

Goodness-of-fit test for the logistic distribution based on multiply type-II censored samples

Suk-Bok Kang¹ · Jun-Tae Han² · Young-Seuk Cho³

¹Department of Statistics, Yeungnam University

²Research Management Team, Korea Student Aid Foundation

³Department of Statistics, Pusan National University

Received 20 November 2013, revised 3 December 2013, accepted 9 December 2013

Abstract

In this paper, we derive the estimators of the location parameter and the scale parameter in a logistic distribution based on multiply type-II censored samples by the approximate maximum likelihood estimation method. We use four modified empirical distribution function (EDF) types test for the logistic distribution based on multiply type-II censored samples using proposed approximate maximum likelihood estimators. We also propose the modified normalized sample Lorenz curve plot for the logistic distribution based on multiply type-II censored samples. For each test, Monte Carlo techniques are used to generate the critical values. The powers of these tests are also investigated under several alternative distributions.

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

1. Introduction

The probability density function (pdf) and the cumulative distribution function (cdf) of the logistic distribution are given by

$$f(x; \sigma, \theta) = \frac{\pi}{\sigma\sqrt{3}} \frac{\exp\left(-\frac{\pi(x-\theta)}{\sigma\sqrt{3}}\right)}{\left[1 + \exp\left(-\frac{\pi(x-\theta)}{\sigma\sqrt{3}}\right)\right]^2}, \quad -\infty < x < \infty, \quad \sigma > 0, \quad (1.1)$$

and

$$F(x; \sigma, \theta) = \frac{1}{1 + \exp\left(-\frac{\pi(x-\theta)}{\sigma\sqrt{3}}\right)}, \quad -\infty < x < \infty, \quad (1.2)$$

where σ and θ are scale and location parameters respectively.

¹ Professor, Department of Statistics, Yeungnam University, Gyeongsan 712-749, Korea

² Corresponding author: Manager, Research Management Team, Korea Student Aid Foundation, Seoul 100-753, Korea. E-mail: hanjt@kosaf.go.kr

³ Professor, Department of Statistics, Pusan National University, Busan 609-735, Korea

The logistic distribution has been one of the most important statistical distributions because of its simplicity. It has been used for growth models and as a substitute for the normal distribution.

Order statistics from the logistic distribution and their moments were first examined by Birnbaum and Dudman (1963). Mathai (2003) studied some general moments of order statistic in the logistic distribution. Mathai (2003) also evaluated survival function and reliability in the logistic distribution.

The most common censoring schemes are type-I and type-II censoring, but the conventional type-I and type-II censoring schemes do not have flexibility. Multiply type-II censoring is a generalization of type-II censoring. 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 like-test. Another situation where multiply censored samples arise naturally is when some units failed between two points of observation with the exact times of failure of these units unobserved.

The approximate maximum likelihood estimating method was first developed by Balakrishnan (1989) for the purpose of providing explicit estimators of the scale parameter in the Rayleigh distribution. It has been noted that in most cases, the maximum likelihood method does not provide explicit estimators based on censored samples. When the sample is multiply censored, the maximum likelihood method does not admit explicit solutions. Therefore, it is desirable to develop which approximations to this maximum likelihood method would provide us with estimators that are explicit functions of order statistics. Balakrishnan *et al.* (1995) obtained the best linear unbiased estimators (BLUEs), the maximum likelihood estimators and the approximate maximum likelihood estimators (AMLEs) in a logistic distribution based on multiply type-II censored samples. Fei *et al.* (1995) studied the estimation for the two-parameter Weibull distribution and extreme-value distribution under multiply type-II censoring. They compared the mean squared errors of the maximum likelihood estimators, AMLEs, and BLUEs of the parameters in the extreme value distribution. Maswadah (2003) derived the conditional confidence intervals for the parameters based on the generalized order statistics. Balakrishnan *et al.* (2004) discussed point and interval estimation for the extreme value distribution under progressively type-II censoring. Balakrishnan and Kateri (2008) proposed a simple graphical solution for the determination of the maximum likelihood estimator of the shape parameter in the Weibull distribution. Kang *et al.* (2008) derived the AMLEs of the scale parameter in the half-logistic distribution under progressive type-II censoring. They also proposed the estimators of the reliability function by using the proposed estimators of the parameters. Han and Kang (2008) derived the AMLEs of the scale parameter and the location parameter in a double Rayleigh distribution based on multiply type-II censored samples. Recently, Shin and Lee (2012) suggested an estimation method of the parameter in an exponential distribution based on a progressive type-I interval censored sample with semi-missing observation. Kim *et al.* (2011a, 2011b) suggested Bayesian estimations under type-I hybrid and progressively type-II censoring.

