TAKSIRAN YANG LEBIH EFISIEN UNTUK PARAMETER PADA DISTRUSI WEIBULL

Erma Kusuma Wati[1], Sigit Sugiarto[2], Bustami[2]

emakusumawati70@yahoo.co.id

[1]Mahasiswa Program S1 Matematika
[2]Dosen Matematika, Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Riau
Kampus Binawidya Pekanbaru (28293), Indonesia

ABSTRACT

In this paper the authors use the method of maximum likelihood and Bayesian methods to estimate parameters of the Weibull distribution. Prior used for the Bayesian method is Jeffery prior information. The loss function used is the quadratic loss function. Then the computing was done by comparing of the maximum likelihood method and Bayesian method to find the smallest Mean Square Error. The comparison among the discussed methods is done by simulation methods. The estimated parameters of Weibull distribution obtained from the maximum likelihood is the best compared to Bayes using Jeffery prior.

Keywords: *Method of maximum likelihood, method of Bayesian, Jeffery prior information.*

1. PENDAHULUAN

Statistika inferensi berkaitan dengan pengambilan kesimpulan tentang parameter populasi yang didasarkan pada informasi data sampel dari populasi yang menjadi perhatian. Parameter yang menjadi perhatian dapat berupa rata-rata, variansi dan parameter lainya. Taksiran untuk parameter ada dua jenis yaitu, taksiran titik dan taksiran interval. Taksiran ini dapat diperoleh dengan menggunakan dua pendekatan yaitu, pendekatan klasik dan pendekatan Bayesian. Pendekatan Bayesian pada dasarnya berbeda dengan pendekatan klasik. Pada pendekatan klasik, model data sampel dinyatakan dalam bentuk fungsi densitas yang distribusinya tergantung pada parameter yang nilainya tidak diketahui. Salah satu teknik yang digunakan dalam pendekatan klasik adalah metode maksimum likelihood. Pada pendekatan Bayesian, parameter pada pendekatan klasik dianggap sebagai suatu variabel random yang distribusinya disebut distribusi prior. Distribusi prior terdiri dari dua yaitu prior informatif dan prior non informative. Pada artikel ini penulis menggunakan prior non informatif yaitu informasi Jeffery prior[3]. Dengan menggunakan Teorema Bayes, distribusi prior yang telah dimiliki diperbaharui dengan mengkombinasikannya dengan fungsi likelihood untuk membentuk distribusi posterior [5].

Penaksir ada dua macam, pertama penaksir tak bias yaitu apabila rata-rata perkiraan sama dengan parameter sebenarnya dan yang kedua penaksir bias yaitu apabila rata-rata perkiraan tidak sama dengan parameter sebenarnya. Penaksir tak bias terbaik adalah penaksir yang mempunyai variansi minimum. Penaksir bias terbaik adalah penaksir yang mempunyai Mean Square Error minimum[2], untuk selanjutnya penulis menggunakan singkatan MSE.

Pada kertas kerja ini dibahas taksiran yang lebih efisien untuk parameter pada distribusi Weibull. Distribusi Weibull biasanya digunakan dalam pembahasan data uji hidup. Data uji hidup merupakan topik dalam bebagai bidang biomedik dan industri [1]. Fungsi densitas distribusi Weibull [1]

$$f(x;\theta,p)=\frac{p}{\theta}x^{p-1}\exp\left(-\frac{x^p}{\theta}\right),\qquad \theta,p,x>0 \tag{1}$$

Dengan θ adalah parameter skala dan p adalah parameter bentuk, x adalah variabel random. Parameter p adalah parameter bentuk yang diketahui yang bernilai konstan. Sehingga parameter yang akan ditaksir adalah parameter θ. Fungsi densitas Weibull pada persamaan (1) memiliki ekspektasi θ dan variansi θ^2. Dalam menaksir parameter dilakukan dengan menggunakan dua metode, yaitu metode maksimum likelihood dan metode Bayes menggunakan informasi Jeffery prior. Selanjutnya akan ditentukan bias atau tak bias dari kedua penaksir tersebut. Setelah itu, menentukan taksiran yang lebih efisien. Karena perhitungan secara analitik sangat rumit, maka dalam menentukan taksiran yang lebih efisien dilakukan dengan bantuan simulasi numerik.

