Communications for Statistical Applications and Methods
2018, Vol. 25, No. 6, 647–658
https://doi.org/10.29220/CSAM.2018.25.6.647
Print ISSN 2287-7843 / Online ISSN 2383-4757

On the maximum likelihood estimation for a normal distribution under random censoring

Namhyun Kim[1,a]

[a]Department of Science, Hongik University, Korea

Abstract

In this paper, we study statistical inferences on the maximum likelihood estimation of a normal distribution when data are randomly censored. Likelihood equations are derived assuming that the censoring distribution does not involve any parameters of interest. The maximum likelihood estimators (MLEs) of the censored normal distribution do not have an explicit form, and it should be solved in an iterative way. We consider a simple method to derive an explicit form of the approximate MLEs with no iterations by expanding the nonlinear parts of the likelihood equations in Taylor series around some suitable points. The points are closely related to Kaplan-Meier estimators. By using the same method, the observed Fisher information is also approximated to obtain asymptotic variances of the estimators. An illustrative example is presented, and a simulation study is conducted to compare the performances of the estimators. In addition to their explicit form, the approximate MLEs are as efficient as the MLEs in terms of variances.

Keywords: Kaplan-Meier estimators, Koziol-Green model, maximum likelihood estimators, normal distribution, random censoring

1. Introduction

In statistical analysis of life time data, some well-known common distributions are exponential, Weibull, lognormal, and gamma distribution. The lognormal distribution is especially useful when a hazard rate is initially increasing and then decreasing. We also need inferences for a normal distribution since the logarithm of a lognormal variable follows a normal. As for censoring types, the most common and simplest censoring schemes are type I or type II censoring. For numerous censoring types, see Tableman and Kim (2004), and Lee and Wang (2003).

Gupta (1952), Cohen (1959, 1961), and Kim (2014b) studied the estimation of a type II censored normal distribution. Balakrishinan *et al.* (2003) dealt with the estimation of a normal distribution on a progressively type II censoring scheme, which is a generalization of traditional type II censoring. Maximum likelihood estimators (MLEs) based on type II or progressive type II censored data from a normal distribution do not have explicit forms, and the situation remains as it is for many distributions when data are censored. Kim (2014b) and Balakrishinan *et al.* (2003) derived approximate MLEs under type II censoring and progressively type II censoring of a normal, respectively.

Random censoring occurs frequently in survival studies. In this paper, we consider the parameter estimation of a normal distribution for randomly censored data by applying the same approximate method used in Kim (2014b) and Balakrishinan *et al.* (2003). As it is mentioned, they dealt with type

[1] Department of Science, Hongik University, 94 Wausan-Ro, Mapo-Gu, Seoul 04066, Korea. E-mail: nhkim@hongik.ac.kr

Published 30 November 2018 / journal homepage: http://csam.or.kr

II or progressively type II censoring, not random censoring. The Kaplan and Meier (1958) estimator is used for the plotting position of randomly censored data.

The approximation method was first developed by Balakrishnan (1989) to find the approximate MLE of the scalar parameter in the Rayleigh distribution with left and right type II censoring. The method approximates the nonlinear part of the likelihood equations; subsequently, many researchers used it for other distributions under several censoring schemes that are most often progressively type II censoring. Balakrishinan *et al.* (2003) did it for a normal distribution. Balakrishnan and Kannan (2001), and Balakrishnan *et al.* (2004) studied the estimation for the logistic distribution and the extreme value distribution, respectively. Asgharzadeh (2006, 2009) dealt with the problem for generalized logistic distribution and generalized exponential distribution, respectively. Balakrishnan and Asgharzadeh (2005), and Kang *et al.* (2008) used the method for the half logistic distribution. Seo and Kang (2007), and Kim and Han (2009) discussed the procedure for the Rayleigh distribution. Sultan *et al.* (2014) used it for inverse Weibull distribution. As for random censoring, Kim (2014a, 2016) applied the approximate method to generalized exponential distributions and Weibull distribution, respectively.

In Section 2, we derive MLEs and approximate MLEs for a normal under random censoring. In Section 3, we provide expressions for observed Fisher information to obtain the approximate variances of the estimators. Section 4 presents simulation results that compare MLEs and approximate MLEs. Section 5 ends the paper with some concluding remarks.

2. MLEs and the approximate MLEs

Let $T_1, \ldots, T_n$ be lifetimes with distribution function F and probability density function (pdf) f, and $C_1, \ldots, C_n$ be random censoring times drawn independently of the $T_1, \ldots, T_n$ from distribution function G and pdf g. The T_i's are censored on the right by C_i. On each of n individuals, we observe n random pairs (X_i, δ_i), $i = 1, \ldots, n$, where

$$X_i = \min(T_i, C_i) \quad \text{and} \quad \delta_i = \begin{cases} 1, & \text{if } T_i \le C_i, \\ 0, & \text{if } T_i > C_i. \end{cases}$$

The observed random pairs (X_i, δ_i) could be written as $(X_{(i)}, \delta_{(i)})$ where $X_{(1)} \le \cdots \le X_{(n)}$ are the ordered observations of $X_1, \ldots, X_n$, and $\delta_{(i)}$ is the δ corresponding to $X_{(i)}$. Then the likelihood function based on the ordered $(X_{(i)}, \delta_{(i)})$ becomes

$$\begin{aligned} L &= n! \prod_{j=1}^{n} \left(f\left(x_{(j)}\right) \bar{G}\left(x_{(j)}\right) \right)^{\delta_{(j)}} \left(g\left(x_{(j)}\right) \bar{F}\left(x_{(j)}\right) \right)^{1-\delta_{(j)}} \\ &= n! \prod_{j=1}^{n} \left(f\left(x_{(j)}\right) \right)^{\delta_{(j)}} \left(\bar{F}\left(x_{(j)}\right) \right)^{1-\delta_{(j)}} \prod_{j=1}^{n} \left(g\left(x_{(j)}\right) \right)^{1-\delta_{(j)}} \left(\bar{G}\left(x_{(j)}\right) \right)^{\delta_{(j)}} \end{aligned} \tag{2.1}$$

with $\bar{F} = 1 - F$, $\bar{G} = 1 - G$. If we assume the distribution of C_i's does not involve any parameters of interest, the last factor in (2.1) plays no role in the maximization. Hence we have

$$L \propto \prod_{j=1}^{n} \left(f\left(x_{(j)}\right) \right)^{\delta_{(j)}} \left(\bar{F}\left(x_{(j)}\right) \right)^{1-\delta_{(j)}}.$$

