Estimation for the Half Logistic Distribution under Progressive Type-II Censoring

Suk-Bok Kang¹⁾, Young-Seuk Cho²⁾, Jun-Tae Han³⁾

Abstract

In this paper, we derive the approximate maximum likelihood estimators (AMLEs) and maximum likelihood estimator of the scale parameter in a half-logistic distribution based on progressive Type-II censored samples. We compare the proposed estimators in the sense of the mean squared error for various censored samples. We also obtain the approximate maximum likelihood estimators of the reliability function using the proposed estimators. We compare the proposed estimators in the sense of the mean squared error.

Keywords: Approximate maximum likelihood estimator; half logistic distribution; progressive Type-II censoring; reliability.

1. Introduction

The cumulative distribution function(cdf) and the probability density function(pdf) of the random variable having the half-logistic distribution are given by

$$F(x) = \frac{1 - \exp\left(-\frac{x - \theta}{\sigma}\right)}{1 + \exp\left(-\frac{x - \theta}{\sigma}\right)}, \quad x \ge \theta, \ \sigma \ge 0$$
 (1.1)

and

$$f(x) = \frac{2\exp\left(-\frac{x-\theta}{\sigma}\right)}{\sigma\left\{1 + \exp\left(-\frac{x-\theta}{\sigma}\right)\right\}^2},\tag{1.2}$$

where θ and σ are the location and the scale parameters, respectively.

Use of this distribution as a possible life-time model has been suggested by Balakrishnan (1985) who had established several recurrence relations for the single and the product moments of order statistics. Adatia (2000) studied the estimation of the location and scale parameters of the half-logistic distribution using generalized ranked-set

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

Assistant professor, Division of Mathematics and Statistics, Pusan National University, Pusan 609-735, Korea.

³⁾ Researcher, Institute for National Health Insurance, National Health Insurance Corporation, Seoul 121-749, Korea. Correspondence: maru@nhic.or.kr

sampling technique. Balakrishnan and Aggarwala (1996) established several recurrence relations satisfied by the single and the product moments for order statistics from the right truncated generalized half-logistic distribution.

A generalization of Type-II censoring is progressive Type-II censoring. In this case, the first failure in the sample is observed and a random sample of size r_1 is immediately drawn from the remaining n-1 unfailed items and removed from the test, leaving $n-1-r_1$ items in test. When the second item has failed, r_2 of the still unfailed items are removed and so on.

Balakrishnan and Wong (1991) derived the approximate maximum likelihood estimator(AMLE) of the scale parameter of the half-logistic distribution with Type-II right censoring. They also studied the bias and variance of proposed estimator. Adatia (1997) obtained an approximate best linear unbiased estimatiors (BLUEs) of the scale and the location parameters of the half-logistic distribution based on fairly large doubly censosred samples. Balakrishnan et al. (2004) studied point and interval estimation for the extreme value distribution based on progressively Type-II censored samples. Kang and Park (2005) derived the AMLEs of the scale parameter of the half-logistic distribution based on multiply Type-II censored samples. Balakrishnan and Asgharzadeh (2005) discussed the maximum likelihood estimator(MLE) of the scale parameter of the halflogistic distribution based on progressive Type-II censored samples. They also provided the AMLE of the scale parameter of the half-logistic distribution based on progressive Type-II censored samples. See and Kang (2007) proposed the AMLEs of the scale parameter when the location parameter is known and the AMLE of the location parameter when the scale parameter is known in the two-parameter Rayleigh distribution based on progressive Type-II censored samples. They also proposed the AMLEs of the location and scale parameters in the two-parameter Rayleigh distribution based on progressive Type-II censored samples when two parameters are unknown.

Recently, Lee et al. (2008) proposed the AMLEs of the scale parameter in a triangular distribution based on multiply Type-II censored samples by the approximate maximum likelihood estimation methods. 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.

In this paper, we derive the AMLEs and the MLE of the scale parameter σ and the location parameter θ under progressive Type-II censoring. The scale parameter is estimated by approximate maximum likelihood estimation method using two different types of Taylor series expansions. We also obtain the AMLEs and the MLE of the reliability function using the proposed estimators. We compare the proposed estimators in the sense of the mean squared error(MSE). We also compare the proposed estimators in the sense of the MSE through Monte Carlo simulation for various censored samples.

