East Asian Math J 13 (1997), No 2, pp 307-318

# APPROXIMATE ESTIMATION IN TWO-PARAMETER EXPONENTIAL DISTRIBUTION BASED ON TYPE-II CENSORED SAMPLE

JE YOON LEE

## 1. Introduction

The random variable $X$ has an exponential distribution if it has a probability density function(pdf) of forms:

$$f(x) = \frac{1}{\sigma}\exp\left(-\frac{x-\mu}{\sigma}\right), \qquad x \geq \mu, \sigma > 0 \tag{1}$$

where $\mu$ and $\sigma$ are the location and scale parameters, respectively.

Lloyd(1952) described a method of obtaining the best linear unbiased estimators(BLUEs) of the parameters of exponential distribution, using order statistics. Gupta(1952) proposed estimation of the mean and standard deviation of a normal population from a censored sample. The approximate maximum likelihood estimation method was first developed by Balakrishnan(1989a,b) for the purpose of providing the explicit estimators of the scale parameter in the Rayleigh distribution and the mean and standard deviation in the normal distribution with censoring. 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 BLUE and the optimum unbiased absolute estimator. Some historical remarks and a good summary of the approximate maximum likelihood estimation may be found in Balakrishnan and Cohen(1991). Recently Kang et al.(1997) studied the minimum risk estimator(MRE) and AMLE of the parameters in two-parameter exponential distribution based on Type-II censored sample.

Received Sep 19, 1997

In this paper, we obtain more simple MRE of the location parameter based on only two order statistics $X_{r+1,n}$ and $X_{r+2.n}$, and the AMLE of the scale parameter by using the proposed MRE of the location parameter. We also compare the proposed estimators with the BLUE which was proposed by Lloyd(1952) and the simplified linear estimator in the sense of the mean squared error (MSE).

## 2. Preliminary

Consider two-parameter exponential distribution with density function (1) and cumulative distribution function(cdf).

$$F(x) = \begin{cases} 1 - \exp\left(-\dfrac{x-\mu}{\sigma}\right), & x \geq \mu \\ 0, & x < \mu. \end{cases} \tag{2}$$

Let us consider an experiment in which $n$ exponential components are put to test simultaneously at time $x = 0$, and the failure times of there components are recorded. Suppose that some initial observations are censored(possibly because of some failures during the time when some checks and adjustments are being made on the devices) and some final observations are also censored(possibly because the experimenter terminates the experiment before all components have failed). Let

$$X_{r+1\ n} \leq X_{r+2\cdot n} \leq \cdots \leq X_{n-s.n} \tag{3}$$

be the available Type-II censored sample from the exponential distribution with pdf (1), where the first $r$ and the last $s$ observations are censored.

It is well known that the expectation, the variance and the covariance of the $i$th order statistic $X_{i.n}$ from the two-parameter exponential distribution are given by

$$E(X_{i:n}) = \mu + \sigma \sum_{j=1}^{i} (n-j+1)^{-1} \tag{4}$$

and

$$\begin{aligned}\mathrm{Var}(X_{i:n}) &= \sigma^2 \sum_{k=1}^{i}(n-k+1)^{-2} \\ &= \mathrm{Cov}(X_{i:n}, X_{j:n}), \quad i < j.\end{aligned} \tag{5}$$

For the standard exponential distribution, we will use the notations as follows;

$$E(X_{i:n}) = \alpha_{i:n} = \sum_{l=n-i+1}^{n} 1/l \quad (1 \leq i \leq n)$$

and

$$\mathrm{Cov}(X_{i:n}, X_{j:n}) = \beta_{i,j:n} = \sum_{l=n-i+1}^{n} 1/l^2 \quad (1 \leq i \leq j \leq n).$$

Lloyd(1952) derived the BLUEs of $\mu$ and $\sigma$ as follows;

$$\widehat{\mu}_{BLUE} = \sum_{i=r+1}^{n-s} a_i X_{i:n} \tag{6}$$

and

$$\widehat{\sigma}_{BLUE} = \sum_{i=r+1}^{n-s} b_i X_{i:n} \tag{7}$$

where

$$a_i = \begin{cases} 1 + \dfrac{(n-r-1)}{(n-r-s-1)} \displaystyle\sum_{l=n-r}^{n} \frac{1}{l}, & \text{for } i = r+1 \\ -\dfrac{1}{(n-r-s-1)} \displaystyle\sum_{l=n-r}^{n} \frac{1}{l}, & \text{for } r+2 \leq i \leq n-s-1 \\ -\dfrac{(s+1)}{(n-r-s-1)} \displaystyle\sum_{l=n-r}^{n} \frac{1}{l}, & \text{for } i = n-s \end{cases}$$

and

$$b_i = \begin{cases} -\dfrac{(n-r-1)}{(n-r-s-1)}, & \text{for } i = r+1 \\ \dfrac{1}{(n-r-s-1)}, & \text{for } r+2 \le i \le n-s-1 \\ \dfrac{(s+1)}{(n-r-s-1)}, & \text{for } i = n-s. \end{cases}$$