Usually lifetimes are positive. When we assume the lifetimes follow a lognormal distribution, the logarithm of the lifetimes should follow a normal distribution. In this case we need the inference for a normal. Let T_i's follow the normal distribution $N(\mu, \sigma^2)$ with the density

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad -\infty < x < \infty.$$

Here, the lifetime itself should be e^{T_i}. Then the log-likelihood function $l = \ln L$ gives

$$l \propto \sum_{j=1}^{n} \delta_{(j)} \left(-\log\sigma - \frac{1}{2}\xi_j^2\right) + \sum_{j=1}^{n} \left(1 - \delta_{(j)}\right) \log \bar{\Phi}(\xi_j),$$

where $\xi_j \equiv \xi_j(\mu, \sigma) = (x_{(j)} - \mu)/\sigma$, $\phi(t) = (1/\sqrt{2\pi})e^{-(1/2)t^2}$, $\Phi(\xi_j) = \int_{-\infty}^{\xi_j} \phi(t)dt$, and $\bar{\Phi}(\xi_j) = 1 - \Phi(\xi_j)$.

Using $\partial\xi_j/\partial\mu = -1/\sigma$, $\partial\xi_j/\partial\sigma = -\xi_j/\sigma$,

$$\frac{\partial}{\partial\mu} \log \bar{\Phi}(\xi_j) = \frac{1}{\sigma} Q(\xi_j),$$
$$\frac{\partial}{\partial\sigma} \log \bar{\Phi}(\xi_j) = \frac{\xi_j}{\sigma} Q(\xi_j),$$

with $Q(\xi_j) = \phi(\xi_j)/\bar{\Phi}(\xi_j)$, we have the likelihood equations

$$\frac{\partial l}{\partial\mu} = \frac{1}{\sigma} \sum_{j=1}^{n} \delta_{(j)}\xi_j + \frac{1}{\sigma} \sum_{j=1}^{n} \left(1 - \delta_{(j)}\right) Q(\xi_j) = 0, \tag{2.2}$$

$$\frac{\partial l}{\partial\sigma} = -\frac{n_u}{\sigma} + \frac{1}{\sigma} \sum_{j=1}^{n} \delta_{(j)}\xi_j^2 + \frac{1}{\sigma} \sum_{j=1}^{n} \left(1 - \delta_{(j)}\right) \xi_j Q(\xi_j) = 0, \tag{2.3}$$

where $n_u = \sum_{j=1}^{n} \delta_{(j)}$ is the number of uncensored data. We let $\hat{\mu}$, $\hat{\sigma}$ be the solutions of the above equations.

The likelihood equations (2.2) and (2.3) do not have an explicit solution except for all $\delta_{(j)} = 1$, a complete sample. Therefore we consider approximate MLEs that give an explicit form and do not need any iterations for computations. As mentioned in Section 1, the approximation method was first developed by Balakrishnan (1989); subsequently, researchers have used it for other distributions under several censoring schemes.

The process can be done by approximating $Q(\xi_i) \cong a_i + b_i\xi_i$. The idea is to expand $Q(\xi_i)$ in the Taylor series by keeping the first two terms around a suitable point ξ_{i0}. We will define a_i, b_i later in this section. By replacing $Q(\xi_i) \cong a_i + b_i\xi_i$ in (2.2), and (2.3), we have the approximate likelihood equations

$$\frac{\partial l}{\partial\mu} \cong \frac{1}{\sigma} \sum_{j=1}^{n} \delta_{(j)}\xi_j + \frac{1}{\sigma} \sum_{j=1}^{n} \left(1 - \delta_{(j)}\right) \left(a_j + b_j\xi_j\right) = 0, \tag{2.4}$$

$$\frac{\partial l}{\partial\sigma} \cong -\frac{n_u}{\sigma} + \frac{1}{\sigma} \sum_{j=1}^{n} \delta_{(j)}\xi_j^2 + \frac{1}{\sigma} \sum_{j=1}^{n} \left(1 - \delta_{(j)}\right) \xi_j \left(a_j + b_j\xi_j\right) = 0. \tag{2.5}$$

Substituting $\xi_j = (x_{(j)} - \mu)/\sigma$ and multiplying σ^2 or σ^3, we have

$$\sum_{j=1}^{n} \delta_{(j)} \left(x_{(j)} - \mu\right) + \sum_{j=1}^{n} \left(1 - \delta_{(j)}\right)\left(a_j\sigma + b_j \left(x_{(j)} - \mu\right)\right) = 0, \tag{2.6}$$

$$-n_u\sigma^2 + \sigma \sum_{j=1}^{n} \left(1 - \delta_{(j)}\right) a_j \left(x_{(j)} - \mu\right) + \sum_{j=1}^{n} \delta_{(j)} \left(x_{(j)} - \mu\right)^2 + \sum_{j=1}^{n} \left(1 - \delta_{(j)}\right) b_j \left(x_{(j)} - \mu\right)^2 = 0. \tag{2.7}$$

From (2.6), we have the solution $\tilde{\mu}$,

$$\tilde{\mu} = D + E\sigma \tag{2.8}$$

with

$$D = \frac{\sum_{j=1}^{n} \delta_{(j)} x_{(j)} + \sum_{j=1}^{n} \left(1 - \delta_{(j)}\right) b_j x_{(j)}}{n_u + \sum_{j=1}^{n} \left(1 - \delta_{(j)}\right) b_j}, \qquad E = \frac{\sum_{j=1}^{n} \left(1 - \delta_{(j)}\right) a_j}{n_u + \sum_{j=1}^{n} \left(1 - \delta_{(j)}\right) b_j}. \tag{2.9}$$