2. PENAKSIR MAKSIMUM LIKELIHOOD

Misalkan $X_1, X_2,..., X_n$ adalah sampel random berukuran *n* yang berasal dari distribusi Weibull dengan fungsi densitas pada persamaan (1). Berdasarkan sampel random $X_1, X_2,..., X_n$ ditaksir parameter dengan menggunakan metode maksimum likelihood. Ambil vektor $\boldsymbol{x}=(x_1,x_2,\dots,x_n)$, asumsikan $x_1, x_2,..., x_n$ adalah saling bebas dalam distribusi Weibull. Maka fungsi likelihood adalah

$$L(\boldsymbol{x};\theta,p)=\prod_{i=1}^{n} f(x_i;\theta,p)$$

$$L(\boldsymbol{x};\theta,p)=\left(\frac{p}{\theta}\right)^n\prod_{i=1}^{n} x_i^{\,p-1}\exp\left(-\frac{\sum_{i=1}^{n} x_i^{\,p}}{\theta}\right). \tag{2}$$

Diasumsikan parameter bentuk p diketahui, maka penaksir maksimum likelihood dari parameter θ adalah dengan mencari turunan pertama $\ln L(\hat{\theta};x_1,x_2,..,x_n)$ pada persamaan (2) terhadap θ dan samakan dengan nol,

$$\frac{d(\ln(L(x_1,x_2,...,x_n;\theta)))}{d\theta}=0.$$

Maka

$$\frac{d}{d\theta}\ln[L(\boldsymbol{x};\theta,p)] = -\frac{n}{\theta}+\frac{1}{\theta^2}\sum_{i=1}^{n} x_i^{\ p} = 0 \Rightarrow \theta = \frac{\sum_{i=1}^{n} x_i^{\ p}}{n}.$$

Sehingga penaksir untuk parameter θ menggunakan metode maksimum likelihood adalah

$$\hat{\theta}_{MLE} = \frac{\sum_{i=1}^{n} x_i^{\ p}}{n}. \qquad (3)$$

Dan $E(\hat{\theta}_{MLE}) = \theta,$ maka penaksir maksimum likelihood bersifat tak bias.

3. PENAKSIR BAYES

Berikut ini diberikan langkah-langkah untuk mendapatkan penaksir Bayes [1] untuk fungsi densitas distribusi Weibull yang ditunjukan pada persamaan (1) adalah sebagai berikut

1. Inputkan sejumlah *n* ukuran sampel untuk menguji data dan waktu hidup dari sampel random dicatat sebagai fungsi densitas peluang $f(x;\theta,p),$ dimana θ adalah parameter skala dan p adalah parameter bentuk yang diketahui.

2. Menentukan fungsi densitas peluang waktu hidup yang diasumsikan sebagai fungsi densitas peluang bersyarat $f(x|\theta)$ dan prior $g(\theta)$ yang digunakan adalah informasi Jeffery prior $g(\theta) = k\sqrt{\mathrm{I}(\theta)},\, k$ adalah konstanta.

3. Menentukan fungsi densitas peluang gabungan yaitu dengan menggabungkan fungsi likelihood dengan informasi Jeffery prior yang dirumuskan dengan

   $$H(x_1, x_2,..., x_n;\theta, p) = L(\boldsymbol{x};\theta, p)g(\theta).$$

4. Menentukan fungsi densitas peluang marginal dari $(x_1, x_2,..., x_n;\theta)$ yaitu integral dari fungsi densitas peluang gabungan yang dirumuskan dengan

   $$P(x_1, x_2,..., x_n) = \int_0^\infty H(x_1, x_2,..., x_n;\theta, p)d\theta.$$

5. Menentukan fungsi densitas peluang bersyarat atau distribusi posterior dari θ yaitu

   $$\pi(\theta|x) = \frac{H(x_1, x_2,..., x_n;\theta, p)}{P(x_1, x_2,..., x_n)}.$$

6. Menentukan Penaksir Bayes dari θ dengan menggunakan fungsi kerugian kuadratik, dengan menggunakan teorema berikut ini

   Teorema 1 [2: hal. 324]

   Penaksir Bayes, $\hat{\theta}_B$ dari θ pada fungsi kerugian kuadratik $\ell(\hat{\theta};\theta) = (\hat{\theta}-\theta)^2$ adalah ekspektasi bersyarat dari θ yang relatif terhadap posterior

$$\hat{\theta}_B = E[\theta|x] = \int_0^\infty \theta\pi(\theta|x)d\theta.$$

Untuk menentukan penaksir Bayes, diperlukan distribusi posterior atau distribusi peluang bersyarat. Distribusi posterior dirumuskan dengan

$$\pi(\theta|x) = \frac{H(x_1, x_2,..., x_n;\theta, p)}{P(x_1, x_2,..., x_n)}. \tag{4}$$

