



Estimasi Parameter Distribusi *Mixture* Eksponensial dan Weibull dengan Metode Bayesian Markov *Chain* Monte Carlo

Ulfa Destiarina ^a, Mustika Hadijati ^b, Desy Komalasari ^c, Nurul Fitriyani ^{d,*}

^a Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Mataram, Mataram, 83125, Indonesia. Email: destya1412@gmail.com

^b Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Mataram, Mataram, 83125, Indonesia. Email: mustika.hadijati@unram.ac.id

^c Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Mataram, Mataram, 83125, Indonesia. Email: desykomalasari@unram.ac.id

^{d,*} Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Mataram, Mataram, 83125, Indonesia. Email: nurul.fitriyani@unram.ac.id

ABSTRACT

In parameter estimation, sometimes there are several problems that require the completion of a mixture distribution. This study aimed to apply the parameter estimation of exponential and Weibull mixture distribution in simulation data using the Bayesian Markov Chain Monte Carlo (MCMC) estimation method. The results obtained indicate that the analytic calculations of parameter estimation were more accurate than the calculations with the help of software, based on the terms of the suitability of the theory and its integration process.

Keywords: Bayesian; Exponential distribution; Markov *Chain* Monte Carlo (MCMC); Mixture distribution; Weibull distribution.

ABSTRAK

Dalam estimasi parameter, kadangkala terdapat beberapa permasalahan yang menuntut penyelesaian dengan suatu distribusi *mixture* atau distribusi campuran. Penelitian ini bertujuan untuk menerapkan estimasi parameter distribusi *mixture* eksponensial dan Weibull pada data simulasi dengan metode estimasi Bayesian Markov *Chain* Monte Carlo (MCMC). Hasil yang diperoleh menunjukkan bahwa perhitungan analitik estimasi parameter lebih akurat dibandingkan perhitungan dengan bantuan perangkat lunak, apabila dipandang dari segi kesesuaian teori serta proses integrasinya.

* Corresponding author.

Alamat e-mail: nurul.fitriyani@unram.ac.id

Keywords: Bayesian; Distribusi campuran; Distribusi eksponensial; Distribusi Weibull; Markov *Chain Monte Carlo* (MCMC).

Diserahkan: 07-06-2019; Diterima: 28-06-2019;

Doi: <https://doi.org/10.29303/emj.v1i1.30>

1. Pendahuluan

Dalam statistika, estimasi parameter digunakan untuk mencari parameter distribusi yang berkaitan dengan data pengamatan yang dimiliki. Namun terkadang terdapat beberapa permasalahan yang menuntut penyelesaian dengan suatu distribusi *mixture*, yang merupakan distribusi baru hasil penggabungan dua atau lebih distribusi. Distribusi ini akan menghasilkan suatu model baru yang berbeda dari model sebelum penggabungan dilakukan. McLachlan dan Basford (1988) menyatakan bahwa model *mixture* merupakan model khusus yang mampu memodelkan sifat multimodal data yang mencerminkan susunan beberapa sub-populasi atau grup.

Distribusi eksponensial merupakan salah satu distribusi yang tergolong sering digunakan untuk menyelesaikan berbagai permasalahan dalam statistika, salah satunya pada analisis reliabilitas. Selain eksponensial, salah satu distribusi yang juga banyak digunakan dalam pemecahan masalah yang berkaitan dengan teori keandalan (reliabilitas) adalah distribusi Weibull. Kedua distribusi ini menggambarkan keseluruhan data secara jelas, terutama dalam pengujian dan permodelan data.

Namun apabila dalam kasus yang komponennya merupakan penggabungan dari beberapa sub-populasi, belum tentu analisisnya dapat diselesaikan menggunakan distribusi eksponensial atau Weibull biasa. Oleh karena itu, diperlukan suatu distribusi baru yang tepat dan sesuai dalam menyelesaikan permasalahan tersebut. Distribusi baru ini dapat berupa distribusi *mixture* eksponensial dan Weibull yang merupakan hasil penggabungan dari kedua distribusi tersebut.

Dalam estimasi parameter suatu distribusi, terdapat beberapa metode yang dapat digunakan, salah satunya metode Bayesian. Metode estimasi non-klasik atau Bayesian ini telah diaplikasikan secara luas dalam teori dan analisis statistik. Dalam konsep Bayesian, parameter dipandang sebagai suatu variabel acak yang tidak diketahui. Hal ini sejalan dengan kenyataan yang terjadi dalam banyak fenomena kehidupan nyata, dimana parameter tidak dapat diperlakukan sebagai suatu hal yang dapat bernilai konstan terus menerus (Ieren dan Oguntunde, 2018). Pada metode ini, parameter yang digunakan merupakan variabel acak yang memiliki distribusi *prior*, yaitu distribusi subyektif yang didasarkan pada keyakinan seseorang dan dirumuskan sebelum data sampel diambil (Walpole

dan Myers, 1995). Distribusi sampel yang digabung dengan distribusi *prior* akan menghasilkan suatu distribusi, yaitu distribusi *posterior* (Horst dalam penelitian Hazhiah, dkk., 2012).

Pemodelan menggunakan metode Bayesian didasarkan pada model *posterior*, yaitu memadukan data masa lalu sebagai informasi untuk distribusi *prior* dan data pengamatan yang digunakan sebagai penyusun fungsi *likelihood*. Distribusi *posterior* dari bentuk Bayesian yang sangat rumit ini dapat diperoleh dengan pendekatan numerik Markov *Chain Monte Carlo* (MCMC). Dalam hal ini, MCMC merupakan proses integrasi hasil perpaduan antara Monte Carlo dengan sifat Markov *Chain*. Penelitian ini bertujuan untuk menerapkan hasil estimasi parameter distribusi *mixture* eksponensial dan Weibull dengan metode Bayesian MCMC pada data simulasi.

