嘉南藥理科技大學專題研究計畫成果報告

計畫編號：CNMI95-02

計畫名稱：在對數常態(Lognormal)分配下於移除的設限資料上預估信賴區間

執行期間：95 年 1 月 1 日至 95 年 12 月 31 日

□整合型計畫　　□個別型計畫

計畫總主持人：　　計畫主持人：薛雅明

子計畫主持人：

An Alternative Method for Prediction Intervals of an Ordered Observation from Weibull Distribution Based on Censored Sample

Hsiaoting Chiu, Hai-Lin Lu
Department of Information Management
Chia-Nan University
886-6-3663994
htchiu@mail.chna.edu.tw

Tsong-Huey Wu
Department of Health Care and Administration
Chia-Nan University
886-6-3660118
thwu@mail.chna.edu.tw

Hai-Wen Lu
Department of E-commerce
Nanhua University
886-6-3663994
hwlu@mail.nhu.edu.tw

ABSTRACT
This paper provides some suitable pivotal quantities for constructing the prediction intervals of the *j*th future order observation from the two-parameter weibull distribution based on censored samples. The method employed is more general in the sense that it can be applied to any data scheme. The precisions of the generated intervals are compared via simulations. Finally two illustrative examples are included.

Categories
G.3 [PROBABILITY AND STATISTICS]: Probabilistic algorithms---Reliability and life testing, Random number generation

General Terms
Reliability

Keywords
Prediction Intervals, Monte Carlo Simulation, Order Statistics

1. INTRODUCTION
In most researching of reliability, the Weibull distribution is widely used as a model of lifetime data (Bain and Engelhardt [2], Agresti [1]). Let us consider the two-parameter Weibull distribution with probability density function (pdf)

$$w(t) = \frac{\beta}{\alpha}(\frac{t}{\alpha})^{\beta-1}\exp\{-(\frac{t}{\alpha})^{\beta}\}, \quad t > 0, \tag{1}$$

and cumulative distribution function(cdf)

$$W(t) = 1 - \exp\{-(\frac{t}{\alpha})^{\beta}\}, \tag{2}$$

where $\beta > 0$ and $\alpha > 0$ are the shape and scale parameters, respectively. Logistic distribution is similar to Weibull distribution. For convenience , we adopt Weibull distribution to explain the process. It is worth noting that if T is a random variable having the Weibull cdf given by formula (2), then the random variable $X = lnT$ is distributed as a smallest Type I extreme value variate with pdf

$$f(x) = \frac{1}{\sigma}\exp(\frac{x-\mu}{\sigma})\exp(-e^{\frac{x-\mu}{\sigma}}) \quad -\infty < \mu < \infty, \sigma > 0, \tag{3}$$

where $\mu = \ln\beta$ and $\sigma = 1/\beta$. Its cdf has the form

$$F(x) = 1 - \exp(-e^{\frac{x-\mu}{\sigma}}). \tag{4}$$

In life testing studies, several lifetimes of units put on test may not be observed due to time limitations or money and material resources restrictions on data collection. Consider an experiment in which n identical components are placed on test simultaneously. Suppose the experiment was terminated when the (n-s)th component failed, thus censoring the last s components. Such a sample is called Type II right censored sample. If some initial r observations are also censored, it is called Type II doubly censored.

The studies of estimating the prediction intervals of the future data are quite important and valuable in lifetime analysis. There have been several studies in the literature dealing with such problems. For the exponential distribution, Lawless[9] and Likes[12] estimated the prediction intervals based on the order statistics, $X_{(j)}\,(r < j \le n)$, of a sample while the first r data of the sample were observed. Mann and Grubbs[15] proposed an alternative method to construct approximate prediction intervals. Kaminsky and Nelson[8] constructed prediction intervals by using the best linear unbiased estimates (BLUE) of the parameters as a pivotal statistic. For the Weibull distribution, Mann and Saunders[16] used three specially selected order statistics to predict the minimum of a single future sample. Engelhardt and Bain[5] constructed the prediction limits for the *j*th smallest of some set of future observations. Fertig *et al.*[6] provided Monte Carlo estimates of percentiles of the distribution of a statistics S for constructing prediction intervals of a future observation. Lawless[10] used a conditional method to obtain a prediction interval for the first order statistic of a set of future observations, based on previous data; Hsieh[7] used the same technique to construct prediction intervals for future observations. Mann and Fertig[14] constructed the tables for obtaining the best linear invariant estimates (BLIE) of parameters. Balakrishnan and Cohen[3] proposed an approximate maximum likelihood estimates (AMLE) of parameters. All these researches are under the scheme that the available data is either right censored or doubly censored.

It is well known that the Type II censored data, the right, left and doubly censored data are all special cases of multiple censored data. In this paper, we consider the general case of

the multiple Type II censored data scheme. Suppose n components are placed on test in life testing. The lifetime of the first r, the middle l, and the last s components are assumed unobserved or missing. That is, we assume $X_{(r+1)} < X_{(r+2)} <\ldots< X_{(r+k)}$ and $X_{(r+k+l+1)} < X_{(r+k+l+2)} <\ldots< X_{(n-s)}$ are observable and no others. In practice, multiple Type II censored problems may arise when some components failed between two points of observation with exact times of these failure unobservable components (Balasubramanian and Balakrishnan [4]).

In next section, following the ideas of Wu *et al*. [18], we present our method of constructing the prediction intervals of the future unknown observations for Type II censored data. We describe the procedure for calculating the percentiles of the distributions of the pivotal quantities, and the simulation results are compared with the existing method in section 3 and 4, respectively. In section 5, we illustrate our method with two examples. A brief discussion is presented in section 6.

