A COMPARISON OF PRIORS FOR THE PARAMETER OF THE TIME-TO-FAILURE MODEL

M. Tahir, and M. Saleem

Department of Statistics, Government College University, Faisalabad 38000, Pakistan

ABSTRACT: The objective of this study was to compare the Bayes estimates of the parameter of the time-to-failure model based on informative and non-informative priors. The Gamma distribution wasassumed as the informative conjugate prior while the Jeffreys prior was assumed as an uninformative prior. The comparison is based on the posterior variance, the Bayesian interval estimate, the coefficient of skewness of the posterior distribution and the Bayes posterior risk.

Keywords: Jeffreys prior; Elicitation; Prior predictive distribution; Bayes posterior risk

INTRODUCTION

The Bayesian approach has several advantages over the classical approach because it can utilize the prior information in a formal way, satisfies the axioms of coherence and utilize decision theory. This study provides a Bayesian analysis of the time-to-failure model using informative (Gamma) and uninformative (Jeffreys) priors. A method is also given to elicit the hyperparameters of the prior density for the parameters of the said model. Kadane *et al*. (1980), Chaloner and Duncan (1983), Gavasakar (1988), Aslam (2003), Al-Awadhi and Gartwaite (1998), Kadane and Wolfson (1998) and Hahn (2006) discussed prior elicitation methods including the one based on prior predictive distribution. The comparison of the informative and non-informative priors with respect to posterior variance, Bayesian interval estimate, coefficient of skewness for posterior distribution and Bayes posterior risk is presented.

MATERIALS AND METHODS

The Posterior Distribution of the Parameter using Informative Prior (IP): The distribution of the time-to-failure system usually follows the exponential distribution

$$f(x) = \lambda\, e^{-\lambda x},\ 0 < x < \infty,\ \lambda > 0 \qquad (1)$$

It is to be assumed that the prior distribution of λ is a Gamma distribution with hyperparameters 'a' and 'b', so the prior distribution of λ is as under.

$$p(\lambda) = \frac{b^a}{\Gamma(a)} \lambda^{a-1} e^{-b\lambda},\ 0 < \lambda < \infty,\ a,b > 0. \qquad (2)$$

Hence the posterior distribution of λ $p(\lambda|x) \propto \lambda^{\alpha-1} e^{-\beta\lambda},\ x > 0,\ \lambda > 0$ for the given data $(\boldsymbol{x} = x_1, x_2, \ldots, x_n)$ is as bellow. (3)

This is the density kernel of the Gamma distribution with parameters $\alpha = a + n$ and $\beta = b + \sum_{i=1}^{n} x_i$. So the posterior distribution of λ given data is $Gamma(\alpha, \beta)$.

The Elicitation of Hyperparameters: The prior predictive distribution (PPD) of a random variable X is defined as under.

$$p(x) = \int_0^\infty p(x|\lambda)\, p(\lambda)\, d\lambda \qquad (4)$$

So PPD for an exponential random variable X takes the following form.

$$p(x) = \frac{ab^a}{(b+x)^{a+1}},\quad x > 0,\quad a, b > 0 \qquad (5)$$

which is $Eg(a, b)$ here Eg is exponential-gamma distribution. Using the exponential-gamma distribution as a PPD and comparing it with the expert's assessment of this distribution, we choose those hyper-parameters which make the assessment agree closely with a member of the family. The prior predictive probabilities of cycles to failure (in ten thousands) over the intervals (0, 3) and (3, 6) for a large number of heater switches subject to an overload voltage, according to an expert, say, are 0.69 and 0.20 respectively. A program written in Mathematica package for eliciting the hyper-parameters of prior density is given in appendix. The elicited values of the hyper-parameters *a* and *b* are 8.9936 and 21.5698 respectively.

The Posterior Distribution using the Informative Prior: Having elicited the hyper-parameters, the prior distribution of λ is $G(8.9936,\ 21.5698)$. Consider the following random sample of cycles to failure (in ten thousands) for 20 heater switches subject to an overload voltage taken from Kapur and Lamberson (1977): 0.0100, 0.0340, 0.1940, 0.5670, 0.6010, 0.7120, 1.2910, 1.3670, 1.9490, 2.3700, 2.4110, 2.8750, 3.1620, 3.2800, 3.4910, 3.6860, 3.8540, 4.2110, 4.3970, 6.4730. Here $n=20$ and $\sum_{i=1}^{20} x_i = 46.9350$. Hence the posterior distribution of parameter λ for given data $(\boldsymbol{x} = x_1, x_2, \ldots, x_{20})$ is $Gamma\left(28.9936,\ 68.5048\right)$.

