INFERENSI BAYESIAN UNTUK σ^2 DARI DISTRIBUSI NORMAL DENGAN BERBAGAI DISTRIBUSI PRIOR

RESTI NANDA YANI, FERRA YANUAR, HAZMIRA YOZZA

*Jurusan Matematika,
Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Andalas,
Kampus UNAND Limau Manis Padang, Indonesia,
email : restinandayani.rny@gmail.com*

Abstrak. Pada penelitian ini dilakukan pendugaan parameter variansi (σ^2) dari distribusi Normal dengan mean (μ) diketahui. Pendugaan parameter variansi (σ^2) tersebut dilakukan secara analitik dengan menggunakan distribusi Invers Gamma sebagai prior konjugat, metode Jeffrey sebagai prior non-informatif dan distribusi Uniform sebagai prior non-konjugat. Pada penelitian ini kriteria evaluasi penduga yang digunakan adalah MSE dan sifat tak bias. Berdasarkan studi analitik diperoleh bahwa distribusi Invers Gamma sebagai prior konjugat merupakan prior terbaik diantara dua distribusi prior lainnya.

Kata Kunci: Inferensi statistika, metode Bayes, distribusi prior, fungsi likelihood, distribusi Normal, Invers Gamma, metode Jeffrey, distribusi Uniform

1. Pendahuluan

Inferensi statistik dapat dikelompokkan dalam dua bidang utama yaitu pendugaan parameter dan pengujian hipotesis. Pendugaan parameter merupakan prosedur yang dilakukan untuk menduga parameter populasi, seperti nilai tengah, ragam, proporsi, dan lain-lain. Parameter adalah sebarang nilai yang menjelaskan ciri populasi [8]. Uji hipotesis merupakan suatu proses untuk memutuskan benar atau salahnya suatu hipotesis berdasarkan hasil dari pengamatan [1].

Pendugaan parameter dapat dilakukan dengan dua metode yaitu metode klasik dan metode Bayes. Metode klasik melakukan pendugaan parameter hanya berdasarkan informasi yang diperoleh dari contoh acak yang diambil dari populasi. Metode Bayes menggabungkan pengetahuan subyektif mengenai distribusi peluang dari parameter yang tidak diketahui, dengan informasi yang diperoleh dari data sampel [7]. Distribusi peluang dari parameter yang tidak diketahui ini dipilih secara subyektif atau berdasarkan hasil penelitian sebelumnya dan idealnya ditentukan sebelum pengumpulan data dimulai. Distribusi ini disebut distribusi prior [6]. Informasi yang diperoleh mengenai fungsi kepekatan peluang dari data sampel disebut fungsi *likelihood*. Dengan menggabungkan informasi dari distribusi prior dan informasi dari data sampel, maka didapatkan distribusi posterior yang selanjutnya menjadi dasar untuk inferensi di dalam metode Bayes [3].

Terdapat beberapa jenis distribusi prior yaitu distribusi prior konjugat, distribusi prior non konjugat dan distribusi prior non-informatif [2]. Distribusi prior

disebut prior konjugat jika distribusi posteriornya memiliki keluarga sebaran yang sama karakternya dengan keluarga sebaran dari distribusi prior. Distribusi prior disebut prior non konjugat jika distribusi posteriornya memiliki keluarga sebaran yang tidak sama karakternya dengan distribusi prior [5]. Jika informasi mengenai distribusi parameter tidak diketahui, maka digunakan prior non-informatif. Pemilihan distribusi prior ini sangat berpengaruh terhadap dugaan parameter yang dihasilkan [2].

Salah satu distribusi dari peubah acak kontinu yang ada dalam bidang statistik adalah distribusi Normal. Distribusi Normal bergantung pada dua parameter yaitu μ dan σ^2 masing-masingnya sebagai nilai tengah dan simpangan baku yang dilambangkan dengan $N(\mu, \sigma^2)$. Dalam tugas akhir ini, akan diduga parameter variansi (σ^2) populasi yang berdistribusi Normal [8].

2. Landasan Teori

2.1. Distribusi Normal

Definisi 2.1. [4] Suatu peubah acak X dikatakan memiliki sebaran Normal dengan nilai tengah μ dan variansi σ^2 jika memiliki fungsi kepekatannya dalam bentuk

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2\right), \quad (2.1)$$

untuk $-\infty < x < \infty$, $-\infty < \mu < \infty$ dan $0 < \sigma < \infty$.

Peubah acak yang berdistribusi Normal dengan parameter μ dan σ^2 dapat dinyatakan dengan $X \sim N(\mu, \sigma^2)$.

