Pendekatan Metode Bayesian untuk Kajian Estimasi Parameter Distribusi Log-Normal untuk Non-Informatif Prior

Evi Noor Diana dan Soehardjoepri
Jurusan Matematika, Fakultas MIPA, Institut Teknologi Sepuluh Nopember (ITS)
Jl. Arief Rahman Hakim, Surabaya 60111 Indonesia
e-mail: djoepri.its@gmail.com

***Abstrak*— Distribusi Log-Normal merupakan salah satu dari beberapa distribusi kontinu. Distribusi Log-Normal tersebut memiliki parameter yang harus di estimasi yaitu parameter mean θ dan varians σ^2. Penelitian dilakukan dengan metode Bayes yaitu penggabungan distribusi sampel dan distribusi prior, sehingga diperoleh distribusi posterior. Distribusi prior yang digunakan adalah non-informatif prior. Teknik penentuan dari non-informatif prior menggunakan metode Jeffrey's dari distribusi Log-Normal univariat. Setelah mendapatkan distribusi posterior, menemukan distribusi marginal dari mean θ dan varians σ^2. Setelah mendapatkan distribusi marginal, langkah selanjutnya adalah mendapatkan nilai harapan (ekspektasi) dari distribusi marginal. Sehingga akan mendapatkan estimasi titik untuk mean θ dan varians σ^2.**

***Kata Kunci*— Distribusi Log-Normal , Fungsi Likelihood, Non-Informatif Prior, Metode Bayes, Distribusi Posterior.**

I. PENDAHULUAN

Estimasi adalah suatu metode dimana kita dapat memperkirakan nilai dari suatu populasi dengan menggunakan nilai dari sampel. Jenis estimasi ada dua yaitu estimasi titik dan estimasi interval. Estimasi titik adalah nilai yang berfungsi untuk suatu pendugaan dari parameter populasi. Estimasi Interval adalah interval yang menyatakan keberadaan dari suatu parameter populasi [1]. Contoh sederhana, apabila kita ingin menyebrang jalan, kita menaksir berapa kecepatan kendaraan yang melewati jalan pada saat itu. Sehingga kita dapat memutuskan akan menyebrang atau tidak. Untuk mengestimasi parameter suatu populasi maka diambil sebuah sampel yang representatif, sebelum estimasi dilakukan, perlu diketahui lebih dulu keadaan populasi variabel acak tersebut secara apriori, seperti bentuk distribusinya, dan karakteristik parameter-parameter lain. Walaupun kerapkali informasi tentang populasinya sangat minimal, informasi yang diperoleh secara apriori itu kemudian dapat ditambahkan pula dengan informasi yang diperoleh dari sampel itu sendiri

Metode Bayes memandang parameter sebagai variabel yang menggambarkan pengetahuan awal tentang parameter sebelum pengamatan dilakukan dan dinyatakan dalam suatu distribusi yang disebut dengan distribusi prior. Pemilihan prior secara umum dilakukan berdasarkan diketahui atau tidaknya informasi mengenai parameter. Jika informasi mengenai parameter diketahui, maka prior informatif, yaitu prior yang memengaruhi hasil distribusi posterior dan bersifat sangat subjektif, sedangkan jika informasi mengenai parameter tidak tersedia, maka digunakan prior non-informatif yang tidak memberikan pengaruh yang signifikan terhadap distribusi posterior, sehingga informasi yang diperoleh dari data amatan bersifat lebih objektif. Setelah pengamatan dilakukan, informasi dalam distribusi prior dikombinasikan dengan informasi dengan data sampel melalui teorema Bayes, dan hasilnya dinyatakan dalam bentuk distribusi yang disebut distribusi posterior [2].

Sebuah peubah dapat dimodelkan sebagai log-normal bila ia dapat dipandang sebagai hasil kali dari banyak peubah acak bebas (independen) masing-masing bernilai positif. Sebagai contoh, dalam keuangan, peubah ini dapat mewakili pendapatan majemuk dari sederetan perdagangan atau faktor diskonto jangka panjang dapat diturunkan dari hasil kali faktor-faktor diskonto jangka pendek. Dalam komunikasi nirkabel, atenuasi yang disebabkan oleh pembayangan atau pengaburan lambat dari benda acak sering dianggap berdistribusi Log-Normal.

