Journal of Reliability and Statistical Studies; ISSN (Print): 0974-8024, (Online): 2229-5666
Vol. 9, Issue 1 (2016): 21-28

BAYES ESTIMATES FOR THE PARAMETERS OF POISSON TYPE LENGTH BIASED EXPONENTIAL CLASS MODEL USING NON-INFORMATIVE PRIORS

Rajesh Singh[1], Pritee Singh[2] and Kailash Kale[3]
[1]Department of Statistics, S. G. B. Amravati University, Amravati, India
[2]Department of Statistics, Institute of Science, Nagpur, India
[3]Department of Statistics, R. D. N. C., Bandra (W), Mumbai, India
E Mail: rsinghamt@hotmail.com, priteesingh25@gmail.com,
kailashkale10@gmail.com

Received September 23, 2015
Modified March 09, 2016
Accepted April 09, 2016

Abstract

In this paper, the failure intensity has been characterized by one parameter length biased exponential class Software Reliability Growth Model (SRGM) considering the Poisson process of occurrence of software failures. This proposed length biased exponential class model is a function of parameters namely; total number of failures θ_0 and scale parameter θ_1. It is assumed that very little or no information is available about both these parameters. The Bayes estimators for parameters θ_0 and θ_1 have been obtained using non-informative priors for each parameter under square error loss function. The Monte Carlo simulation technique is used to study the performance of proposed Bayes estimators against their corresponding maximum likelihood estimators on the basis of risk efficiencies. It is concluded that both the proposed Bayes estimators of total number of failures and scale parameter perform well for proper choice of execution time.

Key Words: Binomial Process, Non-Informative Prior, Maximum Likelihood Estimator (MLE), Rayleigh Class, Software Reliability Growth Model (SRGM), Incomplete Gamma Function, Confluent Hypergeometric Function.

1. Introduction

This paper considers Poisson type length biased exponential class model as per classification scheme of Musa and Okumoto (1984) developed for software reliability models. If we consider the whole program as a single entity to be executed then it takes very long time for execution (i.e. in months or years) for the real time system. Musa et al. (1987) have suggested that it is convenient to divide the whole program into number of runs. The nature and size of the run is depending on the function which is executed by the program. The runs are identical repetitions of each other termed as run type. Therefore, the time required for the execution of run type is depending upon the size of run. Thus, the number of failures observed in single run may vary as the size of run varies. Hence the length biased distribution defined by Fisher (1934) and formalized by Rao (1965) can be introduced as SRGM. These distributions have also been applied in reliability theory (cf. Gupta and Keating (1986), Gupta and Tripathi (1990) and Khatree (1989)).

Here, it is considered that the failure experienced by time t is distributed as Poisson (i.e. Type) and time to failure of an individual fault following length biased exponential distribution (i.e. Class). In other words, the functional form of failure intensity in this class can be described by length biased exponential distribution. In this model, it is assumed that the software failures are independent of each other but depend on length of time interval which contains the same software failures. In this paper, Bayes estimators for the parameters are obtained using the technique of Musa et al. (1987) (see also Singh et al. (2009) and Singh and Andure (2008)) for this model considering non-informative priors and they are compared with MLEs in further sections.

2. Model Formulation

Considering time to failure following length biased exponential Class with parameter θ_1 and Poisson occurrence (type) of software failure i.e.

$$f(t) = \begin{cases} t\theta_1^2 e^{-\theta_1 t} \; ; \; t > 0, \theta_1 > 0, E[t] \neq 0 \\ 0 \qquad ; \qquad otherwise \end{cases}. \tag{1}$$

where $f(t)$ denotes the length biased exponential distribution and θ_1 is scale parameter of the distribution. Assuming that the total number of faults remaining in the program at time $t = 0$ is a Poisson random variable with mean θ_0, the failure intensity $\lambda(t) = \theta_0 f(t)$ (cf. [7]) can be obtained as

$$\lambda(t) = \theta_0 t\theta_1^2 e^{-\theta_1 t} \; ; \qquad t > 0, \theta_1 > 0 \; \theta_0 > 0 \tag{2}$$

The parameter θ_0 can also be define as number of failures present initially in the software i.e. at time $t_0 = 0$. The random variables i.e. failures experienced $M(t)$ with an average $\mu(t_e) = \theta_0 F(t)$ upto execution time t_e can be obtained as

$$\mu(t_e) = \theta_0\left[1 - (1 + \theta_1 t_e)e^{-\theta_1 t_e}\right]; \; t > 0, \theta_1 > 0 \; \theta_0 > 0 \tag{3}$$