Substituting $\mu = D + E\sigma$ into (2.7), we have

$$A\sigma^2 - B\sigma - C = 0 \tag{2.10}$$

with

$$\begin{aligned} A &= n_u - n_u E^2 + \sum_{j=1}^{n} \left(1 - \delta_{(j)}\right)\left(a_j E - b_j E^2\right), \\ B &= -2E \sum_{j=1}^{n} \delta_{(j)} \left(x_{(j)} - D\right) - 2E \sum_{j=1}^{n} \left(1 - \delta_{(j)}\right) b_j \left(x_{(j)} - D\right) + \sum_{j=1}^{n} \left(1 - \delta_{(j)}\right) a_j \left(x_{(j)} - D\right), \\ C &= \sum_{j=1}^{n} \delta_{(j)} \left(x_{(j)} - D\right)^2 + \sum_{j=1}^{n} \left(1 - \delta_{(j)}\right) b_j \left(x_{(j)} - D\right)^2. \end{aligned} \tag{2.11}$$

Using (2.9), we can easily see that $A = n_u$, and the first two terms in B vanish. Hence the quadratic equation (2.10) in σ becomes

$$n_u\sigma^2 - B\sigma - C = 0 \tag{2.12}$$

with

$$B = \sum_{j=1}^{n} \left(1 - \delta_{(j)}\right) a_j \left(x_{(j)} - D\right)$$

and C in (2.11). The solution $\tilde{\sigma}$ of the quadratic equation (2.12) is

$$\tilde{\sigma} = \frac{B + \sqrt{B^2 + 4n_u C}}{2n_u}. \tag{2.13}$$

Since we can show $b_i \geq 0$, it follows $C \geq 0$ (see Remark 1 below). Therefore the other root of the equation (2.12) can not be a solution of σ. Note that the approximate MLEs in (2.8) and (2.13) lead to the MLEs $\hat{\mu} = \bar{x}$, $\hat{\sigma}^2 = \sum_{j=1}^{n}(x_j - \bar{x})^2/n$ of the normal distribution for a complete sample.

Now let us think about the explicit form of a_i and b_i in $Q(\xi_i) \cong a_i + b_i\xi_i$. Let

$$\xi_{i0} = \Phi^{-1}(p_i), \tag{2.14}$$

where Φ^{-1} is the inverse of Φ, and p_i is a plotting position or a quantile probability. By expanding $Q(\xi_i)$ around ξ_{i0} in the Taylor series keeping only the first two terms, $Q(\xi_i)$ can be approximated by

$$Q(\xi_i) \cong Q(\xi_{i0}) + Q'(\xi_{i0})(\xi_i - \xi_{i0}) \equiv a_i + b_i\xi_i$$

with

$$a_i = Q(\xi_{i0}) - Q'(\xi_{i0})\xi_{i0}, \tag{2.15}$$

$$b_i = Q'(\xi_{i0}) = Q(\xi_{i0})(Q(\xi_{i0}) - \xi_{i0}). \tag{2.16}$$

Remark 1. The second equality in (2.16) follows easily using

$$\phi'(\xi_{i0}) = -\xi_{i0}\phi(\xi_{i0}) \quad \text{and} \quad \bar{\Phi}'(\xi_{i0}) = -\phi(\xi_{i0}).$$

From

$$b_i \left(\bar{\Phi}(\xi_{i0})\right)^2 = \phi(\xi_{i0}) \left(\phi(\xi_{i0}) - \xi_{i0}\bar{\Phi}(\xi_{i0})\right),$$

$$\phi(\xi_{i0}) - \xi_{i0}\bar{\Phi}(\xi_{i0}) \geq \phi(\xi_{i0}) - \int_{\xi_{i0}}^{\infty} z\phi(z)dz \geq \phi(\xi_{i0}) + \int_{\xi_{i0}}^{\infty} \phi'(z)dz \geq 0,$$

b_i in (2.16) should be nonnegative, and C in (2.11) is also nonnegative. See also Kim (2014b), and Balakrishnan *et al.* (2003).

Remark 2. Let Z_i, $i = 1, \ldots, n$, be a sample from a normal distribution with $\mu = 0$, $\sigma = 1$, and let $\hat{\mu}_z$, $\hat{\sigma}_z$ be their MLEs. Then we can easily show

$$\hat{\mu} = \sigma\hat{\mu}_z + \mu, \quad \hat{\sigma} = \sigma\hat{\sigma}_z$$

from the likelihood equations (2.2), (2.3). The same relations hold for the approximate MLEs by the equations (2.8), (2.13). Hence we can assume the true value of $\mu = 0$ and $\sigma = 1$ without loss of generality when we simulate the performance of the estimators.

For randomly censored data, the Kaplan-Meier estimator p_i^{KM} is usually used for the plotting position p_i in (2.14),

$$p_i^{\text{KM}} = 1 - \prod_{j \leq i} \left(\frac{n-j}{n-j+1}\right)^{\delta_{(j)}}.$$

The estimator has been studied in Kaplan and Meier (1958), Efron (1967), Breslow and Crowley (1974), and Meier (1975). Michael and Schucany (1986) suggested the modified Kaplan-Meier estimator $p_{i,c}^{\text{MS}}$,

$$p_{i,c}^{\text{MS}} = 1 - \frac{n-c+1}{n-2c+1} \prod_{j \leq i} \left(\frac{n-j-c+1}{n-j-c+2}\right)^{\delta_{(j)}}, \quad 0 \leq c \leq 1,\ i = 1, \ldots, n, \tag{2.17}$$

that reduces to $p_{i,c} = (i - c)/(n - 2c + 1)$ for a complete sample. However $p_{i,c}^{\mathrm{MS}}$ in (2.17) could have a negative value if the smallest data value is censored. Therefore we modify $p_{i,c}^{\mathrm{MS}}$ again as

$$p_{i,c} = 1 - \prod_{j \le i} \left(\frac{n - j - c + 1}{n - j - c + 2} \right)^{\delta_{(j)}} \left(\frac{n - c + 1}{n - 2c + 1} \right)^{\delta_{(1)}} \left(\frac{n - c}{n - 2c + 1} \right)^{1 - \delta_{(1)}}, \quad 0 \le c \le 1,\ i = 1, \ldots, n. \quad (2.18)$$

