Communications of the Korean Statistical Society
2009, Vol. 16, No. 2, 349–361

Goodness-of-fit Test for the Weibull Distribution Based on Multiply Type-II Censored Samples

Suk-Bok Kang[1,a], Jun-Tae Han[b]

[a]Dept. of Statistics, Yeungnam Univ.,
[b]Institute for National Health Insurance, National Health Insurance Co.

Abstract

In this paper, we derive the approximate maximum likelihood estimators of the shape parameter and the scale parameter in a Weibull distribution under multiply Type-II censoring by the approximate maximum likelihood estimation method. We develop three modified empirical distribution function type tests for the Weibull distribution based on multiply Type-II censored samples. We also propose modified normalized sample Lorenz curve plot and new test statistic.

Keywords: Approximate maximum likelihood estimator, goodness-of-fit test, modified normalized sample Lorenz curve, multiply Type-II censored sample, Weibull distribution.

1. Introduction

The Weibull distribution is a widely used and flexible probability distribution that has many applications in many fields such as reliability theory, survival analysis, bioassay, and environmental science. Its cumulative distribution function(cdf) and the probability density function(pdf) of the random variable having the Weibull distribution are given by

$$F_X(x;\sigma,\lambda) = 1 - \exp\left[-\left(\frac{x}{\sigma}\right)^{\lambda}\right], \quad x>0,\ \sigma>0,\ \lambda>0 \tag{1.1}$$

and

$$f_X(x;\sigma,\lambda) = \frac{\lambda}{\sigma^{\lambda}} x^{\lambda-1} \exp\left[-\left(\frac{x}{\sigma}\right)^{\lambda}\right]. \tag{1.2}$$

The popularity of the Weibull distribution for life data analysis is due primarily to its flexibility. Because of its flexibility and popularity, many inferential techniques are found in the literature. In many life test studies, it is common that the lifetimes of test units may not be able to record exactly. For example, multiply Type-II censored sampling arises in a life-testing experiment whenever the experimenter does not observe the failure times of some units placed on a life-test.

Balakrishnan *et al.* (1995) derived the estimators for the location and scale parameters of the extreme value distribution under multiply Type-II censoring. Wu and Yang (2002) proposed the weighted moments estimators of the scale parameter of the exponential distribution based on a multiply Type-II censored sample. Balakrishnan *et al.* (2004) discussed point and interval estimation for

[1] Crresponding author: Professor, Department of Statistics, Yeungnam University, Gyeongsan 712-749, Korea.
E-mail: sbkang@yu.ac.kr

the extreme value distribution under progressive Type-II censoring. Ng *et al.* (2004) computed the expected fisher information and the asymptotic variance covariance matrix of the maximum likelihood estimator based on a progressive Type-II censored sample from a Weibull distribution. Han and Kang (2006) derived approximate maximum likelihood estimators(AMLEs) of the scale parameter and the location parameter in the two-parameter Rayleigh distribution under multiply Type-II censoring by the approximate maximum likelihood estimation method when two parameters are unknown.

Aho *et al.* (1985) studied goodness-of-fit tests for two-parameter Weibull distribution based on Type-II censored samples with both parameters assumed unknown. Gibson and Higgins (2000) proposed a new goodness-of-fit test for Weibull or extreme value distribution with left or right censored data. Kang and Lee (2006a) proposed three modified entropy estimators based on Type-II censored samples and the goodness-of-fit tests of the Weibull distribution based on the proposed entropy estimators. Wang (2008) proposed a test statistic to test whether the progressive Type-II censored samples come from an exponential distribution. Sanjel and Balakrishnan (2008) proposed a Laguerre polynomial approximation for a goodness-of-fit test for exponential distribution based on progressive Type-II right censored data.

In this paper, we derive the AMLEs of the shape parameter λ and the scale parameter σ under multiply Type-II censored sample. We use three modified empirical distribution function(EDF) type test for the Weibull distribution based on multiply Type-II censored samples using proposed AMLEs. We also propose modified normalized sample Lorenz curve plot and new test statistic for the Weibull distribution based on multiply Type-II censored samples.

2. Approximate Maximum Likelihood Estimators

Let X be a random variable with *pdf* (1.2). The distribution of $Y = \ln X$ have the feature, which is the extreme-value distribution with location parameter $\mu = \ln\sigma$ and scale parameter $\theta = 1/\lambda$. The *pdf* and *cdf* of Y given by

$$f(y;\mu,\theta) = \frac{1}{\theta}\exp\left(\frac{y-\mu}{\theta}\right)\exp\left[-\exp\left(\frac{y-\mu}{\theta}\right)\right] \tag{2.1}$$

and

$$F(y;\mu,\theta) = 1-\exp\left[-\exp\left(\frac{y-\mu}{\theta}\right)\right]. \tag{2.2}$$

Let us assume that the multiply Type-II censored sample from a sample of size n is

$$Y_{a_1:n} < Y_{a_2:n} < \cdots < Y_{a_s:n}, \tag{2.3}$$

where $1 \le a_1 < a_2 < \cdots < a_s \le n$, $a_0 = 0$, $a_{s+1} = n+1$, $F(y_{a_0:n}) = 0$ and $F(y_{a_{s+1}:n}) = 1$.

The likelihood function based on the multiply Type-II censored sample (2.1) can be written as

$$L = n!\prod_{j=1}^{s} f(y_{a_j:n})\prod_{j=1}^{s+1}\frac{\left[F\left(y_{a_j:n}\right)-F\left(y_{a_{j-1}:n}\right)\right]^{a_j-a_{j-1}-1}}{\left(a_j-a_{j-1}-1\right)!}. \tag{2.4}$$

The random variable $Z_{i:n} = (Y_{i:n}-\mu)/\theta$, and $f(z)$ and $F(z)$ are the *pdf* and the *cdf* of the standard extreme-value distribution, respectively.

