The Korean Communications
in Statistics Vol. 5, No. 3, 1998
pp. 879-885

Approximate MLE for Singly Truncated Normal Distribution

Suk-Bok Kang[1] and Young-Suk Cho[2]

Abstract

In this paper, we propose the approximate maximum likelihood estimators (AMLE) of the location and the scale parameter of the singly left truncated normal distribution. We compare the proposed estimators with the simpler estimators (SE) in terms of the mean squared error (MSE) through Monte Carlo methods.

1. Introduction

Truncated samples are those from which certain population values are entirely excluded. It is perhaps more accurate to state that truncation occurs to populations, and samples described as being truncated are in fact samples from truncated population. Truncated samples are classified according to whether points of truncation are unknown or known. If these points are unknown, they become additional parameters to be estimated from sample data. Cohen (1959) obtained the simpler estimators for the parameters in the normal distribution when samples are singly censored or truncated. Chiu and Leung (1981) studied a graphical method for estimating the parameters of a truncated normal distribution. The theory and applications of truncated and censored samples are summarized by Cohen (1991).

The normal distribution was first published by Abraham de Moivre in 1733 as an approximation for the distribution of the sum of binomial random variables. It is the single most important distribution in probability and statistic. A random variable X follows the normal distribution with mean μ and variance σ^2 if it has the pdf and the cdf as follows;

$$\begin{aligned} f(x;\mu,\sigma) &= \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], \quad -\infty < x < \infty, 0 < \sigma, \\ F(x;\mu,\sigma) &= \int_{-\infty}^{x} f(y;\mu,\sigma)dy. \end{aligned} \tag{1.1}$$

The pdf and the cdf of the standard normal distribution $Z=(X-\mu)/\sigma$ become

1) Professor, Department of Statistics, Yeungnam University, Kyongsan, 712-749, Korea
2) Adjunct Assistant professor, Department of Statistics, Yeungnam University, Kyongsan, 712-749, Korea

$$\phi(z;0,1) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right), \qquad -\infty < z < \infty, 0 < \sigma,$$
$$\Phi(z;0,1) = \int_{-\infty}^{z} \phi(y;0,1)dy. \tag{1.2}$$

We consider the case that the normal distribution (1.1) is truncated on the left at $x = T$. The pdf of the singly left truncated normal distribution X_T becomes

$$f_{X_T}(x;\mu,\sigma) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}[1-F(T)]} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], & T \le x, \\ 0, & \text{elsewhere}. \end{cases} \tag{1.3}$$

The variance and the expected value of X_T are given by

$$\mathrm{Var}(X_T) = \sigma^2[1 - Q(Q-\xi)],$$
$$E(X_T) = \sigma(Q-\xi) + T \tag{1.4}$$

where $Q = Q(\xi) = \phi(\xi)/(1-\phi(\xi))$ and $\xi = (T-\mu)/\sigma$.

The approximate maximum likelihood estimating method was first developed by Balakrishnan (1989a,b) for the purpose of providing the explicit estimators of the scale parameter in the Rayleigh distribution. He derived AMLE of the scale parameter from Type-II double censored sample in two-dimensional Rayleigh distribution. Kang (1996) obtained the 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 efficienct than the BLUE and the optimum unbiased absolute estimator. Kang and Cho (1997a,b) derived the minimum risk estimator and the AMLE of the scale parameter of the one-parameter exponential distribution under general Type-II censored and progressive Type-II censored sample, and the proposed estimators were compared in terms of the mean squared error through Nelson's data. Kang, et al. (1997) obtained the minimum risk estimator (MRE) and the AMLE of parameters of the two-parameter exponential distribution based on Type-II censoring. Woo, et al. (1998) obtained the AMLE for Rayleigh distribution in singly right censored samples.

We consider the AMLE of the location and the scale parameter of the singly left truncated normal distribution. We compare the proposed estimators with the SEs in terms of the mean squared error through Monte Carlo methods.

