

## Bayesian Inference for Pareto Distribution with Prior Conjugate and Prior Non Conjugate

### Inferensi Bayesian untuk Distribusi Pareto dengan Prior Konjugat dan Prior Non Konjugat

Ferra Yanuar\*<sup>1</sup>, Cici Saputri\*<sup>2</sup>, Dodi Devianto\*<sup>3</sup>

#### Abstract

The purpose of this study is to determine the best estimator for estimating the shape  $\theta$  parameters of the Pareto distribution with the known  $\lambda$  scale parameter. Estimation of these parameters is done by using the Gamma distribution as the prior distribution of the conjugate and the Uniform distribution as the non-conjugate prior distribution. A comparison of the two prior distributions is done through simulation studies with various sample sizes. The best estimator net is a method that produces the smallest posterior variance, absolute bias, and Bayes confidence interval. This study proves that the Bayes estimator by using the prior conjugate distribution produces all indicators of the goodness of the model with a smaller value than the non-conjugate prior distribution. Thus it can be concluded that the estimator with prior conjugate will produce a better predictive value than prior non-conjugate.

**Keywords:** Pareto distribution, posterior variance, absolute bias, Bayes confidence interval.

#### Abstrak

Tujuan penelitian ini adalah menentukan penduga terbaik untuk menduga parameter bentuk  $\theta$  dari distribusi Pareto dengan parameter skala  $\lambda$  diketahui. Pendugaan parameter tersebut dilakukan dengan menggunakan distribusi Gamma sebagai distribusi prior konjugat dan distribusi Uniform sebagai distribusi prior non-konjugat. Perbandingan kedua distribusi prior dilakukan melalui studi simulasi dengan berbagai ukuran sampel. Metode penduga terbaik adalah metode yang menghasilkan nilai posterior variansi, nilai bias, dan lebar selang kepercayaan Bayes yang terkecil. Penelitian ini membuktikan bahwa penduga Bayes dengan menggunakan distribusi prior konjugat menghasilkan semua indikator kebaikan model dengan nilai yang lebih kecil daripada dengan distribusi prior non-konjugat. Dengan demikian dapat disimpulkan bahwa penduga dengan prior konjugat akan menghasilkan nilai dugaan yang lebih baik daripada prior non-konjugat.

**Kata kunci:** distribusi Pareto variansi posterior, bias mutlak, selang kepercayaan Bayes.

\**Program Studi Matematika,*  
*Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Andalas,*  
*Kampus UNAND Limau Manis Padang, Indonesia*  
<sup>1</sup>*ferrayanuar@sci.unand.ac.id*(Corresponding author), <sup>2</sup>*csaputri5@gmail.com*

## 1. Pendahuluan

Statistika inferensia dikelompokkan dalam dua bidang utama yaitu pendugaan parameter dan pengujian hipotesis. Penduga parameter merupakan prosedur yang dilakukan untuk menduga parameter populasi seperti nilai tengah, ragam, proporsi, dan lain sebagainya. Sedangkan pengujian hipotesis adalah suatu prosedur yang dilakukan dengan tujuan untuk memutuskan apakah menerima atau menolak hipotesis mengenai parameter populasi.

Metode yang dapat digunakan untuk menduga parameter adalah metode klasik dan metode Bayes. Metode klasik melakukan pendugaan parameter hanya berdasarkan informasi yang diperoleh dari contoh acak populasi. Metode Bayes menggabungkan pengetahuan subyektif mengenai distribusi peluang dari parameter yang tidak diketahui, dengan informasi yang diperoleh dari data sampel (fungsi *likelihood*) [1]. 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 dengan distribusi prior [1]–[4]. Gabungan fungsi *likelihood* dan distribusi prior akan menghasilkan distribusi baru yang dinamakan dengan distribusi posterior yang selanjutnya akan menjadi dasar inferensi dalam metode Bayes.