The probability of getting $M(t) = m$ number of failures experienced by time t_e can be obtained by Poisson density with mean $\mu(t_e)$ (cf. [7]),

$$P[M(t) = m] = \frac{\{\theta_0[1-(1+\theta_1 t_e)e^{-\theta_1 t_e}]\}^m exp\{\theta_0[1-(1+\theta_1 t_e)e^{-\theta_1 t_e}]\}}{m!} \tag{4}$$

Following figures show the behavior of $\lambda(t)$ and $\mu(t_e)$.

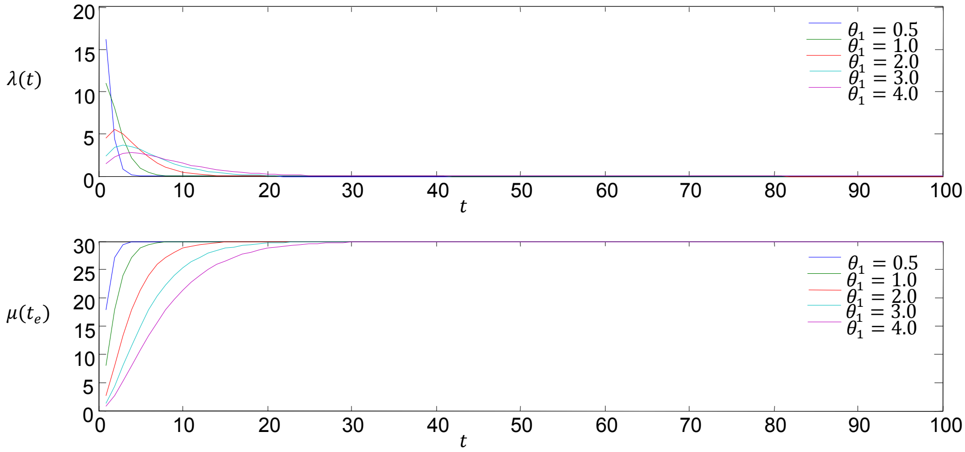

Fig. 1: Behavior of failure intensity $\lambda(t)$ and expected number of failures $\mu(t_e)$ for fixed $\theta_0(=30)$, different values of $\theta_1(=0.5, 1.0(1.0)4.0)$and $t_e(=1(1)100)$

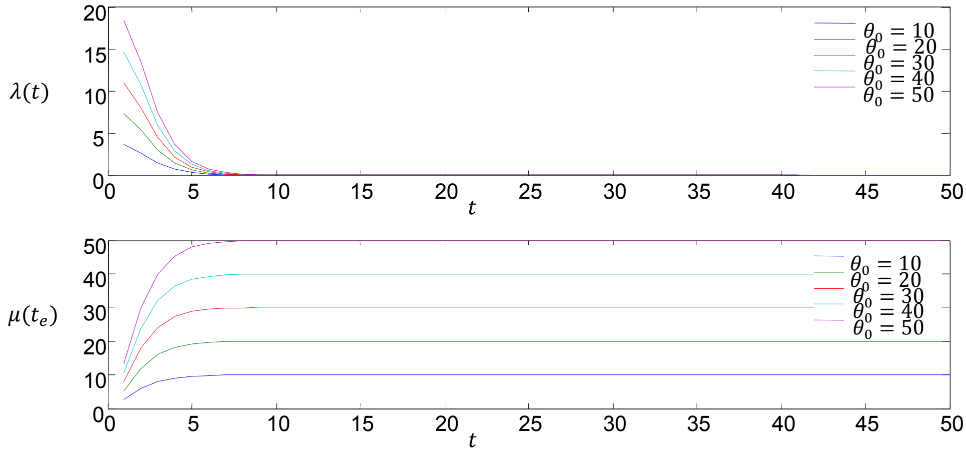

Fig. 2: Behaviour of failure intensity $\lambda(t)$and expected number of failures $\mu(t_e)$for fixed $\theta_1(=1.0)$, different values of $\theta_0(=10(10)50)$ and $t_e(=1(1)50)$.