2.2. Distribusi Gamma dan Invers Gamma

Definisi 2.2. [7] Peubah acak kontinu X dikatakan berdistribusi Gamma, dengan parameter $\kappa > 0$ dan $\theta > 0$, bila fungsi kepekatannya berbentuk:

$$f(x) = \begin{cases} \frac{1}{\theta^\kappa \Gamma(\kappa)} x^{\kappa-1} \exp\left(-\frac{x}{\theta}\right), & x > 0 \\ 0, & \text{selainnya} \end{cases} \quad (2.2)$$

Peubah acak X yang berdistribusi Gamma dengan parameter κ dan θ dapat dinyatakan dengan $X \sim GAM(\kappa, \theta)$.

Definisi 2.3. [5] Jika peubah acak Y berdistribusi Gamma, $Y \sim GAM(\kappa, \theta)$ maka peubah acak X dengan $X = \frac{1}{Y}$ berdistribusi Invers Gamma, $X \sim IG\left(\kappa, \frac{1}{\theta}\right)$ dengan fungsi kepekatannya:

$$f(x) = \begin{cases} \frac{1}{\Gamma(\kappa)} x^{-(\kappa+1)} \exp\left(-\frac{1}{x}\right), & x > 0, \\ 0, & \text{selainnya,} \end{cases} \quad (2.3)$$

bila $\kappa > 0$ dan $\theta > 0$.

Peubah acak X yang berdistribusi Invers Gamma dengan parameter κ dan θ dapat dinyatakan dengan $X \sim IG(\kappa, \theta)$.

2.3. Distribusi Uniform

Definisi 2.4. [1] Fungsi kepekatan peluang dari sebaran Uniform yaitu:

$$f(x; a, b) = \begin{cases} \frac{1}{b-a}, & a < x < b, \\ 0, & \text{selainnya.} \end{cases} \quad (2.4)$$

2.4. Fungsi Likelihood

Definisi 2.5. [1] Fungsi likelihood adalah fungsi kepekatan peluang bersama dari n peubah acak $X_1, X_2, \dots, X_n$ yang dihitung pada $x_1, x_2, \dots, x_n$ dinyatakan dalam bentuk $f(x_1, x_2, \dots, x_n)$. Jika $x_1, x_2, \dots, x_n$ ditetapkan, maka fungsi likelihood adalah fungsi dari parameter dan dinotasikan dengan $f(x|\theta)$. Jika $X_1, X_2, \dots, X_n$ menyatakan suatu peubah acak dari $f(x; \theta)$, maka:

$$\begin{aligned} f(x|\theta) &= f(x_1; \theta)f(x_2; \theta) \cdots f(x_n; \theta) \\ &= \prod_{i=1}^n f(x_i; \theta) \end{aligned}$$

2.5. Distribusi Prior dan Posterior

Dalam metode Bayes, distribusi prior memiliki peran penting dalam pendugaan parameter θ yang tidak diketahui. Distribusi prior merupakan distribusi awal yang memberikan informasi mengenai parameter dan harus ditentukan terlebih dahulu sebelum merumuskan distribusi posteriornya. Pada dasarnya, distribusi prior dipilih secara subyektif oleh peneliti terhadap suatu nilai parameter yang diduga.

Definisi 2.6. [1] Fungsi kepekatan peluang bersyarat dari θ jika diketahui pengamatan sampel $x = (x_1, x_2, \dots, x_n)$ disebut fungsi kepekatan peluang posterior, yang diberikan dari:

$$f(\theta|x) = \frac{f(x|\theta)f(\theta)}{\int_{-\infty}^{\infty} f(x|\theta)f(\theta) d\theta} \quad (2.5)$$

Definisi 2.7. [8] Nilai tengah dari distribusi posterior $f(\theta|x_1, x_2, \dots, x_n)$ dinyatakan dengan T , disebut penduga Bayes untuk $\tau(\theta)$.

2.6. Metode Jeffrey

Salah satu bentuk pendekatan dari prior non-informatif adalah dengan menggunakan metode Jeffrey. Metode ini menyatakan bahwa distribusi prior $f(\theta)$ merupakan akar kuadrat dari informasi Fisher yang dinyatakan sebagai $I(\theta)$.