Since $f'(z) = f(z)(1-e^z)$ and $1-F(z) = f(z)/e^z$ on differentiating the log-likelihood functions as follows;

$$\frac{\partial \ln L}{\partial \mu} = -\frac{1}{\theta}\left[(a_1-1)\frac{f(z_{a_1:n})}{F(z_{a_1:n})} - (n-a_s)e^{z_{a_s:n}} + \sum_{j=1}^{s} e^{z_{a_j:n}} + \sum_{j=2}^{s}(a_j - a_{j-1} - 1)\frac{f(z_{a_j:n}) - f(z_{a_{j-1}:n})}{F(z_{a_j:n}) - F(z_{a_{j-1}:n})}\right] \tag{2.5}$$
$$= 0$$

and

$$\frac{\partial \ln L}{\partial \theta} = -\frac{1}{\theta}\left[(a_1-1)\frac{f(z_{a_1:n})}{F(z_{a_1:n})}z_{a_1:n} - (n-a_s)e^{z_{a_s:n}}z_{a_s:n} + \sum_{j=1}^{s} z_{a_j:n} - \sum_{j=1}^{s} e^{z_{a_j:n}} z_{a_j:n}\right. \tag{2.6}$$
$$\left. + \sum_{j=2}^{s}(a_j - a_{j-1} - 1)\frac{f(z_{a_j:n})z_{a_j:n} - f(z_{a_{j-1}:n})z_{a_{j-1}:n}}{F(z_{a_j:n}) - F(z_{a_{j-1}:n})}\right]$$
$$= 0.$$

Since the above equations are very complicated, the Equation (2.5) and (2.6) do not admit explicit solutions for μ and θ, respectively.

Let $\xi_i = F^{-1}(p_i) = \ln(-\ln q_i)$, where $p_i = i/(n+1)$, $q_i = 1 - p_i$.

The Equation (2.5) does not admit an explicit solution for μ. But we can approximate the following functions by

$$\frac{f(z_{a_1:n})}{F(z_{a_1:n})} \approx \alpha_2 + \beta_2 z_{a_1:n} \tag{2.7}$$

$$e^{z_{a_j:n}} \approx e^{\xi_{a_j}}(1 - \xi_{a_j}) + e^{\xi_{a_j}} z_{a_j:n} \tag{2.8}$$

$$\frac{f(z_{a_j:n})}{F(z_{a_j:n}) - F(z_{a_{j-1}:n})} \approx \alpha_{2j} + \beta_{2j} z_{a_j:n} + \gamma_{2j} z_{a_{j-1}:n} \tag{2.9}$$

$$\frac{f(z_{a_{j-1}:n})}{F(z_{a_j:n}) - F(z_{a_{j-1}:n})} \approx \alpha_{3j} + \beta_{3j} z_{a_j:n} + \gamma_{3j} z_{a_{j-1}:n} \tag{2.10}$$

and

$$\frac{f(z_{a_j:n}) - f(z_{a_{j-1}:n})}{F(z_{a_j:n}) - F(z_{a_{j-1}:n})} \approx \alpha_{4j} + \beta_{4j} z_{a_j:n} + \gamma_{4j} z_{a_{j-1}:n}, \tag{2.11}$$

where

$$\alpha_2 = \frac{f(\xi_{a_1})}{p_{a_1}}\left[1 - (1 + \ln q_{a_1})\xi_{a_1} + \frac{f^2(\xi_{a_1})}{p_{a_1}}\xi_{a_1}\right],$$

$$\beta_2 = \frac{f(\xi_{a_1})}{p_{a_1}}\left[(1 + \ln q_{a_1}) - \frac{f^2(\xi_{a_1})}{p_{a_1}}\right],$$

$$\alpha_{2j} = \frac{f(\xi_{a_j})}{p_{a_j} - p_{a_{j-1}}}\left[1 - (1 + \ln q_{a_j})\xi_{a_j} + K_j\right],$$

$$\beta_{2j} = \frac{f(\xi_{a_j})}{p_{a_j} - p_{a_{j-1}}}\left[(1 + \ln q_{a_j}) - \frac{f(\xi_{a_j})}{p_{a_j} - p_{a_{j-1}}}\right],$$

$$\gamma_{2j} = \frac{f(\xi_{a_j})f(\xi_{a_{j-1}})}{\left(p_{a_j} - p_{a_{j-1}}\right)^2},$$

$$\alpha_{3j} = \frac{f(\xi_{a_{j-1}})}{p_{a_j} - p_{a_{j-1}}}\left(1 - (1 + \ln q_{a_{j-1}})\xi_{a_{j-1}} + K_j\right),$$

$$\beta_{3j} = -\frac{f(\xi_{a_j})f(\xi_{a_{j-1}})}{\left(p_{a_j} - p_{a_{j-1}}\right)^2}, \qquad \gamma_{3j} = \frac{f(\xi_{a_{j-1}})}{p_{a_j} - p_{a_{j-1}}}\left[(1 + \ln q_{a_{j-1}}) + \frac{f(\xi_{a_{j-1}})}{p_{a_j} - p_{a_{j-1}}}\right],$$

$$\alpha_{4j} = \alpha_{2j} - \alpha_{3j}, \qquad \beta_{4j} = \beta_{2j} - \beta_{3j}, \qquad \gamma_{4j} = \gamma_{2j} - \gamma_{3j},$$

$$K_j = \frac{f(\xi_{a_j})\xi_{a_j} - f(\xi_{a_{j-1}})\xi_{a_{j-1}}}{p_{a_j} - p_{a_{j-1}}}.$$