The variances of the BLUEs(see Balakrishnan and Cohen(1991)) are given by

$$\text{Var}(\widehat{\mu}_{BLUE}) = \sigma^2 \left[ \frac{1}{n-r-s-1} \left( \sum_{l=n-r}^{n} 1/l \right)^2 + \sum_{l=n-r}^{n} (1/l)^2 \right] \tag{8}$$

and

$$\text{Var}(\widehat{\sigma}_{BLUE}) = \frac{\sigma^2}{n-r-s-1}. \tag{9}$$

Gupta(1952) proposed the simplified linear estimators(SLE) for $\mu$ and $\sigma$ obtained from the BLUEs of $\mu$ and $\sigma$ simply by replacing the variance-covariance matrix by an identity matrix.

The SLE of $\mu$ is given by(see Balakrishnan and Cohen(1991) )

$$\widehat{\mu}_{SLE} = \sum_{j=r+1}^{n-s} c_j X_{j\,n} \tag{10}$$

where

$$c_j = \frac{\sum_{i=r+1}^{n-s} \alpha_{i:n}^2 - \alpha_{j:n} \sum_{i=r+1}^{n-s} \alpha_{i:n}}{(n-r-s) \sum_{i=r+1}^{n-s} (\alpha_{i:n} - \bar{\alpha})^2}$$

$$= \frac{1}{n-r-s} - \frac{\bar{\alpha}(\alpha_{j:n} - \bar{\alpha})}{\sum_{i=r+1}^{n-s} (\alpha_{i:n} - \bar{\alpha})^2}, \quad r+1 \leq j \leq n-s$$

with $\bar{\alpha}$ as defined in

$$\bar{\alpha} = \frac{1}{(n-r-s)} \sum_{i=r+1}^{n-s} \alpha_{i:n}.$$

Similarly, the BLUE of $\sigma$ simplifies to

$$\widehat{\sigma}_{SLE} = \sum_{j=r+1}^{n-s} d_j X_{j:n} \tag{11}$$

where

$$d_j = \frac{(n-r-s)\alpha_{j:n} - \sum_{i=r+1}^{n-s} \alpha_{i:n}}{(n-r-s) \sum_{i=r+1}^{n-s} (\alpha_{i:n} - \bar{\alpha})^2}$$

$$= \frac{(\alpha_{j:n} - \bar{\alpha})}{\sum_{i=r+1}^{n-s} (\alpha_{i:n} - \bar{\alpha})^2}, \quad r+1 \leq j \leq n-s$$

The variances of these simplified linear estimators may be obtained as

$$\mathrm{Var}(\widehat{\mu}_{SLE}) = \sigma^2 \sum_{i=r+1}^{n-s} \sum_{j=r+1}^{n-s} c_i c_j \beta_{i,j:n} \tag{12}$$

and

$$\mathrm{Var}(\widehat{\sigma}_{SLE}) = \sigma^2 \sum_{i=r+1}^{n-s} \sum_{j=r+1}^{n-s} d_i d_j \beta_{i,j:n} \tag{13}$$

## 3. MRE and approximate estimation

In this section, we propose the MRE of the location parameter and the AMLE of the scale parameter by using the proposed MRE of the location parameter. The likelihood function based on the Type-II censored sample in (3) is given by

$$\begin{aligned} L = & \frac{n!}{r!s!}[F(X_{r+1:n};\mu,\sigma)]^r[1-F(X_{n-s:n};\mu,\sigma)]^s \\ & \times \prod_{i=r+1}^{n-s} f(X_{i:n};\mu,\sigma), \quad X_{r+1:n} \geq \mu \end{aligned} \tag{14}$$

We denote $Z_{i:n} = (X_{i:n}-\mu)/\sigma$. The equality (14) can be written as

$$\begin{aligned} L = & \frac{n!}{r!s!}\sigma^{-A}[F(Z_{r+1:n})]^r[1-F(Z_{n-s:n})]^s \\ & \times \prod_{i=r+1}^{n-s} f(Z_{i:n}), \qquad Z_{r+1:n} \geq 0 \end{aligned} \tag{15}$$

where $A = n - r - s$ is the size of the censored sample (3), and $f(z)$ and $F(z)$ are the pdf and cdf of the standard exponential distribution, respectively.