2. Maximum Likelihood Estimation

We will discuss the maximum likelihood estimation of the scale parameter based on progressive Type-II censored samples. Let $X_{1:m:n}, \ldots, X_{m:m:n}$ denote such a progressive Type-II censored sample with (r_1, \ldots, r_m) being the progressive censoring scheme.

Note that the case m=n, in which case $r_1=\cdots=r_m=0$, corresponds to the complete sample situation, while the case $r_1=\cdots=r_{m-1}=0$, $r_m=n-m$ corresponds

to the usual Type-II censored sample.

The likelihood function based on the progressively Type-II censored sample is given by

$$L = C \prod_{i=1}^{m} f(x_{i:m:n}; \theta, \sigma) \left\{ 1 - F(x_{i:m:n}; \theta, \sigma) \right\}^{r_i},$$
 (2.1)

where $C = n(n-1-r_1)(n-2-r_1-r_2)\cdots(n-m+1-r_1-\cdots-r_m-1)$.

The random variable $Z_{i:m:n} = (X_{i:m:n} - \theta)/\sigma$ has a standard half-logistic distribution with pdf and cdf;

$$f(z_{i:m:n}) = \frac{2\exp(-z_{i:m:n})}{\{1 + \exp(-z_{i:m:n})\}^2}, \quad z \ge 0$$

and

$$F(z_{i:m:n}) = \frac{1 - \exp(-z_{i:m:n})}{1 + \exp(-z_{i:m:n})}, \quad z \ge 0.$$

The $f(z_{i:m:n})$, $f'(z_{i:m:n})$ and $F(z_{i:m:n})$ are satisfied as

$$f'(z_{i:m:n}) = -F(z_{i:m:n})f(z_{i:m:n}), \qquad f(z_{i:m:n}) = \frac{\{1 - F(z_{i:m:n})\}\{1 + F(z_{i:m:n})\}}{2}.$$

From the Equation (2.1), the log-likelihood function may be written as

$$\ln L = K - m \ln \sigma + \sum_{i=1}^{m} \ln f(z_{i:m:n}) + \sum_{i=1}^{m} r_i \ln \left[1 - F(z_{i:m:n}) \right], \tag{2.2}$$

where K is a constant.

On differentiating the log-likelihood function with respect to σ in turn and equation to zero, we obtain the estimating equation as

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{1}{2\sigma} \left\{ 2m - \sum_{i=1}^{m} r_i z_{i:m:n} - \sum_{i=1}^{m} (r_i + 2) F(z_{i:m:n}) z_{i:m:n} \right\} = 0.$$
 (2.3)

We can find the MLE of σ as values $\hat{\sigma}$ that maximize the log-likelihood function in (2.2) by solving the equation $\partial \ln L/\partial \sigma = 0$. Since the Equation (2.3) cannot be solved explicitly, some numerical method must be employed.

From the Equation (2.1), the likelihood function is a monotonically increasing function of θ . Thus the MLE of θ is $\hat{\theta} = X_{1:m:n}$.

3. Approximate Maximum Likelihood Estimators

Since the log-likelihood equation do not admit explicit solutions, it will be desirable to consider an approximation to the likelihood equation which provide us with explicit estimator for the scale parameter.

Let

$$\xi_{i:m:n} = F^{-1}(p_{i:m:n}) = -\ln\left(\frac{q_{i:m:n}}{1 + p_{i:m:n}}\right),$$

where $q_{i:m:n} = 1 - p_{i:m:n}$ and

$$p_{i:m:n} = 1 - \prod_{j=m-i+1}^{m} \frac{j + r_{m-j+1} + \dots + r_m}{j+1 + r_{m-j+1} + \dots + r_m}, \quad i = 1, \dots, m.$$

First, we can approximate the following function by Taylor series expansion as

$$F(z_{i:m:n})z_{i:m:n} \simeq \alpha_i + \beta_i z_{i:m:n}, \tag{3.1}$$

where

$$\alpha_{i} = -\frac{1}{2} q_{i:m:n} (1 + p_{i:m:n}) \left\{ \ln \left(\frac{1 + p_{i:m:n}}{q_{i:m:n}} \right) \right\}^{2},$$

$$\beta_{i} = \frac{1}{2} q_{i:m:n} (1 + p_{i:m:n}) \ln \left(\frac{1 + p_{i:m:n}}{q_{i:m:n}} \right) + p_{i:m:n}.$$