Next, we can approximate the following function by

$$\frac{f(z_{a_1:n})}{F(z_{a_1:n})} z_{a_1:n} \approx \alpha_1 + \beta_1 z_{a_1:n} \tag{2.12}$$

$$e^{z_{a_j:n}} z_{a_j:n} \approx -e^{\xi_{a_j}} \xi_{a_j}{}^2 + e^{\xi_{a_j}}(\xi_{a_j} + 1) z_{a_j:n} \tag{2.13}$$

and

$$\frac{f(z_{a_j:n}) z_{a_j:n} - f(z_{a_{j-1}:n}) z_{a_{j-1}:n}}{F(z_{a_j:n}) - F(z_{a_{j-1}:n})} \approx \alpha_{1j} + \beta_{1j} z_{a_j:n} + \gamma_{1j} z_{a_{j-1}:n}, \tag{2.14}$$

where

$$\alpha_1 = \frac{f(\xi_{a_1})}{p_{a_1}} \xi_{a_1}^2 \left[\frac{f(\xi_{a_1})}{p_{a_1}} - (1 + \ln q_{a_1})\right],$$

$$\beta_1 = \frac{f(\xi_{a_1})}{p_{a_1}}\left[1 + (1 + \ln q_{a_1})\xi_{a_1} - \frac{f(\xi_{a_1})}{p_{a_1}}\xi_{a_1}\right],$$

$$\alpha_{1j} = K_j^2 - \frac{f'(\xi_{a_j})\xi_{a_j}^2 - f'(\xi_{a_{j-1}})\xi_{a_{j-1}}^2}{p_{a_j} - p_{a_{j-1}}},$$

$$\beta_{1j} = \frac{f(\xi_{a_j})}{p_{a_j} - p_{a_{j-1}}}\left[1 + (1 + \ln q_{a_j})\xi_{a_j} - K_j\right],$$

$$\gamma_{1j} = -\frac{f(\xi_{a_{j-1}})}{p_{a_j} - p_{a_{j-1}}}\left[1 + (1 + \ln q_{a_j})\xi_{a_j} - K_j\right].$$

By substituting the Equations (2.7), (2.8) and (2.11) into the Equation (2.5), we can derive an estimator of μ as follows;

$$\hat{\mu} = \frac{A_0 B_1 - A_1 B_0}{A_0 C_1 - A_1 C_0}, \tag{2.15}$$

where

$$A_0 = (a_1 - 1)\alpha_2 - (n - a_s)e^{\xi_{a_s}}(1 - \xi_{a_s}) + s - \sum_{j=1}^{s} e^{\xi_{a_j}}(1 - \xi_{a_j}) + \sum_{j=2}^{s}(a_j - a_{j-1} - 1)\alpha_{4j},$$

$$B_0 = (a_1 - 1)\beta_2 Y_{a_1:n} - (n - a_s)e^{\xi_{a_s}} Y_{a_s:n} - \sum_{j=1}^{s} e^{\xi_{a_j}} Y_{a_j:n} + \sum_{j=2}^{s}(a_j - a_{j-1} - 1)\left(\beta_{4j} Y_{a_j:n} + \gamma_{4j} Y_{a_{j-1}:n}\right),$$

$$C_0 = (a_1 - 1)\beta_2 - (n - a_s)e^{\xi_{a_s}} - \sum_{j=1}^{s} e^{\xi_{a_j}} + \sum_{j=2}^{s}(a_j - a_{j-1} - 1)(\beta_{4j} + \gamma_{4j}),$$

$$A_1 = s + (a_1 - 1)\alpha_1 + (n - a_s)e^{\xi_{a_s}}\xi_{a_s}^2 + \sum_{j=1}^{s} e^{\xi_{a_j}}\xi_{a_j}^2 + \sum_{j=2}^{s}(a_j - a_{j-1} - 1)\alpha_{1j},$$

$$B_1 = (a_1 - 1)\beta_1 Y_{a_1:n} - (n - a_s)e^{\xi_{a_s}}(\xi_{a_s} + 1)Y_{a_s:n} + \sum_{j=1}^{s} Y_{a_j:n} - \sum_{j=1}^{s} e^{\xi_{a_j}}(\xi_{a_j} + 1)Y_{a_j:n}$$
$$+ \sum_{j=2}^{s}\left(a_j - a_{j-1} - 1\right)\left(\beta_{1j}Y_{a_j:n} + \gamma_{1j}Y_{a_{j-1}:n}\right),$$

$$C_1 = (a_1 - 1)\beta_1 - (n - a_s)e^{\xi_{a_s}}(\xi_{a_s} + 1) + s - \sum_{j=1}^{s} e^{\xi_{a_j}}(\xi_{a_j} + 1) + \sum_{j=2}^{s}(a_j - a_{j-1} - 1)(\beta_{1j} + \gamma_{1j}).$$

The Equation (2.6) also does not admit an explicit solution for θ.

First, we can derive an estimator of θ by substituting the Equations (2.12), (2.13) and (2.14) into the Equation (2.6), as follows;

$$\hat{\theta}_1 = \frac{-B_1 + C_1\hat{\mu}}{A_1}. \tag{2.16}$$

Second, we can derive an estimator of by substituting the Equations (2.7), (2.8), (2.9) and (2.10) into the Equation (2.6), as follows;

$$\hat{\theta}_2 = \frac{-B_2 + \sqrt{B_2^2 - 4sC_2}}{2s}, \tag{2.17}$$

where

$$B_2 = (a_1 - 1)\alpha_2 Y_{a_1:n} - (n - a_s)e^{\xi_{a_s}}(1 - \xi_{a_s})Y_{a_s:n} + \sum_{j=1}^{s} Y_{a_j:n} - \sum_{j=1}^{s} e^{\xi_{a_j}}(1 - \xi_{a_j})Y_{a_j:n}$$
$$+ \sum_{j=2}^{s}\left(a_j - a_{j-1} - 1\right)\left(\alpha_{2j}Y_{a_j:n} + \alpha_{3j}Y_{a_{j-1}:n}\right)$$
$$- \left\{(a_1 - 1)\alpha_2 - (n - a_s)e^{\xi_{a_s}}(1 - \xi_{a_s}) + s - \sum_{j=1}^{s} e^{\xi_{a_s}}(1 - \xi_{a_s}) + \sum_{j=2}^{s}(a_j - a_{j-1} - 1)(\alpha_{2j} + \alpha_{3j})\right\}\hat{\mu},$$