Dimana $H(x_1, x_2,..., x_n;\theta, p)$ adalah fungsi densitas peluang gabungan dan $P(x_1, x_2,..., x_n)$ adalah fungsi densitas peluang marginal. Fungsi densitas peluang gabungan diperoleh dengan menggabungan fungsi likelihood $L(\boldsymbol{x};\theta, p)$ dengan informasi Jeffery prior $g(\theta)$. Maka perumusan fungsi densitas peluang gabungan adalah

$$H(x_1, x_2,..., x_n;\theta, p) = L(\boldsymbol{x};\theta, p)g(\theta). \tag{5}$$

Di dalam [4], informasi Jeffery prior didefinisikan dengan

$$g(\theta) = k\sqrt{\mathrm{I}(\theta)},\ k \text{ adalah konstanta.} \tag{6}$$

$\mathrm{I}(\theta)$ adalah informasi Fisher. Informasi Fisher diperoleh dengan mencari

$$\mathrm{I}(\theta) = -E_\theta\left[\frac{d^2 \ln L(\boldsymbol{x};\theta)}{d\theta^2}\right],$$

maka

$$\mathrm{I}(\theta) = \frac{n}{\theta^2}. \tag{7}$$

Dengan mensubtitusikan persamaan (7) kepersamaan (6), maka diperoleh informasi Jeffery prior

$$g(\theta) = k\frac{\sqrt{n}}{\theta}. \tag{8}$$

Dengan mensubtitusikan persamaan (2) dan persamaan (8) kepersamaan (5), maka diperoleh fungsi densitas peluang gabungan

$$H(x_1, x_2,..., x_n;\theta, p) = \left(\frac{p}{\theta}\right)^n \prod_{i=1}^n x_i^{\ p-1} \exp\left(-\frac{\sum_{i=1}^n x_i^{\ p}}{\theta}\right) k\frac{\sqrt{n}}{\theta}$$

$$H(x_1, x_2,..., x_n;\theta, p) = \frac{kp^n\sqrt{n}}{\theta^{n+1}} \exp\left[\ln \prod_{i=1}^n x_i^{\ p-1} - \frac{\sum_{i=1}^n x_i^{\ p}}{\theta}\right]. \tag{9}$$

Fungsi densitas peluang marginal dari θ pada data $(x_1, x_2,..., x_n)$ adalah

$$P(x_1, x_2, ..., x_n) = \int_0^\infty H(x_1, x_2, ..., x_n; \theta, p) d\theta. \tag{10}$$

Dengan mensubtitusikan persamaan (9) kepersamaan (10), maka

$$P(x_1, x_2, ..., x_n) = kp^n \sqrt{n} \exp\left[(p-1) \sum_{i=1}^{n} \ln x_i \right] \int_0^\infty \frac{1}{\theta^{n+1}} \exp\left(-\frac{\sum_{i=1}^{n} x_i^{\,p}}{\theta} \right) d\theta. \tag{11}$$

Dengan menyelesaikan integral $\int_0^\infty \frac{1}{\theta^{n+1}} \exp\left(-\frac{\sum_{i=1}^{n} x_i^{\,p}}{\theta} \right) d\theta$ pada persamaan (11), diperoleh

$$\int_0^\infty \frac{1}{\theta^{n+1}} \exp\left(-\frac{\sum_{i=1}^{n} x_i^{\,p}}{\theta} \right) d\theta = \frac{(n-1)!}{(\sum_{i=1}^{n} x_i^{\,p})^n}. \tag{12}$$

Dengan mensubtitusikan persamaan (12) kepersamaan (11), maka diperoleh fungsi densitas peluang marginal berikut ini

$$P(x_1, x_2, ..., x_n) = \frac{kp^n \sqrt{n}(n-1)!}{\left(\sum_{i=1}^{n} x_i^{\,p} \right)^n} \exp\left[(p-1) \sum_{i=1}^{n} \ln x_i \right]. \tag{13}$$

Selanjutnya akan tentukan fungsi densitas peluang bersyarat atau distribusi posteriornya. Dengan mensubtitusikan persamaan (9) dan persamaan (13) kepersamaan (4), maka diperoleh distribusi posterior berikut ini

$$\pi(\theta | x) = \frac{\dfrac{kp^n \sqrt{n}}{\theta^{n+1}} \exp\left[(p-1) \sum_{i=1}^{n} \ln x_i - \dfrac{\sum_{i=1}^{n} x_i^{\,p}}{\theta} \right]}{\dfrac{kp^n \sqrt{n}(n-1)!}{\left(\sum_{i=1}^{n} x_i^{\,p} \right)^n} \exp\left[(p-1) \sum_{i=1}^{n} \ln x_i \right]}. \tag{14}$$