By substituting the Equation (3.1) into the equation (2.3), we may approximate the likelihood equation in (2.3) by

$$\frac{\partial \ln L}{\partial \sigma} \simeq -\frac{1}{2\sigma} \left\{ 2m - \sum_{i=1}^{m} r_i z_{i:m:n} - \sum_{i=1}^{m} (r_i + 2)(\alpha_i + \beta_i z_{i:m:n}) \right\} = 0.$$
 (3.2)

We can derive an estimator of σ as follows;

$$\tilde{\sigma}_{1} = \frac{\sum_{i=1}^{m} r_{i} \left(X_{i:m:n} - \hat{\theta} \right) + \sum_{i=1}^{m} (r_{i} + 2) \beta_{i} \left(X_{i:m:n} - \hat{\theta} \right)}{2m - \sum_{i=1}^{m} (r_{i} + 2) \alpha_{i}}.$$
(3.3)

Since $\alpha_i < 0$ and $\beta_i > 0$, the estimator $\tilde{\sigma}_1$ is always positive.

Second, we can approximate the following function

$$F(z_{i:m:n}) \simeq \gamma_i + \delta_i z_{i:m:n},\tag{3.4}$$

where

$$\gamma_i = p_{i:m:n} + \frac{1}{2} q_{i:m:n} (1 + p_{i:m:n}) \ln \left(\frac{q_{i:m:n}}{1 + p_{i:m:n}} \right),$$

$$\delta_i = \frac{1}{2} q_{i:m:n} (1 + p_{i:m:n}).$$

Balakrishnan and Asgharzadeh (2005) derived an AMLE of the scale parameter σ using the Equation (3.4) when $\theta = 0$ as follows:

$$\tilde{\sigma}_2 = \frac{-A + \sqrt{A^2 - 8mB}}{4m},\tag{3.5}$$

where

$$A = -\left\{ \sum_{i=1}^{m} r_i X_{i:m:n} + \sum_{i=1}^{m} (r_i + 2) \gamma_i X_{i:m:n} \right\},$$

$$B = -\sum_{i=1}^{m} (r_i + 2) \delta_i X_{i:m:n}^2.$$

In this case, we use this estimator that is modified by

$$A = -\left\{ \sum_{i=1}^{m} r_i \left(X_{i:m:n} - \hat{\theta} \right) + \sum_{i=1}^{m} (r_i + 2) \gamma_i \left(X_{i:m:n} - \hat{\theta} \right) \right\},$$

$$B = -\sum_{i=1}^{m} (r_i + 2) \delta_i \left(X_{i:m:n} - \hat{\theta} \right)^2.$$

4. Estimation of the Reliability

The reliability function of the half-logistic distribution is

$$R(t) = 1 - F(t) = P(X > t) = 1 - \frac{1 - \exp\left(-\frac{t - \theta}{\sigma}\right)}{1 + \exp\left(-\frac{t - \theta}{\sigma}\right)}, \quad t > 0.$$
 (4.1)

For the progressive Type-II censored data, we now propose the AMLEs and MLE of the reliability function R(t) using the proposed AMLEs $\tilde{\sigma}_i$ and MLE $\hat{\sigma}$, $\hat{\theta}$ that can be used for progressively Type-II censored sample as follows.

$$\tilde{R}_{i}(t) = 1 - \frac{1 - \exp\left(-\frac{t - \hat{\theta}}{\tilde{\sigma}_{i}}\right)}{1 + \exp\left(-\frac{t - \hat{\theta}}{\tilde{\sigma}_{i}}\right)}, \quad i = 1, 2 \text{ and } \hat{R}(t) = 1 - \frac{1 - \exp\left(-\frac{t - \hat{\theta}}{\hat{\sigma}}\right)}{1 + \exp\left(-\frac{t - \hat{\theta}}{\hat{\sigma}}\right)}. \quad (4.2)$$

From the Equation (4.2), the mean squared errors of these estimators are simulated by Monte Carlo method(based on 10,000 Monte Carlo runs) for sample size n = 20, m = 15(3*0, 2, 4*0, 3, 6*0)(see, Figure 5.1 and Figure 5.2).

