Jahangirnagar University Journal of Statistical Studies
ISSN 1022-4734

http://www.juniv.edu/department/stat

JUJSS, Vol. 36, 2022, pp. 63-76

On Some Examples of Parametric Prediction Interval for Reliability with Exponential and Weibull Distribution

Abdul Hannan Chowdhury*
School of Business and Economics, North South University, Dhaka, Bangladesh

Abstract

Despite their practical importance, prediction intervals have received little attention in texts on statistics, quality control, and reliability analysis *as well as* in life testing experiments, except in relation to regression analysis. In this p*aper*, an attempt has been made to provide a comprehensive presentation of important prediction intervals and to provide numerical examples of their application in the context of reliability. It is well regarded that the Weibull distribution is a life testing distribution and a widely known distribution in reliability and survival analysis. Nevertheless, exponential distribution is a special case of the Weibull, the case which corresponds to constant failure rate. The purpose of this paper is to discuss parametric prediction intervals for reliability when the form of the distribution is exponential or Weibull. The study would cover some of the important prediction intervals with relevant examples.

1. Introduction

In reliability and life testing experiments, prediction intervals, which use the results of a past sample, provide useful information about the realization of a random variable in a future sample from the same distribution. That is, a prediction interval is an interval which uses the results of a past sample to contain the results of a future sample from the same population with a specified probability that serves different purposes. For example, one might wish to predict the number of product failures which will occur in a future period using past data.

Suppose $X_1, X_2, \ldots, X_n$ denote an ordered random sample of size n drawn from a population of size (n+k), say, X_1, $X_2, \ldots, X_n$, X_{n+1}, $X_{n+2}, \ldots, X_{n+k}$. Now, if we consider a second (future) sample of size k, X_{n+1}, $X_{n+2}, \ldots, X_{n+k}$ from the same population where our interest is to make a probability statement for the future sample based on the information of the past sample. A prediction interval, in contrast to a confidence interval or tolerance interval, could be applicable in such a situation. It is common practice to compute a confidence interval for the population parameter such as for population mean. Sometimes a confidence interval is desired for a future observation itself, rather than its mean. In this case the confidence interval must be somewhat wider to allow for the variation of the variable itself about its mean. Since the interval is for a variable, rather than a parameter, it is sometimes referred to as a prediction interval, instead of a confidence interval. Furthermore, a p-level prediction interval for future observation; may also be

* Corresponding author: hannan.chowdhury@northsouth.edu

interpreted as a p-expectation tolerance interval. A prediction interval can easily be distinguished from a confidence interval for an unknown population parameter (such as the population mean) and a tolerance interval to contain a specified proportion of the population. For further details about confidence interval, tolerance interval and prediction interval consult Hahn [1970, 1972]. The main goal of the study is to demonstrate parametric prediction intervals for reliability when the form of the distribution is either exponential or Weibull. This study would provide a wide-ranging presentation of important prediction intervals with relevant examples in the context of reliability.

2. Review of Literature

One of the earliest papers on prediction intervals is Baker (1935). Since then, a large number of papers on prediction intervals have appeared in the literature. An early review paper on the subject is Hahn & Nelson (1973). A comprehensive review paper by Patel [1989] describes the availability of a large variety of prediction intervals for several life distributions. In the literature, a variety of prediction intervals is available for normal and exponential distributions. For the exponential distribution alone, we figured out prediction intervals by Hahn (1975), Kaminsky (1977), Kaminsky & Nelson (1974), Lawless (1971, 1972, 1977), Hahn & Meeker (1991), Nelson (1970), and others are prominent. For many of these prediction intervals, factors for calculating prediction limits are generally tabulated. However, there are some important prediction intervals for which no such tabled values are available. One of the pioneer papers on prediction intervals for a such case based on one parameter exponential distribution obtained by Lawless (1971). Since then, this model has been investigated extensively, and several prediction intervals covering diverse situations are now available for it.

In different kinds of literature, parametric and non-parametric prediction intervals have been discussed extensively by many authors. Parametric prediction intervals are intervals obtained when the form of the population is known such as normal, exponential and are discussed in papers written by Chew (1969), Hahn (1970), Hahn & Nelson (1973), Hall & Prairie (1973) and Hall, Prairie & Motlagh (1975). On the other hand, non-parametric prediction intervals are intervals obtained when the form of the distribution is unknown and these are discussed in papers written by Danzinger & Davis (1964), Lawless (1971), and Nelson (1963). But, not too many papers have discussed prediction intervals in the context of reliability.

