2} - \dots - \omega_q \varepsilon_{t-q}$$

dimana,

μ	=	<i>mean</i> konstanta proses
$\omega_1, \omega_2, \dots, \omega_q$	=	koefisien regresi
$\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$	=	variabel bebas yang merupakan <i>lag</i> dari residual
q	=	ordo MA

Model Auto Regressive Moving Average (ARMA)

Penggabungan antara model AR dan model MA seringkali dapat menjelaskan perilaku data secara lebih baik. Model yang dinotasikan dengan ARMA(p,q) ini mempunyai bentuk persamaan sebagai berikut:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t - \omega_1 \varepsilon_{t-1} - \omega_2 \varepsilon_{t-2} - \dots - \omega_q \varepsilon_{t-q}$$

Single Exponential Smoothing

Metode *Single Exponential Smoothing* atau pemulusan eksponensial tunggal adalah suatu prosedur yang mengulang perhitungan secara terus menerus dengan menggunakan data terbaru. Metode ini digunakan jika data tidak dipengaruhi secara signifikan oleh faktor tren dan musiman. Jika data stasioner, maka dengan menggunakan metode pemulusan tunggal merupakan pendekatan yang cukup baik, namun bila terdapat tren metode ini tidak cukup baik. Bentuk umum persamaan metode pemulusan eksponensial tunggal sebagai berikut:

$$F_{t+1} = \alpha X_t + (1-\alpha)F_t$$

Keterangan:

- α : koefisien/parameter pemulusan, nilai $0 < \alpha < 1$
- X_t : data aktual waktu ke-t
- F_t : nilai peramalan ke-t

Koefisien pemulusan α berfungsi sebagai faktor penimbang. Jika α mendekati nilai 1, maka ramalan yang baru akan lebih mendekati nilai dari observasi. Sebaliknya, jika α mendekati 0, maka ramalan yang baru akan lebih mendekati hasil peramalan sebelumnya. Nilai α yang menghasilkan tingkat kesalahan yang paling kecil (optimum) adalah yang dipilih dalam proses peramalan. Untuk mengetahui berapa nilai α yang paling optimum, dapat digunakan metode simulasi atau *trial and error*.

Double Exponential Smoothing Dua parameter (Holt's Method)

Metode pemulusan eksponensial linear dari *Holt* pada prinsipnya serupa dengan *Brown* kecuali bahwa *Holt* tidak menggunakan rumus pemulusan ganda secara langsung. Sebagai gantinya, *Holt* memuluskan nilai tren dengan parameter yang berbeda dari parameter yang digunakan pada deret asli. Prediksi dari pemulusan eksponensial didapat dengan menggunakan dua parameter pemulusan yaitu α dan β yang bernilai $0 < \alpha < 1$ dan $0 < \beta < 1$. Pada kasus dimana nilai $\alpha = \beta$ metode ini sama dengan metode *double exponential smoothing Brown*."

$$F_{t+m} = L_t + b_t m$$

Pemulusan Keseluruhan:

$$L_t = \alpha X_t + (1-\alpha)(L_{t-1} + b_{t-1})$$

Pemulusan Tren:

$$b_t = \beta (L_t - L_{t-1}) + (1-\beta)b_{t-1}$$

Keterangan:

- α, β : koefisien/parameter pemulusan, nilai $0 < (\alpha, \beta) < 1$
- X_t : data aktual waktu ke-t
- F_t : nilai peramalan ke-t
- m : jumlah periode yang akan diramalkan

Triple Exponential Smoothing

Metode peramalan ini digunakan jika unsur musiman multiplikatif, maksudnya pola tiap musiman berbeda karena adanya fluktuasi yang besar dan berbeda tiap musimnya. Prediksi dari pemulusan eksponensial didapat dengan menggunakan tiga parameter pemulusan yaitu α, β, γ yang bernilai $0 < \alpha < 1, 0 < \beta < 1$ dan $0 < \gamma < 1$. Metode ini didasarkan atas tiga persamaan pemulusan, masing-masing unsur stasioner, tren dan musiman.

$$F_{t+m} = (L_t + b_t m) S_{t-s+m} \tag{2}$$

Pemulusan level:

$$L_t = \alpha \frac{X_t}{S_{t-s}} + (1-\alpha)(L_{t-1} + b_{t-1})$$

Pemulusan Tren:

$$b_t = \beta(L_t - L_{t-1}) + (1-\beta)b_{t-1}$$

Pemulusan Musiman:

$$S_t = \gamma \frac{X_t}{L_t} + (1-\gamma)S_{t-s}$$

$$L_s = \frac{1}{s}(X_1 + X_2 + \dots + X_s)$$

$$S_t = \frac{X_t}{L_s}, \text{ dimana } t = 1, 2, 3, \dots, s$$

$$b_s = \frac{1}{s} \left[\frac{X_{s+1} - X_1}{s} + \frac{X_{s+2} - X_2}{s} + \dots + \frac{X_{s+s} - X_s}{s} \right]$$

Keterangan:

- α, β, γ : koefisien/parameter pemulusan, nilai $0 < (\alpha, \beta, \gamma) < 1$
- X_t : data aktual waktu ke-t
- F_t : nilai peramalan ke-t
- m : jumlah periode yang akan diramalkan
- s : panjang musiman

Auto Regressive Integrated Moving Average (ARIMA)

Metode ARIMA mensyaratkan bahwa data yang akan diramalkan adalah stasioner. Kestasioneran data dapat dilihat melalui *correlogram* maupun dengan uji formal *Augmented Dickey-Fuller* (ADF). Persamaannya dinyatakan dalam bentuk umum sebagai berikut.

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \alpha_1 \Delta Y_{t-1} + \alpha_2 \Delta Y_{t-2} + \dots + \alpha_m \Delta Y_{t-m} + \varepsilon_t$$

Di mana:

- m = panjangnya *lag* yang digunakan
- β_1 = komponen intersep
- $\beta_2 t$ = komponen tren.

Apabila data tidak stasioner di *level* (data asli) maka perlu dilakukan penurunan (*differencing*). Data yang tidak stasioner akibat adanya tren umumnya dihilangkan dengan melakukan *differencing* pertama dengan formula $Y'_t = Y_t - Y_{t-1}$. Data yang tidak stasioner akibat adanya efek musiman dapat dihilangkan dengan *differencing* musiman dengan formula $Y'_t = Y_t - Y_{t-4}$ (untuk data kuartalan). Dalam penelitian ini dilakukan *differencing* pertama untuk variabel ekspor, dan *differencing* musiman untuk variabel ekspor.

Setelah data memenuhi syarat stasioner, selanjutnya dilakukan identifikasi ordo *autoregressive* dan *moving average* untuk proses estimasi nantinya. Identifikasi ordo dilakukan dengan melihat *correlogram* dengan memerhatikan ACF dan PACF. *Lag* yang signifikan pada plot ACF akan dinyatakan sebagai ordo dari *moving average*, sedangkan *lag* yang signifikan pada plot PACF akan dinyatakan sebagai ordo dari *autoregressive*. Namun demikian, konsep *trial and error* dapat dilakukan untuk mencapai model terbaik dengan tetap memerhatikan *lag* yang signifikan.

Selanjutnya dilakukan estimasi parameter. Penentuan model yang baik pada metode ini berdasarkan kriteria AIC terkecil. Kemudian hasil tersebut didiagnosis, apakah *white noise* atau tidak. Persamaan sederhana dari *white noise*

sederhana adalah sebagai berikut, dimana Y_t dibangun oleh *overall level* (c) dan *random error component* (e_t) yang tidak berkorelasi antar waktu.

$$Y_t = c + e_t$$

Pemeriksaan *white noise* dapat dilakukan dengan melihat *correlogram* melalui ACF dan PACF dari residualnya (Makridakis [5]). Apabila tidak ada *spike* pada *correlogram* maka ini menunjukkan residual series tersebut *white noise*. Pemeriksaan menyeluruh terhadap kelayakan model menggunakan uji *chi-square* (c^2) yang berbasis pada statistik Ljung-Box Q dengan Q menyebar mengikuti distribusi $\chi^2_{(m-r)}$,

$$Q_m = n'(n' + 2) \sum_{k=1}^m \frac{r_k^2(e)}{n' - k}$$

Apabila *white noise* terpenuhi maka model dapat digunakan. Namun apabila tidak maka harus kembali memilih model lain dengan beberapa kriteria. Makridakis (1997) mengatakan bahwa kriteria untuk memilih model ARIMA terbaik adalah yang memberikan *sum of squared errors* terkecil atau nilai *likelihood* terbesar. Namun demikian ukuran yang umum digunakan adalah *Akaike's Information Criterion* (AIC) yang minimum.

$$AIC \approx n(1 + \log(2\pi)) + n \log \sigma^2 + 2m$$

Exponential Smoothing

Untuk melakukan peramalan dengan metode ini maka terlebih dahulu diidentifikasi pola yang ada di dalam deret yang dilihat dari plotnya. Identifikasi ini berfungsi untuk menentukan jenis *exponential smoothing* yang akan digunakan. Apabila plot data tidak memperlihatkan adanya tren sehingga tidak dapat diprediksikan akan meningkat atau menurun dan juga tidak memperlihatkan adanya pola musiman, maka metode yang lebih cocok digunakan untuk data tersebut adalah *single exponential smoothing*. Kemudian apabila plot data memperlihatkan adanya pola tren maka metode yang lebih cocok digunakan adalah *double exponential smoothing*. Sedangkan apabila plot data menunjukkan adanya pola tren dan pola musiman, maka metode yang lebih cocok digunakan adalah *triple exponential smoothing*.

Untuk membuktikan adanya pola tren dalam data maka dilakukan regresi deret dengan konstanta tren. Apabila *p-value* variabel tren kurang dari α (signifikan), maka dapat disimpulkan terdapat tren dalam data. Hasil uji tren untuk variabel produksi dan ekspor kelapa sawit memberikan hasil *p-value* yang signifikan sehingga dapat disimpulkan bahwa data produksi dan ekspor kelapa sawit mengandung tren.

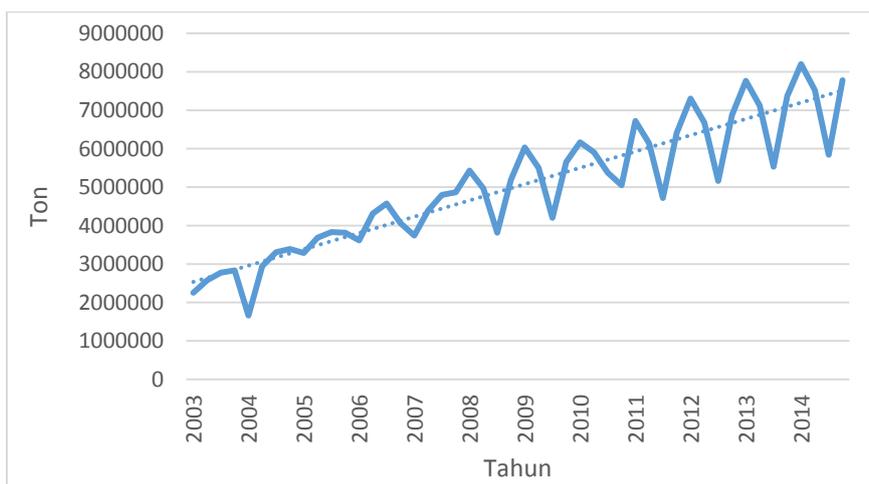
Sedangkan data produksi kelapa sawit terlihat memiliki pola musiman dengan periode triwulanan, indikasi ini dapat dibuktikan dengan melihat grafik data tersebut. Atau dapat juga melalui plot ACF yang menunjukkan *lag* yang berulang pada periode tertentu. Maka diperoleh identifikasi bahwa data produksi kelapa sawit memiliki pola tren dan juga musiman.

Dengan demikian, metode yang digunakan adalah *Double Exponential Smoothing* untuk variabel ekspor kelapa sawit dan *Triple Exponential Smoothing* untuk variabel produksi kelapa sawit.

Produksi

Berkembangnya subsektor perkebunan kelapa sawit di Indonesia tidak lepas dari adanya kebijakan pemerintah yang memberikan berbagai insentif, terutama kemudahan dalam hal perizinan dan bantuan subsidi investasi untuk pembangunan perkebunan rakyat dengan pola PIR-bun dan dalam pembukaan wilayah baru untuk areal perkebunan besar swasta. Kebijakan ini akhirnya menumbuhkan banyak perusahaan yang bergerak dengan komoditas kelapa sawit.

Peramalan untuk variabel produksi pertama dilakukan dengan metode ARIMA (Box-Jenkins), diawali dengan mengidentifikasi pola data. Dalam hal ini variabel produksi memiliki pola musiman dan tren seperti terlihat pada Gambar 1. Data tersebut harus *difference* sehingga memperoleh data yang lebih stasioner.



Gambar 1. Plot Data Produksi CPO pada *Level*

Setelah dilakukan identifikasi plot ACF dan PACF, dapat dikombinasi beberapa model *Seasonal* ARIMA yang memungkinkan diantaranya ARIMA (2,0,4)(0,1,0)₄, ARIMA (2,0,0)(0,1,1)₄, ARIMA (0,0,1)(0,1,1)₄, ARIMA (0,0,2)(0,1,0)₄, ARIMA (0,0,2)(0,1,1)₄, ARIMA (1,0,1)(0,1,1)₄. Model tersebut adalah hasil estimasi parameter yang secara signifikan untuk dapat digunakan melakukan peramalan.

Sebelum digunakan untuk peramalan, terlebih dahulu model diperiksa kelayakannya. Kelayakan model dapat diketahui dengan melihat residualnya apakah bersifat *white noise* atau tidak. Dari hasil pengujian residual diketahui bahwa untuk model ARIMA (2,0,4)(0,1,0)₄, ARIMA (2,0,0)(0,1,1)₄, ARIMA (0,0,2)(0,1,0)₄, ARIMA (0,0,2)(0,1,1)₄ tidak ada *lag* yang *spike*. Kemudian tidak ditemukan pula *lag* yang signifikan terhadap $\alpha=0,05$ sehingga keputusan yang diambil adalah tidak tolak H_0 . Dengan demikian model tersebut memenuhi kriteria *white noise*.

Kemudian dari keempat model tersebut akan dipilih salah satu yang memberikan *error* peramalan paling kecil. Diketahui model ARIMA (2,0,4)(0,1,0)₄ memberikan *error* yang paling kecil dengan nilai RMSE sebesar 223.385,4. Dengan demikian, model *seasonal* ARIMA yang terpilih adalah ARIMA (2,0,4)(0,1,0)₄. Persamaan dari model tersebut dapat dituliskan sebagai berikut.

$$\hat{Y}_t = 965.835,1 + Y_{t-4} - 0,44278 (Y_{t-1} - Y_{t-5}) - 0,635762 (Y_{t-2} - Y_{t-6}) + \\ 0,601988 e_{t-1} + 0,316644 e_{t-2} - 0,142928 e_{t-3} - 0,758097 e_{t-4}$$

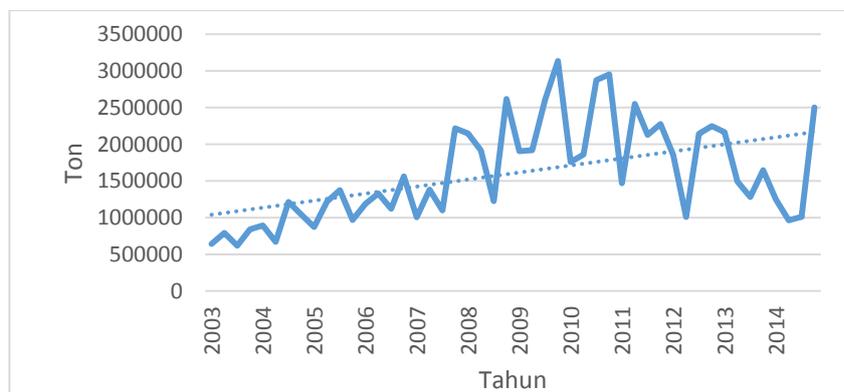
Selanjutnya pada metode *triple exponential smoothing* dilakukan *trial and error* untuk menentukan nilai parameter pemulusan *level* α , tren β , dan musiman γ . Parameter yang optimum diestimasi sehingga didapatkan nilai parameter $\alpha=0,0000$, $\beta=0,0000$, dan $\gamma=0,6001$ dengan RMSE sebesar 498.523,2. Berdasarkan model umum peramalan *triple exponential smoothing* pada persamaan (2). Persamaan dari model tersebut dapat ditulis sebagai berikut.

$$F_{t+m} = \left\{ (0) \frac{X_t}{S_{t-s}} + (1-0) (L_{t-1} + b_{t-1}) + \left[0 (L_t L_{t-1}) + (1-0) b_{t-1} \right] m \right\} S_{t-s+m} \\ S_t = 0,59 \frac{X_t}{L_t} + (1-0,6001) S_{t-s}$$

Dengan demikian diketahui bahwa model yang memberikan kesalahan minimum adalah ARIMA (2,0,4)(0,1,0)₄. Oleh karena itu peramalan volume produksi minyak kelapa sawit akan dilakukan dengan model ARIMA (2,0,4)(0,1,0)₄.

Ekspor

Mulai tahun 2008 hingga 2014 terlihat bahwa kefluktuatifan ekspor kelapa sawit Indonesia semakin ekstrem dengan perubahan-perubahan yang tajam. Kinerja dan daya saing minyak kelapa sawit Indonesia agak lemah di perdagangan dunia terutama periode 2009 hingga 2012⁶. Hal ini disebabkan karena krisis ekonomi global pada tahun 2008 menyebabkan permintaan minyak kelapa sawit beberapa negara pengimpor menurun akibat terjadinya perlambatan pertumbuhan ekonomi di beberapa negara, sehingga kestabilan permintaan dari luar negeri maupun penyediaan dalam negeri menjadi tidak menentu.



Gambar 2. Plot data ekspor pada level

Berdasarkan gambar di atas terlihat bahwa data ekspor belum stasioner pada *level*. Demikian pula apabila dilihat dari *correlogram*-nya, plot ACF masih memperlihatkan pola eksponensial yang berarti data belum stasioner. Dengan demikian harus dilakukan *differencing* untuk data tersebut. Uji akar unit ADF dengan tren dan intersep data ekspor pada perbedaan pertama menunjukkan

bahwa data ekspor telah stasioner pada pembedaan pertama (*1st difference*) sehingga telah dapat digunakan untuk melakukan peramalan.

Setelah dilakukan identifikasi plot ACF dan PACF, dapat dikombinasi beberapa model ARIMA yang memungkinkan diantaranya ARIMA (1,1,0), ARIMA (2,1,0), ARIMA (3,1,0), ARIMA (0,1,1), ARIMA (0,1,2). Model tersebut adalah hasil estimasi parameter yang secara signifikan untuk dapat digunakan melakukan peramalan. Kemudian sebelum digunakan untuk peramalan, terlebih dahulu model diperiksa kelayakannya. Kelayakan model dapat diketahui dengan melihat residualnya apakah bersifat *white noise* atau tidak. Dari hasil pengujian residual diketahui bahwa model ARIMA (3,1,0) memenuhi kriteria *white noise*, sedangkan empat model lainnya tidak.

Peramalan dilakukan dengan menggunakan ARIMA (3,1,0) dan didapatkan peramalan dengan RMSE sebesar 715,169,9. Persamaan dari model ARIMA (3,1,0) dapat dituliskan sebagai berikut.

$$\hat{Y}_t = 105.552,4 + 0,121793 Y_{t-1} + 0,054319 Y_{t-2} + 0,289953 Y_{t-3} - 0,533935 Y_{t-4}$$

Selanjutnya pada metode *double exponential smoothing* dilakukan *trial and error* untuk menentukan nilai parameter pemulusan level α dan tren β yang optimum. Parameter yang optimum diestimasi pada *training data*, kemudian digunakan untuk melakukan peramalan untuk periode *test data*. Parameter tersebut dipilih hingga mendapatkan model yang memberikan *error* terkecil. Dalam hal ini *software* Eviews 8 dapat melakukan iterasi secara otomatis untuk mendapat parameter yang optimum.

Dari hasil iterasi Eviews 8 didapatkan nilai parameter $\alpha=0,2700$ dan $\beta=0,3701$ dengan RMSE sebesar 482.340,3. Berdasarkan persamaan umum model umum peramalan *double exponential smoothing* pada persamaan (1). Persamaan dari model tersebut dapat ditulis sebagai berikut.

$$F_{t+m} = L_t + b_t m$$

$$L_t = 0,2700 X_t + 0,7300 (L_{t-1} + b_{t-1})$$

$$b_t = 0,3701 (L_t - L_{t-1}) + 0,6299 b_{t-1}$$

Model terbaik untuk melakukan peramalan adalah model yang menghasilkan kesalahan yang minimum. Berdasarkan pemaparan sebelumnya diketahui nilai RMSE dari model ARIMA (3,1,0) adalah sebesar 715,169,9. Sedangkan nilai RMSE dari model *double exponential smoothing* $\alpha=0,2700$ dan $\beta=0,3701$ adalah sebesar 482.340,3. Dengan demikian diketahui bahwa model yang memberikan kesalahan minimum adalah model *double exponential smoothing* $\alpha=0,2700$ dan $\beta=0,3701$ untuk variabel ekspor.

Berdasarkan hasil pembahasan diketahui bahwa untuk melakukan peramalan variabel produksi, model yang lebih baik adalah ARIMA (2,0,4)(0,1,0)₄ karena memberikan RMSE yang lebih kecil dibandingkan dengan model *triple exponential smoothing* $\alpha=0,000$, $\beta=0,000$, $\gamma=0,6001$. Kemudian untuk melakukan peramalan variabel ekspor, model yang lebih baik adalah model *double exponential smoothing* $\alpha=0,2700$ dan $\beta=0,3701$ karena memberikan RMSE yang lebih kecil dibandingkan model ARIMA (3,1,0). Tidak dapat ditentukan manakah metode peramalan yang terbaik, karena hasilnya akan bergantung pada pola data. Dengan

demikian metode yang harus digunakan adalah metode yang memberikan nilai kesalahan terkecil, dengan tetap memperhatikan karakteristik data yang ada. Berikut adalah data hasil peramalan produksi dan ekspor CPO dengan metode terbaik.

Tabel 1. Hasil Peramalan Produksi dan Ekspor Minyak Kelapa Sawit Indonesia

No.	Periode	Produksi (ton)	Ekspor (ton)
(1)	(2)	(3)	(4)
1.	Triwulan I 2015	8.724.695	-
2.	Triwulan II 2015	7.948.988	-
3.	Triwulan III 2015	6.744.915	-
4.	Triwulan IV 2015	8.159.922	-
5.	Triwulan I 2016	9.164.050	-
6.	Triwulan II 2016	8.423.436	1.844.592
7.	Triwulan III 2016	7.209.063	1.795.695

Produksi minyak kelapa sawit mengalami peningkatan yang cukup besar di mana pada triwulan I 2003 hanya 2.252.790 ton sedangkan triwulan III 2016 diperkirakan naik menjadi 7.209.063 ton. Hal ini harus menjadi perhatian pemerintah untuk mengontrol peningkatan produksi kelapa sawit tidak hanya berfokus pada peningkatan luas areal perkebunan kelapa sawit dengan pemberian izin pembukaan wilayah baru tetapi juga harus berupaya meningkatkan produksi melalui peningkatan produktivitas agar tidak banyak mengakibatkan kerusakan lingkungan. Kontribusi ekspor kelapa sawit terhadap produksi yang masih besar dan cenderung meningkat yaitu 22% pada Triwulan II 2016 dan 25% pada Triwulan III 2016 juga harus menjadi perhatian pemerintah dalam hal mengatur besaran pajak ekspor minyak kelapa sawit untuk melindungi konsumen domestik.

3. Kesimpulan

Kesimpulan yang dapat diambil antara lain:

1. Perkembangan produksi minyak kelapa sawit (CPO) Indonesia dari tahun 2003 hingga 2014 secara triwulan menunjukkan kenaikan. Selain karena cerahnya prospek komoditas ini, kenaikan ini terjadi juga karena adanya dukungan pemerintah yang memberikan insentif, terutama kemudahan dalam hal perizinan pendirian perusahaan perkebunan. Oleh karena itu, dapat dilihat juga perkembangan perusahaan perkebunan kelapa sawit mengalami kenaikan dari tahun 2003 hingga 2014. Sama halnya dengan produksi, ekspor minyak kelapa sawit (CPO) Indonesia juga menunjukkan tren yang naik dari tahun ke tahun, meskipun pasca terjadi krisis tahun 2008 perkembangannya menjadi kurang stabil.
2. Untuk melakukan peramalan variabel produksi, model yang lebih baik adalah ARIMA (2,0,4)(0,1,0)₄ karena memberikan RMSE yang lebih kecil dibandingkan dengan model *triple exponential smoothing* $\alpha=0,000$, $\beta=0,000$, $\gamma=0,6001$. Kemudian untuk melakukan peramalan variabel ekspor, model yang lebih baik adalah model *double exponential smoothing* $\alpha=0,2700$ dan $\beta=0,3701$ karena memberikan RMSE yang lebih kecil dibandingkan model ARIMA (3,1,0).

3. Dari hasil peramalan diketahui bahwa produksi minyak kelapa sawit (CPO) Indonesia masih menunjukkan kenaikan. Sedangkan ekspor minyak kelapa sawit Indonesia juga menunjukkan kenaikan dengan rata-rata laju pertumbuhan per triwulan yang lebih besar dibandingkan rata-rata laju pertumbuhan produksi.

Referensi

- [1] Kemenperin. (2007). *Gambaran Sekilas Industri Minyak Kelapa Sawit*. Jakarta: Sekretariat Jenderal. Retrieved from <http://www.kemenperin.go.id/download/289/Paket-Informasi-Komoditi-Minyak-Kelapa-Sawit>.
- [2] Hagi, Syaiful Hadi & Ermy Tety. (November 2012). Analisis Daya Saing Ekspor Minyak Sawit Indonesia dan Malaysia di Pasar Internasional. *Pekbis Jurnal*, Vol.4, no.3, 180-191.
- [3] Rambe, M. K. (2009). *Peramalan Hasil PRoduksi Minyak Kelapa Sawit pada PT. Perkebunan Nusantara III (Persero) Sumatera Utara*. Medan: USU.
- [4] Montgomery, D. C, J. S. Gardiner & L. S. Johnson. (1990). *Forecasting and Time Series Analysis* (2nd ed.).
- [5] Makridakis, S., Steven C. Wheelwright, & Rob J. Hyndman. (1997). *Forecasting Methods and Applications 3rd Ed*. Wiley.
- [6] Ermawati, T., & Yuni Saptia. (Desember 2013). Kinerja Ekspor Minyak Kelapa Sawit Indonesia. *Buletin Ilmiah Litbang Perdagangan*, Vol.7 No.2.

PENGARUH TINGKAT KEMANDIRIAN FISKAL, INVESTASI, DAN TENAGA KERJA TERHADAP PERTUMBUHAN EKONOMI KABUPATEN/KOTA DI JAWA TENGAH TAHUN 2010-2014

ANDHIKA RAHMADANI¹, DEWI PURWANTI²

1 Badan Pusat Statistik, 12.7015@stis.ac.id

2 Sekolah Tinggi Ilmu Statistik, dewip@stis.ac.id

Abstrak. Pertumbuhan ekonomi Provinsi Jawa Tengah cenderung tumbuh secara lambat dan lebih rendah jika dibandingkan dengan pertumbuhan ekonomi provinsi lain di Pulau Jawa dan pertumbuhan ekonomi nasional. Penelitian ini bertujuan untuk mengetahui pengaruh variabel tingkat kemandirian fiskal, investasi, dan tenaga kerja terhadap pertumbuhan ekonomi kabupaten/kota di Jawa Tengah tahun 2010-2014. Variabel yang digunakan dalam penelitian ini adalah pertumbuhan ekonomi, derajat kemandirian fiskal, rasio belanja modal, Penanaman Modal Asing (PMA), Penanaman Modal Dalam Negeri (PMDN), dan tenaga kerja. Metode analisis yang digunakan dalam penelitian ini adalah analisis deskriptif dan analisis regresi data panel. Hasil penelitian ini menunjukkan bahwa variabel derajat kemandirian fiskal, rasio belanja modal, PMA, PMDN, dan tenaga kerja memiliki hubungan positif dan berpengaruh signifikan terhadap pertumbuhan ekonomi kabupaten/kota di Jawa Tengah tahun 2010-2014. Pemerintah daerah harus berupaya menggali Pendapatan Asli Daerah (PAD) secara optimal, meningkatkan kualitas belanja daerah, menjaga iklim investasi yang kondusif serta mendorong penciptaan lapangan kerja baru guna mendorong pertumbuhan ekonomi daerah secara optimal.

Kata kunci: pertumbuhan ekonomi, tingkat kemandirian fiskal, investasi, tenaga kerja, regresi data panel

1. Pendahuluan

Kondisi makro ekonomi di Provinsi Jawa Tengah dapat dilihat dari indikator pertumbuhan ekonomi daerah. Rata-rata pertumbuhan ekonomi Provinsi Jawa Tengah tahun 2011-2014 adalah yang terendah jika dibandingkan dengan rata-rata pertumbuhan ekonomi provinsi lain di Pulau Jawa. Rata-rata pertumbuhan ekonomi Provinsi Jawa Tengah tahun 2011-2014 adalah sebesar 5,30 persen, tumbuh di bawah rata-rata pertumbuhan ekonomi nasional yang mencapai sebesar 5,80 persen pada kurun waktu yang sama. Hal tersebut menunjukkan adanya masalah kelesuan pertumbuhan ekonomi di Provinsi Jawa Tengah pada kurun waktu tersebut.

Tabel 1. Laju pertumbuhan PDRB riil (harga konstan 2010) menurut provinsi di Pulau Jawa tahun 2011-2014 (persen).

Provinsi	Tahun				Rata-Rata
	2011	2012	2013	2014	
(1)	(2)	(3)	(4)	(5)	(6)
DKI Jakarta	6,73	6,53	6,11	5,95	6,32
Jawa Barat	6,50	6,50	6,34	5,06	6,07
Jawa Tengah	5,30	5,34	5,14	5,42	5,30
DI Yogyakarta	5,21	5,37	5,49	5,18	5,31
Jawa Timur	6,44	6,64	6,08	5,86	6,25
Banten	7,03	6,83	7,13	5,47	6,58
Nasional	6,16	6,16	5,74	5,21	5,80

Jika dilihat dari segi Sumber Daya Manusia (SDM), Sumber Daya Alam (SDA), dan infrastruktur seharusnya Provinsi Jawa Tengah tidak kalah dengan provinsi lain di Pulau Jawa. Mengacu pada Rencana Pembangunan Jangka Menengah Daerah (RPJMD) tahun 2008-2013 dan 2013-2018, pertumbuhan ekonomi Provinsi Jawa Tengah masih jauh dibawah target yang telah ditetapkan, yaitu pertumbuhan ekonomi berkisar 5,9 hingga 6,66 persen. [6] Hal tersebut mengindikasikan bahwa pertumbuhan ekonomi Provinsi Jawa Tengah masih belum optimal atau lebih rendah dari potensi pertumbuhan ekonomi yang seharusnya dapat dicapai.

Pertumbuhan ekonomi Provinsi Jawa Tengah merupakan hasil kontribusi dari 35 kabupaten/kota yang ada di Jawa Tengah. Pertumbuhan ekonomi kabupaten/kota di Jawa Tengah periode 2011-2014 cukup beragam. Kabupaten Semarang menjadi wilayah dengan rata-rata pertumbuhan ekonomi tertinggi di Provinsi Jawa Tengah, yaitu sebesar 6,28 persen. Sedangkan, kabupaten Cilacap menjadi wilayah dengan rata-rata pertumbuhan ekonomi terendah di Provinsi Jawa Tengah, yaitu sebesar 2,66 persen. Selain itu, terdapat 11 kabupaten/kota dengan rata-rata pertumbuhan ekonomi dibawah rata-rata pertumbuhan ekonomi Provinsi Jawa Tengah dan terdapat 29 kabupaten/kota dengan rata-rata pertumbuhan ekonomi dibawah rata-rata pertumbuhan ekonomi nasional. [6] Hal tersebut mengindikasikan bahwa kinerja perekonomian kabupaten/kota di Provinsi Jawa Tengah juga masih belum optimal dan belum merata.

Selain dilihat dari laju pertumbuhan ekonomi, kondisi makro ekonomi Provinsi Jawa Tengah juga bisa dilihat dari PDRB per kapita. PDRB per kapita Provinsi Jawa Tengah hanya berada di peringkat kelima dari enam provinsi yang ada di Pulau Jawa (BPS [1]).

Berdasarkan latar belakang yang telah diuraikan, dapat disimpulkan bahwa terdapat empat pokok permasalahan terkait pertumbuhan ekonomi di Provinsi Jawa Tengah tahun 2010-2014, yaitu: 1) pertumbuhan ekonomi Provinsi Jawa Tengah cenderung tumbuh secara lambat. 2) pertumbuhan ekonomi Provinsi Jawa Tengah masih lebih rendah jika dibandingkan dengan pertumbuhan ekonomi provinsi lain di Pulau Jawa dan pertumbuhan ekonomi nasional. 3) pertumbuhan ekonomi kabupaten/kota di Jawa Tengah masih cukup beragam, dimana terdapat kabupaten/kota dengan pertumbuhan ekonomi tinggi, tetapi masih banyak kabupaten/kota dengan pertumbuhan ekonomi yang masih rendah. 4) diperlukan

pertumbuhan ekonomi yang lebih tinggi lagi untuk mampu mendorong kinerja perekonomian daerah di Provinsi Jawa Tengah secara optimal.

Sejak era otonomi daerah, desentralisasi fiskal memegang peranan penting terhadap pertumbuhan ekonomi daerah. Oates [9] mengatakan bahwa desentralisasi fiskal berpotensi memberikan peran dan kontribusi dalam pembangunan ekonomi. Sejalan dengan hal tersebut, Mungkasa [8] mengatakan bahwa terlepas dari perdebatan tentang dampak positif dan negatif dari otonomi daerah, setidaknya berbagai kalangan mempercayai terdapat banyak hal positif dari penerapan konsep otonomi daerah di Indonesia, diantaranya semakin meningkatnya tingkat kemandirian dan kemampuan daerah dalam mengelola pembangunan ekonomi daerahnya dan berdampak pada perkembangan perekonomian daerah yang signifikan, ditandai dengan peningkatan Produk Domestik Regional Bruto (PDRB), pendapatan per kapita, dan pertumbuhan ekonomi. Selain itu, pertumbuhan ekonomi juga tidak dapat dilepaskan dari peranan investasi dan tenaga kerja. Todaro [10] mengatakan bahwa pertumbuhan ekonomi merupakan fungsi dari investasi, hal ini dikarenakan tingkat pertumbuhan ekonomi dan investasi merupakan hal yang tidak dapat dipisahkan dan saling membutuhkan. Semakin besar investasi maka semakin besar tingkat pertumbuhan yang dapat dicapai. Sedangkan, pertumbuhan angkatan kerja yang terjadi seiring pertumbuhan penduduk adalah salah satu faktor positif yang dapat memacu pertumbuhan ekonomi. Jumlah tenaga kerja yang lebih besar akan menambah jumlah tenaga kerja produktif.

Adapun tujuan penelitian ini adalah sebagai berikut:

1. Untuk mengetahui gambaran umum pertumbuhan ekonomi, tingkat kemandirian fiskal, investasi, dan tenaga kerja kabupaten/kota di Jawa Tengah tahun 2010-2014.
2. Untuk mengetahui pengaruh variabel tingkat kemandirian fiskal, investasi, dan tenaga kerja terhadap pertumbuhan ekonomi kabupaten/kota di Jawa Tengah tahun 2010-2014.

2. Hasil – Hasil Utama

Beberapa definisi dalam penelitian ini dijelaskan sebagai berikut:

1. Pertumbuhan ekonomi adalah pertumbuhan PDRB riil dari suatu periode ke periode berikutnya.
2. PDRB riil (harga konstan 2010) adalah jumlah nilai tambah yang dihasilkan oleh seluruh unit usaha dalam suatu wilayah tertentu atau merupakan nilai barang dan jasa akhir yang dihasilkan oleh seluruh unit ekonomi dan digunakan untuk menggambarkan perkembangan riil barang dan jasa yang dihasilkan oleh seluruh kegiatan ekonomi di daerah tersebut.
3. Derajat kemandirian fiskal adalah rasio Pendapatan Asli Daerah (PAD) terhadap Total Penerimaan Daerah (TPD) dengan satuan persen. Derajat kemandirian fiskal adalah ukuran yang mencerminkan tingkat kemampuan pemerintah daerah dalam rangka menggali atau meningkatkan PAD guna membiayai pembangunan daerah.
4. Rasio belanja modal adalah rasio belanja modal terhadap total belanja daerah dengan satuan persen. Belanja modal adalah belanja pemerintah daerah yang digunakan untuk pengeluaran yang dilakukan dalam rangka pembelian atau

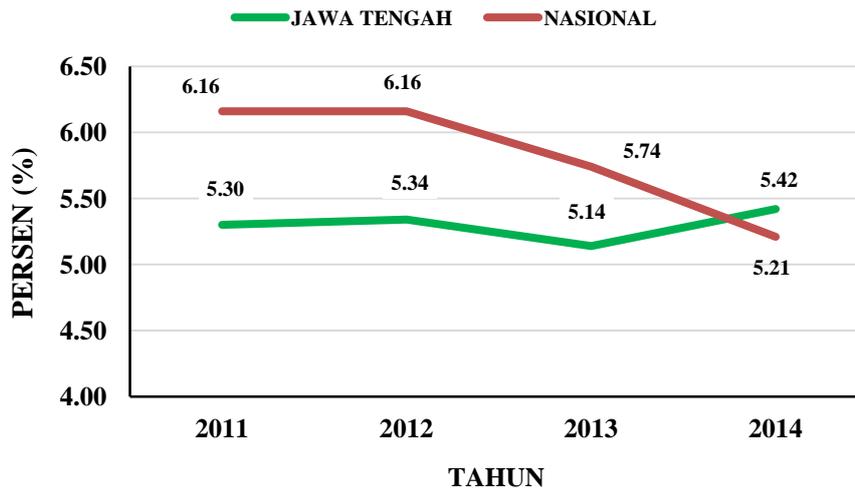
- pengadaan atau pembangunan aset tetap berwujud yang mempunyai nilai manfaat lebih dari dua belas bulan untuk digunakan dalam kegiatan pemerintah, seperti dalam bentuk tanah, peralatan dan mesin, gedung dan bangunan, jalan, irigasi dan jaringan, dan aset tetap lainnya.
5. Penanaman Modal Asing (PMA) adalah kegiatan menanamkan modal untuk melakukan usaha di wilayah negara Republik Indonesia, yang dilakukan oleh penanam modal asing, baik yang menggunakan modal asing sepenuhnya maupun yang berpatungan dengan penanam modal dalam negeri.
 6. Penanaman Modal Dalam Negeri (PMDN) adalah kegiatan menanamkan modal untuk melakukan usaha di wilayah negara Republik Indonesia yang dilakukan oleh penanam modal dalam negeri dengan menggunakan modal dalam negeri.
 7. Tenaga Kerja adalah seluruh penduduk usia kerja (15 tahun dan lebih) yang bekerja.

Analisis Deskriptif

Pertumbuhan Ekonomi

Pada tahun 2013, pertumbuhan ekonomi Jawa Tengah mengalami perlambatan menjadi 5,14 persen dari tahun sebelumnya sebesar 5,34 persen. Hal tersebut menunjukkan bahwa kinerja perekonomian Provinsi Jawa Tengah pada kurun waktu tersebut masih belum optimal. Bank Indonesia [4] mengatakan bahwa perlambatan tersebut dikarenakan pertumbuhan ekspor yang mengalami perlambatan. Sedangkan dari sisi sektoral, melemahnya sektor perdagangan, hotel dan restoran, serta pertanian turut mendorong melambatnya pertumbuhan ekonomi Provinsi Jawa Tengah pada tahun 2013.

Namun demikian, pertumbuhan ekonomi Provinsi Jawa Tengah pada tahun 2014 kembali membaik dan bahkan mencatat pertumbuhan ekonomi tertinggi sejak tahun 2011. Hal tersebut menjadi anomali, dimana ketika pertumbuhan ekonomi nasional melambat, pertumbuhan ekonomi Provinsi Jawa Tengah justru mengalami peningkatan. Bank Indonesia [5] mengatakan bahwa pada saat kondisi perekonomian global masih belum membaik, permintaan domestik justru menjadi penggerak perekonomian daerah. Kondisi tersebut ditunjukkan oleh ekspor luar negeri yang melambat, sedangkan ekspor antar daerah justru mengalami peningkatan. Sedangkan dari sisi sektoral, pertumbuhan sektor industri pengolahan, perdagangan besar-eceran, dan reparasi mobil-sepeda motor, serta sektor konstruksi turut mendorong pertumbuhan ekonomi Provinsi Jawa Tengah pada tahun 2014. Selain itu, adanya hajatan Pemilu Legislatif dan Presiden pada tahun 2014 turut mendorong peningkatan konsumsi Lembaga Non Profit Rumah Tangga (LNPR) yang berdampak pada peningkatan pertumbuhan ekonomi.



Gambar 1. Laju pertumbuhan PDRB riil (harga konstan 2010) Provinsi Jawa Tengah dan rata-rata nasional tahun 2011-2014 (persen).

Tingkat Kemandirian Fiskal

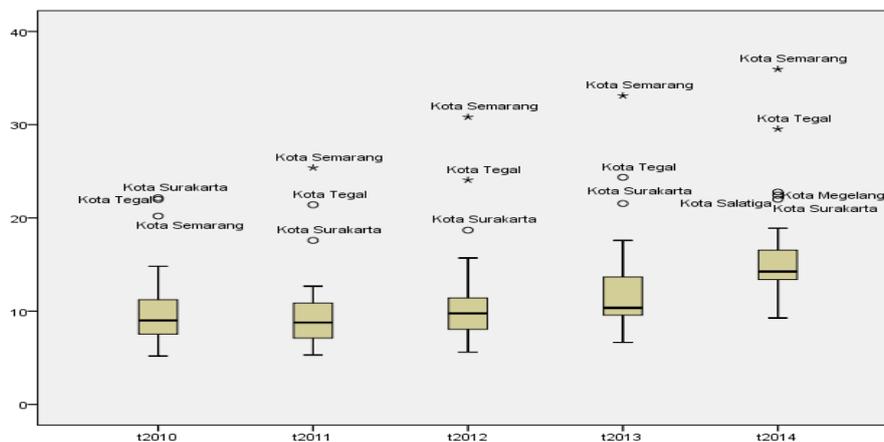
Tabel 2 menunjukkan statistik deskriptif rata-rata kemandirian fiskal kabupaten/kota di Jawa Tengah mengalami peningkatan yaitu dari 10,03 persen pada tahun 2010 menjadi 15,82 persen pada tahun 2014, walaupun sempat mengalami penurunan pada tahun 2011 dari tahun sebelumnya menjadi 9,82 persen. Peningkatan derajat kemandirian fiskal tersebut tidak terlepas dari peningkatan PAD kabupaten/kota di Jawa Tengah secara keseluruhan. Peningkatan PAD tersebut didorong oleh berbagai kebijakan pemerintah daerah dalam menggali sumber-sumber PAD. Hal tersebut dapat dilihat dari peningkatan pajak daerah, hasil pengelolaan kekayaan daerah yang dipisahkan (BUMD), dan lain-lain PAD yang sah setiap tahunnya. Namun demikian, tingkat kemandirian fiskal kabupaten/kota di Jawa Tengah masih tergolong rendah, yang ditunjukkan oleh rata-rata derajat kemandirian fiskal kurang dari 20 persen. Hal tersebut menunjukkan bahwa sebagian besar kabupaten/kota di Jawa Tengah masih belum optimal dalam menggali potensi sumber-sumber PAD daerahnya masing-masing, salah satunya adalah retribusi daerah. Hal tersebut dibuktikan dari sumbangan retribusi daerah yang hanya tumbuh sebesar 3,47 persen pada kurun waktu 2010-2014.

Tabel 2. Statistik deskriptif derajat kemandirian fiskal kabupaten/kota di Jawa Tengah tahun 2011-2014 (persen).

Statistik	2010	2011	2012	2013	2014
(1)	(2)	(3)	(4)	(5)	(6)
Mean	10,03	9,82	11,03	12,35	15,82
Median	9,01	8,78	9,76	10,36	14,25
Minimum	5,18	5,30	5,61	6,65	9,27
Maximum	22,16	25,39	30,82	33,11	35,96

Gambar 2 menunjukkan sebaran derajat kemandirian fiskal kabupaten/kota di Jawa Tengah tahun 2010-2014. Kota Semarang, Kota Tegal, Kota Surakarta, dan Kota Magelang merupakan wilayah dengan derajat kemandirian fiskal yang tinggi

jika dibandingkan dengan kabupaten/kota lainnya di Jawa Tengah. Sedangkan, sebagian besar kabupaten/kota lainnya memiliki derajat kemandirian fiskal kurang dari 20 persen. Artinya sebagian besar kabupaten/kota di Jawa Tengah memiliki derajat kemandirian fiskal yang tergolong kurang dan sangat kurang. Rendahnya derajat kemandirian fiskal sebagian besar kabupaten/kota di Jawa Tengah tersebut menunjukkan masih tingginya tingkat ketergantungan fiskal terhadap pemerintah pusat. Selain itu, hal tersebut juga menunjukkan adanya pola hubungan instruktif antara pemerintah pusat dan daerah dalam pelaksanaan otonomi daerah, dimana peranan pemerintah pusat lebih dominan daripada kemandirian pemerintah daerah.



Gambar 2. Derajat kemandirian fiskal kabupaten/kota di Jawa Tengah tahun 2010-2014 (persen).

Investasi

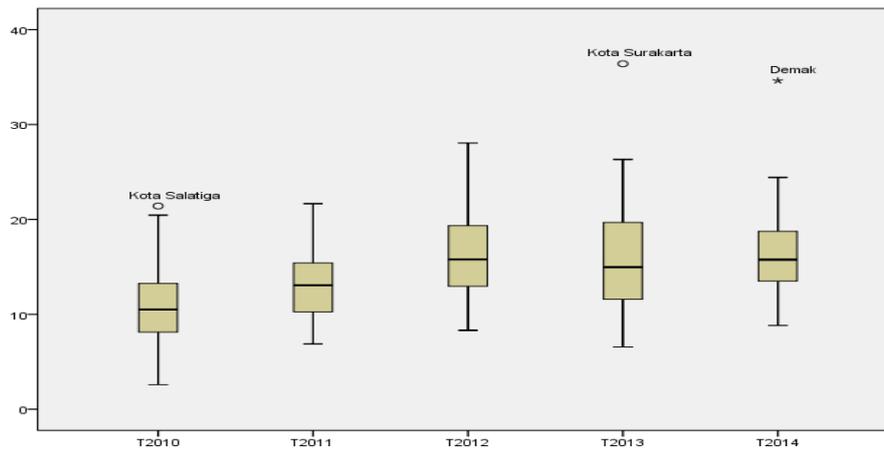
Tabel 3 adalah statistik deskriptif rasio belanja modal terhadap total belanja daerah kabupaten/kota di Jawa Tengah tahun 2010-2014. Rata-rata rasio belanja modal kabupaten/kota di Jawa Tengah mengalami peningkatan yaitu 10,67 persen pada tahun 2010 menjadi 16,75 persen pada tahun 2014. Peningkatan rasio belanja modal tersebut didorong oleh kebijakan pemerintah daerah yang menambah alokasi belanja modal guna percepatan pembangunan infrastruktur. Hal tersebut menunjukkan adanya pergeseran dari penurunan alokasi belanja pegawai kepada peningkatan alokasi belanja modal. Pergeseran tersebut merupakan indikasi positif terhadap perbaikan kualitas struktur belanja daerah kabupaten/kota di Jawa Tengah. Namun demikian, hal tersebut juga menunjukkan bahwa rasio belanja modal kabupaten/kota di Jawa Tengah masih tergolong rendah, yang ditunjukkan oleh rata-rata rasio belanja modal kurang dari 20 persen. Rendahnya rasio belanja modal kabupaten/kota di Jawa Tengah salah satunya disebabkan karena masih tingginya alokasi belanja daerah yang digunakan untuk belanja pegawai. Hal tersebut ditunjukkan oleh rata-rata rasio belanja pegawai terhadap total belanja daerah tahun 2014 yang mencapai 56,56 persen dari total belanja daerah.

Gambar 3 memperlihatkan rasio belanja modal terhadap total belanja daerah kabupaten/kota di Jawa Tengah tahun 2010-2014. Kota Salatiga, Kota Surakarta, dan Kabupaten Demak merupakan wilayah dengan rasio belanja modal yang tinggi jika dibandingkan dengan kabupaten/kota lainnya di Jawa Tengah. Sedangkan, sebagian besar kabupaten/kota lainnya memiliki rasio belanja modal kurang dari 20 persen. Artinya sebagian besar kabupaten/kota di Jawa Tengah memiliki rasio belanja

modal yang tergolong rendah. Hal tersebut menunjukkan bahwa kualitas struktur belanja daerah Kota Salatiga, Kota Surakarta, dan Kabupaten Demak sudah cukup baik. Sebaliknya, struktur belanja daerah kabupaten/kota lainnya di Jawa Tengah masih didominasi belanja pegawai.

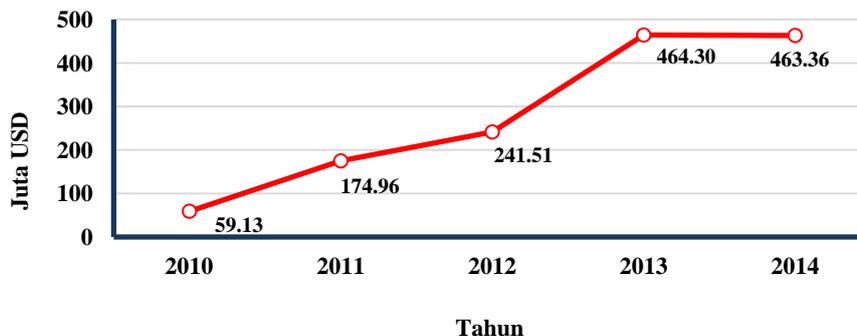
Tabel 3. Statistik deskriptif rasio belanja modal terhadap total belanja daerah kabupaten/kota di Jawa Tengah tahun 2010-2014 (persen).

Statistik	2010	2011	2012	2013	2014
(1)	(2)	(3)	(4)	(5)	(6)
Mean	10,68	13,02	16,20	15,99	16,76
Median	10,51	13,06	15,79	14,97	15,76
Minimum	2,59	6,87	8,32	6,57	8,83
Maximum	21,41	21,66	28,05	36,40	34,62



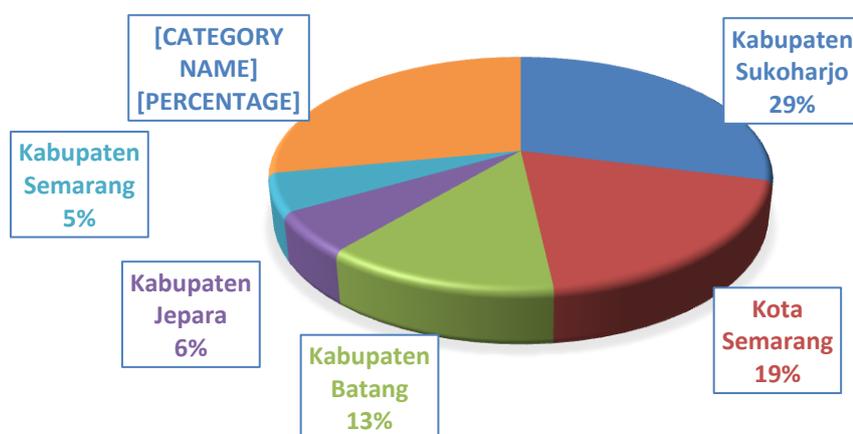
Gambar 3. Rasio belanja modal terhadap total belanja daerah kabupaten/kota di Jawa Tengah tahun 2010-2014 (persen).

Gambar 4 menunjukkan realisasi nilai investasi PMA total kabupaten/kota di Jawa Tengah mengalami peningkatan yang cukup pesat, yaitu 59,13 juta US dolar pada tahun 2010 menjadi 463,36 juta US dolar pada tahun 2014. Peningkatan realisasi nilai investasi PMA di Jawa Tengah dipengaruhi oleh beberapa faktor, antara lain ketersediaan lahan dan Upah Minimum Regional (UMR) yang relatif lebih rendah dibandingkan wilayah lain.



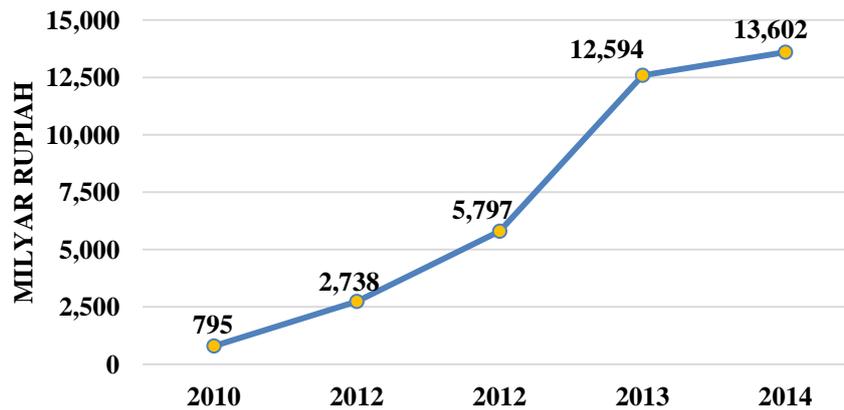
Gambar 4. Perkembangan realisasi nilai investasi Penanaman Modal Asing (PMA) total 35 kabupaten/kota di Jawa Tengah tahun 2010-2014 (juta US dolar).

Gambar 5 menunjukkan bahwa Kabupaten Sukoharjo, Kota Semarang, Kabupaten Batang, Kabupaten Jepara, dan Kabupaten Semarang menjadi lima besar wilayah dengan realisasi nilai investasi PMA terbesar pada kurun waktu 2010 hingga 2014, yaitu mencapai 72 persen dari total realisasi nilai investasi PMA di Jawa Tengah. Sedangkan, kabupaten/kota lainnya hanya menyumbang 28 persen dari total realisasi nilai investasi PMA di Jawa Tengah. Hal tersebut menunjukkan bahwa investasi PMA masih belum merata di kabupaten/kota di Jawa Tengah. Kabupaten Sukoharjo menjadi wilayah dengan realisasi nilai investasi PMA terbesar di Jawa Tengah dikarenakan geliat perekonomian yang berkembang pesat di kawasan bisnis Solo Baru. Semakin padatnya lokasi bisnis dan perdagangan di pusat Kota Surakarta mendorong investor mencari lokasi baru untuk pengembangan usahanya. Oleh karena itu, kawasan bisnis Solo Baru yang terletak di Kabupaten Sukoharjo dipilih para investor asing karena letaknya yang strategis, yaitu dekat dengan pusat Kota Surakarta dan memiliki potensi industri besar dan menengah yang cukup besar.



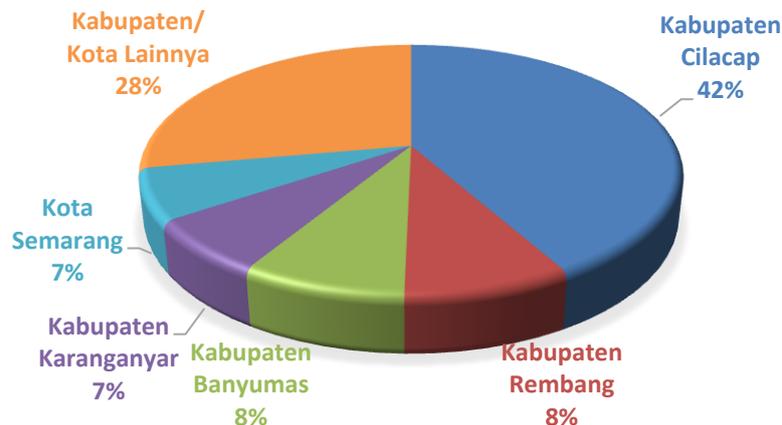
Gambar 5. Distribusi total realisasi nilai investasi Penanaman Modal Asing (PMA) berdasarkan lokasi kabupaten/kota di Jawa Tengah tahun 2010-2014 (persen).

Realisasi nilai investasi PMDN terus mengalami peningkatan pada kurun waktu 2010 hingga 2014. Realisasi nilai investasi PMDN mengalami peningkatan yang cukup pesat yaitu 795 milyar rupiah pada tahun 2010 menjadi 13,6 trilyun rupiah pada tahun 2014. Peningkatan realisasi nilai investasi PMDN di Jawa Tengah juga dipengaruhi oleh faktor ketersediaan lahan dan Upah Minimum Regional (UMR) yang relatif lebih rendah dibandingkan wilayah lain. Selain itu, Badan Penanaman Modal Daerah (BPMD) Jawa Tengah mengatakan bahwa perkembangan investasi PMA dan PMDN dikarenakan adanya kemudahan perizinan melalui Pelayanan Terpadu Satu Pintu (PTSP) yang sudah diterapkan di hampir seluruh kabupaten/kota yang turut menarik para investor yang juga menganggap Jawa Tengah sebagai daerah ramah investasi.



Gambar 6. Perkembangan realisasi nilai investasi Penanaman Modal Dalam Negeri (PMDN) total 35 kabupaten/kota di Jawa Tengah tahun 2010-2014 (milyar rupiah).

Kabupaten Cilacap, kabupaten Rembang, kabupaten Banyumas, kabupaten Karanganyar, dan kota Semarang menjadi lima besar wilayah dengan realisasi nilai investasi PMDN terbesar dalam kurun waktu 2010 hingga 2014 yaitu mencapai 72 persen dari total realisasi nilai investasi PMDN di Jawa Tengah. Sedangkan, kabupaten/kota lainnya hanya menyumbang 28 persen dari total realisasi nilai investasi PMDN di Jawa Tengah. Hal tersebut menunjukkan bahwa investasi PMDN masih belum merata di kabupaten/kota di Jawa Tengah. Kabupaten Cilacap menjadi wilayah dengan realisasi nilai investasi PMDN terbesar di Jawa Tengah dikarenakan adanya pengembangan relokasi bisnis di Provinsi Jawa Tengah dimana yang salah satu tujuannya adalah Kabupaten Cilacap.

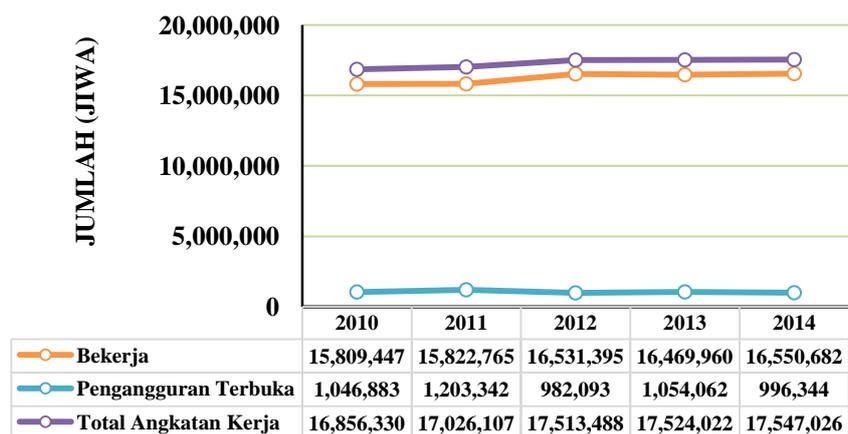


Gambar 7. Distribusi total realisasi nilai investasi Penanaman Modal Dalam Negeri (PMDN) berdasarkan lokasi kabupaten/kota di Jawa Tengah tahun 2010-2014 (persen).

Tenaga Kerja

Jumlah penduduk yang bekerja di Provinsi Jawa Tengah mengalami kenaikan, yaitu sebesar 15,8 juta jiwa pada tahun 2010 menjadi 16,6 juta jiwa pada tahun 2014. Sedangkan, jumlah pengangguran terbuka di Provinsi Jawa Tengah mengalami penurunan, yaitu sebesar 1.046.833 jiwa pada tahun 2010 menjadi 996.344 jiwa pada tahun 2014. Secara keseluruhan, total angkatan kerja di Provinsi

Jawa Tengah terus mengalami kenaikan pada kurun waktu 2010 hingga 2014, yaitu sebesar 16.856.330 jiwa pada tahun 2010 menjadi 17.547.026 jiwa pada tahun 2014. Kenaikan rasio jumlah penduduk yang bekerja terhadap angkatan kerja menunjukkan bahwa terjadi peningkatan penyerapan tenaga kerja di Provinsi Jawa Tengah, dimana pada tahun 2010 penyerapan tenaga kerja sebesar 93,79 persen menjadi 94,32 persen pada tahun 2014. Sejalan dengan hal tersebut, Tingkat Pengangguran Terbuka (TPT) di Provinsi Jawa Tengah juga mengalami penurunan, yaitu sebesar 6,21 persen pada tahun 2010 menjadi 5,68 persen pada tahun 2014. Hal tersebut menunjukkan bahwa kondisi ketenagakerjaan di Provinsi Jawa Tengah mengalami perbaikan pada kurun waktu 2010-2014. Perbaikan kondisi ketenagakerjaan tersebut tidak terlepas dari perkembangan investasi yang pesat dan perkembangan kegiatan usaha di Provinsi Jawa Tengah pada kurun waktu tersebut.



Gambar 8. Perkembangan jumlah penduduk yang bekerja, pengangguran terbuka, dan total angkatan kerja menurut kabupaten/kota di Jawa Tengah tahun 2010-2014 (jiwa).

Analisis Inferensia

Analisis Regresi Data Panel

Dalam model regresi data panel terdapat tiga model estimasi yang dapat dihasilkan. Yang pertama adalah model *common effect*, model ini menghasilkan nilai intersep yang sama untuk seluruh kabupaten/kota sehingga model ini mengasumsikan bahwa pertumbuhan ekonomi setiap kabupaten/kota di Jawa Tengah adalah sama pada kurun waktu 2010 hingga 2014. Yang kedua adalah model *fixed effect*, model ini menghasilkan nilai intersep yang berbeda untuk masing-masing kabupaten/kota sehingga model ini mengasumsikan bahwa terdapat perbedaan pertumbuhan ekonomi masing-masing kabupaten/kota di Jawa Tengah pada kurun waktu 2010 hingga 2014. Yang terakhir adalah model *random effect*, model ini juga menghasilkan nilai intersep yang berbeda untuk masing-masing kabupaten/kota. Perbedaan antara model *fixed effect* dan model *random effect* adalah sifat intersep pada model *fixed effect* selalu tetap, sedangkan pada model *random effect* intersep diasumsikan bersifat acak dan nilainya dimasukkan ke dalam komponen *error*.

Berdasarkan hasil estimasi regresi diperoleh model terbaik yaitu model *fixed effect* dengan *cross section weight* dan *cross section SUR* (PCSE).

Tabel 4. Ringkasan statistik hasil estimasi regresi model *fixed effect* (FEM) dengan *cross section weight* dan *cross section SUR* (PCSE).

<i>Variable</i>	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-Statistic</i>	<i>t-Table</i>	<i>Prob.</i>
(1)	(2)	(3)	(4)	(5)	(6)
DK	0.016634	0.003042	5.468836	1.6562	0,0000 *
BM	0.003576	0.000888	4.026200	1.6562	0,0000 *
LNPMA	0.002530	0.000993	2.548796	1.6562	0,0059 *
LNPMDN	0.001219	0.000687	1.774807	1.6562	0,0390 *
LNTK	0.548808	0.134822	4.070622	1.6562	0,0000 *
C	2.286413	1.735104	1.317739	1.6562	0.0949
Weighted Statistics					
R-squared				0.997389	
Adjusted R-squared				0.996635	
F-statistic				1322.311	
Prob(F-statistic)				0.000000	

Keterangan: Hasil pengolahan program *Eviews 8.0* (satu arah)

(*) Signifikan pada taraf signifikansi 5 persen

Hasil regresi menunjukkan bahwa baik secara simultan maupun secara parsial, pada tingkat signifikansi 5 persen seluruh variabel independen, yaitu derajat kemandirian fiskal, rasio belanja modal, Penanaman Modal Asing (PMA), Penanaman Modal Dalam Negeri (PMDN), dan tenaga kerja memiliki hubungan positif dan berpengaruh signifikan terhadap pertumbuhan ekonomi kabupaten/kota di Jawa Tengah tahun 2010-2014.

Berdasarkan model terbaik di atas dapat dilihat nilai *R-squared* sebesar 0,997389 dan nilai *Adjusted R-squared* sebesar 0,996635. Hal tersebut menunjukkan bahwa variabel independen yang digunakan dalam model dapat menjelaskan keragaman variabel pertumbuhan ekonomi sebesar 99,66 persen, sedangkan sisanya dijelaskan oleh variabel lain yang tidak terdapat dalam model. Nilai statistik uji F atau uji simultan menunjukkan nilai *p-value* kurang dari 0,05 sehingga dapat disimpulkan bahwa pada tingkat signifikansi 5 persen, minimal terdapat satu variabel independen yang berpengaruh signifikan terhadap pertumbuhan ekonomi. Persamaan yang terbentuk dalam penelitian ini adalah sebagai berikut:

$$\ln\widehat{PDRB}_{it}=(2,286413+\mu_i)+0,016634DK_{it}^*+0,003576BM_{it}^*+0,002530\ln PMA_{it}^*+0,001219\ln PMDN_{it}^*+0,548808\ln TK_{it}^*$$

Pada model *fixed effect* terdapat efek individu yang merupakan gambaran heterogenitas setiap kabupaten/kota. Heterogenitas antar kabupaten/kota yang dihasilkan mencerminkan adanya faktor-faktor atau variabel-variabel lain yang dimiliki oleh suatu kabupaten/kota tetapi tidak dimiliki oleh kabupaten/kota lainnya. Dengan kata lain, kabupaten/kota tersebut memiliki karakteristik masing-masing yang tercermin dalam variabel lain di luar variabel independen dalam model. Apabila diasumsikan variabel independen dalam persamaan sama atau tidak

berubah maka pertumbuhan ekonomi kabupaten/kota di Jawa Tengah hanya akan tergantung dari efek individu. Wilayah yang memiliki nilai efek individu terbesar adalah Kabupaten Kudus yaitu sebesar 1,32055. Artinya jika variabel independen dalam persamaan diasumsikan sama atau tidak berubah maka Kabupaten Kudus akan memiliki pertumbuhan ekonomi tertinggi di Provinsi Jawa Tengah yang hanya akan tergantung dari efek individu, yaitu sebesar 1,32055. Efek individu tersebut merupakan cerminan dari variabel yang tidak terobservasi, sehingga untuk menganalisisnya adalah dengan menjelaskan faktor-faktor atau variabel-variabel yang tidak dimasukkan dalam model tetapi masih memiliki korelasi dengan variabel-variabel independen dalam model (Ekananda, 2016). Faktor-faktor tersebut antara lain dikarenakan letak strategis Kabupaten Kudus yang berada di jalur perlintasan ekonomi antar provinsi serta keberadaan industri rokok di Kabupaten Kudus yang berperan besar dalam menggerakkan roda perekonomian Kabupaten Kudus.

3. Kesimpulan

- 1.a. Secara umum pertumbuhan ekonomi Provinsi Jawa Tengah tahun 2010-2014 cenderung tumbuh secara lambat dengan rata-rata pertumbuhan ekonomi sebesar 5,30 persen. Sedangkan, pertumbuhan ekonomi kabupaten/kota di Jawa Tengah tahun 2010-2014 cukup beragam dan berfluktuatif.
- b. Tingkat kemandirian fiskal kabupaten/kota di Jawa Tengah tahun 2010-2014 masih tergolong rendah. Hal tersebut ditunjukkan oleh sebagian besar kabupaten/kota di Jawa Tengah yang memiliki derajat kemandirian fiskal kurang dari 20 persen. Sementara itu, investasi baik oleh pemerintah maupun swasta mengalami peningkatan, khususnya investasi swasta melalui Penanaman Modal Asing (PMA) dan Penanaman Modal Dalam Negeri (PMDN) yang tumbuh pesat pada kurun waktu 2010-2014. Sedangkan, rasio belanja modal kabupaten/kota di Jawa Tengah tahun 2010-2014 masih tergolong rendah, yaitu kurang dari 20 persen dari total belanja daerah. Sedangkan, jumlah tenaga kerja di Provinsi Jawa Tengah juga mengalami peningkatan pada kurun waktu 2010-2014, yang diiringi oleh kenaikan penyerapan tenaga kerja serta penurunan tingkat pengangguran terbuka (TPT).
2. Variabel derajat kemandirian fiskal, rasio belanja modal, PMA, PMDN, dan tenaga kerja memiliki hubungan positif dan berpengaruh signifikan terhadap pertumbuhan ekonomi kabupaten/kota di Jawa Tengah tahun 2010-2014.

Referensi

- [1] Badan Pusat Statistik. (2015). Produk Domestik Regional Bruto Kabupaten/Kota di Indonesia 2010-2014. Jakarta: BPS.
- [2] Badan Pusat Statistik Jawa Tengah. (2011). Jawa Tengah Dalam Angka 2011-2015. Semarang: BPS Jawa Tengah dan Bappeda Jawa Tengah.
- [3] Badan Pusat Statistik Jawa Tengah. (2011-2015). Statistik Keuangan Pemerintah Provinsi dan Kabupaten/Kota di Jawa Tengah 2010-2015. Semarang: BPS Jawa Tengah.
- [4] Bank Indonesia. (2013). *Kajian Ekonomi Regional Jawa Tengah Triwulan IV 2013*. Semarang: Kantor Perwakilan Bank Indonesia Wilayah V (Jateng-DIY).

- [5] Bank Indonesia. (2014). *Kajian Ekonomi Regional Jawa Tengah Triwulan IV 2014*. Semarang: Kantor Perwakilan Bank Indonesia Wilayah V (Jateng-DIY).
- [6] Bappenas. (2015). *Seri Analisis Pembangunan Wilayah Provinsi Jawa Tengah 2015*. Jakarta: Bappenas.
- [7] Ekananda, Mahyus. (2016). *Analisis Ekonometrika Data Panel: Edisi 2*. Jakarta: Mitra Wacana Media.
- [8] Mungkasa, Oswar. 2013. Desentralisasi dan Otonomi Daerah di Indonesia: Konsep, Pencapaian dan Agenda Kedepan. Diakses pada tanggal 30 Agustus 2016 melalui [https://www.academia.edu/2759012/Desentralisasi dan Otonomi Daerah di Indonesia Konsep Pencapaian dan Agenda Kedepan](https://www.academia.edu/2759012/Desentralisasi_dan_Otonomi_Daerah_di_Indonesia_Konsep_Pencapaian_dan_Agenda_Kedepan).
- [9] Oates, Wallace E. (1993). Fiscal Decentralization and Economic Development. *National Tax Journal*, Vol.46 No.2, 237-243.
- [10] Todaro, Michael P dan Stephen C. Smith. (2003). *Pembangunan Ekonomi di Dunia Ketiga Edisi Kedelapan : Jilid 1*. Jakarta: Erlangga.

ANALISIS DISPARITAS PEMBANGUNAN DAERAH TERTINGGAL, TERDEPAN, DAN TERLUAR DI INDONESIA TAHUN 2008-2013

LINDA ANNISA¹, ITA WULANDARI²

¹ Badan Pusat Statistik, 12.7223@stis.ac.id

² Sekolah Tinggi Ilmu Statistik, ita.wulandari@stis.ac.id

Abstrak. Daerah Tertinggal, Terdepan dan Terluar (3T) mencakup daerah perbatasan dan daerah tertinggal. Daerah ini memiliki karakteristik yang sama, yaitu pertumbuhan dan perkembangan wilayah yang lamban dibandingkan daerah lainnya. Hal ini mengakibatkan terjadinya kesenjangan tingkat pembangunan antar wilayah (disparitas) antara daerah 3T dan daerah non-3T. Apabila hal tersebut terus dibiarkan, disparitas ini dapat membahayakan stabilitas dan kedaulatan bangsa. Oleh karena itu, peneliti ini menganalisis tingkat disparitas pembangunan yang terjadi antara kabupaten/kota yang tergolong sebagai daerah 3T dan daerah non-3T. Metode analisis yang digunakan adalah Indeks Williamson, Tipologi Klassen, dan analisis regresi data panel. Hasil analisis Indeks Williamson menunjukkan bahwa terjadi disparitas pembangunan manusia antara daerah 3T dan daerah non-3T dengan kecenderungan menurun dalam periode 2008-2013. Dari analisis tipologi Klassen tahun 2008-2013, terlihat bahwa hampir seluruh daerah 3T mengalami peningkatan kuadran yang berarti pembangunan ekonomi di daerah 3T membaik. Kemudian, dari hasil analisis regresi data panel diperoleh bahwa presentase rumah tangga dengan akses terhadap air bersih dan persentase rumah tangga dengan akses terhadap listrik memiliki pengaruh yang signifikan dan positif terhadap Indeks Pembangunan Manusia (IPM). Sementara itu, variabel persentase penduduk miskin memiliki pengaruh signifikan dan negatif terhadap IPM.

Kata kunci : disparitas pembangunan, daerah 3T, Indeks Williamson, Tipologi Klassen, *random effects*.

1. Pendahuluan

Pembangunan merupakan suatu proses dimana terjadi perubahan pada struktur sosial dan institusi. Proses ini disertai percepatan pada pertumbuhan ekonomi, pengurangan kesenjangan, dan pengentasan kemiskinan.[1] Pembangunan ditujukan untuk keadilan yang merata, dimana seluruh penduduk memiliki pangan, pekerjaan, dan pendapatan yang mencukupi.[2]

Akan tetapi, tujuan tersebut belum tercapai di Indonesia. Ketidakmerataan ini terlihat dari kesenjangan karakteristik sosial-ekonomi antar wilayah yang mengakibatkan adanya daerah yang tergolong maju dan kurang maju. Salah satu daerah yang tergolong kurang maju tersebut adalah daerah Tertinggal, Terdepan, dan Terluar (3T). Daerah yang tergolong Tertinggal, Terdepan, dan Terluar (3T) adalah daerah yang berbatasan dengan negara lain, baik perbatasan darat maupun

perbatasan laut, dan daerah yang tertinggal. Kabupaten/kota yang termasuk daerah 3T tersebar di 27 provinsi di Indonesia. Jumlah kabupaten/kota tersebut adalah 143 termasuk di antaranya adalah Timor Tengah Utara, Bengkayang, Nunukan, dan Pegunungan Bintang. Sementara itu, 371 kabupaten/kota tidak tergolong sebagai daerah 3T, di antaranya adalah Flores Timur, Halmahera Tengah, Mimika, dan Kaimana.

Daerah 3T memiliki karakteristik yaitu ketidakmerataan akses masyarakat terhadap pemanfaatan sumber daya wilayah, tingginya tingkat kemiskinan, dan kesenjangan perkembangan.[3] Akibatnya, masalah di daerah 3T biasanya terkait dengan kondisi sosial, ekonomi, aksesibilitas, serta keterbatasan infrastruktur. Kesenjangan pembangunan tersebut bukan disebabkan karena daerah tersebut tidak memiliki potensi ekonomis. Namun, letak geografis yang sulit dijangkau mengakibatkan sulitnya penyediaan sarana dan prasarana di daerah ini.[4]

Untuk mengatasi kesenjangan pendidikan, pemerintah telah menjalankan program Sarjana Mengajar di Daerah Terdepan, Terluar, dan Tertinggal (SM3T) sejak tahun 2009. Program tersebut mewajibkan sarjana guru yang ingin melamar menjadi PNS mengabdikan diri di daerah terpencil terlebih dahulu. Program lain yang dijalankan pemerintah adalah Nusantara Sehat, Guru Garis Depan, dan E-Sabak. Selain itu, untuk meningkatkan perekonomian, pemerintah telah memberi kesempatan pada investor untuk turut membangun daerah 3T.[5] Namun, pembangunan di daerah 3T masih lebih tertinggal dibandingkan daerah lainnya.

Kesenjangan pembangunan ini jika terus dibiarkan dapat menimbulkan rasa tidak percaya terhadap pemerintah.[6] Selain itu, pengaruh sosial ekonomi dari negara tetangga berpotensi untuk menimbulkan konflik dan isu nasionalisme.

Berdasarkan penjelasan tersebut, kesenjangan pembangunan, jika terus dibiarkan, akan semakin melebar dan dapat membahayakan stabilitas dan kedaulatan bangsa.[7] Oleh karena itu, penelitian ini bertujuan untuk melihat perkembangan disparitas Indeks Pembangunan Manusia (IPM) antar kabupaten/kota di daerah 3T dan kabupaten/kota non-3T serta menganalisis variabel apa saja yang memengaruhi IPM di kabupaten/kota 3T tahun 2008-2013. Peneliti tertarik untuk meneliti tentang “Analisis Disparitas Pembangunan Daerah Tertinggal, Terdepan, dan Terluar di Indonesia Tahun 2008-2013”. Dengan analisis ini, diharapkan akan diperoleh gambaran mengenai kesenjangan pembangunan yang terjadi dan variabel yang memengaruhi pembangunan di daerah 3T.

2. Hasil – Hasil Utama

Perkembangan Tingkat Disparitas

Untuk mengukur tingkat disparitas pembangunan antara wilayah digunakan indeks Williamson. Dari indeks Williamson dapat dilihat perkembangan tingkat disparitas antara daerah 3T dan daerah non-3T. Untuk itu, terdapat 2 indeks Williamson yang dihasilkan, yaitu indeks Williamson untuk seluruh kabupaten/kota di Indonesia dan indeks Williamson untuk kabupaten/kota di Indonesia yang tidak termasuk daerah 3T. Kedua angka tersebut dibandingkan untuk melihat ada tidaknya kesenjangan pembangunan manusia dan perkembangan kesenjangan tersebut dari tahun ke tahun.

Dalam penelitian ini, digunakan periode 2008 hingga 2013. Penentuan batas akhir periode penelitian ini dikarenakan pada tahun 2014 penghitungan IPM telah menggunakan metode yang terbaru, dimana nilai yang dihasilkan lebih rendah sehingga tidak dapat dibandingkan dengan tahun-tahun sebelumnya yang menggunakan IPM metode lama. Alasan dari pemilihan tahun 2008 sebagai awal periode penelitian adalah karena pemerintah telah menjalankan program untuk mengembangkan daerah 3T di tahun 2010. Penelitian ini ingin melihat perkembangan kesenjangan sebelum dan setelah program pemerintah dilaksanakan dimana jarak antara periode awal penelitian hingga program pemerintah dan jarak antara program pemerintah hingga periode akhir penelitian adalah sama.

Tabel 1. Perkembangan Indeks Williamson Seluruh Kabupaten/kota Tahun 2008-2013

Tahun Indikator	2008	2009	2010	2011	2012	2013
(1)	(2)	(3)	(4)	(5)	(6)	(7)
IPM	0,059	0,060	0,061	0,060	0,059	0,058
AHH	0,043	0,043	0,042	0,041	0,041	0,040
AMH	0,074	0,077	0,081	0,080	0,079	0,078
RLS	0,204	0,203	0,201	0,196	0,192	0,190
PPP	0,023	0,023	0,025	0,024	0,024	0,024

Tabel 1 menunjukkan perkembangan indeks

indeks Williamson untuk seluruh kabupaten/kota tahun 2008-2013. Berdasarkan Tabel 1, Indeks Williamson untuk IPM, Angka Melek Huruf (AMH), dan *Purchasing Power Parity* (PPP) mengalami peningkatan dari tahun 2008 hingga 2010. Meskipun IPM daerah 3T terus meningkat, namun peningkatan ini lebih rendah dibandingkan daerah lain yang tidak termasuk 3T. Hal ini dapat terlihat dari perhitungan laju perkembangan IPM tahun 2008-2013 dimana tidak ada daerah 3T yang memiliki laju perkembangan tergolong cepat. Sementara itu, 40,97% daerah non-3T memiliki laju perkembangan tergolong cepat. Sehingga, terjadi peningkatan kesenjangan pada periode tersebut. Namun, setelah dilaksanakannya program pemerintah, indeks Williamson ketiga indikator tersebut terus menurun hingga 2013. Sementara itu, indeks Williamson AMH dan Rata-rata Lama Sekolah (RLS) mengalami penurunan dari awal hingga akhir periode penelitian.

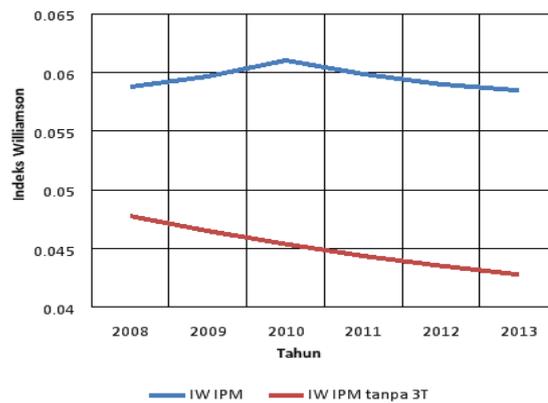
Disparitas tertinggi terjadi pada indikator AMH. Hal ini disebabkan karena beberapa daerah 3T memiliki AMH sangat rendah dibandingkan kabupaten/kota lainnya. Selain itu, capaian ini cenderung stagnan dari tahun ke tahun. Sehingga, ukuran kesenjangan pun membesar.

Tabel 2. Perkembangan indeks Williamson kabupaten/kota non-3T tahun 2008-2013

Tahun Indikator	2008	2009	2010	2011	2012	2013
(1)	(2)	(3)	(4)	(5)	(6)	(7)
IPM	0,048	0,046	0,045	0,044	0,043	0,043

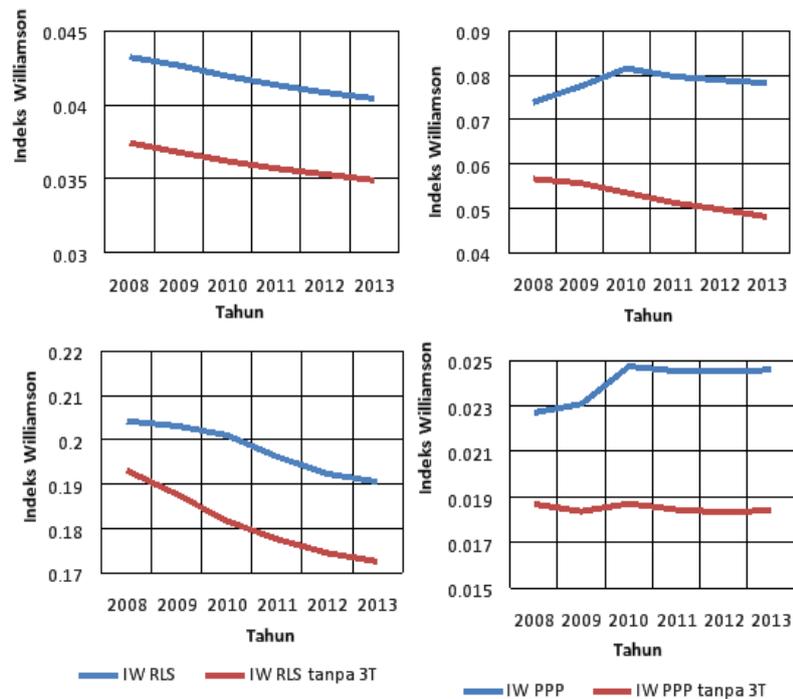
AHH	0,037	0,037	0,036	0,036	0,035	0,035
AMH	0,057	0,056	0,053	0,051	0,050	0,048
RLS	0,193	0,188	0,182	0,177	0,174	0,172
PPP	0,019	0,018	0,019	0,018	0,018	0,018

Kemudian, indeks Williamson dihitung untuk kabupaten/kota yang tidak termasuk daerah 3T. Pada tabel 2, dapat dilihat bahwa indeks Williamson yang dihasilkan lebih rendah dibandingkan ketika daerah 3T diikutsertakan dalam analisis. Hal ini menunjukkan bahwa memang terjadi kesenjangan dalam hal pembangunan manusia antara daerah 3T dan daerah non-3T.



Gambar 1. Perbandingan indeks Williamson IPM antara kabupaten/kota seluruh Indonesia dan kabupaten/kota yang tidak termasuk daerah 3T

Pada indikator IPM, AHH, AMH, dan RLS terjadi penurunan indeks Williamson dari awal hingga akhir periode penelitian. Dengan kata lain, tingkat kesenjangan antar daerah yang tidak termasuk 3T terus menurun. Hasil ini berbeda dengan hasil analisis sebelumnya dimana ketika daerah 3T diikutsertakan dalam analisis, tingkat kesenjangan meningkat hingga tahun 2010 dan mulai mengalami penurunan sejak tahun 2011.



Gambar 2. Perbandingan indeks Williamson indikator IPM antara kabupaten/kota seluruh Indonesia dan kabupaten/kota yang tidak termasuk daerah 3T

Gambar 1 adalah perbandingan indeks Williamson IPM antara kabupaten/kota seluruh Indonesia dan kabupaten/kota yang tidak termasuk daerah 3T. Hal tersebut terlihat bahwa terdapat perbedaan di antara kedua daerah tersebut dimana indeks Williamson untuk kabupaten/kota yang tidak termasuk daerah 3T lebih rendah dibandingkan indeks Williamson seluruh kabupaten/kota Indonesia. Hal ini mengindikasikan bahwa masuknya daerah 3T dalam penghitungan indeks mengakibatkan kesenjangan meningkat. Peningkatan kesenjangan ini dikarenakan kualitas pembangunan manusia yang rendah di daerah 3T.

Kemudian, jika dilihat dari indikator pembentuk IPM pada gambar 2, terlihat bahwa indeks Williamson kabupaten/kota seluruh Indonesia, termasuk 3T, selalu lebih tinggi dibandingkan indeks Williamson kabupaten/kota yang tidak termasuk 3T.

Pada Gambar 2 juga dapat dilihat bahwa pada indikator AHH dan RLS, dari awal hingga akhir periode penelitian, perbedaan antara tingkat kesenjangan kedua daerah tersebut relatif sama tiap tahunnya. Dengan kata lain, peningkatan capaian indikator AHH dan RLS antara daerah 3T dan daerah non-3T relatif sama. Sementara itu, pada indikator AMH dan PPP, indeks Williamson kabupaten/kota seluruh Indonesia meningkat pada periode 2008 hingga 2010.

Tingkat Kemajuan Kabupaten/Kota 3T

Untuk melihat tingkat kemajuan tiap daerah 3T, akan digunakan tipologi Klassen. Penentuan kuadran untuk daerah 3T dilakukan dengan cara membandingkan PDRB per kapita dan laju pertumbuhan ekonomi daerah terhadap PDB per kapita dan laju pertumbuhan ekonomi Indonesia.

Daerah yang tergolong kuadran I adalah daerah yang maju dan tumbuh

cepat dimana PDRB per kapita sudah lebih tinggi dari PDB per kapita Indonesia dan laju pertumbuhan PDRB kabupaten/kota juga lebih cepat dari PDB Indonesia. Kemudian, daerah yang tergolong kuadran II adalah daerah maju tetapi tertekan dimana PDRB per kapita kabupaten/kota lebih tinggi daripada PDB per kapita Indonesia, namun laju pertumbuhan PDRB kabupaten/kota masih di bawah laju pertumbuhan PDB Indonesia. Kuadran III menunjukkan daerah yang berkembang cepat, dimana PDRB per kapita masih di bawah PDB Indonesia, namun laju pertumbuhan PDRB lebih cepat dibandingkan laju pertumbuhan PDB Indonesia. Kuadran IV menunjukkan daerah relatif tertinggal dimana PDRB kabupaten/kota dan laju pertumbuhan PDRB kabupaten/kota lebih rendah dibandingkan Indonesia.

Tabel 3. Persentase daerah 3T berdasarkan perubahan kuadran tipologi Klassen daerah 3T tahun 2008-2013

2013	I	II	III	IV	Total
2008	I	II	III	IV	Total
(1)	(2)	(3)	(4)	(5)	(6)
I	5.66	1.89	1.89	0.00	9.43
II	4.72	0.00	2.83	0.00	7.55
III	0.00	32.08	0.00	0.00	32.08
IV	1.77	45.28	0.00	1.89	50.94
Total	14.15	79.25	4.72	1.89	100.00

Keterangan:

Abu-abu : Kuadran tahun 2013 lebih rendah dibanding kuadran tahun 2008

Diarsir : Kuadran tahun 2013 lebih tinggi dibanding kuadrat tahun 2008

Pada tahun 2013, semakin banyak daerah 3T yang mengalami peningkatan menjadi kuadran I. Sehingga, keadaan ekonomi daerah 3T jika dilihat dari PDRB per kapita dan laju pertumbuhan PDRB semakin membaik dibandingkan tahun 2008, terlihat dari lebih banyaknya daerah yang tergolong sebagai daerah maju dan lebih sedikit daerah yang tergolong relatif tertinggal.

Jika melihat perubahan kuadran tipologi Klassen yang terjadi antara 2008 dan 2013, sebanyak 91 daerah (85.85%) tergolong dalam kuadran yang lebih tinggi dibandingkan tahun 2008, 7 daerah (6.60%) tergolong dalam kuadran yang lebih rendah dibandingkan tahun 2008, yaitu Nias Selatan, Donggala, Bengkalis, Rokan Hilir, Kota Dumai, Sumbawa Barat, dan Sorong, dan 8 daerah (7,55%) tergolong dalam kuadran yang sama dengan tahun 2008 yaitu Indragiri Hilir, Pelalawan, Kota Batam, Teluk Bintuni, Raja Ampat, Kota Jayapura, Bangkalan, dan Seram Bagian Timur.

Pengaruh Beberapa Variabel Independen terhadap IPM

Untuk melihat variabel-variabel yang memengaruhi IPM digunakan regresi data panel. Dalam penelitian ini, terdapat tiga variabel independen, yaitu persentase rumah tangga dengan akses terhadap air bersih, persentase persentase rumah tangga dengan akses terhadap listrik, dan persentase penduduk miskin.

Uji Kesesuaian Model Regresi Data Panel

Uji Chow dilakukan untuk melihat apakah model *fixed effects* lebih baik dibandingkan model *common effects*. Hasil penghitungan menunjukkan bahwa nilai *F-statistics* adalah 227.8605 yang lebih besar daripada $F_{0,05; 105, 527} = 1.2681$. Akibatnya, H_0 ditolak atau *intercept* berbeda antar daerah. Sehingga model *fixed effects* lebih baik dibandingkan model *common effects*.

Karena kesimpulan dari uji Chow menunjukkan bahwa model *fixed effects* lebih baik, maka dilanjutkan dengan melakukan uji Hausman. Uji Hausman dilakukan untuk melihat apakah model *random effects* lebih baik dibandingkan model *fixed effects*.

Hasil penghitungan menunjukkan bahwa nilai *Hausman-statistics* adalah 7.5523 yang lebih kecil daripada $\chi^2_{0,05;3} = 7.8147$. Akibatnya, H_0 tidak ditolak. Sehingga model *random effects* lebih baik dibandingkan model *fixed effects*.