The last term in $p_{i,c}$ is to avoid 0 value. Note that we assumed the data $x_{(1)}, \ldots, x_{(n)}$ are ordered observations and $\delta_{(1)}$ corresponds to the indicator of the minimum. The particular choice of c is of little consequence, and popular values for c is $c = 0$ or $c = 0.5$. We often use the Blom (1958)'s position $c = 3/8 = 0.375$ for a normal distribution, because it approximates the expected value of the order statistics from the standard normal distribution.

Now we have defined $a_i,\ b_i,\ \xi_{i0},\ p_{i,c}$ in (2.15), (2.16), (2.14), (2.18), respectively, and the approximate MLEs $\tilde{\mu}$, $\tilde{\sigma}$ of μ, σ in (2.8), (2.13) have been completely defined.

3. Asymptotic variances and covariance

In this section, the observed Fisher information is computed to give the asymptotic variances and covariance of the MLEs $\hat{\mu}$, $\hat{\sigma}$ or the approximate MLEs $\tilde{\mu}$, $\tilde{\sigma}$. From the likelihood equation (2.2), and

$$\frac{\partial \phi(\xi_i)}{\partial \mu} = \frac{\xi_i}{\sigma} \phi(\xi_i), \qquad \frac{\partial \bar{\Phi}(\xi_i)}{\partial \mu} = \frac{1}{\sigma} \phi(\xi_i), \qquad \frac{\partial Q(\xi_i)}{\partial \mu} = \frac{1}{\sigma} \left(\xi_i Q(\xi_i) - (Q(\xi_i))^2 \right) \equiv \frac{1}{\sigma} J(\xi_i),$$

we obtain

$$-\frac{\partial^2 l}{\partial \mu^2} = \frac{1}{\sigma^2} v_{11} \quad \text{with } v_{11} = n_u - \sum_{j=1}^{n} (1 - \delta_{(j)}) J(\xi_j), \quad (3.1)$$

$$-\frac{\partial^2 l}{\partial \sigma tial \mu} = \frac{1}{\sigma^2} v_{12} \quad \text{with } v_{12} = 2 \sum_{j=1}^{n} \delta_{(j)} \xi_j + \sum_{j=1}^{n} (1 - \delta_{(j)}) Q(\xi_j) - \sum_{j=1}^{n} (1 - \delta_{(j)}) \xi_j J(\xi_j). \quad (3.2)$$

From (2.3) and

$$\frac{\partial \phi(\xi_i)}{\partial \sigma} = \frac{\xi_i^2}{\sigma} \phi(\xi_i), \qquad \frac{\partial \bar{\Phi}(\xi_i)}{\partial \sigma} = \frac{\xi_i}{\sigma} \phi(\xi_i), \qquad \frac{\partial Q(\xi_i)}{\partial \sigma} = \frac{1}{\sigma} \xi_i J(\xi_i),$$

we get

$$-\frac{\partial^2 l}{\partial \sigma^2} = \frac{1}{\sigma^2} v_{22} \quad \text{with } v_{22} = -n_u + 3 \sum_{j=1}^{n} \delta_{(j)} \xi_j^2 + 2 \sum_{j=1}^{n} (1 - \delta_{(j)}) \xi_j Q(\xi_j) - \sum_{j=1}^{n} (1 - \delta_{(j)}) \xi_j^2 J(\xi_j). \quad (3.3)$$

From (3.1), (3.2), and (3.3), the observed Fisher information matrix becomes

$$\mathbf{I} = \frac{1}{\sigma^2} \begin{bmatrix} v_{11} & v_{12} \\ v_{12} & v_{22} \end{bmatrix}. \quad (3.4)$$

By inverting (3.4), we obtain the asymptotic variance-covariance matrix of the estimates $\hat{\mu}$ and $\hat{\sigma}$ as

$$\mathbf{I}^{-1} = \sigma^2 \begin{bmatrix} v_{11} & v_{12} \\ v_{12} & v_{22} \end{bmatrix}^{-1} \cong \begin{bmatrix} \widehat{\mathrm{Var}}(\hat{\mu}) & \widehat{\mathrm{Cov}}(\hat{\mu}, \hat{\sigma}) \\ \widehat{\mathrm{Cov}}(\hat{\mu}, \hat{\sigma}) & \widehat{\mathrm{Var}}(\hat{\sigma}) \end{bmatrix}. \quad (3.5)$$

From the approximate likelihood equations (2.4) and (2.5), we get

$$-\frac{\partial^2 l}{\partial \mu^2} \cong \frac{1}{\sigma^2}\tilde{v}_{11} \quad \text{with } \tilde{v}_{11} = n_u + \sum_{j=1}^{n}\left(1-\delta_{(j)}\right)b_j,$$

$$-\frac{\partial^2 l}{\partial \sigma \partial \mu} \cong \frac{1}{\sigma^2}\tilde{v}_{12} \quad \text{with } \tilde{v}_{12} = 2\sum_{j=1}^{n}\delta_{(j)}\xi_j + \sum_{j=1}^{n}\left(1-\delta_{(j)}\right)\left(a_j + b_j\xi_j\right) + \sum_{j=1}^{n}\left(1-\delta_{(j)}\right)b_j\xi_j,$$

$$-\frac{\partial^2 l}{\partial \sigma^2} \cong \frac{1}{\sigma^2}\tilde{v}_{22} \quad \text{with } \tilde{v}_{22} = -n_u + 3\sum_{j=1}^{n}\delta_{(j)}\xi_j^2 + 2\sum_{j=1}^{n}\left(1-\delta_{(j)}\right)\xi_j\left(a_j + b_j\xi_j\right) + \sum_{j=1}^{n}\left(1-\delta_{(j)}\right)b_j\xi_j^2,$$