The behavior of $\lambda(t)$ and $\mu(t_e)$ are studied by plotting the graphs between failure times $\lambda(t)$ and $\boldsymbol{t_e}$ as well as $\mu(t_e)$ and $\boldsymbol{t_e}$ considering different values of $\theta_1(=0.5,1.0(1.0)5.0)$ for fixed $\theta_0 = 30$ and $\theta_0(=10(10)\,50)$ for fixed $\theta_1 = 1.0$ which are presented in Figure-1 and Figure-2 respectively. Some of the important observations are presented here.

Failure intensity $\lambda(t)$ is very high for the smaller values of θ_1, becoming unimodal positively skewed and for larger values of $\theta_1 > 1.0$ and fixed $\theta_0 = 30$ (see Figure-1). The expected number of failures are large for smaller values of failure rate θ_1. The Figure-2 shows that the failure intensity $\lambda(t)$ is less sensitive for increasing values of $\theta_0(=10(10)50)$ and fixed $\theta_1 = 1.0$. The slope of failure intensity and expected number of failures remains similar for increasing values of θ_0 and fixed $\theta_1 = 1.0$.

3. Maximum Likelihood Estimation

Now, assuming that m_e failures are experienced at times $t_i, i = 1,2,\ldots,m_e$ up to execution time is t_e ($\geq t_{m_e}$) and the likelihood function of θ_0 and θ_1 can be obtained as $L(\theta_0,\theta_1) = \{\prod_{i=1}^{m_e}\lambda(t_i)\}\exp(-\mu(t_e))$ (cf. Musa et al. (1987)). Using the above failure intensity function given in (2) and mean software failures in (3), the likelihood function is

$$L(\theta_0,\theta_1) = \theta_0^{m_e}\theta_1^{2m_e}\left[\prod_{i=1}^{m_e} t_i\right]e^{-T\theta_1}e^{-\theta_0\left[1-\theta' e^{-\theta_1 t_e}\right]} \qquad (5)$$

where

$$\sum_{i=1}^{m_e} t_i = T$$

and

$$\theta' = (1+\theta_1 t_e)$$

The Maximum Likelihood Estimators denoted by $\hat{\theta}_{m0}$ and $\hat{\theta}_{m1}$ for parameters $\theta_0$and θ_1 respectively can be obtained by using standard method, which are

$$\hat{\theta}_{m0} = \left[\frac{m_e}{\left(1-(1+\hat{\theta}_{m1}t_e)e^{-\hat{\theta}_{m1}t_e}\right)}\right] \qquad (6)$$

and

$$\hat{\theta}_{m1} = \left[\frac{(2m_e - T\theta_{m1})e^{\hat{\theta}_{m1}t_e}}{\hat{\theta}_{m0}t_e^2}\right]^{1/2} \qquad (7)$$

4. Bayesian parameter estimation

While testing of the software the tester or experimenter may not have any past experience or may have very little knowledge about the number of failures present in the software initially (at $t = 0$) i.e. θ_0 and scale parameter θ_1. Hence in such case due to unavailability of prior information about both the parameters θ_0 and θ_1, the non informative priors can be considered. The following non-informative prior distributions $g(\theta_0)$ and $g(\theta_1)$ are considered for parameters θ_0 and θ_1

$$g(\theta_0) \propto \begin{cases} \frac{1}{\theta_0} \ ; & \theta_0 \epsilon\ [0,\infty) \\ 0 \ ; & otherwise \end{cases} \qquad (8)$$

and

$$g(\theta_1) \propto \begin{cases} \frac{1}{\theta_1} \ ; & \theta_1 \epsilon\ [0,\infty) \\ 0 \ ; & otherwise \end{cases} \qquad (9)$$

The joint posterior of θ_0 and θ_1 given $\underline{t}(= t_i,\ i = 1,2,\ldots,m_e)$ is obtained by using equations (5), (8) and (9) which is

$$\pi(\theta_0,\theta_1|\underline{t}) \propto \theta_0^{m_e-1}\theta_1^{2m_e-1}e^{-T\theta_1}e^{-\theta_0}e^{\theta_0\theta' e^{-\theta_1 t_e}}$$

$$m_e < \theta_0 < \infty, 0 < \theta_1 < \infty \qquad (10)$$

The marginal posterior of θ_1, say $\pi(\theta_1|\underline{t})$ can be obtained as

$$\pi(\theta_1|\underline{t}) \propto \theta_1^{2m_e-1}e^{-T\theta_1}\left[1-\theta' e^{-\theta_1 t_e}\right]^{-m_e}\Gamma\left(m_e, m_e\theta' e^{-\theta_1 t_e}\right)$$

$$0 < \theta_1 < \infty$$

where

$$\theta' e^{-\theta_1 t_e} \leq 1 \quad \text{since } \theta' \leq e^{\theta_1 t_e}.$$

