Journal of the Korean
Data & Information Science Society
1999, Vol. 10, No.1, pp. 147 ~ 154

Bayesian Prediction Inference for Censored Pareto Model [1]

Jeong Hwan Ko[2] · Yeung-Hoon Kim[3]

Abstract

Using a noninformative prior and an inverted gamma prior, the Bayesian predictive density and the prediction intervals for a future observation or the $p-th$ order statistic of n' future observations from the censord Pareto model have been obtained. In additions, numerical examples are given in order to illustrate the proposed predictive procedure.

Key Words and Phrases: Predictive density function, Prediction Interval, Bayesian Approach, Censored Sample, Pareto Model

1. Introduction

The Pareto distribution has played an important role in the investigations of city population sizes, income distributions, insurance risk, business risk, etc. A few authors have studied the Bayesian inference procedures for the Pareto distribution. Muniruzzaman(1968) was the first to consider the Bayesian approach for classical Pareto distribution. Arnold and Press(1983), Geissor(1984), Nigm and Hamdy(1987) have discussed the Bayesian approaches in the Pareto distribution. Recently, Bayesian estimation of shape parameter of classical Pareto distribution is provided by Pandey, Singh and Mishra(1996), and Tiwari, Yang and Zalkikar(1996).

Problem of predicting a future observation has received much attention and has been dealt mainly in two approaches. One is the usual classical approach and the other is Bayesian approach. Based the Bayesian approach, Chhikara and

---

[1]The paper is supported in the part by an Andong National University Research Grant for 1997
[2]Associate Professor, Department of Statistics, Andong National University, Andong, 760-749, Korea
[3]Professor, Department of Statistics, Andong National University, Andong, 760-749, Korea

Guttman(1982), Sinha(1989), Upadhyay and Pandey(1989) and Nigm and AL- Wahab(1996) suggested the Bayesian inference about prediction for Gaussian, lognormal, exponential and Burr distributions, respectively.

In the paper, we obtain Bayesian predictive distributions and consider the prediction intervals of future observations based upon the random sampling from the censored Pareto model.

In Section 2, we consider the Bayesian predictive density and prediction intervals for a future observation from the pareto distribution. As a prior distribution, we consider a noninformative prior and an inverted gamma prior.

In Section 3, we deal with the problem of Bayesian prediction analysis for the $p-th$ order statistic of n' future observations.

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

2. Prediction of a Future Observation

Let $\underline{X} = (X_1, X_2, \cdots, X_n)$ be a random sample from the Pareto distribution with unknown parameter θ and known parameter σ whose the probability density function is given by

$$f(x|\theta,\sigma) = \theta\sigma^{\theta}x^{-(\theta+1)}, \qquad x > \sigma, \tag{1}$$

where $\sigma > 0$ is a scale parameter and $\theta > 0$ may be called the shape parameter.

Then the likelihood function of $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(r)}$ under the type II censoring at $r - th$ failure is

$$L(\theta,\sigma|\underline{x}) \propto \theta^{r}\exp\left[-\theta\left(\ln\left(x_{(r)}^{n-r}\prod_{i=1}^{r}x_i - n\ln\sigma\right)\right)\right]$$
$$\sigma > 0,\ \theta > 0,\ x_i > \sigma,\ i = 1,\cdots,r. \tag{2}$$

Now, a noninformative prior distribution for θ is given by

$$\pi(\theta) \propto \frac{1}{\theta}, \quad \theta > 0. \tag{3}$$

Then the posterior density of θ given $\underline{X} = \underline{x}$ is

$$\pi(\theta|\underline{x},\sigma) = \frac{\theta^{r-1}}{\Gamma(r)}(\ln K)^{r}\exp(-\theta\ln K), \quad 0 < \theta < \infty, \tag{4}$$

where $K = \sigma^{-n}x_{(r)}^{n-r}\left(\prod_{i=1}^{r}x_i\right)$.

The distribution of a future observation y given θ, σ is

$$f(y|\theta,\sigma,\underline{x}) = \frac{\theta\sigma^{\theta}}{y^{\theta+1}}, \quad y > \sigma, \ \theta > 0, \ \sigma > 0. \tag{5}$$

Therefore the predictive density function of a future observation y can be derived and is given in the following theorem.

