Proceedings of Joint Conference of Korean Data And Information Science Society and The Korean Data Analysis Society,
April 30 - May 1, 2004, pp. 203-210

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

Suk-Bok Kang[1]

Abstract

When the available sample is multiply Type-II censored, the maximum likelihood estimators of the location and the scale parameters of two- parameter exponential distribution do not admit explicitly. In this case, we propose some estimators which are linear functions of the order statistics and also propose some estimators by approximating the likelihood equations appropriately. We compare the proposed estimators by the mean squared errors.

Keywords : Exponential distribution, multiply Type-II censored samples, approximate maximum likelihood estimator, mean squared error

1. Introduction

The exponential distributions have been used as models in analyzing life-time data quite extensively; for example, see Lawless (1982), Kambo (1978), Balakrishnan and Cohen (1991), Kang et al. (1997), and Kang and Cho (1998) etc. Most works are based on the assumption that the sample is either Type-II right censored or Type-II doubly censored, and some are concerned with Type-I censored samples. Balakrishnan (1990), Balasubramanian and Balakrishnan (1992), Fei and Kong (1994), and Upadhyay et al. (1996) considered the inference for the exponential distribution under multiple Type-II censoring. They obtained several point estimation methods for the one-parameter as well as two-parameter exponential distribution. Kang (2003) obtained the approximate maximum likelihood estimators for the exponential distribution under multiple Type-II censoring.

By considering the likelihood equations based on multiply Type-II censored samples and noting that they do not admit explicit maximum likelihood estimators, we propose some explicit estimators derived by appropriately approximating the likelihood equations.

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

Multiply Type-II censored samples may arise in practice in different ways. They may arise in life-testing experiments when the failure times of some units were not observed due to mechanical or experimental difficulties. Another situation where multiply censored samples arise naturally is when some units failed between two points of observation with exact times of failure of these units unobserved.

In this paper, we derive 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.

2. Estimators of the location and the scale parameters

Consider a two-parameter exponential distribution with the probability density function (pdf)

$$f(x;\theta,\sigma)=\frac{1}{\sigma}e^{-(x-\theta)/\sigma},\quad x\geq\theta,\ \sigma>0 \tag{2.1}$$

and the cumulative distribution function (cdf)

$$F(x;\theta,\sigma)=1-e^{-(x-\theta)/\sigma},\quad x\geq\theta,\ \sigma>0 \tag{2.2}$$

where θ is the warranty time, σ is the residual mean life, and $\mu=\theta+\sigma$ is the expected life-time.

Suppose n items are placed on a life-testing experiment and that the a_1th, a_2th, $\ldots$, a_sth failure-times are only made available, where

$$1\leq a_1 < ts < a_s \leq n.$$

That is,

$$X_{a_1:n}\leq X_{a_2:n}\leq\cdots\leq X_{a_s:n} \tag{2.3}$$

is the multiply Type-II censored sample available from (2.1)

Let $a_0=0$, $a_{s+1}=n+1$, $F(x_{a_0:n})=0$, $F(x_{a_{s+1}:n})=1$, then the likelihood function based on the multiply Type-II censored sample (2.3) is given by

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

$$= \frac{1}{\sigma^s} \frac{n!}{\prod_{j=1}^{s+1}(a_j - a_{j-1} - 1)!} [F(Z_{a_1:n})]^{a_1-1} \left[\prod_{j=2}^{s} [F(Z_{a_j:n}) - F(Z_{a_{j-1}:n})]^{a_j - a_{j-1} - 1} \right]$$

$$\times [1 - F(Z_{a_s:n})]^{n-a_s} \prod_{j=1}^{s} f(Z_{a_j:n}) \tag{2.4}$$

where $Z_{i:n} = (X_{i:n} - \theta)/\sigma$, and $f(z) = e^{-z}$ and $F(z) = 1 - e^{-z}$ are the pdf and the cdf of the standard exponential distribution respectively.

