IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Impact Factor: 5.731
Volume 2, Issue 1, January 2022

Estimation of Parameters of PTRC SRGM using Non-informative Priors

Rajesh Singh[1], Pritee Singh[2], Kailash Kale[3]
Director, IIL, RTM Nagpur University, Nagpur, India[1]
Department of Statistics, Institute of Science, Nagpur, India[2]
Department of Statistics, G. N. A. ACS College, Barshitakli, Akola, India[3]
rsinghamt@hotmail.com[1], priteesingh25@gmail.com[2] and kailashkale10@gmail.com[3]

Abstract: *Reliability is an essentially important characteristic of software. The reliability of software has been assessed by considering Poisson Type occurrence of software failures and the failure intensity of one parameter say* (η_1) *Rayleigh class. Here, it is assumed that the software contains fixed number of inherent faults say* (η_0). *The scale parameter of Rayleigh density* (η_1) *and fixed number of inherent faults contained in software are the parameters of interest. The failure intensity and mean failure function of this Poisson Type Rayleigh Class (PTRC) Software Reliability Growth Model (SRGM) have been studied. The estimates of above parameters can be obtained by using maximum likelihood method. Bayesian technique has been used to about estimates of* η_0 *and* η_1 *if prior knowledge about these parameters is available. The prior knowledge about these parameters is considered in the form of non-informative priors for both the parameters. The proposed Bayes estimators are compared with their corresponding maximum likelihood estimators on the basis of risk efficiencies under squared error loss. The Monte Carlo simulation technique is used for calculating risk efficiencies. It is seen that both the proposed Bayes estimators can be preferred over corresponding MLEs for the proper choice of the values of execution time.*

Keywords: Rayleigh probability density function, non- informative priors, squared error loss.

I. Introduction

The research on statistical modeling of software reliability has been started from 1960's as a Markovian processes. Software reliability has been defined by many authors, the definition by Musa et. al. (1987) has been noted here. The probability of failure free operation of computer programs for specified period of time in specified environment. Many models are proposed to characterize reliability of software categorized according to finiteness/infiniteness of software failures (faults)initially present in software.

In this paper, to guess about the behavior of parametersη_0 and η_1 after application prior knowledge or past experience about them, Bayesian estimation technique has been used. The use of Baye's techniqueto obtainestimators for the parameters of SRMs/SRGMs has been noted in Musa et al. (1987).Also, the authors Singh and Andure (2008) and Singh et al. (2009)have used this technique to obtain the Bayes estimates for the parameters of Poisson Type Exponential Class considering gamma prior.

This chapter considers and discusses the Rayleigh model as a SRGM. The Rayleigh model is a member of the family of the Weibull distribution (cf. Sinha (1985)). Yamada et al. (1986) have proposed SRGMs incorporating the quantity of test-effort exhausted on software testing described by the Rayleigh curve. A SRGM with testing effort described by exponential, Rayleigh and Weibull curve have also beenproposed by Yamada et al. (1983, 1986) and others. Schick and Wolverton (1978a, 1978b) have extended the Jelinski-Moranda model which includes the non constant failure rate model assuming Rayleigh distribution.The use of Rayleigh in the study of software reliability has been shown by many authors (see Boehm et al. (2000), DeMarco (1982), Putnam (1978),Kan (2002), Gaffney and Pietrolewicz (1990), Trachtenberg (1982) etc).

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 5.731

Volume 2, Issue 1, January 2022

II. Model Formulation

Let T be the positive random variable having Rayleigh distribution and t is its realization then, its probability density function and cumulative distribution function are is given by

$$f(t) = \begin{cases} t\eta_1^{-2} exp\left(-\frac{1}{2}(t\eta_1^{-1})^2\right), & t > 0, \eta_1 > 0 \\ 0 & , Otherwise \end{cases} \qquad ...(2.1)$$

then

$$F_a(t) = \left[1 - exp\left(-\frac{1}{2}(t\eta_1^{-1})^2\right)\right] \quad , t > 0, \eta_1 > 0 \qquad ...(2.2)$$