Theorem 2.1 If a noninformative prior of θ is used, the predictive density function of a future observation is given by

$$\pi(y|\underline{x}) = \frac{r(\ln \mathrm{K})^{\mathrm{r}}}{y\left[\ln \mathrm{K} - \ln\left(\sigma/\mathrm{y}\right)\right]^{\mathrm{r+1}}}, \quad y \geq \sigma. \tag{6}$$

Under a noninformative prior for θ, the $100(1-\gamma)$ % equal-tail prediction interval (C_{NL}, C_{NU}) for a future observation y is

$$\left(\sigma \exp\left[(\ln K(1-\gamma/2)^{-1/\gamma} - 1)\right], \ \sigma \exp\left[(\ln K(\gamma/2)^{-1/\gamma} - 1)\right] \right)$$

Also, the $100(1-\gamma)$ % most plausible prediction bounds M_{NL} and M_{NU} are the simultaneous solution of

$$\left[\frac{\ln K}{\ln K + \ln\left(M_{NL}/\sigma\right)}\right]^{r} - \left[\frac{\ln K}{\ln K + \ln\left(M_{NU}/\sigma\right)}\right]^{r} = 1-\gamma$$

and

$$\left[\frac{\ln K + \ln\left(M_{NU}/\sigma\right)}{\ln K + \ln\left(M_{NL}/\sigma\right)}\right] = \frac{M_{NL}}{M_{NU}}.$$

Consider an inverted gamma prior distribution with parameter α , β , which is given by

$$\pi(\theta|\alpha,\beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)}\theta^{\alpha-1}\exp\left(-\beta\theta\right), \quad 0 < \theta < \infty, \ \alpha, \ \beta > 0. \tag{7}$$

Then from the equations (2) and (7), the posterior density function of θ given $\underline{X} = \underline{x}$ is given by

$$\pi(\theta|\sigma,\underline{x}) = \frac{\theta^{r+\alpha+1}}{\Gamma(r+\alpha)}(\beta + \ln K)^{r+\alpha}\exp\left[-\theta(\beta + \ln K)\right]. \tag{8}$$

Hence the predictive density of a future observation can be derived and is given by in the following theorem.

Theorem 2.2. For the inverted gamma prior with the parameters α and β for θ, the predictive density function of a future observation y is given by

$$\pi(y|\underline{x}) = \frac{(r+\alpha)(\beta+\ln K)^{r+\alpha}}{y\Big[(\beta+\ln K)-\ln(\sigma/y)\Big]^{r+\alpha+1}}, \quad y \geq \sigma. \tag{9}$$

With the inverted gamma prior distribution with parameters α and β for θ, the $100(1-\gamma)$ % equal-tail prediction limits C_{GL} and C_{GU} for y is as follows:

$$\Bigg(\sigma \exp\Big[(\beta+\ln K)((1-\gamma/2)^{-1/(r+\alpha)}-1)\Big], \quad \sigma \exp\Big[(\beta+\ln K)((\gamma/2)^{-1/(r+\alpha)}-1)\Big] \Bigg).$$

Also the $100(1-\gamma)$ % most plausible prediction interval (M_{GL}, M_{GU}) of a future observation y is obtained by solving simultaneously the following equations:

$$\left[\frac{(\beta+\ln K)}{(\beta+\ln K)+\ln(M_{GL}/\sigma)}\right]^{r+\alpha} - \left[\frac{(\beta+\ln K)}{(\beta+\ln K)+\ln(M_{GU}/\sigma)}\right]^{r+\alpha} = 1-\gamma$$

and

$$\left[\frac{(\beta+\ln K)+\ln l(M_{GU}/\sigma)}{(\beta+\ln K)+\ln(M_{GL}/\sigma)}\right]^{r+\alpha+1} = \frac{M_{GL}}{M_{GU}}.$$

3. Prediction of the Ordered Observation

In this section, we consider the predictive density function of the $p-th$ order statistic, $Y_{(p)}$, of n' future observations when the distribution follows the Pareto distribution. Here the probability density function of $Y_{(p)}$ is as follows:

$$f(y_{(p)}|\theta,\sigma) = \frac{\theta}{y_{(p)}B(p,n'-p+1)} \sum_{j=0}^{p-1} \binom{p-1}{j} (-1)^j \left(\frac{\sigma}{y_{(p)}}\right)^{(n'-p+j+1)\theta},$$

$$y_{(p)} \geq \sigma, \; 1 \geq p \geq n'. \tag{10}$$

Then the following theorem holds.

