Comparison of Non-Informative Priors for Number of Defects (Poisson) Model

Muhammad Tahir

Department of Statistics, Government Degree College, Pindi Gheb, Pakistan.

Email: tahirqaustat@yahoo.com

and

Zawar Hussain

Department of Statistics, Quaid-i-Azam University 45320, Islamabad 44000, Pakistan.

Email: zhlangah@yahoo.com

Abstract

In this study the comparison of non-informative (Jeffreys and Uniform) priors is made. Rossman, et al (1998) and Elfessi and Reineke (2001) studied the relationship between Bayesian and classical estimation using the continuous uniform and exponential distribution In this study, we base our comparison of the two non-informative priors on the posterior variance, the Bayesian point estimates, the coefficients of skewness of posterior and the Bayes estimators.

Key words: Bayesian analysis; Non-Informative prior; Posterior distribution; Prior predictive distribution; Classical estimator; Bayes estimator.

1. INTRODUCTION

This study provides a Bayesian analysis of the model for the number of defects. In this study we considered the two non-informative priors (Jeffrey's and Uniform) and studied their performance using different performance measures. The posterior distribution and posterior predictive distribution for the parameters of the model for the number of defects are also derived using the above said priors.

The posterior distributions of the parameter (number of defects of the system) using the above said priors and relationships between Bayesian & classical counterparts are presented in Sections 2 and 3. The comparison of these non-informative priors with respect to posterior variance is made in Section 4 while the comparison of same non-informative priors based on

Bayesian point estimates is carried out in Section 5. In Sections 6 comparison of these non-informative priors using coefficient of skewness for posterior distribution are given. Finally, the comparison of priors with respect to Bayes estimators is made in Section 7.

2. The Posterior Distribution of the Parameter using the Jeffreys Prior (JP)

Usually, the distribution of the discrete time-to-failure system follows the Poisson distribution. So the probability mass function (p.m.f) of the Poisson distribution for a random variable X having parameter λ is:

$$f\left(x_i\right)=\frac{e^{-\lambda}\lambda^{x}}{x!}, \qquad x=0,1,2,\cdots,\ \lambda>0. \tag{2.1}$$

The likelihood function for a simple random sample of size n is given by:

$$L\left(x,\lambda\right)=\prod_{i=1}^{n} f\left(x_i\right)=\frac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n}x_i}}{\prod_{i=1}^{n}x_i}, \tag{2.2}$$

here the parameter λ is unknown.

In the situations, where one does not have much information about the parameters, Jeffreys (1946, 1961), suggested a non-informative prior (NIP). This defines the density of the parameters proportional to the square root of the determinant of the Fisher information matrix. Symbolically, the Jeffreys prior distribution $\left[p_j\left(\lambda\right)\right]$ is given by

$$p_j\left(\lambda\right)\propto\sqrt{\det\left\{\mathrm{I}\left(\boldsymbol{\lambda}\right)\right\}}, \tag{2.3}$$

where $\boldsymbol{\lambda}=\left(\lambda_1,\lambda_2,...,\lambda_k\right)^{t}$ a vector of parameters, 'det' denotes the determinant and $I\left(\boldsymbol{\lambda}\right)$ is the $\left(k\times k\right)$ Fisher information matrix given by

$$\mathrm{I}_{ij}\left(\lambda\right)=-E\left(\frac{\partial^2 \ln L\left(\lambda\right)}{\partial\lambda_i\partial\lambda_j}\right). \tag{2.4}$$

The Jeffrey's prior (see Berger (1985)) for an exponential distribution with parameter λ is:

$$p_J\left(\lambda\right)\propto\lambda^{-\frac{1}{2}}, \qquad 0<\lambda<\infty \tag{2.5}$$

The posterior distribution of parameter λ for given data $\left(\boldsymbol{x}=x_1,x_{2,}...x_n\right)$ using equations (2.2) and (2.5) is:

$$p(\lambda|\boldsymbol{x}) \propto L(\boldsymbol{x},\lambda)\, p(\lambda)$$
$$\propto \lambda^{\sum_{i=1}^{n} x_i + \frac{1}{2} - 1} e^{-n\lambda}, \qquad (2.6)$$

which is the density kernel of the gamma distribution with parameters $\sum_{i=1}^{n} x_i + \frac{1}{2}$ and n. So the posterior distribution of λ given data is $Gamma\left(\sum_{i=1}^{n} x_i + \frac{1}{2}, n\right)$.

The following data of size 10 is generated from Poisson distribution with parameter λ=2: 1, 2, 3, 1, 2, 0, 5, 3, 1, and 2. The sum of all the 10 observations is 20 (i.e. $n = 10$, $\sum_{i=1}^{n} x_i = 20$). So the posterior distribution of parameter λ for the given data $(\boldsymbol{x} = x_1, x_{2,} \ldots x_{10})$, using equation (2.6) is the Gamma distribution with parameters $\alpha = 20.50$ and $\beta = 10.00$ i.e., $Gamma(20.50, 10)$.

3. The Posterior Distribution of the Parameter using the Uniform Prior (UP)

Laplace (1774, 1812) found for the problems he encountered, that, it worked exceptionally well to simply always choose the prior for λ to be the constant $\{p(\lambda) = 1\}$ on the parameter space.

The Uniform prior (non-informative prior) distribution of λ is:

$$p(\lambda) \propto 1, \qquad 0 < \lambda < \infty. \qquad (3.1)$$

The posterior distribution of parameter λ for given data $(\boldsymbol{x} = x_1, x_{2,} \ldots x_n)$ using equations (2.2) and (3.1) is:

$$p(\lambda|x) \propto \lambda^{\sum_{i=1}^{n} x_i + 1 - 1} e^{-n\lambda}, \qquad (3.2)$$

which is the density kernel of the gamma distribution with parameters $\sum_{i=1}^{n} x_i + 1$ and n. Hence the posterior distribution of λ given data is $Gamma\left(\sum_{i=1}^{n} x_i + 1, n\right)$.

