Pak. j. life soc. Sci. (2010), 8(2): 159-161

Pakistan Journal of
Life and Social Sciences

On Relationship between some Bayesian and Classical Estimators

Muhammad Tahir and Muhammad Saleem
Department of Statistics, Government College University, Faisalabad, Pakistan

Abstract

In this study, we discussed the functional relationship between Bayesian and classical estimators of parameters of some discrete and continuous probability distributions. The maximum likelihood estimator (MLE) is used as representative of classical inference while posterior mean and posterior mode represent Bayesian estimators under squared error loss and zero one loss functions respectively. The posterior means (modes) are expressed as a function of corresponding likelihood estimators and prior means (modes). This functional relationship depicts that maximum likelihood estimator can be considered as a special case of their Bayesian counterpart if values of the hyper-parameters are set to zero. Further the relationship is identical for all the distributions except Maxwell and Rayleigh.

Keywords: Squared error loss function; Zero one loss function; Maximum likelihood estimator; Posterior mean; Posterior mode; Conjugate prior

Introduction

The relationship between Bayesian and classical estimation using the continuous uniform distribution and exponential distribution respectively described by Rossman, *et al*. (1998) and Elfessi and Reineke (2001) . Aslam (2003) and Hahn (2006) discuss prior elicitation while Tahir and Hussain (2008) compare uninformative priors for number of defects model. Aslam and Tahir (2010) focus Bayesian and Classical Analysis of Time-to-Failure Model. The Bayesian approach is preferred to the classical approach because the former can utilize the prior information in a formal way, satisfies the axioms of coherence and utilize decision theory. This study provides the relationship between Bayesian and classical estimators. Bernoulli distribution, Binomial distribution, Geometric distribution, Negative Binomial distribution, Exponential distribution, Poisson distribution, Power distribution, Maxwell distribution and Rayleigh distribution are used as sampling distributions in this paper. Beta distribution, Gamma distribution and Square root inverted gamma distribution are used as prior distributions.

***Corresponding Author:** Muhammad Saleem,
Department of Statistics
Govt. College University, Faisalabad , Pakistan
Email: selim.stat.qau@gmail.com

Materials and Methods

The Likelihood Function and MLE

The likelihood function summarizes the information contained in the sample. Maximum likelihood estimates make use of sample data only and have a number of desirable properties.

Posterior distribution and Bayes estimates

Bayesian Statistics utilizes prior information in a formal way and represents the knowledge about the parameter of sample data prior to observing the data. The priors used in this paper are all conjugate priors. The parameters of the prior distribution are called hyper-parameters. The posterior distribution summarizes two sources of information, the prior information through the prior distribution and the sample information via the likelihood function. Unlike classical Statistics, Bayesian school of thought considers the unknown parameter as a random variable and the inferences and decisions are based on posterior distribution of the unknown parameter.

The prior distribution of parameter θ of Bernoulli distribution is assumed as a Binomial distribution with hyper-parameters 'a' and 'b'. So the posterior distribution of θ is $Beta(\alpha,\ \beta)$ with parameters $\alpha = a + \sum_{i=1}^{n} x_i$ and $\beta = b + n - \sum_{i=1}^{n} x_i$. The posterior distribution of parameter θ of Binomial distribution using Beta distribution as a prior is the $Beta(\alpha,\ \beta)$ with $\alpha = a + \sum_{i=1}^{n} x_i$ and $\beta = b + nr - \sum_{i=1}^{n} x_i$. The posterior distribution of parameter θ of Geometric distribution using Beta distribution as a prior is $Beta(\alpha,\ \beta)$ with parameters $\alpha = a + n$ and $\beta = b + \sum_{i=1}^{n} x_i$. The posterior distribution of parameter θ of Negative Binomial distribution using Beta distribution as a prior is