Hsieh (1997) computed quantiles related to prediction intervals for future Weibull order statistics using a conditional method for two scenarios: (i) if only previous independent failure data are available, and (ii) if both previous independent failure data and early-failure data in ongoing experiment are available. Note that quantiles for constructing prediction intervals depend on ancillary statistics of observed data while using the conditional method. Hsieh (1996) utilized the identical method to get prediction intervals for future observations, based on only early-failure data of a current experiment. Hsieh (1997) extended the prediction problem to the case of using both previous independent data and early-failure data of the ongoing experiment. Comparisons have been made for interval widths of different parameter estimators forming prediction intervals in different ways.

Jiang & Zhang (2002) considered prediction intervals for a future observation in the context of mixed linear models assuming that the future observation is independent of the current ones using a distribution-free method. They showed that for standard mixed linear models, a simple method based on the (regression) residuals works well for constructing prediction intervals. Hahn & Meeker (1991) reviewed and compared three types of statistical intervals as the confidence interval, the prediction interval, and the tolerance interval. Note that distribution-free methods play a more significant role in prediction intervals than they do in confidence intervals, especially for large samples. Wu (2015) suggested the general weighted moments' estimators (GWMEs) of the scale parameter of one-parameter exponential distribution based on a multiply type II censored sample to construct the prediction intervals for future observations. Nevertheless, Wu (2016) proposed the prediction interval for future waiting times or inter-arrival time to demonstrate the prediction intervals based on GWMEs. Here, the objective was to investigate the utilization of GWMEs in constructing a pivotal quantity and to find out the prediction interval of future waiting times or inter-arrival times between the two consecutive future observations.

However, one of the major objectives of this study is to provide a comprehensive presentation of important parametric prediction intervals and to provide numerical examples of their application in the context of reliability when the form of the distribution is exponential or Weibull. Since the natural logarithm of a variable with a Weibull distribution has an extreme value distribution. Therefore, prediction intervals for the Weibull distribution found under this study may also be used for the extreme value distribution. This was discussed by Mann & Saunders (1969) and Antle & Rademacher (1972).

3. Prediction Interval

Suppose in a situation in which a sample of size n, $X_1, X_2,...,X_n$ is taken from the population under consideration with unknown parameters and also suppose that a second (future) sample of size k, $X_{n+1}, X_{n+2},...,X_{n+k}$ is taken from the same population. Now let us suppose that $g(X_{n+1}, X_{n+2},...,X_{n+k})$ be some statistic (function of the observations) for the second sample of size k and $g_1(X_1, X_2,...,X_n)$, $g_2(X_1, X_2,...,X_n)$ which are the functions of the first sample of size n. Thus, a two-sided 100γ % prediction interval to contain the future statistic $g(X_{n+1}, X_{n+2},...,X_{n+k})$ with probability γ (specified) is an interval with lower and upper limit $g_1(X_1, X_2,...,X_n)$ and $g_2(X_1, X_2,...,X_n)$ which are the functions of the observations in the first sample, such that the interval encloses the future statistic with probability γ, that is,

$$P[g_1(X_1, X_2,...,X_n) < g(X_{n+1}, X_{n+2},...,X_{n+k}) < g_2(X_1, X_2,...,X_n)] = \gamma \tag{1}$$

for any possible values of the unknown parameters of the underlying distribution. However, the interpretations of such prediction intervals are as follows: Suppose a 100γ% prediction interval for a future sample statistic is calculated from past samples for many such pairs of past and future samples. Then the interval will enclose the future sample statistic in a fraction γ of the pairs of samples in the long run. That is, such intervals enclose the corresponding future statistic with probability γ.

4. Simultaneous Prediction Interval

Simultaneous prediction intervals may be defined for k statistics, g_l, g_2...,g_k, each a separate function of the observations in k independent future samples from the same distribution as a first sample X_1,...,X_n. Two-sided simultaneous prediction intervals to contain g_l, g_2....,g_k, with probability γ are intervals with lower and upper endpoints $g_{i1}(X_1,...,X_n)$ and $g_{i2}(X_1,..., X_n)$, i = l,..., k, which are functions of the observations in the first sample, such that the intervals each enclose the corresponding future statistics with probability γ; that is,

$$P[g_{11} \leq g_1 \leq g_{l2}, \text{ and} \cdots \text{and } g_{kL} \leq g_k \leq g_{k2}] = \gamma \tag{2}$$

for any possible values of the unknown parameters of the underlying distribution. One-sided simultaneous prediction intervals are similarly defined. However, such simultaneous prediction intervals have the following interpretation: Suppose for many sets of a single past sample and k future samples, 100γ% simultaneous prediction intervals are calculated for sample statistics for each of k future samples. Then all of the k intervals will enclose their respective sample statistics in a fraction γ of the sets of samples in the long run. That is, such intervals enclose all of the corresponding future statistics with probability γ·