2. A GENERAL FORM OF PIVOTAL QUANTITY

The prediction intervals of our method for $X_{(j)}$ are based on a subset $\{X_{(n_i)}\}_{i=1}^{c}$ of $\{X_{(k)}\}_{k=1}^{d}$,where $1 \le n_1 < n_2 <\ldots< n_c \le d < j \le n$. Let $Y_i = \frac{X_i - \mu}{\sigma}$, then Y_i has extreme value distribution with μ=0 and σ=1. And $Y_{(i)} = \frac{X_{(i)} - \mu}{\sigma}$ is the ith order statistic of Y_i. We define some pivotal quantities (proof shown in Appendix I) of the following general forms,

$$\widetilde{U}_h = \frac{X_{(j)} - X_{(n_c)}}{\widetilde{W}_h}, \quad h = 1,\ldots,4, \quad n - s < j \le n,$$

where

$$\widetilde{W}_1 = \sum_{i=2}^{c} \frac{g(E(Y_{(n_i)}))}{\sum_{t=2}^{c} g(E(Y_{(n_t)}))}(X_{(n_i)} - X_{(n_1)}), \tag{5}$$

$$\begin{aligned}\widetilde{W}_2 &= \sum_{i=1}^{n_1-1} \frac{g(E(Y_{(i)}))}{\sum_{t=1}^{n} g(E(Y_{(t)}))}(X_{(n_2)} - X_{(n_1)}) \\ &+ \sum_{i=2}^{c} \frac{g(E(Y_{(i)}))}{\sum_{t=1}^{n} g(E(Y_{(t)}))}(X_{(n_i)} - X_{(n_1)}) \\ &+ \sum_{i=1}^{c-1} \sum_{j=n_i+1}^{n_{i+1}-1} \frac{g(E(Y_{(i)}))}{\sum_{t=1}^{n} g(E(Y_{(t)}))} \times \left(\frac{X_{(n_i)} + X_{(n_{i+1})} - 2X_{(n_1)}}{2} \right) \\ &+ \sum_{i=n_c+1}^{n} \frac{g(E(Y_{(i)}))}{\sum_{t=1}^{n} g(E(Y_{(t)}))}(X_{(n_c)} - X_{(n_1)}),\end{aligned} \tag{6}$$

$$\widetilde{W}_3 = \prod_{i=2}^{c} (X_{(n_i)} - X_{(n_1)})^{\frac{g(E(Y_{(n_i)}))}{\sum_{i=2}^{c} g(E(Y_{(n_t)}))}}, \tag{7}$$

$$\begin{aligned}\widetilde{W}_4 &= \left(\prod_{i=1}^{n_1-1} (X_{(n_2)} - X_{(n_1)})^{\frac{g(E(Y_{(n_i)}))}{\sum_{t=1}^{n} g(E(Y_{(t)}))}} \right) \\ &\times \left(\prod_{i=2}^{c} (X_{(n_i)} - X_{(n_1)})^{\frac{g(E(Y_{(n_i)}))}{\sum_{t=1}^{n} g(E(Y_{(n_t)}))}} \right) \\ &\times \prod_{i=1}^{c-1} \prod_{i=n_i+1}^{n_{i+1}-1} \left(\frac{X_{(n_i)} + X_{(n_{i+1})} - 2X_{(n_1)}}{2} \right)^{\frac{g(E(Y_{(j)}))}{\sum_{t=1}^{n} g(E(Y_{(t)}))}} \\ &\times \left(\prod_{i=n_c+1}^{n} (X_{(n_c)} - X_{(n_1)})^{\frac{g(E(Y_{(i)}))}{\sum_{t=1}^{n} g(E(Y_{(t)}))}} \right)\end{aligned} \tag{8}$$

$\widetilde{W}_1$ and $\widetilde{W}_2$ come from the ideas of arithmetic means, and $\widetilde{W}_3$ and $\widetilde{W}_4$ follow the concepts of geometric means. The equations (5) to (8) are general forms. Wu *et al.* [18] consider the case where every datum of different position has the same weight. In the case of extreme value distribution, it seems reasonable to assume that the weight of each datum point should be different for different position. From the properties of the extreme value distribution, we suggest that the weighted factors are equal to $g(E(Y_{(n_i)})) / \sum_{t=1}^{n} g(E(Y_{(n_t)}))$ in $\widetilde{W}_1$, where $E(Y_{(n_i)})$ is the expected values of $Y_{(n_i)}$. The $E(Y_{(n_i)})$ is defined as $\int_{-\infty}^{\infty} Y_{(n_i)} q(Y_{(n_i)}) dY_{(n_i)}$, where $q(Y_{(n_i)})$ is the pdf of the n_ith order statistic of $Y_{(n_i)}$. Since the parameters μ and σ in (5) will be cancelled (see the proof in Appendix I), without loss of generosity, we simply treat them as standard extreme value distribution. Therefore, $E(Y_{(n_i)})$ will be constants and set $g(z) = e^{z} e^{-e^{z}}$, $z = E(Y_{(n_i)})$. It also showed that the weighted factors do not depend on the parameters μ and σ.