Balasubramanian and Balakrishnan (1992) derived the maximum likelihood estimator of θ and an approximate maximum likelihood estimator of σ as follows;

$$\hat{\theta} = \begin{cases} X_{1:n}, & \text{if } a_1 = 1 \\ X_{a_1:n} + \hat{\sigma} \ln\left(\frac{n - a_1 + 1}{n}\right), & \text{if } a_1 > 1 \end{cases}$$

and

$$\hat{\sigma} = \frac{1}{s - \sum_{i=1}^{s-1} (a_{i+1} - a_i - 1)\gamma_i} \left[\sum_{i=1}^{s} X_{a_i:n} + (n - a_s) X_{a_s:n} \right.$$
$$\left. + \sum_{i=1}^{s-1} (a_{i+1} - a_i - 1)[\delta_i X_{a_i:n} + (1 - \delta_i) X_{a_{i+1}:n}] - (n - a_1 + 1) X_{a_1:n} \right]$$

where

$$p_i = \frac{i}{n+1}, \quad q_i = 1 - p_i$$

$$\delta_i = \frac{q_{a_i}}{q_{a_i} - q_{a_i+1}} - \frac{q_{a_i} q_{a_i+1}}{(q_{a_i} - q_{a_i+1})^2} \ln\left(\frac{q_{a_i}}{q_{a_i+1}}\right)$$

$$\gamma_i = \frac{q_{a_i+1} \ln q_{a_i+1} - q_{a_i} \ln q_{a_i}}{q_{a_i} - q_{a_i+1}} + \delta_i \ln q_{a_i} + (1 - \delta_i) \ln q_{a_i+1}$$

Since this maximum likelihood estimator is the function of the maximum likelihood estimator of the scale parameter σ, we consider a simple estimator and an unbiased estimator of θ that is function of order statistics as follows;

$$\hat{\theta}_1 = X_{a_1:n}$$

and

$$\hat{\theta}_2 = \frac{1}{h(a_2) - h(a_1)} [h(a_2) X_{a_1:n} - h(a_1) X_{a_2:n}]$$

where

$$h(a) = \sum_{j=1}^{a} (n - j + 1)^{-1}$$

Also we can derive the other estimator by minimizing the mean squared error (MSE) among the class of estimators of the form $cX_{a_1:n} + (1-c)X_{a_2:n}$ where c is constant.

$$\hat{\theta}_3 = cX_{a_1:n} + (1-c)X_{a_2:n}$$

where

$$c = \frac{g(a_1) - g(a_2) - h^2(a_2) + h(a_1)h(a_2)}{g(a_1) - g(a_2) - [h(a_1) - h(a_2)]^2}$$

$$g(a) = \sum_{j=1}^{a}(n-j+1)^{-2}$$

Now, we can also derive the other unbiased estimator and biased estimator among the class of estimators of the form $c_1 X_{a_1:n} + c_2 \sum_{j=2}^{s} X_{a_j:n}$ where c_1 and c_2 are constants.

$$\hat{\theta}_4 = \frac{\sum_{j=2}^{s}[h(a_j)X_{a_1:n} - h(a_1)X_{a_j:n}]}{\sum_{j=2}^{s}h(a_j) - (s-1)h(a_1)}$$

$$\hat{\theta}_5 = [1-(s-1)c_1]X_{a_1:n} + c_1\sum_{j=2}^{s}X_{a_j:n}$$

where

$$c_1 = \frac{h(a_1)[(s-1)h(a_1) - \sum_{j=2}^{s}h(a_j)]}{V}$$

and

$$V = (s-1)^2[h^2(a_1) - g(a_1)] + \sum_{j=2}^{s} g(a_j) + 2\sum_{j=2}^{s-1}(s-j)g(a_j) + \left(\sum_{j=2}^{s}h(a_j)\right)^2$$
$$-2(s-1)h(a_1)\sum_{j=2}^{s}h(a_j)$$