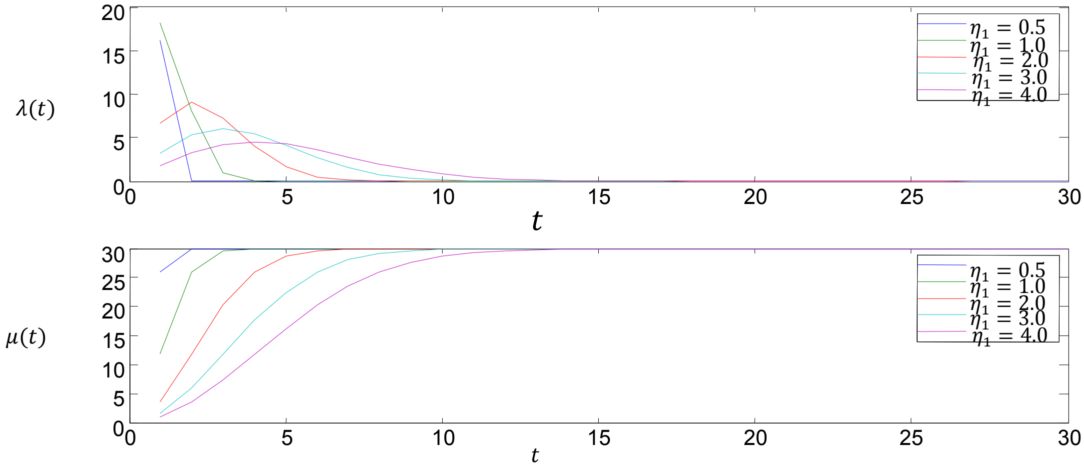
where $f(t)$ denotes the length biased exponential distribution and θ_1 is scale parameter of the distribution. Assuming that the total number of faults remaining in the program at time $t = 0$ is a Poisson random variable with mean θ_0, the failure intensity $\lambda(t) = \theta_0 f(t)$ (cf.Musa et. al. (1987)) can be obtained as

$$\lambda(t) = \eta_0 t\eta_1^{-2} exp\left(-\frac{1}{2}(t\eta_1^{-1})^2\right) \quad , t > 0, \eta_1 > 0, \eta_0 > 0 \qquad ...(2.3)$$

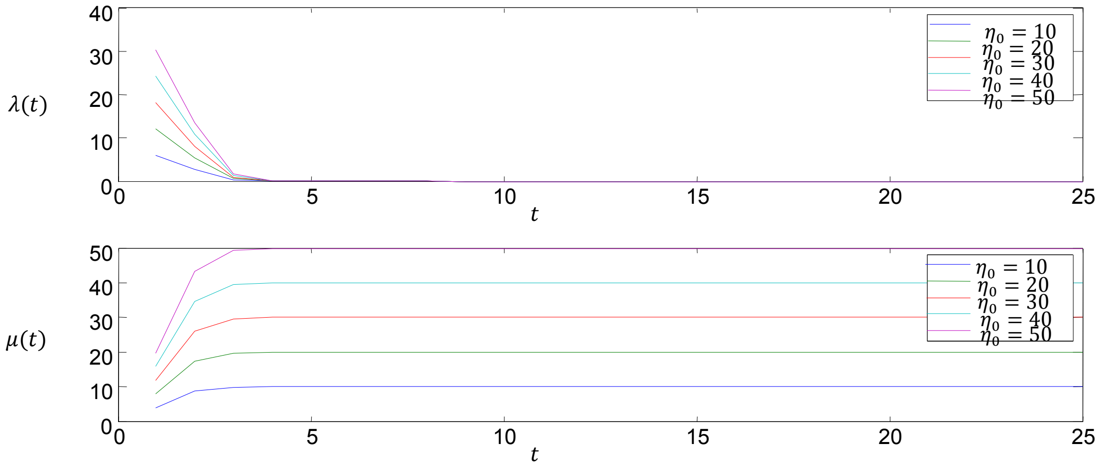
Also, the mean failures function also termed as expected number of failures at time t_e is

$$\mu(t) = \eta_0 \left[1 - exp\left(-\frac{1}{2}(t\eta_1^{-1})^2\right)\right] \qquad ...(2.4)$$

The study of the behavior of failure intensity function $\lambda(t)$ and mean failure function $\mu(t)$ has been shown in following figures.



Graph G 1.1: Behavior of failure intensity $\lambda(t)$and expected number of failures $\mu(t_e)$ upto t_e for fixed $\eta_0(= 30)$ and $t(= 1(1)50)$, $\eta_1(= 0.5(0.5)5.0)$.



