MSEPBurr Distribution: Properties and Parameter Estimation

Achmad Syahrul Choir
Doctoral Student at Department of Statistics
Faculty of Mathematics, Computation, and Science Data
Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
BPS-Statistics Indonesia
madsyair@bps.go.id

Nur Iriawan
Department of Statistics
Faculty of Mathematics, Computation and Science Data
Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
nur_i@statistika.its.ac.id

Brodjol Sutijo Suprih Ulama
Department of Business Statistics, Faculty of Vocational
Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
brodjol_su@statistika.its.ac.id

Mohammad Dokhi
Politeknik Statistika STIS, Jakarta, Indonesia
dokhi@stis.ac.id

Abstract

MSNBurr and MSTBurr distribution have been developed as Neo-Normal distributions that represent a relaxation of normality. The difference between them is that the MSTBurr's peak is below MSNBurr's. In this paper, we propose a MSEPBurr distribution with its peak could be not only lower but also high-er than MSNBurr. Furthermore, we study several properties of MSEPBurr, such as mean, variance, skewness, kurtosis, and quantile. The MSEPBurr parameters are estimated by using the Bayesian approach with the BUGS language implementation for its computation. We employ simulation study and use existing data to illustrate the application of the regression model. In real data, we notice that MSEPBurr has similar performance with MSNBurr and MSTBurr that they outperform Normal and Student-*t* distribution in Australian athlete data because their skewness can accommodate long left tail excellently. However, their performance is less than the Student-*t* model in chemical reaction rate data because their skewness can not accommodate long right tail perfectly. Although in general their perfor-mance is the same, we observe that the MSEPBurr performs better than the MSNBurr and the MSTBurr in some simulated data.

Keywords: Distribution, Bayesian, Normal Relaxation, MSEPBurr, MSNBurr, MSTBurr

1. Introduction

The Normal distribution is commonly used in statistical modeling. However, the use of this distribution is sometimes incompatible with the available data. Therefore, some distributions have been developed as a relaxation of the Normal distribution. Subbotin distribution (Subbotin, 1923) and Exponential Power (EP) distribution (Box and Tiao, 1973) represent a relaxation of the Normal distribution in terms of kurtosis. Both distributions can be mesokurtic like Normal distribution, platykurtic, or lepto-kurtic. Another form of the Normal distribution relaxation was done on its skewness. It has been proposed by Azzalini (1985) called the Skew-Normal distribution. This distribution has symmetrical or skewed properties with an unstable mode at its location parameter. An extended Skew-Normal distribution has been applied to regression analysis by Olosunde Akinlolu (2011). In contrast to Azzalini (1985), Fernandez and Steel (1998) proposed the Skew-Normal distribution that has a stable mode of the location parameter.

The study of the skewed or symmetrical distribution has been carried out by Iriawan (2000) who developed the "Modified to be Stable as Normal from Burr", hereinafter referred to MSNBurr distribution. It was derived from the modification of Burr type II distribution (Burr, 1942). The mode of MSNBurr distribution is stable like that of the Skew-Normal distribution (Fernandez and Steel, 1998). The symmetrical MSNBurr perfectly fits the Normal distribution, but its tails are fatter than that of the Normal distribution.

Iriawan (2000) also developed the "Modified to be Stable as t from Burr", hereinafter referred to MSTBurr distribution with its peak could be below MSNBurr's when their location and scale parameters are the same. In this paper, we propose "Modified to be Stable Exponential Power from Burr", henceforth referred to MSEPBurr distribution with its peak could be not only lower but also higher than MSNBurr distribution.

2. The Neo-Normal Distribution

The Neo-Normal distribution is a distribution, which represents a relaxation of the Normal distribution (Iriawan, 2000). This distribution can be the same or different from the Normal distribution due to the shape parameter that plays a role in assigning the magnitude of kurtosis or skewness. One of the preliminary works of relaxation of the Normal distribution was conducted by Box and Tiao (1973) who investigate EP distribution. The probability density function (pdf) of a random variable Z_1 that follows EP distribution is

$$f(z_1|\mu,\sigma,v)=\frac{C(v)}{\sigma}exp\left(-q(v)\left|\frac{z_1-\mu}{\sigma}\right|^{2/(1+v)}\right), \quad \textbf{(1)}$$

where $-\infty<z_1<\infty,-\infty<\mu<\infty,\sigma>0,-1<v<1$,

$$q(v)=\left(\frac{\Gamma\left(\frac{3}{2}(1+v)\right)}{\Gamma\left(\frac{1}{2}(1+v)\right)}\right)^{1/(1+v)},C(v)=\frac{\Gamma\left(\frac{3}{2}(1+v)\right)^{\frac{1}{2}}}{(1+v)\Gamma\left(\frac{1}{2}(1+v)\right)^{\frac{3}{2}}}.$$

The height of the mode of the EP distribution could be higher or lower than the Normal distribution, depending on the value of parameter v. The Normal distribution is a special form of EP distribution when $v=0$.

Iriawan (2000) has developed the Neo-Normal distribution from modified Burr type II distribution (Burr, 1942). The Burr type II distribution is also known as Generalized Logistic type I (Johnson et al., 1995; Abdelfattah, 2015). The cumulative distribution function (CDF) and pdf of a random variable Z_2 that follows the Burr type II distribution are given by

$$F(z_2)=(1+exp(-z_2))^{-\alpha}, \quad \textbf{(2)}$$

and

$$f(z_2|\alpha)=\alpha\, exp(-z_2)(1+exp(-z_2))^{-(\alpha+1)}, \quad \textbf{(3)}$$

respectively, where $-\infty < z_2 < \infty$, and $\alpha > 0$. The mode of Burr type II distribution is varied according to the value of parameter α. Iriawan (2000) modified Equation (2) by transforming $Z_3 = Z_2 - log\,\alpha$ so that its CDF became as follows

$$F(z_3) = \left(1 + \frac{exp(-z_3)}{\alpha}\right)^{-\alpha}, \quad (4)$$

and the corresponding pdf in Equation (3) became

$$f(z_3|\alpha) = exp(-z_3)\left(1 + \frac{exp(-z_3)}{\alpha}\right)^{-(\alpha+1)}. \quad (5)$$