and the approximate observed Fisher information matrix $\tilde{\mathbf{I}}$,

$$\tilde{\mathbf{I}} = \frac{1}{\sigma^2}\begin{bmatrix} \tilde{v}_{11} & \tilde{v}_{12} \\ \tilde{v}_{12} & \tilde{v}_{22} \end{bmatrix},$$

$$\tilde{\mathbf{I}}^{-1} = \sigma^2\begin{bmatrix} \tilde{v}_{11} & \tilde{v}_{12} \\ \tilde{v}_{12} & \tilde{v}_{22} \end{bmatrix}^{-1} \cong \begin{bmatrix} \widetilde{\mathrm{Var}}(\tilde{\mu}) & \widetilde{\mathrm{Cov}}(\tilde{\mu},\tilde{\sigma}) \\ \widetilde{\mathrm{Cov}}(\tilde{\mu},\tilde{\sigma}) & \widetilde{\mathrm{Var}}(\tilde{\sigma}) \end{bmatrix}. \tag{3.6}$$

4. Simulation study and example

4.1. Simulation results

We compare the performance of the MLEs with the approximate MLEs through a simulation study. To control the ratio of the censored data, we use two different random censoring models. One is the Koziol and Green (1976) censorship model, that is

$$1 - G = (1-F)^{\beta_1} \quad \text{for some } \beta_1 > 0, \tag{4.1}$$

where G is the distribution function of the censoring times $C_1, \ldots, C_n$, and β_1 is called a censoring parameter. Csörgő and Horváth (1981), Chen *et al.* (1982), and Kim (2014a) discussed the motivation and characterization of this model. According to Chen *et al.* (1982), the model could arise in a series system with two components. In this system, it can only work if both components are functioning. Let T_i be the life time of the first component with the distribution function F, and C_i be the second with G. Then $X_i = \min(T_i, C_i)$ should be the lifetime of the system. If the second component is also a series system of β_1 independent and identically distributed subcomponents with F, then C_i should be the minimum of β_1's life times, and the distribution function G of C_i should be (4.1). Under the Koziol-Green model, the expected ratio of the censored data could be

$$P(T_i > C_i) = \int_{-\infty}^{\infty}(1-F(x))dG(x) = \int_0^1 \beta_1(1-x)^{\beta_1}dx = \frac{\beta_1}{\beta_1+1} \equiv \gamma,$$

and we call γ the censoring ratio of the model.

When the second component is parallel with β_2 subcomponents, it works if at least one component is functioning. In this case the censoring distribution G becomes

$$G = F^{\beta_2} \quad \text{for some } \beta_2 > 0. \tag{4.2}$$

Table 1: Averages of the MLEs $\hat{\mu}$, $\hat{\sigma}$, variances of $\hat{\mu}$, $\hat{\sigma}$ and covariances under the Koziol-Green model

γ (β_1)	n	$\hat{\mu}$	$\hat{\sigma}$	Var($\hat{\mu}$)	Var($\hat{\sigma}$)	Cov($\hat{\mu}, \hat{\sigma}$)	$\widehat{\text{Var}}(\hat{\mu})$	$\widehat{\text{Var}}(\hat{\sigma})$	$\widehat{\text{Cov}}(\hat{\mu}, \hat{\sigma})$
$\frac{3}{5}\left(\frac{3}{2}\right)$	20	0.016	0.962	0.122	0.077	0.053	0.126	0.079	0.057
	30	0.009	0.976	0.072	0.045	0.030	0.074	0.048	0.032
	40	0.004	0.974	0.051	0.033	0.021	0.052	0.034	0.022
	50	−0.001	0.982	0.039	0.026	0.016	0.041	0.027	0.017
$\frac{1}{2}$ (1)	20	0.003	0.957	0.082	0.052	0.026	0.084	0.055	0.029
	30	0.010	0.976	0.056	0.035	0.017	0.055	0.036	0.018
	40	0.008	0.989	0.042	0.026	0.012	0.041	0.027	0.013
	50	0.001	0.987	0.032	0.021	0.009	0.032	0.021	0.010
$\frac{1}{3}\left(\frac{1}{2}\right)$	20	−0.003	0.955	0.063	0.039	0.011	0.061	0.038	0.012
	30	0.003	0.974	0.042	0.027	0.008	0.041	0.026	0.008
	40	−0.004	0.981	0.032	0.019	0.005	0.031	0.019	0.006
	50	−0.003	0.984	0.027	0.016	0.005	0.025	0.015	0.004
$\frac{1}{4}\left(\frac{1}{3}\right)$	20	−0.001	0.960	0.056	0.035	0.008	0.057	0.034	0.008
	30	−0.001	0.974	0.039	0.021	0.004	0.038	0.022	0.005
	40	0.000	0.981	0.028	0.017	0.003	0.028	0.017	0.003
	50	−0.005	0.986	0.023	0.014	0.003	0.023	0.013	0.003

MLEs = maximum likelihood estimators.

Table 2: Averages of the approximate MLEs $\tilde{\mu}$, $\tilde{\sigma}$, variances of $\tilde{\mu}$, $\tilde{\sigma}$ and covariances under the Koziol-Green model