Porter III *et al.* (1992) developed three modified Kolmogorov-Smirnov, Anderson-Darling, and Cramer-von Mises tests for the Pareto distribution with unknown parameters of the location and scale and known shape parameter based on the complete samples. Puig and Stephens (2000) studied some tests of fit for the Laplace distribution based on the empirical distribution function (EDF) statistics and the application of the Laplace distribution in the least absolute deviations regression. Choulakian and Stephens (2001) studied the estimation

of parameters and goodness-of-fit tests for the generalized Pareto distribution.

In this paper, we derive the AMLEs of the location parameter θ and the scale parameter σ in a logistic distribution under multiply type-II censored sample. We use four modified EDF type tests, including the modified Kolmogorov-Smirnov test, the modified Anderson-Darling test, the modified Cramer-von Mises test, and the modified Watson test for the logistic distribution with unknown parameters based on multiply type-II censored samples using the proposed estimators and EDF. We propose the modified normalized sample Lorenz curve (NSLC) plot to test for the logistic distribution based on multiply type-II censored samples.

For each test, Monte Carlo techniques are used to generate the critical values. The powers of these tests are also investigated under lognormal, exponential, Rayleigh, double Rayleigh alternative distributions.

2. Approximate maximum likelihood estimators

We assume that n items are put on a life test, but only a_1 th, a_2 th, ..., a_s th failures are observed, the rest are unobserved or missing, where a_1, a_2, \dots, a_s are considered to be fixed.

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

$$X_{a_1:n} \leq X_{a_2:n} \leq \cdots \leq X_{a_s:n}, \quad (2.1)$$

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

The likelihood function for the logistic distribution based on the multiply type-II censored sample (2.1) is given by

$$L = \frac{n!}{\prod_{j=1}^{s+1} (a_j - a_{j-1} - 1)!} \prod_{j=1}^{s+1} [F(z_{a_j:n}) - F(z_{a_{j-1}:n})]^{a_j - a_{j-1} - 1} \left(\frac{\pi}{\sigma\sqrt{3}} \right)^s \prod_{j=1}^s f(z_{a_j:n}), \quad (2.2)$$

where $Z_{i:n} = \pi(X_{i:n} - \theta)/(\sigma\sqrt{3})$, and $f(z)$ and $F(z)$ are the pdf and the cdf of the standard logistic distribution, respectively.

Since $f(z) = F(z)\{1 - F(z)\}$, we can obtain the likelihood equations as follows;

$$\begin{aligned} \frac{\partial \ln L}{\partial \theta} &= -\frac{\pi}{\sigma\sqrt{3}} \left[s + (a_1 - 1) - (a_1 - 1)F(z_{a_1:n}) - (n - a_s)F(z_{a_s:n}) - 2 \sum_{j=1}^s F(z_{a_1:n}) \right. \\ &\quad \left. + \sum_{j=2}^s (a_j - a_{j-1} - 1) \frac{f(z_{a_j:n}) - f(z_{a_{j-1}:n})}{F(z_{a_j:n}) - F(z_{a_{j-1}:n})} \right] \\ &= 0 \end{aligned} \quad (2.3)$$

and

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{1}{\sigma} \left[s + (a_1 - 1)z_{a_1:n} - (a_1 - 1)F(z_{a_1:n})z_{a_1:n} \right]$$

$$\begin{aligned}
& - (n - a_s)F(z_{a_s:n})z_{a_s:n} + \sum_{j=1}^s z_{a_j:n} - 2 \sum_{j=1}^s F(z_{a_1:n})z_{a_j:n} \\
& + \sum_{j=2}^s (a_j - a_{j-1} - 1) \frac{f(z_{a_j:n})z_{a_j:n} - f(z_{a_{j-1}:n})z_{a_{j-1}:n}}{F(z_{a_j:n}) - F(z_{a_{j-1}:n})} \\
& = 0.
\end{aligned} \tag{2.4}$$

Since the likelihood equations are very complicated, the equations (2.4) and (2.5) do not admit explicit solutions for μ and θ , respectively.