5. Problem Statement

Reliability studies and life testing experiments mostly deals problem-related to failure data. Let us suppose that in an experiment some prior failure information is given and one would like to obtain information on the next failure times or to predict the range of the future failure times. Assume that the failure times of a system that occurs from an exponential or from a single Weibull process and the successive failure times x_1, x_2,....,x_n have been recorded. The vital question is that we are concerned about the next failure occur time. In this situation, prediction intervals in contrast to confidence or tolerance intervals are appropriate. That means, a prediction interval for future failure time x_{n+1}, x_{n+2},....,x_{n+k} (in general), would be quite applicable. This implies that we may use the result of a past sample to construct an interval that will contain the results of a future sample from the same population with a specified probability. Throughout it is assumed that both the past and the future samples are obtained with simple random sampling from the same population. The validity of prediction intervals depends strongly on this key assumption.

6. Prediction Interval When Lifetime Follows One Parameter Exponential Distribution

Suppose, in a life testing experiment involving items whose life time's follow an exponential distribution and our problem is to predict the rth ordered observation X_r in a sample of n from the same distribution, based on the observed values of the first k ordered observations from the sample ($k < r \leq n$). Now, suppose that $X_1 \leq X_2 \leq ..... \leq X_n$ are ordered observations in a sample of n from the exponential with mean $1/\theta$, having density,

$$f(x;\theta) = \theta e^{-\theta x}, \qquad \theta > 0, x > 0 \tag{3}$$

Let $S_k = \Sigma X_i + (n-k)X_k$ and the variate $u = u(k,r,n)$ for given $k < r \leq n$ and u is defined by

$$u = (X_r - X_k)/S_k.$$

While deriving density function of u we note two well-known results (see Epstein and Sobel, 1953) concerning ordered observations from an exponential distribution:

(i) the variates $w_1 = nX_1$, $wi = (n-i+1)(X_i - X_{i-l})$, $i = 2,...,n$ are independently distributed with density (3), and (ii) $2\theta S_k = 2\theta\Sigma wi$ is distributed as χ^2 with 2k degrees of freedom.

It then follows rather easily that $\theta(X_r - X_k)$ and θS_k are independently distributed and that $u = (X_r - X_k)/S_k$ has a distribution not involving θ. The probability density function of u is found as

$$f(u) = k/B(r-k,n-r+1) \sum_{i=0}^{r-k-1} (r-k-1)\, (-1)^i[1 + (n-r+i+1)u]^{-k-1}\ ; \quad (u > 0) \tag{4}$$

where , $B(a, b) =(a-1)!\,(b-1)!/(a+b-1)!$ when a, b are positive integers. Integration yields

$$P(u \geq t) = k/B(r-k,n-r+1)[\sum_{i=0}^{r-k-1} (r-k-1)\, (-1)^i/(n-r-+i+1)][1 + (n-r+i+1)t]^{-k} \quad (u > 0)$$

$$i = P(t; k,r,n) \tag{5}$$

when the distribution function of u is given by $F(t) = 1 - P(u \geq t)$.

However, probability statements about u provide prediction statements on X_r, on the basis of observed X_k, S_k,. For example, the statement $P(u \leq t_o) = \alpha$ yields prediction statement,

$$P(X_r, \leq X_k + t_oS_k) = \alpha, \tag{6}$$

giving a (one-sided) 100α% prediction interval for X_r.

However, two additional remarks concerning the evaluation of the above probabilities:

(i) In the important case where r = n (that is, we wish to predict the largest observation on the basis of the k smallest), expression (5) can be expressed as

$$P(t; k,r,n) = 1 - \sum_{i=0}^{n-k} (n-k) (-1)^i [1+it]^{-k} \tag{7}$$

hence, the distribution function of $u_1 = u(k, n, n)$ is given by

$$P(u_1 \le t) = \sum_{i=0}^{n-k} (n-k) (-1)^i [1+it]^{-k} \tag{8}$$

(ii) In the special case where k = r-1, Epstein and Sobel's (1953) results, $u_2 = u(r\text{-}1, r, n) = (r-1)(n-r+1)(X_r-X_{r-1})/S_{r-1}$ is an F variate with (2, 2r-2) degrees of freedom, so that appropriate probability statements can be read from standard tables of the F distribution.

