



PENERBIT ANDI®

# Bayesian Neural Network

dalam Pemodelan Small Area Estimation

Nur Iriawan | Brodjol Sutijo Suprih Ulama  
Kartika Fithriasari | Achmad Syahrul Choir  
Rindang Bangun Prasetyo

# **BAYESIAN NEURAL NETWORK DALAM PEMODELAN SMALL AREA ESTIMATION**

Nur Iriawan  
Brodjol Sutijo Suprih Ulama  
Kartika Fithriasari  
Achmad Syahrul Choir  
Rindang Bangun Prasetyo

**PENERBIT Andi**

# **BAYESIAN NEURAL NETWORK DALAM PEMODELAN SMALL AREA ESTIMATION**

**Disusun oleh : Nur Iriawan**

**Brodjol Sutijo Suprih Ulama**

**Kartika Fithriasari**

**Achmad Syahrul Choir**

**Rindang Bangun Prasetyo**

Hak Cipta ©2020 pada Penulis.

Editor : Adit Christian

Desain Cover : Ofry

Setter : yulius basuki

Hak Cipta dilindungi undang-undang.

Dilarang memperbanyak atau memindahkan sebagian atau seluruh isi buku ini dalam bentuk apa pun, baik secara elektronik maupun mekanis, termasuk memfotokopi, merekam atau dengan sistem penyimpanan lainnya, tanpa izin tertulis dari Penulis.

Penerbit oleh Penerbit ANDI (Anggota IKAPI)

Jl. Beo 38-40, Telp. (0274) 561881 (Hunting), Fax. (0274) 588282 Yogyakarta 55281

Percetakan: CV ANDI OFFSET

Jl. Beo 38-40, Telp. (0274) 561881 (Hunting), Fax. (0274) 588282 Yogyakarta 55281

*Iriawan, Nur*

**Bayesian Neural Network dalam Pemodelan Small Area Estimation/Nur Iriawan,  
Brodjol Sutijo Suprih Ulama, Kartika Fithriasari, Achmad Syahrul Choir, dan  
Rindang Bangun Prasetyo**

**– Ed. I. – Yogyakarta: ANDI**

**24 - 23 - 22 - 21 - 20**

**hlm xviii + 204 ; 16x23 cm.**

**5 4 3 2 1**

**ISBN: 978-623-01-0895-2 (PDF)**

**1. Internet**

**2. Ulama, Brodjol Sutijo Suprih**

**3. Fithriasari, Kartika**

**4. Choir, Achmad Syahrul**

**5. Prasetyo, Rindang Bangun**

**DDC'23 :**

# KATA PENGANTAR

*Alhamdulillah Robbil 'Alamin*, puji syukur kehadiran Allah *Subhanahu wata'ala* Yang Maha Pengasih lagi Maha Penyayang, karena atas anugerah dan rahmat-Nya, buku Bayesian Neural Network dalam Pemodelan Small Area Estimation dapat terselesaikan dengan baik.

Buku ini memadukan dua metode, yaitu Bayesian dan Neural Network yang diimplementasikan dalam pemodelan Small Area Estimation (SAE). Perpaduan tersebut diawali dengan membahas konsep dasar Bayesian dan Neural Network dan bertahap memadukan keduanya melalui pintu Bayesian Regresi yang mendasari pemodelan SAE. Dalam buku ini juga dibahas tentang bagaimana pemanfaatan teknologi informasi untuk penyajian data yang lebih menarik dan interaktif. Materi bahasan disajikan dalam bahasa yang mudah dan dilengkapi dengan contoh aplikasi baik dalam bentuk source code maupun tampilan hasilnya.

Pada kesempatan ini kami ingin mengucapkan terima kasih kepada Kementerian Riset, Teknologi, dan Pendidikan Tinggi Republik Indonesia yang telah mendanai penelitian kami sehingga membuahkan buku ini sebagai salah satu hasilnya, sesuai dengan Surat Perjanjian Penugasan Pelaksanaan Program Penelitian Nomor: 003246.216/ IT2.11/PN.08/2015 Tanggal 5 Februari 2015 dan Nomor: 01830/IT2.11/ PN.08/2016 Tanggal 17 Februari 2016, serta semua pihak yang telah membantu terselesainya buku ini. Kami sangat bangga hati dapat menyumbangkan buku ini sebagai jariah kami untuk dapat dimanfaatkan dengan baik terutama dalam studi Bayesian, Neural Network dan Small Area Estimation.

Akhir kata, dengan segala keterbatasan kemampuan dan pengetahuan yang telah dioptimalkan dalam penyusunan buku ini, kami masih menyadari adanya kekurangsempurnaan. Oleh sebab itu saran dan masukan dari semua pihak sangatlah kami butuhkan.

Surabaya, Desember 2016

Tim Penulis

# DAFTAR ISI

<b>KATA PENGANTAR.....</b>	<b>III</b>
<b>DAFTAR ISI .....</b>	<b>V</b>
<b>DAFTAR TABEL.....</b>	<b>XI</b>
<b>DAFTAR GAMBAR .....</b>	<b>XIII</b>
<b>BAB 1 PENGANTAR BAYESIAN DAN BAYESIAN</b>	
<b>REGRESI .....</b>	<b>1</b>
1.1 Pengantar Bayesian.....	1
1.2 Bayesian dan Frekuentis .....	5
1.2.1 Perbedaan Bayesianis dan Frekuentis.....	6
1.2.2 Kompromi antara Bayesianis dan Frekuentis.....	10
1.2.3 Kritik-Kritik terhadap Bayesian .....	12

1.3	Bayesian Regresi.....	14
1.3.1	<i>Likelihood</i> Gabungan antara $\beta_0$ atau $\beta_{\bar{x}}$ dan $\beta_1$ .....	17
1.3.2	Prior Parameter untuk Model Regresi, $\beta_{\bar{x}}$ dan $\beta_1$ .....	20
1.3.3	Posterior Gabungan untuk Parameter Model Regresi, $\beta_{\bar{x}}$ dan $\beta_1$ .....	20
1.3.4	Marginal Posterior $\beta_1$ .....	21
1.3.5	Marginal Posterior $\beta_{\bar{x}}$ .....	22
1.3.6	Bayesian Credible Interval dan Pengujian Hipotesis .....	23
1.3.7	Distribusi dan Observasi Prediktif .....	24
1.4	Bayes Factor .....	28
1.5	Komputasi Bayesian .....	31
1.5.1	Bayesian Inference Using Gibbs Sampling (BUGS).....	31
1.5.2	Gibbs Sampler .....	42
1.5.3	Adaptive Rejection Sampling .....	43
1.5.4	Slice Sampler .....	44
1.5.5	Metropolis Hasting.....	47
1.5.6	Konvergensi MCMC.....	48

**BAB 2 PENGANTAR NEURAL ETWORK DAN APLIKASINYA.....53**

2.1	Pengantar Neural Network.....	53
2.2	Single Layer Perceptron.....	56
2.3	Feed-Forward Neural Network (FFNN) .....	56

2.3.1	Proses Pembelajaran pada FFNN .....	60
2.3.2	Aplikasi FFNN .....	64
2.4	Radial Basis Function .....	69
2.4.1	Aplikasi RBFNN.....	74
2.5	Neural Network dan Statistika .....	77
<b>BAB 3</b>	<b>SMALL AREA ESTIMATION .....</b>	<b>81</b>
3.1	Pendahuluan .....	81
3.2	Estimasi Langsung .....	82
3.3	Estimasi Tidak Langsung.....	83
<b>BAB 4</b>	<b>BAYESIAN SMALL AREA ESTIMATION .....</b>	<b>89</b>
4.1	Pendahuluan .....	89
4.2	Estimasi Langsung dan Estimator Regresi Linear .....	90
4.3	Model Bayesian Hierarki .....	93
4.3.1	Model Tingkat Unit .....	93
4.3.2	Model Tingkat Area.....	98
4.4	Distribusi Prior Parameter Model .....	99
4.5	Kebaikan Hasil Estimasi .....	101
4.6	SAE dengan Ketidaktersediaan Data di Suatu Area.....	102
4.7	Efek Random dengan Korelasi Spasial.....	103
4.7.1	Spesifikasi Multivariat Normal dalam SAE .....	104
4.7.2	Spesifikasi CAR dalam SAE.....	104
4.8	Penggunaan Informasi dari Level Area yang Lebih Tinggi dalam SAE .....	107



4.9	Klasifikasi Wilayah Hasil SAE dalam Pengambilan Kebijakan .....	109
<b>BAB 5 BAYESIAN NEURAL NETWORK UNTUK SMALL AREA ESTIMATION..... 111</b>		
5.1	Pendahuluan .....	111
5.2	Bayesian Neural Network.....	112
5.3	Hierarchical Bayesian Neural Network dan Penerapannya dalam SAE.....	116
<b>BAB 6 BAYESIAN GENERAL REGRESSION NEURAL NETWORK UNTUK SMALL AREA ESTIMATION ... 119</b>		
6.1	Pendahuluan .....	119
6.2	Arsitektur GRNN.....	120
6.3	Struktur GRNN dalam SMALL AREA ESTIMATION .....	124
6.4	Komputasi HBGRNN .....	127
<b>BAB 7 KASUS I: HIERARCHICAL BAYES SAE UNTUK PROPORSI KEMISKINAN ..... 133</b>		
7.1	Pendahuluan .....	133
7.2	Estimasi Proporsi Kemiskinan Tingkat Kecamatan .....	135
7.3	Kesimpulan.....	142
<b>BAB 8 KASUS II: HBGRNN-SAE UNTUK ANGKA PUTUS SEKOLAH PADA RUMAH TANGGA MISKIN ..... 143</b>		
8.1	Pendahuluan .....	143
8.2	Estimasi Angka Putus Sekolah di Rumah Tangga Miskin .....	144
8.3	Kesimpulan.....	149

<b>BAB 9 PENYAJIAN HASIL ANALISIS MELALUI SISTEM INFORMASI GEOGRAFI DAN DASHBOARD .....</b>	<b>151</b>
9.1 Pendahuluan .....	151
9.2 Pembuatan Peta Tematik .....	153
9.2.1 Tampilan <i>windows</i> utama Pada ArcView .	154
9.2.2 Memulai project baru .....	157
9.2.3 Melakukan Edit Tabel Data.....	162
9.2.4 Membuat Peta Tematik.....	165
9.2.5 Membuat Layout dan Export Peta Tematik....	170
9.3 Sistem Informasi Geografi Angka Putus Sekolah	173
9.3.1 Menu Akses.....	176
9.3.2 Menu Entri Data.....	178
9.3.3 Menu Output .....	180
9.3.4 Menu Pengaturan.....	182
9.4 Pengembangan Dashboard .....	183
<b>DAFTAR PUSTAKA .....</b>	<b>187</b>
<b>TENTANG PENULIS.....</b>	<b>201</b>



# DAFTAR TABEL

Tabel 1.1 Interpretasi <i>Bayes Factor</i> dalam Pemilihan Hipotesis.....	30
Tabel 1.2. Contoh Data Regresi Sederhana.....	38
Tabel 1.3 Metode Pembangkitan sampel sesuai urutan yang digunakan di WinBUGS.....	42
Tabel 2.1 Rata-Rata Variabel Pada Setiap Kelompok .....	74
Tabel 2.2 Standar Deviasi Variabel pada Setiap Kelompok ....	74
Tabel 2.3 Kesamaan Model NN dan Model Statistika .....	78
Tabel 8.1 DIC Model HB dan HBGRNN .....	145
Tabel 9.1 Tombol yang Terdapat pada Toolbar dan Fungsinya.....	160



# DAFTAR GAMBAR

Gambar 1.1	(a) dan (c) Interval Konfidensi;(b) dan (d) Interval Kredibel. ....	9
Gambar 1.2	<i>Scatter Plot</i> antara banyak R. Kelas dan J. Siswa.....	15
Gambar 1.3	<i>Scatter Plot</i> antara R. Kelas dan J. Siswa .....	17
Gambar 1.4	Directed Acylic Graph (DAG) Model Regresi Sederhana.....	32
Gambar 1.5	<b>Syntax</b> BUGS model regresi sederhana. ....	33
Gambar 1.6	Doodle DAG Regresi Berganda yang Dibuat di WinBUGS. ....	34
Gambar 1.7	<i>Window Specification Tool</i> .....	37
Gambar 1.8	<i>Window Sample Monitor Tool</i> .....	40
Gambar 1.9	<i>Window Update Tool</i> .....	40
Gambar 1.10	<b>Adaptive Rejection Sampling</b> dengan Metode Garis Singgung; Garis Tebal: $\log(g(X))$ ; Garis Tipis: Envelope $\log(G_s(x))$ ; Garis Putus: <b>Squeezing Function</b> $\log(b_s(X))$ ; Anak Panah: Absis yang Digunakan. ....	44

Gambar 1.11	Algoritma <i>Slice Sampler</i> . .....	45
Gambar 1.12	Diagram <b>History Output</b> MCMC: (a) Bentuk "ulat bulu gemuk" dan (b) bentuk mengular ...	48
Gambar 2.1	Struktur sederhana sebuah neuron.....	54
Gambar 2.2	Struktur neuron jaringan syaraf tiruan. ....	55
Gambar 2.3	Arsitektur FFNN dengan satu <b>hidden layer</b> , <b>n</b> unit <b>input</b> , <b>p</b> unit neuron/node pada <b>hidden layer</b> , dan <b>m</b> unit neuron/node pada <b>output layer</b> . .....	57
Gambar 2.4	Fungsi <b>logistik sigmoid</b> . .....	59
Gambar 2.5	Fungsi aktivasi tanh. ....	60
Gambar 2.6	Perintah matlab model FFNN untuk klasifikasi.....	66
Gambar 2.7	Arsitektur RBFNN dengan <b>n</b> unit di lapisan tersembunyi.....	71
Gambar 2.8	Arsitektur RBFNN dengan <b>K</b> unit di lapisan tersembunyi dengan fungsi aktivasi Gaussian. ....	73
Gambar 2.9	<i>Syntax</i> aplikasi RBFNN bagian 1.....	75
Gambar 2.10	<i>Syntax</i> aplikasi RBFNN bagian 2.....	75
Gambar 2.11	<b>Single Layer Linear Perceptron</b> fungsi aktivasi linear = regresi multivariat.....	79
Gambar 2.12	<b>Single Layer Perceptron</b> fungsi aktivasi logistik = regresi logistik.....	79
Gambar 6.1	Arsitektur GRNN. ....	123
Gambar 6.2	DAG HBGRNN Tipe 1 yang menggunakan hiperprior. ....	126
Gambar 7.1	<b>Syntax</b> model logit-normal <b>hierarchical</b> Bayes SAE proporsi kemiskinan. ....	137
Gambar 7.2	Statistik (a), <b>DENSITY plot</b> (b), <b>history plot</b> (c), dan <b>autocorrelation plot</b> (d) proporsi kemiskinan pada $i = 1$ dan 2.....	139

Gambar 7.3	Perbandingan hasil estimasi langsung ( <i>direct</i> ) dengan HBSAE proporsi kemiskinan ( <i>p</i> ) pada kecamatan sampel.....	140
Gambar 7.4	Peta kemiskinan level kecamatan di Jawa Timur berdasarkan hasil HB SAE.....	141
Gambar 8.1	Perbandingan hasil estimasi APTS pada rumah tangga miskin dengan estimasi langsung ( <i>Direct</i> ), HB dan HBGRNN.....	146
Gambar 8.2	Peta Tematik Hasil Estimasi HBGRNN APTS pada Rumah Tangga Miskin Jawa Timur Menurut Kabupaten/Kota, Tahun 2014.....	147
Gambar 8.3	Perbandingan nilai CV estimasi langsung ( <i>direct</i> ), HB dan HBGRNN.....	148
Gambar 8.4	<b>Density plot</b> nilai $SE(p)$ metode HB dan HBGRNN. ....	149
Gambar 9.1	Tampilan <b>Window</b> Views. ....	155
Gambar 9.2	Tampilan <b>Window</b> Tables.....	155
Gambar 9.3	Tampilan <b>Window</b> Charts. ....	156
Gambar 9.5	Tampilan <b>Window</b> 'Welcome to ArcView GIS'.....	158
Gambar 9.6	Tampilan <b>Window</b> Utama ArcView.....	158
Gambar 9.7	Tampilan <b>Window</b> 'Save Project As'.....	159
Gambar 9.8	Tampilan <b>Window</b> 'Add Theme'.....	159
Gambar 9.9	Tampilan peta Jawa Timur pada <b>Window</b> View1.....	160
Gambar 9.10	Tampilan menu table dan edit pada <b>Window</b> Table.....	162
Gambar 9.11	Tampilan pembuatan variabel baru pada <b>Window</b> Table.....	163
Gambar 9.12	Tampilan pengisian variabel pada <b>Window</b> Table.....	164



Gambar 9.13 Tampilan penghapusan variabel pada <b>Window</b> Table.....	164
Gambar 9.14 Tampilan penggantian tampilan <b>window</b> yang aktif.....	165
Gambar 9.15 Tampilan <b>Window</b> Legend Editor.....	167
Gambar 9.16 Tampilan <b>Window</b> Classification.....	168
Gambar 9.17 Tampilan <b>Window</b> Fill Palette dan Color Palette.....	168
Gambar 9.18 Tampilan <b>Window</b> Legend Editor yang sudah disesuaikan.....	169
Gambar 9.19 Tampilan <b>Window</b> View1 yang menampilkan peta tematik APTS Rumah Tangga Miskin.....	169
Gambar 9.20 Tampilan <b>Window</b> Template Manager.....	170
Gambar 9.21 Tampilan Layout awal dari Template Manager.....	171
Gambar 9.22 Tampilan Layout setelah diedit sesuai kebutuhan.....	171
Gambar 9.23 Tampilan <b>Window</b> Export.....	172
Gambar 9.24 Tampilan Splash Screen Aplikasi SIG APTS.....	175
Gambar 9.25 Tampilan <b>Window</b> utama aplikasi SIG APTS. .	176
Gambar 9.26 Fasilitas pada Menu Akses Aplikasi SIG APTS.	176
Gambar 9.27 Tampilan <b>Window</b> Login Aplikasi SIG APTS....	177
Gambar 9.28 Tampilan <b>Window</b> Data Pengguna SIG APTS .	177
Gambar 9.29 Tampilan <b>Window</b> peringatan konfirmasi keluar. ....	178
Gambar 9.30 Fasilitas pada Menu Entri Data Aplikasi SIG APTS.....	178
Gambar 9.31 Tampilan <b>Window</b> Entri Data Penelitian SIG APTS.....	179
Gambar 9.32 Tampilan <b>Window</b> Entri Master Data Wilayah SIG APTS.....	179

Gambar 9.33 Fasilitas pada menu <b>output</b> aplikasi SIG APTS.....	180
Gambar 9.34 Tampilan <b>Window</b> Tabulasi SIG APTS.....	180
Gambar 9.35 Tampilan <b>Window</b> Grafik SIG APTS. ....	181
Gambar 9.36 Tampilan <b>Window</b> Pemetaan SIG APTS.....	182
Gambar 9.37 Fasilitas pada menu Pengaturan aplikasi SIG APTS.....	182
Gambar 9.38 Tampilan <b>Window</b> Tentang Program SIG APTS.....	183
Gambar 9.39 Pembuatan Query di DBXtra.....	184
Gambar 9.40 Desain <b>Dashboard</b> APTS Penduduk Miskin Jawa Timur. ....	185
Gambar 9.41 Tampilan <b>Dashboard</b> APTS Penduduk Miskin Jawa Timur.....	186



# bab 01 PENGANTAR BAYESIAN DAN BAYESIAN REGRESI

Bab ini:

1. Mengenalkan metode Bayesian.
2. Menjelaskan perbedaan Bayesian dan Frekuentis.
3. Menjelaskan kompromi Bayesian dan Frekuentis.
4. Menjelaskan metode Bayesian dalam pemodelan regresi.
5. Menjelaskan beberapa kritik terhadap metode Bayesian. Menjelaskan komputasi Bayesian dengan algoritma MCMC.