Persamaan (14) dapat disederhanakan menjadi

$$\pi(\theta | x) = \frac{kp^n \sqrt{n} \exp\left[(p-1) \sum_{i=1}^{n} \ln x_i \right] \exp\left(-\dfrac{\sum_{i=1}^{n} x_i^{\,p}}{\theta} \right) \left(\sum_{i=1}^{n} x_i^{\,p} \right)^n}{\theta^{n+1} kp^n \sqrt{n}(n-1)! \exp\left[(p-1) \sum_{i=1}^{n} \ln x_i \right]}$$

$$\pi(\theta|x)=\frac{\left(\sum_{i=1}^{n}x_i^{\ p}\right)^n}{\theta^{n+1}(n-1)!}\exp\left(-\frac{\sum_{i=1}^{n}x_i^{\ p}}{\theta}\right). \tag{15}$$

Persamaan (15) adalah distribusi posterior untuk parameter θ. Untuk mendapatkan taksiran Bayes dengan menggunakan fungsi kerugian kuadratik, maka akan ditentukan ekspektasi bersyarat dari θ yang relatif terhadap posterior sesuai dengan teorema 1. Berdasarkan teorema 1, maka diperoleh penaksir Bayes menggunakan iformasi Jeffery prior berikut ini

$$\hat{\theta}_{BJ}=E[\theta|x]=\int_0^\infty \theta\pi(\theta|x)d\theta$$

$$\hat{\theta}_{BJ}=\int_0^\infty \theta\frac{\left(\sum_{i=1}^{n}x_i^{\ p}\right)^n}{\theta^{n+1}(n-1)!}\exp\left(-\frac{\sum_{i=1}^{n}x_i^{\ p}}{\theta}\right)d\theta$$

$$\hat{\theta}_{BJ}=\frac{\left(\sum_{i=1}^{n}x_i^{\ p}\right)^n}{(n-1)!}\int_0^\infty \theta^{-n}\exp\left(-\frac{\sum_{i=1}^{n}x_i^{\ p}}{\theta}\right)d\theta. \tag{16}$$

Dengan menyelesaikan integral $\int_0^\infty \theta^{-n}\exp\left(-\frac{\sum_{i=1}^{n}x_i^{\ p}}{\theta}\right)d\theta$ pada persamaan (16), diperoleh

$$\int_0^\infty \theta^{-n}\exp\left(-\frac{\sum_{i=1}^{n}x_i^{\ p}}{\theta}\right)d\theta=\frac{\Gamma(n-1)}{\left(\sum_{i=1}^{n}x_i^{\ p}\right)^{n-1}}. \tag{17}$$

Dengan mensubtitusikan persamaan (17) kepersamaan (16), maka didapat

$$\hat{\theta}_{BJ}=\frac{\left(\sum_{i=1}^{n}x_i^{\ p}\right)}{(n-1)}. \tag{18}$$

Persamaan (18) adalah penaksir Bayes menggunakan informasi Jeffery prior dan $E(\hat{\theta}_{BJ})=\frac{n\theta}{(n-1)}$, maka penaksir Bayes menggunakan informasi Jeffery prior bersifat bias.

4. MSE

Untuk penaksir tak bias, maka akan diperoleh MSE sama dengan variansi[2]. Sedangkan untuk penaksir bias akan diperoleh

$$\text{MSE}(\hat{\theta}) = \text{Var}(\hat{\theta}) + [b(\hat{\theta})]^2, \tag{19}$$

dengan syarat jika $\hat{\theta}$ merupakan penaksir dari θ. Perhitungan MSE dapat dilakukan secara analitik tapi perhitungannya sangat rumit, maka dalam perhitunganya diperlukan bantuan simulasi numerik. Pada simulasi digunakan perumusan MSE berikut ini

$$\text{MSE}(\hat{\theta}) = \frac{\sum_{i=1}^{1000} (\hat{\theta}_i - \theta)^2}{R} \text{ [1: Hal.103].} \tag{20}$$

