

Estimation for the Skewed Exponential Distribution Based on Multiply Type-II Censored Samples

Suk-Bok Kang¹⁾ · Jun-Tae Han²⁾ · Sun-Mi Park³⁾

Abstract

In this paper, we derive the approximate maximum likelihood estimators of the scale and location parameters of the skewed exponential distribution based on multiply Type-II censored samples. We compare the proposed estimators in the sense of the mean squared error for various censored samples.

Keywords : Approximate maximum likelihood estimator, Multiply Type-II censored sample, Skewed exponential distribution

1. Introduction

Perhaps the most important and widely used probability distribution in life-analysis is the two-parameter exponential distribution with the probability density function (p.d.f.) of the form

$$g_X(x; \theta, \sigma) = \frac{1}{\sigma} e^{-(x-\theta)/\sigma}, \quad x \geq \theta, \quad \sigma > 0 \quad (1.1)$$

and the cumulative distribution function (c.d.f.) of the form

$$G_X(x; \theta, \sigma) = 1 - e^{-(x-\theta)/\sigma}, \quad x \geq \theta, \quad \sigma > 0. \quad (1.2)$$

Many authors studied skewness into a symmetric distribution in many ways.

Azzalini (1985) observed the symmetric and independent random variable X with density function g_X and distribution function G_X .

It follows that

$$f_X(x) = 2g_X(x)G_X(\lambda x), \quad x > 0 \quad (1.3)$$

is a valid density function corresponding to possibly skew distribution. If X is normal variables, then the above equation gives the density of the skew normal distribution introduced by many authors.

We consider a random variable X is exponential variable with $\lambda = 1$. Then the

1 Professor, Department of Statistics, Yeungnam University, Gyongsan, 712-749, Korea
E-mail : sbkang@yu.ac.kr

2 Graduate student, Department of Statistics, Yeungnam University, Gyongsan, 712-749, Korea

3 Graduate student, Department of Statistics, Yeungnam University, Gyongsan, 712-749, Korea

density function of a new distribution is

$$f_X(x) = \frac{2}{\sigma} \left[\exp\left\{-\frac{x-\theta}{\sigma}\right\} - \exp\left\{-\frac{2(x-\theta)}{\sigma}\right\} \right] \quad (1.4)$$

and cdf is

$$F_X(x) = \left[1 - \exp\left\{-\frac{x-\theta}{\sigma}\right\} \right]^2. \quad (1.5)$$

Azzalini and Valle (1996) discussed the multivariate version of the skew-normal density. Branco and Dey (2001) proposed a general class of multivariate skew-elliptical distributions, and obtained moments of the skew-elliptical distribution. Gupta (2003) proposed multivariate skew t -distribution and studied including the moments, he also discussed multivariate skew-Cauchy distribution, which is special case of the multivariate skew t -distribution. Gupta and Chang (2003) discussed multivariate skew-symmetric distribution and relationship with the Wishart distribution. Liseo and Loperfido (2004) discussed some peculiar features of default Bayes analysis of the scalar skew normal model, and consider the general scalar case with unknown location and scale parameters. Kozubowski and Panorska (2004) studied the tests of hypothesis about symmetry based on samples from possibly asymmetric Laplace distributions and present exact and limiting distribution of the test statistics. Ma and Genton (2004) proposed a flexible class of skew-symmetric distributions for which the p.d.f. has the form of a product of a symmetric density and a skewing function.

Censoring occurs when we are unable to observe the response variable of interest. In most cases of censored samples, Estimators of parameters may not be obtained explicitly by the maximum likelihood method. The approximate maximum likelihood estimating method was first developed by Balakrishnan (1989) for the purpose of providing the explicit estimators of the scale parameter in Rayleigh distribution.

Kang (1996) obtained the approximate maximum likelihood estimator (AMLE) for the scale parameter of the double exponential distribution based on Type-II censored samples and he showed that the proposed estimator is generally more efficient than the best linear unbiased estimator and the optimum unbiased absolute estimator.

Multiply Type-II censored sampling arises in life-testing experiments when the failure times of some units were not observed due to mechanical or experimental difficulties. Balakrishnan (1990), Fei and Kong (1994) considered the inference for the exponential distribution under multiply Type-II censoring. They obtained several point estimation methods for the one-parameter as well as two-parameter exponential distribution.