2. Distribusi Eksponensial

Definisi 1 (Bonald dan Feuillet, 2011).

Variabel acak X berdistribusi eksponensial, dengan parameter $\beta > 0$, bila fungsi densitasnya berbentuk :

$$f(x; \beta) = \begin{cases} \beta e^{-x\beta} & x \geq 0, x \in R \\ 0 & \text{untuk } x \text{ lainnya} \end{cases}$$

3. Distribusi Weibull Dua Parameter

Definisi 2 (Walpole dan Myers, 1995).

Suatu variabel acak kontinu X berdistribusi Weibull dengan parameter α dan β jika fungsi densitas peluangnya adalah sebagai berikut.

$$f(x) = \begin{cases} \alpha \beta x^{\alpha-1} e^{-\beta x^\alpha} & x \geq 0 \\ 0 & \text{yang lain} \end{cases}$$

dimana $\alpha > 0$ dan $\beta > 0$.

4. Distribusi *Mixture*

McLachlan dan Basford (1988) menyatakan bahwa model *mixture* merupakan model khusus yang dapat memodelkan sifat multimodal data yang mencerminkan susunan beberapa sub populasi atau grup, dimana setiap sub populasi merupakan komponen penyusun dari model *mixture* serta mempunyai proporsi yang bervariasi untuk masing-masing komponennya.

Pola multimodal didekati dengan pola *mixture* yang dapat direpresentasikan sebagai berikut (Iriawan, 2001).

$$f(x|\theta, w) = \sum_{j=1}^n w_j g_j(x|\theta_j) \quad (1)$$

Fungsi $f(t|\theta, w)$ adalah fungsi densitas dari model *mixture*; $g_j(x|\theta_j)$ adalah fungsi dari sebanyak n komponen penyusun model *mixture*; θ_j adalah vektor parameter dengan elemen-elemen $(\theta_1, \theta_2, \dots, \theta_n)$; w adalah parameter proporsi dengan elemen-elemen (w_1, w_2, \dots, w_n) ; w_j adalah parameter proporsi komponen *mixture* dengan $\sum_{j=1}^n w_j = 1$ serta $0 \leq w_j \leq 1, j = 1, 2, \dots, n$; dan n adalah banyaknya komponen dalam *mixture*.

5. Metode Estimasi Bayesian

Misalkan $X_1, X_2, X_3, \dots, X_n$ merupakan sebuah sampel acak berukuran n dari distribusi yang mempunyai fungsi densitas peluang berbentuk $f(x; \theta)$, $\theta \in \Omega \subset \mathfrak{R}$. Langkah-langkah untuk menentukan estimasi Bayesian bagi θ adalah (Herrhyanto, 2011):

1. Penentuan fungsi densitas peluang gabungan dari $X_1, X_2, X_3, \dots, X_n$.

$$g(x_1, x_2, \dots, x_n; \theta) = f(x_1; \theta) \cdot \dots \cdot f(x_n; \theta) \quad (2)$$

2. Penentuan fungsi densitas dari θ , yang besarnya dipilih dan disesuaikan dengan $g(x_1, x_2, \dots, x_n)$. Distribusi yang mempunyai fungsi densitas dari θ , dinotasikan $\lambda(\theta)$, dinamakan distribusi *prior*.
3. Estimator Bayesian untuk θ ditentukan oleh:

$$\delta(x_1, \dots, x_n) = \frac{\int \theta \cdot g(x_1, \dots, x_n; \theta) \cdot \lambda(\theta) d\theta}{\int g(x_1, \dots, x_n; \theta) \cdot \lambda(\theta) d\theta} \quad (3)$$

Distribusi *posterior* ditentukan dengan:

$$p(\theta|x) = \frac{g(x_1, x_2, \dots, x_n; \theta) \cdot \lambda(\theta)}{\int g(x_1, x_2, \dots, x_n; \theta) \cdot \lambda(\theta) d\theta} \quad (4)$$

Definisi 3 (Walpole dan Myers, 1995)

Rata-rata distribusi *posterior* $p(\theta|x_1, x_2, x_3, \dots, x_n)$, dinyatakan dengan θ^* , disebut estimasi Bayesian untuk θ .

Definisi 4 (Walpole dan Myers, 1995)

Selang $a < \theta < b$ akan disebut selang Bayesian $(1-\alpha)$ 100% untuk θ bila,

$$\begin{aligned} \int_{\theta^*}^b f(\theta|x_1, \dots, x_n) d\theta &= \int_a^{\theta^*} f(\theta|x_1, \dots, x_n) d\theta \\ &= \frac{1 - \alpha}{2} \end{aligned}$$

Selanjutnya, estimasi densitas *mixture* dengan metode Bayesian dilakukan untuk menemukan distribusi *posterior* dengan mengalikan distribusi *prior* dengan *likelihood* datanya (Gamerman, 1995).

Hal-hal yang perlu diperhatikan dalam estimasi parameter dengan metode Bayesian, yaitu:

1. Distribusi *prior*
2. Fungsi *likelihood*
3. *Likelihood* model *mixture*, yang merupakan penggabungan dari beberapa fungsi *likelihood* pembentuknya. Dari persamaan (1) dengan $\sum_{i=1}^k n_i = n$, dan w adalah proporsi komponen *mixture*, maka *likelihood* model *mixture* yang akan terbentuk adalah (Paramita dan Iriawan, 2014):

$$\begin{aligned} L_{fmix}(\theta, w|t) &= \prod_{i=1}^{n_1} w_1 f(t_i|\theta_1) + \dots \\ &+ \prod_{i=1}^n w_k f(t_i|\theta_k) \end{aligned} \quad (5)$$

4. Distribusi *posterior*

Dalam analisis Bayesian, semua parameter dalam model *mixture* yang akan diduga diperlakukan sebagai suatu variabel acak dari parameter θ . Dengan menghilangkan pengaruh nilai total peluang, $p(\theta|x)$ sebagai bentuk *posterior* dapat dinyatakan dalam bentuk proporsional sebagai perkalian antara *likelihood* $L(x|\theta)$ dengan *prior* $p(\theta)$ sebagai berikut (Paramita dan Iriawan, 2014).

$$p(\theta|x) \propto L(x|\theta) p(\theta) \quad (6)$$

6. Markov Chain Monte Carlo

Markov Chain Monte Carlo (MCMC) merupakan pendekatan yang digunakan untuk mengatasi kesulitan dalam menentukan integral *posterior* yang membutuhkan proses integrasi yang rumit dan cukup lama. MCMC membangkitkan data sampel parameter θ yang memiliki distribusi tertentu menggunakan gibbs *sampling* (Ntzoufras, 2009).