The marginal posterior of θ_0, say $\pi(\theta_0|\underline{t})$ is

$$\pi\left(\theta_0|\underline{t}\right) \propto \Gamma(2m_e)\theta_0^{m_e-1}e^{-\theta_0}t_e^{-2m_e}\sum_{j=0}^{\infty}\frac{\theta_0^j}{j!}\Psi(2m_e, 2m_e + j + 1, t_e^{-1}T')$$
$$,0 < \theta_1 < \infty$$

where

$$T' = T + jt_e$$

The Bayes estimators for the parameter θ_0 (say $\hat{\theta}_{B0}$) and for θ_1 (say $\hat{\theta}_{B1}$) are posterior mean under squared error loss function and are

$$\hat{\theta}_{B0} = D^{-1}\sum_{j=0}^{\infty}\frac{\Gamma(m_e+j+1,m_e)}{j!}\Psi(2m_e, 2m_e + j + 1, T't_e^{-1}) \qquad (11)$$

and

$$\hat{\theta}_{B1} = 2D^{-1}m_e t_e^{-1}\sum_{j=0}^{\infty}\frac{\Gamma(m_e+j,m_e)}{j!}\Psi(2m_e + 1, 2m_e + j + 2, T't_e^{-1}) \qquad (12)$$

where $\Psi(\alpha, \beta; x)$ is Confluent Hypergeometric Function (cf. Abramowitz and Stegun (1965) and Gradshteyn and Ryzhik (1994)), normalizing constant is

$$D = \sum_{j=0}^{\infty}\frac{\Gamma(m_e+j,m_e)}{j!}\Psi(2m_e, 2m_e + j + 1, T't_e^{-1})$$

Suppose $\hat{\theta}$ be a mle or Bayes estimator of unknown parameter θ, let $L(\theta, \hat{\theta})$ be a squared error loss function then risk of $\hat{\theta}$ is defined as $R = E[\hat{\theta} - \theta]^2$. The risk efficiency of $\hat{\theta}$ over any other estimator $\hat{\theta}'$ is defined as

$$RE = R'R^{-1} \qquad (13)$$

where

$$R' = E[\hat{\theta}' - \theta]^2 \qquad , \forall\, \hat{\theta}' \in \Theta$$

5. Discussion

Here, the Bayes estimators of total number of failures θ_0 and failure rate θ_1 are compared with the corresponding maximum likelihood estimators. The performance of proposed Bayes estimators $\hat{\theta}_{B0}$ and $\hat{\theta}_{B1}$ over maximum likelihood estimators $\hat{\theta}_{m0}$ and $\hat{\theta}_{m1}$ have been compared on the basis of risks efficiencies. The risks efficiencies depend upon the values of total execution time i.e. t_e and m_e i.e. failures experienced at times $t_i, i = 1,2, \dots, m_e$ such that t_e ($\geq t_{m_e}$). To study the performance, a sample of size, say m_e was generated from the length biased exponential distribution subject to condition that t_{m_e} should be less than t_e and it is repeated 10^3 times. Then, using Monte Carlo simulation technique risk efficiencies have been evaluated and are presented in the graphs from Figure 3 to 5.