Salah satu distribusi peubah acak kontinu dalam statistika adalah distribusi Pareto. Distribusi Pareto merupakan distribusi peubah acak kontinu dengan dua parameter yaitu parameter skala yang disimbolkan dengan  $\lambda$ , dan parameter bentuk yang disimbolkan dengan  $\theta$ . Penelitian ini akan membahas pendugaan parameter bentuk ( $\theta$ ) dari distribusi Pareto dimana parameter skala ( $\lambda$ ) diketahui menggunakan metode Bayes. Distribusi prior yang digunakan dalam menduga parameter  $\theta$  dari distribusi Pareto ini adalah distribusi Gamma sebagai distribusi prior konjugat dan distribusi Uniform sebagai distribusi prior non-konjugat.

## 2. Tinjauan Teoritis

### 2.1. Distribusi Pareto

Distribusi Pareto diperkenalkan oleh seorang pakar ekonomi dari Italia yang bernama Vilfredo Pareto. Distribusi Pareto sering dipakai dalam persoalan uji hidup seperti umur suatu komponen yang diukur dari suatu waktu tertentu sampai rusak. Distribusi Pareto merupakan salah satu distribusi peubah acak kontinu dengan dua parameter yaitu  $\theta$  sebagai parameter bentuk dan  $\lambda$  sebagai parameter skala.

**Definisi 1** [5] *Fungsi kepekatan peluang dari distribusi Pareto dengan parameter bentuk  $\theta$  dan parameter skala  $\lambda$  adalah*

$$f(x; \lambda, \theta) = \frac{\theta \lambda^\theta}{x^{\theta+1}}$$

dengan  $x \geq \lambda > 0$  dan  $\theta > 0$ . Sedangkan fungsi distribusi kumulatifnya adalah

$$F(x) = 1 - \left(\frac{\lambda}{x}\right)^\theta, \quad x \geq \lambda$$

Peubah acak  $X$  berdistribusi Pareto dengan parameter  $\lambda$  dan  $\theta$  dinotasikan dengan  $X \sim \text{Pareto}(\lambda, \theta)$ .

**Definisi 2** [6] Fungsi Gamma dilambangkan dengan  $\Gamma(\kappa)$  untuk  $\kappa > 0$  suatu bilangan bulat yang didefinisikan oleh

$$\Gamma(\kappa) = \int_0^{\infty} t^{\kappa-1} e^{-t} dt.$$

**Definisi 3** [7] Suatu peubah acak  $X$  dikatakan berdistribusi Gamma dengan parameter bentuk  $\alpha$  dan parameter skala  $\beta$ , jika fungsi kepekatan peluangnya adalah

$$f(x; \alpha, \beta) = \frac{\beta^\alpha x^{\alpha-1} e^{-x\beta}}{\Gamma(\alpha)}, \quad \text{dengan } x \geq 0 \text{ dan } \alpha, \beta > 0$$

Peubah acak  $X$  berdistribusi Gamma dengan parameter  $\alpha$  dan  $\beta$  dinotasikan dengan  $X \sim \text{Gamma}(\alpha, \beta)$ .

**Definisi 4** [8] Suatu peubah acak kontinu  $X$  dikatakan distribusi Uniform pada selang  $(a, b)$  jika memiliki fungsi kepekatan peluang sebagai berikut:

$$f(x) = \frac{1}{b-a}, \quad a < x < b$$

dan nol untuk  $x$  lainnya.

Notasi yang menyatakan bahwa  $X$  memiliki fungsi kepekatan peluang seperti pada Definisi 4 adalah  $X \sim \text{Unif}(a, b)$ .

## 2.2. Fungsi Likelihood

Konsep dasar lainnya yang digunakan dalam metode Bayes adalah fungsi *likelihood*, distribusi prior dan distribusi posterior. Berikut uraian ringkas terkait konsep tersebut.