$$C_2 = (a_1 - 1)\beta_2(Y_{a_1:n} - \hat{\mu})^2 - (n - a_s)e^{\xi_{a_s}}(Y_{a_s:n} - \hat{\mu})^2 - \sum_{j=1}^{s} e^{\xi_{a_j}}(Y_{a_s:n} - \hat{\mu})^2$$
$$+ \sum_{j=2}^{s}\left(a_j - a_{j-1} - 1\right)\left\{\beta_{2j}\left(Y_{a_j:n} - \hat{\mu}\right)^2 + 2\gamma_{2j}\left(Y_{a_j:n} - \hat{\mu}\right)\left(Y_{a_{j-1}:n} - \hat{\mu}\right) - \gamma_{3j}\left(Y_{a_j:n} - \hat{\mu}\right)^2\right\}.$$

Since $\theta = 1/\lambda$ and $\mu = \ln\sigma$, we can obtain the AMLEs of the shape parameter λ and the scale parameter σ as follows;

$$\hat{\lambda}_i = \frac{1}{\hat{\theta}_i}, \quad i = 1, 2 \tag{2.18}$$

and

$$\hat{\sigma} = e^{\hat{\mu}}. \tag{2.19}$$

The mean squared errors and biases of the proposed AMLEs are simulated by Monte Carlo method(based on 10,000 Monte Carlo runs) for sample size $n = 20, 40$, and various choices of censoring under multiply Type-II censored sample. These values are given in Table 1.

3. Goodness-of-fit Tests

It is important that how well a sample of data agrees with a given distribution as its population. In this section, we consider some goodness-of-fit tests of the Weibull distribution based on multiply Type-II censored samples.

3.1. Modified empirical distribution function type tests

A well known EDF $F_n(x)$ is

$$F_n = \frac{\text{the number of } X_i' s \le x}{n}. \tag{3.1}$$

For complete samples under a simple hypothesis, the Kolmogorov-Smirnov(D), the Cramer-von Mises(W^2), and the Anderson-Darling(A^2) statistics are defined as

$$\begin{aligned} D^+ &= \sup_x \left[F_n(x) - F_0(x)\right], \\ D^- &= \sup_x \left[F_0(x) - F_n(x)\right], \\ D &= \max\left[D^+, D^-\right], \end{aligned} \tag{3.2}$$

$$W^2 = n \int_{-\infty}^{\infty} \left[F_n(x) - F_0(x)\right]^2 dF_0(x) \tag{3.3}$$

and

$$A^2 = n \int_{-\infty}^{\infty} \frac{\left[F_n(x) - F_0(x)\right]^2}{F_0(x)\left[1 - F_0(x)\right]} dF_0(x), \tag{3.4}$$

where $F_0(x)$ is the cdf assumed under H_0.

Recently, Kang and Lee (2006b) developed three EDF type modified Kolmogorov-Smirnov test, modified Anderson-Darling test, and modified Cramer-von Mises test for the two-parameter exponential distribution with unknown parameters based on multiply Type-II censored samples.

Now, we use three EDF type modified Kolmogorov-Smirnov test, modified Cramer-von Mises test, and modified Anderson-Darling test for multiply Type-II censored samples from Weibull distribution using the proposed AMLEs $\hat{\lambda}_k$ and $\hat{\sigma}$ as follows;

$$\begin{aligned} D_k^+ &= \max_{1 \le a_j \le s}\left[\frac{a_j}{s} - F\left(x_{a_j:n}; \hat{\lambda}_k, \hat{\sigma}\right)\right], \\ D_k^- &= \max_{1 \le a_j \le s}\left[F\left(x_{a_j:n}; \hat{\lambda}_k, \hat{\sigma}\right) - \frac{a_j}{s}\right], \\ D_k &= \max_{1 \le a_j \le s}\left[D_k^+, D_k^-\right], \end{aligned} \tag{3.5}$$

$$W_k^2 = \frac{1}{12s} + \sum_{j=1}^{s}\left[F\left(x_{a_j:n}; \hat{\lambda}_k, \hat{\sigma}\right) - \frac{2a_j - 1}{2s}\right]^2 \tag{3.6}$$

and

$$A_k^2 = -s - \frac{1}{s} + \sum_{j=1}^{s} \left(2a_j - 1\right)\left[\ln F\left(x_{a_j:n}; \hat{\lambda}_k, \hat{\sigma}\right) + \ln\left\{1 - F\left(x_{a_{s+1-j}:n}; \hat{\lambda}_k, \hat{\sigma}\right)\right\}\right]. \tag{3.7}$$

3.2. Modified normalized sample lorenz curve

The Lorenz curve is extensively used in the study of income distribution and used to be a powerful tool for the analysis of a variety of scientific problems.

Cho *et al.* (1999) proposed the transformed Lorenz curve that can be used in the study of symmetric distribution. The transformed Lorenz curve is defined by

$$\mathrm{TL}(r_i) = \frac{\sum_{j=1}^{i} X_{j:n}}{\sum_{j=1}^{n} X_{j:n}}, \quad r_i = \frac{i}{n},\ i = 1, 2, \ldots, n. \tag{3.8}$$

Kang and Cho (2001) proposed the normalized sample Lorenz curve(NSLC) using the Equation (3.8) for the complete sample as follows;

$$\mathrm{NSLC}(r_i) = \frac{\mathrm{TSL}(r_i)}{\mathrm{TSL}_F(r_i)}, \quad r_i = \frac{i}{n},\ i = 1, 2, \ldots, n, \tag{3.9}$$

where

$$\mathrm{TSL}(r_i) = \frac{\sum_{j=1}^{i} (X_{j:n} - X_{1:n})}{\sum_{j=1}^{n} (X_{j:n} - X_{1:n})} - r_i + 1,$$

$$\mathrm{TSL}_F(r_i) = \frac{\sum_{j=1}^{i} \left[F^{-1}(p_j) - F^{-1}(p_1)\right]}{\sum_{j=1}^{n} \left[F^{-1}(p_j) - F^{-1}(p_1)\right]} - r_i + 1.$$

Now, we propose modified NSLC based on multiply Type-II censored samples.

