The Korean Communications
in Statistics Vol. 4, No. 2, 1997
pp. 465 - 471

Bayesian and Empirical Bayesian Prediction Analysis for a Future Observation

Jeong Hwan Ko[1)]

Abstract

This paper deals with the problems of obtaining some Bayesian and empirical Bayesian predictive densities and prediction intervals of a future observation $X_{(r+\nu)}$ in the Rayleigh distribution. Using an inverse gamma prior distribution, some predictive densities and prediction intervals are proposed and studied. Also the behaviors of the proposed results are examined via numerical examples.

1. Introduction

Statistical prediction analysis could provide warranty limits for the future performance of systems or could be used in situations where a producer compare the performance of both his product and that of a competitor and wishs to determine the difference in future mean performance of the products. thus this analysis plays a very important role in reliability analysis for some lifetime models, quality control and other applicatin areas.

Serveral distributions have been introduced and discussed for this problem with a Bayesian point of view. Chhikara and Guttman(1982), Nigam and Hamdy(1987), Sinha(1989) and Upadhyay and Pandey(1989) suggested the Bayesian inference about prediction for inverse Gaussian, lognormal, Pareto, exponential distributions, respectively. We also deal here with the prediction analysis based on the parametric empirical Bayesian method. this method was studied by Efron and Morris(1973) and Morris(1983).

In this paper, we deal with the Bayesian and empirical Bayesian predictive density and prediction intervals for the future observtion based upon a censored sample observed from the Rayleigh models.

The probability density function(pdf), denoted by $R(\sigma^2)$, is given by

$$f(x\ ;\ \sigma) = \frac{x}{\sigma^2}\exp\left(-\frac{x^2}{2\sigma^2}\right), \qquad 0 < x < \infty \quad . \tag{1.1}$$

1) Assistant Professor, Department of Statistics, Andong National University, Andong, 760-749, Korea

The properties and application of Rayleigh distribution were discussed by Siddiqui(1962), Dyer and Whisenand(1973a,b). Sinha and Howlader(1983) obtained the Bayes estimator and credible intervals for the reliability function.

In Section 2.2, we propose the Bayesian and empirical Bayesian predictive density function and prediction intervals of a future observation $X_{(r+\nu)}$ based upon the Type II censored sample from the distribution with pdf's in (1.1).

In Section 2.3, we provide numerical examples for the proposed equal-tail and most plausible prediction interval of a future observation with respect to a inverse gamma prior distribution.

2. Prediction Analysis of a Future Observation

Let $\underset{\sim}{X}=(X_1,\cdots,X_n)$ with $X_{(1)}\le X_{(2)}\le\cdots\le X_{(r)}$ be the observed lifetimes of the first r components to fail in a random sample of n components whose lifetimes have the Rayleigh distribution with probability density function in (1.1) and the survival function is

$$\overline{F}(x:\sigma)=\exp\left(-\frac{x^2}{2\sigma^2}\right),\quad \sigma>0,\quad 0<x<\infty. \tag{2.1}$$

Then the likelihood function is given by

$$L(\sigma\mid \underset{\sim}{x})=\frac{n!}{(n-r)!}\frac{\prod_{i=1}^{r}x_i}{\sigma^{2r}}\exp\left[-\frac{\sum_{i=1}^{r}x_i^2+(n-r)x_{(r)}^2}{2\sigma^2}\right], \tag{2.2}$$

$$\sigma>0,\quad 0<x_i<\infty,\quad i=1,2,\cdots,n.$$

Suppose a sample of n components is put on life test and the distribution of their lifetimes is a Rayleigh and that the lifetimes of the first r of these to make inferences about the lifetimes of the remaining $n-r$ components in the sample. Let $X_{(r+\nu)}$ denote the lifetime of the $(r+\nu)^{th}$ component to fail, where $\nu=1,2,\cdots,n-r$. Then the conditional predictive density of $X_{(r+\nu)}$ given $\underset{\sim}{X}=\underset{\sim}{x}$ is

