J. Sci. I. R. Iran

Vol.10 No.1
Winter 1999

BAYES PREDICTION INTERVALS FOR THE BURR TYPE XII DISTRIBUTION IN THE PRESENCE OF OUTLIERS

M.S. Abu-Salih and M.S. Ahmed

Department of Statistics, Yarmouk University, Irbid, Jordan

Abstract

Using a sample from Burr type XII distribution, Bayes prediction intervals are derived for the maximum and minimum of a future sample from the same distribution, but in the presence of a single outlier of the type $\theta_0\theta$. The prior of Θ is assumed to be the gamma conjugate. A real example is given to illustrate the procedure. Also, the comparison between the values of the prediction bounds for different values of θ_0 and different future sample sizes is given.

1. Introduction

The outcome of an experiment may be used to predict some feature of a future experiment under the same or different conditions in the form of prediction intervals for some statistics. The prediction intervals for order statistics of a future sample from the population of the observed sample have been considered by many authors including Lawless [7,8], Likes [9], Dunsmore [3], Lingappaiah [10], Evans and Nijm [4], Abu-Salih et al. [1], Sartawi and Abu-Salih [14], and Abu-Salih and Sartawi [2].

Evans and Ragab [5] and Nigm [13] gave the prediction bounds for the k-th order statistic in a sample from a Burr type XII distribution in the case of a censored sample.

The aforementioned works assumed the future sample to be from the original population under the same conditions. In practice the conditions may change, and it may happen that some values of the future sample are far away from the main group. Such values are called outliers.

Lingappaiah [11,12] obtained Bayes prediction bounds for future observations from the exponential life time distribution in the presence of outliers. The case of Burr type X distribution was treated by Jaheen [6].

In this work we shall be concerned with Bayes prediction bounds for the maximum and minimum of a future sample from a Burr type XII distribution in the presence of a single outlier of the type $\theta_0\theta$, where θ is the unknown parameter of the distribution. This means that the process has undergone some abrupt change which is specified by the new value of the parameter Viz, $\theta_0\theta$, where θ_0 is known.

The probability density function (p.d.f.) of the Burr type XII distribution is given by

$$f(x|c, \theta)= \theta c x^{c-1} (1-x^c)^{-(\theta+1)}, \quad x > 0 \; (c > 0, \theta > 0) \quad (1)$$

The cumulative distribution function (c.d.f.) is given by

$$F(x)= 1-(1+x^c)^{-\theta} \quad , \; x>0$$

2. Main Results

Let $X_1,X_2,\ldots,X_n$ be a random sample of size n from a population defined by (1). Suppose $Y_1,Y_2,\ldots Y_m$ is an independent random sample of future observations from

Keywords: Bayes Prediction Intervals, Burr Type XII Distribution, Outliers

the same distribution. Assuming c to be known, we derive Bayesian prediction bounds for the maximum and for the minimum of the future observations in the presence of a single outiler of the type $\theta_0\theta$.

2.1 Bayesian Prediction Bounds for the Maximum

It is known that the p.d.f. g(y) of the maximum in a sample of size m is given by

$$g(y)= m\,\{F(y)\}^{m-1} f(y) \tag{2}$$

where f(y) and F(y) are the p.d.f. and c.d.f. of Y respectively.

In the presence of a single outlier, Equation (2) becomes

$$h(y)= (m-1)\,\{F(y)\}^{m-2} F^*(y)\, f(y)+\{F(y)\}^{m-1} f^*(y) \tag{3}$$

where $f^*(y)$ and $F^*(y)$ are the p.d.f. and c.d.f. of the outlier Y.