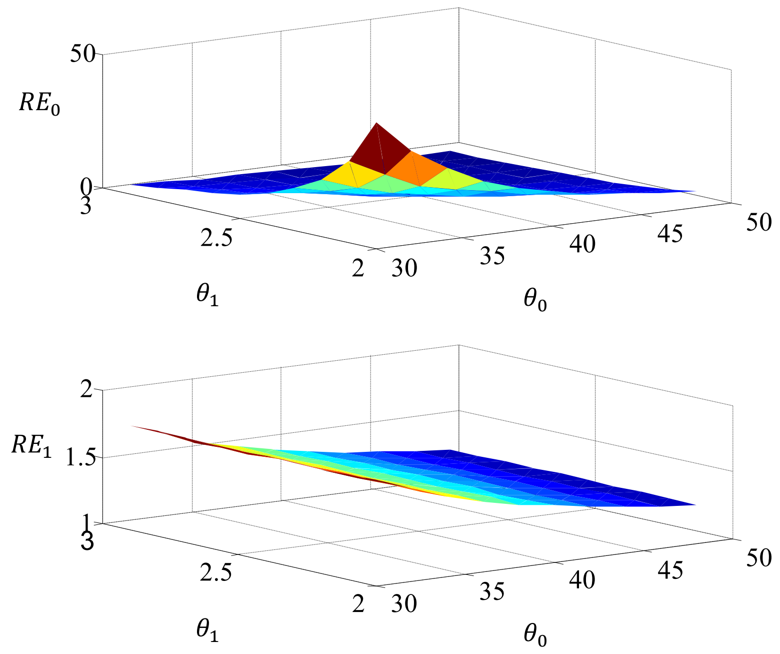

Fig. 3: Risk efficiencies of $\widehat{\theta}_{B0}$ and $\widehat{\theta}_{B1}$ for different values of $\theta_0(=30(2)48)$ and $\theta_1(=1.5(0.1)2.4)$ when $t_e = 75$

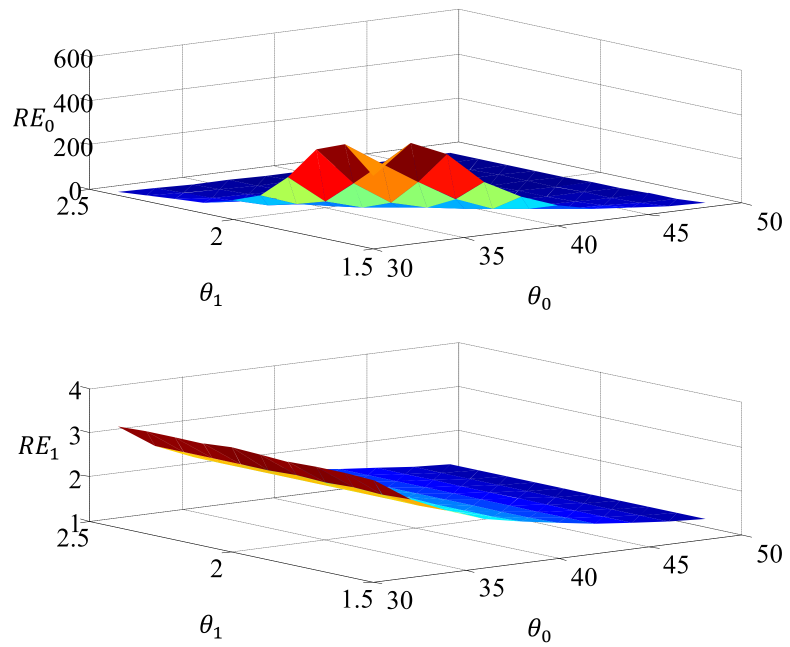

Fig. 4: Risk efficiencies of $\widehat{\theta}_{B0}$ and $\widehat{\theta}_{B1}$ for different values of $\theta_0(=30(2)48)$ and $\theta_1(=1.5(0.1)2.4)$ when $t_e = 100$.