The modified NSLC based on multiply Type-II censored samples is given by

$$\mathrm{MNSLC}(r_i) = \frac{\mathrm{MTSL}(r_i)}{\mathrm{MTSL}_F(r_i)}, \quad r_i = \frac{a_i}{n},\ i = 1, 2, \ldots, s, \tag{3.10}$$

where

$$\mathrm{MTSL}(r_i) = \frac{\sum_{j=1}^{i} (X_{a_j:n} - X_{a_1:n})}{\sum_{j=1}^{s} (X_{a_j:n} - X_{a_1:n})} - r_i + 1,$$

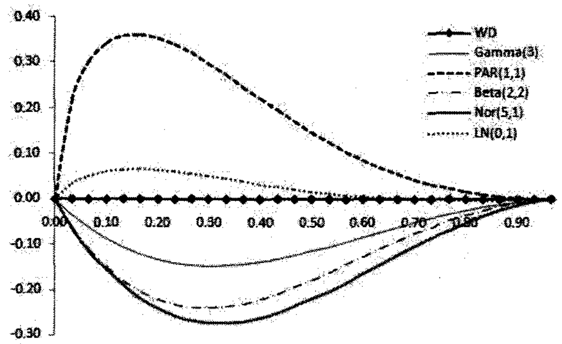

Figure 1: *Modified NSLC plot: Complete data (n=30)*

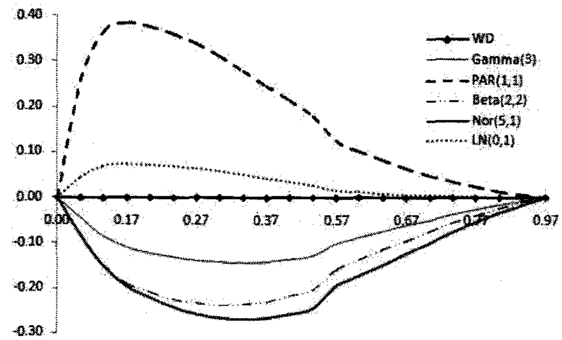

Figure 2: *Modified NSLC plot: Multiply Type-II censored data (n=30, a_j=1, 5~13, 17~25, 28~30)*

$$\mathrm{MTSL}_F(r_i) = \frac{\sum_{j=1}^{i}\left[F^{-1}\left(p_{a_j};\hat{\lambda}_k,\hat{\sigma}\right) - F^{-1}\left(p_{a_1};\hat{\lambda}_k,\hat{\sigma}\right)\right]}{\sum_{j=1}^{s}\left[F^{-1}\left(p_{a_j};\hat{\lambda}_k,\hat{\sigma}\right) - F^{-1}\left(p_{a_1};\hat{\lambda}_k,\hat{\sigma}\right)\right]} - r_i + 1.$$

We also propose the modified NSLC plot for multiply Type-II censored samples using ($1 - r_i$, $1 - \mathrm{MNSLC}(r_i)$). If data come from the Weibull distribution, the modified NSLC plot is $y = 0$ (see, Figure 1 and Figure 2). The value of $1 - \mathrm{MNSLC}_{i,k}$ increases and then decreases as $1 - r_i$ increases when the alternative distribution is Pareto and lognormal distribution. But the value of $1 - \mathrm{MNSLC}_{i,k}$ decreases and then increases as $1 - r_i$ increases when the alternative distributon is gamma, beta and normal distribution.

We propose new test statistics based on modified NSLC for the Weibull distribution under multiply Type-II censored samples. The new test statistic(TS) is defined by

$$\mathrm{TS}_k = \frac{\sum_{j=1}^{s}\left(sX_{a_j:n} - X_{a_1:n}\right)^2}{\sum_{j=1}^{s}\left[sF^{-1}\left(p_{a_j};\hat{\lambda}_k,\hat{\sigma}\right) - F^{-1}\left(p_{a_1};\hat{\lambda}_k,\hat{\sigma}\right)\right]^2}, \quad k = 1, 2. \tag{3.11}$$

Table 1: The relative mean squared errors(biases) for the estimators of the shape parameter λ and the scale parameter σ.

