SIMULASI PENAKSIRAN PARAMETER DISTRIBUSI WEIBULL CAMPURAN UNTUK DATA SURVIVAL HETEROGEN DENGAN PENDEKATAN BAYESIAN

IndoMS Journal on Statistics
Vol. 2, No. 2 (2014), Page 37-46

SIMULASI PENAKSIRAN PARAMETER *DISTRIBUSI* WEIBULL CAMPURAN UNTUK DATA SURVIVAL HETEROGEN DENGAN PENDEKATAN BAYESIAN

Sri Astuti Thamrin[1], Armin Lawi[2] dan Rifaatul Mahmudah[1]
[1]Statistika, Jurusan Matematika Universitas Hasanuddin, Makassar
[2]Ilmu Komputer, Jurusan Matematika Universitas Hasanuddin, Makassar
Email: tuti@unhas.ac.id; armin@unhas.ac.id

Abstract

A Weibull distribution is one of the most popular parametric model that can be used in survival context. However, where observations are taken from a possibly heterogeneous population, the simple Weibull model may not always be appropriate and mixture models may be considered. In this paper, we use the Bayesian approach to fit the Weibull mixture model with unknown number of components with the right censored heterogeneous survival data. The birth-death Markov Chain Monte Carlo (MCMC) algorithm is used for estimating the Weibull mixture model parameters. Based on the simulation, the posterior mean of all the Weibull mixture model estimates are still reasonably accurate.

Kata kunci: Bayesian, Mixtures, MCMC, Survival Analysis, Weibull.

Abstrak

Distribusi Weibull merupakan salah satu dari model parametrik yang sangat popular serta dapat digunakan dalam konteks analisis survival. Namun, ketika pengamatan dilakukan pada populasi yang kemungkinannya heterogen, model Weibull yang sederhana tidak memadai dan karenanya model campuran (*mixture*) menjadi layak untuk dipertimbangkan. Pada tulisan ini, pendekatan Bayesian digunakan untuk menyelesaikan model campuran Weibull (*mixture Weibull*) dengan sejumlah komponen yang tidak diketahui dan dengan data survival heterogen yang tersensor kanan. Algoritma kematian dan kelahiran MCMC (*birth-death Markov Chain Monte Carlo*) digunakan untuk mengestimasi parameter model campuran Weibull. Berdasarkan hasil simulasi, rata-rata posterior dari semua estimasi model campuran Weibull terlihat akurat.

Kata kunci: Analisis survival, Bayesian, Markov Chain Monte Carlo, Mixture, Weibull

1. Pendahuluan

Analisis kelangsungan hidup (*survival*) merupakan analisis yang digunakan untuk menganalisis data kelangsungan waktu hidup (*survival time*). Pada awalnya, analisis kelangsungan hidup sering dilakukan dengan menggunakan metode non parametrik. Namun, seiring dengan perkembangan komputasi terutama disebabkan munculnya algoritma komputasi baru, maka penerapan analisis kelangsungan hidup kini menggunakan metode Bayesian.

Distribusi Weibull merupakan salah satu model parametrik yang paling populer digunakan dalam konteks kelangsungan hidup (Dodson, 1994). Namun, penggunaan model Weibull sederhana tidak selalu tepat digunakan untuk populasi yang heterogen, Oleh karena itu, model Weibull campuran (*mixture Weibull*) perlu dipertimbangkan untuk digunakan dalam populasi heterogen. Dalam prakteknya, penerapan jumlah komponen Weibull yang diketahui seperti dua komponen atau lebih belum cukup mewakili sepenuhnya bentuk distribusi kelangsungan hidup campuran. Oleh karena itu, pada kasus itu distribusi Weibull campuran hanya terbatas pada jumlah komponen tetap (Tsionas, 2002).