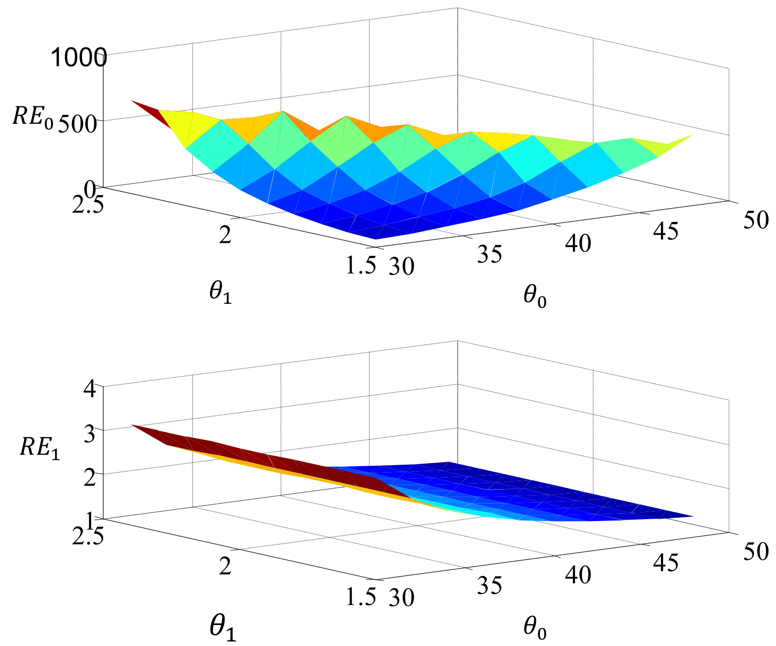

Fig. 5: Risk efficiencies of $\hat{\theta}_{B0}$ and $\hat{\theta}_{B1}$ for different values of $\theta_0(=30(2)48)$ and $\theta_1(=1.5(0.1)2.4)$ when t_e = 150.

From figures 3 to 5, it can be seen that the risk efficiencies RE_0 of $\hat{\theta}_{B0}$ decrease as θ_0 and θ_1 increase. It can also be seen that for small values of θ_1 and $\theta_0$the proposed Bayes estimator of $\theta_0$perform better than MLE. The RE_0 first increases, attains a maxima and then decreases as the value of t_e increases. Similarly, it can be observed that the risk efficiencies of $\hat{\theta}_{B1}$ i.e. RE_1 decrease as the values of θ_1 and θ_0 increase but the risk efficiencies RE_1decrease slowly due to increase of θ_0 and t_e. It is important to note that the proposed Bayes estimator$\hat{\theta}_{B1}$ is always better than MLE. On the basis of better performance of risk efficiencies of $\hat{\theta}_{B0}$ and $\hat{\theta}_{B1}$ over $\hat{\theta}_{m0}$ and $\hat{\theta}_{m0}$ following conclusions can be drawn.

6. Conclusions

- Researchers/software testers may have no information about prior belief about the occurrence of software failures.
- The proposed Bayes estimator of θ_0 can be preferred over MLE if it is felt that total number of failures may not be very large and failure rate may be small.
- The proposed Bayes estimator of θ_1 can be preferred over MLE.
- This model can be used when there are moderately large numbers of software failures in single run.

Acknowledgement

This work is a part of Major Research Project funded by University Grants Commission (UGC), New Delhi (India). The authors would like to express gratitude and thanks to U.G.C. New Delhi.

References

1. Abramowitz M. and Stegun I. A. (1965). Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, New York, Dover publications.
2. Fisher R. A. (1934). The effects of methods of ascertainment upon the estimation of frequencies, Ann. Eugenics, 6, p.13-25.
3. Gradshteyn I. S. and Ryzhik I. M. (1994). Table of Integrals, Series, and Products, Alan Jeffrey (editor) 5th Ed., New York, Academic Press.
4. Gupta R. C. and Keating J. P. (1986). Relations for reliability measures under length biased sampling, Scand Journal of Statistics, 13, p. 49-56.
5. Gupta R. C. and Tripathi R. C. (1990). Effect of length-biased sampling on the modeling error, Communication in statistics –Theory and Methods, 19(4), p. 1483-1491.
6. Khatree R. (1989). Characterization of Inverse-Gaussian and Gamma distributions through their length-biased distributions. IEEE Trans. on Reliability 38(5), p. 610-611.
7. Musa J. D. and Okumoto K. (1984). A logarithmic Poisson execution time model for software reliability measurement, Proceedings of Seventh International conference on software engineering, Orlando, p. 230-238.
8. Musa J. D., Iannino A. and Okumoto K. (1987). Software Reliability: Measurement, Prediction, Application, New York, McGraw-Hill.
9. Rao C. R. (1965). On discrete distributions arising out of methods of ascertainment, In Classical and Contagious Discrete Distributions, Eds. G.P. Patil, Pergamon Press and Statistical Publishing Society, Calcutta, p. 320-332.
10. Singh R. and Andure N. W. (2008). Bayes estimators for the parameters of the Poisson type exponential distribution", IAPQR transactions, 33 (2), p. 121-128.
11. Singh R., Vidhale A. A. and Carpenter M. (2009). Bayes estimators of parameters of Poisson Type Exponential Class Software Model considering generalized Poisson and Gamma priors, Journal of Model Assist. Statis. Appl., 4(2), p. 83-89.