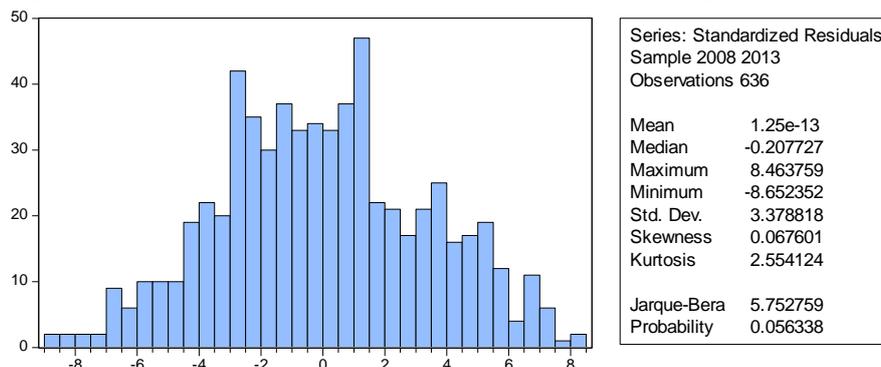
Estimasi model *random effects*

Berdasarkan uji kesesuaian model yang telah dilakukan, model *random effects* lebih baik dibandingkan model regresi panel lainnya. Model ini mempertimbangkan heterogenitas individu.

$$\widehat{IPM}_{it} = 70.7868 + u_i + 0.0099AIRBERSIH_{it} + 0.0262LISTRIK_{it} - 0.2354MISKIN_{it}$$

t	(139.8170)	(3.6055)	(8.4215)	(-23.6216)
p-value	(0.0000)	(0.0003)	(0.0000)	(0.0000)
R ²	= 66,1838%			
Adjusted R ²	= 66,0233%			
F-statistics	= 412,3097			

Model *random effects* telah mengakomodasi autokorelasi dan heteroskedastisitas.[8] Sehingga uji asumsi yang perlu dilakukan adalah uji asumsi normalitas dan tidak multikolinearitas. Uji normalitas menggunakan uji Jarque-Bera yang ditunjukkan oleh Gambar 3. Asumsi normalitas telah terpenuhi karena nilai uji Jarque-Bera memberikan nilai *p-value* sebesar 0,0563 yang lebih besar dari $\alpha = 0,05$. Asumsi tidak multikolinearitas juga telah terpenuhi karena nilai VIF masing-masing variabel independen kurang dari 10 seperti yang terlihat pada Tabel



3.

Gambar 3. Uji normalitas menggunakan Jarque-Bera

Tabel 3. Uji Multikolinearitas
Coefficients^a

Model	Unstandardized Coefficients		Standardized Coefficients	t
	B	Std. Error	Beta	
1 (Constant)	67.920	.771		88.112
airbersih	.043	.008	.186	5.283
listrik	.025	.008	.135	3.213
miskin	-.173	.016	-.432	-11.088

a. Dependent
Variable: ipm

Pengaruh variabel independen terhadap IPM

Berdasarkan model regresi data panel yang diperoleh, nilai *Adjusted R*² adalah sebesar 66.0233%. Dengan kata lain, sebesar 66.0233% variasi pada variabel IPM mampu dijelaskan oleh variabel independen yang terdapat dalam model. Sementara 33.9766% sisanya dijelaskan oleh variabel lain yang tidak terdapat dalam model. Hasil uji simultan memberikan nilai *p-value* kurang dari $\alpha = 0.05$, sehingga seluruh variabel independen memiliki pengaruh secara serentak yang signifikan terhadap IPM.

Uji parsial memberikan nilai *p-value* kurang dari $\alpha = 0.05$ pada masing-masing variabel independen. Sehingga, masing-masing variabel independen memberikan pengaruh signifikan terhadap IPM, dengan asumsi variabel lain konstan.

Berdasarkan model yang terbentuk, dapat dilihat bahwa hanya variabel persentase penduduk miskin yang berpengaruh negatif terhadap IPM. Variabel ini juga memberikan pengaruh terbesar terhadap IPM, yaitu sebesar -0.2354.

Variabel persentase rumah tangga dengan akses terhadap air bersih memiliki koefisien regresi sebesar 0.0099. Artinya, peningkatan persentase rumah tangga dengan akses terhadap air bersih sebesar 1% akan meningkatkan IPM sebesar 0.0099, dengan asumsi variabel lain konstan. Air bersih dapat meningkatkan IPM karena akses terhadap air bersih berkontribusi signifikan dalam mengurangi kematian anak.[9]

Hasil ini sesuai dengan penelitian Sapkota yang meneliti pengaruh ketersediaan infrastruktur terhadap IPM di beberapa negara. Sapkota menemukan bahwa akses terhadap air bersih signifikan dalam meningkatkan capaian IPM.[10]

Variabel persentase rumah tangga dengan akses terhadap listrik memiliki koefisien regresi sebesar 0.0262. Dengan kata lain, semakin banyak rumah tangga yang memiliki akses terhadap listrik (PLN dan non-PLN) maka semakin tinggi

capaian IPM daerah tersebut. Hal ini disebabkan peningkatan konsumsi listrik dapat secara langsung merangsang pertumbuhan ekonomi yang lebih cepat dan secara tidak langsung meningkatkan pembangunan sosial terutama pada daerah yang berkembang.[11] Peningkatan persentase rumah tangga dengan akses terhadap listrik sebesar 1% akan meningkatkan IPM sebesar 0.0262, dengan asumsi variabel lain konstan.

Hasil ini sejalan dengan penelitian Kusharjanto dan Kim yang meneliti pengaruh ketersediaan infrastruktur terhadap IPM kabupaten/kota di Pulau Jawa. Penelitian ini menyimpulkan bahwa listrik memberikan pengaruh terbesar dalam peningkatan IPM dibandingkan variabel lainnya dalam penelitian.[12]

Variabel persentase penduduk miskin memiliki koefisien regresi sebesar -0.2354. Artinya, semakin besar persentase penduduk miskin, maka semakin rendah capaian IPM daerah tersebut. Hal ini dikarenakan seseorang yang berstatus miskin mengalami keterbatasan untuk meningkatkan kualitas pembangunan manusia. Jika seseorang miskin, akan lebih sulit untuk mengakses fasilitas kesehatan dan pendidikan yang layak dan pada akhirnya akan menurunkan kualitas sumber daya manusia di daerah tersebut.[13] Sehingga, nilai IPM akan turun. Peningkatan persentase penduduk miskin sebesar 1% akan menurunkan IPM sebesar 0.1961, dengan asumsi variabel lain bernilai konstan.

Hal ini sejalan dengan penelitian Siddiqui yang meneliti pengaruh pendapatan per kapita, kemiskinan, dan ketersediaan layanan sosial masyarakat terhadap pembangunan pada distrik-distrik di Pakistan. Penelitian tersebut menyimpulkan bahwa kemiskinan memiliki hubungan negatif dengan pembangunan.[14]

3. Kesimpulan

1. Kondisi daerah 3T secara umum tergolong masih tertinggal dibandingkan daerah non 3T. Hal tersebut dapat dilihat dari:
 - a. Aspek kependudukan: Sebagian besar daerah 3T memiliki persentase penduduk miskin lebih tinggi daripada persentase penduduk miskin nasional
 - b. Aspek pendidikan: Hampir separuh daerah 3T memiliki APM Sekolah Menengah/Sederajat di bawah APM Sekolah Menengah nasional
 - c. Aspek kesehatan: Sebagian besar daerah 3T memiliki persentase rumah tangga dengan akses terhadap air bersih dan sanitasi layak lebih rendah daripada persentase nasional
 - d. Aspek ekonomi: Sebagian besar daerah 3T memiliki PDRB per kapita lebih rendah daripada PDB per kapita.
 - e. Aspek pembangunan manusia: Sebagian besar daerah 3T memiliki capaian IPM dan indikator IPM yang lebih rendah dibandingkan capaian nasional.
2. Pembangunan ekonomi daerah 3T tahun 2013 semakin membaik dibandingkan tahun 2008. Hal ini dilihat dari lebih banyaknya daerah yang tergolong maju dan lebih sedikit daerah yang tergolong relatif tertinggal pada Tipologi Klassen.
3. Terdapat pengaruh yang signifikan dari persentase rumah tangga dengan akses terhadap air bersih, persentase rumah tangga dengan akses terhadap listrik, dan persentase penduduk miskin terhadap IPM di kabupaten/kota 3T.

Referensi

- [1] Todaro, Michael P., 1981, *Economic Development in The Third World*. New York: Longman Publishing Group.
- [2] Seers, D., 1969), The Meaning of Development. *Institute of Development Studies Communication Series*, 44.
- [3] Jaya, W. K., 2015, Kedaulatan Ekonomi: Strategi Memajukan Ekonomi Daerah Terdepan, Terluar, dan Tertinggal dalam Pusat Studi Pancasila Universitas Gadjah Mada (Ed.) *Membangun Kedaulatan Bangsa berdasarkan Pancasila: Pemberdayaan Masyarakat Daerah Terluar, Terdepan, dan Tertinggal (3T)*. Jogjakarta: Pusat Studi Pancasila, Universitas Gadjah Mada.
- [4] Huruswati dan Pudjianto., 2009, Refleksi: Strategi Pengembangan Daerah Perbatasan dan Daerah Tertinggal dalam Bambang Rudito (Ed.) *Masalah Kebutuhan, dan Sumber Daya di Daerah Perbatasan dan Daerah Tertinggal*. Jakarta: P3KS Press, 243-254.
- [5] Prasajo, Z. H., 2013, Dinamika Masyarakat Lokal di Perbatasan. *Walisongo*, 21(2), 417-436.
- [6] Qasim dan Chaudhary., 2014, Determinants of Human Development Disparities: A Cross District Analysis of Punjab, Pakistan. *Annual General Meeting and Conference on Poverty, Inequality, and Economic Growth*. Pakistan Society of Development Economists. Pakistan: Islamabad.
- [7] Evianto, E., 2010, *Analisis Disparitas Indeks Pembangunan Manusia Kabupaten/Kota di Provinsi Jawa Barat dan Faktor-Faktor yang Mempengaruhi Capaiannya (Model Regresi Data Panel 25 Kabupaten/Kota Tahun 2003-2007)* [Tesis]. Jakarta: Universitas Indonesia.
- [8] Greene, W. H., 2003, *Econometric Analysis (5th ed)*. New York: Prentice Hall.
- [9] Jalan, J. dan Ravallion, M., 2003, Does Piped Water Reduce Diarrhea for Children in Rural India?. *Journal of Econometrics*, 112 (1), 153-173.
- [10] Sapkota, J.B., 2014, Infrastructure Access and Human Development: Cross Country Evidence and Post-2015 Development Strategies. *JICA Research Institute Working Papers*, 70.
- [11] Leung, C. dan Meisen, P., 2005, *How Electricity Consumption Affects Social and Economic Development by Comparing Low, Medium, and High Human Development Countries*. San Diego, GENI.
- [12] Kusharjanto dan Kim., 2011, Infrastructure and Human Development: The Case of Java, Indonesia. *Journal of the Asia Pacific Economy*, 16(1), 111-124.
- [13] Ruseva, M., D. Genov, dan V. Jekova., 2015, Impact of Poverty on the Human Development Index in Bulgaria (For Example of South Central Region). *Trakia Journal of Sciences*, 13(1).
- [14] Siddiqui, R., 2008, Income, Public Social Services, and Capability Development: A Cross-district Analysis of Pakistan. *PIDE Working Paper*, 43, 1-24.

PENERAPAN REGRESI DATA PANEL UNTUK MELIHAT PENGARUH PEMBANGUNAN INFRASTRUKTUR TERHADAP PERTUMBUHAN EKONOMI DI PULAU SUMATERA

ULFA MARDIYAH¹, WINIH BUDIARTI²

¹Badan Pusat Statistik Kabupaten Buol, ulfa.mardiyah@bps.go.id
² Sekolah Tinggi Ilmu Statistik, winih@stis.ac.id

Abstrak. Regresi data panel memiliki keunggulan dibandingkan regresi data *crosssection* dan regresi data *time series*. Penggunaan data panel yang menggabungkan kedua jenis data dalam satu model akan memberikan model yang lebih realistis dan mampu menjelaskan perbedaan antar individu dalam satu waktu maupun perbedaan perilaku suatu unit individu antar waktu. Untuk melihat pengaruh pembangunan infrastruktur terhadap pertumbuhan ekonomi provinsi-provinsi di Pulau Sumatera tahun 2004-2013 digunakan analisis regresi data panel. Pulau Sumatera merupakan pulau yang dikenal sebagai sentra produksi dan pengolahan hasil bumi serta lumbung energi nasional. Hal tersebut menjadikan Pulau Sumatera sangat berperan penting dalam menunjang pertumbuhan ekonomi di Indonesia. Namun, selama 10 tahun terakhir ini, besar kontribusi PDRB Pulau Sumatera terhadap PDB Indonesia menurun setiap tahunnya. Hal tersebut diduga disebabkan oleh pembangunan infrastruktur yang tidak merata dan kurang memadai antar provinsi di Pulau Sumatera. Variabel infrastruktur yang digunakan adalah infrastruktur jalan, pelabuhan, listrik, dan air bersih. Model estimasi data panel terbaik adalah model *fixed effect* dengan struktur varians kovarians residual bersifat heteroskedastik dan ada *cross sectional correlation* (*Seemingly Uncorelated Regression/SUR*). Dari persamaan regresi data panel, infrastruktur jalan, pelabuhan, dan air bersih terbukti secara signifikan berpengaruh positif terhadap pertumbuhan ekonomi di Pulau Sumatera, sedangkan infrastruktur listrik tidak berpengaruh terhadap pertumbuhan ekonomi di Pulau Sumatera. Infrastruktur jalan memiliki pengaruh terbesar dengan elastisitas 0,277, elastisitas infrastruktur pelabuhan sebesar 0,153 dan infrastruktur air bersih memiliki elastisitas terendah yaitu 0,151.

Kata kunci : analisis regresi data panel, model *fixed effect*, *Seemingly Uncorelated Regression*, infrastruktur, pertumbuhan ekonomi.

1. Pendahuluan

Perkembangan infrastruktur dengan pembangunan ekonomi memiliki hubungan yang sangat erat dan saling ketergantungan satu sama lain. Pemerintah dalam mendukung pembangunan ekonomi suatu daerah berusaha menyediakan sarana dan prasarana sebagai penunjang kegiatan ekonominya. Sarana dan

prasarana tersebut dapat tercermin dari pembangunan infrastruktur, baik berupa prasarana transportasi (jalan, pelabuhan, bandara), jaringan listrik dan komunikasi, maupun bangunan-bangunan lain.

Basri [9] menyebutkan bahwa peranan penyediaan infrastruktur terhadap pembangunan ekonomi yaitu: (1) Mempercepat dan menyediakan barang-barang yang dibutuhkan masyarakat dengan biaya yang lebih murah; (2) Infrastruktur yang baik dapat merangsang adanya stabilitas dan mengurangi disparitas harga antar daerah; (3) Berfungsi meningkatkan nilai tambah barang dan jasa; dan (4) Infrastruktur yang memperlancar transportasi turut mempengaruhi terbentuknya harga yang efisien.

Pulau yang sangat berpotensi untuk mendukung pertumbuhan ekonomi nasional adalah Pulau Sumatera. Wilayah ini diharapkan menjadi gerbang ekonomi nasional ke Eropa, Afrika, Asia Selatan, Asia Timur, serta Australia. Secara umum, Pulau Sumatera berkembang dengan baik di bidang ekonomi dan sosial dengan kegiatan ekonomi utama seperti perkebunan kelapa sawit, karet, serta batubara. Akan tetapi, Pulau Sumatera juga memiliki beberapa hal yang harus dibenahi, salah satunya adalah infrastruktur dasar yang kurang memadai untuk pengembangan industri, antara lain jalan yang sempit dan rusak, rel kereta api yang sudah rusak dan tua, pelabuhan laut yang kurang efisien, serta kurangnya energi listrik yang dapat mendukung industri.

Pulau Sumatera merupakan wilayah penyumbang ekonomi terbesar kedua setelah Pulau Jawa yang juga dikenal sebagai sentra produksi dan pengolahan hasil bumi dan lumbung energi nasional. Pada tahun 2004-2013 kontribusi PDRB Pulau Sumatera terhadap PDB nasional mengalami penurunan dari tahun ke tahun. Hal ini diduga karena ketersediaan infrastruktur yang kurang memadai. Oleh karena itu, perlu dilakukan penelitian tentang pengaruh infrastruktur jalan, pelabuhan, listrik, dan air bersih terhadap pertumbuhan ekonomi di Pulau Sumatera selama periode 2004-2013. Pulau Sumatera terdiri dari sepuluh provinsi yaitu Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Kepulauan Bangka Belitung, dan Kepulauan Riau. Tentunya terdapat perbedaan pertumbuhan ekonomi dan pembangunan infrastruktur antara provinsi-provinsi tersebut.

Oleh karena itu, penelitian ini akan menggunakan regresi data panel, karena data yang dikaji merupakan gabungan dari data *time series* pada variabel dependen maupun beberapa variabel independen, dan data *cross section* antar wilayah/provinsi. Menurut Baltagi dalam Gujarati [12] terdapat beberapa kelebihan penggunaan data panel: (1) Heterogenitas antarindividu dapat dikontrol, panel data mengusulkan bahwa individu bersifat berbeda-beda atau heterogen; (2) Penggunaan panel data dapat memberikan informasi data yang lebih banyak dan beragam, permasalahan multikolinieritas yang minim, derajat bebas yang lebih banyak, dan lebih efisien; (3) Analisis penyesuaian dinamis (*dynamics of change*) lebih baik dilakukan oleh panel data; (4) Data Panel lebih unggul dalam mengidentifikasi dan mengukur efek yang tidak dapat ditemukan pada data *cross section* murni atau *time-series* murni; (5) Model pada panel data dapat digunakan untuk membangun dan menguji model pada perilaku yang lebih kompleks daripada data *cross section* murni atau *time series* murni; (6) Data panel dapat menyediakan data lebih dari ribuan unit observasi, sehingga dapat meminimalkan bias yang mungkin terjadi karena agregasi individu.

Berdasarkan hal di atas, maka penelitian ini bertujuan untuk memberikan gambaran terkini mengenai perkembangan infrastruktur jalan, pelabuhan, listrik,

dan air bersih dan pertumbuhan ekonomi di Pulau Sumatera pada tahun 2004-2013 serta menerapkan regresi data panel untuk melihat pengaruh infrastruktur jalan, pelabuhan, listrik, dan air bersih terhadap pertumbuhan ekonomi di Pulau Sumatera pada tahun 2004-2013.

2. Hasil – Hasil Utama

2.1 Ruang Lingkup Penelitian dan Data yang Digunakan

Penelitian ini akan mengamati bagaimana perkembangan infrastruktur jalan, pelabuhan, listrik, dan air bersih pada 10 provinsi di Pulau Sumatera. Penelitian ini juga mengkaji mengenai hubungan dan pengaruh antara infrastruktur terhadap pertumbuhan ekonomi di Pulau Sumatera. Pulau Sumatera dipilih sebagai lokasi penelitian karena Sumatera merupakan daerah yang memiliki pembangunan infrastruktur yang cukup lengkap namun kurang memadai, khususnya infrastruktur jalan, pelabuhan, listrik, dan air bersih dibandingkan dengan pulau lainnya seperti, Jawa, Bali-Nusa Tenggara, Kalimantan, Sulawesi, dan Maluku-Papua. Analisis yang digunakan terdiri dari analisis deskriptif dengan bantuan tabel dan grafik, serta analisis inferensia dengan menggunakan regresi data panel untuk melihat pengaruh pembangunan infrastruktur jalan, pelabuhan, listrik, dan air bersih terhadap pertumbuhan ekonomi di Pulau Sumatera tahun 2004-2013.

Penelitian ini mencakup seluruh provinsi di Pulau Sumatera yakni sebanyak 10 provinsi dan menggunakan data yang tersedia selama periode 2004-2013. Tahun 2004 dipilih sebagai periode awal penelitian karena pada tahun tersebut sudah terdapat kestabilan data di masing-masing propinsi di Pulau Sumatera.

Data yang digunakan merupakan data sekunder selama periode 2004-2013 dari 10 provinsi di Pulau Sumatera, dengan rinciannya sebagai berikut:

1. PDRB atas dasar harga konstan 2000 dalam satuan mata uang milyar rupiah.
Sumber data: website BPS (www.bps.go.id) diakses pada 7 Mei 2015.
2. Panjang jalan negara, provinsi, dan kabupaten/kotamadya dalam kondisi baik dan sedang dalam satuan kilometer (km).
Sumber data: seri publikasi BPS “*Statistik Transportasi*” (tahun 2004-2013), berdasarkan pengumpulan data yang bersumber dari Direktorat Jenderal Bina Marga, Kementerian Pekerjaan Umum, Dinas Pekerjaan Umum, Dinas Pekerjaan Umum Pemerintah Provinsi dan Dinas Pekerjaan Umum Pemerintah Kabupaten/Kota.
3. Jumlah bongkar dan muat barang di pelabuhan dalam satuan ton.
Sumber data: seri publikasi BPS “*Statistik Transportasi*” (tahun 2004-2013) berdasarkan pengumpulan dari beberapa dokumen, dokumen dari pelabuhan yang diusahakan atau pelabuhan dibawah naungan PT. (Persero) Pelindo yang meliputi 111 cabang pelabuhan adalah SIMOPPEL T II-01 sampai dengan T II-09. Sedangkan dokumen dari pelabuhan yang tidak diusahakan atau pelabuhan di bawah naungan Ditjen Perhubungan Laut yang mencakup 344 Pelabuhan adalah SIMOPPEL T II-UPT, dimana sebelumnya menggunakan dokumen LL I/1 dan LL I/2.
4. Jumlah energi listrik yang disalurkan kepada pelanggan dalam satuan Gwh (giga watt hour),

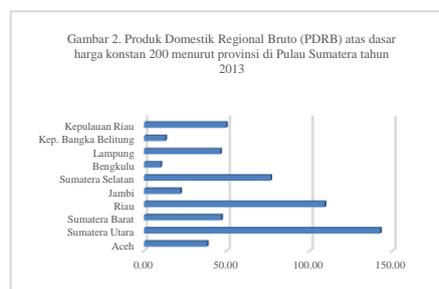
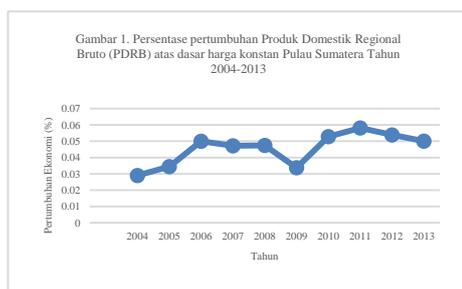
Sumber data: seri publikasi BPS “Statistik Listrik” (tahun 2004-2013) berdasarkan pengumpulan data dari hasil Survei Tahunan Perusahaan Listrik Negara (PLN), pengumpulan data menggunakan kuesioner Listrik.

- Jumlah air bersih yang disalurkan Perusahaan Air Bersih kepada pelanggan dalam satuan ribu m³

Sumber data: seri publikasi BPS “Statistik Air Bersih” (tahun 2004-2013) berdasarkan pengumpulan data dari hasil Survei Tahunan Perusahaan Air Bersih yang dilakukan oleh Koordinator Statistik Kecamatan (KSK) atau staf BPS Kabupaten/Kota dengan cara pencacahan langsung ke setiap perusahaan. Pencacahan Perusahaan Air bersih dilaksanakan secara lengkap di seluruh Indonesia.

2.2 Pertumbuhan Ekonomi dan Perkembangan Infrastruktur di Pulau Sumatera Periode 2004-2013

Pertumbuhan ekonomi merupakan salah satu indikator penting guna melakukan evaluasi dan koreksi terhadap pembangunan ekonomi yang telah dilaksanakan pada masa atau periode yang lalu. Dalam mengukur tingkat pertumbuhan ekonomi digunakan angka PDRB atas dasar harga konstan. Selama kurun waktu 2004-2013, perekonomian Sumatera pernah mengalami kontraksi karena masa krisis ekonomi global yang disebabkan oleh krisis ekonomi Amerika pada tahun 2008. Dengan adanya krisis global tersebut, laju pertumbuhan ekonomi Sumatera berkurang menjadi 3,55 persen dari yang sebelumnya 4,82 persen. Setelah itu (pada Gambar 1) pada dua tahun berikutnya menunjukkan perbaikan. Akan tetapi tahun 2012 dan 2013 mengalami penurunan pertumbuhan ekonomi yang menurut Kemenkeu disebabkan oleh pengeluaran ekspor barang dan jasa pertumbuhannya turun 0,78 persen, melambatnya investasi serta tingginya konsumsi masyarakat terutama di sektor non makanan.

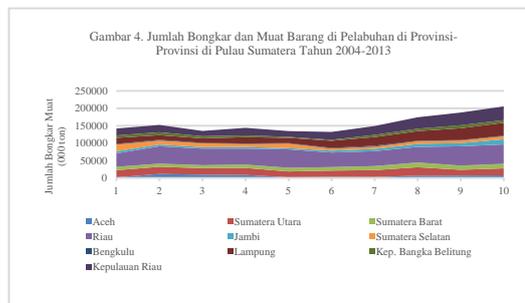


Pertumbuhan ekonomi Sumatera yang cukup tinggi juga terjadi pada tahun 2006 yaitu sebesar 5,26 persen dari tahun sebelumnya yang hanya mencapai 3,56 persen. Hal ini berbanding terbalik dengan pertumbuhan ekonomi Indonesia yang mengalami penurunan dengan angka sebesar 5,19 persen yang dari sebelumnya sebesar 5,37 persen. Penurunan ini terjadi akibat adanya kenaikan harga BBM yang menyebabkan inflasi menjadi cukup tinggi.

Pulau Sumatera juga mengalami perbedaan pendapatan produksi antar provinsi. Hal ini terlihat pada Gambar 2 yang menunjukkan besarnya ketimpangan antara provinsi yang mempunyai PDRB tertinggi dengan provinsi yang mempunyai PDRB terendah. Pada tahun 2013, PDRB tertinggi diraih oleh Sumatera Utara dengan nilai 38.012,97 milyar rupiah, PDRB tertinggi kedua Riau dengan nilai 109.073,14 milyar rupiah dan PDRB terendah adalah Bengkulu dengan nilai sebesar 10.052,31 milyar rupiah.

Jalan merupakan infrastruktur yang penting untuk menghubungkan satu daerah ke daerah lain. Ketersediaan infrastruktur jalan sangat penting dalam perekonomian karena pelayanan dan kapasitas jalan berkaitan dengan kelancaran pengangkutan barang dan jasa yang menggerakkan perekonomian. Kondisi jalan dapat dibedakan menjadi baik, sedang, rusak, dan rusak berat. Jalan dalam kondisi baik dan sedang akan memberikan kontribusi yang lebih tinggi dibandingkan dengan kondisi jalan yang rusak maupun rusak berat.

Panjang jalan di Pulau Sumatera merupakan jalan yang terpanjang dibanding pulau lainnya. Pada tahun 2013, total panjang jalan di Pulau Sumatera mencapai 174,314 km. Berdasarkan Gambar 3, persentase panjang jalan menurut kondisi di Pulau Sumatera menunjukkan bahwa 40,98 persen jalan berada dalam keadaan baik, 23,50 persen kondisi sedang, 20,65 persen dalam kondisi rusak sedangkan jalan dengan kondisi rusak berat adalah sebesar 14,87 persen. Perkembangan infrastruktur jalan dalam kondisi baik dan sedang selama periode 2004-2013 di Pulau Sumatera cenderung meningkat setiap tahunnya, dengan rata-rata tingkat pertumbuhan sebesar 4,97 persen per tahun. Akan tetapi pada tahun 2006 jalan dengan kondisi baik dan sedang terjadi penurunan sebesar 1,75 persen yang berubah menjadi kondisi rusak dan rusak berat. Provinsi yang mengalami penurunan kondisi jalan baik dan sedang adalah provinsi Sumatera Selatan, Riau, Sumatera Barat, dan Aceh. Namun setelah terjadi penurunan tersebut, Sumatera terus menunjukkan perbaikan kondisi jalan tiap tahunnya.



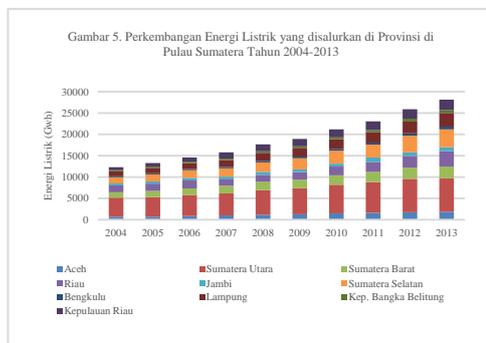
Di Pulau Sumatera, infrastruktur jalan bukanlah satu-satunya infrastruktur yang berperan penting dalam menghubungkan wilayah satu dengan lainnya. Karena Pulau Sumatera merupakan wilayah dengan garis pantai yang luas maka infrastruktur pelabuhan juga sangat berperan penting untuk mendorong pertumbuhan ekonomi di Sumatera.

Grafik 4 menunjukkan bahwa jumlah bongkar muat barang di Pulau Sumatera berfluktuatif, namun selama empat tahun terakhir cenderung meningkat. Hal ini disebabkan oleh penyumbang jumlah bongkar dan muat barang terbanyak di Pulau Sumatera, yaitu provinsi Riau, Kepulauan Riau, dan Lampung dengan rata-rata jumlah bongkar dan muat barang sebesar 47.515 ribu ton, 25.109 ribu ton, dan 22.001 ribu ton pertahun. Ketiga provinsi juga cenderung mengalami peningkatan setiap tahunnya, karena provinsi-provinsi tersebut merupakan wilayah yang strategis untuk proses pengangkutan barang-barang kebutuhan produksi dan hasil produksi antar pulau maupun luar negeri.

Listrik merupakan salah satu energi yang berperan penting dalam perkembangan kehidupan manusia modern. Di Pulau Sumatera, penjualan energi listrik terus meningkat dari tahun ke tahun. Hal ini dapat dilihat pada Gambar 5 yang menunjukkan bahwa setiap tahun terjadi peningkatan energi listrik yang didistribusikan kepada pelanggan. Peningkatan energi listrik tertinggi terjadi pada

tahun 2012 yang pertumbuhannya mencapai 12,32 persen dengan jumlah energi listrik sebesar 25.875,85 Gwh. Peningkatan energi listrik yang terendah terjadi pada tahun 2009 yang pertumbuhannya mencapai 7,06 persen dengan jumlah energi listrik sebesar 18.919,57 Gwh. Hal ini disebabkan adanya krisis global yang mengakibatkan kurangnya alokasi dana untuk peningkatan terhadap energi listrik di koridor Sumatera.

Mulai tahun 2004 hingga 2013, Provinsi Sumatera Utara merupakan wilayah dengan jumlah energi listrik yang didistribusikan kepada pelanggan paling tinggi di Pulau Sumatera dengan rata-rata jumlah energi listrik sebesar 5.889,10 Gwh. Pada tahun 2013, penggunaan energi listrik di Sumatera Utara terutama dimanfaatkan untuk pelanggan rumah tangga sekitar 49 persen, untuk industri sebesar 27 persen dan untuk bisnis sekitar 15 persen. Sedangkan jumlah energi listrik yang didistribusikan kepada pelanggan yang paling rendah terdapat pada Provinsi Bengkulu dengan rata-rata jumlah energi listrik sebesar 374,63 Gwh. Pada tahun 2013 di Bengkulu, pengguna energi listrik paling banyak dimanfaatkan oleh rumah tangga sekitar 75 persen. Adanya perbedaan jumlah energi listrik yang didistribusikan kepada pelanggan di Provinsi Sumatera Utara dan Bengkulu tersebut mencerminkan belum meratanya ketersediaan infrastruktur listrik di Pulau Sumatera.



Pembangunan infrastruktur air bersih yang digunakan dalam penelitian ini adalah ketersediaan air bersih yang diselenggarakan oleh PDAM di seluruh provinsi yang ada di pulau Sumatera. Pada periode 2004-2013, volume air bersih yang disalurkan pada pelanggan di Pulau Sumatera terus meningkat dari tahun ke tahun (Gambar 6), namun pada tahun 2007 mengalami penurunan volume air bersih yang didistribusikan terhadap pelanggan menjadi sebesar 8,45 persen. Hal ini disebabkan karena jumlah pelanggan pada tahun 2007 menurun dari 1,62 juta pelanggan menjadi 1,39 juta pelanggan. Peningkatan volume air bersih tertinggi terjadi pada tahun 2012 dengan tingkat pertumbuhan sebesar 16,94 persen dengan jumlah volume air bersih yang disalurkan sebesar 594,66 juta m³ dan jumlah pelanggan sebesar 2,09 juta pelanggan. Walaupun dari gambar 13 tersebut menunjukkan bahwa volume air bersih cenderung meningkat, namun fakta yang menunjukkan bahwa pelayanan air bersih di Pulau Sumatera masih sangat terbatas. Sehingga pemenuhan kebutuhan air bersih masyarakat sebagian besar bersumber dari sumur dan air hujan [8].

2.3 Regresi Data panel

Data panel merupakan data yang terdiri dari data *cross section* dan data *time series*. Jenis data ini memiliki keunggulan dibandingkan dengan data

crosssection murni dan *time series* murni. Penggunaan data panel dalam suatu model dapat membantu menjelaskan perbedaan antarindividu dalam satu waktu dan juga perbedaan perilaku suatu unit individu antarwaktu. Pada data panel, variabel yang digunakan memiliki analisis antarindividu dan antarwaktu yang ditandai oleh penggunaan indeks i untuk individu ($i= 1, \dots, N$) dan t untuk periode waktu ($t=1, \dots, T$). Dengan demikian, model yang dibangun dengan data panel dapat memberikan model yang lebih realistis daripada *cross section* murni dan *time series* murni [12].

Regresi data panel digunakan untuk menjawab tujuan penelitian tentang pengaruh infrastruktur jalan, pelabuhan, listrik, dan air bersih terhadap pertumbuhan ekonomi di pulau Sumatera. Melalui analisis tersebut, penelitian akan dikaji pada masing-masing provinsi di pulau Sumatera. Adapun langkah-langkah yang harus dilakukan dalam analisis regresi data panel adalah melakukan pemilihan estimasi model data panel terbaik dan melakukan analisis terhadap spesifikasi model yang terbentuk.

Pada penelitian ini yang pertama kali dilakukan dalam pemilihan model terbaik adalah pengujian stationeritas terhadap masing-masing variabel. Melalui pengujian stationeritas *Levin Lin & Chu* (Tabel 1), diperoleh variabel PDRB atas dasar harga konstan 2000, panjang jalan kondisi baik dan sedang, jumlah bongkar muat barang di pelabuhan, jumlah energi listrik yang didistribusikan, dan volume air bersih yang didistribusikan di Pulau Sumatera yang telah ditransformasi dalam bentuk *logaritma natural* adalah stationer pada level, kecuali jumlah energi listrik yang didistribusikan stationer pada different pertama.

Tabel 2. Uji Stasioneritas dengan Metode Levin, Lin & Chu

Variabel	Statistic	Prob*
(1)	(2)	(3)
LN (PDRB)	-4.5702	0.0000
LN (ROAD)	-8.4582	0.0000
LN (PORT)	-6.4978	0.0000
D (LN (ELECTR))	-9.3201	0.0000
LN (WATER)	-8.5214	0.0000

Estimasi Model Regresi Data Panel

Sehubungan dengan digunakannya data panel dalam penelitian ini, maka pengolahan dapat dilakukan dengan menggunakan tiga model, yaitu model pendekatan kuadrat kecil (*common effect*), pendekatan efek tetap (*fixed effect*), dan pendekatan efek acak (*random effect*).

Model *common effect* adalah salah satu tipe model data panel yang memiliki koefisien yang konstan untuk *intercept* dan *slope*. Untuk model data panel ini dapat menggunakan metode *Ordinary Least Squares Regression Model*. Model *common effect* dapat didefinisikan ke dalam model berikut:

$$y_{it} = \alpha + x'_{kit}\beta_k + \varepsilon_{it}$$

untuk $i= 1, 2, \dots, N$ dan $t = 1, 2, \dots, T$

dengan i menunjukkan urutan provinsi yang diobservasi pada data *cross section*, sedangkan t menunjukkan periode pada data *time-series*. Komponen *error* yang terdapat pada model ini hanyalah *error* secara keseluruhan (ε_{it}). Pada metode ini diasumsikan intersep dan koefisien dari setiap variabel sama untuk setiap provinsi yang diobservasi seperti model regresi OLS. Penggunaan asumsi ini secara logika

menyebabkan model *common effect* menjadi tidak cukup tepat menjelaskan model dalam penelitian terutama untuk melihat variasi antar provinsi dan waktu. Untuk itu dilakukan uji kesesuaian model dengan *Chow test* untuk menentukan apakah model yang digunakan adalah model *common effect*.

Model kedua adalah model *Fixed Effect* yang memiliki konstan slope namun memiliki intersep yang bergantung pada data panel dari serangkaian grup observasi (*cross section*) dalam hal ini adalah provinsi yang diobservasi. Model ini dikenal juga sebagai *Least Squares Dummy Variable (LSDV)* atau *covariance model*, karena sebanyak $i-1$ variabel *dummy* digunakan dalam model ini. Persamaan model ini adalah sebagai berikut:

$$y_{it} = \alpha_i + x'_{kit}\beta_k + \varepsilon_{it} \text{ dengan } \alpha_i = \alpha + D_i\alpha_i$$

Secara umum yang sering dilakukan pada model *fixed effect* yaitu dengan memasukkan variabel *dummy* (D_i) untuk menghasilkan nilai koefisien *slope* atau parameter yang berbeda-beda antar unit *cross section* [7]. Dengan demikian, persamaan dapat ditulis sebagai berikut:

$$y_{it} = \alpha + \sum_{i=1}^n D_i \alpha_i + x'_{kit}\beta_k + \varepsilon_{it}$$

Pada model *fixed effect* terdapat 3 metode estimasi yang dapat digunakan berdasarkan asumsi *varians-covarians* residualnya, yaitu:

1. *Ordinary Least Square (OLS/LSDV)*, jika struktur matriks *varians-covarians* residualnya diasumsikan bersifat homoskedastik dan tidak ada *cross sectional correlation*.
2. *Generalized Least Square (GLS)/Weighted Least Square (WLS)*, jika struktur matriks *varians-covarians* residualnya diasumsikan bersifat heteroskedastik dan tidak ada *cross sectional correlation*.
3. *Feasible Generalized Least Square (FGLS)/Seemingly Uncorrected Regression (SUR)* atau *Maximum Likelihood Estimator (MLE)*, jika struktur matriks *varians-covarians* residualnya diasumsikan bersifat heteroskedastik dan ada *cross sectional correlation*

Model terakhir yaitu model *Random Effect*. Dalam model ini terdapat perbedaan intersep untuk setiap provinsi dan intersep tersebut merupakan variabel *random* atau stokastik. Sehingga dalam model *random effects* terdapat dua komponen residual, yakni residual secara menyeluruh (ε_{it}) dan residual secara individu (u_i). Persamaan model *random effects* dapat ditulis sebagai berikut:

$$y_{it} = \alpha + x'_{kit}\beta_k + w_{it} \text{ dengan } w_{it} = u_i + \varepsilon_{it}$$

Gujarati [12] menjelaskan bahwa terdapat beberapa asumsi yang harus dipenuhi dalam model *random effect*. Secara matematis, asumsi-asumsi tersebut adalah:

$$u_i \sim N(0, \sigma_u^2) \quad \varepsilon_{it} \sim N(0, \sigma_\varepsilon^2) \quad E(u_i \varepsilon_{it}) = 0 \text{ untuk semua } i$$

$$E(u_i u_j) = 0 \text{ jika } i \neq j \quad E(\varepsilon_{it} \varepsilon_{is}) = E(\varepsilon_{it} \varepsilon_{js}) = 0 \text{ jika } i \neq j \text{ atau } t \neq s$$

Hal ini menunjukkan bahwa komponen error tidak berkorelasi satu sama lain dan tidak ada autokorelasi antara *cross section* dan *time series*. Komponen *error* w_{it} menjelaskan besarnya deviasi titik potong anggota panel dari nilai rata-rata. Komponen *error* ini tidak dapat diamati. Oleh sebab itu, asumsi sebelumnya juga harus mengikuti $E(w_{it}) = 0$ dan $\text{Var}(w_{it}) = \sigma_u^2 + \sigma_\varepsilon^2$, serta $E(w_{it} w_{is}) = \sigma_u^2$ untuk $t \neq s$. Meskipun komponen error w_{it} bersifat heteroskedastik, nyatanya terdapat korelasi antara w_{it} dan w_{is} , yaitu:

$$\text{Corr}(w_{it} w_{is}) = \frac{\sigma_u^2}{\sigma_\varepsilon^2 + \sigma_u^2}$$

Oleh karena itu, pada model ini metode OLS tidak bisa digunakan untuk mendapatkan estimator yang efisien. Metode yang tepat untuk mengestimasi model *random effect* adalah *Generalized Least Square* (GLS) dengan asumsi homoskedastik dan tidak ada *cross sectional correlation*.

Pemilihan Model Terbaik

Uji Chow atau Likelihood Ratio Test digunakan untuk mengetahui apakah model *fixed effects* lebih baik digunakan daripada model *common effects*. Dari pengujian tersebut dapat dilakukan dengan melihat signifikansi model *fixed effect* melalui uji statistik F. Hipotesisnya adalah:

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_n = 0 \text{ (common effect)}$$

$$H_1 : \text{minimal ada satu } \alpha_i \neq 0 \text{ (fixed effect)}$$

Adapun uji statistiknya adalah sebagai berikut [11]:

$$F_{\text{hitung}} = \frac{(R_{\text{FEM}}^2 - R_{\text{Pool}}^2)/(n-1)}{(1 - R_{\text{FEM}}^2)/(nT - n - k)} \sim F_{\alpha(n-1);(nT-n-k)}$$

Apabila nilai statistik F_{hitung} lebih besar dari nilai F_{tabel} pada tingkat signifikansi α , maka hipotesis null ditolak yang artinya teknik regresi data panel dengan *fixed effects* lebih baik daripada model regresi data panel dengan *common effects*.

Uji Breusch-Pagan *Lagrange Multiplier* (B-P LM) digunakan untuk mengetahui apakah model *random effects* lebih baik digunakan daripada model *common effect*. Model ini didasarkan pada nilai residual dari model *common effects*. Hipotesis *null* yang digunakan adalah intersep bukan merupakan variabel *random* atau stokastik. Dengan kata lain, varians dari efek spesifik individu α_i sama dengan nol, yakni $\sigma_u^2 = 0$ yang artinya model *common effects* lebih baik daripada baik daripada model *random effect*.

$$H_0 : \alpha_u = 0 \text{ (common effect)}$$

$$H_1 : \alpha_u \neq 0 \text{ (random effect)}$$

Adapun nilai statistik Breusch-Pagan *Lagrange Multiplier* (B-P LM) dihitung dengan formula sebagai berikut [11]:

$$LM = \frac{nT}{2(T-1)} \left[\frac{\sum_{i=1}^n (\sum_{t=1}^T \hat{\varepsilon}_{it})^2}{\sum_{i=1}^n \sum_{t=1}^T \hat{\varepsilon}_{it}^2} - 1 \right]^2$$

dengan n adalah jumlah individu, T adalah periode waktu, dan $\hat{\varepsilon}_{it}$ adalah residual/error term metode OLS.

Uji LM ini didasarkan pada distribusi *chi-square* dengan derajat bebas sebesar 1. Apabila hasil statistik LM lebih besar dari nilai kritis statistik *chi-square*, maka hipotesis *null* ditolak yang artinya estimasi yang tepat untuk regresi data panel tersebut adalah metode *random effects*.

Terkadang kita berhadapan dengan kondisi dimana *random effects model* maupun *fixed effect model* lebih baik dari *common effects model*. Untuk itu perlu dilakukan *Hausman Test* untuk menentukan apakah model lebih tepat dijelaskan dengan model *fixed effect* atau model *random effect*. *Hausman test* dilakukan dengan hipotesis sebagai berikut:

$H_0 : E(\epsilon_{it}|X_{it}) = 0$ (*random effect*)

$H_1 : E(\epsilon_{it}|X_{it}) \neq 0$ (*fixed effect*)

Nilai statistik Hausman ini mengikuti distribusi *Chi-Square* dengan formula sebagai berikut (Greene, 2003):

$$W = \chi^2 [k - 1] = \left[\mathbf{b} - \hat{\beta}_{GLS} \right] \hat{\Sigma}^{-1} \left[\mathbf{b} - \hat{\beta}_{GLS} \right]$$

Statistik uji Hausman tersebut mengikuti distribusi *chi-square* dengan derajat bebas sebanyak k (untuk k adalah variabel bebas). Apabila nilai statistik Hausman lebih besar daripada nilai kritis statistik *chi-square* maka hipotesis null akan ditolak yang artinya estimasi yang tepat digunakan untuk regresi data panel adalah *fixed effects*. Kelemahan dari uji Hausman ini adalah perbedaan matriks kovarian memungkinkan diperoleh hasil yang tidak definit positif sehingga dapat disimpulkan bahwa hipotesis *null* tidak ditolak.

Selanjutnya, untuk model estimasi regresi data panel terpilih akan dilanjutkan pengujian untuk memilih estimator dengan struktur *varians covarians* dari *residual* yang lebih baik. Namun, jika terpilih model *random effects* maka pengujian untuk struktur *varians-covarians residual* tidak perlu dilakukan.

Masalah lain dalam estimasi pada analisis regresi data panel adalah menentukan struktur *varians-covarians* dari *residual* yang lebih baik. Terdapat beberapa metode yang sesuai dengan asumsi struktur heteroskedastik dan tidak ada *cross sectional correlation*, asumsi struktur heteroskedastik dan ada *cross sectional correlation* (*Seemingly Uncorrelated Regression/SUR*), dan asumsi struktur adanya autokorelasi antar waktu pada *error term*.

Pengujian asumsi di atas berbeda dengan pengujian dalam persamaan tunggal, dimana dalam analisis persamaan tunggal, dilakukan pengujian apakah terjadi gejala heteroskedastik ataukah autokorelasi untuk satu individu. Selanjutnya, dari hasil pengujian tersebut, dilakukan perbaikan (*remidial*) model agar didapatkan hasil estimasi yang BLUE. Namun, dalam analisis data panel pengujian dilakukan untuk menentukan estimator manakah yang lebih baik untuk melakukan estimasi. Estimator tersebut disesuaikan dengan kondisi matriks *varians-covarians residual*.

Uji LM dilakukan apabila model yang terpilih adalah *fixed effect model*. Pengujian ini berguna untuk melihat struktur *varians-covarians residual*, apakah bersifat homoskedastis atau heteroskedastis. LM *test* dilakukan dengan hipotesis sebagai berikut:

$H_0: \sigma_i^2 = \sigma^2$ (Struktur *varians-kovarians residual* homoskedastis)

$H_1: \sigma_i^2 \neq \sigma^2$ (Struktur *varians-kovarians residual* heteroskedastis)

Secara matematis, statistik uji yang digunakan dapat dirumuskan sebagai berikut:

$$LM = \frac{T}{2} \sum_{i=1}^N \left[\frac{\hat{\sigma}_i^2}{\hat{\sigma}^2} - 1 \right]^2$$

dengan T adalah jumlah observasi, n adalah jumlah individu, $\hat{\sigma}_i^2$ adalah *varians residual* persamaan ke- i pada kondisi homoskedastis, dan $\hat{\sigma}^2$ adalah *sum square residual* persamaan sistem pada kondisi homoskedastis.

Statistik uji LM ini mengikuti distribusi statistik *chi-square* dengan derajat bebas sebanyak $n-1$. Jika nilai statistik LM lebih besar dari nilai kritis statistik *chi square*, maka hipotesis *null* akan ditolak, yang berarti struktur *varians-covarians residual* bersifat heteroskedastis.

Uji λ LM dilakukan apabila hasil pengujian LM menunjukkan bahwa

struktur *varians-covarians residual* bersifat heteroskedastis. Pengujian ini dilakukan dengan hipotesis sebagai berikut:

H_0 : $\text{corr}(\varepsilon_i, \varepsilon_j) = 0$ (tidak ada *cross sectional correlation*)

H_1 : $\text{corr}(\varepsilon_i, \varepsilon_j) \neq 0$ (ada *cross sectional correlation*)

Secara matematis, statistik uji yang digunakan dapat dirumuskan sebagai berikut:

$$\lambda\text{LM} = T \sum_{i=2}^n \sum_{j=1}^{i-1} r_{ij}^2$$

dengan T adalah jumlah observasi, n adalah jumlah individu, r_{ij}^2 adalah *residual correlation coefficient* antara persamaan ke i dan ke j. Ekspresi $\sum_{i=2}^n \sum_{j=1}^{i-1} r_{ij}^2$ menunjukkan jumlah diatas atau di bawah diagonal dari *residual correlation coefficient* pada kondisi persamaan yang diuji untuk melihat apakah struktur off diagonal matriks residual = 0.

Statistik uji λLM ini mengikuti distribusi statistik chi square dengan derajat bebas sebanyak $n(n-1)/2$. Jika nilai statistik LM lebih besar dari nilai kritis statistik *chi-square*, maka hipotesis *null* akan ditolak, yang berarti struktur *varians-covarians residual* bersifat heteroskedastik dan ada *cross sectional correlation* (*Seemingly Uncorrelated Regression/SUR*).

Tabel 2 berikut adalah hasil uji yang dilakukan pada tahapan pemilihan model terbaik dalam penelitian ini. Pertama, uji Chow (Uji F) yang digunakan untuk melihat apakah model yang terbaik *fixed effect model* atau *common effect model*. Dengan tingkat kepercayaan 95 persen diperoleh hasil bahwa model yang terpilih adalah *fixed effect model*. Kedua, uji *Hausman* digunakan untuk memilih mana model yang terbaik antara *fixed effect model* dan *random effect model*. Dengan tingkat kepercayaan 95 persen diperoleh hasil bahwa model yang terpilih adalah *fixed effect model*. Sehingga model estimasi terbaik untuk menganalisis pengaruh infrastruktur jalan, pelabuhan, listrik dan air bersih terhadap pertumbuhan ekonomi Pulau Sumatera adalah model *fixed effect*.

Tabel 2. Hasil Tahapan Uji dalam Pemilihan Model Terbaik

Nama Uji	Nilai Statistik	Prob*	Keputusan
(1)	(2)	(3)	(4)
Uji Chow	F Hitung 67.880	0.0000	FEM lebih baik dari CEM
Uji Hausman	χ^2 Hitung 32.152	0.0000	FEM lebih baik dari REM
Uji LM	LM 488.963	0.0000	Struktur varians-kovarians heteroskedastik
uji λLM	λLM 145.996	0.0000	heteroskedastik dan ada <i>cross sectional correlation</i> (SUR).

Setelah mendapatkan model *fixed effect*, maka pengujian selanjutnya adalah pengujian asumsi struktur *varians-covarians residual* dengan uji *Lagrange Multiplier (LM)*. Dengan tingkat kepercayaan 95 persen diperoleh hasil bahwa pada persamaan struktur *varians-covarians residualnya* adalah heteroskedastik. Karena struktur *varians-covarians residualnya* bersifat heteroskedastik, maka perlu dilakukan pemilihan estimator struktur heteroskedastik dan tidak ada *cross sectional correlation* atau *Seemingly Uncorrelated Regression (SUR)* dengan uji λLM . Hasil yang diperoleh dari persamaan tersebut dengan tingkat kepercayaan 95

persen adalah struktur *varians-covarians residualnya* bersifat heteroskedastik dan ada *cross sectional correlation* atau *SUR*.

Setelah melalui beberapa tahap, maka model terbaik yang terpilih adalah model *fixed effect* dengan struktur *varians-covarians residual* bersifat heteroskedastik dan ada *cross sectional correlation* (*Seemingly Uncorelated Regression/SUR*). Untuk menghasilkan estimasi yang bersifat *BLUE* (*Best Linear Unbiased Estimator*) dari model terbaik tersebut, maka pengujian asumsi klasik yang perlu dilakukan yaitu pengujian normalitas, dan non multikolinieritas. Karena model yang digunakan *fixed effect* dengan struktur *varians-covarians residual* bersifat heteroskedastik dan ada *cross sectional correlation* (*Seemingly Uncorelated Regression/SUR*), maka sudah mengakomodir asumsi homoskedastik dan nonautokorelasi. Untuk pengujian asumsi normalitas dilakukan dengan melihat nilai statistik uji *Jarque Berra* residualnya dan hasilnya menyatakan bahwa asumsi telah terpenuhi. Pengujian terhadap non multikolinieritas dilakukan dengan melihat nilai VIF masing-masing variabel. Hasil nilai VIF masing-masing variabel diperoleh $VIF < 10$, sehingga dapat disimpulkan bahwa model tersebut memenuhi asumsi nonmultikolinieritas. Dari uraian di atas, maka model telah melewati tahapan pengujian asumsi dan hasilnya telah memenuhi semua asumsi, sehingga model tersebut dapat diinterpretasikan.

2.4 Persamaan Regresi Data Panel

Peneliti menggunakan persamaan yang mencakup 10 wilayah pengamatan, yaitu Provinsi Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Kepulauan Bangka Belitung, dan Kepulauan Riau yang masing-masing wilayah mencakup 10 tahun pengamatan (tahun 2004 sampai 2013). Variabel yang digunakan untuk persamaan tersebut adalah PDRB Atas Dasar Harga Konstan (dalam milyar Rp), panjang jalan kondisi baik dan sedang (dalam km), jumlah energi listrik yang disalurkan pada pelanggan (dalam Gwh), dan volume air bersih yang disalurkan pada pelanggan (dalam m³). Dalam proses pengolahan, semua variabel diubah ke dalam bentuk *logaritma natural*.

Model estimasi pengaruh infrastruktur jalan, pelabuhan, listrik, dan air bersih terhadap pertumbuhan ekonomi di pulau Sumatera secara ringkas dapat ditulis dalam bentuk persamaan sebagai berikut:

$$\ln\widehat{PDRB}_{it} = 3,9222* + \hat{\gamma}_i + 0,2771 \lnROAD_{it}* + 0,1529 \lnPORT_{it}* + 0,0610\lnELECTR_{it}+0,1514\lnWATER_{it}*$$

* signifikansi pada taraf uji 5 persen

Keterangan:

- PDRB : PDRB atas dasar harga konstan 2000 (Milyar Rupiah)
- ROAD : Panjang jalan dalam kondisi baik dan sedang (Km)
- PORT : Jumlah bongkar dan muat barang di pelabuhan (Ton)
- ELECTR : Jumlah energi listrik yang disalurkan pada pelanggan (Gwh)
- WATER : Volume air bersih yang disalurkan pada pelanggan (Ribu m³)

Pada saat variabel lain *ceteris paribus*, pertumbuhan ekonomi masing-masing provinsi di Pulau Sumatera dapat dilihat berdasarkan nilai α_i (individual effect). Terlihat pada Tabel 3 pertumbuhan ekonomi provinsi-provinsi di Pulau Sumatera dari yang tertinggi sampai yang terendah yaitu Provinsi Riau, Sumatera Utara, Sumatera Selatan, Kepulauan Riau, Aceh, Lampung, Sumatera Barat,

Kepulauan Bangka Belitung, Jambi, dan Bengkulu. Hal ini sesuai dengan pembahasan sebelumnya yang menyatakan bahwa PDRB Provinsi Riau dan Sumatera Utara berkontribusi besar terhadap PDRB Pulau Sumatera, sedangkan PDRB Provinsi Bengkulu merupakan kontribusinya paling kecil terhadap PDRB Pulau Sumatera.

Tabel 3. Koefisien persamaan regresi data panel di Pulau Sumatera

Provinsi	$\alpha_i = \alpha + \gamma_i$ (Individual Effect)
(1)	(2)
Aceh	4,015829
Sumatera Utara	4,449448
Sumatera Barat	3,833661
Riau	4,665978
Jambi	3,355053
Sumatera Selatan	4,311837
Bengkulu	3,048193
Lampung	3,913637
Kep. Bangka Belitung	3,423705
Kepulauan Riau	4,205239

Tabel 4. Ringkasan statistik hasil estimasi model *fixed effect cross sectional weight with cross sectional/SUR* (PCSE)

Variabel	Koefisien	t-hitung	Prob.
(1)	(2)	(3)	(4)
C	3,9222	5,0715*	0,0000
LOG(ROAD?)	0,2771	6,5374*	0,0000
LOG(PORT?)	0,1529	5,0329*	0,0000
D(LOG(ELECTR?))	0,0160	0,2196	0,8267
LOG(WATER?)	0,1514	3,0578*	0,0031
<i>R-square</i>	0,9895		
<i>Adj R-square</i>	0,9877		
F-hitung	553,1052*		
Prob (F-hitung)	0,0000		

* signifikansi pada taraf uji 5 persen

Hasil dari estimasi model terbaik menunjukkan bahwa nilai F-statistik adalah 553,1052 dan signifikan pada taraf nyata 5 persen (Tabel 4), yang menunjukkan variabel-variabel penjelas yang digunakan pada persamaan regresi data panel yaitu, panjang jalan kondisi baik dan sedang, jumlah bongkar muat barang di pelabuhan, jumlah energi listrik yang disalurkan, dan volume air bersih yang didistribusikan secara simultan (bersama-sama) signifikan dalam memengaruhi variabel tak bebas yaitu pertumbuhan ekonomi di Pulau Sumatera. Nilai *Adjusted R²* dari hasil pengolahan data tersebut mencapai 0,9877, artinya, besarnya proporsi variasi dari variabel tak bebas yang bisa dijelaskan oleh variabel bebas pada model tersebut sebesar 98,77 persen, sedangkan sisanya sebesar 1,23 persen dijelaskan oleh variabel lain di luar model. Kemudian, berdasarkan

pengujian secara parsial diperoleh bahwa dengan tingkat signifikansi 5 persen, variabel panjang jalan dengan kondisi baik dan sedang, jumlah bongkar dan muat barang dipelabuhan, dan volume air bersih yang disalurkan pada pelanggan signifikan secara statistik memengaruhi pertumbuhan ekonomi di Pulau Sumatera. Dari keempat variabel infrastruktur, variabel infrastruktur yang paling besar pengaruhnya terhadap pertumbuhan ekonomi pulau Sumatera adalah panjang jalan, dengan elastisitas sebesar 0,2771. Hal ini menandakan bahwa panjang jalan berpengaruh positif terhadap pertumbuhan ekonomi di pulau Sumatera. Peningkatan panjang jalan sebesar 1 persen akan meningkatkan pertumbuhan ekonomi sebesar 0,2771 persen dengan kondisi variabel lain *ceteris paribus*. Pernyataan ini sesuai dengan hipotesis awal yang menyatakan bahwa panjang jalan akan berhubungan positif dan berpengaruh secara signifikan terhadap pertumbuhan ekonomi. Hal ini dikarenakan panjang jalan mempunyai peranan penting dalam kegiatan perekonomian. Distribusi faktor produksi maupun barang dan jasa hasil produksi sangat tergantung dari keberadaan infrastruktur jalan. Secara spasial, mobilitas manusia dan hasil produksi menentukan kemajuan suatu wilayah karena interaksi dan keterbukaan dengan wilayah lain meningkatkan pangsa pasar baik faktor produksi maupun hasil produksi.

Infrastruktur pelabuhan juga berpengaruh signifikan dan positif terhadap pertumbuhan ekonomi di pulau Sumatera. Elastisitasnya yang bernilai 0,1529 menunjukkan bahwa apabila jumlah bongkar dan muat barang di pelabuhan meningkat sebesar 1 persen maka akan meningkatkan pertumbuhan ekonomi sebesar 0,1529 persen dengan kondisi variabel lain *ceteris paribus*. Hasil penelitian ini sama dengan penelitian yang dilakukan oleh Maryaningsih et al. [14]. Dalam hasil penelitiannya menyatakan bahwa penyediaan infrastruktur pelabuhan yang cukup, merupakan prakondisi untuk mencapai pertumbuhan ekonomi yang tinggi dan berkesinambungan. Selain itu, dengan adanya infrastruktur pelabuhan dapat melancarkan transportasi pengiriman hasil produksi barang atau jasa antar pulau maupun luar negeri.

Infrastruktur listrik tidak berpengaruh terhadap pertumbuhan ekonomi di Pulau Sumatera pada tingkat signifikansi 5 persen. Hal ini disebabkan ketersediaan listrik di Pulau Sumatera belum memenuhi besarnya kebutuhan energi listrik para pelanggan, khususnya untuk kebutuhan industri. Industri merupakan sektor yang mempunyai kontribusi terbesar terhadap PDRB di Pulau Sumatera, tentunya energi listrik akan sangat berperan penting dalam proses produksi di suatu industri. Namun, rata-rata penggunaan listrik yang disalurkan oleh PLN (Perusahaan Listrik Negara) untuk bidang industri tiap tahunnya hanya mencapai 18,67 persen. Hal ini menunjukkan bahwa listrik PLN yang digunakan untuk kegiatan industri masih sedikit, sehingga instansi atau perusahaan yang belum terpenuhi seluruh kebutuhan listriknya mengupayakan pembangkit listrik sendiri secara terbatas (*captive power*). Dengan adanya *captive power* tersebut, jumlah energi listrik dari PLN tidak berpengaruh terhadap pertumbuhan ekonomi di Pulau Sumatera.

Infrastruktur air bersih berpengaruh signifikan dan berhubungan positif terhadap pertumbuhan ekonomi. Elastisitas dari variabel air bersih yang bernilai 0,1514 menunjukkan bahwa apabila volume air bersih yang didistribusikan meningkat sebesar 1 persen, maka akan meningkatkan pertumbuhan ekonomi sebesar 0,1514 persen. Dengan melihat hasil estimasi tersebut, maka dapat disimpulkan bahwa ketersediaan infrastruktur air bersih akan memberikan manfaat bagi masyarakat untuk melakukan kegiatan sosial maupun ekonomi, sehingga hal tersebut dapat menunjang pertumbuhan ekonomi di Pulau Sumatera.

Uraian pengaruh keempat infrastruktur yang diteliti terhadap pertumbuhan ekonomi di Pulau Sumatera menunjukkan bahwa infrastruktur berpengaruh secara positif. Hal ini disebabkan karena infrastruktur merupakan konektivitas untuk mendukung pengembangan kegiatan ekonomi utama di Pulau Sumatera. Menurut MP3E1 [10], kegiatan ekonomi utama di pulau Sumatera antara lain: kelapa sawit, karet, batu bara, perkapalan, besi baja dan Kawasan Strategis Nasional (KSN) selat Sunda. Hasil ini juga sesuai dengan pertumbuhan ekonomi menurut Todaro dan Smith [17], yang menyatakan bahwa pertumbuhan ekonomi salah satunya berasal dari akumulasi modal termasuk didalamnya semua investasi dan modal fisik seperti panjang jalan, jumlah bongkar muat barang dipelabuhan, energi listrik dan volume air bersih.

3. Kesimpulan

Berdasarkan hasil-hasil di atas diperoleh kesimpulan sebagai berikut:

1. Selama periode 2004-2013, pertumbuhan ekonomi di Pulau Sumatera cenderung berfluktuatif. Rata-rata pertumbuhan ekonomi di Pulau Sumatera mencapai 4,79 persen per tahun. Provinsi yang berkontribusi paling tinggi terhadap PDRB Pulau Sumatera adalah Sumatera Utara dan Riau karena provinsi tersebut merupakan daerah yang terkonsentrasikan pada kegiatan produksi yang cukup tinggi, sedangkan provinsi yang berkontribusi paling rendah terhadap PDRB Pulau Sumatera adalah Provinsi Bengkulu. Provinsi yang mengalami perkembangan infrastruktur tercepat di koridor Sumatera adalah Sumatera Utara dan Riau, sedangkan provinsi lain mengalami perkembangan infrastruktur yang masih lambat. Kondisi tersebut menunjukkan belum meratanya ketersediaan infrastruktur jalan, pelabuhan, listrik, dan air bersih di Pulau Sumatera pada periode 2004-2013.
2. Model estimasi data panel terbaik adalah model *fixed effect* dengan struktur varians kovarians residual bersifat heteroskedastik dan ada *cross sectional correlation* (*Seemingly Uncorelated Regression* /SUR). Dari persamaan regresi data panel, infrastruktur jalan, pelabuhan, dan air bersih terbukti secara signifikan berpengaruh positif terhadap pertumbuhan ekonomi di Pulau Sumatera, sedangkan infrastruktur listrik tidak berpengaruh terhadap pertumbuhan ekonomi di Pulau Sumatera.

Referensi

- [1] Badan Pusat Statistik [BPS]. Berbagai Edisi. *Statistik Air Bersih*. Jakarta: BPS.
- [2] Badan Pusat Statistik [BPS]. Berbagai Edisi. *Statistik Indonesia*. Jakarta: Badan Pusat Statistik.
- [3] Badan Pusat Statistik [BPS]. Berbagai Edisi. *Statistik Listrik*. Jakarta: Badan Pusat Statistik.
- [4] Badan Pusat Statistik [BPS]. Berbagai Edisi. *Statistik Transportasi*. Jakarta: Badan Pusat Statistik.
- [5] Badan Pusat Statistik [BPS]. Berbagai Edisi. *Statistik Perhubungan*. Jakarta: Badan Pusat Statistik.
- [6] Badan Pusat Statistik [BPS]. Berbagai Edisi. *Produk Domestik Regional Bruto Provinsi-Provinsi di Indonesia Menurut Lapangan Usaha (PDRB)*. Jakarta: Badan Pusat Statistik.

- [7] Baltagi, Badi H. 2005. *Econometric Analysis of Panel Data*. New York: John Wiley & Sons.
- [8] Bappenas. 2012. *Pembangunan Daerah Dalam Angka 2013*. Jakarta: Kementerian Perencanaan Pembangunan Nasional/Badan Perencanaan Pembangunan Nasional.
- [9] Basri, Faisal H. 2002. *Perekonomian Indonesia: Tantangan dan Harapan bagi Kebangkitan Indonesia*. Jakarta: Erlangga.
- [10] Coordinating Misnistry For Economic Affairs. 2011. *Masterplan Acceleration and Expansion of Indonesia Economix Development 2011-2025 Republik Indonesia*. Jakarta, Coordinating Misnistry For Economic Affairs.
- [11] Greene, William H. 2003. *Econometric Analysis (Fifth Edition)*. New Jersey: Prentice Hall
- [12] Gujarati, Damodar N and Dawn C. Porter. 2012. *Dasar-Dasar Ekonometrika Buku 2*. Edisi Kelima. R. Carlos Mangungsong [Penerjemah]. Salemba Empat: Jakarta.
- [13] Mankiw, N. Gregory. 2007. *Makroekonomi*. Edisi Keenam. Fitria Liza dan Imam Nurmawan [Penerjemah]. Jakarta: Erlangga.
- [14] Maryaningsih, Novi, dkk. 2012. Pengaruh Infrastruktur terhadap Pertumbuhan Ekonomi Indonesia. Bank Indonesia. Diakses pada tanggal 28 Desember 2014 melalui <http://www.bi.go.id/id/publikasi/wp/Pages/Pengaruh-Infrastruktur-terhadap-Pertumbuhan-Ekonomi-Indonesia.aspx>
- [15] Nachrowi, Djalal Nachrowi dan Hardius Usman. 2006. *Pendekatan Populer dan Praktis Ekonometrika untuk Analisis Ekonomi dan Keuangan*. Jakarta: Fakultas Ekonomi Universitas Indonesia.
- [16] Neter, J., Wasserman W., dan Kutner M.H. 1989. *Applied Linier egression Model Second Edition*. United State of Amerika.
- [17] Todaro, Michael P. dan Stephen C. Smith. 2012. *Economic Development*. 11th Edition. New York: Edison-Wesley.

KONVERGENSI EKONOMI ANTARKABUPATEN/KOTA DI PROVINSI PAPUA BARAT

I MADE WAHYU DWI SEPTIKA¹, DEWI PURWANTI²

1 Badan Pusat Statistik, wahyuseptika10@gmail.com

2 Sekolah Tinggi Ilmu Statistik, dewip@stis.ac.id

Abstrak. Provinsi Papua Barat mengalami pertumbuhan ekonomi yang cukup baik selama penyelenggaraan kebijakan otonomi khusus. Akan tetapi, masih terjadi ketimpangan PDRB per kapita yang tinggi antarkabupaten/kota. Padahal, terjadinya pertumbuhan ekonomi yang tinggi diharapkan lebih disebabkan oleh peningkatan perekonomian wilayah miskin yang mampu mengurangi *gap* PDRB per kapita dengan wilayah kaya sehingga terjadi konvergensi ekonomi. Tujuan dari penelitian ini adalah mengidentifikasi terjadinya konvergensi ekonomi antarkabupaten/kota di Provinsi Papua Barat serta menganalisis variabel-variabel yang mempercepat terjadinya konvergensi ekonomi. Data yang digunakan bersumber dari Badan Pusat Statistik (BPS) dengan struktur data panel, yaitu 11 kabupaten/kota di Provinsi Papua Barat selama periode 2010-2014. Penelitian ini menggunakan analisis deskriptif dan analisis inferensia regresi data panel. Hasil penelitian ini menunjukkan bahwa terjadi konvergensi ekonomi antarkabupaten/kota di Provinsi Papua Barat yang diidentifikasi melalui konvergensi sigma maupun konvergensi beta. Variabel yang berpengaruh positif dan signifikan terhadap pertumbuhan ekonomi dalam upaya mempercepat konvergensi ekonomi antarkabupaten/kota di Provinsi Papua Barat, antara lain: persentase penduduk yang bekerja terhadap angkatan kerja, angka harapan hidup, dan persentase penduduk berumur 10 tahun ke atas yang berpendidikan minimal SMA. Sementara itu, dana otonomi khusus dan penyesuaian tidak berpengaruh signifikan dalam upaya mempercepat konvergensi ekonomi.

Kata kunci: Papua Barat, otonomi khusus, *gap* PDRB, konvergensi sigma, konvergensi beta

1. Pendahuluan

Pasca reformasi 1998 terdapat delapan provinsi baru hasil pemekaran wilayah di Indonesia tetapi hanya Provinsi Papua Barat yang menyelenggarakan kebijakan otonomi khusus. Realisasi Dana otonomi khusus dan penyesuaian (DOKP) yang diperoleh oleh Provinsi Papua Barat pada tahun 2010 sebesar 960 miliar rupiah. Nilai tersebut terus meningkat hingga mencapai 1,95 triliun rupiah pada tahun 2014 (BPS [1]). Menurut Huda [2], pemberian kebijakan otonomi khusus bagi Provinsi Papua Barat sebagai upaya percepatan pembangunan ekonomi dalam rangka mewujudkan kesetaraan dengan kemajuan daerah lain maupun mengatasi ketimpangan PDRB per kapita antarkabupaten/kota yang terjadi di wilayah ini.

Percepatan pertumbuhan ekonomi suatu daerah bergantung pada faktor produksi yang digunakan sebagai input dalam proses menghasilkan output. Menurut Irawan dan Suparmoko [3], semakin banyak jumlah penduduk yang terlibat sebagai tenaga kerja dalam kegiatan produksi maka akan meningkatkan jumlah produksi yang mendorong terjadinya pertumbuhan ekonomi. Kuantitas

tenaga kerja harus diimbangi dengan kualitas yang memadai karena perbedaan kualitas tenaga kerja menyebabkan perbedaan produktivitas. Tjokroamidjojo dan Mustopadidjaja [4] menyatakan bahwa produktivitas tenaga kerja dapat ditingkatkan dengan meningkatkan mutu sumber daya manusia dengan melakukan investasi modal manusia melalui dimensi pendidikan, keterampilan, dan kesehatan.

Pendidikan berperan penting dalam hal kemampuan suatu perekonomian mengadopsi perkembangan teknologi dan kecakapan manajerial untuk kegiatan pembangunan berkelanjutan. Peningkatan pendidikan bagi tenaga kerja akan meningkatkan keterampilan pekerja yang pada akhirnya akan meningkatkan output yang dihasilkan. Dimensi lain yang perlu diperhatikan dalam investasi modal manusia adalah berkaitan dengan kesehatan masyarakat. Kesehatan merupakan hal yang amat penting dalam kehidupan masyarakat. Menurut Todaro [5], peningkatan derajat kesehatan penduduk mengakibatkan masyarakat memiliki umur panjang dan hidup sehat yang secara langsung memperbaiki kualitas tenaga kerja sehingga meningkatkan kemampuan belajar dan hasil kerja dalam perekonomian.

Berdasarkan data BPS [6], perekonomian Provinsi Papua Barat tumbuh cukup baik pada periode 2011-2014. Laju pertumbuhan ekonomi Provinsi Papua Barat pada tahun 2011 dan 2012 hanya sebesar 3,64 persen dan 3,63 persen. Selanjutnya, perekonomian Provinsi Papua Barat tumbuh cepat hingga mencapai sebesar 7,36 persen pada tahun 2013 hingga berada di atas laju pertumbuhan ekonomi nasional. Pada tahun 2014, laju pertumbuhan ekonomi Provinsi Papua Barat hanya mencapai sebesar 5,38 persen. Diharapkan, terjadinya laju pertumbuhan ekonomi yang tinggi di Provinsi Papua Barat lebih disebabkan oleh adanya peningkatan perekonomian daerah miskin sehingga mampu mengurangi *gap* PDRB per kapita dengan wilayah maju dan menurunkan ketimpangan antarwilayah.

Akan tetapi, pertumbuhan ekonomi yang tinggi di Provinsi Papua Barat ternyata masih dibarengi dengan ketidakmerataan pendapatan antarkabupaten/kota di dalamnya. Nilai Indeks Williamson sebagai ukuran ketimpangan PDRB per kapita antarkabupaten/kota di Provinsi Papua Barat selama periode 2010-2014 selalu berada di atas nilai 0,8. Nilai Indeks Williamson yang mendekati satu menunjukkan terjadi ketimpangan PDRB per kapita yang tinggi di Provinsi Papua Barat. Dalam jangka panjang, ketimpangan pendapatan antarwilayah tersebut akan memperparah kondisi kemiskinan suatu daerah.

Pelaksanaan kebijakan otonomi khusus diharapkan secara cepat mengurangi *gap* PDRB per kapita antarkabupaten/kota di Provinsi Papua Barat. Oleh karena itu, pemberian DOKP mendorong terjadinya konvergensi ekonomi antarkabupaten/kota, yaitu mendorong wilayah dengan perekonomian tertinggal atau PDRB per kapita yang rendah untuk tumbuh lebih cepat dari wilayah kaya sehingga dapat mengusahakan tingkat PDRB per kapita antarwilayah yang cenderung seragam.

Teori konvergensi menyatakan bahwa tingkat pendapatan per kapita yang dialami oleh daerah maju dan daerah berkembang dan tertinggal pada suatu saat akan konvergen (*convergence*). Hal ini didasarkan bahwa daerah maju akan mengalami kondisi *steady state*, yaitu penambahan investasi tidak menambah pendapatan. Sementara daerah maju mengalami kondisi *steady state*, daerah berkembang dan tertinggal memiliki tambahan investasi sehingga menambah pendapatan dan pada akhirnya suatu saat akan menyamai pendapatan daerah maju (*catching up effect*).

Tujuan Penelitian

Tujuan dari dilakukannya penelitian ini adalah:

1. Memperoleh gambaran umum perekonomian Provinsi Papua Barat periode 2010-2014.
2. Mengidentifikasi terjadinya konvergensi ekonomi antarkabupaten/kota di Provinsi Papua Barat periode 2010-2014.
3. Menganalisis variabel yang mendorong terjadinya konvergensi ekonomi antarkabupaten/kota di Provinsi Papua Barat periode 2010-2014.

2. Hasil-Hasil Utama

Pertumbuhan Ekonomi Neo-Klasik

Pertumbuhan ekonomi menjadi salah satu ukuran dalam pembangunan ekonomi. Penambahan produksi yang berkesinambungan pada periode tertentu di suatu negara menunjukkan kondisi perekonomian yang semakin membaik. Teori Pertumbuhan Neo-Klasik melihat suatu perekonomian dari sudut pandang penawaran atau dikenal sebagai Model Solow-Swan.

Menurut Todaro dan Smith [7], model pertumbuhan Solow berasumsi bahwa perekonomian dengan skala hasil yang terus berkurang (*diminishing return to scale*) dari input tenaga kerja dan modal jika kedua variabel tersebut dianalisis secara terpisah. Namun, jika keduanya dianalisis secara bersamaan maka Solow juga mengasumsikan *constant return to scale*. Fungsi produksi model pertumbuhan Solow-Swan dapat dinyatakan dengan persamaan:

$$Y = K^{\alpha}(AL)^{1-\alpha}, \quad (1)$$

keterangan :

Y : tingkat pertumbuhan ekonomi

K : tingkat pertumbuhan modal

L : tingkat pertumbuhan tenaga kerja

A: produktivitas tenaga kerja.

Berdasarkan asumsi *constant returns to scale*, dalam jangka panjang fungsi produksi ini dapat menganalisis perbandingan variabel ekonomi dengan jumlah pekerja sehingga fungsi produksi neo-klasik menjadi:

$$y=f(k), \quad (2)$$

Kemiringan fungsi produksi Solow-Swan digambarkan oleh *marginal product of capital* ($MPK = f'(k) = f'(k+1) - f'(k)$). MPK menunjukkan banyaknya output tambahan yang dihasilkan seorang pekerja ketika mendapatkan satu unit modal tambahan. Menurut Mankiw [8], ketika modal bertambah menyebabkan rata-rata pekerja memiliki banyak modal dalam proses produksi sehingga satu unit modal tambahan hanya sedikit meningkatkan produksi. Wilayah dengan modal perekonomian yang besar seperti negara maju akan memiliki MPK yang lebih rendah dibanding negara berkembang. Hal ini menjelaskan tingkat pertumbuhan ekonomi wilayah yang sedang berkembang lebih tinggi dibandingkan dengan pertumbuhan ekonomi wilayah maju.

Menurut Snowdon dan Vane [9], model ini memprediksi bahwa nantinya akan terjadi perekonomian yang menuju *steady-state* yang menuju titik

ekuilibrium. Titik ekuilibrium perekonomian jangka panjang akan tercapai apabila kapital per tenaga kerja mencapai suatu tingkat yang stabil. Rasio kapital terhadap tenaga kerja dipengaruhi oleh tabungan dan depresiasi atau penyusutan.

$$\Delta k = sf(k) - \delta k, \quad (3)$$

keterangan:

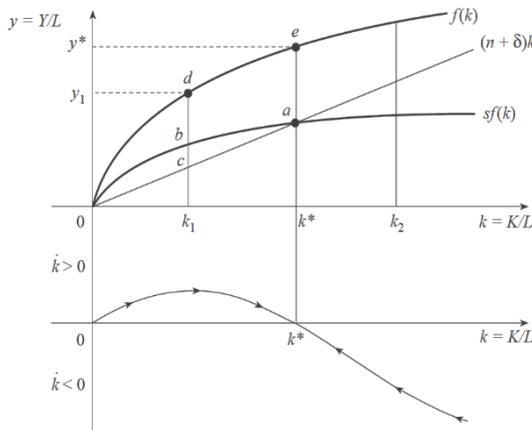
Δk : Perubahan persediaan modal terhadap tenaga kerja antar periode

$s[f(k)]$: Tabungan

δk : Depresiasi

Menurut Fungsi $s[f(k)]$ tidak hanya menjelaskan mengenai tabungan saja, tetapi juga menjelaskan investasi tenaga kerja yang diperlukan agar rasio kapital terhadap tenaga kerja tidak berubah. Kondisi *steady-state* berdasarkan tingkat investasi dan tingkat depresiasi tertentu akan terjadi dalam keadaan:

$$sf(k^*) = \delta k^*, \quad (4)$$



Sumber: Snowdon & Vane (2005)

Berdasarkan Gambar 2, kondisi *steady-state* adalah suatu keadaan ketika rasio kapital terhadap tenaga kerja konstan karena investasi per tenaga kerja sama dengan jumlah depresiasi sehingga hanya mampu menutupi tingkat penyusutan per tenaga kerja. Jika variabel populasi dimasukkan ke dalam model pertumbuhan neo-klasik maka pertumbuhan populasi dapat mengurangi pertumbuhan ekonomi

karena dipandang sama dengan tingkat penyusutan atau depresiasi.

Secara formula hal ini dijelaskan sebagai berikut:

$$\Delta k = sf(k) - (n + \delta)k, \quad (5)$$

Konvergensi Ekonomi Antarwilayah

Teori konvergensi menyatakan bahwa tingkat pendapatan per kapita yang dialami oleh daerah maju dan daerah berkembang dan tertinggal pada suatu saat akan konvergen. Kondisi perekonomian yang konvergen adalah kondisi daerah miskin untuk memacu pertumbuhannya sehingga dapat mengejar ketertinggalan dari perekonomian daerah kaya. Teori pertumbuhan ekonomi Solow-Swan menjadi

kerangka dasar konvergensi ekonomi.

Analisis pada penelitian ini menggunakan bantuan *software Microsoft Excel 2010* dan *Eviews 8*. Analisis konvergensi menurut Barro dan Sala-i-Martin [10]) dilihat dari dua ukuran pendekatan, yaitu konvergensi sigma dan konvergensi beta. Konvergensi sigma adalah analisis disparitas pendapatan antarwilayah yang dihitung melalui ukuran dispersi berupa koefisien variasi dari logaritma natural PDRB per kapita kabupaten/kota di Provinsi Papua Barat dengan formula sebagai berikut:

$$KV_t = S_t / \bar{x}_t, \quad (6)$$

keterangan:

KV_t : Koefisien Variasi pada tahun ke-t.

S_t : Standar deviasi logaritma natural PDRB per kapita dari 11 kabupaten/kota di Provinsi Papua Barat pada tahun ke-t.

\bar{x}_t : Rata-rata logaritma natural PDRB per kapita dari 11 kabupaten/kota di Provinsi Papua Barat pada tahun ke-t.

t : 2010, 2011, ..., 2014.

Nilai koefisien variasi dari logaritma natural PDRB per kapita yang memiliki *trend* menurun mengindikasikan terjadinya konvergensi ekonomi di suatu daerah. Selanjutnya dilakukan analisis konvergensi beta untuk mengidentifikasi terjadinya konvergensi ekonomi antarkabupaten/kota di Provinsi Papua Barat. Konvergensi beta adalah kecenderungan perekonomian daerah miskin tumbuh lebih cepat dari perekonomian daerah kaya yang dilihat dari sisi pendapatan atau output per kapita. Laju pertumbuhan ekonomi daerah kaya akan mengalami perlambatan dari waktu ke waktu dikarenakan tingkat pengembalian modal yang berkurang (*diminishing return on capital*) dalam proses menuju kondisi *steady state*.

Dalam menganalisis konvergensi beta, tahap awal mengidentifikasi konvergensi beta absolut. Analisis konvergensi beta absolut hanya menggunakan satu variabel bebas, yaitu pendapatan per kapita awal (*initial income per capita*) suatu daerah. Menurut Sala-i-Martin [8], model yang digunakan dalam mengidentifikasi terjadinya konvergensi beta absolut adalah sebagai berikut:

$$\ln\left(\frac{y_{it}}{y_{i(t-1)}}\right) = \alpha + \theta \ln(y_{i(t-1)}) + \varepsilon_{it}, \quad (7)$$

keterangan:

α : intersep

θ : Slope

y_{it} : tingkat pendapatan per kapita daerah ke-i tahun ke t

$y_{i(t-1)}$: tingkat pendapatan per kapita daerah ke-i tahun ke t-1

ε_{it} : residual yang mengikuti distribusi normal

i : kabupaten/kota di Provinsi Papua Barat

t : 2010, 2011, ..., 2014

Dalam model konvergensi absolut, nilai *slope* (θ) pada variabel tingkat pendapatan per kapita daerah ke-i tahun ke t-1 ($y_{i(t-1)}$) dapat menjelaskan terjadi atau tidaknya konvergensi. Jika nilai estimasi parameter θ bernilai negatif ($\theta < 0$) dan signifikan maka terdapat kondisi *beta convergence*. *Slope* bertanda negatif menandakan bahwa wilayah yang tertinggal memiliki kecenderungan untuk tumbuh lebih cepat dari wilayah maju dan mengusahakan PDRB yang *convergence*. Sebaliknya, *slope* bertanda positif menunjukkan tidak terjadi konvergensi bahkan mengarah kepada dispersi yang semakin lebar (ketimpangan).

Berdasarkan hal tersebut, penerapan konvergensi beta absolut menjadi relevan jika hanya dilihat dari modal fisik tanpa memandang perbedaan karakteristik dan struktur perekonomian antarwilayah. Namun, pada kenyataannya setiap wilayah memiliki karakteristik perekonomian tersendiri. Hal ini menyebabkan analisis konvergensi beta absolut menjadi lemah sehingga tahapan selanjutnya adalah melakukan estimasi persamaan konvergensi beta kondisional. Analisis konvergensi beta kondisional memperhitungkan perbedaan karakteristik perekonomian. Konvergensi beta absolut menjadi relevan saat parameter-parameter

perekonomian antarwilayah bersifat homogen.

Secara prinsip, mengidentifikasi terjadinya konvergensi beta pada model beta kondisional sama dengan beta absolut, yaitu nilai parameter θ harus secara signifikan bernilai negatif. Model yang digunakan pun sama, hanya saja model konvergensi beta kondisional adalah model pada konvergensi absolut yang ditambah dengan variabel kontrol. Model yang digunakan pada konvergensi beta kondisional adalah:

$$\ln\left(\frac{y_{it}}{y_{i(t-1)}}\right) = \alpha + \theta \ln(y_{i(t-1)}) + \gamma X_{it} + \varepsilon_{it}, \quad (8)$$

keterangan:

α : intersep

y_{it} : tingkat pendapatan per kapita daerah ke-i tahun ke t

$y_{i(t-1)}$: tingkat pendapatan per kapita daerah ke-i tahun ke t-1

ε_{it} : residual yang mengikuti distribusi normal

X_{it} : vektor dari variabel-variabel kontrol di daerah-i pada tahun ke-t

i : kabupaten/kota di Provinsi Papua Barat

t : 2010, 2011, ..., 2014.

Menurut Barro dan Sala-i-Martin dalam Arifin [11], kecepatan konvergensi (*convergence rate*) yang terjadi dapat dihitung dengan menggunakan hasil estimasi yang dihasilkan oleh parameter θ dengan rumus sebagai berikut:

$$\varphi = \ln(1+|\theta|) / T, \quad (9)$$

keterangan:

φ : kecepatan konvergensi per tahun

$|\theta|$: nilai mutlak *slope* variabel $\ln(y_{i(t-1)})$ pada persamaan konvergensi beta (persamaan 8)

T : rentang periode penelitian

Setelah diperoleh kecepatan konvergensi (*convergence rate*) yang terjadi, maka dapat dihitung waktu yang dibutuhkan suatu wilayah untuk menutup setengah ketimpangan PDRB per kapita yang terjadi atau waktu untuk mencapai setengah konvergen (*half time*) dapat menggunakan rumus sebagai berikut:

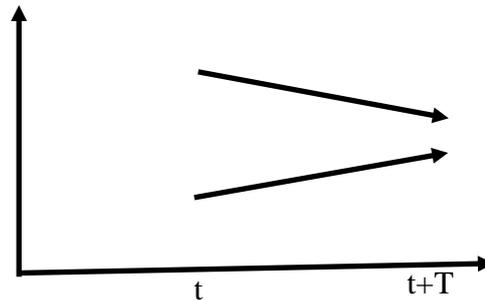
$$\tau = \frac{-\ln(0,5)}{\varphi} = \frac{\ln(2)}{\varphi}, \quad (10)$$

keterangan:

τ : waktu untuk mencapai setengah konvergen (*half time*)

φ : kecepatan konvergensi per tahun

Menurut Sala-i-Martin [12], hubungan antara konvergensi sigma dan konvergensi beta terlihat dari perilaku PDRB dua wilayah seiring berjalannya waktu. Pada Gambar 3, terlihat situasi perekonomian yang berbeda antara dua wilayah. Wilayah A memiliki perekonomian yang lebih maju dibandingkan wilayah B pada titik waktu awal (t). Selanjutnya, dalam rentang waktu antara tahun ke-t hingga tahun ke-t+T, dapat dikatakan bahwa telah terjadi konvergensi beta karena pertumbuhan ekonomi wilayah A lebih kecil dari wilayah B pada periode tersebut. Selain itu, terjadi penurunan disparitas atau ketimpangan pendapatan antara dua wilayah tersebut dalam rentang waktu ke-t hingga ke-t+T maka dapat disimpulkan pula bahwa telah terjadi konvergensi sigma.



Sumber: Sala-i-Martin

Gambar 3. Terjadi konvergensi sigma dan konvergensi beta

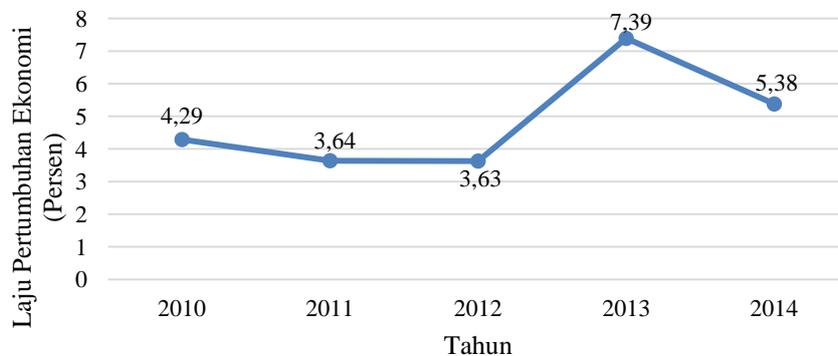
Struktur data yang digunakan adalah data panel dengan *cross section* sejumlah 11 kabupaten/kota di Provinsi Papua Barat dan *time series* periode 2010-2014. Data yang digunakan berupa data sekunder kuantitatif yang seluruhnya diperoleh dari BPS. Beberapa faktor yang diduga mendorong terjadinya konvergensi ekonomi di Provinsi Papua Barat sebagai berikut:

1. Tingkat pendapatan per kapita yang digunakan dalam penelitian ini adalah Pendapatan Domestik Regional Bruto (PDRB) per kapita
2. Dana Otonomi Khusus dan Penyesuaian (DOKP) adalah dana yang dialokasikan untuk membiayai pelaksanaan otonomi khusus suatu daerah dan dana yang dialokasikan untuk membantu daerah dalam rangka melaksanakan kebijakan tertentu pemerintah dan DPR sesuai perundang-undangan.
3. Persentase penduduk yang bekerja terhadap angkatan kerja adalah persentase penduduk dengan status bekerja dibandingkan dengan jumlah penduduk yang termasuk angkatan kerja.
4. **Bekerja** adalah kegiatan ekonomi yang dilakukan oleh seseorang dengan maksud memperoleh atau membantu memperoleh pendapatan atau keuntungan, paling sedikit 1 jam (tidak terputus) dalam seminggu yang lalu.
5. Angkatan Kerja adalah penduduk usia kerja (15 tahun atau lebih) di suatu wilayah yang aktif bekerja, atau punya pekerjaan namun sementara tidak bekerja, dan yang sedang mencari pekerjaan.
6. Angka Harapan Hidup (AHH) merupakan rata-rata perkiraan banyak tahun yang dapat ditempuh oleh seseorang selama hidup.
7. Persentase penduduk berumur 10 tahun ke atas yang berpendidikan minimal SMA adalah persentase dari jumlah penduduk yang berusia 10 tahun ke atas yang memiliki ijazah minimal tingkat pendidikan SMA atau sederajat dibandingkan dengan total jumlah penduduk usia 10 tahun ke atas.

