International Journal of Statistical Sciences *ISSN 1683–5603*
Vol. 6 (Special Issue), 2007, pp 127-136

Bayesian Prediction for the Weibull Distribution

Manisha Pal and Asis Kumar Chattopadhyay
Dept. of Statistics
Calcutta University
Calcutta, India

[Received January 25, 2005; Revised June 17, 2007; Accepted July 8, 2007]

Abstract

In this paper we discuss the problem of predicting the unobserved lifetimes in a type II censored random sample from the Weibull life distribution, using the Bayesian approach. We also attempt to predict the time to the first failure in a future sample based on information from previous samples.

Keywords and Phrases: Bayesian prediction, Weibull distribution, Exponential-exponential prior, Type II censored sample.

AMS Classification: Primary 60G25; Secondary 62F25.

1 Introduction

Prediction problems arise in many real life situations. Various disciplines in which prediction finds importance are medicine, engineering, and business, among others. In the context of reliability theory, one may be interested in predicting the time to the next failure or in predicting the failure times in a new experiment. Statistical prediction uses data from an informative experiment in order to make some statement about the future outcome.

In the problem of estimating the predictive probabilities of an unobserved random variable, whose distribution is dependent on an unknown finite dimensional parameter, a crude procedure would be to substitute a point estimate of the parameter, obtained from past data, in the predictive distribution. However, this fails to take into account any uncertainty that may be present in the parameter. A better method, therefore, would be to adopt the Bayesian approach, which assigns a suitable prior distribution to explain the uncertainty. There have been many studies in the area of predictive inference. The pioneering works include Aitchison (1964), Aitchison and Sculthorpe

(1965) among others. Some recent studies in this area are due to Abd Ellah (2003), Khan and Chattopadhyay (2003), Khan (1996, 2004, 2006), who carried out predictive analyses under different set-ups.
In this paper we attempt to predict the failure times, which are precisely the order statistics in a sample, on the basis of a set of available lifetime data from the Weibull distribution. The Weibull distribution is an important lifetime distribution having the p.d.f.

$$f(t) = \alpha\beta t^{\alpha-1} e^{-\beta t^{\alpha}}, \alpha, \beta > 0, t > 0. \tag{1}$$

where β is the scale parameter and α the shape parameter.
The two-parameter Weibull distribution is one of the most commonly used lifetime distributions in survival analysis. The survival function and failure rate of the distribution have very simple and easy-to-study forms. In recent years, the Weibull distribution has become rather popular in analyzing lifetime data and, in the presence of censoring, it is found to be very easy to handle.
In Section 2 we consider a type II censored sample of size n from the Weibull distribution with r observations available, and predict the remaining (n-r) unobserved lifetimes. In Section 3 we predict the time to the first failure in a future sample based on information from previous samples. The Bayesian approach is used for the predictions.

2 Bayesian one Sample Prediction

Let $x_1, x_2, \ldots, x_r$ be the ordered failure times in a sample of n items whose failure times are independently and identically distributed with p.d.f. given by (1).
Based on the first r observations, the likelihood function becomes

$$\begin{aligned} L(\alpha, \beta, \underline{\mathrm{x}}) &= \frac{n!}{(n-r)!}\alpha^r \beta^r \left(\prod_{i=1}^{r} x_i^{\alpha-1} e^{-\beta x_i^{\alpha}}\right) e^{-(n-r)\beta x_r^{\alpha}}, 0 < x_1 < x_2 < \ldots < x_r, \\ & \qquad \underline{\mathrm{x}} = (x_1, x_2, \ldots, x_r)' \\ &= \frac{n!}{(n-r)!}(\alpha\beta)^r e^{-\beta \sum_{i=1}^{r} x^{\alpha}, -(n-r)x^{\alpha},} \prod_{i=1}^{r} x_i^{\alpha-1}. \end{aligned}$$