MCMC adalah suatu metode simulasi yang menggunakan algoritma integrasi Monte Carlo untuk mencari nilai parameter model yang sesuai dengan proses Markov Chain. Dalam implementasinya, perangkat lunak WinBUGS dapat digunakan untuk memudahkan proses estimasi parameter model *mixture* ini (Iriawan, 2011).

Markov Chain atau Rantai Markov adalah proses stokastik dimana waktu dan ruang keadaan adalah diskrit. Dalam terapannya secara sederhana, sifat Markov adalah:

$$P[X_{n+1} = x_{n+1} | X_n = x_n, \dots, X_0 = x_0] \\ = P[X_{n+1} = x_{n+1} | X_n = x_n] \quad (7)$$

Probabilitas ini menyatakan bahwa probabilitas pada waktu (langkah) ke $(n + 1)$ hanya dipengaruhi oleh langkah ke n (sifat Markov) dan tidak dipengaruhi oleh langkah-langkah sebelumnya. Probabilitas bersyarat $P[X_{n+1} = x_{n+1} | X_n = x_n]$ disebut probabilitas transisi satu langkah (Haryono, 1995).

Pendekatan Monte Carlo adalah metode yang dikembangkan oleh fisikawan menggunakan generasi nomor acak untuk menghitung integral. Misalkan kita ingin menghitung integral kompleks :

$$\int_a^b h(x) dx \quad (8)$$

Jika kita dapat menguraikan $h(x)$ ke dalam fungsi $f(x)$ dan fungsi densitas peluang $p(x)$ yang didefinisikan melalui interval (a, b) , maka diketahui bahwa :

$$\int_a^b h(x) dx = \int_a^b f(x)p(x) dx = E_{p(x)}[f(x)] \quad (9)$$

Oleh karenanya, integral dapat dinyatakan sebagai ekspektasi $f(x)$ terhadap densitas $p(x)$. Jadi, jika kita mengambil sejumlah bilangan x_1, x_2, \dots, x_n dari suatu variabel acak dengan fungsi densitas $p(x)$, maka,

$$\int_a^b h(x) dx = E_{p(x)}[f(x)] \approx \frac{1}{n} \sum_{i=1}^n f(x_i) \quad (10)$$

Hal ini disebut sebagai integrasi Monte Carlo. Integrasi Monte Carlo dapat digunakan untuk perkiraan distribusi *posterior* (atau marginal *posterior*) yang diperlukan untuk analisis Bayesian. Dapat diperoleh hasil integral $I(y)$ dari perkiraan dengan perhitungan berikut.

$$I(y) = \int f(y|x) p(x) dx \\ \hat{I}(y) = \frac{1}{n} \sum_{i=1}^n f(y|x_i) \quad (11)$$

dimana x_i diambil dari fungsi densitas $p(x)$. Untuk estimasi standar *error* Monte Carlo diberikan oleh (Walsh, 2004):

$$SE^2[\hat{I}(y)] = \frac{1}{n} \left(\frac{1}{n-1} \sum_{i=1}^n (f(y|x_i) - \hat{I}(y))^2 \right) \quad (12)$$

7. Gibbs Sampling

Skenario yang digunakan dalam pengambilan data sampel pada umumnya adalah dengan metode *gibbs sampling* (Casella dan George, 1992). *Gibbs*

sampling merupakan generator yang sangat efisien, sehingga sering digunakan sebagai generator variabel acak pada analisis data yang menggunakan metode MCMC (Iriawan, 2000). *Gibbs sampling* adalah teknik terkenal untuk menghasilkan sampel dari distribusi multivariat yang rumit yang sering digunakan dalam prosedur Monte Carlo. Dalam bentuk yang paling sederhana, *gibbs sampling* digunakan untuk memperbarui setiap variabel dari distribusi bersyarat yang diberikan oleh semua variabel lain dalam sistem (Rasmussen, 2000).

Metode *gibbs sampling* dimulai dengan beberapa nilai y_0 awal untuk y dan memperoleh x_1 dengan menghasilkan variabel acak dari distribusi bersyarat $p(x|y = y_0)$. Langkah berikutnya, x_1 digunakan untuk menghasilkan nilai baru dari y_1 yang diperoleh dari distribusi bersyarat berdasarkan nilai x_0 , $p(y|x = x_0)$. Prosesnya dapat dituliskan sebagai berikut.

$$x_i \sim p(x|y = y_{i-1}) \quad (13)$$

$$y_i \sim p(y|x = x_i) \quad (14)$$

dengan mengulang proses ini sebanyak n kali (Walsh, 2004).

8. Data dan Metode

Penelitian ini terdiri dari 2 (dua) tahapan, yaitu estimasi parameter secara analitik, yang kemudian dilanjutkan dengan estimasi parameter dengan data simulasi menggunakan *software WinBUGS*. Estimasi parameter secara analitik, dilakukan dengan langkah-langkah sebagai berikut.

