Jurnal Matematika UNAND
Vol. **IX** No. **2** Hal. 84 – 91
Edisi April 2020
ISSN : 2303–291X
e-ISSN : 2721–9410

PENDUGAAN PARAMETER MIU DARI DISTRIBUSI LOG-NORMAL DENGAN MENGGUNAKAN METODE *MAXIMUM LIKELIHOOD ESTIMATION* (MLE) DAN METODE BAYES

INDAH PRATIWI, FERRA YANUAR,* HAZMIRA YOZZA

Program Studi S1 Matematika,
Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Andalas,
Kampus UNAND Limau Manis Padang, Indonesia.
email : indahpratiwi@gmail.com, ferrayanuar@sci.unand.ac.id, hazmirayoza@sci.unand.ac.id

Diterima 17 Februari 2020 Direvisi 7 Maret 2020 Dipublikasikan 29 April 2020

Abstrak. Penelitian ini membahas tentang pendugaan parameter μ dari distribusi Log-Normal dengan σ^2 diketahui. Penelitian ini menggunakan metode *Maximum Likelihood Estimation* (MLE) dan metode Bayes dengan prior konjugat. Penduga parameter μ dengan metode MLE dinyatakan sebagai $\widehat{\mu}_{MLE} = \frac{\Sigma_{i=1}^{n} ln(X_i)}{n}$ dan penduga parameter μ dengan metode Bayes dinyatakan sebagai $\widehat{\mu}_B = \frac{m\sigma^2 + n\overline{x}^* p}{\sigma^2 + np}$. Pada penelitian ini kriteria evaluasi penduga yang digunakan adalah MSE dan sifat tak bias. Berdasarkan studi analitik dan studi kasus diperoleh bahwa pendugaan μ dari distribusi Log-Normal dengan metode Bayes lebih baik di bandingkan metode *Maximum Likelihood Estimation* (MLE).

Kata Kunci: Metode Bayes, Metode *Maximum Likelihood Estimation* Distribusi Log-Normal, Prior Konjugat

1. Pendahuluan

Statistika adalah metode-metode yang digunakan dalam pengumpulan, penyajian, analisis, dan penafsiran data. Terdapat dua bagian utama dari statistika, yaitu statistika deskriptif dan statistika inferensia. Statistika deskriptif bertujuan untuk menyajikan informasi data sebagai deskripsi fakta dalam bentuk numerik, tabel, grafik atau kurva distribusi. Statistika inferensia menggunakan konsep probabilitas untuk membuat perkiraan ataupun generalisasi dari suatu objek berdasarkan informasi data yang diambil fakta sebagai populasi atau sampel.

Statistik inferensia dapat dibedakan menjadi dua yaitu penduga parameter dan uji hipotesis. Pendugaan parameter dilakukan untuk menduga nilai parameter suatu populasi yang menjadi perhatian peneliti. Terdapat dua cara yang dilakukan untuk

*penulis korespondensi

menduga parameter, yaitu dengan melakukan penduga titik dan penduga selang. Penduga titik merupakan suatu nilai yang diperoleh dari suatu data contoh yang digunakan untuk memperkirakan suatu parameter populasi, sedangkan penduga selang merupakan suatu selang kepercayaan dari suatu parameter yang tergantung pada penduga titik suatu data contoh.

Penduga titik dapat dilakukan dengan beberapa cara. Dua di antaranya yaitu metode klasik dan metode Bayes. Salah satu teknik yang digunakan dalam metode klasik adalah metode *Maximum Likelihood Estimation* (MLE). Pada metode Bayes parameter dipandang sebagai variabel acak. Pengetahuan awal tentang distribusi parameter sebelum pengamatan dilakukan dinyatakan dalam suatu distribusi yang disebut dengan distribusi prior. Pemilihan distribusi prior ini adalah didasarkan pada subjektifitas peneliti yang dapat dipilih berdasarkan pendapat ahli atau model terdahulu. Dengan memasukkan distribusi prior ini menyebabkan pendugaan titik dengan metode Bayes akan menghasilkan nilai dugaan yang lebih menghampiri nilai yang sebenarnya. Dengan kata lain metode bayes menghasilkan nilai dugaan yang lebih baik dari metode klasik (*Maximum Likelihood Estimation*) [3]. Distribusi prior kemudian digabungkan dengan distribusi sampel maka menghasilkan distribusi baru yang disebut distribusi posterior yang selanjutnya akan menjadi dasar untuk inferensi Bayesian.