Recently, Kang (2003) derived several type point estimators of the location and the scale parameters in an explicit form for the general case when the available sample is multiply Type-II censored. Kang et al. (2004) obtained the AMLE for the scale parameter of the Weibull distribution based on multiply Type-II censored

samples.

In this paper, we derive the AMLEs of the scale parameter σ and the location parameter θ based on multiply Type-II censored samples. We also compare the proposed estimators in the sense of the mean squared error (MSE) for various censored samples.

2. Approximate Maximum Likelihood Estimators

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

$$X_{a_1:n} < X_{a_2:n} < \dots < X_{a_s:n} \tag{2.1}$$

where $1 \leq a_1 < a_2 < \dots < a_s \leq n$

$$a_0 = 0, \quad a_{s+1} = n + 1, \quad F(x_{a_0:n}) = 0, \quad F(x_{a_{s+1}:n}) = 1. \tag{2.2}$$

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

$$L = n! \prod_{j=1}^s f(x_{a_j:n}) \prod_{j=1}^{s+1} \frac{[F(x_{a_j:n}) - F(x_{a_{j-1}:n})]^{a_j - a_{j-1} - 1}}{(a_j - a_{j-1} - 1)!}. \tag{2.3}$$

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

$$\begin{aligned} \frac{\partial \ln L}{\partial \sigma} = & -\frac{1}{\sigma} \left[s + (a_1 - 1) \frac{f(Z_{a_1:n})}{F(Z_{a_1:n})} Z_{a_1:n} \right. \\ & - (n - a_s) \frac{f(Z_{a_s:n})}{1 - F(Z_{a_s:n})} Z_{a_s:n} + \sum_{j=1}^s \frac{f'(Z_{a_j:n})}{f(Z_{a_j:n})} Z_{a_j:n} \\ & \left. + \sum_{j=2}^s (a_j - a_{j-1} - 1) \frac{f(Z_{a_j:n})Z_{a_j:n} - f(Z_{a_{j-1}:n})Z_{a_{j-1}:n}}{F(Z_{a_j:n}) - F(Z_{a_{j-1}:n})} \right] = 0. \end{aligned} \tag{2.4}$$

where $Z_{i:n} = (X_{i:n} - \theta) / \sigma$, $f(z) = 2(e^{-z} - e^{-2z})$, $f'(z) = -2(e^{-z} - 2e^{-2z})$, and $F(z) = (1 - e^{-z})^2$.

Since the likelihood equations is very complicated, the equation (2.4) does not admit an explicit solution for σ .

Let

$$\xi_i = F^{-1}(p_i) = -\ln(1 - \sqrt{p_i}), \quad \text{where } p_i = \frac{i}{n+1}, \quad q_i = 1 - p_i$$

We may approximate the following functions in Taylor series around the points ξ_{a_1} , ξ_{a_s} , ξ_{a_j} and $(\xi_{a_j}, \xi_{a_{j-1}})$, respectively.

$$\frac{f(Z_{a_1:n})}{F(Z_{a_1:n})} Z_{a_1:n} \approx \alpha_1 + \beta_1 Z_{a_1:n} \tag{2.5}$$

$$\frac{f(Z_{a_i:n})}{1-F(Z_{a_i:n})} Z_{a_i:n} \simeq x_1 + \delta_1 Z_{a_i:n} \quad (2.6)$$

$$\frac{f'(Z_{a_i:n})}{f(Z_{a_i:n})} Z_{a_i:n} \simeq x_j + \delta_j Z_{a_i:n} \quad (2.7)$$

$$\frac{f(Z_{a_i:n})Z_{a_i:n} - f(Z_{a_{j-1}:n})Z_{a_{j-1}:n}}{F(Z_{a_i:n}) - F(Z_{a_{j-1}:n})} \simeq \alpha_j + \beta_j Z_{a_i:n} + \gamma_j Z_{a_{j-1}:n} \quad (2.8)$$

$$\frac{f(Z_{a_1:n})}{F(Z_{a_1:n})} \simeq \alpha_2 + \beta_2 Z_{a_1:n} \quad (2.9)$$