## 1.1 Pengantar Bayesian

Hal utama yang harus dilalui dalam mempelajari statistika ialah estimasi parameter model. Bayesian adalah salah satu cara estimasi tersebut. Baik model yang diestimasi tersebut akan merepresentasikan data sebagai densitas data maupun model hubungan antarvariabel dalam data yang diskrit/kontinu, *single/multiple* parameter, linear/nonlinear, parametrik/nonparametrik, *cross-section*/stokastik (basis waktu–model *time series*– maupun basis area –model *spatial*–), dan pemodelan berbasis normalitas/neo-normalitas/non-normalitas. Tampak sangat luas sekali kegunaan Bayesian itu. Oleh sebab itu bagian awal buku ini memberikan gambaran tentang bagian mana dari kegunaan-kegunaan Bayesian di atas yang akan dibahas dan diimplementasikan dalam buku ini.

Dalam pendekatan Bayesian, data sampel yang diperoleh dari populasi, juga memperhitungkan suatu distribusi awal yang disebut prior. Menurut Iriawan (2001) untuk mengestimasi suatu parameter model, data dikelompokkan menjadi dua bagian, yaitu data pengamatan saat ini yang bersifat sesaat selama studi dan data yang bersifat *long memory* histogram yang terkumpul sebagai pengetahuan/pengalaman di masa lalu. Berbeda dengan pendekatan statistika klasik (frekuentis) yang memandang parameter sebagai yang bernilai tetap dan tunggal, pendekatan statistika Bayesian memandang parameter sebagai variabel random yang memiliki distribusi yang disebut sebagai distribusi prior. Distribusi prior selanjutnya dipadukan dengan pengamatan yang diperoleh dan dapat ditentukan distribusi posterior sehingga diperoleh estimator Bayesian.

Inferensi Bayesian didasarkan pada distribusi posterior yang merupakan perpaduan antara distribusi prior (informasi masa lalu sebelum dilakukan observasi) dan data observasi yang digunakan untuk menyusun fungsi *likelihood* (Box dan Tiao, 1973). Hubungan distribusi posterior dengan distribusi prior dan *likelihood* dapat dituliskan sebagai berikut:

$$\text{Distribusi posterior} \propto \text{likelihood} \times \text{Distribusi prior}$$

Ilustrasi cara memperoleh estimator Bayesian dapat diberikan sebagai berikut. Apabila terdapat parameter  $\theta$  yang harus diestimasi, yang menggambarkan karakteristik data observasi lama daya tahan seorang anak tetap bersekolah (*survival time*),  $T$ , maka distribusi probabilitas untuk posterior  $\theta$  pada data  $t$  akan proporsional dengan perkalian antara distribusi prior  $\theta$  dan fungsi *likelihood*  $\theta$  yang diberikan oleh data  $t$  yang diamati. Secara matematis dapat dituliskan sebagai berikut:

$$p(\theta|t) = \frac{p(t|\theta)p(\theta)}{p(t)},$$

atau

$$p(\theta|t) \propto p(t|\theta)p(\theta).$$

$p(\theta | t)$  merupakan distribusi posterior dari  $\theta$  yang nilainya proporsional dengan perkalian antara fungsi *likelihood*  $p(t | \theta)$  dan distribusi prior  $p(\theta)$ . Hal ini berarti bahwa pengetahuan tentang  $\theta$  yang telah diperoleh jauh hari sebelum pengamatan (prior) akan diperbarui oleh data pengamatan saat ini (*likelihood*  $p(t | \theta)$ ). Deskripsi dan keputusan tentang karakteristik data  $t$  didasarkan pada hasil perbaruan pola (posterior)  $\theta$  ini.

Apabila  $t$  hanya sebuah variabel yang tunggal dan berparameter  $\theta$ , maka karakteristik  $\theta$  merupakan karakteristik distribusi dari data  $t$ , dan *likelihood*  $p(t | \theta)$  yang dibutuhkan dalam mempelajari  $\theta$  adalah *likelihood* dari distribusi data  $t$ . Namun, apabila  $t$  terdiri atas dua atau lebih variabel dan saling mempunyai keterkaitan hubungan dengan salah satu sebagai respons, maka *likelihood* harus dibangun dengan menyusun *likelihood* model, misalkan model regresi.

Spesifikasi dari distribusi prior sangat penting pada metode Bayesian karena distribusi prior memengaruhi bentuk posterior yang akan digunakan untuk mengambil keputusan. Apabila informasi prior tersedia, maka informasi untuk distribusi prior akan terangkum di dalamnya. Tetapi biasanya informasi prior tidak mudah tersedia, sehingga perlu penetapan distribusi prior yang sesuai dengan pola data pengamatannya. Distribusi tersebut biasa dikenal dengan sebutan prior sekawan (*conjugate*) yang parameterisasi distribusi priornya dapat ditentukan secara informatif atau non-informatif sesuai data pengamatan yang diperolehnya. Dapat juga prior sekawan ini diberikan suasana yang lebih adaptif terhadap datanya, yaitu dengan memberikan reparameterisasi ke dalam bentuk *flat-conjugate prior*. Sebagai contoh, parameternya sekawan dengan distribusi normal, maka priornya diberikan pola normal dengan varians yang cukup besar, sehingga bentuk priornya adalah normal yang datar dan domainnya lebar, dan sering dinamakan sebagai *flat-conjugate*

*prior*. Macam-macam cara penentuan prior dalam estimasi Bayesian dijelaskan di bawah ini.

Distribusi prior merupakan informasi terdahulu mengenai parameter model yang dipelajari sebelum pengamatan data dilakukan. Pemilihan distribusi prior dalam pendekatan Bayesian harus tepat. Berbagai macam distribusi prior yang dapat digunakan dalam Bayesian, antara lain:

1. *Conjugate prior* atau *non-conjugate prior*, yaitu penentuan *prior* didasarkan pada pola *likelihood* dari datanya (Box dan Tiao, 1973). Jika distribusi prior sepola dengan *likelihood*-nya, maka dikatakan prior tersebut sebagai *conjugate prior* dan sebaliknya sebagai *non-conjugate prior*.
2. *Proper prior* atau *improper prior* (*Jeffreys prior*), yaitu *prior* yang terkait dengan pemberian bobot atau densitas di setiap titik apakah terdistribusi secara *uniform* atau tidak (Ntzoufras, 2009).
3. *Informative prior* atau *non-informative prior*, yaitu penentuan *prior* yang didasarkan pada ketersediaan pengetahuan atau informasi sebelumnya mengenai pola distribusi data yang diperoleh dari penelitian sebelumnya (Box dan Tiao, 1973).
4. *Pseudo prior* (Carlin dan Chib, 1995), yaitu penentuan prior berdasarkan pada hasil analisis secara frekuentis. Umumnya, prior jenis ini adalah *conjugate* dengan memberikan nilai parameter priornya berdasarkan pada hasil pengolahan datanya secara frekuentis. Misalnya distribusi dan parameter priornya diberi nilai dari hasil estimasi parameter model dengan metode maksimum *likelihood*.

Tidak menutup kemungkinan spesifikasi distribusi prior diberikan sebagai kombinasi di antara keempat macam kategori prior di atas, misalnya prior didefinisikan sebagai prior *conjugate-informative*. Hal ini berarti bahwa distribusi prior sepola dengan *likelihood* datanya,

tetapi parameter distribusi prior yang *conjugate* tersebut ditentukan berdasarkan informasi dari data yang telah dimiliki untuk membangun *likelihood* tersebut. Seandainya *conjugate prior*-nya adalah normal, maka parameter normal ditentukan dengan *mean* dan varians dari data yang dimiliki.

Jika diinginkan distribusi priornya dijaga tetap *conjugate* tetapi tidak mempunyai informasi terbaru tentang datanya, maka dikatakan sebagai prior *conjugate-non-informative*. Sebagai contoh, seandainya *conjugate prior*-nya adalah normal, maka parameter distribusi Normal ditentukan dengan *mean* NOL dan varians sangat besar. Prior demikian dinamakan *flat-conjugate prior*.

## 1.2 Bayesian dan Frekuentis

Bayesianis dan frekuentis adalah dua kelompok dalam ilmu statistika yang sering menjadi sorotan karena berbagai diskusi dan perdebatan antara keduanya. Bayesianis adalah kumpulan statistikawan yang bekerja dan mendukung pengembangan metode Bayesian dalam analisis statistika. Penggagas metode ini antara lain adalah Bayes (1763), Laplace (1814), Dale (1982), dan Erkenntnis (1996). Di sisi lain, frekuentis adalah sekelompok statistikawan yang masih berpegang teguh pada prinsip-prinsip metode statistika klasik, seperti *uniformly minimum-variance unbiased estimator* (UMVUE), metode momen, maksimum *likelihood estimator* (MLE), dan sebagainya yang digagas oleh Fisher, Neyman, Pearson, Wald, dan lain-lain.

Perbedaan gagasan utamanya adalah bahwa parameter yang selalu menjadi pokok bahasan dalam statistika harus diestimasi. Proses estimasinya yang menargetkan bahwa nilainya akan tunggal, sehingga tidak berdistribusi, menjadi berbeda pada saat parameter tersebut dianggap tidak tunggal dan mempunyai distribusi. Pendapat pertama digolongkan pada kelompok frekuentis dan pendapat kedua digolongkan pada kelompok Bayesianis. Dalam hal estimasi parameter, frekuentis lebih banyak menggantungkan diri pada data yang diperoleh dan



digunakan untuk membangun *likelihood*. Sedangkan Bayesian masih memadukan *likelihood* tersebut dengan distribusi prior dari parameter yang dipelajari. Selain itu, beberapa perbedaan antara Bayesianis dan frekuentis dapat diberikan di Subbab 1.2.1.

## 1.2.1 Perbedaan Bayesianis dan Frekuentis

### 1. “fix dan random”

Bagi para frekuentis, data merupakan sampel random yang bersifat terulang (*repeatable*). Artinya, data dapat diperoleh secara berulang-ulang sehingga akan menghasilkan frekuensi. Dalam pengambilan data yang berulang-ulang tersebut, terkandung parameter yang bersifat konstan. Dengan demikian, dapat disimpulkan bahwa dalam paham frekuentis, parameter bersifat tetap (*fixed*).

Lain halnya bagi para Bayesianis. Data diperoleh dari hasil pengamatan sebenarnya, dianggap sebagai catatan kejadian sebagai ketetapan Sang Maha Pencipta, sehingga tetap. Sementara itu, parameter distribusi data pengamatan yang tidak diketahui dianggap bersifat probabilistik. Oleh karena itu, data bersifat tetap dan tidak terulang lagi. Data sebagai catatan *sunatullah* yang telah terjadi dan tidak akan terulang untuk terjadi lagi; baik besaran, waktu kejadian, maupun tempatnya sedangkan parameter bersifat random. Karena parameter yang berubah-ubah, maka kejadian yang terjadi dan dicatat sebagai data menjadi berubah-ubah dan berbeda dengan kejadian-kejadian masa lalu, walaupun tempat dan pelakunya adalah sama.

### 2. Probabilitas

Para frekuentis mendefinisikan probabilitas sebagai frekuensi relatif suatu kejadian dalam sejumlah percobaan/*trial* tertentu. Dengan demikian probabilitas adalah sesuatu yang riil, objektif dan kuantitas yang dapat diukur. Sementara itu, Bayesianis memandang probabilitas sebagai derajat kepercayaan terhadap kebenaran

suatu proposisi. Oleh karena itu, probabilitas dalam Bayesian lebih seperti konstruksi logika daripada berbentuk fisik, sehingga bersifat normatif dan subjektif. Pengalaman (*prior knowledge*) yang berbeda antara satu orang dengan yang lainnya akan membangun tingkat kepercayaan yang bisa berbeda. Tingkat kepercayaan tersebut diwujudkan dalam bentuk distribusi posterior, yang dapat diperoleh dari  $prior\ knowledge \times likelihood\ data$ .

### 3. Tujuan (*Goal*)

Tujuan dari analisis statistika berbasis frekuentis adalah menghasilkan prosedur dengan jaminan 'frekuensi' yang ada. Sementara itu, tujuan Bayesian adalah untuk melakukan analisis terhadap kepercayaan akan suatu kejadian dalam proses tertentu. Derajat kepercayaannya dibangun dengan memadukan pengamatan sekarang (dimasukkan ke dalam *likelihood*) dengan pengetahuan/pengalaman masa lalu (*prior knowledge*) pengamat. Bangunan itu dinamakan sebagai posterior.

### 4. Interval Konfidensi dan Interval Kredibel

Frekuentis sangat erat dengan interval konfidensi (*confidence interval*) yang dibangun dengan basis normalitas dan simetri. Hasil konsep ini diberikan pada Gambar 1.1 (a). karena simetri, pusat interval bisa di *mean* atau median atau modus,  $X_{mod}$ . Dari pusat interval tersebut, lompat ke kiri sejauh B diperoleh batas bawah,  $X_{bb}$ , dan lompat ke kanan sejauh B diperoleh batas atas,  $X_{ba}$ . Akibatnya, diperoleh ketinggian densitas yang sama di kedua batas tersebut yang ditandai dengan garis densitas setimbang (DS).

Konsep penentuan interval konfidensi ini tidak layak diterapkan untuk menentukan batas interval data berpola menceng. Lompatan ke kiri dan ke kanan sejauh jarak yang sama B, tidak logis. Karena proses lompatan tersebut menghasilkan densitas yang tidak setimbang. Perhatikan Gambar 1.1 (b). Untuk memperoleh batas interval yang adil, dengan densitas di setiap batas interval yang

sama, maka diperlukan jarak lompatan ke batas bawah yang tidak sama dengan ke batas atas. Tetapi, target luasan di antara kedua batas tersebut,  $X_{bb}$ , dan  $X_{ba}$  tetap dapat dijamin sebesar  $1 - \alpha$  dan ketinggian densitasnya adalah sama. Cara penentuan interval inilah yang digunakan dalam Bayesian. Hasil intervalnya dinamakan dengan *credible interval* atau interval yang kredibel. Keuntungan dari adanya algoritma penghitungan *credible interval* ini adalah bahwa metode ini akan memberikan hasil sama dengan hasil hitungan *confidence interval*, apabila *input* distribusi data ke dalam algoritma tersebut adalah simetri. Perbandingan formulasi kedua interval ini dapat dituliskan dalam dua persamaan berikut:

$$P(\bar{x} - B \leq x \leq \bar{x} + B) = 1 - \alpha, \dots\dots\dots (1.1)$$

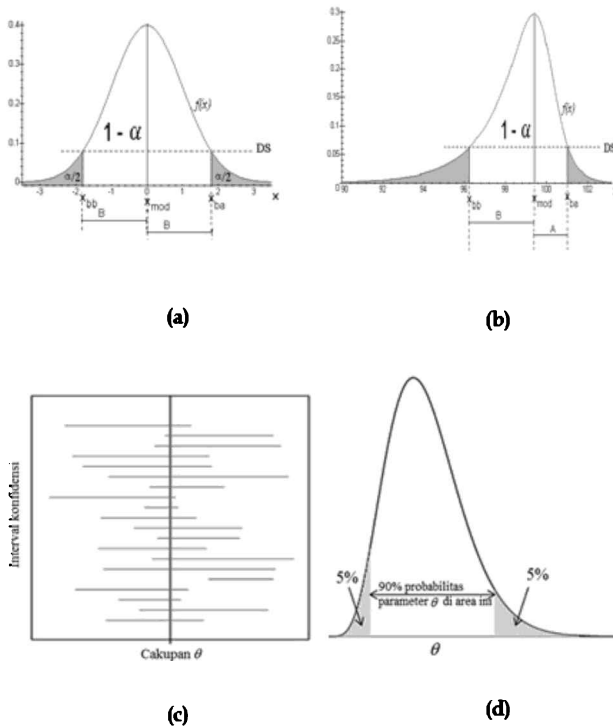
dan

$$\begin{aligned} P(X_{bb} \leq x \leq X_{ba}) &= 1 - \alpha, \\ P(X_{bb}) &= P(X_{ba}) \end{aligned} \quad (1.2)$$

masing-masing untuk interval konfidensi dan interval kredibel. Secara umum persamaan (1.1) digunakan pusat intervalnya adalah  $\bar{x}$  yang bisa juga diubah menjadi  $X_{mod}$  atau  $X_{med}$  karena konsep simetri *input* distribusi data yang akan dicari batas intervalnya. Sedangkan dalam proses persamaan (1.2) untuk menentukan intervalnya, tidak memerlukan informasi di mana pusat datanya.

Ilustrasi lain pengertian kedua interval tersebut adalah sebagai berikut. Andaikan dalam analisis frekuentis didapatkan suatu interval konfidensi 90% untuk *mean* yaitu 136,2 gram sampai 139,6 gram. Interpretasi dari interval ini adalah kita memiliki keyakinan (*confidence*) sebesar 90% bahwa nilai *mean* populasi berada dalam interval 136,2 sampai 139,6 gram. Perhatikan bahwa kesimpulan ini bukanlah suatu pernyataan probabilitas. Interval ini dibuat berdasarkan pernyataan probabilitas, tetapi karena pernyataan yang dimasukkan (yaitu *mean*) bukan merupakan kuantitas yang

bersifat random, sehingga bukan termasuk probabilitas. Bila dijelaskan lebih lanjut, interval konfidensi 90% menyatakan bahwa jika dilakukan pengambilan sampel sebanyak 100 kali, dari masing-masing sampel akan dihasilkan interval konfidensi, maka dapat dikatakan bahwa 90 dari 100 interval konfidensi yang terbentuk akan mengandung nilai parameter (*mean*) yang sebenarnya, lihat Gambar 1.1 (c). Sebaliknya, jika interval konfidensi ini adalah berbasis Bayesian, maka dapat diinterpretasikan bahwa dengan probabilitas 90%, nilai *mean* populasi berada dalam interval 136,2 sampai 139,6 gram. Perhatikan Gambar 1.1 (d)



Gambar 1.1 (a) dan (c) Interval Konfidensi;(b) dan (d) Interval Kredibel.