The posterior mode $\frac{\sum_{i=1}^{n} x_i}{n}$, of the Gamma distribution using Uniform prior is equal to its classical counterparts; the Maximum Likelihood Estimator (MLE) and the Uniformly Minimum Variance Unbiased Estimator (UMVUE). Hence for the data considered above, the posterior distribution of parameter λ for given data $\left(\boldsymbol{x} = x_1, x_{2,} \ldots x_{10}\right)$ using equations (3.2) is a Gamma distribution with parameters $\alpha = 21.00$ and $\beta = 10.00$ i.e. $Gamma\left(21.00,10\right)$.

4. Comparison of Non-Informative Priors with respect to Posterior Variance

The Posterior variances of parameter λ using the two non-informative priors are given in the following Table 1.

Table 1: Posterior Variance of Parameter λ

	Variance using	
	NIP	
	JP	UP
λ	0.2050	0.2100

From Table 1, it is obvious that $Var\left(\lambda\right)$ using Jeffrey's prior is equal to $Var\left(\lambda\right)$using Uniform prior. That is, the Jeffrey's and the Uniform priors are approximately equally efficient. So either of them can be used as a non-informative prior and hence robustness with respect to the choice of non-informative prior is observed. However, it is reasonable to prefer the Uniform prior being simpler as compared to the Jeffreys prior.

5. Comparison based on Bayesian Point Estimates

The Bayesian point estimates of λ are presented in Table 2. Classical counterparts are also given in Table 2.

From Table 2, we conclude that both the posterior mode and posterior mean using the two priors are almost the same as the MLE and UMVUE.

Table 2: Bayesian point estimates using non-informative priors

Bayesian Estimates				Classical Counterpart (MLE & UMVUE)
		JP	2.05	2.00
		UP	2.10	2.00
		JP	1.95	2.00
		UP	2.00	2.00

6. Comparison of Priors using Coefficient of Skewness

This section provides the comparison of priors using coefficient of skewness (c. o. s). The coefficients of skewness are calculated from the posterior distributions and are discussed below. The coefficient of the posterior distribution is given by:

$$\text{Coefficient of Skewness} = \gamma_1 = 2\sqrt{\frac{1}{\alpha}}$$

Table 3**:** Coefficient of Skewness for Posterior Distribution

		Posterior Parameters (α, β)	Coeff of Skewness γ_1
	JP	(20.50, 10.00)	0.4417
	UP	(21.00, 10.00)	0.4365

From Table 3, it is observed that $\gamma_1 > 0$, therefore, the posterior distributions based on the Jeffreys and the Uniform priors are not symmetrical, rather they both are slightly positively and almost equally skewed. However, because of the simplicity the Uniform prior may be preferred to the Jeffreys prior

7. Comparison of Priors using Bayes Estimator

Bayes decision is a decision $'d^*'$ which minimizes risk function and $'d^*'$ is the best decision. If the decision is choice of an estimator then the Bayes decision is a Bayes estimator. Table 5 summarizes the Bayes estimator based on non-informative (Jeffreys and Uniform) priors for different loss functions.

Table 4: Bayes estimator for different loss functions

Loss Function $L(\lambda, d)$	Bayes Estimator d^*		Posterior Parameters (α, β)	Bayes Estimator d^*	
$L_1 = \left(1 - \frac{d}{\lambda}\right)^2$	$\frac{\alpha - 2}{\beta}$	Jeffreys Prior	(20.50, 10.00)	1.85	2.00
		Uniform Prior	(21.00, 10.00)	1.90	2.00
$L_2 = \frac{(\lambda - d)^2}{\lambda}$	$\frac{\alpha - 1}{\beta}$	Jeffreys Prior	(20.50, 10.00)	1.95	2.00
		Uniform Prior	(21.00, 10.00)	2.00	2.00
$L_3 = (\lambda - d)^2$	$\frac{\alpha}{\beta}$	Jeffreys Prior	(20.50, 10.00)	2.05	2.00
		Uniform Prior	(21.00, 10.00)	2.10	2.00

From Table 4, it can be seen that the Bayes estimator for different loss functions L_1, L_2 and L_3, using the two priors are almost equal to UMVUE and MLE.

8. Conclusion

In this article, we examined the relative performance of the two non-informative priors – the Uniform prior and Jeffery's prior—using the different performance measures. Both the priors are equally efficient no matter what is the performance criterion. The Bayes' estimators are also calculated for a given set of data. It is observed that Uniform prior may be preferred due to its simplicity. The Bayes' estimators are almost equal to classical estimators.

References

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*. Second Edition. Springer-Verlag New York, Inc.

Bernardo, J. and A. Smith (1994). *Bayesian Theory*. John Wiley & Sons.

Elfessi, A., and Reineke, D. M. (2001).A Bayesian Look at Classical Estimation: The Exponential Distribution. *Journal of Statistics Education* **9**, 1.

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

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

Laplace, P. (1774). Me'moire sur la Probabilite' des Causes par les Evenements, *Mem Acad. R. Sci*. Presente's par Divers Savans, **6**, 621-656, {translated in Statistical Science (1986), 359-378}.

Laplace, P. (1812). The'orie Analytique des Probabilite's, Paris: courcier.

Lindley, D. V. (1972). *Bayesian Statistics: A Review*, Philadelphia. PA. Society for Industrial and Applied Mathematics.

Rossman, A. J., Short, T. H. and Parks, M. T. (1998). Bayes Estimators for the Continuous Uniform Distribution. *Journal of Statistics Education* **6**, No. 3.