Belum tepatnya penggunaan distribusi Weibull sederhana untuk populasi yang heterogen, memungkinkan pemodelan waktu hidup dapat dilakukan dengan pendekatan Bayesian untuk model Weibull campuran untuk mengetahui banyaknya komponen (*component*) yang tidak diketahui. Dalam estimasi parameter dengan menggunakan metode Bayes diperlukan distribusi posterior. Distribusi posterior ini dibentuk dari distribusi prior. Nilai estimasi parameter diperoleh dengan simulasi pengambilan sampel parameter dari distribusi posterior dengan menggunakan metode *Markov Chain Monte Carlo* (MCMC) (Gilks dan Wild, 1992). Salah satu algoritma dari metode MCMC yang digunakan adalah algoritma *Gibbs sampling*. Algoritma *Gibbs sampling* dapat diterapkan apabila distribusi probabilitas bersama (*joint probility distribution*) dari parameternya tidak diketahui, tetapi distribusi bersyarat (*conditional distribution*) dari tiap-tiap variabel diketahui (Walsh, 2004).

Penelitian tentang distribusi Weibull campuran telah banyak dilakukan. Dalam konteks Bayesian, Chen dkk (1985) telah menggunakan model campuran dua komponen untuk menganalisis data survival kanker yang awalnya telah dilakukan oleh Berkson dan Gage (1952). Quiang (1994) juga telah menggunakan model Weibull campuran dalam konteks penyakit kanker paru-paru. Namun, penelitian yang telah dilakukan tersebut, hanya dapat mengatasi masalah untuk bentuk distribusi survival dengan hanya dua komponen seperti yang telah dilakukan oleh Tsionas (2002).

Dalam tulisan ini akan dilakukan simulasi untuk pemodelan waktu kelangsungan hidup heterogen melalui pendekatan Bayesian pada campuran model survival Weibull dengan jumlah komponen yang tidak diketahui. Algoritma birth-death MCMC akan digunakan dalam mengestimasi berapa sebenarnya jumlah komponen yang ada yang tidak diketahui dalam model Weibull campuran itu.

2. Metode

Distribusi Weibull merupakan distribusi yang menggambarkan waktu hidup dari makhluk hidup. Distribusi Weibull banyak digunakan dalam analisis data kelangsungan hidup. Distribusi ini diperkenalkan oleh fisikawan Swedia Waloddi Weibull pada tahun 1939 (Walpole dan Myers, 1995).

Fungsi kepadatan peluang peubah acak kontinu s berdistribusi Weibull, dengan parameter θ dan γ:

$$W(s|\theta,\gamma) = \theta\gamma s^{\gamma-1}\exp\{-\theta s^{\gamma}\},\ \theta > 0;\ \gamma > 0, \tag{1}$$

dimana θ adalah parameter skala (*scale parameter*) dan γ adalah parameter bentuk (*shape parameter*).

Fungsi kepadatan peluang Weibull campuran komponen k (Marin, et al, 2005), didefinisikan:

$$f(s|k,\omega,\theta,\gamma) = \sum_{j=1}^{k}\omega_j W\left(s|\theta_j,\gamma_j\right), \tag{2}$$

dimana $\theta = (\theta_1,\dots,\theta_k)$, $\gamma = (\gamma_1,\dots,\gamma_k)$ adalah parameter distribusi Weibull, dan $\omega = (\omega_1,\dots,\omega_k)$ vektor nonnegatif dengan jumlah bobot 1.

Bila data waktu hidup suatu subyek yang mengalami sensor kanan untuk n pasien adalah $X = (x_1,\dots,x_n)$ dimana $x_i = (s_i,\delta_i)$, s_i adalah waktu pengamatan dan δ_i adalah fungsi indicator maka:

$$\delta_i = \begin{cases} 1, & \text{jika waktu hidup (lifetimes) tidak tersensor, yaitu } S_i = s_i \\ 0, & \text{jika waktu hidup (lifetimes) tersensor, yaitu } S_i > s_i \end{cases},$$

Fungsi likelihood dari data tersebut adalah:

$$l(k,\omega,\theta,\gamma|\text{data}) \propto \prod_{i=1}^{n}\sum_{j=1}^{k}\omega_j(\theta_j\gamma_j)^{\delta_i}{s_i}^{(\gamma-1)\delta_i}\exp\left\{-\theta_j s^{\gamma_j}\right\} \tag{3}$$