1. Menggabungkan distribusi-distribusi yang akan diteliti, yaitu dua distribusi serumpun distribusi eksponensial dan distribusi Weibull dua parameter.
2. Mengestimasi parameter dari distribusi yang telah digabungkan (distribusi *mixed*).
3. Melakukan estimasi parameter distribusi *mixture*:
 - a. Menentukan distribusi *mixture* hasil penggabungan distribusi eksponensial dan distribusi Weibull, dan diperoleh fungsi densitas *mixture* yang digunakan dalam estimasi parameter
 - b. Menentukan distribusi *prior* berdasarkan pengetahuan dan pertimbangan peneliti, dengan didasarkan pada bentuk fungsi densitas *mixture* yang diperoleh.
 - c. Menentukan fungsi *likelihood*.
 - d. Menentukan distribusi *posterior*.
 - e. Mengestimasi parameter populasi berdasarkan definisi Bayesian, diperoleh dengan mencari nilai rata-rata dari distribusi *posterior*.

4. Menguji estimator yang telah diperoleh dari distribusi gabungan dengan empat kriteria pengujian, yaitu tak bias, varian minimum (effisien), konsisten, dan *Mean Squared Erro* (MSE) minimum.

Selanjutnya, hasil estimasi parameter yang diperoleh secara analitik, dibandingkan dengan hasil estimasi parameter yang diperoleh dengan data simulasi, yang diperoleh melalui langkah-langkah sebagai berikut.

1. Membangkitkan data yang berdistribusi eksponensial dengan parameter awal $\beta_0 = 1$ dan $\beta_0 = 0.5$, dengan beberapa jumlah sampel parameter hasil bangkitan gibbs *sampling*, yaitu 1000, 5000, dan 8000.
2. Membangkitkan data yang berdistribusi Weibull dua parameter dengan beberapa ragam nilai parameter awal, yaitu $\alpha_0 = 2$, $\alpha_0 = 3$ dan $\beta_0 = 1$, $\beta_0 = 0.5$, serta nilai pembobot $w_1 = 0.75$ dan $w_1 = 0.25$.
3. Menggabungkan distribusi eksponensial dan Weibull
4. Mengestimasi parameter model *mixture* eksponensial dan Weibull.

9. Hasil dan Pembahasan

Berdasarkan Definisi 1 dan Definisi 2, diperoleh fungsi distribusi *mixture* (campuran) dari eksponensial dan Weibull untuk dua sub populasi berikut.

$$f(x|w, \alpha, \beta) = w_1 \alpha \beta x^{\alpha-1} e^{-\beta x^\alpha} + w_2 \beta e^{-x^\beta}$$

Parameter yang akan diestimasi adalah parameter skala β , dengan masing-masing parameter bentuk α dan pembobot w_1, w_2 telah diketahui. Diasumsikan bahwa distribusi *prior* untuk parameter β adalah distribusi eksponensial, dengan parameter *prior* β_0 . Dalam distribusi *prior*, variabel acak dari fungsi densitas diubah menjadi parameter, sehingga variabel x diganti dengan parameter β dan diperoleh fungsi *prior* sebagai berikut.

$$P(\beta) = \beta_0 e^{-\beta \beta_0}$$

Hal lain yang dibutuhkan dalam mengestimasi parameter suatu distribusi *mixture* adalah menentukan fungsi *likelihood* berdasarkan fungsi densitas dari distribusi *mixture* eksponensial dan Weibull, yaitu:

$$L(x|\beta) = (w_1)^{n_1} (\alpha \beta)^{n_1} e^{-\beta \sum_{i=1}^{n_1} x_i^\alpha} \prod_{i=1}^{n_1} x_i^{\alpha-1} + (w_2)^{n_2} \beta^{n_2} e^{-\beta \sum_{i=1}^{n_2} x_i}$$

Dan diperoleh bentuk marginal berikut.

$$m(\beta) = \int_{-\infty}^{\infty} \left[\frac{(w_1)^{n_1} (\alpha \beta)^{n_1} e^{-\beta \sum_{i=1}^{n_1} x_i^\alpha} \prod_{i=1}^{n_1} x_i^{\alpha-1}}{(w_1)^{n_1} (\alpha \beta)^{n_1} e^{-\beta \sum_{i=1}^{n_1} x_i^\alpha} \prod_{i=1}^{n_1} x_i^{\alpha-1} + (w_2)^{n_2} \beta^{n_2} e^{-\beta \sum_{i=1}^{n_2} x_i}} \right] \cdot [\beta_0 e^{-\beta \beta_0}] d\beta$$

Berdasarkan definisi Bayesian diperoleh distribusi *posterior*:

$$P(\beta|x) = \frac{[(w_1)^{n_1} (\alpha \beta)^{n_1} e^{-\beta \sum_{i=1}^{n_1} x_i^\alpha} \prod_{i=1}^{n_1} x_i^{\alpha-1} + (w_2)^{n_2} \beta^{n_2} e^{-\beta \sum_{i=1}^{n_2} x_i}] [\beta_0 e^{-\beta \beta_0}]}{\int_{-\infty}^{\infty} [(w_1)^{n_1} (\alpha \beta)^{n_1} e^{-\beta \sum_{i=1}^{n_1} x_i^\alpha} \prod_{i=1}^{n_1} x_i^{\alpha-1} + (w_2)^{n_2} \beta^{n_2} e^{-\beta \sum_{i=1}^{n_2} x_i}] [\beta_0 e^{-\beta \beta_0}] d\beta}$$

Menurut Ntzoufras (2009), fungsi marginal untuk Bayesian *mixture* eksponensial dan Weibull adalah proporsional dari fungsi peluang bersyarat yang

bernilai 1. Oleh karena itu, dapat ditulis distribusi *posterior* berdasarkan metode Bayesian sebagai berikut.