2. Parameter Estimation in Singly Left Truncated

Moment estimators (ME) are obtained by equating the mean $\overline{X}$ and the variance $S^2 = \sum_{i=1}^{n}(X_i - \overline{X})^2/n$ of the truncated sample to the mean and variance of the truncated normal population.

From the expressions in (1.4), the estimating equations become

$$S^2 = \sigma^2[1 - Q(Q-\xi)],$$
$$\overline{X} = \sigma(Q-\xi) + T. \tag{2.1}$$

From $\xi = (T-\mu)/\sigma$, the estimator for μ is given by

$$\hat{\mu}_{ME} = T - \hat{\sigma}_{ME}\ \hat{\xi}. \tag{2.2}$$

From the two equations of (2.1), we can eliminate σ^2 as follows;

$$\frac{S^2}{(\overline{X}-T)^2} = \frac{1-Q(Q-\xi)}{(Q-\xi)^2} \tag{2.3}$$

and from the second equation (2.1), we obtain

$$\sigma = \frac{\overline{X}-T}{Q-\xi}.$$

From the above results, we can obtain the estimator of σ as follows;

$$\hat{\sigma}_{ME} = \frac{\overline{X}-T}{Q(\hat{\xi})-\hat{\xi}} \tag{2.4}$$

where $\hat{\xi}$ is the solution of (2.3) and $Q(\hat{\xi}) = \phi(\hat{\xi})/(1-\phi(\hat{\xi}))$.

Cohen (1959) derived the simpler estimators as follows;

$$\hat{\sigma}_{SE} = \sqrt{S^2 + \theta(\hat{\xi})(\overline{X}-T)^2} \tag{2.5}$$

and

$$\hat{\mu}_{SE} = \overline{X} - \theta(\hat{\xi})(\overline{X}-T) \tag{2.6}$$

where $\hat{\xi}$ is the solution of (2.3) and $\theta(\hat{\xi}) = Q(\hat{\xi})/(Q(\hat{\xi})-\hat{\xi})$. He also obtain the numerical values of the auxiliary estimation function $\theta(\xi)$ to facilitate the practical application of the estimators (2.5) and (2.6).

The likelihood function of a random sample of size n from a singly left truncated normal distribution with pdf (1.1) is

$$L = \frac{1}{[1-\Phi(\xi)]^n}\ \frac{1}{\sigma^n(\sqrt{2\pi})^n}\ \exp\left[-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2\right],\ \xi = (T-\mu)/\sigma. \tag{2.7}$$

We obtain likelihood equations for μ and σ from (2.6) to be

$$\frac{\partial \ln L}{\partial \mu} = -\frac{n}{\sigma}\ \frac{\phi(\xi)}{1-\Phi(\xi)} + \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) \tag{2.8}$$

and

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma}\ \frac{\phi(\xi)}{1-\Phi(\xi)}\ \xi - \frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i-\mu)^2. \tag{2.9}$$

The likelihood equations (2.8) and (2.9) do not admit explicit solutions except in complete sample case (i.e., $T \rightarrow -\infty$). But, we may expand the function $\phi(\xi)/(1-\Phi(\xi))$ in (2.8) and (2.9) in Taylor series around the real point p, and approximate it by

$$\frac{\phi(\xi)}{1-\Phi(\xi)} \approx \alpha + \beta\xi \tag{2.10}$$

where

$$\alpha = \frac{\phi(p)}{1-\Phi(p)} - \left[\frac{-p\phi(p)(1-\Phi(p)) + \phi(p)^2}{(1-\Phi(p))^2} \right] p$$

and

$$\beta = \frac{-p\phi(p)(1-\Phi(p)) + \phi(p)^2}{(1-\Phi(p))^2}.$$

By using equation (2.10), we may approximate the likelihood equations (2.8) and (2.9) by

$$\begin{aligned} \frac{\partial \ln L}{\partial \mu} \approx \frac{\partial \ln L^*}{\partial \mu} &= -\frac{n}{\sigma}(\alpha + \beta\xi) + \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu) \\ &= 0 \end{aligned} \tag{2.11}$$

and

$$\begin{aligned} \frac{\partial \ln L}{\partial \sigma} \approx \frac{\partial \ln L^*}{\partial \sigma} &= -\frac{n}{\sigma}(\alpha + \beta\xi)\xi - \frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^{n} (x_i - \mu)^2 \\ &= 0. \end{aligned} \tag{2.12}$$