Graph G 1.2: Behavior of failure intensity $\lambda(t)$and expected number of failures $\mu(t_e)$ upto t_e for fixed $\eta_1(= 1.0)$ and $t(= 1(1)50)$, $\eta_0(= 10(5)55)$.

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 5.731

Volume 2, Issue 1, January 2022

From Graphs G-1.1 and G-1.2, it has been concluded that the behavior of $\lambda(t)$ is very sensitive for values of η_1 and $\lambda(t)$ is high for the smaller values of η_1, becoming unimodal positively skewed distribution for the large values of η_1. The values of $\mu(t)$ are large for smaller failure rate $\eta_1$and decreases as η_1 increases. Further it is observed that $\lambda(t)$ is less sensitive for increasing values of η_0 and for fixed η_1. The slope of $\lambda(t)$ and $\mu(t)$ remains similar for increasing η_0. The Poisson process provides a good approximation to the occurrence of various existent(real) life situations. Under Poisson type occurrence of software failures, the process is studied using non homogeneous Poisson Process [NHPP] with failure intensity $\lambda(t)$. Let $M\ (t)$denote failure experienced by time$t$&$[M\ (t) = m]$ increases by 1 whenever a failure occurs.According to the Poisson probability law chance of getting $[M(t) = m]$ is

$$P[M\ (t) = m] = \frac{[\mu(t)]^m exp[-\mu(t)]}{m!}$$

where $\mu(t) = \int_0^t f(x)dx$with mean and variance equal to $\mu(t)$.

In present case,

$$P[M\ (t) = m] = \frac{\left\{\eta_0\left[1-exp\left(-\frac{1}{2}(t\eta_1^{-1})^2\right)\right]\right\}^m exp\left[-\eta_0\left[1-exp\left(-\frac{1}{2}(t\eta_1^{-1})^2\right)\right]\right]}{m!} \quad ...(2.5)$$

having mean and variance equal to$\eta_0\left[1 - exp\left(-\frac{1}{2}(t\eta_1^{-1})^2\right)\right]$.

If $[Q\ (t) = q]$ is expected number of remaining failures at time t is

$$P[Q\ (t) = q] = \frac{[\upsilon(t)]^m exp[-\upsilon(t)]}{m!}$$

where $\upsilon(t) = \eta_0 - \mu(t)$ with mean and variance equal to $\upsilon(t)$.

In present case,

$$P[Q\ (t) = q] = \frac{\left\{\eta_0 exp\left(-\frac{1}{2}(t\eta_1^{-1})^2\right)\right\}^m exp\left[-\eta_0 exp\left(-\frac{1}{2}(t\eta_1^{-1})^2\right)\right]}{m!} \quad ...(2.6)$$

with mean and variance $\eta_0 exp\left(-\frac{1}{2}(t\eta_1^{-1})^2\right)$.

III. Maximum Likelihood Estimators

The likelihood function of Poisson type models is depend on η_0 and η_1. Assuming the total execution time is t_e, m_e failures are experienced upto t_e at times $t_i, i = 1,2, \dots, m_e$, η_0 be the number of inherent software failures and η_1 the scale parameter in $f(t)$then the likelihood function of η_0 and η_1 obtained as (cf. Musa et al. (1987)).

$$L(\eta_0, \eta_1|\underline{t}) = \left\{\prod_{i=1}^{m_e} \lambda(t_i)\right\} exp[-\mu(t_e)]$$

Having the values of $\lambda(t)$ and $\mu(t_e)$, the likelihood function takes a form,

$$L(\eta_0, \eta_1|\underline{t}) = \eta_0^{m_e}\eta_0^{-2m_e}[\prod_{i=1}^{m_e} t_i]exp\left(-\frac{1}{2}T\eta_1^{-2}\right) exp\left\{-\eta_0\left[1 - exp\left(-\frac{1}{2}t_e^2\eta_1^{-2}\right)\right]\right\} \quad ...(3.1)$$

where

$$T = \sum_{i=1}^{m_e} t_i^2$$

The maximum likelihood estimators of η_0 and η_1 can be obtained as

$$\sum_{i=1}^{m_e}(\hat{\eta}_{m0} - i + 1)^{-1} = \frac{1}{2}(t_e\eta_1^{-1})^2 \quad ...(3.2)$$

and

$$(\hat{\eta}_{m1})^2 = \frac{1}{2}[(\hat{\eta}_{m0} - m_e)t_e^2 + T]m_e^{-1}. \quad ...(3.3)$$

The values of $\hat{\eta}_{m0}$ and $\hat{\eta}_{m1}$ are the solutions of above simultaneous equations.