**Definisi 5** [6] Fungsi kepekatan peluang bersama dari peubah acak berdimensi  $n$ ,  $X_1, X_2, \dots, X_n$ , yang dihitung pada  $x_1, x_2, \dots, x_n$  dengan parameter  $\theta$  didefinisikan dengan  $f(x_1, x_2, \dots, x_n; \theta)$  dan dikatakan sebagai fungsi *likelihood*.. Untuk  $x_1, x_2, \dots, x_n$  tetap, fungsi *likelihood* adalah fungsi dari  $\theta$  yang dinotasikan dengan  $L(\theta)$ . Jika  $X_1, X_2, \dots, X_n$  merupakan contoh acak dari  $f(x; \theta)$  maka fungsi *likelihood* didefinisikan dengan:

$$\begin{aligned} L(\theta) &= f(x_1; \theta) f(x_2; \theta) \dots f(x_n; \theta) \\ &= \prod_{i=1}^n f(x_i; \theta). \end{aligned}$$

Distribusi prior merupakan distribusi awal dari suatu parameter yang memberikan informasi mengenai parameter yang tidak diketahui nilainya. Dalam metode Bayes, distribusi prior memiliki peran penting dalam pendugaan parameter karena untuk mendapatkan distribusi posterior maka harus ditentukan terlebih dahulu distribusi priornya.

**Definisi 6** [9] Fungsi kepekatan peluang posterior dari  $\theta$  apabila diketahui contoh pengamatan  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  adalah:

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

**Definisi 7** [10] Nilai tengah dari distribusi posterior  $f(\theta|x_1, x_2, \dots, x_n)$  dinotasikan dengan  $\hat{\theta}$ , disebut dengan penduga Bayes untuk  $\theta$ .

### 3. Hasil dan Pembahasan

Pada bagian ini diuraikan hasil studi analitis dan studi simulasi dari pendugaan parameter bentuk  $\theta$  dari distribusi Pareto dengan menggunakan metode Bayes serta pembahasan dari hasil yang diperoleh.

#### 3.1 Fungsi Likelihood dari Distribusi Pareto

Misalkan  $X_1, X_2, \dots, X_n$  adalah peubah acak dari distribusi Pareto dengan parameter skala  $\lambda$  dan parameter bentuk  $\theta$ , atau dapat ditulis dengan  $X_i \sim \text{Pareto}(\lambda, \theta)$ , maka fungsi *likelihood*nya adalah

$$\begin{aligned} L(\theta) &= \prod_{i=1}^n f(x_i; \theta) \\ &= \prod_{i=1}^n \frac{\theta \lambda^\theta}{x_i^{\theta+1}} \\ &= \theta^n \lambda^{n\theta} \prod_{i=1}^n \frac{1}{x_i^{\theta+1}} \end{aligned}$$

#### 3.2. Pendugaan Parameter $\theta$ dari Distribusi Pareto dengan $\lambda$ Diketahui Menggunakan Distribusi Gamma Sebagai Prior Konjugat

Misalkan dipilih suatu peubah acak  $\theta$  berdistribusi Gamma dengan parameter  $(\alpha, \beta)$ , atau ditulis dengan  $\theta \sim \text{Gamma}(\alpha, \beta)$ , sebagai distribusi prior konjugat. Fungsi kepekatan peluang dari  $\theta$  dapat dinyatakan sebagai berikut:

$$f(\theta) = \frac{\beta^\alpha \theta^{\alpha-1} e^{-\theta\beta}}{\Gamma(\alpha)}$$

dengan  $\theta \geq 0$ ,  $\alpha, \beta > 0$ .

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

$$\begin{aligned} f(\theta, \mathbf{x}) &= \frac{\beta^\alpha \theta^{\alpha-1} e^{-\theta\beta}}{\Gamma(\alpha)} \times \theta^n \lambda^{n\theta} \prod_{i=1}^n \frac{1}{x_i^{\theta+1}} \\ &= \theta^{n+\alpha-1} \left( \frac{\lambda^n e^{-\beta}}{\prod_{i=1}^n x_i} \right)^\theta \frac{\beta^\alpha}{\prod_{i=1}^n x_i \Gamma(\alpha)} \end{aligned}$$