$$P(\beta|x) \propto \frac{[(w_1)^{n_1} (\alpha \beta)^{n_1} e^{-\beta \sum_{i=1}^{n_1} x_i^\alpha} \prod_{i=1}^{n_1} x_i^{\alpha-1} + (w_2)^{n_2} \beta^{n_2} e^{-\beta \sum_{i=1}^{n_2} x_i}] [\beta_0 e^{-\beta \beta_0}]}{1}$$

$$\propto \left[(w_1)^{n_1} (\alpha\beta)^{n_1} e^{-\beta \sum_{i=1}^{n_1} x_i^\alpha} \prod_{i=1}^{n_1} x_i^{\alpha-1} + (w_2)^{n_2} \beta^{n_2} e^{-\beta \sum_{i=1}^{n_2} x_i} \right] [\beta_0 e^{-\beta\beta_0}]$$

Selanjutnya, rata-rata dari distribusi *posterior mixture* eksponensial dan Weibull yang merupakan estimator Bayesian untuk parameter β , diperoleh sebagai berikut.

$$E(x) = \int_{-\infty}^{\infty} x f(x) dx$$

Dalam hal ini, parameter β menjadi variabel dan fungsi distribusi *posterior* yang diperoleh sebagai fungsi densitasnya. Jadi, diperoleh rata-rata dari distribusi *posterior mixture* eksponensial dan Weibull berikut.

$$E(\beta) = \hat{\beta} = \int_{-\infty}^{\infty} \beta P(\beta|x) d\beta$$

$$= \beta_0 \left(\frac{(w_1)^{n_1} \frac{\alpha^{n_1}}{(\beta_0 + \sum_{i=1}^{n_1} x_i^\alpha)^{n_1+2}} \prod_{i=1}^{n_1} x_i^{\alpha-1} \Gamma(n_1 + 2)}{+(1 - w_1)^{n_2} \frac{1}{(\beta_0 + \sum_{i=1}^{n_2} x_i)^{n_2+2}} \Gamma(n_2 + 2)} \right)$$

Persamaan rata-rata dari distribusi *posterior mixture* eksponensial dan Weibull yang diperoleh selanjutnya digunakan untuk menentukan persamaan varian distribusi *posterior mixture* eksponensial dan Weibull. diperoleh:

$$\text{Var}(\beta) = \beta_0 \left(\frac{(w_1)^{n_1} \frac{\alpha^{n_1}}{(\beta_0 + \sum_{i=1}^{n_1} x_i^\alpha)^{n_1+3}} \prod_{i=1}^{n_1} x_i^{\alpha-1} \Gamma(n_1 + 3)}{+(1 - w_1)^{n_2} \frac{1}{(\beta_0 + \sum_{i=1}^{n_2} x_i)^{n_2+3}} \Gamma(n_2 + 3)} \right) - \left(\beta_0 \left(\frac{(w_1)^{n_1} \frac{\alpha^{n_1}}{(\beta_0 + \sum_{i=1}^{n_1} x_i^\alpha)^{n_1+2}} \prod_{i=1}^{n_1} x_i^{\alpha-1} \Gamma(n_1 + 2)}{+(1 - w_1)^{n_2} \frac{1}{(\beta_0 + \sum_{i=1}^{n_2} x_i)^{n_2+2}} \Gamma(n_2 + 2)} \right) \right)^2$$

Dengan memperhatikan hasil yang diperoleh, dapat dikatakan bahwa estimator yang diperoleh bersifat bias bagi parameter β . Menurut Herrhyanto (2011), syarat awal terpenuhinya asumsi varian minimum adalah estimator yang diperoleh merupakan estimator yang tak bias, sehingga dapat dikatakan bahwa estimator yang diperoleh tidak memenuhi asumsi varian minimum.

Panik (2005) menyatakan bahwa untuk mengetahui terpenuhi atau tidaknya asumsi konsisten bagi suatu estimator, digunakan salah satu syarat dari asumsi konsisten yaitu varian dari estimator mendekati nol apabila $n \rightarrow \infty$. Hasil yang diperoleh menunjukkan bahwa estimator memenuhi asumsi konsisten.

Lebih jauh lagi, menurut Bolstad (2007), estimator Bayesian tidak menekankan kepada sifat tak bias dari sebuah estimator, karena estimator Bayesian biasanya merupakan estimator yang bias. Namun seringkali estimator Bayesian mempunyai MSE yang lebih kecil daripada estimator klasik. Oleh karenanya, estimator Bayesian dapat dikatakan lebih baik daripada estimator klasik ketika dinilai dengan kriteria MSE. Dengan demikian, dapat disimpulkan

bahwa estimator yang diperoleh merupakan estimator yang baik.

Selanjutnya, estimasi parameter *mixture* eksponensial dan Weibull pada data simulasi, dilakukan dengan menggunakan bantuan perangkat lunak (*software*) WinBUGS 1.4 yang berbasis metode Bayesian Markov Chain Monte Carlo (MCMC). Pada WinBUGS, tidak terdapat *tool* yang dapat digunakan untuk membangkitkan data. Oleh karena itu, data simulasi dibangkitkan dengan bantuan menggunakan perangkat lunak (*software*) Minitab. Data simulasi tersebut merupakan data acak berdistribusi eksponensial dan Weibull, dengan parameter distribusi yang beragam. Estimasi parameter model dilakukan setelah data simulasi dibangkitkan.

Berikut diberikan langkah-langkah estimasi parameter model *mixture* dengan menggunakan WinBUGS:

1. Menentukan fungsi densitas *mixture*.
2. Menentukan distribusi *prior* untuk parameter yang diestimasi (yang dalam hal ini adalah parameter skala β).
3. Membuat *syntax* fungsi *likelihood* dan parameter *prior*.

4. Memasukkan data simulasi bangkitan.
5. Membuat inisial untuk data dan parameter yang akan diestimasi.
6. Menjalankan program MCMC untuk memperoleh nilai estimator parameter.