$Beta(\alpha,\ \beta)$ with parameters $\alpha = a + nr$ and $\beta = b + \sum_{i=1}^{n} x_i$. If the sample observations are taken from the exponential distribution with parameter θ and the prior distribution of θ is a Gamma distribution with hyper-parameters 'a' and 'b', then the posterior distribution of θ is $Gamma(\alpha,\ \beta)$ with $\alpha = a + n$ and $\beta = b + \sum_{i=1}^{n} x_i$. Had we used $f(x) = \frac{1}{\theta}\ e^{-\frac{x}{\theta}}, x > 0$ as probability density function of the exponential distribution we would have assumed Inverted Gamma prior as conjugate prior. The posterior distribution of parameter θ of poisson distribution using Gamma distribution as a prior is $Gamma(\alpha,\ \beta)$ with $\alpha = a + \sum_{i=1}^{n} x_i$ and $\beta = b + n$. The posterior distribution of parameter of Power Function distribution θ using Gamma distribution as a prior distribution for the given data is $Gamma(\alpha,\ \beta)$ with $\alpha = a + n\sum_{i=1}^{n} x_i$ and $\beta = b + \sum_{i=1}^{n} \ln(1/x_i)$. Saleem *et al*. (2010) present Bayesian analysis of power function mixture distribution. It is assumed that the prior distribution of Maxwell parameter θ is a Square Root Inverted Gamma distribution with hyper-parameters 'a' and 'b', hence the posterior distribution of θ is the Square Root Inverted Gamma distribution $Ga^{-\frac{1}{2}}(\alpha,\ \beta)$ with parameters $\alpha = a + \frac{3n}{2}$ and $\beta = b + \frac{1}{2}\sum_{i=1}^{n} x_i^2$. The posterior distribution of parameter θ of Rayleigh distribution using Square Root Inverted Gamma distribution as a prior distribution is $Ga^{-\frac{1}{2}}(\alpha,\ \beta)$ with parameters $\alpha = a + n$ and $\beta = b + \frac{1}{2}\sum_{i=1}^{n} x_i^2$. Saleem and Aslam (2008a, b) worked on Rayleigh mixture.

Results and Discussion

Unlike Rossman, *et al*. (1998) and Elfessi and Reineke (2001), the relationship between Bayesian and classical estimators based on posterior mean and posterior mode are given in Table 1 and Table 2 for parameters of a number of distributions. Table 1 (Table 2) depicts that the posterior means (posterior modes) are obtained by the sum of numerators of the prior mode and of the MLE divided by the sum of denominators of the same except the Rayleigh and Maxwell cases. In Maxwell and Rayleigh cases it is observed that the squared posterior modes are obtained by the sum of squared numerators of the prior mode and of the MLE divided by the sum of squared denominators of the same.

Table 1 Relationship between MLEs, Prior Means and Posterior Means

Sampling Dist.	$f(x\|\theta)$	MLE	Prior Dist.	$p(\theta\|a,\ b)$	Prior Mean	Posterior Mean
Bernoulli	$\theta^x(1-\theta)^{1-x}, x = 0,1$	$\frac{\sum x}{n}$	Beta	$\frac{1}{B(a,\ b)}\theta^{a-1}(1-\theta)^{b-1}$	$\frac{a}{a+b}$	$\frac{a+\sum x}{a+b+n}$
Binomial	$\binom{r}{x}\theta^x(1-\theta)^{r-x}$ $x = 0,1,2,...,n$	$\frac{\sum x}{nr}$	Beta	$\frac{1}{B(a,\ b)}\theta^{a-1}(1-\theta)^{b-1}$	$\frac{a}{a+b}$	$\frac{a+\sum x}{a+b+nr}$
Geometric	$\theta(1-\theta)^x; x = 0,1,2,...$	$\frac{n}{n+\sum x}$	Beta	$\frac{1}{B(a,\ b)}\theta^{a-1}(1-\theta)^{b-1}$	$\frac{a}{a+b}$	$\frac{a+n}{a+b+n+\sum x}$
Negative Binomial	$\binom{r+x-1}{r-1}\theta^r(1-\theta)^x$ $x = 0,1,2,...$	$\frac{nr}{nr+\sum x}$	Beta	$\frac{1}{B(a,\ b)}\theta^{a-1}(1-\theta)^{b-1}$	$\frac{a}{a+b}$	$\frac{a+nr}{a+b+nr+\sum x}$
Exponential	$\theta\ e^{-\theta x}, x > 0$	$\frac{n}{\sum x}$	Gamma	$\frac{b^a}{\Gamma(a)}\theta^{a-1}e^{-b\theta}$	$\frac{a}{b}$	$\frac{a+n}{b+\sum x}$
Poisson	$\frac{\theta^x e^{-\theta}}{x!}, x = 0,1,2,...$	$\frac{\sum x}{n}$	Gamma	$\frac{b^a}{\Gamma(a)}\theta^{a-1}e^{-b\theta}$	$\frac{a}{b}$	$\frac{a+\sum x}{b+n}$
Power Function	$\theta\ x^{\theta-1}, x > 1$	$\frac{n}{\sum \ln(1/x)}$	Gamma	$\frac{b^a}{\Gamma(a)}\theta^{a-1}e^{-b\theta}$	$\frac{a}{b}$	$\frac{a+n}{b+\sum \ln(1/x)}$