For the remaining (n-r) items let $y_s = x_{r+s}$, the time to failure of the sth item, $1 \leq s \leq n-r$.
Then, the conditional p.d.f. of the failure time of the sth item, $1 \leq s \leq n-r$, given that r items have already failed, is given by

$$
\begin{aligned}
h_r(y_s \mid \alpha, \beta) &= D(s)\alpha\beta y_s^{\alpha-1} e^{(n-r)\beta x_r^\alpha - \beta x_r^\alpha - \beta y_s^\alpha m(s)} [e^{-\beta x_r^\alpha} - e^{-\beta y_s^\alpha}]^{s-1} \\
&= D(s)\alpha\beta y_s^{\alpha-1} \sum_{j-0}^{s-1} (-1)^j \binom{s-1}{j} \left(\frac{e^{-\beta y_s^\alpha}}{e^{-\beta x_r^\alpha}} \right)^{k(s,j)}, y_s > x_r
\end{aligned}
$$

where

$$
D(s) = s \binom{n-r}{s}, m(s) = n - s - r + 1, k(s,j) = n - s - r + j + 1.
$$

Now, in choosing the prior for (α, β), we note that for the one parameter exponential family, the most popular choice of prior is the conjugate prior gamma distribution. More generally, one can use a non-conjugate prior, but this does not change the qualitative conclusion that if the value to be predicted is sufficiently large, the maximum likelihood estimator performs better than the Bayes estimator under squared error loss (see Bernardo et al., 1998). Following the above argument, we consider the exponential-exponential prior density for (α, β), given by

$g(\alpha, \beta) = g_1(\beta \mid \alpha) g_2(\alpha)$,

where

$g_1(\beta \mid \alpha) = \frac{1}{\alpha} e^{-\beta/\alpha}, \alpha, \beta > 0$

$g_2(\alpha) = \frac{1}{\theta} e^{-\alpha/\theta}, \alpha, \theta > 0,$

so that

$g(\alpha, \beta) = \frac{1}{\alpha\theta} e^{-\left(\frac{\alpha}{\theta} + \frac{\beta}{\alpha}\right)}, \alpha, \beta, \theta > 0.$

The joint density of $\underline{x}$, α and β therefore is $g(\alpha, \beta) L(\alpha, \beta, \underline{x})$.
Hence the density of $\underline{x}$ comes out as

$$
\frac{n!}{(n-r)!} \frac{1}{\beta \prod_{i=1}^{r} x_i} B^{-1},
$$

where

$$
B^{-1} = \Gamma(r+1) \int_0^\infty \frac{e^{-\alpha/\theta} \alpha^{r-1} \prod_{i=1}^{r} x_i^\alpha}{\left(\sum_{i=1}^{r} x_i^\alpha + (n-r) x_r^\alpha + 1/\alpha \right)^{r+1}} d\alpha.
$$

The posterior joint density of α and β given $\underline{x}$ is therefore given by

$$q(\alpha,\beta \mid \underline{x}) = B\alpha^{r-1}\beta^{r}e^{-\left(\frac{\alpha}{\theta}+\frac{\beta}{\alpha}\right)}e^{-\beta\sum_{i=1}^{r}x_i^{\alpha}-(n-r)\beta x_r^{\alpha}}\prod_{i=1}^{r}x_i^{\alpha}, \alpha,\beta>0.$$

Hence the Bayes predictive density of y_s is

$$\begin{aligned} f^*(y_s \mid \underline{x}) &= \int_0^\infty\int_0^\infty h_r(y_s \mid \alpha,\beta)q(\alpha,\beta \mid \underline{x})d\alpha d\beta \\ &= (r+1)D(s)\sum_{j=0}^{s-1}A_j(s)I_{js}(y_s \mid \underline{x})/I_0(\underline{x}), \end{aligned}$$

where

$$A_j(s) = (-1)^j\binom{s-1}{j}$$

$$I_{js}(y_s \mid \underline{x}) = \int_0^\infty \alpha^r e^{-\alpha/\theta}\frac{y_s^{\alpha-1}\prod_{i=1}^{r}x_i^{\alpha}}{\left[\frac{1}{\alpha}+\sum_{i-1}^{r}x_i^{\alpha}+(s-j-1)x_r^{\alpha}+k(s,j)y_s^{\alpha}\right]^{r+2}}d\alpha$$

$$I_0(\underline{x}) = \int_0^\infty \alpha^{r-1}e^{-\alpha/\theta}\frac{\prod_{i=1}^{r}x_i^{\alpha}}{\left[\frac{1}{\alpha}+\sum_{i-1}^{r}x_i^{\alpha}+(n-r)x_r^{\alpha}\right]^{r+1}}d\alpha$$