Pada penelitian ini, parameter yang diestimasi hanya parameter skala β , dengan asumsi parameter

α_0 , β_0 , w_1 , dan w_2 telah diketahui. Di samping itu, parameter bentuk α yang merupakan parameter lain penyusun model *mixture* dan sekaligus bagian dari distribusi Weibull, dianggap sebagai konstanta dalam model *mixture*-nya. Hal ini mengakibatkan akibat munculnya asumsi baru oleh peneliti, yaitu parameter $\alpha = \alpha_0$.

Tabel 1 – Rata-rata dan Standar Deviasi

| Parameter Awal (Prior) | | | Jumlah Sampel Parameter Bangkitan (n_p) | | | | | |
|------------------------|------------|-----------|---|-----------------|-----------|-----------------|-----------|-----------------|
| Ekspensial | | Weibull | 1000 | | 5000 | | 8000 | |
| β_0 | α_0 | β_0 | Rata-rata | Standar deviasi | Rata-rata | Standar deviasi | Rata-rata | Standar deviasi |
| 1 | 2 | 1 | 0.8114 | 0.1805 | 0.8072 | 0.1760 | 0.8048 | 0.1773 |
| 0.5 | 1 | 0.5 | 0.8739 | 0.1944 | 0.8693 | 0.1895 | 0.8668 | 0.1910 |
| 1 | 3 | 1 | 0.0835 | 0.0186 | 0.0830 | 0.0181 | 0.0828 | 0.0182 |
| 0.5 | 3 | 0.5 | 0.1017 | 0.0226 | 0.1012 | 0.0220 | 0.1009 | 0.0222 |

Proses estimasi parameter β dengan MCMC, dengan beragam kondisi α_0, β_0, w_1 dan w_2 yang diketahui, menghasilkan nilai rata-rata, standar deviasi, variansi, dan MCMC *error*, tertera pada Tabel 1 dan Tabel 2.

Berdasarkan Tabel 1 dan Tabel 2 yang diberikan, terlihat bahwa nilai rata-rata, standar deviasi, dan variansi dari masing-masing parameter menunjukkan hasil yang beragam, namun relatif sama. Begitu pula ketika diestimasi dengan menggunakan beberapa jumlah sampel parameter dengan nilai beragam, baik rata-rata, standar deviasi, dan variansi hanya

memiliki sedikit perbedaan nilai untuk ketiga jumlah sampel parameter tersebut. Hal ini mengindikasikan bahwa dalam metode Bayesian MCMC, jumlah sampel yang berbeda dan keragaman nilai parameter akan menghasilkan nilai estimator yang relatif konsisten

Nilai MCMC *error* yang diperoleh menunjukkan nilai standar *error* untuk proses integrasi MCMC. Hasil estimasi yang tertera pada Tabel 2 menunjukkan bahwa nilai MCMC *error* relatif kecil. Semakin kecil nilai MCMC *error*, maka semakin baik pula hasil integrasinya.

Tabel 2 – Variansi dan MCMC Error

| Parameter Awal (Prior) | | | Jumlah Sampel Parameter Bangkitan (n_p) | | | | | |
|------------------------|------------|-----------|---|------------|----------|------------|----------|------------|
| Ekspensial | | Weibull | 1000 | | 5000 | | 8000 | |
| β_0 | α_0 | β_0 | Variansi | MCMC error | Variansi | MCMC error | Variansi | MCMC error |
| 1 | 2 | 1 | 0.03258 | 0.00506 | 0.03098 | 0.00284 | 0.03143 | 0.00174 |
| 0.5 | 1 | 0.5 | 0.03779 | 0.00545 | 0.03591 | 0.00306 | 0.03648 | 0.00188 |
| 1 | 3 | 1 | 0.00034 | 0.00052 | 0.00032 | 0.00029 | 0.00033 | 0.00017 |
| 0.5 | 3 | 0.5 | 0.00051 | 0.00059 | 0.00048 | 0.00035 | 0.00049 | 0.00021 |

Berdasarkan plot *history* yang tertera Tabel 3, terlihat bahwa plot *history* yang dihasilkan menggambarkan proses perubahan parameter pada setiap iterasi MCMC. Hasil yang diperoleh menunjukkan gambaran proses integrasi MCMC

yang berubah secara teratur dan stabil dalam domain parameter tertentu.

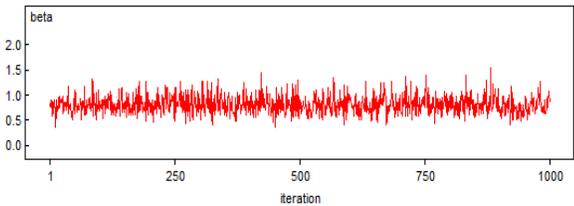
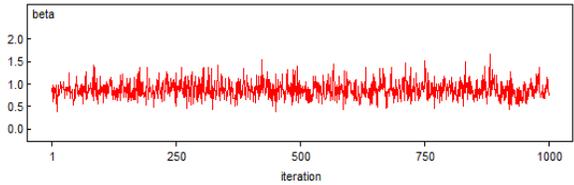
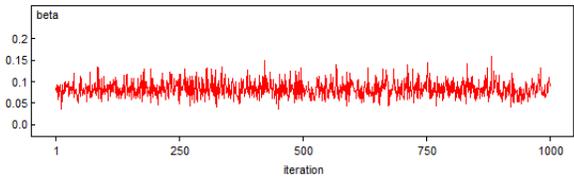
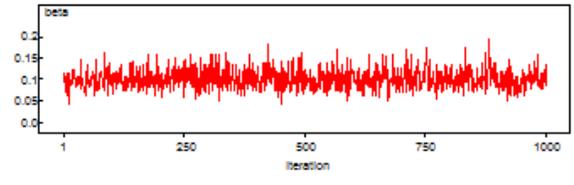
Selanjutnya, plot *autocorrelation* (autokorelasi) pada Tabel 4 menunjukkan autokorelasi antar parameter β . Hasil yang diperoleh menunjukkan

bahwa pada plot *autocorrelation* untuk ketiga kombinasi parameter menunjukkan signifikansi hanya pada *lag* ke-0. Hal ini menunjukkan bahwa sampel parameter yang dibangkitkan bersifat acak dan independen.