$$f(x_{(r+\nu)}\mid\sigma,\underset{\sim}{x})=\frac{\nu\binom{n-r}{\nu}(\overline{F}(x_{(r)})-\overline{F}(x_{(r+\nu)}))^{\nu-1}\overline{F}(x_{(r+\nu)})^{n-r-\nu}f(x_{(r+\nu)})}{\overline{F}(x_{(r)})^{n-r}} \tag{2.3}$$

$$= \frac{\nu x_{(r+\nu)}\binom{n-r}{\nu}}{\sigma^2} \sum_{j=0}^{\nu-1}\binom{n-r}{\nu}(-1)^j \exp\left(-\frac{(n-r-\nu+j+1)(x_{(r+\nu)}^2 - x_{(r)}^2)}{2\sigma^2}\right), \quad x_{(r+\nu)} > x_{(r)} .$$

We consider an inverse gamma prior distribution for σ with probability density function

$$\Pi(\sigma : \alpha, \beta) = \frac{\exp(-1/\beta\sigma^2)}{\Gamma(\alpha)\beta^{\alpha}\sigma^{2(\alpha+1)}}, \qquad \alpha,\ \beta > 0,\ \sigma > 0. \tag{2.4}$$

Combining the likelihood function in (2.2) and an inverse gamma prior density function, the posterior density of σ given $\underset{\sim}{X} = \underset{\sim}{x}$ is

$$\Pi(\sigma \mid \underset{\sim}{x}) = \frac{((\beta Z^2+2)/\beta)^{r+\alpha+1/2}}{\Gamma(r+\alpha+1/2)2^{r+\alpha-1/2}} \frac{1}{\sigma^{2(r+\alpha+1)}} \exp\left(-\frac{\beta Z^2+2}{2\beta\sigma^2}\right), \quad \alpha, \beta > 0,\ \sigma > 0. \tag{2.5}$$

where $Z^2 = \sum_{i=1}^{r} x_i^2 + (n-r)x_{(r)}^2$

If (2.3) and (2.4) are considered, then the predictive pdf $X_{(r+\nu)}$ is given as follows:

Theorem 2.1. Under the conditional predictive density and an inverse gamma prior, the predictive density $X_{(r+\nu)}$ given $\underset{\sim}{X} = \underset{\sim}{x}$ is

$$\Pi(X_{(r+\nu)} \mid \underset{\sim}{x}) = 2\nu x_{(r+\nu)}\, \beta\binom{n-r}{\nu}\left(r+\alpha+\frac{1}{2}\right)^{r+\alpha+1/2} \tag{2.6}$$

$$\times \sum_{j=0}^{\nu-1}\binom{\nu-1}{j}(-1)^j(\beta Z^2+2+(n-r-\nu+j+1)\beta(x_{(r+\nu)}^2 - x_{(r)}^2))^{-(r+\alpha+3/2)}$$

If an inverse gamma prior is used, then the equal-tail $100(1-r)\%$ prediction interval (C_{GL}, C_{GU}) for $X_{(r+\nu)}$ can be constructed by solving the following equations:

$$\frac{r}{2} = \nu\binom{n-r}{\nu}\sum_{j=0}^{\nu-1}\binom{\nu-1}{j}(-1)^j(n-r-\nu+j+1)^{-1}$$

$$\times\left(1-\left(\frac{\beta Z^2+2}{\beta Z^2+2+(n-r-\nu+j+1)\beta(C_{GL}^2 - x_{(r)}^2)}\right)^{r+\alpha+1/2}\right)$$

and

$$\frac{r}{2} = \nu\binom{n-r}{\nu}\sum_{j=0}^{\nu-1}(-1)^j(n-r-\nu+j+1)^{-1}$$

$$\times\left(\frac{\beta Z^2+2}{\beta Z^2+2+(n-r-\nu+j+1)\beta(C_{GU}^2 - x_{(r)}^2)}\right)^{r+\alpha+1/2}$$