Untuk mencari inferensi posterior digunakan algoritma *Gibbs sampling* yang diperkenalkan oleh Diebolt dan Robert (1994). Langkahnya adalah pertama, diperkenalkan variabel indikator Z_i, untuk $i = 1,\dots,n$, yang didefinisikan dari komponen campuran ke- i pengamatan yang telah dihasilkan dengan cara:

$$P(Z_i = j|k,\omega) = \omega_j, \qquad \text{dan } S_i|k,Z_i = j,\gamma,\theta \sim W(.|\theta_j,\gamma_j)$$

Bersyarat pada indikatornya, maka fungsi likelihood adalah:

$$l(k,\omega,\theta,\gamma,z|k,\text{data}) \propto \prod_{j=1}^{k}(\theta_j,\gamma_j)^{\tilde{n}_j}\exp\left(\gamma_j\sum_{i:Z_i=j}^{n}\log S_i - \theta_j\sum_{i:Z_i=j}^{n}s^{\gamma_{z_i}}\right) \tag{4}$$

dimana, $\tilde{n}_j = \#\{i: Z_i = j \ dan\ \delta_i = 1\}$, untuk $j = 1,\dots,k$, merupakan jumlah data yang tidak tersensor untuk komponen j campuran. Distribusi posterior dapat dihitung dengan mengkombinasikan likelihood dengan distribusi prior yaitu:

(i) $\omega\ |k,z,\theta,\gamma,\text{data} \sim Dirichlet\ (\phi_1 + n_1,\dots,\phi_k + n_k)$

(ii) $\theta_j\ |k,z,\gamma,\text{data} \sim Gamma\ (\tilde{n}_j + \gamma_\theta,\beta_\theta + \sum_{\{i:z_i=j\}}{s_i}^{\gamma_j})$

Dengan bentuk fungsi kepadatan peluang sebagai berikut :

$$f(\theta_j) = \frac{\left(\beta_\theta + \Sigma_{\{i:z_i=j\}} s_i^{\gamma_j}\right)^{\tilde{n}_j+\gamma_\theta} \theta^{(\tilde{n}_j+\gamma_\theta)-1} e^{-\theta\left(\beta_\theta + \Sigma_{\{i:z_i=j\}} s_i^{\gamma_j}\right)}}{\Gamma(\tilde{n}_j+\gamma_\theta)} \quad (5)$$

untuk $j = 1,\dots,k$ dan $\tilde{n}_j = \#\{i: z_i = j \text{ dan } \delta_i = 1\}$

(iii) $f(\gamma_j \mid k, z, \theta, \text{data}) \propto g(\gamma_j)$ dimana

$$g(\gamma_j) = \gamma_j^{\tilde{n}_j+\gamma_\theta-1} \exp\{-\gamma_j(\beta_\theta - \sum_{\{i:z_i=j\}}' \delta_i \log s_i) - \theta_j \sum_{\{i:z_i=j\}}' s_i^{\gamma_j}\} \quad (6)$$

(iv) $Pr(Z_i = j \mid k, \theta, \gamma, \omega, \text{data}) \propto \omega_j(\theta_j \gamma_j s_i^{(\gamma_j-1)})^{\delta_i} \exp\{-\theta_j s_i^{\gamma_j}\}$ (7)

untuk $i = 1,\dots,n$

Dengan diberikannya distribusi bersyarat, maka dapat didefinisikan dengan menggunakan algoritma *Gibbs sampling* untuk mensimulasikan sampel dari distribusi posterior bersama dengan prosedur sebagai berikut:

1. $t = 0$, atur nilai awal $\omega^0, \theta^0, \gamma^0$
2. $z_i^{t+1} \sim z_i \mid k, \omega^t, \theta^t, \gamma^t$, data, untuk $i = 1, 2, \dots, n$
3. $\omega^{t+1} \sim \omega \mid k, z^{t+1}, \theta^t, \gamma^t$, data
4. $\theta_j^{t+1} \sim \theta_j \mid k, z^{t+1}, \omega^{t+1}, \gamma^t$, data, untuk $j = 1, 2, \dots, k$
5. $\gamma_j^{t+1} \sim \gamma_j \mid k, z^{t+1}, \omega^{t+1}, \theta^{t+1}$, data, untuk $j = 1, 2, \dots, k$
6. $t = t + 1$. lanjut langkah 2