Adapun kernel *density* pada Tabel 4 menggambarkan bentuk *plot* dari distribusi *posterior*.

Masing-masing kernel *density* untuk ketiga kombinasi parameter menunjukkan bahwa pola distribusi *posterior* yang cenderung membentuk kurva normal dan relatif sama untuk ketiga parameter.

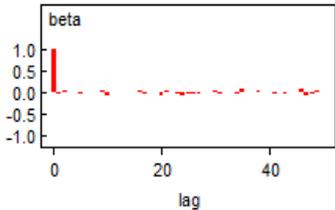
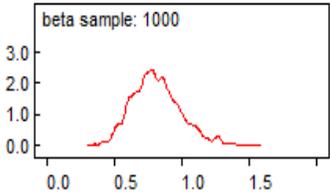
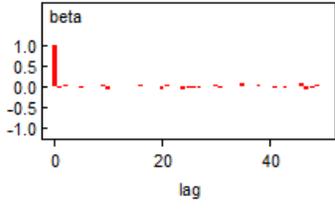
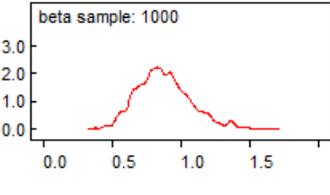
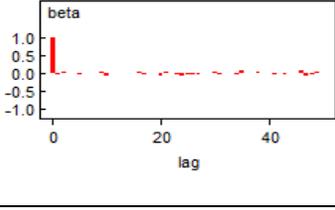
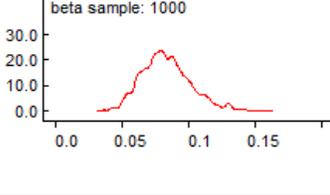
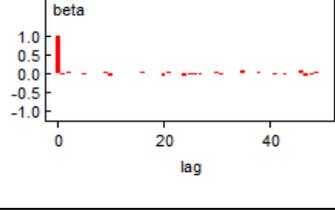
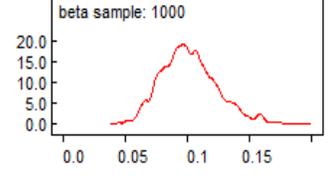
Tabel 3 – Plot History

| Parameter awal (<i>prior</i>) | | | Plot History |
|---------------------------------|------------|-----------|--|
| Ekspensial | Weibull | | |
| β_0 | α_0 | β_0 | |
| 1 | 2 | 1 |  |
| 0.5 | 2 | 0.5 |  |
| 1 | 3 | 1 |  |
| 0.5 | 3 | 0.5 |  |

Hasil estimasi parameter untuk beberapa jumlah sampel populasi lain yang digunakan (1000, 5000, 8000) memberikan hasil yang relatif sama, sehingga

Tabel 3 dan Tabel 4 masing-masing hanya menunjukkan bentuk *plot* untuk sampel parameter dengan jumlah 1000.

Tabel 4 – Plot *Autocorrelation* dan *Kernel Density*

| Parameter awal (<i>prior</i>) | | | <i>Plot Autocorrelation</i> | <i>Kernel density</i> |
|---------------------------------|------------|-----------|---|---|
| Ekspensial | Weibull | | | |
| β_0 | α_0 | β_0 | | |
| 1 | 2 | 1 |  |  |
| 0.5 | 2 | 0.5 |  |  |
| 1 | 3 | 1 |  |  |
| 0.5 | 3 | 0.5 |  |  |

Setelah diperoleh hasil estimasi dengan menggunakan perangkat lunak (*software*) yang diterapkan pada beberapa sampel data simulasi, dilakukan perhitungan secara manual menggunakan persamaan rata-rata secara analitik. Perhitungan ini dilakukan dengan menerapkan masing-masing tiga jenis data simulasi sebagai variabel acak bersamaan dengan parameter awal yang diketahui. Adapun perbandingan hasil perhitungan analitik dengan hasil perhitungan *software* tertera pada Tabel 5.

Tabel 5 menunjukkan hasil estimasi parameter yang diperoleh dari perhitungan analitik dan perhitungan *software* dengan jumlah sampel

parameter yang beragam. Selain itu, hasil perhitungan juga ditentukan berdasarkan kombinasi dari empat parameter awal yang telah ditentukan sebelumnya.

Hasil yang diperoleh menunjukkan perbedaan estimasi parameter untuk hasil analitik dan *software* yang cukup signifikan. Perbedaan ini dapat disebabkan oleh beberapa faktor, seperti standar *error* proses integrasi MCMC dan perhitungan distribusi *posterior* yang diproporsionalkan menjadi lebih sederhana. Dalam perhitungan analitik, nilai yang diproporsionalkan hanya pada persamaan *posterior* dengan menghilangkan fungsi marginalnya.

Lain halnya dengan *software*, tindakan memproporsionalkan fungsi diduga lebih banyak

cakupannya, khususnya untuk nilai-nilai parameter yang diasumsikan sebagai konstanta.

