## 嘉南藥理科技大學專題研究計畫成果報告

計畫編號：CNMI95－04

計畫名稱：在極值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
$\square$ 整合型計畫

計畫總主持人：

子計畫主持人：
$\square$ 個別型計畫

計畫主持人：陸海林

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

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

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

Hai-Wen Lu<br>Department of E-commerce<br>Nanhua University<br>886-6-3663994<br>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 Extreme-value 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)
$w(t)=\frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1} \exp \left\{-\left(\frac{t}{\alpha}\right)^{\beta}\right\}, \quad t>0$,
and cumulative distribution function(cdf)
$W(t)=1-\exp \left\{-\left(\frac{t}{\alpha}\right)^{\beta}\right\}$,
where $\beta>0$ and $\alpha>0$ are the shape and scale parameters, respectively. Extreme-value 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 $X=\ln T$ is distributed as a smallest Type I extreme value variate with pdf
$f(x)=\frac{1}{\sigma} \exp \left(\frac{x-\mu}{\sigma}\right) \exp \left(-e^{\frac{x-\mu}{\sigma}}\right) \quad-\infty<\mu<\infty, \sigma>0$,
where $\mu=\ln \beta$ and $\sigma=1 \beta$. Its cdf has the form

$$
\begin{equation*}
F(x)=1-\exp \left(-e^{\frac{x-\mu}{\sigma}}\right) \tag{4}
\end{equation*}
$$

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, $X_{(j)}(r<j \leq n)$, 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 $\left\{X_{\left(n_{i}\right)}\right\}_{i=1}^{c} \quad$ of $\quad\left\{X_{(k)}\right\}_{k=1}^{d} \quad$,where $1 \leq n_{1}<n_{2}<\ldots<n_{c} \leq d<j \leq n$. Let $Y_{i}=\frac{X_{i}-\mu}{\sigma}$, then $Y_{i}$ has extreme value distribution with $\mu=0$ and $\sigma=1$. And $Y_{(i)}=\frac{X_{(i)}-\mu}{\sigma}$ is the $i$ th order statistic of $Y_{i}$. We define some pivotal quantities (proof shown in Appendix I) of the following general forms,
$\widetilde{U}_{h}=\frac{X_{(j)}-X_{\left(n_{c}\right)}}{\widetilde{W}_{h}}, \quad h=1, \ldots, 4, \quad n-s<j \leq n$,
where

$$
\begin{align*}
& \widetilde{W}_{1}=\sum_{i=2}^{c} \frac{g\left(E\left(Y_{\left(n_{i}\right)}\right)\right)}{\sum_{t=2}^{c} g\left(E\left(Y_{\left(n_{t}\right)}\right)\right)}\left(X_{\left(n_{i}\right)}-X_{\left(n_{1}\right)}\right),  \tag{5}\\
& \widetilde{W}_{2}=\sum_{i=1}^{n_{1}-1} \frac{g\left(E\left(Y_{(i)}\right)\right)}{\sum_{t=1}^{n} g\left(E\left(Y_{(t)}\right)\right)}\left(X_{\left(n_{2}\right)}-X_{\left(n_{1}\right)}\right) \\
& +\sum_{i=2}^{c} \frac{g\left(E\left(Y_{(i)}\right)\right)}{\sum_{t=1}^{n} g\left(E\left(Y_{(t)}\right)\right)}\left(X_{\left(n_{i}\right)}-X_{\left(n_{1}\right)}\right) \\
& +\sum_{i=1}^{c-1} \sum_{j=n_{i}+1}^{n_{i+1}-1} \frac{g\left(E\left(Y_{(i)}\right)\right)}{\sum_{t=1}^{n} g\left(E\left(Y_{(t)}\right)\right)} \times\left(\frac{X_{\left(n_{i}\right)}+X_{\left(n_{i+1}\right)}-2 X_{\left(n_{1}\right)}}{2}\right) \\
& +\sum_{i=n_{c}+1}^{n} \frac{g\left(E\left(Y_{(i)}\right)\right)}{\sum_{t=1}^{n} g\left(E\left(Y_{(t)}\right)\right)}\left(X_{\left(n_{c}\right)}-X_{\left(n_{1}\right)}\right) \tag{6}
\end{align*}
$$