Selanjutnya akan dicari formula untuk  $f(\mathbf{x})$  dengan cara mengintegrasikan hasil kali fungsi likelihood dengan distribusi prior terhadap  $\theta$ , sehingga diperoleh:

$$\begin{aligned} f(\mathbf{x}) &= \int_0^\infty \theta^{n+\alpha-1} \left( \frac{\lambda^n e^{-\beta}}{\prod_{i=1}^n x_i} \right)^\theta \frac{\beta^\alpha}{\prod_{i=1}^n x_i \Gamma(\alpha)} d\theta \\ &= \frac{\beta^\alpha}{\prod_{i=1}^n x_i \Gamma(\alpha)} \left( \frac{1}{\beta + \sum_{i=1}^n \ln \frac{x_i}{\lambda}} \right)^{\alpha+n} \Gamma(\alpha + n) \end{aligned}$$

Dengan demikian, distribusi posterior untuk  $\theta$  bersyarat  $\mathbf{x}$  adalah:

$$\begin{aligned} f(\theta|\mathbf{x}) &= \frac{f(\theta, \mathbf{x})}{f(\mathbf{x})} \\ &= \frac{\left( \beta + \sum_{i=1}^n \ln \frac{x_i}{\lambda} \right)^{\alpha+n} \theta^{n+\alpha-1} e^{-\theta(\beta + \sum_{i=1}^n \ln \frac{x_i}{\lambda})}}{\Gamma(\alpha + n)} \end{aligned}$$

Atau dapat dinyatakan bahwa:

$$\theta_1 | \mathbf{x} \sim \text{Gamma} \left( \alpha + n, \beta + \sum_{i=1}^n \ln \frac{x_i}{\lambda} \right)$$

### 3.3. Pendugaan Parameter $\theta$ dari Distribusi Pareto dengan $\lambda$ Diketahui Menggunakan Distribusi Uniform Sebagai Prior Non-Konjugat

Misalkan dipilih suatu peubah acak  $\theta$  berdistribusi Uniform dengan parameter  $a = 0$  dan  $b = 1$ , atau ditulis  $\theta \sim \text{Unif}(0,1)$ . Fungsi kepekatan peluang dari peubah acak  $\theta$  adalah:

$$f(\theta) = \frac{1}{b-a} = 1, \quad 0 < \theta < 1$$

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

$$f(\theta, \mathbf{x}) = \theta^n \lambda^{n\theta} \prod_{i=1}^n \frac{1}{x_i^{\theta+1}}$$

Selanjutnya akan dicari formula untuk  $f(\mathbf{x})$  dengan cara mengintegalkan hasil kali fungsi likelihood dengan distribusi prior terhadap  $\theta$ , sehingga diperoleh:

$$\begin{aligned} f(\mathbf{x}) &= \int_0^{\infty} \theta^n \lambda^{n\theta} \prod_{i=1}^n \frac{1}{x_i^{\theta+1}} d\theta \\ &= \frac{1}{\prod_{i=1}^n x_i} \left( \frac{1}{\sum_{i=1}^n \ln \frac{x_i}{\lambda}} \right)^{n+1} \Gamma(n+1) \end{aligned}$$

Dengan demikian, distribusi posterior untuk  $\theta$  bersyarat  $\mathbf{x}$  adalah:

$$\begin{aligned} f(\theta|\mathbf{x}) &= \frac{\theta^n \lambda^{n\theta} \prod_{i=1}^n \frac{1}{x_i^{\theta+1}}}{\frac{1}{\prod_{i=1}^n x_i} \left( \frac{1}{\sum_{i=1}^n \ln \frac{x_i}{\lambda}} \right)^{n+1} \Gamma(n+1)} \\ &= \frac{\left( \sum_{i=1}^n \ln \frac{x_i}{\lambda} \right)^{n+1} \theta^{n+1-1} e^{-\theta \left( \sum_{i=1}^n \ln \frac{x_i}{\lambda} \right)}}{\Gamma(n+1)} \end{aligned}$$

Atau dapat dinyatakan bahwa:

$$\theta_2|\mathbf{x} \sim \text{Gamma} \left( n+1, \sum_{i=1}^n \ln \frac{x_i}{\lambda} \right)$$

### 3.4. Studi Simulasi

Pada sub bab ini akan dilakukan pendugaan parameter skala  $\theta$  dengan menggunakan data bangkitan R yang berdistribusi Pareto dimana nilai  $\lambda$  diketahui. Data sampel yang diambil sebanyak tiga buah kelompok data dengan ukuran yang berbeda, yaitu  $n_1 = 30$ ,  $n_2 = 80$ , dan  $n_3 = 200$ . Pada data bangkitan ini ditetapkan  $\lambda = 3$  dan  $\theta = 1$ . Tabel 1 dan Tabel 2 berikut menyajikan rangkuman hasil studi simulasi untuk pendugaan parameter yang dimaksud.