5. The Simulated Results and Illustrative Example

The simulations were carried out for sample size n = 10(5)20, 30, 40 and different choices of the effective sample size m. For simplicity in notation, we will denote the schemes $(0, 0, \ldots, n-m)$ by ((n-m)*0, n-m), for example, (10*0) denotes the progressive censoring scheme $(0, 0, \ldots, 0)$ and (3*0, 2, 2, 0) denotes the progressive censoring scheme (0, 0, 0, 2, 2, 0).

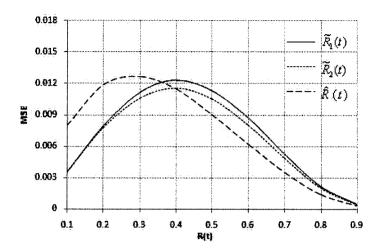


Figure 5.1: The relative MSEs of $\tilde{R}_i(t)$ and $\hat{R}(t)$ when the location parameter is unknown [m = 15(3 * 0, 2, 4 * 0, 3, 6 * 0)]

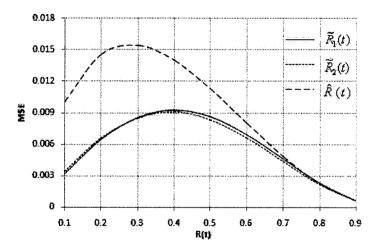


Figure 5.2: The relative MSEs of $\tilde{R}_i(t)$ and $\hat{R}(t)$ when the location parameter is known [m = 15(3*0, 2, 4*0, 3, 6*0)]

The convergence of the Newton-Raphson method depended on the choice of the initial values. For this reason, the proposed AMLE $\tilde{\sigma}_1$ and $\tilde{\sigma}_2$ were used as starting values for the iterations, the MLE is obtained by solving the nonlinear equations (2.3) using Newton-Raphson method. The mean squared errors of the proposed AMLEs and MLE are simulated by Monte Carlo method(based on 10,000 Monte Carlo runs) for sample size $n=10(5)20,\ 30,\ 40$ and various choices of censoring under progressive Type-II censored sample with $\sigma=1$ and $\theta=0$. These values are given in Table 5.1 and 5.2.

	\overline{m}	Scheme	$\hat{ heta}$	$ ilde{\sigma}_1$	$ ilde{\sigma}_2$	ô
10	10	(10 * 0)	0.063196	0.078208	0.076629	0.142314
	6	(3*0,2,2,0)	0.063600	0.138543	0.137045	0.142022
	6	(2*0,4,3*0)	0.063600	0.137122	0.134839	0.167250
	6	(4,5*0)	0.063600	0.132780	0.130675	0.191610
	5	(5, 4 * 0)	0.064016	0.158477	0.155321	0.213622
	5	(4*0,5)	0.064016	0.168801	0.166944	0.194544
	5	(0,5,3*0)	0.064016	0.162799	0.159422	0.200173
15	15	(15 * 0)	0.030344	0.050703	0.049872	0.124611
	10	(5,9*0)	0.029970	0.076263	0.075098	0.146760
	10	(4*0,3,3*0,2,0)	0.029970	0.077959	0.076663	0.080487
	10	(0,3,6*0,2,0)	0.029970	0.076941	0.075696	0.083594
	10	(2*0,1,0,2,0,2,3*0)	0.029970	0.078291	0.076978	0.100750
20	20	(20 * 0)	0.017670	0.036468	0.036000	0.115493
	15	(3*0,2,4*0,3,6*0)	0.017694	0.050724	0.049987	0.093133
	10	(5, 2 * 0, 5, 6 * 0)	0.017479	0.077032	0.075983	0.127705
	10	(2*0,1,0,2,0,2,2*0,5)	0.017479	0.078774	0.077820	0.093016
	10	(2*0,3,0,2,0,2,2*0,3)	0.017479	0.078108	0.076906	0.082563
30	30	(30 * 0)	0.008072	0.024086	0.023841	0.108069
	20	(3*0,5,3*0,5,12*0)	0.008167	0.036586	0.036196	0.097341
	20	(2*0, 10, 17*0)	0.008167	0.036191	0.035843	0.113585
	20	(9*0,10,10*0)	0.008167	0.036880	0.036482	0.078897
	15	(5, 6*0, 10, 7*0)	0.008178	0.050924	0.050364	0.085354
	15	(10, 6*0, 5, 7*0)	0.008178	0.050039	0.049507	0.095634
40	20	(8*0, 10, 10, 10*0)	0.004691	0.037335	0.036992	0.073347
	20	(2*0,5,5,5,5,14*0)	0.004691	0.036505	0.036197	0.102695
	20	(5, 16 * 0, 5, 5, 5)	0.004691	0.036535	0.036284	0.051901
	20	(3*0, 15, 7*0, 5, 8*0)	0.004691	0.036609	0.036305	0.077270