Upon solving equations (2.11) and (2.12), we obtain the AMLEs of μ and σ as follows;

$$\hat{\mu}_{AMLE} = a - b\,\hat{\sigma}_{AMLE} \tag{2.13}$$

and

$$\hat{\sigma}_{AMLE} = \frac{B + \sqrt{B^2 + 4nC}}{2n} \tag{2.14}$$

where

$$a = \frac{\sum_{i=1}^{n} X_i - n\beta T}{n(1-\beta)},$$

$$b = \alpha/(1-\beta),$$

$$B = 2b \sum_{i=1}^{n} (X_i - a) - n\alpha(T-a) - 2n\beta b(T-a),$$

and

$$C = -n\beta(T-a)^2 + \sum_{i=1}^{n} (X_i - a)^2.$$

The conditional bias of $\hat{\mu}_{AMLE}$ can be computed exactly from (2.13). But it is not possible to determine the conditional bias of $\hat{\sigma}_{AMLE}$ exactly. However, it may be evaluated approximately by $E(\partial \ln L^* / \partial\sigma) / E(-\partial^2 \ln L^* / \partial\sigma^2)$ (see Kendall and Stuart (1973)). Moreover, we obtain from equations (2.11) and (2.12)

$$E\left(-\frac{\partial^2 \ln L^*}{\partial\mu^2}\right) = \frac{n(1-\beta)}{\sigma^2}, \quad E\left(-\frac{\partial^2 \ln L^*}{\partial\mu\partial\sigma}\right) = \frac{V_1}{\sigma^2}, \tag{2.15}$$

and

$$E\left(-\frac{\partial^2 \ln L^*}{\partial\sigma^2}\right) = \frac{V_2}{\sigma^2} \tag{2.16}$$

where

$$V_1 = \frac{n}{\sigma}\left[2E(X_T) - 2\mu - \alpha\sigma - 2\beta T + 2\beta\mu\right]$$

and

$$V_2 = \frac{n}{\sigma^2}\left[3(E(X_T^2) - 2\mu E(X_T) + \mu^2) - 2\alpha(T-\mu)\sigma - 3\beta(T-\mu)^2 - \sigma^2\right].$$

From the above expressions, we can obtain the approximate variances of the AMLEs as follows;

$$\mathrm{Var}(\hat{\mu}_{AMLE}) \approx \frac{\sigma^2 V_2}{n(1-\beta)V_2 - V_1^2}, \quad \mathrm{Var}(\hat{\sigma}_{AMLE}) \approx \frac{\sigma^2 n(1-\beta)}{n(1-\beta)V_2 - V_1^2}, \tag{2.17}$$

and

$$\mathrm{Cov}(\hat{\mu}_{AMLE}, \hat{\sigma}_{AMLE}) \approx -\frac{\sigma^2 V_1}{n(1-\beta)V_2 - V_1^2}. \tag{2.18}$$

3. The Simulated Result

Random numbers of the two-parameter normal distribution were generated by IMSL subroutine RNNOA and transformed $\mu + \sigma *$ RNNOA. The SE $\hat{\sigma}_{SE}$ and the SE $\hat{\mu}_{SE}$ was obtained by Fortran program and Newton-Raphson method.

We investigated the MSEs of the $\hat{\mu}_{AMLE}$ and the $\hat{\sigma}_{AMLE}$ for each $\mu = 0.0, 1.0$, $\sigma = 1.0, 2.0$, and truncated on the left at point $T = -3.0, -2.0, -1.5$. The simulation procedure is repeated 1,000 times for each sample sizes $n=20(20)60$.