Case 1: A Life Test where All Units Are Observed Until Failure:

Consider a life test with n units and whose lifetimes follow the same exponential distribution, are put on test simultaneously, and where all units are observed until failure. We can provide a prediction interval for the largest lifetime X_n on the basis of the k smallest lifetimes $X_1 < X_2 < \cdots < X_k$; X_n is in this case the total elapsed time required to complete the test.

Numerical Example 1

Suppose that 10 items, whose lifetimes are distributed according to the same exponential distribution, are on test simultaneously, and that the first four items to fail to do so at times 30, 90, 120, 170 hours. For n = 10, k = 4 we can find $P(u_1 \le 2.10)$ is very nearly 95%. Since $X_4 = 170$ and $S_4 = 1430$, this yields the prediction statement $P(X_{10} \le 170 + (2.10)\ 1430) = P(X_{10} \le 3173) = .95$. That is, we can be (approximately) 95% confident that the total elapsed test time will not exceed 3173 hours.

Case 2: A Life Test Where Testing Is Terminated After the rth Failure:

Consider a life testing situation similar to that in the above section, but suppose that it had been decided beforehand to terminate the test after the fifth failure.

Numerical Example 2

On the basis of the first four failure times, we can compute, say, an upper 95% prediction limit for X_5. With r = 5, k = 4, n = 10, we consider $u = (X_r - X_k)/S_k$; using equation (5) or noting that in this case 24u is an F variate with (2, 8) degrees of freedom, we find that $P(u \leq .1858) = .95$. Given the observed values $X_4 = 170$, $S_4 = 1430$, this yields the prediction statement

$$P(X_5 \leq 436) = .95.$$

We can be 95% confident that the fifth failure will occur before 436 hours.

7. Prediction Interval Based on Ranges when Lifetimes Follows Two Parameter Exponential Distribution.

Let $R_o(n)$ be the sample range of the lifetimes when n items are put on a life test (without replacement). Similarly, let $R_f(k)$ be the future s-independent sample range of the lifetimes when k items are put on a similar life test. Then the first prediction interval that we have obtained could be used to predict $R_f(k)$ on the basis of the observed $R_o(n)$.

Let, the lifetimes of all items follows a two-parameter exponential distribution, then the two parameter exponential distribution is:

$$f(x; \theta, \beta) = (1/ \theta)\exp[(x-\beta)/\theta], \text{ for all } x \geq \beta.$$

If β is known, the distribution becomes, effectively, the one-parameter exponential. Here, θ = scale parameter (unknown), β = location parameter (known or unknown), and $(X_1 < X_2 < ...... < X_n)$, $(X_{n+1} < X_{n+2} < ..... < X_{n+k})$ = ordered failure times from two s-independent samples. Since, $R_o(n)$, $R_f(k)$ = ranges of past and future samples respectively, where, $R_o(n) = X_n - X_1$, and $R_f(k) = X_{n+k} - X_{n+1}$.

Also $(X^*_1 < X^*_2 < ...... < X^*_n)$, $(X^*_{n+1} < X^*_{n+2} < ...... < X^*_{n+k})$ = ordered standardized r.v.'s from two s-independent samples from exponential distribution and $R_o^*(n)$, $R_f^*(k)$ = standardized sample ranges past and future samples respectively. Moreover, $1-\alpha$ = prediction probability.

Now, consider the ratio of two independent ranges, $V = R_f(k)/R_o(n) = R_f^*(k)/R_o^*(n)$. The probability distribution of the r.v. Here, V, does not depend on any parameter(s).The Cdf $H_f(Xj)$ of the r.v. $R_f^*(k)$ is known [see David, 1981, p 12]:

$H_f(X_j) = (1-e^{-xj})^{k-1}$, for $X_j \geq 0$. Similarly, cdf

$H_o(X_i) = (1- e^{-xj})^{n-1}$, for $X_i \geq 0$.

Now consider P $(V \leq v) = P[R_f^*(k) \leq v\, R_o^*(n)]$ (using binomial theorem, see Colangelo and Patel, 1972).