Theorem 3.1 If a noninformative prior distribution for θ is used, the predictive density function of $p-th$ order statistic, $y_{(p)}$, of n' future observations is given by

$$\begin{aligned}\pi(y_{(p)}|\underline{x}) &= \frac{r(\ln K)^r}{y_{(p)}B(p,n'-p+1)}\sum_{j=0}^{p-1}\binom{p-1}{j}(-1)^j \\ &\times \left[\ln K-(n'-p+j+1)\ln\left(\frac{\sigma}{y_{(p)}}\right)\right]^{-(r+1)}, \quad y_{(p)}\geq\sigma. \qquad (11)\end{aligned}$$

With a noninformative prior distribution of θ, the $100(1-\gamma)$ % equal-tail prediction interval (C_{NL}, C_{NU}) for the $p-th$ order statistics, $y_{(p)}$, in a future sample of n' items are the solutions of the equations:

$$\frac{\gamma}{2}=\frac{1}{B(p,n'-p+1)}\sum_{j=0}^{p-1}\binom{p-1}{j}(-1)^j(n'-p+j+1)^{-1} \times\left\{1-\frac{\ln K}{\ln K+(n'-p+j+1)\ln\left(C_{NU}/\sigma\right)}\right\}^r$$

and

$$\frac{\gamma}{2}=\frac{1}{B(p,n'-p+1)}\sum_{j=0}^{p-1}\binom{p-1}{j}(-1)^j(n'-p+j+1)^{-1} \times\left\{\frac{\ln K}{\ln K+(n'-p+j+1)\ln\left(C_{NL}/\sigma\right)}\right\}^r.$$

Furthermore, the $100(1-\gamma)$ % most plausible prediction interval (M_{NL}, M_{MU}) for $y_{(p)}$ is obtained by solving simultaneously the followings:

$$\begin{aligned}&\frac{1}{B(p,n'-p+1)}\sum_{j=0}^{p-1}\binom{p-1}{j}(-1)^j(n-p+j+1)^{-1} \\ &\times\left[1+\frac{(n'-p+j+1)\ln(M_{NL}/\sigma)}{\ln K}\right]^{-r} \\ -\;&\frac{1}{B(p,n'-p+1)}\sum_{j=0}^{p-1}\binom{p-1}{j}(-1)^j(n-p+j+1)^{-1} \\ &\times\left[1+\frac{(n'-p+j+1)\ln(M_{NU}/\sigma)}{\ln K}\right]^{-r} \\ =\;&1-\gamma\end{aligned}$$

and

$$M_{NU}\sum_{j=0}^{p-1}\binom{p-1}{j}(-1)^j(n'-p+j+1)^{-1}\left[1+\frac{(n'-p+j+1)\ln(M_{NL}/\sigma)}{\ln K}\right]^{-(r+1)}$$

$$= M_{NL}\sum_{j=0}^{p-1}\binom{p-1}{j}(-1)^j(n'-p+j+1)^{-1}\left[1+\frac{(n'-p+j+1)\ln(M_{NU}/\sigma)}{\ln K}\right]^{-(r+1)}.$$

Theorem 3.2. With the inverted gamma prior with parameters α and β for θ, the predictive density function of $y_{(p)}$ is

$$\pi(y_{(p)}|\underline{x}) = \frac{(\beta+\ln K)^{r+\alpha}(r+\alpha)}{y_{(p)}B(p,n'-p+1)}\sum_{j=0}^{p-1}\binom{p-1}{j}(-1)^j \times\left[(\beta+\ln K)-(n'-p+j+1)\ln\left(\frac{\sigma}{y_{(p)}}\right)\right]^{-(r+\alpha+1)}, y(p)\ge\sigma. \quad (12)$$

If the inverted gamma prior with parameters α and β for θ is used, then the $100(1-\gamma)$ % equal-tail prediction interval (C_{GL}, C_{GU}) of $y_{(p)}$ can be obtained by solving two equations

$$\frac{\gamma}{2} = \frac{1}{B(p,n'-p+1)}\sum_{j=0}^{p-1}\binom{p-1}{j}(-1)^j(n'-p+j+1)^{-1} \times\left\{1-\frac{\beta+\ln K}{(\beta+\ln K)+(n'-p+j+1)\ln(C_{GL}/\sigma)}\right\}^{r+\alpha}$$

and

$$\frac{\gamma}{2} = \frac{1}{B(p,n'-p+1)}\sum_{j=0}^{p-1}\binom{p-1}{j}(-1)^j(n'-p+j+1)^{-1} \times\left\{\frac{\beta+\ln K}{(\beta+\ln K)+(n'-p+j+1)\ln(C_{GU}/\sigma)}\right\}^{r+\alpha}.$$