According to the general data scheme mentioned in section 1, the pivotal quantities of (5) to (8) are transformed to the following forms

$$\hat{U}_h = \frac{X_{(j)} - X_{(n-s)}}{\hat{W}_h}, \quad h = 1,\ldots,4, \quad n - s < j \le n,$$

where

$$\begin{aligned}\hat{W}_1 &= \sum_{i=r+2}^{r+k} \frac{g(E(Y_{(i)}))}{S_c}(X_{(i)} - X_{(r+1)}) \\ &+ \sum_{i=r+k+l+1}^{r+k+l+m} \frac{g(E(Y_{(i)}))}{S_c}(X_{(i)} - X_{(r+1)}),\end{aligned} \tag{9}$$

$$\hat{W}_2 = \sum_{i=1}^{r} \frac{g(E(Y_{(i)}))}{S_n}(X_{(r+2)} - X_{(r+1)})$$
$$+ \sum_{i=r+2}^{r+k} \frac{g(E(Y_{(i)}))}{S_n}(X_{(i)} - X_{(r+1)})$$
$$+ \sum_{i=r+k+1}^{r+k+l} \frac{g(E(Y_{(i)}))}{S_n}\left(\frac{X_{(r+k)} + X_{(r+k+l+1)} - 2X_{(r+1)}}{2}\right)$$
$$+ \sum_{i=r+k+l+1}^{n-s} \frac{g(E(Y_{(i)}))}{S_n}(X_{(i)} - X_{(r+1)})$$
$$+ \sum_{i=n-s+1}^{n} \frac{g(E(Y_{(i)}))}{S_n}(X_{(n-s)} - X_{(r+1)}), \tag{10}$$

$$\hat{W}_3 = \prod_{i=r+2}^{r+k}(X_{(i)} - X_{(r+1)})^{\frac{g(E(Y_{(i)}))}{S_c}}$$
$$\times \prod_{i=r+k+l+1}^{r+k+l+m}(X_{(i)} - X_{(r+1)})^{\frac{g(E(Y_{(i)}))}{S_c}}, \tag{11}$$

$$\hat{W}_4 = \prod_{i=1}^{r}(X_{(r+2)} - X_{(r+1)})^{\frac{g(E(Y_{(i)}))}{S_n}}$$
$$\times \prod_{i=r+2}^{r+k}(X_{(i)} - X_{(r+1)})^{\frac{g(E(Y_{(i)}))}{S_n}}$$
$$\times \prod_{i=r+k+1}^{r+k+l}\left(\frac{X_{(r+k)} + X_{(r+k+l+1)} - 2X_{(r+1)}}{2}\right)^{\frac{g(E(Y_{(i)}))}{S_n}}$$
$$\times \prod_{i=r+k+l+1}^{n-s}(X_{(i)} - X_{(r+1)})^{\frac{g(E(Y_{(i)}))}{S_n}}$$
$$\times \prod_{i=n-s+1}^{n}(X_{(n-s)} - X_{(r+1)})^{\frac{g(E(Y_{(i)}))}{S_n}}, \tag{12}$$

where

$$S_c = \sum_{t=r+2}^{r+k} g(E(Y_{(t)})) + \sum_{t=r+k+l+1}^{r+k+l+m} g(E(Y_{(t)})),$$
$$S_n = \sum_{t=1}^{n} g(E(Y_{(t)})).$$

For comparison, the other pivotal quantity is $\hat{U}_a$. Let

$$\hat{U}_a = \frac{X_{(j)} - X_{(n-s)}}{\hat{W}_a}, \qquad n - s < j \le n, \qquad \text{where}$$
$$\hat{W}_a = \hat{\sigma}. \tag{13}$$

The $\hat{\sigma}$ is the AMLE of σ, which can be obtained from Balakrishnan and Cohen[3]. The percentiles data of $\hat{U}_a$ distribution are listed in Table 9 (see Appendix II).

From equations (9) to (12), the distributions of $\hat{U}_h$ depend only on *n, r, k, l, m, s, j*, but not on μ and σ. Then, we have

$$1-\alpha = P\{0 < \frac{X_{(j)} - X_{(n-s)}}{\hat{W}_h} < \hat{u}_h(1-\alpha; n, r, k, l, m, s, j)\}$$
$$= P\{X_{(n-s)} < X_{(j)} < X_{(n-s)} + \hat{u}_h(1-\alpha; n, r, k, l, m, s, j) \times \hat{W}_h\}.$$

Table 1. The properties of parameters $\tilde{\mu}$ and $\tilde{\sigma}$ of 60,000 random samples for each combination case.

n	$\tilde{\mu}_{mean} \pm \tilde{\mu}_{sd}$	$\tilde{\sigma}_{mean} \pm \tilde{\sigma}_{sd}$
13	-0.009147 ± 0.353776	0.921192 ± 0.344354
25	-0.001123 ± 0.228614	1.000021 ± 0.201247
30	0.000347 ± 0.208632	0.999757 ± 0.183330
35	-0.001845 ± 0.191873	1.000458 ± 0.169762
40	-0.001771 ± 0.172441	0.999497 ± 0.146308

Therefore, $(X_{(n-s)}, X_{(n-s)} + \hat{u}_h(1-\alpha; n, r, k, l, m, s, j) \times \hat{W}_h)$, h = 1,…,4 are one-sided 100(1-α)% prediction intervals of $X_{(j)}$ based on *m*+*k* observations.