**Tabel 1.** Hasil Pendugaan untuk Parameter  $\theta$

Banyak Data (n)	Distribusi Prior	Posterior Nilai Tengah	Variansi Posterior
30	Konjugat	1,2706	0,0265
	Non-konjugat	1,2914	0,0538
80	Konjugat	1,2643	0,0099
	Non-konjugat	1,2722	0,0199

200	Konjugat	1,0976	0,0030
	Non-konjugat	1,1002	0,0060

**Tabel 2.** Hasil Pendugaan Selang Kepercayaan Bayes 95% untuk Parameter  $\theta$

Banyak Data (n)	Distribusi Prior	Bias*	Selang Kepercayaan Bayes 95%	Lebar Selang
30	Konjugat	0,2706	(0,9515 ; 1,5897)	0,6382
	Non-Konjugat	0,2914	(0,8368 ; 1,7460)	0,9092
80	Konjugat	0,2643	(1,0693 ; 1,4593)	0,3900
	Non-konjugat	0,2722	(0,9957 ; 1,5487)	0,5530
200	Konjugat	0,0976	(0,9902 ; 1,2050)	0,2148
	Non-konjugat	0,1002	(0,9484 ; 1,2520)	0,3036

\*Bias = Nilai dugaan – Nilai yang ditetapkan

Berdasarkan Tabel 1 di atas dapat diketahui bahwa posterior nilai tengah dari distribusi prior konjugat Gamma lebih mendekati nilai sebenarnya daripada distribusi prior non-konjugat. Selain itu, posterior variansi dari distribusi prior konjugat Gamma menghasilkan nilai yang lebih kecil daripada posterior variansi dari distribusi prior non-konjugat. Hasil perhitungan pada Tabel 2 menunjukkan bahwa nilai bias dan lebar selang kepercayaan Bayes 95% dari distribusi prior konjugat Gamma lebih kecil daripada distribusi prior non-konjugat. Jika dilihat dari banyaknya data sampel yang dipilih, dapat disimpulkan bahwa semakin banyak data sampel yang digunakan maka posterior nilai tengah dari distribusi prior konjugat semakin mendekati nilai sebenarnya dan nilai dari posterior variansi, bias serta lebar selang kepercayaannya semakin kecil.

Hasil ini membuktikan bahwa penduga Bayes dengan menggunakan distribusi prior konjugat lebih baik daripada penduga Bayes dengan distribusi prior non-konjugat [6].

#### 4. Kesimpulan

Berdasarkan pembahasan yang telah dilakukan, maka dapat disimpulkan bahwa distribusi prior konjugat dari distribusi Pareto  $(\lambda, \theta)$  dengan  $\lambda$  diketahui adalah distribusi Gamma  $(\alpha', \beta')$  dimana  $\alpha' = n + 1$  dan  $\beta' = \sum_{i=1}^n \ln \frac{x_i}{\lambda}$ . Distribusi posterior yang diperoleh dapat dinyatakan

dalam bentuk distribusi Gamma  $(\alpha^*, \beta^*)$  dimana  $\alpha^* = \alpha' + n$  dan  $\beta^* = \beta' + \sum_{i=1}^n \ln \frac{x_i}{\lambda}$  dengan penduga Bayes untuk  $\theta_1$  adalah  $\hat{\theta}_1 = \frac{\alpha+n}{\beta+\sum_{i=1}^n \ln \frac{x_i}{\lambda}}$ .