IV. Bayesian Analysis: Prior Selection

The Bayesian analysis, utilizes the prior information (past experience) about parameter and produce posterior probability distribution of the parameter. The posterior obtained by this method can be used to produce efficient and consistent estimators for the parameters. Due to the accuracy in computation using computer, the Bayesian techniques

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Volume 2, Issue 1, January 2022

IJARSCT

Impact Factor: 5.731

are widely preferred. [See Martz and Waller (1982), Box and Tiao (1973), Berger (1985), etc..]. Since non-informative priors preserve objectivity, the Bayes procedure of obtaining posterior with non-informative priors is being progressively acknowledged. If one knows how a parameter arises, it is easy to estimate it even before experimentation. Such a priori values when estimated using past experience, can be used to do preliminary planning and resource allocation before testing begins. Generally, the total number of inherent faults (η_0)present in the software is unknown. Hence, due to the lack of prior knowledge about (η_0)a non-informative priors used. Non-informative prior probabilities play an important role in Bayesian analysis. Here it is assumed that η_0 and η_1 are independent. Hence, the non-informative prior for η_0 becomes more suitable and it is

$$g(\eta_0) = \begin{cases} \eta_0^{-1} & ; \quad \eta_0 > 0 \\ 0 & ; Otherwise \end{cases}$$

Sinceη_0 is unknown, the experimenter may not have guess about the how frequently the failures occur i.e. about the scale parameter of Rayleigh density. Hence, the non-informative prior for $\eta_1$becomes more suitable and it is

$$g(\eta_1) = \begin{cases} \eta_1^{-1} & ; \quad \eta_1 > 0 \\ 0 & ; Otherwise \end{cases}$$

The joint prior distribution for η_0 and η_1 becomes

$$g(\eta_0, \eta_1) = \begin{cases} \eta_0^{-1}\eta_1^{-1} & ; \quad \eta_1 > 0 \\ 0 & ; Otherwise \end{cases} \qquad ...(4.1)$$

V. POSTERIOR AND MARGINAL POSTERIOR DISTRIBUTION

The above prior probability distributions can be used to produce posterior probability distribution for the parameters. According to Bayesian procedure of obtaining posterior by combining the likelihood function (3.1) with joint prior distribution (4.1), the joint posterior distribution of η_0 and $\eta_1$given data$\underline{t}$ is,

$$\pi(\eta_0, \eta_1|\underline{t}) \propto \eta_0^{m_e-1}\eta_1^{-2m_e-1} exp\left(-\tfrac{1}{2}T\eta_1^{-2}\right) exp\left\{-\eta_0\left[1 - exp\left(-\tfrac{1}{2}t_e^2\eta_1^{-2}\right)\right]\right\}$$

$$m_e < \eta_0 < \infty, 0 < \eta_1 < \infty \qquad ...(5.1)$$

The constant of proportionality D is,

$$D = 2^{m_e-1}\Gamma(m_e)\sum_{j=0}^{\infty}\Gamma(m_e + j, m_e)\,[T + jt_e^2]^{-m_e}$$

where $\Gamma(.)$ and $\Gamma(.,.)$ are gamma and incomplete gamma function [see Gradshteyn and Ryzik (1996), Arfken (1985), Abramowitz and Stegun (1965), etc..]

The marginal posteriors of $\eta_1$is

$$\pi(\eta_1|\underline{t}) \propto \eta_1^{-2m_e-1} exp\left(-\tfrac{1}{2}T\eta_1^{-2}\right)\frac{\Gamma\left(m_e,\left[1-exp\left(-\frac{1}{2}t_e^2\eta_1^{-2}\right)\right]m_e\right)}{\left[1-exp\left(-\frac{1}{2}t_e^2\eta_1^{-2}\right)\right]^{-m_e}}; \quad 0 < \eta_1 < \infty \qquad ...(5.2)$$