3. CALCULATION AND ALGORITHM

The exact distributions of the pivotal quantities $\hat{U}_h$ (*h*=1,…, 4) can not be derived algebraically, but we can approximate the distributions of $\hat{U}_h$ (*h*=1,…, 4) by using large quantities of the Monte Carlo sampling with some programming algorithms to generate the percentiles of $\hat{U}_h$. All the simulations were run with the aid Microsoft Quick Basic 4.5 program and Foxbase database software package. The procedures for generating the percentiles of $\hat{U}_h$ are as follows:

a. Give and set μ=0, σ=1. (For providing the properties of parameters $\tilde{\mu}$ and $\tilde{\sigma}$ of the random samples generated by computer, 60,000 Monte Carlo runs are done for each combination of *n*, *r*, *k*, *l*, *m*, *s*, *j* (some selected cases). The results are presented in Table I. Using Table 5.3 in Mann *et al.*[17] for case of *n*=13 and Table 1 in Mann *et al.*[14] for remaining cases to obtain the necessary weights, we can calculate their BLIE's of μ and σ respectively. The $\tilde{\mu}_{mean}$ and $\tilde{\sigma}_{mean}$ of those random samples are very close to 0 and 1, respectively.)

b. Calculate the following statistics: $\hat{U}_1$ in (9), $\hat{U}_2$ in (10), $\hat{U}_3$ in (11), $\hat{U}_4$ in (12).

c. In the Step a and Step b, 600,000 replicates are used to compute the percentiles of $\hat{U}_h$ (*h*=1,…, 4) for each combination of *n*, *r*, *k*, *l*, *m*, *s*, *j*.

d. Sort 600,000 results of each combination of *n*, *r*, *k*, *l*, *m*, *s*, *j* in ascending order.

e. Retrieve the value of $\hat{U}_h$ (*h*=1,…, 4) under different significance levels of α.

From the above procedures, we obtain the values of $\hat{U}_h$ (*h*=1,…, 4) according to the exact position of $\hat{U}_h$ (*h*=1,…, 4) in Step d.

In our simulation, 600,000 replicates are done for each combination of *n*, *r*, *k*, *l*, *m*, *s*, *j*. To save space, we only list part of the percentiles of $\hat{U}_h$ (*h*=1,…, 4) in Table 5 to 8 (see Appendix II).

4. COMPARISON

In this section, we compare the performance of our method with $\hat{U}_a$. We calculate their average lengths of 95% prediction intervals, and coverage probabilities for some selected combinations of n, r, k, l, m, s, j. Referring to the data scheme mentioned in section 1, the simulation is computed by the following procedures:

a. Give and set $\mu=0$, $\sigma=1$.

b. Generate n (n=10,13,20,40) random samples from the standard extreme value distribution.

c. Calculate the values of $\hat{W}_h$ (h=1,…, 4,a), and then make a multiply of $\hat{W}_h$ by $\hat{U}_h$ (h=1,…, 4, a) (from Table 5 to 9) for each combination of n, r, k, l, m, s, j.

d. Repeat steps b to c, execute 10,000 runs and record all upper bounds of the confidence intervals of $X_{(n-s+1)}$ and $X_{(n-s+2)}$.

e. From the results in steps c and d, calculate the average length of the 10,000 confidence intervals, and coverage probabilities for all methods

The results of simulation are listed in Table 2. It is clear that the 95% estimated expected lengths of $\hat{U}_2$ or $\hat{U}_a$ are shorter. The difference among $\hat{U}_1$, $\hat{U}_2$, and $\hat{U}_a$ is not significant (about 0% to 3%). It is also shown that the confidence intervals of $\hat{U}_h$ (h=1,…, 4, a) have almost 95% coverage probabilities. It is interesting to note that if the sample size n is larger, then the difference of average lengths among $\hat{U}_h$ (h=1,…, 4, a) will be smaller. And it also showed that simulation has the property of convergence.

Table 2. The average length of 95% prediction intervals , and coverage probabilities for $X_{(j)}$ by difference statistics: $\mu=0$, $\sigma=1$.

n	$\hat{U}_1$	$\hat{U}_2$	$\hat{U}_3$	$\hat{U}_4$	$\hat{U}_a$
10	3.034146 (95.23%)	2.834668 (95.11%)	3.553925 (95.22%)	2.868328 (95.18%)	2.729171 (95.01%)
10	1.26180 (95.04%)	1.208341 (94.97%)	1.462240 (95.34%)	1.341985 (95.18%)	1.174512 (94.92%)
10	1.156878 (94.40%)	1.123969 (94.52%)	1.335957 (94.58%)	1.196265 (94.46%)	1.094455 (94.40%)
13	0.868976 (95.04%)	0.846851 (95.01%)	0.917785 (94.93%)	0.885132 (94.99%)	0.762706 (95.15%)
13	0.882302 (94.69%)	0.852831 (94.86%)	0.999311 (94.37%)	0.945090 (94.53%)	0.831682 (94.85%)
13	0.853402 (94.49%)	0.808929 (94.56%)	0.982991 (94.50%)	0.869487 (94.49%)	0..788812 (94.58%)
20	0.573943 (94.70%)	0.545609 (94.71%)	0.601589 (94.64%)	0.555719 (94.66%)	0.500980 (94.87%)
20	0.570727 (95.14%)	0.544830 (95.08%)	0.638461 (95.00%)	0.556134 (95.05%)	0.590064 (95.22%)
40	0.248098 (94.79%)	0.244359 (94.82%)	0.251093 (94.76%)	0.247359 (94.78%)	0.224264 (94.68%)
40	0.236455 (94.99%)	0.233026 (94.99%)	0.244821 (94.86%)	0.250343 (94.79%)	0.230777 (95.04%)
40	0.336696 (94.96%)	0.335886 (94.94%)	0.349590 (94.95%)	0.356296 (94.90%)	0.330274 (94.95%)

5. EXAMPLES

5.1 Example 1

Consider the following 13 components were placed on test, and the test was terminated at the time of the 10th failure (Mann and Fertig[13]). The first 10 observations are given below:

0.22, 0.50, 0.88, 1.00, 1.32, 1.33, 1.54, 1.76, 2.50, 3.00.