A. MSE dari Penaksir Maksimum Likelihood

Karena penaksir maksimum likelihood pada persamaan (4) bersifat tak bias, maka

$$\begin{aligned}
\text{MSE}(\hat{\theta}_{MLE}) &= \text{Var}(\hat{\theta}_{MLE}) \\
&= \text{Var}\left(\sum_{i=1}^{n} \frac{X_i^{\ p}}{n} \right) \\
&= \frac{\theta^2}{n}.
\end{aligned}$$

B. MSE dari Penaksir Bayes Menggunakan Informasi Jeffery Prior

Karena penaksir pada persamaan (18) bersifat bias, berdasarkan persamaan (19) maka MSE dari penaksir Bayes menggunakan informasi Jeffery prior adalah

$$\begin{aligned}
\text{MSE}(\hat{\theta}_{BJ}) &= \text{Var}(\hat{\theta}_{BJ}) + (b(\hat{\theta}_{BJ}))^2 \\
&= \frac{1}{(n-1)^2} \text{Var}\left(\sum_{i=1}^{n} X_i^{\ p} \right) + \frac{\theta^2}{(n-1)^2}
\end{aligned}$$

$$\text{MSE}(\hat{\theta}_{BJ}) = \frac{1}{(n-1)^2} [n\theta^2] + \frac{\theta^2}{(n-1)^2}$$

$$\text{MSE}(\hat{\theta}_{BJ}) = \frac{\theta^2 (n+1)}{(n-1)^2}.$$

5. Studi Simulasi dan Pembahasan

Dalam simulasi ini akan ditunjukan nilai MSE antara metode maksimum likelihood dan metode Bayes menggunakan informasi Jeffery prior. Dalam perhitungan simulasi program yang digunakan adalah program Matlab 7.0.1. Dengan mengambil ukuran sampel yang besar dan ukuran sampel yang kecil yaitu $n = 25, 50, 75, 100, 125, 150, 175, 200$, nilai parameter skala (θ) yang digunakan yaitu $\theta = 0.5, 1.5$ dan nilai parameter bentuk (p) yaitu $p = 0.8, 1.2$ jumlah pengulangan (R) adalah R=1000 dan dengan menggunakan perumusan MSE pada persamaan (20), maka akan diperoleh nilai MSE dari penaksir maksimum likelihood dan penaksir Bayes menggunakan informasi Jeffery prior. Dengan menghitung nilai MSE dari kedua metode yang telah ditaksir, dapat dilihat metode mana yang mempunyai nilai MSE yang lebih kecil. Hasil dari studi simulasi dirangkum dan ditabulasikan pada Tabel dan Grafik.

Tabel 1 Nilai MSE dari taksiran parameter untuk $\theta = 0.5$ dan $p = 0.8$

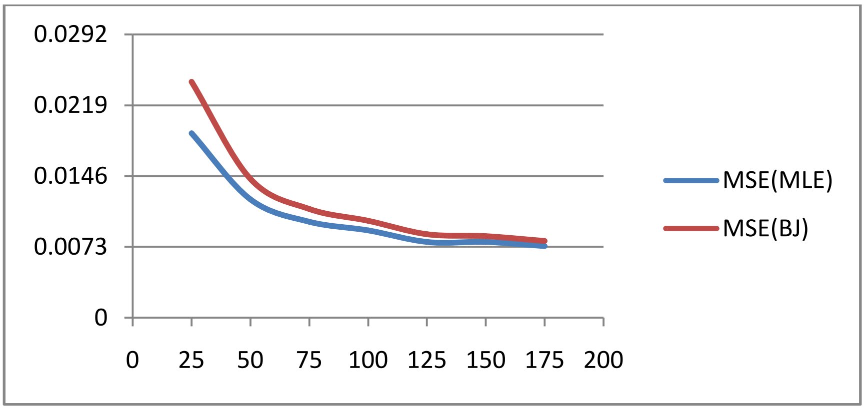
n	MSE(MLE)	MSE(BJ)	SELISIH NILAI MSE
25	0.0190	0.0243	0.0053
50	0.0122	0.0143	0.0021
75	0.0099	0.0112	0.0013
100	0.0090	0.0100	0.001
125	0.0078	0.0086	0.0008
150	0.0078	0.0084	0.0006
175	0.0074	0.0079	0.0005
200	0.0073	0.0078	0.0005