Let $U= \max(Y_1,Y_2,\ldots,Y_m)$, then in the presence of a single outlier of the type $\theta_0\theta$, Equation (3) becomes

$$\begin{aligned} h(u|c,\theta) &= (m-1)\, c\theta u^{c-1} (1+u^c)^{-(q+1)} [1- (1+u^c)^{-\theta}]^{m-2}. [1-(1+u^c)^{-\theta_0\theta}] + c\theta_0\theta u^{c-1}(1+u^c)^{-(\theta_0\theta+1)} [1-(1+u^c)^{-\theta})]^{m-1}, \quad u>0 \\ &= (m-1)c\theta\psi(u)e^{-0\phi(u)}(1-e^{-0\phi(u)})^{m-2}(1-e^{-\theta_0\theta\phi(u)}) \\ &\quad +c\theta_0\theta\psi(u)e^{-\theta_0\theta\phi(u)}(1-e-^{\theta\phi}(u))^{m-1}, \quad u > 0 \end{aligned} \tag{4}$$

where

$$\phi(u)= \ln (1+u^c)$$

and

$$\psi(u)= \frac{u^{c-1}}{1+u^c}$$

Using the binomial expansion, we write (4) in the form

$$h (u \mid \theta) = c\theta\psi (u) \sum_{j=0}^{m-2} (-1)^j \binom{m-2}{j} [m-1)\, e^{-\theta(j+1)\phi(u)} + \theta_0 e^{-\theta(j+\theta 0)\phi(u)}-(m+\theta_0-1)e^{-\theta(j+\theta_0+1)\phi(u)}] \tag{5}$$

For the first random sample $X_1,X_2,\ldots,X_n$, the likely function will be

$$L(\theta|\underline{x})= c^n\theta^n\psi^* e^{-\theta T}, \quad q > 0 \tag{6}$$

where $T= \sum_{i=1}^{n} f(x_i)$, $\psi^* = \prod_{i=1}^{n} \psi(x_i)$, and ψ and ϕ are as defined in (4).

Let the prior p.d.f. of Θ be' the conjugate prior, namely, Gamma (α, $1/\beta$):

$$\prod (\theta,\alpha,\beta)= \frac{\beta^\alpha}{\Gamma(\alpha)}\, \theta^{\alpha-1}e^{-\beta\theta}, \ \theta > 0, (\alpha > 0, \beta > 0) \tag{7}$$

Combining equations (6) and (7) the posterior p.d.f. of Θ is given by:

$$\prod{}^*(\theta|\underline{x})= \frac{(T+\beta)^{n+\alpha}\, \theta^{n+\alpha-1}}{\Gamma(n+\alpha)}\, e^{-(T+\beta)\theta}, \ \theta>0 \tag{8}$$

The Bayes predictive p.d.f. $p(u|\underline{x})$ of U given $\underline{x}$ is given by

$$p(u|\underline{x})= \int_\theta h(u|\theta) \prod{}^*(\theta|x)d\theta \tag{9}$$

Substituting (5) and (8) in (9), we get

$$\begin{aligned} p(u|\underline{x}) &= c(n+\alpha)\, (\beta+T)^{n+\alpha}\psi(u) \times \sum_{j=0}^{m-2} (-1)^j \binom{m-2}{j} \\ &\left[\frac{(m-1)}{\{\beta+T+(j+1)\, \phi(u)\}^{n+\alpha+1}} + \frac{\theta_0}{\{\beta+T+(j+\theta_0)\phi(u)\}^{n+\alpha+1}} - \frac{(m+\theta_0-1)}{\{\beta+T+(j+\theta_0+1)\, \phi(u)\}^{n+\alpha+1}}\right], \ u > 0 \end{aligned} \tag{10}$$

To construct prediction intervals for U based on the observations $x_1,x_2,\ldots,x_n$ we evaluate

$$\begin{aligned} P(U \geq a |\underline{x}) &= \int_a^\infty p(u|\underline{x})\, du = \sum_{j=0}^{m-2} (-1)^j \binom{m-2}{j} \\ &\left[\frac{m-1}{j+1}\left(\frac{\beta+T}{\beta+T+(j+1)\phi(a)}\right)^{n+\alpha} + \left(\frac{\theta_0}{j+\theta_0} - \frac{m+\theta_0-1}{j+\theta_0+1}\right) \left(\frac{\beta+T}{\beta+T+(j+\theta_0)\phi(a)}\right)^{n+\alpha}\right] \end{aligned} \tag{11}$$

where $\phi(a)= \ln(1+a^c)$

Since $\phi(0)= 0$, it is easy to check that $P(U\geq 0|\underline{x})= 1$ by using the following formula

$$\sum_{j=0}^{n} (-1)^j \binom{n}{j} (j+\varepsilon)^{-1}= \frac{n!}{\prod_{j=0}^{n} (j+\varepsilon)} \tag{12}$$

given by Lingappaiah [11,12].