Aturan Jeffrey

Distribusi prior $f(\theta)$ dikatakan distribusi prior non-informatif dari parameter θ jika distribusi prior tersebut proporsional dengan akar dari informasi Fisher [2].

$$f(\theta) \propto \sqrt{I(\theta)},$$

dengan informasi Fisher dari parameter θ untuk suatu peubah acak $x = (x_1, x_2, \dots, x_n)$ didefinisikan dengan:

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

3. Pembahasan

3.1. Likelihood dari Distribusi Normal

Jika diketahui $X_1, X_2, \dots, X_n$ peubah acak dari distribusi Normal, dengan $X_i \sim N(\mu, \sigma^2)$, dengan nilai tengah μ diketahui, maka fungsi *likelihood*-nya adalah:

$$\begin{aligned} f(x|\sigma^2) &= \prod_{i=1}^n f(x_i; \sigma^2) \\ &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left(-\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right) \\ &= (2\pi\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right] \end{aligned}$$

3.2. Analisis Bayesian untuk σ^2 dengan Berbagai Prior

3.2.1. Distribusi Invers Gamma sebagai Prior Konjugat

Misalkan dipilih peubah acak σ^2 berdistribusi Invers Gamma dengan parameter $\kappa = 1$ dan $\theta = 1$, ditulis dengan $\sigma^2 \sim IG(1, 1)$ sebagai distribusi prior konjugat. Fungsi kepekatan peluang dari σ^2 dapat dinyatakan dalam bentuk:

$$f(\sigma^2) = (\sigma^2)^{-2} \exp \left(-\frac{1}{\sigma^2} \right), \sigma^2 > 0.$$

Dengan demikian hasil kali fungsi *likelihood* dengan distribusi prior dapat ditulis sebagai berikut:

$$f(\sigma^2, \mathbf{x}) = (\sigma^2)^{-\frac{n}{2}} \exp \left(-\frac{1}{\sigma^2} \right) \times (2\pi\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma^2} \right] \quad (3.1)$$

$$= (2\pi)^{-\frac{n}{2}} (\sigma^2)^{-\left(\frac{n}{2}+2\right)} \exp \left[-\left(\frac{2 + \sum_{i=1}^n (x_i - \mu)^2}{2\sigma^2} \right) \right] \quad (3.2)$$

Selanjutnya, integralkan hasil kali fungsi *likelihood* dengan distribusi prior terhadap σ^2 , sehingga diperoleh:

$$f(\mathbf{x}) = (2\pi)^{-\frac{n}{2}} \left(\frac{2 + \sum_{i=1}^n (x_i - \mu)^2}{2} \right) \Gamma \left(\frac{n}{2} + 1 \right). \quad (3.3)$$

Dengan demikian, distribusi posterior untuk σ^2 yang bersyarat $\mathbf{x}$ dapat ditulis:

$$\begin{aligned} f(\sigma^2|\mathbf{x}) &= \frac{(2\pi)^{-\frac{n}{2}} (\sigma^2)^{-(\frac{n}{2}+2)} \exp \left[- \left(\frac{2 + \sum_{i=1}^n (x_i - \mu)^2}{2\sigma^2} \right) \right]}{(2\pi)^{-\frac{n}{2}} \left(\frac{2 + \sum_{i=1}^n (x_i - \mu)^2}{2} \right) \Gamma \left(\frac{n}{2} + 1 \right)} \\ &= \frac{\left(\frac{2 + \sum_{i=1}^n (x_i - \mu)^2}{2} \right)^{\left(\frac{n}{2} + 1 \right)}}{\Gamma \left(\frac{n}{2} + 1 \right)} (\sigma^2)^{-(\frac{n}{2}+2)} \exp \left[- \left(\frac{2 + \sum_{i=1}^n (x_i - \mu)^2}{2\sigma^2} \right) \right] \end{aligned}$$

Atau dapat dinyatakan bahwa $\sigma_1^2|\mathbf{x} \sim IG \left(\frac{n}{2} + 1, \frac{2 + \sum_{i=1}^n (x_i - \mu)^2}{2} \right)$.

3.2.2. Metode Jeffrey sebagai Prior Non-Informatif

Misalkan suatu peubah acak ϑ yang beranggotakan μ dan σ^2 yang saling bebas atau ditulis $\vartheta = (\mu, \sigma^2)$. Pada bagian ini distribusi dari peubah acak ϑ inilah yang dijadikan sebagai prior non-informatif, sehingga fungsi kepekatan peluang dari ϑ dapat ditulis $f(\vartheta) = f(\mu, \sigma^2) = f(\mu)f(\sigma^2)$.