From (2.4), the likelihood equation for σ is obtained as

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{1}{\sigma}\Bigg[s + (a_1-1)\frac{f(Z_{a_1:n})}{F(Z_{a_1:n})}Z_{a_1:n} - (n-a_s)Z_{a_s:n} - \sum_{j=1}^{s}Z_{a_j:n}$$
$$+\sum_{j=2}^{s}(a_j - a_{j-1} - 1)\frac{f(Z_{a_j:n})Z_{a_j:n} - f(Z_{a_{j-1}:n})Z_{a_{j-1}:n}}{F(Z_{a_j:n}) - F(Z_{a_{j-1}:n})}\Bigg]$$
$$= 0 \qquad (2.5)$$

Equation (2.5) does not admit an explicit solution for σ. But we can expand the functions $\frac{f(Z_{a_1:n})}{F(Z_{a_1:n})}$ and $\frac{f(Z_{a_j:n})Z_{a_j:n} - f(Z_{a_{j-1}:n})Z_{a_{j-1}:n}}{F(Z_{a_j:n}) - F(Z_{a_{j-1}:n})}$ as follows;

$$\frac{f(Z_{a_1:n})}{F(Z_{a_1:n})} \approx \alpha - \beta Z_{a_1:n} \tag{2.6}$$

$$\frac{Z_{a_j:n} f(Z_{a_j:n}) - Z_{a_{j-1}:n} f(Z_{a_{j-1}:n})}{F(Z_{a_j:n}) - F(Z_{a_{j-1}:n})} \approx \alpha_j + \beta_j Z_{a_j:n} + \gamma_j Z_{a_{j-1}:n} \tag{2.7}$$

where

$$\xi_r = F^{-1}(p_\gamma) = -\ln q_\gamma$$

$$\alpha = \frac{f(\xi_{a_1})}{p_{a_1}}\left[1 + \xi_{a_1} + \frac{f(\xi_{a_1})}{p_{a_1}}\xi_{a_1}\right]$$

$$\beta = \frac{f(\xi_{a_1})}{p_{a_1}{}^2}[p_{a_1} + f(\xi_{a_1})]$$

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

$$\beta_j = \frac{f(\xi_{a_j})}{p_{a_j} - p_{a_{j-1}}}\left[1 - \xi_{a_j} - \frac{\xi_{a_j} f(\xi_{a_j}) - \xi_{a_{j-1}} f(\xi_{a_{j-1}})}{p_{a_j} - p_{a_{j-1}}}\right]$$

$$\gamma_j = \frac{-f(\xi_{a_{j-1}})}{p_{a_j} - p_{a_{j-1}}}\left[1 - \xi_{a_{j-1}} - \frac{\xi_{a_j} f(\xi_{a_j}) - \xi_{a_{j-1}} f(\xi_{a_{j-1}})}{p_{a_j} - p_{a_{j-1}}}\right]$$

Upon using the approximation of (2.6) and (2.7) in the likelihood equation of (2.5), we derive two approximate maximum likelihood estimators of σ to be

$$\widehat{\sigma}_i = \frac{-B + \sqrt{B^2 - 4AC}}{2A}, \quad i = 1, 2 \tag{2.8}$$

where

$$A = s + \sum_{j=2}^{s}(a_j - a_{j-1} - 1)\alpha_j$$

$$\begin{aligned} B = \ & (a_1 - 1)\alpha X_{a_1:n} - (n - a_s)X_{a_s:n} - \sum_{j=1}^{s} X_{a_j:n} \\ & + \sum_{j=2}^{s}(a_j - a_{j-1} - 1)(\beta_j X_{a_j:n} + \gamma_j X_{a_{j-1}:n}) \\ & - [(a_1 - 1)\alpha - (n - a_s) - s + \sum_{j=2}^{s}(a_j - a_{j-1} - 1)(\beta_j + \gamma_j)]\ \widehat{\theta}_i \end{aligned}$$

$$C = (a_1 - 1)\beta(X_{a_1:n} - \widehat{\theta}_i)$$

3. Biases and mean squared errors

It is well known that the expectation and the variance of the ith order statistic, and the covariance of the ith and the jth order statistics from the two-parameter exponential distribution with pdf (2.1) are given by

$$E(X_{i:n}) = \mu + \sigma \sum_{j=1}^{i} (n-j+1)^{-1} \tag{2.9}$$

$$\mathrm{Var}(X_{i:n}) = \sigma^2 \sum_{k=1}^{i} (n-k+1)^{-2} = \mathrm{Cov}(X_{i:n}, X_{j:n}), \quad i \le j \tag{2.10}$$