The $100(1-\gamma)\%$ most plausible prediction interval (H_{GL}, H_{GU}) for $X_{(r+\nu)}$ are obtained by solving simultaneously the followings;

$$\nu \binom{n-r}{\nu} \sum_{j=0}^{\nu-1} \binom{\nu-1}{j} (-1)^j (n-r-\nu+j+1)^{-1} \times \left(1+\frac{(n-r-\nu+j+1)\beta(H_{GL}^2-X_{(r)}^2)}{(\beta Z^2+2)}\right)^{-(r+\alpha+1/2)}$$

$$-\nu \binom{n-r}{\nu} \sum_{j=0}^{\nu-1} \binom{\nu-1}{j} (-1)^j (n-r-\nu+j+1)^{-1} \times \left(1+\frac{(n-r-\nu+j+1)\beta(H_{GU}^2-X_{(r)}^2)}{(\beta Z^2+2)}\right)^{-(r+\alpha+1/2)}$$

$$= 1 - \gamma$$

and

$$H_{GL} \sum_{j=0}^{\nu-1} \binom{\nu-1}{j} (-1)^j (n-r-\nu+j+1)^{-1} \left(1+\frac{(n-r-\nu+j+1)\beta(H_{GL}^2-X_{(r)}^2)}{(\beta Z^2+2)}\right)^{-(r+\alpha+3/2)}$$

$$= H_{GU} \sum_{j=0}^{\nu-1} \binom{\nu-1}{j} (-1)^j (n-r-\nu+j+1)^{-1} \left(1+\frac{(n-r-\nu+j+1)\beta(H_{GU}^2-X_{(r)}^2)}{(\beta Z^2+2)}\right)^{-(r+\alpha+3/2)}$$

Now, we consider empirical Bayesian approach of the predictive density of a future observation $X_{(r+\nu)}$ for the Rayleigh distribution under the type II censoring.

If a prior for σ is an inverse gamma distribution with parameters α and β in (2.4), then the posterior density function of σ given $\underset{\sim}{X}=\underset{\sim}{x}$ is given in (2.5). Therefore, under the distribution of a future observation of $X_{(r+\nu)}$ in (2.3), the Bayesian predictive density function is given in (2.6). We can estimate the unknown parameters α and β for an inverse gamma prior distribution by using the parametric empirical Bayesian approach.

The likelihood function under the type II censoring is given by

$$L(\alpha,\beta \mid \underset{\sim}{x}) = \frac{\prod_{i=1}^{r} x_i (\Gamma(\alpha+\frac{5}{2}))^r 2^{r(\alpha+3/2)} (\Gamma(\alpha+\frac{3}{2}))^{n-r} 2^{(n-r)(\alpha+1/2)}}{(\Gamma(\alpha))^n \beta^{n\alpha} \prod_{i=1}^{r} \left(\frac{\beta x_i^2+2}{\beta}\right)^{\alpha+5/2} \left(\frac{\beta x_{(r)}^2+2}{\beta}\right)^{(n-r)(\alpha+3/2)}} \qquad (2.7)$$

In order to obtain the maximum likelihood estimators of α , β , we must compute the first partial derivative of log-likelihood function.

Therefore, the MLE's $\hat{\alpha}$ and $\hat{\beta}$ of α and β can be obtained by simultaneously solving.

$$\hat{\alpha}=\frac{\frac{5}{2}\left(\sum_{i=1}^{r}\left(\frac{x_i^2}{\hat{\beta}x_i^2+2}\right)-\frac{r}{\hat{\beta}}\right)+\frac{3}{2}(n-r)\left(\left(\frac{x_{(r)}^2}{\hat{\beta}x_{(r)}^2}+2\right)-\frac{1}{\hat{\beta}}\right)}{\left[-\frac{n}{\hat{\beta}}-\left(\sum_{i=1}^{r}\left(\frac{x_i^2}{\hat{\beta}x_i^2+2}\right)-\frac{r}{\hat{\beta}}\right)-(n-r)\left(\left(\frac{x_{(r)}^2}{\hat{\beta}x_{(r)}^2}+2\right)-\frac{1}{\hat{\beta}}\right)\right]}$$

and (2.8)

$$r\frac{\Gamma'(\hat{\alpha}+\frac{5}{2})}{\Gamma(\hat{\alpha}+\frac{5}{2})}+(n-r)\frac{\Gamma'(\hat{\alpha}+\frac{3}{2})}{\Gamma(\hat{\alpha}+\frac{3}{2})}+n\log 2-n\frac{\Gamma'(\hat{\alpha})}{\Gamma(\hat{\alpha})}$$

$$=\sum_{i=1}^{r}\log(\hat{\beta}x_i^2+2)+(n-r)\log(\hat{\beta}x_{(r)}^2+2)$$

Then the following theorem is obtained.