Let $v_\gamma = v(\gamma;m,n)$ be the γ lower quantile for the Cdf(K).

$$\text{Then, } P\{v_{\alpha 1} \leq R_f(k)/R_o(n) \leq v_{1-\alpha 2}\} = l - \alpha \tag{10}$$

with $0 < \alpha_1 < 1$, $0 < \alpha_2 < 1$ and $\alpha_1 + \alpha_2 = \alpha$. The $(1-\alpha)$ two-sided prediction interval of the future sample range $R_f(k)$ on the basis of the past sample range $R_o(n)$ is:

$$[v_{\alpha 1} R_o(n),\ v_{1-\alpha 2} R_o(n)]. \tag{11}$$

Similarly, $(1-\alpha)$ one-sided prediction intervals of the future sample ranges $R_f(k)$ are:

lower prediction limits for $R_f(k)$ is: $[v_\alpha R_o(n),\ \infty]$,

and upper prediction limits for $R_f(k)$ is: $[O,\ v_{1-\alpha 2} R_o(n)]$

For computation of prediction factors consult Colangelo and Patel (1972).

Numerical Example 3

In an accelerated life test, consider the following failure times (in weeks) of 10 transistors having a 2-parameter exponential life distribution: 7, 9, 9, 10, 13, 14, 16, 17, 19, 25. One would like to predict the range of the failure times of a future such test of 15 transistors using a 90% prediction interval.

Then, for n = 10, $R_o(10) = 25-7 = 18$, k = 15, and the equal-tail case of $\alpha_1 = \alpha_2 = 0.05$, we find $v_{o.o5} = 0.464797$ and $v_{o.95} = 3.01029$ from table 1 of Colangelo and Patel, (1972). This provides a 90% two-sided prediction interval for $R_f(15)$ as: $[0.464797 \cdot 18, 3.01029 \cdot 18] = [8.37, 54.19]$.

Similarly, 90% one-sided lower and upper prediction intervals for future sample ranges are:

lower prediction intervals: $[0.569004 \cdot 18, 90) = [10.24, \infty)$

and upper prediction intervals is: $[0, 2.43540 \cdot 18] = [0, 43.84]$.

8. Prediction Interval Based on Waiting Time when Lifetime follows Two Parameter Exponential Distribution:

Let W(i) be the waiting time between failures (i-1) and i when n items are put on a life test (without replacement), (i= 1,2,...,n). Now, we want to obtain a prediction interval which can be used to predict the future waiting time W(s) on the basis of the observed (past) waiting time W(r),$(1 \leq r < s \leq n)$.

Let, W(i) = waiting time between failures (i-1) and i: $W(i) = X_i - X_{i-1}$, i=1,2,....n and w*(i) = standardized waiting time between failures (i-1) and i: $w^*(i) = X_i - X_{i-1}$, i=1,2,....n, also 1-α = prediction probability. Let us consider the ratio,

$$F= (n-s+l)W(s)/(n-r+l)W(r) = 2(n-s+l)W^*(s)/2(n-r+l)W^*(r); (1 \leq r < s \leq n). \tag{12}$$

The (n-s+l)W*(s) is s-independent of (n-r+ I)W*(r), and both have the two-parameter exponential distribution [see David, 1981]. The random variable F has an F-distribution with 2 degrees of freedom in both numerator and denominator. Let $f\gamma = f(\gamma;2,2)$ be its γ lower quantile: $P(F \leq f\gamma) = \gamma$. Since

$$P\{f_{\alpha 1} \leq (n-s+l)W(s)/(n-r+l)W(r) \leq f_{1-\alpha 2}\}= l- \alpha \tag{13}$$

with $\alpha_1+\alpha_2 = \alpha$ defined as in previous section, we have a (l- α) two-sided prediction interval of a future waiting time W(s) on the basis of the past waiting time W(r):

$$[cW(r)f_{\alpha 1,} cW(r)\ f_{1-\alpha 2}] \tag{14}$$

where, c =(n-r+l)/(n-s+l). Similarly (l- α) one-sided lower and upper prediction limits W(s) are:

$$[cW(r)f_{\alpha 1}, \infty) \text{ and } [0, cW(r)\ f_{1-\alpha 2}] \tag{15}$$

Since $f\gamma = f(\gamma;2,2)$ can be found from the F-distribution tables. It should be mentioned that necessary prediction factors can be obtained from known tables and to get these prediction factors see Colangelo and Patel, (1972).

Numerical Example 4

In an accelerated life test, let 15 transistors be put on the test (with replacement) and let failure times have a two-parameter Exponential distribution. Let the failure times for transistors #4 and #5, be 10 and 13 weeks, respectively. One would like to find a 90% prediction interval for the waiting time between future failures #9 and #10.