Table 5.1: The relative mean squared errors for the estimators of the scale parameter σ and the location parameter θ .

From Table 5.1 and 5.2, AMLE $\tilde{\sigma}_1$ and $\tilde{\sigma}_2$ are more efficient than the maximum likelihood estimator $\hat{\sigma}$ and the AMLEs are the function of the available progressive Type-II censored sample $X_{1:m:n}, \ldots, X_{m:m:n}$ and progressive censored scheme (r_1, \ldots, r_m) but the MLE does not have the closed form. So the MLE can be evaluated by some numerical method.

The AMLE $\tilde{\sigma}_2$ is more efficient than the AMLE $\tilde{\sigma}_1$ and MLE $\hat{\sigma}$ in the sense of the MSE when the location parameter is unknown. But the estimator $\tilde{\sigma}_1$ is more efficient than the estimator $\tilde{\sigma}_2$ and $\hat{\sigma}$ in the sense of the MSE when the location parameter is known except some right censoring.

As expected, the MSE of all estimators decreases as sample size n increases. For fixed sample size, the MSE increases generally as the number of unobserved or missing data n-m increases.

From Figure 5.1 and 5.2, the MSEs of all estimators increase and then decrease as R(t) increase.

When the location parameter is unknown (n = 20, m = 15(3*0, 2, 4*0, 3, 6*0)), the estimator $\tilde{R}_2(t)$ is generally more efficient than $\tilde{R}_1(t)$, but the estimator $\hat{R}(t)$ is generally more efficient than $\tilde{R}_1(t)$ and $\tilde{R}_2(t)$ when R(t) > 0.4. Among the three estimators $\hat{R}(t)$, $\tilde{R}_1(t)$, $\tilde{R}_2(t)$, $\hat{R}(t)$ is no good estimator when R(t) is small but $\hat{R}(t)$ is good estimator

	•					
n	m	Scheme	$ ilde{\sigma}_1$	$ ilde{\sigma}_2$	$\hat{\sigma}$	
	10	(10 * 0)	0.067899	0.070844	0.206526	
	6	(3*0,2,2,0)	0.112634	0.121093	0.143336	
	6	(2*0,4,3*0)	0.113954	0.122874	0.210921	
10	6	(4, 5 * 0)	0.112904	0.121141	0.263509	
	5	(5, 4 * 0)	0.134356	0.144943	0.289342	
	5	(4*0,5)	0.134512	0.136355	0.127671	
	5	(0, 5, 3 * 0)	0.135790	0.146976	0.255676	
	15	(15 * 0)	0.045611	0.046968	0.171901	
	10	(5, 9 * 0)	0.068368	0.071424	0.200552	
15	10	(4*0,3,3*0,2,0)	0.067931	0.070796	0.091272	
	10	(0,3,6*0,2,0)	0.067531	0.070275	0.100976	
	10	(2*0,1,0,2,0,2,3*0)	0.068766	0.072055	0.128421	
	20	(20*0)	0.033982	0.034767	0.153287	
	15	(3*0, 2, 4*0, 3, 6*0)	0.046048	0.047541	0.122460	
20	10	(5, 2 * 0, 5, 6 * 0)	0.069529	0.072922	0.166058	
	10	(2*0,1,0,2,0,2,2*0,5)	0.068688	0.069538	0.068801	
	10	(2*0,3,0,2,0,2,2*0,3)	0.068335	0.069512	0.067060	
	30	(30 * 0)	0.022623	0.022950	0.133918	
	20	(3*0,5,3*0,5,12*0)	0.034597	0.035464	0.123326	
30	20	(2*0, 10, 17*0)	0.034364	0.035203	0.143582	
30	20	(9*0, 10, 10*0)	0.034738	0.035647	0.099994	
	15	(5, 6*0, 10, 7*0)	0.047038	0.048738	0.106608	
	15	(10, 6*0, 5, 7*0)	0.046444	0.048009	0.121403	
	20	(8*0, 10, 10, 10*0)	0.035481	0.036481	0.090305	
40	20	(2*0,5,5,5,5,14*0)	0.034912	0.035818	0.126938	
40	20	(5, 16 * 0, 5, 5, 5)	0.034224	0.034440	0.040073	
	20	(3*0,15,7*0,5,8*0)	0.034888	0.035830	0.096083	