Similarly, the marginal posteriors of $\eta_0$is

$$\pi(\eta_0|\underline{t}) \propto \eta_0^{m_e+j-1} exp(-\eta_0) I_{5.1}; \quad m_e < \eta_0 < \infty,$$

where

$$I_{5.1} = \int_0^{\infty} \eta_1^{-2m_e-1} exp\left(-\tfrac{1}{2}T\eta_1^{-2}\right) exp\left\{-\eta_0\left[1 - exp\left(-\tfrac{1}{2}t_e^2\eta_1^{-2}\right)\right]\right\} d\eta_1$$

The integrand $I_{5.1}$ can be solved as

$$I_{5.1} = \eta_0^j exp(-\eta_0)\sum_{j=0}^{\infty} 2^{m_e-1}\Gamma(m_e)[T + jt_e^2]^{-m_e}$$

VI. BAYES ESTIMATORS

The Bayes estimators can be obtained by using marginal posterior obtained in previous section after choosing loss function. In present study, square error loss has been chosen and under this loss function the Bayes estimator becomes posterior mean. The Bayes estimator i.e. posterior mean ofη_0 is

$$\hat{\eta}_{B0} \propto 2^{m_e-1}\Gamma(m_e)\sum_{j=0}^{\infty}\Gamma(m_e + j + 1, m_e)\,[T + jt_e^2]^{-m_e} \qquad ...(6.1)$$

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Impact Factor: 5.731
Volume 2, Issue 1, January 2022

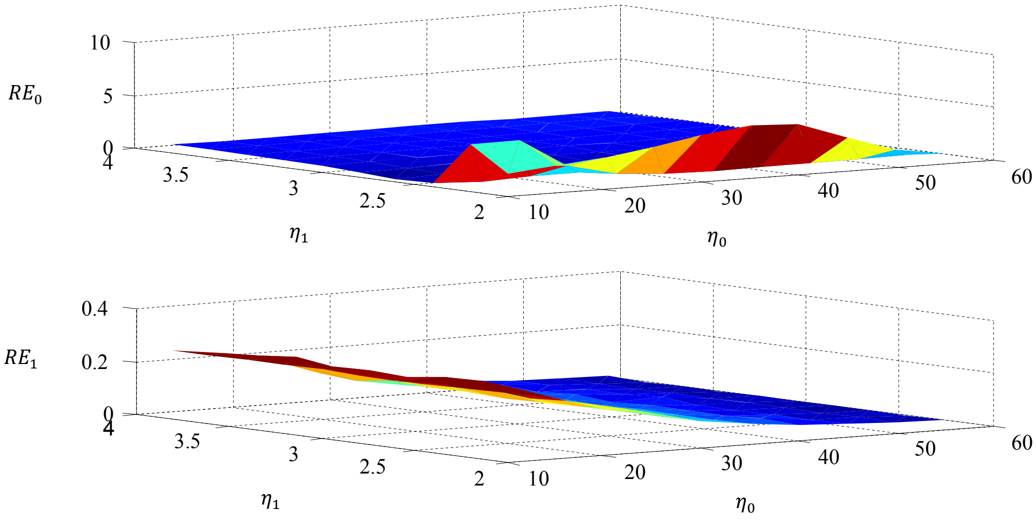
Similarly, The Bayes estimator i.e. posterior mean ofη_1 is

$$\hat{\eta}_{B1} \propto 2^{(m_e^{'}-1)}\Gamma(m_e^{'})\sum_{j=0}^{\infty}\Gamma(m_e+j,m_e)\,[T+jt_e^2]^{-m_e^{'}} \qquad \text{...(6.2)}$$