$$\frac{f(Z_{a_i:n})}{1-F(Z_{a_i:n})} \simeq x_2 + \delta_2 Z_{a_i:n} \quad (2.10)$$

$$\frac{f'(Z_{a_i:n})}{f(Z_{a_i:n})} \simeq x_{2j} + \delta_{2j} Z_{a_i:n} \quad (2.11)$$

$$\frac{f(Z_{a_i:n})}{F(Z_{a_i:n}) - F(Z_{a_{j-1}:n})} \simeq \alpha_{1j} + \beta_{1j} Z_{a_i:n} + \gamma_{1j} Z_{a_{j-1}:n} \quad (2.12)$$

$$\frac{f(Z_{a_{j-1}:n})}{F(Z_{a_i:n}) - F(Z_{a_{j-1}:n})} \simeq \alpha_{2j} + \beta_{2j} Z_{a_i:n} + \gamma_{2j} Z_{a_{j-1}:n} \quad (2.13)$$

where

$$\alpha_1 = -\frac{\xi_{a_1}^2}{p_{a_1}} \left[f'(\xi_{a_1}) - \frac{f^2(\xi_{a_1})}{p_{a_1}} \right]$$

$$\beta_1 = \frac{1}{p_{a_1}} \left[f(\xi_{a_1}) + \xi_{a_1} f'(\xi_{a_1}) - \frac{f^2(\xi_{a_1})}{p_{a_1}} \xi_{a_1} \right]$$

$$x_1 = -\frac{\xi_{a_i}^2}{q_{a_i}} \left[f'(\xi_{a_i}) + \frac{f^2(\xi_{a_i})}{q_{a_i}} \right]$$

$$\delta_1 = \frac{1}{q_{a_i}} \left[f(\xi_{a_i}) + \xi_{a_i} f'(\xi_{a_i}) - \frac{f^2(\xi_{a_i})}{q_{a_i}} \xi_{a_i} \right]$$

$$x_j = -\frac{\xi_{a_i}^2}{f(\xi_{a_i})} \left[f'(\xi_{a_i}) - \frac{[f'(\xi_{a_i})]^2}{f(\xi_{a_i})} \right]$$

$$\delta_j = \frac{1}{f(\xi_{a_i})} \left[f'(\xi_{a_i}) + \xi_{a_i} f'(\xi_{a_i}) - \frac{[f'(\xi_{a_i})]^2}{f(\xi_{a_i})} \xi_{a_i} \right]$$

$$\alpha_j = K^2 - \frac{\xi_{a_i}^2 f'(\xi_{a_i}) - \xi_{a_{j-1}}^2 f'(\xi_{a_{j-1}})}{p_{a_i} - p_{a_{j-1}}}$$