Langkah 5 merupakan langkah yang rumit pada prosedur diatas yaitu pengambilan sampel distribusi bersyarat γ_j, untuk $j = 1, \dots, k$ dalam (iii). Metode yang biasanya digunakan untuk menyelesaikan langkah yang rumit tersebut adalah dengan menggunakan algoritma Metropolis-Hastings (Tsionas, 2002). Akan tetapi, pada tulisan ini digunakan algoritma slice sampling. Alasannya, algoritma ini lebih cepat dan lebih efisien. Proses algoritma ini menggunakan skema yang mengikuti simulasi berikut:

5.a. pertama, variabel acak seragam; $y \sim Uniform(0, g(\gamma_j^{(t)}))$, dimana $\gamma_j^{(t)}$ merupakan nilai γ_j saat ini,

5.b. kedua, mensimulasikan $\gamma_j^{(t+1)}$ dari distribusi seragam (*Uniform*) dengan $S_{(y)} = \{\gamma_j : g(\gamma_j) \geq y\}$ dimana $g(\gamma_j)$ dalam (iii)

Dalam prakteknya satu-satunya kesulitan dalam algoritma ini adalah dalam mengevaluasi *S(y)*.

Inferensi ketika *k* tidak diketahui langkah-langkahnya sama dengan inferensi ketika k diketahui, hanya saja dalam mensimulasikan sampelnya perlu memperpanjang algoritma *Gibbs sampling* (jika mengubah nilai k, jumlah parameter juga diubah), sehingga berbagai macam algoritma perlu dipertimbangkan seperti algoritma *reversible jump*, dan atau algoritma *birth-death* MCMC. Seperti yang ditunjukkan dalam Cappe dkk (2003) kedua metode dari algoritma tersebut pada dasarnya sama. Algoritma *birth-death* biasanya lebih mudah diterapkan, sedangkan algoritma *reversible jump* memerlukan waktu eksekusi yang lebih pendek. Dengan mengikuti Stephens (2000) dapat digunakan algoritma *birth-death*.

Untuk menerapkan algoritma ini, algoritma pengambilan sampel *Gibbs* dimodifikasi dengan menggganti k dengan $k^{(t)}$ sepanjang algoritmanya dan mengubah langkah 6 dari algoritma tersebut menjadi:

6. Menghasilkan k^{t+1} dan memodifikasi parameter yang tersisa
7. $t = t+1$, lanjut langkah 2

Untuk mensimulasikan realisasi dari proses *birth-death*, diberikan algoritma sederhana berikut:

a. Mulai dari nilai awal (k,ω,θ,γ)
b. Hitung angka kematian $\delta_j(k,\omega,\theta,\gamma)$, untuk j = 1,2, …, *k*
c. Menghasilkan waktu untuk kelahiran pertama atau kematian dari distribusi eksponensial dengan rata-rata $\frac{1}{\beta+\delta(k,\omega,\theta,\gamma)}$
d. Pilih kelahiran dengan peluang $\frac{\beta}{\beta+\delta(k,\omega,\theta,\gamma)}$ atau kematian komponen *j* dengan peluang $\frac{\delta_j(k,\omega,\theta,\gamma)}{\beta+\delta(k,\omega,\theta,\gamma)}$ untuk $j = 1,2,\dots,k$
e. Modifikasi model parameter k,ω,θ,γ yang sesuai.
f. Ulangi langkah b