## 5. Banyaknya Data Pengamatan

Banyaknya data pengamatan yang digunakan untuk parameter dengan membangun *likelihood*, sangat berbeda tuntutan antara frekuentis dan Bayesianis. Estimasi model regresi misalnya, tuntutan pemenuhan derajat bebas (*degrees of freedom* atau 'df') model sangat tinggi sekali, bahkan menjadi wajib dalam frekuentis. Dalam estimasi model regresi linear sederhana yang mempunyai dua parameter model, misalnya,  $\beta_0$  sebagai intersep dan  $\beta_1$  sebagai slop model regresi, data pengamatannya harus minimal tiga pasang respons dan prediktornya, agar  $df > 0$ . Atau dengan kata lain  $n > d$ , dengan  $n$  banyaknya pasang data dan  $d$  banyaknya parameter model regresi yang diestimasi. Dalam Bayesian pemodelan regresi linear sederhana seperti ini dapat dilakukan estimasi parameternya dengan baik, walaupun datanya hanya dua pasang ( $n = 2$ ). Analisis Bayesian pada keadaan data demikian, diperlukan memasang prior dengan kategori kuat dalam pemodelannya, sehingga prior lebih mendominasi *likelihood* datanya dan posterior parameternya masih dapat dipercaya.

### 1.2.2 Kompromi antara Bayesianis dan Frekuentis

Baik Bayesian maupun frekuentis pada dasarnya masing-masing memiliki kelebihan dan kekurangan. Little (2006) meringkas kelebihan dan kelemahan dari kedua metode ini menurut dua ukuran, yaitu (1) inferensi berdasarkan model yang diasumsikan; dan (2) formulasi model. Untuk ukuran pertama, jelas bahwa Bayesian lebih superior daripada frekuentis. Namun untuk ukuran kedua, frekuentis lebih unggul dari Bayesian. Tulisan Little (2006) ini sangat terkait dengan pernyataan Box (1980), yaitu bahwa teori *sampling* sangat dibutuhkan untuk mengeksplorasi dan mengkritisi suatu model agar dapat menjelaskan suatu data. Sementara itu, teori Bayes dibutuhkan untuk mengestimasi

parameter bersyarat sesuai kecukupan model. Box (1980) menuangkan idenya ini ke dalam bentuk faktorisasi, yaitu:

$$P(Y, \theta | M) = P(Y | M)P(\theta | Y, M). \quad (1.3)$$

dengan elemen kedua di ruas kanan dalam persamaan (1.3) adalah distribusi posterior dari parameter  $\theta$  yang diberikan data  $Y$  dan model  $M$ , yang juga merupakan dasar untuk membuat inferensi. Sementara itu, elemen pertama di ruas kanan dalam persamaan (1.3) adalah distribusi marginal dari data  $Y$  berdasarkan model  $M$  dan digunakan untuk mendapatkan validitas dari  $M$ , dengan bantuan pemahaman frekuentis. Rubin (1984) menambahkan bahwa penerapan statistika seharusnya menggunakan prinsip-prinsip Bayesian untuk menyelaraskan inferensi dengan kondisi data sebenarnya, akan tetapi, perhitungan frekuentis juga akan sangat bermanfaat untuk menjadikan pernyataan Bayesian menjadi berkaidah *scientific*. *Scientific* yang dimaksudkan adalah kemampuan untuk membuktikan kebenaran/kesalahan berdasarkan uji empiris.

Babu (2012) menjabarkan ide kombinasi antara frekuentis dan Bayesian, yang mana kebanyakan prosedur frekuentis dimanfaatkan dalam analisis Bayesian sebagai berikut:

1. **Komputasi Bayesian:** *Gibbs sampler* dan *Markov Chain Monte Carlo* (MCMC) adalah metode-metode yang digunakan untuk menduga distribusi posterior. Kunci utama dari metode ini adalah kekonvergenan yang sering mengacu pada prosedur kekonvergenan dalam frekuentis.
2. **Pengembangan distribusi prior:** penentuan distribusi prior adalah salah satu bagian yang cukup sulit dalam analisis Bayesian. Sangat jarang ditemukan distribusi prior yang secara akurat dapat merefleksikan semua *prior knowledge* peneliti. Untuk mengatasi hal ini, maka dapat digunakan metode frekuentis untuk membangun distribusi prior. Chib dan Carlin menyebutnya dengan *pseudo prior*.

3. **Konsistensi:** Konsistensi merupakan konsep statistika yang menyatakan bahwa suatu prosedur akan mendekati proses/kondisi yang sebenarnya jika data yang digunakan berukuran besar. Hal ini merupakan salah satu alat estimasi paling sederhana dalam frekuentis yang dapat digunakan oleh Bayesian. Penduga Bayes dikatakan konsisten jika ruang parameter bersifat terhingga. Suatu hal yang tidak dibenarkan jika dalam iterasi komputasi Bayesiannya mempunyai ruang parameter yang bersifat tak hingga.
4. **Estimasi:** dalam permasalahan estimasi parameter standar, metode objektif Bayesian dan frekuentis sering kali menghasilkan jawaban yang sama atau identik. Untuk model linear Gaussian standar, estimasi dan interval konfidensi dari frekuentis sama persis dengan estimasi dan interval kredibel (*credible interval*) dari metode objektif Bayesian standar. Keadaan ini secara umum terjadi jika terdapat sesuatu yang disebut struktur invarian; yaitu ketika suatu estimator distandarisasi, maka distribusinya menjadi bebas dari parameter yang tidak diketahui (termasuk juga adanya *nuisance parameter* –parameter yang tidak menjadi interes untuk diestimasi–).

### 1.2.3 Kritik-Kritik terhadap Bayesian

Bayesian sebagai ilmu/metode yang berkembang pesat dalam beberapa dekade terakhir harus menghadapi tantangan dari para pendukung metode klasik (frekuentis) sebagai pendahulunya. Menurut Sivaganesan (2008), terdapat beberapa kritik yang ditujukan terhadap metode Bayesian ini, yaitu:

1. Dibutuhkan spesifikasi distribusi prior untuk semua parameter,

2. Metode Bayesian sering kali melibatkan integral berdimensi besar, atau dengan kata lain memerlukan proses komputasi yang kompleks.

Berkaitan dengan poin 1: Dalam analisis Bayesian, distribusi prior diatasi dengan menentukannya berdasarkan pada hasil penelitian sebelumnya, publikasi, intuisi peneliti, pendapat para ahli, aspek kenyamanan (*convenience*) dan sumber-sumber lainnya. Namun demikian, sering kali ditemui bahwa pengetahuan untuk menentukan distribusi prior sering kali kurang jelas (kurang meyakinkan –*vague prior*) atau bahkan sama sekali tidak ada informasi yang tersedia (*non-informative prior*) terutama bagi peneliti pemula di bidang yang diteliti dan/atau sebagai pemula pengguna Bayesian. Hal ini mengakibatkan penentuan distribusi prior menjadi dirasa sangat sulit. Selain itu, penentuan prior juga cenderung menjadi sangat subjektif, sebab antarindividu pengguna Bayesian bisa memiliki opini tentang parameter model yang dihadapi berbeda, sehingga distribusi prior yang digunakan juga beda dan pada akhirnya berimbas pada hasil (*distribusi posterior*) yang berbeda pula.

Alasan ini kemudian dibantah oleh para pendukung Bayesian. Bagi Bayesianis, untuk menangani subjektivitas ini sering digunakan “*flat prior*” maupun “*flat conjugate prior*”. Skenario ini memberikan kesempatan yang sama bagi semua kemungkinan nilai parameter dalam domainnya untuk berekspresi mewarnai distribusi atau model yang sedang dipelajari. Cara ini sudah sangat umum digunakan dalam metode Bayesian kontemporer. Akan tetapi, pendekatan “*flat prior*” ini juga bukan tanpa cela. Ketika ruang parameter bersifat tak hingga (*infinite domain*) maka *flat prior* pada hakikatnya sudah tidak menjadi distribusi karena integralnya menjadi tak hingga. Padahal, suatu fungsi distribusi jika diintegrasikan terhadap batas-batasnya seharusnya bernilai 1. Kondisi ini dapat diatasi dengan metode prior non-informatif, *Jeffrey's prior*, *maximum entropy prior* dan *reference prior*.



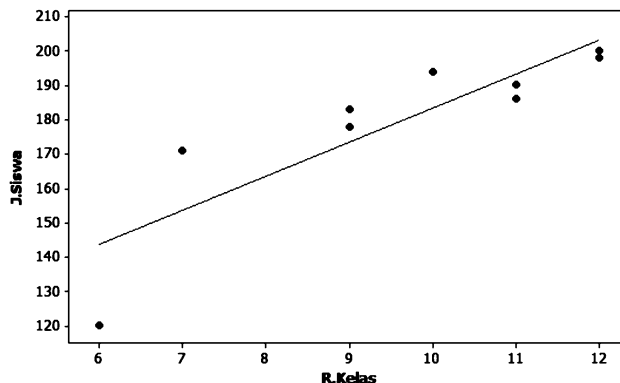
Untuk poin 2: Perlu disadari bahwa banyak fungsi distribusi yang terkadang bersifat tidak *closed form*, sehingga dibutuhkan metode numerik. Kelompok frekuentis pun juga menghadapi masalah untuk mengestimasi parameter distribusi jenis ini. Semakin berkembangnya teknologi komputer, baik perangkat keras maupun dukungan perkembangan perangkat lunaknya, permasalahan semacam ini sudah tidak perlu lagi diperdebatkan. Metode MCMC yang diperkenalkan oleh Geman dan Geman (1984) adalah suatu metode yang dapat digunakan untuk memperoleh integral eksak melalui simulasi berbasis komputasi. MCMC bekerja dengan simulasi stokastik yang menirukan proses Markov dari sebuah distribusi gabungan posterior parameter, dengan *state* prosesnya adalah capaian estimasi parameter satu dan memberikan lingkup suasana (bersyarat) untuk estimasi parameter yang lainnya, dan dilakukan silih berganti secara iteratif sampai semua parameter mencapai kondisi estimasi yang konvergen. Jadi tujuan MCMC adalah melakukan simulasi untuk membakitkan sejumlah besar sampel dari distribusi posterior, lalu menggunakannya untuk menduga nilai-nilai parameter distribusi posterior yang diinginkan.

### 1.3 Bayesian Regresi

Sering dua variabel, namakan  $x$  dan  $y$  berstruktur berpasangan yang diperoleh dari suatu pengamatan sebanyak  $N$ , mengindikasikan adanya hubungan. Generalisasi dari hubungan tersebut ingin dibuatkan persamaan yang mampu menggambarkan keterkaitan antara kedua variabel tersebut. Sering kali besaran nilai variabel  $x$  direncanakan dan selanjutnya digunakan untuk memprediksikan nilai variabel  $y$  dengan menggunakan persamaan sebagai generalisasi hubungan tersebut. Anggap bahwa nilai variabel  $x$  yang direncanakan tersebut telah diketahui dan tidak mempunyai kesalahan pengukuran. Karena nilai variabel  $y$  akan diprediksikan berdasarkan nilai variabel  $x$  yang telah diketahui, maka variabel  $y$  dinamakan sebagai respons atau variabel

dependen, dan variabel  $x$  sebagai variabel prediktor/penentu respons atau variabel independen.

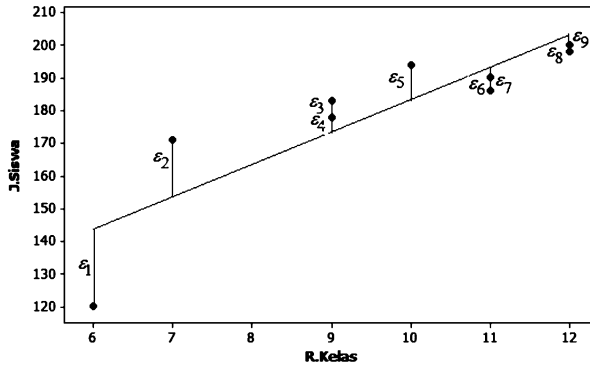
Dengan adanya  $n$  pasang pengamatan, namakan  $(x_i, y_i)$ , dengan  $i = 1, 2, \dots, n$ , maka  $n$  pasang data tersebut dapat di-plot-kan untuk memperlihatkan dugaan adanya kecenderungan hubungannya. Gambar 1.2 memberikan contoh yang menunjukkan keterkaitan hubungan antara nilai-nilai variabel jumlah ruang kelas (diberi nama R. Kelas) di suatu sekolah dasar,  $x$ , dan nilai-nilai variabel jumlah siswa (diberi nama J. Siswa) yang aktif di suatu sekolah dasar,  $y$ . Ketergantungan nilai  $y$  terhadap  $x$  dapat dideteksi melalui besaran nilai korelasi antara dua variabel tersebut yang positif sesuai dengan cerminan arah garis yang mengikuti pergerakan *plot* pasangan data,  $(x_i, y_i)$ ,  $i = 1, 2, \dots, n$ , dan  $n$  adalah jumlah sampel, adalah naik ke kanan atas. Garis ini dapat direpresentasikan sebagai bentuk garis regresi linear sederhana,  $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$  atau dalam bentuk parameterisasi lain  $y_i = \beta_{\bar{x}} + \beta_1 (x_i - \bar{x}) + \varepsilon_i$  dengan  $\beta_{\bar{x}}$  adalah rata-rata nilai  $y$  diberikan nilai  $x = \bar{x}$ ,  $\beta_1$  adalah *slope*, dan  $\varepsilon_i$  adalah error.



**Gambar 1.2** Scatter Plot antara banyak R. Kelas dan J. Siswa

Persamaan garis ini yang akan menerangkan keterkaitan hubungan ketergantungan antara variabel jumlah siswa yang aktif di suatu sekolah dasar,  $y$ , dengan variabel jumlah ruang kelas,  $x$ . Sesuai dengan Gambar 1.3, persamaan tersebut tidak mampu menerangkan secara menyeluruh variabilitas datanya. Setiap titik masih memungkinkan menyimpang dari garis regresi tersebut, karena garis lurus atau linear yang dibangun tidak mungkin melewati setiap titik datanya. Ada tiga asumsi dalam membangun model regresi linear sederhana ini, yaitu (Draper dan Smith, 1998; Gelman dkk., 2004):

1. Asumsi rata-rata model (*mean model*). Asumsi ini memberikan kerangka pikir bahwa  $\mu_{y|x} = \beta_0 + \beta_1 x$ , atau dalam bentuk parameterisasi lain sebagai  $\mu_{y|x} = \beta_{\bar{x}} + \beta_1(x - \bar{x})$ , tidak diketahui dan harus diestimasi dengan menaksir nilai  $\beta_0$  atau  $\beta_{\bar{x}}$  dan  $\beta_1$ .
2. Asumsi setiap penyimpangan data terhadap model (eror model). Besaran penyimpangan setiap titik ke model diasumsikan mengikuti pola distribusi Normal( $0, \sigma^2$ ) atau  $\varepsilon \sim IIDN(0, \sigma^2)$ .
3. Asumsi setiap penyimpangan data terhadap model (eror model) yang saling bebas. Besarnya setiap penyimpangan data pada model adalah diasumsikan saling bebas.



Gambar 1.3 Scatter Plot antara R. Kelas dan J. Siswa

Pemilihan nilai  $\beta_0$  atau  $\beta_{\bar{x}}$  dan  $\beta_1$  harus tepat sehingga nilai penyimpangan,  $\varepsilon_i$ , akan minimum. Oleh sebab itu, cara *Ordinary Least Square* (OLS) atau *Maximum Likelihood Estimator* (MLE) dalam frekuentis akan digantikan dengan cara Bayesian di bawah ini.

### 1.3.1 Likelihood Gabungan antara $\beta_0$ atau $\beta_{\bar{x}}$ dan $\beta_1$

Dengan mengacu pada ketiga asumsi ke dua di atas, maka setiap data ke- $i$  akan membentuk *likelihood* sebagai berikut:

$$L_i(\beta_{\bar{x}}, \beta_1) \propto \exp\left(-\frac{1}{2\sigma^2} [y_i - (\beta_{\bar{x}} + \beta_1(x_i - \bar{x}))]^2\right). \quad (1.4)$$

Penyajian persamaan (1.4) adalah dalam bentuk proporsional (ditulis dengan symbol  $\propto$ ), karena diabaikannya konstanta normalitas dalam distribusi Normal yang tidak memuat  $\beta_{\bar{x}}$  dan  $\beta_1$  pada penyusunan *likelihood* gabungan ini. Selanjutnya, dengan menggunakan asumsi ketiga, bahwa setiap observasi mempunyai penyimpangan,  $\varepsilon_i$ , yang saling bebas, maka bentuk *likelihood* untuk seluruh data sampel adalah sebagai berikut:

$$\begin{aligned}
L_{\text{sampel}}(\beta_{\bar{x}}, \beta_1) &\propto \prod_{i=1}^n \exp\left(-\frac{1}{2\sigma^2} [y_i - (\beta_{\bar{x}} + \beta_1(x_i - \bar{x}))]^2\right) \\
&= \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n [y_i - (\beta_{\bar{x}} + \beta_1(x_i - \bar{x}))]^2\right) \quad (1.5)
\end{aligned}$$

Faktor jumlahan dalam eksponen, namakan sebagai JML, dapat disederhanakan menjadi:

$$\begin{aligned}
\text{JML} &= \sum_{i=1}^n [y_i - \bar{y} + \bar{y} - (\beta_{\bar{x}} + \beta_1(x_i - \bar{x}))]^2 \\
&= \sum_{i=1}^n (y_i - \bar{y})^2 + 2 \sum_{i=1}^n (y_i - \bar{y})(\bar{y} - (\beta_{\bar{x}} + \beta_1(x_i - \bar{x}))) \quad (1.6) \\
&\quad + \sum_{i=1}^n [\bar{y} - (\beta_{\bar{x}} + \beta_1(x_i - \bar{x}))]^2
\end{aligned}$$

dan JML dalam (1.6) dapat disederhanakan dalam bentuk sebagai berikut:

$$SS_y - 2\beta_1 SS_{xy} + \beta_1^2 SS_x + n(\beta_{\bar{x}} - \bar{y})^2,$$

dengan mendefinisikan  $SS_x = \sum_{i=1}^n (x_i - \bar{x})^2$ ,  $SS_y = \sum_{i=1}^n (y_i - \bar{y})^2$ , dan  $SS_{xy} = \sum_{i=1}^n ((y_i - \bar{y})(x_i - \bar{x}))$ . Akibatnya *likelihood* gabungan di persamaan (1.5) dapat dituliskan menjadi sebagai berikut:

$$L_{\text{sampel}}(\beta_{\bar{x}}, \beta_1) \propto \exp\left(-\frac{1}{2\sigma^2} [SS_y - 2\beta_1 SS_{xy} + \beta_1^2 SS_x + n(\beta_{\bar{x}} - \bar{y})^2]\right) \quad (1.7)$$

Persamaan (1.7) dapat didekomposisi dalam bentuk perkalian dua faktor berstruktur eksponen sebagai berikut:

$$L_{\text{sampel}}(\beta_{\bar{x}}, \beta_1) \propto \exp\left(-\frac{1}{2\sigma^2} P\right) \times \exp\left(-\frac{1}{2\sigma^2} Q\right) \quad (1.8)$$

dengan

$$\begin{aligned}
P &= -\frac{1}{2\sigma^2} [SS_y - 2\beta_1 SS_{xy} + \beta_1^2 SS_x] \\
&= -\frac{1}{2\sigma^2 / SS_x} \left[ \beta_1 - \frac{SS_{xy}}{SS_x} \right]^2
\end{aligned}$$

dan

$$Q = -\frac{1}{2\sigma^2} \left[ n(\beta_{\bar{x}} - \bar{y})^2 \right]$$

$$= -\frac{1}{2\sigma^2/n} \left[ (\beta_{\bar{x}} - \bar{y})^2 \right].$$

Untuk lebih menyederhanakan model *likelihood* dan untuk memudahkan mengingat akan proses penghitungan *likelihood* gabungannya, serta dengan mengambil simbol baru bahwa  $SS_{xy} / SS_x = B$  dan  $\bar{y} = B_{\bar{x}}$ , maka persamaan (1.8) dapat dituliskan sebagai berikut:

$$L_{\text{sampel}}(\beta_{\bar{x}}, \beta_1) \propto L_{\text{sampel}}(\beta_{\bar{x}}) \times L_{\text{sampel}}(\beta_1), \quad (1.9)$$

dengan

$$L_{\text{sampel}}(\beta_{\bar{x}}) \propto -\frac{1}{2\sigma^2/n} \left[ (\beta_{\bar{x}} - B_{\bar{x}})^2 \right],$$

dan

$$L_{\text{sampel}}(\beta_1) \propto -\frac{1}{2\sigma^2/SS_x} \left[ \beta - B \right]^2.$$

Persamaan (1.9) yang menampilkan struktur *likelihood* sebagai dua faktor yang terhubung dengan perkalian. Hal ini menyatakan bahwa kedua faktor adalah *likelihood* yang masing-masing adalah independen.