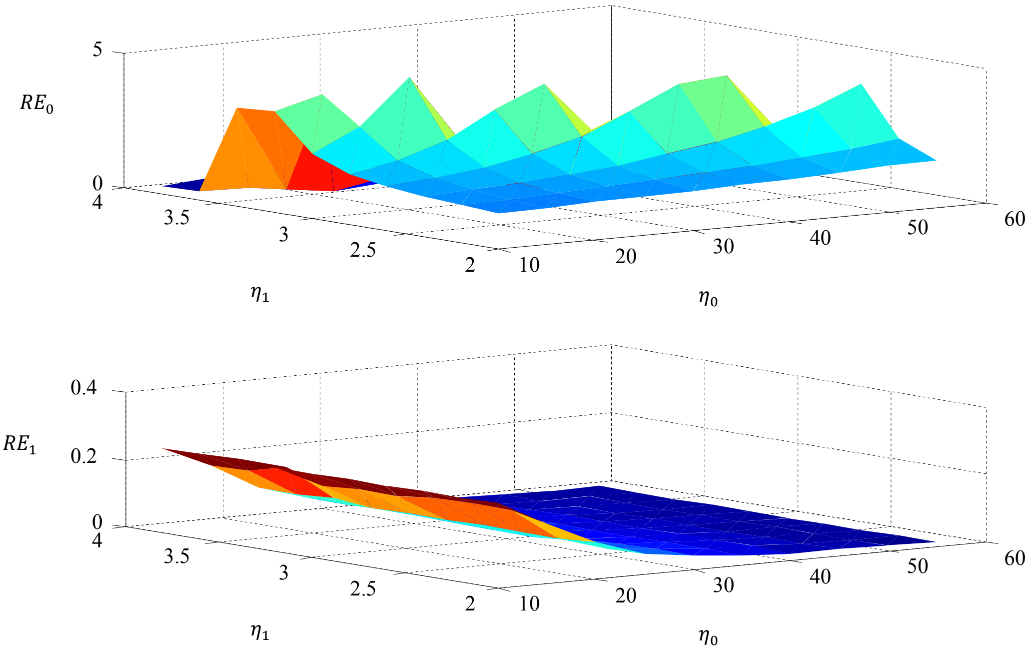
where $m_e^{'} = \left(m_e - {}^{1}\!/_{2}\right)$

VII. Discussion

The proposed Bayes estimators i.e. $\hat{\eta}_{B0}$ and $\hat{\eta}_{B1}$are compared with corresponding maximum likelihood estimators i.e. $\hat{\eta}_{m0}$ and $\hat{\eta}_{m1}$ resp. on the basis of risk efficiencies $RE_j = R_j^{'}R_j^{-1}$ where $R_j = E\left[\hat{\eta}_{Bj} - \eta_j\right]^2$ and $R_j^{'} = E\left[\hat{\eta}_{mj} - \eta_j\right]^2$; $j = 0{,}1$. The proposed Bayes estimators are depends upon the execution time t_e. The performance of proposed estimators has been checked by generating sample of m_e failures generated 10^3 times using inverse transformation method. The risk efficiencies are evaluated by Monte Carlo Simulation technique. The performance of proposed Bayes estimators by means of risk efficiencies has been summarized in the form of graphs from G-7.1 to G-7.8.



Graph (G-7.1): Risk Efficiencies of $\hat{\eta}_{B0}$ and $\hat{\eta}_{B1}$,$t_e = 50$; $\eta_1 = 2.0(0.2)3.8$ and $\eta_0 = 10(5)55$



Graph (G-7.2): Risk Efficiencies of $\hat{\eta}_{B0}$ and $\hat{\eta}_{B1}$,$t_e = 125$; $\eta_1 = 2.0(0.2)3.8$and $\eta_0 = 10(5)55$