A $100\delta\%$ prediction interval for $y_s = x_{r+s}, 1 \le s \le (n-r)$, will therefore be given by $[l(\underline{x}), u(\underline{x})]$ where $Pr[l(\underline{x}) \le x_{r+s} \le u(\underline{x})] = \delta$, and l($\underline{x}$) and $u(\underline{x})$ satisfy

$$Pr[y_s \ge l(\underline{x})] = \frac{1+\delta}{2} \text{ and } Pr[y_s \ge u(\underline{x})] = \frac{1-\delta}{2}. \tag{2}$$

As the Bayes predictive distribution of y_s is absolutely continuous, $l(\underline{x})$ and $u(\underline{x})$ will be uniquely determined from (2).

To find $l(\underline{x})$ and $u(\underline{x})$ we note that for any $\lambda > 0$,

$$
\begin{aligned}
Pr(y_s \geq \lambda \mid \underline{x}) &= \int_{\lambda}^{\infty} f^*(y_s \mid \underline{x}) dy_s \\
&= (r+1)D(s)\sum_{j=0}^{s-1}\frac{A_j(s)}{I_0(\underline{x})}\int_{\lambda}^{\infty} I_{js}(y_s \mid \underline{x}) dy_s, \\
&= D(s)\int_{0}^{\infty}\frac{\alpha^{r-1}e^{\alpha/\theta}\prod_{i=1}^{r}x_i^{\alpha}}{k(s,j)}.\frac{1}{[a_{js}(\alpha, \underline{x}) + k(s,j)\lambda^{\alpha}]^{r+1}}d\alpha,
\end{aligned}
$$

where

$$
a_{js}(\alpha, \underline{x}) = \frac{1}{\alpha} + \sum_{i=1}^{r} x_i^{\alpha} + (s-j-1)x_r^{\alpha}.
$$

Special Cases

(i) For $s = 1$,

$$
\begin{aligned}
Pr(x_{r+1} \geq \lambda \mid \underline{x}) &= (y_1 \geq \lambda \mid \underline{x}) \\
&= \int_{\lambda}^{\infty} f^*(y_1 \mid \underline{x}) dy_1 \\
&= \frac{1}{I_0(\underline{x})}\int_{0}^{\infty}\frac{\alpha^{r-1}e^{-\alpha/\theta}\prod_{i=1}^{r}x_i^{\alpha}}{[a_{01}(\alpha, \underline{x}) + (n-r)\lambda^{\alpha}]^{r+1}}d\alpha \\
&= \frac{W(r,n, \underline{x})}{I_0(\underline{x})}, \text{ say}
\end{aligned}
$$

where

$$
W(r,n, \underline{x}) = \int_{0}^{\infty}\frac{\alpha^{r-1}e^{-\alpha/\theta}\prod_{i=1}^{r}x_i^{\alpha}}{[a_{01}(\alpha, \underline{x}) + (n-r)\lambda^{\alpha}]^{r+1}}d\alpha.
$$

(ii) For $s = n - r$,

$$\begin{aligned} Pr(x_n \geq \lambda \mid \underline{x}) &= (y_{n-r} \geq \lambda \mid \underline{x}) \\ &= \int_{\lambda}^{\infty} f^*(y_{n-r} \mid \underline{x}) dy_{n-r} \\ &= D(n-r) \sum_{j=0}^{n-r-1} \frac{\binom{n-r-1}{j}}{I_0(\underline{x})} \int_0^{\infty} \frac{\alpha^{r-1} e^{-\alpha/\theta} \prod_{i=1}^{r} x_i^{\alpha}}{(j+1)[a_{j,n-r}(\alpha, \underline{x}) + (j+1)\lambda^{\alpha}]^{r+1}} d\alpha. \end{aligned}$$