It is assumed that the 10 observed data are from the same Weibull distribution. We transform the data to extreme value form: the logs of the 10 observations are

-1.541, -0.693, -0.128, 0.000, 0.278, 0.285, 0.432, 0.565, 0.916, 1.099.

Table 3. The one-sided 90% prediction intervals, the percentiles of $\hat{U}_k$ (*k*=1,…,4, *a*), and CMd using MLE and BLIE (Hsieh[7]).

	$\hat{U}_k(11)$	$\hat{U}_k(12)$	$X_{(11)}$	$X_{(12)}$
$\hat{U}_1$	0.243	0.456	(1.099,1.579)	(1.099, 1. 998)
$\hat{U}_2$	0.224	0.418	(1.099,1.573)	(1.099, 1.987)
$\hat{U}_3$	0.268	0.507	(1.099,1.612)	(1.099, 2.068)
$\hat{U}_4$	0.244	0.458	(1.099,1.600)	(1.099, 2.038)
$\hat{U}_a$	0.616	1.134	(1.099,1.535)	(1.099, 1.902)
(C, M)	-	-	(1.099,5.263)	(1.099, 8.002)
(C, BLIE)	-	-	(1.099,5.276)	(1.099, 8.022)

Note: CMd=Conditional Method, M=MLE

Table 4. The one-sided 95% prediction intervals and the percentiles of $\hat{U}_k$ (k=1,…,4, *a*)

	$\hat{U}_k(11)$	$\hat{U}_k(12)$	$X_{(11)}$	$X_{(12)}$
$\hat{U}_1$	0.148	0.240	(-0.45, 0.122)	(-0.45, 0.490)
$\hat{U}_2$	0.152	0.245	(-0.45, 0.099)	(-0.45, 0.444)
$\hat{U}_3$	0.162	0.264	(-0.45, 0.227)	(-0.45, 0.690)
$\hat{U}_4$	0.165	0.268	(-0.45, 0.126)	(-0.45, 0.496)
$\hat{U}_a$	0.542	0.880	(-0.45, 0.135)	(-0.45, 0.514)

In this case, we have n = 13, r = 0, k = 10, l = 0, m = 0, and s = 3. Applying our method to estimate the one-sided 90% prediction intervals of $X_{(11)}$ and $X_{(12)}$. The results are presented in Table III. It is clear that the shorter prediction intervals are obtained by the pivotal quantities $\hat{U}_2$ and $\hat{U}_a$.

5.2 Example 2

The following are 10 observations data from Lawless[11].

-3.57, -2.55, -2.02, -1.66, -1.36, -1.15, -0.95, -0.77, -0.61, -0.45.

It assumed that above data were obtained from a sample of 20, which are distributed according to extreme value distribution, and the last 50% data were censored. And only from the *3*rd to the 6th failure times and from the 9th to the 10th failure times are available. In other words, this is the case of n = 20, r = 2, k = 4, l = 2, m = 2, and s = 10, The one-sided 95% prediction intervals of $X_{(11)}$ and $X_{(12)}$ are listed in Table 4. It is obvious that the pivotal quantity $\hat{U}_2$ has the shortest prediction intervals.

6. DISCUSSION

From Table 2, the average length of prediction intervals of $\hat{U}_1$, $\hat{U}_2$, and $\hat{U}_a$ are shorter than $\hat{U}_3$ and $\hat{U}_4$. Since $\hat{U}_3$ and $\hat{U}_4$ are longer than $\hat{U}_2$, $\hat{U}_1$ and $\hat{U}_2$ are preferred to both of them. The average lengths of prediction intervals of $\hat{U}_3$ and $\hat{U}_4$ are longer than $\hat{U}_1$ and $\hat{U}_2$. It may be the reason that the power operations in geometric means will cause the results extended unexpectedly. Intuitively, our method produces good result because we have given different weight to each datum point. Following the algorithm of section 3, it is straightforward to construct prediction intervals for the future failure time by the pivotal quantities $\hat{U}_1$ and $\hat{U}_2$. Note also that $\hat{U}_1$ and $\hat{U}_2$ can be applied to any kind of data scheme.

Comparing with the existing methods, it is true that calculation procedures of $\hat{U}_a$ are simpler than $\hat{U}_1$ and $\hat{U}_2$. But since the computations of $\hat{U}_1$ and $\hat{U}_2$ can be easily done by computers, it seems to be not an important consideration. Furthermore, following the some algorithm, it is not difficult for generating and simulating larger sample size *n*. Thus it makes this method to be potentially more useful than the existing ones. For further study, this simulation scheme can be easily applied to other family of location and scale distributions.

7. APPENDICES

7.1 Appendix I

Theorem: If $\widetilde{W}_1$ in (5) is an estimator of μ and σ based on multiple type II censored sample $X_{(n_0)} \le X_{(n_1)} \le \cdots \le X_{(n)}$ from two-parameter Weibull distribution, then $\widetilde{U}_i = \frac{X_{(j)} - X_{(n_c)}}{\widetilde{W}_1}$ is a pivotal (parameter-free) quantity

Proof:

Define $Y_i = \frac{X_i - \mu}{\sigma}$, then Y_i does not depend on μ and σ. And $Y_{(i)} = \frac{X_{(i)} - \mu}{\sigma}$ is the *i*th order statistics of Y_i.