Analisis Deskriptif

Gambaran awal mengenai perekonomian Provinsi Papua Barat dilihat dari pertumbuhan ekonomi yang menunjukkan perkembangan PDRB ADHK suatu wilayah pada waktu tertentu dibanding periode sebelumnya. Gambar 4 menunjukkan laju pertumbuhan ekonomi Provinsi Papua Barat periode 2010-2014. Menurut BPS, Provinsi Papua Barat mengalami laju pertumbuhan sebesar 4,29 persen pada tahun 2010. Laju pertumbuhan tertinggi terjadi pada tahun 2013 mencapai 7,39 persen.

Perekonomian Provinsi Papua Barat yang tumbuh cepat tidak terlepas dari adanya upaya pemerintah dalam hilirisasi sektor pertambangan minyak dan gas. Hal tersebut ditunjukkan oleh laju pertumbuhan sektor industri pengolahan yang mencapai 8,46 persen pada tahun 2013 dengan kontribusi terhadap total PDRB sebesar 30,27 persen. Selanjutnya, laju pertumbuhan ekonomi Provinsi Papua Barat pada tahun 2014 sebesar 5,38 persen.



Sumber : Badan Pusat Statistik

Gambar 4. Laju pertumbuhan ekonomi Provinsi Papua Barat periode 2010- 2014

Setelah mengetahui perkembangan laju pertumbuhan ekonomi Provinsi Papua Barat, perlu diketahui mengenai sektor-sektor ekonomi yang menunjang perekonomiannya. Pada tahun 2010, *leading sector* atau sektor dengan kontribusi terbesar terhadap PDRB Provinsi Papua Barat adalah sektor industri pengolahan (32,70 persen) kemudian diikuti oleh sektor pertambangan dan penggalian (27,13 persen), serta sektor pertanian, kehutanan, dan perikanan (11,82 persen). Namun, sektor yang menjadi *leading sector* perekonomian Provinsi Papua Barat pada tahun 2014 ialah sektor industri pengolahan (30,07 persen), sektor pertambangan dan penggalian (20,73 persen), serta sektor konstruksi (12,87 persen) sedangkan sektor pertanian, kehutanan, dan perikanan (10,79 persen) menempati posisi keempat.

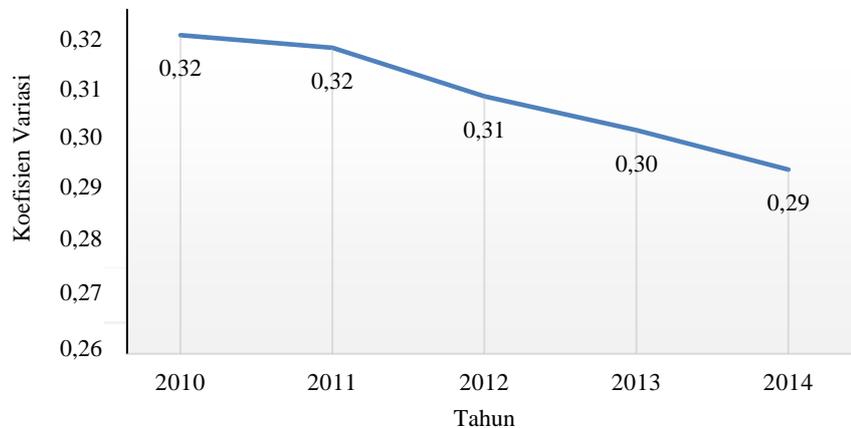
Analisis Konvergensi

Penelitian ini bertujuan untuk melihat apakah terjadi kecenderungan PDRB per kapita antarkabupaten/kota di Provinsi Papua Barat untuk seragam di masa yang akan datang (konvergensi ekonomi) dengan menggunakan analisis konvergensi. Dalam analisis konvergensi ada dua jenis konvergensi yaitu konvergensi sigma dan konvergensi beta. Konvergensi sigma adalah proses suatu daerah mengejar ketertinggalan dengan daerah lain dapat dilihat dari berkurangnya ketimpangan pendapatan per kapita sedangkan konvergensi beta adalah kecenderungan perekonomian daerah miskin tumbuh lebih cepat dari perekonomian daerah kaya yang dilihat dari sisi PDRB per kapita sehingga mampu mengurangi *gap* antarkabupaten/kota di Provinsi Papua Barat.

Konvergensi Sigma

Analisis konvergensi sigma pada penelitian ini menggunakan ukuran tingkat dispersi linier yaitu koefisien variasi logaritma natural PDRB per kapita antarkabupaten/kota di Provinsi Papua Barat periode 2010-2014. Gambar 5 menunjukkan ketimpangan PDRB per kapita antarkabupaten/kota di Provinsi Papua Barat cenderung menurun secara linier. Koefisien variasi logaritma natural PDRB per kapita antarkabupaten/kota di Provinsi Papua Barat pada tahun 2010 adalah sebesar 0,32. Selanjutnya, nilai koefisien variasi terus menurun secara

perlahan pada periode berikutnya dan hanya mencapai sebesar 0,29 pada tahun 2014. Penurunan nilai koefisien variasi logaritma natural PDRB per kapita antarkabupaten/kota mengindikasikan bahwa telah terjadi telah konvergensi sigma atau terdapat kecenderungan PDRB per kapita antarkabupaten/kota di Provinsi Papua Barat untuk seragam (konvergen) di masa yang akan datang.



Sumber: Badan Pusat Statistik

Gambar 5. Koefisien variasi logaritma natural PDRB per kapita antarkabupaten/kota di Provinsi Papua Barat periode 2010-2014

Konvergensi Beta

Analisis konvergensi beta pada penelitian ini dimulai dengan konvergensi beta absolut kemudian dilanjutkan dengan konvergensi beta kondisional.

a) Konvergensi Beta Absolut

Fixed effects model dengan metode estimasi FGLS dengan *cross section* SUR-PCSE untuk mengestimasi model konvergensi absolut di Provinsi Papua Barat dengan persamaan konvergensi beta absolut dapat dituliskan sebagai berikut:

$$\ln \left(\frac{Y_{it}}{Y_{i(t-1)}} \right) = (1,905739 + u_i) - 0,549689 \ln(Y_{i(t-1)})^* \quad (11)$$

* Signifikan pada taraf uji 5 persen

Tabel 1. Hasil estimasi regresi data panel konvergensi beta absolut dengan *fixed effects model*

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1,905739	0,502345	3,793684	0,0002
$\ln(Y_{i(t-1)})$	-0,549689	0,146123	-3,761838	0,0002

Persamaan 11 menunjukkan bahwa telah terjadi konvergensi beta absolut pada PDRB per kapita antarkabupaten/kota di Provinsi Papua Barat periode 2010-2014. Hal tersebut ditunjukkan oleh koefisien dari PDRB per kapita periode sebelumnya (*initial income*) yang bernilai negatif dan signifikan secara statistik yang menandakan bahwa wilayah dengan kondisi perekonomian yang lebih tertinggal, yaitu wilayah dengan PDRB per kapita rendah, memiliki kecenderungan secara absolut untuk tumbuh lebih cepat dibanding wilayah-wilayah yang lebih

maju sehingga memperkecil *gap* PDRB per kapita antarwilayah di Provinsi Papua Barat.

b) Konvergensi Beta Kondisional

Analisis konvergensi beta kondisional adalah analisis lanjutan setelah melakukan analisis konvergensi beta absolut. Pada analisis konvergensi beta kondisional, terdapat variabel-variabel penjelas lain yang dimasukkan ke dalam model selain variabel PDRB per kapita periode sebelumnya (*initial income*) masing-masing kabupaten/kota.

Analisis ini bertujuan untuk melihat pengaruh variabel-variabel penjelas lainnya terhadap terjadinya konvergensi ekonomi di Provinsi Papua Barat. Adapun variabel-variabel penjelas tersebut adalah Dana Otonomi Khusus dan Penyesuaian (DOKP), Pendapatan Asli Daerah (PAD), persentase penduduk yang bekerja terhadap jumlah angkatan kerja, Angka Harapan Hidup (AHH), dan persentase penduduk berumur 10 tahun ke atas yang berpendidikan minimal SMA. Konvergensi beta kondisional menggunakan *fixed effects model* dengan metode estimasi FGLS dengan *cross section* SUR-PCSE dengan persamaan sebagai berikut:

$$\ln\left(\frac{Y_{it}}{Y_{i(t-1)}}\right) = (-3,887478 + u_i) - 0,963505 \ln(Y_{i(t-1)})^* + 0,007524 \ln(\text{DOKP}_{it}) + 0,006746 \ln(\text{PAD}_{it})^* + 0,006746 \text{Bekerja}_{it}^* + 0,102523 \text{AHH}_{it}^* + 0,001633 \text{SMA}_{it}, \quad (12)$$

* Signifikan pada taraf uji 5 persen

Tabel 2. Hasil estimasi regresi data panel konvergensi beta kondisional dengan *fixed effects model* (lampiran 9)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2,229638	0,967279	-2,305063	0,0134
ln(PDRB1)	-0,963505	0,082960	-11,61408	0,0000
ln(DOKP)	0,004282	0,015297	0,286460	0,3881
ln(PAD)	0,024865	0,014101	1,763337	0,0429
BEKERJA?	0,007208	0,002098	3,435328	0,0007
AHH?	0,071271	0,017561	4,058388	0,0001
SMA?	0,001162	0,000880	1,320283	0,0973
Fixed Effects (Cross)				
_FAKFAK—C	-0,246141		_KABSORONG--C	0,949602
_KAIMANA—C	-0,160962		_RAJAAMPAT--C	0,268484
_TELWONDAMA—C	0,218929		_TAMBRAUW--C	-0,999118
_TELBINTUNI—C	2,749520		_MAYBRAT--C	-0,925321
_MANOKWARI—C	-0,608797		_KOTSORONG--C	-0,406758
_SORONGSEL—C	-0,511541			
Ringkasan Statistik				
R-squared			0,899239	

Salah satu keunggulan *fixed effects model* pada regresi data panel adalah diperolehnya efek individu yang merupakan gambaran keragaman (*heterogeneity*) antarindividu yang bersifat *time invariant*. Keragaman antarkabupaten/kota yang dihasilkan pada efek individu mencerminkan adanya faktor-faktor lain di luar model yang bersifat tetap sepanjang waktu pada suatu wilayah namun tidak

dimiliki oleh wilayah lain (Ekananda [13]). Faktor-faktor lain tersebut dapat berupa keterampilan manajerial pimpinan daerah, adat istiadat, dan kemampuan suatu daerah dalam mengelola sumber pendapatan daerah.

Berdasarkan efek individu dan konstanta bersama pada Tabel 2, maka dapat disimpulkan bahwa Kabupaten Teluk Bintuni mempunyai efek individu terbesar yaitu 2,75. Hal ini menggambarkan bahwa ketika semua variabel bebas pada model konvergensi beta kondisional bernilai sama untuk setiap kabupaten/kota maka wilayah tersebut memiliki tingkat pertumbuhan PDRB per kapita tertinggi atau merupakan wilayah yang relatif maju dari wilayah lainnya. Sebaliknya, Kabupaten Tambrauw memiliki efek individu terkecil yaitu sebesar -0,99 yang menandakan bahwa wilayah ini relatif tertinggal dari kabupaten/kota lain di Provinsi Papua Barat.

PDRB per kapita tahun sebelumnya (*initial income*) berpengaruh negatif dan signifikan secara statistik. Hal ini menandakan bahwa terdapat kecenderungan wilayah dengan kondisi perekonomian yang lebih tertinggal, yaitu wilayah dengan PDRB per kapita yang rendah, untuk tumbuh lebih cepat dibanding wilayah-wilayah yang lebih maju sehingga memperkecil ketimpangan PDRB per kapita antarwilayah di Provinsi Papua Barat. Berdasarkan hal tersebut, dapat disimpulkan bahwa telah terjadi konvergensi ekonomi antarkabupaten/kota di Provinsi Papua Barat periode 2010-2014 berdasarkan analisis konvergensi beta kondisional.

Laju konvergensi (ϕ) yang terjadi antarkabupaten/kota di Provinsi Papua Barat pada model konvergensi beta kondisional periode 2010-2014 adalah sebesar 13,49 persen per tahun. Nilai tersebut menandakan bahwa ketimpangan PDRB per kapita antarkabupaten kota akan berkurang 13,49 persen tiap tahunnya. Hal ini menandakan bahwa dengan adanya variabel penjelas lain dapat mempercepat laju konvergensi di Provinsi Papua Barat sehingga membutuhkan waktu 5,02 tahun untuk mencapai setengah konvergensi. Variabel penjelas yang digunakan pada konvergensi beta kondisional berperan penting dalam mendorong perekonomian wilayah miskin untuk tumbuh lebih cepat daripada daerah maju atau mendukung terjadinya konvergensi ekonomi antarkabupaten/kota di Provinsi Papua Barat.

Berdasarkan Tabel 2 dapat dilihat bahwa variabel lain yang berpengaruh signifikan terhadap pertumbuhan PDRB adalah Pendapatan Asli Daerah (PAD), persentase penduduk yang bekerja terhadap angkatan kerja, dan angka harapan hidup. Variabel Pendapatan Asli Daerah (PAD) berpengaruh signifikan dan positif terhadap pertumbuhan PDRB per kapita dalam rangka konvergensi ekonomi antarkabupaten/kota di Provinsi Papua Barat. Setiap peningkatan 1 persen PAD maka akan meningkatkan pertumbuhan PDRB per kapita sebesar 0,024865 persen dengan asumsi *ceteris paribus*. PAD merupakan gambaran mengenai kemandirian daerah dilihat dari sisi keuangan daerah itu sendiri. Peran PAD dalam rangka mempercepat konvergensi ekonomi antarkabupaten/kota adalah menjadi sumber dana yang dapat digunakan daerah dalam melaksanakan program kerja demi menciptakan kesejahteraan masyarakat.

Variabel persentase penduduk yang bekerja terhadap angkatan kerja berpengaruh signifikan dan berhubungan positif terhadap pertumbuhan ekonomi dalam upaya mempercepat terjadinya konvergensi ekonomi. Koefisien regresi sebesar 0,007208 menandakan bahwa setiap peningkatan 1 persen persentase penduduk yang bekerja terhadap angkatan kerja maka akan meningkatkan pertumbuhan ekonomi sebesar 0,007208 persen dengan asumsi *ceteris paribus*. Hal tersebut berarti bahwa semakin banyak penduduk yang aktif dalam kegiatan produksi maka akan meningkatkan jumlah produksi pada suatu wilayah sehingga

akan mendorong perkembangan perekonomian dan mempercepat konvergensi ekonomi. BPS [15] mencatat bahwa sebesar 45,26 persen penduduk bekerja pada sektor pertanian. Nilai tersebut menunjukkan bahwa banyak penduduk yang masih menggantungkan penghasilannya dari sektor pertanian. Wilayah yang relatif tertinggal dan berkembang cenderung digerakkan oleh sektor pertanian namun masih dengan tingkat produktivitas yang rendah.

Variabel angka harapan hidup dapat mewakili indikator kesehatan penduduk. Penelitian ini menghasilkan variabel angka harapan hidup berhubungan positif terhadap pertumbuhan ekonomi dalam upaya mempercepat konvergensi ekonomi. Setiap peningkatan 1 tahun angka harapan hidup maka akan meningkatkan pertumbuhan ekonomi sebesar 0,071271 persen dengan asumsi *ceteris paribus*. Hal tersebut juga menandakan bahwa penurunan derajat kesehatan dapat menyebabkan penurunan produksi karena kemampuan bekerja penduduk sebagai tenaga kerja menurun jika berada pada kondisi yang kurang sehat sehingga produktivitasnya menurun.

3. Kesimpulan

Kesimpulan yang diperoleh berdasarkan hasil dan pembahasan sebelumnya dalam penelitian ini sebagai berikut:

1. Pertumbuhan ekonomi di Provinsi Papua Barat periode 2010-2014 cenderung meningkat. Perekonomian Provinsi Papua Barat periode 2010-2014 digerakkan oleh sektor Industri Pengolahan dengan kategori industri pengolahan minyak dan gas bumi yang memberikan sumbangan terbesar pada sektor ini. Selain itu, sektor Pertambangan dan Penggalan juga menjadi sektor penunjang nilai PDRB Provinsi Papua Barat.
2. Telah terjadi konvergensi sigma antarkabupaten/kota di Provinsi Papua Barat periode 2010-2014 yang terlihat dari nilai koefisien variasi logaritma natural PDRB per kapita antarkabupaten/kota yang memiliki tren menurun selama periode 2010-2014. Hasil ini menunjukkan bahwa ketimpangan PDRB per kapita antarkabupaten/kota di Provinsi Papua Barat cenderung menurun.
3. Telah terjadi konvergensi beta antarkabupaten/kota di Provinsi Papua Barat periode 2010-2014. Identifikasi konvergensi beta dibagi menjadi:
 - a. Konvergensi beta absolut menunjukkan terjadinya konvergensi ekonomi antarkabupaten/kota di Provinsi Papua Barat periode 2010-2014 yang menandakan bahwa terdapat kecenderungan wilayah dengan kondisi perekonomian tertinggal untuk mengejar ketertinggalannya dibanding wilayah yang maju dan mewujudkan pemerataan pendapatan.
 - b. Konvergensi beta kondisional yaitu konvergensi beta dengan memperhatikan perbedaan karakteristik juga menunjukkan terjadinya konvergensi ekonomi antarkabupaten/kota di Provinsi Papua Barat. Adanya variabel penjelas yang merupakan perbedaan karakteristik antarwilayah mempercepat terjadinya konvergensi ekonomi di Provinsi Papua Barat periode 2010-2014.
4. Berdasarkan hasil konvergensi beta kondisional, variabel-variabel penjelas yang berpengaruh secara positif dan signifikan terhadap pertumbuhan PDRB per kapita dalam upaya mempercepat konvergensi ekonomi antarkabupaten/kota di Provinsi Papua Barat antara lain: Pendapatan Asli Daerah (PAD), persentase penduduk yang bekerja terhadap jumlah angkatan

kerja, dan Angka Harapan Hidup (AHH). Sementara itu, variabel Dana Otonomi Khusus dan Penyesuaian (DOKP) dan persentase penduduk berumur 10 tahun ke atas yang minimal berpendidikan SMA tidak signifikan mempengaruhi pertumbuhan PDRB per kapita.

Referensi

- [1] Badan Pusat Statistik. 2015. Statistik Daerah Provinsi Papua Barat. Jakarta: BPS.
- [2] Huda, Ni'matul. 2014. *Desentralisasi Asimetris dalam NKRI*. Bandung: Nusa Media.
- [3] Irawan., dan Suparmoko. (2002). *Ekonomi Pembangunan*. Yogyakarta: BPFE UGM
- [4] Tjokromidjojo, H. Bintoro., dan Mustopadidjaja, A.R. (1990). *Pengantar Pemikiran Tentang Teori & Strategi Pembangunan Nasional*. Jakarta: C.V. Haji Masagung.
- [5] Todaro, Michael P. (2004). *Ekonomi Regional, Teori dan Aplikasi: Edisi Ketujuh*. Jakarta: Bumi Aksara.
- [6] Badan Pusat Statistik. (2015). Provinsi Papua Barat Dalam Angka. Jakarta: BPS.
- [7] Todaro, Michael P. dan Stephen C. Smith (2003). *Pembangunan Ekonomi di Dunia Ketiga: Edisi Kedelapan*. Jakarta: Erlangga.
- [8] Mankiw, N. Gregory. (2003). *Macroeconomics: Fourth Edition*. New York: Worth Publisher. Inc.
- [9] Snowdown, Brian., dan Vane, Howard. (2005). *Modern Macroeconomics*. Northampton: Massachusetss Institute of Technology.
- [10] Barro, Robert J., dan Sala-i-Martin, Xavier. (2004). *Economic Growth: Second Edition*. Massachusetts London, England: The MIT Press Cambridge.
- [11] Arifin, Zainal. 2009. Kesenjangan dan Konvergensi Ekonomi Antar Kabupaten Pada Empat Koridor di Provinsi Jawa Timur. *HUMANITY*. Vol. 4, No. 2, Hal. 154-164.
- [12] Sala-i-Martin, Xavier X. (1996). The Classical Approach to Convergence Analysis. *The Economic Journal*, Vol. 106, Hal. 1019-1036.
- [13] Ekananda, Mahyus. (2016). *Analisis Ekonometrika Data Panel: Edisi Kedua*. Jakarta: Mitra Wacana Media.
- [14] Haryono. (2014). *Pengaruh Kebijakan Otonomi Khusus Terhadap Pertumbuhan Ekonomi Provinsi Papua Barat Tahun 2009-2012*. [Skripsi]. Sekolah Tinggi Ilmu Statistik.
- [15] Badan Pusat Statistik. (2015). Provinsi Papua Barat Dalam Angka. Jakarta: BPS.

GENERALIZED LINIER MODELS (GLM) KEJADIAN KECELAKAAN LALU LINTAS DI KOTA MERAUKE

HARIANI FITRIANTI¹, MARKUS PALOBO².

¹Universitas Musamus, hariani_fitrianti@yahoo.co.id

² Universitas Musamus, markunmusmath@gmail.com

Abstrak. Meningkatnya kepadatan lalu lintas disebabkan oleh peningkatan sarana dan prasarana yang lambat dibandingkan dengan pertumbuhan lalu lintas, sehingga tingkat kecelakaan lalu lintas semakin tinggi. Faktor-faktor penyebab tingkat kecelakaan lalu lintas yaitu manusia, lingkungan, dan kendaraan. Organisasi Kesehatan Dunia (WHO) meramalkan bahwa pada tahun 2030 kecelakaan lalu lintas akan menjadi faktor pembunuh terbesar di dunia. Tingkat kecelakaan untuk Kabupaten Merauke yang paling tinggi adalah tingkat kecelakaan yang disebabkan oleh faktor manusia. Oleh sebab itu perlu dilakukan penelitian tentang masalah tersebut. Tujuan dari penelitian ini yaitu memodelkan kejadian kecelakaan lalu lintas yang disebabkan oleh faktor manusia. Metode yang digunakan adalah metode pendekatan *Generalized Linier Models* (GLM) dengan distribusi dari variabel respon adalah *Poisson* dan fungsi penghubung variabel respon dan variabel predikat adalah fungsi *logaritma*. Hasil pemodelan yang diperoleh menggunakan GLM empat faktor yang mempengaruhi tingkat kecelakaan lalu lintas yaitu faktor lengah, faktor tidak tertib lalu lintas, faktor pengaruh alkohol, dan faktor melebihi batas kecepatan. Dengan hasil prediksi yang diperoleh dari model tersebut yaitu 31 kejadian kecelakaan lalu lintas dan mengalami peningkatan banyaknya kejadian kecelakaan lalu lintas dari periode sebelumnya sebesar 5,4%.

Kata kunci: Kecelakaan Lalu Lintas, Faktor Manusia, *Generalized Linear Models*

1. Pendahuluan

A. Latar Belakang

Transportasi merupakan bagian yang sangat penting dari kehidupan manusia, khususnya transportasi dengan kendaraan bermotor, baik untuk kebutuhan pergerakan manusia maupun angkutan barang. Kemajuan ilmu pengetahuan dan teknologi mendukung perkembangan alat transportasi, sehingga menyebabkan laju pertumbuhan alat transportasi semakin meningkat. Alat transportasi bertujuan untuk mewujudkan lalu lintas dan angkutan jalan dengan selamat, aman, cepat, lancar, tertib dan teratur, nyaman dan efisien, mampu memadukan moda transportasi lainnya, menjangkau seluruh pelosok wilayah daratan, untuk menunjang pemerataan, pertumbuhan dan stabilitas sebagai pendorong, penggerak dan penunjang pembangunan nasional dengan biaya yang

terjangkau oleh daya beli masyarakat [1]. Peningkatan jumlah kendaraan selain berdampak positif, juga memberikan dampak negatif, salah satunya meningkatnya jumlah kecelakaan lalu lintas yang melibatkan kendaraan.

Kota Merauke saat ini mengalami pertumbuhan yang sangat cepat, baik dari segi ekonomi, pendidikan, maupun sarana dan prasarana. Jumlah penduduk kota Merauke setiap tahun mengalami pertumbuhan yang cukup pesat, dikarenakan pertumbuhan ekonomi kota Merauke yang semakin baik. Pertumbuhan penduduk turut mempengaruhi peningkatan jumlah kendaraan di kota Merauke, baik itu kendaraan roda empat, maupun kendaraan roda dua. Demikian juga dengan prasarana transportasi berupa jalan setiap tahun mengalami peningkatan dari segi kuantitas dan kualitas. Dengan segala perkembangan dan pertumbuhan yang terjadi, muncul berbagai masalah, salah satunya masalah transportasi. Masalah yang paling disorot adalah tentang masalah keselamatan lalu lintas yang mengalami peningkatan kejadian kecelakaan dari tahun ke tahun di Kota Merauke, mulai dari kecelakaan ringan sampai kecelakaan yang menimbulkan korban jiwa yang disebabkan oleh berbagai macam faktor diantaranya yaitu faktor jalan, faktor pengemudi, faktor kendaraan, ataupun kondisi lingkungan dan alam.

Kecelakaan sering terjadi kota Merauke sebagai pusat kegiatan ekonomi, sosial, dan budaya. Penduduk merupakan faktor utama dalam perkembangan suatu kota yang diiringi dengan pertumbuhan wilayah dan perekonomian dan kota-kota pendukung sekitarnya serta kota-kota yang memiliki pusat-pusat kegiatan tertentu. Penyebaran kegiatan ekonomi tidak terpusat di satu tempat saja. Hal ini dikarenakan kebutuhan tersebut tidak dapat dipenuhi di sekitar tempat tinggal. Pemerintah telah banyak melakukan usaha penanggulangan kecelakaan lalu lintas, diantaranya pemberian rambu – rambu lalu lintas namun masalah tersebut tidak dapat terselesaikan dengan mudah.

Kecelakaan lalulintas adalah suatu peristiwa di jalan yang tidak diduga dan tidak disengaja melibatkan kendaraan dengan atau tanpa pengguna jalan lain yang mengakibatkan korban manusia dan/atau kerugian harta benda [2]. Menurut hasil studi *Transport Research Laboratory* tingkat kematian akibat kecelakaan lalulintas di Indonesia jauh lebih tinggi dibandingkan dengan negara-negara maju di Eropa dan Amerika Utara [3]. Berdasarkan data kecelakaan lalu lintas pada tahun 2006, tercatat 36.000 orang meninggal dunia karena kecelakaan di jalan, 19.000 di antaranya melibatkan pengendara sepeda motor [4]. Menurut data kecelakaan lalulintas di Kepolisian Republik Indonesia, kecelakaan transportasi jalan pada tahun 2012 tercatat sebanyak 109.038 kasus dengan jumlah korban meninggal dunia 25.131 orang [5].

Penyebab kecelakaan antara lain adalah: manusia, sarana dan

prasarana (kendaraan dan jalan), alam atau lingkungan.

1. Faktor Manusia

Faktor manusia yang dicatat oleh kepolisian, meliputi jenis kelamin korban, usia korban, profesi korban, dan peran korban dalam berkendara.

2. Faktor Kendaraan

Faktor kendaraan yang paling sering terjadi adalah ban pecah, rem tidak berfungsi sebagaimana seharusnya, mesin yang tiba-tiba mati, lampu mati, dan berbagai penyebab lainnya.

3. Faktor Jalan

Jalan merupakan bagian dari sistem transportasi darat yang memegang peranan penting untuk kelancaran transportasi. Selain peran tersebut, jalan juga memegang peran penting sebagai salah satu penyebab kecelakaan lalu lintas.

4. Faktor Alam atau Lingkungan

Faktor lingkungan ini bisa berupa pengaruh cuaca yang tidak menguntungkan, kondisi lingkungan jalan, benda-benda asing yang ada di jalan dan lain sebagainya.

Dalam penelitian ini, faktor manusia menjadi fokus perhatian dari peneliti. Hal ini dikarenakan hampir seluruh kejadian kecelakaan lalu lintas yang terjadi di kota Merauke semuanya disebabkan oleh karena faktor manusia. Penyebab kecelakaan lalu lintas di Indonesia paling banyak disebabkan oleh faktor manusia (91%). Faktor kedua kecelakaan sebanyak 5% adalah faktor kendaraan, dan faktor jalan sebanyak 3% serta faktor lingkungan sebesar 1% [6]. Faktor manusia dapat dikelompokkan menjadi dua yaitu kondisi pengemudi dan usia pengemudi.

a. Kondisi Pengemudi.

Lima faktor yang menyebabkan kecelakaan yaitu: fisik pengemudi, tingkat kedisiplinan dan pemahaman berlalu lintas masih rendah, kecakapan pengemudi, jarak pandang yang kurang (dalam mengambil jarak aman antar kendaraan) dan pelanggaran nilai batas kecepatan maksimum kendaraan (*speeding*).

b. Usia Pengemudi.

Berdasarkan usia pelaku kecelakaan lalu lintas, sebagian besar berusia antara 22 s.d 30 tahun kemudian disusul usia antara 31 s.d 40 tahun, di mana pada rentang usia tersebut tergolong sebagai usia tingkat emosinya paling stabil, tingkat kecekatan dan reflek yang lebih baik dibanding golongan usia lainnya, namun biasanya pada usia golongan ini tingkat mobilitasnya di jalan juga sangat tinggi. Jika pelaku kecelakaan golongan ini juga sekaligus menjadi korban, maka hal ini sekaligus merupakan golongan usia yang paling produktif. *World Health Organization* (WHO) mencatat hampir 1,2 juta orang di

seluruh dunia setiap tahun tewas akibat kecelakaan di jalan. Dari jumlah itu, 40 persen berusia di bawah 25 tahun. Jutaan lagi mengalami cedera dan sebagian lagi mengalami cacat seumur hidup [7].

B. Tujuan Penelitian

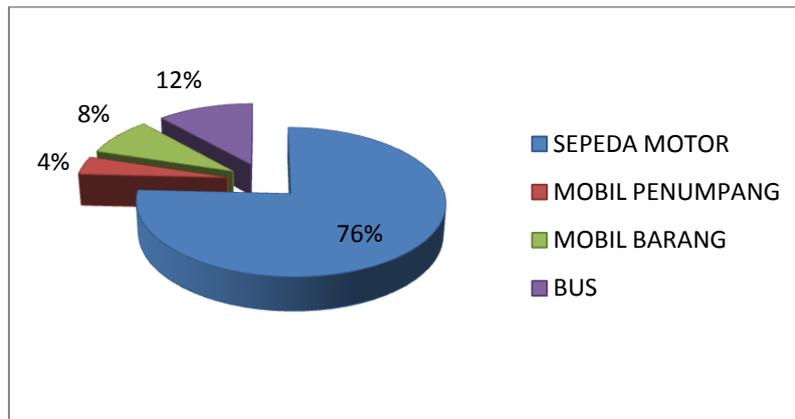
Tujuan dari penelitian ini adalah untuk memperoleh informasi tentang kecelakaan lalu lintas di kota Merauke yang meliputi:

1. Menentukan jenis kendaraan yang memiliki tingkat kecelakaan lalu lintas tertinggi di Kota Merauke.
2. Menentukan pengaruh dan signifikansi banyaknya kejadian kecelakaan lalu lintas yang disebabkan oleh faktor manusia. di Kota Merauke.
3. Menentukan model faktor-faktor banyaknya kejadian kecelakaan lalu lintas kejadian kecelakaan lalu lintas menggunakan pendekatan GLM.
4. Meramalkan banyaknya kejadian kecelakaan lalu lintas untuk satu periode kedepan.

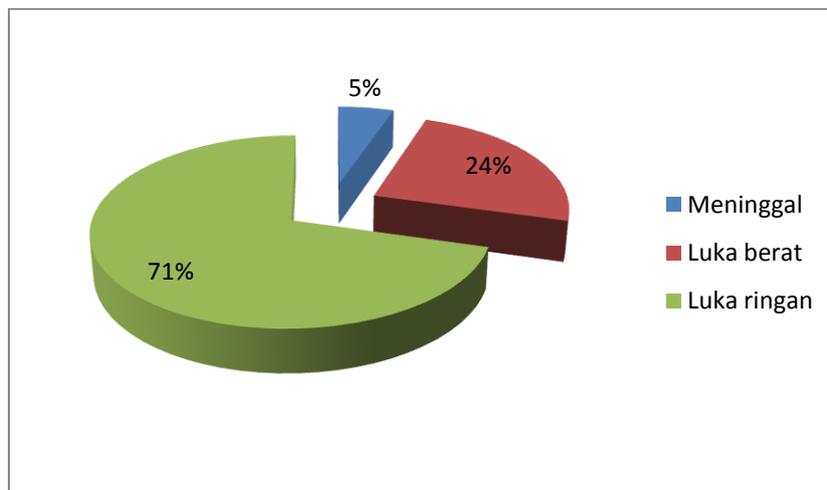
2. Hasil – Hasil Utama

A. Pengumpulan Data dan Statistik Deskriptif

Data yang diperlukan pada penelitian ini adalah data banyaknya kejadian kecelakaan lalu lintas dan data penyebab kecelakaan oleh faktor manusia yang diperoleh dari Unit Kecelakaan Lalu Lintas Kepolisian Resor Merauke pada bulan Januari tahun 2014 sampai dengan Oktober tahun 2016. Dari data yang diperoleh yang tercatat adalah data kecelakaan lalu lintas yang terdiri dari kecelakaan sepeda motor, kecelakaan bus, kecelakaan pada mobil penumpang, dan kecelakaan pada mobil barang. Sedangkan data penyebab kecelakaan oleh faktor manusia yaitu kecelakaan lalu lintas yang disebabkan oleh lengah, lelah, mengantuk, sakit, tidak tertib, tekanan psiko, pengaruh obat, pengalkohol, dan batas kecepatan. Untuk lebih jelasnya dapat dilihat pada Gambar 1 dan 2 berikut ini:

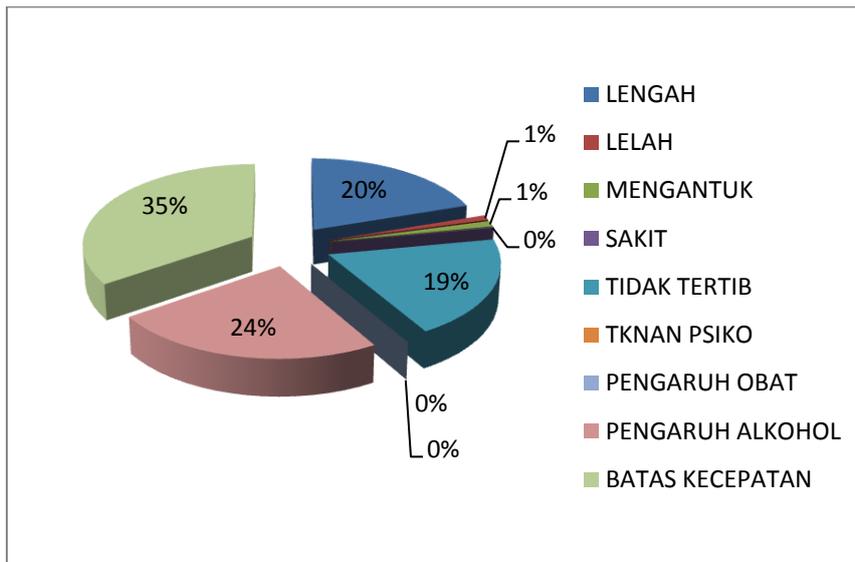


Gambar 1. Kendaraan yang Terlibat Kejadian kecelakaan lalu lintas



Gambar 2. Kecelakaan Lalu Lintas Berdasarkan Tingkat Keparahan

Faktor penyebab banyaknya kecelakaan kendaraan sepeda motor yang diakibatkan oleh manusia terbagi dalam beberapa katagori yaitu lengah, lelah, mengantuk, sakit, tidak tertib, tekanan psiko, pengaruh obat, pengalkohol, dan batas kecepatan. berdasarkan data yang diperoleh bahwa tidak semua kategori tersebut menyebabkan kecelakaan kendaraan sepeda motor. Teridentifikasi penyebab kecelakaan kendaraan sepeda motor di Kabupaten Merauke yaitu lengah, lelah, mengantuk, tata tertib, alkohol, dan batas kecepatan. Untuk lebih jelasnya dapat dilihat pada Gambar 3 berikut ini:



Gambar 3. Kecelakaan Lalu Lintas Berdasarkan Penyebab

Berdasarkan Gambar 3 terlihat bahwa pengaruh sakit, pengaruh obat, pengaruh lelah, pengaruh mengantuk, dan pengaruh tekanan psiko tidak berpengaruh signifikan terhadap penyebab kecelakaan lalu lintas, sehingga pada penelitian ini difokuskan pada banyaknya kejadian kecelakaan lalu lintas yang disebabkan oleh empat faktor yaitu lengah, pengaruh tidak tertib, pengaruh alkohol, dan pengaruh melebihi batas kecepatan.

B. Pemodelan Kejadian Kecelakaan Lalu Lintas

Pemodelan banyaknya kejadian kecelakaan lalu lintas pada penelitian ini menggunakan pendekatan *generalized linear model* (GLM), dengan data banyaknya kejadian kecelakaan lalu lintas sebagai variabel respon (Y) dan data penyebab terjadinya kecelakaan oleh faktor manusia sebagai variabel bebas (X). Data variabel respon secara teori adalah tipe data diskrit. Oleh sebab itu distribusi untuk banyaknya kejadian kecelakaan lalu lintas diasumsikan berasal dari distribusi keluarga eksponensial yaitu distribusi Poisson, karena distribusi Poisson merupakan salah satu jenis distribusi data diskrit, sehingga pemodelan banyaknya kejadian kecelakaan lalu lintas menggunakan pendekatan GLM dengan distribusinya *Poisson* dan fungsi penghubungnya yaitu fungsi logaritma[8]. Untuk distribusi *Poisson* mean dan variansi sama, sehingga persamaannya dapat ditulis sebagai berikut:

$$\ln \mu_i = \exp(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_6 x_{6i}) \quad (1)$$

dengan :

- μ_i = rata-rata banyaknya kejadian kecelakaan lalu lintas
- x_{1i} = faktor lengah
- x_{2i} = faktor lelah
- x_{3i} = faktor mengantuk
- x_{4i} = faktor tidak tertib lalu lintas
- x_{5i} = faktor pengaruh alkohol
- x_{6i} = faktor melebihi batas kecepatan

Langkah-langkah pemodelan GLM yaitu sebagai berikut:

1. Uji Asumsi Variabel Respon

pada penjelasannya sebelumnya telah diasumsikan bahwa variabel respon berdistribusi *Poisson*, maka akan diuji asumsi tersebut menggunakan uji khi-kuadrat. Hasil uji distribusi variabel respon dengan menggunakan bantuan *sofware* R.3.2.1 diperoleh perhitungan $\chi^2 = 149,4226$ dan $\chi_{tabel} = 47,4$. Hipotesa nol yaitu H_0 variabel respon berdistribusi *Poisson*, maka diperoleh $\chi^2 > \chi_{tabel}$ sehingga hipotesa tersebut tidak ditolak artinya variabel respon dapat diasumsikan berdistribusi *Poisson*. Hasil estimasi parameter distribusi *Poisson* menggunakan *metode maksimum likelihood* (MLE) yaitu $\mu = 32,94$.

2. Menaksir parameter model banyaknya kejadian kecelakaan lalu lintas

Langkah selanjutnya setelah memperoleh distribusi dari variabel respon yaitu menaksir parameter model banyaknya kejadian kecelakaan lalu lintas dengan fungsi penghubung variabel respon dan variabel prediktor yaitu fungsi logaritma. Untuk menaksir parameter model banyaknya kecelakaan kendaraan bermotor yaitu menggunakan metode *weighted least square* [9]. Dengan bantuan *sofware* R.3.2.1 diperoleh hasil pada Tabel 1.

Tabel 1. Estimasi Parameter Model Banyaknya Kejadian Kecelakaan Lalu Lintas

Variabel prediktor	Estimasi	Standar error
Intercep	2.682830	0.091997
Lengah	0.032473	0.005396

Tidak tertib	0.051404	0.006716
Pengaruh alkohol	0.044206	0.014507
Melebihi batas kecepatan	0.037933	0.007480

3. Uji signifikansi parameter

Uji signifikansi digunakan untuk menguji parameter variabel bebas berpengaruh signifikan terhadap banyaknya kejadian kecelakaan lalu lintas dengan tingkat kepercayaan 95%. Uji signifikansi menggunakan uji-Wald [9]. Hasil uji signifikansi model ditunjukkan pada Tabel 2.

Tabel 2. Hasil Uji Signifikan Parameter GLM

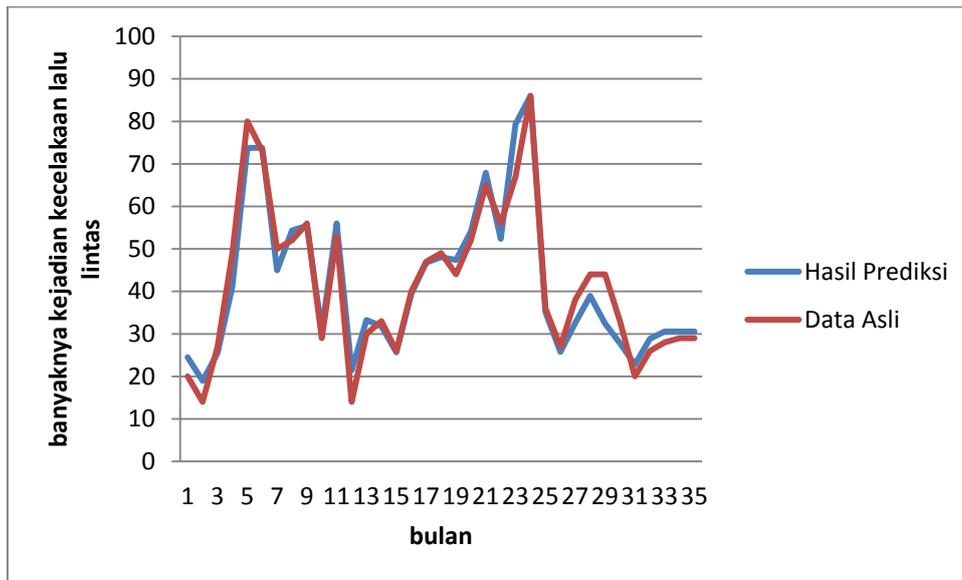
Variabel bebas	<i>p-value</i>
Intercep	< 2e-16 ***
Lengah	1.76e-09 ***
Tidak tertib	1.94e-14 ***
Pengaruh alkohol	0.00231 **
Melebihi batas kecepatan	3.94e-07 ***

Berdasarkan Tabel 2 dengan $\alpha = 0,05$ dengan hipotesa nol $\beta = 0$, maka faktor lengah, faktor tidak tertib, faktor pengaruh alkohol, dan faktor melebihi batas kecepatan diperoleh nilai $p\text{-value} > \alpha$, sehingga hipotesa nol ditolak artinya faktor lengah, faktor tidak tertib, faktor pengaruh alkohol, dan faktor melebihi batas kecepatan berpengaruh signifikan terhadap banyaknya kejadian kecelakaan lalu lintas.

$$\ln \hat{\mu}_i = \exp(2,6828 + 0,0325 x_1 + 0,0514x_4 + 0,0442x_5 + 0,0379x_6) \quad (2)$$

C. Hasil Prediksi Banyaknya Kecelakaan Lalu Lintas

Untuk memprediksi banyaknya kejadian kecelakaan lalu lintas di kota Merauke untuk bulan november dapat menggunakan model GLM pada persamaan (2). Hasil prediksi banyaknya kejadian kecelakaan lalu lintas yang disebabkan oleh faktor manusia untuk bulan november dapat dilihat pada Gambar berikut:



Gambar 4. Perbandingan hasil prediksi dan data asli banyaknya kejadian kecelakaan lalu lintas

Gambar 4 terlihat bahwa hasil prediksi dan realita banyaknya kejadian kecelakaan lalu lintas memiliki pola tren yang sama dan hasil prediksi banyaknya kejadian kecelakaan lalu lintas cukup dekat dengan realita banyaknya kejadian kecelakaan lalu lintas, sehingga dapat disimpulkan model banyaknya kejadian kecelakaan lalu lintas pada persamaan (2) dapat digunakan untuk memprediksi banyaknya kejadian kecelakaan lalu lintas pada Kota Merauke. Hasil prediksi banyaknya kejadian kecelakaan lalu lintas menggunakan persamaan (2) untuk bulan november yaitu 31 kejadian kecelakaan lalu lintas dan peningkatan banyaknya kejadian kecelakaan lalu lintas yaitu sebesar 5,4%.

3. Kesimpulan

Berdasarkan hasil penelitian dapat disimpulkan banyaknya kejadian kecelakaan lalu lintas untuk Kota Merauke yaitu:

1. Tingkat kecelakaan lalu lintas menggunakan kendaraan sepeda motor lebih tinggi dibandingkan menggunakan bus, mobil angkutan barang, dan mobil penumpang.
2. Tidak ada kecelakaan lalu lintas yang disebabkan oleh faktor tekanan psikolog dan faktor sakit.
3. Faktor lengah, faktor tidak tertib, faktor pengaruh alkohol, dan faktor melebihi batas kecepatan signifikan mempengaruhi banyaknya kejadian kecelakaan lalu lintas.
4. Model faktor-faktor banyaknya kecelakaan kendaraan sepeda motor

menggunakan pendekatan GLM yaitu:

$$\ln \hat{\mu}_i = \exp(2,6828 + 0,0325 x_1 + 0,0514 x_4 + 0,0442 x_5 + 0,0379 x_6)$$

5. Banyaknya kejadian kecelakaan lalu lintas di Kota Merauke untuk satu periode kedepan yaitu 31 kejadian kecelakaan lalu lintas dan mengalami peningkatan dari periode sebelumnya sebesar 5,4%

Referensi

- [1]. Warpani 1992. Undang-Undang Nomor 14 Tahun 1992 Tentang Lalu Lintas dan Angkutan Jalan.
- [2]. Undang-Undang No. 22 Tahun 2009 tentang Lalu Lintas dan Angkutan Jalan
- [3]. *Transport Research Laboratory (TRL). (1995). Costing Road Accident in Developing Countries, Overseas Road Note 10.* United Kingdom: Overseas Centre, Crowthorne, Beshire
- [4]. Sugiyanto, G. (2010). Kajian Karakteristik dan Estimasi Biaya kecelakaan Lalu Lintas Jalan di Indonesia dan Vietnam. *Jurnal Berkala Transportasi FSTPT*, 10(2), 135-148.
- [5]. Puslitbang Kemenhub. (2013). Diskusi Litbang: Keselamatan Jalan menjadi Tanggung Jawab Bersama. Jakarta: Pusat Komunikasi Publik Litbang Kemenhub.
- [6]. Direktorat Keselamatan Transportasi Darat (DKTD). (2006). Manajemen Keselamatan Transportasi Jalan, Naskah Workshop Manajemen Keselamatan Transportasi Darat, Batam: Direktorat Jenderal Perhubungan Darat, 13 Desember 2006.
- [7]. Direktorat Jenderal Perhubungan Darat (Ditjen Hubdat). (2004). Cetak Biru Keselamatan Lalu Lintas dan Angkutan Jalan. Jakarta: Direktorat Lalu Lintas dan Angkutan Jalan, Direktorat Perhubungan Darat.
- [8]. Cahyandari, R. (2014). Pengujian Overdispersi pada Model Regresi Poisson. *Statistik*, Vol.14 No.2, November, 69-76
- [9]. Jong, P. D., & Heller, G. Z. (2008). *Generalized Linear Models For Insurance Data.* Cambridge: Cambridge University

APLIKASI *STOCHASTIC PRODUCTION FRONTIER* UNTUK MENGUKUR EFISIENSI TEKNIS USAHA TANI PADI DI PROVINSI KEPULAUAN BANGKA BELITUNG

FITRI KARTIASIH¹ DAN ADI SETIAWAN²

¹Sekolah Tinggi Ilmu Statistik, fkartiasih@stis.ac.id

²Badan Pusat Statistik Kota Jakarta Pusat, adise@bps.go.id

Abstrak: Industri beras atau padi merupakan salah satu industri strategis di Provinsi Kep. Bangka Belitung. Akan tetapi selama ini usaha tani padi belum efisien padahal usaha tani padi merupakan sumber mata pencaharian sebagian besar masyarakat di Provinsi Kep. Bangka Belitung. Tujuan penelitian ini antara lain untuk memberikan gambaran usaha tani padi di Provinsi Kep. Bangka Belitung, menganalisis efisiensi teknis usaha tani padi serta faktor-faktor yang mempengaruhinya dan menganalisis tingkat penerimaan usaha tani padi di Provinsi Kep. Bangka Belitung. Data yang digunakan dalam penelitian ini bersumber dari data hasil Survei Rumah Tangga Usaha Tanaman Padi 2014 (SPD 2014). Metode analisis yang digunakan adalah analisis deskriptif dan analisis *Stochastic Production Frontier*. Berdasarkan hasil penelitian menunjukkan bahwa faktor-faktor yang mempengaruhi produksi padi di Kep. Bangka Belitung adalah penggunaan benih, pupuk, pestisida dan penggunaan pekerja dibayar. Rata-rata tingkat efisiensi teknis petani padi di Kep. Bangka Belitung adalah 70 persen dari produksi maksimum. Hal ini menunjukkan bahwa usaha tani padi cukup efisien. Faktor-faktor yang berpengaruh negatif terhadap efisiensi teknis usaha tani padi adalah faktor umur petani, alat pengolahan lahan, status alat pengolahan lahan dan sistem tanam. Hasil penelitian juga menunjukkan bahwa semakin efisien usaha tani padi maka pendapatan petani juga semakin besar.

Kata kunci: efisiensi teknis, *Stochastic Production Frontier*, padi

1. Pendahuluan

Industri beras atau padi merupakan salah satu industri strategis di Indonesia. Sumbangan padi atau beras terhadap Produk Domestik Bruto (PDB) pertanian mencapai 28,8% yang menyerap tenaga kerja (*employment*) sebesar 28,79% dari total pekerja di pertanian (*agriculture employment*). Jumlah orang yang bekerja pada industri padi atau beras mencapai 12,05 juta orang, terbesar dibandingkan dengan industri manapun yang ada di tanah air (Sawit, [12]).

Akan tetapi sangat disayangkan karena selama ini usaha tani padi belum efisien padahal usaha tani padi menjadi sumber mata pencaharian sebagian besar masyarakat di Indonesia. Hal ini terlihat dari penggunaan tenaga kerja dalam jumlah banyak dan membutuhkan waktu relatif lama (Pitojo [11]). Luas lahan yang dimiliki juga relatif sempit, pengadaan sarana produksi dilakukan secara perorangan sehingga harganya menjadi lebih tinggi. Kendala tersebut mendorong perlunya usaha tani padi yang lebih efisien melalui penerapan berbagai teknologi. Efisiensi merupakan perbandingan output dan input yang digunakan dalam suatu

proses produksi (Nicholson [10]). Produksi pertanian yang efisien akan menurunkan biaya produksi yang dikeluarkan oleh petani. Hal ini tentu saja akan meningkatkan kesejahteraan petani tersebut.

Tabel 1 Produktivitas Padi Menurut Provinsi di Indonesia Tahun 2013-2015
(Kuintal/Hektar)

Provinsi	2013	2014	2015	Provinsi	2013	2014	2015
Aceh	46.68	48.39	50.56	Nusa Tenggara Barat	50.08	48.8	51.71
Sumatera Utara	50.17	50.62	51.74	Nusa Tenggara Timur	32.8	33.5	35.61
Sumatera Barat	49.82	50.06	50.25	Kalimantan Barat	31.01	30.4	29.4
Riau	36.63	36.35	36.63	Kalimantan Tengah	32.84	34.6	35.07
Jambi	43.36	45.53	44.31	Kalimantan Selatan	42.34	42.1	41.87
Sumatera Selatan	45.96	45.26	48.67	Kalimantan Timur	42.7	42.6	41.2
Bengkulu	42.17	40.2	44.92	Kalimantan Utara	34.72	36.1	27.27
Lampung	50.26	51.18	51.49	Sulawesi Utara	50.1	48.9	49.05
Kep. Bangka Belitung	27.83	23.62	22.85	Sulawesi Tengah	45.98	46.5	48.57
Kep. Riau	36.15	36.44	36.46	Sulawesi Selatan	51.22	52.2	52.41
Dki Jakarta	58.88	53.86	55.95	Sulawesi Tenggara	42.23	46.8	47.07
Jawa Barat	59.53	58.82	61.22	Gorontalo	52.01	50.2	55.51
Jawa Tengah	56.06	53.57	60.25	Sulawesi Barat	48.8	47.7	49.41
Di Yogyakarta	57.88	57.87	60.65	Maluku	41.74	47.5	55.72
Jawa Timur	59.15	59.81	61.13	Maluku Utara	37.57	34	35.11
Banten	52.92	52.95	56.61	Papua Barat	39.76	40.2	42.12
Bali	58.66	60.12	62.14	Papua	41.3	43.1	43.95

Sumber: BPS [2]

Produktivitas padi provinsi-provinsi di Indonesia sangat bervariasi. Produktivitas padi di Indonesia secara umum mengalami peningkatan selama periode 2013-2015. Produktivitas padi tertinggi terdapat di Provinsi Jawa Timur yaitu sebesar 61,13 kuintal per hektar pada tahun 2015. Sedangkan produktivitas padi di Kepulauan Bangka Belitung menunjukkan angka yang terendah dibandingkan dengan provinsi-provinsi lain di Indonesia selama tahun 2013 sampai dengan 2015 yaitu berturut-turut 27,83; 23,62; 22,85 kuintal per hektar. Nilai produktivitas padi di provinsi ini mengalami penurunan selama tiga tahun terakhir. Hal ini menjadi menarik untuk dikaji mengapa produktivitas padi di Kepulauan Bangka Belitung tidak sampai setengahnya produktivitas padi di Pulau Jawa dan bahkan yang terendah di Indonesia.

Berdasarkan latar belakang dan permasalahan yang telah diuraikan sebelumnya, maka tujuan penelitian ini adalah: 1) Menganalisis gambaran usaha tani padi; 2) Menganalisis efisiensi teknis usaha tani padi serta faktor-faktor yang mempengaruhinya; 3) Menganalisis tingkat penerimaan usaha tani padi di Provinsi

Kep. Bangka Belitung.

Lipse. *et al.* [9] menjelaskan bahwa fungsi produksi adalah hubungan fungsi yang memperlihatkan output maksimum yang dapat diproduksi oleh setiap input dan oleh kombinasi berbagai input. Namun dalam perkembangannya, sebagian besar penelitian empiris menggunakan metode *least square* untuk memperkirakan fungsi produksi dan lebih sering disebut fungsi produksi rata-rata. Coelli, Rao dan Battese [5] menyatakan bahwa fungsi produksi *frontier* adalah fungsi produksi yang menggambarkan output maksimum yang dapat dicapai dari setiap tingkat penggunaan input. Jadi apabila suatu usaha tani berada pada titik di fungsi produksi *frontier* artinya usaha tani tersebut efisiensi secara teknis.

Data yang digunakan dalam penelitian ini bersumber dari data hasil Survei Rumah Tangga Usaha Tanaman Padi 2014 (SPD 2014) Badan Pusat Statistik. Metode analisis yang digunakan untuk menjawab tujuan penelitian ini adalah 1) Analisis deskriptif untuk menganalisis gambaran usaha tani padi di Provinsi Kep. Bangka Belitung, 2) Analisis *Stochastic Production Frontier* untuk mengetahui tingkat efisiensi teknis usaha tani padi dan faktor-faktor yang mempengaruhinya serta untuk menganalisis tingkat penerimaan usaha tani padi di Provinsi Kep. Bangka Belitung. Data penelitian diolah dengan menggunakan program Frontier 4.1.

Metode efek inefisiensi teknis yang digunakan dalam penelitian ini mengacu pada model efek inefisiensi teknis yang dikembangkan oleh Coelli [4] dan Battese dan Coelli [5] dapat dirumuskan dalam persamaan berikut :

$$Y = \beta_0 B^{\beta_1} U^{\beta_2} P^{\beta_3} PB^{\beta_4} PTB^{\beta_5} e^{v_i - u_i} \quad (1)$$

Untuk memudahkan pendugaan maka persamaan di atas ditransformasikan dalam bentuk logaritma natural dan dapat dituliskan sebagai berikut :

$$\ln Y = \beta_0 + \beta_1 \ln B + \beta_2 \ln U + \beta_3 \ln P + \beta_4 \ln PB + \beta_5 \ln PTB + v_i - u_i \quad (2)$$

dengan

Y : Rata-rata produksi padi (kg gabah kering panen/hektar)

B : Rata-rata benih padi (kg/hektar)

U : Rata-rata pupuk urea (kg/hektar)

P : Rata-rata pestisida cair (ml/hektar)

PB : Rata-rata pekerja dibayar (HOK/hektar)

PTB : Rata-rata pekerja tidak dibayar (HOK/hektar)

β_0 : Intersep

β_i : Koefisien Parameter, dimana $i = 1, 2, 3, \dots, 5$, $0 < \beta_i < 1$ (*Diminishing Return*)

$v_i - u_i$: Error term (u_i = efek inefisiensi teknis dalam model)

v_i : variabel acak yang berkaitan dengan faktor-faktor eksternal (iklim, hama/penyakit dan kesalahan permodelan) sebarannya simetris dan menyebar normal ($v_{ij} \sim N(0, \sigma_v^2)$)

u_i : variabel acak non negatif dan diasumsikan mempengaruhi tingkat inefisiensi teknis dan berkaitan dengan faktor-faktor internal dan sebarannya bersifat setengah normal ($u_{it} \sim \left| N(0, \sigma_u^2) \right|$)

Variabel sisa (random shock) v_i merupakan variabel acak yang bebas dan secara identik terdistribusi normal (independent-identically distributed/i.i.d) dengan rata-rata (mathematical expectation/ u_i) bernilai nol dan ragamnya konstan, σ_v^2 ($N(0, \sigma_v^2)$), serta bebas dari u_i . Variabel kesalahan u_i adalah variabel yang menggambarkan efek inefisiensi di dalam produksi, diasumsikan terdistribusi secara bebas di antara setiap observasi dan nilai v_i . Variabel acak u_i tidak boleh bernilai negatif dan

distribusinya normal dengan nilai distribusi $N(\mu_i, \sigma_u^2)$ (Coelli dan Battese [5]). Variabel u_i yang digunakan untuk mengukur efek inefisiensi teknis, diasumsikan bebas dan distribusinya terpotong normal dengan $N(\mu_i, \sigma^2)$.

Faktor-faktor yang diperkirakan mempengaruhi tingkat inefisiensi teknis petani padi dalam penelitian ini adalah umur petani (Z_1), tingkat pendidikan formal (Z_2), *dummy* alat/sarana pengolahan lahan yang utama (Z_3), *dummy* status alat/sarana pengolahan lahan yang utama (Z_4), *dummy* penyuluhan (Z_5), *dummy* keanggotaan dalam kelompok tani (Z_6), *dummy* status lahan (Z_7), dan *dummy* system penanaman (Z_8). Dengan demikian parameter distribusi (u_i) efek inefisiensi teknis dalam penelitian ini adalah :

$$u_i = \delta_0 + \delta_1 Z_1 + \delta_2 Z_2 + \delta_3 Z_3 + \delta_4 Z_4 + \delta_5 Z_5 + \delta_6 Z_6 + \delta_7 Z_7 + w_{it} \quad (3)$$

u_i = Efek inefisiensi teknis

δ_0 = Konstanta

Z_1 = Umur petani (tahun)

Z_2 = Tingkat pendidikan formal

Z_3 = *Dummy* alat/sarana pengolahan yang utama ($D_1 = 1$ jika traktor roda 2, 4 atau lebih, $D_1 = 0$ jika lainnya)

Z_4 = *Dummy* mengikuti penyuluhan ($D_2 = 1$ jika mengikuti, $D_2 = 0$ jika tidak mengikuti)

Z_5 = *Dummy* keanggotaan dalam kelompok tani ($D_3 = 1$ jika anggota, $D_3 = 0$ jika bukan anggota)

Z_6 = *Dummy* status kepemilikan lahan ($D_4 = 1$ milik sendiri, $D_4 = 0$ jika sewa/bebas sewa)

Z_7 = *Dummy* sistem penanaman dalam kelompok tani ($D_5 = 1$ jika tunggal, $D_5 = 0$ lainnya)

Pengujian hipotesis hanya dilakukan untuk hasil output efek efisiensi teknis *frontier*. Untuk mengetahui apakah ada efek inefisiensi di dalam model menggunakan nilai *LR test* galat satu sisi, Kodde dan Palm [6]. Sedangkan untuk masing-masing variabel penduga apakah koefisien dari masing-masing parameter bebas (δ_i) yang dipakai secara terpisah berpengaruh nyata atau tidak terhadap parameter tidak bebas (μ_i) dengan menggunakan t-hitung.

Hipotesis 1 :

$$H_0 : \gamma = \delta_0 = \delta_1 = \delta_2 = \delta_3 = \delta_4 = \dots \delta_8 = 0$$

$$H_1 : \gamma = \delta_0 = \delta_1 = \delta_2 = \delta_3 = \delta_4 = \dots \delta_8 > 0$$

Hipotesis nol artinya efek inefisiensi teknis tidak ada dalam model. Jika hipotesis ini diterima (gagal tolak H_0), maka model fungsi produksi rata-rata sudah cukup mewakili data empiris. Uji statistik yang digunakan adalah uji *chi-square*.

$$LR = -2 \{ \ln[L(H_0)/L(H_1)] \}$$

Dimana $L(H_0)$ dan $L(H_1)$ adalah nilai dari fungsi *likelihood* di bawah hipotesa H_0 dan H_1 .

Kriteria uji :

LR galat satu sisi $> \chi^2_{\text{restriksi}}$ (table Kodde Palm) maka tolak H_0

LR galat satu sisi $< \chi^2_{\text{restriksi}}$ (table Kodde Palm) maka gagal tolak H_0

Tabel *chi-square* Kodde dan Palm adalah *table upper and lower bound* dari nilai kritis untuk uji bersama persamaan dan pertidaksamaan restriksi.

Hipotesis 2 :

$$H_0 : \delta_i = 0$$

$$H_1 : \delta_i \neq 0$$

$$i = 1, 2, \dots, 7$$

Hipotesis nol berarti koefisien dari masing-masing variabel di dalam model efek inefisiensi sama dengan nol. Jika hipotesis nol ini diterima maka masing-masing variabel penjelas dalam model efek inefisiensi tidak memiliki pengaruh terhadap tingkat inefisiensi di dalam proses produksi.

Uji statistik yang digunakan yaitu :

$$t\text{-hitung} = \frac{\delta_i - 0}{S(\delta_i)}$$

$$t\text{-tabel} = t_{(\alpha/2, n-k-1)}$$

Kriteria uji :

$$|t\text{-hitung}| > t\text{-tabel}_{(\alpha/2, n-k-1)} : \text{tolak } H_0$$

$$|t\text{-hitung}| < t\text{-tabel}_{(\alpha/2, n-k-1)} : \text{gagal tolak } H_0$$

dimana : k = jumlah variabel bebas

n = jumlah pengamatan (responden)

S (δ_i) = simpangan baku koefisien efek inefisiensi

2. Hasil – Hasil Utama

Berdasarkan Survei Rumah Tangga Usaha Tanaman Padi 2014, jumlah petani padi di Provinsi Kep. Bangka Belitung menunjukkan bahwa sekitar 90,94 persen petani tanaman padi adalah laki-laki, dan hanya 9,06 persen perempuan. Sementara itu, berdasarkan kelompok umur, sebagian besar petani tanaman padi (92,48 persen) berumur 30 tahun atau lebih, dan hanya sekitar 7,52 persen yang berumur di bawah 30 tahun. Hal yang cukup menarik adalah ternyata sebanyak 40,51 persen petani tanaman pangan berumur 50 tahun atau lebih .

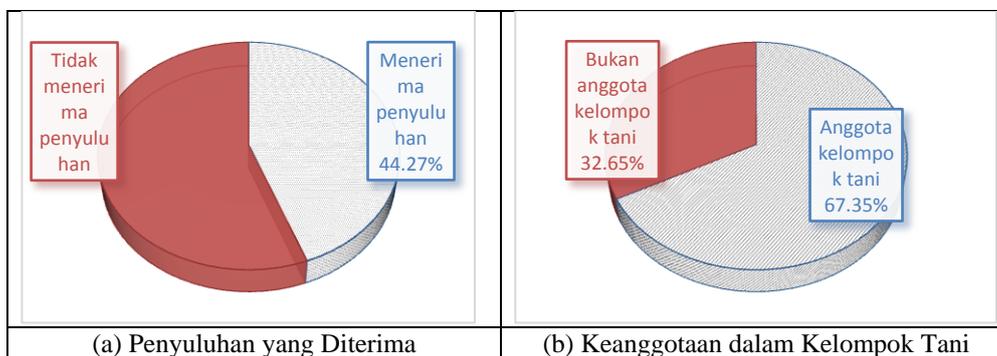
Jika dilihat menurut tingkat pendidikan petani menunjukkan bahwa sebagian besar petani tanaman pangan berpendidikan sekolah dasar (40,51 persen) dan tidak tamat sekolah dasar (34,53 persen). Sebanyak 12,99 persen berpendidikan SLTP, dan hanya 9,57 persen berpendidikan SLTA. Sementara itu, proporsi petani yang berpendidikan D1/D2, akademi/D3, dan setingkat sarjana atau lebih relatif kecil, masing-masing hanya di bawah 1 persen. Tingkat pendidikan petani merupakan indikator demografi yang juga berpengaruh terhadap produktivitas usaha tani tanaman pangan yang dijalankan oleh petani, karena indikator ini berkorelasi secara positif dengan kemampuan petani dalam menyerap teknologi dalam bidang pertanian, serta tingkat efisiensi usaha tani yang mereka jalankan.

Penyuluhan/bimbingan mengenai pengelolaan usaha tani, baik yang berkaitan dengan kegiatan budidaya tanaman padi (*on-farm*) maupun di luar kegiatan budidaya (*off-farm*), sangat penting dalam meningkatkan efisiensi usaha

tani yang dijalankan oleh petani. Dengan demikian, produktivitas usaha tani mereka dapat ditingkatkan. Dalam hal ini, keberadaan para penyuluh pertanian memegang peranan yang sangat krusial.

Pada Gambar 1(a) dapat dilihat bahwa mayoritas rumah tangga padi (55,73 persen) tidak memperoleh penyuluhan/bimbingan mengenai pengelolaan usaha tanaman padi selama setahun yang lalu. Sedangkan sebanyak 44,27 persen rumah tangga yang memperoleh penyuluhan/bimbingan mengenai pengelolaan usaha tani padi selama setahun yang lalu.

Keberadaan kelompok tani sangat penting bagi para petani. Kelompok tani pada dasarnya merupakan wadah belajar mengajar bagi para anggotanya untuk meningkatkan pengetahuan dan keterampilan dalam melakukan kegiatan usaha tani. Dengan demikian, produktivitas usaha tani dapat ditingkatkan. Selain sebagai wadah belajar mengajar para anggotanya, keberadaan kelompok tani juga dapat memperkuat kerjasama antara sesama petani yang menjadi anggota. Melalui kerja sama tersebut efisiensi usaha tani dapat ditingkatkan, dan petani memiliki kemampuan yang lebih dalam menghadapi berbagai hambatan dan kendala dalam melakukan kegiatan usaha tani.



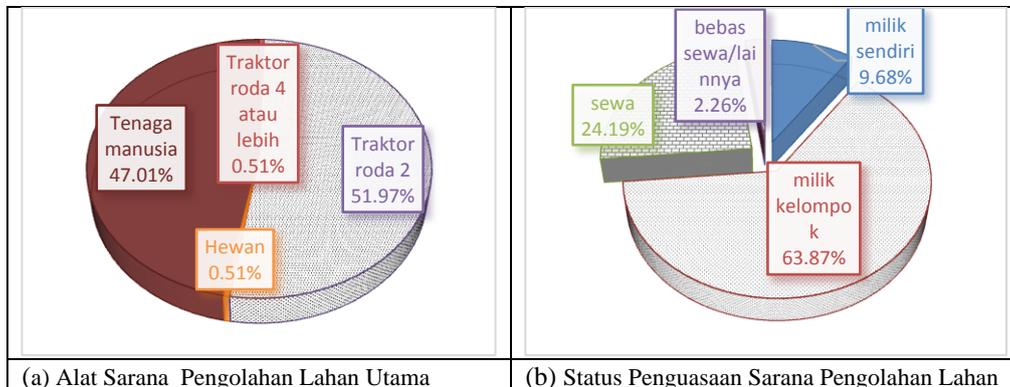
Gambar 1. Persentase Rumah tangga Padi Menurut Penyuluhan yang Diterima Selama Setahun yang Lalu dan Keanggotaan dalam Kelompok Tani

Suatu rumah tangga dianggap menjadi anggota kelompok tani jika pada saat pencacahan minimal salah satu anggota rumah tangga menjadi anggota kelompok tani. Pada Gambar 1(b) terlihat bahwa sebanyak 67,35 persen rumah tangga padi merupakan anggota kelompok tani pada saat pencacahan. Sementara 32,65 persen sisanya tidak menjadi anggota kelompok tani pada saat pencacahan.

Mekanisasi pertanian atau penggunaan mesin dalam kegiatan budidaya pertanian memiliki pengaruh yang signifikan dalam meningkatkan produktivitas dan efisiensi kegiatan usaha tani. Salah satu bentuk mekanisasi pertanian adalah penggunaan traktor, baik roda empat atau lebih maupun roda dua (*hand tractor*) dalam kegiatan pengolahan lahan.

Hasil Survei Rumah Tangga Usaha Tanaman Padi (ST2013 SPD) memperlihatkan bahwa sebagian besar petani padi menggunakan traktor sebagai alat pengolahan lahan yang utama. Pada Gambar 2(a) dapat dilihat bahwa persentase rumah tangga yang menggunakan traktor roda empat atau lebih dan traktor roda dua (*hand*

tractor) sebagai alat pengolahan lahan yang utama masing-masing mencapai 0,51 persen dan 51,97 persen dari jumlah total rumah tangga padi. Sementara itu, rumah tangga yang menggunakan hewan dan tenaga manusia dalam kegiatan pengolahan lahan masing-masing hanya sebesar 0,51 persen dan 47,01 persen.



Gambar 2. Persentase Rumah Tangga Padi Menurut Jenis Alat dan Status Penguasaan Sarana Pengolahan Lahan Utama

Berdasarkan Gambar 2(b) status penguasaan sarana pengolahan lahan, sebagian besar traktor yang digunakan rumah tangga padi adalah milik kelompok (63,87 persen), sedangkan 24,19 persennya menyewa, 9,68 persen milik sendiri dan sisanya adalah bebas sewa/lainnya.

Benih merupakan input produksi yang sangat krusial dalam kegiatan usaha tani. Penggunaan benih/bibit unggul akan meningkatkan produktivitas usaha tani. Sebagian besar rumah tangga membudidayakan padi inbrida menggunakan benih varietas lokal yaitu Balok Ketutu sebesar 14,11 persen, Mikongga 15,87 persen. Selain varietas lokal tersebut, rumah tangga padi sawah juga menggunakan varietas benih Ciherang, Ciliwung dan IR-64 masing-masing sebesar 18,52 persen, 5,11 persen dan 4,76 persen.

Pada umumnya (sekitar 71,13 persen) tanaman padi diusahakan oleh rumah tangga pada lahan milik sendiri. Sementara itu, sebanyak 3,78 persen rumah tangga mengusahakan padi pada lahan sewa, dan sebanyak 25,09 persen pada lahan bebas sewa serta lainnya. Sistem penanaman padi dapat dilakukan dengan cara tunggal, tumpang sari/tanaman sela, maupun campuran. Sebagian besar (sekitar 94,33 persen) rumah tangga usaha tani padi menggunakan sistem penanaman tunggal, sedangkan 5,67 persen lainnya menggunakan sistem penanaman tumpang sari atau tanaman sela.

Model yang digunakan untuk mengestimasi fungsi produksi usaha tani padi adalah model fungsi *Cobb-Douglas Stochastic Production Frontier* menggunakan parameter *Maximum Likelihood Estimated* (MLE). Metode MLE digunakan untuk menggambarkan hubungan antara produksi maksimum yang dapat dicapai dengan

penggunaan faktor-faktor produksi yang ada. Faktor-faktor yang diduga mempengaruhi produksi padi adalah penggunaan benih, penggunaan pupuk urea, penggunaan pestisida cair, penggunaan tenaga kerja dibayar dan tidak dibayar yang semuanya dalam satuan per hektar luas panen padi. Fungsi *Cobb-Dougllass Stochastic Production Frontier* mengikuti kaidah *diminishing return*, sehingga agar relevan dengan analisa ekonomi maka pencarian model menghendaki nilai koefisien yang positif. Nilai koefisien yang positif dapat memberikan informasi untuk melakukan upaya agar setiap penambahan input dapat menghasilkan tambahan output yang lebih besar (Soekartawi [13]).

Hasil estimasi produksi *Stochastic Frontier* padi di Kepulauan Bangka Belitung ditampilkan dalam tabel berikut:

Tabel 2 Pendugaan Parameter dengan Metode MLE untuk Fungsi Produksi Cobb-Douglas Sochastic Frontier Padi di Kepulauan Bangka Belitung

Variabel	Koefisien	t-hitung
Intercept	7,9866	0,3134
Ln Benih	0,1484	1,9739*
Ln Pupuk (Urea)	-0,0089	-0,1142
Ln Pestisida	0,0244	0,3543
Ln Pekerja dibayar	0,1386	2,4597*
Ln Pekerja tidak dibayar	0,0441	0,7052
R ²		0,2830
σ ²		9,6183
γ		0,01738
LR test of one side error		56,6553*

Keterangan: * signifikan pada alpha 5%

Tabel 2 adalah output dari pendugaan parameter dengan metode MLE untuk fungsi produksi Cobb-Douglas Sochastic Frontier padi di Kepulauan Bangka Belitung. Model yang dibentuk tidak memiliki masalah multikolinearitas dan memiliki nilai R² sebesar 0,2830. Nilai ini menunjukkan bahwa keragaman produksi yang dapat dijelaskan oleh variabel bebas dalam model sebesar 28,30 persen. Model ini memiliki nilai LR galat satu sisi sebesar 56,6553 yang lebih besar dari *Chi Square* Kodde dan Palm pada $\alpha = 0,05$ yaitu 55,190; sehingga terdapat inefisiensi teknis di Kepulauan Bangka Belitung. Output di atas (Tabel 2) juga dapat ditulis dalam bentuk persamaan sebagai berikut:

$$\widehat{\ln Y} = 7,9866 + 0,1484 \ln B - 0,0089 \ln U + 0,0244 \ln P + 0,1386 \ln PB + 0,0441 \ln PTB$$

Penggunaan benih berpengaruh positif dan signifikan dengan tingkat kepercayaan 95 persen terhadap produksi padi. Nilai elastisitas benih terhadap produksi sebesar 0,1484 menunjukkan bahwa penambahan jumlah benih sebesar satu persen akan meningkatkan produksi padi sebesar 0,1484 persen, *ceteris paribus* (diasumsikan variabel yang lain konstan/tetap). Hal ini menunjukkan bahwa jumlah benih yang digunakan petani selama ini masih memungkinkan ditambah untuk meningkatkan hasil produksi. Elastisitas penggunaan benih adalah yang paling besar dibandingkan dengan variabel lainnya.

Penggunaan pekerja dibayar berpengaruh positif dan signifikan dengan tingkat kepercayaan 95 persen terhadap produksi padi. Nilai elastisitas jumlah penggunaan pekerja dibayar terhadap produksi sebesar 0,1386 menunjukkan bahwa penambahan jumlah pekerja dibayar sebesar satu persen akan meningkatkan produksi padi sebesar 0,1386 persen, *ceteris paribus*. Hal ini menunjukkan bahwa jumlah penggunaan pekerja dibayar yang digunakan petani selama ini masih memungkinkan ditambah untuk meningkatkan hasil produksi. Penggunaan pupuk, pestisida maupun pekerja yang tidak dibayar tidak signifikan berpengaruh terhadap produksi padi. Hal ini dikarenakan nilai elastisitasnya relatif kecil sehingga pengaruhnya pun juga kecil atau sedikit.

Model efek inefisiensi teknis dianalisis secara simultan dalam model *stochastic production frontier*. Variabel-variabel bebas yang digunakan adalah umur, tingkat pendidikan, alat /sarana pengolahan lahan yang utama, penyuluhan, keanggotaan dalam kelompok tani, status lahan dan sistem penanaman.

Tabel 3 Persentase Rumahtangga Padi Menurut Tingkat Efisiensi Teknis di Kepulauan Bangka Belitung

Efisiensi Teknis	Persentase (%)
< 0,20	44,86
0,20 – 0,30	47,57
> 0,30	7,57
Rata-Rata TE	0,2011
Minimum TE	0,0680
Maksimum TE	0,3502

Berdasarkan Tabel 3 dapat dilihat bahwa rata-rata tingkat efisiensi teknis petani padi di Kepulauan Bangka Belitung adalah 0,2011 atau 20,11 persen dari produksi maksimum. Hal ini menunjukkan bahwa usaha tani padi di Kepulauan Bangka Belitung belum efisien dan masih terdapat peluang meningkatkan produksi sebesar 79,89 persen untuk mencapai produksi maksimum. Sebagian besar rumahtangga padi memiliki tingkat efisiensi teknis yang rendah yaitu dibawah 0,3.

Tabel 4 Pendugaan Parameter Maximum-Likelihood Model Inefisiensi Teknis Produksi Padi di Kepulauan Bangka Belitung

Variabel	Koefisien	t-hitung
Intercept	2,5007	0,0981
Umur	0,0023	0,5193
Tingkat pendidikan	-0,0167	-0,3756
Alat pengolahan lahan	-0,7536	-6,2254*
Penyuluhan	0,0346	0,3047
Kelompok tani	-0,4305	-3,0866*
Status lahan	-0,1178	-1,1069
Sistem tanam	0,0372	0,1491

Faktor-faktor yang mempengaruhi efisiensi teknis dianalisis dengan model efek inefisiensi teknis dengan variabel-variabel yang telah diterangkan sebelumnya. Hasil dari analisis model inefisiensi teknis di Kepulauan Bangka Belitung yang terlihat pada Tabel 4 menunjukkan bahwa variabel alat pengolahan

lahan dan keanggotaan dalam kelompok tani berpengaruh negatif dan signifikan terhadap inefisiensi teknis produksi. Hal ini sesuai dengan teori bahwa penggunaan mesin yaitu traktor baik roda 2 maupun lebih akan mengurangi inefisiensi teknis. Keikutsertaan rumahtangga dalam kelompok tani juga akan mengurangi tingkat inefisiensi teknis. Sedangkan variable lainnya tidak signifikan berpengaruh terhadap inefisiensi teknis.

Tabel 5 Rata-rata Nilai Produksi Petani Padi Menurut Tingkat Efisiensi Teknis di Kepulauan Bangka Belitung

Efisiensi Teknis	Rata-rata Nilai Produksi (Rp)
< 0,20	3.206.289
0,20 – 0,30	7.317.989
> 0,30	12.401.600

Berdasarkan Tabel 5 dapat dilihat bahwa rata-rata nilai produksi padi pada tahun 2014 baik produksi utama maupun produksi ikutan (sampingan) dari petani yang memiliki nilai efisiensi teknis kurang dari 0,20 adalah sebesar Rp. 3.206.289,-. Sedangkan rata-rata nilai produksi padi dari petani yang memiliki nilai efisiensi teknis 0,20 – 0,30 adalah sebesar Rp. 7.317.989,-. Dan rata-rata nilai produksi padi dari petani yang memiliki nilai efisiensi teknis lebih dari 0,30 adalah sebesar Rp. 12.401.600,-. Hal ini menunjukkan bahwa semakin efisien usaha tani padi maka pendapatan petani juga semakin besar.

3. Kesimpulan

Berdasarkan hasil dan pembahasan di atas, maka dapat diambil beberapa kesimpulan sebagai berikut:

1. Petani padi di Provinsi Kep. Bangka Belitung sebagian besar sudah menggunakan mesin atau traktor dalam pengolahan lahan. Pada umumnya (sekitar 72,43 persen) padi sawah diusahakan oleh rumah tangga pada lahan milik sendiri. Sementara itu, sebanyak 16,14 persen rumah tangga mengusahakan padi sawah pada lahan sewa, dan sebanyak 11,42 persen pada lahan bebas sewa serta lainnya. Sebagian besar (sekitar 98,46 persen) rumah tangga usaha tani padi sawah menggunakan sistem penanaman tunggal, sedangkan 0,89 persen lainnya menggunakan sistem penanaman tumpang sari atau tanaman sela.
2. Faktor-faktor yang mempengaruhi produksi padi di Provinsi Kep. Bangka Belitung adalah penggunaan benih dan penggunaan pekerja dibayar. Penggunaan benih memiliki nilai elastisitas yang paling tinggi dibandingkan dengan empat variabel lainnya. Variabel pupuk, pestisida dan penggunaan tenaga kerja tidak dibayar tidak signifikan berpengaruh terhadap produksi padi. Rata-rata tingkat efisiensi teknis petani padi di Provinsi Kep. Bangka Belitung masih rendah yaitu sebesar 0,2011 atau 20,11 persen dari produksi maksimum. Hal ini menunjukkan bahwa usaha tani padi di Provinsi Kep. Bangka Belitung belum efisien. Faktor-faktor yang berpengaruh negatif terhadap inefisiensi teknis usaha tani padi adalah alat pengolahan lahan dan keanggotaan dalam kelompok tani.

3. Semakin efisien usaha tani padi maka pendapatan petani juga semakin besar.

Referensi

- [1] Aigner, D.J., C.A.K. Lovell, and P. Schmidt. 1977. Formulation and Estimation of Stochastic Frontier Production Function Models. *Journal of Econometrics* 6, pp. 21–37.
- [2] Badan Pusat Statistik. 2016. *Statistik Indonesia*. Jakarta : BPS
- [3] Battese GE, Coelli TJ. 1998. Prediction of Firm-Level Technical Efficiencies with a General Frontier Production Function and Panel Data. *Journal of Econometrics* No. 38: 387-399.
- [4] Coelli TJ. 1997. *A Guide to Frontier Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation*. Armidale: CEPA University of New England.
- [5] Coelli, Tom, Prasada Rao dan George Battese. 1998. *An Introduction to Efficiency and Production Analysis*. Kluwer Academic Publishers. United States of America.
- [6] Kodde, DA., Palm, FC. 1986. Wald Criteria For Jointly Testing Equality and Inequality Restriction. *The Econometric Society*. Vol 54. No 5: 1243 – 1248.
- [7] Kumbhakar SC. 2002. Specification of Production Risk, Risk Preference and Technical Efficiency. *American Journal of Agricultural Economics*. 84 (1): 8-22
- [8] Lau LJ, Yotopoulos PA. 1971. A Test for Relative Efficiency and Application to Indian Agriculture. *The American Economic Review*. 61(1): 94-109.
- [9] Lipsey *et all*. 1995. *Pengantar Mikroekonomi*. Edisi Kesepuluh. Jilid 1. Binarupa Aksara. Jakarta.
- [10] Nicholson W. 1994. *Teori Ekonomi Mikro*. Volume ke-1. Wirajaya D, penerjemah. Jakarta (ID): Binarupa Aksara. Terjemahan dari: *Microeconomic Theory Basic Principles and Extensions*. Ed ke-1.
- [11] Pitojo S. 1997. *Budi Daya Padi Sawah TABELA*. Jakarta (ID): PT Penebar Swadaya.
- [12] Sawit MH, 2009. *Respons Negara Berkembang dan Indonesia dalam Menghadapi Krisis Pangan Global 2007-2008*, *Majalah Pangan*, 18(54).
- [13] Soekartawi. 2002. *Analisis Usaha tani*. Jakarta : UI Press.

MODEL PERSAMAAN STRUKTURAL UNTUK MENGETAHUI FAKTOR-FAKTOR YANG MEMPENGARUHI PRESTASI BELAJAR MATEMATIKA SISWA SMAN KECAMATAN PASAR REBO BERDASARKAN TINGKAT KEDISIPLINAN

ISMAH

Universitas Muhammadiyah Jakarta, Jl. KH. Ahmad Dahlan Cirendeu Ciputat
ismah.fr@gmail.com

Abstrak. Tujuan dari penelitian ini adalah untuk mengetahui faktor-faktor yang mempengaruhi prestasi belajar matematika siswa, serta apakah terdapat perbedaan faktor-faktor yang mempengaruhi prestasi belajar matematika siswa berdasarkan tingkat kedisiplinan belajar siswa. Metode penelitian yang digunakan adalah kuantitatif survei dengan sampel 260 siswa SMA yang berada di kecamatan Pasar Rebo. Analisis data terdiri dari pengujian instrumen angket dengan uji validitas dan reliabilitas, analisis deskriptif data penelitian, uji prasyarat analisis kenormalan, analisis data menggunakan model persamaan struktural dengan metode *maximum likelihood* dan uji kelayakan model (*goodness of fit*). Hasil analisis data diperoleh faktor psikologis, kelelahan dan lingkungan sekolah mempengaruhi prestasi belajar matematika siswa. Hal ini dibuktikan dengan hasil pengujian terhadap model struktural yang dibangun nilai probabilitas dari ketiga faktor tersebut masing-masing kurang dari 0,05 sehingga ketiga faktor dinyatakan signifikan mempengaruhi prestasi belajar matematika siswa. Selain itu diperoleh hasil tidak terdapat perbedaan faktor-faktor yang mempengaruhi prestasi belajar matematika siswa berdasarkan tingkat kedisiplinan belajar dengan kategori sedang dan tinggi. Berdasarkan kriteria *goodness of fit* diketahui bahwa model yang dibangun dapat dikatakan baik dikarenakan memenuhi batas standar baku untuk setiap kriteria *goodness of fit*.

Kata kunci: Model Persamaan Struktural, Prestasi Belajar Matematika, Kedisiplinan Belajar.

1. Pendahuluan

Pendidikan dengan kualitas terbaik selalu menjadi tujuan dari salah satu strategi dalam membangun bangsa di negara-negara berkembang, seperti Indonesia. Hal tersebut dikarenakan melalui pendidikan, Negara dapat mencetak sumber daya manusia yang akan memberikan kontribusi positif terhadap pembangunan Negara. Kualitas pendidikan dapat dicerminkan oleh peningkatan prestasi belajar peserta didik. Sekolah memiliki peran penting dalam menggali, menumbuhkan dan mengembangkan potensi siswa.

Borovik & Gardiner [1] matematika merupakan mata pelajaran yang membutuhkan kemampuan tingkat tinggi untuk dapat memenuhi standar kelulusan di sekolah. Tugas guru matematika mencetak siswa memiliki kemampuan tingkat

tinggi untuk dapat menyelesaikan permasalahan dalam matematika. Oleh karena itu, mahasiswa matematika sebagai calon guru matematika perlu dibina dengan baik guna mencapai tujuan pendidikan yang baik pula. Prestasi siswa akan menjadi ukuran keberhasilan proses binaan terhadap siswa selama berada di lingkungan sekolah.

Keberhasilan proses belajar mengajar dapat dicerminkan dari prestasi siswa yang dicapai. Akan tetapi, proses belajar mengajar bukanlah satu-satunya faktor penentu prestasi siswa. Banyak faktor yang mempengaruhi prestasi belajar siswa yang berasal dari dalam diri (internal) dan luar diri siswa (eksternal), hal ini yang menjadi landasan dilakukan penelitian-penelitian untuk mengkaji faktor-faktor tersebut. Hasil akhir penelitian diharapkan dapat memberikan informasi faktor apa yang dominan mempengaruhi prestasi belajar siswa, sehingga faktor tersebut akan menjadi perhatian pendidik diantaranya guru untuk mencapai tujuan pendidikan dengan kualitas terbaik. Selain itu, pengetahuan terhadap faktor-faktor yang mempengaruhi prestasi belajar akan menjadi salah satu upaya menghindari gangguan perilaku dan kesalahan dalam belajar.

Penelitian-penelitian yang dilakukan memiliki permasalahan yang semakin kompleks, yang menunjukkan perkembangan ilmu pengetahuan. Jenis variabel yang bervariasi dalam penelitian mengakibatkan teknik analisis statistik yang ada terlihat kelemahannya sehingga tidak dapat digunakan. Salah satu teknik analisis yang diadopsi dari beberapa kelemahan teknik analisis yang ada adalah *Struktural Equation Modelling* (SEM) yang telah dibakukan dalam bahasa Indonesia menjadi Model Persamaan Struktural.

Model persamaan struktural merupakan teknik analisis multivariat yang mampu mengetahui hubungan antar variabel yang kompleks. Jenis variabel yang mampu dianalisis oleh model persamaan struktural adalah variabel laten dan *manifest*. Variabel laten merupakan variabel yang tidak dapat diukur secara langsung, yang membutuhkan indikator-indikator sebagai proksi, sedangkan variabel *manifest* adalah indikator yang mengukur variabel tersebut. Analisis terhadap variabel laten dapat dilakukan menggunakan teknik analisis faktor konfirmatori (*confirmatory factor analysis*), namun untuk menganalisis pengaruh yang melibatkan beberapa variabel laten dan variabel *manifest* dapat menggunakan model persamaan struktural. Oleh karena itu, istilah model persamaan struktural dikatakan sebagai teknik analisis yang menggabungkan model pengukuran analisis faktor konfirmatori dengan model struktural analisis regresi atau analisis jalur.

Berdasarkan keunggulan yang dimiliki oleh model persamaan struktural menjadikan metode tersebut dapat diaplikasikan ke berbagai bidang sosial seperti pendidikan, psikologi, ekonomi, kesehatan dan lainnya yang melibatkan variabel-variabel yang tidak dapat diukur secara langsung dalam penelitiannya. Oleh karena ini untuk mengetahui faktor-faktor yang mempengaruhi prestasi belajar siswa dapat diterapkan model persamaan struktural untuk menganalisis datanya, dikarenakan banyak variabel laten yang dilibatkan. Seperti penelitian terdahulu yang pernah dilakukan oleh Siti Nurmawati [3], Muchamad Sapto Adi Wibowo [7] dan Muhammad Amin Paris [4].

Kedisiplinan siswa dalam belajar sebagai cermin proses belajar yang baik, sehingga keterampilan siswa akan terasah dan muncul kebiasaan baik. Kedisiplinan dalam belajar dapat berdampak pada pola kebiasaan baik siswa ketika memasuki jenjang pendidikan yang lebih tinggi, bahkan sampai ke dunia kerja. Selain itu, kedisiplinan dapat mempengaruhi prestasi siswa. Oleh karena perlu dilakukan

analisis faktor yang mempengaruhi prestasi belajar siswa berdasarkan tingkat kedisiplinan.

Tujuan penelitian ini dilakukan, antara lain untuk menentukan faktor-faktor internal dan eksternal yang berpengaruh terhadap prestasi belajar siswa SMAN Kecamatan Pasar Rebo berdasarkan perbedaan kedisiplinan belajar, dan membangun model kausalitas yang mampu menghubungkan faktor-faktor yang mempengaruhi prestasi belajar siswa SMAN Kecamatan Pasar Rebo.