Salah satu contoh penelitian yang membahas estimasi parameter dengan metode Bayes adalah penelitian yang dilakukan oleh Sultan dan Ahmad. Penelitiannya menjelaskan tentang mendapatkan perbandingan nilai estimasi titik metode *Maksimum Likelihood Estimator (MLE)* dengan metode *Bayesian*. Menggunakan distribusi Log-Normal dengan diketahui distribusi prior (*informative prior).* Kemudian dicari fungsi likelihood lalu dikalikan dengan distribusi prior menjadi distribusi posterior. Setelah mendapatkan distribusi posterior, menemukan distribusi marginal dari mean μ dan varian σ^2. Langkah selanjutnya mendapatkan nilai harapan (*ekspektasi)* dari distribusi marginal sehingga mendapatkan estimasi titik untuk mean μ dan varians σ^2 [3].

Berdasarkan penelitian tersebut, penelitian ini akan membahas tentang estimasi parameter distribusi Log-Normal untuk non-informatif prior. Estimasi titik didapatkan dengan menggunakan metode Bayes.

II. TINJAUAN PUSTAKA

A. Distribusi Log-Normal

Misalkan sebuah peubah acak X bilangan real positif $(0 < x < \infty)$. Sedemikian sehingga $Y = \ln x$ merupakan distribusi Normal dengan rata-rata θ dan variansi σ^2 . $X = e^y$ merupakan distribusi Log-Normal atau dapat ditulis dengan $LN(\theta, \sigma^2)$ dan $Y \approx N(\theta, \sigma^2)$. Karena X dan Y dihubungkan oleh relasi $Y = \ln x$, maka fungsi kepekatannya adalah sebagai berikut [1] :

$$f(x) = \begin{cases} \frac{1}{x\,\sigma\sqrt{2\pi}}\, exp\left[\frac{-(\ln x - \theta)^2}{2\sigma^2}\right], & x > 0 \\ 0\ , & untuk\ x\ yang\ lainnya \end{cases} \quad (1)$$

Variabel acak positif X yang berdistribusi Log-Normal mempunyai mean $\exp\left(\theta + \frac{\sigma^2}{2}\right)$ dan varians $\exp(2\theta + \sigma^2)\{\exp(\sigma^2) - 1\}$ [1].

B. Fungsi Likelihood

Fungsi Likelihood adalah misalkan $X_1, X_2, \dots, X_n$ sampel acak dengan fungsi peluang $f(x_i; \theta)$, i = 1, 2, ..., n. Apabila L yaitu fungsi peluang bersama dari $X_1, X_2, \dots, X_n$ dipandang sebagai fungsi dari θ dan $x_1, x_2, \dots, x_n$ sebagai bentangan tertentu maka [4]

$$L(\theta) = \prod_{i=1}^{n} f(x_i; \theta) \quad (2)$$

C. Non-Informatif Prior

Non-informatif berhubungan dengan situasi dimana distribusi prior tidak memiliki basis populasi. Hanya terdapat sedikit informasi prior sehingga distribusi prior berperan minimal dalam distribusi posterior. Salah satu bentuk pendekatan dari prior non-informatif adalah dengan menggunakan metode Jeffrey's. Distribusi prior untuk satu parameter θ adalah sekitar non informatif jika diambil proposional maka akar kuadrat dari informasi Fisher [2].

$$f(\theta) = [I(\theta)]^{\frac{1}{2}} \quad (3)$$

dimana $I(\theta)$ merupakan nilai harapan informasi Fisher

$$I(\theta) = -E_\theta\left[\frac{\partial^2 \log f(x;\theta)}{\partial\theta^2}\right] \quad (4)$$

D. Metode Bayes

Pendekatan Bayes dalam mengestimasi estimator dari parameter menggabungkan informasi dari sampel dengan informasi lain telah tersedia sebelumnya.

Besaran parameter θ dalam sebuah populasi yang memuat variabel acak x dapat digunakan dalam bentuk aturan peluang, yaitu sebagai berikut [5]:

$$f(\theta|x) = \frac{f(x|\theta) f(\theta)}{f(x)} \quad (5)$$

E. Distribusi Posterior

Distribusi posterior adalah penggabungan distribusi sampel dan distribusi prior [2].

$$posterior \sim likelihood\ x\ prior$$

$$f(\theta|\,x) = f(x\,|\,\theta)\,.\,f(\theta) \quad (6)$$

dimana $f(x\,|\,\theta)$ merupakan fungsi likelihood dan $f(\theta)$ merupakan distribusi prior.