Let $Z_i = Y_{(j)} - Y_{(n_c)}$, $Z_2 = \sum_{i=2}^{c} \frac{g(E(Y_{(n_i)}))}{\sum_{t=2}^{c} g(E(Y_{(n_t)}))}(Y_{(n_i)} - Y_{(n_1)})$,

where

$\sum_{i=2}^{c} \frac{g(E(Y_{(n_i)}))}{\sum_{t=2}^{c} g(E(Y_{(n_t)}))}$ are constant.

Define $Z_3 = Z_1 / Z_2$. Both Z_1 and Z_2 are pivotal; therefore, Z_3 is also pivotal. It follows that

$$\widetilde{U}_1 = \frac{\frac{X_{(j)} - \mu}{\sigma} - \frac{X_{(n_c)} - \mu}{\sigma}}{\sum_{i=2}^{c} \frac{g(E(Y_{(n_i)}))}{\sum_{t=2}^{c} g(E(Y_{(n_t)}))}\left(\frac{X_{(n_i)} - \mu}{\sigma} - \frac{X_{(n_1)} - \mu}{\sigma}\right)}$$

$$= \frac{Y_{(j)} - Y_{(n_c)}}{\sum_{i=2}^{c} \frac{g(E(Y_{(n_i)}))}{\sum_{t=2}^{c} g(E(Y_{(n_t)}))}(Y_{(n_i)} - Y_{(n_1)})} = \frac{Z_1}{Z_2} = Z_3.$$

Similarly, it is easy to show that the estimators $\widetilde{U}_2$, $\widetilde{U}_3$, and $\widetilde{U}_4$ are also pivotal quantities.

7.2 Appendix II

Table 5. The percentiles data of $\hat{U}_1$ distribution for $X_{(j)}$.

n	*r*	*k*	*l*	*m*	*s*	*j*	0.90	0.95
10	0	3	0	0	7	4	1.355785	2.201880
10	0	3	0	0	7	5	2.333157	3.681949
10	0	8	0	0	2	9	0.368497	0.5.3914
10	0	8	0	0	2	10	0.751464	0.972094
10	2	5	0	0	3	8	0.996771	1.426663
10	2	5	0	0	3	9	1.885445	2.568596
10	3	4	0	0	3	8	0.999649	1.437188
10	3	4	0	0	3	9	1.879692	2.569213
10	1	3	1	2	3	8	1.618781	0.864230
10	1	3	1	2	3	9	1.154031	1.533077
10	1	2	2	1	4	7	0.722657	1.030894
10	1	2	2	1	4	8	1.303315	1.769215
13	0	10	0	0	3	11	0.243662	0.330549
13	0	10	0	0	3	12	0.456420	0.589753
13	0	8	0	0	5	9	0.264716	0.369639
13	0	8	0	0	5	10	0.471422	0.624253
13	3	6	0	0	4	10	0.785583	1.106489
13	3	6	0	0	4	11	1.443556	1.935043
13	4	7	0	0	2	12	0.912021	1.239771
13	4	7	0	0	2	13	1.908133	2.469798
13	2	5	1	1	4	10	0.606973	0.843070
13	2	5	1	1	4	11	1.111241	1.469986
13	3	6	1	1	2	12	0.773738	1.050499
13	3	6	1	1	2	13	1.621465	2.075764
20	0	10	0	0	10	11	0.164343	0.224058
20	0	10	0	0	10	12	0.283892	0.371838
20	0	16	0	0	4	17	0.132946	0.177885
20	0	16	0	0	4	18	0.245659	0.309447
20	2	4	2	2	10	11	0.378516	0.520569
20	2	4	2	2	10	12	0.651078	0.854504
40	0	30	0	0	10	31	0.050056	0.066743
40	0	30	0	0	10	32	0.087866	0.109813
40	10	20	0	0	10	31	0.203558	0.269217
40	10	20	0	0	10	32	0.345067	0.442121
40	8	15	3	10	4	37	0.207640	0.270819
40	8	15	3	10	4	38	0.382493	0.469227

Table 6. The percentiles data of $\hat{U}_2$ distribution for $X_{(j)}$.

n	*r*	*k*	*l*	*m*	*s*	*j*	0.90	0.95
10	0	3	0	0	7	4	1.125945	1.807184
10	0	3	0	0	7	5	1.926988	3.009241
10	0	8	0	0	2	9	0.341619	0.463249
10	0	8	0	0	2	10	0.690662	0.889023
10	2	5	0	0	3	8	0.854753	1.209239
10	2	5	0	0	3	9	1.607269	2.163889
10	3	4	0	0	3	8	0.857231	1.219519
10	3	4	0	0	3	9	1.602540	2.171974
10	1	3	1	2	3	8	0.535858	0.744663
10	1	3	1	2	3	9	0.994515	1.312484
10	1	2	2	1	4	7	0.626326	0.886672
10	1	2	2	1	4	8	1.119622	1.518678
13	0	10	0	0	3	11	0.224299	0.302451
13	0	10	0	0	3	12	0.418464	0.537054
13	0	8	0	0	5	9	0.228817	0.317286
13	0	8	0	0	5	10	0.404464	0.530687
13	3	6	0	0	4	10	0.678069	0.950597
13	3	6	0	0	4	11	1.239673	1.644747
13	4	7	0	0	2	12	0.925312	1.257150
13	4	7	0	0	2	13	1.925821	2.487037
13	2	5	1	1	4	10	0.476881	0.655567
13	2	5	1	1	4	11	0.864392	1.130554
13	3	6	1	1	2	12	0.683757	0.921912
13	3	6	1	1	2	13	1.423511	1.812982
20	0	10	0	0	10	11	0.138502	0.187599
20	0	10	0	0	10	12	0.238964	0.307889
20	0	16	0	0	4	17	0.125104	0.166941
20	0	16	0	0	4	18	0.230551	0.289223
20	2	4	2	2	10	11	0.289980	0.396636
20	2	4	2	2	10	12	0.496145	0.645765
40	0	30	0	0	10	31	0.046921	0.062344
40	0	30	0	0	10	32	0.082149	0.102351
40	10	20	0	0	10	31	0.187509	0.248863
40	10	20	0	0	10	32	0.326648	0.407463
40	8	15	3	10	4	37	0.207977	0.272013
40	8	15	3	10	4	38	0.384794	0.467108