γ (β_1)	n	$\tilde{\mu}$	$\tilde{\sigma}$	$\hat{\mu}-\tilde{\mu}$	$\hat{\sigma}-\tilde{\sigma}$	Var($\tilde{\mu}$)	Var($\tilde{\sigma}$)	Cov($\tilde{\mu}, \tilde{\sigma}$)	$\widetilde{\text{Var}}(\tilde{\mu})$	$\widetilde{\text{Var}}(\tilde{\sigma})$	$\widetilde{\text{Cov}}(\tilde{\mu}, \tilde{\sigma})$
$\frac{3}{5}\left(\frac{3}{2}\right)$	20	0.009	0.970	0.0069	−0.0075	0.120	0.073	0.052	0.129	0.079	0.059
	30	0.004	0.979	0.0051	−0.0034	0.071	0.043	0.029	0.075	0.048	0.033
	40	0.000	0.976	0.0036	−0.0021	0.051	0.032	0.020	0.053	0.034	0.022
	50	−0.004	0.983	0.0030	−0.0013	0.039	0.026	0.015	0.041	0.027	0.017
$\frac{1}{2}$ (1)	20	−0.003	0.961	0.0056	−0.0039	0.081	0.051	0.026	0.086	0.055	0.030
	30	0.006	0.978	0.0037	−0.0019	0.055	0.034	0.016	0.056	0.036	0.018
	40	0.006	0.990	0.0027	−0.0011	0.041	0.026	0.012	0.042	0.027	0.013
	50	−0.001	0.988	0.0022	−0.0007	0.032	0.021	0.009	0.032	0.021	0.010
$\frac{1}{3}\left(\frac{1}{2}\right)$	20	−0.006	0.957	0.0031	−0.0015	0.062	0.039	0.012	0.062	0.038	0.012
	30	0.002	0.975	0.0019	−0.0009	0.042	0.027	0.008	0.042	0.026	0.008
	40	−0.005	0.981	0.0015	−0.0004	0.032	0.019	0.005	0.031	0.019	0.006
	50	−0.004	0.985	0.0010	−0.0005	0.027	0.016	0.005	0.025	0.015	0.004
$\frac{1}{4}\left(\frac{1}{3}\right)$	20	−0.003	0.961	0.0022	−0.0009	0.055	0.035	0.008	0.057	0.034	0.008
	30	−0.003	0.974	0.0013	−0.0004	0.039	0.021	0.004	0.038	0.022	0.005
	40	−0.001	0.981	0.0009	−0.0004	0.028	0.017	0.003	0.029	0.017	0.003
	50	−0.006	0.986	0.0008	−0.0002	0.023	0.014	0.003	0.023	0.013	0.003

MLEs = maximum likelihood estimators.

We call it P model. The model is closely related to generalized exponential distribution. See Gupta and Kundu (2007), and Kim (2014a). In this case, the expected censoring ratio is

$$P(T_i > C_i) = \int_{-\infty}^{\infty} (1 - F(x)) dG(x) = \frac{1}{\beta_2 + 1} \equiv \gamma.$$

The model is also considered in Kim (2011). The two models are the same when $\beta_1 = \beta_2 = 1$.

Tables 1–4 present the simulation results for the sample sizes $n = 20, 30, 40, 50$, and the censoring ratio $\gamma = 3/5, 1/2, 1/3, 1/4$ with $N = 2000$ repetitions for each of the censoring model in (4.1) and (4.2). The data are generated by S-plus package. They are generated with the true value of $\mu = 0$, $\sigma = 1$ without loss of generality by Remark 2 in Section 2. Note that the MLEs of μ and σ are derived under the assumption that the distribution of the censoring time does not involve any parameters of

Table 3: Averages of the MLEs $\hat{\mu}$, $\hat{\sigma}$, variances of $\hat{\mu}$, $\hat{\sigma}$ and covariances under the P model

γ (β_2)	n	$\hat{\mu}$	$\hat{\sigma}$	$\widehat{\mathrm{Var}}(\hat{\mu})$	$\widehat{\mathrm{Var}}(\hat{\sigma})$	$\widehat{\mathrm{Cov}}(\hat{\mu},\hat{\sigma})$	$\widehat{\mathrm{Var}}(\hat{\mu})$	$\widehat{\mathrm{Var}}(\hat{\sigma})$	$\widehat{\mathrm{Cov}}(\hat{\mu},\hat{\sigma})$
$\frac{3}{5}\left(\frac{2}{3}\right)$	20	0.015	0.964	0.107	0.070	0.040	0.120	0.073	0.050
	30	0.008	0.975	0.075	0.047	0.030	0.074	0.046	0.029
	40	0.005	0.979	0.049	0.033	0.019	0.052	0.032	0.020
	50	−0.001	0.983	0.041	0.026	0.016	0.041	0.026	0.015
$\frac{1}{2}$ (1)	20	0.020	0.966	0.088	0.053	0.027	0.088	0.058	0.031
	30	0.012	0.971	0.055	0.035	0.017	0.055	0.036	0.018
	40	0.003	0.978	0.042	0.026	0.013	0.040	0.026	0.013
	50	0.002	0.986	0.034	0.022	0.011	0.032	0.021	0.010
$\frac{1}{3}$ (2)	20	0.010	0.966	0.064	0.041	0.011	0.062	0.042	0.014
	30	0.017	0.985	0.043	0.028	0.009	0.042	0.028	0.009
	40	0.002	0.984	0.030	0.021	0.006	0.031	0.020	0.006
	50	0.004	0.982	0.023	0.016	0.005	0.024	0.016	0.005
$\frac{1}{4}$ (3)	20	0.001	0.965	0.056	0.037	0.008	0.056	0.036	0.009
	30	0.008	0.983	0.038	0.025	0.006	0.038	0.024	0.006
	40	0.006	0.988	0.026	0.019	0.003	0.028	0.018	0.004
	50	0.001	0.985	0.022	0.014	0.003	0.022	0.014	0.003

MLEs = maximum likelihood estimators.