Theorem 2.2. For an inverse gamma prior with parameters α and β, the empirical Bayesian predictive density function of $X_{(r+\nu)}$ is given by

$$\Pi(X_{(r+\nu)}\mid \underset{\sim}{x})=2\nu x_{(r+\nu)}\hat{\beta}\binom{n-r}{\nu}(r+\hat{\beta}+\frac{1}{2})(\hat{\beta}Z^2+2)^{r+\hat{\alpha}+1/2}$$

$$\times\sum_{j=0}^{\nu-1}\binom{\nu-1}{j}(-1)^j(\hat{\beta}Z^2+2+(n-r-\nu+j+1)\hat{\beta}(x_{(r+\nu)}^2-x_{(r)}^2))^{-(r+\hat{\alpha}+1/2)} \quad (2.10)$$

If an inverse gamma prior is used, then the equal-tail $100(1-r)\%$ prediction interval (C_{EL}, C_{EU}) for $X_{(r+\nu)}$ can be constructed by solving the following equations:

$$\frac{r}{2}=\nu\binom{n-r}{\nu}\sum_{j=0}^{\nu-1}\binom{\nu-1}{j}(-1)^j(n-r-\nu+j+1)^{-1}$$

$$\times\left(1-\left(\frac{\hat{\beta}Z^2+2}{\hat{\beta}Z^2+2+(n-r-\nu+j+1)\hat{\beta}(C_{EL}^2-x_{(r)}^2)}\right)^{r+\hat{\alpha}+1/2}\right)$$

and

$$\frac{r}{2}=\nu\binom{n-r}{\nu}\sum_{j=0}^{\nu-1}(-1)^j(n-r-\nu+j+1)^{-1}$$

$$\times\left(\frac{\hat{\beta}Z^2+2}{\hat{\beta}Z^2+2+(n-r-\nu+j+1)\hat{\beta}(C_{EU}^2-x_{(r)}^2)}\right)^{r+\hat{\alpha}+1/2}$$

The $100(1-\gamma)\%$ most plausible prediction interval (H_{EL}, H_{EU}) for $X_{(r+\nu)}$ are obtained by solving simultaneously the followings :

$$\nu\binom{n-r}{\nu}\sum_{j=0}^{\nu-1}\binom{\nu-1}{j}(-1)^j(n-r-\nu+j+1)^{-1}$$
$$\times\left(1+\frac{(n-r-\nu+j+1)\hat{\beta}(H_{EL}^2-x_{(r)}^2)}{(\hat{\beta}Z^2+2)}\right)^{-(r+\hat{\alpha}+1/2)}$$

$$-\ \nu\binom{n-r}{\nu}\sum_{j=0}^{\nu-1}\binom{\nu-1}{j}(-1)^j(n-r-\nu+j+1)^{-1}$$
$$\times\left(1+\frac{(n-r-\nu+j+1)\hat{\beta}(H_{EU}^2-x_{(r)}^2)}{(\hat{\beta}Z^2+2)}\right)^{-(r+\hat{\alpha}+1/2)}$$

$$= 1-\gamma$$

and

$$H_{EL}\sum_{j=0}^{\nu-1}\binom{\nu-1}{j}(-1)^j(n-r-\nu+j+1)^{-1}\left(1+\frac{(n-r-\nu+j+1)\hat{\beta}(H_{EL}^2-x_{(r)}^2)}{(\hat{\beta}Z^2+2)}\right)^{-(r+\hat{\alpha}+3/2)}$$

$$= H_{EU}\sum_{j=0}^{\nu-1}\binom{\nu-1}{j}(-1)^j(n-r-\nu+j+1)^{-1}\left(1+\frac{(n-r-\nu+j+1)\hat{\beta}(H_{EU}^2-x_{(r)}^2)}{(\hat{\beta}Z^2+2)}\right)^{-(r+\hat{\alpha}+3/2)}$$