Table 7. The percentiles data of $\hat{U}_3$ distribution for $X_{(j)}$.

n	*r*	*k*	*l*	*m*	*s*	*j*	0.90	0.95
10	0	3	0	0	7	4	1.608883	2.721678
10	0	3	0	0	7	5	2.821951	4.645655
10	0	8	0	0	2	9	0.415438	0.576626
10	0	8	0	0	2	10	0.856595	1.138579
10	2	5	0	0	3	8	1.287732	1.902335
10	2	5	0	0	3	9	2.484045	3.534942
10	3	4	0	0	3	8	1.292677	1.923213
10	3	4	0	0	3	9	2.481441	3.525552
10	1	3	1	2	3	8	0.788329	1.142375
10	1	3	1	2	3	9	1.503843	2.067356
10	1	2	2	1	4	7	1.036246	1.569153
10	1	2	2	1	4	8	1.934792	2.786440
13	0	10	0	0	3	11	0.268639	0.369248
13	0	10	0	0	3	12	0.507211	0.665167
13	0	8	0	0	5	9	0.294301	0.415125
13	0	8	0	0	5	10	0.529261	0.710224
13	3	6	0	0	4	10	1.010156	1.470474
13	3	6	0	0	4	11	1.893738	2.624943
13	4	7	0	0	2	12	1.219110	1.723033
13	4	7	0	0	2	13	2.596685	3.509463
13	2	5	1	1	4	10	0.783438	1.131463
13	2	5	1	1	4	11	1.455241	2.007740
13	3	6	1	1	2	12	1.024048	1.446484
13	3	6	1	1	2	13	2.188790	2.919691
20	0	10	0	0	10	11	0.178347	0.245756
20	0	10	0	0	10	12	0.309257	0.410631
20	0	16	0	0	4	17	0.142240	0.191015
20	0	16	0	0	4	18	0.263603	0.334755
20	2	4	2	2	10	11	0.485984	0.685521
20	2	4	2	2	10	12	0.851258	1.153520
40	0	30	0	0	10	31	0.051966	0.069362
40	0	30	0	0	10	32	0.091289	0.114727
40	10	20	0	0	10	31	0.251929	0.337695
40	10	20	0	0	10	32	0.442873	0.559259
40	8	15	3	10	4	37	0.257641	0.340726
40	8	15	3	10	4	38	0.479742	0.589852

Table 8. The percentiles data of $\hat{U}_4$ distribution for $X_{(j)}$

n	*r*	*k*	*l*	*m*	*s*	*j*	0.90	0.95
10	0	3	0	0	7	4	1.151481	1.851704
10	0	3	0	0	7	5	1.970969	3.083778
10	0	8	0	0	2	9	0.379575	0.521806
10	0	8	0	0	2	10	0.776022	1.019825
10	2	5	0	0	3	8	1.069152	1.543437
10	2	5	0	0	3	9	2.031816	2.827447
10	3	4	0	0	3	8	1.067482	1.554482
10	3	4	0	0	3	9	2.026169	2.824255
10	1	3	1	2	3	8	0.616590	0.866977
10	1	3	1	2	3	9	1.156301	1.545349
10	1	2	2	1	4	7	0.689432	0.985376
10	1	2	2	1	4	8	1.244403	1.689974
13	0	10	0	0	3	11	0.244019	0.332777
13	0	10	0	0	3	12	0.458090	0.594576
13	0	8	0	0	5	9	0.244909	0.340996
13	0	8	0	0	5	10	0.435766	0.575554
13	3	6	0	0	4	10	0.882208	1.260661
13	3	6	0	0	4	11	1.636582	2.243070
13	4	7	0	0	2	12	1.435964	2.086102
13	4	7	0	0	2	13	3.087689	4.285415
13	2	5	1	1	4	10	0.575818	0.805406
13	2	5	1	1	4	11	1.055841	1.408208
13	3	6	1	1	2	12	0.952730	1.339496
13	3	6	1	1	2	13	2.032407	2.687218
20	0	10	0	0	10	11	0.144297	0.195873
20	0	10	0	0	10	12	0.248815	.0322439
20	0	16	0	0	4	17	0.132960	0.178270
20	0	16	0	0	4	18	0.246340	0.310967
20	2	4	2	2	10	11	0.319509	0.437563
20	2	4	2	2	10	12	0.547889	0.717535
40	0	30	0	0	10	31	0.048600	0.064684
40	0	30	0	0	10	32	0.085220	0.106602
40	10	20	0	0	10	31	0.282031	0.379259
40	10	20	0	0	10	32	0.499665	0.634700
40	8	15	3	10	4	37	0.291018	0.385185
40	8	15	3	10	4	38	0.544320	0.673119