Simulated Illustration

We generated a random sample of size 20 from Weibull (α, β) with (α, β) generated from exponential-exponential prior having prior parameter $\theta = 2$. The ordered observations are 0.1454, 0.1579, 0.1636, 0.1745, 0.2013, 0.2295, 0.2375, 0.25583, 0.2646, 0.2671, 0.2710, 0.2725, 0.2746, 0.2804, 0.2959, 0.3207, 0.3261, 0.3261, 0.3582, 0.3748. Suppose censoring is done at $r = 8$ so that the available ordered failure times are 0.1454, 0.1579, 0.1636, 0.1745, 0.2013, 0.2295, 0.2375, 0.25583.

Then a 95% Bayesian prediction interval for the next failure time viz. x_9 is obtained as (0.257, 0.401), and a 95% Bayesian prediction interval for the time to last failure viz. x_{20} is (0.378, 0.700). We note that the actual failure times lie within these intervals.

3 Bayesian Multi-Sample Prediction

Let $x_1, x_2, \ldots, x_{r_0}$ be the ordered failure times of the first r_0 items failing in an initial sample of size n_0 taken from Weibull (α, β) distribution, given by (1). Let subsequent samples of sizes $n_1, n_2, \ldots, n_k$ respectively be chosen. We shall use the Bayesian approach to predict the first failure time in a sample based on the earlier samples.

Let $x_{j(1)}$ denote the smallest order statistic in the jth sample and, for simplicity sake, let us write

$$y_j = x_{j(1)}, j = 1, 2, \ldots, k.$$

The p.d.f. of y_j is given by

$$h_j(y_j \mid \alpha, \beta) = n_j \alpha \beta y_j^{\alpha-1} e^{-n_j \beta y_j^{\alpha}}, y_j \geq 0, j = 1, 2, \ldots, k.$$

Based on available data $\underline{x} = (x_1, x_2, \ldots, x_{r_0})'$ from initial sample of size n_0, the posterior density of (α, β) is given by

$$q_0(\alpha, \beta \mid \underline{x}) = B_0 \alpha^{r_0-1} \beta^{r_0} e^{-\left(\frac{\alpha}{\theta} + \frac{\beta}{\alpha}\right)} e^{-\beta \sum_{i=1}^{r_0} x_i^{\alpha} - (n_0 - r_0)\beta x_{r_0}^{\alpha}} \prod_{i=1}^{r_0} x_i^{\alpha}, \alpha, \beta > 0,$$

where

$$B_0^{-1} = \Gamma(r_0+1)I_0(\underline{\mathrm{x}}),$$

$$I_0(\underline{\mathrm{x}}) = \int_0^{\infty} \alpha^{r_0-1} e^{-\alpha/\theta} \frac{\prod_{i=1}^{r} x_i^{\alpha}}{\left[\frac{1}{\alpha} + \sum_{i=1}^{r_0} x_i^{\alpha} + (n_0-r_0)x_{r_0}^{\alpha}\right]^{r_0+1}} d\alpha.$$

Therefore, for sample 1, Bayes predictive density of y_1 is

$$\begin{aligned} f_1^*(y_1 \mid \underline{\mathrm{x}}) &= \int_0^{\infty}\int_0^{\infty} h_1(y_1 \mid \alpha,\beta) q_0(\alpha,\beta \mid \underline{\mathrm{x}}) d\beta d\alpha \\ &= B_0 n_1 \Gamma(r_0+2) \int_0^{\infty} \alpha^{r_0} e^{-\alpha/\theta} \frac{y_1^{\alpha-1}\prod_{i=1}^{r_0} x_i^{\alpha}}{\left[\frac{1}{\alpha} + \sum_{i=1}^{r_0} x_i^{\alpha} + (n_0-r_0)x_{r_0}^{\alpha} + n_1 y_1^{\alpha}\right]^{r_0+2}} d\alpha \\ &= B_1^{-1}, \text{ say} \end{aligned}$$

