Reliability Estimation of a Two Mixture Exponential Model Using Gibbs sampler

Heecheul Kim and Pyongkoo Kim
Department of Statistics, Dongguk University
Seoul, Korea
Department of Quality Control, Chung Cheong College
Chung Buk, Korea

Abrstract

A Markov Chain Monte Carlo method with data augmentation is developed to compute the features of the posterior distribution. This data augmentation approach facilitates the specification of the transitional measure in the Markov Chain. Bayesian analysis of the mixture exponential model discusses using the Gibbs sampler. Parameter and reliability estimators are obtained. A numerical study is provided.

1.Introduction

In the problem of life testing and reliability analysis, the exponential distribution plays a central role as a useful statistical model. The problem of estimating reliability in the mixture model has been considered on some papers. For example, the Mixtures of exponential(Mendenhall and Hader(1958), Tallis and Light(1968), Weibull(Kao(1959), Falls(1970)), inverse Gaussian distribution(Akman and Huwang(1997)) and Pareto distributions(Upadhyay and Shastri(1997)) have also been considered by numerous authors. Mixture models are generally rather difficult to handle statistically. Bayesian estimation of complicated functions requires simpler estimation techniques due to the mathematical difficulties involved in the classical Bayes approach.

This paper considers the estimation of reliability of systems. We use weighted sum of two exponential distributions as a life time distribution. The maximum likelihood estimator and the Gibbs estimator of reliability of the system are derived. By simulation risk behaviors of derived estimators are compared. Section 3 develops the reliability estimation

of the mixture exponential distribution. Section 2 develops the Gibbs algorithms for evaluating the posterior distributions.
A numerical example is given in Section 4. Some concluding remarks are given in Section 5.

2. Gibbs Sampling and Model

2.1 Gibbs Sampling

In this section, We use the Gibbs sampling originally introduced in Geman and Geman(1984), and more recently popularized by Gelfand and Smith(1990). Also Gelman and Rubin(1992) introduced iterative simulation using multiple sequences. In this section, We use Gelman and Rubin's method as follows.
First, simulate $m \geq 2$ sequences independently, each of the length $2n$, with starting points drawn from an over-dispersed distribution. To diminish the effect of the starting distribution, discard the first n iterations of each sequence, and focus attention on the last n.
Let us start with initial value $(u_1^{(0)}, u_2^{(0)}, \cdots, U_p^{(0)})$.
Generate a variate

$$
\begin{aligned}
u_1^{(1)} &\sim f(U_1 \mid U_2^{(0)}, U_3^{(0)}, \cdots, U_p^{(0)}, D_n), \\
u_2^{(1)} &\sim f(U_2 \mid U_1^{(0)}, U_3^{(0)}, \cdots, U_p^{(0)}, D_n), \\
&\vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
u_p^{(1)} &\sim f(U_p \mid U_1^{(0)}, U_2^{(0)}, \cdots, U_{p-1}^{(0)}, D_n)
\end{aligned}
$$

Thus each variable is visited in the natural order and a cycle in this scheme requires p random variate generations. After $2n$ such iterations, one arrives at

$$(U_1^{(2n)}, U_2^{(2n)}, \cdots, U_p^{(2n)}).$$

Under mild conditions(Geman and Geman(1984)),

$$(U_1^{(n)}, \cdots, U_p^{(n)}) \xrightarrow{d} (U_1, U_2, \cdots, U_p)$$

as $n \to \infty$.

The Gibbs sampling through m replications of the last n-iteration generates mn i.i.d p -tuples

$$(U_{1j}^{(l)}, \cdots, U_{pj}^{(l)})\ (j=1,2,\cdots,m,\ l=n+1,\cdots,2n);$$

$U_1, \cdots, U_p$ may possibly be vectors in the above scheme.

To obtain a pdf estimate(of any posterior), we use the Rao-Blackwell argument:

$$f(\widehat{U_s}) \simeq (mn)^{-1} \sum_{j=1}^{m} \sum_{l=n+1}^{2n} f(U_s \mid U_{rj}^{l},\ r \neq s).$$

2.2 Gibbs sampling of Two Mixture Exponential Model

The following two mixture Exponential Model is considered on this paper.

$$f(x \mid \lambda_1, \lambda_2, p) = p\lambda_1 \exp(-x\lambda_1) + q\lambda_2 \exp(-x\lambda_2) \tag{2.1}$$

where, $p+q=1$.