From Table 1, the proposed estimators $\hat{\mu}_{AMLE}$ and $\hat{\sigma}_{AMLE}$ are more efficient than the SEs in terms of the mean squared error except ξ and n are large. We can also see that the MSE increases as ξ increases and is same when ξ and σ are fix. Cohen's simpler esrimators are very complicated but the proposed estimators $\hat{\mu}_{AMLE}$ and $\hat{\sigma}_{AMLE}$ have explicit forms. So the proposed estimators are useful to estimate the location and scale parameter in normal distribution that is truncated on the left at point T.

Table 1. The mean square errors of several estimators

$\xi=(T-\mu)/\sigma$		n	MSE($\hat{\mu}_{SE}$)	MSE($\hat{\mu}_{AMLE}$)	MSE($\hat{\sigma}_{SE}$)	MSE($\hat{\sigma}_{AMLE}$)
-3.0	$T=-3.0$	20	.052444	.050276	.027574	.024883
	$\mu=0.0$	40	.023367	.022502	.014015	.012548
	$\sigma=1.0$	60	.017117	.016549	.008770	.007935
-3.0	$T=-2.0$	20	.052444	.050276	.027574	.024883
	$\mu=1.0$	40	.023367	.022502	.014015	.012548
	$\sigma=1.0$	60	.017117	.016549	.008770	.007935
-2.0	$T=-2.0$	20	.071704	.048195	.041060	.029425
	$\mu=0.0$	40	.031715	.025496	.019016	.015883
	$\sigma=1.0$	60	.020513	.017139	.013238	.011644
-1.5	$T=-1.5$	20	.166028	.057987	.057502	.041184
	$\mu=0.0$	40	.051089	.038045	.026064	.027895
	$\sigma=1.0$	60	.029486	.032386	.017033	.022730
-1.5	$T=-3.0$	20	.664114	.232132	.230008	.164915
	$\mu=0.0$	40	.204354	.152375	.104257	.111754
	$\sigma=2.0$	60	.117945	.129744	.068132	.091086
-1.25	$T=-1.5$	20	.881132	.303036	.281598	.218471
	$\mu=1.0$	40	.356036	.244741	.142084	.167900
	$\sigma=2.0$	60	.162887	.216499	.082812	.137465

References

[1] Balakrishnan, N. (1989a). Approximate MLE of the scale parameter of the Rayleigh distribution with censoring, *IEEE Transactions on Reliability*, Vol. 38, 355-357.

[2] Balakrishnan, N. (1989b). Approximate maximum likelihood estimation of the mean and standard deviation of the normal distribution based on Type-II censored sample, *Journal of Statistic-Computation and Simulation*, Vol. 32, 137-148.

[3] Chiu, W. K. and Leung, M. P. Y. (1981). Agraphical method for estimating the parameters

of a truncated normal distribution, *Journal of Quality Technology*, Vol. 13(1), 42-45.

[4] Cohen, A. C. (1959). Simplified estimators for the normal distribution when samples are sigly censored or truncated, *Technometrics,* Vol. 1, 217-237.

[5] Cohen, A. C. (1991). *Truncated and Censored Samples*, Marcel Deker, New York.

[6] Kendall, M. G. and Stuart, A. (1973). *The Advanced Theory of Statistics*, 2, Charles Griffin and Co., London; Hafner, New York.

[7] Kang, S. B. (1996). Approximate MLE for the scale parameter of the double exponential distribution based on Type-II censoring, *Journal of the Korean Mathematical Society*, Vol. 33(1), 69-79.

[8] Kang, S. B. and Cho, Y. S. (1997a). Estimation of the parameters in an exponential distribution with Type-II censoring, *The Korean Communications in Statistics*, Vol. 4(3), 929-941.

[9] Kang, S. B. and Cho, Y. S. (1997b). Estimation for exponential distribution under general progressive Type-II censored samples, *Journal of Statistical & Theory Methods*, Vol. 8(2), 239-245.

[10] 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 Communication in Statistics*, Vol. 4(3), 929-941.

[11] Woo, J. S., Kang, S. B., Cho, Y. S., and Jeon, S. C. (1998). Approximate MLE for Rayleigh distribution in singly right censored samples, *The Korean Communication in Statistics*, Vol. 5(1), 225-230.