*Likelihood* untuk  $\beta_1$  adalah berbentuk normal dengan rata-rata  $B$  dan varians  $\sigma^2/SS_x$ . Dengan cara yang sama, *likelihood* untuk  $\beta_{\bar{x}}$  adalah berbentuk normal dengan rata-rata  $B_{\bar{x}}$  dan varians  $\sigma^2/n$ .

### 1.3.2 Prior Parameter untuk Model Regresi, $\beta_{\bar{x}}$ dan $\beta_1$

Mengacu pada penjabaran bentuk *likelihood* gabungan antara  $\beta_{\bar{x}}$  dan  $\beta_1$  yang menyatakan saling independen, maka disini akan digunakan prior kedua parameter yang saling independen. Anggap prior untuk  $\beta_{\bar{x}}$  adalah  $p(\beta_{\bar{x}})$  dan prior untuk  $\beta_1$  adalah  $p(\beta_1)$ , maka prior gabungan yang akan digunakan dalam pemodelan Bayesian regresi linear disini adalah sebagai berikut:

$$p(\beta_{\bar{x}}, \beta_1) = p(\beta_{\bar{x}}) \times p(\beta_1), \quad (1.10)$$

dengan  $p(\beta_{\bar{x}})$  maupun  $p(\beta_1)$  dapat menggunakan prior Normal( $a, b^2$ ) atau prior Normal *flat* dengan mengambil nilai parameter  $b^2$  yang sangat besar.

### 1.3.3 Posterior Gabungan untuk Parameter Model Regresi, $\beta_{\bar{x}}$ dan $\beta_1$

Mengacu pada model dasar pembentukan posterior sebagai berikut:

$$\text{Posterior} \propto \text{Likelihood} \times \text{Prior}$$

maka posterior gabungan untuk parameter regresi linear sederhana dapat dibentuk dengan mengalikan prior gabungan dalam persamaan (1.10) dengan *likelihood* gabungan dalam persamaan (1.9) yang dapat dituliskan sebagai berikut:

$$p(\beta_{\bar{x}}, \beta_1 | \text{data\_reg}) \propto L_{\text{sampel}}(\beta_{\bar{x}}, \beta_1) \times p(\beta_{\bar{x}}, \beta_1), \quad (1.11)$$

dengan  $\text{data\_reg}$  adalah himpunan pasangan pengamatan  $(x_1, y_1)$ ,  $(x_2, y_2)$ , ...,  $(x_n, y_n)$ .

Dengan kembali mengacu pada sifat *likelihood* gabungan dan prior

gabungan untuk kedua parameter regresi linear sederhana  $\beta_{\bar{x}}$  dan  $\beta_1$  adalah independen, posterior gabungan dalam persamaan (1.11) dapat disusun oleh perkalian dua posterior marginal untuk  $\beta_{\bar{x}}$  dan  $\beta_1$ , namakan  $p(\beta_{\bar{x}} | \text{data\_reg})$  dan  $p(\beta_1 | \text{data\_reg})$ . Oleh sebab itu, bentuk posterior gabungan dalam persamaan (1.11) dapat dituliskan sebagai berikut:

$$p(\beta_{\bar{x}}, \beta_1 | \text{data\_reg}) \propto p(\beta_{\bar{x}} | \text{data\_reg}) \times p(\beta_1 | \text{data\_reg}). \quad (1.12)$$

### 1.3.4 Marginal Posterior $\beta_1$

Sesuai dengan persamaan (1.10), apabila diambil prior  $p(\beta_1)$  adalah Normal( $a, b^2$ ), maka akan diperoleh posterior marginal untuk  $\beta_1$ , yaitu  $p(\beta_1 | \text{data\_reg})$ , yang berdistribusi Normal( $m_{\beta_1}, s_{\beta_1}^2$ ), dengan

$$\frac{1}{s_{\beta_1}^2} = \frac{1}{b^2} + \frac{SS_x}{\sigma^2} \quad (1.13)$$

dan

$$m_{\beta_1} = \left( \frac{1/b^2}{1/s_{\beta_1}^2} \times a \right) + \left( \frac{SS_x/\sigma^2}{1/s_{\beta_1}^2} \times B \right). \quad (1.14)$$

Persamaan (1.13) menyatakan bahwa presisi posterior  $p(\beta_1 | \text{data\_reg})$

, yaitu  $1/s_{\beta_1}^2$ , adalah senilai dengan presisi prior yang digunakan,

yaitu  $1/b^2$ , ditambah dengan presisi *likelihood* datanya, yaitu  $SS_x/\sigma^2$



. Sementara itu, rata-rata posterior  $p(\beta_1 | \text{data\_reg})$  terbentuk oleh dua suku, yaitu:

1. Suku Rata-Rata Prior

Suku ini senilai dengan rata-rata parameter prior, yaitu  $a$ , yang diboboti oleh proporsi antara presisi prior, yaitu  $1/b^2$ , dengan presisi posteriornya, yaitu  $1/s_{\beta_1}^2$ , sehingga, suku pertama ini adalah

$$\left( \frac{1/b^2}{1/s_{\beta_1}^2} \times a \right).$$

2. Suku Rata-Rata Posterior

Suku ini senilai dengan rata-rata parameter posterior, yaitu  $B$  di persamaan (1.9), yang diboboti oleh proporsi antara presisi *likelihood* rata-ratanya, yaitu  $SS_x/\sigma^2$ , dengan presisi posteriornya,

yaitu  $1/s_{\beta_1}^2$ . Sehingga, suku kedua ini adalah

$$\left( \frac{SS_x/\sigma^2}{1/s_{\beta_1}^2} \times B \right).$$

### 1.3.5 Marginal Posterior $\beta_{\bar{x}}$

Dengan cara yang sama, penghitungan posterior parameter *intercept*,  $\beta_{\bar{x}}$ , dapat dilakukan dengan menentukan priornya terlebih dahulu dan kemudian dihitung posteriornya. Jika prior dipilih Normal  $(c, d^2)$ , maka

akan diperoleh posterior untuk  $\beta_{\bar{x}}$  berupa Normal  $(m_{\beta_{\bar{x}}}, s_{\beta_{\bar{x}}}^2)$ , dengan nilai presisi dan rata-rata posteriornya adalah sebagai berikut:

$$\frac{1}{s_{\beta_{\bar{x}}}^2} = \frac{1}{d^2} + \frac{n}{\sigma^2} \quad (1.15)$$

dan

$$m_{\beta_{\bar{x}}} = \left( \frac{1/b^2}{1/s_{\beta_{\bar{x}}}^2} \times c \right) + \left( \frac{n/\sigma^2}{1/s_{\beta_{\bar{x}}}^2} \times B_{\bar{x}} \right). \quad (1.16)$$

Nilai  $B_{\bar{x}}$  mengacu pada persamaan (1.9).

### 1.3.6 Bayesian Credible Interval dan Pengujian Hipotesis

Distribusi posterior  $\beta_1$  memberikan hasil perpaduan antara *likelihood* data dengan prior yang dipasangkannya. Posterior ini memberikan kepercayaan rentang nilai  $\beta_1$  setelah peng-*update*-an prior oleh data dalam *likelihood*. Ringkasan rentang posterior  $\beta_1$  ini diberikan dalam bentuk  $(1 - \alpha) \times 100\%$  Bayesian *credible interval*, yaitu berupa:

$$m_{\beta_1} \pm z_{\frac{\alpha}{2}} \times \sqrt{s_{\beta_1}^2}. \quad (1.17)$$

Sebenarnya nilai  $\sigma^2$  tidak diketahui dengan pasti. Oleh sebab itu, nilai ini ditaksir melalui nilai residual model, yaitu  $\varepsilon_i = y_i - (B_{\bar{x}} + B(x_i - \bar{x}))$ .

Oleh sebab itu taksiran nilai  $\sigma^2$  menjadi:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n [y_i - (B_{\bar{x}} + B(x_i - \bar{x}))]^2}{n - 2}.$$

Dengan tidak diketahuinya sebaran data yang sebenarnya, maka desain rentang *credible interval*, agar mampu menerangkan variabilitas posterior, menggunakan distribusi *t-Student* pada derajat bebas  $n-2$ , sehingga *credible interval* di persamaan (1.17) menjadi

$$m_{\beta_1} \pm t_{(n-2), \frac{\alpha}{2}} \times \sqrt{s_{\beta_1}^2}. \quad (1.18)$$

Perbedaannya jika interval nilai estimasi  $\beta_1$  ini dilakukan menggunakan *confidence interval* dalam metode frekuentis akan diperoleh sebagai berikut:

$$B \pm t_{(n-2), \frac{\alpha}{2}} \times \frac{\hat{\sigma}^2}{\sqrt{SS_x}}, \quad (1.19)$$

dengan  $\hat{\sigma}^2$  dihitung dari varians residual setiap data terhadap garis yang dibangun dengan metode kuadrat terkecil (*least square*).

Pengujian hipotesis terhadap nilai estimasi parameter dalam regresi linear sederhana menggunakan *credible interval* ini adalah dengan melihat apakah nilai NOL berada di dalam *credible interval* tersebut. Jika nilai NOL berada di dalam interval tersebut, maka nilai estimasi parameter tersebut secara statistik dapat dianggap bernilai NOL. Hal ini dapat pula dikatakan bahwa hipotesis NULL gagal ditolak.

### 1.3.7 Distribusi dan Observasi Prediktif

Setelah model regresi antara respons dan independen variabel terbentuk, selanjutnya pemanfaatan model antara lain adalah untuk menentukan distribusi dan observasi prediktifnya. Hal ini dapat dilakukan jika diberikan suatu nilai pengamatan independen variabel

$x_{n+1}$  tertentu, maka prediktif respons  $\tilde{y}_{n+1}$  adalah

$$\tilde{y}_{n+1} = \hat{\beta}_x + \hat{\beta}_1 (x_{n+1} - \bar{x}), \quad (1.20)$$

dengan  $\hat{\beta}_1$  adalah nilai taksiran slop dan  $\hat{\beta}_{\bar{x}}$  adalah nilai taksiran intersep pada saat  $x = \bar{x}$ .

Bagaimanakah ketepatan nilai prediktif  $\tilde{y}_{n+1}$ ? Pernyataan jawaban dari pertanyaan ini dapat diberikan dengan dua penjelasan sebagai berikut. Pertama, dengan melihat symbol yang digunakan adalah  $\tilde{y}_{n+1}$  bukan  $y_{n+1}$ . Hal ini memberikan arti bahwa dengan memanfaatkan nilai estimasi parameter  $\hat{\beta}_1$  dan  $\hat{\beta}_{\bar{x}}$  dalam persamaan (1.20) diperoleh nilai prediksi respons  $\tilde{y}_{n+1}$ , bukan nilai respons yang sebenarnya yang tidak diketahui. Selain itu, bahwa kedua nilai estimasi parameter  $\hat{\beta}_1$  dan  $\hat{\beta}_{\bar{x}}$  adalah hasil penghitungan dari distribusi posterior yang bersifat random dan mempunyai nilai probabilitas untuk muncul, sehingga nilai prediksi respons  $\tilde{y}_{n+1}$  merupakan respons yang mempunyai probabilitas untuk diperoleh. Kedua, bahwa setiap pemberian nilai  $x_{n+1}$  ke dalam persamaan (1.20), model akan mempunyai kontribusi error sebesar  $\mathcal{E}_{n+1}$  yang bersifat random pula. Oleh sebab itu nilai prediksi  $\tilde{y}_{n+1}$  dapat digunakan sebagai nilai prediksi respons, jika pengamatan  $x_{n+1}$  diberikan.

Observasi baru, respons ke  $(n+1)$  atau  $y_{n+1}$  diberikan nilai estimasi parameter modelnya  $\hat{\beta}_1$  dan  $\hat{\beta}_{\bar{x}}$  serta diketahui pengamatan  $x_{n+1}$  dan 'data' observasi yang digunakan untuk membangun modelnya, yaitu  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ , akan mempunyai distribusi, namakan  $f(y_{n+1} | x_{n+1}, data)$ . Distribusi ini dapat diperoleh dengan

menghilangkan parameter pengganggu dalam modelnya,  $\hat{\beta}_1$  dan  $\hat{\beta}_{\bar{x}}$ , yaitu dengan mengintegalkan terhadap parameter pengganggu tersebut sebagai berikut:

$$f(y_{n+1} | x_{n+1}, data) = \iint f(y_{n+1}, \beta_{\bar{x}}, \beta_1 | x_{n+1}, data) d\beta_{\bar{x}} d\beta_1. \quad (1.21)$$

Proses demikian ini merupakan salah satu keuntungan cara Bayesian yang dinamakan sebagai proses marginalisasi. Dalam proses marginalisasi ini, parameter model diposisikan sebagai parameter pengganggu.

Langkah yang harus ditempuh untuk mencari distribusi observasi baru  $y_{n+1}$  adalah dengan cara membentuk distribusi posterior gabungan antara  $y_{n+1}$  dengan parameter  $\hat{\beta}_1$  dan  $\hat{\beta}_{\bar{x}}$ , yaitu:

$$f(y_{n+1}, \beta_{\bar{x}}, \beta_1 | x_{n+1}, data) = f(y_{n+1} | \beta_{\bar{x}}, \beta_1, x_{n+1}, data) \times g(\beta_{\bar{x}}, \beta_1 | x_{n+1}, data). \quad (1.22)$$

Perhatikan kembali bahwa observasi baru  $y_{n+1}$  adalah random yang hanya tergantung pada nilai  $x_{n+1}$  serta nilai random posterior  $\hat{\beta}_1$  dan  $\hat{\beta}_{\bar{x}}$  dan tidak tergantung pada 'data'. Tetapi, nilai random posterior  $\hat{\beta}_1$  dan  $\hat{\beta}_{\bar{x}}$  sangat tergantung pada 'data'. Oleh sebab itu, bentuk distribusi gabungan posterior antara observasi baru  $y_{n+1}$  dengan parameter  $\hat{\beta}_1$  dan  $\hat{\beta}_{\bar{x}}$  dapat disusun sebagai berikut:

$$f(y_{n+1}, \beta_{\bar{x}}, \beta_1 | x_{n+1}) = f(y_{n+1} | \beta_{\bar{x}}, \beta_1, x_{n+1}) \times g(\beta_{\bar{x}}, \beta_1 | data) \quad (1.23)$$

yaitu senilai dengan perkalian antara probabilitas observasi baru diketahui parameter model dikalikan dengan distribusi posterior parameter

model diberikan 'data' pengamatan sebelumnya. Jadi,  $y_{n+1}$  adalah observasi baru dari suatu model regresi yang random dan mengikuti distribusi Normal dengan parameter *mean*  $\mu_{n+1} = \beta_{\bar{x}} + \beta_1(x_{n+1} - \bar{x})$  dan varians  $\sigma^2$  atau dapat dituliskan sebagai berikut

$$\begin{aligned} y_{n+1} &\sim N(\beta_{\bar{x}} + \beta_1(x_{n+1} - \bar{x}), \sigma^2) \\ &\sim N(\mu_{n+1}, \sigma^2). \end{aligned} \quad (1.24)$$

Jika posterior  $\hat{\beta}_{\bar{x}}$  berdistribusi Normal( $m_{\beta_{\bar{x}}}, s_{\beta_{\bar{x}}}^2$ ) dan posterior  $\hat{\beta}_1$  berdistribusi Normal( $m_{\beta_1}, s_{\beta_1}^2$ ), maka  $\mu_{n+1}$  akan berdistribusi Normal( $m_{\mu_{n+1}}, s_{\mu_{n+1}}^2$ ), dengan

$$\begin{aligned} m_{\mu_{n+1}} &= m_{\beta_{\bar{x}}} + (x_{n+1} - \bar{x}) \times m_{\beta_1} \\ s_{\mu_{n+1}}^2 &= s_{\beta_{\bar{x}}}^2 + (x_{n+1} - \bar{x}) \times s_{\beta_1}^2 \end{aligned} \quad (1.25)$$

Jika distribusi posterior nilai observasi baru diproses dari persamaan (1.24), maka harus dilakukan proses marginalisasi dari distribusi gabungan posterior antara  $y_{n+1}$  dan  $\mu_{n+1}$  sebagai berikut:

$$\begin{aligned} f(y_{n+1} | x_{n+1}, data) &= \int f(y_{n+1} | \mu_{n+1}, x_{n+1}, data) \\ &\quad \times g(\mu_{n+1} | x_{n+1}, data) d\mu_{n+1} \\ &= \int f(y_{n+1} | \mu_{n+1}) \times g(\mu_{n+1} | x_{n+1}, data) d\mu_{n+1} \\ &\propto \int e^{-\frac{1}{2\sigma^2}(y_{n+1} - \mu_{n+1})^2} \times e^{-\frac{1}{2s_{\mu_{n+1}}^2}(\mu_{n+1} - m_{\mu_{n+1}})^2} d\mu_{n+1} \quad (1.26) \\ &\propto \int e^{-\frac{\sigma^2 s_{\mu_{n+1}}^2}{2\sigma^2 s_{\mu_{n+1}}^2} \left( \mu_{n+1} - \frac{y s_{\mu_{n+1}}^2 + m_{\mu_{n+1}} \sigma^2}{s_{\mu_{n+1}}^2 + \sigma^2} \right)^2} \\ &\quad \times e^{-\frac{1}{2(s_{\mu_{n+1}}^2 + \sigma^2)}(y_{n+1} - m_{\mu_{n+1}})^2} d\mu_{n+1} \end{aligned}$$