Distribusi Log-Normal merupakan sebuah distribusi dari peubah acak kontinu dengan parameter meannya μ dan ragam σ^2. Dalam penelitian ini dilakukan pengkajian mengenai pendugaan parameter μ dengan σ^2 diketahui dari distribusi Log-Normal dengan metode *Maximum Likelihood Estimation* dan metode Bayes. Dalam metode Bayes distribusi prior yang digunakan adalah prior konjugat. Hasil pendugaan parameter μ dari distribusi Log-Normal dengan metode MLE dan metode Bayes ini akan dibandingkan dengan menggunakan studi simulasi. Perbandingannya dilakukan berdasarkan ketakbiasan, kekonsistenan dan keefisienan penduga pada kedua metode.

2. Landasan Teori

2.1. *Teorema Bayes*

Teorema 2.1. ***(Teorema Bayes)*** [10] *Jika kejadian $B_1, \cdots B_k$ merupakan suatu partisi ruang sampel S dengan $P(B_j) \neq 0$ untuk $j = 1, 2, \cdots, k$, maka untuk setiap kejadian A dalam S sedemikian sehingga $P(A) \neq 0$ berlaku*

$$P(B_j \mid A) = \frac{P(A \mid B_j)P(B_j)}{\sum_{i=1}^{k} P(A \mid B_i)P(B_i)},$$

untuk $j = 1, 2, \cdots, k$.

2.2. *Distribusi Log-Normal*

Distribusi Log-Normal dalam teori peluang adalah distribusi peluang kontinu dari variabel acak yang logaritmanya terdistribusi normal. Bain menuliskan dalam bukunya bahwa suatu peubah acak X bernilai real positif $(0 < x < \infty)$, sedemikian sehingga $Y = \ln(X)$ merupakan peubah acak berdistribusi normal dengan nilai harapan μ dan variansi σ^2, maka $X = e^Y$ dikatakan memenuhi distribusi Log- Normal. Karena X dan Y dihubungkan oleh relasi $Y = \ln X$, maka fungsi kepekatan peluang dari X yang memiliki distribusi Log-Normal adalah [1]:

$$f(x) = \begin{cases} \dfrac{1}{\sqrt{2\Pi x^2\sigma^2}} \exp\left(-\dfrac{1}{2}\left(\dfrac{\ln(x)-\mu}{\sigma}\right)^2\right), 0 < x < \infty, -\infty < \mu < \infty, \sigma^2 > 0 \\ 0 \qquad\qquad\qquad\qquad\qquad\qquad lainnya \end{cases}$$

2.3. *Metode* Maximum Likelihood Estimation

Definisi 2.2. [1] *Fungsi kepekatan peluang bersama peubah acak $X_1, X_2, \cdots, X_n$ yang dihitung pada $x_1, x_2, \cdots, x_n$ adalah $f(x_1, x_2, \cdots, x_n|\mu)$, dan ini dirujuk oleh fungsi likelihood untuk $x_1, x_2, \cdots, x_n$ tetap,fungsi likelihood adalah fungsi dari μ yang dinotasikan dengan $L(\mu)$. Jika $X_1, X_2, \cdots, X_n$ merupakan contoh acak dari $f(x;\mu)$, maka:*

$$L(\mu) = \prod_{i=1}^{n} f(x_i;\mu). \tag{2.1}$$

Langkah-langkah untuk menentukan penduga kemungkinan maksimum dari μ adalah

(1) Menentukan fungsi *likelihood* $L(\mu) = f(x_1, x_2, \cdots, x_n;\mu)$, dengan $X_1, X_2, \cdots, X_n$ saling bebas.
(2) Bentuk logaritma natural dari fungsi *likelihood*, $\ln L(\mu)$.
(3) Menentukan turunan dari logaritma natural fungsi *likelihood* terhadap μ dan menyamakannya dengan 0.

$$\frac{\partial \ln(\mu)}{\partial \mu} = 0.$$

Penyelesaian dari persamaan pada langkah 3, yakni $\widehat{\mu}_{MLE}$ merupakan calon penduga kemungkinan maksimum untuk μ
(4) Tentukan turunan kedua dari logaritma natural fungsi *likelihood* terhadap μ. Jika $\dfrac{\partial^2 \ln L(\mu)}{\partial^2 \mu} < 0$, maka akan membuktikan bahwa $\widehat{\mu}_{MLE}$ benar-benar memaksimumkan fungsi *likelihood* $L(\mu)$, sehingga $\widehat{\mu}_{MLE}$ merupakan penduga kemungkinan maksimum bagi μ.