n	m	a_j	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\sigma}$
20	0	1~20	0.0575(0.1097)	0.0479(0.0749)	0.0547(−0.0053)
	1	1~19	0.0625(0.1078)	0.0541(0.0781)	0.0553(−0.0080)
		2~20	0.0623(0.1171)	0.0521(0.0829)	0.0552(−0.0025)
	2	1~18	0.0696(0.1110)	0.0613(0.0834)	0.0566(−0.0107)
		3~20	0.0675(0.1250)	0.0565(0.0904)	0.0557(0.0004)
		2~19	0.0680(0.1159)	0.0591(0.0870)	0.0557(−0.0056)
	3	1~17	0.0776(0.1172)	0.0690(0.0908)	0.0579(−0.0145)
		4~20	0.0742(0.1329)	0.0621(0.0977)	0.0565(0.0034)
		2~18	0.0761(0.1199)	0.0674(0.0934)	0.0569(−0.0086)
		3~19	0.0745(0.1245)	0.0649(0.0957)	0.0561(−0.0030)
	4	2~17	0.0857(0.1273)	0.0767(0.1022)	0.0582(−0.0127)
		4~19	0.0827(0.1332)	0.0722(0.1042)	0.0568(−0.0004)
		3~18	0.0840(0.1295)	0.0745(0.1033)	0.0572(−0.0063)
		2~4 7~14 16~20	0.0621(0.1159)	0.0466(0.0507)	0.0554(−0.0025)
	5	3~17	0.0954(0.1383)	0.0856(0.1135)	0.0584(−0.0107)
		4~18	0.0939(0.1394)	0.0836(0.1130)	0.0578(−0.0040)
		2~6 10~19	0.0680(0.1151)	0.0553(0.0670)	0.0559(−0.0050)
	6	4~17	0.1081(0.1501)	0.0974(0.1253)	0.0588(−0.0087)
40	0	1~40	0.0210(0.0509)	0.0189(0.0346)	0.0274(−0.0023)
	1	1~39	0.0217(0.0481)	0.0200(0.0341)	0.0276(−0.0028)
		2~40	0.0216(0.0524)	0.0194(0.0363)	0.0275(−0.0016)
	2	1~38	0.0229(0.0481)	0.0212(0.0350)	0.0278(−0.0035)
		3~40	0.0221(0.0538)	0.0199(0.0376)	0.0276(−0.0010)
		2~39	0.0224(0.0497)	0.0206(0.0359)	0.0276(−0.0022)
	3	1~37	0.0240(0.0487)	0.0224(0.0363)	0.0282(−0.0043)
		4~40	0.0228(0.0555)	0.0205(0.0394)	0.0277(−0.0003)
		2~38	0.0236(0.0497)	0.0219(0.0370)	0.0279(−0.0030)
		3~39	0.0230(0.0510)	0.0212(0.0373)	0.0277(−0.0017)
	4	2~37	0.0248(0.0504)	0.0232(0.0384)	0.0282(−0.0038)
		4~39	0.0237(0.0528)	0.0218(0.0391)	0.0278(−0.0010)
		3~38	0.0242(0.0511)	0.0225(0.0385)	0.0279(−0.0025)
	5	3~37	0.0255(0.0518)	0.0238(0.0400)	0.0283(−0.0033)
		4~38	0.0250(0.0529)	0.0233(0.0404)	0.0280(−0.0018)
		2~6 10~19 21~40	0.0216(0.0521)	0.0183(0.0188)	0.0275(−0.0017)
	15	6~10 16~25 31~40	0.0242(0.0577)	0.0203(0.0213)	0.0282(0.0014)
		6~25 31~35	0.0318(0.0590)	0.0287(0.0356)	0.0293(−0.0026)

Above procedure was repeated 10,000 times for the Weibull distribution based on multiply Type-II censored samples, sample size $n = 20, 40$. The 95^{th} percentile were found and these critical values ($\alpha = 0.05$) are given in Table 2. The probabilities of Type I error the four tests under the Weibull distribution (Wei(1, 1)) are given in Table 3.

4. The Simulated Results

From Table 1, $\hat{\lambda}_2$ is more efficient than $\hat{\lambda}_1$ in the sense of the MSE. As expected, the MSE of all estimators decreases as sample size n in creases. For fixed sample size, the MSE increases generally as m ($m = n - s$ was the number of unobserved or missing data) increases. The biases of $\hat{\lambda}_2$ are smaller than biases of $\hat{\lambda}_1$.

The powers of four tests with significance level 0.05 for the Weibull distribution based on multiply Type-II censored samples are investigated under 4 alternative distributions. These values are given in Table 4~7.

Table 2: Critical values ($\alpha = 0.05$)

n	m	a_j	D_1	D_2	A_1^2	A_2^2	W_1^2	W_2^2	TS_1	TS_2
20	0	1~20	0.179	0.173	0.795	0.751	0.129	0.124	11.864	11.875
	1	2~20	0.208	0.204	4.810	4.691	0.144	0.139	11.924	11.951
	2	1~18	0.238	0.233	1.000	0.979	0.188	0.181	6.865	6.868
	5	2~6 10~19	0.409	0.414	8.623	8.517	0.792	0.790	8.768	8.830
	6	4~17	0.449	0.447	10.883	10.840	0.969	0.951	6.185	6.182
40	0	1~40	0.128	0.126	0.759	0.738	0.123	0.122	10.126	10.131
	1	2~40	0.140	0.138	4.734	4.683	0.131	0.129	10.124	10.132
	2	1~38	0.154	0.152	0.900	0.901	0.155	0.153	6.889	6.890
	12	6~10 16~25 31~40	0.643	0.651	34.308	33.808	4.077	4.094	11.973	12.032
	15	6~25 31~35	0.567	0.572	19.309	19.194	2.632	2.598	5.557	5.542

Table 3: The probabilities of Type I error for four tests under the Weibull distribution.

n	m	a_j	D_1	D_2	A_1^2	A_2^2	W_1^2	W_2^2	TS_1	TS_2
20	0	1~20	0.048	0.050	0.050	0.049	0.049	0.049	0.046	0.046
	1	2~20	0.050	0.049	0.047	0.046	0.047	0.047	0.046	0.046
	2	1~18	0.048	0.049	0.049	0.049	0.047	0.048	0.053	0.053
	5	2~6 10~19	0.049	0.051	0.054	0.053	0.044	0.044	0.051	0.051
	6	4~17	0.051	0.053	0.045	0.046	0.047	0.046	0.053	0.053
40	0	1~40	0.053	0.053	0.055	0.054	0.055	0.053	0.045	0.045
	1	2~40	0.055	0.055	0.056	0.057	0.055	0.056	0.045	0.045
	2	1~38	0.055	0.057	0.056	0.056	0.056	0.058	0.047	0.047
	12	6~10 16~25 31~40	0.045	0.049	0.058	0.054	0.049	0.046	0.043	0.043
	15	6~25 31~35	0.052	0.051	0.061	0.063	0.062	0.059	0.045	0.044