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

First, the equation (2.4) does not admit an explicit solution for θ . But we can expand the following function by Taylor series expansion as follows;

$$F(z_{a_j:n}) \simeq \kappa_{1j} + \delta_{1j}z_{a_j:n} \tag{2.5}$$

and

$$\frac{f(z_{a_j:n}) - f(z_{a_{j-1}:n})}{F(z_{a_j:n}) - F(z_{a_{j-1}:n})} \simeq \alpha_{1j} + \beta_{1j}z_{a_j:n} + \gamma_{1j}z_{a_{j-1}:n}, \tag{2.6}$$

where

$$\begin{aligned}
\kappa_{1j} &= p_{a_j}q_{a_j}\xi_{a_j}, \quad \delta_{1j} = p_{a_j}q_{a_j}, \\
\alpha_{1j} &= \frac{(1+K_j)(p_{a_j}q_{a_j} - p_{a_{j-1}}q_{a_{j-1}})}{p_{a_j} - p_{a_{j-1}}} - \frac{\xi_{a_j}p_{a_j}q_{a_j}(1-2p_{a_j}) - \xi_{a_{j-1}}p_{a_{j-1}}q_{a_{j-1}}(1-2p_{a_{j-1}})}{p_{a_j} - p_{a_{j-1}}}, \\
\beta_{1j} &= \frac{1}{p_{a_j} - p_{a_{j-1}}} \left[p_{a_j}q_{a_j}(1-2p_{a_j}) - \frac{(p_{a_j}q_{a_j})^2}{p_{a_j} - p_{a_{j-1}}} + \frac{p_{a_j}q_{a_j}p_{a_{j-1}}q_{a_{j-1}}}{p_{a_j} - p_{a_{j-1}}} \right], \\
\gamma_{1j} &= \frac{1}{p_{a_j} - p_{a_{j-1}}} \left[\frac{p_{a_{j-1}}q_{a_{j-1}}(p_{a_j}q_{a_j} - p_{a_{j-1}}q_{a_{j-1}})}{p_{a_j} - p_{a_{j-1}}} - p_{a_{j-1}}q_{a_{j-1}}(1-2p_{a_{j-1}}) \right], \\
K_j &= \frac{f(\xi_{a_j})\xi_{a_j} - f(\xi_{a_{j-1}})\xi_{a_{j-1}}}{p_{a_j} - p_{a_{j-1}}}.
\end{aligned}$$

Next, we can approximate the following function by

$$F(z_{a_j:n})z_{a_j:n} \simeq \kappa_{2j} + \delta_{2j}z_{a_j:n} \tag{2.7}$$

and

$$\frac{f(z_{a_j:n})z_{a_j:n} - f(z_{a_{j-1}:n})z_{a_{j-1}:n}}{F(z_{a_j:n}) - F(z_{a_{j-1}:n})} \simeq \alpha_{2j} + \beta_{2j}z_{a_j:n} + \gamma_{2j}z_{a_{j-1}:n}, \tag{2.8}$$

where

$$\begin{aligned}
\kappa_{2j} &= -\xi_{a_j}^2 p_{a_j}q_{a_j}, \quad \delta_{2j} = \xi_{a_j}p_{a_j}q_{a_j} + p_{a_j}, \\
\alpha_{2j} &= K_j^2 - \frac{\xi_{a_j}p_{a_j}q_{a_j}(1-2p_{a_j}) - \xi_{a_{j-1}}p_{a_{j-1}}q_{a_{j-1}}(1-2p_{a_{j-1}})}{p_{a_j} - p_{a_{j-1}}}, \\
\beta_{2j} &= \frac{(1-K_j)p_{a_j}q_{a_j} + \xi_{a_j}p_{a_j}q_{a_j}(1-2p_{a_j})}{p_{a_j} - p_{a_{j-1}}}, \\
\gamma_{2j} &= -\frac{(1-K_j)p_{a_{j-1}}q_{a_{j-1}} + \xi_{a_{j-1}}p_{a_{j-1}}q_{a_{j-1}}(1-2p_{a_{j-1}})}{p_{a_j} - p_{a_{j-1}}}.
\end{aligned}$$