Faktor kedua dalam persamaan (1.26) tidak memuat faktor integral  $\mu_{n+1}$ , sehingga penyelesaian integral persamaan tersebut menjadi sebagai berikut

$$f(y_{n+1} | x_{n+1}, data) \propto \exp\left(-\frac{1}{2(s_{\mu_{n+1}}^2 + \sigma^2)}(y_{n+1} - m_{\mu_{n+1}})^2\right) \quad (1.27)$$

atau dapat dituliskan

$$y_{n+1} \sim N(m_{\mu_{n+1}}, s_{\mu_{n+1}}^2 + \sigma^2). \quad (1.28)$$

## 1.4 Bayes Factor

Di dalam metode Bayesian, *Bayes Factor* merupakan metode alternatif hipotesis frekuentis. *Bayes Factor* mengukur bukti dukungan suatu model terhadap model lainnya. Sebagai contoh, terdapat suatu data  $\mathbf{y}$  yang diasumsikan berasal dari salah satu hipotesis  $H_1$  untuk model  $M_1$  dan  $H_2$  untuk model  $M_2$ . Masing-masing hipotesis ini memiliki peluang densitas  $p(\mathbf{y} | H_1)$  dan  $p(\mathbf{y} | H_2)$  serta diketahui peluang *prior*  $p(H_1)$  dan  $p(H_2)$ . Berdasarkan teorema Bayes, diperoleh peluang *posterior*  $p(H_1 | \mathbf{y})$  dan  $p(H_2 | \mathbf{y})$

$$p(H_k | \mathbf{y}) = \frac{p(\mathbf{y} | H_k)p(H_k)}{\sum_{k=1}^2 p(\mathbf{y} | H_k)p(H_k)}$$

dengan  $k=1,2$ , sehingga rasio  $p(H_2 | \mathbf{y})$  terhadap  $p(H_1 | \mathbf{y})$  adalah

$$\frac{p(H_2 | \mathbf{y})}{p(H_1 | \mathbf{y})} = \frac{p(\mathbf{y} | H_2)}{p(\mathbf{y} | H_1)} \times \frac{p(H_2)}{p(H_1)}$$

dengan:

$$\frac{p(H_2 | \mathbf{y})}{p(H_1 | \mathbf{y})} = \text{Posterior Odds,}$$

$$\frac{p(\mathbf{y} | H_2)}{p(\mathbf{y} | H_1)} = \frac{\int_{\Theta_2} p(\theta_2 | H_2) p(\mathbf{y} | \theta_2, H_2) d\theta_2}{\int_{\Theta_1} p(\theta_1 | H_1) p(\mathbf{y} | \theta_1, H_1) d\theta_1} \quad (1.29)$$

$$= BF_{21} = \text{Bayes Factor}$$

dan

$$\frac{p(H_2)}{p(H_1)} = \text{Prior Odds.}$$

$p(\mathbf{y} | H_1)$  dan  $p(\mathbf{y} | H_2)$  pada persamaan (1.29) adalah *marginal likelihood* model  $M_1$  dan  $M_2$ . Penghitungan *marginal likelihood* ini secara langsung sulit dilakukan. Beberapa peneliti menghitung nilai *marginal likelihood* ini berdasarkan *output* MCMC yaitu di antaranya Chib (1995) berdasarkan *output Gibbs sampler*, Chib dan Jeliazkov (2001) berdasarkan *output* Metropolis Hasting serta Bartolucci dkk (2006) berdasarkan *output reversible jump*. Pendekatan lain dilakukan oleh Carlin dan Chib (1995) dengan menggunakan *pseudoprior* dan Iriawan dan Wright (1997) mengusulkan Struktur Perkalian Distribusi (SPD). Carlin dan Chib (1995) serta Iriawan dan Wright (1997) menghitung *Bayes Factor* menggunakan algoritma *Gibbs sampler*, tetapi Carlin dan Chib (1995) hanya untuk perbandingan dua model sedangkan SPD dapat untuk membandingkan beberapa model. Interpretasi *Bayes Factor* telah dikemukakan oleh Jeffrey (1961) dan disarikan kembali oleh Kass dan Raftery (1995) sesuai Tabel 1.1.



**Tabel 1.1** Intepretasi *Bayes Factor* dalam Pemilihan Hipotesis

Log $B_{21}$	$B_{21}$	Bukti Keandalan $H_2$ terhadap $H_1$ Berdasarkan Data
0-0,5	1-3,2	Tidak Ada
0.5-1	3,2 -10	CukupKuat
1-2	10-100	Kuat
>2	>100	SangatKuat

Sumber: Kass dan Raftery (1995)

Nilai *Bayes Factor* dapat didekati dengan *Pseudo Bayes Factor*

$$PBE_{21} = \exp(LPML_2 - LPML_1)$$

dengan

$$LPML_1 = \sum_{i=1}^n \log(p_i(y_i | y_{-i}, H_1))$$

dan

$$LPML_2 = \sum_{i=1}^n \log(p_i(y_i | y_{-i}, H_2)).$$

$p_i(y_i | y_{-i}, H_1)$  disebut sebagai *Conditional Predictive Ordinate (CPO)* ke- $i$  pada model  $M_1$  yang merupakan densitas prediktif berdasarkan semua data kecuali pengamatan ke- $i$  ( $y_i$ ). Gelfand dan Dey (1994) menunjukkan bahwa  $CPO_i$  dan  $LPML_1$  atau *Log Pseudo-Marginal Likelihood* pada model  $M_1$  mudah diestimasi dari *output* MCMC posterior,

$$CPO_1^{-1} = \frac{1}{s} \sum_{k=1}^s \frac{1}{p_i(y_i | \theta^k, H_1)}.$$

Demikian juga  $CPO_2$  dan  $LPML_2$  pada model  $M_2$ .

## 1.5 Komputasi Bayesian

Dalam banyak pemodelan dengan pendekatan Bayesian, distribusi posterior gabungan adalah berdimensi banyak, kompleks dan tidak bersifat *close-form*. Metode Bayesian memerlukan berbagai karakteristik posterior, seperti momen dan kuantil. Ini memerlukan operasi integral yang tidak selalu mudah diselesaikan. Metode MCMC sering digunakan untuk menangani permasalahan ini.

MCMC adalah metode yang populer untuk membuat simulasi dari distribusi *posterior*. Strategi pengambilan sampel MCMC menghasilkan Markov Chain yang *positive recurrent*, *irreducible* dan *aperiodic* yang berarti nilai parameter yang dibangkitkan konvergen pada distribusi stasioner. Jika kondisi ini terpenuhi, distribusi parameter yang dihasilkan MCMC sama dengan distribusi *posterior* yang diinginkan (Robert, 1996).

### 1.5.1 Bayesian Inference Using Gibbs Sampling (BUGS)

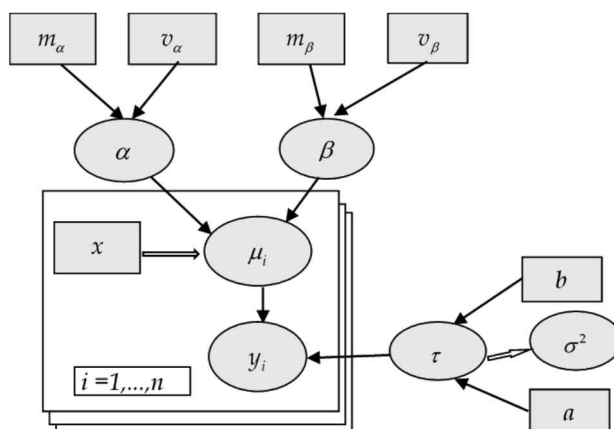
Salah satu algoritma MCMC yang banyak dipakai adalah algoritma *Gibbs sampler*. Algoritma ini digunakan oleh perangkat lunak BUGS. BUGS mengimplementasikan metode Bayesian ke dalam bentuk bahasa program sehingga pengguna dimudahkan ketika menggunakan metode Bayesian. Karena kemudahannya tersebut, BUGS menjadi populer, baik dalam komunitas akademik maupun nonakademik. Dalam perkembangannya, BUGS telah diimplementasikan ke berbagai perangkat lunak, yaitu WinBUGS, OpenBUGS, JAGS dan *Nimble*, yang merupakan perangkat lunak “Open Source”. Selain itu, BUGS menginspirasi pembuatan perangkat lunak serupa, misalnya *stan* ([mc-stan.org/](https://mc-stan.org/)) dan *mamba* (<https://github.com/brian-j-smith/Mamba.jl/>).

Kemudahan BUGS ditunjang oleh keluwesannya (*flexibility*), yang disebabkan oleh penggunaan teori pemodelan grafis (*graphical modelling theory*). Misalkan ditentukan model regresi linear sederhana:

$$y_i \sim N(\mu_i, \sigma^2), \quad \mu_i = \alpha + \beta x_i, \quad i = 1, \dots, n$$

untuk nilai respons  $y_i$  dan nilai variabel prediktor  $x_i$  ( $i = 1, \dots, n$ ). Berdasarkan metode Bayesian, distribusi prior parameter  $\alpha$ ,  $\beta$  dan  $\sigma^2$  harus ditentukan, misalkan *node*

$\alpha \sim N(m_\alpha, v_\alpha)$ ,  $\beta \sim N(m_\beta, v_\beta)$ , dan  $\sigma^2 \sim \text{InversGamma}(a, b)$  (atau  $\tau \sim \text{Gamma}(a, b)$  dengan  $\tau = 1 / \sigma^2$ ), dengan  $m_\alpha, v_\alpha, m_\beta, v_\beta, a$  dan  $b$  suatu bilangan konstan.



**Gambar 1.4 Directed Acyclic Graph (DAG) Model Regresi Sederhana**

Penggambaran model dapat dijelaskan ke dalam bentuk *directed acyclic graph* (DAG) seperti ditunjukkan oleh Gambar 1.4.

Setiap kuantitas di dalam model dinyatakan dalam *node*, sedangkan *link* antar-*node* menunjukkan ketergantungan langsung.  $\alpha$ ,  $\beta$ , dan  $\tau$  adalah *stochastic node* setiap  $y_i$ , sedangkan  $\mu_i$  merupakan *logical node*. Diagram ini disebut *directed* karena setiap *link* adalah sebuah anak panah (ada anak panah tunggal sebagai *directed stochastic* dan ada anak panah ganda sebagai *directed logical*); diagram ini adalah *acyclic* karena hanya satu arah mengikuti arah anak panah dan tidak mungkin kembali ke sebuah *node* setelah meninggalkannya (Lunn dkk., 2009).

*Node* persegi panjang merupakan *node* yang bernilai konstan yang diketahui. *Node* elips menunjukkan hubungan deterministik (fungsi/persamaan) atau kuantitas stokastik, yaitu kuantitas yang memerlukan asumsi distribusi tertentu. Hubungan stokastik dan hubungan deterministik masing-masing ditunjukkan dengan anak panah bergaris tunggal dan anak panah bergaris ganda. Struktur repetitif seperti *loop* dari  $i = 1$  sampai dengan  $n$  ditunjukkan oleh *plate* yang mungkin bisa bersarang jika modelnya berhierarki.

Seperti telah disebutkan sebelumnya, BUGS merupakan bahasa program. Bahasa ini dirancang untuk meniru spesifikasi matematis suatu model. Hubungan stokastik dilambangkan dengan “~” dan hubungan deterministik/logikal dilambangkan dengan “<-“. Struktur repetitif ditunjukkan dengan “for-loop” yang mungkin bersarang jika modelnya berhierarki. Syntax pada Gambar 1.5 menjelaskan model regresi linear sederhana dengan asumsi *error* berdistribusi Normal.

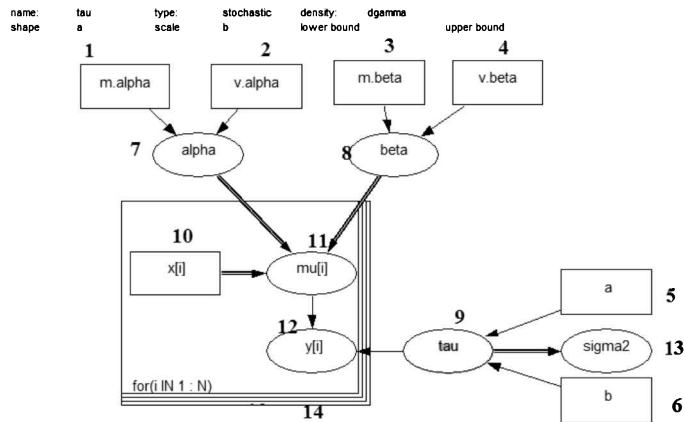
```

model;
{
  for( i in 1 : n ) {
    y[i] ~ dnorm(mu[i],tau)
  }
  for( i in 1 : n ) {
    mu[i] <- alpha + beta * x[i]
  }
  alpha ~ dnorm( m.alpha,v.alpha)
  beta ~ dnorm( m.beta,v.beta)
  tau ~ dgamma(a,b)
  sigma2 <- 1 / tau
}

```

**Gambar 1.5** *Syntax* BUGS model regresi sederhana.






DAG di WinBUGS serta OpenBUGS dibuat dalam doodle. Sebagai contoh, berikut ini cara pembuatan doodle yang menggambarkan model regresi berganda dengan asumsi eror berdistribusi Normal  $(0, \sigma^2)$ . DAG hasil dari Doodle-nya WinBUGS disajikan di Gambar 1.6.



**Gambar 1.6 Doodle DAG Regresi Berganda yang Dibuat di WinBUGS.**

Berikut penjelasan mengenai tahapan membuat doodle di Gambar 1.6:

1. Buat *node* `m.alpha` sebagai *node* parameter *mean* dari *node* `alpha`.
  - a. Klik *mouse* kiri di *window* doodle, sehingga muncul *node* kotak .
  - b. Isi **name** dengan **m.alpha** sehingga tampil `m.alpha`.
  - c. Klik **type** dan kemudian pilih **constant**.
2. Buat *node* `v.alpha` dengan cara seperti tahap 1 sebagai *node* parameter *precisian* dari *node* `alpha`.
3. Buat *node* `m.beta` dengan cara seperti langkah sebelumnya.
4. Buat *node* `v.beta`.
5. Buat *node* `a`.
6. Buat *node* `b`.

7. Buat *node* prior **alpha**.
  - a. Klik *mouse* kiri di *window* doodle, sehingga muncul *node* elips .
  - b. Isi **name** dengan **alpha**
  - c. Isian **type** adalah **stochastic**
  - d. Klik **density** dan kemudian pilih **dnorm**
  - e. Tekan dan tahan tombol keyboard **Ctrl** dan klik *node* **m.alpha** untuk *mean*-nya. Setelah itu, tekan dan tahan tombol keyboard **Ctrl** dan klik *node* **v.alpha** untuk presisinya.
8. Buat *node* prior **beta** dengan cara seperti langkah 7 dengan **m.beta** untuk *mean*-nya dan **v.beta** untuk presisinya.
9. Buat *node* **tau** yang merupakan presisi model dengan cara:
  - a. Klik *mouse* kiri di *window* doodle, sehingga muncul *node* elips .
  - b. Isi **name** dengan **tau**
  - c. Klik **type** dan kemudian pilih **stochastic**
  - d. Pilih **density**-nya dengan **dgamma**
  - e. Tekan dan tahan tombol keyboard **Ctrl** dan klik *node* **a** untuk *shape*-nya. Setelah itu, tekan dan tahan tombol keyboard **Ctrl** dan klik *node* **b** untuk *scale*-nya.
10. Buat *node* data variabel bebas .
  - a. Klik *mouse* kiri di *window* doodle, sehingga muncul *node* kotak .
  - b. Isi **name** dengan **x[i]**
  - c. Klik **type** dan kemudian pilih **constant**
11. Buat *node* **mu[i]** dengan cara:
  - a. Klik *mouse* kiri di *window* doodle, sehingga muncul *node* elips .


- b. Isi **name** dengan **mu[i]**
- c. Klik **type** dan kemudian pilih **logical**
- d. Isikan **value** dengan **alpha+beta\*x[i]**

**Catatan:** Isian ini merupakan persamaan regresi linear


$$\hat{y}_i = \mu_i = \alpha + \beta x_i$$

- e. Koneksikan **node mu[i]** dengan **node-node** yang ada di **value** (persamaan regresi), sehingga muncul anak panah, dengan mengklik **node mu[i]** sehingga berwarna abu-abu dan menekan dan tahan tombol keyboard **Ctrl** , kemudian klik pada **node alpha** dan **beta**.

12. Buat **node y[i]**

- a. Klik **mouse** kiri di **window** doodle, sehingga muncul **node** elips .
- b. Isi **name** dengan **y[i]**
- c. Isian **type** adalah **stochastic**
- d. Klik **density** dan kemudian pilih **dnorm**. Isian **mean** dan **precision** dengan cara: (i) Klik **node y[i]** sehingga berwarna abu-abu, (ii) Tekan dan tahan tombol keyboard **Ctrl** dan klik pada **node mu[i]** dan **tau** secara berurutan

13. Buat **node sigma2** yang merupakan varians.

- a. Klik **mouse** kiri di **window** doodle, sehingga muncul **node** elips .
- b. Isi **name** dengan **sigma2**
- c. Klik **type** dan kemudian pilih **logical**
- d. Isikan **value** dengan **1/tau**
- e. Koneksikan **node sigma2** dengan **node** yang ada di **value (tau)** sehingga muncul anak panah: (i) Klik **node**

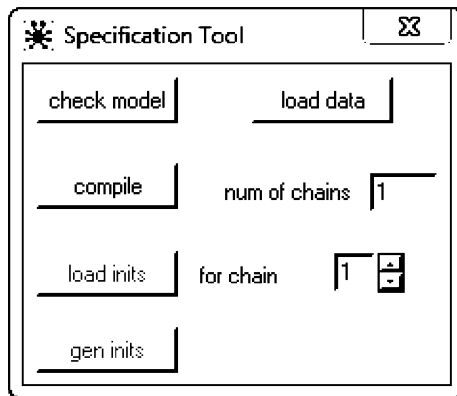
**sigma2** sehingga berwarna abu-abu, (ii) Tekan dan tahan tombol keyboard **Ctrl** dan klik pada **node tau**

14. Buat *plate* dengan cara:

- a. Tekan dan tahan tombol keyboard **Ctrl** dan klik *mouse* kiri di area *window* doodle. Isikan **indeks** dengan **i**, **from** dengan **1** dan **to** dengan **n**.
- b. Ubah ukuran *plate* dengan cara klik *mouse* kiri, tahan pada pojok kanan bawah *plate* dan geser *mouse* sehingga lebarnya mencakup semua *node* yang terdapat indeks-i.
- c. Geser *plate* sehingga semua *node* yang berindeks **i** berada di dalam *plate*. Ini dilakukan dengan cara: klik *mouse* kiri pada pinggir *plate* yang tebal lalu geser ke arah yang diinginkan.

BUGS dijalankan pada program WinBUGS dengan tahapan sebagai berikut:

1. Pemeriksaan *syntax* (kode) model yang dibuat
  - a. Pilih menu **Model** dan pilih **Spesification** sehingga muncul *window* di Gambar 1.7



**Gambar 1.7** *Window Specification Tool*



- b. Blok kata "model" atau tempatkan kursor di bagian kode.
- c. Tekan tombol **check model**. Jika kode sudah benar, di bagian kiri bawah (*status bar*) WinBUGS ada kalimat.

model is syntactically correct

2. Pembacaan data yang digunakan:

Misalnya akan digunakan data sebagaimana disajikan pada Tabel 1.2 berikut.