Adapun sistematika penelitian ini adalah : (1) Penyusunan kuesioner, pengambilan data melalui pengisian kuesioner oleh responden, yaitu 260 siswa yang dipilih secara acak yang berasal dari SMAN Kecamatan Pasar Rebo Jakarta Timur tahun akademik 2015-2016. (2) Klasifikasi data berdasarkan tingkat kedisiplinan belajar siswa. (3) Analisis data menggunakan model persamaan struktural. Ghazali [2] mengajukan tahapan permodelan dan analisis persamaan struktural menjadi tujuh langkah yaitu (a) Pengembangan model secara teoritis. (b) Menyusun diagram jalur (*path diagram*). (c) Mengubah diagram jalur menjadi persamaan struktural. (d) Memilih matrik input untuk analisis data. (e) Identifikasi model, teknik estimasi model persamaan struktural dengan metode *maximum likelihood estimation* (MLE) yang lebih efisien dan *unbiased* jika asumsi normalitas multivariat terpenuhi. (f) Menilai kriteria *Goodness of fit*. (g) Interpretasi terhadap model.

2. Hasil – Hasil Utama

Data pada variabel kedisiplinan belajar diperoleh melalui angket yang bersifat tertutup dengan jumlah item sebanyak 12 butir. Skor yang digunakan dalam angket tersebut adalah 1 sampai 4. Berdasarkan data yang diperoleh dari angket yang disebarkan kepada responden menunjukkan bahwa variabel kedisiplinan memperoleh skor tertinggi 44 dari skor maksimal yang dicapai (X_{max}) sebesar $4 \times 12 = 48$ dan skor terendah yang diperoleh sebesar 27 dari skor minimum yang dicapai (X_{min}) $1 \times 12 = 12$. Dari skor tersebut kemudian dianalisis, diperoleh mean (M) sebesar 37,16 dan standar deviasi (s) sebesar 3,99.

Frekuensi kategori variabel kedisiplinan berdasarkan ketentuan sebagai berikut:

Diketahui X_{max} sebesar $4 \times 12 = 48$ dan X_{min} sebesar $1 \times 12 = 12$.

Mean ideal (M_i) = $\frac{1}{2} (X_{max} + X_{min}) = \frac{1}{2} (48 + 12) = 30$

Standar deviasi ideal (SD_i) = $\frac{1}{6} (X_{max} - X_{min}) = \frac{1}{6} (48 - 12) = 6$

a. Tinggi : $> M_i + 1 SD_i = > 30 + 6 = > 36$

b. Sedang : $M_i - 1 SD_i$ sampai $M_i + 1 SD_i = 30 - 6$ s/d $30 + 6 = 24$ s/d 36

c. Rendah : $< M_i - 1 SD_i = < 30 - 6 = < 24$

Tabel 1. Frekuensi kategori variabel kedisiplinan belajar

No.	Skor	Frekuensi	Persentase (%)	Kategori
1	37 - 48	165	63	Tinggi
2	24 - 36	95	37	Sedang
3	12-23	0	0	Rendah
TOTAL		260	100	

Berdasarkan tabel 1, dapat diketahui kedisiplinan pada kategori rendah

sebanyak 0 siswa (0%), kategori sedang sebanyak 95 siswa (37%) dan kategori tinggi sebanyak 165 siswa (63%).

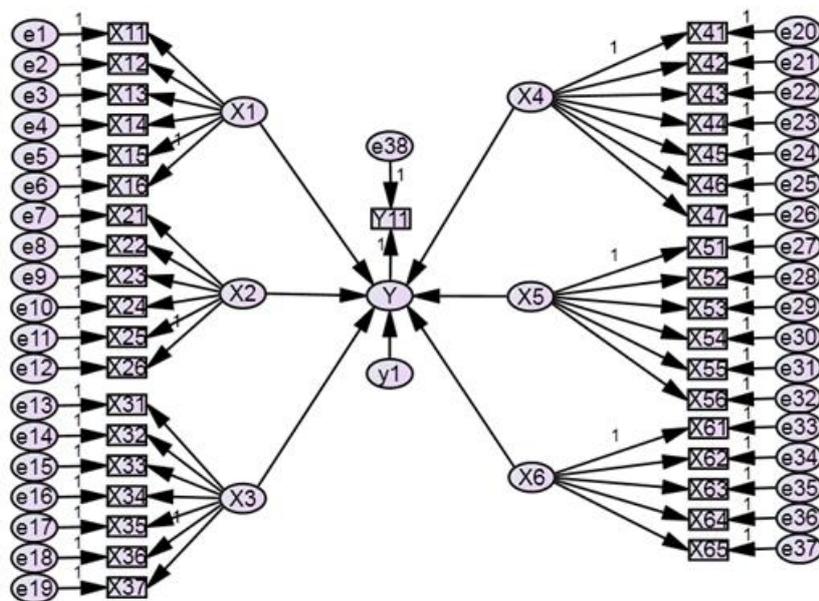
Tahapan dalam analisis SEM adalah sebagai berikut:

1. Pengembangan model berdasarkan teori.

Dalam penelitian ini model yang dikembangkan mengadopsi kutipan Slameto [6], mengenai faktor-faktor yang mempengaruhi prestasi belajar yaitu faktor internal yang terdiri dari jasmaniah (X1), psikologis (X2) dan kelelahan (X3), sedangkan faktor eksternal terdiri dari lingkungan keluarga (X4), lingkungan sekolah (X5) dan lingkungan tempat tinggal (X6).

2. Menyusun *path diagram*

Path diagram model SEM yang dirancang terdiri dari 37 indikator dari variabel laten eksogen dan 1 indikator dari variabel laten endogen dalam model, serta 83 parameter yang diestimasi. Berikut ini *path diagram* yang dirancang.



Gambar 1. *Path Diagram*

3. Uji normalitas data

Asumsi yang harus dipenuhi dalam SEM adalah data berdistribusi normal. Uji normalitas pada AMOS dilihat pada tabel *assessment of normality* (tabel 2) pada kolom *critical ratio* (c.r). Kriteria kenormalan adalah apabila nilai c.r < $\pm 2,58$ pada tingkat 5% maka data berdistribusi normal.

Tabel 2. *Assessment of normality*

Variable	skew	c.r.	Kurtosis	c.r.
X11	-.359	-.324	.363	.740
X12	-.540	-1.205	.963	1.007
X13	-.792	-1.459	.321	.655
X14	-.133	-.709	.231	.472
X15	-.987	-1.627	.805	1.643
X16	.030	.123	.234	.602
X21	.175	.715	.197	.525
X22	.29	1.590	.816	-2.228
X23	-.532	-2.406	1.405	1.299
X24	-.744	-1.066	.664	2.393
X25	-.754	-1.121	.990	1.089
X26	-.730	-1.987	1.306	.028
X31	-.351	-1.920	1.864	.090
X32	-.213	-1.166	.891	.895
X33	1.350	-2.373	1.444	.406
X34	-.996	-2.438	1.714	.680
X35	1.304	-.122	1.452	.428
X36	1.112	-.074	1.795	.903
X37	1.036	-.660	.234	.100

Variable	skew	c.r.	Kurtosis	c.r.
X41	1.180	-.444	.728	1.182
X42	1.134	-.195	.193	.720
X43	-.846	-.622	1.527	.169
X44	1.262	-.894	.789	.618
X45	-.954	-.211	1.228	.352
X46	1.765	-.638	1.239	1.577
X47	1.064	-.809	1.064	.368
X51	-.084	-.343	-.545	-1.113
X52	.462	1.887	-.544	-1.110
X53	-.285	-1.162	-1.080	-2.204
X54	-.226	-.921	-.867	-1.770
X55	.215	.876	.021	.043
X56	.486	1.984	.042	.086
X61	.436	2.871	-1.079	-1.551
X62	.038	.250	-1.800	-.925
X63	.431	2.037	-1.170	-1.850
X64	-.432	-2.043	-1.134	-1.731
X65	-.180	-1.182	-1.296	-2.265
Multivariate			-1.278	-1.232

Berdasarkan tabel 2 terlihat nilai c.r yang diperoleh untuk semua variabel berada antara -2.58 dan 2.58. Oleh karena itu dapat disimpulkan bahwa semua variabel yang masuk dalam model berdistribusi normal, sehingga asumsi kenormalan data terpenuhi.

Ukuran sampel dalam penelitian ini adalah 260 dengan asumsi normalitas terpenuhi, sehingga digunakan teknik estimasi model dengan metode *maximum likelihood*.

2.1. Analisis faktor-faktor yang mempengaruhi prestasi belajar siswa berdasarkan tingkat kedisiplinan rendah

Tingkat kedisiplinan belajar siswa dikategorikan ke dalam tiga bagian yaitu, rendah, sedang dan tinggi. Namun berdasarkan hasil analisis deskripsi pada tabel 1 diperoleh hasil bahwa tingkat kedisiplinan siswa SMAN kecamatan Pasar Rebo Jakarta Timur adalah kategori sedang dan tinggi, tidak ada satupun siswa yang masuk kategori tingkat kedisiplinan rendah. Oleh karena itu, analisis data untuk mengetahui faktor-faktor yang mempengaruhi prestasi belajar siswa di bagi dalam dua kategori tingkat kedisiplinan belajar yaitu sedang dan tinggi.

Hasil analisis faktor-faktor yang mempengaruhi prestasi belajar berdasarkan tingkat kedisiplinan rendah dapat dilihat pada hasil pengolahan menggunakan AMOS.

Nilai *chi-square* yang diperoleh 401,221 dengan *degree of freedom* $\frac{1}{2}(p + q)(p + q + 1) - t$

p = Banyaknya indikator dari variabel laten eksogen dalam model

q = Banyaknya indikator dari variabel laten endogen dalam model

t = Banyaknya parameter model yang akan diestimasi

sehingga diperoleh *degree of freedom* $\frac{1}{2}(37 + 1)(37 + 1 + 1) - 83 = 658$,

dengan *probability level* = .971 yang lebih besar dari 0.05 maka dapat disimpulkan model yang dibangun secara keseluruhan *overall fit* (signifikan).

Tabel 3. *Regression Weights*

			Estimate	S.E.	C.R.	P	Label
Y	<---	X1	.017	.030	.580	.062	
Y	<---	X2	.147	.074	1.982	.047	
Y	<---	X3	-.098	.061	1.603	.049	
Y	<---	X4	.009	.051	.176	.060	
Y	<---	X5	.921	.058	.358	.020	
Y	<---	X6	.071	.075	.948	.343	
X11	<---	X1	-.329	.204	.102	.053	
X12	<---	X1	-.011	.099	.114	.109	
X13	<---	X1	-.098	.245	.401	***	
X14	<---	X1	.435	.202	.149	***	
X15	<---	X1	1.000				
X16	<---	X1	-.295	.190	1.555	.120	
X21	<---	X2	.879	.024	1.002	***	
X22	<---	X2	-.011	.065	.378	.090	
X23	<---	X2	.308	.045	.506	***	
X24	<---	X2	.175	.011	2.122	***	
X25	<---	X2	1.000				
X26	<---	X2	.020	.190	-1.555	***	
X31	<---	X3	.865	.071	.278	.909	
X32	<---	X3	.779	.030	.231	***	
X33	<---	X3	.899	.202	1.177	***	
X34	<---	X3	.067	.962	.856	***	
X35	<---	X3	1.000				
X36	<---	X3	.036	.038	.962	***	
X37	<---	X3	.189	.562	.100	.689	
X41	<---	X4	1.000				
X42	<---	X4	1.206	.079	.485	.601	
X43	<---	X4	.435	.200	1.046	.055	
X44	<---	X4	.227	.256	.397	***	
X45	<---	X4	.759	.230	.755	.210	
X46	<---	X4	1.038	.590	.431	.103	
X47	<---	X4	.538	.461	.566	.061	
X51	<---	X5	1.000				
X52	<---	X5	.002	.789	.537	***	
X53	<---	X5	.892	.003	.050	***	
X54	<---	X5	.094	.018	.302	***	
X55	<---	X5	1.053	.090	.480	***	
X56	<---	X5	.024	.073	.468	***	
X61	<---	X6	1.000				
X62	<---	X6	1.173	.007	.803	.619	
X63	<---	X6	.311	.731	.176	.083	
X64	<---	X6	.922	.204	.622	.601	
X65	<---	X6	.688	.009	1.555	.057	

*** signifikan

Tabel 3 menunjukkan nilai signifikansi setiap variabel dalam menjelaskan variabel lainnya berdasarkan pada nilai *probability* pada kolom P (***). Dari tabel 3 dapat diketahui variabel psikologis (X2), kelelahan (X3) dan lingkungan sekolah (X5) mampu menjelaskan variabel prestasi belajar (Y) secara signifikan dikarenakan nilai *probability* yang diperoleh kurang dari $\alpha = 0.05$.

Sedangkan indikator yang mampu menjelaskan variabel eksogen secara signifikan diantaranya adalah:

- a. X13, X14 dan X15 mampu menjelaskan X1
- b. X21, X23, X24, X25 dan X26 mampu menjelaskan X2
- c. X32, X33, X34, X35 dan X36 mampu menjelaskan X3
- d. X41 dan X44 mampu menjelaskan X4
- e. X51, X52, X53, X54, X55, dan X56 mampu menjelaskan X5
- f. X61 mampu menjelaskan X6

Banyaknya indikator yang tidak signifikan dalam menjelaskan variabel jasmaniah (X1), lingkungan keluarga (X4) dan lingkungan tempat tinggal (X6) menjadikan ketiga variabel tersebut tidak signifikan dalam menjelaskan variabel prestasi belajar (Y). Besar hubungan antara kedua variabel dapat dilihat pada tabel 4 berikut ini.

Tabel 4 *Standardized Regression Weights*

	Estimate		Estimate
Y <--- X1	.520	X35 <--- X3	0.69
Y <--- X2	.799	X36 <--- X3	0.886
Y <--- X3	-.828	X37 <--- X3	0.408
Y <--- X4	.465	X41 <--- X4	0.641
Y <--- X5	.971	X42 <--- X4	0.565
Y <--- X6	.498	X43 <--- X4	0.634
X11 <--- X1	-.363	X44 <--- X4	0.831
X12 <--- X1	-.565	X45 <--- X4	0.639
X13 <--- X1	-.681	X46 <--- X4	0.447
X14 <--- X1	.821	X47 <--- X4	0.466
X15 <--- X1	.622	X51 <--- X5	0.816
X16 <--- X1	-.599	X52 <--- X5	0.791
X21 <--- X2	.879	X53 <--- X5	0.983
X22 <--- X2	.501	X54 <--- X5	0.905
X23 <--- X2	.808	X55 <--- X5	0.867
X24 <--- X2	.875	X56 <--- X5	0.929
X25 <--- X2	.711	X61 <--- X6	0.783
X26 <--- X2	0.681	X62 <--- X6	0.539
X31 <--- X3	0.432	X63 <--- X6	0.685
X32 <--- X3	0.779	X64 <--- X6	0.357
X33 <--- X3	0.841	X65 <--- X6	0.561
X34 <--- X3	0.978		

Berdasarkan tabel 4 terlihat bahwa beberapa indikator memiliki hubungan yang lemah di bawah 0.60 dengan variabel eksogen. Sedangkan terdapat hubungan yang kuat antara variabel eksogen dengan variabel endogen yaitu X2 dengan Y sebesar 79,7%, X3 dengan Y sebesar 82,8% dan X5 dengan Y sebesar 97,1%. Sedangkan variabel X1, X4 dan X6 memiliki hubungan yang sangat lemah dengan variabel Y, terlihat hasil estimasi pada tabel 4 bahwa ketiga variabel tersebut tidak signifikan dalam menjelaskan variabel Y.

Dengan kata lain, terdapat tiga faktor yang mampu menjelaskan prestasi belajar matematika siswa sekolah menengah atas kecamatan Pasar Rebo Jakarta Timur yaitu faktor psikologis, kelelahan dan lingkungan sekolah berdasarkan

tingkat kedisiplinan rendah.

2.2. Analisis faktor-faktor yang mempengaruhi prestasi belajar siswa berdasarkan tingkat kedisiplinan tinggi

Hasil analisis faktor-faktor yang mempengaruhi prestasi belajar berdasarkan tingkat kedisiplinan tinggi dapat dilihat pada tabel berikut ini hasil pengolahan menggunakan AMOS.

Dikarenakan diperoleh *probability level* = .956 lebih besar dari 0.05 maka dapat disimpulkan model yang dibangun secara keseluruhan/*overall fit* (signifikan). Adapun untuk mengetahui faktor-faktor yang mempengaruhi prestasi belajar matematika siswa dapat di lihat pada tabel 5 berikut ini.

Tabel 5. Regression Weights: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
Y <--- X1	.012	.024	.633	.102	
Y <--- X2	.142	.055	.978	***	
Y <--- X3	-.251	.037	.836	***	
Y <--- X4	.026	.084	.367	.088	
Y <--- X5	.894	.017	.457	***	
Y <--- X6	.140	.052	.838	.079	
X11 <--- X1	-.026	.001	.543	.067	
X12 <--- X1	-.242	.152	.454	.221	
X13 <--- X1	-.283	.325	.379	.060	
X14 <--- X1	.295	.106	.094	.061	
X15 <--- X1	1.000				
X16 <--- X1	-.077	.143	.406	.072	
X21 <--- X2	.792	.011	.893	***	
X22 <--- X2	.152	.055	.204	***	
X23 <--- X2	.286	.037	.612	***	
X24 <--- X2	.109	.017	1.026	***	
X25 <--- X2	1.000				
X26 <--- X2	.063	.084	.738	.075	
X31 <--- X3	.865	.071	.278	.909	
X32 <--- X3	.635	.056	.739	***	
X33 <--- X3	.923	.167	1.004	***	
X34 <--- X3	.141	.203	.592	***	
X35 <--- X3	1.000				
X36 <--- X3	.284	.072	.376	***	
X37 <--- X3	.603	.340	.692	.071	
X41 <--- X4	1.000				
X42 <--- X4	.837	.071	.662	.368	
X43 <--- X4	.072	.064	.941	.057	
X44 <--- X4	.403	.063	.940	.067	
X45 <--- X4	.821	.079	.258	.118	
X46 <--- X4	.902	.075	.874	***	
X47 <--- X4	1.203	.076	.047	.082	
X51 <--- X5	1.000				
X52 <--- X5	.053	.055	.923	***	
X53 <--- X5	1.052	.003	.137	***	
X54 <--- X5	.280	.022	.538	***	

	Estimate	S.E.	C.R.	P	Label
X55 <--- X5	.872	.036	.700	***	
X56 <--- X5	.092	.051	.892	***	
X61 <--- X6	1.000				
X62 <--- X6	.473	.030	.605	.158	
X63 <--- X6	.961	.502	.324	.077	
X64 <--- X6	.394	.059	.829	.544	
X65 <--- X6	.720	.013	.062	.061	

*** signifikan

Tabel 5 menunjukkan nilai signifikansi setiap variabel dalam menjelaskan variabel lainnya berdasarkan pada nilai *probability* pada kolom P (***). Dari tabel 5, hasil analisis berdasarkan kategori tingkat kedisiplinan tinggi dapat diketahui bahwa variabel psikologis (X2), kelelahan (X3) dan lingkungan sekolah (X5) mampu menjelaskan variabel prestasi belajar matematika (Y) secara signifikan dikarenakan nilai *probability* yang diperoleh kurang dari $\alpha = 0.05$.

Sedangkan indikator yang mampu menjelaskan variabel eksogen secara signifikan diantaranya adalah:

- X15 mampu menjelaskan X1
- X21, X22, X23, X24, dan X25 mampu menjelaskan X2
- X32, X33, X34, X35 dan X36 mampu menjelaskan X3
- X41 dan X46 mampu menjelaskan X4
- X51, X52, X53, X54, X55, dan X56 mampu menjelaskan X5
- X61 mampu menjelaskan X6

Secara keseluruhan hasil analisis terhadap indikator dalam menjelaskan variabel laten eksogen, yang diperoleh dengan kategori disiplin belajar tinggi hampir sama dengan hasil analisis dengan kategori disiplin belajar rendah. Adapun perbedaannya dapat diamati pada tabel 6 berikut.

Tabel 6. Indikator variabel eksogen berdasarkan tingkat kedisiplinan belajar rendah dan tinggi

Variabel eksogen	Indikator	
	Disiplin Belajar Rendah	Disiplin Belajar Tinggi
X1	X13, X14, X15	X15
X2	X21, X23, X24, X25, X26	X21, X22, X23, X24, X25
X3	X32, X33, X34, X35, X36	X32, X33, X34, X35, X36
X4	X41, X44	X41, X46
X5	X51, X52, X53, X54, X55, X56	X51, X52, X53, X54, X55, X56
X6	X61	X61

Sedangkan besar hubungan antara kedua variabel dapat dilihat pada tabel 7 berikut ini.

Tabel 7 Standardized Regression Weights

			Estimate			Estimate	
Y	<---	X1	.461	X35	<---	X3	.732
Y	<---	X2	.826	X36	<---	X3	.910
Y	<---	X3	-.904	X37	<---	X3	.520
Y	<---	X4	.522	X41	<---	X4	.772
Y	<---	X5	.939	X42	<---	X4	.481
Y	<---	X6	.435	X43	<---	X4	.567
X11	<---	X1	-.397	X44	<---	X4	.492
X12	<---	X1	-.578	X45	<---	X4	.525
X13	<---	X1	-.502	X46	<---	X4	.682
X14	<---	X1	.550	X47	<---	X4	.520
X15	<---	X1	.622	X51	<---	X5	.824
X16	<---	X1	-.497	X52	<---	X5	.801
X21	<---	X2	.881	X53	<---	X5	.955
X22	<---	X2	.738	X54	<---	X5	.923
X23	<---	X2	.970	X55	<---	X5	.975
X24	<---	X2	.873	X56	<---	X5	.895
X25	<---	X2	.828	X61	<---	X6	.632
X26	<---	X2	.581	X62	<---	X6	.497
X31	<---	X3	.467	X63	<---	X6	.522
X32	<---	X3	.820	X64	<---	X6	.492
X33	<---	X3	.880	X65	<---	X6	.574
X34	<---	X3	.952				

Hubungan yang kuat antara variabel eksogen dengan variabel endogen terlihat pada variabel X2 dengan Y sebesar 82,6%, X3 dengan Y sebesar 90,4% dan X5 dengan Y sebesar 93,9%. Sedangkan variabel X1, X4 dan X6 memiliki hubungan yang sangat lemah dengan variabel Y, terlihat hasil estimasi pada tabel 7 bahwa ketiga variabel tersebut tidak signifikan dalam menjelaskan variabel Y.

Dengan kata lain, terdapat tiga faktor yang mampu menjelaskan prestasi belajar matematika siswa sekolah menengah atas kecamatan Pasar Rebo Jakarta Timur yaitu faktor psikologis, kelelahan dan lingkungan sekolah berdasarkan tingkat kedisiplinan tinggi.

2.3. Evaluasi Model *Goodness of Fit*

Hasil evaluasi terhadap faktor-faktor yang mempengaruhi prestasi belajar matematika siswa sekolah menengah atas kecamatan Pasar Rebo Jakarta Timur dapat diamati berdasarkan pada beberapa kriteria seperti yang dituliskan Schumacker & Lomax [5]. Berikut ini adalah hasil *goodness of fit* yang diperoleh.

Tabel. 8. *Goodness of Fit*

Kriteria	<i>Cut off</i>	Nilai yang diperoleh
Chisquare	Diharapkan kecil	Nilai chi square yang diperoleh dengan prob lebih besar dari 0,05
RMSEA (<i>Root Mean Square Error of Approximation</i>)	$\leq 0,08$	0,005
GFI (<i>Goodness of Fit Index</i>)	$\geq 0,90$	0,902
AGFI (<i>Adjusted Goodness of Fit</i>)	$\geq 0,90$	0,907
NFI (<i>Normed Fit Index</i>)	$\geq 0,90$	0,833
CFI (<i>Comparative Fit Index</i>)	$\geq 0,95$	0,901

Tabel 8 menunjukkan bahwa hasil uji kelayakan ada yang tidak memenuhi nilai ideal yaitu NFI dan CFI, hal tersebut dikarenakan terdapat beberapa parameter yang tidak signifikan. Namun, sebagian besar uji kelayakan terhadap parameter yang dimasukkan dalam model telah memenuhi nilai ideal dari kriteria suatu model yang baik.

2.4. Interpretasi Hasil

Berdasarkan hasil survey 260 siswa, diketahui tingkat kedisiplinan siswa sekolah menengah atas kecamatan Pasar Rebo Jakarta Timur di kategorikan sedang dan tinggi. Adanya perbedaan tingkat kedisiplinan belajar siswa menjadikan analisis terhadap faktor-faktor yang mempengaruhi prestasi belajar siswa di bagi dalam dua kategori berdasarkan tingkat kedisiplinan belajar. Hal ini dilakukan untuk melihat apakah terdapat perbedaan faktor-faktor yang mempengaruhi prestasi belajar siswa yang memiliki tingkat disiplin belajar rendah dengan siswa yang memiliki tingkat disiplin belajar tinggi.

Adapun faktor-faktor yang dilibatkan dalam model SEM untuk melihat pengaruh terhadap prestasi belajar siswa adalah faktor internal (jasmaniah, psikologis, kelelahan) dan eksternal (lingkungan keluarga, lingkungan sekolah, lingkungan tempat tinggal). Analisis dilakukan menggunakan *software* AMOS, untuk mempermudah dalam pengembangan model yang melibatkan banyak indikator dalam menjelaskan faktor-faktor yang berpengaruh terhadap prestasi belajar matematika siswa.

Berdasarkan hasil analisis menggunakan SEM dengan metode *maximum likelihood* diperoleh faktor-faktor yang signifikan mempengaruhi prestasi belajar matematika siswa Sekolah Menengah Atas kecamatan Pasar Rebo Jakarta Timur yaitu psikologis, kelelahan dan lingkungan sekolah. Ketiga faktor tersebut berpengaruh secara signifikan terhadap prestasi hasil belajar matematika siswa ditinjau berdasarkan tingkat disiplin rendah dan tinggi. Dengan kata lain bahwa tingkat disiplin belajar siswa tidak memberikan perbedaan terhadap faktor-faktor yang mempengaruhi prestasi belajar matematika siswa.

Adapun indikator signifikan yang mampu menjelaskan variabel-variabel eksogen psikologis, kelelahan dan lingkungan sekolah ditinjau dari tingkat disiplin belajar siswa rendah dan tinggi dapat dilihat pada tabel 9 berikut.

Tabel 9. Indikator berdasarkan tingkat kedisiplinan belajar

Faktor yang signifikan	Indikator	
	Disiplin Belajar Rendah	Disiplin Belajar Tinggi
Psikologis	<ol style="list-style-type: none"> 1. Asal SLTP, 2. Nilai UN matematika, 3. Gaya belajar, 4. Lama belajar matematika diluar sekolah, 5. Cita-cita yang ingin dicapai 	<ol style="list-style-type: none"> 1. Asal SLTP, 2. Status SLTP 3. Nilai UN matematika, 4. Gaya belajar, 5. Lama belajar matematika diluar sekolah,
Kelelahan	<ol style="list-style-type: none"> 1. Kegiatan diluar sekolah 2. Keikutsertaan organisasi sekolah 3. Keikutsertaan ekstrakurikuler 4. Jenis kegiatan ekstrakurikuler 5. Jarak tempat tinggal dengan sekolah 	<ol style="list-style-type: none"> 1. Kegiatan diluar sekolah 2. Keikutsertaan organisasi sekolah 3. Keikutsertaan ekstrakurikuler 4. Jenis kegiatan ekstrakurikuler 5. Jarak tempat tinggal dengan sekolah
Lingkungan Sekolah	<ol style="list-style-type: none"> 1. Kenyamanan ruang kelas 2. Kelengkapan fasilitas sekolah 3. Kesesuaian materi pelajaran dengan ujian 4. Penyampaian materi 5. Intensitas pertemuan 6. Kesesuaian keahlian guru dengan matapelajaran yang diampu 	<ol style="list-style-type: none"> 1. Kenyamanan ruang kelas 2. Kelengkapan fasilitas sekolah 3. Kesesuaian materi pelajaran dengan ujian 4. Penyampaian materi 5. Intensitas pertemuan 6. Kesesuaian keahlian guru dengan matapelajaran yang diampu

Indikator-indikator signifikan yang mampu menjelaskan variabel psikologis ada perbedaan sedikit berdasarkan tingkat kedisiplinan rendah dan tinggi yaitu status SLTP dan cita-cita yang ingin dicapai. Sedangkan Indikator-indikator signifikan yang mampu menjelaskan variabel kelelahan dan lingkungan sekolah adalah sama untuk siswa yang memiliki tingkat kedisiplinan rendah dan tinggi. Dengan kata lain secara umum analisis faktor-faktor yang mempengaruhi prestasi belajar matematika siswa berdasarkan tingkat kedisiplinan rendah dan tinggi tidak memiliki perbedaan.

Variabel jasmaniah tidak signifikan mempengaruhi prestasi belajar, kemungkinan terjadi karena dari survei yang dilakukan terhadap 260 siswa, kurang dari 5% siswa yang mengidap penyakit kategori berat, sehingga mengakibatkan kurangnya informasi (data) prestasi belajar matematika bagi siswa yang mengidap penyakit berat atau yang mengidap beberapa penyakit. Selain itu kuantitas siswa yang memiliki kebiasaan merokok, sarapan pagi dan berolahraga sebagai indikator yang digunakan dalam menjelaskan variabel jasmaniah juga sangat sedikit (tidak proporsional).

Variabel lingkungan keluarga dan lingkungan tempat tinggal juga tidak signifikan mempengaruhi prestasi belajar matematika siswa Sekolah Menengah Atas kecamatan Pasar Rebo Jakarta Timur. Hal tersebut dapat dianalisa berdasarkan salah satu indikator variabel lingkungan keluarga yaitu kelengkapan orang tua. Kelengkapan orang tua tidak mempengaruhi prestasi belajar siswa, terlihat diantara 260 siswa yang telah disurvei, terdapat siswa berprestasi yang tidak memiliki bapak atau ibu. Sedangkan untuk indikator variabel lingkungan tempat tinggal, terdapat siswa yang bertempat tinggal dekat dengan keramaian seperti terminal bis dan pasar, namun mereka memiliki prestasi yang baik.

3. Kesimpulan

Faktor yang mempengaruhi prestasi belajar matematika siswa sekolah menengah atas kecamatan Pasar Rebo Jakarta Timur yang dilibatkan dalam model SEM dalam penelitian ini diantaranya adalah faktor internal yang meliputi jasmaniah, psikologis dan kelelahan, serta faktor eksternal meliputi lingkungan keluarga, lingkungan sekolah dan lingkungan tempat tinggal.

Faktor-faktor tersebut dianalisa berdasarkan tingkat kedisiplinan belajar siswa sedang dan tinggi. Hasil analisa dapat disimpulkan bahwa :

1. Tidak terdapat perbedaan faktor-faktor yang mempengaruhi prestasi belajar matematika siswa berdasarkan tingkat kedisiplinan belajar siswa sedang dan tinggi.
2. Faktor-faktor yang signifikan mempengaruhi prestasi belajar matematika siswa diantaranya adalah psikologis, kelelahan dan lingkungan sekolah
3. Tingkat kedisiplinan belajar tidak mempengaruhi prestasi hasil belajar matematika siswa

Model SEM yang dibangun dikategorikan model yang baik berdasarkan uji kelayakan model (*Goodness of Fit*)

Referensi

- [1] Borovik, A.V. and Gardiner, T., 2006, Mathematical Abilities and Mathematical Skills, *World Federation of National Mathematics Competitions Conference (2006)*, Cambridge, England, <http://www.maths.manchester.ac.uk/avb/pdf/abilities.pdf>
- [2] Ghozali, I., 2008. *Model Persamaan Struktural Konsep & Aplikasi dengan Program AMOS 16.0*. Semarang : Badan Penerbit Univ Diponegoro
- [3] Nurmawati, S., 2013, Penerapan Model Persamaan Struktural, *Skripsi*, Yogyakarta: UIN Sunan Kalijaga Yogyakarta
- [4] Paris, M.A., 2012, Model Persamaan Struktural Pengaruh Motivasi, Kapabilitas dan Lingkungan terhadap Prestasi Belajar Mahasiswa Tahun Pertama Program Studi S1 Matematika FMIPA-IPB, *Al-Adzka*, vol 2 no 2
- [5] Schumacker, R.E., 2010, *A Beginner's Guide to Structural Equation Modelling*, New York : Routledge Taylor & Francis Group
- [6] Slameto, 2003, *Belajar dan Faktor-Faktor yang mempengaruhinya*, Jakarta: PT Asdi Mahasatya.
- [7] Wibowo, M.S.A., 2010, Pemodelan Persamaan Struktural Faktor-Faktor yang Mempengaruhi Prestasi Belajar Mahasiswa IPB, *Skripsi*, Bogor: IPB.

PENELUSURAN RUMUS UJI KUBINGER SEBAGAI PADANAN ANOVA DUA ARAH (STUDI KASUS: PENGARUH FAKTOR METODE PREPARASI DAN FAKTOR MEREK TERHADAP KADAR TIMBAL PADA JAMU CINA)

FITRI CATUR LESTARI

Sekolah Tinggi Ilmu Statistik, fitricaturlestari@stis.ac.id

Abstrak. Metode statistik parametrik dapat digunakan jika data memenuhi asumsi tertentu. Jika asumsi tersebut tidak dapat dipenuhi maka solusi alternatifnya yaitu dengan menggunakan metode statistik nonparametrik. Beberapa uji dalam metode statistik parametrik mempunyai padanan dalam metode statistik nonparametrik. Salah satu uji dalam metode statistik parametrik yang biasa digunakan dalam *randomized block design (RBD)* dan *factorial design (FD)* adalah anova dua arah. Anova dua arah dalam RBD tidak menguji interaksi antara dua faktor sedangkan dalam RF mengujinya. Anova dua arah dalam RBD mempunyai padanan dalam metode statistik nonparametrik yang disebut uji Friedman dan anova dua arah dalam FD adalah uji Kubinger. Rumus dari uji Kubinger diturunkan dari uji dalam anova dua arah. Tujuan dari penelitian ini adalah untuk menelusuri asal rumus tersebut secara teori dan mengaplikasikannya dalam studi kasus di bidang farmasi yang ingin mengetahui pengaruh faktor metode preparasi dan faktor merek terhadap kadar timbal jamu cina. Berdasarkan hasil uji Kubinger dengan tingkat signifikansi 5% dapat disimpulkan bahwa interaksi antara faktor metode preparasi dan faktor merek terhadap kadar timbal tidak signifikan, perbedaan rata-rata kadar timbal jamu cina berdasarkan faktor metode preparasi tidak signifikan dan perbedaan rata-rata kadar timbal jamu cina berdasarkan faktor merek tidak signifikan juga.

Kata kunci: ANOVA dua arah, rancangan faktorial, Kubinger

1. Pendahuluan

Metode statistik parametrik dapat digunakan jika data memenuhi asumsi tertentu. Jika asumsi tersebut tidak dapat dipenuhi maka solusi alternatifnya yaitu dengan menggunakan metode statistik nonparametrik. Kelebihan dari metode statistik nonparametrik adalah menganalisis data yang tidak memenuhi asumsi dalam metode statistik parametrik. Kelebihan lain adalah dapat digunakan untuk sampel kecil, lebih sedikit asumsi yang dibutuhkan, dan dapat digunakan untuk data kategorik dalam Siegel [1]. Djarwanto [2] menyebutkan bahwa metode statistik nonparametrik digunakan untuk beberapa kondisi sebagai berikut: pola distribusi populasinya tidak diketahui, data nominal atau ordinal, ukuran sampel kecil dan karakteristik distribusi populasinya tidak diketahui dengan pasti.

Beberapa uji dalam metode statistik parametrik mempunyai padanan dalam metode statistik nonparametrik. Contohnya dalam metode statistik parametrik uji t digunakan untuk menguji beda rata-rata dari dua populasi independen, uji t berpasangan untuk menguji beda rata-rata dua populasi dependen dan analisis variansi yang padanannya di dalam metode statistik nonparametrik berturut-turut uji Wilcoxon-Mann-Whitney, uji Wilcoxon Signed Ranks dan uji Kruskal-Wallis dalam Siegel [1].

Dalam rancangan percobaan, ada dua rancangan yang populer: *Randomized Block Design* (RBD) dan rancangan faktorial atau *Factorial Design* (FD). RBD mempunyai karakteristik adanya blok atau kelompok, Setiaini blok atau kelompok dipengaruhi oleh beberapa perlakuan dalam Gaspersz [3]. Selain perlakuan, pengaruh *error* juga disebabkan oleh perbedaan blok. Sedangkan FD mempunyai karakteristik menguji interaksi antara faktor-faktor melalui kombinasi level atau taraf di samping juga menguji efek utamanya secara terpisah. Kedua rancangan tersebut menggunakan analisis yang disebut sebagai analisis dua arah.

Salah satu uji dalam metode statistik parametrik yang biasa digunakan dalam *randomized block design* (RBD) dan *factorial design* (FD) adalah anova dua arah. Anova dua arah dalam RBD tidak menguji interaksi antara dua faktor sedangkan dalam RF mengujinya. Anova dua arah dalam RBD mempunyai padanan dalam metode statistik nonparametrik yang disebut uji Friedman dan anova dua arah dalam FD adalah uji Kubinger. Kelebihan uji ini adalah dapat mengetahui efek interaksi yang tidak dapat ditemukan di uji Friedman. Rumus dari uji Kubinger diturunkan dari uji dalam anova dua arah. Tujuan dari penelitian ini adalah untuk menelusuri asal rumus tersebut secara teori dan mengaplikasikannya dalam studi kasus di bidang farmasi.

2. Teori dan Aplikasi

I. Teori Anova Dua Arah

Anova dua arah adalah suatu uji yang digunakan untuk mengetahui efek baris, efek kolom, dan efek interaksi terhadap variabel respon.

A. Model

Dalam Neter et al. [4] model yang digunakan dalam anova dua arah adalah:

$$Y_{ijk} = \mu_{...} + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk} \quad (2.1)$$

Y_{ijk} = observasi ke-k untuk faktor A ke-i dan faktor B ke-j

$\mu_{...}$ = konstanta

α_i = efek faktor A

β_j = efek faktor B

$(\alpha\beta)_{ij}$ = efek interaksi antara faktor A dan faktor B

ε_{ijk} = error

$i=1,2,\dots,a; j=1,2,\dots,b; k=1,2,\dots,n; N=abn$; dengan asumsi:

$$\sum_{i=1}^a \alpha_i = 0 \quad \sum_{j=1}^b \beta_j = 0 \quad \sum_{i=1}^a (\alpha\beta)_{ij} = 0 \quad \sum_{j=1}^b (\alpha\beta)_{ij} = 0 \quad ; \varepsilon_{ijk} \sim \text{NID}(0, \sigma^2) \quad (2.2)$$

B. Notasi

$$Y_{ij.} = \sum_{k=1}^n Y_{ijk} \rightarrow \bar{Y}_{ij.} = \frac{Y_{ij.}}{n}$$

$$Y_{i..} = \sum_{j=1}^b \sum_{k=1}^n Y_{ijk} \rightarrow \bar{Y}_{i..} = \frac{Y_{i..}}{bn}$$

$$Y_{.j.} = \sum_{i=1}^a \sum_{k=1}^n Y_{ijk} \rightarrow \bar{Y}_{.j.} = \frac{Y_{.j.}}{an}$$

$$Y_{...} = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n Y_{ijk} \rightarrow \bar{Y}_{...} = \frac{Y_{...}}{nab}$$

$$Y_{...} = \sum_{i=1}^a \sum_{j=1}^b Y_{ij.}$$

C. Estimasi Parameter

Berdasarkan model (2.1) maka dapat dituliskan:

$$\varepsilon_{ijk} = Y_{ijk} - \mu_{...} - \alpha_i - \beta_j - (\alpha\beta)_{ij}$$

Jika jumlahan *error* pada (2.2) dikuadratkan, maka persamaan yang terbentuk menjadi:

$$Q = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n \varepsilon_{ijk}^2 = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (Y_{ijk} - \mu_{...} - \alpha_i - \beta_j - (\alpha\beta)_{ij})^2 \quad (2.3)$$

Dengan metode *Least Square*, penduga parameternya sebagai berikut:

Tabel 2.1 Parameter dan Penduganya

Parameter	Penduga
$\mu_{...}$	$\hat{\mu}_{...} = \bar{Y}_{...}$
$\alpha_i = \mu_{i..} - \mu_{...}$	$\hat{\alpha}_i = \bar{Y}_{i..} - \bar{Y}_{...}$
$\beta_j = \mu_{.j.} - \mu_{...}$	$\hat{\beta}_j = \bar{Y}_{.j.} - \bar{Y}_{...}$
$(\alpha\beta)_{ij} = \mu_{ij.} - \mu_{i..} - \mu_{.j.} + \mu_{...}$	$(\hat{\alpha}\hat{\beta})_{ij} = \bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...}$

D. Jumlahan Kuadrat dan Rumus Komputasi

Berdasarkan model anova dua arah (2.1), estimasi modelnya menjadi:

$$\begin{aligned} \hat{Y}_{ijk} &= \hat{\mu}_{...} + \hat{\alpha}_i + \hat{\beta}_j + (\hat{\alpha}\hat{\beta})_{ij} \\ \hat{Y}_{ijk} &= \bar{Y}_{...} + (\bar{Y}_{i..} - \bar{Y}_{...}) + (\bar{Y}_{.j.} - \bar{Y}_{...}) + (\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...}) \\ \hat{Y}_{ijk} &= \bar{Y}_{ij.} \end{aligned} \quad (2.4)$$

Error merupakan perbedaan antara parameter dan penduganya. Sehingga *error* adalah $\varepsilon_{ijk} = Y_{ijk} - \hat{Y}_{ijk}$. Berdasarkan persamaan (2.4), *error* menjadi

$\varepsilon_{ijk} = Y_{ijk} - \bar{Y}_{ij.}$ Deviasi total adalah data dikurangi dengan rata-ratanya. Notasi dari deviasi total adalah $Y_{ijk} - \bar{Y}_{...}$. Deviasi ini dapat diuraikan sebagai berikut:

$$Y_{ijk} - \bar{Y}_{...} = Y_{ijk} - \bar{Y}_{ij.} + \bar{Y}_{ij.} - \bar{Y}_{...} \quad (2.5)$$

$$Y_{ijk} - \bar{Y}_{...} = \text{deviasi total} \quad (2.6)$$

$$Y_{ijk} - \bar{Y}_{ij.} = \text{deviasi error} \quad (2.7)$$

$$\bar{Y}_{ij.} - \bar{Y}_{...} = \text{deviasi perlakuan} \quad (2.8)$$

Deviasi perlakuan dapat diuraikan menjadi:

$$\bar{Y}_{ij.} - \bar{Y}_{...} = (\bar{Y}_{i..} - \bar{Y}_{...}) + (\bar{Y}_{.j.} - \bar{Y}_{...}) + (\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} - \bar{Y}_{...}) \quad (2.9) \bar{Y}_{ij.} - \bar{Y}_{...} =$$

deviasi perlakuan

$$(\bar{Y}_{i..} - \bar{Y}_{...}) = \text{deviasi efek faktor A} \quad (2.10)$$

$$(\bar{Y}_{.j.} - \bar{Y}_{...}) = \text{deviasi efek faktor B} \quad (2.11)$$

$$(\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} - \bar{Y}_{...}) = \text{deviasi efek interaksi faktor A dan B} \quad (2.12)$$

Jika deviasi yang telah disebutkan dalam (2.6) sampai (2.8) dan (2.10) sampai (2.12) dikuadratkan dan dijumlahkan, jumlahan kuadrat dan rumus komputasinya sebagai berikut:

Tabel 2.2 Rumus Jumlahan Kuadrat (*Sum Square*) dan Rumus Komputasi

<i>Sum Square</i>	Notasi/Rumus Umum	Rumus Umum
<i>Sum Square Total (SST)</i>	$\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (Y_{ijk} - \bar{Y}_{...})^2$	$\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n Y_{ijk}^2 - \frac{Y_{...}^2}{abn}$
<i>Sum Square Error (SSE)</i>	$\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (Y_{ijk} - \bar{Y}_{ij.})^2$	$\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n Y_{ijk}^2 - \frac{\sum_{i=1}^a \sum_{j=1}^b Y_{ij.}^2}{n}$
<i>Sum Square Treatment (SSTr)</i>	$\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{Y}_{ij.} - \bar{Y}_{...})^2$	$\frac{\sum_{i=1}^a \sum_{j=1}^b Y_{ij.}^2}{n} - \frac{Y_{...}^2}{abn}$
<i>Sum Square of A factor effect (SSA)</i>	$\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{Y}_{i..} - \bar{Y}_{...})^2$	$\frac{\sum_{i=1}^a Y_{i..}^2}{bn} - \frac{Y_{...}^2}{abn}$
<i>Sum Square of B factor effect (SSB)</i>	$\sum_{i=1}^a \sum_{b=1}^b \sum_{k=1}^n (\bar{Y}_{.j.} - \bar{Y}_{...})^2$	$\frac{\sum_{j=1}^b Y_{.j.}^2}{an} - \frac{Y_{...}^2}{abn}$
<i>Sum Square of A and B interaction effect (SSAB)</i>	$\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} - \bar{Y}_{...})^2$	$\frac{\sum_{i=1}^a \sum_{j=1}^b Y_{ij.}^2}{n} - \frac{\sum_{i=1}^a Y_{i..}^2}{bn} - \frac{\sum_{j=1}^b Y_{.j.}^2}{an} + \frac{Y_{...}^2}{abn}$

E. Derajat Bebas

Teorema yang digunakan dalam menentukan derajat bebas adalah:

Teorema 2.1.5.1

Central Limit Theorem menyatakan bahwa jika Y_1, \dots, Y_n adalah sampel acak dari sebuah distribusi dengan rata-rata μ dan $\sigma^2 < \infty$, maka pendekatan distribusi

$Z_n = \frac{\sum_{i=1}^n Y_i - n\mu}{\sqrt{n}\sigma}$ adalah normal standar, $Z_n \xrightarrow{d} Z \sim N(0,1)$ dimana $n \rightarrow \infty$. Selanjutnya Z_n dapat dituliskan sebagai $Z_n = \frac{\sqrt{n}(\bar{Y} - \mu)}{\sigma}$ dalam Bain dan Engelhardt [5].

Beberapa teorema dari Bain dan Engelhardt [5] dalam distribusi *Chi Square* distribution adalah:

Teorema 2.1.5.2

Jika $Z \sim N(0,1)$, maka $Z^2 \sim \chi_1^2$

Teorema 2.1.5.3

Jika $Y_i \sim \chi_{v_i}^2, i=1, \dots, n$, adalah variabel *Chi Square* yang independen maka

$$V = \sum_{i=1}^n Y_i \sim \chi_{\sum_{i=1}^n v_i}^2$$

Teorema 2.1.5.4

$\frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$ adalah efek dari teorema sebelumnya

$$V_1 = \sum_{i=1}^n \frac{(Y_i - \mu)^2}{\sigma^2} \sim \chi_n^2 \quad \text{dan} \quad V_3 = \frac{n(\bar{Y} - \mu)^2}{\sigma^2} \sim \chi_1^2$$

Dengan memeriksa hubungan ini:

$$\begin{aligned}
 V_1 &= V_2 + V_3 \\
 \sum_{i=1}^n \frac{(Y_i - \mu)^2}{\sigma^2} &= \sum_{i=1}^n \frac{(Y_i - \bar{Y})^2}{\sigma^2} + \sum_{i=1}^n \frac{(\bar{Y} - \mu)^2}{\sigma^2} \\
 \sum_{i=1}^n \frac{(Y_i - \mu)^2}{\sigma^2} &= \frac{(n-1)S^2}{\sigma^2} + \frac{n(\bar{Y} - \mu)^2}{\sigma^2}
 \end{aligned}$$

derajat bebas $V_1 =$ derajat bebas $V_2 +$ derajat bebas V_3

maka derajat bebas untuk $V_2 = \frac{(n-1)S^2}{\sigma^2}$ adalah $n-1$ dimana derajat bebas V_1 dikurangi dengan derajat bebas V_3 . Jika teorema diaplikasikan pada jumlahan

kuadrat, maka $\frac{JKP}{\sigma^2} \sim \chi_{ab-1}^2$, $\frac{JKA}{\sigma^2} \sim \chi_{a-1}^2$ dan $\frac{JKB}{\sigma^2} \sim \chi_{b-1}^2$, dan $\frac{JKAB}{\sigma^2} \sim \chi_{(a-1)(b-1)}^2$.

II. Rumus Tambahan Untuk Uji Kubinger

Rumus tambahan yang digunakan untuk menelusuri rumus dalam uji Kubinger didasari dari Conover [6] :

$$1. \quad \sum_{i=a}^N i = \frac{(N+a)(N-a+1)}{2}$$

$$2. \quad \sum_{i=1}^N i = \frac{(N+1)N}{2}$$

$$3. \quad \sum_{i=1}^N i^2 = \frac{N(N+1)(2N+1)}{6}$$

$$4. \quad \bar{R} = \frac{\sum_{i=1}^N R_i}{N} = \frac{1}{N} \frac{(N+1)N}{2} = \frac{(N+1)}{2}$$

$$5. \quad S^2 = \frac{\sum_{i=1}^N (R_i - \bar{R})^2}{N-1} = \frac{N(N+1)}{12}$$

6. S^2 adalah penduga tak bias untuk σ^2 atau $E(S^2) = \sigma^2$

Diketahui $E(R_i) = \mu$; $\text{var}(R_i) = \sigma^2$

$$E(R_i^2) - (E(R_i))^2 = \sigma^2$$

$$E(R_i^2) - \mu^2 = \sigma^2$$

$$E(R_i^2) = \sigma^2 + \mu^2$$

Sehingga:

$$E(\bar{R}) = E\left(\frac{\sum_{i=1}^N R_i}{N}\right) = \frac{E\left(\sum_{i=1}^N R_i\right)}{N} = \frac{\sum_{i=1}^N E(R_i)}{N} = \frac{N\mu}{N} = \mu$$

$$\text{var}(\bar{R}) = \text{var}\left(\frac{\sum_{i=1}^N R_i}{N}\right) = \frac{\text{var}\left(\sum_{i=1}^N R_i\right)}{N^2} = \frac{\sum_{i=1}^N \text{var}(R_i)}{N^2} = \frac{N\sigma^2}{N^2} = \frac{\sigma^2}{N}$$

$$\text{var}(\bar{R}) = \frac{\sigma^2}{N}$$

$$E(\bar{R}^2) - (E(\bar{R}))^2 = \frac{\sigma^2}{N}$$

$$E(\bar{R}^2) - \mu^2 = \frac{\sigma^2}{N}$$

$$E(\bar{R}^2) = \frac{\sigma^2}{N} + \mu^2$$

Rumus lain dari $S^2 = \frac{\sum_{i=1}^N (R_i - \bar{R})^2}{N-1}$ dapat dicari dengan:

$$\begin{aligned}
S^2 &= \frac{\sum_{i=1}^N (R_i - \bar{R})^2}{N-1} = \frac{\sum_{i=1}^N (R_i^2 - 2R_i\bar{R} + \bar{R}^2)}{N-1} \\
&= \frac{\sum_{i=1}^N R_i^2 - 2\bar{R}\sum_{i=1}^N R_i + \sum_{i=1}^N \bar{R}^2}{N-1} \\
&= \frac{\sum_{i=1}^N R_i^2 - 2N\bar{R}^2 + N\bar{R}^2}{N-1} = \frac{\sum_{i=1}^N R_i^2 - N\bar{R}^2}{N-1} \\
E(S^2) &= \sigma^2 \\
E(S^2) &= E\left(\frac{\sum_{i=1}^N (R_i - \bar{R})^2}{N-1}\right) = E\left(\frac{\sum_{i=1}^N R_i^2 - N\bar{R}^2}{N-1}\right) \\
&= \frac{E\left(\sum_{i=1}^N R_i^2\right) - NE(\bar{R}^2)}{N-1} = \frac{\sum_{i=1}^N E(R_i^2) - NE(\bar{R}^2)}{N-1} \\
&= \frac{NE(R_i^2) - NE(\bar{R}^2)}{N-1} = \frac{NE(R_i^2) - NE(\bar{R}^2)}{N-1} \\
&= \frac{N(\sigma^2 + \mu^2) - N\left(\frac{\sigma^2}{N} + \mu^2\right)}{N-1} = \frac{N\sigma^2 + N\mu^2 - \sigma^2 - N\mu^2}{N-1} \\
&= \frac{(N-1)\sigma^2}{N-1} = \sigma^2
\end{aligned}$$

Terbukti bahwa S^2 adalah penduga tak bias dari σ^2 .

III. Teori Uji Kibinger

Layout data untuk uji Kubinger secara umum terdapat pada tabel 2.3 berikut ini:

Tabel 2.3
Layout Data

		Faktor Kolom (j)			
		1	2	...	b
Faktor Baris (i)	1	Y_{11}	Y_{12}	...	Y_{1b}
	...	\vdots	\vdots		\vdots
	n	Y_{1n}	Y_{2n}		Y_{bn}

		Y_2	Y_2	...	Y_2
	11	\vdots	\vdots		\vdots
		Y_2	Y_2		Y_2
	$1n$				bn
		\vdots	\vdots		\vdots
		Y_a	Y_a	...	Y_a
	11	\vdots	\vdots		\vdots
		Y_a	Y_a		Y_a
	$1n$		$2n$		bn

A. Uji Hipotesis untuk Efek Baris (Faktor A)

- Hipotesis
Ho: efek baris tidak signifikan
Ha: efek baris signifikan
- Alpha= α

$$3. \text{ Statistik uji} = \frac{12}{a(N+1)} \sum_{i=1}^a (\bar{R}_{i..}^* - \bar{R}_{...}^*)^2$$

berdistribusi χ_{a-1}^2 dengan $N=abn$ dalam Hühn and Léon [7]

Semua nilai Y_{ijk} ditransformasi menjadi ranking tunggal $Y_{ijk} \rightarrow R_{ijk}$. Lalu ditransformasi menjadi $R_{ijk} \rightarrow R_{ijk}^t = R_{ijk} - \bar{R}_{ij.} + \bar{R}_{i..}$. R_{ijk}^t diurutkan menjadi R_{ijk}^* . Ketika membentuk ranking melalui transformasi tersebut, jika ada nilai yang sama pada Y_{ijk} , maka rankingnya adalah rata-rata dari ranking Y_{ijk} yang memiliki nilai yang sama tersebut. Sehingga terbentuk ranking R_{ijk}^t menjadi R_{ijk}^* .

Rumus statistik ujinya berasal dari:

$$\begin{aligned} \frac{JKA}{S^2} &= \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{R}_{i..}^* - \bar{R}_{...}^*)^2}{N(N+1)} \\ &= \frac{12}{N(N+1)} bn \sum_{i=1}^a (\bar{R}_{i..}^* - \bar{R}_{...}^*)^2 \\ &= \frac{12}{a(N+1)} \sum_{i=1}^a (\bar{R}_{i..}^* - \bar{R}_{...}^*)^2 \end{aligned}$$

Berdasarkan persamaan (2.10), deviasi faktor A adalah $(\bar{Y}_{i..} - \bar{Y}_{...})$. Jika diuraikan maka rumusnya menjadi:

$$\begin{aligned} (\bar{Y}_{i..} - \bar{Y}_{...}) &= (\bar{Y}_{i..} - \mu) + (\mu - \bar{Y}_{...}) \\ (\bar{Y}_{i..} - \bar{Y}_{...}) &= (\bar{Y}_{i..} - \mu) - (\bar{Y}_{...} - \mu) \end{aligned}$$

Jika dibuat jumlahan kuadrat dan dibagi dengan σ^2 pada kedua sisi, maka:

$$\begin{aligned}
& \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{Y}_{i..} - \bar{Y}_{...})^2}{\sigma^2} = \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{Y}_{i..} - \mu)^2}{\sigma^2} - \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{Y}_{...} - \mu)^2}{\sigma^2} \\
& = \frac{bn \sum_{i=1}^a (\bar{Y}_{i..} - \mu)^2}{\sigma^2} - \frac{abn(\bar{Y}_{...} - \mu)^2}{\sigma^2} \\
& = \sum_{i=1}^a \left(\frac{\bar{Y}_{i..} - \mu}{\frac{\sigma}{\sqrt{bn}}} \right)^2 - \left(\frac{\bar{Y}_{...} - \mu}{\frac{\sigma}{\sqrt{abn}}} \right)^2
\end{aligned}$$

Berdasarkan teorema 2.1.5.4,

$$\sum_{i=1}^a \left(\frac{\bar{Y}_{i..} - \mu}{\frac{\sigma}{\sqrt{bn}}} \right)^2 \sim \chi_a^2 \quad \text{dan} \quad \left(\frac{\bar{Y}_{...} - \mu}{\frac{\sigma}{\sqrt{abn}}} \right)^2 \sim \chi_1^2$$

maka derajat bebas untuk $\frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{Y}_{i..} - \bar{Y}_{...})^2}{\sigma^2}$ adalah $a - 1$. Sehingga,

$$\frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{Y}_{i..} - \bar{Y}_{...})^2}{\sigma^2} \sim \chi_{a-1}^2 \Leftrightarrow \frac{JKA}{\sigma^2} \sim \chi_{a-1}^2$$

- Daerah kritis: H_0 ditolak jika statistik hitung lebih dari nilai Chi Square tabel (χ_{a-1}^2)

B. Uji Hipotesis untuk Efek Kolom (Faktor B)

- Hipotesis
 H_0 : efek kolom tidak signifikan
 H_a : efek kolom signifikan
- Alpha= α

- Statistik uji = $\frac{12}{b(N+1)} \sum_{j=1}^b (\bar{R}^*_{.j} - \bar{R}^*_{...})^2$
berdistribusi χ_{b-1}^2 dengan $N=abn$ dalam Hühn and Léon [7]

Semua nilai Y_{ijk} ditransformasi menjadi ranking tunggal $Y_{ijk} \rightarrow R_{ijk}$. Lalu ditransformasi menjadi $R_{ijk} \rightarrow R^t_{ijk} = R_{ijk} - \bar{R}_{ij.} + \bar{R}_{.j.}$. R^t_{ijk} diurutkan menjadi R^*_{ijk} . Ketika membentuk ranking melalui transformasi tersebut, jika ada nilai yang sama pada Y_{ijk} maka rankingnya adalah rata-rata dari ranking Y_{ijk} yang memiliki nilai yang sama tersebut. Sehingga terbentuk ranking R^t_{ijk} menjadi R^*_{ijk} .

Rumus statistik ujinya berasal dari:

$$\begin{aligned} \frac{JKB}{S^2} &= \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{R}^*_{.j.} - \bar{R}^*_{...})^2}{N(N+1)} \\ &= \frac{12}{N(N+1)} an \sum_{j=1}^b (\bar{R}^*_{.j.} - \bar{R}^*_{...})^2 \\ &= \frac{12}{b(N+1)} \sum_{j=1}^b (\bar{R}^*_{.j.} - \bar{R}^*_{...})^2 \end{aligned}$$

Berdasarkan persamaan (2.11), deviasi faktor A adalah $(\bar{Y}_{.j.} - \bar{Y}_{...})$. Jika diuraikan maka rumusnya menjadi:

$$\begin{aligned} (\bar{Y}_{.j.} - \bar{Y}_{...}) &= (\bar{Y}_{.j.} - \mu) + (\mu - \bar{Y}_{...}) \\ (\bar{Y}_{.j.} - \bar{Y}_{...}) &= (\bar{Y}_{.j.} - \mu) - (\bar{Y}_{...} - \mu) \end{aligned}$$

Jika dibuat jumlahan kuadrat dan dibagi dengan σ^2 pada kedua sisi, maka:

$$\begin{aligned} \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{Y}_{.j.} - \bar{Y}_{...})^2}{\sigma^2} &= \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{Y}_{.j.} - \mu)^2}{\sigma^2} - \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{Y}_{...} - \mu)^2}{\sigma^2} \\ &= an \sum_{j=1}^b \frac{(\bar{Y}_{.j.} - \mu)^2}{\sigma^2} - \frac{abn(\bar{Y}_{...} - \mu)^2}{\sigma^2} \\ &= \sum_{j=1}^b \left(\frac{\bar{Y}_{.j.} - \mu}{\frac{\sigma}{\sqrt{an}}} \right)^2 - \left(\frac{\bar{Y}_{...} - \mu}{\frac{\sigma}{\sqrt{abn}}} \right)^2 \end{aligned}$$

Berdasarkan teorema 2.1.5.4,

$$\begin{aligned} \sum_{j=1}^b \left(\frac{\bar{Y}_{.j.} - \mu}{\frac{\sigma}{\sqrt{an}}} \right)^2 &\sim \chi_b^2 \quad \text{dan} \quad \left(\frac{\bar{Y}_{...} - \mu}{\frac{\sigma}{\sqrt{abn}}} \right)^2 \sim \chi_1^2 \\ &\frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{Y}_{.j.} - \bar{Y}_{...})^2}{\sigma^2} \end{aligned}$$

maka derajat bebas untuk σ^2 adalah $b-1$. Sehingga,

$$\frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{Y}_{.j.} - \bar{Y}_{...})^2}{\sigma^2} \sim \chi_{b-1}^2 \Leftrightarrow \frac{JKB}{\sigma^2} \sim \chi_{b-1}^2$$

4. Daerah kritis: H_0 ditolak jika statistik hitung lebih dari nilai Chi Square tabel (χ_{b-1}^2)

C. Uji Hipotesis untuk Efek Interaksi (Faktor A dan B)

1. Hipotesis

Ho: efek interaksi faktor baris dan faktor kolom tidak signifikan

Ha: efek interaksi faktor baris dan faktor kolom signifikan

2. Alpha= α

$$3. \text{ Statistik uji} = \frac{12}{ab(N+1)} \sum_{i=1}^a \sum_{j=1}^b (\bar{R}_{ij}^* - \bar{R}_{i..}^* - \bar{R}_{.j.}^* - \bar{R}_{...}^*)^2$$

berdistribusi $\chi^2_{(a-1)(b-1)}$ dengan $N=abn$ dalam Hühn and Léon [7]

Semua nilai Y_{ijk} ditransformasi menjadi ranking tunggal $Y_{ijk} \rightarrow R_{ijk}$. Lalu ditransformasi menjadi $R_{ijk} \rightarrow R'_{ijk} = R_{ijk} - \bar{R}_{i..} + \bar{R}_{.j.}$. R'_{ijk} diurutkan menjadi R^*_{ijk} . Ketika membentuk ranking melalui transformasi tersebut, jika ada nilai yang sama pada Y_{ijk} , maka rankingnya adalah rata-rata dari ranking Y_{ijk} yang memiliki nilai yang sama tersebut. Sehingga terbentuk ranking R'_{ijk} menjadi R^*_{ijk} .

Rumus statistik ujinya berasal dari:

$$\begin{aligned} \frac{JKAB}{S^2} &= \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{R}_{ij}^* - \bar{R}_{i..}^* - \bar{R}_{.j.}^* - \bar{R}_{...}^*)^2}{N(N+1)} \\ &= \frac{12}{N(N+1)} n \sum_{i=1}^a \sum_{j=1}^b (\bar{R}_{ij}^* - \bar{R}_{i..}^* - \bar{R}_{.j.}^* - \bar{R}_{...}^*)^2 \\ &= \frac{12}{ab(N+1)} \sum_{i=1}^a \sum_{j=1}^b (\bar{R}_{ij}^* - \bar{R}_{i..}^* - \bar{R}_{.j.}^* - \bar{R}_{...}^*)^2 \end{aligned}$$

Berdasarkan persamaan (2.14), (2.23), (2.24), dan (2.25), maka

$$JKP = JKA + JKB + JKAB$$

$$\frac{JKP}{\sigma^2} = \frac{JKA}{\sigma^2} + \frac{JKB}{\sigma^2} + \frac{JKAB}{\sigma^2}$$

$$\frac{JKAB}{\sigma^2} = \frac{JKP}{\sigma^2} - \frac{JKA}{\sigma^2} - \frac{JKB}{\sigma^2}$$

Karena $\frac{JKP}{\sigma^2} \sim \chi^2_{ab-1}$, $\frac{JKA}{\sigma^2} \sim \chi^2_{a-1}$ dan $\frac{JKB}{\sigma^2} \sim \chi^2_{b-1}$, maka derajat bebas

$$\begin{aligned} \text{untuk } \frac{JKAB}{\sigma^2} \text{ adalah} \\ &= ab - 1 - (a - 1) - (b - 1) \\ &= ab - 1 - a + 1 - b + 1 \\ &= ab - a - b + 1 \end{aligned}$$

$$= (a - 1)(b - 1). \text{ Sehingga, } \frac{JKAB}{\sigma^2} \sim \chi^2_{(a-1)(b-1)}$$

$$\Leftrightarrow \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} - \bar{Y}_{...})^2}{\sigma^2} \sim \chi^2_{(a-1)(b-1)}$$

5. Daerah kritis: Ho ditolak jika statistik hitung lebih dari nilai Chi Square tabel ($\chi^2_{(a-1)(b-1)}$)

IV. Aplikasi Uji Kubinger

Penelitian ini merupakan aplikasi dari metode statistik nonparametrik pada bidang farmasi. Percobaan dilakukan oleh Ajeng Sukma Ratnapani [8] dari jurusan farmasi di Universitas Islam Indonesia. Tujuan dari penelitiannya adalah untuk mengetahui efek interaksi antara faktor metode preparasi dengan faktor merek terhadap kadar timbal pada jamu cina serta efek utama dari kedua faktor tersebut.

Ukuran sampel dalam percobaan tersebut sangat kecil yaitu hanya dua untuk setiap kombinasi perlakuan dikarenakan harga sampel yang sangat mahal. Jamu cina lebih mahal dibandingkan dengan jamu yang lain dan pengukuran kadar timbal dengan Spektrofotometri Serapan Atom (SSA) juga memerlukan biaya yang besar.

Informasi mengenai distribusi populasi tidak diketahui dengan pasti meskipun uji asumsi normalitas sudah dilakukan. Hasil uji normalitas bisa jadi diragukan jika ukuran sampelnya kecil dalam Conover [6]. Oleh karena itu uji Kubinger tepat digunakan pada kasus ini.

Tabel berikut ini adalah data hasil percobaan di bidang farmasi:

Tabel 2.4
Kadar Timbal dalam Jamu Cina

		Faktor Merek (j)		
		1	2	3
Faktor Metode Preparasi (i)	Destruksi Basah	13,9056	7,5989	8,1710
		17,6000	6,7123	6,5648
	Destruksi Kering	12,5650	11,7000	8,2400
		15,1600	9,1050	7,3750

Berdasarkan hasil uji Kubinger, dengan tingkat signifikansi 5% dapat disimpulkan bahwa:

1. interaksi antara faktor merek dan faktor metode preparasi terhadap kadar timbal jamu cina tidak signifikan karena nilai statistik hitung (0.782051) kurang dari nilai Chi Square tabel (5.99146).
2. perbedaan rata-rata kadar timbal jamu cina berdasarkan faktor merek tidak signifikan karena nilai statistik hitung (1.88462) kurang dari nilai Chi Square tabel (5.99146), dan
3. perbedaan rata-rata kadar timbal jamu cina berdasarkan faktor metode preparasi tidak signifikan karena nilai statistik hitung (0.025641) kurang dari nilai Chi Square tabel (3.84146).

3. Kesimpulan

Uji Kubinger dapat menguji efek interaksi antara faktor A dan B dengan

statistik uji $\frac{12}{ab(N+1)} \sum_{i=1}^a \sum_{j=1}^b (\bar{R}_{ij}^* - \bar{R}_{i..}^* - \bar{R}_{.j.}^* - \bar{R}_{...}^*)^2$ yang berdistribusi

$\chi_{(a-1)(b-1)}^2$, efek faktor A dengan statistik uji $\frac{12}{a(N+1)} \sum_{i=1}^a (\bar{R}_{i..}^* - \bar{R}_{...}^*)^2$ yang

berdistribusi χ_{a-1}^2 , efek faktor B dengan statistik uji $\frac{12}{b(N+1)} \sum_{j=1}^b (\bar{R}_{.j.}^* - \bar{R}_{...}^*)^2$

yang berdistribusi χ_{b-1}^2 . Berdasarkan hasil uji Kubinger dengan tingkat signifikansi 5% dapat disimpulkan bahwa interaksi antara faktor metode preparasi dan faktor merek terhadap kadar timbal tidak signifikan, perbedaan rata-rata kadar timbal jamu cina berdasarkan faktor metode preparasi tidak signifikan dan perbedaan rata-rata kadar timbal jamu cina berdasarkan faktor merek tidak signifikan juga.

Pernyataan terima kasih. Segala puji bagi Allah SWT yang telah memberikan segala yang terbaik. Terima kasih kepada orang tua dan keluarga dan semua orang yang ada di hati dan pikiran penulis. Semoga kita semua selamat dunia akhirat. Aamiin.

Referensi

- [1] Siegel, S., Castellan, N.J. 1988. *Nonparametric Statistics for The Behavioral Science*, 2nd Edition, McGraw-Hill, Singapore.
- [2] Djarwanto, Ps. 2003. *Statistik Nonparametrik Edisi 2003/2004*, BPFE, Yogyakarta.
- [3] Gaspersz, V. 1995. *Teknik Analisis dalam Penelitian Percobaan*, Edisi 1, Tarsito, Bandung.
- [4] Neter, J., Wasserman, W., Kutner, M.H. 1990. *Applied Linier Statistical Models: Regression, Analysis of Variance, and Experimental Design*, 3rd Edition, Rhicard D. Irwin, Inc., Boston.
- [5] Bain, L.J., Engelhardt, M. 1991. *Introduction to Probability and Mathematical Statistics*, 2nd Edition, Duxbury Press, California.
- [6] Conover, W.J. 1980. *Practical Nonparametric Statistics*, 2nd Edition, John Wiley&Sons, Inc., Canada.
- [7] Hühn, M., Léon, J. 1995. *Nonparametric Analysis of Cultivar Performance Trials: Experimental Results and Comparison of Different Procedures Based on Ranks*, Agronomy Journal, vol. 87, 627-632. Atau Hühn, M., Léon, J., *Nonparametric Analysis of Cultivar Performance Trials - Experimental Results And Comparison of Different Procedures Based On Ranks*, URL: <http://www.scilib.univ.kiev.ua/article.php?1091042> .1995.
- [8] Ratnapani, A.S.. 2004. *Penetapan Kadar Logam Timbal (Pb) dan Khromium (Cr) pada Jamu Cina Berbagai Merek dalam Kemasan Kaleng yang Beredar di Pasaran dengan Metode Spektrofotometer Serapan Atom*, UII Yogyakarta.

EFEKTIVITAS ORGANISASI TATA KELOLA MINYAK DAN GAS BUMI DENGAN PEMODELAN STATISTIK “*STRUCTURAL EQUATION MODELLING (SEM)*” (STUDI KASUS DI KABUPATEN KUTAI KARTANEGARA)

RADITA ARINDYA¹

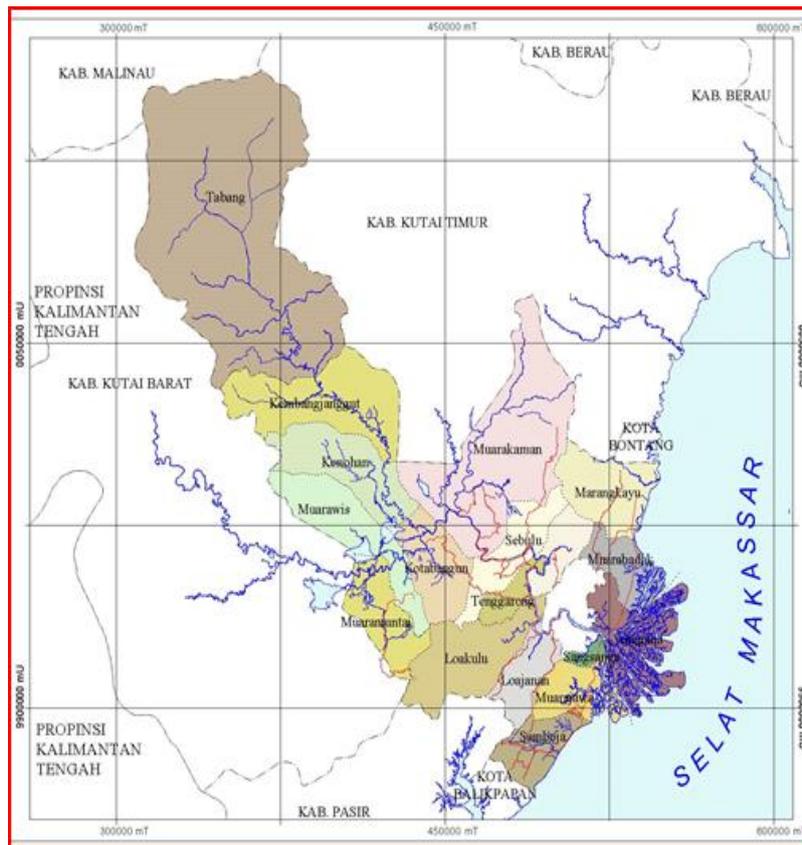
¹ Universitas Satyagama, Jl. Kamal Raya No. 2A Jakarta Barat, email: raditatch@yahoo.com

Abstrak. Migas merupakan komoditas vital yang menguasai hajat hidup orang banyak tetapi pengelolaan migas belum dapat meningkatkan kemakmuran dan kesejahteraan rakyat terutama di daerah penghasil Kabupaten Kutai Kartanegara. Kabupaten Kutai Kartanegara merupakan daerah yang kaya akan sumber daya alam terutama minyak dan gas bumi. Ketersediaan migas seharusnya dapat dinikmati oleh penduduk Kutai Kartanegara sebagai daerah penghasil migas. Tujuan penelitian ini adalah meneliti dan menganalisis besar faktor yang berpengaruh terhadap efektifitas organisasi dalam tata kelola minyak dan gas bumi, yaitu pengaruh kepemimpinan, koordinasi, pengawasan dan partisipasi masyarakat, penelitian tersebut merupakan penelitian eksplanasi dengan pendekatan kuantitatif, metode penelitian yang digunakan adalah metode survei dengan pendekatan deskriptif dan verifikatif atau hipotesis testing. Alokasi sampel dilakukan secara proporsional dengan rumus Slovin agar dapat mewakili berbagai unsur masyarakat yang termasuk dalam populasi (proportionate sampling) sebanyak 260 orang dari populasi 275 orang. Jenis data yang digunakan berbentuk skala ordinal, kemudian dilakukan pengujian hipotesis dengan menggunakan alat uji model persamaan struktural “Structural Equation Modeling (SEM)”, melalui analisis CFA (Confirmatory Factor Analysis). Hipotesis membuktikan terdapat pengaruh signifikan dan positif Peranan Kepemimpinan, Koordinasi, Pengawasan dan Partisipasi Masyarakat terhadap Efektifitas Organisasi dalam Tata Kelola Migas di Kabupaten Kutai Kartanegara. Berdasarkan hasil penelitian dapat disimpulkan bahwa terdapat pengaruh signifikan dan positif pada semua yang secara parsial Peranan Kepemimpinan terhadap Efektifitas Organisasi sebesar 58%, Koordinasi terhadap Efektifitas Organisasi sebesar 77%, Pengawasan terhadap Efektifitas Organisasi sebesar 94%, Partisipasi Masyarakat terhadap Efektifitas Organisasi sebesar 97%.

Kata Kunci: Kepemimpinan, koordinasi, pengawasan, partisipasi masyarakat dan efektivitas manajemen.

1. Pendahuluan

Kabupaten Kutai Kartanegara merupakan sebuah kabupaten di Kalimantan Timur, Indonesia. Ibukota berada di Kecamatan Tenggarong. Kabupaten Kutai Kartanegara memiliki luas wilayah 27.263,10 km² dan luas perairan sekitar 4.097 km² yang secara geografis terletak antara 115°26'28" BT – 117°36'43" BT dan 1°28'21" LU – 1°08'06" berdasarkan data BPS Provinsi Kalimantan Timur. [7]



Gambar 1. Peta Kabupaten Kutai Kartanegara

Pemakai gas alam dari kabupaten Kutai Kartanegara pada tahun 2015, terdiri dari 54.9% untuk ekspor dalam bantu gas alam yang dicairkan (*Natural Gas Liquefaction*), untuk pupuk atau petrokimia sebesar 24.4%, ekspor ke region lain/ Jawa Barat (melalui RFSU) sebesar 9.5% dan sisanya dipakai untuk listrik, pemakaian produksi LNG dan pemakaian gas rumah tangga. Total kapasitas produksi dari semua sumur migas mencapai 134.626 barrel per hari, yang terdiri dari 60,331 barrel minyak mentah dan 74,925 barrel gas alam setara minyak bumi. Pemanfaatan migas untuk listrik di Kabupaten Kutai kartanegara masih rendah karena rasio elektrifikasi atau rasio kepala keluarga yang menikmati listrik hanya 68,14 persen. Kelangkaan dalam penyediaan energi dan 310ketidakterjangkauan harga energi oleh masyarakat akan menciptakan instabilitas sosial. Ketika suatu daerah mengalami situasi kelangkaan energi, baik karena keterbatasan komoditas maupun karena kenaikan harga energi, maka sektor lain seperti ekonomi, sosial dan politik juga akan terpengaruh. Hal tersebut ditulis oleh Demirbas (2004), “*Energy affects all aspects of modern life. The demand for energy is increasing at an exponential rate due to the significant growth of the world population*”. Dengan kata lain, migas sebagai sumber energi memiliki keterkaitan erat terhadap aspek kehidupan lainnya.

2. Metode Penelitian

Penelitian ini merupakan penelitian eksplanasi (*eksplanatory research*) dengan pendekatan kuantitatif. Penelitian eksplanasi merupakan penelitian yang masih pada tahap identifikasi. Peneliti akan dihadapkan pada masalah-masalah yang belum familiar dan bertujuan untuk lebih mengetahui dan mengembangkan konsep sesuai dengan keadaan dilapangan.

Tahapan tersebut menggunakan pendekatan yang dikemukakan oleh Sugiyono (2014), yaitu menentukan :

1. Rancangan penelitian (*research design*)
2. Populasi, sampel, besar sampel dan teknik pengambilan sampel.
3. Variabel penelitian, klasifikasi variabel dan definisi operasional variabel.
4. Instrumen penelitian.
5. Lokasi dan waktu penelitian.
6. Prosedur pengumpulan data.
7. Teknik analisis data.

Tujuan dari sebuah penelitian sosial memiliki beberapa alternatif, tiga diantaranya yaitu melakukan eksplorasi atas topik yang baru, mendeskripsikan suatu fenomena sosial dan menjelaskan mengapa suatu fenomena itu terjadi.

Berkenaan dengan hal di atas tersebut, penelitian ini menelaah efektivitas alternatif upaya manajemen sumber daya energi migas, melalui faktor *good governance* dan partisipasi masyarakat. Diduga bahwa efektivitas organisasi dalam tata kelola migas dipengaruhi oleh empat variabel utama, yaitu variabel kepemimpinan, koordinasi, pengawasan (*good governance*) dan variabel partisipasi masyarakat. Dalam penelitian ini, variabel kepemimpinan, koordinasi, pengawasan (*good governance*) dan variabel partisipasi masyarakat merupakan variabel bebas (*independent*), sedangkan efektivitas organisasi dalam tata kelola migas merupakan variabel terikat (*dependent*).

3. Structural Equation Model (SEM).

Ada 4 variabel yang digunakan dalam penelitian ini. Keempat variabel tersebut adalah kepemimpinan, koordinasi, pengawasan (*good governance*) dan variabel partisipasi masyarakat yang merupakan gambaran dari fenomena sosial, maka dianggap sebagai konstruk atau konsep abstrak yang berupa variabel laten, yang untuk pengukurannya harus didekati dengan variabel manifest berupa indikator-indikator yang merupakan refleksi dari variabel laten di atas. Demikian pula halnya dengan efektivitas organisasi dalam tata kelola migas merupakan variabel laten yang harus didekati dengan variabel manifest yang merupakan refleksi dari variabel laten tersebut. Untuk mencapai tujuan dari penelitian ini, ada dua permasalahan utama yang akan diperhatikan, yaitu masalah pengukuran dan masalah hubungan kausal antar variabel (antara variabel laten dan antara variabel laten dengan variabel terukurannya). Kedua permasalahan tersebut dapat dipecahkan dengan model persamaan struktural atau *Structural Equation Model (SEM)*.

Karakteristik SEM menurut Gazali (2008) dapat diuraikan kedalam komponen-komponen model SEM yang terdiri dari :

1. Dua jenis Variabel, yaitu Variabel Laten dan Variabel Terukur.
2. Dua jenis Model, yaitu Model Struktural (*Structural Model*) dan Model Pengukuran (*Measurement Model*).
3. Dua jenis kesalahan, yaitu Kesalahan Struktural (*Structural Error*) dan

Kesalahan Pengukuran (*Measurement Error*).

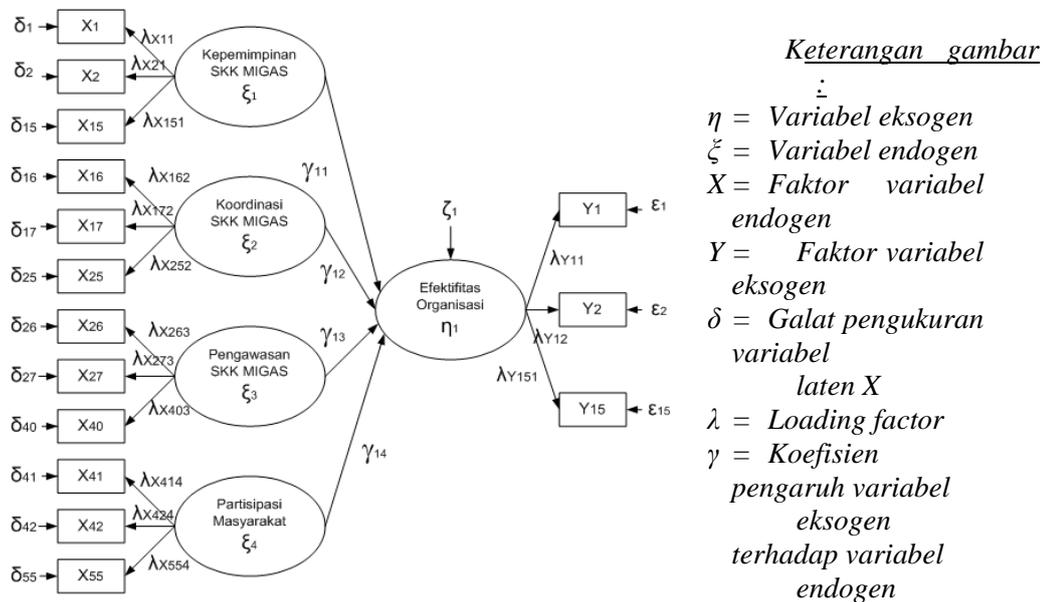
Menurut Slovin dalam Sugiyono (2007), untuk ukuran sampel dapat menggunakan rumus Solvin berikut :

$$n = \frac{N}{1 + Ne^2}$$

Dimana :
 n = ukuran sampel terpilih
 N = banyaknya unit dalam populasi
 e = besarnya *relative error*

Kedua model di atas, yaitu model struktural dan model pengukuran mempunyai karakteristik yang berbeda dengan model regresi biasa. Regresi biasa umumnya men-spesifikasikan hubungan kausal antara variabel-variabel teramati, sedangkan pada model variabel laten SEM, hubungan kausal terjadi antara variabel-variabel yang tidak teramati atau variabel laten. Gazali (2008) menunjukkan bahwa penggunaan variabel-variabel laten pada regresi berganda menimbulkan kesalahan-kesalahan pengukuran yang berpengaruh pada estimasi parameter dan besar kecilnya variansi (*variance*).

Gambar 2 adalah diagram yang memperlihatkan hubungan kausal antar variabel laten (variabel eksogen) dan antara variabel laten dan variable terukur (variabel endogen).



Gambar 2. *Structural Equation Model*

Persamaan umum regresi linear sederhana adalah:

- a. $\hat{Y} = a_1 + \beta_1 X_1$
- b. $\hat{Y} = a_2 + \beta_2 X_2$
- c. $\hat{Y} = a_3 + \beta_3 X_3$
- d. $\hat{Y} = a_4 + \beta_4 X_4$

Persamaan regresi linear berganda adalah :

$$\hat{Y} = a + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4$$

Dimana :

- \hat{Y} = variable terikat, yaitu efektifitas manajemen penyediaan sumber daya energi minyak dan gas bumi di Kabupaten Kutai Kartanegara
- α = konstanta
- β_1 = koefisien regresi variabel X_1
- β_2 = koefisien regresi variabel X_2
- β_3 = koefisien regresi variabel X_3
- β_4 = koefisien regresi variabel X_4
- X_1 = Variabel bebas Kepemimpinan
- X_2 = variabel bebas Koordinasi
- X_3 = variabel bebas Pengawasan
- X_4 = variabel bebas Partisipasi masyarakat

Pada model di atas, keempat variabel exogeneous ξ_1, ξ_2, ξ_3 dan ξ_4 diduga hanya berpengaruh langsung pada variabel endogeneous η_1 (efektivitas organisasi dalam tata kelola migas). Tujuan model persamaan struktural (SEM) adalah untuk menguji apakah model yang diusulkan dalam diagram jalur di atas (model teoritis) sesuai/cocok atau tidak sesuai/tidak cocok dengan data empiris yang diperoleh dari penelitian. Evaluasi terhadap kinerja model tersebut dilakukan secara menyeluruh, dimana secara teknis statistik harus dilakukan pengujian, apakah matriks kovarians teoritik identik dengan matriks kovarians empiris. Yang dimaksud dengan pengujian disini adalah uji kesesuaian (*goodness of fit*), yaitu kesesuaian antara struktur matriks varians-kovarians model teoritis dengan struktur matriks varians-kovarians model empiris. Jika kedua matriks varians-kovarians tersebut identik, maka model teoritis tersebut disimpulkan diterima/sesuai/fit secara sempurna. Apabila ternyata hasil uji kesesuaian tidak menghasilkan kecocokan yang diharapkan maka harus dilakukan re-spesifikasi atau pemodelan kembali, agar diperoleh model yang sesuai dengan hasil penelitian empiris.

4. Pengujian dan Analisis Regresi Linier

4.1. Analisis Determinasi

Untuk menguji kesesuaian model regresi linier digunakan Koefisien Determinasi (R^2) untuk menerangkan seberapa besar variabilitas pada variabel terikat (*dependent*) dari model persamaan regresi linear, terutama dari variabel bebas (*independent*) yang dipergunakan. Perhitungan tersebut dilakukan dengan komputer melalui program atau perangkat lunak “SPSS for Window versi 8.0”.

a. Uji Hipotesis

1. Uji F

Uji F dilakukan untuk menguji ada atau tidaknya pengaruh variabel X_1, X_2 dan X_3 secara parsial maupun bersama-sama terhadap variabel Y. Pengujian tersebut dilakukan dengan komputer melalui program atau perangkat lunak “SPSS for Window versi 8.0”, baik perhitungannya maupun kriteria keputusan dari uji hipotesis (H_0 ditolak, bila $\text{Sig} < \alpha$, yaitu terdapat pengaruh variabel bebas (*independent*) secara parsial maupun bersama-sama terhadap variabel terikat (*dependent*), atau H_1 diterima: H_0 diterima, bila nilai $\text{Sig} \geq \alpha$, yaitu tidak terdapat pengaruh variabel bebas secara parsial maupun bersama-sama terhadap variabel terikat atau H_1 ditolak.

2. Uji t

Uji t dilakukan untuk menguji pengaruh variabel (koefisien regresi) X1, X2 dan X3 terhadap variabel Y secara individual (parsial). Pengujian tersebut dilakukan dengan komputer melalui program atau perangkat lunak “SPSS for Window versi 8.0”, baik untuk perhitungan maupun kriteria keputusan dari uji hipotesis (H_0 ditolak, bila $Sig \leq \alpha$, yaitu koefisien regresi signifikan atau H_1 diterima; H_0 diterima, bila nilai $Sig \geq \alpha$, yaitu koefisien regresi tidak signifikan atau H_1 ditolak).

b. Rancangan Uji Hipotesis.

Hipotesis yang akan diuji dalam penelitian ini dapat dirumuskan sebagai berikut :

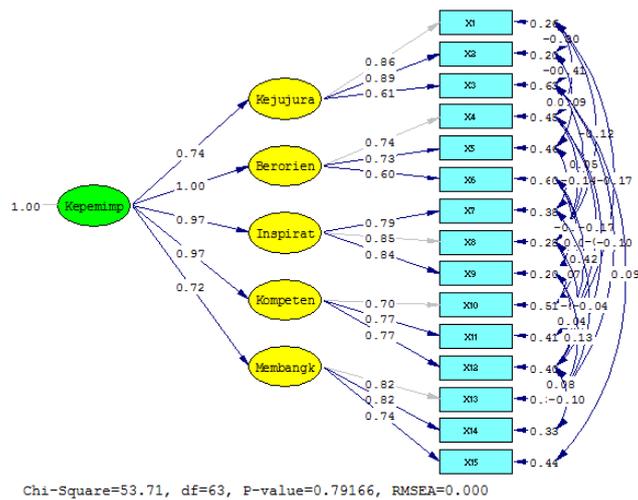
1. $H_0 : \beta_1 = 0$, tidak terdapat pengaruh kepemimpinan terhadap efektifitas manajemen penyediaan sumber daya energi migas di Kabupaten Kutai Kartanegara.
 $H_1 : \beta_1 \neq 0$, terdapat kepemimpinan terhadap efektifitas manajemen penyediaan sumber daya energi migas di Kabupaten Kutai Kartanegara.
2. $H_0 : \beta_2 = 0$, tidak terdapat pengaruh koordinasi terhadap efektifitas manajemen penyediaan sumber daya energi migas di Kabupaten Kutai Kartanegara.
 $H_2 : \beta_2 \neq 0$, terdapat pengaruh koordinasi terhadap efektifitas manajemen penyediaan sumber daya energi migas di Kabupaten Kutai Kartanegara.
3. $H_0 : \beta_3 = 0$, tidak terdapat pengaruh pengawasan terhadap efektifitas manajemen penyediaan sumber daya energi migas di Kabupaten Kutai Kartanegara.
 $H_3 : \beta_3 \neq 0$, terdapat pengaruh pengawasan terhadap efektifitas manajemen penyediaan sumber daya energi migas di Kabupaten Kutai Kartanegara.
4. $H_0 : \beta_4 = 0$, tidak terdapat pengaruh partisipasi masyarakat terhadap efektifitas manajemen penyediaan sumber daya energi migas di Kabupaten Kutai Kartanegara.
 $H_4 : \beta_4 \neq 0$, terdapat pengaruh partisipasi masyarakat terhadap efektifitas manajemen penyediaan sumber daya energi migas di Kabupaten Kutai Kartanegara.
5. $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$, tidak terdapat pengaruh kepemimpinan, koordinasi, pengawasan dan partisipasi masyarakat terhadap efektifitas manajemen penyediaan sumber daya energi migas di Kabupaten Kutai Kartanegara
 $H_1-H_4 : \text{Salah satu atau keempat-empatnya } \beta_i \neq 0$, terdapat pengaruh kepemimpinan, koordinasi, pengawasan dan partisipasi masyarakat terhadap efektifitas manajemen penyediaan sumber daya energi migas di Kabupaten Kutai Kartanegara.