3. Numerical Examples

In this section, to predict a future observation $X_{(r+\nu)}$, the data were generated artifically from the Rayleigh model with parameter $\sigma^2=4$ under the Type II censoring. It is assumed that only the first 14 (thirty percent censoring) ordered failue times are available, and they are given as follows:

0.5244, 0.5523, 0.7027, 0.8034, 1.2393, 1.7997, 1.8306,
1.9144, 2.5410, 2.6403, 2.7548, 3.0523, 3.4957, 3.8314

Under the data, to see the difference between Bayesian and empirical Bayesian approach in the equal-tail and the most plausible prediction interval, the 95% equal-tail and 95% most plausible prediction intervals with respect to an inverse gamma prior $IG(1,2)$ are derived and are showen in Table1.

Table1 : Prediction Interval of $X_{(r+\nu)}$

prior	$X_{(r+\nu)}$	15	16	17	18
IG(1, 2)	M. P.	(3.841, 4.987)	(3.907, 5.512)	(4.026, 6.144)	(4.206, 7.008)
	E. T	(3.899, 5.190)	(3.990, 5.741)	(4.139, 6.404)	(4.354, 7.320)
E. B.	M. P.	(3.841, 4.997)	(3.907, 5.527)	(4.028, 6.163)	(4.209, 7.034)
	E. T.	(3.899, 5.202)	(3.991, 5.758)	(4.142, 6.425)	(4.358, 7.348)

$IG(1,2)$: Inverse gamma prior E.B. : Empirical Bayesian approach
E.T. : Equal-tail pediction bound M.T. : Most plausible prediction bound

References

[1] Chhikara, R. S. and Guttman, I. (1982), Prediction Limits for the Inverse Gaussian distribution, *Technometrics,* 24, 319-324.

[2] Dyer, D. D. and Whisenand, C. W. (1973a), Best Linear Unbiased Estimator of the Parameter of the Rayleigh Distribution - Part I: Small Sample Theory for Censored Order Statistics, *IEEE Transactions on Reliability*, R-22, 27-34.

[3] Dyer, D. D. and Whisenand, C. W. (1973b), Best Linear Unbiased Estimator of the Parameter of the Rayleigh Distribution - Part II: Optimum Theory for Selected Oder Statistics, *IEEE Transactions on Reliability*, R-22, 229-231.

[4] Efron, B. and Morris, C. (1973), Stein's Estimation Rule and Its Competitors-An Empirical Bayes Approach, *Journal of the American Statistical Association,* 68, 117-130.

[5] Morris, C. (1983), Parametric Empirical Bayes Inference:Theoryand Applications, *Journal of the American Statistical Association, 78, 47-59.*

[6] Nigm, A. M. and Hamdy, H. I. (1987), Bayesian Prediction Bounds the Pareto Lifetime Model, *Communicationsin Statistics Part A- Theory and Methods,* 16, 1761-1772.

[7] Siddiqui, M. M. (1962), Some Problems connected with Rayleigh Distribution, *Journal of Research of the National Bureau of Standards-D,* 66D, 170-174.

[8] Sinha, S. K. (1989), Bayesian Inference about the Prediction Credible Intervals and Reliability Function for Lognormal Distribution, *Journal of the Indian Statistical Association*, 27, 73-78.

[9] Sinha, S. K. and Howlader, H. A. (1983), Credible and HPD Intervals of the Parameter and Reliability of Rayleigh Distribution, *IEEE Transactionson Reliability,* R-32, 217-220.

[10] Upadhyay, S. K. and Pandey, M. (1989), Prediction Limits for an Exponential Distribution: A Bayes Predictive Distribution Approach, *IEEE Transactions on Reliability,* R-38, 599-602.