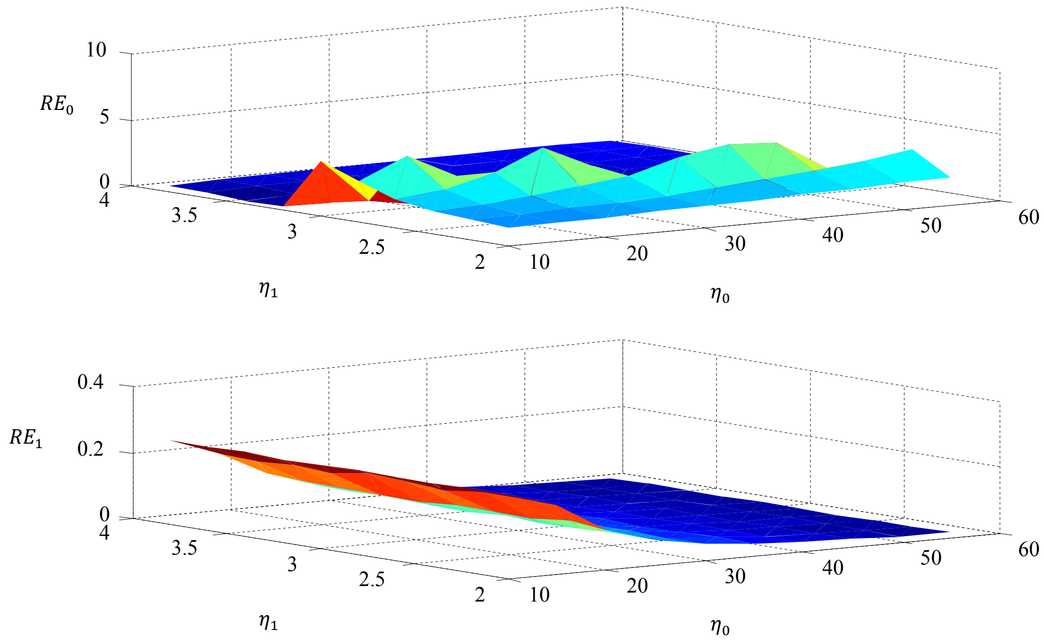
IJARSCT

ISSN (Online) 2581-9429

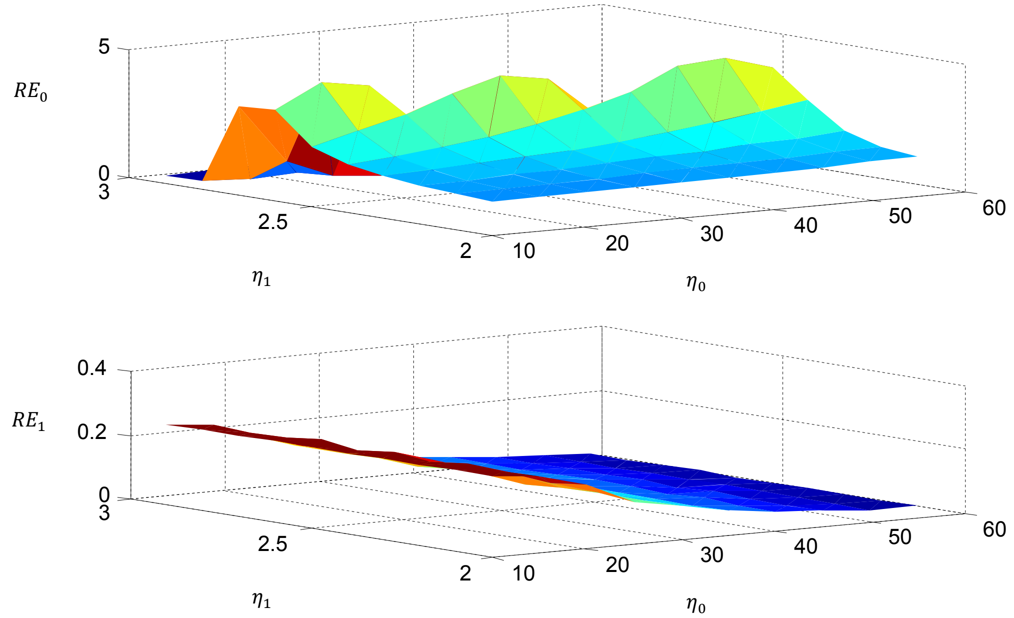
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Volume 2, Issue 1, January 2022

Impact Factor: 5.731



Graph (G-7.3): Risk Efficiencies of $\hat{\eta}_{B0}$ and $\hat{\eta}_{B1}$,$t_e\, =\, 100;\, \eta_1\, =2.0(0.1)2.9$and $\eta_0\, =10(5)55$



Graph (G-7.4): Risk Efficiencies of $\hat{\eta}_{B0}$ and $\hat{\eta}_{B1}$,$t_e\, =\, 125;\, \eta_1\, =2.0(0.1)2.9$and $\eta_0\, =10(5)55$

From graph G-7.1 to G-7.4, it can be seen that the relative efficiency of proposed Bayes estimator $\hat{\eta}_{B0}$ is always better than MLE whenever η_0 and η_1 are small. It can be seen that as the value of η_1 increases the value of RE_0 first increases and then starts decreasing after attaining maxima. That is the performance of $\hat{\eta}_{B0}$ will be better for optimum true value of η_1. As η_0 increases, the risk efficiency RE_0 again has the similar trend as RE_0 for increase of η_1. This means, for proper choice of η_0 the performance of $\hat{\eta}_{B0}$ is best. It can be concluded that for larger value of η_0 and η_1, the Bayes estimator $\hat{\eta}_{B0}$ cannot be preferred. The Bayes estimator should be preferred when the value of η_0 is small. As t_e increases, the risk efficiency decreases. The Bayes estimator $\hat{\eta}_{B1}$ is not better than MLE for almost all the considered choices of η_0, η_1 and t_e as far as the prior is uniform.