By substituting the equations (2.5) and (2.6) into the equation (2.4), and the equations (2.7) and (2.8) into the equation (2.5), we can derive an estimator of θ as follows;

$$\hat{\theta}_1 = \frac{B_0 A_1 - A_0 C_1}{C_0 A_1 - A_0 B_1}, \quad (2.9)$$

where

$$\begin{aligned} A_0 &= s + (a_1 - 1) - (a_1 - 1)\kappa_{11} - (n - a_s)\kappa_{1s} - 2 \sum_{j=1}^s \kappa_{1j} + \sum_{j=2}^s (a_j - a_{j-1} - 1)\alpha_{1j}, \\ B_0 &= -(a_1 - 1)\delta_{11} X_{a_1:n} - (n - a_s)\delta_{1s} X_{a_s:n} - 2 \sum_{j=1}^s \delta_{1j} X_{a_j:n} \\ &\quad + \sum_{j=2}^s (a_j - a_{j-1} - 1)(\beta_{1j} X_{a_j:n} + \gamma_{1j} X_{a_{j-1}:n}), \\ C_0 &= -(a_1 - 1)\delta_{11} - (n - a_s)\delta_{1s} - 2 \sum_{j=1}^s \delta_{1j} + \sum_{j=2}^s (a_j - a_{j-1} - 1)(\beta_{1j} + \gamma_{1j}), \\ A_1 &= s - (a_1 - 1)\kappa_{21} - (n - a_s)\kappa_{2s} - 2 \sum_{j=1}^s \kappa_{2j} + \sum_{j=2}^s (a_j - a_{j-1} - 1)\alpha_{2j}, \\ B_1 &= (a_1 - 1)(1 - \delta_{21}) - (n - a_s)\delta_{2s} + s - 2 \sum_{j=1}^s \delta_{2j} + \sum_{j=2}^s (a_j - a_{j-1} - 1)(\beta_{2j} + \gamma_{2j}), \\ C_1 &= (a_1 - 1)(1 - \delta_{21}) X_{a_1:n} - (n - a_s)\delta_{2s} X_{a_s:n} + \sum_{j=1}^s X_{a_j:n} - 2 \sum_{j=1}^s \delta_{2j} X_{a_j:n} \\ &\quad + \sum_{j=2}^s (a_j - a_{j-1} - 1)(\beta_{2j} X_{a_j:n} + \gamma_{2j} X_{a_{j-1}:n}). \end{aligned}$$

By substituting the equations (2.5) and (2.8) into the equation (2.5), we can derive an estimator of σ as follows;

$$\hat{\sigma}_1 = \frac{\pi}{\sqrt{3}} \left(\frac{-B_2 + \sqrt{B_2^2 - 4A_2 C_2}}{2A_2} \right), \quad (2.10)$$

where

$$\begin{aligned} A_2 &= s + \sum_{j=2}^s (a_j - a_{j-1} - 1)\alpha_{2j}, \\ B_2 &= (a_1 - 1)(X_{a_1:n} - \hat{\theta})(1 - \kappa_{11}) - (n - a_s)\kappa_{11}(X_{a_s:n} - \hat{\theta}) + \sum_{j=1}^s (X_{a_j:n} - \hat{\theta}) \\ &\quad - 2 \sum_{j=1}^s \kappa_{1j}(X_{a_j:n} - \hat{\theta}) + \sum_{j=2}^s (a_j - a_{j-1} - 1)\{\beta_{2j}(X_{a_j:n} - \hat{\theta}) + \gamma_{2j}(X_{a_{j-1}:n} - \hat{\theta})\}, \\ C_2 &= -(a_1 - 1)\delta_{11}(X_{a_1:n} - \hat{\theta})^2 - (n - a_s)\delta_{11}(X_{a_s:n} - \hat{\theta})^2 - 2 \sum_{j=1}^s \delta_{1j}(X_{a_j:n} - \hat{\theta})^2. \end{aligned}$$

It should be mentioned that upon solving equation (2.5) for σ , we obtain a quadratic equation in σ that has two roots. However, one of them drops out, since $\delta_{1j} > 0$.