The distribution with CDF in the Equation (4) and pdf in Equation (5) was called the "Modified Stable Burr" or $MSBurr(\alpha)$. The mode of the MSBurr would be stable at $z_3 = 0$ for any value of the parameter α. Similar to Burr type II distribution, howe-ver, the density of its mode always lower than that of the Standard Normal distribu-tion. For comparison with $Normal(\mu, \sigma^2)$, Iriawan (2000), therefore, added a parameter ω and defined a transformation as follows

$$Z_4 = \tilde{\mu} + \frac{\tilde{\sigma}}{\omega} Z_3.$$

The CDF in Equation (4) would be as follows:

$$F(z_4) = \left(1 + \frac{exp(-\omega\left(\frac{z_4-\tilde{\mu}}{\tilde{\sigma}}\right))}{\alpha}\right)^{-\alpha}, \quad (6)$$

and its pdf in Equation (5) is transformed into:

$$f(z_4|\omega, \alpha, \tilde{\mu}, \tilde{\sigma}) = \frac{\omega}{\tilde{\sigma}} exp\left(-\omega\left(\frac{z_4-\tilde{\mu}}{\tilde{\sigma}}\right)\right)\left(1 + \frac{exp\left(-\omega\left(\frac{z_4-\tilde{\mu}}{\tilde{\sigma}}\right)\right)}{\alpha}\right)^{-(\alpha+1)}, \quad (7)$$

where $-\infty < z_4 < \infty, -\infty < \tilde{\mu} < \infty, \tilde{\sigma} > 0, \alpha > 0$.

As described above, Iriawan (2000) has derived MSBurr distribution from modified Burr type II distribution, such that its mode is stable at its location parameter μ either it is symmetric or skewed. Further, the MSBurr distribution could be modified, in such that its peak as high as certain symmetric unimodal distribution. The step for the last modification is described in the Theorem 1.

Theorem 1. Making the density of MSBurr's mode the same as the density of other symmetrical unimodal distributions' mode

Suppose Z_4 follows $MSBurr(\omega, \alpha, \tilde{\mu}, \tilde{\sigma})$ and Z^* follows a symmetric uni-modal distribution with pdf as

$$h(z^*|\mu^*, \sigma^*, \theta^*) = g\left(\frac{z^* - \mu^*}{\sigma^*}|\theta^*\right),$$

where $z^* \in R$, h and g are pdf of unstandardized and standardized Z^*respectively, μ^* is a location parameter; σ^* is a scale parameter, and θ^* is a shape parameter. If it is given that

i. $h(\mu^*|\sigma^*,\theta^*) = \frac{1}{\tilde{C}(\theta^*)\sigma^*}$, where $\tilde{C}(\theta^*)$ is the function of shape parameter that is a normalizing constant of g,

ii. $\tilde{\sigma} = \sigma^*$, and

iii. $f(\tilde{\mu}|\omega,\alpha,\tilde{\sigma}) = h(\mu^*|\sigma^*,\theta^*)$, where *f*(.) is pdf of MSBurr distribution,

then

$$\omega = \frac{\left(1+\frac{1}{\alpha}\right)^{(\alpha+1)}}{\tilde{C}(\theta^*)}.$$

Proof:

Suppose $Z_4 \sim MSBurr(\omega,\alpha,\tilde{\mu},\tilde{\sigma})$, so pdf of Z_4 on its location parameter $\tilde{\mu}$ follows

$$f(\tilde{\mu}|\omega,\alpha,\tilde{\sigma}) = \frac{\omega}{\tilde{\sigma}} exp\left(-\omega\left(\frac{\tilde{\mu}-\tilde{\mu}}{\tilde{\sigma}}\right)\right)\left(1+\frac{exp\left(-\omega\left(\frac{\tilde{\mu}-\tilde{\mu}}{\tilde{\sigma}}\right)\right)}{\alpha}\right)^{-(\alpha+1)},$$

$$= \frac{\omega}{\tilde{\sigma}}\left(1+\frac{1}{\alpha}\right)^{-(\alpha+1)}.$$

Given the mode of Z_4 is equal to Z^*,

$$f(\tilde{\mu}|\omega,\alpha,\tilde{\sigma}) = h(\mu^*|\sigma^*,\theta^*),$$

$$\frac{\omega}{\tilde{\sigma}}\left(1+\frac{1}{\alpha}\right)^{-(\alpha+1)} = \frac{1}{\tilde{C}(\theta^*)\sigma^*}.$$

If $\tilde{\sigma} = \sigma^*$, then we get

$$\omega = \frac{\left(1+\frac{1}{\alpha}\right)^{(\alpha+1)}}{\tilde{C}(\theta^*)}.$$

Corollary 1. MSNBurr distribution

The density of $MSBurr(\omega,\alpha,\tilde{\mu},\tilde{\sigma})$ on its mode will be equal to the density of $Normal(\tilde{\mu},\tilde{\sigma})$'s mode when

$$\omega = \frac{\left(1+\frac{1}{\alpha}\right)^{(\alpha+1)}}{\sqrt{2\pi}}. \quad \textbf{(8)}$$

Corollary 2. MSTBurr distribution

The density of $MSBurr(\omega,\alpha,\tilde{\mu},\tilde{\sigma})$ on its mode will be equal to the density of $t(\tilde{\mu},\tilde{\sigma},\ddot{v})'s$ mode when

$$\omega = \frac{\Gamma\left(\frac{\ddot{v}+1}{2}\right)\left(1+\frac{1}{\alpha}\right)^{(\alpha+1)}}{\sqrt{\ddot{v}\pi}\,\Gamma\left(\frac{\ddot{v}}{2}\right)}. \quad \textbf{(9)}$$

When the MSBurr distribution has ω as in Equation (8), it is called the "Modified to be Stable to Normal from Burr" or $MSNBurr(\alpha,\tilde{\mu},\tilde{\sigma})$. Meanwhile, the MSBurr dis-tribution is called the

"Modified to be Stable to t from Burr" or MSTBurr$(\ddot{v}, \alpha, \tilde{\mu}, \tilde{\sigma})$ when ω satisfies Equation (9) (Iriawan, 2000).