5. Hasil Penelitian

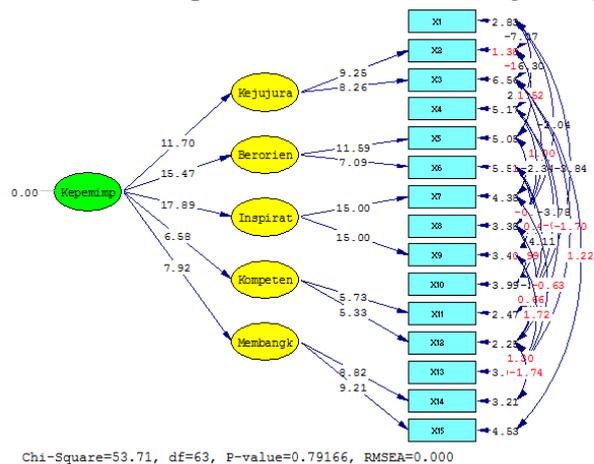
1. Confirmatory Factor Analysis (CFA) Variabel Laten Kepemimpinan

Kepemimpinan adalah fenomena sosial yang merupakan konstruk atau konsep yang bersifat laten, oleh karena itu disebut sebagai variabel laten yang tidak terukur. Agar dapat merefleksikan variabel yang tidak terukur tersebut maka harus didekati menggunakan variabel yang terukur, yang disebut variabel manifest. Seberapa baik pendekatan yang dilakukan oleh variabel manifest terhadap variabel

laten, maka harus dibentuk model pengukuran yang dapat menguji kedekatan atau kesesuaian variabel manifest terhadap variabel laten. Variabel manifest terdiri dari dimensi-dimensi, dan masing-masing dimensi diukur oleh indikatornya. Model pengukuran variabel kepemimpinan bersifat *second order* yaitu pertama variabel ini diukur oleh lima dimensi yaitu Kejujuran, Berpandangan Kedepan, Menginspirasi, Kompeten dan Membangkitkan semangat. Untuk melihat kesesuaian hubungan antara dimensi dengan indikatornya, maka muatan faktor dari masing-masing indikator untuk setiap dimensinya harus dihitung, kemudian dilakukan uji statistik berdasarkan besarnya nilai t_{hitung} .



Gambar 3. Model Pengukuran Variabel Laten Kepemimpinan



Gambar 4. T_{hitung} Variabel Laten Kepemimpinan

Pada t_{hitung} terdapat indikator yang tidak ada nilainya dan disebut *variable reference*. Dari Gambar 3, perhitungan *first order* faktor loading dan t_{hitung} variabel kepemimpinan diatas diatas dapat dilihat bahwa :

1. Indikator dengan faktor loading tertinggi dimensi kejujuran adalah X2 ($\lambda=0,89$) yaitu “Berani untuk mengatasi dirinya sendiri, berani menolak dan bertindak melawan segala kebatilan yang bertentangan dengan suara hati”. Indikator ini menunjukkan sifat terpenting dalam sifat dimensi kejujuran.
2. Indikator dengan faktor loading tertinggi dimensi berpandangan kedepan adalah X4 ($\lambda= 0,74$) yaitu “mempunyai pandangan arah yang realistis”.

3. Indikator dengan faktor *loading* tertinggi dimensi menginspirasi adalah X8 ($\lambda=0,85$) yaitu “Solusi menyelesaikan kesulitan”.
4. Indikator dengan faktor *loading* tertinggi dimensi kompeten adalah X11($\lambda=0,77$) yaitu “mampu menganalisa dan memecahkan berbagai permasalahan”.
5. Indikator dengan faktor *loading* tertinggi dimensi membangkitkan semangat adalah X13 ($\lambda=0,82$) yaitu “membangkitkan sifat optimis”.
6. Semua nilai faktor *loading* (λ) > 0,5 untuk indikator yang mengukur dimensi kepemimpinan dengan nilai $t_{hitung} > 2,0$ maka dikatakan bahwa indikator yang mengukur masing-masing dimensi tersebut memiliki tingkat validitas yang baik.

Hasil *Composite Reliability* dan *AVE* pada *1st order* dari gambar 3 dan gambar 4, dimensi kepemimpinan adalah sebagai berikut :

1. Dimensi kejujuran memiliki nilai $CR=0,84$ dan nilai $AVE=0,64$.
2. Dimensi berpandangan kedepan memiliki nilai $CR=0,76$ dan $AVE=0,51$
3. Dimensi menginspirasi memiliki nilai $CR=0,88$ dan nilai $AVE=0,70$, sehingga memiliki tingkat reliabilitas yang tertinggi, dibandingkan dimensi lainnya dalam dimensi kepemimpinan.
4. Dimensi Kompeten memiliki $CR = 0,79$ dan $AVE = 0,56$
5. Dimensi membangkitkan semangat memiliki $CR = 0,83$ dan $AVE=0,62$
6. Kelima dimensi memiliki nilai $CR > 0,70$ nilai $AVE > 0,50$ sehingga memiliki tingkat keandalan yang tinggi dan mewakili keragaman atas varians indikator yang baik.

Selanjutnya melihat tingkat validitas dari pengukuran *second order*, yaitu variabel laten yang diukur oleh kelima dimensinya dengan menggunakan Lisrel 8.7.

Hasil Faktor *loading* dan t_{hitung} (*2nd order*) indikator dimensi kepemimpinan adalah sebagai berikut :

1. Dimensi kejujuran memiliki $\lambda = 0,74$ dan $t_{hitung} = 11,74$
2. Berpandangan kedepan memiliki nilai $\lambda = 1$ dan $t_{hitung} = 15,42$
3. Memberikan inspirasi memiliki nilai $\lambda = 0,97$ dan $t_{hitung} = 17,89$
4. Kompeten memiliki nilai $\lambda = 0,97$ dan $t_{hitung} = 6,58$
5. Membangkitkan semangat memiliki nilai $\lambda = 0,72$ dan $t_{hitung} = 7,59$
6. Kelima dimensi pada *second order*, memiliki nilai $\lambda > 0,50$ dan $t_{hitung} > 2,0$. maka dikatakan valid.

Selanjutnya pada *2nd order* dari perhitungan pada gambar 3 dan gambar 4 , dihitung reliabilitas dan *AVE* dari variabel laten dengan kelima dimensi dapat digunakan formula :

$$\begin{aligned} (\sum \lambda_i)^2 &= (0,74 + 1 + 0,97 + 0,97 + 0,72)^2 = 19,360 \\ \sum \varepsilon_i &= 0,08 + 0,07 + 0,05 + 0,15 + 0,09 = 0,440 \\ \sum \lambda_i^2 &= 0,74^2 + 1^2 + 0,97^2 + 0,97^2 + 0,72^2 = 3,948 \\ CR &= \frac{(\sum \lambda_i)^2}{(\sum \lambda_i)^2 + \sum \varepsilon_i} = \frac{19,360}{19,360 + 0,440} = 0,977 \\ AVE &= \frac{\sum \lambda_i^2}{\sum \lambda_i^2 + \sum \varepsilon_i} = \frac{3,948}{3,948 + 0,440} = 0,899 \end{aligned}$$

Dari perhitungan *CR* dan *AVE* dimensi kepemimpinan pada *second order* didapatkan bahwa :

1. Dimensi kepemimpinan pada *second order* memiliki $CR = 0,977$ dan $AVE=0,899$.
2. Hasil *second order* dimensi kepemimpinan memiliki nilai $CR > 0,70$ nilai $AVE > 0,50$ sehingga memiliki tingkat keandalan yang tinggi dan mewakili keragaman atas varians indikator yang baik.

Hasil evaluasi akhir model pengukuran adalah berdasarkan *goodness of fit*.

Evaluasi terhadap kesesuaian atau kecocokan dari model pengukuran dengan menggunakan absolut *Goodness of Fit (GOF)* yang umum digunakan yaitu *Chi-Square* dan *RMSEA*.

Hasil analisa terakhir dari variabel laten kepemimpinan ini adalah sebagai berikut :

1. Nilai untuk *Goodness of fit* model kepemimpinan, $(\chi^2) = 53,71$ dan $RMSEA = 0,000$
2. Hasil $(\chi^2) \geq 0.05$ dan $RMSEA \leq 0.08$ maka model menunjukkan baik (*fit*).

Berdasarkan hasil pengujian, terlihat koefisien jalur kepemimpinan terhadap efektivitas organisasi adalah 0,76 dengan arah positif atau berpengaruh 58%. Koefisien jalur bertanda positif menunjukkan semakin efektif kepemimpinan, maka akan meningkatkan efektivitas organisasi. Selanjutnya nilai absolut t_{hitung} variabel kepemimpinan (6,07) lebih besar dari t_{kritis} (1,96) diputuskan untuk menolak H_0 sehingga H_a diterima. Jadi berdasarkan hasil pengujian dengan tingkat kepercayaan 95% disimpulkan kepemimpinan berpengaruh signifikan terhadap efektivitas organisasi, karena nilai koefisien korelasinya 0,76 termasuk dalam katagori hubungan tinggi/kuat (Hipotesis pertama diterima, karena koefisiennya signifikan).

Berdasarkan hasil pengujian koefisien jalur antara koordinasi terhadap efektivitas organisasi dalam tata kelola migas adalah 0,88 dengan arah positif atau berpengaruh 77%. Koefisien jalur yang bertanda positif menunjukkan semakin efektif koordinasi maka akan meningkatkan efektivitas organisasi dalam tata kelola migas. Selanjutnya nilai absolut t_{hitung} variabel koordinasi (5,13) lebih besar dari t_{kritis} (1,96), diputuskan untuk menolak H_0 , sehingga H_a diterima. Jadi berdasarkan hasil pengujian dengan tingkat kepercayaan 95% disimpulkan koordinasi berpengaruh signifikan terhadap efektivitas organisasi dalam tata kelola migas, karena nilai koefisien korelasinya 0,88, termasuk dalam katagori hubungan sangat tinggi/sangat kuat (Hipotesis kedua diterima, karena koefisiennya signifikan).

Koefisien jalur antara pengawasan terhadap efektivitas organisasi dalam tata kelola migas adalah 0,97 dengan arah positif atau berpengaruh 94%. Koefisien jalur yang bertanda positif menunjukkan semakin efektif pengawasan maka akan meningkatkan efektivitas organisasi dalam tata kelola migas. Selanjutnya nilai absolut t_{hitung} variabel pengawasan (3,08) lebih besar dari t_{kritis} (1,96), diputuskan untuk menolak H_0 , sehingga H_a diterima. Jadi berdasarkan hasil pengujian dengan tingkat kepercayaan 95% disimpulkan pengawasan berpengaruh signifikan terhadap efektivitas organisasi dalam tata kelola migas karena nilai koefisien korelasinya 0,97 termasuk dalam katagori hubungan sangat tinggi/sangat kuat. (Hipotesis ketiga diterima, karena koefisiennya signifikan).

Partisipasi masyarakat berpengaruh sebesar 94% terhadap efektivitas organisasi dalam tata kelola migas di kabupaten Kutai Kartanegara, ditunjukkan oleh perencanaan sebagai faktor determinan partisipasi masyarakat, pelaksanaan sebagai faktor pendukung partisipasi masyarakat dan evaluasi sebagai faktor penunjang partisipasi masyarakat

6. Kesimpulan

1. Kepemimpinan di SKK MIGAS berpengaruh signifikan dan positif terhadap efektivitas manajemen sumber daya energi migas di kabupaten Kutai Kartanegara, dengan koefisien jalur 0,76. Koefisien jalur yang bertanda positif menunjukkan bahwa semakin baik kepemimpinan maka akan meningkatkan efektivitas manajemen. Nilai absolut t_{hitung} variabel kepemimpinan (6,07) lebih besar dari t_{kritis} (1,96) sehingga kepemimpinan berpengaruh signifikan terhadap efektivitas manajemen, karena nilai koefisien korelasinya 0,76 maka termasuk dalam katagori hubungan tinggi, kuat..
2. Koordinasi di SKK MIGAS berpengaruh signifikan dan positif terhadap efektivitas manajemen sumber daya energi migas di kabupaten Kutai Kartanegara, dengan koefisien jalur 0,88 dengan arah positif. Koefisien jalur yang bertanda positif menunjukkan bahwa semakin baik koordinasi maka akan meningkatkan efektivitas manajemen. Nilai absolut t_{hitung} variabel koordinasi (5,13) lebih besar dari t_{kritis} (1,96) sehingga koordinasi berpengaruh signifikan terhadap efektivitas manajemen, karena nilai koefisien korelasinya 0,88 maka termasuk dalam katagori hubungan sangat tinggi, sangat kuat.
3. Pengawasan di SKK MIGAS berpengaruh signifikan dan positif terhadap efektivitas manajemen sumber daya energi migas di kabupaten Kutai Kartanegara, dengan koefisien jalur 0,97 dengan arah positif. Koefisien jalur yang bertanda positif menunjukkan bahwa semakin baik pengawasan maka akan meningkatkan efektivitas manajemen. Nilai absolut t_{hitung} variabel pengawasan (3,08) lebih besar dari t_{kritis} (1,96) sehingga pengawasan berpengaruh signifikan terhadap efektivitas manajemen, karena nilai koefisien korelasinya 0,97 maka termasuk dalam katagori hubungan sangat tinggi, sangat kuat.
4. Partisipasi masyarakat berpengaruh signifikan dan positif terhadap efektivitas manajemen sumber daya energi migas di kabupaten Kutai Kartanegara, dengan koefisien jalur 0,97 dengan arah positif. Koefisien jalur yang bertanda positif menunjukkan bahwa semakin baik partisipasi masyarakat maka akan meningkatkan efektivitas manajemen. Nilai absolut t_{hitung} variabel koordinasi (1,98) lebih besar dari t_{kritis} (1,96) sehingga koordinasi berpengaruh signifikan terhadap efektivitas manajemen, karena nilai koefisien korelasinya 0,97 maka termasuk dalam katagori hubungan sangat tinggi, sangat kuat.

7. Daftar Pustaka

- (1) Budi, S. (2014). Pemerintahan dan Manajemen Sekitar Publik (*Government and management about public*). Yogyakarta: Published by CAPS (*Center of Academic Publishing Service*).
- (2) Gazali, I.D.F. (2008). *Structural Equation Modelling: Teori, Konsep dan Aplikasi dengan Lisrel 9.10*, edisi 4. Semarang: Published by Universitas Diponegoro.
- (3) Isbandi, R.A. (2013). Intervensi Komunitas & Pengembangan Masyarakat: Sebagai Upaya Pemberdayaan Masyarakat [*Community Asset Based*

- Participatory Planning: from thought to Application*], Edisi Revisi 2012. Jakarta: Published by PT Raja Grafindo Persada.
- (4) Jones, G.R. (2017). *Organizational Theory, Design, and Change*. New Jersey: Pearson Prentice Hall.
 - (5) Kouzes, J.M. and dan Barry, Z.P. (2007). *Tantangan Kepemimpinan (Leadership Challenge)*, Jakarta: Penerbit Erlangga.
 - (6) Siagian, S.P. (2014). *Filsafat Administrasi (Administration Philosophy)*, cetakan ke-7. Jakarta: Penerbit PT. Bumi Aksara.
 - (7) ---, (2015). *Provinsi Kalimantan Timur Dalam Angka 2016 [Kalimantan Timur in Figures]*, No. Katalog : 1102001.64, No. Publikasi : 64.000.1503, No. ISBN: 0215-2266, Penerbit BPS Provinsi Kalimantan Timur.

PENERAPAN METODE ORDERING DUALY IN TRIANGLES (ORDIT) PADA MASALAH TINDAK KEJAHATAN DI INDONESIA

TARUGA RUNADI¹, YEKTI WIDYANINGSIH²

^{1,2}Departemen Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Indonesia, Depok, 16424, Indonesia, ¹taruga.runadi@sci.ui.ac.id,
²yekti@sci.ui.ac.id

Abstrak. Berdasarkan data yang dihimpun oleh POLRI, jumlah tindak kejahatan yang terjadi di Indonesia dari tahun ke tahun cenderung berfluktuasi dan untuk beberapa jenis tindak kriminal tertentu justru cenderung naik. Pemerintah perlu mengetahui daerah-daerah yang berpotensi tinggi terhadap tindak kejahatan untuk mendapatkan skala prioritas penanggulangan tindak kejahatan. Penelitian ini bertujuan mendapatkan peringkat untuk tiap provinsi di Indonesia berkaitan dengan jumlah terjadinya tindak kejahatan. Ada lima kelompok kasus kejahatan yang akan dianalisis yaitu kelompok I: kejahatan terhadap kesusilaan, kelompok II: kejahatan terhadap nyawa dan fisik, kelompok III: kejahatan terhadap kemerdekaan seseorang, kelompok IV: kejahatan terhadap barang, serta kelompok V: kejahatan penipuan dan korupsi. Tiap kelompok kejahatan terdiri dari beberapa jenis kejahatan. Analisis data menggunakan metode Ordering Dually in Triangles dan prosesnya dilakukan menggunakan bantuan software R. Hasil analisis data menunjukkan provinsi terburuk untuk kelompok I sampai V berturut-turut adalah Gorontalo, Jambi, Kalimantan Timur, Kalimantan Timur, Papua. Hasil penelitian diharapkan dapat membantu pemerintah dalam menentukan daerah prioritas penanganan tindak kejahatan.

Kata Kunci: Peringkat, Kejahatan, ORDIT, Indonesia

1. Pendahuluan

1.1. Latar Belakang

Menciptakan keamanan terhadap warganya adalah salah satu tugas utama suatu Negara. Warga maupun pemerintah dapat meluangkan tenaganya untuk meningkatkan kesejahteraan ketika kebutuhan akan keamanan telah terpenuhi. Di Indonesia, kewajiban pemerintah untuk menciptakan rasa aman terhadap warga Negara tertuang dalam UUD 1945 pasal 28G ayat 1 yang berbunyi “Setiap orang berhak atas perlindungan diri pribadi, keluarga, kehormatan, martabat, dan harta benda yang di bawah kekuasaannya, serta berhak atas rasa aman dan perlindungan dari ancaman ketakutan untuk berbuat atau tidak berbuat sesuatu yang merupakan hak asasi”

Terciptanya rasa aman dalam lingkungan masyarakat tidak lepas dari upaya penanganan terhadap tindak kriminal. Tindak kriminal merupakan perbuatan seseorang yang dapat diancam hukuman berdasarkan KUHP atau Undang-undang serta peraturan lainnya yang berlaku di Indonesia. Dari tahun ke tahun jumlah tindak kriminal di Indonesia cenderung berfluktuasi. Pada tahun 2011 jumlah kejadian kejahatan adalah sebanyak 347.605, turun pada tahun 2012 menjadi

341.159, dan kembali meningkat hingga tahun 2015 menjadi 352.936. Dalam rentang tahun tersebut tidak terlihat adanya kecenderungan penurunan yang konsisten pada jumlah tindak kejahatan.

Jumlah tindak kejahatan di atas hanya menggambarkan peristiwa kejahatan secara umum sedangkan tiap jenis tindak kejahatan yang terjadi memiliki latar belakang kemunculan dan cara penanggulangan yang berbeda. Dalam upaya mengurangi tindak kejahatan dengan efektif, pemerintah perlu mengetahui wilayah-wilayah terbaik dan terburuk berdasarkan kelompok jenis kejahatan. Dalam penelitian ini akan dihasilkan peringkat wilayah provinsi di Indonesia berdasarkan lima kelompok tindak kejahatan. Pengelompokan jenis tindak kejahatan dalam penelitian ini secara umum sejalan dengan pengelompokan dalam Kitab Undang-Undang Hukum Pidana (KUHP). Untuk mendapatkan peringkat tiap provinsi, metode *Ordering Dually In Triangles* (ORDIT) telah sesuai dengan tujuan penelitian.

1.2. Tujuan

Tujuan dari penelitian ini adalah menganalisis lima provinsi terbaik dan terburuk di Indonesia berdasarkan: 1) Tindak kejahatan terhadap kesusilaan, 2) Tindak kejahatan terhadap nyawa dan fisik, 3) Tindak kejahatan terhadap kemerdekaan seseorang, 4) Tindak kejahatan terhadap barang, 5) Tindak kejahatan penipuan dan korupsi.

1.3. Data

Penelitian ini akan mengolah data jumlah tindak kejahatan per provinsi selama tahun 2015 yang dihimpun dari laporan evaluasi data kriminalitas POLRI dan dipublikasikan oleh Badan Pusat Statistik (BPS). Angka jumlah tindak kejahatan tersebut diolah menjadi angka kejahatan per 100.000 penduduk yang disebut *crime rate*. Angka *crime rate* (Cr) dihitung dengan rumus berikut. Angka *crime rate* tiap tindak kejahatan akan digunakan untuk mendapatkan peringkat provinsi menggunakan metode ORDIT.

$$Cr = \frac{\text{jumlah kejahatan}}{\text{jumlah penduduk}} \times 100.000.$$

Data *crime rate* tersebut dikelompokkan menjadi lima kelompok tindak kriminal yaitu: I) kelompok kejahatan kesusilaan yang terdiri dari perkosaan dan pencabulan; II) kelompok kejahatan terhadap nyawa dan fisik yang terdiri dari pembunuhan, penganiayaan berat, penganiayaan ringan, kekerasan dalam rumah tangga; III) kelompok kejahatan terhadap kemerdekaan seseorang yang terdiri dari penculikan dan mempekerjakan anak dibawah umur; IV) kelompok kejahatan terhadap barang yang terdiri dari pencurian tanpa kekerasan, pencurian dengan kekerasan, pencurian dengan senjata api, pencurian dengan senjata tajam, pencurian dengan pemberatan, pencurian kendaraan bermotor, penadahan, pengrusakan, dan pembakaran dengan sengaja; V) kelompok kejahatan penipuan dan korupsi yang terdiri dari penipuan, penggelapan, dan korupsi.

1.4. Metode ORDIT

Ordit adalah suatu metode perbandingan antar objek dimana tiap objek memiliki beberapa (lebih dari satu) indikator yang dapat diamati. Dalam penelitian ini objek yang diamati adalah provinsi-provinsi di Indonesia sedangkan indikator yang diamati adalah nilai *crime rate* untuk tiap tindak kejahatan dalam satu kelompok.

Misalkan akan dilakukan perankingan terhadap n objek berdasarkan dua buah variabel, yaitu x dan y . Misalkan sepasang objek Θ (dengan nilai variabel x_{Θ} dan y_{Θ}) dan ϵ (dengan nilai variabel x_{ϵ} dan y_{ϵ}) akan dibandingkan dan akan dilihat apakah salah satu objek lebih unggul dari objek lainnya. Akan ada tiga kemungkinan hubungan antara sepasang objek tersebut yaitu:

- Θ memiliki hubungan *ascribe advantage* terhadap ϵ (dilambangkan $\Theta aa \epsilon$) jika $x_{\Theta} > x_{\epsilon}$ dan $y_{\Theta} > y_{\epsilon}$.
- Θ memiliki hubungan *subordinate status* terhadap ϵ (dilambangkan $\Theta ss \epsilon$) jika $x_{\Theta} < x_{\epsilon}$ dan $y_{\Theta} < y_{\epsilon}$.
- Θ memiliki hubungan *indefinite instances* terhadap ϵ (dilambangkan $\Theta ii \epsilon$) jika $x_{\Theta} > x_{\epsilon}$ dan $y_{\Theta} < y_{\epsilon}$ atau $x_{\Theta} < x_{\epsilon}$ dan $y_{\Theta} > y_{\epsilon}$.

Secara sederhana, *ascribe advantage* diartikan sebagai lebih unggul sedangkan *subordinate status* dapat diartikan kalah unggul, sehingga $\Theta aa \epsilon$ artinya Θ lebih unggul dibanding ϵ .

Setiap anggota dari n objek dapat dibandingkan terhadap seluruh objek pembanding yang merupakan anggota dari *deleted domain* $DD = n - 1$. Persentase kemunculan dari relasi-relasi di atas dihitung sebagai berikut.

$$AA = 100\% \times ff(aa)/DD$$

$$SS = 100\% \times ff(ss)/DD$$

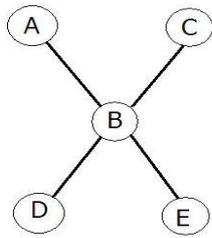
$$II = 100\% \times ff(ii)/DD$$

dimana $ff()$ adalah notasi dari *focal frequency* yang didefinisikan sebagai jumlah kemunculan jenis hubungan (aa, ss, atau ii) antara suatu objek yang diamati dengan objek lain yang merupakan anggota DD .

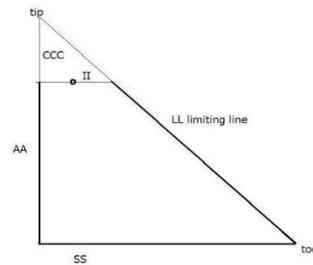
Diagram Hasse dalam Gambar 1 mengilustrasikan hubungan antar lima buah objek (A, B, C, D, dan E). A memiliki *ascribe advantage* terhadap B, D, dan E, serta bersifat *indefinite* terhadap C. Objek B memiliki *subordinate status* terhadap A dan C juga *ascribe advantage* terhadap D dan E. Deleted Domain pada Gambar 1 berjumlah 4. Dapat diketahui bahwa $AA + SS + II = 100\%$. Didefinisikan juga *complement of case condition relative* terhadap *ascribe advantage* sebagai berikut $CCC = 100 - AA$.

Pendekatan ini dapat disimbolisasikan secara skematik oleh skema subordinasi pada Gambar 2. Pada skema tersebut, titik merepresentasikan sebuah objek yang diamati dimana titik tersebut akan membagi skema menjadi dua bagian yaitu triplet trapesium (terdiri dari AA, SS, dan II) di bagian bawah serta atap segitiga (terdiri dari CCC, SS, dan II) di bagian atas. Gabungan bagian atas dan bawah membentuk sebuah segitiga siku-siku. Bagian “tip” menunjukkan situasi ketika nilai $AA=100\%$ dan “toe” ketika nilai $SS=100\%$. Garis hipotenusa menunjukkan garis pembatas untuk merencanakan posisi objek.

Ketika sebuah objek memiliki *ascribe advantage* terhadap semua objek yang lain maka objek ini memiliki AA sebesar 100%, dengan kata lain frekuensi *ascribe advantage* sama dengan jumlah anggota DD . Saat kondisi ini tercapai maka triplet trapesium menjadi sebuah segitiga. Ketika objek yang dianalisis tidak memiliki AA sebesar 100% maka akan muncul atap segitiga sebagaimana diilustrasikan pada Gambar 2.



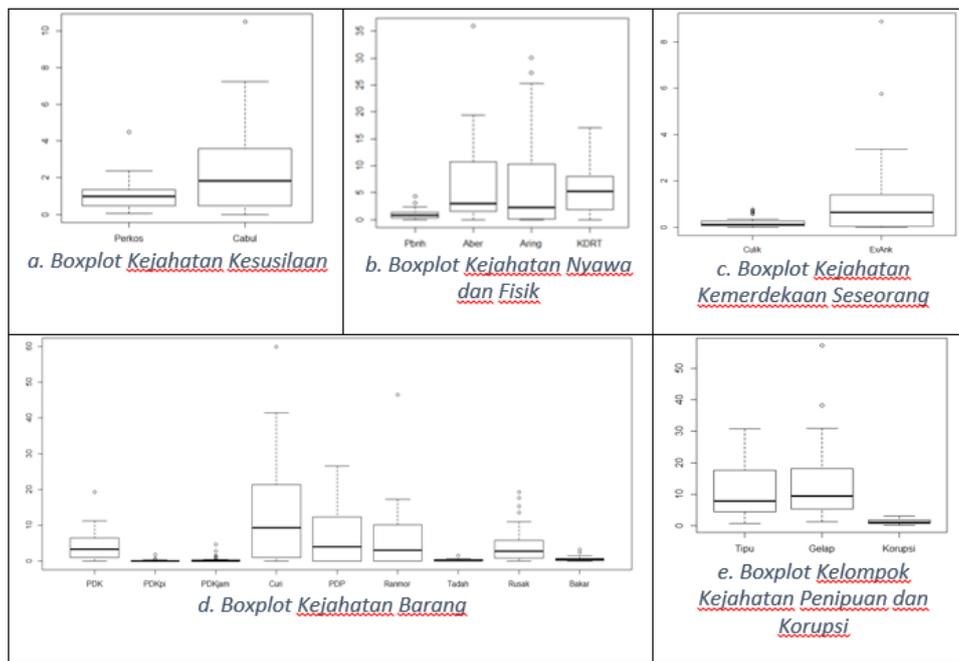
Gambar 1. Diagram Hasse



Gambar 2. Skema subordinasi

Kaki-kaki dari atap segitiga digabungkan menjadi sebuah nilai decimal **ccc.bbb**. Nilai ccc didapatkan dari pembulatan CCC ke dua digit decimal kemudian dikalikan dengan 100. Sedangkan bbb didapatkan dari membagi SS dengan CCC. Nilai dari **ccc.bbb** akan dikonversi menjadi nilai *Saliant* berdasarkan besar atau kecilnya nilai **ccc.bbb** suatu objek dibandingkan dengan objek yang lain dan digunakan untuk menentukan peringkat tiap objek yang dimanati.

2. Hasil dan Pembahasan



Gambar 3. Boxplot Tiap Data Kelompok Kejahatan

Distribusi jumlah kejahatan yang dikategorikan dalam satu kelompok memiliki perbedaan. Hal ini ditunjukkan oleh boxplot pada Gambar 3, dimana terlihat bahwa tiap tindak kejahatan dalam satu kelompok memiliki nilai maksimum, nilai minimum, nilai kuartil yang bervariasi. Oleh sebab itu, untuk mendapatkan peringkat tiap objek tidak dapat dilakukan dengan berdasarkan rata-

rata nilai indikator tiap objek. Metode ORDIT akan lebih baik dalam menentukan peringkat.

Tabel 1. Output R Kelompok I

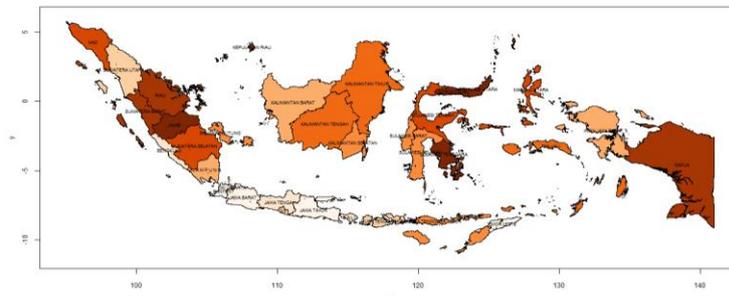
Case IDs	AA	SS	ORDIT	Salnt
1	12.9	3.23	8710.037	15.5
2	35.48	0	6452	4
3	16.13	3.23	8387.039	12.5
4	6.45	29.03	9355.31	23
5	3.23	29.03	9677.3	27
6	6.45	3.23	9355.035	18

Tabel 2. Output R kelompok II

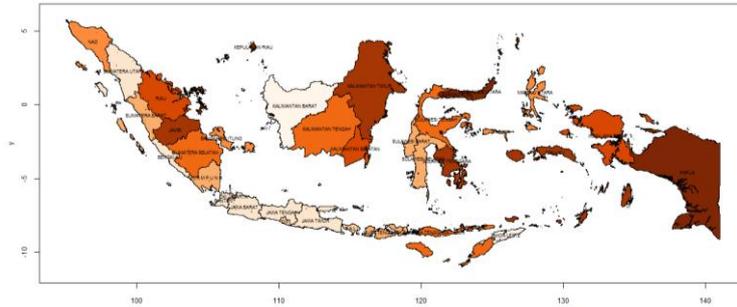
Case IDs	AA	SS	ORDIT	Salnt
1	25.81	35.48	7419.478	14
2	38.71	3.23	6129.053	10
3	22.58	22.58	7742.292	15
4	29.03	32.26	7097.455	12
5	9.68	32.26	9032.357	26
6	12.9	3.23	8710.037	22

Proses analisis data dilakukan menggunakan *software* R. Tabel 1 menunjukkan hasil output (6 objek pertama) dari *software* R untuk mendapatkan hasil ranking dengan metode ORDIT pada kelompok kejahatan terhadap nyawa dan fisik, sedangkan table 2 merupakan output untuk kelompok kejahatan terhadap barang. Case ID pada table output tersebut adalah nama atau identitas dari objek yang dilakukan perankingan dimana dalam penelitian ini CaseID merepresentasikan identitas provinsi di Indonesia. Kolom AA berisi nilai dari persentase *ascribe advantage* sedangkan kolom SS berisi persentase dari subordinate status pada tiap CaseID yang bersesuaian. Kolom ORDIT berisi hasil nilai **ccc.bbb** yang telah dijelaskan pada bagin 1.4. Kolom Salnt berisi nilai *Saliant* yang menunjukkan peringkat tiap provinsi yang ditentukan berdasarkan nilai pada kolom ORDIT. Semakin kecil nilai pada kolom Salnt berarti semakin tinggi tingkat keparahan provinsi yang bersesuaian dalam kaitannya dengan tingkat kejahatan.

Hasil perankingan disajikan secara visual menggunakan peta yang dihasilkan menggunakan *software* R. Pada peta hasil ranking akan terlihat warna pada tiap area provinsi berupa gradasi dari putih cerah hingga coklat gelap. Semakin gelap warna suatu provinsi menunjukkan nilai *Saliant* yang semakin kecil yang artinya semakin tinggi ranking tindak kejahatannya (semakin buruk).

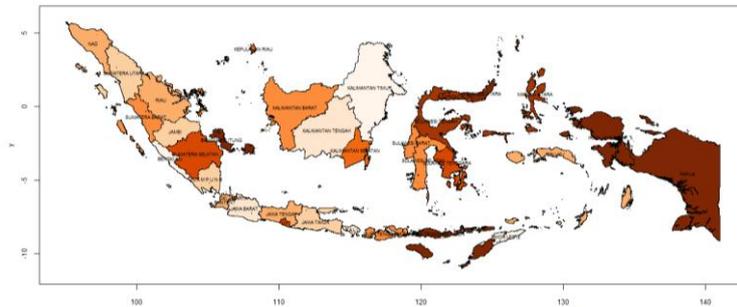


Gambar 4. Peta Peringkat I (Kejahatan Terhadap Kesusilaan)



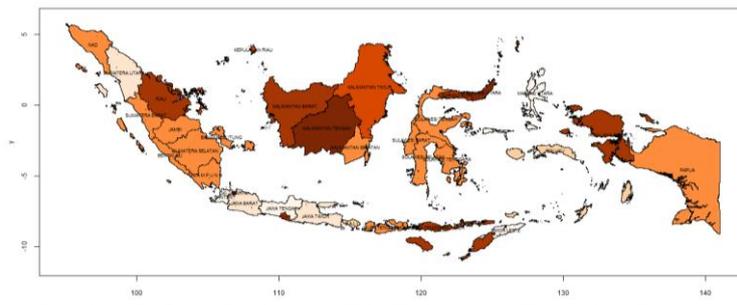
Gambar 5. Peta Peringkat II (Kejahatan Terhadap Nyawa dan Fisik)

Dilihat dari kelima gambar, yaitu Gambar 4 sampai Gambar 8, secara umum, daerah yang berwarna gelap tersebar dari wilayah Indonesia barat sampai Indonesia timur kecuali pada kelompok kejahatan terhadap kemerdekaan seseorang (Gambar 6) dimana terlihat bahwa bagian gelap lebih cenderung terjadi di wilayah Indonesia timur.



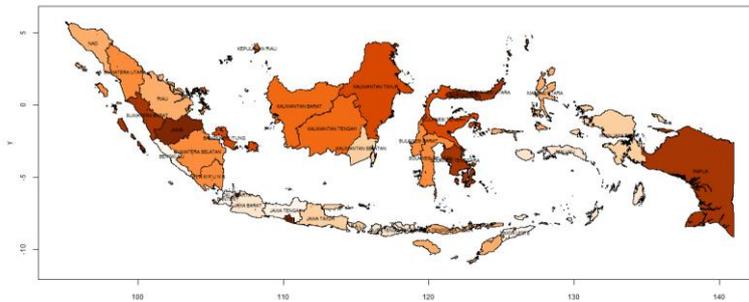
Gambar 6. Peta Peringkat III (Kejahatan Kemerdekaan Seseorang)

Di antara lima pulau terbesar di Indonesia (Sumatra, Jawa, Kalimantan, Sulawesi, Papua), pulau Jawa terlihat lebih terang dibanding lainnya yang mengindikasikan secara umum provinsi-provinsi di pulau Jawa tingkat kejahatan tergolong rendah. Namun terlihat provinsi DKI Jakarta berwarna jauh lebih gelap dibanding provinsi lainnya di pulau Jawa, yang artinya tingkat kejadian kejahatan, untuk lima kelompok tindak kejahatan yang dianalisis, masih tergolong tinggi. Provinsi Daerah Istimewa Yogyakarta juga terlihat berwarna gelap untuk kelompok kejahatan terhadap kemerdekaan seseorang, terhadap barang, serta penipuan dan korupsi.



Gambar 7. Peta Peringkat IV (Kejahatan Terhadap Barang)

Di pulau Sumatra, area yang berwarna gelap sering terlihat di bagian tengah, tepatnya di wilayah provinsi Jambi, Riau, dan Sumatra Barat untuk kelompok kejahatan I, III, IV dan V. Sedangkan dilihat dari Gambar 5 untuk kelompok kejahatan II terlihat merata. Di pulau Kalimantan sebaran tindak kejahatan kelompok I, III, IV dan V terlihat lebih merata walaupun warna yang ditunjukkan gelap. Untuk kejahatan kelompok II, kejadian kejahatan cenderung terpusat di wilayah Kalimantan Timur.



Gambar 8. Peta Peringkat V (Kejahatan Penipuan dan Korupsi)

Di pulau Sulawesi, provinsi Gorontalo dan Sulawesi Utara selalu berwarna gelap untuk kelima kelompok kejahatan. Provinsi Papua selalu berwarna gelap kecuali untuk kelompok kejahatan terhadap barang sedangkan provinsi Papua Barat tergolong rawan terhadap kejahatan terhadap barang dan kejahatan terhadap kemerdekaan seseorang.

3. Kesimpulan

Lima provinsi paling buruk berdasarkan nilai *Saliant* (kolom salnt pada Tabel 1) paling kecil, untuk kelompok kejahatan terhadap kesusilaan adalah Gorontalo, Sulawesi Utara, Kepulauan Riau, Sulawesi Tenggara, dan Jambi. Sedangkan untuk 5 provinsi yang paling baik (ranking rendah) adalah Jawa Barat, Banten, Jawa timur, Bengkulu dan Jawa Tengah. Untuk empat kelompok kejahatan yang lain, provinsi-provinsi terbaik dan terburuk dapat dilihat pada Table 3 sampai Tabel 6.

Tabel 3. Provinsi terbaik dan terburuk kelompok II

Kelompok Kejahatan Terhadap Nyawa dan Fisik		
No	Baik	Buruk
1	Kal. Barat	Jambi
2	Sumatera	Kal. Timur
3	Jawa Barat	Kepulauan Riau
4	Jawa Tengah	Gorontalo
5	Jawa Timur	Sulawesi Utara

Tabel 4. Provinsi terbaik dan terburuk kelompok III

Kelompok Kejahatan Kemerdekaan Terhadap Seseorang		
No	Baik	Buruk
1	Banten	Kal. Timur
2	Jawa Tengah	Gorontalo
3	Jawa Barat	NTT
4	Jawa Timur	Papua
5	Kal. Barat	Sulawesi Utara

Tabel 5. Provinsi terbaik dan terburuk kelompok IV

Kelompok Kejahatan Terhadap Barang		
No	Baik	Buruk
1	Banten	Kal. Timur
2	Jawa Tengah	Gorontalo
3	Jawa Barat	NTT
4	Jawa Timur	Papua
5	Kal. Barat	Sulawesi Utara

Tabel 6. Provinsi terbaik dan terburuk kelompok V

Kelompok Kejahatan Penipuan dan Korupsi		
No	Baik	Buruk
1	Jawa Tengah	Papua
2	Banten	DI Yogyakarta
3	Bengkulu	Jakarta
4	Maluku	Jambi
5	Jawa Barat	Gorontalo

Referensi

- [1] Badan Pusat Statistik, 2015, *Penduduk Indonesia: Hasil Survei Penduduk Antar Sensus 2015*, Jakarta: Badan Pusat Statistik.
- [2] Badan Pusat Statistik, 2016, *Statistik Kriminal 2016*, Jakarta: Badan Pusat Statistik.
- [3] Bruggemann, R. dan Patil, G., 2011, *Ranking and Prioritization for Multi-indicator Systems*, Springer.
- [4] Myers, W. dan Patil, G., 2010, Preliminary Prioritization Based on Partial Order Theory and Software for Compositional Complexes in Landscape Ecology, with Applications to Restoration, Remediation, and Enhancement, *Environ and Ecol Stat*, 4 (2010), 411–436.
- [5] Myers, W. dan Patil, G., 2012, *Multivariate Methods of Representing Relations in R for Prioritization Purposes*, Springer.

APLIKASI ANALISIS KONJOIN DESAIN *FULLPROFILE* RESPON *RATING* PADA PREFERENSI MAHASISWA TERHADAP KUALITAS DOSEN SEKOLAH TINGGI ILMU STATISTIK

FITRI CATUR LESTARI

Sekolah Tinggi Ilmu Statistik, fitricaturlestari@stis.ac.id

Abstrak. Preferensi mahasiswa terhadap kualitas dosen dari berbagai aspek dianggap penting untuk dikaji secara khusus. Dosen dengan segala karakteristik dan kualitas yang dimilikinya sangat berperan dalam menentukan kualitas suatu perguruan tinggi. Pengukuran preferensi dapat dilakukan dengan analisis konjoin yang merupakan analisis yang khusus mengkaji tentang preferensi. Rancangan dalam analisis konjoin yang digunakan dalam penelitian ini adalah rancangan kombinasi lengkap dengan jenis data respon *rating* (nilai). Adapun prosedur analisisnya adalah regresi linear berganda dengan variabel *dummy*. Metode penarikan sampel yang digunakan adalah metode *purposive sampling* dengan pemilihan sampel didasarkan pemerataan pada segmen yang diduga terdapat perbedaan preferensi mahasiswa terhadap kualitas dosen yaitu IPK, tingkat/kelas, asal SMA, jenis kelamin, dan kemampuan ekonomi (pendapatan) orang tuanya. Berdasarkan Nilai Relatif Penting (NRP), preferensi mahasiswa terhadap kualitas dosen yaitu penguasaan materi kuliah, pembawaan diri, penyampaian materi, kemampuan memotivasi mahasiswa, dan metode pembelajaran. Preferensi mahasiswa terhadap kualitas dosen berbeda-beda menurut segmen tingkat/kelas, asal SMA, jenis kelamin, dan kemampuan ekonomi (pendapatan) orang tuanya. Sedangkan preferensi terhadap kualitas dosen tidak berbeda menurut segmen IPK.

Kata kunci: analisis konjoin, preferensi, kombinasi lengkap, *rating*

1. Pendahuluan

Preferensi mahasiswa terhadap kualitas dosen dari berbagai aspek dianggap penting untuk dikaji secara khusus. Dosen dengan segala karakteristik dan kualitas yang dimilikinya sangat berperan dalam menentukan kualitas suatu perguruan tinggi. Pentingnya peran dosen dinyatakan oleh Sudiana [1] bahwa dosen sebagai salah satu komponen perguruan tinggi berperan sangat besar dalam mewujudkan kualitas perguruan tinggi. Faktor kualitas dosen juga terbukti berpengaruh signifikan terhadap motivasi belajar mahasiswa dengan tingkat signifikansi 0,01 dan memiliki korelasi yang paling kuat di antara faktor yang lain dalam Pujadi [2]. Bidang kegiatan dosen melaksanakan pendidikan dan pengajaran merupakan bidang yang lebih utama di antara 2 (dua) bidang lainnya yaitu penelitian, dan

pengabdian kepada masyarakat dalam DPPM [3]. Kualitas dosen pada bidang pendidikan dan pengajaran merupakan faktor yang menentukan kualitas perguruan tinggi.

Mahasiswa sebagai konsumen atau pihak yang berkepentingan (*stakeholders*) terhadap dosen pada proses belajar mengajar, layak untuk memberikan evaluasi terhadap kualitas dosen. Suatu penilaian berkaitan dengan kepuasan dan preferensi. Suatu penilaian akan tinggi ketika kepuasan tinggi dan kepuasan tinggi sangat dipengaruhi oleh preferensi. Chaplin [4] mendefinisikan preferensi sebagai suatu sikap yang lebih menyukai sesuatu benda daripada benda lainnya. Dalam penelitian ini, preferensi diartikan sebagai suatu sikap yang lebih menyukai karakteristik dosen tertentu dibandingkan dengan karakteristik dosen yang lainnya. Selanjutnya akan dikaji preferensi tersebut berdasarkan segmen karakteristik mahasiswa: Indeks Prestasi Kumulatif (IPK), tingkat/kelas, asal SMA, jenis kelamin, dan kemampuan ekonomi (pendapatan) orang tua.

Hasil penelitian ini dapat dimanfaatkan sebagai bahan acuan bagi dosen dalam kegiatannya melaksanakan pendidikan dan pengajaran. Selain itu, mahasiswa juga dapat menggunakan hasil penelitian ini untuk mengoptimalkan proses pembelajarannya di perguruan tinggi. Pihak perguruan tinggi juga dapat menyusun suatu kegiatan evaluasi kualitas dosen berdasarkan preferensi mahasiswa yang dipaparkan dalam penelitian ini.

2. Metode Penelitian

Langkah-langkah dalam melakukan analisis konjoin adalah sebagai berikut:

2.1 Perumusan masalah

Langkah awal dalam melakukan analisis konjoin yaitu perumusan masalah dalam Aaker et. al. [5]. Perumusan masalah dimulai dari mendefinisikan produk sebagai kumpulan dari atribut-atribut dimana setiap atribut terdiri atas beberapa taraf/level. Informasi mengenai atribut yang mewakili preferensi konsumen bisa diperoleh melalui diskusi dengan pakar, eksplorasi data sekunder, atau melakukan tes awal dalam Rosada [6].

Kualitas dosen diwujudkan dalam suatu karakteristik atau atribut yang terdiri atas taraf-taraf. Atribut dan taraf dosen dalam studi ini ditetapkan dengan melakukan tes awal atau penelitian pendahuluan. Berikut ini tabel hasil penelitian pendahuluan dalam penentuan atribut dan taraf:

Tabel 1 Atribut yang Mempengaruhi Preferensi Mahasiswa terhadap Kualitas Dosen

No.	Atribut	Persentase
1	Penyampaian materi	24.00%
2	Metode mengajar kreatif dan inovatif	20.00%
3	Penguasaan materi	19.00%
4	Pembawaan diri	11.50%
5	Suka memotivasi	9.00%
6	Penilaian obyektif	5.50%
7	Sistematika	3.00%
8	Pendidikan	2.00%
9	Wibawa	2.00%
10	Pengelolaan kelas	2.00%
11	Penampilan	1.50%
12	Penggunaan bahan ajar	1.00%
13	Umur	0.00%

Berdasarkan tabel 1, terdapat 5 (lima) atribut yang persentasenya relatif tinggi yaitu berkisar 9,00% sampai 24,00%. Adapun kelima atribut ini masing-masing terdiri atas 2 (dua) taraf dengan perincian : penyampaian materi kuliah (komunikasi dua dan satu arah), metode pembelajaran (kreatif dan tidak kreatif), penguasaan materi kuliah (menguasai dan tidak menguasai), pembawaan diri (humoris dan serius), dan kemampuan memotivasi mahasiswa (memotivasi dan tidak memotivasi).

2.2 Perancangan Stimuli

Rancangan kombinasi yang akan digunakan dalam penelitian ini disusun berdasarkan kombinasi lengkap (*full profile*) atau evaluasi banyak faktor dalam Sawtooth [7]. Karena terdapat 5 (lima) atribut maka terdapat 2^5 atau 32 kartu. Pada tahap ini juga disusun pertanyaan yang berkaitan dengan peubah demografi dalam bentuk kuesioner untuk mengetahui karakteristik mahasiswa. Hal ini memungkinkan adanya temuan tentang segmen mahasiswa berdasarkan preferensinya terhadap kualitas dosen.

2.3 Pengumpulan Data

Data pada penelitian ini merupakan data primer yang dikumpulkan melalui survei berbasis wawancara dengan menggunakan kuesioner dan kartu-kartu stimuli yang berisi tentang karakteristik atau profil yang menggambarkan kualitas dosen (dalam hal ini pengajar mata kuliah eksak). Kartu-kartu stimuli tersebut dievaluasi oleh mahasiswa dalam bentuk *rating* (nilai). Penelitian dilakukan di STIS pada Rabu-Kamis, 20-21 Juli 2011 dengan melibatkan 125 mahasiswa STIS tahun akademik 2010/2011 yang dipilih berdasarkan metode *purposive sampling*. Metode *purposive sampling* adalah salah satu metode penarikan contoh tak berpeluang dengan mengumpulkan informasi dari sumber yang tepat di antaranya anggota masyarakat yang dipandang dapat memberikan informasi yang dibutuhkan atau hanya mereka yang dirasa dapat memberikan informasi yang kita butuhkan dalam Wibisono [8]. Pemilihan sampel juga didasarkan pemerataan pada segmen yang diduga terdapat perbedaan preferensi mahasiswa terhadap kualitas dosen yaitu IPK, tingkat/kelas, asal SMA, jenis kelamin, dan kemampuan ekonomi (pendapatan) orang tuanya.

2.4 Prosedur Analisis

Prosedur analisis yang digunakan dalam penelitian ini adalah metode regresi dengan variabel *dummy*. Adapun secara umum model dasar analisis konjoin dalam Kuhfeld [9] dengan prosedur analisis tersebut adalah:

$$Y_{ij} = \beta_0 + \sum_{i=1}^m \sum_{j=1}^k \beta_{ij} X_{ij} + \varepsilon_{ij}$$

keterangan :

Y_{ij} = Peringkat seluruh responden

β_0 = Intersep

k = Banyak taraf dari atribut ke- i

m = Jumlah atribut

X_{ij} = Peubah boneka atau *dummy variable* dari atribut ke- i taraf ke- j

β_{ij} = *Part worth* atau nilai kegunaan atribut ke- i taraf ke- j

ε_{ij} = Galat

Dengan model regresi tersebut, maka dapat ditentukan nilai kegunaan dari taraf-taraf tiap atribut (NKT) untuk menentukan nilai pentingnya suatu taraf relatif terhadap taraf yang lain pada suatu atribut. Setelah menentukan NKT, maka Nilai Relatif Penting (NRP) dapat dihitung dengan formula sebagai berikut:

$$NRP_i = \frac{UT_i - UR_i}{\sum_{i=1}^k (UT_i - UR_i)}$$

Keterangan:

NRP_i = NRP atribut ke-i

UT_i = NKT tertinggi atribut ke-i

UR_i = NKT terendah atribut ke-i

k = Jumlah atribut

Data preparation dan analisis tersebut dilakukan dengan menggunakan *software* pengolah data: SPSS 13.0, Microsoft Excel 2003, dan SAS versi 9.1.

3. Hasil – Hasil Utama

3.1 Preferensi Mahasiswa terhadap Kualitas Dosen

Berdasarkan hasil pengolahan data, atribut yang paling mempengaruhi preferensi mahasiswa terhadap kualitas dosen adalah penguasaan materi kuliah dengan Nilai Relatif Penting (NRP) sebesar 32.79%. Atribut berikutnya adalah pembawaan diri, penyampaian materi, kemampuan memotivasi mahasiswa, dan metode pembelajaran dengan NRP berturut-turut 17.82%, 17.20%, 16.75% dan 15.44%. Dosen yang paling disukai oleh mahasiswa adalah dosen yang menguasai materi kuliah karena Nilai Kegunaan Taraf (NKT) paling tinggi di antara taraf lainnya yaitu 2.15. Dosen berikutnya yang disukai secara berturut-turut adalah dosen yang humoris (NKT=1.17), menyampaikan materi dengan komunikasi dua arah (NKT=1.13), mampu memberi motivasi kepada mahasiswanya (NKT=1.10), dan metode pembelajarannya kreatif (NKT=1.01). Berikut ini tabel yang berkaitan dengan NKT dan NRP yang dihasilkan:

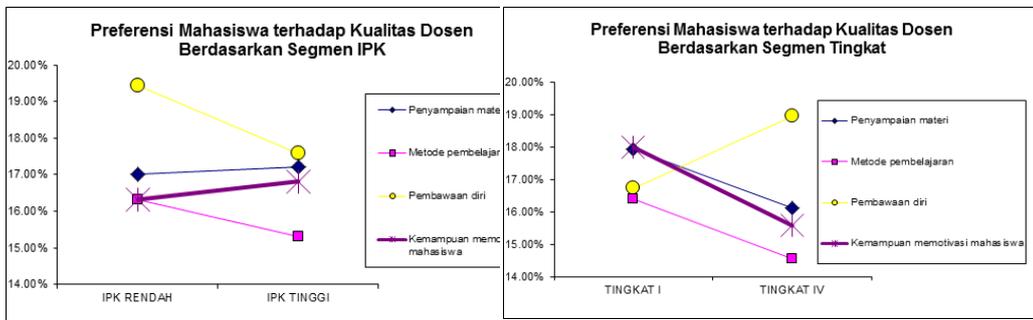
Tabel 2 Hasil Prosedur Analisis

No.	Atribut	Taraf	Nilai Kegunaan Taraf	Jarak	Nilai Relatif Penting	Urutan
1.	Penyampaian materi	a. Komunikasi dua arah	1.13	2.26	17.20%	3
		b. Komunikasi satu arah	-1.13			
2.	Metode pembelajaran	a. Kreatif	1.01	2.03	15.44%	5
		b. Monoton	-1.01			
3.	Penguasaan materi kuliah	a. Menguasai materi	2.15	4.30	32.79%	1
		b. Tidak menguasai materi	-2.15			
4.	Pembawaan diri	a. Humoris	1.17	2.34	17.82%	2
		b. Serius	-1.17			
5.	Kemampuan memotivasi mahasiswa	a. Memotivasi mahasiswa	1.10	2.20	16.75%	4
		b. Tidak memotivasi mahasiswa	-1.10			
Total				13.12	100%	

Keterangan: Urutan terkecil=paling disukai

3.2 Segmentasi Preferensi Mahasiswa terhadap Kualitas Dosen

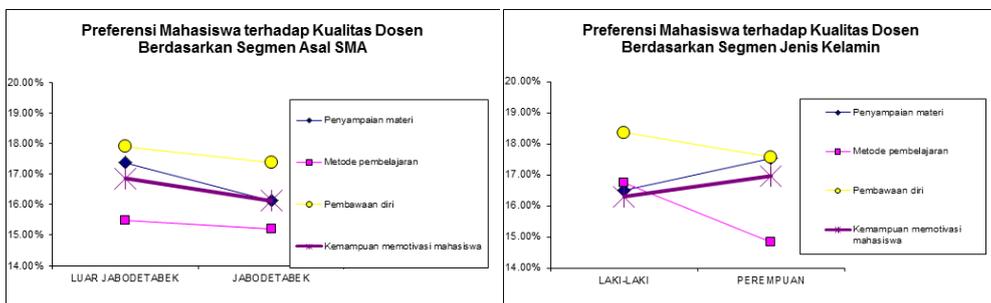
Segmentasi preferensi mahasiswa terhadap kualitas dosen ditinjau menurut karakteristik IPK, tingkat/kelas, asal SMA, jenis kelamin, dan kemampuan ekonomi (pendapatan) orang tuanya. Berdasarkan tabel pada lampiran, dapat disimpulkan pada setiap kelompok mahasiswa, Nilai Relatif Penting (NRP) yang tertinggi selalu terjadi pada atribut penguasaan materi kuliah. Sehingga pada pembahasan selanjutnya, segmentasi preferensi mahasiswa yang diamati adalah berdasarkan 4 (empat) atribut lain selain atribut penguasaan materi kuliah.



Gambar 1 Preferensi mahasiswa berdasarkan IPK

Gambar 2 Preferensi mahasiswa berdasarkan tingkat

Gambar 1 menunjukkan tidak adanya perbedaan preferensi mahasiswa terhadap kualitas dosen pada kelompok mahasiswa dengan IPK rendah maupun IPK tinggi. Urutan atribut yang dianggap penting adalah sama di kedua kelompok mahasiswa tersebut (semakin tinggi NRP menunjukkan semakin penting atribut). Sedangkan gambar 2 menunjukkan adanya kecenderungan preferensi mahasiswa tingkat I terhadap kualitas dosen berbeda dengan preferensi mahasiswa tingkat IV. Mahasiswa tingkat IV cenderung lebih menyukai dosen dengan pembawaan diri humoris sedangkan mahasiswa tingkat I lebih menyukai dosen yang mampu memotivasi mahasiswanya.

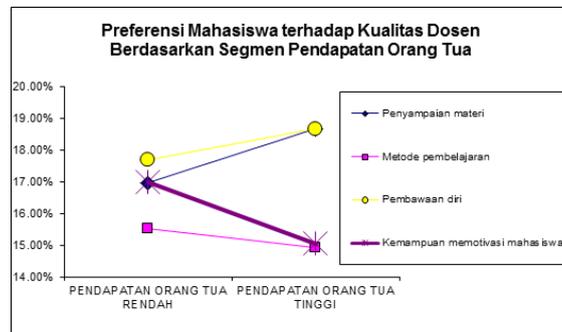


Gambar 3 Preferensi mahasiswa berdasarkan asal SMA

Gambar 4 Preferensi mahasiswa berdasarkan jenis kelamin

Pada gambar 3 dapat dideskripsikan bahwa mahasiswa yang berasal dari Jabodetabek maupun luar Jabodetabek menyukai dosen yang menyampaikan materi (komunikasi) secara 2 arah. Namun, dapat diketahui pula pada gambar tersebut ada indikasi bahwa mahasiswa yang berasal dari Jabodetabek juga menyukai dosen yang mempunyai kemampuan memotivasi mahasiswa. Hal ini dapat dilihat pada lampiran yang menunjukkan NRP atribut penyampaian materi dan NRP atribut kemampuan memotivasi mahasiswanya adalah sama pada kelompok mahasiswa yang berasal dari Jabodetabek. Salah satu hal menarik pada

gambar 4 adalah adanya perbedaan preferensi yang menyolok pada kelompok mahasiswa laki-laki (mahasiswa) dan mahasiswa perempuan (mahasiswi). Mahasiswi cenderung lebih menyukai dosen yang memiliki kemampuan memotivasi mahasiswanya daripada dosen yang metode pembelajarannya kreatif. Sedangkan mahasiswa sebaliknya, cenderung menyukai dosen yang metode pembelajarannya kreatif daripada dosen yang memiliki kemampuan memotivasi mahasiswa.



Gambar 5 Preferensi mahasiswa berdasarkan pendapatan orang tua

Perbandingan preferensi berdasarkan perbedaan pendapatan orang tua dapat diamati pada gambar 5 dan lampiran. Mahasiswa dengan pendapatan orang tua yang rendah, relatif sama dalam menyukai dosen yang komunikatif dan dosen yang mampu memotivasi mahasiswa. Sedangkan mahasiswa yang orang tuanya berpendapatan tinggi, relatif berbeda dalam menyukai dosen yang komunikatif dan dosen yang mampu memotivasi mahasiswa. Mahasiswa pada kelompok ini lebih menyukai dosen yang komunikatif daripada dosen yang mampu memotivasi mahasiswanya.

4. Kesimpulan

Urutan preferensi mahasiswa terhadap kualitas dosen mencakup penguasaan materi kuliah, pembawaan diri, penyampaian materi, kemampuan memotivasi mahasiswa, dan metode pembelajaran. Tidak ada perbedaan preferensi mahasiswa dengan IPK tinggi maupun IPK rendah terhadap kualitas dosen. Mahasiswa tingkat IV cenderung lebih menyukai dosen dengan pembawaan diri humoris sedangkan mahasiswa tingkat I lebih menyukai dosen yang mampu memotivasi mahasiswanya. Mahasiswa yang berasal dari Jabodetabek maupun luar Jabodetabek menyukai dosen yang menyampaikan materi (komunikasi) secara dua arah. Namun, ada indikasi bahwa mahasiswa yang berasal dari Jabodetabek juga menyukai dosen yang mempunyai kemampuan memotivasi mahasiswa dengan NRP yang relatif sama dengan NRP penyampaian materi. Mahasiswi cenderung lebih menyukai dosen yang memiliki kemampuan memotivasi mahasiswanya daripada dosen yang metode pembelajarannya kreatif. Sedangkan mahasiswa cenderung menyukai dosen yang metode pembelajarannya kreatif daripada dosen yang memiliki kemampuan memotivasi mahasiswa. Mahasiswa dengan pendapatan orang tua yang rendah, relatif sama (NRP relatif sama) dalam menyukai dosen yang komunikatif dan dosen yang mampu memotivasi mahasiswa. Mahasiswa dengan pendapatan orang tua yang tinggi lebih menyukai dosen yang komunikatif daripada dosen yang mampu memotivasi mahasiswanya.

Referensi

- [1] Sudiana IN. 2003. *Peranan Profesionalisme Tenaga Pengajar (Dosen) terhadap Proses Pembelajaran di Perguruan Tinggi*. Jurnal Pendidikan dan Pengajaran IKIP Negeri Singaraja, No. 4 TH. XXXVI.
- [2] Pujadi A. 2007. *Faktor-Faktor yang Mempengaruhi Motivasi Belajar Mahasiswa: Studi Kasus pada Fakultas Ekonomi Universitas Bunda Mulia*. *Business & Management Journal* Bunda Mulia, Vol: 3, No. 2.
- [3] [DPPM] Direktorat Penelitian dan Pengabdian pada Masyarakat. 2002. *Kebijakan dan Program Penelitian dan Pengabdian kepada Masyarakat*. Jakarta: Departemen Pendidikan Nasional
- [4] Chaplin JP. 2002. *Kamus Lengkap Psikologi*. Edisi kelima. Terjemahan: dr. Kartini Kartono. Jakarta: PT. Raja Grafindo Persada.
- [5] Aaker DA, Day GS. 1980. *Marketing Research*. New York : John Willey&Son, Inc.
- [6] Rosada R. 2002. *Perbandingan Metode Pairwise Comparison dan Full profile dalam Pengumpulan Data untuk Analisis Konjoin*. Skripsi. Jurusan Statistika IPB. Bogor.
- [7] <http://www.sawtooth.com>, *Conjoint Analysis*, 2001.
- [8] Wibisono D. 2003. *Riset Bisnis*. Jakarta : PT. Gramedia Pustaka Utama.
- [9] Kuhfeld WF. 2000. “*Conjoint Analysis Examples*”, *SAS Institut, Inc.* <http://www.sawtooth software.com>. [28 Januari 2004]

PENDEKATAN DISTRIBUSI GUMBEL DALAM BAGAN KENDALI RANGE – EXPONENTIALLY WEIGHTED MOVING AVERAGE (R-EWMA) PADA PROSES PRODUKSI COKLAT DI INDONESIA PADA TAHUN 2009-2013

ERNA TRI HERDIANI¹, SITI AISYAH ALDILLAH², SRI ASTUTI
THAMRIN³

1 Universitas Hasanuddin, herdiani.erna@gmail.com
2 Universitas Hasanuddin, sitiaisyahaldillah1@gmail.com
3 Universitas Hasanuddin, tutithamrin@yahoo.com

Abstrak. Bagan Kendali *Range-Exponentialy Weighted Moving Average (R-EWMA)* merupakan tipe bagan kendali yang digunakan untuk memonitor perubahan rentang pada suatu proses serta untuk mendeteksi penurunan dan peningkatan variabilitas. Salah satu aplikasi bagan kendali R-EWMA adalah penentuan bagan kendali pada proses produksi coklat di Indonesia, ternyata pendekatan distribusi yang cocok untuk kasus ini adalah distribusi Gumbel. Hasilnya menunjukkan semakin kecil nilai dari bobot dari R-EWMA maka jarak dari kedua batas tersebut semakin kecil.

Kata kunci: Range, Eksponensial Weighting Moving Average, Bagan Kendali, variabilitas

1. Pendahuluan

Bagan kendali EWMA yang diperkenalkan oleh Robert [5] adalah salah satu bagan kendali yang dapat mengontrol produksi saat adanya pergeseran-pergeseran nilai rata-rata yang kecil dalam produksi. Penggunaan bagan kendali EWMA untuk memonitor proses statistik lain telah dilakukan oleh beberapa ahli, diantaranya adalah Montgomery [4] yang membahas mengenai bagan kendali EWMA untuk memonitor pergeseran kecil pada *mean*. Crowder dan Hamilton [3] juga membahas mengenai bagan kendali EWMA untuk standar deviasi. Castagliola [2] menggunakan bagan kendali S^2 -EWMA untuk variansi dengan menggunakan transformasi logaritma. Selanjutnya dikembangkan untuk memonitor *range*. Sehingga, Castagliola [2] mengembangkan bagan kendali S^2 -EWMA menjadi bagan kendali R-EWMA adapun distribusi pendekatan yang digunakan dengan menggunakan distribusi normal.

Bagan Kendali R-EWMA adalah tipe bagan kendali yang digunakan untuk memonitor perubahan rentang pada suatu proses serta untuk mendeteksi penurunan dan peningkatan variabilitas. Adapun distribusi pendekatan yang akan dilakukan

dalam paper ini selain normal adalah distribusi gumbel. Oleh karena itu, pada penelitian ini akan dikaji mengenai bagan kendali R-EWMA untuk memonitoring rentang proses dengan menggunakan pendekatan distribusi gumbel dan normal. Dan hasilnya akan diaplikasikan pada data proses produksi coklat di Indonesia.

2.2. Hasil dan Diskusi

Misalkan $X_{k,1}, \dots, X_{k,n}$ adalah sampel dari n normal independen $N(\mu_0, \sigma_0)$ variabel acak, dimana μ_0 adalah rata-rata proses, σ_0 adalah standar deviasi proses, dan k adalah subgrup. Jika diasumsikan bahwa σ_0 diketahui, apabila misalnya parameter ini diestimasikan berada dalam kontrol populasi, yaitu

$$\hat{\sigma}_0 = S/K_S(n) \text{ atau } \hat{\sigma}_0 = R/K_R(n)$$

dimana S dan R adalah standar deviasi dan rentang sampel yang berada dibawah kontrol populasi dan $K_S(n)$ serta $K_R(n)$ adalah konstanta yang membuat $\hat{\sigma}_0$ memiliki nilai. Masalah desain statistik pada tipe bagan kendali EWMA dengan estimasi parameter telah diselesaikan, Castagliola[2]. Dalam memantau proses variansi sampel. Crowder dan Hamilton [3], mengikuti rekomendasi dari Box [1] menyarankan aplikasi dari EWMA klasik pada logaritma dari sampel variansi berturut-turut $T_k = \ln S_k^2$ yaitu,

$$Z_k = (1 - \lambda)Z_{k-1} + \lambda T_k \quad (1)$$

dengan λ adalah konstanta $0 < \lambda \leq 1$. Tujuan utama dari pendekatan ini adalah bahwa $T_k = \ln S_k^2$, memiliki distribusi log-Gamma, cenderung lebih berdistribusi normal daripada sampel variansi S_k^2 . Castagliola [2] mengembangkan pendekatan ini dengan menggunakan transformasi tiga parameter untuk sampel standar deviasi

$$T_k = a + b \ln(S_k + c) \quad (2)$$

dengan $c > 0$ kemudian gunakan statistik EWMA pada persamaan (1). Dapat dilihat bahwa transformasi ini termasuk dalam transformasi original yang diusulkan oleh Johnson. Harapan utama dari pendekatan ini adalah jika a,b,c dipilih langsung, maka transformasi ini dapat mengakibatkan perkiraan normalitas untuk T_k lebih baik daripada pendekatan Crowder dan Hamilton [3], Untuk membuat T_k berdistribusi normal maka ini berhubungan dengan membuat distribusi dari T_k simetris. Salah satu keuntungan lain dari pendekatan ini adalah adanya perubahan (penurunan) pada ARL yang berada diluar kendali dibandingkan dengan bagan kendali standar deviasi klasik. Gagasan utama yang dikembangkan pada penelitian ini adalah untuk menerapkan penggunaan dari transformasi tiga parameter untuk range R_k dari sampel standar deviasi yaitu,

$$T_k = a + b \ln(R_k + c) \quad (3)$$

Dengan R_k adalah range dari subgrup k yang dinyatakan dengan :

$$R_k = \max_{j=1, \dots, n} X_{k,j} - \min_{j=1, \dots, n} X_{k,j} \quad (4)$$

karena $E(R_k) = K_R(n)\sigma_0$, dimana $K_R(n)$ adalah

$$K_R(n) = 2 \int_0^{+\infty} 1 - \{\Phi(x)\}^n - \{1 - \Phi(x)\}^n dx$$

Maka nilai awal Z_0 menjadi $Z_0 = a + b \ln\{K_R(n)\sigma_0 + c\}$. Jika nilai rata-rata $E(T_k)$ dan standar deviasi $\sigma(T_k)$ dari T_k sesuai dengan parameter a,b dan c diketahui maka batas kontrol EWMA untuk perubahan range akan menjadi :

$$LCL = E(T_k) - K \left(\frac{\lambda}{2 - \lambda} \right)^{1/2} \sigma(T_k)$$

$$UCL = E(T_k) + K \left(\frac{\lambda}{2 - \lambda} \right)^{1/2} \sigma(T_k)$$

dengan K adalah bilangan konstan positif. Pada bagian berikutnya akan ditunjukkan bahwa apapun nilai dari n dan σ_0 , nilai $E(T_k)$ akan tetap mendekati 0 maka dapat diasumsikan = 0. Sehingga, dapat menggunakan batas kontrol dari bagan kendali R -EWMA.

$$LCL = -K \left(\frac{\lambda}{2 - \lambda} \right)^{1/2} \sigma(T_k)$$

$$UCL = +K \left(\frac{\lambda}{2 - \lambda} \right)^{1/2} \sigma(T_k)$$

dengan,

K, λ = Parameter penghalus dan faktor skala dari bagan kendali EWMA

σ = Nominal standar deviasi dari proses

T_k = "Transformasi" range dalam kontrol dari subgrup K

Misalkan x_1, x_2, \dots, x_n adalah sampel acak dari peubah acak yang berdistribusi normal dengan mean μ_0 dan variansi 1. Apabila sampel diurutkan maka akan menjadi $y_1 < y_2 < \dots < y_n$. Akan dicari distribusi dari jangkauan R di mana :

$$R = R_{max} - R_{min}$$

$$R = Y_n - Y_1$$

berdasarkan statistik terurut, yaitu :

$$g_{1,n}(y_1, y_n) = \begin{cases} n(n-1) (\Phi(y_n) - \Phi(y_1))^{n-2} \phi(y_1) \phi(y_n) \\ 0 \end{cases}$$

fkp dari sampel jangkauan $Z_1 = Y_n - Y_1$. $F(x) = x$ dimana $0 < x < 1$

$$g_{1,n}(y_1, y_n) = \begin{cases} n(n-1) (\Phi(y_n) - \Phi(y_1))^{n-2} \phi(y_1) \phi(y_n) \\ 0 \end{cases}$$

Sebagai tambahan dari $Z_1 = Y_n - Y_1$, misalkan $Z_2 = Y_1$. Fungsi $z_1 = y_n - y_1$, $z_2 = y_1$, memiliki invers $y_n = z_1 + z_2$, $y_1 = z_2$, sehingga transformasi *Jacobian* :

$$J = \begin{vmatrix} \frac{\partial y_1}{\partial z_1} & \frac{\partial y_1}{\partial z_2} \\ \frac{\partial y_n}{\partial z_1} & \frac{\partial y_n}{\partial z_2} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 1 & 1 \end{vmatrix} = |1|$$

Dan fkp dari Z_1 dan Z_2 ialah

$$h(z_1, z_2) = \begin{cases} |1| n(n-1) (\Phi(z_1 + z_2) - \Phi(z_2))^{n-2} \phi(z_2) \phi(z_1 + z_2) \\ 0 \end{cases}$$

berdasarkan fkp dari jangkauan $Z_1 = Y_n - Y_1$ dari sampel acak sebanyak n :

$$h(z_1) = \begin{cases} \int_{-\infty}^{\infty} n(n-1) (\Phi(z_1 + z_2) - \Phi(z_2))^{n-2} \phi(z_2) \phi(z_1 + z_2) \\ 0 \end{cases}$$

$$= n(n-1) \int_{-\infty}^{\infty} (\Phi(z_1 + z_2) - \Phi(z_2))^{n-2} \phi(z_2) \phi(z_1 + z_2)$$

dimana $h(z_1)$ merupakan fungsi kepadatan dari $Z_1 = Y_n - Y_1$. Untuk kasus ini

$R = Z_1$, $Y_n = \max(x_n)$ dan $Y_i = \min(x_i)$, oleh karena itu, fungsi kepadatan kumulatif dari jangkauan R adalah :

$$R = F_R \left(n(n-1) \int_{-\infty}^{\infty} (\Phi(z_1 + z_2) - \Phi(z_2))^{n-2} \phi(z_2) \phi(z_1 + z_2) \right),$$

Jadi fungsi kepadatan peluang dari jangkauan R adalah :

$$R = n(n-1) \int_{-\infty}^{\infty} (\Phi(z_1 + z_2) - \Phi(z_2))^{n-2} \phi(z_2) \phi(z_1 + z_2)$$

R juga memiliki nilai pemusatan, nilai pemusatan dari R -EWMA merupakan nilai mean. Dengan menggunakan *Simpson Rule* pada persamaan akan dicari nilai mean dari R , dengan fkp dari R ialah

$$f_R = \int_{-\infty}^{\infty} n(n-1) (\Phi(z_1 + z_2) - \Phi(z_2))^{n-2} \phi(z_2) \phi(z_1 + z_2) dZ_2$$

Dengan menggunakan persamaan (2.) akan didapatkan

$$f(0) = [\Phi(z_1) - \Phi(0)]^{n-2} \phi(0) \phi(z_1)$$

$$f(h) = [\Phi(z_1 + h) - \Phi(h)]^{n-2} \phi(h) \phi(z_1 + h)$$

$$f(2h) = [\Phi(z_1 + 2h) - \Phi(2h)]^{n-2} \phi(2h) \phi(z_1 + 2h)$$

Substitusi nilai dari f_0 , f_h dan f_{2h} pada persamaan (2.), sehingga akan menjadi

$$\int_0^{2h} f(x) dx = 2h \left(\frac{1}{6} f(0) + \frac{2}{3} f(h) + \frac{1}{6} f(2h) \right)$$

Jika diketahui $h = \frac{b-a}{n}$ memiliki batas $[-1,1]$, maka $h = \frac{2}{n}$. Dengan substitusi nilai h ke persamaan (4.1) diperoleh

$$\begin{aligned} \int_0^{\frac{4}{n}} f(x) dx &= \frac{4}{n} \left(\frac{1}{6} [[\Phi(z_1) - \Phi(0)]^{n-2} \phi(0) \phi(z_1)] \right. \\ &\quad + \frac{2}{3} \left[\left[\Phi\left(z_1 + \frac{2}{n}\right) - \Phi\left(\frac{2}{n}\right) \right]^{n-2} \phi\left(\frac{2}{n}\right) \phi\left(z_1 + \frac{2}{n}\right) \right] \\ &\quad + \frac{1}{6} \left[\left[\Phi\left(z_1 + \frac{4}{n}\right) - \Phi\left(\frac{4}{n}\right) \right]^{n-2} \phi\left(\frac{4}{n}\right) \phi\left(z_1 + \frac{4}{n}\right) \right] \left. \right) \end{aligned} \quad (4.2)$$

fkp dari Z_1 diasumsikan sama dengan fkp dari R , sehingga

$$\begin{aligned} f_{Z_1} &= n(n-1) \frac{4}{n} \left(\frac{1}{6} [[\Phi(z_1) - \Phi(0)]^{n-2} \phi(0) \phi(z_1)] \right. \\ &\quad + \frac{2}{3} \left[\left[\Phi\left(z_1 + \frac{2}{n}\right) - \Phi\left(\frac{2}{n}\right) \right]^{n-2} \phi\left(\frac{2}{n}\right) \phi\left(z_1 + \frac{2}{n}\right) \right] \\ &\quad + \frac{1}{6} \left[\left[\Phi\left(z_1 + \frac{4}{n}\right) - \Phi\left(\frac{4}{n}\right) \right]^{n-2} \phi\left(\frac{4}{n}\right) \phi\left(z_1 + \frac{4}{n}\right) \right] \left. \right) \end{aligned}$$

Untuk mencari nilai mean dari R dapat dilakukan dengan cara sebagai berikut :

$$E(Z_1) = \int_{-\infty}^{\infty} z_1 f_{Z_1} dz_1$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} z_1 \cdot n(n-1) \frac{4}{n} \left(\frac{1}{6} [[\Phi(z_1) - \Phi(0)]^{n-2} \phi(0)\phi(z_1)] \right. \\
&\quad + \frac{2}{3} \left[\left[\Phi\left(z_1 + \frac{2}{n}\right) - \Phi\left(\frac{2}{n}\right) \right]^{n-2} \phi\left(\frac{2}{n}\right) \phi\left(z_1 + \frac{2}{n}\right) \right] \\
&\quad \left. + \frac{1}{6} \left[\left[\Phi\left(z_1 + \frac{4}{n}\right) - \Phi\left(\frac{4}{n}\right) \right]^{n-2} \phi\left(\frac{4}{n}\right) \phi\left(z_1 + \frac{4}{n}\right) \right] \right) dZ_1
\end{aligned}$$

Untuk nilai $n = 3$, maka :

$$\begin{aligned}
E(Z_1) &= 6 \int_{-\infty}^{\infty} z_1 \cdot \frac{4}{3} \left(\frac{1}{6} [[\Phi(z_1) - \Phi(0)]^{n-2} \phi(0)\phi(z_1)] \right. \\
&\quad + \frac{2}{3} \left[\left[\Phi\left(z_1 + \frac{2}{3}\right) - \Phi\left(\frac{2}{3}\right) \right]^1 \phi\left(\frac{2}{3}\right) \phi\left(z_1 + \frac{2}{3}\right) \right] \\
&\quad \left. + \frac{1}{6} \left[\left[\Phi\left(z_1 + \frac{4}{3}\right) - \Phi\left(\frac{4}{3}\right) \right]^1 \phi\left(\frac{4}{3}\right) \phi\left(z_1 + \frac{4}{3}\right) \right] \right) dZ_1 \\
&= 0,4349
\end{aligned}$$

sehingga didapatkan nilai dari $E(Z_1)$ untuk $n = 3$ adalah 0,4349, Sedangkan variansinya adalah sebagai berikut :

$$\text{var}(Z_1) = E(Z_1^2) - [E(Z_1)]^2$$

$$\begin{aligned}
E(Z_1^2) &= \int_{-\infty}^{\infty} z_1^2 f_{Z_1} dZ_1 \\
&= \int_{-\infty}^{\infty} z_1^2 \cdot n(n-1) \frac{4}{n} \left(\frac{1}{6} [[\Phi(z_1) - \Phi(0)]^{n-2} \phi(0)\phi(z_1)] \right. \\
&\quad + \frac{2}{3} \left[\left[\Phi\left(z_1 + \frac{2}{n}\right) - \Phi\left(\frac{2}{n}\right) \right]^{n-2} \phi\left(\frac{2}{n}\right) \phi\left(z_1 + \frac{2}{n}\right) \right] \\
&\quad \left. + \frac{1}{6} \left[\left[\Phi\left(z_1 + \frac{4}{n}\right) - \Phi\left(\frac{4}{n}\right) \right]^{n-2} \phi\left(\frac{4}{n}\right) \phi\left(z_1 + \frac{4}{n}\right) \right] \right) dZ_1
\end{aligned}$$

Untuk nilai $n = 3$,

$$\begin{aligned}
E(Z_1^2) &= 6 \int_{-\infty}^{\infty} z_1^2 \cdot \frac{4}{3} \left(\frac{1}{6} [[\Phi(z_1) - \Phi(0)]^{n-2} \phi(0)\phi(z_1)] \right. \\
&\quad + \frac{2}{3} \left[\left[\Phi\left(z_1 + \frac{2}{3}\right) - \Phi\left(\frac{2}{3}\right) \right]^{n-2} \phi\left(\frac{2}{3}\right) \phi\left(z_1 + \frac{2}{3}\right) \right] \\
&\quad \left. + \frac{1}{6} \left[\left[\Phi\left(z_1 + \frac{4}{3}\right) - \Phi\left(\frac{4}{3}\right) \right]^{n-2} \phi\left(\frac{4}{3}\right) \phi\left(z_1 + \frac{4}{3}\right) \right] \right) dZ_1 \\
&= 0,4349
\end{aligned}$$

Jika $[E(Z_1)]^2 = \left(\int_{-\infty}^{\infty} z_1 f_{Z_1} dZ_1 \right)^2$, maka untuk nilai $n = 3$, $[E(Z_1)]^2 = (0,4349)^2 = 0,1891$. Sehingga didapatkan nilai variansi untuk $n = 3$, $\text{var}(Z_1) = E(Z_1^2) - [E(Z_1)]^2 = 0,2458$. Dengan cara yang sama maka akan diperoleh nilai R dengan n yang berbeda seperti pada Tabel 1.

Tabel 1. Rata-rata dan variansi dari R

n	E \textcircled{R}	Var \textcircled{R}
3	0,4349	0,2458
4	0,4149	0,2428
5	0,3938	0,2387
6	0,3776	0,2350
7	0,3657	0,2320
8	0,3571	0,2296
9	0,3507	0,2277
10	0,3458	0,2262
11	0,3421	0,2251
12	0,3392	0,2242
13	0,3369	0,2234
14	0,3350	0,2228
15	0,3335	0,2223

Selanjutnya, akan ditentukan nilai dari mean dan variansi dari T_k , terlebih dahulu akan dicari nilai dari $A(n), B(n), C(n)$ dan w dengan menggunakan nilai mean dan variansi dari R serta menggunakan nilai *skewness*, sebagai berikut :

$$w = \left[\sqrt{\gamma_3 \left(\frac{R}{2}\right)^2 + 1 + \gamma_3 \left(\frac{R}{2}\right)} \right]^{\frac{1}{3}} - \left[\sqrt{\gamma_3 \left(\frac{R}{2}\right)^2 + 1 - \gamma_3 \left(\frac{R}{2}\right)} \right]^{\frac{1}{3}}$$

selanjutnya,

$$B(n) = \frac{1}{\sqrt{\ln(w^2 + 1)}}$$

$$A(n) = \frac{B(n)}{2} \ln \left(\frac{w^2(w^2 + 1)}{V(R)} \right)$$

$$C(n) = \frac{\sqrt{V(R)}}{w} - E(R)$$

Untuk $n = 3$,

$$w = \left[\sqrt{\gamma_3(R)/2)^2 + 1 + \gamma_3(R)/2} \right]^{\frac{1}{3}} - \left[\sqrt{\gamma_3(R)/2)^2 + 1 - \gamma_3(R)/2} \right]^{\frac{1}{3}}$$

$$= 1,1117 - 0,8995 = 0,2122$$

$$B(n) = \frac{1}{\sqrt{\ln(w^2 + 1)}} = \frac{1}{\sqrt{\ln((0,2122)^2 + 1)}} = \frac{1}{0,2099} = 4,7653$$

$$A(n) = \frac{B(n)}{2} \ln \left(\frac{w^2(w^2+1)}{V(R)} \right) = 2,3826 \times \ln(0,0596) = -6,7187$$

$$C(n) = \frac{\sqrt{V(R)}}{w} - E(R) = 4,1868 - 1,6926 = 2,4942$$

Selanjutnya untuk $n=4, 5, \dots, 15$ akan ditampilkan dalam Tabel 2.

Tabel 2. Nilai parameter dari R – EWMA

N	w	A(n)	B(n)	C(n)
3	0,2122	-67,187	47,653	24,942
4	0,1726	-94,189	58,359	30,382
5	0,1540	-111,944	65,338	32,867
6	0,1440	-123,046	69,800	33,547
7	0,1383	-129,754	72,641	33,193
8	0,1349	-133,630	74,438	32,279
9	0,1329	-135,706	75,565	31,077
10	0,1318	-136,525	76,217	29,713
11	0,1311	-136,596	76,576	28,305
12	0,1309	-136,154	76,728	26,891
13	0,1309	-135,354	76,728	25,500
14	0,1310	-134,405	76,652	24,170
15	0,1313	-133,301	76,500	22,888

Bagan Kendali R –EWMA dinyatakan dengan :

$$LCL = E(T_k) - K \left(\frac{\lambda}{2 - \lambda} \right)^{1/2} \sigma(T_k)$$

$$UCL = E(T_k) + K \left(\frac{\lambda}{2 - \lambda} \right)^{1/2} \sigma(T_k)$$

Oleh karena itu, terlebih dahulu akan dicari nilai dari T_k , $E(T_k)$ dan $Var(T_k)$. Adapun data yang akan digunakan adalah data cokelat bulanan di Indonesia pada tahun 2009-2013, lihat Tabel 3.

Tabel 3. Data Cokelat Bulanan di Indonesia pada Tahun 2009 - 2013

Bulan	2013	2012	2011	2010	2009
Januari	3,4	3,3	4,94	4,60	4,50
Februari	3,3	3,3	4,60	4,20	3,70
Maret	3,3	3,2	5,33	4,80	4,30
April	5,3	5,2	3,47	6,10	6,20
Mei	4,5	5,8	3,89	6,80	6,90
Juni	3,8	4,9	6,43	5,80	5,70
Juli	5,6	5,5	6,68	5,00	5,50
Agustus	7,2	7,1	6,60	3,70	4,10
September	4,9	4,8	6,52	5,70	6,80
Oktober	4,7	4,1	6,44	6,40	7,20
November	4,5	3,6	6,36	6,00	6,30
Desember	4,0	2,6	6,28	6,20	6,40

Berdasarkan Data cokelat bulanan pada tabel 4.5 akan dihitung nilai dari T_k , $E(T_k)$ dan $Var(T_k)$ sebagai berikut :

Untuk nilai $k = 1$,

$$R_k = X_{k,j\max} - X_{k,j\min}$$

$$R_1 = X_{1,3} - X_{1,4\min} = 4,94 - 3,30 = 1,64$$

Kemudian untuk nilai $k = 2$,

$$R_2 = X_{2,3} - X_{2,4\min} = 4,60 - 3,3 = 1,30$$

Setelah didapatkan nilai dari R_k untuk nilai $k = 1$ dan $k = 2$ selanjutnya akan dihitung nilai dari T_k untuk $k = 1$ ialah

$$T_1 = a + b \ln(R_1 + c) = -11,194 + 6,5338 \ln(1,64 + 3,2867) = 0,78$$

Untuk nilai $k = 2$,

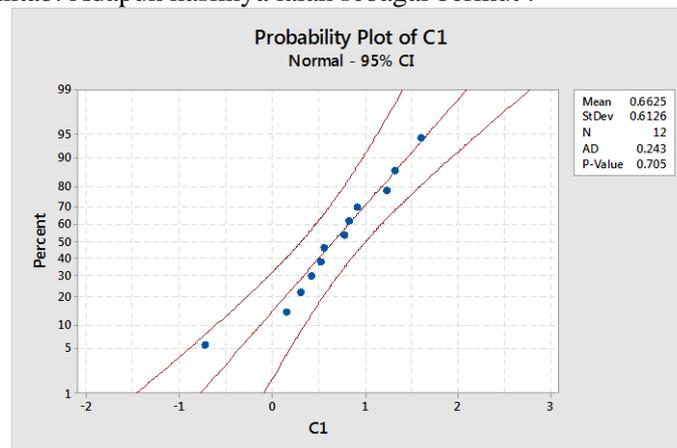
$$T_2 = a + b \ln(R_2 + c) = -11,194 + 6,5338 \ln(1,30 + 3,2867) = 1,24$$

Dan dengan cara yang sama akan di peroleh :

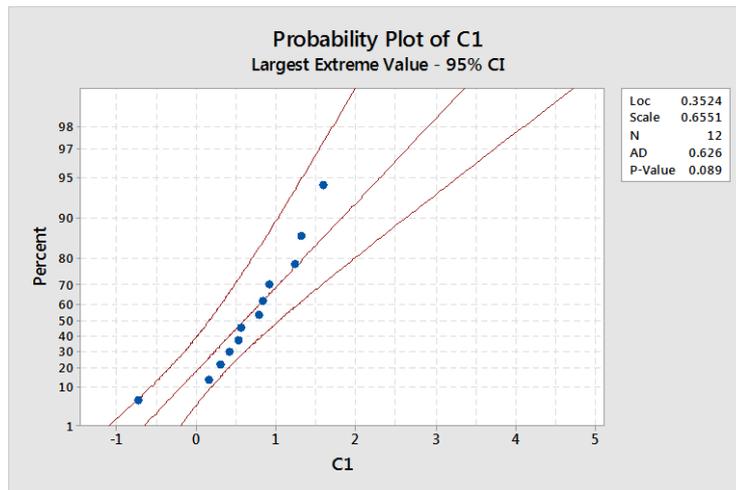
Tabel 4. Nilai R_k dan T_k untuk Cokelat Bulanan, Indonesia

k	Rk	Tk
1	1,64	0,78
2	1,30	1,24
3	2,13	0,16
4	2,73	0,53
5	3,01	0,83
6	2,63	0,42
7	1,68	-0,72
8	3,50	1,32
9	2,00	0,31
10	3,10	0,92
11	2,76	0,56
12	3,80	1,60

Selanjutnya akan diselidiki distribusi T_k melalui *probability plot* dari software minitab. Adapun hasilnya ialah sebagai berikut :



Gambar 1.a. Distribusi Normal



Gambar 1.b. Distribusi Largest Extreme Value (Gumbel)

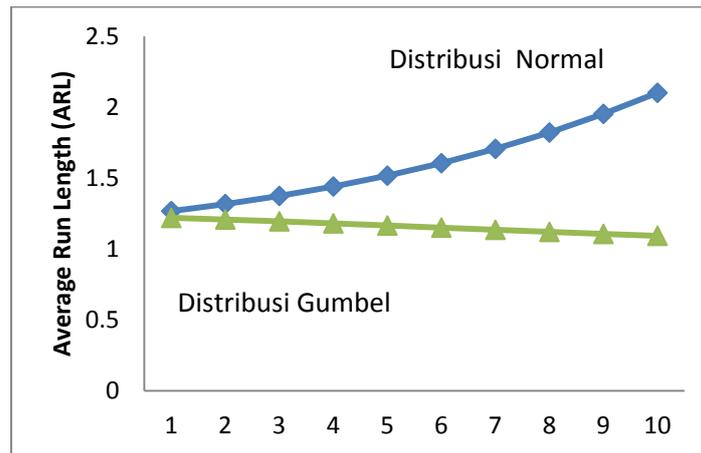
Setelah distribusi T_k diselidiki dengan menggunakan *probability* plot pada software minitab didapatkan hasil seperti pada gambar 1.a. dan 1.b. Pada gambar 1.a. dan 1.c. nilai *p-value* yang didapatkan lebih besar dari 0,05 , sehingga *probability* plot T_k yang dapat disarankan untuk digunakan adalah distribusi normal, logistik dan Gumbel. Apabila yang diambil adalah nilai *p* value yang terbesar maka sebaiknya digunakan adalah distribusi normal.

Selanjutnya akan dihitung nilai mean dan variansi dari T_k dengan menggunakan ketiga distribusi tersebut dan akan dicari batas control dari R EWMA dari data produksi coklat di Indonsia, dengan mengambil nilai $\lambda = 0,2$.