III. ANALISIS DAN PEMBAHASAN

A. Fungsi Likelihood

Jika $X_1, X_2, \dots, X_n$ adalah sampel acak berdistribusi Log-Normal dengan densitas $f(x_1;\, \theta, \sigma^2)$ maka fungsi likelihoodnya didefinisikan dengan [5]:

$$L(\theta, \sigma^2) = \prod_{i=1}^{n} f(x_i; \theta, \sigma^2)$$

$$\left(\frac{1}{\sqrt{2\pi\,\sigma^2 x_i^2}}\right)^n exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\ln x_i - \theta)^2\right] \quad (7)$$

B. Non-informatif Prior

Distribusi non-informatif prior $f(\vartheta)$ dimana $\vartheta = (\theta, \sigma^2)$, diasumsikan bahwa θ dan σ^2 adalah independen (bebas) sehingga $f(\vartheta) = f(\theta) f(\sigma^2)$. Menentukan distribusi non-informatif prior $f(\sigma^2)$ adalah sebagai berikut:

$$f(x; \theta, \sigma^2) = \frac{1}{x\sigma\sqrt{2\pi}}\, exp\left[\frac{-(\ln x - \theta)^2}{2\sigma^2}\right]$$

$$\log f(x; \theta, \sigma^2) = -\frac{1}{2}\log 2\pi - \frac{1}{2}\log\sigma^2 - \frac{1}{2}\log x^2 - \left(\frac{(\ln x - \theta)^2}{2\sigma^2}\right)$$

Jika $u = \sigma^2$ maka

$$\log f(x; \theta, u) = -\frac{1}{2}\log 2\pi - \frac{1}{2}\log u - \frac{1}{2}\log x^2 - \left(\frac{(\ln x - \theta)^2}{2u}\right)$$

$$\frac{d\ \log f(x; \theta, u)}{du} = -\frac{1}{2u} + \frac{(\ln x - \theta)^2}{2u^2}$$

$$\frac{d^2\ \log f(x; \theta, u)}{du^2} = \frac{1}{2u^2} - \frac{(\ln x - \theta)^2}{u^3}$$

$$\frac{d^2\ \log f(x; \theta, u)}{du^2} = \frac{1}{2\sigma^4} - \frac{(\ln x - \theta)^2}{\sigma^6}$$

$$I(\sigma^2) = -E\left[\frac{d^2\ \log f(x; \theta, u)}{du^2}\right]$$

$$I(\sigma^2) = \frac{1}{2\sigma^4}$$

$$f(\sigma^2) = \sqrt{I(\sigma^2)} \propto \frac{1}{\sigma^2}$$

Sedangkan nilai non-informatif prior untuk $f(\theta)$= c (konstan) sehingga diperoleh

$$f(\vartheta) = f(\theta) f(\sigma^2) = c\ x\ \frac{1}{\sigma^2} \propto \frac{1}{\sigma^2} \quad (8)$$

Jadi nilai non-informatif prior

$$f(\vartheta) = \boldsymbol{\frac{1}{\sigma^2}} \quad (9)$$

C. Distribusi Posterior

Setelah mencari fungsi likelihood dan menentukan distribusi prior dari distribusi Log-Normal dapat dicari distribusi posteriornya. Kepadatan posterior bersama dari θ dan σ^2 adalah diberikan

$$posterior\ \propto prior * fungsi\ likelihood$$

$$f(\theta, \sigma^2|x) \propto\ f(\vartheta) *\ L(\theta, \sigma^2) \quad (10)$$

Dari persamaan (1) dan (2) Kepadatan posterior bersama dari θ dan σ^2 adalah diberikan

$$f(\theta, \sigma^2|x) \propto \left(\left(\frac{1}{\sqrt{2\pi x_i}}\right)^n \left(\frac{1}{\sigma^2}\right)^{\left(\frac{n}{2}+1\right)}\right) x$$
$$\left(\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\ln x_i - \theta)^2\right]\right)$$
$$f(\theta, \sigma^2|x) \propto \left(\frac{c}{(\sigma^2)^{\left(\frac{n}{2}+1\right)}}\, exp\left[-\frac{\beta}{2\sigma^2}\right]\right) \boldsymbol{x}$$
$$\left(\exp\left[\frac{-n}{2\sigma^2}\left(\theta - \frac{\sum_{i=1}^{n}\ln x_i}{n}\right)^2\right]\right) \quad (11)$$