Here n =15, r = 5, W (5) =13-10 =3, s = 10; and $f_{0.90} = f(0.90; 2, 2) = 9.0$. This provides a 90% one-sided upper prediction limit of: (15-5+1 / 15-10+1) 3 · 9 = 49.5.

9. Prediction Intervals when Lifetimes Follows Weibull Distribution

The procedures for obtaining prediction intervals for a future sample from an exponential distribution can be readily extended to obtain prediction intervals for a sample from a Weibull distribution with a known value of the shape parameter. Mann and Saunders (1969) provided one-sided lower prediction limits for the smallest value in a future sample when samples are taken from a Weibull distribution. They used such a limit as the warranty time for the life of a product. Mann (1970) extended tabulations for such a prediction limit based on a linear combination of three selected order statistics from the first sample.

Suppose the lifetimes of all items follow a three parameter Weibull distribution, then

$$f_x(x) = \beta/\theta[(x\text{-}n)/\theta]^{\beta-1} \exp[-(x\text{-}n)/\theta)^{\beta}]. \tag{16}$$

The parameters β, θ and n are referred to as shape, scale and location parameters respectively. This can also be expressed as $X \rightarrow W(\theta, \beta, n)$. Consider a single Weibull process such as the failure times of a system, and suppose the successive failure times $x_1,...,x_n$ have been recorded. Perhaps the most natural question concerns when the next failure will occur. This suggested that a prediction interval for x_{n+1}, or more generally for x_{n+m} would be quite useful and meaningful in this framework. A prediction interval is a confidence interval for a future observation. Thus a γ level lower prediction limit for x_{n+m} is a statistic T_L (n, m, y) such that P [T_L (n, m, y) $\leq$ x_{n+m}] = γ.

Consider first the case m = 1. The limit TL should be a function of the sufficient statistics and the probability must be free of parameters.

Theorem: *Suppose X_n ,...,X_{n+1} denote the first n+1 successive times of occurrence of a Weibull process, and suppose the observed values x_1, ... x_n are available. Then, a lower γ level prediction limit for X_{n+1} is $T_L(n, 1, \gamma) = x_n \exp[(\gamma^{-1/(n-1)} -1)/\beta]$* (see Bain and Engelhardt, 1991).

Numerical Example 5

Crow (1974) provided the following simulated data for k = 3 systems with true common β = 0.5 and common θ = 2.778. The data are actually obtained using time truncation at time 200, but for illustrative purposes suppose failure truncation had been employed.

System 1:
4.3, 4.4, 10.2, 23.5, 23.8, 26.4, 74.0, 77.1, 92.1, 197.2

System 2:
0.1, 5.6, 18.6, 19.5, 24.2, 26.7, 45.1, 45.8, 75.7, 79.7, 98.6, 120.1, 161.8, 180.6, 190.8

System 3:
8.4, 32.5, 44.7, 48.4, 50.6, 73.6, 98.7, 112.2, 129.8, 136.0, 195.8

Now, consider the system 1 data. A lower 90% predicted failure time will be

$$T_L(10, 1, 0.90) = 197.2 \ \exp[(0.90^{-1/9} -1) / 0.51] = 201.8$$

10. Prediction Intervals of Waiting Time with GWME of the Scale Parameter of the One-Parameter Exponential Distribution

To predict the waiting time, the pivotal quantity is measured as $V = (Y_{(j)} - Y_{(j-1)})/\hat{\theta}$, $n-s < j \leq n$ based on the GWME $\hat{\theta}$ of the scale parameter of the one-parameter exponential distribution where the lifetimes Y with pdf given by $f(y) = 1/\theta \exp(-y/\theta)$, $y \geq 0$, $\theta > 0$.

Suppose that $Y_{(r+1)} < \ldots < Y_{(r+k)} < Y_{(r+k+l+1)} < \ldots < Y_{(n-s)}$ be the available multiply type II censored sample from the above distribution. The GWME to estimate the scale parameter θ is defined as

$$\hat{\theta} = W_{r+1} Y_{(r+1)} + \ldots + W_{r+k} Y_{(r+k)} + W_{r+k+l+1} Y_{(r+k+l+1)} + \ldots + W_{n-s} Y_{(n-s)} = W_{\sim}^{T} Y_{\sim},$$

where $W_{\sim} = [W_{r+1}, \ldots, W_{r+k}, W_{r+k+l+1}, \ldots, W_{n-s}]^T$ and $Y_{\sim} = [Y_{(r+1)}, \ldots, Y_{(r+k)}, Y_{(r+k+l+1)}, \ldots, Y_{(n-s)}]^T$.