Balakrishnan *et al.* (1995) derived an AMLE of the location parameter θ and the scale parameter σ as follows;

$$\hat{\theta}_2 = B_3 - \frac{\pi}{\sqrt{3}} \hat{\sigma}_2 C_3 \quad (2.11)$$

and

$$\hat{\sigma}_2 = \frac{\pi}{\sqrt{3}} \left(\frac{-D + \sqrt{D^2 + 4sE}}{2s} \right), \quad (2.12)$$

where

$$\begin{aligned} B_3 &= \frac{1}{g} \left[(a_1 - 1)\delta_{11}X_{a_1:n} + (n - a_s)\delta_{1s}X_{a_s:n} + 2 \sum_{j=1}^s \delta_{1j}X_{a_j:n} \right. \\ &\quad \left. + \sum_{j=2}^s (a_j - a_{j-1} - 1)(\delta_{1j}X_{a_j:n} + \delta_{1j-1}X_{a_{j-1}:n}) \right], \\ C_3 &= \frac{1}{g} \left[s + (a_1 - 1)(1 - \kappa_{11}) - (n - a_s)\kappa_{1s} + 2 \sum_{j=1}^s \kappa_{1j} + \sum_{j=2}^s (a_j - a_{j-1} - 1)L_j \right], \\ D &= (a_1 - 1)(1 - \kappa_{11})X_{a_1:n} - (n - a_s)\kappa_{1s}X_{a_s:n} + 2 \sum_{j=1}^s (1 - 2\kappa_{1j})X_{a_j:n} \\ &\quad + \sum_{j=2}^s (a_j - a_{j-1} - 1)L_j^* - gB_3C_3, \\ E &= (a_1 - 1)\delta_{11}(X_{a_1:n} - B_3)^2 + (n - a_s)\delta_{1s}(X_{a_s:n} - B_3)^2 + 2 \sum_{j=1}^s \delta_{1j}(X_{a_j:n} - B_3)^2, \\ &\quad + \sum_{j=2}^s (a_j - a_{j-1} - 1) \frac{p_{a_j}q_{a_j}p_{a_{j-1}}q_{a_{j-1}}}{(p_{a_j} - p_{a_{j-1}})^2} (X_{a_j:n} - X_{a_{j-1}:n})^2, \\ g &= (a_1 - 1)\delta_{11} + (n - a_s)\delta_{1s} + 2 \sum_{j=1}^s \delta_{1j} + \sum_{j=2}^s (a_j - a_{j-1} - 1)(\delta_{1j-1} - \delta_{1j}), \\ L_j &= \delta_{1j} [(p_{a_j} - p_{a_{j-1}})^{-1} + \xi_{a_j}] - \delta_{1j-1} [(p_{a_j} - p_{a_{j-1}})^{-1} + \xi_{a_{j-1}}], \\ L_j^* &= \delta_{1j} [(p_{a_j} - p_{a_{j-1}})^{-1} + \xi_{a_j}] X_{a_j:n} - \delta_{1j-1} [(p_{a_j} - p_{a_{j-1}})^{-1} + \xi_{a_{j-1}}] X_{a_{j-1}:n}. \end{aligned}$$

3. Goodness of fit tests

In this section, we consider some goodness-of-fit tests of the logistic distribution based on multiply type-II censored samples.

3.1. Modified empirical distribution function type tests

A well known EDF $F_n(x)$ is

$$F_n(x) = \frac{\text{the number of } X' s \leq x}{n}. \quad (3.1)$$

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

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

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

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

and

$$U^2 = W^2 - n \left(\bar{F}_n(x) - \frac{1}{2} \right)^2, \quad (3.5)$$

where $F_0(x)$ is the cdf assumed under H_0 and $\bar{F}_n(x)$ is the mean of the $F_n(x)$.

In the form given above, test statistics can only be used with complete samples, i.e. no censoring. Modification of the test statistics for censored samples and for composite hypothesis H_0 with unspecified parameters has been studied by Pettitt and Stephens (1976). Kang and Lee (2006) developed three EDF type modified Kolmogorov-Smirnov test, modified Anderson-Darling test, and modified Cramer-von Mises test for the two-parameter exponential distribution based on multiply type-II censored samples.