Kang et al.(1997) proposed the estimator $\widehat{\mu} = X_{r+1:n}$ of the location parameter and they also proposed the combined estimater of $X_{r+1:n}, \cdots X_{n-s:n}$.

First, we propose more simple estimator of the form $aX_{r+1:n} + (1-a)X_{r+2:n}$ ,where $a$ is a constant which minimize the MSE of the proposed estimator $aX_{r+1:n} + (1-a)X_{r+2:n}$.

We can obtain the MRE as follows;

$$\widehat{\mu}_{MRE} = aX_{r+1\cdot n} + (1-a)X_{r+2\cdot n} \tag{16}$$

where

$$a = \frac{1}{2}[1 + (n-r-1)h(r+2)]$$

and

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

We can obtain the mean and variance of the MRE as follows;

$$E(\widehat{\mu}_{MRE}) = \mu + \sigma[ah(r+1) + (1-a)h(r+2)] \tag{17}$$

and

$$\mathrm{Var}(\widehat{\mu}_{MRE}) = \sigma^2[a(2-a)g(r+1) + (1-a)^2 g(r+2)] \tag{18}$$

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

Also, we will obtain the AMLE of the scale parameter. First, we differentiate the logarithm of the likelihood function(15) for $\sigma$ as follows;

$$\begin{aligned}\frac{\partial \ln L}{\partial \sigma} = -\frac{1}{\sigma}\Bigg[A + rZ_{r+1\cdot n}\frac{f(Z_{r+1\cdot n})}{F(Z_{r+1\cdot n})} - sZ_{n-s\cdot n} \\ \times \frac{f(Z_{n-s\cdot n})}{1-F(Z_{n-s\cdot n})} + \sum_{i=r+1}^{n-s}\frac{f'(Z_{i\cdot n})}{f(Z_{i\cdot n})}\Bigg] = 0.\end{aligned} \tag{19}$$

Equation(19) does not admit an explicit solution for $\sigma$.

But since $\frac{f(Z_{n-s\cdot n})}{1-F(Z_{n-s\cdot n})} = 1$ and $\frac{f'(Z_{i\cdot n})}{f(Z_{i\cdot n})} = -1$, we can expand the function $\frac{f(Z_{r+1\cdot n})}{F(Z_{r+1\cdot n})}$ appearing in (19) to Taylor series around the point $\xi_{r+1} = F^{-1}(p_{r+1}) = -\ln(q_{r+1})$ and then approximate it by

$$\frac{f(Z_{r+1\cdot n})}{F(Z_{r+1\cdot n})} \simeq \alpha - \beta Z_{r+1\cdot n} \tag{20}$$

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

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

and

$$\beta = \frac{f(\xi_{r+1})}{p_{r+1}^2}[p_{r+1} + f(\xi_{r+1})].$$

Now making use of the approximate expression in (20), we obtain the approximate likelihood equation of (19) as follows;

$$\begin{aligned}\frac{\partial \ln L}{\partial \sigma} \simeq \frac{\partial \ln L^*}{\partial \sigma} = -\frac{1}{\sigma}\Bigg[A + rZ_{r+1\,n}(\alpha - \beta Z_{r+1\,n}) \\ - sZ_{n-s.n} - \sum_{i=r+1}^{n-s} Z_{i.n}\Bigg] = 0.\end{aligned} \tag{21}$$

Since $Z_{i.n} = (X_{i\,n} - \mu)/\sigma$, using the MRE $\widehat{\mu}_{MRE}$ of $\mu$ and solving equation(21) for $\sigma$, we can derive the AMLE of $\sigma$ as follows;

$$\widehat{\sigma}_{AMLE} = \frac{1}{2A}(-B + \sqrt{B^2 + 4AC}) \tag{22}$$

where

$$\begin{aligned}B =& [r\alpha(1-a) + a(A+s)]X_{r+1\,n} \\ &+ (1-a)(A+s-r\alpha)X_{r+2.n} \\ &- sX_{n-s\,n} - \sum_{i=r+1}^{n-s} X_{i\,n}\end{aligned}$$

and

$$C = r\beta(1-a)^2(X_{r+1.n} - X_{r+2.n})^2$$

Now, we can obtain the MSE of $\widehat{\mu}_{MRE}$ as follows;

$$\begin{aligned}\text{MSE}(\widehat{\mu}_{MRE}) =& \sigma^2[a(2-a)g(r+1) + (1-a)^2 g(r+2) \\ &+ a^2h^2(r+1) + 2a(1-a)h(r+1)h(r+2) \\ &+ (1-a)^2h^2(r+2)]\end{aligned}$$