Following Corollary 1 and Corollary 2, the new modified MSBurr distribution with its peak as high as the mode of EP distribution is proposed. Equation (1) showed that the normalizing constant has a function of shape parameter as follows

$$\tilde{C}(v) = \frac{(1+v)\Gamma\left(\frac{1}{2}(v+1)\right)^{\frac{3}{2}}}{\Gamma\left(\frac{3}{2}(v+1)\right)^{\frac{1}{2}}}.$$

By employing Theorem 1, MSBurr's peak would be as high as EP's when

$$\omega = \frac{\Gamma\left(\frac{3}{2}(v+1)\right)^{\frac{1}{2}}\left(1+\frac{1}{\alpha}\right)^{(\alpha+1)}}{(1+v)\Gamma\left(\frac{1}{2}(v+1)\right)^{\frac{3}{2}}}. \qquad \textbf{(10)}$$

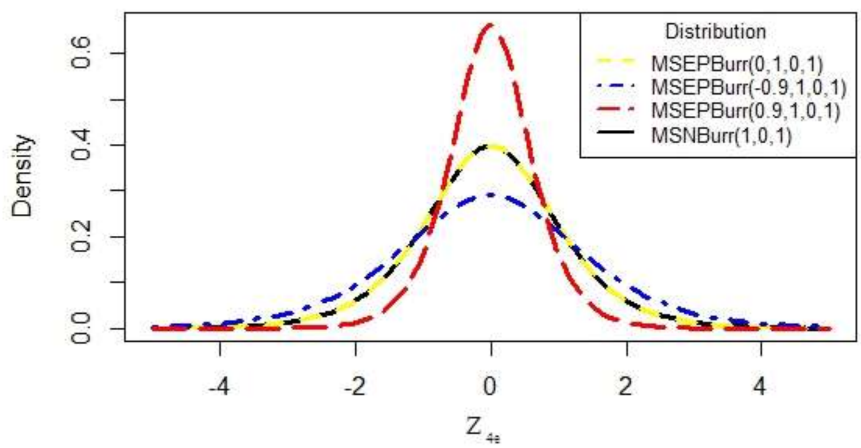


Figure 1. The comparison of MSEPBurr(0,1,0,1), MSEPBurr(-0.9,1,0,1), MSEPBurr (0.9,1,0,1), and MSNBurr(1,0,1)

The MSBurr distribution with ω satisfies Equation (10) is referred to as the MSEP-Burr distribution. Because it was derived from EP distribution, it is natural if its peak could be either lower or higher than that of MSNBurr distribution when their location and scale parameters are the same. The comparison of MSEPBurr distribution and MSNBurr distribution was shown in Figure 1**.** This figure shows that MSEPBurr distribution close to MSNBurr distribution when $v = 0$.

3. Properties of MSEPBurr Distribution

The MSEPBurr distribution is a particular form of MSBurr distribution, then its pro-perties, such as mean, variance, skewness, and kurtosis, could be measured from the central moment deriving from the cumulant generating function of the MSBurr. The cumulant generating function of $MSBurr(\omega, \alpha, \tilde{\mu}, \tilde{\sigma})$ is defined as follows

$$K_{Z_4}(t) = t\tilde{\mu} - \frac{\tilde{\sigma}t}{\omega}\log\alpha + \log\Gamma\left(\alpha + \frac{\tilde{\sigma}t}{\omega}\right) + \log\Gamma\left(1 - \frac{\tilde{\sigma}t}{\omega}\right) - \log\Gamma(\alpha). \quad (11)$$

The first moment of $MSBurr(\omega, \alpha, \tilde{\mu}, \tilde{\sigma})$ deriving from the first cumulant (κ_1) is

$$E(Z_4) = \kappa_1, \quad = \tilde{\mu} + \frac{\tilde{\sigma}}{\omega}\left(\psi_0(\alpha) - \psi_0(1) - \log\alpha\right), \quad (12)$$

where $\psi_0(.)$is digamma function. Furthermore, the *r*-th central moment deriving from *r*-th cumulant is

$$E(Z_4 - E(Z_4))^r = \kappa_r, \quad = \frac{\tilde{\sigma}^r}{\omega^r}\left(\psi_{(r-1)}(\alpha) + (-1)^r \psi_{(r-1)}(1)\right), \quad (13)$$

where $\psi_{(r-1)}(.)$is an *r*-th derivative of $\log\Gamma(.), r = 2,3, \ldots$. Based on moment in Equ-ation (12) and central moment in Equation (13) where ω follows Equation (10), the mean $(E(Z_{4e}))$, variance $(Var(Z_{4e}))$, skewness $(\gamma_1(Z_{4e}))$, and excess kurtosis $(\gamma_2(Z_{4e}))$ of Z_{4e} which follows the MSEPBurr distribution is defined as

$$E(Z_{4e}) = \tilde{\mu} + \frac{\tilde{\sigma}(1+v)\Gamma\left(\frac{1}{2}(v+1)\right)^{\frac{3}{2}}}{\Gamma\left(\frac{3}{2}(v+1)\right)^{\frac{1}{2}}\left(1+\frac{1}{\alpha}\right)^{(\alpha+1)}}\left(\psi_0(\alpha) - \psi_0(1) - \log\alpha\right), \quad (14)$$

$$Var(Z_{4e}) = \frac{\tilde{\sigma}^2(1+v)\Gamma\left(\frac{1}{2}(v+1)\right)^3}{\Gamma\left(\frac{3}{2}(v+1)\right)\left(1+\frac{1}{\alpha}\right)^{2(\alpha+1)}}\left(\psi_1(\alpha) + \psi_1(1)\right), \quad (15)$$

$$\gamma_1(Z_{4e}) = \frac{(\psi_2(\alpha) - \psi_2(1))}{(\psi_1(\alpha) + \psi_1(1))^{\frac{3}{2}}}, \quad (16)$$

$$\gamma_2(Z_{4e}) = \frac{(\psi_3(\alpha) + \psi_3(1))}{(\psi_1(\alpha) + \psi_1(1))^2}, \quad (17)$$

respectively, where $\psi_1(.), \psi_2(.), \psi_3(.)$are trigamma, tetragamma, and pentagamma functions respectively .