The weights $W_{\sim}$ are determined so that the MSE of the proposed GWME is minimized and the GWME with minimum MSE is obtained as $\theta = W_{\sim}^{T} Y_{\sim}$.

Since $Y_{(1)}/\theta$, …., $Y_{(n)}/\theta$ are the n order statistics from a standard exponential distribution and $\hat{\theta}/\theta = W_{\sim}^{T}/Y_{\sim}\theta$ is a linear combination of n order statistics from a standard exponential distribution, the distribution of pivotal quantity $V = (Y_{(j)} - Y_{(j-1)}/\theta) / \hat{\theta}/\theta$ is independent of θ, $n - s < j \leq n$.

Let $V(\delta; n, j, r, k, l, s)$ be the δ percentile of the distribution of V satisfying $P(V \leq V(\delta; n, j, r, k, l, s)) = \delta$. Make use of the pivotal quantity, and the prediction interval of waiting time $Y_{(j)} - Y_{(j-1)}$, $n - s < j \leq n$ is proposed in the following theorem.

Theorem: For multiply type II censored sample $Y_{(r+1)} < \ldots < Y_{(r+k)} < Y_{(r+k+l+1)} < \ldots < Y_{(n-s)}$, the prediction interval of waiting time $Y_{(j)} - Y_{(j-1)}$, $n - s < j \leq n$ is $(V(\alpha/2; n, j, r, k, l, s)\hat{\theta}, V(1-\alpha/2; n, j, r, k, l, s)\hat{\theta})$ (for further details see Wu 2016).

Numerical Example 6

Suppose that the time to breakdown of an insulating fluid between electrodes is assumed to be exponentially distributed and recorded at 5 different voltages [6]. To illustrate the prediction interval of waiting time assuming 35 kV, the data with n = 12, r = 2, k = 3, l = 1 and s = 5 and the multiply type-II censored failure times (seconds) are: –, –, 41, 87, 93, –, 116, –, –, –, –, –. The weights are 0.23568, 0.12544, 0.19776, and 0.8058.

Here, the estimated scale parameter is $\hat{\theta} = W_{(3)} Y_{(3)} + W_{(4)} Y_{(4)} + W_{(5)} Y_{(5)} + W_{(7)} Y_{(7)}$
$= 41 * 0.23568 + 87 * 0.12544 + 93 * 0.19776 + 116 * 0.8058 = 132.4406$.

We obtained 95% prediction intervals for future waiting times $Y_{(8)} - Y_{(7)}$, $Y_{(9)} - Y_{(8)}$, $Y_{(10)} - Y_{(9)}$, $Y_{(11)} - Y_{(10)}$, and $Y_{(12)} - Y_{(11)}$ are obtained for V (0.025; 12, j, 2, 3, 1, 5), V (0.975; 12, j, 2, 3, 1, 5) corresponds to [0.0058 & 1.1098] with prediction interval is (0.76816, 146.9826); [0.0073 & 1.3875] with prediction interval is (0.9668, 183.7613); [0.0098 & 1.8605] with prediction interval is (1.2979, 246.4057); [0.0146 & 2.7797] with prediction interval is (1.9336, 368.1451); and [0.0292 & 5.5634] with prediction interval is (3.8673, 736.8200), (See for details Wu 2016).

11. Concluding Remarks

For a single future observation, a prediction interval is an interval that will contain a future observation from a population with a specified coverage probability. Many practical problems require that a past sample be used to construct a prediction interval to contain the results of a future sample. In this paper, the author presented most of the available prediction intervals for life testing experiments especially when the lifetimes follow exponential or Weibull distribution, and illustrated their use to provide a guide to those who require such methods in practical applications. Since the natural logarithm of a variable with a Weibull distribution has an extreme value distribution. Therefore, prediction intervals for the Weibull distribution found under this study may also be used for the extreme value distribution.

References

Antle, C. E., and Rademaker, F., (1972). “An Upper Confidence Limit on the Maximum of m Future Observations from a Type I Extreme Value Distribution,” *Biometrika*, 59, 475–477.

Baker, G. A., (1935). “The Probability that the Mean of a Second Sample will differ from the Mean of a First Sample by Less than a Certain Multiple of the Standard Deviation of the First Sample,” *Annals of Mathematical Statistics*, 6, 197-201.

Bain, L. J., and Engelhardt, M., (1991). “Statistical Analysis of Reliability and Life-Testing Models-Theory and Methods,” *Marcel Dekker, Inc.* NY.