Also the $100(1-\gamma)$% most plausible prediction bounds M_{GL} and M_{GU} for the $p-th$ order statistics, $y_{(p)}$, in a future sample of n' items can be obtained by solving the following equations simultaneously:

$$\frac{1}{B(p,n'-p+1)}\sum_{j=0}^{p-1}\binom{p-1}{j}(-1)^j(n-p+j+1)^{-1}$$
$$\times\left[1+\frac{(n'-p+j+1)\ln(M_{GL}/\sigma)}{(\beta+\ln K)}\right]^{-(r+\alpha)}$$
$$-\frac{1}{B(p,n'-p+1)}\sum_{j=0}^{p-1}\binom{p-1}{j}(-1)^j(n-p+j+1)^{-1}$$
$$\times\left[1+\frac{(n'-p+j+1)\ln(M_{GU}/\sigma)}{(\beta+\ln K)}\right]^{-(r+\alpha)}$$
$$=1-\gamma$$

and

$$M_{GU}\sum_{j=0}^{p-1}\binom{p-1}{j}(-1)^j(n'-p+j+1)^{-1}\left[1+\frac{(n'-p+j+1)\ln(M_{GL}/\sigma)}{(\beta+\ln K)}\right]^{-(r+\alpha+1)}$$
$$=M_{GL}\sum_{j=0}^{p-1}\binom{p-1}{j}(-1)^j(n'-p+j+1)^{-1}\left[1+\frac{(n'-p+j+1)\ln(M_{GU}/\sigma)}{(\beta+\ln K)}\right]^{-(r+\alpha+1)}.$$

4. Numerical Examples

In this Section, to predict a future observation and the $p-th$ order statistic of n' future observation, the data were generated artifically from the Pareto model with unkown parameter θ and kown parameter σ under the Type II censoring. It is assumed that only the first 20(twenty percent cesoring) ordered failure times are available, and they are given as follows:

2.0329 2.2015 2.4660 2.0385 5.0060 3.2459 2.9942 4.3439 5.4538 2.0346
5.1906 2.5148 2.0626 2.4487 3.2138 4.2926 2.3897 3.5802 2.0823 3.0921

Under a noninformative prior and an inverted gamma prior, the 95 % equal-tail prediction intervals and the 95 % highest posterior density prediction intervals are given by

$$(C_{NL},\ C_{NU}) = (2.0328,\ 26.9943),\quad (C_{GL},\ C_{GU}) = (2.0361,\ 34.6239)$$

and

$$(M_{NL},\ M_{NU}) = (2.0000,\ 15.9463),\quad (M_{GL},\ M_{GU}) = (2.0000,\ 19.4891)$$

respectively. From this result, one can see that both prediction intervals sensitive to the changes of the prior.

References

1. Arnold, B. C. and Press, S. J. (1983). Bayesian Inference for Pareto Populations, *Journal of Econometrics*, 21, 287-306.

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

3. Geisser, S. (1984). Predicting Pareto and Exponential Observables, *The Canadian J. of Statistics*— 12, 143-152.

4. Muniruzzaman, A. N. M. (1968). Estimation of Parameters of Pareto Distribution. A Bayesian approach in : Statistics for Development, *Preceedings (Institute of Statistical Research and Training, University of Dacca).*

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

6. Nigm, A. M. and AL-Wahab, N. Y. ABD (1996). Bayesian Prediction With Sample Size for the Burr Lifetime Distribution. *Communications in Statistics Part A - Theory and Methods*, 25, 1289-1303.

7. Pandey, B. N., Singh, B. P. and Mishra, C. S. (1996). Bayes Estimation of Shape Parameter of Classical Pareto Distribution under LINEX Loss Function, *Communications in Statistics Part A - Theory and Methods*, 25, 3125-3145.

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. Tiwari, R. C., Yang, Y. and Zalkikar, J. N. (1996). Bayes Estimation for the Pareto Failue - Model Using Gibbs Sampling, *IEEE Transactions on Reliability*, R-45, 471-476.

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.