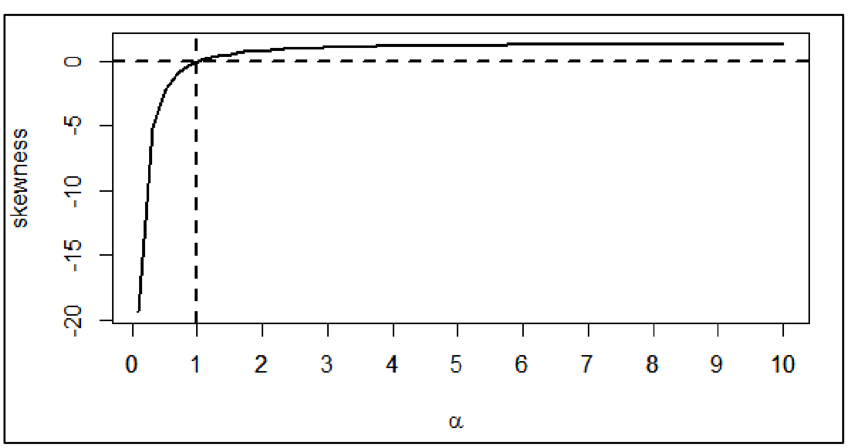


Figure 2. Plot of skewness of MSEPBurr where $0.1 \leq \alpha < 10$

It is easy to see that α is a parameter which has a role in determining the magnitude of the skewness and kurtosis. Figure 2 shows that MSEPBurr distribution is symme-tric when $\alpha = 1$. Otherwise, this distribution is left skew if $\alpha < 1$, and is right skew if $\alpha > 1$. It is shown that the magnitude of negative skewness is greater than positive skewness. It means that the MSEPBurr distribution more adaptively accommodate the left skew data than right skew one, in particular when $\alpha < 1$. Moreover, Figure 3 shows that the minimum value of excess kurtosis in MSEPBurr distribution is 1.2, which is when $\alpha = 1$. This shows that the MSEPBurr distribution is leptokurtic. The kurtosis is influenced by the value of skewness. The left skew MSEPBurr distribution has a sharper peak than the right skew one.

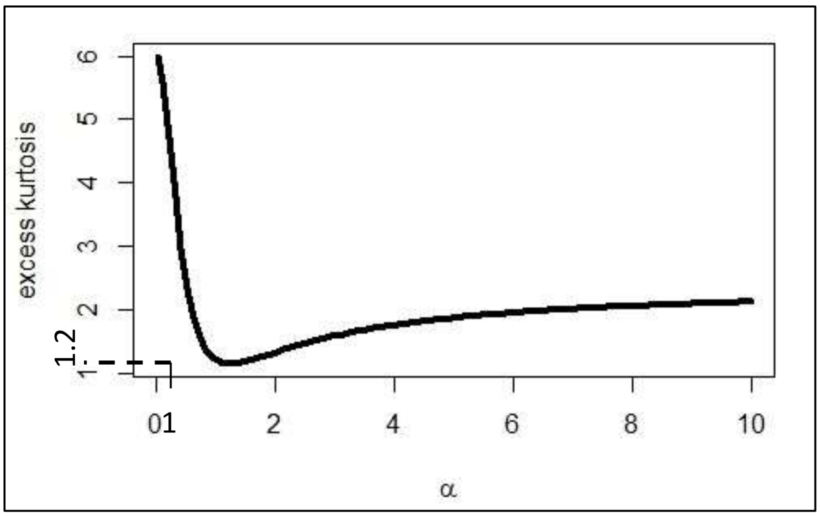


Figure 3. Plot of MSEPBurr's excess kurtosis

Another property discussed in this paper is the quantile of the MSEPBurr distribution. We obtain the quantile by using inverse of CDF in Equation (6), where ω follows Equation (10), that leads to

$$Q(u) = \tilde{\mu} - \frac{\tilde{\sigma}(1+v)\Gamma\left(\frac{1}{2}(v+1)\right)^{\frac{3}{2}}}{\Gamma\left(\frac{3}{2}(v+1)\right)^{\frac{1}{2}}\left(1+\frac{1}{\alpha}\right)^{(\alpha+1)}}\left(log\,\alpha + log\left(u^{-\frac{1}{\alpha}} - 1\right)\right) \qquad \textbf{(18)}$$

Based on the Equation (18), the random numbers that have MSEPBurr distribution could be drawn by using the invers transform as in Algorithm 1.

Algorithm 1. Generating the MSEPBurr random number

Step 1. Generate $u \sim U(0,1)$,
Step 2. Calculate $z_{4e} = Q(u)$ from Equation (18),
Step 3. Return z_{4e} (as MSEPBurr random number).

4. Parameter Estimation of MSEPBurr Using Bayesian

We estimate the MSEPBurr distribution parameters using the Bayesian approach. The parameter estimator is obtained from the posterior distribution, which is proportional to the likelihood times the prior distribution. Let Z_{4i} follows the $MSEPBurr(v, \alpha, \tilde{\mu}, \tilde{\sigma})$ distribution, i=1, 2, ...,n, where n is the sample size, then the likelihood of the MSEPBurr distribution is

$$f(\mathbf{z_{4e}}|v,\alpha,\tilde{\mu},\tilde{\sigma}) = \prod_{i=1}^{n} \frac{\Gamma\left(\frac{3}{2}(v+1)\right)^{\frac{1}{2}}\left(1+\frac{1}{\alpha}\right)^{(\alpha+1)}}{\tilde{\sigma}(1+v)\Gamma\left(\frac{1}{2}(v+1)\right)^{\frac{3}{2}}} \frac{E_i}{\left(1+\frac{E_i}{\alpha}\right)^{(\alpha+1)}},$$

$$= \frac{\Gamma\left(\frac{3}{2}(v+1)\right)^{\frac{n}{2}}\left(1+\frac{1}{\alpha}\right)^{n(\alpha+1)}}{\tilde{\sigma}^n(1+v)^n\Gamma\left(\frac{1}{2}(v+1)\right)^{\frac{3n}{2}}} \prod_{i=1}^{n} \frac{E_i}{\left(1+\frac{E_i}{\alpha}\right)^{(\alpha+1)}}, \tag{19}$$

where $\mathbf{z_{4e}} = (z_{4e1}, z_{4e2}, \ldots, z_{4en})^T$, and

$$E_i = exp\left(\frac{\Gamma\left(\frac{3}{2}(v+1)\right)^{\frac{1}{2}}\left(1+\frac{1}{\alpha}\right)^{(\alpha+1)}}{(1+v)\Gamma\left(\frac{1}{2}(v+1)\right)^{\frac{3}{2}}}\left(\frac{z_{4ei}-\tilde{\mu}}{\tilde{\sigma}}\right)\right).$$