Now from (2.9), we can obtain the means and variances of the estimators of the location parameter θ as follows

$$E(\hat{\theta}_1) = \theta + \sigma h(a_1)$$

$$E(\hat{\theta}_2) = \theta$$

$$E(\hat{\theta}_3) = \theta + [c h(a_1) + (1-c) h(a_2)]\sigma$$

$$E(\hat{\theta}_4) = \theta$$

$$E(\hat{\theta}_5) = \theta + \left[\{1-(s-1)c_1\} h(a_1) + c_1 \sum_{j=2}^{s} h(a_j) \right] \sigma$$

$$\mathrm{Var}(\hat{\theta}_1) = \sigma^2 g(a_1)$$

$$\mathrm{Var}(\hat{\theta}_2) = \mathrm{MSE}(\hat{\theta}_2)$$

$$= \frac{1}{[h(a_2) - h(a_1)]^2} [\{h(a_2) - 2h(a_1)\} h(a_2) g(a_1) + h^2(a_1) g(a_2)] \sigma^2$$

$$\mathrm{Var}(\hat{\theta}_3) = [c^2 g(a_1) + (1-c)^2 g(a_2) + 2c(1-c) g(a_1)] \sigma^2$$

$$\mathrm{Var}(\hat{\theta}_4) = \frac{1}{\left[\sum_{j=2}^{s} h(a_j) - (s-1) h(a_1) \right]^2} \left[g(a_1) \left(\sum_{j=2}^{s} h(a_j) \right)^2 \right.$$
$$+ h^2(a_1) \left\{ \sum_{j=2}^{s} g(a_j) + 2 \sum_{j=2}^{s-1} (s-j) g(a_j) \right\}$$
$$\left. - 2(s-1) h(a_1) g(a_1) \sum_{j=2}^{s} h(a_j) \right] \sigma^2$$

$$\mathrm{Var}(\hat{\theta}_5) = \left[[1-(s-1)c_1]^2 g(a_1) + c_1^2 \left(\sum_{j=2}^{s} g(a_j) + 2 \sum_{j=2}^{s-1} (s-j) g(a_j) \right) \right.$$
$$\left. + 2c_1 [1-(s-1)c_1](s-1) g(a_1) \right] \sigma^2$$

From the means and the variances, we can easily obtain the biases and the mean squared errors. The numerical values of the biases and the mean squared errors for various sample sizes n and censored samples $X_{a_1:n}, \ldots, X_{a_s:n}$ from the above equations about the means and the variances are given in Table 1.

From Table 1, $\hat{\theta}_4$ is an unbiased estimator and better than the other proposed estimators in the sense of the mean squared error.

Table 1. The biases and the mean squared errors for the proposed estimators of the location parameter θ.

$n=5$	$\widehat{\theta}_1$		$\widehat{\theta}_2$		$\widehat{\theta}_3$		$\widehat{\theta}_4$		$\widehat{\theta}_5$	
a_i	bias	MSE	bias	MSE	bias	MSE	bias	MSE	bias	MSE
1, 2, 3, 4, 5	0.200	0.080		0.080	0.100	0.060		0.045	0.043	0.048
1, 2, 3, 4	0.200	0.080		0.080	0.100	0.060		0.044	0.058	0.050
2, 3, 4, 5	0.450	0.305		0.305	0.225	0.204		0.129	0.129	0.153
1, 3, 4, 5	0.200	0.080	0	0.060	0.068	0.054	0	0.045	0.043	0.048
2, 3, 4	0.450	0.305		0.305	0.225	0.204		0.107	0.202	0.175
2, 3, 5	0.450	0.305		0.305	0.225	0.204		0.120	0.139	0.157
1, 3, 5	0.200	0.080		0.060	0.068	0.054		0.044	0.046	0.049