Given data set $D_n = \{x_1, x_2, \cdots, x_n\}$, $n = n_1 + n_2$, the likelihood function is given by

$$L_m(p, \lambda_1, \lambda_2 \mid D_n) \propto p^{n_1} q^{n_2} \lambda_1^{n_1} \lambda_2^{n_2} \cdot \exp\left(-\lambda_1 \sum_{i=1}^{n_1} x_{i1} - \lambda_2 \sum_{i=1}^{n_2} x_{i2}\right) \tag{2.2}$$

We assume the following prior densities for λ_1, λ_2, p

$$\pi(\lambda_1, \lambda_2, p) = \pi_1(p \mid \lambda_2, \lambda_1) \cdot \pi_2(\lambda_1 \mid \lambda_2) \cdot \pi_3(\lambda_2)$$

where,

(1) $\pi_1(p \mid \lambda_1, \lambda_2) \sim Beta(\lambda_1, \lambda_2)$

(2) $\pi_2(\lambda_1 \mid \lambda_2) \sim \mathrm{Gamma}(a_1, \lambda_2)$

(3) $\pi_3(\lambda_2) \sim \mathrm{Gamma}(a_2, b_2)$

where, a_1, a_2 and b_2 are known positive constants. We use $Gamma(a, b)$ to denote the gamma distribution with the mean a/b and $Beta(c, d)$ to denote the beta distribution with the mean $c/(c+d)$.
The joint posterior density is,

$$P(p, \lambda_1, \lambda_2, \mid D_n) \propto p^{n_1} q^{n_2} \lambda_1^{n_1} \lambda_2^{n_2} \exp\left(-\lambda_1 \sum_{i=1}^{n_1} x_{i1} - \lambda_2 \sum_{i=1}^{n_2} x_{i2}\right)$$
$$\cdot p^{\lambda_1 - 1} q^{\lambda_2 - 1} \cdot \lambda_1^{a_1 - 1} \exp(-\lambda_2 \lambda_1) \cdot \lambda_2^{a_2 - 1} \exp(-b_2 \lambda_2)$$

In this case, the marginal posterior densities for the Gibbs algorithm are given by

$$\begin{aligned}(1)\ f(p \mid \lambda_1, \lambda_2) &\propto p^{\lambda_1 + n_1 - 1} \cdot (1-p)^{\lambda_2 + n_2 - 1} \\ &\propto Beta(\lambda_1 + n_1, \lambda_2 + n_2)\end{aligned} \tag{2.3}$$

$$\begin{aligned}(2)\ f(\lambda_1 \mid p, \lambda_2) &\propto \lambda_1^{n_1 + a_1 - 1} \exp(-\lambda_1 (\sum_{i=1}^{n_1} x_{i1} - \ln p + \lambda_2)) \\ &\propto Gamma(n_1 + a_1, \sum_{i=1}^{n_2} x_{i1} + \lambda_2 - \ln p)\end{aligned} \tag{2.4}$$

$$\begin{aligned}(3)\ f(\lambda_2 \mid \lambda_1, p) &\propto \lambda_2^{n_2 + a_2 - 1} \exp(-\lambda_2 (\sum_{i=1}^{n_2} x_{i2} - \ln(1-p) + \lambda_1 + b_2)) \\ &\propto Gamma(a_2 + n_2, \sum_{i=1}^{n_2} x_{i2} - \ln(1-p) + \lambda_1 + b_2)\end{aligned} \tag{2.5}$$

3. Reliability Estimation

The reliability function for the time t given by

$$\begin{aligned}R(t) &= P_r(X > t) = \int_t^{\infty} f(x \mid \lambda_1, \lambda_2, p) dx \\ &= p \exp(-t\lambda_1) + q \exp(-t\lambda_2)\end{aligned} \tag{3.1}$$

To obtain estimators, we need the MLE of $(\lambda_1, \lambda_2, p)$.
For a random sample size n, If $p\lambda_1 \exp(-\lambda_1 x_i) \geq q\lambda_2 \exp(-\lambda_2 x_i)$ regards it as a subgroup $(x_{i1},\ i = 1, 2, \cdots, n_1)$, otherwise consider other subgroup

$(x_{i2},\ i=1,2,\cdots,n_2)$, the likelihood function is

$$L_m(\lambda_1,\lambda_2,p\mid D_n)\propto p^{n_1}q^{n_2}\lambda_1^{n_1}\lambda_2^{n_2}\cdot\exp\left(-\lambda_1\sum_{i=1}^{n_1}x_{i1}-\lambda_2\sum_{i=1}^{n_2}x_{i2}\right)$$