In this research, the prior distributions for the MSEPBurr parameters are set to:

a. $\alpha \sim GSBeta(q^*, l_b, u_b), 0 < l_b < u_b < \infty$,
where GSBeta or Generalized Symmetrical Beta is a Beta distribution wich its domain is widened in the interval $l_b < \alpha < u_b$, and it has pdf as

$$f(\alpha) = \frac{\Gamma(2q^*)((\alpha-l_b)(u_b-\alpha))^{(q^*-1)}}{\Gamma(q^*)^2(u_b-l_b)^{(2q^*-1)}}, \tag{20}$$

where $q^* \geq 1$ (Box and Tiao, 1973),

b. $v \sim GSBeta(q_v{}^*, l_{bv}, u_{bv}), -1 < l_{bv} < u_{bv} < 1$, where its pdf is

$$f(v) = \frac{\Gamma(2q_v{}^*)((v-l_{bv})(u_{bv}-v))^{(q_v^*-1)}}{\Gamma(q_v{}^*)^2(u_{bv}-l_{bv})^{(2q_v{}^*-1)}}, \tag{21}$$

where $q_v{}^* \geq 1$

c. $\tilde{\mu} \sim Normal(v_0, \varphi_0{}^2)$ where its pdf is

$$f(\tilde{\mu}) = \frac{1}{\varphi_0\sqrt{2\pi}} exp\left(-\frac{1}{2}\frac{(\tilde{\mu}-v_0)^2}{\varphi_0{}^2}{}_0\right) \tag{22}$$

d. $\tilde{\sigma} \sim Inverse-gamma(a_0, b_0)$ where its pdf is

$$f(\tilde{\sigma}) = \frac{b_0{}^{a_0}}{\Gamma(a_0)} \tilde{\sigma}^{(-a_0-1)} exp\left(-\frac{b_0}{\tilde{\sigma}}\right). \tag{23}$$

The joint posterior distribution of MSEPBurr parameters, obtained by multiplying the likelihood in Equation (19) and the independent prior distributions in Equation (20) to Equation (23), is defined as follows

$$f(v,\alpha,\tilde{\mu},\tilde{\sigma}|\mathbf{z_{4e}}) \propto f(\mathbf{z_{4e}}|v,\alpha,\tilde{\mu},\tilde{\sigma})f(v)f(\alpha)f(\tilde{\mu})f(\tilde{\sigma}). \tag{24}$$

The full conditional distribution of each parameters derived from Equation (24) are

a. $$f(v|\mathbf{z_{4e}},\alpha,\tilde{\mu},\tilde{\sigma}) \propto \frac{\Gamma\left(\frac{3}{2}(v+1)\right)^{\frac{n}{2}}}{(1+v)^n\Gamma\left(\frac{1}{2}(v+1)\right)^{\frac{3n}{2}}}\prod_{i=1}^{n}\frac{E_i}{\left(1+\frac{E_i}{\alpha}\right)^{(\alpha+1)}} \times ((v-l_{bv})(u_{bv}-v))^{(q_v^*-1)} \tag{25}$$

b. $$f(\alpha|\mathbf{z_{4e}},v,\tilde{\mu},\tilde{\sigma}) \propto \left(1+\frac{1}{\alpha}\right)^{n(\alpha+1)}\prod_{i=1}^{n}\frac{E_i}{\left(1+\frac{E_i}{\alpha}\right)^{(\alpha+1)}} \times ((\alpha-l_b)(u_b-\alpha))^{(q^*-1)} \tag{26}$$

c. $$f(\tilde{\mu}|\mathbf{z_{4e}},\alpha,v,\tilde{\sigma}) \propto exp\left(-\frac{1}{2}\frac{(\tilde{\mu}-v_0)^2}{\varphi_0{}^2}{}_0\right)\prod_{i=1}^{n}\frac{E_i}{\left(1+\frac{E_i}{\alpha}\right)^{(\alpha+1)}} \tag{27}$$

d. $$f(\tilde{\sigma}|\mathbf{z_{4e}},\alpha,v,\tilde{\mu}) \propto \tilde{\sigma}^{-(n+a_0+1)}\,exp\left(-\frac{b_0}{\tilde{\sigma}}\right).\prod_{i=1}^{n}\frac{E_i}{\left(1+\frac{E_i}{\alpha}\right)^{(\alpha+1)}} \tag{28}$$

We employ Markov Chain Monte Carlo (MCMC) algorithm, particularly Gibbs Sam-pler algorithm in the computation of the MSEPBurr parameters estimation. This algo-rithm is described in Algorithm 2. Algorithm 2 could be applied into Bayesian In-ference Using Gibbs Sampler (BUGS) language (Lunn et al., 2000), that employs Just Another Gibbs Sampling (JAGS) software (Plummer, 2003). This program is run by the *runjags* package (Denwood, 2016) in R software (R Core Team, 2017). The MSEPBurr had been added in the *runjags* module as a new distribution in JAGS.

Algorithm 2. The Gibbs Sampler algorithm for the MSEPBurr parameter estimation

1. Set the initial value of $v_{(0)}, \alpha_{(0)}, . \tilde{\mu}_{(0)}, \tilde{\sigma}_{(0)}$.
2. For each t-th iteration, where t = 1,2, ..., T, T is number of samples,
    a. Generate $v_{(t)}$from $f\left(v_{(t)}|\mathbf{z_{4e}}, \alpha_{(t-1)}, \tilde{\mu}_{(t-1)}, \tilde{\sigma}_{(t-1)}\right)$ in Equation (25),
    b. Generate $\alpha_{(t)}$ from $f(\alpha_{(t)}|\mathbf{z_{4e}}, v_{(t)}, \tilde{\mu}_{(t-1)}, \tilde{\sigma}_{(t-1)})$ in Equation (26),
    c. Generate $\tilde{\mu}_{(t)}$from $f(\tilde{\mu}_{(t)}|\mathbf{z_{4e}}, \alpha_{(t)}, v_{(t)}, \tilde{\sigma}_{(t-1)})$ in Equation (27),
    d. Generate $\tilde{\sigma}_{(t)}$ from $f(\tilde{\sigma}_{(t)}|\mathbf{z_{4e}}, \alpha_{(t)}, v_{(t)}, \tilde{\mu}_{(t)})$ in Equation (28).
3. Return $v_{(1)}, \ldots, v_{(T)}, \alpha_{(1)}, \ldots, \alpha_{(T)}, \tilde{\mu}_{(1)}, \ldots, \tilde{\mu}_{(T)}, \tilde{\sigma}_{(1)}, \ldots, \tilde{\sigma}_{(T)}$.