Studi ini juga menghasilkan bahwa distribusi prior non-konjugat dari distribusi Pareto  $(\lambda, \theta)$  dengan  $\lambda$  diketahui adalah distribusi Uniform  $(0,1)$  dimana nilai dari distribusi prior non-konjugat yang dihasilkan adalah 1. Distribusi posterior yang diperoleh dapat dinyatakan dalam bentuk distribusi Gamma  $(\alpha^*, \beta^*)$  dimana  $\alpha^* = n + 1$  dan  $\beta^* = \sum_{i=1}^n \ln \frac{x_i}{\lambda}$  dengan penduga Bayes untuk  $\theta_2$  adalah  $\hat{\theta}_2 = \frac{n+1}{\sum_{i=1}^n \ln \frac{x_i}{\lambda}}$ .

Berdasarkan studi simulasi dengan menggunakan tiga buah kelompok data yang telah dibangkitkan menggunakan program R untuk kedua jenis distribusi prior, diperoleh kesimpulan bahwa nilai posterior variansi, bias, dan selang kepercayaan dari distribusi prior konjugat Gamma menghasilkan nilai yang lebih kecil daripada distribusi prior non-konjugat. Sehingga, dapat disimpulkan bahwa penduga parameter Bayes dengan distribusi prior konjugat lebih baik dari distribusi prior non-konjugat. Selain itu, penduga parameter menggunakan ukuran sampel yang besar akan menghasilkan nilai posterior variansi, bias, dan selang kepercayaan yang lebih kecil pada jenis prior yang sama. Kajian ini juga membuktikan bahwa dengan penggunaan ukuran sampel yang lebih besar maka nilai dugaan yang dihasilkan akan lebih mendekati nilai yang sebenarnya.

## Daftar Pustaka

- [1] Y. Feng, Y. Chen, and X. He, "Bayesian quantile regression with approximate likelihood," *Bernoulli*, vol. 21, no. 2, pp. 832–850, 2015.
- [2] C. Muharisa, F. Yanuar, and D. Devianto, "Simulation Study of the Using of Bayesian Quantile Regression in Non-normal Error," *Cauchy - J. Mat. Murni dan Apl.*, vol. 5, no. 3, pp. 121–126, 2018.
- [3] F. Yanuar, K. Ibrahim, and A. Aziz Jemain, "Bayesian structural equation modeling for the health index," *J. Appl. Stat.*, vol. 40, no. 6, pp. 1254–1269, 2013.
- [4] F. Yanuar, A. Zetra, C. Muharisa, and D. Devianto, "Bayesian Quantile Regression Method to Construct the Low Birth Weight Model," *J. Phys. Confrence Ser.*, vol. 1245, pp. 1–7, 2019.
- [5] D.-H. Kim, S.-G. Kang, and W.-D. Lee, "Noninformative priors for Pareto distribution," *J. Korean Data Inf. Sci. Soc.*, vol. 20, no. 6, pp. 1213–1223, 2009.
- [6] F. Yanuar, H. Yozza, and R. V. Rescha, "Comparison of Two Priors in Bayesian Estimation for Parameter of Weibull Distribution," *Sci. Technol. Indones.*, vol. 4, no. 3, pp. 82–87, 2019.
- [7] M. Maswadah, A. M. Seham, and M. Ahsanullah, "Bayesian Inference on the Generalized Gamma Distribution Based on Generalized Order Statistics," *J. Stat. Theory Appl.*, vol. 12, no. 4, pp. 356–377, 2013.
- [8] S. A. Thamrin, M. N. Zoraida, A. K. Jaya, and Ansariadi, "Performance of bayesian using conjugate prior estimator for Weibull right censored survival data," *Asian J. Sci. Res.*, vol. 11, no. 3, pp. 376–382, 2018.

- [9] T. G. Ieren and P. E. Oguntunde, "A Comparison between Maximum Likelihood and Bayesian Estimation Methods for a Shape Parameter of the Weibull-Exponential Distribution," *Asian J. Probability Stat.*, vol. 1, no. 1, pp. 1–12, 2018.
- [10] T. G. Ieren and P. Oguntunde, "A Comparison between Maximum Likelihood and Bayesian Estimation Methods for a Shape Parameter of the Weibull-Exponential Distribution," *Asian J. Probability Stat.*, vol. 1, no. 1, pp. 1–12, 2018.