$$
\begin{equation*}
\widetilde{W}_{3}=\prod_{i=2}^{c}\left(X_{\left(n_{i}\right)}-X_{\left(n_{1}\right)}\right)^{\frac{g\left(E\left(Y_{\left(n_{i}\right)}\right)\right)}{\Sigma_{i=2}^{c} g\left(E\left(Y_{\left(n_{t}\right)}\right)\right)}} \tag{7}
\end{equation*}
$$

$$
\begin{align*}
& \widetilde{W}_{4}=\left(\prod_{i=1}^{n_{1}-1}\left(X_{\left(n_{2}\right)}-X_{\left(n_{1}\right)}\right)^{\frac{g\left(E\left(Y_{\left(n_{i}\right)}\right)\right)}{\sum_{t=1}^{n} g\left(E\left(Y_{(t)}\right)\right)}}\right) \\
& \times\left(\prod_{i=2}^{c}\left(X_{\left(n_{i}\right)}-X_{\left(n_{1}\right)}\right)^{\frac{g\left(E\left(Y_{\left(n_{i}\right)}\right)\right)}{\sum_{t=1}^{n} g\left(E\left(Y_{\left(n_{t}\right)}\right)\right)}}\right) \\
& \times \prod_{i=1}^{c-1} \prod_{i=n_{i}+1}^{n_{i+1}^{-1}}\left(\frac{X_{\left(n_{i}\right)}+X_{\left(n_{i+1}\right)}-2 X_{\left(n_{1}\right)}}{2}\right)^{\frac{g\left(E\left(Y_{(j)}\right)\right)}{\sum_{t=1}^{n} g\left(E\left(Y_{(t)}\right)\right)}} \\
& \times\left(\prod_{i=n_{c}+1}^{n}\left(X_{\left(n_{c}\right)}-X_{\left(n_{1}\right)}\right)^{\frac{g\left(E\left(Y_{(i)}\right)\right)}{\sum_{t=1}^{n} g\left(E\left(Y_{(t)}\right)\right)}}\right) \tag{8}
\end{align*}
$$

$\tilde{W}_{1}$ and $\tilde{W}_{2}$ 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 $g\left(E\left(Y_{\left(n_{i}\right)}\right)\right) / \sum_{t=1}^{n} g\left(E\left(Y_{\left(n_{t}\right)}\right)\right)$ in $\widetilde{W}_{1}$, where $E\left(Y_{\left(n_{i}\right)}\right)$ is the expected values of $Y_{\left(n_{i}\right)}$. The $E\left(Y_{\left(n_{i}\right)}\right)$ is defined as $\int_{-\infty}^{\infty} Y_{\left(n_{i}\right)} q\left(Y_{\left(n_{i}\right)}\right) d Y_{\left(n_{i}\right)}$, where $q\left(Y_{\left(n_{i}\right)}\right)$ is the pdf of the $n_{i}$ th order statistic of $Y_{\left(n_{i}\right)}$. Since the parameters $\mu$ and $\sigma$ 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, $E\left(Y_{\left(n_{i}\right)}\right)$ will be constants and set $g(z)=e^{z} e^{-e^{z}}, z=E\left(Y_{\left(n_{i}\right)}\right)$. It also showed that the weighted factors do not depend on the parameters $\mu$ and $\sigma$.

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

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

where

$$
\begin{align*}
& \hat{W}_{1}=\sum_{i=r+2}^{r+k} \frac{g\left(E\left(Y_{(i)}\right)\right)}{S_{c}}\left(X_{(i)}-X_{(r+1)}\right) \\
& +\sum_{i=r+k+l+1}^{r+k+l+m} \frac{g\left(E\left(Y_{(i)}\right)\right)}{S_{c}}\left(X_{(i)}-X_{(r+1)}\right) \tag{9}
\end{align*}
$$

$$
\begin{align*}
& \hat{W}_{2}=\sum_{i=1}^{r} \frac{g\left(E\left(Y_{(i)}\right)\right)}{S_{n}}\left(X_{(r+2)}-X_{(r+1)}\right) \\
& +\sum_{i=r+2}^{r+k} \frac{g\left(E\left(Y_{(i)}\right)\right)}{S_{n}}\left(X_{(i)}-X_{(r+1)}\right) \\
& +\sum_{i=r+k+1}^{r+k+1} \frac{g\left(E\left(Y_{(i)}\right)\right)}{S_{n}}\left(\frac{X_{(r+k)}+X_{(r+k+l+1)}-2 X_{(r+1)}}{2}\right) \\
& +\sum_{i=r+k+l+1}^{n-s} \frac{g\left(E\left(Y_{(i)}\right)\right)}{S_{n}}\left(X_{(i)}-X_{(r+1)}\right) \\
& +\sum_{i=n-s+1}^{n} \frac{g\left(E\left(Y_{(i)}\right)\right)}{S_{n}}\left(X_{(n-s)}-X_{(r+1)}\right), \tag{10}
\end{align*}
$$