Now, we use three EDF type modified Kolmogorov-Smirnov test, modified Anderson-Darling test, and modified Cramer-von Mises test for multiply type-II censored samples from the logistic distribution using the proposed estimators $\hat{\sigma}_k$ and $\hat{\theta}_k$, $k = 1, 2$ as follows;

$$\begin{aligned} D_k^+ &= \max_{1 \leq a_j \leq s} \left[\frac{a_j}{s} - F(x_{a_j:n}; \hat{\sigma}_k, \hat{\theta}_k) \right], \\ D_k^- &= \max_{1 \leq a_j \leq s} \left[F(x_{a_j:n}; \hat{\sigma}_k, \hat{\theta}_k) - \frac{a_j}{s} \right], \\ D_k &= \max_{1 \leq a_j \leq s} [D_k^+, D_k^-], \end{aligned} \quad (3.6)$$

$$W_k^2 = \frac{1}{12s} + \sum_{j=1}^s \left[F(x_{a_j:n}; \hat{\sigma}_k, \hat{\theta}_k) - \frac{2a_j - 1}{2s} \right]^2, \quad (3.7)$$

and

$$A_k^2 = -s - \frac{1}{s} + \sum_{j=1}^s (2a_j - 1) \left[\ln F(x_{a_j:n}; \hat{\sigma}_k, \hat{\theta}_k) + \ln \{1 - F(x_{a_{s+1-j}:n}; \hat{\sigma}_k, \hat{\theta}_k)\} \right]. \quad (3.8)$$

We also obtained the modified Watson test for multiply type-II censored samples from the logistic distribution using the proposed estimators $\hat{\sigma}_k$ and $\hat{\theta}_k$ as follows;

$$U_k^2 = W_k^2 - s \left(\frac{1}{s} \sum_{j=1}^s F(x_{a_j:n}; \hat{\sigma}_k, \hat{\theta}_k) - \frac{1}{2} \right)^2. \quad (3.9)$$

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

$$TL(r_i) = \frac{\sum_{j=1}^i X_{j:n}}{\sum_{j=1}^n X_{j:n}}, \quad r_i = \frac{i}{n}, \quad i = 1, 2, \dots, n. \quad (3.10)$$

Kang and Cho (2001) proposed the NSLC for the complete sample as follows;

$$NSLC(r_i) = \frac{TSL(r_i)}{TSL_F(r_i)}, \quad r_i = \frac{i}{n}, \quad i = 1, 2, \dots, n, \quad (3.11)$$

where

$$\begin{aligned} TSL(r_i) &= \frac{\sum_{j=1}^i (X_{j:n} - X_{1:n})}{\sum_{j=1}^n (X_{j:n} - X_{1:n})} - r_i + 1, \\ TSL_F(r_i) &= \frac{\sum_{j=1}^i [F^{-1}(p_j) - F^{-1}(p_1)]}{\sum_{j=1}^n [F^{-1}(p_j) - F^{-1}(p_1)]} - r_i + 1. \end{aligned}$$

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

$$MNSLC(r_i) = \frac{MTSL(r_i)}{TSL_F(r_i)}, \quad r_i = \frac{a_i}{n}, \quad i = 1, 2, \dots, s, \quad (3.12)$$

where

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

We also propose the modified NSLC plot for multiply type-II censored samples using $(X, Y) = (1 - r_i, 1 - MNSLC(r_i))$. If data come from the logistic distribution, the modified NSLC plot is $y = 0$ (see, Figure 3.1 and Figure 3.2). The value of $1 - MNSLC(r_i)$ increases

and then decreases as $1 - r_i$ increases for all alternative distributions. The modified NSLC plot is a left skewed form when the alternative distributions are the lognormal (LN), exponential (Exp), and Rayleigh distributions (Rayleigh). However, the modified NSLC plot is a symmetrical form when the alternative distribution is the double Rayleigh (DR) distribution.

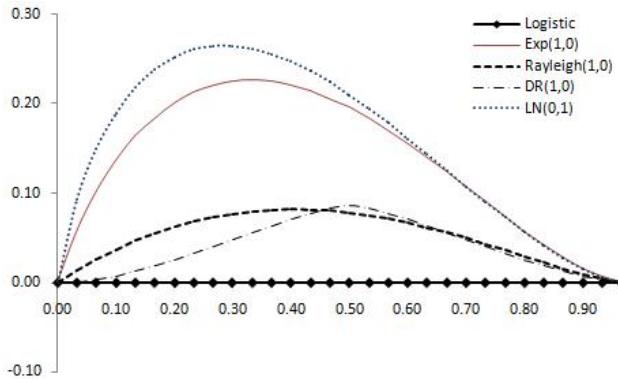


Figure 3.1 Modified NSLC plot : Complete data ($n=30, m=0$)