where

$$B_1^{-1} = n_1(r_0+1)\frac{I_1(\underline{\mathrm{x}})}{I_0(\underline{\mathrm{x}})},$$

$$I_1(\underline{\mathrm{x}}) = \int_0^{\infty} \alpha^{r_0} e^{-\alpha/\theta} \frac{y_1^{\alpha-1}\prod_{i=1}^{r_0} x_i^{\alpha}}{\left[\frac{1}{\alpha} + \sum_{i=1}^{r_0} x_i^{\alpha} + (n_0-r_0)x_{r_0}^{\alpha} + n_1 y_1^{\alpha}\right]^{r_0+2}} d\alpha$$

For sample 2, the posterior density of (α, β) given $\underline{\mathrm{x}}$ and y_1 is given by

$$q_1(\alpha,\beta \mid \underline{\mathrm{x}}, y_1) \infty h_1(y_1 \mid \alpha,\beta) q_0(\alpha,\beta \mid \underline{\mathrm{x}}).$$

Therefore,

$$q_1(\alpha,\beta \mid \underline{\mathrm{x}}, y_1) = B_1 h_1(y_1 \mid \alpha,\beta) q_0(\alpha,\beta \mid \underline{\mathrm{x}}).$$

Hence, Bayes predictive density for y_2 is

$$f_2^*(y_2 \mid \underline{x}, y_1)$$

$$= n_2 n_1 B_1 \Gamma(r_0+3) \int_0^\infty \alpha^{r_0+1} e^{-\alpha/\theta} \frac{\left(\prod_{i=1}^{2} y_i^{\alpha-1}\right)\left(\prod_{i=1}^{r_0} x_i^\alpha\right)}{\left[\frac{1}{\alpha}+\sum_{i=1}^{r_0} x_i^\alpha+(n_0-r_0)x_{r_0}^\alpha+\sum_{i=1}^{2} n_i y_1^\alpha\right]^{r_0+3}} d\alpha$$

$$= B_2^{-1}, \text{ say}$$

where

$$B_2^{-1} = n_2(r_0+2)\frac{I_2(\underline{x})}{I_1(\underline{x})}$$

$$I_2(\underline{x}) = \int_0^\infty \alpha^{r_0+1} e^{-\alpha/\theta} \frac{\left(\prod_{i=1}^{2} y_i^{\alpha-1}\right)\left(\prod_{i=1}^{r_0} x_i^\alpha\right)}{\left[\frac{1}{\alpha}+\sum_{i=1}^{r_0} x_i^\alpha+(n_0-r_0)x_{r_0}^\alpha+\sum_{i=1}^{2} n_i y_1^\alpha\right]^{r_0+3}} d\alpha.$$

Proceeding this way we get, in general, Bayes predictive p.d.f. of y_i as

$$f_j^*(y_j \mid \underline{x}, y_1, y_2, \ldots, y_{j-1}) = n_j(r_0+j)\frac{I_j(\underline{x})}{I_{j-1}(\underline{x})},$$

where

$$I_j(\underline{x}) = \int_0^\infty \alpha^{r_0+j-1} e^{-\alpha/\theta} \frac{\left(\prod_{i=1}^{j} y_i^{\alpha-1}\right)\left(\prod_{i=1}^{r_0} x_i^\alpha\right)}{\left[\frac{1}{\alpha}+\sum_{i=1}^{r_0} x_i^\alpha+(n_0-r_0)x_{r_0}^\alpha+\sum_{i=1}^{j} n_i y_1^\alpha\right]^{r_0+j+1}} d\alpha, \quad j = 1, 2, \ldots, k.$$

