Journal of Statistical
Theory & Methods
1995, Vol. 6, pp 53 ~ 61

A Study of Bayesian and Empirical Bayesian Prediction Analysis for the Rayleigh Model under the Random Censoring[1]

Jeong Hwan Ko[2]

Abstract This paper deals with problems of predicting, based on the random censored sampling, a future observation and the p-th order statistic of n' future observations for the Rayleigh model. We consider the prediction intervals for the Rayleigh model with respect to an inverse gamma prior distribution. In additions, numerical examples are given in order to illustrate the proposed predictive procedure.

1. Introduction

An important problem in life testing which has received much attention in recent years is that of predicting the lifetimes of unused components and the lifetime of a system.

Several distributions have been introduced and discussed for this problem with a Bayesian point of view. Dunsmore(1974) discussed this problem when the underlying distribution is one- or two-parameter exponential distribution. Lingappaiah(1986) proposed the Bayesian prediction in exponential life-testing when sample size is a random variable. Chhikara and Guttman(1982), Nigm and Hamdy(1987), Sinha(1989) suggested the Bayesian inference about prediction for inverse Gaussian, lognormal and Pareto distribution, respectively. Clarotti and Spizzichino(1989) proposed the Bayesian predictive approach in reliability theory. We also deal here with the prediction analysis based on the parametric empirical Bayesian method. This method was studied by Efron and Morris(1973a, 1973b) and Morris(1983). Also Miller(1989) worked on a

---

[1] This paper was supported by NON DIRECTED RESEARCH FUND, Korea Research Foundation in 1994

[2] Department of Statistics, Andong National University, Andong 760-749, Korea.

parametric empirical Bayesian analysis for tolerance bound.

Let $T_1, T_2, \cdots, T_n$ be independent and identically distributed (i.i.d.) random samples lifetimes of n items with the probability density function $f(\cdot)$ and reliability function $\overline{F}(\cdot)$. Let $C_1, C_2, \cdots, C_n$ be i.i.d. censoring times of the items with probability density function $g(\cdot)$ and continuous distribution function $G(\cdot)$. It is assumed that T_i and C_i are mutually independent. We only are able to observe the time X_i and the censoring indicator δ_i on the ith trial, where

$$X_i = \min(T_i, C_i)$$

and

$$\delta_i = I[T_i \le C_i] = \begin{cases} 1, & \text{if } T_i \le C_i \\ 0, & \text{if } T_i > C_i \,. \end{cases}$$

In this paper, we consider problems of the Bayesian and the empirical Bayesian prediction analyses for future observations under random censoring scheme.

In Section 2, we derive the predictive density and the prediction intervals of a future observation or the p-th order statistic of n' future observations for the Rayleigh model under the Bayesian approach.

In Section 3, we deal with the empirical Bayesian prediction approach for the Rayleigh model with respect to an inverse gamma prior distribution.

In Section 4, Numerical examples are given in order to illustrate the proposed predictive procedure.

2. BAYESIAN PREDICTION APPROACH

Let X be a random variable following the Rayleigh distribution, $R(\sigma^2)$,whose probability density function(pdf) is given by

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

Now to predict for a future observation y, it is necessary to derive the Bayesian predictive density function for y. Under the random censoring, the likelihood function for a sample size n is

$$L(\sigma|\underline{x}) = \frac{\prod_{i\in D_1} x_i}{\sigma^{2|D_1|}}\exp\left(-\frac{\sum_{i\in D_1} x_i^2 + \sum_{i\in D_2} x_i^2}{2\sigma^2}\right), \quad \sigma > 0, \tag{2.2}$$

where $|D_1|$ is the number of elements in the set of individuals whose lifetimes are observed D_1 and D_2 is the set of individuals whose only censoring times are available.