The estimators of MSEPBurr parameter are computed from output in Algorithm 2. They are

$$\hat{v} = \frac{\sum_t^T v_{(t)}}{T},\ \hat{\alpha} = \frac{\sum_t^T \alpha_{(t)}}{T},\ \hat{\tilde{\mu}} = \frac{\sum_t^T \tilde{\mu}_{(t)}}{T},\ \hat{\tilde{\sigma}} = \frac{\sum_t^T \tilde{\sigma}_{(t)}}{T},$$

respectively.

5. Simulation Study

We do a simulation study to investigate the performance of the MSEPBurr distri-bution when it is applied to regression modeling. This simulation is started by ge-nerating data *y* and *x* which has a linear relationship as follows

$$\tilde{y}_i = \tilde{\beta}_0 + \tilde{\beta}_1 x_i + \tilde{\varepsilon}_i, i = 1,2,\ldots,n, \quad \textbf{(29)}$$

where n is the number of observations, $\tilde{\beta}_0$ was set to 1, $\tilde{\beta}_1$ was set to 2, and $\tilde{\varepsilon}_i$ follows *MSEPBurr*(0.8,10,0,1) that represents right-skew data. There are 4 generated numbers of data (*n*), i.e.: 10, 30, 100, and 1000. Moreover, the generating simulated data is repeated in 10 iterations.

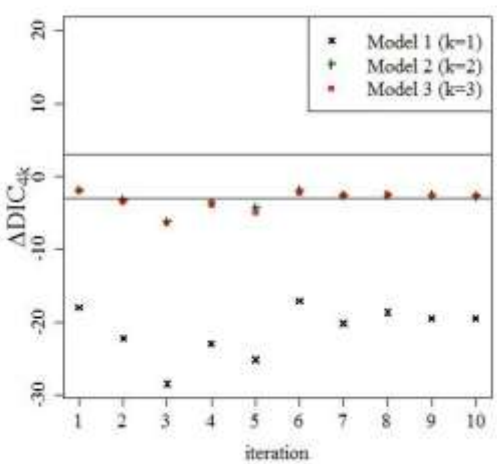
We define 4 regression models for each data simulation. These models are

- Model 1 is a simple linear regression with errors follow the Normal distribution,
- Model 2 is a simple linear regression with errors follow the MSNBurr distribution,
- Model 3 is a simple linear regression with errors follow the MSTBurr distribution,
- Model 4 is a simple linear regression with errors follow the MSEPBurr distribution.

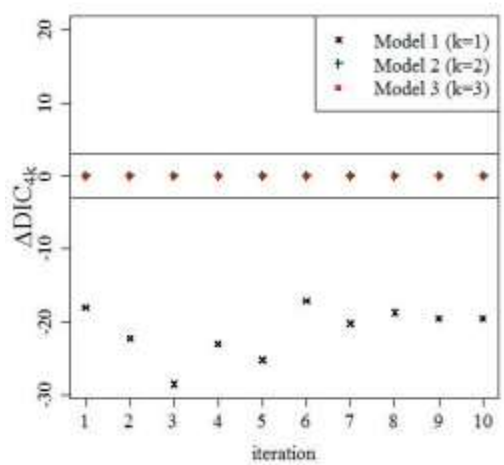
Next, the model parameters in each generated data are estimated by using a Bayesian approach. Each prior of $\tilde{\beta}_0$ and $\tilde{\beta}_1$ is *Normal*(0,0.1). In Model 1, the prior of the pre-cision parameter (τ) follows *Gamma*(1,1). The prior of the parameter α and $\tilde{\sigma}$ in Maudel 2, Model 3, and Model 4 follow *GSBeta*(1,0.1,15) and *Inverse-gamma*(1,1), respectively. The prior of the shape parameter $\nu°$ in Model 3 follows *Uniform*(1,100). The prior of the shape parameter ν in Model 4 follows *GSBeta*(1,-0.99,1). In addi-tion, the performance of each model is compared using its Deviance Information Cri-teria (DIC) (Spiegelhalter et al., 2002; Spiegelhalter et al., 2014). Carlin and Louis (2008) stated that when the DIC differences lies between 3 and 5, usually could be considered the smallest DIC is a better model. The difference between DIC of each model is denoted by

$$\Delta DIC_{4k} = DIC_4 - DIC_k,$$

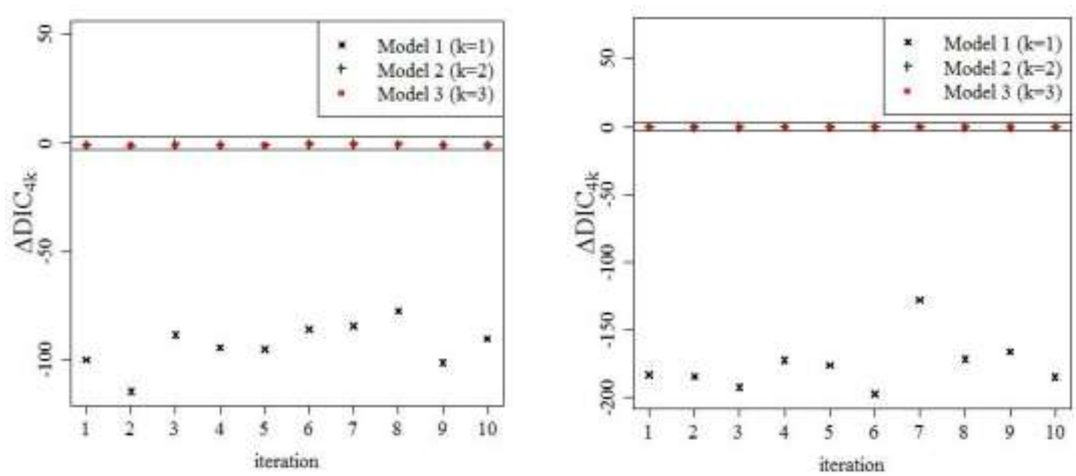
where DIC_4 is DIC of Model 4 and DIC_k is DIC of Model *k*, *k*=1, 2, 3. When $|\Delta DIC_{4k}| \leq 3$ then there is no evidence that Model 4 is better than Model *k*. If $\Delta DIC_{4k} > 3$, the performance of Model k can be considered better than the Model 4. Otherwise, if $\Delta DIC_{4k} < -3$, the performance of Model 4 is considered better than Model *k*.