Dimana $\beta = (n-1)\frac{\sum_{i=1}^{n} \ln {x_i}^2}{n}$ dan c adalah konstanta normal. [2] menjelaskan jika $P(\theta)$ menjadi prior dan $P(x|\theta)$ menjadi fungsi likelihood, fungsi padat peluang (pdf) posterior $P(\theta|x)$ diberikan $P(\theta|x) = c\,P(\theta)P(x|\theta)$, dimana c adalah konstanta normal. Lalu nilai dari c diperoleh dari $c = [\int P(\theta)P(x|\theta)d\theta]^{-1}$ dimana c dapat menjadi

$$c^{-1} = \int_0^\infty \int_{-\infty}^{\infty} f(\theta,\sigma^2|x)\, d\theta\, d\sigma^2$$

$$c^{-1} = \int_0^\infty \int_{-\infty}^{\infty} \left(\frac{1}{(\sigma^2)^{\left(\frac{n}{2}+1\right)}} exp\left[-\frac{\beta}{2\sigma^2}\right]\right) x$$

$$\left(\exp\left[\frac{-n}{2\sigma^2}\left(\theta - \frac{\sum_{i=1}^{n}\ln x_i}{n}\right)^2\right]\right) d\theta\, d\sigma^2 \quad (12)$$

Menggunakan transformasi $t = \frac{\sqrt{n}}{\sigma}\left(\theta - \frac{\sum_{i=1}^{n}\ln x_i}{n}\right)$

$$c^{-1} = \sqrt{\frac{2\pi}{n}} \int_0^\infty \frac{exp\left[-\frac{\beta}{2\sigma^2}\right]}{(\sigma^2)^{\left(\frac{n+1}{2}\right)}}\, d\sigma^2$$

$$c^{-1} = \sqrt{\frac{2\pi}{n}} \left(\frac{\beta}{2}\right)^{-\left(\frac{n-1}{2}\right)} \Gamma\left(\frac{n-1}{2}\right)$$

$$c = \sqrt{\frac{n}{2\pi}} \left(\frac{\beta}{2}\right)^{\left(\frac{n-1}{2}\right)} \frac{1}{\Gamma\left(\frac{n-1}{2}\right)} \quad (13)$$

Persamaan (13) di subtitusikan ke persamaan (11) menjadi

$$f(\theta,\sigma^2|x) = \left(\sqrt{\frac{n}{2\pi}} \frac{(\beta)^{\left(\frac{n-1}{2}\right)}}{\Gamma\left(\frac{n-1}{2}\right)(\sigma^2)^{\left(\frac{n}{2}+1\right)}} exp\left[\frac{-\beta}{2\sigma^2}\right]\right) x$$

$$\left(\exp\left[\frac{-n}{2\sigma^2}\left(\theta - \frac{\sum_{i=1}^{n}\ln x_i}{n}\right)^2\right]\right) \quad (14)$$

D. Distribusi Posterior Marginal

Distribusi posterior marginal dari θ adalah integral terhadap σ^2 dari persamaan (14) menjadi

$$f(\theta|x) = \int_{\sigma^2=0}^{\infty} f(\theta,\sigma^2|x)\, d\sigma^2$$

$$f(\theta|x) = \sqrt{\frac{n}{\beta}} \frac{1}{B\left(\frac{1}{2},\frac{n-1}{2}\right)\left[1+\frac{n}{\beta}\left(\theta-\frac{\sum_{i=1}^{n}\ln x_i}{n}\right)^2\right]^{\frac{n}{2}}} \quad (15)$$

Distrribusi posterior marginal dari σ^2adalah integral terhadap θ dari persamaan (14) menjadi

$$f(\sigma^2|x) = \int_{\theta=-\infty}^{\infty} f(\theta,\sigma^2|x)\, d\theta$$

$$f(\sigma^2|x) = \frac{(\beta)^{\left(\frac{n-1}{2}\right)} exp\left[\frac{-\beta}{2\sigma^2}\right]}{(\sigma^2)^{\left(\frac{n+1}{2}\right)}(2)^{\left(\frac{n-1}{2}\right)}\Gamma\left(\frac{n-1}{2}\right)} \quad (16)$$

E. Estimasi Posterior

Hasil dari distribusi posterior marginal dari θ adalah diberikan pada persamaan (6). Sehingga estimasi posterior dari θ adalah sebagai berikut:

$$\theta^* = E(\theta|x) = \sqrt{\frac{n}{\beta}} \frac{1}{B\left(\frac{1}{2},\frac{n-1}{2}\right)} \int_{-\infty}^{\infty} \frac{\theta\, d\theta}{\left[1+\frac{n}{\beta}\left(\theta-\frac{\sum_{i=1}^{n}\ln x_i}{n}\right)^2\right]^{\frac{n}{2}}} \quad (17)$$