Hence, for any $\lambda \geq 0$,

$$\begin{aligned} Pr(y_j \mid \underline{x}, y_1, y_2, \ldots, y_{j1}) &= \int_\lambda^\infty f_j^*(y_j \mid \underline{x}, y_1, y_2, \ldots, y_{j-1}) dy_j \\ &= \frac{T_{j-1}(\underline{x})}{I_{j-1}(\underline{x})}, \end{aligned}$$

where

$$T_{j-1}(\underline{x}) = \int_0^\infty \alpha^{r_0+j-1} e^{-\alpha/\theta} \frac{\left(\prod_{i=1}^{j-1} y_i^{\alpha-1}\right)\left(\prod_{i=1}^{r_0} x_i^\alpha\right)}{\left[\frac{1}{\alpha}+\sum_{i=1}^{r_0} x_i^\alpha+(n_0-r_0)x_{r_0}^\alpha+\sum_{i=1}^{j-1} n_i y_1^\alpha+n_j\lambda^\alpha\right]^{r_0+j}} d\alpha, \quad j = 1, 2, \ldots, k.$$

Clearly, $\quad T_{j-1}(\underline{x}) = I_{j-1}(\underline{x})$ at $\lambda = 0$.

To obtain 100% Bayesian prediction interval for the first failure time in the jth sample we therefore solve for λ in the equations

$$\frac{T_{j-1}(\underline{x})}{I_{j-1}(\underline{x})} = \frac{1+\delta}{2} \text{ and } \frac{T_{j-1}(\underline{x})}{I_{j-1}(\underline{x})} = \frac{1-\delta}{2}.$$

It is noteworthy that based on the available information, prediction is done at any stage by just knowing the sample size at that stage.

Simulated Illustration

We generate 8 independent samples each of size 20 from Weibull (α, β) distribution, with (α, β) generated from the exponential-exponential prior distribution having $\theta = 2$. For the first sample, which is marked as the initial sample, suppose censoring is done at $r_0 = 8$. The following table shows the first r_0 ordered observations of the initial sample and the first order statistic in each of the subsequent 7 samples :

Sample (0)	0.2029	0.2069	0.2080	0.2139	0.2208	0.2211	0.2251	0.2301
Sample (1)	0.1881							
Sample (2)	0.1810							
Sample (3)	0.1838							
Sample (4)	0.1968							
Sample (5)	0.1666							
Sample (6)	0.1941							
Sample (7)	0.2080							

The 95% Bayesian prediction bounds for $y_j = x_{j(1)}, j = 1(1)8$, are shown in Table 1 below :

Table 1

j	Lower Bound	Upper Bound
1	0.153	0.361
2	0.183	0.355
3	0.132	0.365
4	0.173	0.402
5	0.162	0.321
6	0.172	0.421
7	0.210	0.382
8	0.165	0.362

Acknowledgement

The authors thank the anonymous referee for fruitful suggestions, which immensely helped to improve the presentation of the paper.

References

Abd Ellah, A. H. (2003) : Bayesian one sample prediction bounds for the Lomax distribution, *Indian Journal of Pure & Applied Maths.,* **34(1)**, pp. 101-110.

Aitchison, J. (1964) : Bayesian tolerance regions, *Journal of Royal Statistical Society, B,* **26**, pp. 161-175.

Aitchison, J. and Sculthorpe, D. (1965) : Some problems of statistical prediction, *Biometrika,* **55**, pp. 469-483.

Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (1992) : *A First Course in Order Statistics.* John Wiley & Sons, New York.

Bernardo, J. M., Berger, J. O., Dawid, A. P. and Smith, A. F. M. (1998) : *Bayesian Statistics.* Oxford University Press.

Khan, S. and Chattopadhyay, A. K. (2003) : Predictive analysis of occupational mobility based on number of jobs offered, *Journal of Applied Statistical Sciences,* **12(1),** pp. 11-22.

Khan, S. (2004) : Predictive distribution of regression vector and residual sum of squares for normal multiple regression model, *Communications in Statistics : Theory and Methods,* **33(10)**, pp. 2423-2443.

Khan, S. (2006) : Prediction distribution of future regression and residual sum of squares matrices for multivariate simple regression model with correlated normal responses, *Journal of Applied Probability and Statistics,* **1,** pp. 15-30.

Lingappaiah, G. S. (1986) : Bayes prediction in exponential life-testing when sample size in random variable. *IEEE Transactions on Reliability,* **35(1),** pp. 106-110.