The Posterior Distribution using the Non-Informative Priors (NIP): A non-informative prior has been suggested by Jeffreys (1946, 1961), which is frequently used in situation where one does not have much information about the parameters. The Jeffreys prior for the parameter λ is as under.

$$p_J\left(\lambda\right) \propto 1/\lambda,\ \ 0<\lambda<\infty \tag{6}$$

The posterior distribution of parameter λ is $Gamma\left(n,\ \sum_{i=1}^{n} x_i\right)$. Using the data given in Section 2.3, the posterior distribution of parameter λ is $Gamma\left(20,\ 46.9350\right)$.

Posterior Variance: Prior information is incorporated with the likelihood to find posterior distribution which is the basis of Bayesian inference. Posterior variance based on IP and NIP are compared to assess relative efficiency of the Bayes estimates.

Bayesian Interval Estimate: When $X_1, \ldots, X_n$ are independent and identically distributed exponential random variables, Kapur and Lamberson (1977) showed that $2\lambda\left(b+\sum_{i=1}^{n} x_i\right)$ has a chi-square distribution with $2\left(a+n\right)$ degrees of freedom. By using the posterior distribution, a $100\left(1-C\right)\%$ highest density region (HDR) for parameter λ is as under.

$$\frac{\chi^2_{2(a+n)\ (1-\frac{C}{2})}}{2\left(b+\sum_{i=1}^{n} x_i\right)} < \lambda < \frac{\chi^2_{2(a+n)\ (\frac{C}{2})}}{2\left(b+\sum_{i=1}^{n} x_i\right)},\ \lambda>0,\ x\ >\ 0 \tag{7}$$

Coefficient of Skewness: The coefficients of skewness is calculated from the posterior distribution as $\gamma_1 = 2\sqrt{1/\alpha}$.

Bayes Posterior Risk: The expected value of loss function for a given decision with respect to the posterior distribution is known as posterior risk function and if d^* is a Bayes estimator then $\rho\ (d^*)$ is known as Bayes posterior risk and is defined as bellow.

$$\rho\left(d^*\right) = E_{\lambda|\boldsymbol{x}}\left[L\left(\lambda,\ d^*\right)\right]. \tag{8}$$

RESULTS AND DISCUSSION

Results are presented in Tables 1-4. Hyper-parameters assumed are 8.9936 and 21.5698.

Comparison of Priors with respect to Posterior Variance: The Posterior variance of parameter λ is given in Table 1 which reveals that informative prior provides more efficient estimates.

Comparison of Priors based on Bayesian Interval Estimate: The Bayesian interval estimate of parameter λ is presented in the following Table 2 which depicts that the posterior estimate based on informative prior is more efficient.

Comparison of Priors using Coefficient of Skewness: This section provides the comparison of priors using coefficient of skewness. Table 3 shows that $\gamma_1>0$, therefore, the posterior distribution based on informative and non-informative priors are slightly positively skewed but the skewness is least in case of informative prior.

Table 1: Posterior Variances with IP and NIP

Parameter	Variance using	
	IP	NIP
λ	0.0062	0.0091

Table 2: Bayesian Interval Estimates

Prior Distribution	95% HDR	99% HDR
using IP	(0.2835, 0.5907)	(0.2482, 0.6531)
using NIP	(0.2603, 0.6322)	(0.2206, 0.7113)

Comparison of Priors using Bayes Posterior Risk: Bayes posterior risks using informative prior for different loss functions are lesser than their corresponding Bayes posterior risks using uninformative (Jeffreys) prior as shown in Table 4.

Hence, once again, the superiority of the informative prior is established. The comparison of the informative and non-informative priors with respect to posterior variance, the Bayesian interval estimate, the coefficient of skewness of the posterior distribution and the Bayes posterior risk shows that the informative prior is more advantageous than the uninformative prior.

Table 3: Coefficient of Skewness for Posterior Distribution.