**Tabel 1.2. Contoh Data Regresi Sederhana**

Y	X
10	8,04
8	6,95
13	7,58
9	8,81
11	8,33
14	9,96
6	7,24
4	4,26
12	10,84
7	4,82
5	5,68

Sumber: Anscombe (1973)

Data tersebut selanjutnya disusun dalam format WinBUGS seperti berikut:

list(y=c(10,8,13,9,11,14,6,4,12,7,5), x=c(8.04,6.95, 7.58,8.81,8.33,9.96,7.24,4.26,10.84,4.82,5.68))

Data tersebut diproses dengan cara:

- a. Blok kata **list** di data (klik ganda *mouse* kiri dengan cepat) di kata **list**.
- b. Klik tombol **load data** di *window* Gambar 1.7.

Jika format data yang dimasukkan sudah benar, maka muncul kalimat **data loaded** di *status bar* program WinBUGS.

### 3. *Compiling* model BUGS

*Compiling* berarti menyusun suatu representasi internal dan menentukan metode pembangkitan sampel yang digunakan untuk setiap node. Apabila diinginkan ada lebih dari satu **chain** agar bisa dilakukan pengecekan konvergensi Gelman dan Rubin, pada *window* Gambar 1.7, isikan jumlah **chain** yang diinginkan. Jika tidak, isian **chain** dibiarkan 1 saja. Setelah itu klik tombol **compile**.

### 4. Menentukan inisial

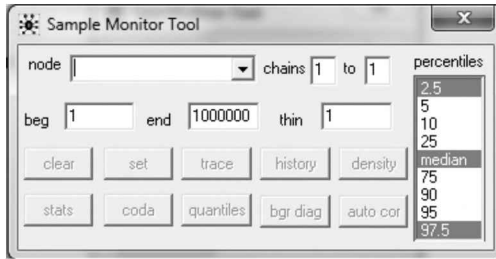
Inisial merupakan angka sembarang yang merupakan nilai awal dari nilai parameter yang akan dibangkitkan. Jika nilai awal ini jauh dari kisaran nilai parameter, proses *burn in* perlu dilakukan. Hasil bangkitan nilai parameter yang belum mencapai *burn in* tidak dipakai untuk inferensi parameter. Misalkan ditentukan inisial sebagai berikut:

`list(alpha=0, beta=0),tau=1)`

Inisial tersebut dimasukkan dengan cara:

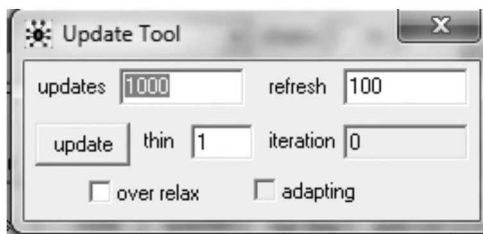
- a. blok kata `list`
- b. klik tombol **load init** di *window*

5. Menentukan parameter yang akan dilakukan inferensi
  - a. Pilih menu **Inference-Sample** sehingga muncul *window* **Sample Monitor Tool** yang disajikan di Gambar 1.8.



**Gambar 1.8** *Window* **Sample Monitor Tool**

- b. Di *node*, isikan/ketikkan node parameter yang akan dikaji, misalkan dalam contoh di atas, ketik **alpha**, lalu tekan tombol set. Apabila akan dilakukan proses burn in, tahap ini dilewati dulu. Setelah *update* pada iterasi tertentu yang dianggap *burn in*, tahapan ini baru dilakukan.
6. Menjalankan kode BUGS (*update*) sesuai jumlah iterasi yang ditentukan
  - a. Pilih menu **Model-Update** sehingga muncul *window* **Update Tool** yang disajikan pada Gambar 1.9.



**Gambar 1.9** *Window* **Update Tool**

- b. Pada updates, isikan jumlah iterasi yang diinginkan, misalkan 10.000.
- c. Isikan nilai **thin**. **Thinning** berguna mengurangi *autocorrelation* hasil MCMC sehingga hasil MCMC tersebut bisa konvergen. Nilai **thin** dapat mengikuti isian *default* (angka 1). Namun, apabila hasil MCMC tidak konvergen, nilai **thin** ini dapat diganti sesuai *lag* iterasi yang *autocorrelation*-nya rendah.
- d. *Over relax* dipilih jika konvergensi tidak dicapai, tetapi pilihan ini tidak selalu membantu untuk mendapatkan konvergensi. *Adapting* diaktifkan (oleh WinBUGS) jika *full conditional distribution posterior* tidak *close form* sehingga menggunakan algoritma Metropolis atau *slice* di dalam bagian algoritma Gibbs Sampler.
- e. Klik **update** untuk menjalankan kode BUGS yang sudah di-*compile*.

Sesuai dengan namanya, BUGS menggunakan algoritma *Gibbs sampler*, yang merupakan salah satu algoritma MCMC. Namun, algoritma ini tidak berdiri sendiri. Beberapa algoritma lain juga digunakan ketika distribusi posterior bersyarat penuh tidak *close-form*.

Beberapa algoritma yang digunakan oleh BUGS dapat dilihat di Tabel 1.3. Ketika distribusi target *closed form*, algoritma yang standar dapat diaplikasikan. Algoritma ARS digunakan ketika distribusi target tidak *closed form* dan *log-concave*. Algoritma *slice sampler* digunakan ketika distribusi target tidak *log-concave* dan intervalnya memiliki batas. Selain itu, algoritma Metropolis Hasting digunakan ketika distribusi target tidak *log-concave* dan intervalnya tidak terbatas.

**Tabel 1.3 Metode Pembangkitan sampel sesuai urutan yang digunakan di WinBUGS.**

Distribusi Target	Metode Pembangkitan Sampel
Diskrit	<i>Invers cumulative distribution function</i>
Closed form ( <i>Conjugate</i> )	<i>Direct sampling</i> menggunakan algoritma standar.
Log-Concave	<i>Adaptive rejection sampling (ARS)</i>
Intervalnya yang memiliki batas	<i>Slice sampler</i>
Interval tak terbatas	Metropolis Hasting

Sumber: Lunn dkk. (2000)

## 1.5.2 Gibbs Sampler

Algoritma *Gibbs sampler* diperkenalkan oleh Geman dan Geman (1984). Algoritma ini merupakan kasus khusus algoritma *Blockwise Metropolis Hasting* dengan peluang penerimaan  $\alpha^* = 1$ . Salah satu kelebihan dari *Gibbs sampler* adalah pada setiap tahapan, sampel dibangkitkan dari distribusi posterior bersyarat penuh (*full conditional posterior distribution*) berdimensi tunggal yang sering kali bentuknya sudah dikenal (*close form*), contohnya distribusi Normal, distribusi *t*, distribusi eksponensial, distribusi gamma dan lain-lain, serta aplikasinya sudah ada di perangkat lunak statistik (Ntzoufras, 2009).

**Algoritma 1.1** Gibbs Sampler:

1. Tentukan nilai inisial  $\theta^{(0)}$
2. Ulangi tahapan berikut ini pada setiap  $t=1,2,\dots,T$ 
  - a. Tentukan  $\theta = \theta^{(t-1)}$
  - b. Untuk setiap  $j=1,2,\dots,d$ , perbarui  $\theta_j$  dengan membangkitkan  $\theta_j$  yang berdistribusi  $f(\theta_j | \theta_{-j}, \mathbf{y})$
  - c. Tetapkan  $\theta^{(t)} = \theta$  dan simpan nilai ini untuk dipakai pada iterasi  $t+1$  di dalam algoritma ini.

Kadang kala sampel tidak dapat langsung dibangkitkan dari distribusi posterior bersyarat penuh karena bentuknya yang tidak *close form*. Pada kondisi ini, algoritma Metropolis dapat digunakan untuk membangkitkan sampel tersebut. Gabungan algoritma Metropolis dan *Gibbs sampler* ini disebut “Metropolis Within Gibbs”(Albert, 2009).

### 1.5.3 Adaptive Rejection Sampling

Algoritma *Adaptive Rejection Sampling* (ARS) merupakan pengembangan dari algoritma *rejection sampling*, khususnya ketika fungsinya *log-concave*. Suatu fungsi  $g(x)$  adalah *log-concave* ketika

determinan  $\frac{d^2 g(X)}{dX dX^T}$  adalah nonpositif. Gilks dan Wild (1992) menunjukkan

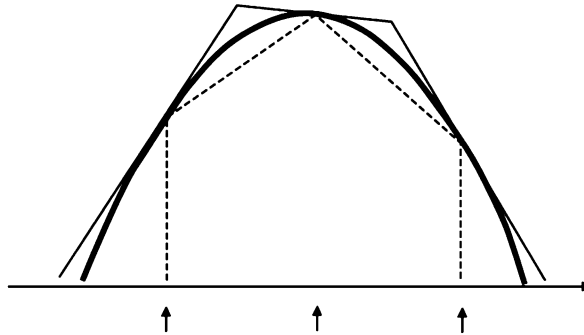
bahwa untuk univariat  $X$ , suatu fungsi *envelope*  $\log(G_s(Y))$  untuk

$\log(g(X))$  dapat diperoleh dengan membuat garis singgung ke  $\log g$  di setiap absis dalam suatu himpunan absis  $S$  yang diberikan. *Envelope* tersebut adalah *piecewise exponential*. Fungsi ini mudah dalam hal pembangkitan sampel. Selain itu, secara otomatis, dalam prosesnya,

terdapat *fungsi squeezing*  $\log(b_s(X))$ . Garis singgung yang dibuat bersifat adaptif sehingga *envelope* juga bersifat adaptif. Oleh karena itu, algoritma ini disebut *adaptive rejection sampling* (ARS).

**Algoritma 1.2** Algoritma ARS:

1. Membuat inisialisasi  $S$
2. Ulangi tahapan berikut:
  - a. Bangkitkan  $X$  dari  $G_s(X)$
  - b. Bangkitkan  $U$  dari  $\text{Uniform}(0,1)$
  - c. Jika  $U \leq g(X) / G_s(X)$  terima  $X$
  - d. Masukkan  $X$  dalam  $S$  hingga satu  $X$  diterima



Sumber: Gilks (1996).

**Gambar 1.10 Adaptive Rejection Sampling dengan Metode Garis**

**Singgung; Garis Tebal:  $\log(g(X))$ ; Garis Tipis: Envelope  $\log(G_S(x))$ ;**

**Garis Putus: Squeezing Function  $\log(b_S(X))$ ; Anak Panah: Absis yang Digunakan.**

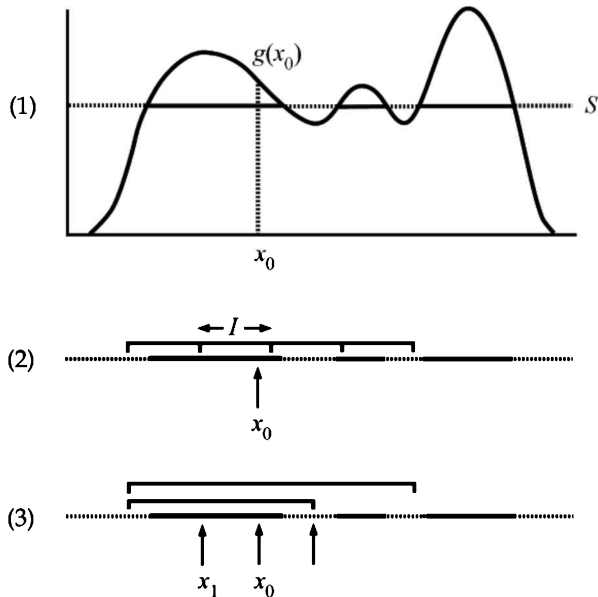
### 1.5.4 Slice Sampler

Algoritma *slice sampler* diusulkan oleh Neal (2003) dan algoritma ini memiliki kemiripan dengan algoritmanya Higdon (1998) dan Damien dkk. (1999). Ini digunakan ketika *full conditional posterior distribution* tidak *closed form*. Algoritma ini paling sederhana ketika digunakan untuk membangkitkan sampel dari satu variabel.

**Algoritma 1.3** Algoritma *Slice Sampler* (Neal, 2003)

1. Bangkitkan  $u \sim U(0, g(x_0))$  dengan demikian “slice” horizontal  $S = \{x : u < g(x)\}$ . Ingat bahwa  $x_0$  selalu di dalam S.
2. Cari interval  $I = (L, R)$  di sekitar  $x_0$  yang berisi semua atau banyak *slice*.
3. Bangkitkan sampel baru  $x_1$  dari bagian *slice* di dalam interval.

Algoritma tersebut ditunjukkan oleh Gambar 1.11.



Sumber: Neal (2003)

**Gambar 1.11** Algoritma *Slice Sampler*.

Algoritma *slice sampler* pada dasarnya dibangun berdasarkan *Gibbs Sampler* (Ntzoufras, 2009). Algoritma ini menambah ruang parameter dengan himpunan variabel random yang menjaga distribusi posterior marjinal tidak berubah tetapi semua bersyarat pada distribusi yang memiliki bentuk standar.

Sebagai contoh suatu fungsi target  $g(x)$  yang sampelnya tidak mudah dibangkitkan (misalnya adanya sifat multimodalitas). Diperkenalkan variabel tambahan  $u$  dengan  $f(u|x)$  sehingga distribusi gabungannya dapat ditulis:

$$f(u, x) = f(u|x)g(x)$$



Distribusi marginal  $f(x)$  sama dengan distribusi target  $g(x)$  karena

$$f(x) = \int f(u, x) du = \int f(u | x) g(x) d(u) = g(x)$$

Dengan cara ini, kita dapat menggunakan *Gibbs sampler* yang membangkitkan nilai  $f(u, x)$  yang berhubungan dengan  $f(u)$  dan  $f(x) = g(x)$

1. Bangkitkan  $u \sim f(u | x)$
2. Bangkitkan  $x \sim f(u | x) g(x)$

Biasanya  $f(u | x)$  yang digunakan adalah distribusi uniform  $U(0, g(x))$  karena

$$f(u, x) = \frac{1}{g(x)} g(x) I(0 < u < g(x)) = I(0 < u < g(x))$$

dan

$$f(x) = \int I(0 < u < g(x)) du = \int_0^{g(x)} du = g(x)$$

dengan  $I(x)$  adalah fungsi indikator yang nilainya sama dengan satu jika  $x$  benar dan nol jika lainnya. Algoritma *slice sampler* dengan pendekatan ini disajikan dalam Algoritma 1.4.

**Algoritma 1.4** Algoritma *Slice Sampler* dalam Bentuk *Gibbs Sampler* (Ntzoufras, 2009, pp.96-101)

1. Bangkitkan  $u^{(t)} \sim U(0, g(x^{(t-1)}))$
2. Bangkitkan  $x^{(t)} \sim U(x : 0 \leq u^{(t)} \leq g(x))$

## 1.5.5 Metropolis Hasting

Algoritma Metropolis Hasting merupakan salah satu algoritma MCMC yang banyak dipakai. Algoritma ini diperkenalkan oleh Hasting (1970) yang merupakan pengembangan dari algoritma yang diperkenalkan Metropolis dkk. (1953).

Berikut ini algoritma Metropolis Hasting untuk membangkitkan parameter yang berdistribusi  $f(\theta | \mathbf{y})$ :

### Algoritma 1.5 Algoritma Metropolis Hasting

1. Tentukan nilai inisial  $\theta^{(0)}$
2. Untuk  $t=1,2,\dots,T$

a. Bangkitkan kandidat nilai  $\theta^*$  dari distribusi proposal  $q(\theta^* | \theta^{(t-1)})$

b. Hitung rasio:

$$R = \frac{f(\theta^* | \mathbf{y})q(\theta^{(t-1)} | \theta^*)}{f(\theta^{(t-1)} | \mathbf{y})q(\theta^* | \theta^{(t-1)})}$$

c. Hitung peluang penerimaan:

$$\alpha(\theta^{(t-1)}, \theta^*) = \min\{R, 1\}$$

d. Bangkitkan bilangan random  $u$  dari  $U(0,1)$ .

e. Jika  $u < \alpha(\theta^{(t-1)}, \theta^*)$  maka tentukan  $\theta^{(t)} = \theta^*$  dan jika

$$u \geq \alpha(\theta^{(t-1)}, \theta^*) \text{ maka } \theta^{(t)} = \theta^{(t-1)}.$$

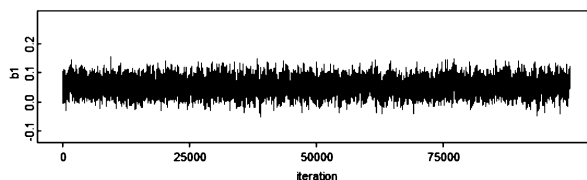
Apabila parameter yang dibangkitkan adalah multiparameter  $\theta$  algoritmanya lebih mudah jika  $\theta$  dipartisi menjadi  $j$  blok

$$\theta = \theta_1, \theta_2, \dots, \theta_j$$

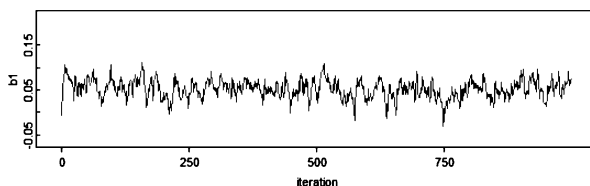
Kemudian algoritma Metropolis Hasting dijalankan pada setiap blok dengan distribusi proposal  $q(\theta_j | \theta_{-j})$  dengan  $\theta_{-j} = \theta_1^{(t)}, \theta_2^{(t)}, \dots, \theta_{j-1}^{(t)}, \theta_{j+1}^{(t)}, \dots, \theta_j^{(t)}$  dan disebut dengan algoritma Blockwise Metropolis Hasting (Bolstad, 2010).

### 1.5.6 Konvergensi MCMC

MCMC yang konvergensi mengandung cukup informasi untuk dipakai sebagai inferensi yang reliabel. Oleh karena itu, pemeriksaan konvergensi *output* MCMC perlu dilakukan. Konvergensi ini dapat diperiksa melalui grafik atau melalui uji statistik.



(a)



(b)

**Gambar 1.12** Diagram *History Output* MCMC: (a) Bentuk “ulat bulu gemuk” dan (b) bentuk mengular

Cara yang mudah untuk memeriksa konvergensi adalah secara visual melalui diagram *history* atau *trace*. *Output* MCMC yang sudah mencapai konvergensi, seharusnya terlihat seperti *random scatter* dengan nilai *mean* yang stabil dan diagram *history*-nya terlihat seperti “ulat bulu yang gemuk/fat hairy caterpillar” (Gambar 1.12 (a)). Apabila bentuknya “mirip ular” (Gambar 1.12 (b)), ini berarti konvergensinya belum tercapai. Bentuk seperti ini disebabkan korelasi yang tinggi di

antara sampel *output* MCMC. Banyaknya iterasi yang dipakai untuk membangkitkan sampel belum cukup untuk mendemonstrasikan konvergensi sehingga iterasinya perlu diperbanyak. Ketika iterasinya diperbanyak, diagram *history*-nya akan nampak seperti “ulat bulu yang gemuk” karena *periodicity* akan “dimampatkan” ke titik tempat *periodicity* ini terlihat seperti *random scatter* (Lunn dkk. 2013).

Selain melalui cara visual, pengecekan konvergensi dapat pula dilakukan melalui uji statistik. Beberapa di antaranya telah dibuat dalam package CODA di perangkat lunak R (Ntzoufras, 2009), yaitu metode Gelman-Brook-Rubin, Geweke, Raftery-Lewis, dan Heidelberger-Weich. Berikut adalah penjelasannya lebih lanjut.