As a prior distribution for σ, we consider an inverse gamma prior distribution, $IG(\alpha,\beta)$, with the probability density function

$$\Pi(\sigma|\alpha,\beta)=\frac{\exp(-1/\beta\sigma^2)}{\Gamma(\alpha)\beta^{\alpha}\sigma^{2(\alpha+1)}},\quad \alpha,\beta>0,\ \sigma>0. \tag{2.3}$$

Then one can obtain easily the posterior density of σ given $\underline{X}=\underline{x}$ which is given by

$$\Pi(\sigma|\underline{x})=\frac{\{(\beta T^2+2)/\beta\}^{|D_1|+\alpha+1/2}}{\Gamma(|D_1|+\alpha+1/2)2^{|D_1|+\alpha-1/2}}\frac{1}{\sigma^{2(|D_1|+\alpha+1)}}\exp\left(-\frac{\beta T^2+2}{2\beta\sigma^2}\right), \tag{2.4}$$
$$\alpha,\beta>0\quad \sigma>0,$$

where $T^2=\sum_{i\in D_1}x_i^2+\sum_{i\in D_2}x_i^2$.

The distribution of a future observation y given σ is

$$f(y|\sigma)=\frac{y}{\sigma^2}\exp\left(-\frac{y^2}{2\sigma^2}\right),\quad \sigma>0,\ 0<y<\infty. \tag{2.5}$$

Thus the predictive density function for y can be obtained and is given in the following theorem.

Theorem 2.1. With an inverse gamma prior with parameters α and β for σ, the predictive density function of y given $\underline{X}=\underline{x}$ is

$$\Pi(y|\underline{x})=\frac{2(|D_1|+\alpha+1/2)y\beta(\beta T^2+2)^{|D_1|+\alpha-1/2}}{(\beta T^2+2+\beta y^2)^{|D_1|+\alpha+1/2}},\quad \alpha,\beta>0\quad y>0. \tag{2.6}$$

Now we want to construct the prediction intervals for a future observation. If an inverse gamma prior with parameters α and β for σ is used, the 100(1-r)% equal-tail prediction interval (C_{GL},C_{GU}) for y is

$$\left(\sqrt{\frac{\beta T^2+2}{\beta}\{(1-r/2)^{-1/|D_1|+\alpha+1/2}-1\}},\ \sqrt{\frac{\beta T^2+2}{\beta}\{(r/2)^{-1/|D_1|+\alpha+1/2}-1\}}\right).$$

The 100(1-r)% most plausible Bayesian prediction interval (M_{GL},M_{GU}) for y can be obtained by solving simultaneously the followings :

$$\left(\frac{\beta T^2+2}{M_{GL}^2+\beta T^2+2}\right)^{|D_1|+\alpha+1/2}-\left(\frac{\beta T^2+2}{M_{GU}^2+\beta T^2+2}\right)^{|D_1|+\alpha+1/2}=1-r$$

and

$$\left(\frac{\beta T^2+2+\beta M_{GU}^2}{\beta T^2+2+\beta M_{GL}^2}\right)^{|D_1|+\alpha+3/2}=\frac{M_{GU}^2}{M_{GL}^2}.$$

Also the distribution of the p-th order statistic, $Y_{(p)}$ of n' future observations is

$$f(y_{(p)}|\sigma)=\frac{1}{B(p,n'-p+1)}\frac{y_{(p)}}{\sigma^2}\exp\left(-\frac{(n'-p+1)y_{(p)}^2}{2\sigma^2}\right)$$
$$\times\left\{1-\exp\left(-\frac{y_{(p)}^2}{2\sigma^2}\right)\right\}^{p-1},\quad y_{(p)}\geq 0,\ 1\leq p\leq n'. \tag{2.7}$$

Hence the predictive density function of $Y_{(p)}$ can be obtained and is given as follows :

Theorem 2.2. With an inverse gamma prior for σ, $IG(\alpha,\beta)$, the predictive density function of the p-th order statistic, $Y_{(p)}$ of n' future observations is given by

$$\Pi(y_{(p)}|\underline{x})=\frac{2\beta(|D_1|+\alpha+1/2)y_{(p)}(\beta T^2+2)^{|D_1|+\alpha+1/2}}{B(p,n'-p+1)}\sum_{i=0}^{p-1}\binom{p-1}{i}(-1)^i$$
$$\times\{\beta T^2+2+(n'-p+1+i)\beta y_{(p)}^2\}^{-(|D_2|+\alpha+3/2)},\quad y_{(p)}>0. \tag{2.8}$$