Tabel 2 Nilai MSE dari taksiran parameter untuk $\theta = 1.5$ dan $p = 1.2$

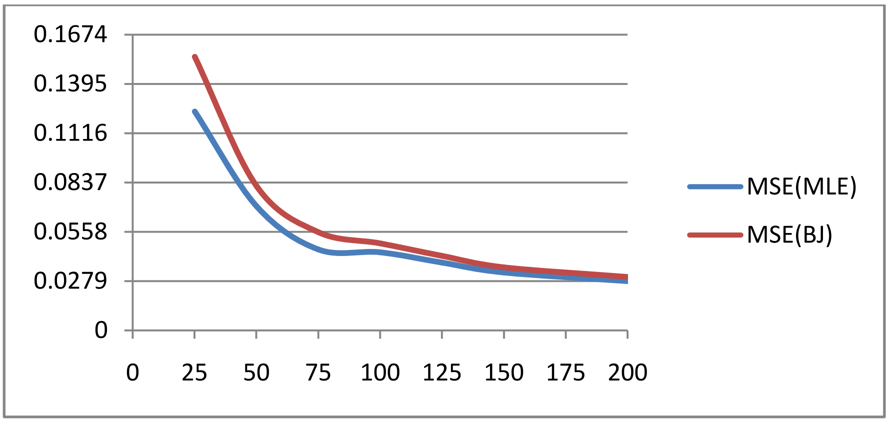
n	MSE(MLE)	MSE(BJ)	SELISIH NILAI MSE
25	0.1238	0.1549	0.0311
50	0.0702	0.0817	0.0115
75	0.0458	0.0555	0.0097
100	0.0442	0.0492	0.005
125	0.0383	0.0422	0.0039
150	0.0327	0.0357	0.003
175	0.0303	0.0329	0.0026
200	0.0279	0.0302	0.0023

Dapat dilihat pada Tabel 1 dan Tabel 2, untuk semua ukuran sampel yang digunakan metode maksimum likelihood mempunyai nilai MSE yang lebih kecil dibandingkan dengan metode Bayes menggunakan informasi Jeffery prior. Sehingga metode maksimum likelihood merupakan penaksir yang lebih baik dibandingakan penaksir Bayes menggunakan informasi Jeffery prior. Dari tabel 1 dan tabel 2 dapat disimpulkan bahwa MSE dan selisih nilai MSE dari penaksir maksimum likelihood dan penaksir Bayes menggunakan informasi Jeffery prior menurun dengan meningkatnya *n* ukuran sampel. Semakin besar *n* ukuran sampel maka taksiran Bayes menggunakan informasi Jeffery prior yang diperoleh semakin dekat dengan nilai taksiran maksimum likelihood. Untuk lebih jelasnya, taksiran nilai MSE dari metode maksimum likelihood dan metode Bayes menggunakan informasi Jeffery prior dapat dilihat pada Gambar 1 dan Gambar 2

Gambar 1 Grafik MSE untuk taksiran parameter dengan $\theta = 0.5$ dan $p = 0.8$



Gambar 2 Grafik MSE untuk taksiran parameter dengan $\theta = 1.5$ dan $p = 1.2$



Dapat dilihat dari Gambar 1 dan Gambar 2, grafik MSE dari metode maksimum likelihood lebih rendah dibandingkan grafik metode Bayes menggunakan informasi Jeffery prior. Ketika ukuran sampel yang diambil 100 maka keadaan grafik dari kedua metode mulai menurun hingga pada ukuran sampel 200. Dari Gambar 1 dan Gambar 2 dapat diimpulkan bahwa semakin besar ukuran sampel yang diambil maka nilai MSE dari kedua metode tersebut juga semakin kecil, sehingga kedua grafik terlihat semakin berhimpit. Semakin besar *n* ukuran sampel maka taksiran Bayes menggunakan informasi Jeffery prior yang diperoleh semakin dekat dengan nilai taksiran maksimum likelihood.

DAFTAR PUSTAKA

[1] Ahmed, A.O.M, H.S. Al-Kutubi & N.A. Ibrahim. 2010. " Comparison of the Bayesian and Maximum Likelihood Estimation for Weibull Distribution ", *journal of mathematics and statistics*, **6** (2): 100-104.

[2] Bain, L.J. & Engelhardt, Max. 1993. *Introduction to Probability and Mathematical Statistics*. *Second Edition*. Duxbury Press. Belmont, California.

[3] Berger, J.O. 1985. *Statistical Decision Theory and Bayesian Analisis*. *Second Edition*. Springer-verlag, Indiana.

[4] Lee, P.M. 1997. *Bayesian Statistics*: An Introduction. *Second Edition*. Arnold. Euston Road, Landon.

[5] Ramachandran, K.M. & Tsokos, C.P. 2009. *Mathematical Statistics with Aplications*. British Library Catalouging.