Misalkan: $t = \sqrt{\frac{n}{(n-1)s^2}}(\theta - \bar{x})\sqrt{n-1}$

$$\theta^* = \frac{\sum_{i=1}^{n}\ln x_i}{n\sqrt{n-1}B\left(\frac{1}{2},\frac{n-1}{2}\right)} \int_{-\infty}^{\infty} \frac{dt}{\left[1+\frac{t^2}{n-1}\right]^{\frac{n}{2}}}$$

$$\theta^* = \frac{\sum_{i=1}^{n}\ln x_i}{n} \quad (8)$$

Sekarang distribusi posterior marginal dari σ^2 adalah diberikan pada persamaan (4.7). Sehingga estimasi posterior dari σ^2 adalah sebagai berikut:

$$\sigma^{*2} = E(\sigma^2|x) = \int_0^\infty \frac{(\beta)^{\left(\frac{n-1}{2}\right)} exp\left[\frac{-\beta}{2\sigma^2}\right]\sigma^2}{(\sigma^2)^{\left(\frac{n+1}{2}\right)}(2)^{\left(\frac{n-1}{2}\right)}\Gamma\left(\frac{n-1}{2}\right)} d\sigma^2$$

$$\sigma^{*2} = \frac{(\beta)^{\left(\frac{n-1}{2}\right)}}{(2)^{\left(\frac{n-1}{2}\right)}\Gamma\left(\frac{n-1}{2}\right)} \int_0^\infty \frac{exp\left[\frac{-\beta}{2\sigma^2}\right]}{(\sigma^2)^{\left(\frac{n-1}{2}\right)}} d\sigma^2$$

$$\sigma^{*2} = \frac{\boldsymbol{\beta}}{\boldsymbol{n-1}} \quad (9)$$

Kemudian estimasi posterior dari ${\theta_1}^*$ dan ${\sigma_1}^{*2}$ adalah

$$\theta = \exp\left[\theta^* + \frac{\sigma^{*2}}{2}\right]$$

$${\theta_1}^* = \exp\left[\frac{\sum_{i=1}^{n}\ln x_i}{n} + \frac{\beta}{2(n-1)}\right] \quad (10)$$

Dan

$$\sigma^2 = \exp[2\theta^* + \sigma^{*2}](\exp[\sigma^{*2}] - 1)$$

$${\sigma_1}^{*2} = \exp\left[2\frac{\sum_{i=1}^{n}\ln x_i}{n} + \frac{\beta}{n-1}\right] x$$

$$\left(\exp\left[\frac{\beta}{n-1}\right] - 1\right) \quad (11)$$

IV. PENUTUP

Dari analisa yang telah dilakukan dapat disimpulkan bahwa mengestimasi parameter dengan menggunakan metode Estimasi Bayesian untuk mendapatkan estimasi titik. Nilai estimasi titik dari parameter mean θ varians σ^2diperoleh dengan bentuk

$${\theta_1}^* = \exp\left[\frac{\sum_{i=1}^{n}\ln x_i}{n} + \frac{\beta}{2(n-1)}\right]$$

$${\sigma_1}^{*2} = \exp\left[2\frac{\sum_{i=1}^{n}\ln x_i}{n} + \frac{\beta}{n-1}\right]\left(\exp\left[\frac{\beta}{n-1}\right] - 1\right)$$

DAFTAR PUSTAKA

[1] Supranto, J. 2001. **Statistika Teori dan Aplikasi Edisi ke-6**. Jakarta: Erlangga.

[2] George E. P. Box dan George C. Tiao. 1973. **Bayesian Inference in Statistical Analysis**. Boston: Addison-Wesley Publishing Company, 1973.

[3] Sultan, R. dan Ahmad, S. P. 2013. **Comparison of Parameters of Lognormal Distribution Based On the Classical and Posterior Estimates**. Journal of Modern Applied Statistical Methods. <http://digitalcommons.wayne.edu/cgi/>. Diakses pada tanggal 04 Maret 2016 pada pukul 08.10.

[4] Walpole, R. E. 1995. **Pengantar Statistika Edisi ke-3**. Jakarta: Gramedia Pustaka Utama.

[5] Sugito dan Ispriyanti, D. 2010. **Distribusi Invers Gamma Pada Inferensi Bayesian**. Media Statistik, (2010, Desember), 59-68.