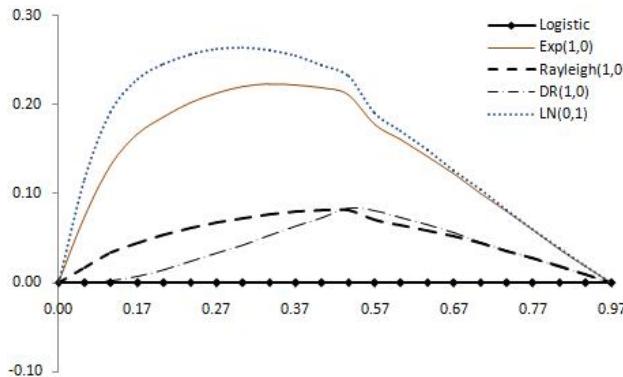


Figure 3.2 Modified NSLC plot : Multiply type-II censored data ($n=30, a_j=1, 5 \sim 13, 17 \sim 25, 28 \sim 30$)

4. Simulation study and illustrative example

In order to evaluate the performance of the proposed estimators, the mean squared errors of all proposed estimators were simulated by a Monte Carlo method for sample size $n = 20, 40$ and various choices of censoring ($m = n - s$ is the number of unobserved or missing data). The simulation procedure was repeated 10,000 times.

From Table 4.2, $\hat{\theta}_1$ and $\hat{\sigma}_1$ are more efficient than $\hat{\theta}_2$ and $\hat{\sigma}_2$ in the sense of the MSE. As expected, the MSEs of all estimators decreases as sample size n increases. For fixed sample size, the mean squared errors increases generally as m increases.

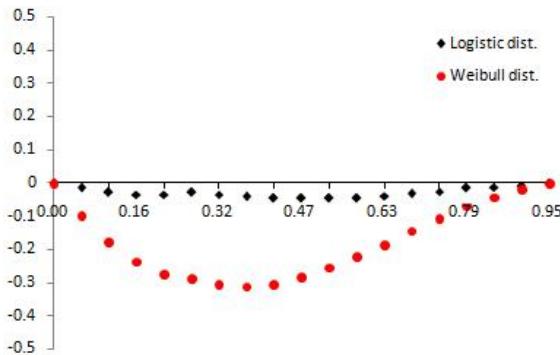


Figure 4.1 Modified NSLC plot : Complete data ($n=19, m=0$)

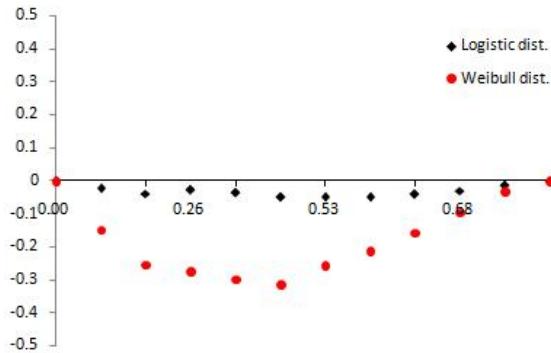


Figure 4.2 Modified NSLC plot : Multiply type-II censored data ($n=19, a_j=1, 2, 6\sim 9, 12\sim 15, 18, 19$)

In order to evaluate the performance of the proposed goodness-of-fit tests, the powers were simulated by a Monte Carlo method for sample size $n = 20, 40$ and various choices of censoring. This simulation procedure was repeated 10,000 times. We compute critical value ($\alpha = 0.05$) and type I error for four tests under the logistic distribution (see Table 4.3 ~ 4.4).

The powers of five tests with significance level 0.05 for the logistic distribution based on multiply type-II censored samples are investigated under 4 alternative distributions. These values are presented in Table 4.5 ~ 4.8.

The modified EDF type tests that use the estimators $\hat{\theta}_1$ and $\hat{\sigma}_1$ are generally more powerful than the tests that use the estimators $\hat{\theta}_2$ and $\hat{\sigma}_2$ when multiply type-II censored data.

For the exponential, lognormal, double Rayleigh alternative distribution, the modified Anderson-Darling test is more powerful than other tests. Especially, for the double Rayleigh alternative distribution, the modified Watson test showed good performance when complete data.

All computation were programmed in Microsoft Visual C++ 6.0 and random numbers for simulations were generated by IMSL subroutines.

We present an example to illustrate the inference procedures discussed in this paper.

We apply the proposed graphical methods to a real data set. Let us consider the following data, which represent failure times to breakdown of an insulating fluid tested at 34 kilovolts.