Tabel 5. Nilai Bagan Kendali R EWMA untuk Distribusi T_k Normal, Logistik dan Gumbel

No	$E(T_k)$	$var(T_k)$	LCL	UCL	Keterangan
1	0.663	0.376	0.050	1,276	Distribusi Normal
2	0.577	1.282	-0.555	1.709	Distribusi Gumbel

Berdasarkan Gambar 2, terdapat titik pengamatan yang berada diluar batas kendali yakni titik 7, 8 dan 12 berdasarkan distribusi normal sedangkan untuk distribusi gumbel hanya mendeteksi satu titik yaitu titik ke-7. Sedangkan melalui average run length (ARL) dapat dilihat performance dari kedua bagan kendali tersebut, lihat Gambar 2.



Gambar 2. Average Run Length (ARL) dari Distribusi Normal dan Largest Extreme Value (Gumbel) dengan $n = 10$

Berdasarkan Gambar 2, dapat disimpulkan bahwa untuk pada $n = 10$ dengan pergeseran 0,1; 0,2 ; ...; 1 menunjukkan distribusi gumbel memiliki nilai yang lebih kecil jika dibandingkan dengan menggunakan distribusi normal.

3. Kesimpulan

Pada bagan kendali R – EWMA untuk data produksi coklat di Indonesia tahun 2009-2013 dapat menggunakan dua asumsi distribusi, dalam menentukan batas atas dan bawah yaitu salah satu dari berdistribusi normal atau gumbel, tetapi berdasarkan nilai ARL yang diperoleh, distribusi gumbel lebih tepat dipilih sebagai asumsi distribusi, karena memiliki nilai yang lebih kecil dibandingkan dengan asumsi distribusi normal. Sehingga bagan kendali yang bagus digunakan adalah berdistribusi gumbel dengan batas sebagai berikut : $UCL = 1,709$ dan $LCL = -0,555$.

Pernyataan terima kasih. Pernyataan terima kasih disampaikan kepada Universitas Hasanuddin yang telah memberikan kesempatan untuk meneliti dan mempublikasikan hasilnya.

Referensi

- [1] Box, G. E., W. G. Hunter dan J.S. Hunter, 1978, *Statistics for Experimenters*. John Wiley & Sons, New York
- [2] Castagliola, P., 2005, A new S2-EWMA Control Chart for Monitoring the Process Variance, *Quality and Reliability Engineering International*, 21:781-794
- [3] Crowder, S. V. dan M. D. Hamilton, 2005, An EWMA for monitoring a process sample median. *Int. J. Reliability Engineering International*.
- [4] Montgomery, D. C., 2009, *Introduction to Statistical Quality Control Sixth Edition*. New York: John Wiley & Sons, Inc.
- [5] Roberts, S.W., 1959, Control Chart Test Based on Geometric Moving Averages. *Technometrics*, Vol 1 no 3, 239-250.

PENGELOMPOKAN KABUPATEN/KOTA DI JAWA BARAT MENGGUNAKAN *HYBRID CLUSTERING* BERDASARKAN INDIKATOR PEMBANGUNAN MANUSIA BERBASIS *GENDER*

MARIA ULPAH¹, ZULHANIF², TITI PURWANDARI²

1 Program Studi Statistika FMIPA Universitas Padjadjaran,
mariaulpah003@gmail.com

2 Program Studi Statistika FMIPA Universitas Padjadjaran, dzulhanif@yahoo.com

2 Program Studi Statistika FMIPA Universitas Padjadjaran,
titipurwandari@yahoo.com

Abstrak. Ketidakadilan *gender* terjadi akibat adanya sistem sosial salah satu jenis kelamin (laki-laki maupun perempuan) mengalami diskriminasi yang mempengaruhi kualitas pembangunan manusia. Analisis *cluster* merupakan teknik pengelompokan yang sudah dikenal, dimana dalam teknik ini bertujuan untuk mengelompokan data ke dalam *cluster* sehingga setiap *cluster* berisi data yang semirip mungkin serta mengetahui karakteristik dari setiap *cluster* yang terbentuk. Pada penelitian ini, dengan memanfaatkan data dari Badan Pusat Statistik guna meningkatkan kualitas pembangunan manusia berbasis *gender*, maka dilakukan pengelompokan kabupaten/kota di Provinsi Jawa Barat berdasarkan indikator pembangunan manusia berbasis *gender* menggunakan *hybrid clustering* yaitu gabungan *k-means* dengan metode *ward*. Metode *hybrid clustering* diusulkan untuk mengatasi kelemahan *k-means* dalam menentukan pusat awal *cluster* (*centroid*) dan menentukan jumlah *cluster* optimum. Penelitian ini menghasilkan 2 *cluster* optimum untuk data laki-laki dimana untuk *cluster* 1 beranggotakan 19 kabupaten/kota, dan *cluster* 2 beranggotakan 8 kabupaten/kota dengan nilai *silhouette coefficient* bernilai 0,34 untuk *cluster* 1 dan 0,21 untuk *cluster* 2. Data perempuan menghasilkan 2 *cluster* optimum dimana untuk *cluster* 1 beranggotakan 18 kabupaten/kota, dan *cluster* 2 beranggotakan 9 kabupaten/kota dengan nilai *silhouette coefficient* bernilai 0,36 untuk *cluster* 1 dan 0,27 untuk *cluster* 2. Hasil pengujian analisis profil untuk data laki-laki dan perempuan didapatkan bahwa profil populasi *cluster* pertama dan kedua tidak sejajar (tidak paralel).

Kata kunci : *Hybrid Clustering*, *K-means*, Metode *Ward*, *Silhouette Coefficient*, Pembangunan Manusia berbasis *Gender*.

1. Pendahuluan

1.1 Latar Belakang

Pembangunan manusia berbasis *gender* merupakan salah satu indikator yang menjadi perhatian dunia [3]. Ketidakadilan *gender* terjadi akibat adanya sistem sosial salah satu jenis kelamin (laki-laki maupun perempuan) mengalami diskriminasi. Dalam mencapai sasaran pembangunan manusia berbasis *gender* [3], terdapat dua indikator penting untuk mengevaluasi pembangunan berbasis *gender*, yaitu indeks pembangunan *gender* (IPG) dan indeks pemberdayaan *gender* (IDG) yang memiliki indikatornya masing-masing.

Secara umum, pembangunan manusia Jawa Barat terus mengalami kemajuan selama periode 2010 hingga 2015 rata-rata tumbuh sebesar 0,99 persen per tahun. Meskipun IPM Jawa Barat menunjukkan kemajuan yang besar, status pembangunan manusia Jawa Barat masih sedang [3]. Oleh karena itu peneliti ingin mengetahui variabel dalam indikator pembangunan manusia berbasis *gender* mana yang menyebabkan pembangunan manusia di Jawa Barat yang masih perlu diperhatikan agar pembangunan manusia di kabupaten/kota di Provinsi Jawa Barat memiliki IPM yang tinggi dan meneliti apakah persebarannya merata. Salah satu sarana untuk mendapatkan suatu pola persebaran kondisi sosial ekonomi di Kabupaten/Kota di Provinsi Jawa Barat adalah dengan melakukan pengelompokan. Pada penelitian ini, peneliti akan menggunakan teknik *clustering* yang bertujuan untuk menemukan pengelompokan dan mengidentifikasi pendistribusian dan pola yang menarik dari data yang digunakan.

Hingga saat ini, para ilmuwan masih terus melakukan berbagai usaha untuk melakukan perbaikan model *cluster*, dan menghitung jumlah *cluster* yang optimal sehingga dapat menghasilkan *cluster* yang paling baik. Dalam membentuk segmentasi *clustering*, terdapat beberapa metode yang umumnya dapat digunakan diantaranya metode *hierarchical agglomerative clustering* yang merupakan salah satu bagian dari metode hirarki, sedangkan metode *partitioning* yang sering digunakan yaitu *k-means*.

Menurut Santosa [7], *k-means* merupakan metode *clustering* yang paling sederhana dan umum digunakan. *K-means* memiliki ketelitian yang cukup tinggi terhadap ukuran objek, sehingga algoritma ini lebih terukur dan efisien untuk pengolahan objek dalam jumlah besar [9], dengan waktu yang relatif cepat dan efisien [2]. Namun *k-means* mempunyai kelemahan yang diakibatkan oleh penentuan pusat awal *cluster*. Hasil *cluster* yang terbentuk dari metode *k-means* sangat bergantung pada inisiasi nilai pusat awal *cluster* yang diberikan. Hal ini menyebabkan hasil *cluster*-nya berupa solusi yang sifatnya *local optimal*.

Dalam upaya menangani kelemahan *k-means*, ada beberapa cara memberi nilai awal (*centroid*) diantaranya, dengan mengambil nilai awal menggunakan hasil *cluster* hirarki dengan jumlah *cluster* yang sesuai [7]. Kelemahan *k-means* metode ini dikombinasikan dengan metode hirarki untuk menentukan jumlah *cluster* dan penentuan pusat awal *cluster* yang akan di sertakan pada metode *k-means* [1] dan diharapkan penggunaan metode ini dapat memberikan hasil pengelompokan yang lebih baik.

Berdasarkan permasalahan diatas, maka pendekatan *hybrid clustering* diusulkan untuk menutupi kekurangan-kekurangan dengan mencoba melakukan pengelompokan data berdasarkan kombinasi metode *k-means* dengan metode

hierarchical agglomerative yaitu metode *ward*.

1.2 Tujuan Penelitian

Tujuan dari penelitian ini adalah memperoleh mengelompokkan kabupaten/kota di Provinsi Jawa Barat dan mengetahui karakteristik dari setiap kelompok yang terbentuk.

1.3 Sistematika Penulisan

1. Pendahuluan

- 1.1 Latar Belakang
- 1.2 Tujuan Penelitian
- 1.3 Sistematika Penulisan

2. Hasil-hasil Utama

- 2.1 Pendahuluan
- 2.2 Standarisasi Variabel
- 2.3 Hasil *Dendogram* menggunakan Metode *Ward*
- 2.4 Penentuan Banyak *Cluster* Optimum
- 2.5 Hasil Pengelompokan Metode *Ward*
- 2.6 Penentuan Nilai Pusat *Cluster*
- 2.7 Hasil Pengelompokan *K-means Clustering*
- 2.8 Validasi *Cluster*
- 2.9 Interpretasi *Cluster*
- 2.10 Analisis Profil

3. Kesimpulan

2. Hasil – Hasil Utama

2.1 Pendahuluan

Pada penelitian ini akan dilakukan analisis data pembangunan manusia berbasis gender di Provinsi Jawa Barat yaitu sebanyak 27 kabupaten/kota menggunakan *software R 3.1.2* pada data untuk *gender* laki-laki dan *gender* perempuan menggunakan metode *hybrid clustering*.

2.2 Standarisasi Variabel

Standarisasi variabel dilakukan karena variabel penelitian memiliki satuan yang berbeda-beda, oleh karena itu diperlukan standarisasi variabel sebagai berikut menggunakan *Z-score* [8]:

$$Z_i = \frac{X_i - \mu}{\sigma} \quad (1)$$

dengan,

- Z_i : nilai standarisasi variabel ke- i
- X_i : data ke- i
- μ : rata-rata
- σ : simpangan baku

Tabel 2.1 Standarisasi Variabel untuk Data Laki-Laki

Kode	Kabupaten	AHHL	HLSL	RLSL	PpKL	KLdP	LsTP
3201	Bogor	68.5	12.69	8.5	13228	82	64.69
3202	Sukabumi	67.76	12.35	6.89	11863	88	59.01
3203	Cianjur	67.12	12.08	6.57	10115	84	51.34
3204	Bandung	70.94	11.31	8.66	12661	72	67.81
3205	Garut	68.49	11.91	7.2	9739	86	60.84
3206	Tasikmalaya	66.02	12.41	7.15	10381	84	61.2
3207	Ciamis	68.35	14.51	7.8	12713	90	53.52
3208	Kuningan	70.22	12.72	7.65	12536	78	52.06
3209	Cirebon	69.28	11.54	6.86	14360	72	71.9
3210	Majalengka	66.69	11.91	7.06	12637	86	56.54
3211	Sumedang	69.88	12.86	8.01	12361	78	59.65
3212	Indramayu	68.31	11.57	6.19	13014	78	59.8
3213	Subang	69.22	11.34	6.83	13791	86	68.32
3214	Purwakarta	67.98	11.02	7.88	16209	75.56	64.66
3215	Karawang	69.45	11.54	7.43	13888	80	59.79
3216	Bekasi	71.12	11.96	9.08	15818	90	60.68
3217	Bandung Barat	69.57	11.05	7.97	11606	84	66.24
3218	Pangandaran	67.87	11.85	7.46	11944	85.71	70.79
3271	Kota Bogor	70.56	12.09	10.61	15748	84.44	64.39
3272	Kota Sukabumi	69.75	14.25	9.25	13583	85.71	55.92
3273	Kota Bandung	71.76	13.36	10.84	20743	94	56.7
3274	Kota Cirebon	69.76	12.71	10.08	14527	74.29	64.59
3275	Kota Bekasi	72.12	13.63	11.07	20494	84	55.38
3276	Kota Depok	71.91	13.41	11.25	20999	62	54.68
3277	Kota Cimahi	71.52	14.15	10.93	14888	73.91	54.45
3278	Kota Tasikmalaya	68.97	13.29	8.85	11598	95.56	60.8
3279	Kota Banjar	68.26	12.73	8.36	14759	96	60.05

dimana,

- AHHL : Angka harapan hidup Kabupaten/Kota Provinsi Jawa Barat tahun 2014 untuk laki-laki (tahun).
- HLSL : Angka harapan lama sekolah Kabupaten/Kota Provinsi Jawa Barat tahun 2014 untuk laki-laki (tahun).
- RLSL : Angka rata-rata lama sekolah Kabupaten/Kota Provinsi Jawa Barat tahun 2014 untuk laki-laki (tahun).

- PpKL : Pengeluaran per kapita yang disesuaikan Kabupaten/Kota Provinsi Jawa Barat tahun 2014 (Rp).
 Proporsi keterwakilan laki-laki
- KLdP : diparlemen Kabupaten/Kota Provinsi Jawa Barat tahun 2014.
- LsTP : Proporsi laki-laki sebagai tenaga profesional Kabupaten/Kota Provinsi Jawa Barat 2014.

Tabel 2.2 Standarisasi Variabel untuk Data Perempuan

Kode	Kabupaten/Kota	AHHP	HLSP	RLSP	PpKP	KPdP	PsTP
3201	Bogor	72.4	11.7	6.39	7527	18	35.3
3202	Sukabumi	71.6	11.9	5.98	5765	12	41
3203	Cianjur	70.9	11.5	6.04	4186	16	48.7
3204	Bandung	74.9	12.1	8	8127	28	32.2
3205	Garut	72.4	11.4	6.31	3884	14	39.2
3206	Tasikmalaya	69.8	12.4	6.59	4475	16	38.8
3207	Ciamis	72.2	13.5	7.1	5718	10	46.5
3208	Kuningan	74.1	11.7	6.88	5948	22	47.9
3209	Cirebon	73.2	12.3	5.74	4630	28	28.1
3210	Majalengka	70.5	11.3	6.31	5339	14	43.5
3211	Sumedang	73.8	13.5	7.31	8790	22	40.4
3212	Indramayu	72.2	11.6	4.72	6586	22	40.2
3213	Subang	73.1	12	5.67	7500	14	31.7
3214	Purwakarta	71.8	11.7	6.78	6659	24.4	35.3
3215	Karawang	73.3	11.7	6.13	7995	20	40.2
3216	Bekasi	75.1	11.5	7.76	7439	10	39.3
3217	Bandung Barat	73.5	11.3	7.33	3398	16	33.8
3218	Pangandaran	71.7	12.1	6.86	6239	14.3	29.2
3271	Kota Bogor	74.5	12.4	9.38	8836	15.6	35.6
3272	Kota Sukabumi	73.7	12.8	8.15	9210	14.3	44.1
3273	Kota Bandung	75.7	13.2	10.2	14945	6	43.3
3274	Kota Cirebon	73.7	13	9.01	10112	25.7	35.4
3275	Kota Bekasi	76.1	13	10	14245	16	44.6
3276	Kota Depok	75.9	13.2	10.2	12848	38	45.3
3277	Kota Cimahi	75.5	13.4	10.4	9677	26.1	45.6
3278	Kota Tasikmalaya	72.9	13.4	7.97	6816	4.44	39.2
3279	Kota Banjar	72.1	12.3	7.2	6488	4	40

Dimana,

- AHHP : Angka harapan hidup Kabupaten/Kota Provinsi Jawa Barat tahun 2014 untuk perempuan (tahun).
- HLSP : Angka harapan lama sekolah Kabupaten/Kota Provinsi Jawa Barat tahun 2014 untuk perempuan (tahun).
- RLSP : Angka rata-rata lama sekolah Kabupaten/Kota Provinsi Jawa Barat tahun 2014 untuk perempuan (tahun).
- PpKP : Pengeluaran per kapita yang disesuaikan Kabupaten/Kota Provinsi Jawa Barat tahun 2014 (Rp).
- KPdP : Proporsi keterwakilan perempuan diparlemen Kabupaten/Kota Provinsi Jawa Barat tahun 2014.
- PsTP : Proporsi perempuan sebagai tenaga profesional Kabupaten/Kota Provinsi Jawa Barat tahun 2014.

2.3 Hasil *Dendrogram* menggunakan Metode *Ward*

Dalam metode *ward*, jarak antara dua *cluster* adalah jumlah kuadrat antara dua *cluster* untuk seluruh variabel. Metode ini cenderung digunakan untuk mengkombinasi *cluster-cluster* dengan jumlah kecil.

Berdasarkan hasil analisis standarisasi variabel, maka data akan dianalisis menggunakan metode *ward* dengan mencari nilai *SSE* [4]:

$$SSE = \sum_{j=1}^N (X_j - \bar{X})' (X_j - \bar{X}) \quad (2)$$

dengan,

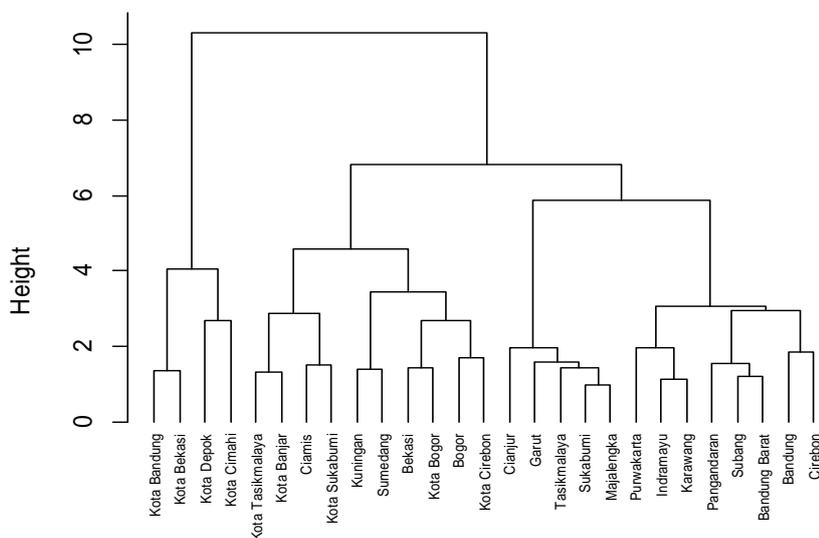
X_j : vektor objek ke- j

\bar{X} : Vektor rata-rata semua objek

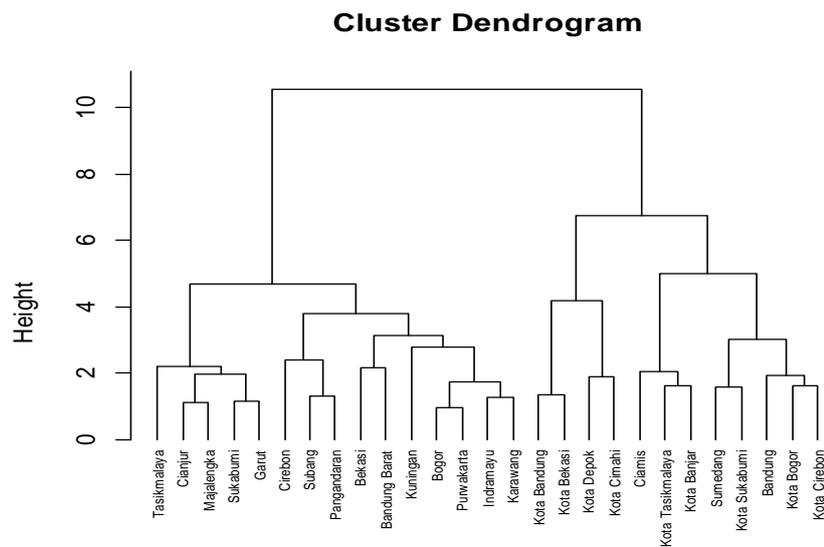
N : banyak objek

Dengan menggunakan persamaan (2), didapatkan hasil pengelompokan menggunakan metode *ward* untuk data laki-laki dan perempuan sebagai berikut:

Cluster Dendrogram



Gambar 2.1 *Dendrogram Ward Method* untuk Data Laki-laki



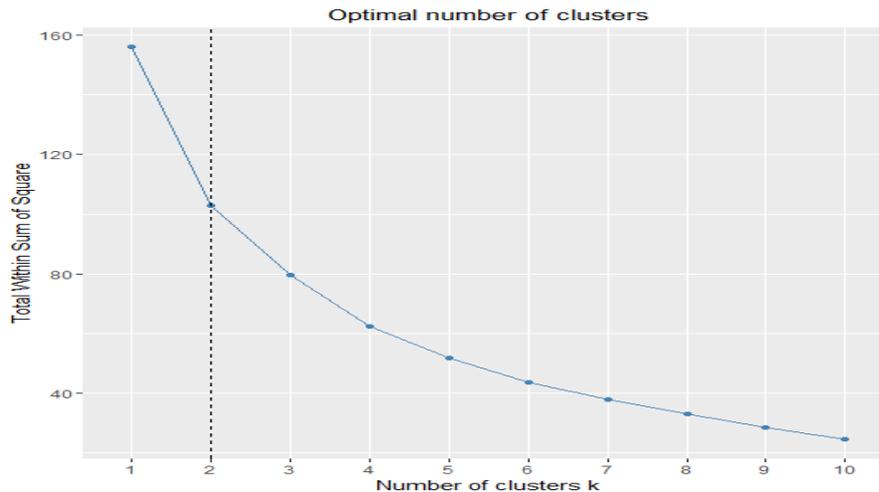
Gambar 2.2 Dendogram Ward Method untuk Data Perempuan

2.4 Penentuan Banyak Cluster Optimum

Metode *elbow* merupakan suatu metode yang digunakan untuk menghasilkan informasi dalam menentukan jumlah *cluster* terbaik dengan cara melihat persentase hasil perbandingan antara jumlah *cluster* yang akan membentuk siku atau yang mengalami penurunan nilai SSE terbesar pada suatu titik dan pada *cluster* selanjutnya nilai SSE akan menurun secara landai. Metode ini memberikan gagasan dengan cara memilih nilai *cluster* dan kemudian menambah nilai *cluster* tersebut untuk dijadikan model data dalam penentuan jumlah *cluster* terbaik. Selain itu, persentase perhitungan yang dihasilkan menjadi pembandingan antara jumlah *cluster* yang ditambahkan.

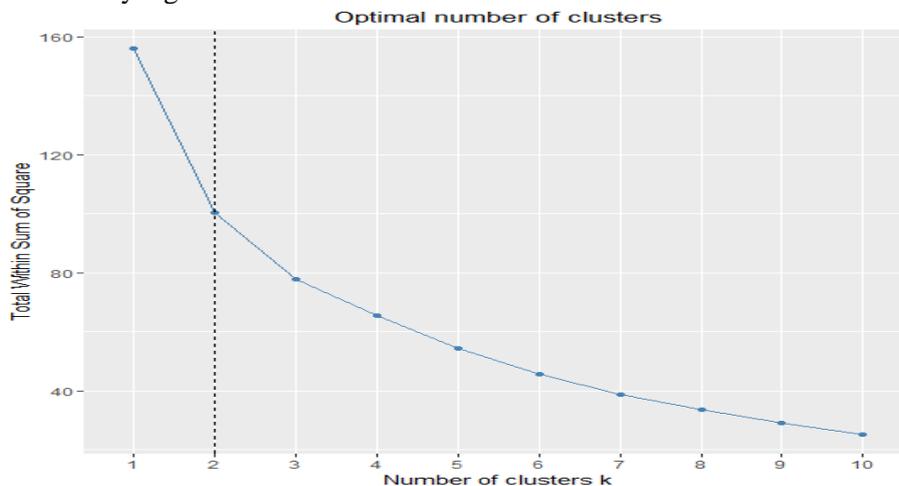
Hasil persentase yang berbeda dari setiap nilai *cluster* dapat ditunjukkan dengan menggunakan grafik sebagai sumber informasinya. Jika nilai *cluster* satu dengan yang lainnya memberikan sudut dalam grafik atau nilainya mengalami penurunan paling besar maka nilai *cluster* tersebut yang terbaik (Bholowalia, 2014).

Kodinariya dan Makwana [6] mengusulkan beberapa cara untuk menentukan nilai k sebagai jumlah *cluster* yang optimum, salah satunya adalah dengan cara metode *elbow* dengan menghitung nilai SSE per-*cluster* menggunakan persamaan 2, untuk menentukan banyak *cluster*, diperoleh banyak *cluster* sebagai berikut:



Gambar 2.3 Jumlah *Cluster* Optimum untuk Data Laki-laki

Banyak *cluster* optimum ditunjukkan oleh penurunan nilai SSE terbesar dari beberapa solusi *k cluster* yang dilakukan. Berdasarkan hasil metode elbow, dapat dilihat bahwa penurunan nilai SSE untuk data laki-laki terbesar yaitu untuk dua *cluster*, dimana hal tersebut menunjukkan jumlah *cluster* yang paling optimal dari *k cluster* yang dilakukan.

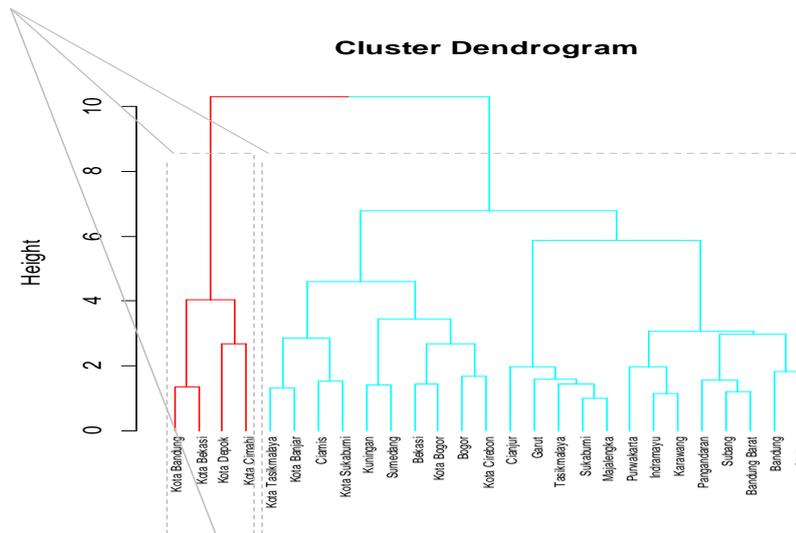


Gambar 2.4 Jumlah *Cluster* Optimum untuk Data Perempuan.

Banyak *cluster* optimum ditunjukkan oleh penurunan nilai SSE terbesar dari beberapa solusi *k cluster* yang dilakukan. Berdasarkan hasil analisis metode elbow, dapat dilihat bahwa penurunan nilai SSE terbesar untuk data perempuan yaitu untuk dua *cluster*, dimana hal tersebut menunjukkan jumlah *cluster* yang paling optimal dari *k cluster* yang dilakukan.

2.5 Hasil Pengelompokan Metode Ward

Berdasarkan hasil metode *elbow* untuk menentukan jumlah *cluster* optimum, didapatkan jumlah *cluster* optimum untuk data laki-laki dan perempuan adalah sebanyak 2 (dua) *cluster*. Berikut hasil dari metode *ward* dengan *cut-off* 2 (dua) *cluster*:



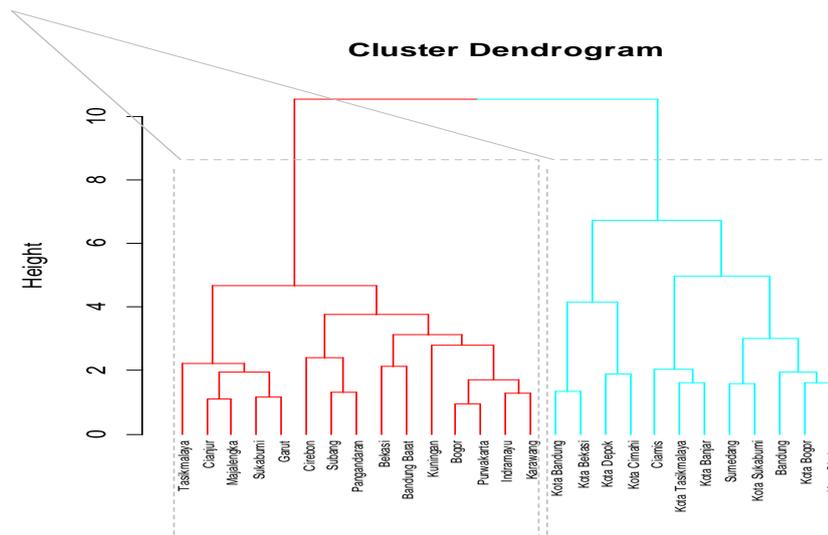
Gambar 2.5 Hasil Metode Ward untuk Data Laki-laki

Dalam Gambar 4.5, pengelompokan menggunakan metode ward dengan banyak cluster adalah 2 (dua) yang ditunjukkan dengan perbedaan warna. Garis dendrogram berwarna merah menunjukkan anggota cluster pertama dan garis dendrogram berwarna biru menunjukkan anggota cluster kedua.

Tabel 2.3 Hasil Metode Ward untuk Data Laki-laki

Cluster 1	
Kabupaten/Kota	Kabupaten/Kota
Bogor	Subang
Sukabumi	Purwakarta
Cianjur	Karawang
Bandung	Bekasi
Garut	Bandung Barat
Tasikmalaya	Pangandaran
Ciamis	Kota Bogor
Kuningan	Kota Sukabumi
Cirebon	Kota Cirebon
Majalengka	Kota Tasikmalaya
Sumedang	Kota Banjar
Indramayu	
Cluster 2	
Kabupaten/Kota	Kabupaten/Kota
Kota Bandung	Kota Depok
Kota Bekasi	Kota Cimahi

Dari Tabel 2.3 dapat diketahui bahwa banyak anggota cluster pada cluster pertama adalah sebanyak 23 kabupaten/kota (85,19%), dan banyaknya anggota cluster pada cluster kedua adalah sebanyak 4 kabupaten/kota (14,81%).



Gambar 2.6 Hasil Metode Ward untuk Data Perempuan

Dalam Gambar 2.6, pengelompokan menggunakan metode *ward* dengan banyak *cluster* adalah 2 (dua) yang ditunjukkan dengan perbedaan warna. Garis *dendrogram* berwarna merah menunjukkan anggota *cluster* pertama dan garis *dendrogram* berwarna biru menunjukkan anggota *cluster* kedua

Tabel 2.4 Hasil Ward untuk Data Perempuan

<i>Cluster 1</i>	
Bogor	Indramayu
Sukabumi	Subang
Cianjur	Purwakarta
Garut	Karawang
Tasikmalaya	Bekasi
Kuningan	Bandung Barat
Cirebon	Pangandaran
Majalengka	

<i>Cluster 2</i>	
Bandung	Kota Cirebon
Ciamis	Kota Bekasi
Sumedang	Kota Depok
Kota Bogor	Kota Cimahi
Kota Sukabumi	Kota Tasikmalaya
Kota Bandung	Kota Banjar

Dari Tabel 2.4 dapat diketahui bahwa banyak anggota pada *cluster* pertama adalah sebanyak 15 kabupaten/kota (55,56%), dan banyaknya anggota pada *cluster* kedua adalah sebanyak 12 kabupaten/kota (44,44%).

2.6 Penentuan Nilai Pusat Cluster

Penentuan nilai pusat *cluster* dilakukan untuk perhitungan pada pengelompokan menggunakan metode *k-means clustering*, dimana nilai pusat *cluster* berasal dari nilai rata-rata tiap *cluster* pada metode *clustering* menggunakan metode *ward*.

Tabel 2.5 Nilai Pusat Cluster Untuk Data Laki-Laki

Cluster	AHHL	HLSL	RLSL	PpKL	KLdP	LsTP
1	-0.2709	-0.2096	-0.3011	-0.314	0.09069	0.16497
2	1.55782	1.20535	1.73138	1.80544	-0.5215	-0.9486

Berdasarkan Tabel 2.5 didapatkan nilai pusat *cluster* untuk data laki-laki.

Tabel 2.6 Nilai Pusat Cluster Untuk Data Perempuan

Cluster	AHHP	HLSP	RLSP	PpKP	KPdP	PsTP
1	-0.4983	-0.7408	-0.6802	-0.5771	-0.0071	-0.2284
2	0.6229	0.9260	0.8502	0.7213	0.0089	0.2855

Berdasarkan Tabel 2.6 didapatkan nilai pusat *cluster* untuk data perempuan.

2.7 Hasil Pengelompokan K-means Clustering

K-means clustering merupakan salah satu metode data *clustering* non-hirarki yang mengelompokkan data dalam bentuk satu atau lebih *cluster*. Data yang memiliki karakteristik yang sama dikelompokkan dalam satu *cluster*, sedangkan data yang memiliki karakteristik yang berbeda dikelompokkan dengan *cluster* yang lain sehingga data yang berada dalam satu *cluster* memiliki tingkat variasi yang kecil (Agusta, 2007).

Pada penelitian ini, ukuran jarak yang akan digunakan pada metoda *k-means* adalah *euclidean distance*. *Euclidean distance* merupakan *dissimilarity measure* yang paling umum digunakan untuk data yang berskala interval atau rasio (Timm, 2002). Persamaan jarak *euclidean* secara umum adalah sebagai berikut (Johnson, 2002):

$$d_{ik} = \sqrt{\sum_{j=1}^J (y_{ij} - y_{kj})^2} \quad (3)$$

dengan,

d_{ik} : jarak *euclidean* objek ke-*i* dengan objek ke-*k*

y_{ij} : objek ke-*i* pada variabel ke-*j*

y_{kj} : objek ke-*k* pada variabel ke-*j*

Berdasarkan analisis *k-means clustering* dengan menggunakan metode *elbow* dihasilkan jumlah *cluster* optimal yaitu dua *cluster* untuk data laki-laki dan perempuan dan penentuan nilai pusat *cluster* (*centroid*) berasal dari nilai rata-rata tiap *cluster* yang diperoleh dari metode *ward* untuk pengelompokan kabupaten/kota di Provinsi Jawa Barat baik untuk data laki-laki maupun perempuan dengan keanggotaan setiap *cluster* sebagai berikut:

Tabel 2.7 Hasil *K-means Clustering* untuk Data Laki-laki

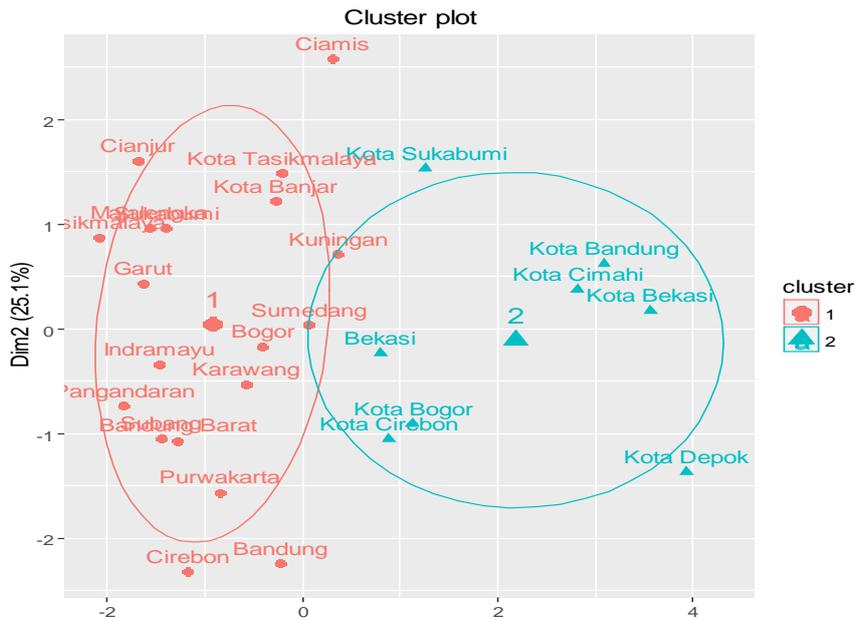
<i>Cluster 1</i>	
Bogor	Sumedang
Sukabumi	Indramayu
Cianjur	Subang
Bandung	Purwakarta
Garut	Karawang
Tasikmalaya	Bandung Barat
Ciamis	Pangandaran
Kuningan	Kota Tasikmalaya
Cirebon	Kota Banjar
Majalengka	
<i>Cluster 2</i>	
Bekasi	Kota Cirebon
Kota Bogor	Kota Bekasi
Kota Sukabumi	Kota Depok
Kota Bandung	Kota Cimahi

Dari Tabel 2.7 dapat diketahui bahwa banyak anggota pada *cluster* pertama adalah sebanyak 19 kabupaten/kota (70,37%), dan banyaknya anggota pada *cluster* kedua adalah sebanyak 8 kabupaten/kota (29,63%).

Tabel 2.8 Hasil *K-means Clustering* untuk Data Perempuan

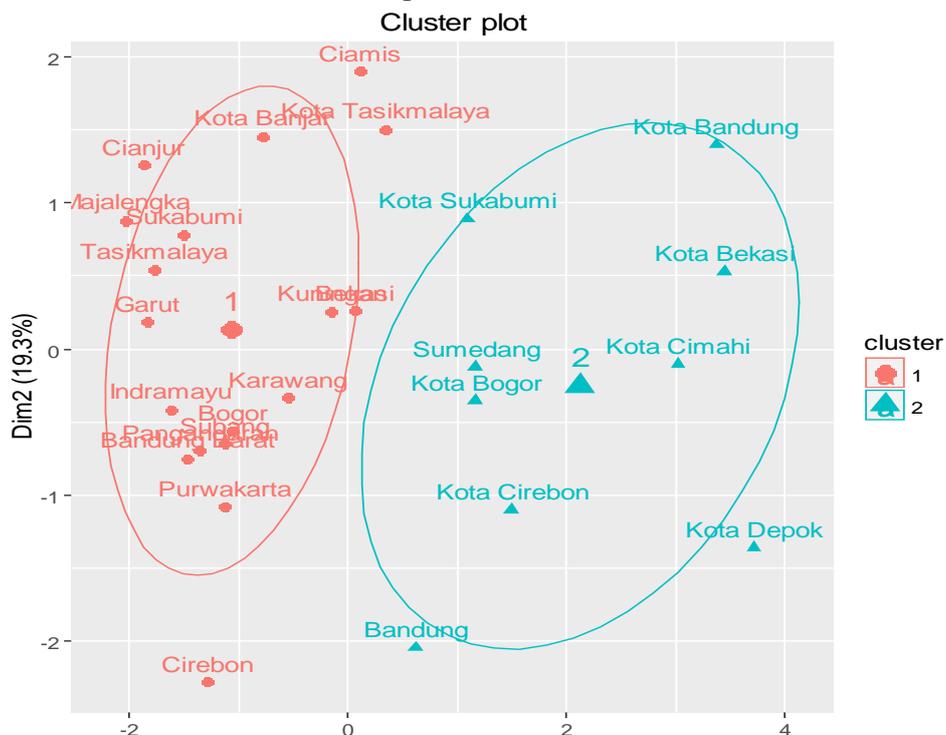
<i>Cluster 1</i>	
Bogor	Indramayu
Sukabumi	Subang
Cianjur	Purwakarta
Garut	Karawang
Tasikmalaya	Bekasi
Ciamis	Bandung Barat
Kuningan	Pangandaran
Cirebon	Kota Tasikmalaya
Majalengka	Kota Banjar
<i>Cluster 2</i>	
Bandung	Kota Cirebon
Sumedang	Kota Bekasi
Kota Bogor	Kota Depok
Kota Sukabumi	Kota Cimahi
Kota Bandung	

Dari Tabel 2.8 dapat diketahui bahwa banyak anggota pada *cluster* pertama adalah sebanyak 18 kabupaten/kota (66,67%), dan banyaknya anggota pada *cluster* kedua adalah sebanyak 9 kabupaten/kota (33,33%).



Gambar 2.7 Cluster Plot untuk Data Laki-Laki

Berdasarkan Gambar 2.7 dapat dilihat *cluster* plot untuk data laki-laki yang terdiri dari dua *cluster*. *Cluster* pertama disimbolkan dengan warna merah muda dan *cluster* kedua disimbolkan dengan warna biru.



Gambar 2.8 Cluster Plot untuk Data Perempuan

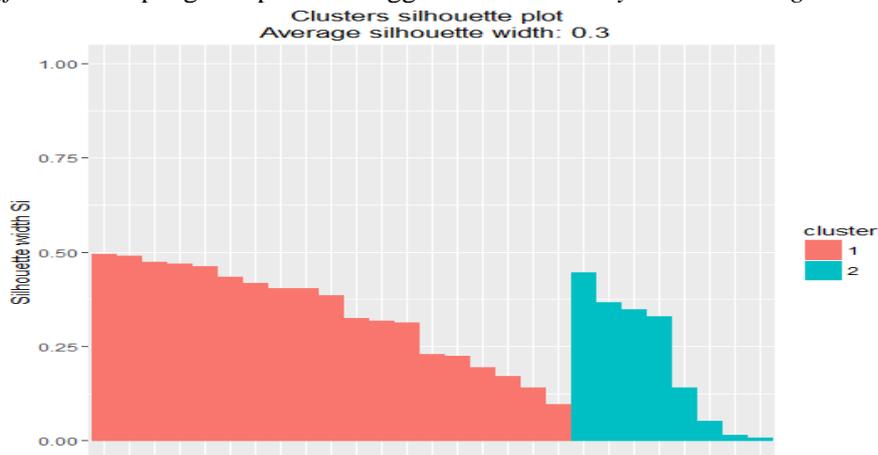
Berdasarkan Gambar 2.8 dapat dilihat *cluster* plot untuk data perempuan yang terdiri dari dua *cluster*. *Cluster* pertama disimbolkan dengan warna merah muda dan *cluster* kedua disimbolkan dengan warna biru.

2.8 Validasi Cluster

Berdasarkan pengelompokan menggunakan *hybrid clustering*, untuk mengetahui apakah hasil *cluster* yang didapatkan berada pada *cluster* yang tepat atau tidak [5], maka peneliti menggunakan metode *silhouette coefficient*.

$$S(i) = \frac{(b(i)-a(i))}{\text{Max}\{a(i),b(i)\}} \quad (4)$$

Nilai *silhouette* berada pada interval $-1 \leq s(i) \leq 1$. Berikut hasil *silhouette coefficient* dari pengelompokan menggunakan metode *hybrid clustering*:



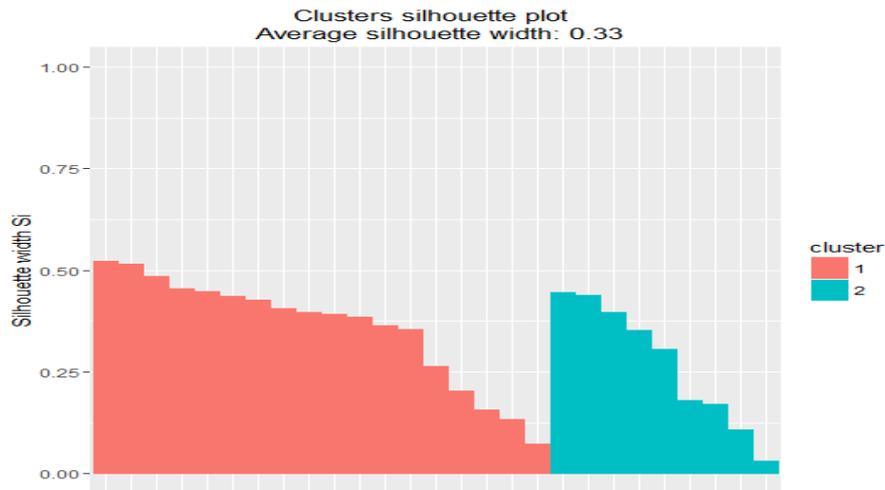
Gambar 2.9 Cluster Silhouette Plot untuk Data Laki-laki

Berdasarkan Gambar 2.9 dapat dilihat *cluster silhouette* plot untuk data laki-laki dimana nilai validasi untuk kedua *cluster* berada diatas nilai 0 yang artinya bahwa data tepat dikelompokkan kedalam *cluster* tersebut.

Tabel 2.9 Nilai *Silhouette Coefficient* untuk Data Laki-laki

<i>Cluster</i>	<i>Size</i>	<i>Ave.sil.width</i>
1	19	0.34
2	8	0.21

Berdasarkan Tabel 2.9 didapatkan nilai rata-rata *silhouette coefficient* untuk *cluster* pertama yaitu sebesar 0,34 dengan jumlah anggota pada *cluster* pertama sebanyak 19 kabupaten/kota dan nilai rata-rata *silhouette coefficient* untuk *cluster* kedua yaitu sebesar 0,21 dengan jumlah anggota pada *cluster* kedua sebanyak 8 kabupaten/kota.



Gambar 2.10 Cluster Silhouette Plot untuk Data Laki-laki

Berdasarkan Gambar 2.10 dapat dilihat *cluster silhouette* plot untuk data laki-laki dimana nilai validasi untuk kedua *cluster* berada diatas nilai 0 yang artinya bahwa data tepat dikelompokkan kedalam *cluster* tersebut.

Tabel 2.10 Nilai *Sillhouette Coefficient* untuk Data Perempuan

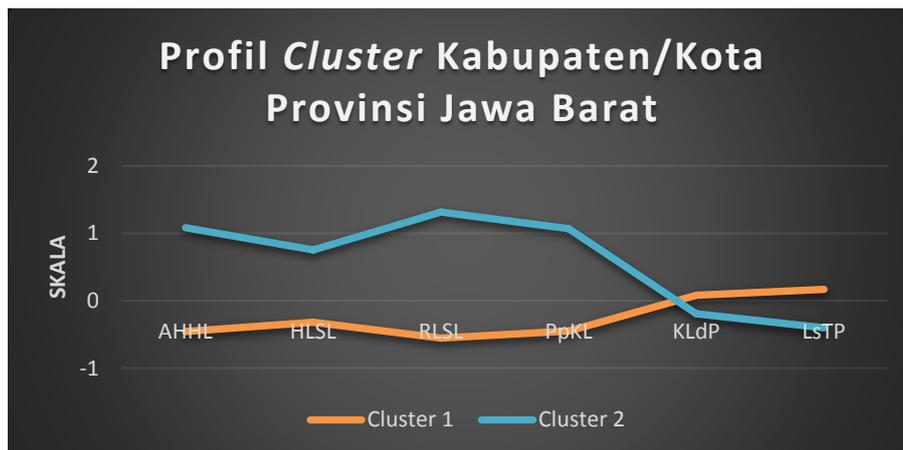
<i>Cluster</i>	<i>Size</i>	<i>Ave.sil.width</i>
1	18	0.36
2	9	0.27

Berdasarkan Tabel 2.10 didapatkan nilai rata-rata *sillhouette coefficient* untuk *cluster* pertama yaitu sebesar 0,36 dengan jumlah anggota pada *cluster* pertama sebanyak 18 kabupaten/kota dan nilai rata-rata *sillhouette coefficient* untuk *cluster* kedua yaitu sebesar 0,27 dengan jumlah anggota pada *cluster* kedua sebanyak 9 kabupaten/kota.

2.9 Interpretasi Cluster

Tahapan interpretasi *cluster* adalah untuk mencari karakter setiap kelompok yang khas dengan tujuan untuk memberi label pada masing-masing *cluster* tersebut, salah satunya dapat dilakukan dengan membandingkan *mean* pada masing-masing kelompok kabupaten/kota. Pada penelitian ini baik untuk *gender* laki-laki maupun perempuan didapatkan banyak *cluster* optimum adalah sebanyak dua *cluster*.

Berdasarkan karakteristik pembangunan manusia berbasis *gender* hasil dari output analisis *cluster* pada bahasan sebelumnya, perlu dilakukan *profiling* untuk memberikan label pada masing-masing tipe *cluster* yang terbentuk berdasarkan karakteristiknya dengan mengacu pada nilai *mean* tiap variabel yang telah distandarisasi dimana jika nilai variabel berada diatas rata-rata (nilai 0) akan dikatagorikan bernilai tinggi dan nilai berada dibawah rata-rata akan dikatagorikan bernilai rendah.



Gambar 2.11 Nilai Rata-rata *Cluster* Data Laki-Laki

Tabel 2.11 Nilai Rata-rata *Cluster* Data Laki-Laki

Variabel	<i>Cluster</i> 1	<i>Cluster</i> 2	Rata-rata
Angka harapan hidup untuk laki-laki	68.57	71.06	69.31
Angka harapan lama sekolah untuk laki-laki	12.14	13.20	12.45
Angka rata-rata lama sekolah untuk laki-laki	7.54	10.39	8.39
Pengeluaran per kapita yang disesuaikan untuk laki-laki	12600.16	17100	13933.44
Proporsi keterwakilan laki-laki diparlemen	83.20	81.04	82.56
Proporsi laki-laki sebagai tenaga profesional	61.53	58.35	60.59

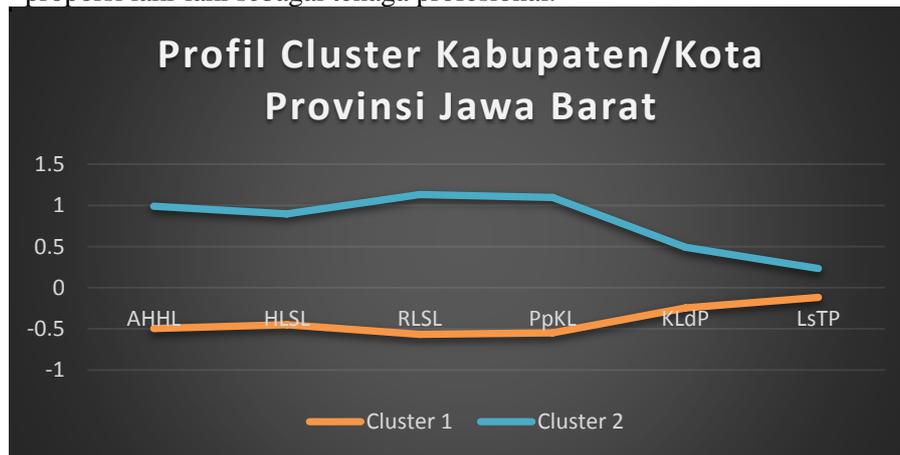
Berdasarkan hasil analisis yang telah dilakukan pada Gambar 2.11 dan Tabel 2.11 untuk data laki-laki, terlihat bahwa didapatkan dua *cluster* optimal. *Cluster* pertama terdiri atas 19 kabupaten/kota, yaitu Bogor, Sukabumi, Cianjur, Sumedang, Indramayu, Subang, Bandung, Garut, Tasikmalaya, Purwakarta, Karawang, Bandung Barat, Ciamis, Pangandaran, Kuningan, Kota Tasikmalaya, Cirebon, Kota Banjar, dan Majalengka. *Cluster* kedua terdiri atas 8 kabupaten/kota, yaitu Bekasi, Kota Bogor, Kota Sukabumi, Kota Bandung, Kota Cirebon, Kota Bekasi, Kota Depok, dan Kota Cimahi. Pada *cluster* pertama memiliki nilai IPM sedang dan nilai indikator IDG yang lebih tinggi dibanding *cluster* kedua sehingga *cluster* pertama dinamakan kabupaten/kota yang memiliki IPM sedang dan indikator IDG tinggi dan *cluster* kedua dinamakan kabupaten/kota yang memiliki IPM tinggi dan indikator IDG rendah.

Adapun karakteristik pada masing-masing *cluster* adalah sebagai berikut:

1. ***Cluster* pertama** merupakan kabupaten/kota dengan karakteristik pembangunan manusia berbasis *gender* memiliki ciri daerah dengan nilai angka harapan hidup rendah, angka harapan lama sekolah rendah, angka rata-rata lama sekolah rendah dan pengeluaran per kapita yang disesuaikan rendah akan tetapi proporsi keterwakilan laki-laki diparlemen dan proporsi laki-laki sebagai tenaga profesional memiliki angka yang tinggi. Oleh karena itu pada

cluster ini diperlukan perhatian terhadap variabel yang memiliki nilai yang rendah diantaranya angka harapan hidup, angka harapan lama sekolah dan angka rata-rata lama sekolah serta pengeluaran perkapita.

2. **Cluster kedua** merupakan kabupaten/kota dengan karakteristik pembangunan manusia berbasis *gender* memiliki ciri daerah dengan nilai angka harapan hidup tinggi, angka harapan lama sekolah tinggi, angka rata-rata lama sekolah tinggi dan pengeluaran per kapita yang disesuaikan tinggi akan tetapi proporsi keterwakilan laki-laki diparlemen dan proporsi laki-laki sebagai tenaga profesional memiliki angka yang rendah. Oleh karena itu pada *cluster* ini diperlukan perhatian dalam aspek keterwakilan laki-laki diparlemen dan proporsi laki-laki sebagai tenaga profesional.



Gambar 2.12 Nilai Rata-rata *Cluster* Data Perempuan

Tabel 2.12 Nilai Rata-rata *Cluster* Data Perempuan

Variabel	Cluster 1	Cluster 2	Rata-rata
Angka harapan hidup untuk perempuan	72.37	74.85	73.20
Angka harapan lama sekolah untuk perempuan	11.95	12.97	12.29
Angka rata-rata lama sekolah untuk perempuan	6.54	9.17	7.42
Pengeluaran per kapita yang disesuaikan untuk perempuan	5921.78	10754.44	7532.67
Proporsi keterwakilan perempuan diparlemen	15.51	21.29	17.44
Proporsi perempuan sebagai tenaga profesional	38.77	40.71	39.41

Berdasarkan hasil analisis yang terlampir pada Gambar 2.12 dan Tabel 2.12 untuk data perempuan, terlihat bahwa didapatkan dua *cluster* optimal. Kedua *cluster* memiliki perbedaan yang sangat terlihat tajam. *Cluster* pertama terdiri atas 18 kabupaten/kota, yaitu Bogor, Sukabumi, Cianjur, Garut, Tasikmalaya, Ciamis, Kuningan, Cirebon, Majalengka, Indramayu, Subang, Purwakarta, Karawang, Bekasi, Bandung Barat, Pangandaran, Kota Tasikmalaya, dan Kota Banjar. *Cluster* kedua terdiri atas 9 kabupaten/kota, yaitu Bandung, Sumedang, Kota Bogor, Kota

Sukabumi, Kota Bandung, Kota Cirebon, Kota Bekasi, Kota Depok, Dan Kota Cimahi. Pada *cluster* pertama memiliki nilai IPM rendah dan nilai indikator IDG lebih rendah dibanding *cluster* kedua sehingga *cluster* pertama dinamakan kabupaten/kota yang memiliki IPM dan indikator IDG rendah. Pada *cluster* kedua memiliki nilai IPM sedang dan nilai indikator IDG lebih tinggi dibanding *cluster* pertama sehingga *cluster* kedua dinamakan kabupaten/kota yang memiliki IPM sedang dan indikator IDG tinggi.

Adapun karakteristik pada masing-masing *cluster* adalah sebagai berikut:

1. **Cluster pertama** merupakan kabupaten/kota dengan karakteristik pembangunan manusia berbasis *gender* memiliki ciri daerah dengan nilai angka harapan hidup rendah, angka harapan lama sekolah rendah, angka rata-rata lama sekolah rendah dan pengeluaran per kapita yang disesuaikan rendah, proporsi keterwakilan perempuan diparlemen dan proporsi perempuan sebagai tenaga profesional memiliki angka yang rendah pula. Hal ini menunjukkan bahwa pada data perempuan untuk *cluster* pertama perlu adanya perbaikan dalam semua variabel pembangunan manusia berbasis *gender*.
2. **Cluster kedua** merupakan kabupaten/kota dengan karakteristik pembangunan manusia berbasis *gender* memiliki ciri daerah dengan nilai angka harapan hidup tinggi, angka harapan lama sekolah tinggi, angka rata-rata lama sekolah tinggi, pengeluaran per kapita yang disesuaikan untuk perempuan tinggi, proporsi keterwakilan perempuan diparlemen yang tinggi dan proporsi perempuan sebagai tenaga profesional memiliki angka yang tinggi. Hal ini menunjukkan anggota kabupaten/kota pada *cluster* ini untuk data perempuan sudah menunjukkan hasil yang sangat baik karena berdasarkan variabel penelitian menunjukkan untuk semua variabel penelitian memiliki nilai yang tinggi.

2.10 Analisis Profil

Dalam analisis profil terdapat 3 tahapan dalam pengujian, yaitu uji kesejajaran, berhimpit dan menguji apakah populasi kelompok memiliki nilai tengah yang sama [7]. Pada penelitian ini akan digunakan analisis profil untuk mengetahui apakah terdapat perbedaan populasi *cluster* yang dihasilkan pada analisis *cluster* menggunakan *hybrid clustering* untuk data kabupaten/kota di Pulau Jawa Barat berdasarkan indikator pembangunan manusia berbasis *gender*.

1. Uji keparalelan profil

Uji kesejajaran (paralel) untuk dua populasi yang menyebar normal dapat dituliskan sebagai berikut:

Hipotesis Uji:

$H_{01} : C_{\mu 1} = C_{\mu 2}$, artinya profil populasi kelompok pertama dan kedua sejajar (paralel)

$H_{11} : C_{\mu 1} \neq C_{\mu 2}$, artinya profil populasi kelompok pertama dan kedua tidak sejajar (tidak paralel)

$\alpha : 5\%$

Dimana C merupakan matriks kontras sedemikian, sehingga membuat persamaan seperti di bawah ini. C adalah matriks kontras

$$C_{(p-1) \times p} = \begin{bmatrix} -1 & 1 & \dots & 0 \\ 0 & -1 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

(5)

Statistik Uji: T^2

$$T^2 = (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)' \mathbf{C}' \left[\left(\frac{1}{n_1} + \frac{1}{n_2} \right) \mathbf{C} \mathbf{S}_{pooled} \mathbf{C}' \right]^{-1} \mathbf{C} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) > c^2$$

(6)

dengan

 $\bar{\mathbf{x}}_1$: matriks rata-rata variabel indikator pembangunan manusia berbasis *gender* pada populasi pertama $\bar{\mathbf{x}}_2$: matriks rata-rata variabel indikator pembangunan manusia berbasis *gender* pada populasi kedua n_1 : jumlah objek pada populasi pertama n_2 : jumlah objek pada populasi kedua \mathbf{S}_{pooled} : matriks kovarians gabungan

$$c^2 = \frac{(n_1 + n_2 - 2)(p - 1)}{n_1 + n_2 - p} F_{p-1, n_1 + n_2 - p}(\alpha)$$

(7)

Kriteria Uji: Tolak H_0 jika nilai dari $T^2 > c^2$ dengan nilai dari c^2 tergantung dari nilai sebaran F dengan db1 = p - 1 dan db2 = n1 + n2 - p pada (α).2. Uji Keberhimpitan profil (*Coincident*)

Ketika profil telah paralel, nilai pertama berada di atas nilai kedua di bawah kondisi ini, profil akan berhimpit hanya jika jumlah $\mu_{11} = \mu_{12} = \dots = \mu_{1p} = \mathbf{1}'\mu_1$ dan $\mu_{21} = \mu_{22} = \dots = \mu_{2p} = \mathbf{1}'\mu_2$ adalah sama. Pengujian hipotesis ini baru dapat dilakukan setelah uji pada kesejajaran dapat diterima. Maka hipotesis nol untuk langkah kedua dapat dituliskan sebagai berikut:

 H_{02} : $\mathbf{1}'\mu_1 = \mathbf{1}'\mu_2$, artinya profil populasi kelompok pertama dan kedua berhimpit H_{12} : $\mathbf{1}'\mu_1 \neq \mathbf{1}'\mu_2$, artinya profil populasi kelompok pertama dan kedua tidak berhimpit α : 5%Statistik Uji: T^2

$$T^2 = \mathbf{1}'(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) \left[\left(\frac{1}{n_1} + \frac{1}{n_2} \right) \mathbf{1}' \mathbf{S}_{pooled} \mathbf{1} \right]^{-1} \mathbf{1}'(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)$$

$$= \left(\frac{\mathbf{1}'(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2} \right) \mathbf{1}' \mathbf{S}_{pooled} \mathbf{1}}} \right) > t_{n_1 + n_2 - 2, \frac{\alpha}{2}}^2 = F_{1, n_1 + n_2 - 2}(\alpha) \quad (8)$$

Kriteria Uji: Tolak H_0 jika nilai dari $T^2 > t_{n_1 + n_2 - 2, \frac{\alpha}{2}}^2 = F_{1, n_1 + n_2 - 2}(\alpha)$.

3. Uji Kesamaan (Level)

Apabila profil-profil tersebut berhimpit (H_0 keberhimpitan diterima), maka seluruh observasi tersebut berasal dari populasi normal yang sama. Maka langkah selanjutnya adalah apakah seluruh peubahnya memiliki nilai rata-rata yang sama.

Asumsi bahwa profil sama, maka nilai tengah pasti sama. Sehingga hipotesisnya dapat dirumuskan sebagai berikut:

 H_{03} : $\mathbf{C}\mu = 0$, artinya profil populasi kelompok pertama dan kedua memiliki nilai tengah yang sama H_{13} : $\mathbf{C}\mu \neq 0$, artinya profil populasi kelompok pertama dan kedua tidak memiliki nilai tengah yang sama α : 5%

Statistik Uji:

$$S = (n_1 + n_2) \bar{\mathbf{x}}' \mathbf{C}' [\mathbf{C} \mathbf{S}_{pooled} \mathbf{C}']^{-1} \mathbf{C} \bar{\mathbf{x}} > F_{p-1, n_1 + n_2 - p}(\alpha) \quad (9)$$

Kriteria Uji: Tolak H_0 pada tingkat α , jika nilai $S > F_{p-1, n_1+n_2-p}(\alpha)$.

Hasil Analisis Profil

1) Uji kesejajaran untuk Data Laki-laki

Uji kesejajaran (paralel) untuk dua populasi *cluster* yang menyebar normal dengan menggunakan Persamaan (5) maka didapatkan hasil $T^2 = 4429.293$ dibandingkan dengan $c^2 = 15.98084$, maka H_0 ditolak, artinya profil populasi kelompok pertama dan kedua tidak sejajar (tidak paralel). Karena H_0 ditolak maka tidak adanya kesejajaran untuk dua *cluster* yang terbentuk pada data laki-laki sehingga tidak perlu dilakukan analisis berikutnya karena sudah pasti tidak berhimpit dan tidak ada kesamaan.

2) Uji kesejajaran untuk Data Perempuan

Uji kesejajaran (paralel) untuk dua populasi *cluster* yang menyebar normal dengan menggunakan Persamaan (5) maka didapatkan hasil $T^2 = 4896.038$ dibandingkan dengan $c^2 = 15.98084$, maka H_0 ditolak, artinya profil populasi kelompok pertama dan kedua tidak sejajar (tidak paralel). Karena H_0 ditolak maka tidak adanya kesejajaran untuk dua *cluster* yang terbentuk pada data perempuan sehingga tidak perlu dilakukan analisis berikutnya karena sudah pasti tidak berhimpit dan tidak ada kesamaan.

3. Kesimpulan

Berdasarkan hasil pembahasan yang telah dilakukan dapat ditarik kesimpulan sebagai berikut:

1. Analisis *hybrid clustering* digunakan untuk mengelompokkan 27 kabupaten/kota di Provinsi Jawa Barat berdasarkan indikator pembangunan berbasis *gender* dapat mengatasi kelemahan *k-means* dalam menentukan jumlah *cluster* dan penentuan pusat awal *cluster* (*centroid*) dan terbentuk dua *cluster* optimum untuk data laki-laki maupun perempuan dengan menggunakan metode ward untuk menentukan pusat awal *cluster*.
2. Ketepatan validasi *cluster* pada penelitian ini menggunakan *hybrid clustering* yang diukur oleh nilai *sillhouette coefficient* didapatkan nilai rata-rata *sillhouette coefficient* untuk data laki-laki yaitu sebesar 0,34 untuk *cluster* pertama dan 0,21 untuk *cluster* kedua. Untuk data perempuan didapatkan nilai rata-rata *sillhouette coefficient* yaitu sebesar 0,36 untuk *cluster* pertama dan 0,27 untuk *cluster* kedua dimana nilai $s(i)$ bernilai positif atau mendekati 1, hal ini menunjukkan bahwa objek tepat dikelompokkan pada *cluster* tersebut.
3. Dalam analisis profil pada penelitian ini digunakan untuk mengetahui kesamaan karakteristik profil populasi penduduk *cluster* pertama dan kedua pada masing-masing kabupaten/kota di Provinsi Jawa Barat berdasarkan indikator pembangunan manusia berbasis *gender* untuk data laki-laki dan perempuan berdasarkan enam variabel indikator pembangunan manusia berbasis *gender* didapatkan bahwa profil populasi kelompok pertama dan kedua tidak sejajar (tidak paralel).

Referensi

- [1] Alfina, Tahta., dkk. 2012. Analisa Perbandingan Metode Hierarchical Clustering, K-means dan Gabungan Keduanya dalam Cluster Data. *Skripsi*.

Surabaya: Jurusan Teknik Industri ITS.

- [2] Arai, Kohei., and et al. 2007. *Hierarchical K-means: an algorithm for centroids initialization for K- means*. Japan.
- [3] Badan Pusat Statistik. 2015. *Pembangunan Manusia berbasis Gender 2015*. Jakarta.
- [4] Johnson, R. A. 2002. *Applied Multivariate Analysis 5th Ed*. New Jersey: Prentice Hall Inc.
- [5] Kaufmann, L. 1987. *Silhouettes: a Graphical Aid to the Interpretation and Validation of Cluster Analysis. Computational and Applied Mathematics*.
- [6] Kodinariya, Trupti., Makwana, Prashant. 2013. *Review on determining number of Clustering K-Means Clustering*. India: *Department of Computer Engineering*.
- [7] Santoso, Budi. 2007. *Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu.
- [8] Sudjana. 2005. *Metoda Statistika*. Bandung: PT. Tarsito Bandung.
- [9] Sumadyo, Malikus. 2016. *Koefisien Silhouette, Metode Pengukuran Efektifitas Algoritma Klastering*. *Skripsi*. Bekasi: Universitas Islam.

METODE BOOTSTRAP BAGI MODEL POISSON DENGAN PENCILAN DAN OVERDISPERSI

KUTHA ARDANA¹, KHAIRIL ANWAR NOTODIPUTRO²,
KUSMAN SADIK³

1 Dep. Matematika – Institut Pertanian Bogor, kuthaardana@gmail.com
2 Dep. Statistika - Institut Pertanian Bogor, ² khairilnotodiputro@gmail.com
3 Dep. Statistika - Institut Pertanian Bogor, kusmansadik@gmail.com

Abstrak. Keberadaan pencilan pada gugus data cacahan menghasilkan penduga yang tak reliabel dan berbias. Penggunaan metode kekar yang dikombinasikan dengan bootstrap mampu mengatasi masalah pencilan pada model Poisson. Hasil simulasi numerik menunjukkan implementasi bootstrap dari hasil metode kekar CUBIF pada model Poisson efektif mengatasi pengaruh pencilan pada gugus data berukuran kecil. Bootstrap hasil pendugaan MLE tidak banyak berbeda dengan pendugaan MLE tanpa bootstrap. Hasil bootstrap dari penduga CUBIF pada kasus overdispersi efektif menurunkan bobot nilai respons yang besar. Metode pendugaan CUBIF tanpa bootstrap menghasilkan penduga di antara penduga MLE dan penduga CUBIF dengan bootstrap.

Kata kunci: bootstrap, model Poisson, pencilan, overdispersi, metode kekar.

1. Pendahuluan

Kehadiran pencilan (*outliers*) pada suatu gugus data dapat menimbulkan pengaruh yang tak diinginkan pada prosedur-prosedur statistik. Pada kasus seperti ini, penggunaan penduga kekar (*robust*) akan mengurangi dampak dari pengaruh pencilan. Namun, pada kebanyakan situasi, kesulitan terjadi dalam hal penurunan sebaran yang pasti bagi penduga kekar. Di sisi lain, bahkan ketika sebaran asimtotiknya dapat diperoleh, namun memiliki kekonvergenan yang cukup lambat. Dalam hal ini penggunaan metode bootstrap akan menghasilkan penduga yang lebih reliabel, bahkan pada kasus dengan ukuran sampel yang kecil (Amado *et al*, [1]).

Gagasan utama pada bootstrap adalah melakukan resampling terhadap data pengamatan secara berulang untuk membuat gugus data yang mirip dengan gugus data semula, kemudian melakukan pengepresan sebaran suatu penduga atau statistik tertentu. Berbagai teknik bootstrap telah diimplementasikan pada model linear (Efron, [3]; Davison dan Hinkley, [2]). Moulton & Zeger [8] mengimplementasikan metode bootstrap untuk model linear terampat (regresi

logistik). Teknik-teknik klasik bootstrap pada model linear dimodifikasi untuk memperhitungkan pengaruh acak pada model linear campuran (Hoai *et al*, [5]).

Masalah overdispersi lazim dijumpai pada model cacahan. Dalam kasus ini, $Var(Y_i) > \mu_i$. Hal ini dapat diakibatkan antara lain oleh adanya korelasi antar respons individu, sampling gerombol, proses agregat pada level data. Pada makalah ini metode bootstrap akan diimplementasikan pada model linear terampat khususnya model Poisson dengan mengombinasikan dengan metode kekar yang ada. Implementasi teknik bootstrap dan teknik kekar diharapkan dapat mengevaluasi akurasi maupun efisiensi pendugaan parameter model Poisson masalah keberadaan pencilan dan overdispersi.

2. Hasil-hasil Utama

2.1 Model Poisson

Misalkan y_i , untuk $i = 1, \dots, n$ bebas dan mengikuti sebaran Poisson. Fungsi kepekatan peluangnya diberikan oleh

$$P(Y = y) = \frac{e^{-\mu} \mu^y}{y!}, y = 0, 1, 2, \dots \quad (1)$$

dengan μ merupakan nilai harapan dan ragam bagi y_i .

Pandang suatu model linear terampat dengan peubah respons yang menyebar secara Poisson, $y_i \sim \text{Po}(\mu_i)$, dan kovariat $x_i \in \mathbb{R}^p$. Misalkan $g(\cdot)$ merupakan suatu fungsi monoton naik yang terdefinisi pada bilangan real positif. Fungsi yang lazim digunakan adalah “log”. Parameter model Poisson $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$ memenuhi hubungan

$$g(\mu_i) = \log(\mu_i) = \mathbf{x}_i^T \boldsymbol{\beta} = \eta_i. \quad (2)$$

Fungsi log-likelihood model Poisson bagi suatu vektor pengamatan bebas y_i adalah

$$l(\boldsymbol{\beta}) = \log L(\boldsymbol{\beta}) = \sum_{i=1}^n (y_i \log \mu_i - \mu_i). \quad (3)$$

Penurunan (3) terhadap $\boldsymbol{\beta}$ menghasilkan sistem persamaan likelihood maksimum

$$\sum_{i=1}^n h'(\eta_i) \frac{(y_i - \mu_i)}{\mu_i} = \mathbf{0}, \quad (4)$$

dengan $\mu_i = h(\eta_i)$ dan $h'(\eta_i) = \partial \mu_i / \partial \eta_i$. Untuk fungsi hubung kanonik $\mu_i = e^{\mathbf{x}_i^T \boldsymbol{\beta}}$, sistem persamaan (4) menghasilkan

$$\sum_{i=1}^n (y_i - e_i^{\mathbf{x}_i \boldsymbol{\beta}}) \mathbf{x}_i = \mathbf{0} \quad (5)$$

(McCullagh & Nelder, [7]).

2.2 Metode Kekar CUBIF

Metode kekar yang digunakan adalah metode CUBIF (*conditionally unbiased bounded-influence estimate*) yang dikenalkan Künsch *et al.* [6]. CUBIF termasuk subklas penduga-M dengan fungsi skor takbias bersyarat yang memenuhi

$$\mathbf{E}_{\boldsymbol{\beta}}(\psi(y, \mathbf{x}, \boldsymbol{\beta} | \mathbf{x})) = \int \psi(y, \mathbf{x}, \boldsymbol{\beta}) P_{\boldsymbol{\beta}}(dy | \mathbf{x}) = 0 \quad (6)$$

Mereka mengenalkan fungsi skor optimal sebagai

$$\psi_{cond}(\mathbf{y}, \mathbf{x}, \boldsymbol{\beta}, b, B) = W \left\{ \mathbf{y} - \boldsymbol{\mu} - c \left(\mathbf{x}^T \boldsymbol{\beta}, \frac{b}{(\mathbf{x}^T B \mathbf{x})^{1/2}} \right) \right\} \mathbf{x},$$

dengan $b > 0$ merupakan batas yang memenuhi $S(\psi) \leq b$, fungsi $c(\cdot)$ merupakan koreksi bias, B adalah matriks dispersi yang didefinisikan sebagai

$$B(\boldsymbol{\beta}) = \mathbf{E}\{\psi_{cond} \psi_{cond}^T\},$$

dan $(\mathbf{x}^T B \mathbf{x})^{1/2}$ menyatakan suatu ukuran leverage. Fungsi bobot, W , diberikan oleh

$$W = W_b((\mathbf{y}, \mathbf{x}, \boldsymbol{\beta}, b, B) = W_b(r(\mathbf{y}, \mathbf{x}, \boldsymbol{\beta}, b, B) (\mathbf{x}^T B \mathbf{x})^{1/2}),$$

dengan $r(\mathbf{y}, \mathbf{x}, \boldsymbol{\beta}, b, B) = \mathbf{y} - \boldsymbol{\mu} - c(\frac{b}{(\mathbf{x}^T B \mathbf{x})^{1/2}})$, $W_b(a) = H_b(a)/a$ dan $H_b(a)/a$ merupakan fungsi Huber.

2.3 Data

2.3.1 Data Bangkitan

Pembangkitan data untuk simulasi bootstrap dan metode kekar model Poisson diimplementasikan dalam suatu percobaan simulasi numerik menggunakan software R dengan spesifikasi sebagai berikut.

- Banyaknya data $n = 50$
- Parameter model
 - slope, $\beta_0 = 0.5$, intersep $\beta_1 = 0.5$
- Kovariat
 - $x \sim \text{Unif}(0, 2)$
- Respons

- Poisson
 - $y_1 \sim \text{Pois}(\lambda = \exp(\beta_0 + \beta_1 x))$
 - $y_2 \sim \text{Pois}(\lambda = \exp(\beta_0 + \beta_1 x))$. Sebanyak 10% atau 5 data y_2 bagi pasangan 5 data x terbesar diganti dengan data pencilan 20. Dengan kata lain, $y_2(x_{(46)}) = y_2(x_{(47)}) = y_2(x_{(48)}) = y_2(x_{(49)}) = y_2(x_{(50)}) = 20$, dengan $x_{(i)}$ menyatakan statistik tataan ke- i , $i = 46, 47, \dots, 50$.
- Overdispersi, d
 - $y.od_1 \sim \text{NegBinom}(size = \frac{\lambda}{d-1}, \mu = \lambda)$,
 $\lambda = \exp(\beta_0 + \beta_1 x_1), d = 5$
 - $y.od_2 \sim \text{NegBinom}(size = \frac{\lambda}{d-1}, \mu = \lambda), \lambda = \exp(\beta_0 + \beta_1 x_1), d = 5$. Sebanyak 10% atau 5 data $y.od_2$ bagi pasangan 5 data x terbesar diganti dengan data pencilan 50.

2.3.2 Data Riil

Pendugaan model Poisson dengan menggunakan metode bootstrap dan kekar juga diimplementasikan terhadap data kecelakaan pesawat udara sedunia dalam kurun waktu 1976 – 1985 (Hinde & Demetrio, [4]).

2.4 Bootstrap

Metode bootstrap yang digunakan pada studi ini adalah bootstrap nonparametrik. Dengan bootstrap nonparametrik, sebaran awal yang tidak diketahui dapat digantikan dengan sebaran empirik dari sampel. Jenis bootstrap nonparametrik yang digunakan adalah bootstrap kasus (*case bootstrap*). Bootstrap kasus merupakan resampling vektor pengamatan dengan pemulihan. Proses resampling dilakukan sebanyak B kali, kemudian dapat ditentukan rata-rata penduga bootstrap beserta bias dan simpangan baku serta selang kepercayaannya.

Misalkan $\hat{\theta}_b^*$ merepresentasikan penduga parameter pada sampel bootstrap ke- b dan B menyatakan banyaknya sampel bootstrap yang akan ditarik dari gugus data awal.

- Nilai harapan penduga bootstrap

$$\hat{\theta}_B = \frac{1}{B} \sum_{b=1}^B (\hat{\theta}_b^*)$$

- Simpangan baku bootstrap

$$\widehat{SE}_B = \sqrt{\frac{1}{B-1} \sum_{b=1}^B (\hat{\theta}_b^* - \hat{\theta}_B)^2}$$

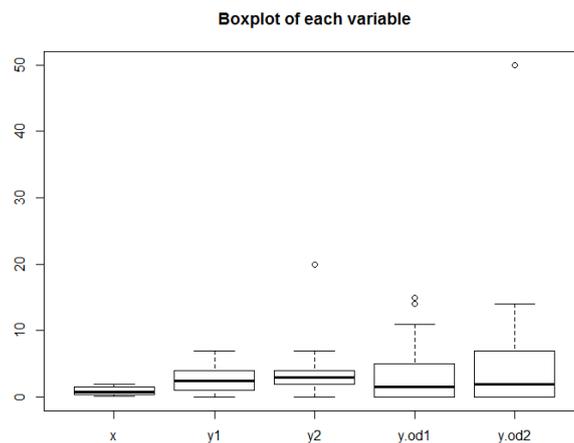
- Selang kepercayaan bootstrap

$$\hat{\theta}_{(\alpha.B)}^* \leq \theta \leq \hat{\theta}_{((1-\alpha).B)}^*$$

2.5 Hasil dan Pembahasan

2.5.1 Profil data bangkitan

Diagram kotak baris masing-masing peubah ditampilkan pada Gambar 1 berikut.



Gambar 1. Diagram kotak garis kovariat x dan peubah respons $y_1, y_2, y.od_1$ dan $y.od_2$

Kovariat x merupakan peubah kontinu yang menyebar Uniform (0,2). Peubah respons $y_1 \sim \text{Po}(\mu = \exp(0.5 + 0.5x))$ dengan kisaran $[0,7]$, sedangkan y_2 adalah y_1 dengan lima pencilan bernilai 20. Peubah respons $y.od_1$ berkisar dari 0 sampai 15 dengan rata-rata 3.28 dan ragam 19.31 ($\approx 6 \times$ rata-rata)], sedangkan $y.od_2$ adalah $y.od_1$ dengan lima pencilan bernilai 50.

2.5.2 Bootstrap dan metode kekar

Kombinasi dari keberadaan faktor pencilan (P), overdispersi (D), dan metode kekar/robust (R), menghasilkan delapan jenis skenario bootstrap. Angka 0 menyatakan absennya suatu faktor, sedangkan angka 1 menyatakan hadirnya suatu faktor. Sebagai contoh, P0.D0.R0 menyatakan hasil bootstrap pada kasus data tanpa pencilan, tanpa overdispersi, dan diduga tanpa metode kekar (menggunakan MLE). Skenario P1.D0.R1, misalnya, menyatakan hasil bootstrap pada kasus data dengan pencilan, tanpa overdispersi, dan diduga dengan metode kekar CUBIF.

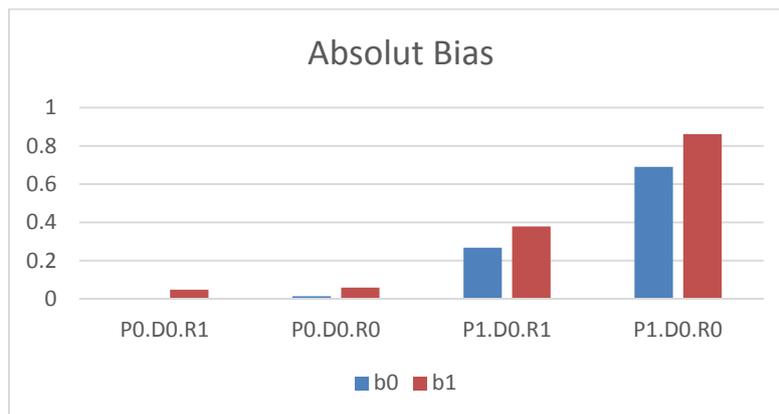
Tabel 1 menampilkan rata-rata bootstrap (dari 2000 resampel) bagi penduga intersep (b_0) dan slope (b_1) beserta biasnya pada kelompok skenario tanpa overdispersi. Pada kasus tanpa pencilan (P0.D0.R0, P0.D0.R1), baik metode MLE maupun metode kekar CUBIF menghasilkan penduga maupun bias yang relatif sama.

Kehadiran pencilan (P1.D0.R0, P1.D0.R1), mulai berpengaruh terhadap b_0 dan b_1 , namun metode kekar CUBIF menghasilkan penduga yang lebih resisten ($b_0 = 0.232$, $b_1 = 0.879$) dibandingkan pendugaan dengan MLE.

Tabel 1. Rataan bootstrap bagi penduga intersep (b_0) dan slope (b_1) beserta biasnya pada kelompok skenario tanpa overdispersi.

Skenario	Rataan Bootstrap		Bias	
	b_0	b_1	b_0	b_1
P0.D0.R0	0.514	0.441	0.014	-0.059
P0.D0.R1	0.505	0.452	0.005	-0.048
P1.D0.R0	-0.190	1.361	-0.690	0.861
P1.D0.R1	0.232	0.879	-0.268	0.379

Bias yang dihasilkan divisualisasikan pada Gambar 2. Tampak bahwa kehadiran pencilan yang diduga dengan MLE (P1.D0.R0) menghasilkan bias terbesar. Hal ini menunjukkan MLE sangat rentan dengan kehadiran pencilan. Bias penduga tereduksi sekitar 50% dengan menggunakan metode kekar CUBIF. Pada kasus tanpa pencilan, baik MLE maupun CUBIF menghasilkan bias yang relatif sama.

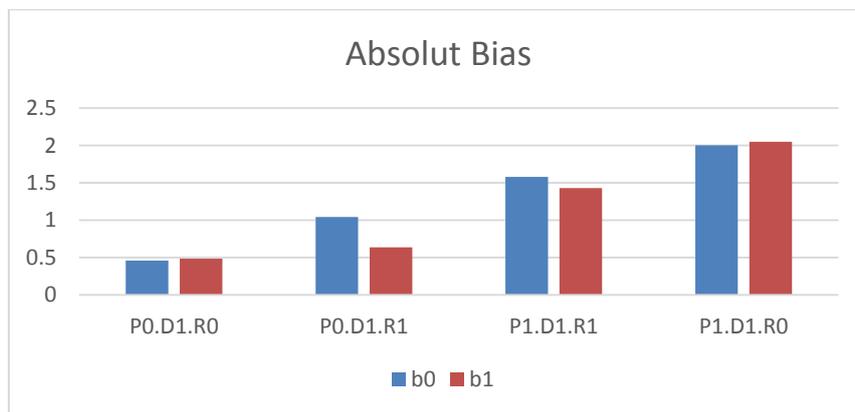


Gambar 2. Absolut bias bootstrap pada berbagai skenario untuk kasus tanpa overdispersi.

Hasil rataaan bootstrap beserta biasnya untuk kasus overdispersi dapat dilihat pada Tabel 2 dan visualisasi absolut biasnya ditampilkan pada Gambar 3. Overdispersi menghasilkan penduga rataaan bootstrap dengan bias yang cukup besar. Pola yang serupa dengan kasus tanpa dispersi juga tampak pada kasus overdispersi.

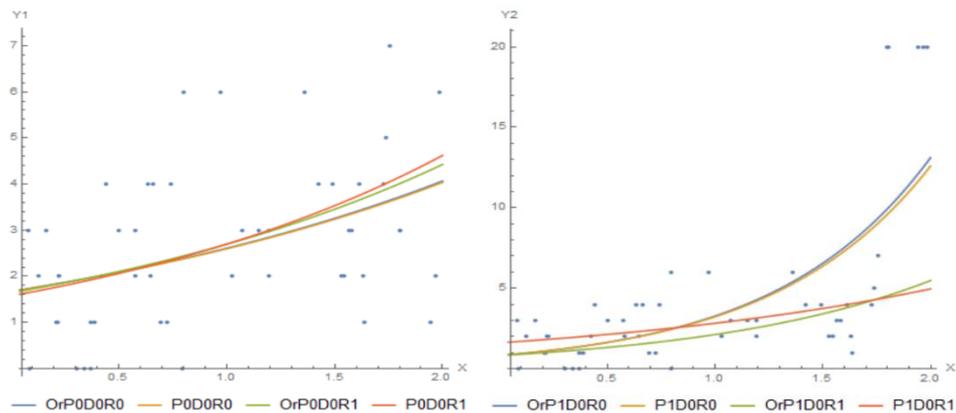
Tabel 2. Rataan bootstrap bagi penduga intersep (b_0) dan slope (b_1) beserta biasnya pada kelompok skenario dengan overdispersi.

Skenario	Rataan Bootstrap		Bias	
	b_0	b_1	b_0	b_1
P0.D1.R0	0.041	0.984	-0.459	0.484
P0.D1.R1	-0.542	1.137	-1.042	0.637
P1.D1.R0	-1.503	2.549	-2.003	2.049
P1.D1.R1	-1.078	1.929	-1.578	1.429



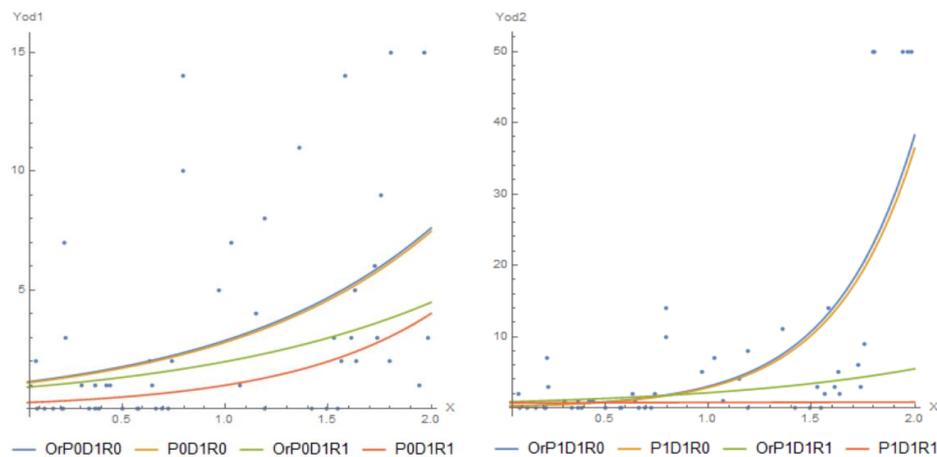
Gambar 3. Absolut bias bootstrap pada berbagai skenario untuk kasus overdispersi.

Efek pendugaan parameter model Poisson tanpa dan dengan bootstrap dapat dilihat pada Gambar 4 (tanpa overdispersi) dan Gambar 5 (dengan overdispersi). Skenario yang didahului “Or” menunjukkan pendugaan tanpa bootstrap. Tanpa kehadiran pencilan (Gambar 4 sebelah kiri), pendugaan dengan bootstrap maupun tanpa bootstrap menghasilkan penduga yang serupa, baik untuk MLE maupun CUBIF. Pada kasus dengan pencilan (Gambar 4 sebelah kanan), bootstrap tidak banyak berpengaruh pada pendugaan MLE. Pada pendugaan dengan metode kekar CUBIF, penduga dengan bootstrap cenderung menjauhi pencilan.



Gambar 4. Pendugaan parameter tanpa bootstrap (“Or”) dan dengan bootstrap untuk kasus tanpa overdispersi. Kiri: tanpa pencilan, kanan: dengan pencilan.

Pada kasus dengan overdispersi, bootstrap juga tampak tidak berpengaruh pada penduga MLE (Gambar 5). Pola pemisahan yang jelas tampak pada penduga CUBIF, dimana hasil bootstrap pada penduga CUBIF efektif menurunkan bobot nilai respons yang besar. Metode pendugaan CUBIF tanpa bootstrap menghasilkan penduga di antara penduga MLE dan penduga CUBIF dengan bootstrap.



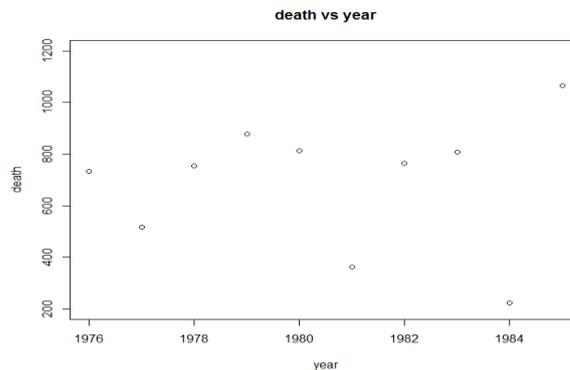
Gambar 5. Pendugaan parameter tanpa bootstrap (“Or”) dan dengan bootstrap untuk kasus dengan overdispersi. Kiri: tanpa pencilan, kanan: dengan pencilan.

2.5.3 Implementasi pada data riil

Pendugaan model Poisson dengan menggunakan metode bootstrap dan kekar diimplementasikan terhadap data kecelakaan pesawat udara sedunia dalam kurun waktu 1976 – 1985 (Hinde, [4]). Tabel 3 menampilkan data kecelakaan pesawat beserta plot tebaran korban kecelakaan (Gambar 6).

Tabel 3. Data kecelakaan pesawat udara sedunia, 1976 – 1985

year	accident	death	mile
1976	24	734	3863
1977	25	516	4300
1978	31	754	5027
1979	31	877	5481
1980	22	814	5814
1981	21	362	6033
1982	26	764	5877
1983	20	809	6223
1984	16	223	7433
1985	22	1066	7107



Gambar 6. Plot tebaran korban kecelakaan pesawat udara sedunia, 1976 – 1985.

Model yang digunakan

- $Y_i \sim \text{Pois}(m_i \lambda_i)$, $m_i = \text{passanger miles}$ sebagai peubah *offset*
- $\log \lambda_i = \beta_0 + \beta_1 \text{year}_i$
- Y_1 : *fatal accident*
- Y_2 : *passanger deaths*

Model *accident* sebagai peubah respons menghasilkan *residual deviance* = 5.4551 (df = 8) yang mengindikasikan model Poisson, sedangkan *passanger deaths* sebagai peubah respons menghasilkan *residual deviance* = 1051.4 (df = 8) yang mengindikasikan overdispersi. Hasil bootstrap dari penduga MLE dan metode kekar CUBIF terhadap model *passanger death* ditampilkan pada Tabel 4.

Tabel 4. Metode MLE dan CUBIF dengan dan tanpa bootstrap untuk model *passanger death*.