Table 4: The powers of the tests for gamma distribution(GAM(3))

n	m	a_j	D_1	D_2	A_1^2	A_2^2	W_1^2	W_2^2	TS_1	TS_2
20	0	1~20	0.711	0.751	0.742	0.772	0.760	0.767	1.000	1.000
	1	2~20	0.791	0.814	0.390	0.480	0.968	0.968	1.000	1.000
	2	1~18	0.662	0.703	0.973	0.977	0.870	0.889	1.000	1.000
	5	2~6 10~19	0.898	0.849	0.334	0.367	0.999	0.999	1.000	1.000
	6	4~17	0.766	0.761	0.529	0.554	0.347	0.403	1.000	1.000
40	0	1~40	0.938	0.944	0.997	0.998	0.982	0.982	1.000	1.000
	1	2~40	0.959	0.964	1.000	1.000	1.000	1.000	1.000	1.000
	2	1~38	0.960	0.963	1.000	1.000	1.000	1.000	1.000	1.000
	12	6~10 16~25 31~40	0.733	0.560	0.169	0.330	1.000	1.000	1.000	1.000
	15	6~25 31~35	0.897	0.834	0.592	0.665	0.641	0.727	1.000	1.000

Table 5: The powers of the tests for lognormal distribution(LN(0,1))

n	m	a_j	D_1	D_2	A_1^2	A_2^2	W_1^2	W_2^2	TS_1	TS_2
20	0	1~20	0.952	0.966	0.980	0.985	0.977	0.979	0.800	0.808
	1	2~20	0.975	0.981	0.780	0.837	1.000	1.000	0.795	0.801
	2	1~18	0.883	0.906	1.000	1.000	0.992	0.994	0.654	0.669
	5	2~6 10~19	0.983	0.968	0.582	0.654	1.000	1.000	0.736	0.740
	6	4~17	0.877	0.879	0.665	0.698	0.682	0.728	0.588	0.599
40	0	1~40	1.000	1.000	1.000	1.000	1.000	1.000	0.954	0.958
	1	2~40	1.000	1.000	0.995	0.997	1.000	1.000	0.953	0.957
	2	1~38	1.000	1.000	1.000	1.000	1.000	1.000	0.929	0.935
	12	6~10 16~25 31~40	0.943	0.916	0.789	0.885	1.000	1.000	0.934	0.931
	15	6~25 31~35	0.943	0.904	0.892	0.933	0.956	0.975	0.841	0.854

Table 6: The powers of the tests for normal distribution(N(5,1))

n	m	a_j	D_1	D_2	A_1^2	A_2^2	W_1^2	W_2^2	TS_1	TS_2
20	0	1~20	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	1	2~20	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	2	1~18	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	5	2~6 10~19	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	6	4~17	1.000	1.000	1.000	1.000	0.925	0.945	1.000	1.000
40	0	1~40	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	1	2~40	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	2	1~38	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	12	6~10 16~25 31~40	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	15	6~25 31~35	1.000	1.000	1.000	1.000	0.993	0.997	1.000	1.000

Table 7: The powers of the tests for Pareto distribution(PAR(1,1))

n	m	a_j	D_1	D_2	A_1^2	A_2^2	W_1^2	W_2^2	TS_1	TS_2
20	0	1~20	0.996	0.997	0.999	0.999	0.999	0.999	0.998	0.999
	1	2~20	0.996	0.997	0.979	0.986	1.000	1.000	0.999	0.999
	2	1~18	0.957	0.965	1.000	1.000	0.999	1.000	0.999	0.999
	5	2~6 10~19	0.992	0.988	0.909	0.929	1.000	1.000	0.999	0.999
	6	4~17	0.912	0.918	0.876	0.893	0.884	0.904	0.999	0.999
40	0	1~40	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	1	2~40	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	2	1~38	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	12	6~10 16~25 31~40	0.999	0.999	0.996	0.998	1.000	1.000	1.000	1.000
	15	6~25 31~35	0.962	0.951	0.998	0.999	0.996	0.998	1.000	1.000

The modified EDF type tests that use the estimator $\hat{\lambda}_2$ is generally more powerful than the tests that use the estimator $\hat{\lambda}_1$ when complete data. But the modified Kolmogorov-Smirnov and Cramer-von Mises tests that use the estimator $\hat{\lambda}_1$ is generally more powerful than the tests that use the estimator $\hat{\lambda}_2$ when censored data.

For the gamma and Pareto alternative distribution, the TS is more powerful than the modified EDF type test. But the modified EDF type test more powerful than the TS when the alternative distribution is lognormal distribution.

For normal alternative, all the tests show good performance.

The power generally decrease as m increase or sample size n decrease.

5. Illustrative Example

In this section, we present an example to illustrate the inference procedures discussed in this paper.

The data given here arose in tests on the endurance of deep groove ball bearings. They were originally discussed by Lieblein and Zelen (1956), who assumed that the data came from a Weibull distribution. The data are the number of million revolutions before failure for each of the 23 ball bearings in the life test.

17.88, 28.92, 33.00, 41.52, 42.12, 45.60, 48.40, 51.84, 51.96, 54.12, 55.56, 67.80, 68.64, 68.64, 68.88, 84.12, 93.12, 98.64, 105.12, 105.84, 127.92, 128.04, 173.40.