For an inverse gamma prior with parameters α and β for σ, the 100(1-r)% equal-tail prediction interval of the p-th order statistic of n' future observations can be obtained by solving the following equations :

$$\frac{r}{2}=\frac{1}{B(p,n'-p+1)}\sum_{i=0}^{p-1}\binom{p-1}{i}(-1)^i(n'-p+1+i)^{-1}$$
$$\left\{1-\left(1+\frac{(n'-p+1+i)\beta C_{GL}^2}{\beta T^2+2}\right)^{(|D_1|+\alpha+1/2)}\right\}$$

and

$$\frac{r}{2}=\frac{1}{B(p,n'-p+1)}\sum_{i=0}^{p-1}\binom{p-1}{i}(-1)^i(n'-p+1+i)^{-1}\left(1+\frac{(n'-p+1+i)\beta C_{GU}^2}{\beta T^2+2}\right)^{(|D_1|+\alpha+1/2)}$$

Also the 100(1-r)% most plausible Bayesian prediction interval $(M_{GL}, M_{GU'})$ for $Y_{(p)}$ can be obtained by solving the following equations :

$$\frac{1}{B(p,n'-p+1)}\sum_{i=0}^{p-1}\binom{p-1}{i}(-1)^i(n'-p+1+i)^{-1}\left(1+\frac{(n'-p+1+i)\beta M_{GL}^2}{\beta T^2+2}\right)^{-(|D_1|+\alpha+1/2)}$$
$$-\frac{1}{B(p,n'-p+1)}\sum_{i=0}^{p-1}\binom{p-1}{i}(-1)^i(n'-p+1+i)^{-1}\left(1+\frac{(n'-p+1+i)\beta M_{GU}^2}{\beta T^2+2}\right)^{-(|D_1|+\alpha+1/2)}$$
$$=1-r$$

and

$$M_{GL}\sum_{i=0}^{p-1}\binom{p-1}{i}(-1)^i\left(1+\frac{(n'-p+1+i)\beta M_{GL}^2}{\beta T^2+2}\right)^{-(|D_1|+\alpha+3/2)}$$
$$=M_{GL}\sum_{i=0}^{p-1}\binom{p-1}{i}(-1)^i\left(1+\frac{(n'-p+1+i)\beta M_{GU}^2}{\beta T^2+2}\right)^{-(|D_1|+\alpha+3/2)}$$

3. Empirical Bayesian Approach

For the Rayleigh distribution under the random censoring, $\Pi(\sigma|\underline{x})$ is derived in (2.4) when the prior distribution is an inverse gamma with parameters α and β given in (2.3). Also the Bayesian predictive density function for y is obtained in (2.5). Under the parametric empirical Bayesian approach, it is necessary to estimate the unknown parameters from the past data $X_1, X_2, \cdots, X_n$. The estimation methods are various and here we will use the maximum likelihood estimation.

Let $X_1, X_2, \cdots, X_n$ be n past random samples from the Rayleigh model whose probability density function is given by (2.1). Now the likelihood function under the random censoring is

$$L(\alpha,\beta|\underline{x})=\frac{\prod_{i\in D_1}x_i\{\Gamma(\alpha+5/2)\}^{|D_1|}\{\Gamma(\alpha+3/2)\}^{|D_2|}2^{|D_1|(\alpha+3/2)}2^{|D_2|(\alpha+1/2)}}{\{\Gamma(\alpha)\}^n\beta^{n\alpha}\prod_{i\in D_1}\left(\frac{\beta x_i^2+2}{\beta}\right)^{\alpha+5/2}\prod_{i\in D_2}\left(\frac{\beta x_i^2+2}{\beta}\right)^{\alpha+3/2}}, \tag{3.1}$$

where $|D_i|$ is the number of observations in the set D_i, $i=1,2$.

In order to obtain the maximum likelihood estimators of α and β, one must compute the first partial derivative of the log-likelihood function. Therefore, the MLE's $\hat{\alpha}$ and $\hat{\beta}$ of α and β can be obtained by simultaneously solving

$$\hat{\alpha}=\frac{\frac{5}{2}\{\sum_{i\in D_1}(\frac{x_i^2}{\hat{\beta}x_i^2+2})-\frac{|D_1|}{\hat{\beta}}\}+\frac{3}{2}\{\sum_{i\in D_2}(\frac{x_i^2}{\hat{\beta}x_i^2+2})-\frac{|D_2|}{\hat{\beta}}\}}{-\frac{n}{\hat{\beta}}-\{\sum_{i\in D_1}(\frac{x_i^2}{\hat{\beta}x_i^2+2})-\frac{|D_1|}{\hat{\beta}}\}-\{\sum_{i\in D_2}(\frac{x_i^2}{\hat{\beta}x_i^2+2})-\frac{|D_1|}{\hat{\beta}}} \tag{3.2}$$