Prior Distribution	Posterior Parameters (α, β)	Coeff of Skewness γ_1
using IP	(28.9936, 68.5048)	0.3714
using NIP	(20.0000, 46.9350)	0.4472

Table 4: Bayes Posterior Risks for Different Loss Functions

Loss Function $L(\lambda, d)$	Bayes Posterior Risk $\rho\ (d^*)$	Prior Distribution using	Bayes Posterior Risk $\rho\ (d^*)$
$\left(1-\frac{d}{\lambda}\right)^2$	$\frac{1}{\alpha-1}$	IP	0.0357
		NIP	0.0526
$\frac{(\lambda-d)^2}{\lambda}$	$\frac{1}{\beta}$	IP	0.0146
		NIP	0.0213
$(\lambda-d)^2$	$\frac{\alpha}{\beta^2}$	IP	0.0062
		NIP	0.0091

APPENDIX

Programme to Elicit the Hyperparameters of Prior Density

Abs[Function[a,a=Function[b,b=(

$$\sum_{x=1}^{3}\left(\frac{1}{x!}\frac{b^a}{Gamma[a]}\frac{Gamma[a+x]}{[b+1]^{a+x}}\right)$$ **- 0.54)+(**

$$\sum_{x=4}^{6}\left(\frac{1}{x!}\frac{b^a}{Gamma[a]}\frac{Gamma[a+x]}{[b+1]^{a+x}}\right)$$ **-0.25)]**

[Table[{j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j},
{j,55.0079,55.0100,.0001}]]][Range[87.0080,87.0100,
{.0001,.0001}]]]

Min[Abs[Function[a,a=Function[b,b=(

$$\sum_{x=1}^{3}\left(\frac{1}{x!}\frac{b^a}{Gamma[a]}\frac{Gamma[a+x]}{[b+1]^{a+x}}\right)$$ **-0.54)+(**

$$\sum_{x=4}^{6}\left(\frac{1}{x!}\frac{b^a}{Gamma[a]}\frac{Gamma[a+x]}{[b+1]^{a+x}}\right)$$ **-0.25)]**

[Table[{j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j},
{j,55.0079,55.0100,.0001}]]][Range[87.0080,87.0100,
{.0001,.0001}]]]]

Position[Abs[Function[a,a=Function[b,b=(

$$\sum_{x=1}^{3}\left(\frac{1}{x!}\frac{b^a}{Gamma[a]}\frac{Gamma[a+x]}{[b+1]^{a+x}}\right)$$

-0.54)+($\sum_{x=4}^{6}\left(\frac{1}{x!}\frac{b^a}{Gamma[a]}\frac{Gamma[a+x]}{[b+1]^{a+x}}\right)$ **-0.25)]**

[Table[{j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j},
{j,55.0079,55.0100,.0001}]]][Range[87.0080,87.0100,
{.0001,.0001}]]],Min[Abs[Function[a,a=Function[b,

b=($\sum_{x=1}^{3}\left(\frac{1}{x!}\frac{b^a}{Gamma[a]}\frac{Gamma[a+x]}{[b+1]^{a+x}}\right)$ **-0.54)+(**

$$\sum_{x=4}^{6}\left(\frac{1}{x!}\frac{b^a}{Gamma[a]}\frac{Gamma[a+x]}{[b+1]^{a+x}}\right)$$ **-0.25)]**

[Table[{j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j},
{j,55.0079,55.0100,.0001}]]][Range[87.0080,87.0100,
{.0001,.0001}]]]]]

Table[{j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j,j},
{j,55.0079,55.0100,.0001}]
%[[17,19]]
Range[87.0080,87.0100,{.0001,.0001}]
%[[17,19]]

REFERENCES

Aslam, M. An Application of Prior Predictive Distribution to Elicit the Prior Density. Journal of Statistical Theory and Applications, 2: 70-83. (2003).

Al-Awadhi, S. A. and P. H. Gartwaite. Prior distribution assessment for a multivariate normal distribution: an experimental study. Communication in Statistics, Part A-Theory and Methods, 27: 1123-1142. (1998).

Chaloner, K. M. and G. T. Duncan. Assessment of beta prior distribution: PM elicitation. The Statistician, 32: 174-180. (1983).

Gavasakar, U. A comparison of two elicitation methods for a prior distribution for a binomial parameter. Management Science, 34: 784-790. (1988).

Hahn, E. D. Re-examining informative prior elicitation through the lens of Markov chain Monte Carlo methods. Journal of the Royal Statistical Society, Series A, 169: 37-48. (2006).

Jeffreys, H. An Invariant Form for the Prior Probability in Estimation Problems, Proceeding of the Royal Society of London, Series A, 186: 453-461. (1946).

Jeffreys, H. *Theory of Probability*. Claredon Press, Oxford, UK. (1961).

Kadane, J. B., J. M. Dickey, R. L.Winkler, W. Smith, and S. C. Peter. Interactive elicitation of opinion for a normal linear model. Journal of the American Statistical Association, 75: 845-854. (1980).

Kadane, J. B. and L. J. Wolfson. Experience in elicitation. The Statistician 47**(**1): 3-19. (1998).

Kapur, K. C. and L. R. Lamberson. Reliability in Engineering Design. Wiley & Sons, Inc. New York. (1977).