2.4. *Metode Estimasi Bayes*

2.4.1. *Distribusi Prior*

Terdapat beberapa tipe distribusi *prior*, berkaitan dengan bentuk distribusi hasil identifikasi pola datanya distribusi *prior* dibagi menjadi dua bagian yaitu distribusi Prior konjugat dan distribusi prior non konjugat. Berkaitan dengan informasi yang terkait dengan penentuan masing-masing parameter, distribusi *prior* dibagi menjadi dua bagian yaitu distribusi prior informatif dan distribusi prior non informatif.

2.4.2. *Distribusi Posterior*

Fungsi kepekatan peluang posterior dari μ jika diketahui sampel pengamatan $x_1, x_2, \cdots, x_n$ adalah [1]:

$$f(x|\mu) = \frac{f(x|\mu)f(\mu)}{\int f(x|\mu)f(\mu)d\mu} \propto f(x|\mu)f(\mu). \tag{2.2}$$

Definisi 2.3. [11] ***Penduga Bayes***. *Nilai harapan dari distribusi posterior $f(\mu|x_1, x_2, \cdots, x_n)$ dinyatakan dengan T, disebut penduga Bayes untuk $\tau(\mu)$.*

3. Data dan Hasil

3.1. *Penduga Parameter μ (σ^2 diketahui) dari Distribusi Log-Normal dengan* Metode Maximum Likelihood Estimation *(MLE)*

Misalkan $X_1, X_2, \cdots, X_n$ adalah contoh acak yang memiliki distribusi Log- Normal (μ, σ^2) dengan σ^2 diketahui. Fungsi *likelihood* bagi parameter tersebut adalah

$$\begin{aligned} L(\mu) &= \prod_{i=1}^{n} f(x_i|\mu, \sigma^2) \\ &= \prod_{i=1}^{n} \left((2\pi\sigma^2)^{-1/2} x_i^{-1} \exp\left[\frac{-(\ln(x_i) - \mu)^2}{2\sigma^2} \right] \right) \\ &= (2\pi\sigma^2)^{-n/2} \left(\prod_{i=1}^{n} x_i^{-1} \right) \exp\left[\sum_{i=1}^{n} \frac{-(\ln(x_i - \mu)^2}{2\sigma^2} \right] \end{aligned} \tag{3.1}$$

Logaritma natural dari fungsi *likelihood* adalah

$$\ln L(\mu) = -\frac{n}{2} \ln(2\pi\sigma^2) - \sum_{i=1}^{n} \ln(x_i) - \frac{\sum_{i=1}^{n} \ln(x_i)^2}{2\sigma^2} + \frac{\sum_{i=1}^{n} \mu \ln(x_i)}{\sigma^2} - \frac{n\mu^2}{2\sigma^2}.$$

Dengan mendiferensialkan terhadap μ dan selanjutnya menyamakan dengan 0, maka diperoleh

$$\begin{aligned} \frac{\partial \ln L(\mu)}{\partial \mu} &= 0 \\ \frac{\sum_{i=1}^{n} \ln(x_i)}{\sigma^2} - \frac{2n\mu}{2\sigma^2} &= 0 \\ \mu &= \frac{\sum_{i=1}^{n} \ln(x_i)}{n} \end{aligned}$$

Selanjutnya akan dilakukan uji turunan kedua untuk menunjukkan bahwa $\widehat{\mu}$ benar-benar memaksimumkan fungsi *likelihood* L(μ)

$$\begin{aligned} \frac{\partial^2 \ln L(\mu, \sigma^2|\mathbf{x})}{\partial \mu^2} &= \frac{\partial}{\partial \mu} \left[\frac{\sum_{i=1}^{n} \ln(x_i)}{\sigma^2} - \frac{n\mu}{\sigma^2} \right] \\ &= -\frac{n}{\sigma^2} < 0 \end{aligned}$$