For complete data, we can obtain the AMLEs $\hat{\lambda}_1 = 2.187829$, $\hat{\lambda}_2 = 2.121929$ and $\hat{\sigma} = 81.450162$. For this example of $n = 23$, $s = 17$ ($a_j = 1, 2, 6\sim9, 13\sim23$), and the multiply Type-II censored samples are

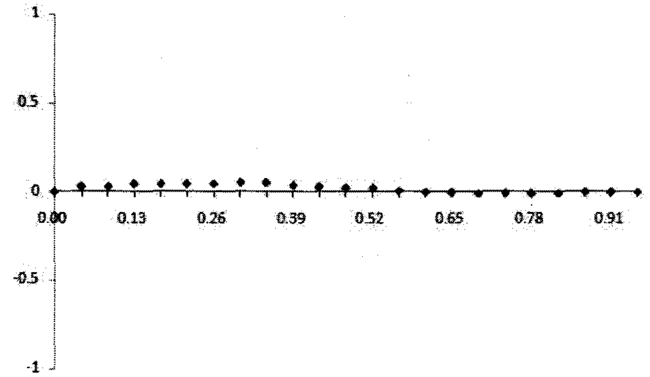

Figure 3: *Modified NSLC plot : Complete data*

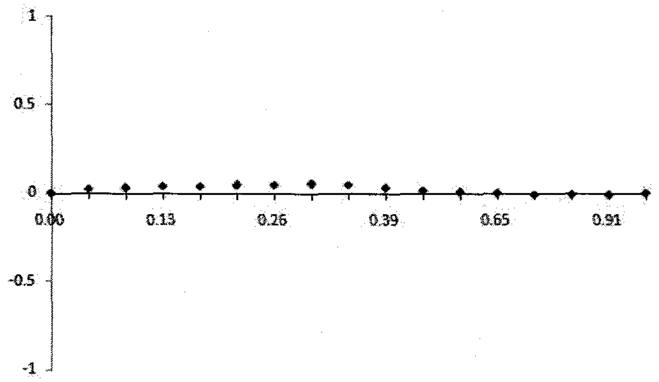

Figure 4: *Modified NSLC plot : Multiply Type-II censored data*

17.88, 28.92, -, -, -, 45.60, 48.40, 51.84, 51.96, -, -, -, 68.64, 68.64, 68.88, 84.12, 93.12, 98.64, 105.12, 105.84, 127.92, 128.04, 173.40.

We can obtain the AMLEs $\hat{\lambda}_1 = 2.179173$, $\hat{\lambda}_2 = 2.108290$ and $\hat{\sigma} = 81.485143$.

We can picture the proposed plots for multiply Type-II censored samples using the AMLEs $\hat{\lambda}_2$ and $\hat{\sigma}$ (see Figure 3 and Figure 4). It is easy to see that the modified NSLC plot is good performance when complete data or multiply Type-II censored samples.

We also compute critical value $(\alpha = 0.05)$ for compete data and the multiply Type-II censored sample.

	D_1	D_2	A_1^2	A_2^2	W_1^2	W_2^2	TS_1	TS_2
Critical Value								
Compete data:	0.168	0.163	0.783	0.740	0.126	0.123	1.121	1.107
Censored sample:	0.431	0.444	13.492	13.233	1.091	1.104	1.138	1.110
Test Statistics								
Compete data:	0.152	0.148	0.345	0.325	0.060	0.056	1.085	1.077
Censored sample:	0.382	0.378	11.808	11.748	0.990	0.994	1.109	1.095

So, test statistics can be computed by (3.5), (3.6), (3.7) and (3.10) which provides strong evidence that the real data set is from a Weibull distribution(Accept H_0).

We will need further study on the various graphical methods or test statistics for testing the distributions based on multiply Type-II censored samples.

References

Aho, M., Bain, L. J. and Engelhardt, M. (1985). Goodness-of-fit tests for the Weibull distribution with unknown parameters and heavy censoring, *Journal of Statistical Computation and Simulation,* **21**, 213–225.

Balakrishnan, N., Gupta, S. S. and Panchapakesan, S. (1995). Estimation of the location and scale parameters of the extreme value distribution based on multiply Type-II censored samples, *Communications in Statistics-Theory and Methods,* **24**, 2105–2125.

Balakrishnan, N., Kannan, N., Lin, C. T. and Wu, S. J. S. (2004). Inference for the extreme value distribution under progressive Type-II censoring, *Journal of Statistical Computation and Simulation,* **74**, 25–45.

Cho, Y. S., Lee, J. Y. and Kang, S. B. (1999). A study on distribution based on the transformed Lorenz curve, *The Korean Journal of Applied Statistics,* **12**, 153–163.

Gibson, E. W. B. and Higgins, J. J. (2000). Gap-ratio goodness of fit tests for Weibull or extreme value distribution assumptions with left or right censored data, *Communications in Statistics-Simulation and Computation,* **29**, 541–557.

Han, J. T. and Kang, S. B. (2006). Estimation for two-parameter Rayleigh distribution based on multiply Type-II censored sample, *Journal of the Korean Data & Information Science Society,* **17**, 1319–1328.

Kang, S. B. and Cho, Y. S. (2001). A study on distribution based on the normalized sample Lorenz curve, *The Korean Communications in Statistics,* **8**, 185–192.

Kang, S. B. and Lee, H. J. (2006a). Goodness-of-fit tests for the Weibull distribution based on the sample entropy, *Journal of the Korean Data & Information Science Society,* **17**, 259–268.

Kang, S. B. and Lee, S. K. (2006b). Test for the exponential distribution based on multiply Type-II censored samples, *The Korean Communications in Statistics,* **13**, 537–550.

Lieblein, J. and Zelen, M. (1956). Statistical investigation of the fatigue life of deep-groove ball bearings, *Journal of Research of the National Bureau of Standards,* **57**, 273–316.

Ng, H. K. T., Chan, P. S. and Balakrishnan, N. (2004). Optimal progressive censoring plan for the Weibull distribution, *Technometrics,* **46**, 470–481.

Sanjel, D. and Balakrishnan, N. (2008). A Laguerre polynomial approximation for a goodness-of-fit test for exponential distribution based on progressively censored data, *Journal of Statistical Computation and Simulation,* **78**, 503–513.

Wang, B. (2008). Goodness-of-fit test for the exponential distribution based on progressively Type-II censored sample, *Journal of Statistical Computation & Simulation,* **78**, 125–132.

Wu, J. W. and Yang, C. C. (2002). Weighted moments estimation of the scale parameter of the exponential distribution based on a multiply Type -II censored sample, *Quality and Reliability Engineering International,* **18**, 149–154.

Received December 2008; Accepted January 2009