Table 2 Relationship between MLEs, Prior Modes and Posterior Modes

Sampling Dist.	$f(x\|\theta)$	MLE	Prior Dist.	$p(\theta\|a,\ b)$	Prior Mode	Posterior Mode
Bernoulli	$\theta^x(1-\theta)^{1-x}, x=0,1$	$\frac{\sum x}{n}$	Beta	$\frac{1}{B(a,\ b)}\theta^{a-1}(1-\theta)^{b-1}$	$\frac{a-1}{a+b-2}$	$\frac{a-1+\sum x}{a+b-2+n}$
Binomial	$\binom{r}{x}\theta^x(1-\theta)^{r-x}$ $x=0,1,2,...,n$	$\frac{\sum x}{nr}$	Beta	$\frac{1}{B(a,\ b)}\theta^{a-1}(1-\theta)^{b-1}$	$\frac{a-1}{a+b-2}$	$\frac{a-1+\sum x}{a+b-2+nr}$
Geometric	$\theta(1-\theta)^x; x=0,1,2,...$	$\frac{n}{n+\sum x}$	Beta	$\frac{1}{B(a,\ b)}\theta^{a-1}(1-\theta)^{b-1}$	$\frac{a-1}{a+b-2}$	$\frac{a-1+n}{a+b-2+n+\sum x}$
Negative Binomial	$\binom{r+x-1}{r-1}\theta^r(1-\theta)^x$ $x=0,1,2,...$	$\frac{nr}{nr+\sum x}$	Beta	$\frac{1}{B(a,\ b)}\theta^{a-1}(1-\theta)^{b-1}$	$\frac{a-1}{a+b-2}$	$\frac{a-1+nr}{a+b-2+nr+\sum x}$
Exponent-ial	$\theta\ e^{-\theta x}, x>0$	$\frac{n}{\sum x}$	Gamma	$\frac{b^a}{\Gamma(a)}\theta^{a-1}e^{-b\theta}$	$\frac{a-1}{b}$	$\frac{a-1+n}{b+\sum x}$
Poisson	$\frac{\theta^x e^{-\theta}}{x!}, x=0,1,2,...$	$\frac{\sum x}{n}$	Gamma	$\frac{b^a}{\Gamma(a)}\theta^{a-1}e^{-b\theta}$	$\frac{a-1}{b}$	$\frac{a-1+\sum x}{b+n}$
Power Function	$\theta\ x^{\theta-1}, x>1$	$\frac{n}{\sum\ln(1/x)}$	Gamma	$\frac{b^a}{\Gamma(a)}\theta^{a-1}e^{-b\theta}$	$\frac{a-1}{b}$	$\frac{a-1+n}{b+\sum\ln(1/x)}$
Maxwell	$\sqrt{\frac{2}{\Pi}}\frac{x^2e^{-\frac{x^2}{2\theta^2}}}{\theta^3}, x>0$	$\sqrt{\frac{\sum x^2}{3n}}$	Square Root Inverted Gamma	$\frac{2b^a}{\Gamma(a)}\theta^{-(2a+1)}e^{-\frac{b}{\theta^2}}$	$\sqrt{\frac{2b}{2a+1}}$	$\sqrt{\frac{2b+\sum x^2}{2a+1+3n}}$
Rayleigh	$\frac{x}{\theta^2}e^{-\frac{x^2}{2\theta^2}}, x>0$	$\sqrt{\frac{2\sum x^2}{2n}}$	Square Root Inverted Gamma	$\frac{2b^a}{\Gamma(a)}\theta^{-(2a+1)}e^{-\frac{b}{\theta^2}}$	$\sqrt{\frac{2b}{2a+1}}$	$\sqrt{\frac{2b+2\sum x^2}{2a+1+2n}}$

The Bayes estimates, posterior mean and posterior mode, reduce to the Classical estimates, maximum likelihood estimates, if values of the hyper-parameters are set to zero. Hence the maximum likelihood estimates can be considered as special case of the Bayes estimates.

References

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

Rossman, A. J., Short, T. H. and Parks, M. T. . Bayes Estimators for the Continuous Uniform Distribution. Journal of Statistics Education, 1998, 6 (3).

Aslam, M. and Tahir, M. Bayesian and Classical Analysis of Time-to-Failure Model with Comparison of Non-Informative Priors. Pakistan Journal of Statistics. 2010, 26: 407-415.

Tahir, M. and Hussain, Z. Comparison of Non-Informative Priors for Number of Defects (Poisson) Model. Inter Stat, #2, April, 2008.

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

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. 2006, 169: 37-48.

Saleem, M. and Aslam, M. Bayesian analysis of the two component mixture of the Rayleigh distribution assuming the uniform and the Jeffreys priors. Journal of Applied Statistical Science, 2008(a)16(4): 105-113.

Saleem, M. and Aslam, M. On prior selection for the mixture of Rayleigh distribution using predictive Intervals. Pakistan Journal of Statistics, 2008 (b) 24(1): 21-35.

Saleem M. and Aslam M. and Economou, P. On the Bayesian analysis of the mixture of Power function distribution using the complete and the censored sample. Journal of Applied Statistics, 2010, 37 (1): 25-40.