If p is known constant, the MLE's are given as

$$\frac{1}{\widehat{\lambda}_1}=\frac{\sum_{i=1}^{n_1}x_{i1}}{n_1},\quad \frac{1}{\widehat{\lambda}_2}=\frac{\sum_{i=1}^{n_2}x_{i2}}{n_2}$$

thus, MLE of R is given as

$$\widehat{R}_{MLE}=p\ \exp(-t\ \widehat{\lambda}_1)+q\ \exp(-t\ \widehat{\lambda}_2).$$

If p is unknown constant, we can be a prior distribution (example, Beta distribution) to obtain estimator of p. Thus we obtain the alternative estimator of reliability function $R(t)$,

$$\widehat{R}_{GIBBS}=\hat{p}\ \exp(-t\ \widehat{\lambda}_1)+\hat{q}\ \exp(-t\ \widehat{\lambda}_2)$$

using the Gibbs sampler procedure.

4. A Numerical Example

In implementing the Gibbs sampler, we can be able to draw samples from the full conditional densities given in (2.3)-(2.5) by using IMSL software. In our simulated data, we take sample size $n=30$ and $\lambda_1=5$, $\lambda_2=2$, $p=0.5$ using IMSL RNEXT function. Table 1 gives the generated data. To diffuse prior stage, we take $a_1=10$, $a_2=30$ and $b_2=0.1$.

In Gibbs sampler, we use 10 sequence and 3000 iterations for each sequence. We can easily compute the posterior means for parameters $\widehat{\lambda}_1=4.928$, $\widehat{\lambda}_2=1.983$ and $\widehat{p}=0.487$. Each parameter estimators are summarized in Table 2, quite close to each other.

Reliability of $R(t) = \hat{p}\exp(-\hat{\lambda}_1 \cdot t) + (1-\hat{p})\exp(-\hat{\lambda}_2 \cdot t)$ and relative errors are given Table 3. In this Table, as following pass time t, relative errors are having a decreasing pattern and then estimators of reliability, for example the MLE and Gibbs method, quite close to each other. Reliability using the MLE and Gibbs methods are very close to the true value.

Table 1. Simulated Data

5.74	5.97	2.58	5.49	1.57
3.59	1.90	2.49	5.52	4.71
3.80	4.86	2.06	2.47	2.19
1.03	5.18	5.40	1.10	2.40
5.38	1.97	1.70	2.15	4.04
5.52	2.08	4.06	1.68	2.34

Table 2. Estimators of Parameters

	True Value	Gibbs Estimator (Posterior Mean)	MLE
$\hat{p}$	0.5	0.487	0.5 (Known Value)
$\hat{\lambda}_1$	5	4.928	4.972
$\hat{\lambda}_2$	2	1.983	1.948

5. Conclusion Remark

We show how Bayesian inference for mixture exponential model is accessible now by using a Marcov Chain Monte Carlo technique with data augmentation.

A Markov Chain Monte Carlo method with data augmentation is developed to compute the features of the posterior distribution. This data augmentation approach facilitates specification of the transitional measure in the Markov Chain. Bayesian analysis of a mixture of exponential model is discussed using Gibbs sampler and MLE. We show that relative errors are having a decreasing pattern and then estimators of reliability, for example the MLE and Gibbs method ,quite close to each other.

Table 3. Reliability of $R(t) = \hat{p}\exp(-\hat{\lambda}_1 \cdot t) + (1-\hat{p})\exp(-\hat{\lambda}_2 \cdot t)$ and Relative Error