Tabel 5 – Perbandingan Hasil Analitik dan Hasil *Software* untuk Estimasi Parameter

| No | Parameter awal (<i>prior</i>) | | | Hasil analitik | Hasil <i>software</i> | | |
|----|---------------------------------|------------|-----------|-------------------------|-----------------------|--------|--------|
| | Ekspensial | Weibull | | | 1000 | 5000 | 8000 |
| | β_0 | α_0 | β_0 | | | | |
| 1 | 1 | 2 | 1 | 3.3064×10^{-7} | 0.8114 | 0.8072 | 0.8048 |
| 2 | 0.5 | 2 | 0.5 | 7.5601×10^{-6} | 0.8739 | 0.8693 | 0.8668 |
| 3 | 1 | 3 | 1 | 8.8454×10^{-7} | 0.0835 | 0.0830 | 0.0828 |
| 4 | 0.5 | 3 | 0.5 | 2.8405×10^{-4} | 0.1017 | 0.1012 | 0.1009 |

Pada dasarnya, dalam estimasi parameter suatu distribusi, dapat dikatakan hasil analitik lebih akurat dibanding dengan hasil perangkat lunak (*software*). Hal ini dikarenakan proses estimasi dalam *software* WinBUGS 1.4 telah baku dan sistematis, sehingga tidak selalu dapat memenuhi asumsi yang ditetapkan peneliti. Sebagai contoh dalam penelitian ini, dilakukan estimasi pada persamaan *mixture* yang mengandung tiga parameter, yaitu α , β , dan pembobot w . Namun, hanya parameter β yang diestimasi, sedangkan parameter α dan w diasumsikan telah diketahui. Hal ini mengakibatkan terjadinya perbedaan dalam fungsi *posterior*-nya.

Menurut Ntzoufras (2009), dalam *software* WinBUGS 1.4, setiap parameter diasumsikan memiliki distribusi *prior*, dan dilakukan estimasi terhadap seluruh parameter tersebut. Namun dalam perhitungan analitiknya, mengestimasi ketiga parameter secara sekaligus akan menjadi sangat sulit, khususnya dalam penyelesaian integrasinya. Oleh karena itu, penelitian ini hanya dilakukan dengan mengestimasi salah satu parameter, dengan asumsi parameter yang lain diketahui.

Selain itu, dalam proses integrasi MCMC, akan selalu terdapat *error* pada nilai yang dihasilkan. Meskipun demikian, adanya *software* akan mempermudah peneliti dalam perhitungan estimasi parameter untuk fungsi-fungsi yang rumit. Dengan demikian, apabila perhitungan estimasi yang akan dianalisis sesuai dengan teori yang ditetapkan pada *software*, maka akan diperoleh nilai estimator yang relatif akurat.

10. Kesimpulan

Hasil penerapan analitik dan *software* pada data simulasi memiliki perbedaan nilai yang cukup signifikan. Dapat disimpulkan dari penelitian ini bahwa perhitungan analitik lebih akurat dibandingkan

perhitungan *software* apabila dipandang dari segi kesesuaian teori serta proses integrasinya.

DAFTAR PUSTAKA

- McLachlan, G.J., dan Basford, K., 1988, *Mixture Models: Inference and Application to Clustering*, New York, Marcel and Decker Inc.
- Ieren, T.G. dan Oguntunde, P.E., 2018, A Comparison between Maximum Likelihood and Bayesian Estimation Methods for a Shape Parameter of the Weibull-Exponential Distribution, *Asian Journal of Probability and Statistics*, 1(1): 1-12, 2018.
- Walpole, R.E., dan Myers, R.H., 1995, *Ilmu Peluang dan Statistika untuk Insinyur dan Ilmuwan, diterjemahkan oleh RK Sembiring*, Bandung, Institut Teknologi Bandung.
- Hazhiah, I.T., Sugito, S., dan Rahmawati, R., 2012, Estimasi Parameter Weibull Dua Parameter dengan Metode Bayes, *Jurnal Gaussian*, 1(1), 103-102.
- Bonald, T., dan Feuillet, M., 2011, *Network Performance Analysis*, United States, Wiley.
- Iriawan, N., 2001, Penaksiran Model Mixture Normal Univariabel: Suatu Pendekatan Metode Bayesian dengan MCMC, *Prosiding Seminar Nasional dan Konferda VII Matematika Wilayah DIY dan Jawa Tengah, Yogyakarta*.
- Herrhyanto, N., 2011, *Statistika Matematis Lanjutan*, Bandung, CV Pustaka Setia.
- Gamerman, D., 1995, *Markov Chain Monte Carlo*, London, Chapman & Nall.
- Paramita, Z. H., dan Iriawan, N., 2014, Analisis Reliabilitas Transformator (Trafo) di PT PLN APJ Surabaya Barat dengan Pendekatan Bayesian Mixture, *Jurnal Sains dan Seni Pomits*, 3(2), 2337-3520.
- Ntzoufras, I., 2009, *Bayesian Modelling using WinBUGS*, USA, Wiley.
- Iriawan, N., 2011, Pemodelan Mixture of Mixture dalam Pemilihan Portofolio, *Prosiding Seminar Nasional Statistika Universitas Diponegoro, Semarang*.

- Haryono, 1995, *Proses Stokastik Terapan*, Surabaya, Institut Teknologi Sepuluh Nopember.
- Walsh, B., 2004, Markov Chain Monte Carlo and Gibbs Sampling, *Lecture Note for EEB 581, version (26) April 2004*.
- Casella, G., dan George, E. I., 1992, Explaining the Gibbs Sampler, *The American Statistician*, 46(3), 167-174.
- Iriawan, N., 2000, Computationally Intensive Approaches to Inference in Neo-Normal Linier Models, *Doctoral Thesis*, Australia, CUT.
- Rasmussen, C.E., 2000, *The Infinite Gaussian Mixture Model*, Denmark, MIT Press.
- Panik, M. J., 2005, *Advanced Statistics from An Elementary Point of View*, United States, Elsevier Academic Press.
- Bolstad, W. M., 2007, *Introduction to Bayesian Statistics 2nd Edition*, Amerika, John Wiley & Sons. Inc.