Simulasi Data

Untuk melakukan simulasi mengestimasi parameter distribusi Weibull campuran ini, maka dilakukan dalam beberapa langkah sebagai berikut :

a. Membangkitkan data
   1. Menetapkan porsi bobot ω untuk distribusi Weibull campuran
   2. Menetapkan nilai parameter skala dan bentuk
   3. Menetapkan banyak data, yaitu 150 data
   4. Membangkitkan data berdistribusi Weibull untuk komponen pertama,kedua dan ketiga
   5. Menggabungkan data tiga komponen berdistribusi Weibull
b. Mengestimasi jumlah komponen dan nilai parameter
   1. Menetapkan nilai parameter untuk prior yakni $\boldsymbol{k},\boldsymbol{\omega},\boldsymbol{\gamma}$, dan $\boldsymbol{\theta}$
   2. Menentukan nilai awal posterior berdasarkan nilai parameter pada bagian (1) untuk posterior $\boldsymbol{k_0},\boldsymbol{\omega_0},\boldsymbol{\gamma_0}$, dan $\boldsymbol{\theta_0}$
   3. Menjalankan algoritma Gibss Sampling berdasarkan proses *birth- death*, yakni algoritma *birth death MCMC.*

Untuk melakukan estimasi parameter pada distribusi Weibull campuran, maka digunakan data sejumlah 150 dengan ditetapkan sebanyak 3 komponen sebagai berikut :

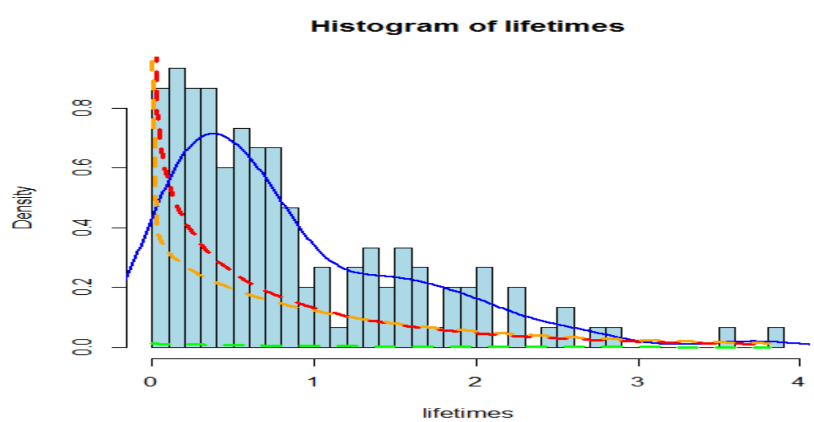
$$f(s|\theta,\gamma) = 0.6\,W(s|1,1.2) + 0.3\,W(s|0.8,1) + 0.1\,W(s|0.3,0.5) \quad (8)$$

Dalam mensimulasi data dan menaksir parameter digunakan bantuan perangkat lunak R 3.0.3.

3. Hasil dan Pembahasan

Berdasarkan hasil simulasi data distribusi Weibull campuran persamaan (8), maka secara sederhana penyebaran data dapat dilihat pada Gambar 1. Selanjutnya, untuk keperluan analisis kelangsungan hidup maka dengan menggunakan konsep tersensor kanan diperoleh sejumlah 18% data yang tersensor dan estimasi parameter hanya akan diberlakukan untuk data yang tidak tersensor sejumlah 82%.

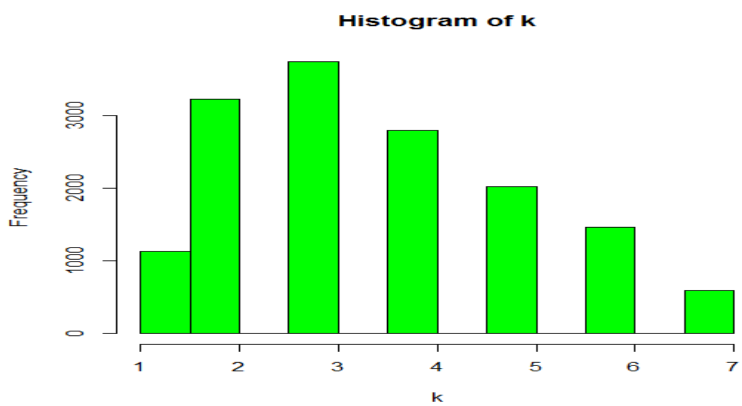
Nilai-nilai 0.6, 0.3 dan 0.1 pada persamaan (8) menunjukkan proporsi data untuk tiap komponennya. Sebesar 60% data berasal dari komponen pertama, 30% berasal dari komponen kedua dan hanya 10% berasal dari komponen ketiga. Data pada komponen pertama digambarkan dengan garis berwarna hijau pada Gambar 1. Sedangkan untuk komponen kedua dan ketiga berturut turut ialah garis berwarna merah dan garis berwarna jingga. Garis biru pada Gambar 1 merupakan garis yang menggambarkan plot dari gabungan ketiga komponen Weibull (Weibull campuran).