Bain, P. T., and Patel, J. K., (1991). “Factors for Calculating Prediction Intervals for Samples from a One-Parameter Exponential Distribution,” *Journal Quality Technology*, 23, 48-52.

Chew, V., (1966). “Confidence, Prediction and Tolerance Regions for the Multivariate Normal Distribution,” *Journal of the American Statistical Association*, 61, 605–617.

Colangelo, L and Patel, J.K., (1972). “Prediction Interval Based on Ranges and Waiting Times for an Exponential Distribution,” *IEEE Transactions on Reliability*, 41, 3, 469-472.

David, H. A., (1981). “Order Statistics,” (2nd edition), John Wiley & Sons.

Danzinger, L., and Davis, S., (1964). “Tables of Distribution-Free Tolerance Limits,” *Annals of Mathematical Statistics*, 35, 1361-1365.

Epstein, B and Sobel, M., (1953). “Life Testing,” *Journal of American Statistical Association*, 48, 486-502.

Hahn, G. J., (1970). "Statistical Intervals for a Normal Population," *Journal of Quality Technology*, 2, 115-125, and 195-206.

Hahn, G. J., (1970). "Understanding Statistical Intervals," *Industrial Engineering*, 2, 45-48.

Hahn, G. J., (1972). "Simultaneous Prediction Intervals to Contain the Standard Deviation or Ranges of Future Samples from a Normal Distribution," *Journal of American Statistical Association*, 67, 938-942.

Hahn, G. J., and Nelson, W., (1973). "A Survey of Prediction Intervals and Their Applications," *Journal of Quality Technology*, 5, 4, 173-188.

Hahn, G. J., and Meeker, W. Q., (1991). Statistical Intervals- A Guide for Practitioners. *John Wiley*, NY.

Hall, I. J., and Prairie, R. R., (1973). "One-sided Prediction Intervals to Contain at Least *M* of k Future Observations," *Technometrics*, 15.

Hall, I. J., Prairie, R. R., and Motlagh, C. K., (1975). "Non-Parametric Prediction Intervals," *Journal of Quality Technology*, 7, 3, 109-114.

Hsieh, H. K., (1997). "Prediction Intervals for Weibull Order Statistics," *Statistica Sinica*, 7, 1039-1051.

Hsieh, H. K., (1996). "Prediction Intervals for Weibull Observations, Based on Early-failure Data." *IEEE Transactions on Reliability*, 45, 666-670.

Kaminsky, K. S., (1977). "Comparison of Prediction Intervals of Failure Times when Life is Exponential," *Technometrics*, 19(1), 83-86.

Kaminsky, K. S., and Nelson, P. I., (1973). "Prediction Intervals for a Sample Quantile," *Statistical Paper*, 16, 1, 57-59.

Lawless, J. F., (1971). "A Prediction Problem Concerning Samples from the Exponential Distribution, with Applications in Life Testing," *Technometrics*, 13, 725-730.

Lawless, J. F., (1972). "On Prediction Intervals for Samples from the Exponential Distribution and Prediction Limits for System Survival," *Sankhya*, 34, 1-14.

Lawless, J. F., (1977). "Prediction Intervals for the Two Parameter Exponential Distribution," *Technometrics*, 19, 469-47.

Mann, N. R., (1970). "Warranty Periods Based on Three Ordered Sample Observations from a Weibull Population," *IEEE Transactions on Reliability*, R-19, 167-171.

Mann, N. R.and Saunders, S· C·, (1969). "On Evaluation of Warranty Assurance when Life has Weibull Distribution," *Biometrika*, 56, 615-625.

Nelson, W. "A Statistical Prediction Interval for Availability," *IEEE Transactions on Reliability*, R-19, 179-182, (1970).

Patel, J. K, "Prediction Intervals- A Review," *Communications in Statistics – Theory and Methods*, 18, 2393-2465, (1989).

Jiang, J., and Zhang, W., "Distribution-Free Prediction Intervals In Mixed Linear Models," *Statistica Sinica*, 12, 537-553, (2002).

Wu, S. F., “The General Weighted Moment Estimator of the Scale Parameter of the Exponential Distribution Under Multiply Type II Censoring,” *ICIC Express Letters,* Part B: Applications, 6, 2, 453-457, (2015).

Wu, S. F., “Prediction Interval of Future Waiting Times of the Exponential Distribution under Multiply Type II Censoring,” *ICIC Express Letters,*10, 9, 2115-2119, (2016).