Table 4: Averages of the approximate MLEs $\tilde{\mu}$, $\tilde{\sigma}$, variances of $\tilde{\mu}$, $\tilde{\sigma}$ and covariances under the P model

γ (β_2)	n	$\tilde{\mu}$	$\tilde{\sigma}$	$\hat{\mu}-\tilde{\mu}$	$\hat{\sigma}-\tilde{\sigma}$	$\mathrm{Var}(\tilde{\mu})$	$\mathrm{Var}(\tilde{\sigma})$	$\mathrm{Cov}(\tilde{\mu},\tilde{\sigma})$	$\widetilde{\mathrm{Var}}(\tilde{\mu})$	$\widetilde{\mathrm{Var}}(\tilde{\sigma})$	$\widetilde{\mathrm{Cov}}(\tilde{\mu},\tilde{\sigma})$
$\frac{3}{5}\left(\frac{2}{3}\right)$	20	0.010	0.987	0.0044	−0.0229	0.108	0.063	0.042	0.126	0.074	0.053
	30	0.004	0.987	0.0041	−0.0122	0.075	0.044	0.030	0.076	0.046	0.030
	40	0.002	0.987	0.0030	−0.0082	0.049	0.031	0.019	0.053	0.032	0.020
	50	−0.003	0.989	0.0025	−0.0059	0.041	0.025	0.016	0.041	0.025	0.015
$\frac{1}{2}$ (1)	20	0.014	0.970	0.0058	−0.0040	0.087	0.051	0.027	0.090	0.058	0.032
	30	0.008	0.974	0.0037	−0.0021	0.054	0.034	0.017	0.055	0.036	0.018
	40	0.001	0.979	0.0026	−0.0011	0.042	0.026	0.013	0.041	0.026	0.013
	50	0.000	0.987	0.0022	−0.0007	0.034	0.022	0.011	0.033	0.021	0.010
$\frac{1}{3}$ (2)	20	0.007	0.965	0.0029	0.0006	0.063	0.040	0.011	0.063	0.042	0.014
	30	0.015	0.985	0.0018	0.0005	0.043	0.027	0.009	0.042	0.028	0.009
	40	0.001	0.983	0.0012	0.0002	0.030	0.021	0.006	0.031	0.020	0.006
	50	0.003	0.981	0.0010	0.0003	0.023	0.016	0.005	0.024	0.016	0.005
$\frac{1}{4}$ (3)	20	−0.001	0.965	0.0018	0.0006	0.056	0.036	0.008	0.056	0.036	0.009
	30	0.007	0.982	0.0011	0.0005	0.038	0.024	0.006	0.038	0.024	0.006
	40	0.005	0.988	0.0008	0.0003	0.026	0.019	0.003	0.028	0.018	0.004
	50	0.000	0.985	0.0006	0.0002	0.022	0.014	0.003	0.022	0.014	0.003

MLEs = maximum likelihood estimators.

interest, and the two models in (4.1) and (4.2) do not satisfy the assumption. However we usually do not know the true censoring model in a real situation, and the ratio of the simulated censored data could hardly be controlled without a censoring model.

Table 1 gives the averages of the MLEs of μ and σ, their variances and covariance by the simulation, and the asymptotic variances and covariance from the inverse of the observed Fisher information given in (3.5) under the Koziol-Green model in (4.1). The MLEs are computed by the S-plus function `survReg`. Table 2 gives the same statistics for the approximate MLEs under the same model. The last three columns are from the inverse of the approximated observed Fisher information given in (3.6). It also provides the average differences of the two estimates. In Tables 3 and 4, the same statistics are given under the P model in (4.2). For true values of σ not equal to 1, the variances and covariances of μ and σ should be multiplied by σ^2 by (3.5) and (3.6).

From Tables 1–4, we observe the following. The average differences between the MLEs and the approximate MLEs are negligible. The differences tend to decrease as the sample size increases or if the censoring ratio becomes smaller. The variances of the estimators, not the mean squared errors, are provided since the biases of the parameters look very small. All the variances and covariances of the MLEs and the approximate MLEs are almost identical. The variances and covariances from the simulation and from the observed Fisher information are also similar. Apparently the approximate MLEs are as efficient as the MLEs. The variances and covariances also reduce considerably as the sample size becomes bigger or the censoring ratio becomes smaller. All the same phenomena happen in both the Koziol-Green model and P model. It seems that the censoring models do not have any clear influence on the estimation process, since we assumed that the censoring distribution does not have any information for the parameters. Kim (2016) studied the influence of the assumption for the censoring model on the estimation for the parameters of a Weibull distribution.

4.2. Example

As an illustrative example, we consider the tumor-free time data of the 30 rats fed with saturated diets. The data set is to investigate the relationship between diet and the development of tumors. The study divided 90 rats into three groups and fed them low fat, saturated fat, and unsaturated fat diets. The data were originally reported by King *et al.* (1979) and discussed in Lee and Wang (2003).

The data are sorted in order and given below. The other two groups are not shown.

43	46	56	58	68	75	79	81	86	86
89	96	98	105	107	110	117	124	126	133
142	142	165	170+	200+	200+	200+	200+	200+	200+

Lee and Wang (2003, Chapter 8, 9) checked the Cox-Snell residual plot for the fitted lognormal model, and showed the goodness-of-fit tests based on the asymptotic likelihood inference and the BIC, AIC. As a result, the lognormal distribution would be selected rather than the exponential or the Weibull distribution by the procedures.

When we compute the MLEs of the parameters, we have

$$\hat{\mu} = 4.764583, \quad \hat{\sigma} = 0.5605291,$$
$$\tilde{\mu} = 4.762847, \quad \tilde{\sigma} = 0.5593185.$$

The variance-covariance matrices in (3.5) and (3.6) are

$$\begin{bmatrix} \widehat{\mathrm{Var}}(\hat{\mu}) & \widehat{\mathrm{Cov}}(\hat{\mu},\hat{\sigma}) \\ \widehat{\mathrm{Cov}}(\hat{\mu},\hat{\sigma}) & \widehat{\mathrm{Var}}(\hat{\sigma}) \end{bmatrix} = \begin{bmatrix} 0.01127 & 0.001401 \\ 0.001401 & 0.007777 \end{bmatrix},$$
$$\begin{bmatrix} \widetilde{\mathrm{Var}}(\tilde{\mu}) & \widetilde{\mathrm{Cov}}(\tilde{\mu},\tilde{\sigma}) \\ \widetilde{\mathrm{Cov}}(\tilde{\mu},\tilde{\sigma}) & \widetilde{\mathrm{Var}}(\tilde{\sigma}) \end{bmatrix} = \begin{bmatrix} 0.01140 & 0.001500 \\ 0.001500 & 0.007818 \end{bmatrix}.$$

We can see that the results are very close each other.