### 1. Metode Gelman-Brook-Rubin (GBR)

Pengecekan konvergensi yang dikembangkan oleh Gelman dan Rubin (1992) ini dilakukan pada rantai lebih dari satu. Konvergensi dilihat dari nilai statistic  $R$  yang diestimasi dari

$$\hat{R} = \frac{\hat{V}}{WSS} = \frac{T'-1}{T'} + \frac{BSS/T'}{WSS} \frac{\kappa+1}{\kappa}$$

dengan  $\kappa$  adalah jumlah sampel bangkitan/rantai (*chain*),  $T'$  adalah jumlah iterasi yang disimpan di sampel/rantai,  $BSS/T'$  adalah varians nilai *mean* posterior pada semua sampel/rantai yang dibangkitkan (variens antar-sampel),  $WSS$  adalah rata-rata dari varians dalam setiap sampel dan

$$\hat{V} = \frac{T'-1}{T'} WSS + \frac{BSS}{T'} \frac{\kappa+1}{\kappa}$$

adalah gabungan estimasi varians *posterior*. Ketika konvergensi tercapai dan ukuran sampel bangkitannya besar, maka  $\hat{R} \rightarrow 1$ . Brook dan Gelman (1998) mengoreksi statistiknya menjadi:

$$\hat{R}_c = \frac{d^* + 3}{d^* + 1} \hat{R}$$

dengan  $d^*$  adalah estimasi derajat bebas gabungan estimasi varian posterior  $\hat{V}$ .

## 2. Metode Geweke

Geweke (1992) melakukan uji pengecekan konvergensi rata-rata (*mean*) setiap parameter secara terpisah berdasarkan nilai sampel pada satu rantai. Uji ini dibuat dengan memandang *output* MCMC sebagai runtun waktu (*time series*). Metode ini menerapkan uji Z untuk memeriksa apakah estimasi rata-rata dari dua subsampel yang berbeda dari semua *output* MCMC adalah sama. Jika rata-rata di awal dan akhir dari keseluruhan *output* MCMC ditolak (*rejected*), maka tidak bisa diasumsikan *output* MCMC tersebut konvergens. Standar di program CODA adalah 10% *output* MCMC di bagian awal dan 50% akhir dari seluruh iterasi. Untuk asimtotik varians *mean* sampel, metode ini mengadopsi nilai estimasi dari teori *spectral density*. Oleh karena itu, untuk suatu himpunan nilai bangkitan parameter

$$\theta_1^{(1)}, \dots, \theta_1^{(T)}$$

rata-rata sampel  $\bar{\theta}_1$  dihitung, diikuti dengan penghitungan *spectral density*-nya runtun waktu  $s_\theta(\omega)$ , sehingga standar eror rata-rata dihitung dengan  $\sqrt{S_\theta(0)/T}$  dan

$$Z = \frac{\bar{\theta}^A - \bar{\theta}^B}{\sqrt{S_\theta^A(0)/T_A + S_\theta^B(0)/T_B}}$$

yang secara asimtotis berdistribusi Normal Standar dengan  $\bar{\theta}^A, \bar{\theta}^B$  rata-rata dua sub sampel (awal dan akhir);  $T_A$  dan  $T_B$  adalah jumlah data sampel bangkitan awal dan akhir; dan  $S_\theta^A(0)/T_A$  dan  $S_\theta^B(0)/T_B$  adalah varians subsampel. Karena nilai Z yang secara asimtotis berdistribusi Normal Standar, nilainya yang berada di daerah

penolakan (dengan taraf signifikansi ( $\alpha$ ) tertentu) mengindikasikan tidak konvergens. Misalkan ditentukan nilai taraf signifikansi,  $\alpha = 0,05$ . Jika nilai  $Z$  lebih dari 1,96 atau kurang dari -1,96, maka hasil MCMC tersebut diindikasikan tidak konvergen..

### 3. Metode Raftery-Lewis

Metode pengecekan konvergensi yang ketiga adalah metode Raftery dan Lewis (1992). Metode ini dapat diterapkan pada *output* MCMC rantai tunggal. Metode ini memfokuskan pada pencapaian tingkat akurasi kuantil yang spesifik, tidak berfokus pada konvergensi rata-rata. Ukuran *default* pada CODA adalah 2,5% persentil yang diestimasi dengan akurasi 0,05 dan peluang 0,95. CODA menghasilkan nilai  $N_{\min}$ ,  $M$ ,  $B$ , dan  $I$ ; yaitu:

- a.  $N_{\min}$  adalah jumlah iterasi minimal yang diperlukan untuk mengestimasi kuantil yang diinginkan dengan spesifikasi awal akurasi diasumsikan saling bebas.
- b.  $N$  total iterasi yang diperlukan.
- c.  $M$  adalah jumlah sampel *burn-in*.
- d.  $I$  adalah faktor independen yang dihitung dengan  $N/N_{\min}$  yang mengindikasikan kenaikan relatif yang disebabkan otokorelasi.  $I$  bernilai 1 mengindikasikan sampel bangkitan saling bebas.

### 4. Metode Heidelberger-Weich

Metode pengecekan konvergensi yang lainnya adalah metodenya Heidelberger dan Weich (1992). Metode ini menggunakan analisis rantai tunggal dari pengamatan univariat dan terdiri atas dua bagian serta berdasarkan ide dari *Brownian bridge theory*. Bagian pertama, metode ini menguji apakah kestasioneran rantai Markov tercapai ( $H_0$ ). Jika hipotesis "null" ditolak, selanjutnya 10% awal dari total iterasi diabaikan dan mengulangi pengujian pada sampel yang tersisa. Prosedur ini diulang dengan mengabaikan 10% sampel

hingga uji kestasioneran tidak ditolak atau lebih dari 50% sampel diabaikan. Jika lebih dari 50% sampel diabaikan,  $H_0$  ditolak dan perlu dilakukan iterasi yang lebih banyak. Jika  $H_0$  tidak ditolak, CODA melaporkan jumlah iterasi yang digunakan untuk melakukan uji ini, jumlah iterasi yang diabaikan dan statistik Cramer-von Mises. Pada bagian kedua, *halfwidth test* diterapkan. Sebagian dari rantai yang sudah lolos uji kestasioneran diperlakukan sebagai runtun waktu, yaitu dengan mengestimasi *spectral density* di nol  $S(0)$ . Kemudian standar eror *mean* asimtotis dihitung dengan  $\sqrt{S(0)/N_p}$  seperti di metode Geweke (1992) dengan  $N_p$  adalah panjang rantai yang lolos uji kestasioneran. Jika *halfwidth* 95% interval kepercayaan *mean* ( $1,96 * \sqrt{S(0)/N_p}$ ) kurang dari  $\varepsilon$  kali *mean* sampel (dengan  $\varepsilon$  adalah bilangan pecahan yang kecil), berarti sudah lulus dari *halfwidth test*. Jika tidak lulus *halfwidth test*, perlu dipertimbangkan rantai yang lebih panjang untuk mencapai presisi parameter yang sesuai dengan nilai yang diinginkan.

## DAFTAR PUSTAKA

- Andrieu, C., N. de Freitas., and A. Doucet. 2001. "Robust Full Bayesian Learning for Radial Basis Network". *Neural Computing*, Vol. 13, hal. 2359-2407.
- Aronoff. 1989. *Geographic Information Sistem: A Management Perspective*. Ottawa, Canada: WDL Publication.
- Arora, V., and P. Lahiri. 1997. "On the Superiority of the Bayesian Method over the BLUP in Small Area Estimation Problems". *Statistica Sinica*, Vol. 7, hal. 1053–1063.
- Arora, V., P. Lahiri., and K. Mukherjee. 1997. "Empirical Bayes Estimation of Finite Population Means from Complex Surveys". *Journal of the American Statistical Association*, Vol. 92, No. 440, hal. 1555–1562.
- Babu, G.J. 2012. "Bayesian and Frequentist Approaches". *Online Proceedings of the Astronomical Data Analysis Conference (ADA VII)*, Cargèse, Corsica.



- Badan Pusat Statistik. 2016. *Data dan Informasi Kemiskinan Kabupaten Kota 2015*. Jakarta: © Badan Pusat Statistik.
- Banerjee, S., B.P Carlin., and A.E Gelfand. 2004. *Hierarchical Modeling and Analysis for Spatial Data*. Florida, Boca Raton: Chapman & Hall/CRC.
- Bartolucci, F., L. Scaccia., and A. Mira. 2006. "Efficient Bayes Factor Estimation from the Reversible Jump Output". *Biometrika*, Vol. 93, No. 1, hal. 41-52.
- Battese, G.E., R.M Harter., and W.A Fuller. 1988. "County Crop Areas Using Survey Data and Satellite Data". *Journal of the American Statistical Association*, Vol. 83, No. 401, hal. 28-36.
- Bayes, T. 1763. "A letter to John Canton". *Philosophical Transactions of the Royal Society of London*, Vol. 53, hal. 269–271.
- Benavent, R., and D. Morales. 2016. "Multivariate Fay–Herriot Models for Small Area Estimation". *Computational Statistics & Data Analysis*, Vol. 94, hal. 372-390.
- Besag, J., J.C York., and A. Mollié. 1991. "Bayesian Image Restoration, with Two Applications in Spatial Statistics". *Annals of the Institute of Statistical Mathematics*, Vol. 43, hal. 1–59.
- Best, N., S. Richardson., and A. Thomson. 2005. "A Comparison of Bayesian Spatial Models for Disease Mapping". *Statistical Methods in Medical Research*, Vol. 14, No. 1, hal. 35-59.
- Bhadra, D., M. Ghosh., and D. Kim. 2012. "Estimation of Median Household Income for Small Areas: A Bayesian Semiparametric Approach". *Calcutta Statistical Association Bulletin*, Vol. 64, hal. 253-254.
- Bishop, C. M. 1995. *Neural Networks for Pattern Recognition*. New York: Oxford University Press.

- Box, G.E.P. 1980. "Sampling and Bayes' Inference in Scientific Modelling and Robustness". *Journal of the Royal Statistical Society*. Series A, Vol. 143, No. 4, hal. 383-430.
- Box, G.E.P., and G.C Tiao. 1973. *Bayesian Inference in Statistical Analysis*, Addison Wesley. Massachusetts: Reading.
- Breslow, N.E., and D.G Clayton. 1993. "Approximate Inference in Generalized Linear Models". *Journal of the American Statistical Association*, Vol. 88, No. 421, hal. 9-25.
- Brooks, S., and A. Gelman. 1998. "Alternative Methods for Monitoring Convergence of Iterative Simulations". *Journal of Computational and Graphical Statistics*, Vol. 7, hal. 434-455.
- Buntine, W.L., and A.S Weigand. 1991. "Bayesian Back-propagation". *Complex Systems*, Vol. 5, hal. 603-643.
- Carlin, B.P., and S. Chib. 1995. "Bayesian Model Choice via Markov Chain Monte Carlo Methods". *Journal of Royal Statistical Society Ser. B*, Vol. 57, No. 3, hal. 473-484.
- Chib, S. 1995. "Marginal Likelihood from the Gibbs Output". *Journal of the American Statistical Association*, Vol. 90, No. 432, hal. 1313-1321.
- Chib, S. and Jeliazkov, I. 2001. "Marginal Likelihood from the Metropolis-Hastings Output". *Journal of the American Statistical Association*, Vol. 96, No. 453, hal. 270-281.
- Choir, A.S., R.B Prasetyo., B.S.S Ulama., N. Iriawan., K. Fithriasari., and M. Dokhi. 2016. "Parameter Estimation of General Regression Neural Network Using Bayesian Approach". *Proceeding of The 7th SEAMS-UGM International Conference on Mathematics and Applications 2015: Enhancing the Role of Mathematics in Interdisciplinary Research*. Yogyakarta: AIP Publishing.

- Congdon, P. 2006. *Bayesian Statistical Modeling*, 2nd ed. New York: John Wiley & Son.
- Dale, A.I. 1982. "Bayes or Laplace? An Examination of The Origin and Early Application of Bayes' Theorem". *Archive for History of Exact Sciences*, Vol. 27, No. 1, hal. 23-47.
- Damien, P., J. Wakefield., and S. Walker. 1999. "Gibbs Sampling for Bayesian Non-Conjugate and Hierarchical Models by Using Auxiliary Variable". *Journal of the Royal Statistical Society, Series B*, Vol. 61, hal. 331-344.
- Datta, G.S. and M. Ghosh. 1991. "Bayesian Prediction in Linear Models: Applications to Small Area Estimation". *The Annals of Statistics*, Vol. 19, No. 4, hal. 1748–1770.
- Denison, D.G.T., C.C Holmes., B.K Mallick., and A.F.M. Smith. 2002. *Bayesian Methods for Nonlinear Classification and Regression*. Chichester, West Sussex, England: John Wiley and Sons.
- Diggle, P., R. Moyeed., B. Rowlingson., and M. Thomson. 2002. "Childhood Malaria in the Gambia: A Case-Study in Model-Based Geostatistics". *Applied Statistics*, Vol. 51, hal. 493–506.
- Diggle, P., J. Tawn., and R. Moyeed. 1998. "Model-Based Geostatistics". *Applied Statistics*, Vol. 47, hal. 299–350.
- Draper, N.R. and H. Smith. 1998. *Applied Regression Analysis*, 3rd ed. New York: John Wiley & Sons.
- Eberley, L.E. and B.P Carlin. 2000. "Identifiability and Convergence Issues for Markov Chain Monte Carlo Fitting of Spatial Models". *Statistics in Medicine*, Vol. 19, hal. 2279–2294.
- Erkenntnis, R.E. 1996. "De Finetti's Reconstruction of The Bayes-Laplace Paradigm". *An International Journal of Scientific Philosophy*, Vol. 45, No. 2, hal. 159–176.

- Fausett, L.V. 1994. *Fundamentals of Neural Networks: Architectures, Algorithms And Applications*. Englewood Cliffs, NJ: Prentice-Hall.
- Fithriasari, K., N. Iriawan., B.SS Ulama., and Sutikno. 2013. "On The Multivariate Time Series Rainfall Modeling Using Time Delay Neural Network". *Journal of Applied Mathematics and Statistics*, Vol. 44, No. 14, hal. 193-201.
- Gelfand, A.E., and D.K Dey. 1994. "Bayesian Model Choice: Asymptotics and Exact Calculations". *Journal of the Royal Statistical Society, Series B*, Vol. 56, hal. 501–514.
- Gelman, A. 2006. "Prior Distributions for Variance Parameters in Hierarchical Models (Comment on Article by Browne and Draper)". *Bayesian Analysis*, Vol. 1, No. 3, hal. 515-534.
- Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B. 2004. *Bayesian Data Analysis*, 2nd ed. Boca Raton, Fla: Chapman and Hall/ CRC Press.
- Gelman, A., and D. Rubin. 1992. "Inference from Iterative Simulation Using Multiple Sequences". *Statistical Science*, Vol. 7, hal. 457-511.
- Geweke, J. 1992. "Evaluating The Accuracy of Sampling-Based Approaches to Calculating Posterior Vol. 4", dalam J. Bernardo, J. Berger, A. Dawid and A. Smith (Eds.) *Bayesian Statistics*. Oxford: Claredon Press (hal. 169-194).
- Ghosh, M., Natarajan, K., Stroud, T.W.F. and Carlin, B.P. 1998. "Generalized Linear Models for Small Area Estimation". *Journal of the American Statistical Association*, Vol. 93, No. 441, hal. 273–282.
- Ghosh, M., and J.N.K. Rao. 1994. "Small Area Estimation: An Appraisal". *Statistical Science*, Vol. 9, No. 1, hal. 55-93.

- Ghosh, M., K. Sinha., and D. Kim. 2006. "Empirical and Hierarchical Bayesian Estimation in Finite Population Sampling under Structural Measurement Error Models". *Scandinavian Journal of Statistics*, Vol. 33.
- Gilks, W.R. 1996. "Full Conditional Distribution", dalam W.R. Gilks, S. Richardson and D.J. Spiegelhalter (Eds.) *Markov Chain Monte Carlo in Practice*. London: Chapman and Hall/CRC (hal. 75-88).
- Gonzalez, M.E. 1973. "Use and Evaluation of Synthetic Estimators", *Proceedings of the Social Statistics Section*. Washington, D.C: American Statistical Association (hal. 33-36).
- Gosh, M., T. Maiti., D. Kim., S. Chakraborty., and A. Tewari. 2004. "Hierarchical Bayesian Neural Networks: An Application to a Prostate Cancer Study". *Journal of the American Statistical Association*, Vol. 99, No. 467, hal. 601-608.
- Goulermas, J.Y., P. Liatsis., X.J. Zeng., and P.Cook. 2007. "Density-Driven Generalized Regression Neural Networks (DD-GRNN) for Function Approximation". *IEEE Transactions on Neural Networks*, Vol. 18, No. 6, hal. 1683-1696.
- Greenland, S. and J. Robins. 1994. "Ecologic Studies—Biases, Misconceptions, and Counterexamples". *American Journal of Epidemiology*, Vol. 139, No. 8, hal. 747–760.
- Hasting, W.K. 1970. "Monte Carlo Sampling Methods Using Markov Chains and Their Application". *Biometrika*, Vol. 57, hal. 97-109.
- Heidelberger, P. and P. Welch. 1992. "Simulation Run Length Control in The Presence of An Initial Transient". *Operations Research*, Vol. 31, hal. 1109-1144.
- Higdon, D. 1998. "Auxiliary Variable Methods for Markov Chain Monte Carlo with Applications". *Journal of the American Statistical Association*, Vol. 93, hal. 585-595.

- Holmes, C.C. and B.K Mallick. 1998. "Bayesian Radial Basis Function of Variable Dimension". *Neural Computing*, Vol. 10, hal. 1217-1233.
- Hwang, J. and D.H Kim. 2015. "Bayesian Curve-Fitting in Semiparametric Small Area Models with Measurement Errors". *Communications for Statistical Applications and Methods*, Vol. 22, No. 4, hal. 349-359.
- Iriawan, N. 2001. "Implementing Bayesian Inference Using MCMC on Minitab". *Forum Statistika dan Komputasi*, Vol. 2, No. 2, hal. 1-6.
- Iriawan, N. and I.W. Wright. 1997. "Multiplicative Joint Density for Model Selection: MCMC Approach". *Australian Statistical Conference (ASC-14)*, Gold Coast, Queensland, Australia.
- Jeffreys, H. 1961. *Theory of Probability*, 3rd ed. London: Oxford University Press.
- Jiang, J. and Lahiri, P. 2006. "Mixed Model Prediction and Small Area Estimation", *Test*, Vol. 15, No. 1, hal. 1-96, doi:10.1007/BF02595419.
- Jiang, J., T. Nguyen., and J.S Rao. 2010. "Fence Method for Nonparametric Small Area Estimation". *Survey Methodology*, Vol. 36, No. 1, hal. 3-11.
- Kass, R., and A. Raftery. 1995. "Bayes Factor". *Journal of American Statistical Association*, Vol. 90, hal. 773-795.
- Kleffe, J., and J.H.K Rao. 1992. "Estimation of Mean Square Error of Empirical Bests Linear Unbiased Predictors under a Random Variance Linear Model". *Journal of Multivariate Analysis*, Vol. 43, hal. 1-15.
- Lampinen, J., and A. Vehtari. 2001. "Bayesian Approach for Neural Networks – Review and Case Studies". *Neural Networks*, Vol. 14, No. 3, hal. 7-24.