Diperoleh turunan kedua dari $L(\mu)$ terhadat μ bernilai kurang dari nol. Hal ini berarti $\widehat{\mu}$ memaksimumkan fungsi kemungkinan L(μ), maka penduga untuk parameter μ menggunakan metode *Maximum Likelihood Estimation* (MLE) adalah

$$\widehat{\mu}_{MLE} = \frac{\sum_{i=1}^{n} \ln(X_i)}{n}.$$

3.2. *Pendugaan Parameter μ (σ^2 diketahui) dari Distribusi Log-Normal dengan Metode Bayes*

Diketahui bahwa $X_1, X_2, \cdots, X_n$ adalah contoh acak yang memiliki distribusi Log-Normal dengan σ^2 diketahui. Berdasarkan Persamaan (3.1) fungsi *likelihood*nya adalah

$$L(\mu) \propto \exp\left[-\frac{n}{2\sigma^2}\left(\frac{\Sigma \ln x_i}{n} - \mu\right)^2\right]. \tag{3.2}$$

Prior konjugat untuk distribusi Log-Normal dengan parameter μ dan σ^2 diketahui adalah distribusi Normal dengan nilai harapan(m) dan variansi(p), sehingga fungsi kepekatan peluang untuk μ adalah

$$f(\mu) \propto \exp\left[-\frac{1}{2p}(\mu - m)^2\right]. \tag{3.3}$$

Dalam teorema Bayes setelah data diambil dan prior telah ditentukan,maka kemudian dicari distribusi posteriornya dengan menggunakan Persamaan (3.2) dan Persamaan (3.3), yaitu

$$\begin{aligned} f(\mu|\mathbf{x}) &\propto f(\mu)L(\mu) \\ &\propto \exp\left[-\frac{1}{2p}(\mu - m)^2\right] . \exp\left[-\frac{n}{2\sigma^2}\left(\frac{\Sigma \ln x_i}{n} - \mu\right)^2\right]. \end{aligned}$$

Bila dinyatakan $\overline{x} = \dfrac{\Sigma \ln X_i}{n}$ maka

$$\begin{aligned} f(\mu|\mathbf{x}) &\propto \exp\left[-\frac{1}{2p}(\mu - m)^2\right] . \exp\left[-\frac{n}{2\sigma^2}(\overline{x} - \mu)^2\right] \\ &\propto \exp\left\{-\frac{1}{2}\left[\mu^2\left(\frac{1}{p} + \frac{n}{\sigma^2}\right) - 2\mu\left(\frac{m}{p} + \frac{n\overline{x}}{\sigma^2}\right) + \frac{m^2}{p} + \frac{\overline{x}^2 n}{\sigma^2}\right]\right\}. \end{aligned} \tag{3.4}$$

Distribusi posterior untuk persamaan (3.4) adalah berdistribusi Normal dengan pembuktian sebagai berikut. Misalkan peubah acak μ berdistribusi Normal dengan parameter μ_n dan σ_n^2, ditulis $\mu \sim N(\mu_n, \sigma_n^2)$. Fungsi kepekatan peluang untuk μ adalah

$$f(\mu) \propto \exp\left(-\frac{1}{2\sigma_n^2}(\mu^2 - 2\mu\mu_n + \mu_n^2)\right). \tag{3.5}$$

Dengan demikian, berdasarkan Persamaan (3.4) dan Persamaan (3.5) maka

$$\begin{aligned} -\frac{1}{2\sigma^2}\mu^2 &= -\frac{\mu^2}{2}\left(\frac{1}{p} + \frac{n}{\sigma^2}\right), \\ \sigma_n^2 &= \frac{p\sigma^2}{\sigma^2 + pn}. \end{aligned} \tag{3.6}$$

dan

$$\frac{\mu\mu_n}{\sigma_n^2} = \mu\left(\frac{m}{p} + \frac{n\overline{x}}{\sigma^2}\right),$$

$$\mu_n = \frac{m\sigma^2 + pn\overline{x}}{\sigma^2 + pn}. \quad (3.7)$$

Dengan demikian distribusi posterior untuk μ pada persamaan (4.2.5) terbukti berdistribusi Normal dengan parameter mean nya adalah $\frac{m\sigma^2 + pn\overline{x}}{\sigma^2 + pn}$ dan parameter variansi nya adalah $\frac{p\sigma^2}{\sigma^2 + pn}$.