$\hat{W}_{3}=\prod_{i=r+2}^{r+k}\left(X_{(i)}-X_{(r+1)}\right)^{\frac{g\left(E\left(Y_{(i)}\right)\right)}{S_{c}}}$

$$
\begin{equation*}
\times \prod_{i=r+k+l+1}^{r+k+l+m}\left(X_{(i)}-X_{(r+1)}\right)^{\frac{g\left(E\left(Y_{(i)}\right)\right)}{S_{c}}} \tag{11}
\end{equation*}
$$

$$
\hat{W}_{4}=\prod_{i=1}^{r}\left(X_{(r+2)}-X_{(r+1)}\right)^{\frac{g\left(E\left(Y_{(i)}\right)\right)}{S_{n}}}
$$

$$
\times \prod_{i=r+2}^{r+k}\left(X_{(i)}-X_{(r+1)}\right)^{\frac{g\left(E\left(Y_{(i)}\right)\right)}{S_{n}}}
$$

$$
\times \prod_{i=r+k+1}^{r+k+l}\left(\frac{X_{(r+k)}+X_{(r+k+l+1)}-2 X_{(r+1)}}{2}\right) \frac{g\left(E\left(Y_{(i)}\right)\right)}{S_{n}}
$$

$$
\times \prod_{i=r+k+l+1}^{n-s}\left(X_{(i)}-X_{(r+1)}\right)^{\frac{g\left(E\left(Y_{(i)}\right)\right)}{S_{n}}}
$$

$$
\times \prod_{i=n-s+1}^{n}\left(X_{(n-s)}-X_{(r+1)}\right)^{\frac{g\left(E\left(Y_{(i)}\right)\right)}{S_{n}}}
$$

where
$S_{c}=\sum_{t=r+2}^{r+k} g\left(E\left(Y_{(t)}\right)\right)+\sum_{t=r+k+l+1}^{r+k+l+m} g\left(E\left(Y_{(t)}\right)\right)$,
$S_{n}=\sum_{t=1}^{n} g\left(E\left(Y_{(t)}\right)\right)$.
For comparison, the other pivotal quantity is $\hat{U}_{a}$. Let
$\hat{U}_{a}=\frac{X_{(j)}-X_{(n-s)}}{\hat{W}_{a}}, \quad n-s<j \leq n, \quad$ where
$\hat{W}_{a}=\hat{\sigma}$.
The $\hat{\sigma}$ is the AMLE of $\sigma$, 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 $\mu$ and $\sigma$. Then, we have
$1-\alpha=P\left\{0<\frac{X_{(j)}-X_{(n-s)}}{\hat{W}_{h}}<\hat{u}_{h}(1-\alpha ; n, r, k, l, m, s, j)\right\}$
$=P\left\{X_{(n-s)}<X_{(j)}<X_{(n-s)}+\hat{u}_{h}(1-\alpha ; n, r, k, l, m, s, j) \times \hat{W}_{h}\right\}$.

Table 1. The properties of parameters $\widetilde{\mu}$ and $\widetilde{\sigma}$ of $\mathbf{6 0 , 0 0 0}$ random samples for each combination case.

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

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

## 3. CALCULATION AND ALGORITHM

The exact distributions of the pivotal quantities $\hat{U}_{h}(h=1, \ldots, 4)$ can not be derived algebraically, but we can approximate the distributions of $\hat{U}_{h}(h=1, \ldots, 4)$ 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 $\mu=0, \sigma=1$. (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 $\mu$ and $\sigma$ respectively. The $\tilde{\mu}_{\text {mean }}$ and $\widetilde{\sigma}_{\text {mean }}$ 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 $\hat{U}_{h}(h=1, \ldots, 4)$ 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 $\hat{U}_{h}(h=1, \ldots, 4)$ under different significance levels of $a$.