The (1-γ) 100% Bayes prediction interval for U is $(L_1(\underline{x}), L_2(\underline{x}))$ such that $P(L_1(\underline{x}) \le U \le L_2(\underline{x}) \mid \underline{x}) = 1-\gamma$.

By taking equal tail probabilities, L_1 and L_2 are found by the use of (11) such that

$$P(U \ge L_1(\underline{x})|\underline{x}) - 1 + \gamma/2$$
$$P(U \ge L_2(\underline{x})|\underline{x}) - \gamma/2.$$

L_1 and L_2 are obtained by the bisection method.

2.2 Bayesian Prediction Bound for the Minimum

The distribution of the minimum in a sample of size m is given by:

$$q(y) = m(1-Fy)^{m-1}).\ f(y) \qquad (13)$$

In the presence of a single outlier, Equation (13) becomes

$$q(y) = (1-F(y))^{m-1} f^*(y) + (m-1)(1-F(y))^{m-2}(1-F^*(y))\ f(y) \qquad (14)$$

where f, F, f* and F* are as defined in section (2-1).

For the Burr XII distribution given in (1), with an outlier of the type $\theta_0\theta$, and V= min Y_i, Equation (14) becomes

$$\theta(v|\theta) = (1+v^c)^{-\theta(m-1)}.\ c\theta_0\theta v^{c-1}\ (1+v^c)^{-(\theta_0\theta+1)} + (m-1)\ (1-v^c)^{-\theta(m-2)}.\ (1+v^c)^{-\theta_0\theta}.\ c\theta v^{c-1}(1+v^c)^{-(\theta+1)},\quad v>0 \qquad (15)$$

Equation (15) simplifies to

$$q(v|\theta) = c\theta\psi(v)\ (m+\theta_0-1)\ \exp\{-(m+0_0-1)\phi(v)0\},\ v>0 \qquad (16)$$

The Bayes predictive p.d.f. p(v|x) can be defined using equation (8) to be

$$p(v|\underline{x}) = \int q(v|\theta)\ .\ \Pi^*(\theta|\underline{x})\ d\theta$$

Carrying out the integration, we get

$$p(v|\underline{x}) = c(n+\alpha)\ (\beta+T)n+\alpha\ \psi(v) \times \frac{(m+\theta_0-1)}{\{\beta+T+(m+\theta_0-1)\ \phi\ (v)\}^{n+\alpha+1}}\ v>0 \qquad (17)$$

To construct Bayes prediction intervals for V based on the observations $x_1, x_2, \dots x_n$, we evaluate

$$P(V \ge b|\underline{x}) = \int_b^\infty p(v|\underline{x})dv = \left(\frac{\beta+T}{\beta+T+(m+\theta_0-1)\ \phi\ (b)}\right)^{n+\alpha},\ b\ge 0$$

(18)

Since $\phi(b) = \ln(1+b^c)$, it is obvious that $P(V \ge 0|\underline{x}) = 1$.

For the equal tail probability (1-γ) 100% prediction intervals for V is $(L_1(\underline{x}), L2(x))$ such that $P(V \ge L_1|\underline{x}) = 1-\frac{\gamma}{2}$ and $P(V \ge L_2|\underline{x}) = \frac{\gamma}{2}$.

Using (18), we get

$$L_i(\underline{x}) = \left[\exp\left(\frac{(1-\zeta_i)\ (\beta+T)}{(m+\theta_0-1)\ \zeta_i}\right) - 1\right]^c,\ \text{for } i=1,2 \text{ where}$$

$$\zeta_1 = \left(1-\frac{\gamma}{2}\right)^{\frac{1}{n+\alpha}},\quad \zeta_2 = \left(\frac{\gamma}{2}\right)^{\frac{1}{n+\alpha}}$$

3. Example

Wingo [15] reported the relief times (in hours) of 50 arthritis patients receiving a fixed dosage of an analgesic. The data is listed in Table 1. Past experience with similar data suggests that the patients relief times could be adequately described by a Burr Type XII distribution (c.f. Wingo [15]).