and

$$|D_1|\frac{\Gamma(\hat{\alpha}+5/2)}{\Gamma(\hat{\alpha}+5/2)}+|D_2|\frac{\Gamma(\hat{\alpha}+3/2)}{\Gamma(\hat{\alpha}+3/2)}+n\log 2-n\frac{\Gamma(\hat{\alpha})}{\Gamma(\hat{\alpha})}$$
$$=\sum_{i\in D_1}\log(\hat{\beta}x_i^2+2)+\sum_{i\in D_2}\log(\hat{\beta}x_i^2+2).$$

Now the empirical Bayesian predictive density function of y can be obtained via replacing α and β by $\hat{\alpha}$ and $\hat{\beta}$.

Theorem 3.1. If an inverse gamma prior with parameters α and β for σ is used, then the empirical Bayesian predictive density function of a future observation y is given by

$$\Pi(y|\underline{x})=\frac{2(|D_1|+\hat{\alpha}+1/2)\hat{\beta}y(\hat{\beta}T^2+2)^{|D_1|+\hat{\alpha}-1/2}}{(\hat{\beta}T^2+2+\hat{\beta}y^2)^{|D_1|+\hat{\alpha}+1/2}},\quad y>0, \tag{3.3}$$

where $\hat{\alpha}$ and $\hat{\beta}$ are the solutions of the equations (3.2).

Theorem 3.2. For an inverse gamma prior with parameters α and β for σ in (2.3), we obtain the predictive density function for the p-th order statistic, $Y_{(p)}$, of a future observations given by

$$\Pi(y_{(p)}|\underline{x}) = \frac{2\hat{\beta}(|D_1|+\alpha+1/2)y_{(p)}(\hat{\beta}T^2+2)^{|D_1|+\hat{\alpha}+1/2}}{B(p,n'-p+1)}\sum_{i=0}^{p-1}\binom{p-2}{i}(-1)^i$$
$$\times\{\hat{\beta}T^2+2+(n'-p+1+i)\hat{\beta}y_{(p)}^2\}^{|D_2|+\hat{\alpha}+3/2},\ \ y_{(p)}>0 \tag{3.4}$$

Therefore one can obtain the following remark by using Theorem 3.1. and Theorem 3.2.

Remark. Under the empirical Bayesian approach, the 100(1-r)% equal-tail prediction interval and most plausible Bayesian prediction interval are identical with the cases of Bayesian approach by substituting α and β with $\hat{\alpha}$ and $\hat{\beta}$, respectively.

4. Numerical Examples

The predictive density function and prediction intervals may be changed due to the selection of prior distributions and the two approaches. Thus in this section, we study the difference of two approaches and the prior-robustness.

To predict a future observation or the p-th order statistic of n' future observations, the data were generated artificially from the Rayleigh model with parameter $\sigma^2 = 4$ under the random censoring. They are listed in Table 4.1.

Table 4.1. Random Sample of size 25 from $R(\sigma)$ under the Random Censoring.

2.240287	.765290	1.163300	2.659954	1.982586*
3.425865	1.683482	3.097933	.644166	1.371086*
4.141332	.594661*	1.795982	3.954805	4.027264
3.598400	1.028340	.753993	4.383654	2.118047
.533241	4.152646	.930757	2.253767*	1.841201*

* denotes a censored observation

Under the data in Table 4.1, to see the difference between Bayesian and empirical Bayesian approach in the equal-tail and the most plausible prediction intervals, the 95% equal-tail and the 95% most plausible prediction intervals with respect to an inverse gamma prior $IG(1,2)$ are derived and are given by

$$(C_{GL},C_{GU})=(.4589,\ 5.7719),\quad (C_{EL},C_{EU})=(.4541,\ 5.7623)$$
$$(M_{GL},M_{GU})=(.2067,\ 5.2594),\quad (M_{EL},M_{EU})=(.2036,\ 52238)$$

respectively. Also both prediction intervals for the p-th order statistic, $Y_{(p)}$, of

n' future observations are computed under two approaches and are listed in Table 4.2. From these results, we can say that there are no much differences between Bayesian and empirical Bayesian approaches for the equal-tail and the most plausible prediction intervals. Also, for study on the prior-robustness with the same data, we consider several values of the parameters α and β of an inverse gamma prior distribution, *i.e.* α = 0.5, 1.0, 1.5, 2.0, 2.5 and β = 1.0, 2.0, 3.0, 4.0, 5.0. With these values, the prediction intervals are computed and are given in Table 4.3. From Table 4.3, as both the values of α and β change, one can say that the prediction intervals are relatively insensitive to the changes of the prior distribution.