(a) *n*=10



(b) *n*=30



(c) *n*=100 (d) *n*=1000

Figure 4. ΔDIC_{4k} **where** k**=1,2,3 in scenario 2 (a)** n**=10, (b)** n**=30, (c)** n**=100, (d)** n**=1000**

The comparison of each model performance based on simulation data is presented in Figure 4. Two horizontal lines in the middle of these figures are created as $\Delta DIC = 3$ and $\Delta DIC_4 = -3$, respectively. Figure 4 (a) shows that when *n*=10, Model 4 outper-forms over other models in 3 of 10 simulation data. However, Figure 4 (b), (c), and (d) shows that Model 4 has the same performance as Model 2 and Model 3. In addi-tion, Figure 4 shows that Model 1 always has the lowest performance because of Nor-mal distribution can not handle asymmetric residuals.

6. Application

In this section, the MSEPBurr regression was applied to two real data sets. The MSEPBurr distribution was compared to Normal, Student-t, MSNBurr, and MSTBurr distribution. In the first example, the regression model was applied using popular "Australian Athletes" data set that has been studied by Rubio and Genton (2016). In the second example, we employ the chemical reaction rate in Box and Tiao (1973) that has been analyzed by Albert et al. (1991). The DIC was used for model performance comparison. The computation of model parameters was also performed using JAGS software that is run using *runjags* package in R software. The posterior samples of each parameter are obtained by 5,000 burn-in in 255,000 iterations. Moreover, we used 25 thin to reduce autocorrelation in MCMC output. Using the *autorun* function in this package, the iteration could be automatically added when convergence has not been achieved. Furthermore, the convergence of MCMC was checked using potential scale reduction factor (PSRF) (Gelman and Rubin, 1992; Brooks and Gelman, 1997).

6.1 Australian athletes data

The model for the first data is as follows (Rubio and Genton, 2016)

$$y_i^* = \beta_1^* x_{1i}^* + \beta_2^* x_{2i}^* + \varepsilon^*{}_i, i = 1,2,\ldots,102, \quad \textbf{(30)}$$

where y_i^*, x_{1i}^*, and x_{2i}^*, denoted the lean body mass, height, and weight, respectively. The errors on the this model (ε_i^*) are assumed following Normal, Student-*t*, MSN-Burr, MSTBurr or MSEPBurr distribution, respectively. The prior of parameters in this model was specified by vague prior. The prior distributions of $\beta_1^*$and β_2^* are assigned to $Normal(0{,}10^6)$. The priors of the precision (σ^{-2}) in the Normal and Student-t distributions were specified to Gamma(0.001,0.001). The scale parameters (σ) in other distributions were set to Inverse-Gamma(0.001,0.001). The prior of degrees of freedom (v_t) in the Student-*t* model was determined to *Uniform*(1,50). Other than that, we set the prior of the parameters α of MSNBurr, MSTBurr, and MSEPBurr model as *GSBeta*(1,0,10) and the prior of the parameters $\ddot{v}$ in MSTBurr model and $\ddot{v}$ in MSEPBurr model were set to *Uniform*(1,50) and *GSBeta*(1,-0.9,1) respectively.

Table 1. Male Australian Athletes Data: Posterior mean, credible interval (CI), PSRF, and DIC

Parameter		Errors Distribution				
		Normal	Student-*t*	MSNBurr	MSTBurr	MSEPBurr
β_1^*	Mean	0.077	0.061	0.042	0.042	0.042
	CI	(0.06,0.10)	(0.04,0.08)	(0.02,0.06)	(0.02,0.07)	(0.02,0.06)
	PSRF	1,000	1,000	1.008	1,000	1,004
β_2^*	Mean	0.732	0.772	0.826	0.826	0.826
	CI	(0.69,0.78)	(0.73,0.82)	(0.77,0.88)	(0.77,0.88)	(0.77,0.88)
	PSRF	1,000	1.000	1.006	1,000	1,004
σ	Mean	2.299	1.259	1.486	1.459	1.600
	CI	(1.99,2.62)	(0.90,1.64)	(1.19,1.78)	(1.17,1.75)	(0.92,2.56)
	PSRF	1.000	1.000	1.000	1.000	1.000

Tabel 1. (continued)

v_t	Mean	-	2.517	-	-	-
	CI	-	(1.20,4.20)	-	-	-
	PSRF	-	1.000	-	-	-
$\ddot{v}$	Mean	-	-	-	25.516-	
	CI	-	-	-	(2.81,49.16)	-
	PSRF	-	-	-	1.000	-
v	Mean	-	-	-	-	-0.002
	CI	-	-	-	-	(-0.98,0.90)
	PSRF	-	-	- -		1.000
α	Mean	-	-	0.224	0.224	0.224
	CI	-	-	(0.06,0.4)	(0.07,0.4)	(0.07,0.40)
	PSRF	-	-	1.001	1.000	1.003
DIC		461,979	438,665	430,301	430,769	430,432

The posterior mean, credible interval (CI), PSRF, and the DIC based on the male Australian athletes data are shown in Table 1. The PSRF of all parameters are 1. It is shown that the MCMC output is convergence. The DIC of the Normal model is largest that it indicates Normal model have lowest performance. The DIC of the Student-*t* model show that the Student-*t* model is better than the Normal model. However, both distributions have higher DIC than those of MSNBurr, MSTBurr and MSEPBurr models, which the last three have almost similar DIC. This result indicates that MSNBurr, MSTBurr and MSEPBurr models are better than Normal and Student-*t* models to capture the pattern of male Australian athletes' data into a linear regression model. It is because the skewness parameters of their error distributions (α) are less than one, which cannot be accommodated by Normal and Student-*t*.

6.2 Chemical reaction rate data

The second data is modeled by the Box and Tiao formula (Box and Tiao, 1973)

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i,\ i = 1,2,\ldots,20, \qquad \textbf{(31)}$$

where

$$y_i = \log L_i, \beta_0 = log\, A - \frac{E}{S}\bar{T}^{-1},\ \beta_1 = \frac{E}{50{,}000S}, x_i = \left(T_i^{-1} - \bar{T}^{-1}\right) \times 50{,}000,$$

$$L_i = log\, A - \frac{E}{S}\frac{1}{T_i}, \text{ and } \bar{T}^{-1} = \sum_{i=1}^{20} T_i^{-1},$$

where S is the known gas constant, T is the temperature, and A and E is a constant to be estimated. The errors in this model (ε_i) are also assumed following Normal, Student-*t*, MSNBurr, MSTBurr or MSEPBurr distribution, respectively. We specified prior of these model parameters as vague prior. These priors are the same as the priors in the first example.