Table 9. The percentiles data of $\hat{U}_a$ distribution for $X_{(j)}$

n	*r*	*k*	*l*	*m*	*s*	*j*	0.90	0.95
10	0	3	0	0	7	4	2.429833	3.884042
10	0	3	0	0	7	5	4.138215	6.444420
10	0	8	0	0	2	9	0.878576	1.169411
10	0	8	0	0	2	10	1.732412	2.181317
10	2	5	0	0	3	8	1.008783	1.420359
10	2	5	0	0	3	9	1.885141	2.518690
10	3	4	0	0	3	8	1.013620	1.423626
10	3	4	0	0	3	9	1.875346	2.521363
10	1	3	1	2	3	8	0.857669	1.191790
10	1	3	1	2	3	9	1.591620	2.115796
10	1	2	2	1	4	7	1.169301	1.808686
10	1	2	2	1	4	8	2.153508	3.249308
13	0	10	0	0	3	11	0.616219	0.818953
13	0	10	0	0	3	12	1.134672	1.423598
13	0	8	0	0	5	9	0.617341	0.839394
13	0	8	0	0	5	10	1.072864	1.385328
13	3	6	0	0	4	10	0.704400	0.981970
13	3	6	0	0	4	11	1.278436	1.695298
13	4	7	0	0	2	12	0.806103	1.085826
13	4	7	0	0	2	13	1.665137	2.113019
13	2	5	1	1	4	10	0.898447	0.815059
13	2	5	1	1	4	11	1.075171	1.399081
13	3	6	1	1	2	12	0.697125	0.929402
13	3	6	1	1	2	13	1.430224	1.800205
20	0	10	0	0	10	11	0.403657	0.542627
20	0	10	0	0	10	12	0.686688	0.871494
20	0	16	0	0	4	17	0.391437	0.514785
20	0	16	0	0	4	18	0.710061	0.870636
20	2	4	2	2	10	11	0.457084	0.643474
20	2	4	2	2	10	12	0.792695	1.060140
40	0	30	0	0	10	31	0.175387	0.229385
40	0	30	0	0	10	32	0.301198	0.370072
40	10	20	0	0	10	31	0.182691	0.241560
40	10	20	0	0	10	32	0.317068	0.393065
40	8	15	3	10	4	37	0.242214	0.315173
40	8	15	3	10	4	38	0.446239	0.541538

8. REFERENCES

[1] Agresti, A. *Categorical Data Analysis*. Wiley: New York, 1990.

[2] Bain, L. J. and Engelhardt, M. *Statistical Analysis of Reliability and Life-testing Models*. Marcel Dekker: New York, 1991.

[3] Balakrishnan, N. and Cohen, C. A. *Order statistics and inference*. Academic Press, Inc. 1991.

[4] Balasubramanian, K. and Balakrishnan, N. Estimation for one- and two-parameter exponential distributions under multiple type-II censoring. *Statistische Hefte*1992; **33**:203-216.

[5] Engelhardt, M. and Bain, L. J. Prediction limits and two-sample problems with complete or censored Weibull data. *Technometrics* 1979; **21**:233-237.

[6] Fertig, K. W., Meyer, M. E., and Mann, N. R. On constructing prediction intervals for samples from a Weibull or Extreme Value Distribution. *Technometrics* 1980; **22**:567-573.

[7] Hsieh, H. K. Prediction intervals for Weibull observations, based on early-failure data. *IEEE Trans. Rel.* 1996; **45**:666-670.

[8] Kaminsky, K.S. and Nelson, P. I. Prediction intervals for the exponential distribution using subsets of the data. *Technometrics* 1974; **16**:57-59.

[9] Lawless, J. F. A prediction problem concerning samples from the exponential distribution, with application to life testing. *Technometrics* 1971; 13:725-730.

[10] Lawless, J. F. On the estimation of safe life when the underlying life distribution is Weibull. *Technometrics* 1973; **15**:857-865.

[11] Lawless, J. F. *Statistical models & methods for lifetime data*. John Wiley & Sons: New York, 1982.

[12] Likes J. Prediction of *s*th ordered observation for the two-parameter exponential distribution. *Technometrics*, 1974; **16**:241-244.

[13] Mann, N. R. and Fertig, K. W. Tables for obtaining Weibull confidence bounds and tolerance bounds based on best linear invariant estimates of parameters of the Extreme-Value distribution. *Technometrics* 1973 **15**:345-361

[14] Mann, N. R. and Fertik, K. W. Simplified efficient point and interval estimators for Weibull parameters. *Technometrics* 1975;**17**:361-368.

[15] Mann, N. R. and Grubbs F. E. Chi-square approximation for exponential parameters, prediction intervals and beta percentiles. *Journal of the American Statistical Association* 1974; **69**:654-661.

[16] Mann, N. R. and Saunders, S. C. On evaluation of warranty assurance when life has a Weibull distribution. *Biometrika* 1969; **56**:615-625.

[17] Mann, N. R., Schafer, R. E., and Singpurwalla, N. D. *Methods for statistical analysis of reliability and life data*. John Wiley & Sons: New York, 1974

[18] Wu, J. W., Lu, H. L., Chen, C. H., and Yang, C. H. A note on the prediction intervals for a future ordered observation from a Pareto distribution. *Quality & Quantity* 2004; **38**:217-233