Table 4.2. Prediction Intervals of $Y_{(p)}$ $(\gamma = 0.05)$

Prior	$Y_{(p)}$	1	2	3	4	5
IG(1,2)	M.P.	(.0639,1.6493)	(.3760,2.2003)	(.6633,2.6345)	(.9278,3.0382)	(1.1823,3.4431)
	E.T.	(.1433,1.8124)	(.4508,2.3144)	(.7361,2.7386)	(1.0023,3.1424)	(1.2608,3.5521)
E.B.	M.P.	(.0654,1.6632)	(.3831,2.2168)	(.6756,2.6524)	(.9451,3.0572)	(1.2046,3.4632)
	E.T.	(.1451,1.8253)	(.4569,2.3282)	(.7466,2.7527)	(1.0172,3.1566)	(1.2801,3.5666)

Table 4.2. Prediction Intervals of y for the Prior α and β $(\gamma = 0.05)$

α	β	1	2	3	4	5
0.5	M.P.	(.2092,5.3369)	(.2086,5.3227)	(.2084,5.3180)	(.2084,5.3156)	(.2083,5.3142)
	E.T.	(.4653,5.8584)	(.4040,5.8428)	(.4636,5.8376)	(.4634,5.8349)	(.4633,5.8335)
1.0	M.P.	(.2073,5.2734)	(.2067,5.2594)	(.2065,5.2547)	(.2064,5.2524)	(.2064,5.2510)
	E.T.	(.4601,5.7874)	(.4589,5.7719)	(.4584,5.7668)	(.4582,5.7642)	(.4581,5.7627)
1.5	M.P.	(.2054,5.2122)	(.2048,5.1983)	(.2046,5.1937)	(.2045,5.1914)	(.2045,5.1900)
	E.T.	(.4550,5.7189)	(.4538,5.7036)	(.4534,5.6986)	(.4532,5.6960)	(.4531,5.6945)
2.0	M.P.	(.2035,5.1530)	(.2030,5.1393)	(.2028,5.1348)	(.2027,5.1324)	(.2027,5.1311)
	E.T.	(.4502,5.6528)	(.4490,5.6377)	(.4486,5.6327)	(.4484,5.6302)	(.4483,5.6286)
2.5	M.P.	(.2018,5.0958)	(.2012,5.0823)	(.2010,5.0777)	(.2019,5.0755)	(.2009,5.0741)
	E.T.	(.4455,5.5889)	(.4443,5.5740)	(.4439,5.5691)	(.4437,5.5665)	(.4436,5.5650)

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

References

Aitchison, J. and Dunsmore, I. R. (1975), *Statistical Prediction Analysis*, Cambridge University Press, London.

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

Clarotti, C. A. and Spizzichino, F. (1989), The Bayes Predictive Approach in Reliability Theory, *IEEE Transactions on Reliability*, R-38, 379-382.

Cohen, A. C. and Whitten, B. J. (1988), *Parameter Estimation in Reliability and Life Span Models*, Marcel Dekker, New York.

Dunsmore, I. R. (1974), The Bayesian Predictive Distribution in Testing Models,*Technometrics*, 16, 455-460.

Efron, B. and Morris, C. (1973a), Combining Possibly Related Estimation Problems, *Journal of the Royal Statistical Society*, B35, 379-421.

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

Lingappaiah, G. S. (1986), Bayes Prediction in Exponential Life-testing when Sample Size is a Random Variable, *IEEE Transactions on Reliability*, R-35, 106-110.

Miller, R. W. (1989), Parameteric Empirical Bayes Tolerance Intervals, *Technometrics*, 31, 449-459.

Morris, C. N. (1983), Parametric Empirical Bayes Inference : Theory and Applications, *Journal American Statistical Association*, 78, 47-59.

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

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.