Table 2. Chemical Reaction Data Rate: Posterior mean, credible interval, PSRF, and DIC

Parameter		Errors Distribution				
		MSEPBurr	Student-t	MSNBurr	MSTBurr	MSEPBurr
β_0	Mean	-3.910	-3.994	-4.005	-4.005	-4.005
	CI	(-4.12,-3.68)	(-4.08,-3.91)	(-4.16,-3.87)	(-4.15,-3.86)	(-4.15,-3.86)
	PSRF	1.000	1.001	1.000	1.000	1.000
β_1	Mean	-0.115	-0.196	-0.176	-0.176	-0.175
	CI	(-0.21,-0.02)	(-0.24,-0.15)	(-0.24,-0.11)	(-0.24,-0.10)	(-0.24,-0.10)
	PSRF	1.001	1.003	1.000	1.000	1.000
σ	Mean	0.440	0.139	0.298	0.292	0.322
	CI	(0.30,0.60)	(0.05,0.31)	(0.18,0.43)	(0.18,0.43)	(0.14,0.55)
	PSRF	1.000	1.001	1.000	1.000	1.000
v_t	Mean	-	2.497	-	-	-
	CI	-	(1.00,4.90)	-	-	-
	PSRF	-	1.009	-	-	-
$\ddot{v}$	Mean	-	-	-	25.543	-
	CI	-	-	-	(2.31,48.81)	-
	PSRF	-	-	-	1.000	-
V	Mean	-	-	-	-	0.008
	CI	-	-	-	-	(-0.90,0.99)

	PSRF	-	-	-	-	1.000
α	Mean	-	-	6.577	6.558	6.546
	CI	-	-	(2.59,10)	(2.60,10)	(2.56,10)
	PSRF	-	-	1.000	1.000	1.000
DIC		26.796	13.903	18.919	18.852	18.889

Table 2 shows that the MCMC samples are converging on all parameters because their PSRF value is 1. It also shows that MSNBurr, MSTBurr and MSEPBurr have skewness parameter $\alpha > 1$. This parameter means they have a right-skew residuals. In addition, Student-*t* model have a degree of freedom $v_t = 0$ which it shows long tail residuals. Based on the DIC value, the Normal model seem have lowest performance in chemical reaction rate data. The MSNBurr, MSTBurr and MSEPBurr models have similar performance and they outperform Normal model. However, their performance is less than the Student-*t* model because their skewness can not well capture the long right tail.

7. Conclusion

This paper has presented MSEPBurr distribution as a general form of MSNBurr distribution. We also have studied the properties of this distribution. The mean and variance of MSEPBurr are affected by *v* parameter, but not for the skewness and kurtosis which they are only influenced by the α parameter. The simulation study showed that the MSEPBurr has better performance in some data, but in general, the performances of MSEPBurr, MSNBurr, and MSTBurr are almost the same. The MSEPBurr, MSNBurr, and MSTBurr models have similar performance when they are applied to male Australian athletes data and chemical reaction rate data. The MSEPBurr, MSNBurr and MSTBurr models outperform Normal and Student-*t* models in Australian athletes data because they perfectly accommodate left skew residuals. However, performance of MSEPBurr, MSNBurr and MSTBurr is lower than Stu-dent-*t* model in chemical reaction rate data because their skewness are not perfectly accommodate long right tail.

Acknowledgements

The first author gratefully acknowledges the support received from BPS-Statistics In-donesia in terms of doctoral scholarship. The authors are grateful to the reviewers for their significant advices to improve this paper. The authors also thank Sarni Maniar Berliana for some comments and corrections.

References

1. Abdelfattah, A. M. (2015). Skew-Type I Generalized Logistic Distribution and its Properties. *Pakistan Journal of Statistics and Operation Research*, 11(3), 267–282.
2. Albert, J., Delampady, M., and Polasek, W. (1991). A Class of Distributions for Robustness Studies. *Journal of Statistical Planning and Inference*, 28(3), 291–304.

3. Azzalini, A. (1985). A Class of Distribution which Includes the Normal Ones. *Scandinavian Journal of Statistics*, 12(2), 171–178.
4. Box, G. E., and Tiao, G. C. (1973). *Bayesian Inference in Statistical Analysis*. Reading, Massachusetts: Addison Wesley.
5. Brooks, S. P., and Gelman, A.(1997). General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics*, 7, 434–455.
6. Burr, I. W. (1942). Cumulative Frequency Functions. *The Annals of Mathematical Statistics*, 13(2), 215–232.
7. Carlin, B. P., and Louis, T. A. (2008). *Bayesian Methods for Data Analysis* (Third ed.). Boca Raton: Chapman and Hall/CRC.
8. Denwood, M. J. (2016). runjags: An R Package Providing Interface Utilities, Model Templates, Parallel Computing Methods and Additional Distributions for MCMC Models in JAGS. *Journal of Statistical Software*, 71(9), 1–25.
9. Fernandez, C., and Steel, M. (1998). On Bayesian Modelling of Fat Tails and Skewness. *Journal of the American Statistical Association*, 93(441), 359–371.
10. Gelman, A., and Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science,* 7, 457–511.
11. Iriawan, N. (2000). *Computationally Intensive Approaches to Inference in Neo-Normal Linear Models*. Ph.D. Thesis, CUT-Australia.
12. Johnson, N. L., Kotz, S., and Balakhrisnan, N. (1995). *Continuos Univariate Distribution* (2 ed., Vol. II). New York: John Wiley and Son.
13. Lunn, D., Thomas, A., Best, N., and Spiegelhalter, D. (2000). WinBUGS-A Bayesian Modelling Framework: Concepts, Structure, and Extensibility. *Statistics and Computing*, 10, 325–337.
14. Olosunde Akinlolu,A. (2011). On Skew-Normal Model for Economically Active Population. *Pakistan Journal of Statistics and Operation Research*, 7(2), 233–243.
15. Plummer, M. (2003). JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, March 20-22. Vienna, Austria.
16. R Core Team. (2017). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
17. Rubio, F. J., and Genton, M. G. (2016). Bayesian Linear Regression with Skew-Symmetric Error Distributions with Applications to Survival Analysis. *Statistics in Medicine*, 35(14), 2441–2454.
18. Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4), 583--639.
19. Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and van der Linde, A. (2014). The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(3), 485–493.
20. Subbotin, M. T. (1923). On the Law of Frequency of Errors. *Matematicheskii Sbornik*, 31, 296–301.