Wingo [15] found the maximum likelihood estimate of c to be $\hat{c} = 5.000$.

Table 1. Relief times (in hours) of 50 arthritis patients

5	10	15	20	25	30	35	40	45	50
0.70	0.84	0.58	0.50	0.55	0.82	0.59	0.71	0.72	0.61
0.62	0.49	0.54	0.72	0.36	0.71	0.35	0.64	0.85	0.55
0.59	0.29	0.75	0.53	0.46	0.60	0.60	0.36	0.52	0.68
0.80	0.55	0.84	0.70	0.34	0.70	0.49	0.56	0.71	0.61
0.57	0.73	0.75	0.58	0.44	0.81	0.80	0.87	0.29	0.50

Table 2. 95% Bayesian prediction bounds for the minimum in the presence of a single outlier of type $\theta_0\theta$

(α, β)	θ_0	m= 10		m= 20	
		L_1	L_2	L_1	L_2
(0.8, 0.1)	0.1	0.232	0.545	0.200	0.469
	0.3	0.231	0.543	0.200	0.468
	0.5	0.230	0.540	0.199	0.467
	1	0.228	0.535	0.198	0.465
	3	0.220	0.515	0.195	0.456
	5	0.213	0.499	0.191	0.448
(2.4, 0.3)	0.1	0.202	0.545	0.174	0.469
	0.3	0.201	0.543	0.174	0.468
	0.5	0.200	0.541	0.174	0.467
	1	0.198	0.535	0.173	0.465
	3	0.191	0.516	0.169	0.456
	5	0.185	0.500	0.166	0.448
(4, 0.5)	0.1	0.202	0.545	0.174	0.470
	0.3	0.201	0.543	0.174	0.469
	0.5	0.200	0.541	0.174	0.468
	1	0.198	0.535	0.173	0.465
	3	0.191	0.516	0.169	0.456
	5	0.185	0.500	0.166	0.448
(5.6, 0.7)	0.1	0.202	0.546	0.174	0.470
	0.3	0.201	0.543	0.174	0.469
	0.5	0.200	0.541	0.174	0.468
	1	0.198	0.535	0.173	0.465
	3	0.191	0.516	0.169	0.457
	5	0.185	0.500	0.167	0.449
(7.2, 0.9)	0.1	0.202	0.546	0.174	0.470
	0.3	0.201	0.544	0.174	0.469
	0.5	0.200	0.541	0.174	0.468
	1	0.198	0.536	0.173	0.466
	3	0.191	0.516	0.169	0.457
	5	0.185	0.500	0.167	0.449

Table 3. 95% Bayesian prediction bounds for the maximum in the presence of a single outlier of type $\theta_0\theta$

(α, β)	θ_0	m= 10 L_1	m= 10 L_2	m= 20 L_1	m= 20 L_2
(0.8, 0.1)	0.1	0.649	0.944	0.708	0.981
	0.3	0.658	0.961	0.726	1.004
	0.5	0.663	0.976	0.733	1.021
	1	0.668	0.999	0.738	1.047
	3	0.671	1.022	0.740	1.061
	5	0.671	1.023	0.740	1.062
(2.4, 0.3)	0.1	0.650	0.945	0.709	0.982
	0.3	0.659	962	727	1.005
	0.5	0.664	0.976	0.733	1.023
	1	0.669	1.000	0.739	1.047
	3	0.672	1.021	0.741	1.061
	5	0.672	1.023	0.741	1.061
(4, 0.5)	0.1	0.651	0.946	0.711	0.936
	0.3	0.660	0.964	0.727	1.006
	0.5	0.664	0.977	0.734	1.023
	1	0.669	1.000	0.739	1.047
	3	0.672	1.021	0.741	1.061
	5	0.672	1.023	0.741	1.061
(5.6, 0.7)	0.1	0.653	0.947	0.712	0.985
	0.3	0.660	0.965	0.728	1.007
	0.5	0.665	0.978	0.735	1.024
	1	0.670	1.001	0.740	1.048
	3	0.672	1.021	0.741	1.061
	5	0.672	1.023	0.741	1.061
(7.2, 0.9)	0.1	0.653	0.948	0.714	0.986
	0.3	0.661	0.966	0.729	1.008
	0.5	0.665	0.979	0.735	1.024
	1	0.670	1.000	0.740	1.048
	3	0.672	1.021	0.742	1.060
	5	0.672	1.022	0.742	1.061