Gambar 1. Histogram Data Distribusi Weibull Campuran (biru) dengan 3 Komponen (hijau, merah dan jingga).

Pada Gambar 2 diperlihatkan bahwa distribusi Weibull yang membentuk distribusi Weibull campuran terdiri dari 3 komponen merupakan jumlah komponen yang paling cocok dengan jumlah iterasi 15.000. Proses estimasi jumlah komponen tersebut tidak terlepas dari nilai- nilai bentuk, skala, serta bobot yang didapatkan pada tiap iterasi, yang nantinya nilai-nilai tersebut mempengaruhi besarnya peluang kematian ataupun kelahiran dari tiap iterasi pada proses menjalankan algoritma birth-death MCMC.

Selain itu, pemilihan nilai λ yang tepat, juga akan menentukan terjadinya proses kelahiran dan kematian yang terjadi pada setiap iterasi. Nilai λ yang sangat besar, akan menyebabkan nilai peluang kelahiran yang selalu condong besar, sehingga pada saat iterasi dijalankan, akan menyebabkan kesalahan dan tidak tepatnya estimasi jumlah komponen. Sedangkan, untuk λ yang terlalu kecil, maka justru peluang kematian yang condong akan menjadi sangat besar sehingga kesimpulan bisa menjadi tidak valid. Pada beberapa penelitian lainnya pemilihan nilai λ dilakukan dengan beberapa nilai yang diujikan.



Gambar 2. Histogram Estimasi Jumlah Komponen Distribusi Weibull Campuran dengan algoritma birth-death MCMC.

Tabel 1. Hasil taksiran rata-rata posterior dari parameter distribusi Weibull campuran

Parameter	Nilai sebenarnya	Rata-rata post 1	Rata-rata post 2	Rata-rata post 3	Rata-rata post 4	Rata-rata post 5	Rata-rata semua post	MSE
Skala 1 (θ_1)	1	0.97	0.99	0.98	1.02	0.97	0.99	0.0004
Skala 2 (θ_2)	0.8	0.91	0.90	0.79	0.75	0.85	0.84	0.0056
Skala 3 (θ_3)	0.3	0.73	0.62	0.42	0.34	0.59	0.54	0.0799
Bentuk 1 (γ_1)	1.2	0.92	1.01	1.07	0.91	1.04	0.99	0.0464
Bentuk 2 (γ_2)	1	0.85	0.85	0.84	0.72	0.88	0.83	0.0309
Bentuk 3 (γ_3)	0.5	0.72	0.59	0.48	0.34	0.56	0.54	0.0181

Bobot 1 (ω_1)	0.6	0.47	0.52	0.64	0.72	0.57	0.58	0.0079
Bobot 2 (ω_2)	0.3	0.24	0.24	0.22	0.20	0.22	0.19	0.0190
Bobot 3 (ω_3)	0.1	0.13	0.12	0.08	0.06	0.10	0.10	0.0007