Since $\lim_{n\to\infty} a = \frac{1}{2}(3+r)$, $\lim_{n\to\infty} g(r+1) = 0$ and $\lim_{n\to\infty} h(r+1) = 0$, we can obtain the limit of the MSE of $\widehat{\mu}_{MRE}$ as follows;

$$\lim_{n\to\infty} \mathrm{MSE}(\widehat{\mu}_{MRE}) = 0$$

Therefore, $\widehat{\mu}_{MRE}$ is a consistent estimator.

We can not obtain the closed form of the MSE for the AMLE of the scale parameter, so we can simulate the values of the relative MSE of the proposed estimators based on 10,000 Monte Carlo runs for $n = 5, 6, 9$ and over various choices of censoring IMSL subroutine. These values are given in Table 3.1 and 3.2

From Table 3.1 and 3.2, we obtain the following results in the sense of MSE;

(i) The MRE $\widehat{\mu}_{MRE}$ is more efficient than the SLE $\widehat{\mu}_{SLE}$.

(ii) The BLUE $\widehat{\mu}_{BLUE}$ is more efficient than the other estimator, but $\widehat{\mu}_{MRE}$ is more simple in formula.

(iii) The AMLE of the scale parameter $\widehat{\sigma}_{AMLE}$ is more efficient than the BLUE $\widehat{\sigma}_{BLUE}$ and SLE $\widehat{\sigma}_{SLE}$. So the proposed AMLE $\widehat{\sigma}_{AMLE}$ of the scale parameter is not only closed form of the random sample, but also very good estimator.

**Table 3.1.** The relative mean squared errors for the estimators of the location parameter $\mu$

| $n$ | $r$ | $s$ | $\widehat{\mu}_{BLUE}$ | $\widehat{\mu}_{SLE}$ | $\widehat{\mu}_{MRE}$ |
|---|---|---|---|---|---|
| 5 | 0 | 0 | 0.05173 | 0.13840 | 0.06210 |
| 5 | 0 | 1 | 0.05492 | 0.08609 | 0.06079 |
| 5 | 1 | 0 | 0.16518 | 0.30617 | 0.19874 |
| 5 | 1 | 1 | 0.20228 | 0.24527 | 0.20615 |
| 6 | 0 | 0 | 0.03270 | 0.11400 | 0.04050 |
| 6 | 0 | 1 | 0.03476 | 0.06739 | 0.04131 |
| 6 | 0 | 2 | 0.03608 | 0.05046 | 0.04044 |
| 6 | 1 | 0 | 0.10120 | 0.22350 | 0.13428 |
| 6 | 1 | 1 | 0.11259 | 0.15792 | 0.13463 |
| 6 | 1 | 2 | 0.13933 | 0.15605 | 0.13765 |
| 6 | 2 | 0 | 0.26236 | 0.43678 | 0.32297 |
| 6 | 2 | 1 | 0.31776 | 0.36534 | 0.31954 |
| 6 | 2 | 2 | 0.52525 | 0.52525 | 0.31762 |
| 9 | 0 | 0 | 0.01425 | 0.08120 | 0.01867 |
| 9 | 0 | 1 | 0.01400 | 0.04445 | 0.01830 |
| 9 | 0 | 2 | 0.01387 | 0.03334 | 0.01793 |
| 9 | 0 | 3 | 0.01546 | 0.02852 | 0.01880 |
| 9 | 1 | 0 | 0.03663 | 0.12227 | 0.05579 |
| 9 | 1 | 1 | 0.03775 | 0.07686 | 0.05544 |
| 9 | 1 | 2 | 0.03904 | 0.06256 | 0.05483 |
| 9 | 1 | 3 | 0.04370 | 0.05867 | 0.05743 |
| 9 | 2 | 0 | 0.07305 | 0.19332 | 0.12159 |
| 9 | 2 | 1 | 0.07951 | 0.13382 | 0.12288 |
| 9 | 2 | 2 | 0.08611 | 0.11331 | 0.12138 |
| 9 | 2 | 3 | 0.09570 | 0.11225 | 0.12029 |
| 9 | 3 | 0 | 0.13331 | 0.27652 | 0.22610 |
| 9 | 3 | 1 | 0.14538 | 0.20732 | 0.22244 |
| 9 | 3 | 2 | 0.17438 | 0.20328 | 0.22508 |
| 9 | 3 | 3 | 0.22441 | 0.24109 | 0.22317 |