Metode	Penduga	Bootstrap	
		Tidak	Ya
<i>death</i> -MLE	b0	117.768	132.332
	b1	-0.061	-0.068
<i>death</i> -CUBIF	b0	105.428	72.182
	b1	-0.271	-0.043

Kedua metode menghasilkan slope negatif yang menunjukkan makin amannya penerbangan pada periode tersebut. Hasil ini juga sejalan dengan hasil yang diperoleh pada simulasi numerik. Penduga bootstrap dari penduga MLE relatif sama dengan tanpa bootstrap. Penduga slope dan intersep yang berbeda tampak pada penduga kekar CUBIF.

3. Kesimpulan

Implementasi bootstrap dari hasil metode kekar CUBIF pada model Poisson efektif mengatasi pengaruh pencilan pada gugus data berukuran kecil. Bootstrap hasil penduga MLE tidak banyak berbeda dengan tanpa bootstrap serta tetap rentan terhadap keberadaan pencilan. Hasil numerik mengindikasikan bahwa hasil bootstrap dari penduga CUBIF pada kasus overdispersi efektif menurunkan bobot nilai respons yang besar. Metode pendugaan CUBIF tanpa bootstrap menghasilkan penduga di antara penduga MLE dan penduga CUBIF dengan bootstrap.

Referensi

- [1] Amado, C., Bianco, A., Boente, G. and Pires, A. (2014). Robust bootstrap: an alternative to bootstrapping robust estimators. *Revsat*, 12 (2), pp. 169-197.
- [2] Davison, A.C. and Hinkley, D.V. (1997). *Bootstrap methods and their application*. Cambridge University Press.
- [3] Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Statist.* 7 1-26.
- [4] Hinde, J. and Demetrio, C. (2010). *Overdispersion: Models and Estimation*. Chapman & Hall/CRC.
- [5] Hoai, T.T., Mentre, F., Holford, N., Follet, C.F. and Comets, E. A comparison of bootstrap approaches for estimating uncertainty of parameters linear mixed-effects models, in *Pharmaceutical Statistics*. Wiley, 2013, 12 (3), pp.129-40.
- [6] Künsch, H. R., Stefanski, L. A. and Carroll, R. J. (1989). Conditionally unbiased bounded-influence estimation in general regression models, with applications to generalized linear models. *J. Amer. Statist. Assoc.* 84(406), 460–466.
- [7] McCullagh, P. and Nelder, J. (1989). *Generalized Linear Models* (second ed.). London: Chapman and Hall.
- [8] Moulton, L.H. and Zeger, S.L. (1991). Bootstrapping Generalized Linear Models. *Computational Statistics & Data Analysis* 11 53-6

KELUARGA PENAKSIR MEAN POPULASI YANG MEMANFAATKAN *AUXILIARY ATTRIBUTE* BESERTA PARAMETERNYA

DJULIO RAHMAWAN¹ • IDA FITHRIANI² • SITI NURROHMAH³

1. Departemen Matematika, FMIPA, Universitas Indonesia, Kampus Baru UI, Depok 16424, Indonesia, djulio.rahmawan@sci.ui.ac.id
2. Departemen Matematika, FMIPA, Universitas Indonesia, Kampus Baru UI, Depok 16424, Indonesia, ida.fithriani@gmail.com
3. Departemen Matematika, FMIPA, Universitas Indonesia, Kampus Baru UI, Depok 16424, Indonesia, snurrohmah@sci.ui.ac.id

Abstrak. Makalah ini membahas mengenai keluarga penaksir *mean* populasi. Penaksir merupakan statistik dari sampel yang digunakan untuk menaksir parameter populasi. Salah satu kriteria penaksir yang baik adalah penaksir dengan presisi yang tinggi. Dalam teori sampling, *auxiliary information* sering digunakan untuk meningkatkan presisi penaksir. *Auxiliary information* yang tersedia dalam bentuk kualitatif disebut *auxiliary attribute*. Nursel Koyuncu memperkenalkan keluarga penaksir *mean* populasi yang dapat memanfaatkan *auxiliary attribute* beserta parameternya. *Mean square error* dari keluarga penaksir Nursel Koyuncu dirumuskan dan ditunjukkan jika penaksir *mean* populasi yang dibentuk dari keluarga penaksir Nursel Koyuncu lebih efisien dibandingkan *mean* sampel. Pembuktian teori ini didukung dengan contoh numerik dengan menggunakan data harga apel di Amerika Serikat.

Kata kunci : penaksir, presisi, *auxiliary attribute*, *mean square error*.

1. Pendahuluan

Populasi adalah jumlah keseluruhan dari unit analisis yang ciri-cirinya akan diduga. Ciri atau karakteristik suatu populasi digambarkan dalam sebuah parameter populasi. Untuk mengetahui nilai dari parameter populasi tersebut, dilakukan pengambilan sampel. statistik dari sampel ini yang akan digunakan untuk menaksir parameter populasi. Dalam menaksir parameter populasi, salah satu hal yang harus diperhatikan adalah penaksir yang digunakan. Penaksir yang digunakan harus memiliki presisi yang tinggi agar hasil taksirannya menjadi lebih akurat. Salah satu cara meningkatkan presisi penaksir adalah dengan menggunakan *auxiliary information*.

Auxiliary information merupakan informasi tambahan yang memiliki korelasi dengan variabel observasi. Penggunaan *Auxiliary information* telah dikembangkan oleh ahli statistika mulai tahun 1930an. Awalnya, Bowle dan Neyman memiliki intuisi mengenai pemanfaatan variabel tambahan yang memiliki korelasi dengan variabel observasi. Mereka beranggapan penggunaan variabel tambahan dapat meningkatkan keakuratan hasil taksiran dalam menaksir *mean* atau total suatu populasi. Penelitian oleh ahli statistika seperti Watson(1937), Cochran (1940), Tripathi (1970) menunjukkan bahwa penggunaan variabel tambahan yang

memiliki korelasi dengan variabel observasi dapat dimanfaatkan dalam prosedur penaksiran. Akhirnya pada saat ini penggunaan *Auxiliary information* biasa digunakan dalam teori sampling.

Auxiliary information tidak selalu tersedia dalam bentuk data yang kuantitatif tetapi juga dapat tersedia dalam bentuk data yang kualitatif atau disebut dengan *auxiliary attribute*. Contohnya adalah hasil produksi susu sapi perah yang mungkin dipengaruhi oleh keturunan sapi perah. Dalam kasus ini, keturunan sapi dapat dijadikan sebagai *auxiliary attribute* untuk mendapatkan taksiran rata-rata produksi susu sapi yang lebih akurat. Meskipun dalam bentuk data kualitatif, *auxiliary attribute* tersebut tetap dapat dimanfaatkan dalam prosedur penaksiran selama variabel memiliki korelasi dengan variabel observasi. Namun, disamping keuntungannya, penggunaan *auxiliary information* memiliki kelemahan yaitu bahwa penaksir yang memanfaatkan *auxiliary information* baik dalam bentuk data yang kuantitatif atau kualitatif merupakan penaksir yang bias (Zahoor, 2013).

Pentingnya penggunaan *auxiliary information* dalam meningkatkan keakuratan hasil taksiran, membuat para ahli statistika terus berusaha mengembangkan pemanfaatan *auxiliary information*. Hasilnya, beberapa ahli statistika mulai melibatkan parameter dari variabel tambahan dalam prosedur penaksiran untuk meningkatkan presisi penaksir. Sen (1978), Sisodiya dan Dwivedi (1981) dan Upadhyaya dan Singh (1984) adalah ahli statistika yang memperkenalkan penggunaan koefisien variasi dari variabel tambahan dalam menaksir *mean* populasi. Singh dan Kakaran (1993) memperkenalkan penggunaan koefisien kurtosis dari variabel tambahan dalam menaksir *mean* populasi. Zaizai Yan dan Bing Tian (2010) memperkenalkan penggunaan koefisien skewness dalam menaksir *mean* populasi. Dalam jurnalnya, para ahli statistika tersebut telah membuktikan jika penggunaan parameter dari variabel tambahan seperti koefisien variasi, koefisien kurtosis dan yang lainnya juga dapat meningkatkan presisi penaksir *mean* populasi. Oleh karena itu, pada saat ini para ahli statistika mulai memperkenalkan keluarga penaksir *mean* populasi yang dapat memanfaatkan seluruh parameter dari variabel tambahan. Nursel Koyuncu adalah salah satunya. Dia memperkenalkan keluarga penaksir yang dapat memanfaatkan *auxiliary attribute* dalam *simple random sampling*. Keluarga penaksir tersebut dapat menyertakan parameter dari variabel *attribute* seperti koefisien variasi, koefisien kurtosis dan yang lainnya untuk meningkatkan efisiensi penaksir. Dalam makalah ini, akan dibahas mengenai bentuk keluarga penaksir *mean* populasi yang diperkenalkan oleh Nursel Koyuncu. Selanjutnya akan dibahas bagaimana membentuk penaksir *mean* populasi dari keluarga tersebut dan membandingkan efisiensinya dengan *mean* sampel dalam *simple random sampling*.

2. Hasil - Hasil Utama

Definisikan Y sebagai notasi variabel random untuk populasi dan y sebagai notasi variabel random untuk sampel. Misalkan (Y_1, Y_2, \dots, Y_N) adalah nilai-nilai dari populasi berhingga berukuran N dengan distribusi *uniform* dan \bar{Y} adalah *mean* populasi yang didefinisikan sebagai

$$\bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i$$

Misalkan (y_1, y_2, \dots, y_n) sampel acak berukuran n yang diambil dari populasi

dengan menggunakan *simple random sampling* tanpa pengembalian dimana setiap anggota populasi memiliki peluang yang sama untuk dipilih. *Mean* sampel didefinisikan sebagai

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

Misalkan akan dicari taksiran *mean* populasi untuk variabel observasi Y . Sampel berukuran n diambil dari populasi berukuran N dengan menggunakan *simple random sampling* tanpa pengembalian. Pandang suatu statistik sampel yaitu *mean* sampel. Akan dibuktikan jika *mean* sampel adalah penaksir yang tak bias untuk *mean* populasi. *Mean* sampel merupakan penaksir yang tak bias untuk *mean* populasi jika $E(\bar{y}) = \bar{Y}$.

$$\begin{aligned} \bar{y} &= \frac{1}{n} \sum_{i=1}^n y_i \\ &= \frac{1}{n} (y_1 + \dots + y_n), \end{aligned}$$

Sebelumnya akan dibuktikan jika $E(y_i) = \bar{Y}$ dimana $i = 1, 2, \dots, n$

$$E(y_i) = \sum_{i=1}^N p_i Y_i = \sum_{i=1}^N \frac{1}{N} Y_i = \bar{Y}$$

maka

$$\begin{aligned} E(\bar{y}) &= \frac{1}{n} E(y_1 + \dots + y_n) \\ &= \frac{1}{n} [E(y_1) + \dots + E(y_n)] \\ &= \frac{1}{n} [\bar{Y} + \dots + \bar{Y}] \\ &= \frac{n\bar{Y}}{n} \\ &= \bar{Y}, \end{aligned}$$

Terbukti jika \bar{y} merupakan penaksir yang tak bias untuk \bar{Y} . *Mean* sampel juga merupakan *Unbiased Minimum Variance Estimator* (UMVE), artinya *mean* sampel memiliki variansi yang minimum. Untuk membuktikan jika *mean* sampel \bar{y} memiliki variansi yang minimum, akan digunakan teorema Lehmann and Scheffe.

Teorema 2.1 (Teorema Lehmann and Scheffe) Misalkan (X_1, X_2, \dots, X_n) , n suatu bilangan bulat positif, menyatakan suatu sampel acak dari suatu distribusi yang mempunyai pdf $f(x; \theta)$, $\theta \in \Omega$. Misalkan $Y_1 = u_1(X_1, X_2, \dots, X_n)$ adalah suatu statistik cukup untuk θ dan misalkan famili pdf $\{g_1(y_1; \theta): \theta \in \Omega\}$ adalah *complete*. Jika terdapat suatu fungsi dari Y_1 yaitu $\psi(Y_1)$ yang merupakan penaksir *unbiased* dari θ , maka $\psi(Y_1)$ adalah penaksir *unbiased* dengan variansi minimum yang unik dari θ . (Hogg, & Craig, 2012)

BUKTI. Misalkan $Y_2 = u_2(X_1, X_2, \dots, X_n)$ adalah penaksir tak bias yang lain dari θ . Berdasarkan Teorema Rao-Blackwell, maka $E(Y_2|y_1)$ adalah penaksir tak bias untuk θ dengan $Var(Y_2|Y_1) < Var(Y_2)$. Karena $E(Y_2|y_1)$ adalah fungsi yang mengandung Y_1 , berdasarkan definisi *completeness*, maka $E(Y_2|y_1) = \psi(Y_1)$ sehingga $Var(\psi(Y_1)) < Var(Y_2)$ dan $\psi(Y_1)$ merupakan *Unbiased Minimum Variance Estimator* (UMVE) yang unik.

Langkah pertama, akan dibuktikan jika *mean* sampel \bar{y} adalah penaksir

yang tak bias untuk *mean* populasi \bar{Y} . Pada pembuktian sebelumnya, telah ditunjukkan jika *mean* sampel \bar{y} merupakan penaksir yang tak bias untuk *mean* populasi \bar{Y} .

Selanjutnya, akan dibuktikan jika *mean* sampel merupakan statistik cukup untuk *mean* populasi \bar{Y} .

Teorema 2.2 (Faktorisasi Neyman) Misalkan (Y_1, Y_2, \dots, Y_n) , n suatu bilangan bulat positif, menyatakan suatu sampel acak dari suatu distribusi yang mempunyai pdf $f(y; \theta)$, $\theta \in \Omega$. Statistik $Z_1 = u_1(Y_1, Y_2, \dots, Y_n)$ adalah suatu statistik cukup untuk θ jika dan hanya jika dapat ditemukan dua fungsi non negatif k_1 dan k_2 sedemikian sehingga

$$f(y_1; \theta)f(y_2; \theta)\dots f(y_n; \theta) = k_1[(y_1, y_2, \dots, y_n); \theta]k_2(y_1, y_2, \dots, y_n),$$

dimana $k_2(y_1, y_2, \dots, y_n)$ tidak bergantung pada θ . (Hogg, & Craig, 2012)

BUKTI. Teorema ini akan dibuktikan untuk variabel acak yang berjenis kontinu. Asumsikan jika teorema Faktorisasi Neyman terpenuhi yaitu

$$f(y_1; \theta)f(y_2; \theta)\dots f(y_n; \theta) = k_1[(y_1, y_2, \dots, y_n); \theta]k_2(y_1, y_2, \dots, y_n).$$

Akan dibuktikan jika $Z_1 = u_1(y_1, y_2, \dots, y_n)$ adalah suatu statistik cukup untuk θ . Misalkan Z_1 mempunyai pdf $g_1(z_1, \theta) = g_1[u_1(y_1, y_2, \dots, y_n; \theta)]$. Statistik Z_1 adalah suatu statistik cukup untuk θ jika

$$\frac{f(y_1; \theta)f(y_2; \theta)\dots f(y_n; \theta)}{g_1[u_1(y_1, y_2, \dots, y_n)]} = H(y_1, y_2, \dots, y_n),$$

dimana $H(y_1, y_2, \dots, y_n)$ tidak bergantung pada θ .

Untuk membuktikan hal di atas terlebih dahulu akan dicari pdf dari Z_1 yaitu $g_1(z_1, \theta) = g_1[u_1(y_1, y_2, \dots, y_n; \theta)]$ yang merupakan pdf marginal dari $g(z_1, z_2, \dots, z_n; \theta)$. Buat transformasi 1-1 :

$$\begin{aligned} z_1 &= u_1(y_1, y_2, \dots, y_n) \text{ dengan fungsi invers } y_1 = w_1(z_1, z_2, \dots, z_n) \\ z_2 &= u_2(y_1, y_2, \dots, y_n) \text{ dengan fungsi invers } y_2 = w_2(z_1, z_2, \dots, z_n) \\ &\dots \end{aligned}$$

$$\begin{aligned} z_n &= u_n(y_1, y_2, \dots, y_n) \text{ dengan fungsi invers } y_n = \\ &w_n(z_1, z_2, \dots, z_n) \end{aligned}$$

dengan *Jacobian* J , maka

$$\begin{aligned} g(z_1, z_2, \dots, z_n; \theta) &= f(y_1; \theta)f(y_2; \theta)\dots f(y_n; \theta)|J| \\ &= k_1(z_1; \theta)k_2(w_1, w_2, \dots, w_n)|J| \end{aligned}$$

dimana $w_i = w_i(z_1, z_2, \dots, z_n)$, $i = 1, 2, \dots, n$

pdf dari Z_1 yaitu $g_1[u_1(y_1, y_2, \dots, y_n; \theta)] = g_1(z_1; \theta)$ didapat dari

$$\begin{aligned} g_1(z_1; \theta) &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g(z_1, z_2, \dots, z_n; \theta) dy_2 \dots dy_n \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} k_1(z_1; \theta)k_2(w_1, w_2, \dots, w_n)|J| \\ &= k_1(z_1; \theta) \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} |J|k_2(w_1, w_2, \dots, w_n) dy_2 \dots dy_n, \end{aligned}$$

k_2 tidak bergantung pada θ , *Jacobian* J dan batas-batas integral pun tidak mengandung θ sehingga hasil integral persamaan diatas tidak mengandung θ , dan dapat dinyatakan sebagai fungsi dari z_1 yaitu $m(z_1)$ sehingga

$$g_1(z_1; \theta) = k_1(z_1; \theta)m(z_1),$$

jika $m(z_1) = 0$, maka $g_1(z_1; \theta) = 0$ sedangkan jika $m(z_1) > 0$, dapat dituliskan

$$k_1[u_1(y_1, y_2, \dots, y_n); \theta] = \frac{g_1[u_1(y_1, y_2, \dots, y_n)]}{m[u_1(y_1, y_2, \dots, y_n)]},$$

dan faktorisasi menjadi

$$f(y_1; \theta) f(y_2; \theta) \dots f(y_n; \theta) = g_1[u_1(y_1, y_2, \dots, y_n); \theta] \frac{k_2(y_1, y_2, \dots, y_n)}{m[u_1(y_1, y_2, \dots, y_n)]}$$

karena fungsi k_2 dan m tidak mengandung θ maka berdasarkan definisi Z_1 merupakan statistik cukup untuk θ .

Misalkan (Y_1, Y_2, \dots, Y_N) adalah nilai-nilai dari populasi berhingga berukuran N dengan distribusi *uniform* dengan

$$f(y) = \frac{1}{N}, \text{ untuk } y = Y_1, Y_2, \dots, Y_N$$

maka *mean* dan variansi dari Y adalah

$$\mu = \frac{1}{N} \sum_{i=1}^N Y_i \text{ dan } \sigma^2 = \frac{1}{N} \sum_{i=1}^N (Y_i - \mu)^2$$

misalkan y_1, y_2, \dots, y_n adalah sampel acak berukuran n yang diambil dari populasi tersebut. Pandang statistik mean sampel \bar{y} sebagai Z_1 dimana $Z_1 = \frac{1}{n} \sum_{i=1}^n y_i$. Akan dibuktikan jika Z_1 merupakan statistik cukup untuk $\mu = \bar{Y}$.

Perhatikan jika

$$\frac{1}{N} = \frac{\mu}{\sum_{i=1}^N Y_i} \text{ dan } \frac{1}{n} = \frac{\bar{y}}{\sum_{i=1}^n y_i}$$

sehingga

$$\begin{aligned} f(y_1; \theta) f(y_2; \theta) \dots f(y_n; \theta) &= \left(\frac{1}{N}\right)^n \\ &= \left(\frac{\mu}{\sum_{i=1}^N Y_i}\right)^{\frac{\sum_{i=1}^n y_i}{\bar{y}}} \\ &= (\mu)^{\frac{\sum_{i=1}^n y_i}{\bar{y}}} \left(\frac{1}{\sum_{i=1}^N Y_i}\right)^{\frac{\sum_{i=1}^n y_i}{\bar{y}}} \end{aligned}$$

pilih fungsi

$$k_1 = (\mu)^{\frac{\sum_{i=1}^n y_i}{\bar{y}}} \text{ dan } k_2 = \left(\frac{1}{\sum_{i=1}^N Y_i}\right)^{\frac{\sum_{i=1}^n y_i}{\bar{y}}}$$

karena k_1 adalah fungsi yang mengandung $(y_1, y_2, \dots, y_n; \mu)$ dan k_2 adalah fungsi dari (y_1, y_2, \dots, y_n) yang tidak bergantung μ , maka berdasarkan teorema Faktorisasi Neyman, Z_1 merupakan statistik cukup untuk μ sehingga *mean* sampel merupakan statistik cukup untuk *mean* populasi \bar{Y} .

Terakhir, akan dibuktikan jika *family* dari pdf *mean* sampel \bar{y} merupakan *family* yang *complete*.

Definisi 2.1 (Statistik Complete) Misalkan variabel acak Z yang berjenis kontinu atau diskrit mempunyai suatu pdf yang merupakan salah satu anggota dari family $\{h(z; \theta): \theta \in \Omega\}$. Jika kondisi $E[u(Z)] = 0$ untuk setiap $\theta \in \Omega$ mengakibatkan $u(z) = 0$ kecuali pada himpunan titik-titik yang mempunyai probabilitas nol untuk pdf $h(z; \theta): \theta \in \Omega$, maka family $\{h(z; \theta): \theta \in \Omega\}$ disebut family pdf yang *complete*. (Hogg, & Craig, 2012).

Dari definisi tersebut, untuk membuktikan *mean* sampel \bar{y} memiliki keluarga pdf yang *complete*, perlu diketahui pdf dari *mean* sampel \bar{y} .

Misalkan $\bar{y} = \frac{y_1 + y_2 + \dots + y_n}{n}$ adalah *mean* sampel acak yang diambil dari

populasi berhingga berukuran N yang berdistribusi *uniform* dengan Y_1, Y_2, \dots, Y_N adalah nilai-nilai populasi tersebut. Jika sebanyak n sampel diambil dari populasi tersebut dan nilainya tambahkan, maka distribusi yang dihasilkan tidak selalu *uniform* karena tidak selalu semua nilai memiliki probabilitas yang sama. Oleh karena itu, untuk mencari pdf dari *mean* sampel akan digunakan teorema berikut.

Teorema 2.3 (Teorema Limit Pusat) Misalkan y_1, y_2, \dots, y_n menyatakan suatu sampel random dari suatu distribusi dengan mean μ dan variansi positif σ^2 . Maka variabel acak

$$\begin{aligned} Z_n &= \frac{(\sum_{i=1}^n y_i - n\mu)}{\sqrt{n}\sigma} \\ &= \frac{(\bar{y} - \mu)}{\frac{\sigma}{\sqrt{n}}} \end{aligned}$$

mempunyai limit distribusi normal standar. (Hogg, & Craig, 2012)

BUKTI. Asumsikan bahwa MGF yaitu $M(t) = E(e^{tX})$, $-h < t < h$. Fungsi $m(t) = E[e^{t(X-\mu)}] = E[e^{(tX-t\mu)}] = e^{-t\mu}M(t)$ juga untuk $-h < t < h$. Karena $m(t)$ adalah MGF untuk $X - \mu$, maka $m(0) = 1, m'(0) = E[X - \mu] = 0$, dan $m''(0) = E(x - \mu)^2 = \sigma^2$. Dengan formula Taylor, terdapat suatu bilangan ξ antara 0 dan t sedemikian sehingga

$$\begin{aligned} m(t) &= m(0) + m'(0)t + m''(\xi)t^2 \\ &= 1 + \frac{m''(\xi)t^2}{2}, \end{aligned}$$

Jika $\sigma^2 t^2 / 2$ ditambahkan dan dikurangi maka diperoleh

$$m(t) = 1 + \frac{\sigma^2 t^2}{2} + \frac{[m''(\xi) - \sigma^2]t^2}{2}$$

Berikutnya perhatikan $M(t; n)$ dimana

$$\begin{aligned} M(t; n) &= E[e^{tY_n}] \\ &= E[\exp(t \frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}})] \\ &= E[\exp(t \frac{X_1 - \mu}{\sigma\sqrt{n}}) \exp(t \frac{X_2 - \mu}{\sigma\sqrt{n}}) \dots \exp(t \frac{X_n - \mu}{\sigma\sqrt{n}})] \\ &= E[\exp(t \frac{X_1 - \mu}{\sigma\sqrt{n}})] \dots E[\exp(t \frac{X_n - \mu}{\sigma\sqrt{n}})] \\ &= \{E[\exp(t \frac{X - \mu}{\sigma\sqrt{n}})]\}^n \\ &= [m(\frac{t}{\sigma\sqrt{n}})]^n, -h < \frac{t}{\sigma\sqrt{n}} < h. \end{aligned}$$

Dalam persamaan (2.2), ganti t dengan $\frac{t}{\sigma\sqrt{n}}$ untuk mendapatkan

$$m(\frac{t}{\sigma\sqrt{n}}) = 1 + \frac{t^2}{2n} + \frac{[m''(\xi) - \sigma^2]t^2}{2n\sigma^2}$$

dimana sekarang ξ antara 0 dan $\frac{t}{\sigma\sqrt{n}}$ dengan $-h\sigma\sqrt{n} < t < h\sigma\sqrt{n}$. Dengan demikian $M(t; n)$ dapat ditulis sebagai berikut

$$M(t; n) = \{1 + \frac{t^2}{2n} + \frac{[m''(\xi) - \sigma^2]t^2}{2n\sigma^2}\}^n$$

karena $m''(t)$ kontinu pada $t = 0$ dan karena $\xi \rightarrow 0$ untuk $n \rightarrow \infty$, maka $\lim_{n \rightarrow \infty} [m''(\xi) - \sigma^2] = 0$. Sehingga $\lim_{n \rightarrow \infty} [1 + \frac{t^2}{2n}]^n = e^{\frac{t^2}{2}}$ untuk setiap nilai dari t . Hal ini membuktikan bahwa variabel random $Y_n = \sqrt{n}(\bar{X}_n - \mu)/\sigma$ konvergen dalam distribusi ke distribusi normal standar. (Terbukti)

Interpretasi dari teorema di atas mengatakan bahwa, jika n bilangan bulat positif yang besar maka variabel acak \bar{y} mempunyai pendekatan distribusi normal dengan *mean* μ dan variansi σ^2/n . Oleh karena itu, dengan mengasumsikan nilai n besar, maka berdasarkan teorema limit pusat, *mean* sampel \bar{y} dapat diasumsikan berdistribusi normal dengan *mean* μ dan variansi σ^2/n yang memiliki pdf

$$f(\bar{y}, \theta) = \frac{1}{\sqrt{2\pi\sigma^2/n}} \exp\left[-\frac{(\bar{y} - \mu)^2}{2\sigma^2/n}\right], \bar{y} = \bar{y}_1, \bar{y}_2, \dots$$

Misalkan terdapat fungsi sembarang $g(\bar{y})$. Keluarga pdf *mean* sampel \bar{y} dikatakan *complete* jika untuk kondisi $E[g(\bar{y})] = 0$ untuk setiap $\theta \in \Omega$ mengakibatkan $g(\bar{y}) = 0$ kecuali pada himpunan titik-titik yang mempunyai probabilitas nol untuk pdf $f(\bar{y}, \theta): \theta \in \Omega$. Akan dicari nilai dari $E[g(\bar{y})]$ yaitu:

$$\begin{aligned} E[g(\bar{y})] &= \sum_{\bar{y}=\bar{y}_1}^{\infty} g(\bar{y})f(\bar{y}, \theta) \\ &= \sum_{\bar{y}=0}^{\infty} g(\bar{y}) \frac{1}{\sqrt{2\pi\sigma^2/n}} \exp\left[-\frac{(\bar{y} - \mu)^2}{2\sigma^2/n}\right] \\ &= g(\bar{y}_1) \frac{1}{\sqrt{2\pi\sigma^2/n}} \exp\left[-\frac{(\bar{y}_1 - \mu)^2}{2\sigma^2/n}\right] + g(\bar{y}_2) \frac{1}{\sqrt{2\pi\sigma^2/n}} \exp\left[-\frac{(\bar{y}_2 - \mu)^2}{2\sigma^2/n}\right] + \dots \end{aligned}$$

perhatikan jika

$$\frac{1}{\sqrt{2\pi\sigma^2/n}} \exp\left[-\frac{(\bar{y} - \mu)^2}{2\sigma^2/n}\right] \neq 0 \text{ untuk } \bar{y} = \bar{y}_1, \bar{y}_2, \dots$$

Akibatnya, jika $E[g(\bar{y})] = 0$ maka $g(\bar{y}_1) = g(\bar{y}_2) = \dots = 0$ sehingga mengimplikasikan jika $g(\bar{y}) = 0$. Berdasarkan definisi statistik *complete*, terbukti *mean* sampel \bar{y} memiliki keluarga pdf yang *complete*. Telah dibuktikan jika *mean* sampel \bar{y} merupakan penaksir yang tak bias untuk \bar{Y} , *mean* sampel \bar{y} adalah statistik cukup untuk \bar{Y} dan keluarga pdf dari *mean* sampel \bar{y} *complete* sehingga dengan Teorema Lehmann and Scheffe, *mean* sampel \bar{y} merupakan penaksir UMVE.

Definisikan Y sebagai variabel observasi. Misalkan (Y_1, Y_2, \dots, Y_N) adalah nilai-nilai dari populasi berhingga berukuran N dengan distribusi *uniform*. Misalkan pula terdapat sebuah *attribute* φ dimana *attribute* φ memiliki korelasi *point biserial* dengan variabel observasi Y . *Attribute* φ merupakan variabel dikotomi dimana nilainya

$$\varphi_i = \begin{cases} 1, & \text{jika unit ke } i \text{ dari populasi memproses attribute } \varphi, \\ 0, & \text{yang lainnya} \end{cases}$$

Misalkan P adalah proporsi populasi yang memproses *attribute* φ yang didefinisikan sebagai

$$P = \frac{A}{N}$$

Misalkan p adalah proporsi sampel yang memproses *attribute* φ yang didefinisikan sebagai

$$p = \frac{a}{n}$$

Nursel Koyuncu memperkenalkan keluarga penaksir *mean* populasi yang dapat memanfaatkan *auxiliary attribute* dalam *simple random sampling* dengan bentuk:

$$\bar{y}_N = [w_1\bar{y} + w_2(P - p)] \left(\frac{\eta P + \lambda}{\eta p + \lambda} \right)$$

dimana w_1 dan w_2 adalah konstanta atau bobot yang ditentukan, \bar{y} adalah *mean* sampel dari variabel observasi Y sedangkan p dan P adalah proporsi sampel dan populasi yang memproses *attribute* φ . Variabel η dan λ merupakan bilangan real atau parameter dari *attribute* φ . Dalam kasus ini, parameter *attribute* φ yang digunakan adalah koefisien variasi dan koefisien kurtosis. Untuk membentuk penaksir *mean* populasi dari keluarga Koyuncu ini, dilakukan dengan cara mengganti variabel η dan λ menjadi bilangan real atau parameter *attribute* φ . Dari kedua parameter ini, dapat dibentuk empat penaksir Koyuncu *mean* populasi yaitu: Penaksir Koyuncu pertama dibentuk dari $\eta = 1$ dan $\lambda = C_p$

$$\bar{y}_{N_1} = [w_1\bar{y} + w_2(P - p)] \left(\frac{P + C_p}{p + C_p} \right)$$

Penaksir Koyuncu kedua dibentuk dari $\eta = 1$ dan $\lambda = B_2(\varphi)$

$$\bar{y}_{N_2} = [w_1\bar{y} + w_2(P - p)] \left(\frac{P + B_2(\varphi)}{p + B_2(\varphi)} \right)$$

Penaksir Koyuncu ketiga dibentuk dari $\eta = C_p$ dan $\lambda = B_2(\varphi)$

$$\bar{y}_{N_3} = [w_1\bar{y} + w_2(P - p)] \left(\frac{C_p P + B_2(\varphi)}{C_p p + B_2(\varphi)} \right)$$

Penaksir Koyuncu keempat dibentuk dari $\eta = B_2(\varphi)$ dan $\lambda = C_p$

$$\bar{y}_{N_4} = [w_1\bar{y} + w_2(P - p)] \left(\frac{B_2(\varphi)P + C_p}{B_2(\varphi)p + C_p} \right)$$

Keluarga penaksir *mean* populasi \bar{y}_N memiliki *mean square error* yaitu:

$$\begin{aligned} MSE(\bar{y}_N) = & \bar{Y}^2 + w_1^2 \bar{Y}^2 \left\{ 1 + \frac{1-f}{n} (C_Y^2 - 4\tau C_{Y\varphi} + 3\tau^2 C_\varphi^2) \right\} + w_2^2 P^2 \frac{1-f}{n} C_\varphi^2 \\ & + 2w_1 \bar{Y}^2 \left\{ 1 + \frac{1-f}{n} (\tau^2 C_\varphi^2 - \tau C_{Y\varphi}) \right\} - 2w_2 \bar{Y} P \tau \frac{1-f}{n} C_\varphi^2 \\ & - 2w_1 w_2 \bar{Y} P \frac{1-f}{n} (C_{Y\varphi} - 2\tau C_\varphi^2) \end{aligned}$$

dimana

$$\tau = \frac{\eta P}{\eta P + \lambda}$$

dan f merupakan *finite population correction* yang dirumuskan dengan $f = \frac{n}{N}$.

Pemilihan nilai konstanta atau bobot w_1 dan w_2 dilakukan secara tepat agar dapat meminimumkan nilai $MSE(\bar{y}_N)$. Pandang $MSE(\bar{y}_N)$ sebagai fungsi dua variabel yaitu variabel w_1 dan w_2 . $MSE(\bar{y}_N)$ akan mencapai nilai minimum ketika

$$\frac{\partial}{\partial w_1} [MSE(\bar{y}_N)] = 0 \quad \text{dan} \quad \frac{\partial}{\partial w_2} [MSE(\bar{y}_N)] = 0$$

dengan menyelesaikan persamaan tersebut, didapatkan nilai konstanta atau bobot w_1 dan w_2 yang dirumuskan sebagai

$$\begin{aligned} w_1 = & \frac{1 - \left(\frac{1-f}{n}\right)\tau^2 C_\varphi^2}{1 + \left(\frac{1-f}{n}\right)(C_Y^2 - \rho_{Y\varphi} C_Y^2 - \tau^2 C_\varphi^2)} \\ w_2 = & \frac{\bar{Y}}{P} \left(\tau + \frac{\left(1 - \left(\frac{1-f}{n}\right)\tau^2 C_\varphi^2\right)(C_{Y\varphi} - 2\tau C_\varphi^2)}{C_\varphi^2 + \left(\frac{1-f}{n}\right)(C_\varphi^2 C_Y^2 - C_{Y\varphi}^2 - \tau^2 C_\varphi^4)} \right) \end{aligned}$$

Setelah itu, nilai w_1 dan w_2 akan disubstitusikan ke dalam *mean square error* \bar{y}_N sehingga didapatkan *mean square error minimum* \bar{y}_N yang dirumuskan sebagai

$$MSE(\bar{y}_N)_{min} = \frac{\left(\frac{1-f}{n} C_Y^2 \bar{Y}^2\right)(1 - \rho_{Y\varphi}^2)}{1 + \frac{\left(\frac{1-f}{n} C_Y^2\right)(1 - \rho_{Y\varphi}^2)}{1 - \left(\frac{1-f}{n}\right)\tau^2 C_\varphi^2}}$$

Untuk mengetahui apakah penaksir yang dibentuk dari keluarga penaksir Koyuncu ini memiliki efisiensi yang lebih baik dibanding dengan *mean* sampel, akan dilakukan perbandingan *mean square error* keluarga penaksir Koyuncu dengan *mean* sampel. Sebelumnya telah dibuktikan jika *mean square error mean* sampel dirumuskan sebagai

$$MSE(\bar{y}) = \frac{1-f}{n} C_Y^2 \bar{Y}^2$$

Keluarga penaksir Koyuncu dikatakan lebih efisien dibandingkan *mean* sampel jika

$$MSE(\bar{y}) > MSE(\bar{y}_N)$$

maka

$$\begin{aligned} MSE(\bar{y}) - MSE(\bar{y}_N) &> 0 \\ \frac{1-f}{n} C_Y^2 \bar{Y}^2 - \frac{\left(\frac{1-f}{n} C_Y^2 \bar{Y}^2\right)(1 - \rho_{pb}^2)}{1 + \frac{\left(\frac{1-f}{n} C_Y^2\right)(1 - \rho_{pb}^2)}{\bar{Y}^2(1 - \left(\frac{1-f}{n}\right)\tau^2 C_\varphi^2)}} &> 0 \end{aligned}$$

persamaan diatas akan berlaku jika

$$1 - \left(\frac{1-f}{n}\right)\tau^2 C_\varphi^2 \geq 0$$

Jika persamaan tersebut terpenuhi, maka $MSE(\bar{y}) > MSE(\bar{y}_N)$ sehingga hal ini dapat membuktikan jika keluarga penaksir Koyuncu memiliki efisiensi yang lebih baik dibandingkan *mean* sampel \bar{y} . Hal ini juga mengimplikasikan jika penaksir yang dibentuk dari keluarga penaksir tersebut memiliki efisiensi yang lebih baik dibandingkan *mean* sampel \bar{y} .

Untuk mendukung teori yang telah dijelaskan sebelumnya, akan dilakukan contoh perhitungan numerik dengan menggunakan data asli. Data yang digunakan adalah data tentang rata-rata harga apel per pound di daerah-daerah yang ada di negara Amerika Serikat dimana harga apel tahun 2012 adalah karakter yang akan diukur dan harga apel tahun 2011 dijadikan sebagai karakter atau variabel tambahan. Definisikan *attribute* φ sebagai

Y : Harga apel per pound tahun 2012

φ : Harga apel per pound lebih dari 0.378310345 tahun 2011

Dari pendefinisian *attribute* φ , didapatkan statistika deskriptif dari data sebagai berikut:

$N = 29$	$n = 7$
$\bar{Y} = 0.452$	$S_Y^2 = 0.035627786$
$C_Y = 0.417595664$	$\rho_{pb} = 0.843360031$
$P = 0.379310345$	$S_\varphi^2 = 0.243842365$
$C_\varphi = 1.30184684$	$B_2(\varphi) = 1.247476162$

Dengan menggunakan nilai-nilai di atas, akan dilakukan perhitungan *mean square error* penaksir *mean* populasi dan dihasilkan

Tabel 3: Nilai *mean square error* penaksir *mean* populasi

Penaksir	<i>Mean square error</i>
\bar{y}_{N_1}	0.001113626988
\bar{y}_{N_2}	0.001113626185
\bar{y}_{N_3}	0.001113620100
\bar{y}_{N_4}	0.001113622293
\bar{y}	0.003861139333

Dari tabel perhitungan *mean square error* diatas terbukti jika penaksir *mean* populasi yang dibentuk dari keluarga Koyuncu memiliki efisiensi yang lebih baik dibandingkan dengan *mean* sampel. Penaksir *mean* populasi yang memanfaatkan koefisien variasi dan koefisien kurtosis yaitu \bar{y}_{N_3} dan \bar{y}_{N_4} memiliki efisiensi yang lebih baik dibandingkan penaksir *mean* populasi yang lain.

3. Kesimpulan

Dari pembahasan di atas, maka dapat disimpulkan beberapa hal, yaitu

1. Pemanfaatan *auxiliary attribute* dalam prosedur penaksiran dapat meningkatkan presisi penaksir.
2. Penaksir *mean* populasi yang dibentuk dari keluarga penaksir Koyuncu memiliki efisiensi yang lebih baik dibandingkan *mean* sampel dalam *simple random sampling*.
3. Penaksir *mean* populasi dari keluarga penaksir Koyuncu yang lebih banyak melibatkan parameter dari *attribute* φ memiliki efisiensi lebih baik dibanding yang lain.

Referensi

- [1] Koyuncu, Nursel. (2012). Efficient estimators of population mean using auxiliary attributes. *Applied Mathematics and Computation* 218,pp 10900-10905.
- [2] Ahmad, Zahoor, Shahbaz, M.Q. dan Hanif, Muhammad (2013). *Two Phase Sampling*. Newcastle : Cambridge Scholars Publishing..
- [3] Brown, J.D. (2001). Point-biserial Correlation Coefficients. *Shiken: JLT Testing & Evaluation SIG Newsletter*, 5(3),pp 13-17.
- [4] Hogg,R.V., dan A.T.Craig.(2012). *Introduction to Mathematical Statistics* 5th ed. United States of America : Prentice Hall, Inc.
- [5] Sukhatme,P.V.(1954). *Sampling Theory of Surveys with Applications*. New Delhi, India : The Indian Society of Agricultural Statistics.

PENGGUNAAN METODE *JACKKNIFE EMPIRICAL LIKELIHOOD* UNTUK MEMBENTUK INTERVAL KEPERCAYAAN DENGAN PARAMETER *MEAN ABSOLUTE DEVIATION*

REVITA WIDASARI¹, IDA FITHRIANI², DAN SASKYA MARY SOEMARTOJO³

1. Departemen Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Indonesia, Depok, 16424, Indonesia, revita.widasari@sci.ui.ac.id
2. Departemen Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Indonesia, Depok, 16424, Indonesia, ida.fithriani@gmail.com
3. Departemen Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Indonesia, Depok, 16424, Indonesia, saskya@sci.ui.ac.id

Abstrak. Metode *jackknife empirical likelihood* adalah suatu metode statistika untuk membentuk interval kepercayaan bagi suatu parameter. Metode *jackknife empirical likelihood* merupakan pengembangan dari metode *empirical likelihood* ketika parameternya berbentuk fungsi nonlinier. Metode *empirical likelihood* sendiri merupakan metode statistika yang digunakan untuk membentuk interval kepercayaan untuk suatu parameter berdasarkan data yang diperoleh dari suatu pengamatan. Pada makalah ini, akan dibahas mengenai pembentukan interval kepercayaan untuk *mean absolute deviation* dari suatu populasi menggunakan metode *jackknife empirical likelihood*. *Mean absolute deviation* itu sendiri merupakan salah satu dari ukuran penyebaran data, yaitu rata-rata harga mutlak dari simpangan data terhadap meannya. Distribusi asimtotik dari fungsi *jackknife empirical likelihood ratio* adalah suatu distribusi chi-kuadrat dengan derajat bebas satu. Fungsi *jackknife empirical likelihood ratio* ini akan digunakan untuk membentuk interval kepercayaan untuk *mean absolute deviation*. Untuk ilustrasi akan digunakan himpunan data asli sebagai contoh..

Kata kunci: interval kepercayaan, *mean absolute deviation*, metode empirical likelihood, metode *Jackknife*, *Jackknife empirical likelihood*.

1. Pendahuluan

Dalam suatu penelitian statistika, diperlukan interpretasi data untuk menarik kesimpulan mengenai karakteristik dari populasi. Populasi adalah himpunan keseluruhan data yang diteliti dengan sembarang nilai karakteristik yang menjadi ciri dari suatu populasi disebut parameter. Untuk mengetahui karakteristik dari suatu populasi dilakukan penaksiran parameter. Dalam makalah ini akan dibahas parameter *mean absolute deviation*. Menurut Zhao [10], *mean absolute deviation* didefinisikan sebagai rata-rata harga mutlak simpangan setiap pengamatan terhadap *mean*.

Pada tahun 1974, Gastwirth [3] telah memeriksa bahwa penaksir dari *mean absolute deviation* berdistribusi normal secara asimtotik. Berdasarkan aproksimasi normal tersebut diperoleh interval kepercayaan untuk *mean absolute deviation*. Menurut Zhao [10], untuk data yang tidak diketahui distribusinya, interval kepercayaan untuk *mean absolute deviation* yang diperoleh berdasarkan aproksimasi normal tidak dapat diharapkan akan memberikan hasil yang akurat. Dalam hal tersebut, dikembangkan suatu teknik inferensi yang tidak memerlukan asumsi mengenai distribusi suatu data, yaitu metode nonparametrik.

Salah satu metode nonparametrik adalah metode *empirical likelihood*. Metode *empirical likelihood* dapat mengatasi data dengan baik jika asumsi independen pada data pengamatannya terpenuhi. Menurut Owen [9], prinsip kerja metode *empirical likelihood* adalah memaksimumkan fungsi *nonparametric likelihood* terhadap beberapa kendala. Kendala yang ada dapat berbentuk persamaan linier atau nonlinier. Untuk kendala yang linier dapat diselesaikan dengan metode pengali Lagrange, namun untuk kendala yang nonlinier menjadi sulit diselesaikan dengan metode pengali Lagrange.

Untuk mengatasi hal tersebut, Jing *et al.*[7] pada tahun 2009 mengajukan suatu metode baru yaitu metode *jackknife empirical likelihood* (JEL). Menurut Jing *et al.*, gagasan utama dari metode JEL adalah mengubah statistik yang diamati menjadi suatu mean dari suatu sampel berdasarkan *jackknife pseudo-values*. Dalam penulisan ini akan dibahas teori mengenai metode *jackknife empirical likelihood* dalam membentuk interval kepercayaan untuk *mean absolute deviation*.

2. Hasil – Hasil Utama

2.1. Large Sample Theory untuk Mean Absolute Deviation

Secara matematis, *mean absolute deviation* (MAD) didefinisikan sebagai berikut. Misalkan X menyatakan variabel random dari suatu fungsi distribusi $F(x)$ dengan mean dari X dinyatakan sebagai $E(X) = \mu$, variansi dari X sebagai $E(X - \mu)^2 = \sigma^2$, dan *mean absolute deviation* dari X sebagai berikut

$$E|X - E(X)| = \theta. \quad (10)$$

Untuk mengetahui karakteristik dari θ akan ditaksir θ dengan cara mengambil suatu sampel random dari distribusi tersebut. MAD sampel merupakan penaksir empiris untuk θ , dan didefinisikan sebagai berikut,

$$\sum_{i=1}^n |X_i - \bar{X}|/n = \hat{\theta}. \quad (10)$$

Berdasarkan Gastwirth [3], jika X_1, X_2, \dots, X_n menyatakan variabel-variabel random iid dengan fungsi distribusi $F(x)$ yang memiliki mean μ dan variansi σ^2 , serta mempunyai fungsi probabilitas $f(x)$ kontinu di lingkungan dari μ , maka statistik $\hat{\theta}$ berdistribusi normal secara asimtotik, yaitu $\sqrt{n}(\hat{\theta} - \theta) \rightarrow N(0, v^2)$, dengan

$$v^2 = 4 * \{p^2 \int_{\mu}^{\infty} (x - \mu)^2 dF(x) + (1 - p)^2 \int_{-\infty}^{\mu} (x - \mu)^2 dF(x) - \frac{\theta^2}{4}\}. \quad (3)$$

Dengan $p = F(\mu)$.

2.2. Metode Empirical Likelihood

2.2.1 Nonparametric Maximum Likelihood

Definisi 2.1. Misal $x_1, \dots, x_n \in \mathbb{R}$. Fungsi distribusi kumulatif empiris atau

empirical cumulative distribution function (ECDF) dari X_1, \dots, X_n adalah

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I_{X_i \leq x}, \text{ untuk } -\infty < x < \infty,$$

dengan $I_{X_i \leq x} = \begin{cases} 1, & \text{jika } X_i \leq x \\ 0, & \text{jika yang lain} \end{cases}$

(Owen [9])

Definisi 2.2. Misalkan X_1, \dots, X_n menyatakan variabel-variabel random yang diasumsikan independen dengan CDF awal F_0 , *nonparametric likelihood* dari CDF F adalah

$$L(F) = \prod_{i=1}^n (F(X_i) - F(X_i -)).$$

(Owen [9])

Teorema 2.1. Misal X_1, \dots, X_n menyatakan variabel-variabel random independen dengan CDF masing-masing F_0 . Misal F_n adalah ECDF dan F adalah sembarang CDF. Jika $F \neq F_n$, maka $L(F) < L(F_n)$.

(Owen [9])

BUKTI. Misalkan $z_1 < z_2 < \dots < z_m$ adalah nilai-nilai observasi yang berbeda-beda pada $\{X_1, X_2, \dots, X_n\}$. Misalkan $n_j \geq 1$ adalah banyaknya nilai-nilai dari X_i yang sama dengan z_j . Misalkan $p_j = F(z_j) - F(z_j -)$ menyatakan probabilitas mendapatkan nilai observasi z_j . Misalkan $\hat{p}_j = n_j/n$ adalah probabilitas empiris munculnya z_j . Jika ada suatu $p_j = 0$ untuk $j = 1, 2, \dots, m$, maka berdasarkan definisi *nonparametric likelihood* didapatkan $L(F) = \prod_{j=1}^m p_j = 0 < L(F_n)$. Untuk itu seharusnya semua nilai $p_j > 0$, dan paling sedikit untuk satu nilai j , $p_j \neq \hat{p}_j$, dengan kata lain probabilitas yang berkaitan dengan F tidak persis sama dengan probabilitas yang berkaitan dengan F_n . Berdasarkan penjelasan di atas, diperoleh:

$$\log \left(\frac{L(F)}{L(F_n)} \right) = \sum_{j=1}^m n_j \log \left(\frac{p_j}{\hat{p}_j} \right) = n \sum_{j=1}^m \hat{p}_j \log \left(\frac{p_j}{\hat{p}_j} \right).$$

Diketahui bahwa $\log(x) < x - 1$ untuk $x > 0$ dan $x \neq 1$. Dengan demikian,

$$\begin{aligned} \log \left(\frac{L(F)}{L(F_n)} \right) &< n \sum_{j=1}^m \hat{p}_j \left(\frac{p_j}{\hat{p}_j} - 1 \right) \leq 0 \\ \frac{L(F)}{L(F_n)} &\leq 1 \\ L(F) &\leq L(F_n). \end{aligned}$$

Dari teorema diketahui $F \neq F_n$, maka tidak mungkin $L(F) = L(F_n)$. Dengan demikian haruslah $L(F) < L(F_n)$. Terbukti bahwa $L(F) < L(F_n)$. Dengan demikian ECDF F_n merupakan NPMLE dari F , dapat ditulis $\hat{F} = F_n$. \square

Teorema 2.2. (i) Misalkan terdapat fungsi likelihood $L(\eta) = f(X_1, \dots, X_n | \eta)$, η adalah suatu parameter. Misal suatu statistik $\hat{\eta}$ adalah MLE dari η . Jika $g(\eta)$ adalah fungsi satu-satu, maka suatu fungsi dari suatu statistik tersebut yaitu $g(\hat{\eta}) = \hat{\theta}$ adalah MLE untuk suatu fungsi dari parameter η yaitu $g(\eta) = \theta$. (ii) Untuk $g(\eta)$ bukan fungsi satu-satu, maka $g(\hat{\eta}) = \hat{\theta}$ tetap merupakan MLE untuk $g(\eta) = \theta$.

(Dudewicz [1])

BUKTI. (i) Untuk $g(\eta)$ adalah fungsi satu-satu, maka setiap η bersesuaian dengan satu nilai dari $g(\eta) = \theta$. Misalkan nilai dari $g(\eta)$ yang memaksimumkan $L(\eta)$ adalah $\widehat{g(\eta)}$. Perhatikan bahwa $L(\eta) = L(g^{-1}(\theta))$. Diketahui bahwa $\hat{\eta}$ adalah MLE dari η , maka $\hat{\eta}$ memaksimumkan fungsi likelihood $L(\eta)$, sehingga $\hat{\eta}$ juga memaksimumkan $L(g^{-1}(g(\hat{\eta}))) = L(g^{-1}(\theta))$. Dengan demikian diperoleh $\hat{\eta} = g^{-1}(\widehat{g(\eta)}) = g^{-1}(\hat{\theta})$, dengan kata lain $g(\hat{\eta}) = \widehat{g(\eta)} = \hat{\theta}$. **(ii)** Misalkan $g(\eta)$ bukan fungsi satu-satu. Untuk setiap η , image dari g adalah $g(\theta)$. Untuk setiap θ pada *range* dari g , didefinisikan *preimage* dari g , sebagai berikut, $g^{-1}(\theta) = \{\eta: g(\eta) = \theta\}$. Maksimum muncul di $\hat{\eta}$ dan domain dari g mencakup $\hat{\eta}$. Dengan demikian, $\hat{\eta}$ berada pada *preimage-preimage* tersebut. Sehingga untuk memaksimumkan $L(\theta)$, pilih $\hat{\theta}$ sedemikian sehingga $g^{-1}(\hat{\theta})$ adalah *preimage* yang mengandung $\hat{\eta}$. Dengan kata lain, $g^{-1}(\hat{\theta}) = \{\hat{\eta}: g(\hat{\eta}) = \hat{\theta}\}$. Dengan demikian terbukti bahwa $\hat{\theta} = g(\hat{\eta})$. \square

Pada *nonparametric likelihood*, telah diketahui bahwa F_n adalah NPMLE dari F . Misalkan $\theta = T(F)$ adalah suatu parameter dari sampel random dengan fungsi distribusi F , di mana T adalah fungsi dari distribusi-distribusi. Misal parameter sebenarnya $\theta_0 = T(F_0)$. Menurut Owen [9], berdasarkan analoginya, maka NPMLE dari θ adalah $\hat{\theta} = T(F_n)$.

2.2.2 Nonparametric Likelihood Ratio

Untuk suatu fungsi distribusi F , didefinisikan fungsi *nonparametric likelihood ratio* sebagai berikut,

$$R(F) = L(F)/L(F_n). \quad (9)$$

Pada *empirical likelihood*, misalkan x_1, x_2, \dots, x_n adalah observasi-observasi yang independen dari suatu distribusi F dan $\theta = T(F)$ adalah parameter yang diamati. Menurut Efron [2], fungsi profil *likelihood* dan fungsi profil *likelihood ratio* untuk θ didefinisikan sebagai berikut

$$L_{pro}(\theta) = \sup_{F: T(F)=\theta} L(F) \quad (2)$$

$$\mathcal{R}_{pro}(\theta) = \sup\{R(F) | T(F) = \theta\}. \quad (9)$$

Berdasarkan Owen [9], *empirical likelihood* mengasumsikan tidak terdapat nilai yang sama pada observasi-observasi X_i . Misalkan $w_i > 0$ menyatakan probabilitas dalam mendapatkan nilai observasi dari X_i yang berbeda-beda, maka diperoleh bentuk fungsi *empirical likelihood ratio* dan fungsi *empirical log-likelihood ratio* sebagai berikut,

$$R(F) = \prod_{i=1}^n nw_i \quad (9)$$

$$\log R(F) = \sum_{i=1}^n \log(nw_i).$$

2.3. Metode Jackknife Empirical Likelihood untuk Mean Absolute Deviation

Misalkan didefinisikan $F_q(x) = \sum_{i=1}^n q_i I\{X_i \leq x\}$, dengan q_i menyatakan probabilitas mendapatkan nilai-nilai observasi x_i . Telah diketahui EL melibatkan pemaksimuman fungsi *nonparametric likelihood* terhadap beberapa kendala, maka fungsi *empirical likelihood* untuk MAD (θ) didefinisikan sebagai berikut,

$$L(\theta) = \max \left\{ \prod_{i=1}^n q_i : \sum_{i=1}^n q_i = 1, \theta(F_q) = \theta \right\}. \quad (7)$$

Selanjutnya diperoleh bentuk fungsi *empirical likelihood ratio* untuk MAD (θ) adalah sebagai berikut,

$$R(\theta) = \frac{L(\theta)}{n^{-n}} = \max \left\{ \prod_{i=1}^n (nq_i) : \sum_{i=1}^n q_i = 1, \theta(F_q) = \theta \right\}. \quad (7)$$

Diketahui bahwa MAD merupakan fungsi nonlinier dan metode EL sulit diaplikasikan untuk fungsi nonlinier. Untuk mengatasi kesulitan tersebut, Jing *et al.* [7] mengajukan suatu metode baru yang disebut sebagai metode *jackknife empirical likelihood* (JEL). Metode ini mengkombinasikan metode *jackknife* dengan metode *empirical likelihood* yang telah dikenal sebelumnya.

2.3.1. Metode Jackknife

Metode *jackknife* pertama kali diperkenalkan oleh Quenouille pada tahun 1949 sebagai metode untuk mereduksi bias dari suatu penaksir. Prinsip kerja metode *jackknife* menurut Jiang [6] adalah sebagai berikut.

1. Misal $\hat{\theta} = \hat{\theta}(x_1, x_2, \dots, x_n)$ menyatakan suatu penaksir untuk parameter θ . Untuk mendapatkan $\hat{\theta}$, lakukan penaksiran parameter dengan menggunakan sampel awal (X_1, \dots, X_n) .
2. Kemudian menghapus satu observasi dari sampel awal yang berukuran n , sampel baru ini disebut sebagai *jackknife samples*, yaitu $\mathbf{X}_{(i)} = (X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)$.
3. Menaksir parameter dengan menggunakan *jackknife sample* tersebut. Penaksir dari *jackknife sample* disebut dengan *partial estimate* dan dinotasikan dengan $\hat{\theta}_{n-1}^{(-i)} = \hat{\theta}(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$.

Berdasarkan Govindarajulu [4], John Tukey mendefinisikan *jackknife pseudo-values* untuk memperoleh penaksir *jackknife*. *Jackknife pseudo-values* didefinisikan sebagai berikut,

$$\hat{V}_i = n\hat{\theta} - (n-1)\hat{\theta}_{n-1}^{(-i)}, \text{ untuk } 1 \leq i \leq n. \quad (4)$$

Pseudo-values tersebut akan dianggap sebagai observasi baru untuk menaksir parameter. Tukey mendefinisikan penaksir *jackknife* untuk θ sebagai rata-rata dari *pseudo-values*, yaitu dapat ditulis sebagai berikut,

$$\widehat{\theta}_{n,jack} := \frac{1}{n} \sum_{i=1}^n \hat{V}_i. \quad (4)$$

Berdasarkan Govindarajulu [4], Tukey menganggap bahwa *pseudo-values* \hat{V}_i dapat diperlakukan seolah-olah saling independen. Berdasarkan Jing *et al* [7], pada tahun 1984, Shi telah membuktikan bahwa *pseudo-values* independen secara asimtotik.

2.3.2. Metode JEL

Metode EL dapat diaplikasikan untuk mean dengan variabel-variabel random yang independen, maka hal tersebut menjadi alasan untuk mengaplikasikan metode EL ke mean dari *pseudo-values* \hat{V}_i . Berikut ini akan dijelaskan mengenai prosedur metode JEL. Misal $\mathbf{p} = (p_1, \dots, p_n)$ adalah suatu vektor probabilitas sedemikian sehingga $\sum_{i=1}^n p_i = 1$ dan $p_i > 0$. Misal $G_{\mathbf{p}}(x) = \sum_{i=1}^n p_i I\{\hat{V}_i \leq x\}$. Perhatikan *mean functional* $\theta(G_{\mathbf{p}}) = \sum_{i=1}^n p_i \hat{V}_i$, dan misal $\theta_{\mathbf{p}} = \sum_{i=1}^n p_i E(\hat{V}_i)$. Fungsi *empirical likelihood* untuk θ diberikan sebagai berikut,

$$L(\theta) = \max \left\{ \prod_{i=1}^n p_i : \sum_{i=1}^n p_i = 1, \theta(G_{\mathbf{p}}) = \theta_{\mathbf{p}} \right\}. \quad (7)$$

Kemudian didefinisikan *jackknife empirical likelihood ratio* dan *jackknife empirical log-likelihood ratio* untuk θ sebagai berikut,

$$R(\theta) = \max \left\{ \prod_{i=1}^n (np_i) : \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i \hat{V}_i = \theta_{\mathbf{p}} \right\} \quad (7)$$

$$\log R(\theta) = \max \left\{ \sum_{i=1}^n \log(np_i) : \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i \hat{V}_i = \theta_{\mathbf{p}} \right\}. \quad (7)$$

Selanjutnya, dari bentuk fungsi *jackknife empirical log-likelihood ratio* tersebut, akan dicari bentuk p_i yang dapat menghasilkan nilai maksimum dari fungsi *jackknife empirical log-likelihood ratio* dengan kendala-kendala yang ada. Kendala-kendala pada bentuk persamaan di atas adalah linier, maka persamaan tersebut dapat diselesaikan menggunakan metode pengali Lagrange.

Berdasarkan Jing *et al* [7], persamaan di atas masih perlu dicek apakah hal berikut ini terpenuhi, yaitu, $-2 \log R(\theta) \xrightarrow{d} \chi_1^2$. Dari persamaan tersebut dapat dibentuk interval kepercayaan pada tingkat $(1 - \alpha)$ untuk θ sebagai berikut, $\mathfrak{R}_c = \{\theta: -2 \log R(\theta) \leq c\}$, dengan c dipilih memenuhi $P(\chi_1^2 \leq c) = 1 - \alpha$.

2.3.3. Metode JEL untuk Mean Absolute Deviation (MAD)

Dengan menggunakan metode *jackknife*, diperoleh *partial estimate* untuk θ sebagai berikut,

$$\hat{\theta}_{n-1}^{(-i)} = \hat{\theta}(X_1, X_2, \dots, X_{i-1}, X_{i+1}, \dots, X_n) = \frac{1}{n-1} \sum_{j \neq i}^n |X_j - \bar{X}_{(-i)}| \quad (10)$$

dengan $\bar{X}_{(-i)} = \frac{1}{n-1} \sum_{j \neq i}^n X_j$. Selanjutnya didefinisikan *jackknife pseudo-values*-nya yaitu,

$$\hat{V}_i = n\hat{\theta} - (n-1)\hat{\theta}_{n-1}^{(-i)}. \quad (10)$$

Dapat dibuktikan bahwa $E(\hat{V}_i) = \theta$, maka dapat dikatakan bahwa $E(\hat{V}_i)$ merupakan parameter yang akan diamati, sehingga dapat ditulis $\theta_{\mathbf{p}} = \sum_{i=1}^n p_i E(\hat{V}_i) = \theta$. Selanjutnya diperoleh penaksir *jackknife* untuk θ sebagai berikut,

$$\widehat{\theta}_{n, \text{jack}} = \frac{1}{n} \sum_{i=1}^n \hat{V}_i. \quad (10)$$

Dengan alasan bahwa metode EL dapat diaplikasikan untuk mean dari \hat{V}_i , maka dapat didefinisikan fungsi *empirical likelihood* untuk θ sebagai berikut,

$$L(\theta) = \max \left\{ \prod_{i=1}^n p_i : \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i \hat{V}_i = \theta \right\}.$$

Berdasarkan metode JEL, diperoleh fungsi *jackknife empirical log-likelihood ratio* untuk θ adalah $\log R(\theta) = -\sum_{i=1}^n \log\{1 + \lambda(\hat{V}_i - \theta)\}$, dengan λ memenuhi

$$g(\lambda) = \frac{1}{n} \sum_{i=1}^n \frac{\hat{V}_i - \theta}{1 + \lambda(\hat{V}_i - \theta)} = 0. \quad (10)$$

Selanjutnya akan diperoleh λ sebagai berikut,

$$\lambda = \frac{(\widehat{\theta}_{n,jack} - \theta)}{\sigma_{n,jack}^2},$$

dengan $\sigma_{n,jack}^2 = 1/n \sum_{i=1}^n (\widehat{V}_i - \theta)^2$ merupakan *jackknife variance estimate*.

Dengan menggunakan aproksimasi Taylor, maka diperoleh,

$$-2 \log R(\theta) = \frac{n (\widehat{\theta}_{n,jack} - \theta)^2}{\sigma_{n,jack}^2}.$$

Selanjutnya notasikan $l(\theta) = -2 \log R(\theta)$. Berdasarkan Zhao [10], teorema 2.3. berikut ini akan digunakan untuk membentuk interval kepercayaan untuk θ . Sebelum itu diberikan kondisi-kondisi regularitas dan lemma-lemma berdasarkan Zhao [10] sebagai berikut:

R.1: pdf $f(x)$ kontinu di lingkungan sekitar μ ;

R.2: terdapat mean μ dan variansi σ^2 dari X_i .

Lemma 2.1. Berdasarkan kondisi-kondisi regularitas di atas, diperoleh

$$n^{\frac{1}{2}} (\widehat{\theta}_{n,jack} - \theta) \rightarrow N(0, v^2).$$

(Zhao [10])

Lemma 2.2. Berdasarkan kondisi-kondisi regularitas, diperoleh $\sigma_{n,jack}^2 \xrightarrow{P} v^2$.

(Zhao [10])

Teorema 2.3. Asumsikan bahwa kondisi regularitas terpenuhi. $l(\theta)$ konvergen dalam distribusi ke χ^2 , di mana χ^2 adalah variabel random chi-square dengan derajat bebas 1.

(Zhao [10])

BUKTI. Berdasarkan lemma 2.1 diperoleh bahwa $\sqrt{n} \{\widehat{\theta}_{n,jack} - \theta\} \xrightarrow{D} N(0, v^2)$,

maka berdasarkan *central limit theorem* $\frac{\sqrt{n} \{\widehat{\theta}_{n,jack} - \theta\}}{v} \xrightarrow{D} N(0,1)$. Berdasarkan

lemma 2.2 diperoleh bahwa $\sigma_{n,jack}^2 \xrightarrow{P} v^2$, sehingga $\sigma_{n,jack} \xrightarrow{P} v$, dan $\frac{\sigma_{n,jack}}{v} \xrightarrow{P} 1$.

Dengan demikian berdasarkan *slutsky's theorem* dapat diperoleh bahwa,

$$\frac{1}{\sigma_{n,jack}/v} \frac{\sqrt{n} \{\widehat{\theta}_{n,jack} - \theta\}}{v} = \frac{\sqrt{n} \{\widehat{\theta}_{n,jack} - \theta\}}{\sigma_{n,jack}} \xrightarrow{D} N(0,1).$$

Misalkan $W_n = \frac{\sqrt{n} \{\widehat{\theta}_{n,jack} - \theta\}}{\sigma_{n,jack}}$, maka $W_n \xrightarrow{D} W$, dengan W berdistribusi $N(0,1)$.

Misalkan $V = W^2$, cdf dari V adalah

$$G(v) = \int_0^v \frac{1}{\sqrt{2\pi}\sqrt{y}} e^{-y/2} dy, \quad v \geq 0.$$

dan cdf dari V untuk $v < 0$ adalah $G(v) = 0$. Selanjutnya pdf dari V dapat diperoleh dengan mencari turunan dari cdf $G(v)$, yaitu $g(v) = G'(v)$ sebagai berikut,

$$g(v) = \begin{cases} \frac{1}{\sqrt{\pi} 2^{1/2}} v^{1/2-1} e^{-v/2}, & 0 < v < \infty, \\ 0 & , \text{lainnya.} \end{cases}$$

Diketahui bahwa $g(v)$ adalah pdf dan $\int_0^\infty g(v) dv = 1$, dan juga telah diketahui

bahwa $\sqrt{\pi} = \Gamma\left(\frac{1}{2}\right)$, maka distribusi dari V adalah $\chi_{(1)}^2$.

Telah diketahui $W_n \xrightarrow{D} N(0,1)$, maka $W_n^2 \xrightarrow{D} \chi_{(1)}^2$, yaitu $\frac{n(\widehat{\theta}_{n,jack} - \theta)^2}{\sigma_{n,jack}^2} \rightarrow \chi_{(1)}^2$.

Dengan demikian terbukti bahwa $l(\theta) = -2 \log R(\theta) = \frac{n(\widehat{\theta}_{n,jack} - \theta)^2}{\sigma_{n,jack}^2}$ konvergen dalam distribusi ke $\chi_{(1)}^2$. \square

Menurut Zhao [10], selanjutnya dapat dicari interval kepercayaan untuk θ berdasarkan teorema 2.3. Dari teorema 2.3. diperoleh bentuk $l(\theta) = -2 \log R(\theta) \rightarrow \chi_{(1)}^2$. Dari bentuk tersebut dapat digunakan untuk membuat interval kepercayaan untuk θ , yaitu untuk sembarang $0 \leq \alpha \leq 1$, misalkan $\Pr(\chi_{(1)}^2 \leq c_\alpha) = 1 - \alpha$, maka interval kepercayaan JEL $100(1 - \alpha)\%$ untuk θ adalah sebagai berikut,

$$R^J = \{\theta: -2 \log R(\theta) \leq c_\alpha\}. \quad (10)$$

2.4. Aplikasi Data untuk Interval Kepercayaan MAD

Diberikan dua himpunan data asli untuk mengilustrasikan interval kepercayaan untuk MAD. Himpunan data pertama adalah hasil analisis laboratorium dari kalori yang terkandung pada 54 buah *hotdog* yang diperoleh dari *Data and Story Library* (DASL) pada *Carnegie Mellon University*. Himpunan data kedua adalah data 38 individu yang menggunakan mesin dialisis ginjal dengan variabel yang diamati adalah waktu kekambuhan infeksi (dalam hari). Data diperoleh dari R *datasets* pada program R. Berdasarkan histogramnya, himpunan data pertama berdistribusi normal, sedangkan himpunan data kedua tidak berdistribusi normal. Berikut merupakan tabel interval kepercayaan dan panjang dari interval kepercayaan berdasarkan taksiran aproksimasi normal (AN) dan taksiran JEL.

Tabel 1. Panjang dari interval kepercayaan dengan himpunan data pertama.

Nominal level		AN		JEL	
		UB	LB	UB	LB
0.99	panjang	29.5681	16.5059	31.1504	18.0336
		13.0622		13.1168	
0.95	panjang	27.940	18.134	29.515	19.669
		9.806		9.846	
0.90	panjang	27.129	18.945	28.701	20.483
		8.184		8.218	

Tabel 2. Panjang dari interval kepercayaan dengan himpunan data kedua.

Nominal level		AN		JEL	
		UB	LB	UB	LB
0.99	panjang	128.213	45.011	145.013	32.879
		83.202		112.134	
0.95	panjang	117.654	55.570	130.782	47.110
		62.084		83.672	
0.90	panjang	112.459	60.765	123.780	54.111
		51.694		69.669	

Dari tabel 1 dan tabel 2 diperoleh *Upper Bound* (UB) dan *Lower Bound* (LB) dan panjangnya untuk setiap interval kepercayaan berdasarkan taksiran aproksimasi normal dan taksiran dari metode JEL. UB dan LB diperoleh untuk tingkat nominal yang berbeda-beda yaitu 99%, 95%, dan 90%. Berdasarkan tabel 1, panjang dari interval kepercayaan menggunakan aproksimasi normal dan JEL tidak terlalu berbeda jauh. Berdasarkan tabel 2, panjang interval kepercayaan dengan metode JEL selalu lebih besar daripada yang berdasarkan aproksimasi normal.

3. Kesimpulan

Dari pembahasan di atas, maka dapat disimpulkan beberapa hal, yaitu:

1. Pembentukan interval kepercayaan untuk suatu parameter menggunakan metode EL akan melibatkan pemaksimuman fungsi *nonparametric likelihood* terhadap beberapa kendala.
2. *Mean absolute deviation* (MAD) merupakan fungsi nonlinier, sehingga pembentukan interval kepercayaan untuk MAD menggunakan metode *empirical likelihood* (EL) sulit dilakukan, sehingga digunakan metode *jackknife empirical likelihood* (JEL).
3. Metode JEL merupakan kombinasi dari metode *jackknife* dan metode EL. Metode JEL ini dapat digunakan untuk mengatasi masalah-masalah pembentukan interval kepercayaan untuk parameter yang berbentuk fungsi nonlinier, seperti MAD.
4. Metode JEL membentuk interval kepercayaan untuk MAD dengan memanfaatkan fungsi *jackknife empirical likelihood ratio* dari MAD-nya.
5. Distribusi asimtotik dari *jackknife empirical likelihood ratio* untuk MAD adalah suatu distribusi chi-kuadrat dengan derajat bebas satu. Berdasarkan distribusi asimtotik tersebut, dapat diperoleh bentuk interval kepercayaan untuk MAD.

Referensi

- [1] Dudewicz, E. J. (1988). *Modern Mathematical Statistics*. Canada: John Wiley & Sons, Inc.
- [2] Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. New York: Chapman and Hall.
- [3] Gaswirth, J. L. (1974). Large sample theory of some measures of income inequality. *Econometrica*, 42, 191-196.
- [4] Govindarajulu, Z., (1999). *Elements of Sampling Theory and Methods*. United States of America: Prentice Hall.
- [5] Hogg, R. V., McKean, J. W. Craig, A. T., (2013). *Introduction to Mathematical Statistics*. United States of America: Pearson Education., Inc.
- [6] Jiang, J., (2010). *Large Sample Techniques for Statistics*. New York: Springer.
- [7] Jing, B., Yuan, Q., Zhou, W. (2009). Jackknife empirical likelihood. *Journal of the American Statistical Association*. 104, 1224-1232.
- [8] Owen, A. B. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika*, 75, 237-249.
- [9] Owen, A. B. (2001). *Empirical Likelihood*. London: Chapman and Hall.
- [10] Zhao, Y., Meng, X., Yang, H. (2015). Jackknife empirical likelihood inference for the mean absolute deviation. *Computational Statistics and Data Analysis*. 91, 92-101.

PROPORTIONAL REDUCTION IN UNCERTAINTY (PRU) DALAM SELEKSI VARIABEL CALON DEBITURKREDIT MIKRO KOMERSIAL

ANNA CHADIDJAH¹, ACHMAD ZANBAR SOLEH²

1. Departemen Statistika FMIPA Unpad, annachadidjah@gmail.com
2. Departemen Statistika FMIPA Unpad, a.zanbar.soleh@unpad.ac.id

Abstrak, Dalam membentuk *Credit Scoring Model* pada Kredit Mikro Komersial di suatu bank, terlebih dahulu harus dilakukan seleksi terhadap variabel yang dominan mempengaruhi tingkat kelancaran pembayaran calon debitur. Informasi mengenai karakteristik calon debitur diperlukan untuk memutuskan menerima atau menolak kredit yang diajukan. Banyaknya informasi karakteristik debitur yang harus dipertimbangkan dalam pengambilan keputusan menimbulkan ketidakpastian layak tidaknya seseorang diberi kredit. Oleh karena itu, diperlukan analisis statistik untuk memilih karakteristik dominan yang mempengaruhi penilaian terhadap calon debitur tersebut. Metode seleksi variabel untuk variabel kategori [1], secara statistik dapat digunakan *Proportional Reduction in Uncertainty (PRU)* yang memanfaatkan nilai ketidakaturan (*entropi*) [2] sebagai ukuran ketidakpastian jawaban dari calon debitur dari data kategori. Dari hasil analisis menggunakan metode PRU, variabel yang terlibat dalam pembentukan *Credit Scoring Model* adalah Plafon, Angsuran, Nilai Likuidasi Agunan, Baki Debet, Status Perkawinan dan Status Debitur.

Kata kunci : Entropi, Variabel Dominan, PRU, Credit scoring

1. Pendahuluan

Informasi mengenai karakteristik calon debitur diperlukan untuk memutuskan menerima atau menolak kredit yang diajukan. Banyaknya informasi karakteristik debitur yang harus dipertimbangkan dalam pengambilan keputusan menimbulkan ketidakpastian layak tidaknya seseorang diberi kredit. Oleh karena itu, diperlukan analisis statistik untuk memilih karakteristik dominan yang mempengaruhi penilaian terhadap calon debitur tersebut. Metode yang digunakan untuk penelitian ini adalah *Proportional Reduction in Uncertainty (PRU)*. Perhitungan nilai PRU tersebut didasarkan pada nilai entropi yang dikembangkan Shannon. Perumusan nilai Entropi Shannon akan didasarkan pada tabel kontingensi yang dibentuk berdasarkan banyaknya variabel dan banyaknya kategori untuk setiap variabel pada data historis profil debitur, atau dengan kata lain tabel kontingensi yang dibentuk oleh multinomial multivariat.

2. Seleksi Variabel dengan PRU

PRU merupakan metode seleksi variabel yang bersifat kategorik. Adapun tahapan yang harus dilakukan adalah sebagai berikut :

1. Buat tabel kontingensi untuk entropi sederhana dari masing-masing karakteristik debitur, sehingga didapat Tabel 1 sebagai berikut :

Tabel 1. Tabel Kontingensi untuk Entropi Sederhana

X_j	1	n_1
	2	n_2

	k	n_k
Jumlah		n

dengan :

k : kategori ke - k dari variabel X_j ($j = 1, 2, \dots, m$)

n_1 : Frekuensi jawaban debitur untuk kategori ke - 1 dari variabel X_j

n_2 : Frekuensi jawaban debitur untuk kategori ke - 2 dari variabel X_j

n_k : Frekuensi jawaban debitur untuk kategori ke - k dari variabel X_j

n : Total frekuensi jawaban debitur dari variabel X_j

2. Hitung entropi sederhana untuk variabel X_j dengan rumus sebagai berikut:

$$H(X_j) = - \sum_{i=1}^k \frac{n_i}{n} \log \frac{n_i}{n} \quad (1)$$

dengan:

n_i : Frekuensi jawaban debitur pada kategori ke- i ($i = 1, 2, \dots, k$) untuk variabel X_j ($j = 1, 2, \dots, m$)

n : Total frekuensi jawaban debitur untuk variabel X_j ($j = 1, 2, \dots, m$)

3. Buat tabel kontingensi untuk entropi variabel X_r bersyarat X_p

Tabel 2. Tabel Kontingensi Frekuensi untuk Entropi Bersyarat

Variabel		X_r				Jumlah
		1	2	...	r	
X_p	1	n_{11}	n_{12}	...	n_{1r}	$n_{1.}$
	2	n_{21}	n_{22}	...	n_{2r}	$n_{2.}$

	p	n_{p1}	n_{p2}	...	n_{pr}	$n_{p.}$
Jumlah		$n_{.1}$	$n_{.2}$...	$n_{.r}$	n

dengan :

n_{11} : Frekuensi banyaknya debitur yang menjawab kategori 1 pada variabel X_p dan kategori 1 pada variabel X_r

n_{12} : Frekuensi banyaknya debitur yang menjawab kategori 1 pada variabel X_p dan kategori 2 pada variabel X_r

n_{21} : Frekuensi banyaknya debitur yang menjawab kategori 2 pada variabel X_p dan kategori 1 pada variabel X_r

- n_{pr} : Frekuensi banyaknya debitur yang menjawab kategori p pada variabel X_p dan kategori r pada variabel X_r
- $n_{1.}$: Jumlah frekuensi banyaknya debitur yang menjawab kategori 1 pada variabel X_p
- $n_{.1}$: Jumlah frekuensi banyaknya debitur yang menjawab kategori 1 pada variabel X_r
- $n_{p.}$: Jumlah frekuensi banyaknya debitur yang menjawab kategori p pada variabel X_p
- $n_{.r}$: Jumlah frekuensi banyaknya debitur yang menjawab kategori r pada variabel X_r
- n : Total jumlah frekuensi jawaban debitur untuk variabel X_r bersyarat X_p

4. Hitung entropi variabel X_r bersyarat X_p dengan rumus sebagai berikut:

$$H(X_r | X_p) = - \sum \sum \frac{n_{pr}}{n} \log \frac{n_{pr}}{n_i} \quad (2)$$

dengan:

- n_{pr} : Frekuensi jawaban debitur pada baris ke - p , kolom ke - r ($p = 1, 2, \dots, k$ dan $r = 1, 2, \dots, k$)
- n : Total jumlah frekuensi jawaban debitur untuk variabel X_r bersyarat X_p
- n_i : Jumlah frekuensi jawaban debitur pada baris ke - i

5. Hitung PRU dengan langkah sebagai berikut:

a. Buat matriks M dengan ukuran $p \times p$ dengan

$$M_{p \times p} = \begin{pmatrix} PRU(X_1|X_1) & PRU(X_2|X_1) & \dots & PRU(X_p|X_1) \\ PRU(X_1|X_2) & PRU(X_2|X_2) & \dots & PRU(X_p|X_2) \\ \dots & \dots & \dots & \dots \\ PRU(X_1|X_p) & PRU(X_2|X_p) & \dots & PRU(X_p|X_p) \end{pmatrix}$$

dengan :

$$PRU(X_j | X_i) = \frac{H(X_j) - H(X_j | X_i)}{H(X_j)} \quad (3)$$

dimana :

$$PRU(X_i | X_j) \neq PRU(X_j | X_i) \quad (4)$$

- b. Hitung jumlah setiap baris pada matriks M
- c. Cari variabel dengan jumlah baris yang maksimum
- d. Hapus baris dan kolom dari variabel yang memiliki jumlah baris yang maksimum
- e. Ulang tahapan b hingga d sampai tersisa dua baris

- f. Buat tabel SPRU yang disusun dari variabel yang memiliki SPRU yang tertinggi hingga terendah yang ditampilkan pada Tabel 3 berikut.

Tabel 3. Tabel SPRU masing-masing variabel

X_j	$SPRU(X_j)$
X_1	$SPRU(X_1)$
X_2	$SPRU(X_2)$
X_3	$SPRU(X_3)$
X_4	$SPRU(X_4)$
...	...
X_m	$SPRU(X_m)$

dengan:

X_j : Variabel dominan terpilih ke- j

$SPRU(X_j)$: Jumlah PRU dari variabel X_j

6. Hitung nilai kumulatif dari variabel terpilih pada tahapan ke- k dengan rumus sebagai berikut:

$$CI(k) = CI(k-1) + [1 - CI(k-1)] \frac{SPRU(k)}{(m-k+1)} \quad (5)$$

dengan:

$CI(k)$ = nilai informasi kumulatif ke- k ($k=1,2,\dots,n$)

$SPRU(k)$ = Jumlah PRU dari variabel pada tahap ke- k ($k=1,2,\dots,n-1$)

m = banyaknya variabel

$CI(0)$ = 0

Setelah melakukan tahapan-tahapan diatas maka akan didapatkan tabel informasi kumulatif dari setiap variabel X_j seperti pada Tabel 4 berikut:

Tabel 4. Tabel Informasi Kumulatif masing-masing variabel

X_j	$CI(X_j)$
X_1	$CI(X_1)$
X_2	$CI(X_2)$
X_3	$CI(X_3)$
X_4	$CI(X_4)$
...	...
X_m	$CI(X_m)$

dengan:

X_j : Variabel dominan terpilih ke- j

$CI(X_j)$: Nilai informasi kumulatif dari variabel X_j

Hasil dari Tabel 4 di atas merupakan nilai kumulatif dari masing-masing variabel dominan terpilih yang diurutkan dari nilai tertinggi hingga terendah. Setelah didapat nilai informasi kumulatif dapat juga dihitung nilai kontribusi masing-masing variabel. Dengan demikian, dapat diketahui besarnya kontribusi masing-masing variabel dari yang tertinggi sampai terendah terhadap pengambilan keputusan layak atau tidaknya seorang calon debitur mendapatkan kredit.

3. Hasil Penelitian dan Pembahasan

Metode PRU ini diterapkan pada data debitur Kredit Mikro Komersial Bank X yang diperkirakan berpengaruh terhadap kelancaran pembayaran angsuran dan bunga dari calon debitur seperti tercantum pada Tabel 5.

Tabel 5. Variabel Penelitian

Variabel	Skala Ukur
Take Home Pay	Rasio
Plafon	Rasio
Besar Angsuran	Rasio
Baki Debet	Rasio
Nilai Likuiditas Agunan	Rasio
Usia	Rasio
Jenis Kepemilikan Agunan	Nominal
Status Debitur	Nominal
Lama Usaha	Nominal
Status Perkawinan	Ordinal
Tujuan Penggunaan Kredit	Nominal

Tabel 6. Menampilkan hasil perhitungan SPRU dan nilai Kumulatif Informasi CI) dari data debitur Kredit Komersial Bank X

Tahap	Variabel	SPRU	CI(%)
1	Plafon	1,9605	17,82
2	Besar Angsuran	1,6340	31,25
3	Nilai Likuidasi Agunan	1,2404	40,73
4	Baki Debet	1,1275	49,08
5	Status Perkawinan	1,0950	57,05
6	Status Debitur	1,0680	64,69
7	Lama Usaha	1,0324	71,98
8	Jenis Kepemilikan Agunan	1,0197	79,13
9	Tujuan Penggunaan Kredit	1,0024	86,10
10	Usia	1,0004	93,05
11	Take Home Pay	1	100

Dari Tabel 6 diatas dapat diinformasikan bahwa pada tahap satu pemilihan variabel dominan, terpilih variabel Plafon (batas maksimal besar pinjaman) artinya Bank X menjadikan variabel Plafon sebagai prioritas utama dalam proses pemberian kredit kepada calon debitur dengan sumbangan informasi sebesar 17,82 %. Dengan demikian hal pertama yang harus diperhatikan oleh Bank X dalam pemberian kredit kepada calon debitur Kredit Mikro Komersial adalah Plafon . Selanjutnya variabel dominan kedua yang menentukan adalah Besar Angsuran dengan besarnya sumbangan informasi yang diberikan oleh variabel tersebut adalah 13,43 %. Variabel dominan ketiga yang harus diprhatikan adalah Nilai Likuiditas dengan sumbangan informasi sebesar 9,48% begitu sterusnya hingga variabel yang memberikan sumbangan informasi terendah yaitu variabel Usia dan variabel Take Home Pay.

4. Kesimpulan

PRU memberikan rumusan dalam seleksi variabel dominan dengan memperhatikan ketidakaturan (entropi) berupa ketidakpastian jawaban yang diberikan responden. Nilai SPRU sejalan dengan nilai *Cumulatif Informasi (CI)* yang diberikan oleh variabel tersebut. Variabel dengan nilai SPRU terbesar merupakan variabel dominan pertama yang harus diperhatikan yang juga memberikan informasi terbesar terhadap permasalahan yang dihadapi. Dari hasil analisis menggunakan metode PRU, variabel yang terlibat dalam pembentukan *Credit Scoring Model* adalah Plafon, Angsuran, Nilai Likuidasi Agunan, Baki Debet, Status Perkawinan dan Status Debitur.

Referensi

- [1] Agresti, A. 2002. *Categorical Data Analysis Second Edition*. New York: John Wiley and Sons, Inc. USA.
- [2] De Falguerolles, AJS, 1993, Un Modele Graphique pour la Selection de Variables Qualitatives, *Revue de Statistiques Appliquess*, XLI(2), 23-41
- [3] Klir, G.J. and T.A. Folger. 1988. *Fuzzy Set, Uncertainty, and Information*. New Jersey 07632: Prentice Hall, Englewood Cliffs
- [4] Reynold , H.T., 1977, *Analysis of Nominal Data*, Sage Publication,

DETERMINAN BAYI DENGAN BERAT BADAN LAHIR RENDAH DI NUSA TENGGARA TIMUR DENGAN METODE GWNBR

LIZA KURNIA SARI¹

¹Sekolah Tinggi Ilmu Statistik, lizakurnia@stis.ac.id

Abstrak. Bayi berat badan lahir rendah (BBLR) adalah bayi lahir hidup dengan berat badan saat lahir kurang dari 2.500 gram tanpa memandang usia kehamilan. Risiko kematian bayi BBLR 20 kali lebih tinggi dibandingkan bayi dengan berat lahir normal sehingga mengurangi kejadian bayi lahir dengan BBLR akan mendukung target pemerintah untuk mengurangi angka kematian bayi (AKB). Banyak faktor yang diduga memengaruhi berat badan bayi saat lahir yaitu usia ibu, paritas, status ibu bekerja, komplikasi selama kehamilan, konsumsi zat besi selama ibu hamil, dan pemeriksaan kehamilan. Faktor yang memengaruhi ini berbeda antar wilayah sesuai karakteristik wilayah pengamatan. Penelitian ini bertujuan mengetahui faktor yang memengaruhi kejadian bayi lahir dengan BBLR, yang berupa data cacahan, di setiap kabupaten/kota di Nusa Tenggara Timur. GWNBR (*Geographically Weighted Negative Binomial Regression*) digunakan ketika data cacahan mengalami heterogenitas spasial dan overdispersi. Metode ini menghasilkan model-model lokal dengan taksiran parameter yang berlaku lokal. Parameter model GWNBR dihitung dengan metode iterasi IRWLS dan Newton Raphson secara bergantian hingga nilainya konvergen dimana pembobot dihitung berdasarkan fungsi kernel adaptif kuadrat ganda. Dengan metode GWNBR provinsi Nusa Tenggara Timur dibagi menjadi 7 kelompok kabupaten/kota berdasarkan faktor yang berpengaruh signifikan terhadap kejadian BBLR, yaitu banyak ibu berusia kurang dari 20 tahun, banyak ibu dengan status bekerja, banyak ibu yang mengonsumsi zat besi sesuai aturan (minimal 90 hari), dan banyak ibu yang memeriksakan kehamilan dengan indikator K4. Berdasarkan nilai AIC, model GWNBR dengan parameter overdispersi bersifat global lebih baik daripada model GWNBR dengan parameter overdispersi bersifat lokal maupun dengan model GWPR.

Kata kunci: BBLR, model lokal, overdispersi, GWNBR, pembobot.

1. Pendahuluan

Tujuan 3 SDGs (*Sustainable Development Goals*) adalah menjamin kehidupan yang sehat dan mendorong kesejahteraan bagi semua orang di segala usia dengan target pada tahun 2030 mengakhiri kematian bayi dan balita yang dapat dicegah, dengan seluruh negara berusaha menurunkan angka kematian neonatal setidaknya

hingga 12 per 1.000 kelahiran hidup dan angka kematian balita 25 per 1.000 kelahiran hidup. Langkah untuk menurunkan angka kematian bayi berkaitan dengan menurunkan banyak bayi yang lahir dengan berat badan lahir rendah (BBLR) karena risiko kematian bayi BBLR 20 kali lebih tinggi dibandingkan bayi dengan berat lahir normal (UNICEF & WHO [9]).

Bayi BBLR adalah bayi lahir hidup dengan berat badan saat lahir kurang dari 2.500 gram tanpa memandang usia kehamilan. Banyak faktor yang mempengaruhi berat badan bayi saat lahir. Faktor internal dari sisi ibu hamil terdiri dari usia, jarak antar kehamilan, jumlah anak yang dilahirkan, penambahan berat ibu selama kehamilan, frekuensi pemeriksaan kehamilan, konsumsi zat besi, dan ibu merokok atau tidak. Sedangkan faktor eksternal terdiri dari lingkungan tempat tinggal dan pendidikan ibu (Ohlsson [8], Jaya dan Siza dalam Nina dkk, [7]).

Dalam berbagai studi, faktor-faktor yang memengaruhi bayi dengan BBLR dianggap sama di semua wilayah penelitian. Wilayah penelitian diasumsikan mempunyai proses stasioner sehingga dapat dibuat sebuah model regresi global untuk menggambarkan hubungan antara variabel respon dengan sekumpulan variabel bebas. Di titik mana pun data diambil, nilai taksiran parameter regresi sama untuk keseluruhan wilayah penelitian. Analisis ini tidak sesuai untuk data spasial yang mengalami ketidakstasioneran spasial karena karakteristik, kondisi geografis, demografi, dan sosial ekonomi setiap wilayah berbeda. Dampak ketidakstasioneran spasial adalah parameter regresi akan bervariasi secara spasial. Akibatnya faktor yang berpengaruh terhadap kejadian bayi lahir dengan BBLR di setiap wilayah berbeda (Lusi [5]).

Apabila sebuah model global dibangun dalam suatu wilayah yang tidak stasioner, taksiran parameter yang dihasilkan akan bersifat bias, tidak efisien, dan mempunyai galat baku yang besar. Akibatnya model tidak *fit* karena tidak bisa mengidentifikasi parameter yang signifikan (Charlton [1]).