$$\begin{aligned} \beta_j &= \frac{1}{p_{a_i} - p_{a_{i-1}}} [(1 - K)f(\xi_{a_i}) + \xi_{a_i}f'(\xi_{a_i})] \\ \gamma_j &= -\frac{1}{p_{a_i} - p_{a_{i-1}}} [(1 - K)f(\xi_{a_{i-1}}) + \xi_{a_{i-1}}f'(\xi_{a_{i-1}})] \\ \alpha_2 &= \frac{1}{p_{a_i}} \left[f(\xi_{a_i}) - \xi_{a_i}f'(\xi_{a_i}) + \frac{f^2(\xi_{a_i})}{p_{a_i}} \xi_{a_i} \right] \\ \beta_2 &= \frac{1}{p_{a_i}} \left[f'(\xi_{a_i}) - \frac{f^2(\xi_{a_i})}{p_{a_i}} \right] \\ x_2 &= \frac{1}{q_{a_i}} \left[f(\xi_{a_i}) - \xi_{a_i}f'(\xi_{a_i}) - \frac{f^2(\xi_{a_i})}{q_{a_i}} \xi_{a_i} \right] \\ \delta_2 &= \frac{1}{q_{a_i}} \left[f'(\xi_{a_i}) + \frac{f^2(\xi_{a_i})}{q_{a_i}} \right] \\ x_{2j} &= \frac{1}{f(\xi_{a_i})} \left[f'(\xi_{a_i}) - \xi_{a_i}f''(\xi_{a_i}) + \frac{[f'(\xi_{a_i})]^2}{f(\xi_{a_i})} \xi_{a_i} \right] \\ \delta_{2j} &= \frac{1}{f(\xi_{a_i})} \left[f''(\xi_{a_i}) - \frac{[f'(\xi_{a_i})]^2}{f(\xi_{a_i})} \right] \\ \alpha_{1j} &= \frac{1}{p_{a_i} - p_{a_{i-1}}} [(1 + K)f(\xi_{a_i}) - \xi_{a_i}f'(\xi_{a_i})] \\ \beta_{1j} &= \frac{1}{p_{a_i} - p_{a_{i-1}}} \left[f'(\xi_{a_i}) - \frac{f^2(\xi_{a_i})}{p_{a_i} - p_{a_{i-1}}} \right] \\ \alpha_{2j} &= \frac{1}{p_{a_i} - p_{a_{i-1}}} [(1 + K)f(\xi_{a_{i-1}}) - \xi_{a_{i-1}}f'(\xi_{a_{i-1}})] \\ \beta_{2j} &= -\frac{f(\xi_{a_i})f(\xi_{a_{i-1}})}{[p_{a_i} - p_{a_{i-1}}]^2} = -\gamma_{1j} \\ \gamma_{2j} &= \frac{1}{p_{a_i} - p_{a_{i-1}}} \left[f'(\xi_{a_{i-1}}) - \frac{f^2(\xi_{a_{i-1}})}{p_{a_i} - p_{a_{i-1}}} \right] \\ K &= \frac{\xi_{a_i}f(\xi_{a_i}) - \xi_{a_{i-1}}f(\xi_{a_{i-1}})}{p_{a_i} - p_{a_{i-1}}} \end{aligned}$$

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

$$\hat{\sigma}_1 = \frac{B_1 + C_1 \hat{\theta}}{A_1} \quad (2.14)$$

where

$$A_1 = s + (a_1 - 1)\alpha_1 - (n - a_s)x_1 + \sum_{j=1}^s x_j + \sum_{j=2}^s (a_j - a_{j-1} - 1)\alpha_j$$

$$B_1 = -(a_1 - 1)\beta_1 X_{a_1:n} - (n - a_s)\delta_1 X_{a_s:n} - \sum_{j=1}^s \delta_j X_{a_j:n} \\ - \sum_{j=2}^s (a_j - a_{j-1} - 1)(\beta_j X_{a_j:n} + \gamma_j X_{a_{j-1}:n})$$

$$C_1 = (a_1 - 1)\beta_1 - (n - a_s)\delta_1 + \sum_{j=1}^s \delta_j + \sum_{j=2}^s (a_j - a_{j-1} - 1)(\beta_j + \gamma_j).$$

By substituting (2.9), (2.10), (2.11), (2.12), and (2.13) into (2.4), we can derive another estimator of σ as follows;

$$\hat{\sigma}_2 = \frac{-B_2 + \sqrt{B_2^2 - 4sC_2}}{2s} \quad (2.16)$$

where

$$B_2 = (a_1 - 1)\alpha_2 X_{a_1:n} - (n - a_s)x_2 X_{a_s:n} + \sum_{j=1}^s x_{2j} X_{a_j:n} \\ + \sum_{j=2}^s (a_j - a_{j-1} - 1)(\alpha_{1j} X_{a_j:n} + \alpha_{2j} X_{a_{j-1}:n}) - \left[(a_1 - 1)\alpha_2 \right. \\ \left. - (n - a_s)x_2 + \sum_{j=1}^s x_{2j} + \sum_{j=2}^s (a_j - a_{j-1} - 1)(\alpha_{1j} + \alpha_{2j}) \right] \hat{\theta}$$

$$C_2 = (a_1 - 1)\beta_2 (X_{a_1:n} - \hat{\theta})^2 - (n - a_s)\delta_2 (X_{a_s:n} - \hat{\theta})^2 \\ + \sum_{j=1}^s \delta_{2j} (X_{a_j:n} - \hat{\theta})^2 + \sum_{j=2}^s (a_j - a_{j-1} - 1) [\beta_{1j} (X_{a_j:n} - \hat{\theta})^2 \\ + 2\gamma_{1j} (X_{a_j:n} - \hat{\theta})(X_{a_{j-1}:n} - \hat{\theta}) - \gamma_{2j} (X_{a_{j-1}:n} - \hat{\theta})^2].$$