**Table 3.2.** The relative mean squared errors for the estimators of the scale parameter $\sigma$

| $n$ | $r$ | $s$ | $\widehat{\sigma}_{BLUE}$ | $\widehat{\sigma}_{SLE}$ | $\widehat{\sigma}_{AMLE}$ |
|---|---|---|---|---|---|
| 5 | 0 | 0 | 1.01369 | 1.36052 | 0.90021 |
| 5 | 0 | 1 | 1.32022 | 1.46961 | 1.12116 |
| 5 | 1 | 0 | 1.30958 | 1.69339 | 1.10757 |
| 5 | 1 | 1 | 1.89817 | 2.07676 | 1.44092 |
| 6 | 0 | 0 | 0.80994 | 1.15057 | 0.72611 |
| 6 | 0 | 1 | 0.96648 | 1.11889 | 0.85846 |
| 6 | 0 | 2 | 1.33928 | 1.41935 | 1.13671 |
| 6 | 1 | 0 | 0.99467 | 1.34930 | 0.88587 |
| 6 | 1 | 1 | 1.35041 | 1.50941 | 1.13308 |
| 6 | 1 | 2 | 2.09668 | 2.18157 | 1.54865 |
| 6 | 2 | 0 | 1.36524 | 1.73402 | 1.17102 |
| 6 | 2 | 1 | 1.95271 | 2.06531 | 1.49538 |
| 6 | 2 | 2 | 4.18445 | 4.18445 | 2.27575 |
| 9 | 0 | 0 | 0.49678 | 0.77011 | 0.46714 |
| 9 | 0 | 1 | 0.55307 | 0.67700 | 0.51771 |
| 9 | 0 | 2 | 0.65861 | 0.76261 | 0.61124 |
| 9 | 0 | 3 | 0.82387 | 0.90964 | 0.73843 |
| 9 | 1 | 0 | 0.57007 | 0.84893 | 0.54329 |
| 9 | 1 | 1 | 0.68545 | 0.82345 | 0.63490 |
| 9 | 1 | 2 | 0.78209 | 0.87261 | 0.71416 |
| 9 | 1 | 3 | 1.00611 | 1.09314 | 0.88324 |
| 9 | 2 | 0 | 0.67281 | 1.00082 | 0.64679 |
| 9 | 2 | 1 | 0.81095 | 0.97364 | 0.75364 |
| 9 | 2 | 2 | 0.98981 | 1.07459 | 0.90044 |
| 9 | 2 | 3 | 1.32770 | 1.40656 | 1.10857 |
| 9 | 3 | 0 | 0.79314 | 1.08609 | 0 77193 |
| 9 | 3 | 1 | 0.98117 | 1.11523 | 0.92105 |
| 9 | 3 | 2 | 1.33998 | 1.42328 | 1.15761 |
| 9 | 3 | 3 | 2.01258 | 2.07485 | 1.48761 |

## References

1. Balakrishnan, N., *Approximate MLE of the scale parameter of the Rayleigh distribution with censoring*, IEEE Transactions on the Reliability **38** (1989a), 355–357.
2. ———, *Approximate maximum likelihood estimation of the mean and standard deviation of the normal distribution based on Type-censored sample*, Journal of Statistic-Computation and Simulation **32** (1989b), 137–148.
3. ———, *On the maximum likelihood estimation of the location and scale parameters of exponential distribution based on multiplying Type-II censored samples*, Journal of Applied Statistics **17** (1990), 55–61.
4. Balakrishnan, N and Cohen, A C , *Order Statistics and Inference:Estimation methods*, Academic Press, San Diego, 1991.
5. Gupta, A. K., *Estimation of the mean and standard deviation of a normal population from a censored sample*, Biometrika **39** (1952), 260–273.
6. Kang, S B , *Approximate MLE for the scale parameters of the double exponential distribution based on Type-II censoring*, Journal of the Korean Mathematical Society **33** no. 1 (1996), 69–79
7. Kang, S B ,Suh, Y S. and Cho, Y S., *Minimum risk estimator in an exponential distribution with Type-II censoring*, To appear in American Statistician (1997).
8. Lloyd, E H , *Least-squares estimation of location and scale parameters using order statistics*, Biometrika **39** (1952), 88–95

Department of Mathematics
University of Ulsan
Ulsan 680-749, Korea