VIII. CONCLUSION

From the above conclusion can be drawn for preferring Bayes estimator $\hat{\eta}_{B0}$ for a proper choice of t_e and suitable expected true value of η_0 and η_1.The proposed Bayes estimator $\hat{\eta}_{B0}$ can be preferred if η_1 is little high and t_e is also large.The proposed Bayes estimator $\hat{\eta}_{B1}$ cannot preferred over MLE.

REFERENCES

[1]. Abramowitz, M., and Stegun, I. A., (1965), “Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables”, Dover publications, New York.

[2]. Arfken, G., (1985), "The Gamma Function (Factorial Function)", in Mathematical Methods for Physicists, 3rd ed. Academic Press, FL.

[3]. Berger, J. O. (1985), “Statistical decision theory and Bayesian analysis”, Springer-Verlag, New York.

[4]. Boehm, B., Abts, C., Winsor Brown, A., Chulani, S., Clark, B. K., Horowitz, E., Madachy, R., Reifer, D. J. and Steece, B. (2000), “Software Cost Estimation with COCOMO II (with CD-ROM)”, Englewood Cliffs, Prentice-Hall, New Jersey.

[5]. Box, G. E. P. and Tiao, G. C. (1973), “A Bayesian inference in statistical analysis”, Addison Wesley reading, Massachusetts.

[6]. DeMarco, T. (1982), “Controlling Software Projects: Management, Measurement and Estimation”, Englewood Cliffs, Prentice-Hall, New Jersey.

[7]. Gaffney, J. and Pietrolewicz, J. (1990), “An Automated Model for Early Error Prediction in Software Development Process,” Proceeding of IEEE Software Reliability Symposium Colorado Spring-1990.

[8]. Gradshteyn, I. S. and Ryzhik, I. M. (1994), “Table *of Integrals, Series, and Products”,* Alan Jeffrey (editor) Fifth edition, Academic Press, New York.

[9]. Kan, S. H. (2002), “Metrics and Models in Software Quality Engineering”, Second Edition, Addison Wesley, New York.

[10]. Martz, H. F. and Waller, R. A. (1982), “Bayesian reliability analysis”, Wiley, New York.

[11]. Musa, J. D., Iannino, A. and Okumoto K. (1987), “Software Reliability: Measurement, Prediction, Application”, McGraw-Hill, New York.

[12]. Putnam, L. H. (1978), “A general empirical solution to the macro software sizing and estimating problem,” IEEE Transactions on Software Engineering, Vol.- 4, pp. 345–367.

[13]. Schick, G. J. and Wolverton, R. W. (1978a), "Assessment of software reliability," Presented at the 11th Annual Meeting of the German Operations Research Society, DGOR, Hamburg, Germany,In Proceeding of Operations Research, Physica-Verlag,Wurzburg-Wien, 1973, pp. 395-422.

[14]. Schick, G. J. and Wolverton, R. W. (1978b), “An Analysis of Competing Software Reliability Models”, IEEE Transactions on Software Engineering, Vol.SE-4, NO.-2,pp. 104-120.

[15]. Singh, R. and Andure, N. W. (2008), “Bayes estimators for the parameters of the Poisson type exponential distribution”, IAPQR transactions, Vol.- 33, No.-2, pp. 121-128.

[16]. 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., Vol-4, No.-2, pp 83-89.

[17]. Sinha, S. K. (1985), “Reliability and life testing”, Wiley, New Delhi.

[18]. Trachtenberg, M. (1982), "Discovering How to Ensure Software Reliability," RCA Engineer, pp. 53–57.

[19]. Yamada, S., Ohba, M. and Osaki, S. (1983), “S-shaped reliability growth modeling for software error detection”, IEEE Transactions on Reliability, Vol.-R-32, No.-5, pp. 475–484.

[20]. Yamada, S., Ohtera, H. and Narihisa, H. (1986), “Software Reliability Growth Models with Testing- Effort”, IEEE Transactions on Reliability, Vol.-35, pp. 19 – 23.