Hasil taksiran parameter untuk 5 kali ulangan proses perhitungan dapat disajikan pada Tabel 1. Nilai-nilai pada Tabel 1 menjelaskan karaterisktik data pada distribusi Weibull campuran ini. Berdasarkan Tabel 1, maka diketahui bahwa rata-rata 58,7% data berasal dari komponen pertama, 19,2% berasal dari komponen kedua, dan 10,3% berasal dari komponen ketiga. Selanjutnya diperoleh rata-rata nilai hasil taksiran parameter skala untuk komponen pertama yaitu 0,9936, parameter skala untuk komponen kedua 0.8435, dan rata-rata hasil taksiran parameter skala untuk komponen ketiga yaitu 0.5453. Pada Tabel 1 juga diketahui rata-rata nilai taksiran parameter bentuk untuk komponen pertama sebesar 0.9950, rata-rata nilai taksiran parameter bentuk untuk komponen kedua yaitu 0.8323, dan rata-rata nilai hasil taksiran parameter bentuk untuk komponen ketiga yaitu sebesar 0.5457. Jika dibandingkan nilai sebenarnya (nilai yang digunakan untuk membangkitkan data) dengan nilai hasil taksiran (rata- rata nilai taksiran dari 5 kali ulangan), maka akan terlihat bahwa nilai hasil taksiran tidak jauh berbeda dengan nilai sebenarnya. Adapun pergeseran nilai dapat dijelaskan sebagai akibat dari pergeseran nilai bobot yang disebakan oleh adanya data yang tersensor. Ukuran kebaikan hasil taksiran ditunjukkan dengan menggunakan nilai *Mean Square Error* (MSE) yang dihitung dari perulangan hasil taksiran. Berdasarkan nilai MSE pada Tabel 1 tersebut maka dapat disimpulkan bahwa hasil taksiran yang diperoleh cukup baik.

4. Kesimpulan

Berdasarkan pada hasil dan pembahasan maka dapat disimpulkan bahwa:

1. Estimasi nilai bobot mempunyai peran sangat besar terhadap taksiran nilai khususnya pada saat proses parameterisasi karena nilai parameter yang ditaksir telah terbarui dengan kondisi data yang tersensor sehingga banyak mempengaruhi hasil taksiran dari tiap parameter yang ditaksir.
2. Hasil estimasi parameter diperoleh setelah mengalikan likelihood dengan distribusi priornya, sehingga dengan kata lain hasil taksiran tersebut merupakan hasil taksiran yang telah diperbarui berdasarkan informasi prior yang ada.

Daftar Pustaka

[1] Berkson, J and Gage, R. P., 1952. Survival curve for cancer patients following treatment, *Journal of the American Statistical Association,* **47**(259), 501–515.

[2] Cappé, O., Robert, C. P., Rydén, T., 2003. Reversible jump birth-and-death and more general continuous time samplers. *J. Roy. Statist*. Soc. B **65**:679–700.

[3] Chen, W. C., Hill, B. M., Greenhouse, J. B., Fayos, J. V. , 1985. Bayesian analysis of survival curves for cancer patients following treatment. In: Bernardo, J. M., DeGroot, M. H., Lindley, D. V., Smith, A. F. M., eds. *Bayesian Statistics 2*. Amsterdam: North Holland, 299–328.
[4] Diebolt, J., Robert, C. P., 1994, Estimation of finite mixture models through Bayesian sampling. *Journal of Royal Statisics Society. B* **56**:363–375.
[5] Dodson, B., 1994, *Weibull Analysis*. Milwaukee: ASQ.
[6] Gilks, W. R. and Wild, P., 1992, Adaptive Rejection Sampling for Gibbs Sampling. *Applied Statisics*. **41**:337-348.
[7] Marín, J. M., Rodríguez-Bernal, M. T., Wiper, M. P., 2005, Using Weibull Mixture Distributions to Model Heterogeneous Survival Data, *Communication in Statistics - Simulation and Computation*, **34(**3):673-684.
[8] Quiang, J., 1994. *A Bayesian Weibull Survival Model*. Unpublished Ph.D. Thesis, Institute of Statistical and Decision Science, Duke University : North Corolina
[9] Stephens, M., 2000, Bayesian analysis of mixture models with unknown number of components - An alternative to reversible jump methods. *The Annals of Statistics,* 28(1):40–74.
[10] Tsionas, E. G., 2002, Bayesian analysis of finite mixtures of Weibull distributions. *Communication in Statistics Theory and Mathematics,* **31**:37–48.
[11] Walpole, R. E dan Myers, R. H., 1995, *Ilmu Peluang dan Statistika untuk Insinyur dan Ilmuwan*. ITB: Bandung.
[12] Walsh, B., 2004, *Markov Chain Monte Carlo and Gibbs Sampling*. Lecture Notes for EEB. 581.