- Laplace, P.S. 1814. *Essai Philosophique sur les Probabilités*. Paris: Courcier. diakses dari: <https://archive.org/details/essaiphilosophi00laplgoog>.
- LeSage, J.P. and R.K Pace. 2004. "Models for Spatially Dependent Missing Data". *Journal of Real Estate Finance and Economics*, Vol. 29, No. 2, hal. 233–254.
- Little, R.J.A. 2006. "Calibrated Bayes: A Bayes/Frequentist Roadmap". *The American Statistician*, Vol. 60, No. 3, hal. 213–223.
- Lunn, D., D. Spiegelhalter., A. Thomas., and N. Best. 2009. "The BUGS Project: Evolution, Critique and Future Directions", *Statistics in Medicine*, Vol. 28, No. 25, hal. 3049-3067.
- Lunn, D.J., A. Thomas., N Best., and D. Spiegelhalter. 2000. "WinBUGS—A Bayesian Modelling Framework: Concepts, Structure, and Extensibility". *Statistics and Computing*, Vol. 10, hal. 325–337.
- Mackay, D.J.C. 1992. "A Practical Bayesian Framework for Backprop Networks". *Neural Computation*, Vol. 4, hal. 448-472.
- Mackay, D.J.C. 1994a. "Bayesian Methods for Backpropagation Networks" dalam E. Domany, J.L.V. Hemmen, and K. Schulten (Eds). *Models of Neural Network III*. New York: Springer-Verlag (Bab 6).
- Mackay, D.J.C. 1994b. "Bayesian Non-linear Modelling for Prediction Competition". *ASHRAE Transaction*, Vol. 100, No. 2, hal. 1053-1062.
- Maiti, T. 1997. "Hierarchical Bayes Estimation of Mortality Rates for Disease Mapping". *Journal of Statistical Planning and Inference*, Vol. 69, hal. 339-348.
- Maiti, T., S. Sinha., and P. Zhong. 2016. "Functional Mixed Effects Model for Small Area Estimation". *Scandinavian Journal of Statistics*, Vol. 43, No. 3, hal. 886-903.

- Malec, D., J. Sedransk., C.L. Moriarity., and F.B. LeClere. 1997. "Small Area Inference for Binary Variables in the National Health Interview Survey". *Journal of the American Statistical Association*, Vol. 92, hal. 815–826.
- Metropolis, N.A., A.W. Rosenbluth., M.N. Rosenbluth., A.H. Teller., and E. Teller. 1953. "Equations of State Calculations by Fast Computing Machines". *The Journal of Chemical Physics*, Vol. 21, No. 6, hal. 1087-1092.
- Müller, P. and D. Rios Insua. 1998. "Feedforward Neural Networks for Nonparametric Regression" dalam D. Dey, P. Mülle, and D. Sinha Practical (Eds.) *Nonparametric and Semiparametric Bayesian Statistics*. New York: Springer, hal. 181-193.
- Neal, R.M. 1996. *Bayesian Learning for Neural Network*. New York: Springer.
- Neal, R. 2003. "Slice Sampling". *The Annals of Statistics*, Vol. 31, hal. 705-767.
- Noviani, A. 2016. "Small Area Estimation dengan Pendekatan Hierarchical Bayesian Neural Network untuk Kasus Anak Putus Sekolah dari Rumah Tangga Miskin di Provinsi Jawa Timur". *Tesis Magister*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Ntzoufras, I. 2009. *Bayesian Modeling Using Winbugs*. New Jersey: John Wiley & Sons.
- Opsomer, J.D., G. Claeskens., M.G. Ranall., and F.J. Breidt. 2008. "Non-Parametric Small Area Estimation Using Penalized Spline Regression". *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, Vol. 70, No. 1, hal. 265-286.
- Orr, M.J.L. 1996. *Introduction to Radial Basis Function Network*. Scotland: Edinburg University.



- Parzen, E. 1962. "On Estimation of a Probability Density Function and Mode". *The Annals of Mathematical Statistics*, Vol. 33, No. 3, hal. 1065-1076.
- Petrucci, A., and N. Salvati. 2005. "Small Area Estimation for Spatial Correlation in Watershed Erosion Assessment". *Journal of Agricultural, Biological & Environmental Statistics*, Vol. 11, No. 2, hal. 169–182.
- Pfeffermann, D. 2002. "Small Area Estimation—New Developments and Directions". *International Statistical Review*, Vol. 70, hal. 125-143.
- Pfeffermann, D. 2013. "New Important Developments in Small Area Estimation". *Statistical Science*, Vol. 28, No. 1, hal. 40-68.
- Prasetyo, R.B., B.S.S. Ulama., N. Iriawan., H. Kuswanto., and K. Fithriasari. 2016. "A Hierarchical Bayesian Binomial-Poisson Model for Small Area Estimation of Dropout Rate in Poverty". *International Conference on Theoretical and Applied statistics 2016*. Surabaya: Jurusan Statistika, ITS.
- Raftery, A. and S. Lewis. 1992. "How Many Iterations in The Gibbs Sampler? Vol. 4" dalam J. Bemardo, J. Berger, A. Dawid and A. Smith (Eds.). *Bayesian Statistics*. Oxford: Claredon Press, (hal. 763-774).
- Rao, J.N.K. 2003. *Small Area Estimation*. New York: John Wiley and Sons, Inc.
- Rao, J.N.K. and I. Molina. 2015. *Small Area Estimation*, 2nd ed. Hoboken, New Jersey: John Wiley & Sons.
- Ripley, B.D. 1996. *Pattern Recognition and Neural Network*. Cambridge: Cambridge University Press.
- Robert, G.O. 1996. "Markov Chain Concept Related to Sampling Algorithm" dalam W.R. Gilks, and S. Richardson (Eds). *Markov Chain Monte Carlo in Practice*. Boca Raton, Florida: CRC Press, hal. 45-57.

- Robinson, G.K. 1991. "That BLUP is a Good Thing: The Estimation of Random Effects (with Discussion)". *Statistical Science*, Vol. 6, hal. 15–51.
- Rubin, D.B. 1984. "Bayesianly Justifiable and Relevant Frequency Calculations for the Applied Statistician". *The Annals of Statistics*, Vol. 12, No. 4, hal. 1151-1172.
- Saei, A., and R. Chambers. 2005. "Working Paper M05/03: Empirical Best Linear Unbiased Prediction for out of Sample Areas". *Technical Report, Southampton Statistical Sciences Research Institute*. Southampton: University of Southampton.
- Salvati, N., H. Chandra., M.G Ranalli., and R. Chambers. 2010. "Small Area Estimation Using a Nonparametric Model-Based Direct Estimator". *Computational Statistics & Data Analysis*, Vol. 54, No. 9, hal. 2159–2171.
- Sánchez-Borrego, I., M. Rueda., and J.F Muñoz. 2012. "Nonparametric Methods in Sample Surveys. Application to the Estimation of Cancer Prevalence". *Quality & Quantity*, Vol. 46, No. 2, hal. 405–414.
- Sarle, W.S. 1994a. "Neural Network Implementation in SAS Software". *Proceedings of the Nineteenth Annual SAS Users Group International*. NC: Cary, hal. 1551-1573.
- Sarle, W.S. 1994b. "Neural Networks and Statistical Models". *Proceedings of the Nineteenth Annual SAS Users Group International Conference*. North Carolina: SAS Institute Inc. hal. 1-13.
- Sarndal, C.E., B. Swensson., and J. Wretman. 1992. *Model Assisted Survey Sampling*. New York: Springer-Verlag Inc.
- Shim, J. and C. Hwang. 2014. "Estimating Small Area Mean with Mixed and Fixed Effects Support Vector Median Regressions". *Neurocomputing*, Vol. 145, hal. 174–181.

- Singh, B.B., G.K Shukla., and D. Kundu. 2005. "Spatio-Temporal Models in Small Area Estimation". *Survey Methodology*, Vol. 31, No. 2, hal. 183–196.
- Sinha, S.K. and A. Sattar. 2015. "Inference in Semi-parametric Spline Mixed Models for Longitudinal Data". *METRON*, Vol. 73, No. 3, hal. 377–395.
- Sivaganesan, S. 2008. "Criticism/Disadvantages of Bayesian Approach" diakses dari <https://math.uc.edu/~siva/mbayes/chap2p.pdf>.
- Specht, D.F. 1991. "A General Regression Neural Network". *IEEE Transactions on Neural Network*, Vol. 2, hal. 568-576.
- Spiegelhalter, D.J., N.G Best., B.P Carlin., and A. van der Linde. 2002. "Bayesian Measures of Model Complexity and Fit (with Discussion)". *Journal of the Royal Statistical Society: Series B*, Vol. 64, hal. 583–639.
- Staubach, C., V. Schmid., L. Knorr-Held., and M. Ziller., 2002. "A Bayesian Model for Spatial Wildlife Disease Prevalence Data". *Preventive Veterinary Medicine*, Vol. 56, No. 1, hal. 75-87.
- Swenson, B., and J.H Wretman., 1992. "The Weighted Regression Technique for Estimating the Variance of Generalized Regression Estimator". *Biometrika*, Vol. 76, hal. 527-537.
- Titterington, D.M. 2004. "Bayesian Methods for Neural Networks and and Related Models". *Statistical Science*, Vol. 19, No. 1, hal. 128–139.
- Torabi, M., and F. Shokoohi. 2015. "Non-parametric Generalized Linear Mixed Models in Small Area Estimation". *Canadian Journal of Statistics*, Vol. 43, No. 1, hal. 82-96.
- Ubaidillah, A., N. Iriawan., B.S Ulama., and K. Fithriasari. 2013. "Pemodelan Kemiskinan dengan Menggunakan Metode Hierarchical Bayesian Neural Network (Studi Kasus Pada

- Sampel Rumah Tangga Survei Sosial Ekonomi Nasional Tahun 2011 di Kota Jambi". *Proceedings of the Jenderal Soedirman University National Conference*. Purwokerto: Universitas Jenderal Soedirman.
- Ugarte, M.D., T. Goicoa., A.F Militino., and M. Durbánb. 2009. "Spline Smoothing in Small Area Trend Estimation and Forecasting". *Computational Statistics and Data Analysis*, Vol. 53, hal. 3616-3629.
- Ulama, B.S.S. 2008. *Jaringan Syaraf Tiruan Fungsi Radial Basis untuk Pemodelan Data Runtun Waktu. Disertasi Doktor*. Yogyakarta: Universitas Gadjah Mada.
- Wong, B.K., T.A Bodnovich., and Y. Selvi. 1997. "Neural Network Applications in Business: A Review and Analysis of The Literature (1988-95)". *Decision Support Systems*, Vol. 19, hal. 301-320.
- You, Y., and J.N.K Rao. 2000. "Hierarchical Bayes Estimation of Small Area Means Using Multi-level Models". *Survey Methodology*, Vol. 26, hal. 173-181.
- Zhang, X., R.D Brooks., and M.L King. 2009. "A Bayesian Approach to Bandwidth Selection for Multivariate Kernel Regression with an Application to State-Price Density Estimation". *Journal of Econometrics*, Vol. 153, hal. 21-32.
- Zhou, Q.M., and Y. You. 2008. "Hierarchical Bayes Small Area Estimation for The Canadian Community Health Survey". *Proceedings of the Survey Methods Section SSC Annual*.



## TENTANG PENULIS



**Nur Iriawan**, setelah menyelesaikan pendidikan dasar dan menengah tahun 1982 dari kota kelahirannya di Blitar, tahun 1986 menyelesaikan studi Sarjana Statistika di FMIPA, Institut Teknologi Sepuluh Nopember (ITS), Surabaya. Sejak tahun 1988, penulis memilih untuk menjadi dosen tetap di almaternya, Jurusan Statistika FMIPA-ITS. Tahun 1989 memperoleh kesempatan studi lanjut di Computer Science dalam Sandwich Program antara Universitas Indonesia (UI) dan University of Maryland (UM) – Baltimore, USA. Tahun 2000 berhasil menyelesaikan studi Doktor dalam bidang statistika dari School of Mathematics and Statistics, Curtin University of Technology (CUT), Perth - Western Australia. Di tengah penyelesaian studinya di Australia, penulis pernah memenangkan sebuah kompetisi penelitian dan memperoleh penghargaan Young Statistician Awards di Australia tahun 1997. Pengalaman mengelola manajemen pendidikan di tingkat

departemen/jurusan satu periode 1999-2003, dan sebagai Pembantu Dekan III (PD3) di FMIPA-ITS tahun 2003-2007. Di sela waktu menekuni pelayanan kemahasiswaan sebagai PD3, penulis dikukuhkan sebagai Profesor di bidang Statistika Komputasi dan Proses Stokastik di ITS pada tahun 2005. Tiga tahun, 2009 s/d 2012, menjadi Tim *Ad-hoc* penyusunan pedoman awal dan pelaksanaan program Bidikmisi di Dikti-Kemdikbud. Mengemban amanah sebagai Wakil Rektor ITS bidang Sumber Daya Manusia, Organisasi, dan Teknologi Sistem Informasi (WR-SDMO-TSI) - ITS dalam periode jabatan 2011-2015. Sekarang diamanahi mengelola Laboratorium Statistika Komputasi di Departemen Statistika, Fakultas Matematika, Komputasi, dan Sains Data (FMKSD), ITS.



**Brodjol Sutijo Suprih Ulama** dilahirkan di Malang 25 Januari 1966. Jenjang pendidikan SD dan SMP ditempuh di Walikukun Kabupaten Ngawi, menempuh pendidikan SMA di Madiun dan melanjutkan kuliah di ITS Surabaya mengambil Jurusan Statistika pada tahun 1984 dan lulus pada tahun 1989. Mulai tahun 1990 menjadi staf pengajar di jurusan Statistika ITS. Tahun 1994-1999 ditunjuk sebagai koordinator Kerja Praktik di Jurusan Statistika ITS. Tahun 1999 melanjutkan pendidikan S2 di IPB lulus tahun 2001, sedangkan pendidikan S3 di tempuh di UGM lulus tahun 2008. Pada tahun 2010 penulis terpilih sebagai dosen teladan III di FMIPA ITS. Pada tahun 2012–2016 dipercaya sebagai sekretaris Pusat pengembangan Karir–SAC ITS. Sekarang, penulis dipercaya meenjadi Sekretaris Departemen Statistika Bisnis, Fakultas Vokasi, ITS. Pernah menjadi instruktur berbagai workshop statistika dan mengikuti berbagai seminar baik nasional maupun internasional. Bidang penelitian yang dikaji adalah Analisis Time Series, Neural Network dan Statistika Komputasi.



**Kartika Fithriasari**, telah menyelesaikan pendidikan menengah atas pada tahun 1987 di Surabaya. Pada tahun 1992 telah menyelesaikan tahap Sarjana di Jurusan Statistika FMIPA ITS. Sejak tahun 1993, penulis memutuskan untuk bekerja sebagai dosen tetap di Jurusan Statistika FMIPA ITS. Tahap magister sudah dilampai oleh penulis sejak tahun 1993 dan lulus pada tahun 1997 di Statistika, IPB. Tahun 2010, penulis melanjutkan studi ke jenjang Doktor di jurusan Statistika ITS dan lulus tahun 2014. Pengalaman mengelola manajemen pendidikan di tingkat jurusan sebagai Sekretaris program studi Pascasarjana Jurusan Statistika ITS tahun 2016 dan sekarang tahun 2017 sebagai Sekretaris Departemen Statistika FMKSD ITS. Bidang penelitian yang ditekuni oleh penulis adalah Statistika Komputasi khususnya Neural Network dan Bayesian.



**Achmad Syahrul Choir**, menyelesaikan pendidikan dasar hingga menengah di Jombang, Jawa Timur. Jenjang DIV ditamatkan penulis di Jurusan Statistika, Sekolah Tinggi Ilmu Statistik (STIS), Jakarta pada 2003. Gelar M.Si. diraih dari Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember (ITS), Surabaya, tahun 2011 dan saat ini melanjutkan ke jenjang S3 di Departemen Statistika, Fakultas Matematika, Komputasi, dan Sains Data, ITS Surabaya. Bekerja di Badan Pusat Statistik (BPS) Kabupaten Sinjai tahun 2003 dan sejak tahun 2009 bertugas di BPS Provinsi Jawa Timur. Pernah mengikuti berbagai *workshop*, seminar nasional dan internasional yang terkait dengan statistika dan neural network. Bidang penelitian yang digeluti adalah Bayesian, General Regression Neural Network (GRNN) dan Statistika Komputasi.





**Rindang Bangun Prasetyo**, menamatkan sekolah dasar di Wonogiri tahun 1992 dengan prestasi mendapatkan juara ke-3 lomba matematika tingkat Kabupaten. Setelah menyelesaikan sekolah menengah di Karanganyar Jawa Tengah, penulis melanjutkan pendidikan di Sekolah Tinggi Ilmu Statistik (STIS) jurusan Komputasi Statistik. Pada tahun 2010, menamatkan S2 di Jurusan Ilmu Ekonomi, Institut Pertanian Bogor (IPB). Saat ini penulis sedang mengambil S3 Statistika di ITS Surabaya. Bidang penelitian yang diminati yaitu Small Area Estimation, Analisis Spasial dan Statistika Komputasi. Selain pendidikan formal, penulis pernah mengikuti beberapa pelatihan, salah satunya, yaitu “In Country Training for The Small Area Statistics” kerja sama BPS dan Japan International Cooperation Agency (JICA) dengan prestasi menjadi peserta terbaik pertama. Riwayat pekerjaan dimulai pada tahun 2003 sebagai PNS di BPS Kabupaten Morowali dan BPS Propinsi Sulawesi Tengah pada tahun 2005, selanjutnya dipindahtugaskan ke BPS RI sejak tahun 2013.

# Bayesian Neural Network

dalam Pemodelan Small Area Estimation

Buku ini memadukan dua metode, yaitu Bayesian dan Neural Network yang diimplementasikan dalam pemodelan *Small Area Estimation* (SAE). Perpaduan tersebut diawali dengan membahas konsep dasar Bayesian dan Neural Network dan bertahap memadukan keduanya melalui pintu Bayesian Regresi yang mendasari pemodelan SAE. Dalam buku ini juga dibahas tentang bagaimana pemanfaatan teknologi informasi untuk penyajian data yang lebih menarik dan interaktif. Materi bahasan disajikan dalam bahasa yang mudah dan dilengkapi dengan contoh aplikasi baik dalam bentuk *source code* maupun tampilan hasilnya.

**Penerbit ANDI**

Jl. Beo 38-40 Yogyakarta

Telp.(0274) 561881 Fax.(0274) 588282

✉ : [andipenerbitan@gmail.com](mailto:andipenerbitan@gmail.com)

🌐 : [www.andipublisher.com](http://www.andipublisher.com)

<http://ebukune.my.id>

**Dapatkan Info Buku Baru, Kirim e-mail: [info@andipublisher.com](mailto:info@andipublisher.com) | [andipublishercom@yahoo.com](mailto:andipublishercom@yahoo.com)**

**PEMROGRAMAN**

ISBN 978-623-01-0895-2 (PDF)



9 786230 108952 | 1 2 3 0 1

Harga Pulau Jawa Rp64.000