Maka penduga nilai harapan Bayes dari μ adalah

$$\widehat{\mu}_B = \frac{m\sigma^2 + pn\overline{X}}{\sigma^2 + pn}. \quad (3.8)$$

3.3. *Evaluasi Sifat Penduga*

3.3.1. *Sifat Tak Bias*

Sifat tak bias dari nilai dugaan μ dengan menggunakan metode *Maximum Likelihood Estimation* adalah

$$E\left(\widehat{\mu}_{MLE}\right) = \mu.$$

Karena E($\widehat{\mu}_{MLE}$) = μ maka $\widehat{\mu}_{MLE}$ merupakan penduga tak bias bagi μ.

Sifat tak bias dari nilai dugaan μ dengan menggunakan metode Bayes adalah

$$E\left(\widehat{\mu}_B\right) = \frac{m\sigma^2 + n\overline{x}^* p}{\sigma^2 + np}.$$

Karena $E(\widehat{\mu}_B) \neq \mu$, maka $\widehat{\mu}_B$ merupakan penduga bias bagi μ. Akan ditentukan apakah $\widehat{\mu}_B$ merupakan penduga tak bias secara asimtotik bagi μ dengan cara berikut.

$$\begin{aligned} Lim_{n\to\infty} E(\widehat{\mu}_B) &= Lim_{n\to\infty}\left(\frac{m\sigma^2 + n\overline{x}^* p}{\sigma^2 + np}\right) \\ &= \mu. \end{aligned}$$

Diperoleh $Lim_{n\to\infty}$ E($\widehat{\mu}_B$) = μ, maka terbukti bahwa $\widehat{\mu}_B$ merupakan penduga tak bias secara asimtotik bagi μ.

3.3.2. *Mean Square Error (MSE)*

Mean Square Error (MSE) dari Nilai Duga μ dengan Menggunakan Metode Maximum Likelihood Estimation (MLE) adalah

$$MSE\left(\widehat{\mu}_{MLE}\right) = \frac{\sigma^2}{n}.$$

Mean Square Error (MSE) dari Nilai Duga μ dengan Menggunakan Metode Bayes adalah

$$Var(\widehat{\mu}_B) = \left(\frac{np}{np + \sigma^2}\right)^2 \frac{\sigma^2}{n}.$$

3.4. *Studi Kasus*

Pada bagian ini akan digunakan data bangkitan sebanyak 30,50, dan 100 buah dari distribusi Log-Normal dengan nilai harapan adalah 5 dan standar deviasi yang telah ditetapkan yaitu 1, sehingga di peroleh hasil nilai seperti berikut:

Tabel 1. Hasil perhitungan penduga μ dari Data Kasus

Kriteria	Banyak Data	Metode MLE	Metode Bayes
Mean	30	4.808156	4.808077
	50	4.958331	4.958316
	100	5.042803	5.042821
Variansi	30	0.033333	0.032491
	50	0.02	0.017832
	100	0.01	0.009345
MSE	30	0.033333	0.032491
	50	0.02	0.017832
	100	0.01	0.009345

Dari Tabel diketahui bahwa nilai dugaan mean dengan metode *Maximum Likelohood Estimation* (MLE) dan metode Bayes sama-sama menghasilkan nilai yang dekat dengan nilai yang ditetapkan, yaitu $\mu = 5$. Tetapi nilai dugaan dengan metode MLE menghasilkan nilai yang lebih dekat dibanding hasil metode Bayes, karena MLE adalah penduga tak bias sedangkan Bayes adalah penduga bias. Pada kedua metode, terlihat bahwa semakin besar ukuran contoh, nilai MSE yang dihasilkan semakin kecil dan mendekati 0. Metode Bayes menghasilkan nilai MSE yang lebih kecil dibandingkan dengan metode MLE untuk ketiga ukuran data contoh.