Fotheringham membangun model *Geographically Weighted Regression* (GWR) sebagai solusi untuk menganalisis data yang mengalami ketidakstasioneran spasial. GWR menggunakan pembobot berdasarkan letak lintang dan bujur dari titik pengamatan. Model GWR digunakan untuk data berdistribusi normal sehingga model ini tidak sesuai jika variabel respon berupa data cacahan. Regresi Poisson dan binomial negatif lebih sesuai digunakan jika variabel respon berupa data cacahan (da Silva & Rodrigues [2]). Nakaya dkk [6] mengembangkan model *Geographically Weighted Poisson Regression* (GWPR) untuk memodelkan variabel respon berupa data cacahan yang mengalami ketidakstasioneran spasial. GWPR harus memenuhi asumsi distribusi Poisson yaitu varians data sama dengan rataan. Jika asumsi ini tidak dipenuhi maka dapat terjadi gejala overdispersi. Overdispersi terjadi jika varians lebih besar daripada rataan. Model *Geographically Weighted Negative Binomial Regression* (GWNBR) digunakan untuk menangani gejala overdispersi pada data cacahan yang mengalami ketidakstasioneran spasial.

Penelitian ini bertujuan mendapatkan model terbaik untuk menggambarkan hubungan antara banyak kejadian BBLR dengan faktor-faktor penyebabnya di Nusa Tenggara Timur dan kemudian mengelompokkan kabupaten/kota di Nusa Tenggara Timur berdasarkan faktor yang berpengaruh signifikan.

2. Hasil – Hasil Utama

2.1 Geographically Weighted Regression

Analisis regresi bertujuan menaksir nilai variabel respon y jika variabel bebas x diberikan. Besar pengaruh perubahan nilai x terhadap y dinyatakan oleh suatu besaran yang disebut parameter. Dalam analisis regresi kita menaksir nilai parameter dan kemudian menentukan apakah nilai parameter tersebut berbeda dari nol secara signifikan. Jika parameter tersebut signifikan maka perubahan variabel bebas akan berpengaruh terhadap variabel respon.

Fotheringham dkk [3] mengembangkan *Geographically Weighted Regression* (GWR), suatu model regresi yang berbasis letak geografis untuk memodelkan ketidakstasioneran spasial yang terdapat pada data. GWR mampu melihat adanya hubungan nonstasioner antar lokasi pengamatan yang tidak bisa dilakukan oleh model global. GWR menghasilkan galat yang lebih kecil dibandingkan model global dan tidak berautokorelasi spasial. Model GWR yang dibangun Fotheringham dkk adalah

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1} \beta_k(u_i, v_i)x_{ik} + \varepsilon_i \quad (1)$$

(u_i, v_i) menunjukkan koordinat titik i dan $\beta_k(u_i, v_i)$ menunjukkan taksiran parameter k di lokasi (u_i, v_i) . Titik pengamatan yang lebih dekat ke- i mempunyai pengaruh lebih besar dalam menaksir $\beta_k(u_i, v_i)$ daripada titik pengamatan yang lebih jauh sehingga diberi pembobot lebih besar. Taksiran parameter di daerah penaksiran yang berpusat di titik regresi i yaitu $\beta(u_i, v_i)$ pada persamaan (1) dihitung menggunakan metode kuadrat terkecil terboboti atau *Weighted Least Square* (WLS) dengan rumus berikut

$$\hat{\beta}(u_i, v_i) = [\mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{X}]^{-1} \mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{y}$$

$\mathbf{W}(u_i, v_i)$ adalah matriks diagonal berukuran n yang berisi pembobot relatif untuk titik i dari n titik pengamatan. Matriks ini spesifik untuk titik i .

2.2 Geographically Weighted Poisson Regression

GWR hanya sesuai untuk memodelkan data berdistribusi normal sehingga berkembang model *Geographically Weighted Poisson Regression* (GWPR) jika

variabel respon merupakan data cacahan yang berdistribusi Poisson dan mengalami ketidakstasioneran spasial. Data tersebut berasal dari percobaan binomial tetapi frekuensi terjadinya jarang.

Model regresi Poisson dapat diperluas untuk data spasial menjadi model GWPR yaitu

$$\ln \mu_i = \beta_0(u_i, v_i) + \sum_{k=1} \beta_k(u_i, v_i)x_{ik}$$

Parameter model $\beta(u_i, v_i)$ ditaksir mengikuti prinsip penaksiran pada GWR.

2.3 Geographically Weighted Negative Binomial Regression

Overdispersi merupakan kondisi di mana varians data lebih besar daripada rataan. Overdispersi terjadi karena adanya korelasi positif antar variabel respon, adanya variasi dalam peluang respon, atau adanya pengelompokan dalam populasi. Dalam praktiknya, overdispersi terjadi karena ada variabel bebas yang penting, tetapi tidak masuk dalam model, data mengandung pencilan, variabel bebas perlu ditransformasi, atau terjadi kesalahan dalam menentukan fungsi penghubung (Hilbe [4]).

Apabila gejala overdispersi pada data cacahan yang mengalami ketidakstasioneran spasial diabaikan maka taksiran parameter yang dihasilkan bias ke bawah. Taksiran galat baku akan terlalu kecil sehingga menyebabkan kesalahan dalam statistik inferensi karena variabel bebas yang semestinya tidak berpengaruh terhadap variabel respon diidentifikasi sebagai variabel yang berpengaruh (Lusi [5]). Model *Geographically Weighted Negative Binomial Regression* (GWNBR) dibangun untuk mengatasi fenomena overdispersi dan ketidakstasioneran spasial pada data cacahan.

GWNBR merupakan pengembangan dari model binomial negatif dengan parameter β_k dan α bervariasi secara spasial. Model GWNBR dideskripsikan da Silva dan Rodrigues sebagai berikut:

$$y_j \sim \text{Binomneg}[t_j \exp(\beta_k(u_j, v_j)x_{jk}), \alpha_j] \quad (2)$$

dengan y_j adalah variabel respon ke- j untuk $j = 1, 2, \dots, n$, t_j adalah variabel *offset*, β_k adalah parameter yang berhubungan dengan variabel bebas x_k untuk $k = 1, 2, \dots, K$, α_j adalah parameter overdispersi, dan (u_j, v_j) adalah koordinat titik pengamatan j .

2.3.1 Penaksiran Parameter

Fungsi peluang terjadinya y_j dituliskan sebagai

$$f(y_j | \mu_j, \alpha_j) = \exp \left[y_j \ln \frac{\alpha_j \mu_j}{1 + \alpha_j \mu_j} - \frac{1}{\alpha_j} \ln(1 + \alpha_j \mu_j) + \ln \left[\Gamma \left(y_j + \frac{1}{\alpha_j} \right) \right] \right. \\ \left. - \ln[\Gamma(y_j + 1)] - \ln \left[\Gamma \left(\frac{1}{\alpha_j} \right) \right] \right]$$

dengan $\mu_j = t_j \exp[\beta_k(u_j, v_j)x_{jk}]$

Parameter β dan α ditaksir secara bergantian hingga nilainya konvergen. Langkah pertama adalah menentukan nilai awal α dan β . Nilai awal ini digunakan untuk menaksir $\beta(u, v)$ dengan metode IRWLS. Karena $\beta(u, v)$ ditaksir untuk setiap daerah penaksiran yang berpusat di titik regresi i dengan $i = 1, 2, \dots, p$ maka logaritma natural dari fungsi kemungkinan lokal model GWNBR adalah

$$l[\beta(u_i, v_i) | x_{jk}, y_j, \alpha_i] \\ = \sum_{j=1}^n \left\{ y_j \ln[\alpha_i \mu_j(\beta(u_i, v_i))] - \left(y_j + \frac{1}{\alpha_i} \right) \ln \left(1 + \alpha_i \mu_j(\beta(u_i, v_i)) \right) \right. \\ \left. + \ln \left[\Gamma \left(y_j + \frac{1}{\alpha_i} \right) \right] - \ln \left[\Gamma \left(\frac{1}{\alpha_i} \right) \right] - \ln[\Gamma(y_j + 1)] \right\} w_{ij} \\ (3)$$

dengan

$$\mu_j(\beta(u_i, v_i)) = t_j \exp[\sum_k \beta_k(u_i, v_i)x_{jk}]$$

dengan $j = 1, 2, \dots, n$. $\mu_j(\beta(u_i, v_i))$ adalah prediksi rata-rata di titik pengamatan j dengan parameter-parameter di titik regresi i . w_{ij} adalah pembobot relatif titik pengamatan j di titik regresi i yang nilainya ditentukan berdasarkan jarak antara j dan i .

Taksiran $\beta(u_i, v_i)$ ditentukan dengan memaksimumkan persamaan (3). Nakaya dkk [6] dan da Silva & Rodrigues [2] menyebutkan solusi persamaan (3) pada iterasi ke- $(m + 1)$ adalah

$$\hat{\beta}(u_i, v_i)^{(m+1)} = [\mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{A}(u_i, v_i)^{(m)} \mathbf{X}]^{-1} \mathbf{X}^T \\ \times \mathbf{W}(u_i, v_i) \mathbf{A}(u_i, v_i)^{(m)} \mathbf{z}(u_i, v_i)^{(m)}$$

\mathbf{X} adalah matriks yang berisi nilai variabel bebas, berdimensi $n \times (k + 1)$ dengan n menyatakan banyak titik pengamatan dalam daerah penaksiran dengan pusat titik regresi i dan k menyatakan banyak variabel bebas

$$\mathbf{X} = \begin{pmatrix} 1 & x_{11} & \dots & x_{1k} \\ 1 & x_{21} & \dots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{nk} \end{pmatrix}$$

$\mathbf{W}(u_i, v_i)$ adalah matriks pembobot diagonal GWR untuk titik i , berdimensi $n \times n$

$$\mathbf{W}(u_i, v_i) = \begin{pmatrix} w_{i1} & 0 & \dots & 0 \\ 0 & w_{i2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_{in} \end{pmatrix}$$

(4)

$\mathbf{A}(u_i, v_i)^{(m)}$ adalah matriks pembobot diagonal GLM untuk titik i pada iterasi ke- m , berdimensi $n \times n$

$$\mathbf{A}(u_i, v_i)^{(m)} = \begin{pmatrix} a_{i1}^{(m)} & 0 & \dots & 0 \\ 0 & a_{i2}^{(m)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{in}^{(m)} \end{pmatrix}$$

(5)

$$\text{dengan } a_{ij}^{(m)} = \frac{\mu_j(\hat{\boldsymbol{\beta}}(u_i, v_i)^{(m)})}{1 + \alpha_i \mu_j(\hat{\boldsymbol{\beta}}(u_i, v_i)^{(m)})} + \frac{[y_j - \mu_j(\hat{\boldsymbol{\beta}}(u_i, v_i)^{(m)})][\alpha_i \mu_j(\hat{\boldsymbol{\beta}}(u_i, v_i)^{(m)})]}{1 + 2\alpha_i \mu_j(\hat{\boldsymbol{\beta}}(u_i, v_i)^{(m)}) + \alpha_i^2 \mu_j^2(\hat{\boldsymbol{\beta}}(u_i, v_i)^{(m)})}$$

$\mathbf{z}(u_i, v_i)^{(m)}$ adalah vektor variabel respon yang disesuaikan, berdimensi $n \times 1$

$$\mathbf{z}(u_i, v_i)^{(m)} = \mathbf{X} \hat{\boldsymbol{\beta}}(u_i, v_i)^{(m)} + \frac{y_j - \mu_j(\hat{\boldsymbol{\beta}}(u_i, v_i)^{(m)})}{a_{ij}^{(m)} [1 + \alpha_i \mu_j(\hat{\boldsymbol{\beta}}(u_i, v_i)^{(m)})]}$$

Setelah taksiran $\boldsymbol{\beta}(u_i, v_i)$ ditentukan, langkah selanjutnya adalah menaksir α_i dengan metode Newton Raphson berdasarkan logaritma natural dari fungsi kemungkinan lokal pada persamaan (3). Untuk menyederhanakan penghitungan, ditaksir nilai r_i di mana $r_i = \frac{1}{\alpha_i}$. Persamaan (3) dituliskan ulang menjadi

$$\begin{aligned} & l[r_i | y_j, \boldsymbol{\beta}(u_i, v_i)] \\ &= \sum_{j=1}^n \left\{ y_j \ln[\mu_j(\boldsymbol{\beta}(u_i, v_i))] - (y_j + r_i) \ln(r_i + \mu_j(\boldsymbol{\beta}(u_i, v_i))) + r_i \ln r_i \right. \\ & \quad \left. + \ln[\Gamma(y_j + r_i)] - \ln[\Gamma(r_i)] - \ln[\Gamma(y_j + 1)] \right\} w_{ij} \end{aligned}$$

(6)

Dengan memaksimumkan persamaan (6), da Silva & Rodrigues [2] menyebutkan taksiran r_i pada iterasi ke-(m+1) adalah

$$r_i^{(m+1)} = r_i^{(m)} - [H_i^{(m)}]^{-1} U_i^{(m)}$$

dengan $U_i^{(m)}$ dan $H_i^{(m)}$ masing-masing adalah turunan pertama dan kedua persamaan (6) terhadap r_i .

$$U_i^{(m)} = \sum_{j=1}^n \left\{ \frac{d \ln \Gamma(r_i^{(m)} + y_i)}{dr_i} - \frac{d \ln \Gamma(r_i^{(m)})}{dr_i} + \ln(r_i^{(m)}) + 1 \right. \\ \left. - \ln \left(r_i^{(m)} + \mu_j(\boldsymbol{\beta}(u_i, v_i)) \right) - \frac{r_i^{(m)} + y_i}{r_i^{(m)} + \mu_j(\boldsymbol{\beta}(u_i, v_i))} \right\} w_{ij}$$

$$H_i^{(m)} = \sum_{j=1}^n \left\{ \frac{d^2 \ln \Gamma(r_i^{(m)} + y_i)}{dr_i^2} - \frac{d^2 \ln \Gamma(r_i^{(m)})}{dr_i^2} + \frac{1}{r_i^{(m)}} \right. \\ \left. - \frac{2}{r_i^{(m)} + \mu_j(\boldsymbol{\beta}(u_i, v_i))} + \frac{r_i^{(m)} + y_i}{[r_i^{(m)} + \mu_j(\boldsymbol{\beta}(u_i, v_i))]^2} \right\} w_{ij}$$

Setelah mendapatkan taksiran r_i , kembali ditaksir $\boldsymbol{\beta}_i$. Proses diulangi sampai taksiran kedua parameter konvergen.

da Silva juga memperkenalkan model GWNBR dengan parameter overdispersi α bersifat stasioner dan hanya $\boldsymbol{\beta}(u_i, v_i)$ yang bervariasi secara spasial. Model ini diberi nama GWNBRg sehingga persamaan (1) menjadi

$$y_j \sim \text{Binomneg}[t_j \exp(\beta_k(u_j, v_j) x_{jk}), \alpha] \quad (7)$$

Taksiran parameter α yang berlaku global digunakan untuk menaksir $\boldsymbol{\beta}(u_i, v_i)$ yang berlaku lokal. Karena α ditaksir global, taksiran $\boldsymbol{\beta}(u_i, v_i)$ di setiap lokasi hanya menggunakan metode IRWLS saja.

2.3.2 Fungsi Kernel

Taksiran parameter dari model GWNBR tidak hanya bergantung pada nilai variabel bebas, tetapi juga pada *bandwidth* yang digunakan dalam menghitung pembobot atau fungsi kernel. *Bandwidth* adalah radius daerah penaksiran yang berpusat di titik i . Titik pengamatan yang berada dalam daerah tersebut memiliki pengaruh dalam penaksiran parameter di titik i . Elemen matriks pembobot GWR yaitu w_{ij} pada persamaan (4) ditentukan berdasarkan kedekatan titik regresi i dengan titik pengamatan j . Titik pengamatan yang dekat ke titik regresi diberi bobot lebih besar daripada titik pengamatan yang lebih jauh.

Ada 2 jenis *bandwidth* yang digunakan dalam menghitung pembobot atau fungsi kernel yaitu *bandwidth* jarak dan *bandwidth* tetangga terdekat. Untuk *bandwidth* jarak digunakan fungsi kernel tetap Gaussian yaitu

$$w_{ij} = \exp \left[-\frac{1}{2} \left(\frac{d_{ij}}{b} \right)^2 \right]$$

sedangkan untuk *bandwidth* tetangga terdekat digunakan fungsi kernel adaptif kuadrat ganda yaitu

$$w_{ij} = \begin{cases} \left[1 - \left(\frac{d_{ij}}{b} \right)^2 \right] & \text{jika } j \text{ adalah satu dari } N \text{ titik terdekat dari } i \\ 0 & \text{untuk lainnya} \end{cases}$$

dengan $d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2}$ adalah jarak Euclian antara titik regresi i dan titik pengamatan j dan b adalah *bandwidth*.

Penaksiran parameter tidak sensitif pada pemilihan kernel, tetapi sensitif pada *bandwidth* dari kernel. *Bandwidth* seperti memuluskan parameter, *bandwidth* yang lebih besar membuat parameter lebih mulus. Model yang terlalu mulus akan menghasilkan parameter yang hampir sama di semua daerah sehingga biasanya besar, sedangkan model yang tidak mulus menghasilkan parameter yang sangat bersifat lokal dengan bias kecil tetapi memiliki varians yang sangat besar (Fotheringham dkk. [3]). *Bandwidth* optimal ditentukan berdasarkan nilai validasi silang (CV) yang paling kecil.

$$CV(b) = \sum_{i=1}^n [y_i - \hat{y}_{\neq i}(b)]^2$$

dengan $\hat{y}_{\neq i}(b)$ adalah taksiran dari pengamatan y_i dengan *bandwidth* b tetapi taksiran pengamatan di titik regresi i dihilangkan dari proses penaksiran. Prosesnya dilakukan secara numerik dengan teknik pencarian *Golden Section*. Teknik ini mengevaluasi nilai CV pada jarak minimum sampai jarak maksimum antar lokasi pengamatan.

2.3.3 Pengujian Signifikansi Parameter dan Kesesuaian Model

Signifikansi taksiran parameter β diuji dengan hipotesis $H_0 : \beta_k = 0$ melawan $H_1 : \beta_k \neq 0$. Alat uji yang digunakan adalah statistik Wald, yaitu

$$G_{\beta} = \left(\frac{\hat{\beta}_k}{se(\hat{\beta}_k)} \right)^2 \sim \chi^2_{(\alpha;1)}$$

Kriteria uji yang digunakan adalah menolak H_0 jika $G_{\beta} > \chi^2_{(\alpha;1)}$. $se(\hat{\beta}_k)$ adalah galat baku taksiran parameter yang merupakan akar kuadrat dari varians. Nilai varians β diperoleh dari diagonal utama matriks varians kovarians berikut:

$$\text{Cov}(\hat{\beta}(u_i, v_i)) = \mathbf{C}(u_i, v_i)\mathbf{A}(u_i, v_i)^{-1}\mathbf{C}^T(u_i, v_i)$$

dengan $\mathbf{C}(u_i, v_i) = [\mathbf{X}^T\mathbf{W}(u_i, v_i)\mathbf{A}(u_i, v_i)\mathbf{X}]^{-1}\mathbf{X}^T\mathbf{W}(u_i, v_i)\mathbf{A}(u_i, v_i)$ dimana $\mathbf{W}(u_i, v_i)$ dan $\mathbf{A}(u_i, v_i)$ masing-masing diperoleh dari persamaan (4) dan (5).

Kesesuaian model GWNBR diuji dengan hipotesis H_0 : model *fit* melawan H_1 : model tidak *fit*. Alat uji yang digunakan adalah statistik devians. Devians merupakan perbandingan antara model yang dibangun dengan sebuah model penuh yang sempurna dalam menjelaskan data. Devians dirumuskan sebagai berikut:

$$\begin{aligned} D &= -2 \ln \frac{L(\mu_j(\beta(u_i, v_i))|y_j)}{L(y_j|y_j)} \\ &= 2 \sum_{j=1}^n \left[y_j \ln \left(\frac{y_j}{\mu_j(\beta(u_i, v_i))} \right) - \left(y_j + \frac{1}{\alpha_i} \right) \ln \left(\frac{1 + \alpha_i y_j}{1 + \alpha_i \mu_j(\beta(u_i, v_i))} \right) \right] w_{ij} \\ &\sim \chi^2_{(\alpha; n-K)} \end{aligned}$$

Kriteria uji yang digunakan adalah menolak H_0 jika $D > \chi^2_{(\alpha; n-K)}$.

2.4 Hasil dan Pembahasan

2.4.1 Data

Data yang digunakan bersumber dari Riset Kesehatan Dasar (Riskesdas) yang dilakukan oleh Badan Penelitian dan Pengembangan Kesehatan (Balitbangkes) Kementerian Kesehatan RI tahun 2013. Variabel dalam penelitian ini meliputi variabel respon, variabel bebas, dan variabel geografis. Variabel respon (Y) dalam penelitian ini adalah banyaknya kejadian BBLR di setiap kabupaten/kota di Provinsi Nusa Tenggara Timur. Variabel bebas yang diduga memengaruhi banyaknya kejadian BBLR adalah status ibu bekerja atau tidak, usia ibu saat melahirkan, nomor urut kehamilan (paritas), komplikasi selama kehamilan, konsumsi zat besi, dan frekuensi memeriksakan kehamilan. Variabel bebas dijabarkan menjadi banyak ibu dengan status bekerja (X_1), banyak ibu berusia kurang dari 20 tahun (X_2), banyak ibu dengan kelahiran pertama atau paritas 1 (X_3), banyak ibu yang mengalami komplikasi selama kehamilan (X_4), banyak ibu yang mengonsumsi zat besi sesuai aturan (minimal 90 hari) (X_5), dan banyak ibu yang memeriksakan kehamilan dengan indikator K4 yaitu minimal 1 kali pada trimester pertama, minimal 1 kali pada trimester kedua, dan minimal 2 kali pada trimester ketiga (X_6). Sedangkan variabel geografis adalah koordinat kabupaten/kota. Setiap kabupaten/kota diwakili oleh satu koordinat.

Variabel bebas yang digunakan harus memenuhi asumsi tidak terdapat multikolinieritas. Kriteria yang digunakan untuk memeriksa adanya multikolinieritas antar variabel bebas adalah *variance inflation factor* (VIF).

Keenam variabel bebas yang digunakan sudah memenuhi asumsi tidak terdapat multikolinieritas karena nilai VIF kurang dari 10 seperti dapat dilihat pada Tabel 1.

Tabel 1 Nilai VIF Variabel Bebas

Variabel	X_1	X_2	X_3	X_4	X_5	X_6
VIF	2,963	1,389	2,789	2,500	3,009	1,872

2.4.2 Model GWNBR

Untuk membangun model lokal diperlukan nilai *bandwidth* dan matriks pembobot. *Bandwidth* yang digunakan dalam penelitian ini adalah *bandwidth* tetangga terdekat karena titik pengamatan tidak tersebar merata. *Bandwidth* optimal untuk model GWNBR adalah 13, artinya sebanyak 13 titik pengamatan digunakan untuk menentukan matriks pembobot di setiap lokasi. Sedangkan *bandwidth* optimal untuk model GWNBRg adalah 21.

Prosedur selanjutnya adalah menaksir parameter model GWNBR dengan menggunakan keenam variabel bebas melalui iterasi numerik. Parameter β ditaksir dengan metode IRWLS dan parameter α ditaksir dengan metode Newton-Raphson. Penaksiran dilakukan secara bergantian sampai nilai kedua parameter konvergen. Proses iterasi dilakukan per kabupaten/kota, dimulai dari Sumba Barat sampai yang terakhir Kota Kupang. Jangkauan taksiran setiap parameter dapat dilihat pada Tabel 2.

Tabel 2 Statistik Taksiran Parameter Model GWNBR

Parameter	Taksiran Model GWNBR		Taksiran Model GWNBRg	
	Min	Maks	Min	Maks
$\hat{\alpha}$	0	2,5	1,078	
$\hat{\beta}_0$	-3,26	13,361	6,286	7,567
$\hat{\beta}_1$	-0,114	1,55	-0,046	-0,022
$\hat{\beta}_2$	-4,06	-0,030	-0,309	-0,109
$\hat{\beta}_3$	0,008	1,84	0,028	0,054
$\hat{\beta}_4$	-9,08	0,099	0,021	0,069
$\hat{\beta}_5$	-0,006	3,12	-0,002	0,028
$\hat{\beta}_6$	-0,069	3,66	-0,027	-0,018

Nilai AIC (Akaike Information Criterion) untuk model GWNBR, GWNBRg, dan GWPR masing-masing sebesar 285,587; 268,215; dan 485,175. Berdasarkan nilai AIC dipilih model terbaik adalah GWNBRg karena memiliki nilai AIC paling kecil.

Berikut ini dijelaskan pengujian signifikansi parameter untuk model terbaik, GWNBRg, untuk mengetahui variabel bebas yang berpengaruh signifikan terhadap kejadian BBLR di setiap kabupaten/kota. Pengujian dilakukan menggunakan statistik

Wald dengan taraf signifikansi 10% dan titik kritis 2,706. Hasil pengujian model dengan statistik Wald mengelompokkan kabupaten/kota di Nusa Tenggara Timur menjadi 7 kelompok berdasarkan variabel yang mempengaruhi kejadian BBLR seperti dapat dilihat pada Tabel 3. Variabel banyak ibu bekerja, banyak ibu berusia kurang dari 20 tahun, dan banyak ibu yang memeriksakan kehamilan dengan indikator K4 berpengaruh negatif terhadap banyak kejadian BBLR. Sedangkan banyak ibu yang mengonsumsi zat besi sesuai aturan berpengaruh positif terhadap banyak kejadian BBLR. Variabel banyak ibu dengan paritas 1 dan banyak ibu yang mengalami komplikasi tidak signifikan di kabupaten/kota mana pun. Sementara di Kabupaten Sikka tidak ada variabel yang signifikan memengaruhi banyak kejadian BBLR.

Tabel 3 Pengelompokan Kabupaten/Kota di Nusa Tenggara Timur berdasarkan Variabel yang Memengaruhi Kejadian BBLR

No	Kabupaten/Kota	Variabel yang Berpengaruh
1	Sumba Barat, Sumba Timur, Ngada, Manggarai, Manggarai Barat, Sumba Tengah, Sumba Barat Daya, Nagekeo, Manggarai Timur	X_1, X_5, X_6
2	Kupang, Timor Tengah Selatan, Timor Tengah Utara, Belu, Alor, Kota Kupang	X_2, X_6
3	Lembata, Flores Timur	X_2
4	Ende	X_1
5	Rote Ndao	X_2, X_5, X_6
6	Sabu Raijua	X_1, X_6
7	Sikka	-

Model GWNBR untuk banyak kejadian BBLR di Sumba Barat adalah

$$\ln \mu_1 = 7,540^* - 0,046^* x_{11} - 0,110 x_{12} + 0,036 x_{13} + 0,021 x_{14} + 0,027^* x_{15} - 0,027^* x_{16} \quad (8)$$

Artinya penambahan satu orang ibu bekerja akan mengurangi resiko BBLR menjadi $e^{-0,046} = 0,955$ kali jika variabel lain dalam model tetap. Penambahan satu orang ibu yang mengonsumsi zat besi sesuai aturan akan meningkatkan resiko BBLR $e^{0,027} = 1,024$ kali jika variabel lain dalam model tetap. Sedangkan penambahan satu orang ibu yang memeriksakan kehamilan dengan indikator K4 akan menurunkan resiko BBLR $e^{-0,027} = 0,974$ kali jika variabel lain dalam model tetap. Meskipun model pada persamaan (8) ditulis lengkap hanya variabel X_1, X_5 dan X_6 yaitu banyak ibu yang bekerja, banyak ibu yang mengonsumsi zat besi sesuai aturan, dan banyak ibu yang memeriksakan kehamilan dengan indikator K4 yang berpengaruh signifikan terhadap kejadian BBLR di Sumba Barat.

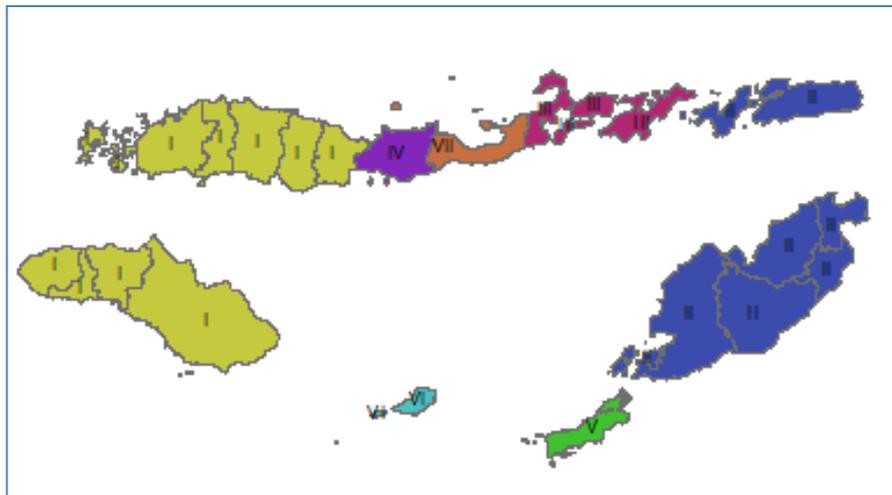
Model GWNBR untuk banyak kejadian BBLR di Kota Kupang adalah

$$\ln \mu_{21} = 6,791^* - 0,023 x_{211} - 0,270^* x_{212} + 0,049 x_{213}$$

$$+ 0,060x_{21\ 4} - 0,001x_{21\ 5} - 0,022*x_{21\ 6}$$

Artinya penambahan satu orang ibu melahirkan berusia kurang dari 20 tahun akan mengurangi resiko BBLR menjadi $e^{-0,270} = 0,763$ kali jika variabel lain dalam model tetap. Sedangkan penambahan satu orang ibu yang memeriksakan kehamilan dengan indikator K4 akan menurunkan resiko BBLR $e^{-0,022} = 0,978$ kali jika variabel lain dalam model tetap.

Hasil pengelompokan kabupaten/kota di Nusa Tenggara Timur berdasarkan variabel yang mempengaruhi kejadian BBLR sesuai dengan hukum pertama Tobler tentang geografi. Dalam Gambar 1 terlihat kabupaten/kota yang berdekatan berada dalam kelompok yang sama.



Gambar 1 Pengelompokan Kabupaten/Kota di Nusa Tenggara Timur Berdasarkan Variabel yang Memengaruhi BBLR

Berdasarkan statistik Wald, hipotesis $H_0: \alpha = 0$ ditolak. Artinya parameter overdispersi global pada model GWNBRg signifikan.

Kesesuaian model GWNBRg yang diperoleh diuji dengan hipotesis H_0 : model fit melawan H_1 : model tidak *fit* menggunakan statistik devians. Devians untuk model GWNBRg adalah 19,735. Nilai ini lebih kecil daripada nilai $\chi^2_{(0,1;14)}$ sebesar 21,064. Hal ini berarti model yang diperoleh sesuai dengan data.

3. Kesimpulan

Berdasarkan hasil penelitian diperoleh kesimpulan berikut:

1. Model terbaik yang menjelaskan determinan bayi dengan berat badan lahir rendah di Nusa Tenggara Timur adalah model GWNBRg yaitu model GWNBR dengan parameter overdispersi bersifat global. Nilai AIC model ini lebih kecil dibandingkan model GWNBR maupun GWPR.
2. Faktor penyebab BBLR bervariasi antar kabupaten/kota sehingga berdasarkan variabel yang berpengaruh signifikan, model GWNBRg mengelompokkan kabupaten/kota di Nusa Tenggara Timur menjadi 7 kelompok.
3. Banyak ibu dengan paritas 1 dan banyak ibu yang mengalami komplikasi tidak signifikan di kabupaten/kota mana pun.
4. Tidak ada variabel yang digunakan yang signifikan memengaruhi banyak kejadian BBLR di Kabupaten Sikka.

Referensi

- [1] Charlton, M. and Fotheringham, A.S., 2009, *Geographically Weighted Regression: White Paper*. Irlandia.
- [2] da Silva, A.R. and Rodrigues, T.C.V., 2013, Geographically Weighted Negative Binomial Regression—incorporating overdispersion, *Journal Statistics and Computing* DOI10.1007/s11222-013-9401-9. Springer US.
- [3] Fotheringham, A.S., Brunson, C., Charlton, M., 2002, *Geographically Weighted Regression, The Analysis of Spatially Varying Relationships*, Chichester: John Wiley & Sons.
- [4] Hilbe, J.M., 2011, *Negative Binomial Regression*, Second Edition, Cambridge: Cambridge University Press.
- [5] Lusi Eka Afri, 2013, Model Regresi Binomial Negatif Terboboti Geografis untuk Data Kematian Bayi, *Jurnal Ilmiah Edu Research* Vol. 2 No. 1: 15–26.
- [6] Nakaya, T, Fotheringham, A.S., Brunson, C., Charlton, M., 2005, Geographically Weighted Poisson Regression for Disease Association Mapping, *Statistics in Medicine*; 24: 2695–2717.
- [7] Nina Haryati, Winita Sulandari, Muslich, 2012, Analisis Faktor-faktor yang Mempengaruhi Berat Badan Bayi Saat Lahir di Kota Surakarta Menggunakan Metode Pohon Regresi, *Prosiding Seminar Nasional Matematika*: 189–195.
- [8] Ohlsson, A. and Shah, P, 2008, *Determinants and Prevention of Low Birth Weight: A Synopsis of the Evidence*, Alberta: IHE Report.
- [9] United Nations Children’s Fund and World Health Organization, 2004, *Low Birthweight: Country, Regional and Global Estimates*, New York: UNICEF.

ANALISIS CLUSTER MENGGUNAKAN METODE FUZZY C-MEANS BERDASARKAN SIMPLE RANDOM SAMPLING

ESTI RAMADITIA MULATSIH¹, IDA FITHRIANI², SITI
NURROHMAH³

1. Departemen Matematika, FMIPA, Universitas Indonesia, Kampus Baru UI, Depok 16424, Indonesia, esti.r@sci.ui.ac.id
2. Departemen Matematika, FMIPA, Universitas Indonesia, Kampus Baru UI, Depok 16424, Indonesia, ida.fithriani@gmail.com
3. Departemen Matematika, FMIPA, Universitas Indonesia, Kampus Baru UI, Depok 16424, Indonesia, snurrohmah@sci.ui.ac.id

Abstrak.. Analisis *cluster* merupakan teknik multivariat yang digunakan untuk mengelompokkan objek berdasarkan karakteristik yang dimilikinya. Salah satu teknik dalam analisis *cluster* adalah metode *Fuzzy K-Means* (lebih dikenal dengan *Fuzzy C-Means*), yang merupakan versi *fuzzy* dari metode *K-Means clustering*. Seperti pada metode *K-Means*, *Fuzzy C-Means* (FCM) juga sangat sensitive terhadap penentuan pusat-pusat awal *cluster*. Dalam paper ini, kami mengusulkan modifikasi dari metode FCM dengan menggunakan metode *sampling* dengan probabilitas. Metode *sampling* digunakan untuk menaksir lokasi pusat-pusat awal *cluster* untuk digunakan ke dalam proses *clustering*. Metode *sampling* yang digunakan adalah *simple random sampling* yang merupakan metode pengambilan sampel yang paling sederhana. Modifikasi dari metode FCM dengan menggunakan metode *simple random sampling* disebut dengan *simple random sampling-Fuzzy C-Means*. Pendekatan *clustering* akan diuji pada himpunan data kanker payudara Wisconsin. Hasil eksperimen menunjukkan bahwa *Fuzzy C-Means* berdasarkan *simple random sampling* lebih efisien dibandingkan metode *Fuzzy C-Means*.

Kata kunci: Analisis *cluster*, *Fuzzy C-Means*, *Simple Random Sampling*.

1. Pendahuluan

Menurut Sharma [1], analisis *cluster* merupakan suatu teknik yang digunakan untuk menggabungkan objek-objek ke dalam kelompok atau *cluster* berdasarkan kesamaan karakteristik diantara objek-objek dalam data, sedemikian sehingga objek-objek yang berada dalam masing-masing *cluster* akan mirip satu sama lain, dan berbeda dengan objek-objek dalam *cluster* lain.

Secara umum, teknik dalam analisis *cluster* dapat dibagi menjadi dua, yaitu *clustering* hirarki dan *clustering* non-hirarki. Salah satu metode *clustering* non hirarki adalah *K-Means clustering*. *K-Means* merupakan suatu metode *clustering* yang membagi titik data atau objek-objek ke dalam *Kcluster*, dimana setiap *cluster* diwakili oleh suatu pusat *cluster* yang disebut *centroid*. Pada metode ini, suatu titik data akan menjadi

anggota dari satu dan hanya satu *cluster*. Namun, terkadang suatu objek tidak dapat ditempatkan tepat pada satu *cluster*, karena sebenarnya objek tersebut terletak di antara dua atau lebih *cluster* yang lain. Dengan memaksakan suatu objek untuk sepenuhnya menjadi anggota suatu *cluster*, akan diperoleh hasil *clustering* yang kurang alami. Untuk mengatasi permasalahan tersebut, pada tahun 1973 Dunn [2] mengusulkan suatu metode pengelompokan data yang memungkinkan suatu titik data berada pada lebih dari satu *cluster*. Metode ini merupakan metode *Fuzzy K-Means* yang lebih dikenal dengan *Fuzzy C-Means*.

Fuzzy C-Means (FCM) merupakan pengembangan dari metode *K-Means* dengan memanfaatkan teori himpunan *fuzzy* dalam mengelompokkan suatu titik data ke dalam suatu *cluster*. Seperti pada metode *K-Means*, FCM juga sensitive terhadap penentuan pusat-pusat awal *cluster*. Untuk mengatasi permasalahan tersebut, dalam skripsi ini diusulkan modifikasi FCM menggunakan metode *sampling* dengan probabilitas.

Metode *sampling* yang digunakan adalah metode *simple random sampling* (SRS), yang merupakan metode pengambilan sampel yang paling sederhana. Modifikasi dari metode FCM dengan menggunakan metode SRS disebut dengan *Simple random sampling-Fuzzy C-Means*.

Dalam makalah ini, pertama akan dibahas mengenai *Fuzzy C-Means* beserta permasalahan pemilihan pusat *cluster* awal yang digunakan, kemudian dilanjutkan dengan penjelasan mengenai proses pengelompokan FCM berdasarkan *Simple Random*. Terakhir, kedua metode *clustering*, yaitu FCM dan *Simple random sampling-Fuzzy C-Means* akan diuji pada himpunan data pasien kanker payudara di Rumah Sakit Univeritas Wisconsin.

2. Hasil – Hasil Utama

Analisis *cluster* merupakan suatu teknik yang digunakan untuk menggabungkan objek-objek ke dalam kelompok atau *cluster* sedemikian sehingga objek-objek dalam masing-masing kelompok akan mirip satu sama lain dan berbeda dengan objek-objek dalam kelompok lain [1]. Secara umum, analisis *cluster* dapat dibagi menjadi dua, yakni hirarki dan non hirarki. Salah satu metode *clustering* non hirarki adalah *Fuzzy C-Means*.

Metode *Fuzzy C-Means* (FCM) merupakan versi *fuzzy* dari metode *K-Means clustering*. Metode ini pertama kali diperkenalkan oleh Ruspini dan Dunn pada tahun 1973 [2], dan kemudian dikembangkan oleh Bezdek pada tahun 1981 [3]. Berbeda dengan metode *K-Means*, Metode FCM memungkinkan sebuah objek untuk memiliki berbagai derajat keanggotaan ke dalam beberapa *cluster*. Derajat keanggotaan dari objek ke- i , x_i dalam *cluster* ke- c adalah u_{ic} . Derajat keanggotaan ini harus memenuhi kondisi berikut:

$$u_{ic} \in [0,1], 1 \leq i \leq N, 1 \leq c \leq C \quad (2.1)$$

$$\sum_{c=1}^C u_{ic} = 1, 1 \leq i \leq N \quad (2.2)$$

$$\sum_{i=1}^N u_{ic} > 0, 1 \leq c \leq C \quad (2.3)$$

dengan N adalah jumlah objek-objek dalam himpunan data X dan C adalah jumlah *cluster*.

Metode FCM menghasilkan suatu himpunan yang elemennya adalah pusat-pusat *cluster* dengan meminimumkan fungsi tujuan sebagai berikut:

$$J = \sum_{c=1}^C \sum_{i=1}^N (u_{ic})^a d^2(\mathbf{x}_i, \mathbf{m}_c) \quad (2.4)$$

dengan \mathbf{x}_i merupakan titik data ke- i dari \mathbf{X} , C menyatakan banyaknya *cluster* yang dibentuk, \mathbf{m}_c menyatakan pusat *cluster* ke- c , u_{ic} menyatakan derajat keanggotaan dari titik data ke- i , \mathbf{x}_i , untuk *cluster* ke- c , $d^2(\mathbf{x}_i, \mathbf{m}_c)$ menyatakan kuadrat jarak Euclidean antara \mathbf{x}_i dan \mathbf{m}_c , dan a menyatakan *weighting exponent* dari derajat keanggotaan, $1 < a < \infty$. Nilai a yang sering digunakan adalah 2.

Melalui optimisasi, fungsi tujuan FCM dapat diminimumkan terhadap u_{ic} dan \mathbf{m}_c . Nilai dari derajat keanggotaan u_{ic} harus memenuhi kondisi yang diberikan pada persamaan (2.1) – (2.3). Derajat keanggotaan dari titik data ke- i , \mathbf{x}_i , untuk *cluster* ke- c dan pusat *cluster* untuk *cluster* ke- c adalah sebagai berikut:

$$u_{ic} = \left(\sum_{h=1}^C \left(\frac{d(\mathbf{x}_i, \mathbf{m}_c)}{d(\mathbf{x}_i, \mathbf{m}_h)} \right)^{\frac{2}{a-1}} \right)^{-1} \quad (2.5)$$

$$\mathbf{m}_c = \frac{\sum_{i=1}^N (u_{ic})^a \mathbf{x}_i}{\sum_{i=1}^N (u_{ic})^a} \quad (2.6)$$

Algoritma Fuzzy C-Means

1. Tentukan jumlah *cluster* C , a , MaksimumIterasi, dan ε (batas *error*).
2. Pilih secara acak C titik data dari himpunan data yang diberikan untuk dijadikan sebagai pusat-pusat awal *cluster*.
3. Hitung derajat keanggotaan menggunakan persamaan (2.5).
4. Perbaiki pusat-pusat *cluster* menggunakan persamaan (2.6).
5. Apabila $\|M^t - M^{t-1}\| \leq \varepsilon$ maka *stop*, sebaliknya kembali ke langkah 3.

Metode FCM merupakan metode *clustering* yang sederhana dan mudah dipahami. Namun, seperti pada metode *K-Means*, hasil *clustering* yang diperoleh dari metode FCM juga sangat tergantung pada inisialisasi pusat awal *cluster* yang diberikan. Metode yang biasanya digunakan untuk mengatasi permasalahan tersebut adalah dengan melakukan *running* FCM berkali-kali dengan pusat-pusat awal *cluster* yang berbeda-beda pada himpunan data yang sama. Meskipun sederhana, metode ini memiliki kelemahan, yaitu tidak ada batasan yang jelas berapa kali inisialisasi pusat-pusat awal *cluster* yang diperlukan.

Pada makalah ini, kami mengusulkan modifikasi dari metode FCM dengan menggunakan metode *sampling* dengan probabilitas. Metode *sampling* yang digunakan adalah metode *simple random sampling* (SRS),

dimana metode ini merupakan metode pengambilan sampel yang paling sederhana dan paling sering digunakan. Modifikasi dari metode FCM dengan menggunakan metode SRS disebut dengan *Simple random sampling-Fuzzy C-Means*.

Metode *Simple random sampling-Fuzzy C-Means* diawali dengan mengambil sampel acak sederhana dari himpunan data yang diberikan. Sampel yang diperoleh kemudian dikelompokkan ke dalam C cluster menggunakan metode FCM. Pusat-pusat cluster yang diperoleh dari pengelompokkan data sampel akan dijadikan sebagai pusat-pusat awal cluster untuk mengelompokkan himpunan data yang diberikan secara keseluruhan.

Berdasarkan penjelasan diatas, dapat dikatakan bahwa metode SRS-*Fuzzy C-Means* merupakan algoritma dua langkah. Langkah pertama disebut sebagai langkah estimasi (*E-step*) dan langkah kedua disebut sebagai langkah *clustering* (*C-step*). Pada langkah estimasi, himpunan data \mathbf{X}_{SRS} dikelompokkan ke dalam C cluster dengan menggunakan metode FCM. Dengan demikian, fungsi tujuan dari metode SRS-*Fuzzy C-Means* pada langkah estimasi dapat didefinisikan sebagai berikut:

$$JE_{SRS} = \sum_{c=1}^C \sum_{i=1}^n (u_{ic})^a d^2(\mathbf{x}_i, \mathbf{m}_c) \quad (2.7)$$

dengan \mathbf{x}_i menyatakan titik data ke- i dari sampel \mathbf{X}_{SRS} , dan $C, u_{ic}, a, \mathbf{m}_c, a$, dan $d^2(\mathbf{x}_i, \mathbf{m}_c)$ masing-masing telah didefinisikan sebelumnya. Melalui optimisasi, JE_{SRS} dapat diminimumkan terhadap u_{ic} dan \mathbf{m}_c . Nilai derajat keanggotaan u_{ic} harus memenuhi kondisi yang diberikan pada persamaan (3.1)-(3.3). Derajat keanggotaan titik data ke- i untuk cluster ke- c dan pusat cluster untuk cluster ke- c adalah sebagai berikut:

$$u_{ic} = \left(\sum_{h=1}^C \left(\frac{d(x_i, m_c)}{d(x_i, m_h)} \right)^{\frac{2}{a-1}} \right)^{-1} \quad (2.8)$$

$$\mathbf{m}_c = \frac{\sum_{i=1}^n (u_{ic})^a \mathbf{x}_i}{\sum_{i=1}^n (u_{ic})^a} \quad (2.9)$$

Pada langkah *clustering*, C pusat cluster yang diperoleh dari pengelompokkan himpunan data \mathbf{X}_{SRS} akan dijadikan sebagai pusat-pusat awal cluster untuk mengelompokkan keseluruhan titik data dari \mathbf{X} ke dalam C cluster.

Algoritma Simple random sampling-Fuzzy C Means

E-step

1. Input \mathbf{X}_{SRS}
2. Tentukan jumlah cluster C, a , Maksimum Iterasi, dan ε .
3. Pilih secara acak C titik data dari \mathbf{X}_{SRS} untuk dijadikan sebagai pusat-pusat awal cluster.
4. Hitung derajat keanggotaan menggunakan persamaan (2.8).
5. Perbaiki pusat-pusat cluster menggunakan persamaan (2.9).
6. Apabila $\|M^t - M^{t-1}\| \leq \varepsilon$ maka *stop*, sebaliknya kembali ke langkah 3.

C-step

1. Input X
2. Input C pusat awal *cluster* hasil *E-step*.
3. Tentukan jumlah *cluster* C , a , Maksimum Iterasi, dan ϵ .
4. Hitung derajat keanggotaan menggunakan persamaan (2.5).
5. Perbaiki pusat-pusat *cluster* menggunakan persamaan (2.6).
6. Apabila $\|M^t - M^{t-1}\| \leq \epsilon$ maka *stop*, sebaliknya kembali ke langkah 4.

Pada makalah ini, metode *Simple-random sampling Fuzzy C-Means* akan diterapkan pada himpunan data pasien kanker payudara di Rumah Sakit Universitas Wisconsin. Himpunan data ini diperoleh dari mesin pembelajaran (*machine learning*) *repository* Universitas California. Himpunan data pasien kanker payudara di Rumah sakit Universitas California terdiri dari 699 titik data yang ditandai dengan sembilan fitur untuk mengklasifikasikan titik data menjadi dua kelompok, yakni *benign* dan *malignant*. Pendekatan *clustering* kami dibandingkan dengan *Fuzzy C-Means* standar.

Kedua teknik *clustering*, *Simple random sampling- Fuzzy C-Means* dan *Fuzzy C-Means* dilakukan dengan menggunakan pusat-pusat awal *cluster* yang sama. Hasil *clustering* yang diperoleh dari kedua metode tersebut kemudian dievaluasi berdasarkan uji validitas *cluster*. Uji validitas *cluster* yang digunakan pada makalah ini adalah uji validitas *cluster* indeks Davies Bouldin.

Indeks Davies Bouldin (DBI) diperkenalkan oleh David L. Davies dan Donald W Bouldin pada tahun 1979 [4], dan digunakan untuk mengevaluasi algoritma *clustering*. Penggunaan indeks validitas Davies Bouldin dapat mengidentifikasi *cluster* yang padat (*compact cluster*) dan terpisah dengan baik terhadap *cluster* yang lain (*Well-Separated*) seperti halnya pada indeks validitas Dunn. Indeks validitas Davies Bouldin dapat dihitung dengan menggunakan persamaan berikut:

$$DBI = \frac{1}{C} \sum_{c=1}^C \max_{i \neq j} \left(\frac{\gamma(m_i) + \gamma(m_j)}{\theta(m_i, m_j)} \right)$$

dengan m_i adalah pusat *cluster* ke- i , $\gamma(m_i)$ adalah rata-rata jarak dari seluruh titik dalam *cluster* ke- i ke pusat *cluster* m_i , $\theta(m_i, m_j)$ adalah jarak antara dua pusat *cluster* m_i dan m_j , dan C adalah jumlah *cluster*.

Skema *clustering* yang optimal menurut indeks Davies Bouldin ialah yang memiliki skore indeks Davies Bouldin terkecil [5].

Tabel 2.1 Hasil Evaluasi Kinerja *Fuzzy C-Means* dan SRS *Fuzzy C-Means*

Algoritma	Indeks Davies-Bouldin			
	n=105	n=210	n=350	n=699
<i>Fuzzy C-Means</i>				0.55121
SRS <i>Fuzzy C-Means</i>	0.56785	0.56174	0.56134	0.55121

Tabel 2.1 diatas, menampilkan hasil evaluasi kinerja kedua algoritma *clustering*, yakni *Fuzzy C-Means*, dan *Simple random sampling-Fuzzy C-Means* yang diterapkan pada himpunan data kanker payudara di Rumah sakit Universitas Wisconsin. Nilai indeks Davies-Bouldin yang ditampilkan pada tabel ini merupakan nilai rata-rata indeks Davies-Bouldin yang diperoleh dengan melakukan dua puluh kali *run* kedua metode *clustering* tersebut, dengan menggunakan pusat-pusat awal *cluster* yang berbeda pada himpunan data yang sama.

3. Kesimpulan

Dari pembahasan di atas, dapat disimpulkan bahwa untuk penerapan pada data yang sama, metode *Simple random sampling-Fuzzy C-Means* lebih efisien dibandingkan dengan metode FCM, dimana pada metode FCM *running* dilakukan berkali-kali dengan pusat-pusat awal *cluster* yang berbeda-beda.

Referensi

- [1] Sharma, S. (1996). Applied Multivariate Techniques. *John Wiley and Sons, Inc.*
- [2] Dunn, J.C.: A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters. *J. Cybernet*, Vol. 3, 1973, pp. 32–57
- [3] Bezdek, J.C. (1981) Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press, New York. <http://dx.doi.org/10.1007/978-1-4757-0450-1>.
- [4] Davies, David L.; Bouldin, Donald W. (1979). "A Cluster Separation Measure". *IEEE Transactions on Pattern Analysis and Machine Intelligence*. *PAMI-1* (2): 224–227. [doi:10.1109/TPAMI.1979.4766909](https://doi.org/10.1109/TPAMI.1979.4766909).
- [5] Salazar GEJ, Veles AC, Parra MCM, Ortega LO. 2002. A Cluster Validity Index for Comparing Non-hierarchical Clustering Methods.

DISTRIBUSI KUMARSWAMY-GEOMETRIK

DEANA ROSAN¹, SITI NURROHMAH², IDA FITHRIANI³

1. Departemen Matematika, FMIPA, Universitas Indonesia, Kampus Baru UI, Depok 16424, Indonesia, deanarosan@sci.ui.ac.id
2. Departemen Matematika, FMIPA, Universitas Indonesia, Kampus Baru UI, Depok 16424, Indonesia, snurrohmah@sci.ui.ac.id
3. Departemen Matematika, FMIPA, Universitas Indonesia, Kampus Baru UI, Depok 16424, Indonesia, ida.fithriani@gmail.com

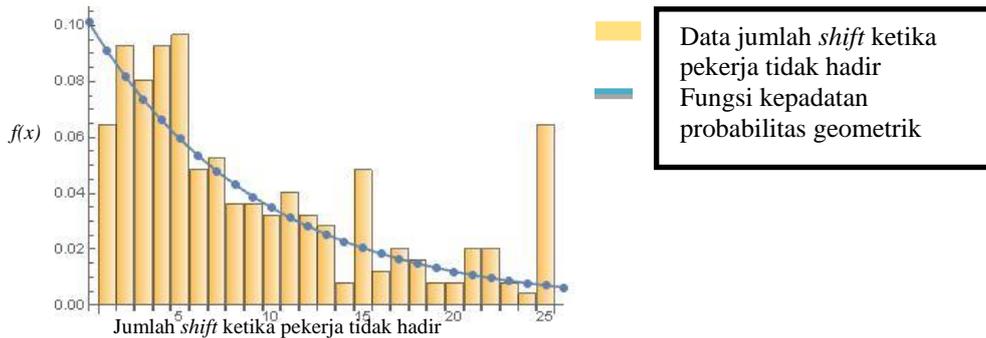
Abstrak. Makalah ini membahas tentang distribusi Kumaraswamy-geometrik yang merupakan distribusi probabilitas dari peubah acak diskrit yang dibangun dengan menggunakan metode *Transformed-Transformer*. Distribusi Kumaraswamy dapat membuat distribusi geometrik menjadi lebih fleksibel. Pembahasan meliputi fungsi distribusi, fungsi kepadatan probabilitas, perilaku limit, serta kasus khusus dari distribusi Kumaraswamy-geometrik. Karakteristik-karakteristik dari distribusi Kumaraswamy-geometrik yang meliputi modus, persentil, momen, fungsi pembangkit momen, dan fungsi pembangkit probabilitas juga akan dibahas pada makalah ini. Selanjutnya, Metode Maksimum *Likelihood* digunakan dalam makalah ini untuk mencari penaksir parameter dari distribusi Kumaraswamy-geometrik. Pada bagian akhir, akan digunakan data tentang jumlah klaim suatu asuransi kendaraan bermotor sebagai ilustrasi penggunaan distribusi Kumaraswamy-geometrik.

Kata kunci: Distribusi Kumaraswamy-geometrik, metode *Transformed-Transformer*, metode maksimum *Likelihood*.

1. Pendahuluan

Suatu perusahaan atau industri dapat menetapkan *lay off* pada pekerjanya. *Lay off* dapat berarti penangguhan sementara atau pemberhentian selamanya. Contoh industri yang menerapkan *lay off* adalah suatu industri baja di Afrika Selatan. Pekerja yang terkena *lay off* berarti pekerja tersebut tidak hadir ditempat kerja. Ketidakhadiran pekerja tersebut dapat membawa dampak negatif, salah satunya target pembangunan yang harus dicapai industri menjadi terhambat karena kekurangan pekerja. Oleh karena itu, permasalahan ini harus segera diatasi, yaitu dengan cara memodelkan permasalahan ini dengan distribusi probabilitas.

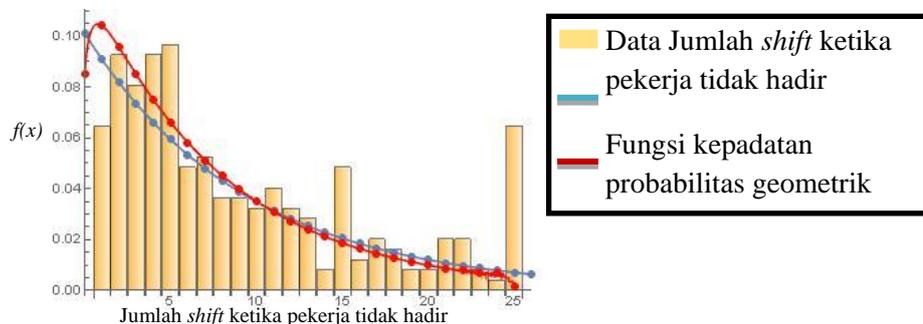
Permasalahan ketidakhadiran pekerja karena *lay off* tersebut dapat dihubungkan dengan percobaan Bernoulli. Percobaan Bernoulli yang berhubungan dengan kasus ini memiliki hasil sukses yaitu *shift* ketika pekerja hadir dan hasil gagal yaitu *shift* ketika pekerja tidak hadir. Kemudian akan dihitung jumlah *shift* ketika pekerja tidak hadir sampai ditemukan *shift* ketika pekerja tersebut hadir untuk yang pertama kali dalam periode *lay off*. Distribusi probabilitas yang digunakan pada masalah tersebut adalah distribusi geometrik.



Gambar 1: Grafik fungsi kepadatan probabilitas empiris data jumlah *shift* ketika pekerja tidak hadir dan fungsi kepadatan probabilitas distribusi geometrik

Pada Gambar 1 terdapat fungsi kepadatan probabilitas empiris data jumlah *shift* ketika pekerja tidak hadir tersebut dan grafik fungsi kepadatan probabilitas distribusi geometrik. Namun, pada gambar tersebut terlihat bahwa ternyata masih banyak observasi dari data jumlah *shift* tersebut yang tidak dapat diakomodir oleh distribusi geometrik, sehingga penggunaan distribusi geometrik saja masih kurang fleksibel dalam memodelkan data tersebut. Berdasarkan hal tersebut, dilakukan perluasan terhadap distribusi geometrik dengan cara membentuk distribusi baru yang didasari oleh distribusi geometrik. Salah satu metode yang dapat digunakan untuk membentuk suatu distribusi baru yaitu metode *Transformed-Transformer*.

Metode *Transformed-Transformer* mengkomposisikan dua peubah acak, yaitu peubah acak T yang bertindak sebagai *transformed* dan peubah acak X yang bertindak sebagai *transformer*. Dalam kasus ini, distribusi yang digunakan untuk peubah acak X adalah distribusi geometrik. Sedangkan, distribusi yang diajukan untuk peubah acak T adalah distribusi Kumaraswamy. Distribusi Kumaraswamy dipilih karena merupakan distribusi yang fleksibel untuk berbagai macam data, contohnya distribusi Kumaraswamy dapat menggambarkan data yang memiliki grafik fungsi kepadatan probabilitas yang berbentuk huruf J, L ataupun huruf U. Jika distribusi Kumaraswamy digunakan sebagai distribusi untuk peubah acak T dan jika X merupakan peubah acak yang berdistribusi geometrik, maka dengan metode *Transformed-Transformer* akan menghasilkan distribusi Kumaraswamy-geometrik.



Gambar 2: Grafik fungsi probabilitas empiris data jumlah *shift* ketika pekerja tidak hadir, fungsi kepadatan probabilitas geometrik, dan fungsi kepadatan probabilitas Kumaraswamy-geometrik

Berdasarkan Gambar 2 terlihat lebih banyak observasi dari data jumlah *shift* ketika pekerja tidak hadir tersebut yang bisa diakomodir oleh distribusi

Kumaraswamy-geometrik dibandingkan dengan distribusi geometrik, sehingga distribusi Kumaraswamy-geometrik memodelkan data tersebut lebih baik dari distribusi geometrik. Dengan demikian, distribusi Kumaraswamy-geometrik dinyatakan lebih fleksibel dari distribusi geometrik.

Dalam makalah ini akan dibahas mengenai proses pembentukan distribusi Kumaraswamy-geometrik dan karakteristik-karakteristik dari distribusi tersebut. Pada akhir pembahasan akan dibahas ilustrasi data menggunakan distribusi Kumaraswamy-geometrik dan dilakukan perbandingannya dengan distribusi geometrik.

2. Hasil – Hasil Utama

2.1 Metode

Berikut beberapa metode yang digunakan dalam makalah ini

1. Metode *Transformed-Transformer*

Metode ini digunakan untuk membangun suatu distribusi baru dengan cara mengkomposisikan dua peubah acak. Misalkan peubah acak yang akan dikomposisikan yaitu peubah acak T dan peubah acak X , maka melalui metode *Transformed-Transformer* akan dihasilkan keluarga $T-X$ dan distribusi yang dihasilkan yaitu keluarga distribusi $T-X$. Pada metode ini, T bertindak sebagai peubah acak yang ditransformasi sedangkan X bertindak sebagai peubah acak yang mentransformasi.

Misalkan $r(t)$ adalah *probability density function* (pdf) dari peubah acak T di mana $t \in [a, b]$, dengan $-\infty < a < b < \infty$. Misalkan $W(F(x))$ adalah fungsi dari *cumulative density function* (cdf) peubah acak X atau $F(x)$ sedemikian sehingga $W(F(x))$ memenuhi kondisi berikut.

- $W(F(x)) \in [a, b]$,
- $W(F(x))$ terturunkan dan monoton tak turun,
- $W(F(x)) \rightarrow a$ jika $x \rightarrow -\infty$ dan $W(F(x)) \rightarrow b$ jika $x \rightarrow \infty$.



Misalkan X adalah suatu peubah acak dengan pdf $f(x)$ dan cdf $F(x)$. Misalkan T adalah suatu peubah acak kontinu dengan pdf $r(t)$ yang terdefinisi pada $[a, b]$ dan cdf $R(t)$. Cdf untuk distribusi dari keluarga $T-X$ didefinisikan sebagai berikut.

$$G(x) = \int_a^{W(F(x))} r(t) dt. \\ = R\{W(F(x))\} = (R \circ W \circ F)(x), \quad (2)$$

dimana $W(F(x))$ memenuhi kondisi (1). Peubah acak X bisa merupakan peubah acak diskrit dan jika X adalah peubah acak diskrit, maka akan dihasilkan keluarga diskrit $T-X$. Jika X merupakan peubah acak diskrit, maka pdf untuk distribusi dari keluarga $T-X$ yaitu sebagai berikut.

$$g(x) = G(x) - G(x - 1) = R(W(F(x))) - R(W(F(x - 1))). \quad (3)$$

2. Metode Maksimum Likelihood

Misalkan sampel acak ukuran n diambil dari suatu distribusi diskrit yang memiliki pdf $f(x; \theta)$, di mana θ adalah parameter yang tidak diketahui. Sampel acak ini memiliki frekuensi observasi n_x , $x = 0, 1, 2, \dots, k$, di mana $\sum_{x=0}^k n_x = n$.

Maka fungsi *Likelihood* untuk kasus ini adalah sebagai berikut

$$L(\theta; x) = \prod_{x=0}^k [f(x; \theta)]^{n_x}$$

dan fungsi *loglikelihood* yaitu

$$\ln[L(\theta; x)] = \sum_{x=0}^k n_x \ln f(x; \theta)$$

Kemudian, untuk mencari nilai taksiran dari parameter dapat dilakukan dengan melakukan turunan pertama dari $\ln(L(\theta))$ terhadap θ dan menyamakannya dengan nol. Sehingga, dengan menyelesaikan persamaan tersebut, akan diperoleh $\hat{\theta}$ atau taksiran dari θ .

3. Uji Kolmogorov-Smirnov

Sebelum membahas uji Kolmogorov – Smirnov, akan diberikan definisi mengenai fungsi distribusi empiris yang dinyatakan sebagai berikut

$$F_n(x) = \frac{\text{banyaknya observasi} \leq x}{n}$$

di mana n adalah total banyaknya observasi.

Uji Kolmogorov – Smirnov berguna dalam menguji kecocokan dugaan distribusi dari data dengan distribusi sebenarnya berlandaskan fungsi distribusi yang diduga dan fungsi distribusi empiris. Prosedurnya menggunakan jarak vertikal maksimum antara dua fungsi distribusi tersebut. Berikut ini merupakan langkah-langkahnya.

- Data
Data terdiri dari sampel acak X_1, X_2, \dots, X_n berukuran n yang berasal dari suatu distribusi dengan fungsi distribusinya $G(x)$ yang tidak diketahui.
 - Hipotesis
Misalkan $G(x)$ adalah fungsi distribusi yang diduga dari data. Maka hipotesisnya adalah
 $H_0: F_n(x) = G(x)$ untuk setiap x ,
 $H_1: F_n(x) \neq G(x)$ untuk paling tidak satu nilai x .
 - Tingkat Signifikansi
Menentukan tingkat signifikansi α
 - Statistik Uji
Misalkan $F_n(x)$ adalah fungsi distribusi empiris. Maka statistik ujinya adalah sebagai berikut
- $$T = \sup_x |G(x) - F_n(x)|$$
- Aturan Keputusan
 H_0 pada tingkat signifikansi α akan ditolak jika T lebih besar dari kuantil ke $(1 - \alpha)$, yaitu $w_{1-\alpha}$.

4. Akaike Information Criteria (AIC)

Pemilihan model terbaik dilakukan ketika terdapat beberapa model distribusi untuk suatu data. Akaike mengusulkan *Akaike Information Criteria* (AIC) berdasarkan data dan mendefinisikannya sebagai berikut

$$AIC = -2\text{Log Likelihood} + 2k$$

di mana k adalah banyaknya parameter dalam model. Model yang terbaik dan paling sesuai dengan data adalah yang memiliki nilai AIC yang terkecil.

2.2 Pembahasan

Distribusi Kumaraswamy-geometrik dibangun dengan metode *Transformed-*

Transformer jika dimisalkan peubah acak T berdistribusi Kumraswamy dan peubah acak X berdistribusi geometrik dan dengan menggunakan $W(F(x)) = F(x)$.

Jika T adalah peubah acak kontinu yang berdistribusi Kumaraswamy dengan parameter α dan β , maka cdf dari T adalah

$$R(t) = 1 - (1 - t^\alpha)^\beta, \quad 0 < t < 1 \quad (4)$$

Dan pdf dari T adalah

$$r(t) = \alpha\beta t^{\alpha-1}(1 - t^\alpha)^{\beta-1}, \quad 0 < t < 1 \quad (5)$$

$$= 0, \quad t \text{ lain,}$$

di mana nilai $\alpha, \beta > 0$.

Misalkan peubah acak X berdistribusi geometrik dengan parameter p , maka pdf dari X adalah

$$f(x) = pq^x, \quad x = 0,1,2, \dots, \quad (6)$$

$$= 0, \quad x \text{ lain.}$$

di mana $p + q = 1, \quad 0 < p < 1, \quad 0 < q < 1$.

dan cdf dari X adalah

$$F(x) = 1 - q^{x+1}, \quad x = 0,1,2, \dots \quad (7)$$

Dengan menggunakan persamaan (2), dimana peubah acak T berdistribusi Kumaraswamy dengan cdf pada persamaan (4) dan peubah acak X berdistribusi geometrik dengan cdf pada persamaan (7) dan menggunakan $W(F(x)) = F(x)$, maka diperoleh

$$G(x) = R(W(F(x))) = R(F(x))$$

$$= R(1 - q^{x+1}) = 1 - [1 - (1 - q^{x+1})^\alpha]^\beta, \quad x = 0,1,2, \dots \quad (8)$$

dan dengan menggunakan persamaan (3) diperoleh

$$g(x) = R(W(F(x))) - R(W(F(x-1)))$$

$$= R(F(x)) - R(F(x-1))$$

$$= R(1 - q^{x+1}) - R(1 - q^x)$$

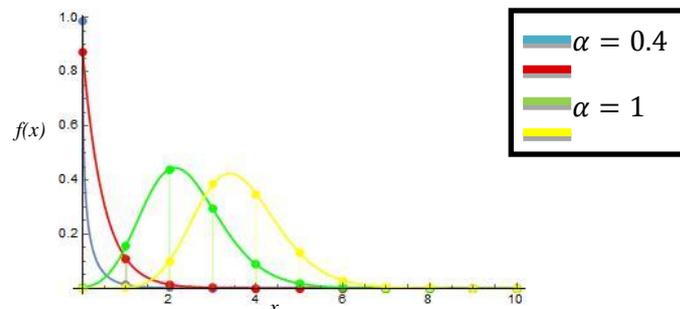
$$= [1 - (1 - q^x)^\alpha]^\beta - [1 - (1 - q^{x+1})^\alpha]^\beta, \quad x = 0,1,2, \dots \quad (9)$$

dimana $\alpha, \beta > 0, p + q = 1, 0 < q < 1$.

Persamaan (8) dan (9) menyatakan cdf dan pdf dari distribusi Kumaraswamy-geometrik (KGD) dengan parameter α, β , dan q . Dengan demikian, peubah acak X tersebut dikatakan memiliki distribusi Kumaraswamy-geometrik dengan parameter α, β , dan q atau secara sederhana dinyatakan dengan $X \sim KGD(\alpha, \beta, q)$.

Analisis grafik pdf KGD untuk beberapa kemungkinan nilai parameter α, β , dan q , yaitu sebagai berikut.

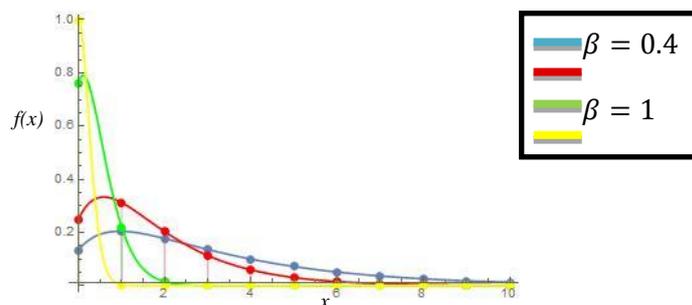
1. Saat nilai parameter α beragam, $\beta = 3$, dan $q = 0.5$



Gambar 3: Grafik pdf distribusi Kumaraswamy-geometrik untuk kemungkinan 1

Dari Gambar 3, dapat dilihat bahwa parameter α merupakan parameter bentuk (*shape*). Hal ini dikarenakan untuk setiap nilai parameter α yang berbeda akan menghasilkan bentuk grafik pdf KGD yang berbeda, di mana parameter lainnya dianggap tetap.

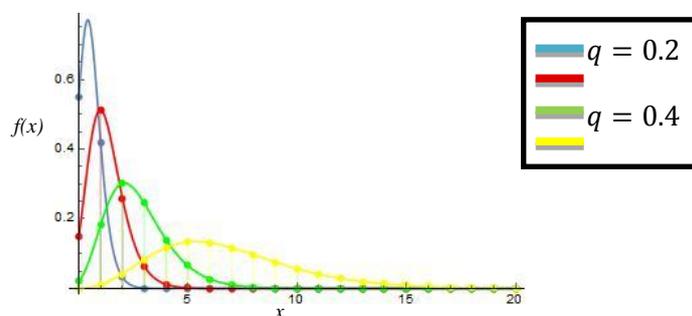
2. Saat nilai parameter β beragam, $\alpha = 2$ dan $q = 0.5$



Gambar 4: Grafik pdf distribusi Kumaraswamy-geometrik untuk kemungkinan 2

Dari Gambar 4, dapat dilihat bahwa parameter β merupakan parameter bentuk (*shape*). Hal ini dikarenakan untuk setiap nilai parameter β yang berbeda akan menghasilkan bentuk grafik pdf yang berbeda. Selain itu, dari Gambar 4 juga diperoleh bahwa KGD bersifat unimodal dan bentuk grafik distribusi tersebut cenderung menurun.

3. Saat nilai parameter q beragam, $\alpha = 5$ dan $\beta = 2$



Gambar 5: Grafik pdf distribusi Kumaraswamy-geometrik untuk kemungkinan 3

Dari Gambar 5, dapat dilihat bahwa parameter q merupakan parameter bentuk (*shape*). Hal ini dikarenakan untuk setiap nilai parameter q yang berbeda akan menghasilkan bentuk grafik pdf distribusi Kumaraswamy-geometrik yang berbeda, di mana parameter lainnya dianggap tetap.

Perilaku limit distribusi Kumaraswamy-geometrik yaitu sebagai berikut.

- a. Limit untuk pdf dari KGD dengan parameter α, β , dan q pada saat x mendekati tak hingga yaitu sebagai berikut

$$\begin{aligned} \lim_{x \rightarrow \infty} g(x) &= \lim_{x \rightarrow \infty} \{ [1 - (1 - q^x)^\alpha]^\beta - [1 - (1 - q^{x+1})^\alpha]^\beta \} \\ &= [1 - (1 - q^\infty)^\alpha]^\beta - [1 - (1 - q^\infty)^\alpha]^\beta = 0 \end{aligned}$$

- b. Limit untuk pdf dari KGD dengan parameter α, β , dan q pada saat q mendekati

satu yaitu sebagai berikut

$$\begin{aligned}\lim_{q \rightarrow 1} g(x) &= \lim_{q \rightarrow 1} \{ [1 - (1 - q^x)^\alpha]^\beta - [1 - (1 - q^{x+1})^\alpha]^\beta \} \\ &= [1 - (1 - 1)^\alpha]^\beta - [1 - (1 - 1)^\alpha]^\beta = 0\end{aligned}$$

- c. Limit untuk pdf dari KGD dengan parameter α, β , dan q pada saat α mendekati tak hingga yaitu sebagai berikut

$$\begin{aligned}\lim_{\alpha \rightarrow \infty} g(x) &= \lim_{\alpha \rightarrow \infty} \{ [1 - (1 - q^x)^\alpha]^\beta - [1 - (1 - q^{x+1})^\alpha]^\beta \} \\ &= [1 - (1 - q^x)^\infty]^\beta - [1 - (1 - q^{x+1})^\infty]^\beta = 0\end{aligned}$$

Karakteristik-karakteristik dari KGD adalah sebagai berikut.

1. Modus

Modus dari KGD merupakan nilai x yang memaksimumkan pdf $g(x)$. Dari hasil analisis pdf KGD telah disimpulkan bahwa KGD bersifat unimodal atau hanya memiliki satu nilai modus untuk berapapun nilai α, β , dan q . Perhatikan grafik pdf dari KGD pada Gambar 3. Dari gambar tersebut, dapat dilihat bahwa modus akan bernilai sama dengan nol ketika $\alpha \leq 1$ dan lebih besar dari nol ketika $\alpha > 1$. Hal ini dapat diperkuat dengan memperhatikan Gambar 5, pada gambar tersebut juga terlihat bahwa untuk nilai $\alpha > 1$ modus akan bernilai lebih besar dari nol. Kesimpulan yang dapat diperoleh yaitu,

- untuk nilai $\alpha \leq 1$ dan sebarang nilai β dan q , modus akan selalu bernilai nol,
- untuk nilai $\alpha > 1$ dan sebarang nilai β dan q , modus pada umumnya akan bernilai lebih besar dari nol.

2. Momen

Momen ke- k dari KGD diperoleh dengan menggunakan rumus untuk momen ke- k , yaitu

$$E[X^k] = \mu'_k = \sum_j x_j^k g(x)$$

sehingga akan diperoleh momen ke- k untuk KGD yaitu sebagai berikut.

$$E[X^k] = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (-1)^{i+j} \frac{\beta^{(i)}}{i!} \frac{(\alpha i)^{(j)}}{j!} (1 - q^j) L_{-k}(q^j)$$

dimana

$$\beta^{(i)} = \beta(\beta - 1)(\beta - 2) \dots (\beta - i + 1),$$

$$(\alpha i)^{(j)} = (\alpha i)(\alpha i - 1)(\alpha i - 2) \dots (\alpha i - j + 1), \text{ dan}$$

$$L_{-k}(q^j) = \sum_{x=0}^{\infty} \frac{q^{jx}}{x^{-k}}$$

$L_{-k}(q^j)$ merupakan fungsi *polylogarithm* yang didefinisikan sebagai berikut

$$L_{-k}(u) = \sum_{x=0}^{\infty} \frac{u^x}{x^{-k}}$$

dengan $u = q^j$. Fungsi *polylogarithm* juga dapat didefinisikan sebagai berikut

$$L_{-n}(z) = \frac{1}{(1-z)^{n+1}} \sum_{k=0}^{n-1} \binom{n}{k} z^{n-k}, \quad n = 1, 2, 3, \dots$$

3. Mean dan Variansi

Untuk mencari mean dan variansi dari KGD akan dicari terlebih dahulu momen pertama dan kedua dari distribusi tersebut.

- Momen pertama dari KGD adalah sebagai berikut.

$$\begin{aligned}\mu'_1 &= E[X] = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (-1)^{i+j} \frac{\beta^{(i)}}{i!} \frac{(\alpha i)^{(j)}}{j!} (1-q^j) L_{-1}(q^j) \\ &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (-1)^{i+j} \frac{\beta^{(i)}}{i!} \frac{(\alpha i)^{(j)}}{j!} \frac{q^j}{(1-q^j)}\end{aligned}\quad (10)$$

- Momen kedua dari KGD yaitu sebagai berikut

$$\begin{aligned}\mu'_2 &= E[X^2] = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (-1)^{i+j} \frac{\beta^{(i)}}{i!} \frac{(\alpha i)^{(j)}}{j!} (1-q^j) L_{-2}(q^j) \\ &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (-1)^{i+j} \frac{\beta^{(i)}}{i!} \frac{(\alpha i)^{(j)}}{j!} \frac{q^j(1+q^j)}{(1-q^j)^2}\end{aligned}$$

Mean dari distribusi Kumarawamy-geometrik adalah momen pertama dari distribusi tersebut, sehingga dari persamaan (10) diperoleh mean yaitu,

$$\mu'_1 = E[X] = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (-1)^{i+j} \frac{\beta^{(i)}}{i!} \frac{(\alpha i)^{(j)}}{j!} \frac{q^j}{(1-q^j)}$$

Variansi untuk KGD yaitu

$$Var(X) = \sigma^2 = E[X^2] - \mu^2$$

$$\sigma^2 = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (-1)^{i+j} \frac{\beta^{(i)}}{i!} \frac{(\alpha i)^{(j)}}{j!} \frac{q^j(1+q^j)}{(1-q^j)^2} - \left[\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (-1)^{i+j} \frac{\beta^{(i)}}{i!} \frac{(\alpha i)^{(j)}}{j!} \frac{q^j}{(1-q^j)} \right]^2$$

Berikut ini akan dicari nilai taksiran dari KGD menggunakan metode Maksimum *Likelihood*. Misalkan sampel acak ukuran n diambil dari KGD dengan frekuensi setiap observasi yaitu n_x , $x = 0, 1, 2, \dots, k$, di mana $\sum_{x=0}^k n_x = n$. Untuk menaksir parameter α, β , dan q akan dicari terlebih dahulu fungsi *likelihood*, yaitu

$$L(x|\alpha, \beta, q) = \prod_{x=0}^k [g(x|\alpha, \beta, q)]^{n_x}$$

sehingga diperoleh fungsi *likelihood* sebagai berikut

$$L(x|\alpha, \beta, q) = \prod_{x=0}^k \{ [1 - (1 - q^x)^\alpha]^\beta - [1 - (1 - q^{x+1})^\alpha]^\beta \}^{n_x}$$

Fungsi tersebut diberi logaritma, sehingga menjadi

$$\ln L(x|\alpha, \beta, q) = \ln \left(\prod_{x=0}^k \{ [1 - (1 - q^x)^\alpha]^\beta - [1 - (1 - q^{x+1})^\alpha]^\beta \}^{n_x} \right)$$

$$l(x|\alpha, \beta, q) =$$

$$n_0 \ln \{ 1 - [1 - (1 - q)^\alpha]^\beta \} + \sum_{x=1}^k n_x \ln \{ [1 - (1 - q^x)^\alpha]^\beta - [1 - (1 - q^{x+1})^\alpha]^\beta \}$$

Nilai parameter yang memaksimumkan $l(x|\alpha, \beta, q)$ diperoleh dengan cara menentukan turunan parsial dari $l(x|\alpha, \beta, q)$ terhadap masing-masing parameter dan hasilnya adalah sebagai berikut.

$$\frac{\partial l(x|\alpha, \beta, q)}{\partial \alpha} = \frac{n_0 \beta (1-q)^\alpha [1 - (1-q)^\alpha]^{\beta-1} \ln(1-q)}{1 - [1 - (1-q)^\alpha]^\beta} + \sum_{x=1}^k \frac{n_x [A_x + B_x]}{[1 - (1-q^x)^\alpha]^\beta - [1 - (1-q^{x+1})^\alpha]^\beta} \quad (11)$$

$$\frac{\partial l(x|\alpha, \beta, q)}{\partial \beta} = \frac{-n_0 [1 - (1-q)^\alpha]^\beta \ln[1 - (1-q)^\alpha]}{1 - [1 - (1-q)^\alpha]^\beta} + \sum_{x=1}^k \frac{n_x [C_x - D_x]}{[1 - (1-q^x)^\alpha]^\beta - [1 - (1-q^{x+1})^\alpha]^\beta} \quad (12)$$

$$\frac{\partial l(x|\alpha, \beta, q)}{\partial q} = \frac{-n_0 \alpha \beta [1 - (1-q)^\alpha]^{\beta-1} (1-q)^{\alpha-1}}{1 - [1 - (1-q)^\alpha]^\beta} + \sum_{x=1}^k \frac{n_x [E_x - F_x]}{[1 - (1-q^x)^\alpha]^\beta - [1 - (1-q^{x+1})^\alpha]^\beta} \quad (13)$$

di mana

$$A_x = -\beta [1 - (1 - q^x)^\alpha]^{\beta-1} (1 - q^x)^\alpha \ln(1 - q^x)$$

$$B_x = \beta [1 - (1 - q^{x+1})^\alpha]^{\beta-1} (1 - q^{x+1})^\alpha \ln(1 - q^{x+1})$$

$$C_x = [1 - (1 - q^x)^\alpha]^\beta \ln[1 - (1 - q^x)^\alpha]$$

$$D_x = [1 - (1 - q^{x+1})^\alpha]^\beta \ln[1 - (1 - q^{x+1})^\alpha]$$

$$E_x = \alpha \beta x q^{x-1} [1 - (1 - q^x)^\alpha]^{\beta-1} (1 - q^x)^{\alpha-1}$$

$$F_x = \alpha \beta (x + 1) q^x [1 - (1 - q^{x+1})^\alpha]^{\beta-1} (1 - q^{x+1})^{\alpha-1}$$

Untuk mendapatkan taksiran Maksimum *Likelihood* dari parameter α, β , dan q ,

persamaan (11), (12), dan (13) disamakan dengan nol, kemudian ketiga persamaan tersebut diselesaikan dengan iterasi numerik menggunakan *software Mathematica*.

2.3 Ilustrasi Data

Data yang digunakan adalah data dari banyaknya klaim asuransi kendaraan bermotor di Irlandia pada tahun 2013 dengan periode asuransi satu tahun. Data tersebut terdapat pada tabel berikut ini.

Tabel 1: Data jumlah klaim asuransi kendaraan bermotor di Irlandia tahun 2013

Jumlah klaim (x)	0	1	2	3	4	5	6	7	8	9	10+
Frekuensi (n_x)	15	5	2	7	1	1	1	1	0	0	3

Selanjutnya, akan diuji secara terpisah apakah data tersebut dapat dimodelkan oleh distribusi geometrik dan distribusi Kumaraswamy-geometrik dengan menggunakan uji Kolmogorov-Smirnov. Kemudian, dengan menggunakan *Akaike Information Criteria* (AIC) akan diketahui distribusi mana yang paling bagus dalam memodelkan data.

1. Uji Kolmogorov-Smirnov untuk distribusi geometrik

Taksiran parameter dari distribusi geometrik dicari terlebih dahulu berdasarkan data yang diberikan. Taksiran parameter dari distribusi geometrik dapat diperoleh dengan menggunakan metode *Maximum Likelihood* dan diperoleh taksiran nilai parameter $\hat{p} = 0.3956$. Setelah diperoleh nilai taksiran parameter, maka dilakukan uji Kolmogorov-Smirnov. Misalkan $F_n(x)$ adalah fungsi distribusi empiris untuk data jumlah klaim dan $G(x)$ adalah fungsi distribusi dari distribusi geometrik dengan taksiran nilai parameter yang telah diperoleh sebelumnya, yaitu

$$G(x) = 1 - (1 - p)^{x+1} = 1 - (0.6044)^{x+1}$$

- Hipotesis
 $H_0: F_n(x) = G(x)$, untuk setiap x
 $H_1: F_n(x) \neq G(x)$, untuk paling tidak satu nilai x
 Atau
 H_0 : Data dapat dimodelkan dengan distribusi geometrik
 H_1 : Data tidak dapat dimodelkan dengan distribusi geometrik

- Tingkat signifikansi

$$\alpha = 0.05$$

- Statistik Uji

$$T = \sup_x |G(x) - F_n(x)|$$

Berdasarkan perhitungan, statistik ujinya adalah

$$T = \sup_x |G(x) - F_n(x)| = 0.1681$$

- Aturan Keputusan

H_0 pada $\alpha = 0.05$ akan ditolak jika T lebih besar dari kuantil ke 0.95, yaitu $w_{0.95}$. Untuk $n = 36$ dan $\alpha = 0.05$, diperoleh

$$w_{0.95} = \frac{1,36}{\sqrt{36}} = 0.2267$$

- Kesimpulan

Statistik uji $T < w_{0.95}$, maka H_0 tidak ditolak
 Sehingga dapat disimpulkan bahwa data jumlah klaim tahun 2013 berdistribusi geometrik.

2. Uji Kolmogorov-Smirnov untuk distribusi Kumaraswamy-geometrik

Taksiran parameter dari distribusi Kumaraswamy-geometrik dicari terlebih dahulu berdasarkan data yang diberikan. Taksiran parameter dari distribusi geometrik dapat diperoleh dengan menggunakan metode *Maximum Likelihood* dan diperoleh taksiran nilai parameter sebagai berikut

$$\hat{\alpha} = 0.094955, \hat{\beta} = 0.104819, \text{ dan } \hat{q} = 0.059823.$$

Setelah diperoleh nilai taksiran parameter, maka dilakukan uji Kolmogorov-Smirnov. Misalkan $F_n(x)$ adalah fungsi distribusi empiris untuk data jumlah klaim dan $G(x)$ adalah fungsi distribusi dari distribusi Kumaraswamy-geometrik dengan taksiran nilai parameter yang telah diperoleh, yaitu

$$\begin{aligned} G(x) &= 1 - [1 - (1 - q^{x+1})^\alpha]^\beta \\ &= 1 - [1 - (1 - 0.059823^{x+1})^{0.094955}]^{0.104819} \end{aligned}$$

- Hipotesis

$$H_0: F_n(x) = G(x), \text{ untuk setiap } x$$

$$H_1: F_n(x) \neq G(x), \text{ untuk paling tidak satu nilai } x$$

Atau

$$H_0: \text{Data dapat dimodelkan dengan distribusi Kumaraswamy-geometrik}$$

$$H_1: \text{Data tidak dapat dimodelkan dengan distribusi Kumaraswamy-geometrik}$$

- Tingkat signifikansi

$$\alpha = 0.05$$

- Statistik Uji

$$T = \sup_x |G(x) - F_n(x)|$$

Berdasarkan perhitungan, statistik ujinya adalah

$$T = \sup_x |G(x) - F_n(x)| = 0.0666$$

- Aturan Keputusan

H_0 pada $\alpha = 0.05$ akan ditolak jika T lebih besar dari kuantil ke 0.95, yaitu $w_{0.95}$. Untuk $n = 36$ dan $\alpha = 0.05$, diperoleh

$$w_{0.95} = \frac{1,36}{\sqrt{36}} = 0.2267$$

- Kesimpulan

Statistik uji $T < w_{0.95}$, maka H_0 tidak ditolak

Sehingga dapat disimpulkan bahwa data jumlah klaim tahun 2013 berdistribusi Kumaraswamy-geometrik.

3. Pemilihan Model Terbaik

Selanjutnya akan dilakukan perbandingan antara distribusi geometrik dan distribusi Kumaraswamy-geometrik untuk mengetahui distribusi mana yang terbaik dalam memodelkan data jumlah klaim. Metode *Akaike Information Criteria* (AIC) akan digunakan dalam melakukan perbandingan tersebut. Distribusi yang memiliki nilai AIC yang terkecil merupakan distribusi yang terbaik untuk memodelkan data.

Tabel 2: Nilai AIC untuk distribusi geometrik dan Kumaraswamy-geometrik

	distribusi geometrik	distribusi Kumaraswamy-geometrik
Nilai AIC	151.346	148.216

Model terbaik adalah model yang memiliki nilai AIC terkecil. Dari Tabel 2, diperoleh nilai AIC terkecil yaitu pada distribusi Kumaraswamy-geometrik, distribusi Kumaraswamy-geometrik dapat memodelkan data jumlah klaim tersebut lebih baik dari distribusi geometrik.

3. Kesimpulan

Kesimpulan yang diperoleh dari makalah ini adalah

1. Distribusi Kumaraswamy-geometrik (KGD) merupakan distribusi dari peubah acak diskrit dengan tiga parameter *shape*, yaitu α , β dan q .
2. KGD merupakan distribusi unimodal, dengan modus bernilai sama dengan nol ketika nilai parameter $\alpha \leq 1$ dan modus akan cenderung bernilai lebih besar dari nol ketika nilai parameter $\alpha > 1$.
3. Jika parameter dari KGD, yaitu α menuju tak terhingga dan nilai parameter q menuju satu, maka pdf dari distribusi ini akan menuju nol.
4. KGD memiliki rumus eksplisit untuk momen, sehingga dapat ditemukan nilai dari setiap momen. Contohnya adalah dapat ditemukan nilai mean dan standar deviasi melalui rumus momen ini.
5. Penaksiran parameter dari KGD diselesaikan secara numerik dengan bantuan *software Mathematica*.
6. KGD mampu memodelkan data jumlah klaim asuransi kendaraan bermotor di Irlandia dengan lebih baik daripada distribusi geometrik.

Referensi

- [1] Akinsete, A., Famoye, F., dan Lee, C. (2014). The Kumaraswamy-geometric distribution. *Journal of Statistical Distributions and Applications*.
- [2] Alzaatreh, A., Lee, C., dan Famoye, F. (2011). A New Method for Generating Families of Continuous Distributions.
- [3] Alzaatreh, A., Lee, C., dan Famoye, F. (2011). On the discrete analogues of continuous distributions. *Statistical Methodology*, 6, 589-603.
- [4] Conover, W.J. (1971). *Practical Nonparametric Statistics*. John Willey & Sons, Inc.
- [5] Hogg, R.V., dan A.T. Craig. (2012). *Introduction to Mathematical Statistics*, 5th ed. United States of America: Prentice Hall, Inc.
- [6] Jones, M.C. (2009). Kumaraswamy's Distribution: A Beta-Type Distribution with Some Tractability Advantages. *Statistical Methodology*, 6, 70-81.
- [7] Klugman, Stuart A., Panjer, Harry H., Willmot, Gordon E. (2012). *Loss Models from Data to Decisions*, 4th ed. New York: Wiley.
- [8] Kumaraswamy, P. (1980). A Generalized Probability Density Functions for Double-Bounded Random Processes. *Journal of Hydrology*, 46, 79-88
- [9] Private Motor Insurance Statistics. (2013). Retrieved from <http://www.centralbank.ie/polstats/stats/motorins/Pages/Data.aspx>
- [10] Workplace Law. *Metal & Engineering Industries Conditions of Employment*. January 6, 2017.

Didukung oleh:

**PT Reasuransi Indonesia Utama
Universitas Gunadarma
PT Tokio Marine Life Insurance Indonesia
PT AIA Financial Indonesia
PT BNI Life Insurance
BPJS Ketenagakerjaan
Persatuan Aktuaris Indonesia (PAI)
PT Asuransi Cigna**

**Vol. 1, September 2017
ISSN : 1907 - 2562**

**Departemen Matematika FMIPA Universitas Indonesia
Jurusan Matematika FMIPA Universitas Padjadjaran**

<http://www.math.ui.ac.id/snm2017>