Prior non-informatif untuk $f(\sigma^2)$:

$$\begin{aligned} f(X; \mu, \sigma^2) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left(-\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right) \\ \log f(X; \mu, \sigma^2) &= -\frac{1}{2} \log(2\pi) - \frac{1}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} (X - \mu)^2 \\ \frac{d \log f(X; \mu, \sigma^2)}{d\sigma^2} &= -\frac{1}{2\sigma^2} + \frac{1}{2\sigma^2} (X - \mu)^2 \\ \frac{d^2 \log f(X; \mu, \sigma^2)}{d(\sigma^2)^2} &= \frac{1}{2\sigma^4} - \frac{1}{\sigma^6} (X - \mu)^2 \\ I(\sigma^2) &= -E \left[\frac{1}{2\sigma^4} - \frac{1}{\sigma^6} (X - \mu)^2 \right] \\ &= -\frac{1}{2\sigma^4} + \frac{1}{\sigma^6} \sigma^2 \\ &= \frac{1}{2\sigma^4} \\ &= \frac{1}{\sqrt{2}\sigma^2} \end{aligned}$$

Dengan demikian didapatkan distribusi prior non-informatif untuk $f(\sigma^2)$ yaitu:

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

Karena μ nilainya diketahui atau dengan kata lain μ merupakan suatu konstanta, sehingga $f(\mu) = c$ (konstan). Dengan demikian diperoleh:

$$f(\vartheta) = (\mu)f(\sigma^2) \propto c \times \frac{1}{\sigma^2} \propto \frac{1}{\sigma^2}.$$

Jadi distribusi prior non-informatif untuk $f(\vartheta) \propto \frac{1}{\sigma^2}$.

Setelah didapatkan fungsi *likelihood* dan distribusi prior non-informatif, maka dapat dicari distribusi posteriornya:

$$f(\vartheta|\mathbf{x}) = \frac{f(\mathbf{x}|\vartheta)f(\vartheta)}{\int_{\sigma^2=0}^{\infty} \int_{\mu=-\infty}^{\infty} f(\mathbf{x}|\vartheta)f(\vartheta) d\mu d\sigma^2}.$$

Penjabaran untuk hasil kali fungsi likelihood dengan distribusi prior non-informatif akan proporsional sama dengan:

$$f(\mathbf{x}|\vartheta)f(\vartheta) \propto (2\pi)^{-\frac{n}{2}}(\sigma^2)^{-(\frac{n}{2}+1)} \exp \left[-\frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma^2} \right]$$

Pertama integralkan $f(\mathbf{x}|\vartheta)f(\vartheta)$ terhadap μ sehingga diperoleh:

$$\int_{\mu=-\infty}^{\infty} f(\mathbf{x}|\vartheta)f(\vartheta) d\mu \propto (2\pi)^{-\frac{n}{2}} \left(\frac{2\pi}{n} \right)^{\frac{1}{2}} (\sigma^2)^{-(\frac{n-1}{2}+1)} \exp \left[-\frac{(n-1)s^2}{2\sigma^2} \right].$$

Selanjutnya, integralkan terhadap σ^2 , sehingga diperoleh $f(x)$ seperti berikut:

$$f(x) \propto (2\pi)^{-\frac{n}{2}} \left(\frac{2\pi}{n} \right)^{\frac{1}{2}} \left(-\frac{(n-1)s^2}{2} \right)^{-(\frac{n-1}{2})} \Gamma \left(\frac{n-1}{2} \right).$$

Dengan demikian distribusi posterior bersama untuk μ dan σ^2 bersyarat $\mathbf{x}$ adalah sebagai berikut:

$$f(\vartheta|\mathbf{x}) \propto \left(\frac{2\pi}{n} \right)^{-\frac{1}{2}} \left(\frac{(n-1)s^2}{2} \right)^{(\frac{n-1}{2})} \left(\Gamma \left(\frac{n-1}{2} \right) \right)^{-1} (\sigma^2)^{-(\frac{n}{2}+1)} \exp \left[-\frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma^2} \right].$$

Dengan mengintegalkan $f(\vartheta|x)$ terhadap μ , maka diperoleh distribusi posterior marginal untuk σ^2 bersyarat x yaitu:

$$f(\sigma^2|\mathbf{x}) \propto \frac{\left(\frac{(n-1)s^2}{2} \right)^{(\frac{n-1}{2})}}{\Gamma \left(\frac{n-1}{2} \right)} (\sigma^2)^{-(\frac{n-1}{2}+1)} \exp \left[-\frac{(n-1)s^2}{2\sigma^2} \right].$$

Atau dapat dinyatakan bahwa

$$\sigma_2^2|\mathbf{x} \sim IG \left(\frac{n-1}{2}, \frac{(n-1)s^2}{2} \right).$$