From the above procedures, we obtain the values of $\hat{U}_{h}(h=1, \ldots, 4)$ according to the exact position of $\hat{U}_{h}(h=1, \ldots$, 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 $\hat{U}_{h}(h=1, \ldots, 4)$ 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 $\hat{W}_{h}(h=1, \ldots, 4, a)$, and then make a multiply of $\hat{W}_{h}$ by $\hat{U}_{h}(h=1, \ldots, 4, a)$ (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 $\hat{U}_{h}(h=1, \ldots, 4, a)$ 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 $\hat{U}_{h}(h=1, \ldots, 4, a)$ 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: $\mu=0, \sigma=1$.

## 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 $\mathbf{9 0} \%$ prediction intervals, the percentiles of $\hat{U}_{k}(k=1, \ldots, 4, a)$, and CMd using MLE and BLIE (Hsieh[7]).

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

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 3rd 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 $\mu$ and $\sigma$ based on multiple type II censored sample $X_{\left(n_{0}\right)} \leq X_{\left(n_{1}\right)} \leq \cdots \leq X_{(n)}$ from two-parameter Weibull distribution, then $\tilde{U}_{i}=\frac{X_{(j)}-X_{\left(n_{c}\right)}}{\tilde{W}_{1}}$ is a pivotal (parameter-free) quantity

## Proof:

Define $Y_{i}=\frac{X_{i}-\mu}{\sigma}$, then $Y_{i}$ does not depend on $\mu$ and $\sigma$. And $Y_{(i)}=\frac{X_{(i)}-\mu}{\sigma}$ is the $i$ th order statistics of $Y_{i}$.
Let $Z_{i}=Y_{(j)}-Y_{\left(n_{c}\right)}, \quad Z_{2}=\sum_{i=2}^{c} \frac{g\left(E\left(Y_{\left(n_{i}\right)}\right)\right)}{\sum_{t=2}^{c} g\left(E\left(Y_{\left(n_{t}\right)}\right)\right)}\left(Y_{\left(n_{i}\right)}-Y_{\left(n_{1}\right)}\right)$, where
$\sum_{i=2}^{c} \frac{g\left(E\left(Y_{\left(n_{i}\right)}\right)\right)}{\sum_{t=2}^{c} g\left(E\left(Y_{\left(n_{t}\right)}\right)\right)}$ 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
$\widetilde{U}_{1}=\frac{\frac{X_{(j)}-\mu}{\sigma}-\frac{X_{\left(n_{c}\right)}-\mu}{\sigma}}{\sum_{i=2}^{c} \frac{g\left(E\left(Y_{\left(n_{i}\right)}\right)\right)}{\sum_{t=2}^{c} g\left(E\left(Y_{\left(n_{t}\right)}\right)\right)}\left(\frac{X_{\left(n_{i}\right)}-\mu}{\sigma}-\frac{X_{\left(n_{1}\right)}-\mu}{\sigma}\right)}$
$=\frac{Y_{(j)}-Y_{\left(n_{c}\right)}}{\sum_{i=2}^{c} \frac{g\left(E\left(Y_{\left(n_{i}\right)}\right)\right)}{\sum_{t=2}^{c} g\left(E\left(Y_{\left(n_{t}\right)}\right)\right)}\left(Y_{\left(n_{i}\right)}-Y_{\left(n_{1}\right)}\right)}=\frac{Z_{1}}{Z_{2}}=Z_{3}$.
Similarly, it is easy to show that the estimators $\tilde{U}_{2}, \tilde{U}_{3}$, and $\widetilde{U}_{4}$ are also pivotal quantities.

### 7.2 Appendix II

Table 5. The percentiles data of $\hat{U}_{1}$ distribution for $X_{(j)}$.
Table 6. The percentiles data of $\hat{U}_{2}$ distribution for $X_{(j)}$.
Table 7. The percentiles data of $\hat{U}_{3}$ distribution for $X_{(j)}$.
Table 8. The percentiles data of $\hat{U}_{4}$ distribution for $X_{(j)}$
Table 9. The percentiles data of $\hat{U}_{a}$ distribution for $X_{(j)}$

## 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 Hefte1992; 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