Table 5.2: The relative mean squared errors for the estimators of the scale parameter σ when the location parameter θ is known.

when R(t) is large.

When the location parameter is known(n = 20, m = 15(3*0, 2, 4*0, 3, 6*0)), the estimator $\tilde{R}_1(t)$ and $\tilde{R}_2(t)$ are generally more efficient than $\hat{R}(t)$, but the MSEs of three estimators are almost same when R(t) > 0.8.

Let us consider the following data, which represent failure times, in minutes, for a specific type of electrical insulation that was subjected to a continuously increasing voltage stress as given by Lawless (1982)

12.3, 21.8, 24.4, 28.6, 43.2, 46.9, 70.7, 75.3, 95.5, 98.1, 138.6, 151.9.

This data has been utilized earlier by Balakrishnan and Wong (1991). For complete data, we can obtain the MLEs $\hat{\theta} = 12.3$ and $\hat{\sigma} = 53.083927$ and the AMLEs $\tilde{\sigma}_1 = 40.368197$ and $\tilde{\sigma}_2 = 40.773586$.

For this example of n=12, m=6(0, 2, 0, 2, 2, 0) and the progressive Type-II censored sample is 12.3, 21.8, 28.6, 46.9, 75.3, 98.1, we can obtain the MLEs $\hat{\theta}=12.3$ and $\hat{\sigma}=46.683830$ and the AMLEs $\tilde{\sigma}_1=45.299069$ and $\tilde{\sigma}_2=45.989116$.

References

- Adatia, A. (1997). Approximate BLUEs of the parameters of the half logistic distribution based on fairly large doubly censored samples, *Computational Statistics & Data analysis*, **24**, 179–191.
- Adatia, A. (2000). Estimation of parameters of the half-logistic distribution using generalized ranked set sampling, *Computational Statistics & Data analysis*, **33**, 1–13.
- Balakrishnan, N. (1985). Order statistics from the half logistic distribution, *Journal of Statistical Computation & Simulation*, **20**, 287–309.
- Balakrishnan, N. and Aggarwala, R. (1996). Relationships for moments of order statistics from the right-truncated generalized half logistic distribution, *Annals of the Institute of Statistical Mathematics*, **48**, 519–534.
- Balakrishnan, N. and Asgharzadeh, A. (2005). Inference for the scaled half-logistic distribution based on progressively Type-II censored samples, *Communications in Statistics Theory & Methods*, **34**, 73-87.
- 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 Wong, K. H. T. (1991). Approximate MLEs for the location and scale parameters of the half-logistic distribution with Type-II right-censoring, *IEEE Transactions on Reliability*, **40**, 140–145.
- 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. and Park, Y. K. (2005). Estimation for the half-logistic distribution based on multiply Type-II censored samples, *Journal of the Korean Data & Information Science Society*. **16**, 145–156.
- Lawless, J. F. (1982). Statistical Models and Methods for Lifetime Data, John Wiley & Sons, New York.
- Lee, H. J., Han, J. T. and Kang, S. B. (2008). Estimation for a triangular distribution based on multiply Type-II censored samples, *Journal of the Korean Data & Information Science Society*, **19**, 319–330.
- Seo, E. H. and Kang, S. B. (2007). AMLEs for Rayleigh distribution based on progressive Type-II censored data, *The Korean Communications in Statistics*, **14**, 329–344.

[Received June 2008, Accepted August 2008]