3.2.3. Distribusi Uniform sebagai Prior Non-Konjugat

Misalkan dipilih peubah acak σ^2 berdistribusi Uniform dengan parameter $a = 0$ dan $b = 1$ atau ditulis dengan $\sigma^2 \sim UNIF(0, 1)$. Bentuk fungsi kepekatan peluangnya adalah:

$$f(\sigma^2) = 1, \quad 0 < \sigma^2 < 1.$$

Dengan demikian hasil kali fungsi *likelihood* dengan distribusi prior akan sama dengan fungsi *likelihood*, sehingga integralnya sebagai berikut:

$$\int_0^1 f(\sigma^2) f(\mathbf{x}|\sigma^2) d\sigma^2 = (2\pi)^{-\frac{n}{2}} \left(\frac{\sum_{i=1}^n (x_i - \mu)^2}{2} \right)^{-\left(\frac{n}{2}-1\right)} \Gamma\left(\frac{n}{2} - 1\right).$$

Distribusi posterior untuk σ^2 yang bersyarat x dapat diperoleh sebagai berikut:

$$f(\sigma^2|\mathbf{x}) = \frac{\left(\frac{\sum_{i=1}^n (x_i - \mu)^2}{2} \right)^{\left(\frac{n}{2}-1\right)}}{\Gamma\left(\frac{n}{2} - 1\right)} (\sigma^2)^{-\left(\frac{n}{2}\right)} \exp \left[-\frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma^2} \right].$$

Atau dapat dinyatakan

$$\sigma_3^2|\mathbf{x} \sim IG \left(\frac{n}{2} - 1, \frac{\sum_{i=1}^n (x_i - \mu)^2}{2} \right).$$

<u>Kriteria</u>	<u>Jenis Distribusi Prior</u>		
	Invers Gamma	Jeffrey	Uniform
<u>Distribusi posterior</u>	$IG \left(\frac{n}{2} + 1, \frac{2 + \sum_{i=1}^n (x_i - \mu)^2}{2} \right)$	$IG \left(\frac{n-1}{2}, \frac{(n-1)s^2}{2} \right)$	$IG \left(\frac{n}{2} - 1, \frac{\sum_{i=1}^n (x_i - \mu)^2}{2} \right)$
Nilai Tengah posterior	$\frac{2 + \sum_{i=1}^n (x_i - \mu)^2}{n}$	$\frac{(n-1)s^2}{n-3}$	$\frac{\sum_{i=1}^n (x_i - \mu)^2}{n-4}$
<u>Variansi posterior</u>	$\frac{2(2 + \sum_{i=1}^n (x_i - \mu)^2)^2}{n^2(n-2)}$	$\frac{2(n-1)^2 s^4}{(n-3)^2(n-5)}$	$\frac{2(\sum_{i=1}^n (x_i - \mu)^2)^2}{(n-4)^2(n-6)}$
<u>Mean Square Error (MSE)</u>	$\frac{2n\sigma^2 + 4}{n^2}$	$\frac{(2n+2)\sigma^2}{(n-3)^2}$	$\frac{(2n+16)\sigma^2}{(n-4)^2}$

Gambar 1. Hasil Uraian Secara Analitik Distribusi Posterior dengan Menggunakan Tiga Jenis Distribusi Prior

4. Kesimpulan

Berdasarkan pembahasan diketahui bahwa, untuk ketiga distribusi prior yang berbeda diperoleh distribusi posterior yang proporsional sama dengan distribusi invers Gamma.

Gambar 2. Distribusi prior dari distribusi Invers Gamma, Jeffrey dan Uniform

Gambar 3. Distribusi Posterior dari Prior Invers Gamma, Prior Jeffrey dan Prior Uniform

Daftar Pustaka

- [1] Bain, L.J and Engelhardt, M. 1992. *Introduction to Probability and Mathematical Statistic*. Second Edition. Duxbury Press, California
- [2] Box, G.E.P. dan Tiao, G.C. 1973. *Bayesian Inference in Statistical Analysis*. Second Edition. Addison-Wesley Publishing Company, London
- [3] Casella, G and R.L. Berger. 2001. *Statistical Inference*. Second Edition. Pacific Grove, California
- [4] Hogg, R.V. and Craig, A.T. 1905. *Introduction to Mathematical Statistic*. Seventh Edition. Pearson Education, United States of America
- [5] Koch, K. R. 2007. *Introduction to Bayesian Statistics* Second Edition. Springer, New York
- [6] Mukhopadhyay, N. 2000. *Probability and Statistical Inference*. Marcel Dekker, New York
- [7] Walpole, E.R. 1986. *Pengantar Statistika*. Edisi Ketiga. PT Gramedia Pustaka Utama, Jakarta
- [8] Walpole, E.R. dan Myers, R.H. 1995. *Ilmu Peluang dan Statistika untuk Insinyur dan Ilmuwan*. Edisi Keempat. ITB, Bandung