Using the given 50 observations we calculate Bayes prediction intervals for U and V, the maximum and minimum of a future sample of the relief times (in hours) for a sample of size m of arthritis patients receiving the same medication as the patients of the first sample, under the assumption that one outlier of type $\theta_0\theta$ is present. The results for m= 10,20 and θ_0= 0.1(0.2) 0.5; 1(2) 5; γ= 0.05 are reported in Tables 2 and 3.

4. Comments

(1) It appears from Table (2) that for fixed values of α and β the bounds L_1 and L_2 decrease as θ_0 increases. Also, when $\alpha \geq 4$ and the corresponding $\beta \geq 0.5$ the values of L_1 and L_2 are stablized for each value of θ_0.

(2) It appears from Table (3) that for fixed values of α and β the bounds L_1 and L_2 increase as θ_0 increases. Also, when a increases and the corresponding β decreases the values of L_1 and L_2 increase also.

(3) It appears that the length of the confidence interval is little larger in the case of the minimum when compared to that of the maximum.

Also the length is more or less a constant.

There is no sense in comparing these lengths with those obtained by other methods because the models used are different.

References

1. Abu-Salih, M.S., Ali Khan, M.S. and Husein, A. Prediction intervals of order statistics for the mixture of two exponential distributions. *Aligarh, J. Statist.*, **7**, 11-22, (1987).
2. Abu-Salih M.S. and Sartawi, H.A., Bayesian prediction bounds for the Burr type III model. *Dirasat, 19B* (Pure and Applied Sciences), **3**, 117-200, (1992).
3. Dunsmore, I.R. The Bayesian predictive distribution in life testing models. *Technometrics*. **16**, 455-460, (1974).
4. Evans, I.G. and Nigm, A.M. Bayesian prediction for two-parameter Weibull lifetime models. *Comm., in Statisti*, **A9**, 649-658, (1980).
5. Evans, I.G. and Ragab, A.S., Bayesian inferences given a type-2 censored sample from a Burr distribution. *Ibid.*, **A12**, 1569-1580, (1983).
6. Jaheen, Z.F. Bayesian approach to prediction with outliers from the Burr type X model. *Microelectron Reliab.*, **35**, 4, 703-705, (1995).
7. Lawless, J.F. A prediction probelm concerning samples from the exponential distribution with application in life testing. *Technometrics* **13**, 725-730, (1971).
8. Lawless, J.F. On the estimation of safe life when the underlying distribution is Weibull. *Ibid.*, **15**, 857-65, (1972).
9. Likes, J. Prediction of s-th order observation from two parameter exponential distribution. *Ibid.*, **16**, 241-244, (1974).
10. Lingappaiah, G.S. Prediction in samples from the gamma distribution as applied to life testing. *Austral. J. Statist.*, **16**, 30-32, (1974).
11. Lingappaiah, G.S. Bayes prediction of maxima and minima in exponential life tests in the presence of outliers. *J. Indust. Math. Soc.* **39**(2), 169-182, (1989).
12. Lingappaiah, G.S. Bayes approach to prediction in samples from gamma population when outliers are present. *Indian J. Pure Appl. Math* **20**(9), 858-870, (1989).
13. Nigm, A.M. Prediction bounds for the Burr model. *Comm., in Statist.*, **17**, 287-297, (1988).
14. Sartawi, H.A., and Abu-Salih, M.S. Bayesian prediction bounds for the Burr type X model. *Commun. Statisti Theory Meth.*, **20(7)**, 2307-2330, (1991).
15. Wingo, D.R. Maximum likelihood methods for fitting the Burr type XII distributions to life test data. *Biometric Journal*, **25**, 77-84, (1983).