4. Kesimpulan

Berdasarkan pembahasan yang telah dilakukan, diperoleh beberapa kesimpulan sebagai berikut:

(1) Penduga parameter μ dengan metode *Maximun Likelihood Estimation* (MLE) untuk distribusi Log-Normal (μ, σ^2) jika dinyatakan sebagai $\widehat{\mu}$ dapat dirumuskan sebagai

$$\widehat{\mu}_{MLE} = \frac{\sum_{i=1}^{n} \ln(X_i)}{n}.$$

Sedangkan Penduga parameter μ dengan metode Bayes untuk distribusi Log-Normal (μ, σ^2) jika dinyatakan sebagai $\widehat{\mu}_B$ dapat dirumuskan sebagai

$$\widehat{\mu}_B = \frac{m\sigma^2 + n\overline{x}^* p}{\sigma^2 + np}$$

(2) Secara teoritis, pendugaan parameter dengan metode MLE adalah penduga tak bias untuk parameter μ dari distribusi Log-Normal dan metode Bayes adalah penduga bias bagi parameter μ dari distribusi Log-Normal. Namun penduga Bayes adalah penduga tak bias asimtotik bagi parameter μ distribusi Log-Normal. Karena masing-masing penduga adalah penduga tak bias dan penduga

bias, sehingga tidak bisa dibandingkan keefisienan dari penduga kedua metode, karena keefisienan penduga berlaku untuk penduga yang tak bias.

(3) Berdasarkan hasil simulasi data yang diperoleh, nilai *Mean Square Error* (MSE) yang didapatkan terlihat bahwa kedua metode memiliki pola yang sama, yaitu semakin besar ukuran contoh, nilai MSE yang dihasilkan semakin kecil dan mendekati 0. Karena nilai dugaan yang didapatkan mendekati nilai parameter μ ketika ukuran contoh bertambah besar dan nilai MSE yang dihasilkan mendekati 0 ketika ukuran contoh bertambah besar, maka penduga pada kedua metode merupakan penduga yang konsisten. Metode Bayes menghasilkan nilai MSE yang lebih kecil dibandingkan dengan metode MLE. Sehingga pendugaan parameter μ dari distribusi Log-Normal dengan metode Bayes lebih konsisten dibandingkan dengan metode MLE.

Daftar Pustaka

[1] Bain, L.J. and Engelhardt, M. 1992. *Introduction to Probability and Mathematical Statistic*. Second Edition. Duxbury Press, California.

[2] Berger, J.O. 1993. *Statistical Decision Theory And Bayesian Analysis*. Second Edition.Sringer-Verlug, New York.

[3] Bolstad, W.M. 2007. *Introduction to Bayesian Statistics*. Second Edition. America: A John Wiley And Sons. Inc.

[4] Box, G.E.P. and Tiao, G.C. 1973.*Bayesian Inference In Statistical Analysis*. Addison-Wesley Publishing Company,Inc. Phillipines.

[5] Casella, G dan R.L. Berger. 2001. *Statistical Inference*. Second Edition. California: Pacific Grove.

[6] Hogg, R.V. dan Craig, A.T. 1905.*Introduction to Mathematical Statistic*. Seventh Edition. United States of America: Pearson Education, Inc.

[7] Koch, K. R. 2007. *Introduction to Bayesian Statistics*. Second Edition. New York: Springer.

[8] Mukhopadhyay, N. 2000. *Probability and Statistical Inference*. New York: Marcel Dekker.

[9] Mustafid. 2003. *Statistika Elementer*. Semarang: Universitas Diponegoro.

[10] Walpole,R.E. 1993. *Pengantar Statistik*. Edisi-3. Jakarta: PT Gramedia Pustaka Utama.

[11] Walpole,R.E., Myers, R.H., Myers,S.L and Ye,K. 2002. *Probability and Statistics for Engineers and Scientist*. Ninth Edition. New York: Pearson Education,Inc.

[12] Yani, R.N., Yanuar, F., dan Yozza, H. 2018. Inferensi Bayesian Untuk σ^2 dari Distribusi Normal dengan Berbagai Distribusi Prior. Jurnal Matematika UNAND VII (2) : 132 – 139.

[13] Yanuar, F., Yozza,H. and Rescha, R.V.2019. Comparison of Two Priors in Bayesian Estimation for Parameters of Weibull Distribusion. *Science and Technology Indonesia*.