From (2.3), the likelihood equation for θ is obtained as

$$\frac{\partial \ln L}{\partial \theta} = -\frac{1}{\sigma} \left[(a_1 - 1) \frac{f(Z_{a_1:n})}{F(Z_{a_1:n})} - (n - a_s) \frac{f(Z_{a_s:n})}{1 - F(Z_{a_s:n})} \right. \\ \left. + \sum_{j=1}^s \frac{f'(Z_{a_j:n})}{f(Z_{a_j:n})} + \sum_{j=2}^s (a_j - a_{j-1} - 1) \frac{f(Z_{a_j:n}) - f(Z_{a_{j-1}:n})}{F(Z_{a_j:n}) - F(Z_{a_{j-1}:n})} \right] = 0. \quad (2.17)$$

Equation (2.17) does not admit an explicit solution for θ . But we can expand the function $\frac{f(Z_{a_j:n}) - f(Z_{a_{j-1}:n})}{F(Z_{a_j:n}) - F(Z_{a_{j-1}:n})}$ as follows;

$$\frac{f(Z_{a_j:n}) - f(Z_{a_{j-1}:n})}{F(Z_{a_j:n}) - F(Z_{a_{j-1}:n})} \simeq \alpha_{3j} + \beta_{3j} Z_{a_j:n} + \gamma_{3j} Z_{a_j:n} \quad (2.18)$$

where

$$\alpha_{3j} = \alpha_{1j} - \alpha_{2j}, \quad \beta_{3j} = \beta_{1j} - \beta_{2j}, \quad \gamma_{3j} = \gamma_{1j} - \gamma_{2j}$$

By substituting (2.9), (2.10), (2.11), and (2.18) into (2.17), we can derive an estimator of θ as follows:

$$\hat{\theta} = \frac{E}{D} \tag{2.19}$$

where

$$D = \frac{A_\theta C_1}{A_1} - C_\theta$$

$$E = -\frac{A_\theta B_1}{A_1} - B_\theta$$

$$A_\theta = (a_1 - 1)\alpha_2 - (n - a_s)\alpha_2 + \sum_{j=1}^s x_{2j} + \sum_{j=2}^s (a_j - a_{j-1} - 1)\alpha_{3j}$$

$$B_\theta = (a_1 - 1)\beta_2 X_{a_1:n} - (n - a_s)\delta_2 X_{a_s:n} + \sum_{j=1}^s \delta_{2j} X_{a_j:n} + \sum_{j=2}^s (a_j - a_{j-1} - 1)(\beta_{3j} X_{a_j:n} + \gamma_{3j} X_{a_{j-1}:n})$$

$$C_\theta = (a_1 - 1)\beta_2 - (n - a_s)\delta_2 + \sum_{j=1}^s \delta_{2j} + \sum_{j=2}^s (a_j - a_{j-1} - 1)(\beta_{3j} - \gamma_{3j})$$

From the above formulae, the mean squared errors of four estimators are simulated by Monte Carlo method for sample size $n = 20, 50$, and various choices

of censoring. These values are given in Table 1.

From Table 1, the estimator $\hat{\sigma}_1$ is more efficient than $\hat{\sigma}_2$ in the sense of MSE when location parameter θ is known. But the estimator $\hat{\sigma}_2$ is more efficient than $\hat{\sigma}_1$ in the sense of MSE when location parameter θ is unknown.