5. Conclusions

In most cases of censored and truncated samples, the maximum likelihood method does not provide explicit estimators. So we discuss another method for the purpose of providing the explicit estimators. We use four modified EDF types test for the logistic distribution based on multiply type-II censored samples using AMLEs. We propose the modified NSLC plot to test for the logistic distribution based on multiply type-II censored samples.

Based on the simulation study and illustrative example, the modified EDF type tests that use the estimators $\hat{\theta}_1$ and $\hat{\sigma}_1$ are generally more powerful than the tests that use the estimators $\hat{\theta}_2$ and $\hat{\sigma}_2$ when multiply type-II censored data. In addition we believe that the modified NSLC plot can be extended to cases with various censoring.

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

- Balakrishnan, N. (1989). Approximate MLE of the scale parameter of the Rayleigh distribution with censoring. *IEEE Transactions on Reliability*, **38**, 355-357.
- 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 & Simulation*, **74**, 25-45.
- Balakrishnan, N. and Kateri, M. (2008). On the maximum likelihood estimation of parameters of Weibull distribution based on complete and censored data. *Statistics & Probability Letters*, **78**, 2971-2975.
- Balakrishnan, N., Gupta, S. S., and Panchapakesan, S. (1995). Estimation of the mean and standard deviation of the logistic distribution based on multiply type-II censored samples. *Statistics*, **27**, 127-142.
- Birnbaum, A. and Dudman, J. (1963). Logistic order statistics. *The Annals of Mathematical Statistics*, **34**, 658-663.
- Choulakian, V. and Stephens, M. A. (2001). Goodness-of-fit tests for the generalized Pareto distribution. *Technometrics*, **43**, 478-484.
- 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.
- Fei, H., Kong, F. and Tang, Y. (1995). Estimation for two-parameter Weibull distribution and extreme-value distribution under multiply type-II censoring. *Communications in Statistics-Theory and Methods*, **24**, 2087-2104.
- Han, J. T. and Kang, S. B. (2008). Estimation for the double Rayleigh distribution based on multiply type-II censored samples. *Communications of the Korean Statistical Society*, **15**, 367-378.
- Kang, S. B., Cho, Y. S. and Han, J. T. (2008). Estimation for the half logistic distribution under progressive type-II censoring. *Communications of the Korean Statistical Society*, **15**, 815-823.
- 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, S. K. (2006). Test for the exponential distribution based on multiply type-II censored samples. *The Korean Communications in Statistics*, **13**, 537-550.
- Kim, Y., Kang, S. B. and Seo, J. I. (2011a). Bayesian estimations on the exponentiated half triangle distribution under type-I hybrid censoring. *Journal of Korean Data & Information Science Society*, **22**, 565-574.
- Kim, Y., Kang, S. B. and Seo, J. I. (2011b). Bayesian estimation in the generalized half logistic distribution under progressively type-II censoring. *Journal of Korean Data & Information Science Society*, **22**, 977-987.
- Kuş, C. and Kaya, M. F. (2006). Estimation of parameters of the loglogistic distribution based on progressive censoring using the EM algorithm. *Hacettepe Journal of Mathematics and Statistics*, **35**, 203-211.
- Mathai, A. M. (2003). Order statistics from a logistic distribution and applications to survival and reliability analysis. *IEEE Transaction on Reliability*, **52**, 200-206.
- Maswadah, M. (2003). Conditional confidence interval estimation for the inverse Weibull distribution based on censored generalized order statistics. *Journal of Statistical Computation & Simulation*, **12**, 887-898.

- Nelson, W. (1982). *Applied life data analysis*, New York: Willy.
- Pettitt, A. N. and Stephens, M. A. (1976). Modified Cramer-von Mises statistics for censored data. *Biometrika*, **63**, 291-298.
- Porter III, J. E., Coleman, J. W. and Moore, A. H. (1992). Modified KS, AD, and C-vM tests for the Pareto distribution with unknown location & scale parameters. *IEEE Transactions on Reliability*, **41**, 112-117.
- Puig, P. and Stephens, M. A. (2000). Tests of fit for the Laplace distribution with applications. *Technometrics*, **42**, 417-424.
- Shin, H. and Lee, K. (2012). Estimation in the exponential distribution under progressive type I interval censoring with semi-missing data. *Journal of Korean Data & Information Science Society*, **23**, 1271-1277.