5. Concluding remarks

In this study, the approximate MLEs for a normal distribution under random censoring are proposed by linearizing the nonlinear functions in the likelihood equations. As results, they give an explicit form and need no iterations contrary to the MLEs. In addition, the approximate MLEs are as efficient as the MLEs in terms of biases and variances from the simulation study. As for censoring models, we considered the Koziol-Green model and P model. Apparently, the censoring models do not have clear influence on the estimation process since we assumed that the censoring distribution does not involve any parameters of interest and ignore the information contained in the censoring model.

This paper assumes no covariates, and the problem with covariates remains a good research topic for studying.

Acknowledgement

This work was supported by 2018 Hongik University Research Fund.

References

Asgharzadeh A (2006). Point and interval estimation for a generalized logistic distribution under progressive type II censoring, *Communications in Statistics-Theory and Methods*, **35**, 1685–1702.

Asgharzadeh A (2009). Approximate MLE for the scaled generalized exponential distribution under progressive type-II censoring, *Journal of the Korean Statistical Society*, **38**, 223–229.

Balakrishnan N (1989). Approximate MLE of the scale parameter of the Rayleigh distribution with censoring, *IEEE Transactions on Reliability*, **38**, 355–357.

Balakrishnan N and Asgharzadeh A (2005). Inference for the scaled half-logistic distribution based on progressively type II censored samples, *Communications in Statistics-Theory and Methods*, **34**, 73–87.

Balakrishnan N and Kannan N (2001). Point and interval estimation for the logistic distribution base on progressive type-II censored samples, in *Handbook of Statistics*, Balakrishnan, N. and Rao, C. R. Eds., **20**, 431–456.

Balakrishnan N, Kannan N, Lin CT, and Ng HKT (2003). Point and interval estimation for gaussian distribution, based on progressively type-II censored samples, *IEEE Transactions on Reliability*, **52**, 90–95.

Balakrishnan N, Kannan N, Lin CT, and Wu SJS (2004). Inference for the extreme value distribution under progressive type-II censoring, *Journal of Statistical Computation and Simulation*, **74**, 25–45.

Blom G (1958). *Statistical Estimates and Transformed Beta Variates*, Wiley, New York.

Breslow N and Crowley J (1974). A large sample study of the life table and product limit estimates under random censorships, *The Annals of Statistics*, **2**, 437–453.

Chen YY, Hollander M, and Langberg NA (1982). Small-sample results for the Kaplan Meier estimator, *Journal of the American statistical Association*, **77**, 141–144.

Cohen AC (1959). Simplified estimators for the normal distribution when samples are singly censored or truncated, *Technometrics*, **1**, 217–237.

Cohen AC (1961). Tables for maximum likelihood estimates: singly truncated and singly censored samples, *Technometrics*, **3**, 535–541.

Csörgő S and Horváth L (1981). On the Koziol-Green Model for random censorship, *Biometrika*, **68**, 391–401.

Efron B (1967). The two sample problem with censored data. *Proc. 5th Berkeley Symp.*, **4**, 831–853.

Gupta AK (1952). Estimation of the mean and standard deviation of a normal population from a censored sample, *Biometrika*, **39**, 260–273.

Gupta RD and Kundu D (2007). Generalized exponential distribution: existing results and some recent developments. *Journal of Statistical Planning and Inference*, **137**, 3537–3547.

Kang SB, Cho YS, and Han JT (2008). Estimation for the half logistic distribution under progressively type-II censoring, *Communications of the Korean Statistical Society*, **15**, 367–378.

Kaplan EL and Meier P (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, **53**, 457–481

Kim C and Han K (2009). Estimation of the scale parameters of the Rayleigh distribution under general progressive censoring, *Journal of the Korean Statistical Society*, **38**, 239–246.

Kim N (2011). Testing log normality for randomly censored data. *The Korean Journal of Applied statistics*, **24**, 883-891.

Kim N (2014a). Approximate MLE for the scale parameter of the generalized exponential distribution under random censoring, *Journal of the Korean Statistical Society*, **43**, 119–131.

Kim N (2014b). Estimation for mean an standard deviation of normal distribution under type II censoring, *Communications for Statistical Applications and Methods*, **21**, 529–538.

Kim N (2016). On the maximum likelihood estimators for parameters of a Weibull distribution under random censoring, *Communications for Statistical Applications and Methods*, **23**, 241–250.

King M, Bailey DM, Gibson DG, Pitha JV, and McCay PB (1979). Incidence and growth of mammary tumors induced by 7,12-dimethylbenz(α)antheacene as related to the dietary content of fat and antioxidant. *Journal of the National Cancer Institute*, **63**, 656–664.

Koziol JA and Green SB (1976). A Cramér-von Mises statistic for randomly censored data, *Biometrika*, **63**, 465–474.

Lee ET and Wang JW (2003). *Statistical Methods for Survival Data Analysis*. John Wiley & Sons, Inc. New Jersey.

Meier P (1975). Estimation of a distribution function from incomplete observations. In *Perspectives in Probability and Statistics*, Ed. J. Gani, 67-87, London, Academic Press.

Michael JR and Schucany WR (1986). Analysis of data from censored samples, In *Goodness of fit techniques*, (Edited by D'Agostino, R. B. and Stephens, M. A.), Chapter 11, New York, Marcel Dekker.

Seo EH and Kang SB (2007). AMLEs for Rayleigh distribution based on progressively type-II censored data, *The Korean Communications in Statistics*, **14**, 329–344.

Sultan KS, Alsadat NH, and Kundu D (2014). Bayesian and maximum likelihood estimation of the inverse Weibull parameters under progressive type-II censoring, *Journal of Statistical Computation and Simulation*, **84**, 2248–265.

Tableman M and Kim JS (2004). *Survival Analysis Using S*, Chapman & Hall/CRC, Boca Raton.

Received June 28, 2018; Revised August 16, 2018; Accepted September 26, 2018