Table 1. The relative MSEs for the estimators of the location parameter θ and the scale parameter σ ,

n	k	a_j	θ is known.		θ is unknown.		
			$\hat{\sigma}_1$	$\hat{\sigma}_2$	$\hat{\theta}$	$\hat{\sigma}_1$	$\hat{\sigma}_2$
20	0	1~20	0.027575	0.029978	0.042931	0.042132	0.037636
	2	1~18	0.029225	0.031639	0.044082	0.046289	0.041024
		3~20	0.027657	0.030184	0.047536	0.044759	0.044137
		2~19	0.028448	0.031036	0.040609	0.043904	0.042529
	4	2~17	0.030462	0.033129	0.042088	0.048915	0.047265
		4~19	0.028582	0.030927	0.059747	0.051219	0.050726
		3~18	0.029309	0.031818	0.049506	0.049539	0.048764
		2~4 7~14 16~20	0.027668	0.034769	0.039691	0.041774	0.041375
	5	3~17	0.030519	0.033093	0.050793	0.053033	0.052169
		4~18	0.029387	0.031728	0.061286	0.054083	0.053540
		2~6 10~19	0.028521	0.033717	0.040399	0.043882	0.042822
	6	4~17	0.030603	0.032991	0.063241	0.058221	0.057621
		1 2 6~9 12~15 17~20	0.027722	0.038166	0.040917	0.041430	0.038740
	50	0	1~50	0.010951	0.011503	0.011308	0.015393
2		1~48	0.011198	0.011750	0.011421	0.015935	0.014891
		3~50	0.010957	0.011537	0.011407	0.015778	0.015601
		2~49	0.011070	0.011657	0.010332	0.015440	0.015098
4		2~47	0.011326	0.011917	0.010435	0.015966	0.015601
		4~49	0.011071	0.011608	0.012810	0.016741	0.016622
		3~48	0.011205	0.011785	0.011534	0.016326	0.016135
		2~4 7~14 16~50	0.010955	0.012388	0.010240	0.015173	0.014911
5		3~47	0.011329	0.011905	0.011601	0.016605	0.016403
		4~48	0.011204	0.011748	0.012884	0.017042	0.016922
		2~6 10~19 21~50	0.010958	0.012574	0.010251	0.015178	0.015045
6		4~47	0.011327	0.011867	0.012967	0.017344	0.017214
		1 2 6~9 12~15 17~50	0.010955	0.012918	0.010919	0.015237	0.014163

References

1. Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scand. J. Statist.*, 12, 171-178.
2. Azzalini, A. and Valle, A. D. (1996). The multivariate skew-normal distribution, *Biometrika*, 83, 715-726.
3. Balakrishnan, N. (1989). Approximate MLE of the scale parameter of the Rayleigh distribution with censoring, *IEEE Transactions on Reliability*, 38, 355-357.
4. Balakrishnan, N. (1990). On the maximum likelihood estimation of the location and scale parameters of exponential distribution based on multiply type-II censored samples, *Journal of Applied Statistics*, 17, 55-61.
5. Branco, M. D. and Dey, D. K. (2001). A general class of multivariate skew-elliptical distribution, *Journal of Multivariate analysis*, 79, 99-113.
6. Fei, H. and Kong, F. (1994). Interval estimations for one-and two-parameter exponential distributions under multiply type-II censored samples, *Communications in Statistics-Theory and Methods*, 23, 1717-1733.
7. Gupta, A. K. (2003). Multivariate skew t -distribution, *Statistics*, 37, 359-363.
8. Gupta, A. K. and Ghang, F. C. (2003). Multivariate skew symmetric distribution, *Applied Mathematics Letters*, 16, 643-646.
9. Kang, S. B. (1996). Approximate MLE for the scale parameter of the double exponential distribution based on type-II censored samples, *Journal of the Korean Mathematical Society*, 33, 69-79.
10. Kang, S. B. (2003). Approximate MLEs for exponential distribution under multiple type-II censoring, *Journal of the Korean Data & Information Science Society*, 14, 983-988.
11. Kang, S. B., Lee, H. J. and Han, J. T. (2004). Estimation of Weibull scale parameter based on multiply type-II censored samples, *Journal of the Korean Data & Information Science Society*, 15, 593-603.
12. Kozubowski, T. J. and Panorska, A. K. (2004) Testing symmetry under a skew Laplace model, *Journal of Statistical Planning and Inference*, 120, 41-63.
13. Liseo, B. and Loperfido, N. (2004). A note on reference priors for the scalar skew-normal distribution, *Journal of Statistical Planning and Inference*, In Press, Corrected Proof, Available online 6 October.
14. Ma, Y. and Genton, M. G. (2004). Flexible class of skew-symmetric distributions, *Board of the Foundation of the Scandinavian Journal of Statistics*, 31, 459-468.