t	$R_{True}(t)$	$\hat{R}_{Gibbs}(t)$	$\hat{R}_{MLE}(t)$	$\frac{[R_{True}(t) - \hat{R}_{Gibbs}(t)]^2}{R_{True}(t)}$	$\frac{[R_{True}(t) - \hat{R}_{MLE}(t)]^2}{R_{True}(t)}$
0.25	0.44651772	0.45453750	0.45149221	0.000244041	0.000150235
0.50	0.22498222	0.23177385	0.23040600	0.000205022	0.000130754
0.75	0.12332395	0.12802317	0.12801049	0.000179062	0.000178098
1.00	0.07103661	0.07414370	0.07474409	0.000135902	0.000193497
1.25	0.04200772	0.04404267	0.04479848	$9.85776 \cdot 10^{-05}$	0.000185402
1.50	0.02517007	0.02650050	0.02720136	$7.03228 \cdot 10^{-05}$	0.000163931
1.75	0.01517792	0.01604658	0.01662033	$4.97147 \cdot 10^{-05}$	0.000137078
2.00	0.00918051	0.00974641	0.01018552	$3.48819 \cdot 10^{-05}$	0.00011002
2.25	0.00556100	0.00592857	0.00625084	$2.42953 \cdot 10^{-05}$	$8.55754 \cdot 10^{-05}$
2.50	0.00337083	0.00360881	0.00383868	$1.68004 \cdot 10^{-05}$	$6.49329 \cdot 10^{-05}$
2.75	0.00204392	0.00219748	0.00235809	$1.15381 \cdot 10^{-05}$	$4.82916 \cdot 10^{-05}$
3.00	0.00123952	0.00133831	0.00144878	$7.87345 \cdot 10^{-06}$	$3.53252 \cdot 10^{-05}$
3.25	0.00075176	0.00081513	0.00089017	$5.34116 \cdot 10^{-06}$	$2.54835 \cdot 10^{-05}$
3.50	0.00045595	0.00049649	0.00054696	$3.60384 \cdot 10^{-06}$	$1.81674 \cdot 10^{-05}$
3.75	0.00027654	0.00030241	0.00033608	$2.41969 \cdot 10^{-06}$	$1.28202 \cdot 10^{-05}$
4.00	0.00016773	0.00018420	0.00020651	$1.61735 \cdot 10^{-06}$	$8.96684 \cdot 10^{-06}$

References

1. Akman, O., and Huwang, L. (1997), "Bayes computation for reability estimation," IEEE Transactions on Reability, Vol.46 52-55.
2. Aitkin, M. (1991), "Posterior bayes factors," Journal of the Royal Statistical Society, Ser. B, 111-142.
3. Dawid, A. P. (1984), "Statistical theory : The prequential approach," Journal of the Royal Statistical Society, Ser. A, 147, 278-292.
4. Falls, L. W.(1970), " Estimation of parameters in compound Weibull distributions," Technometrics, Vol.12, 399-407.
5. Gelfand, A. E., and Smith, A. F. M. (1990), "Sampling-based approaches to calculating marginal densities," Journal of the American Statistical Association, Vol. 85, 398-409.
6. Gelman, A. E., and Rubin D. (1992), "Inference from iterative simulation using multiple sequences," Statistical Science, Vol.7, 457-472.
7. Geman, S., and Geman, D. (1984), "Stocastic relation, Gibbs distribution and the

Bayesian restoration of images," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.6, 721-741.

8. Goel, A. L., and Okumoto, K. (1979), "Time dependent error detection rate model for software reliability and other performance measures," IEEE Transactions on Reliability, Vol.28, 206-211.
9. Johnson, N. L., and Kotz, S.(1970), Continuous univariate distribution-1, John Wiley and Sons, inc. 207-227.
10. Johnson, R. A., and Wichern, D. W.(1992), Applied Multivariate Statistical Analysis, Prentice and Hall, inc. 493-525.
11. Kao, J. H. K.(1959), " A graphical estimation of mixed Weibull parameters in life testing of electron tubes," Technometrics, Vol.1, 389-407.
12. Kuo, L., and Yang, T. Y. (1995), "Bayesian computation of software reliability", Journal of Computational and Graphical Statistics, 65-82.
13. Mendenhall, W., and Hader,R. J.(1958), "Estimation of parameters of mixed exponential distributed failure times from censored life test dat," Biometrika, Vol.45, 504-520.
14. Musa,J.D., and., Iannino, A., and Okumoto, K. (1987), Software reliability: Measurement, Prediction, Application, New York: McGraw Hill.
15. Pettit, L. I., and Young, K. D. S. (1990), "Measuring the effects of bservations on Bayes factors," Biometrika, Vol.77, 455-466.
16. Tallis, G. M., and Light, R.(1968), "The use of fractional moments for estimating the parameters of a mixed exponential distribution," Technometrics, Vol.10, 161-175.
17. Upadhyay, S, K., and Shastri, V.(1997), "Bayes result for classical Pareto distribution via Gibbs sampler, with doubly-censored observations," IEEE Transactions on Reliability, Vol.46 56-59.
18. User'Manual STAT/LIBRARY(1987) Fortran Subroutines for statistical analysis, IMSL, Vol.3.