$n=10$	$\widehat{\theta}_1$		$\widehat{\theta}_2$	$\widehat{\theta}_3$		$\widehat{\theta}_4$	$\widehat{\theta}_5$	
a_i	bias	MSE	MSE	bias	MSE	MSE	bias	MSE
1, 2, 3. 4. 5. 6. 7. 8. 9. 10	0.100	0.020	0.020	0.050	0.015	0.011	0.010	0.011
1, 2, 3. 4. 5. 6. 7. 8. 9	0.100	0.020	0.020	0.050	0.015	0.011	0.012	0.011
1, 2, 3. 5. 6. 7. 8. 9. 10	0.100	0.020	0.020	0.050	0.015	0.011	0.010	0.011
1, 2, 3. 4. 5. 6. 7. 8	0.100	0.020	0.020	0.050	0.015	0.011	0.014	0.011
2, 3. 4. 6. 7. 8. 9. 10	0.211	0.067	0.067	0.106	0.045	0.026	0.024	0.027
1, 2, 3. 4. 7. 8. 9. 10	0.100	0.020	0.020	0.050	0.015	0.011	0.010	0.011
1, 2, 3. 4. 5. 6. 7	0.100	0.020	0.020	0.050	0.015	0.011	0.016	0.012
2, 3. 4. 5. 6. 7. 8	0.211	0.067	0.067	0.106	0.045	0.027	0.035	0.029
3. 4. 5. 7. 8. 9. 10	0.336	0.151	0.151	0.168	0.094	0.049	0.045	0.052
2, 3. 4. 5. 6. 7	0.211	0.067	0.067	0.106	0.045	0.027	0.043	0.030
1, 3. 5. 7. 9. 10	0.100	0.020	0.015	0.033	0.013	0.011	0.011	0.011
2, 3. 4. 7. 8. 9	0.211	0.067	0.067	0.106	0.045	0.026	0.029	0.028
2, 3. 4. 5. 6	0.211	0.067	0.067	0.106	0.045	0.026	0.056	0.032
2, 4. 6. 8. 10	0.211	0.067	0.045	0.071	0.037	0.026	0.026	0.028
2. 3. 4. 7. 8	0.211	0.067	0.067	0.106	0.045	0.025	0.035	0.029

References

1. 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

2. Balakrishnan, N. and Cohen, A. C.(1991). *Order Statistics and Inference: Estimation Methods.* Academic Press. Boston.

3. Balasubramanian, K. and Balakrishnan, N. (1992). Estimation for one- and two-parameter exponential distributions under multiple type-II censoring. *Statistical Papers,* 33, 203-216.

4. David, H. A. (1981). Order Statistics, Second Edition. John Wiley & Sons, New York.

5. Fei, H. and Kong, F. (1994). Interval Estimations for one- and two-parameter exponential distributions under multiple type-II censoring, *Communications in Statistics-Theory and Methods,*

23, 1717-1733.

6. Kambo, N. S. (1978). Maximum likelihood estimators of the location and scale parameters of the exponential distribution from a censored sample. *Communications in Statistics-Theory and Methods,* A7(12), 1129-1132.

7. Kang, S. B. (2003). Approximate MLEs for exponential distribution under multiple Type-II censoring, *Journal of Korean Data & Information Science Society,* 14, 983-988.

8. Kang, S. B. and Cho, Y. S. (1998). Approximate MLE of exponential distribution for truncated samples, *Journal of Information & Optimization Sciences,* 19, 403-409.

9. Kang, S. B., Suh, Y. S., and Cho, Y. S. (1997). Estimation of the parameters in an exponential distribution with type-II censoring, *The Korean Communications in Statistics,* 4, 929-941.

10. Lawless, J. F.(1982). *Statistical Models and Methods for Lifetime Data.* John Wiley & Sons, New York.

11. Upadhyay, S. K., Singh, U., and Shastri, V. (1996). Estimation of exponential parameters under multiply type II censoring, *Communications in Statistics-Simulation,* 25, 801-815.