Maximum Likelihood Estimation of Location and Scale Parameters from Multi-censored Samples

Maximum Likelihood Estimation of Location and Scale Parameters from Multi-censored Samples

Hidetoshi NAKAYASU*, Ken'ichi MORI** and Shigeo KASE**

(Received November 15, 1977)

This paper is concerned with the problem of maximum likelihood estimation of the location and scale parameters in some distribution from multi-censored samples. The estimation procedure performed here gives:

(1) general formulae of calculating maximum likelihood estimators for the location-scale type distribution from multi-censored samples,
(2) asymptotic variance-covariance matrix of ML-estimates, and
(3) method of determining the confidence region for location and scale parameters by the likelihood ratio test theory.

A numerical example using CFRP fatigue test data illustrates the proposed method in case where it is applied to multi-censored samples.

1. Introduction

Studies on the maximum likelihood estimation problem from censored samples have been made by Cohen[1], Wingo[2] and many other researchers. Their approaches, however, are based on some specified distributions, i.e., normal, log-normal, Weibull and others, and the procedures derived are of limited use. In terms of censored type, few approaches have treated multi-censored samples.

The principal reasons why we discuss maximum likelihood estimation of location and scale parameters from multi-censored samples are as follows:

(1) In life testing the censored sample arises frequently at various stages, and
(2) The general formulae of calculating ML-estimates for location-scale type distribution do not prevail which appears frequently in the field of life testing.

As a numerical example for illustration of the proposed method, its application to multi-censored sample from CFRP fatigue test data is discussed.

2. Multi-censored samples

Let N be the total number of specimens, and n the number of failure specimens. Suppose that censoring occurs in k stages at time $T_j\ (> T_{j-1}),\ (j=1, 2, 3, \cdots, k)$ and r_j surviving specimens are removed (censored) from testing at j-th stage. Then we have

* Graduate Student, Department of Industrial Engineering, College of Engineering.
** Department of Industrial Engineering, College of Engineering.

$$N = n + \sum_{j=1}^{k} r_j . \tag{1}$$

There are two types of censoring: In type I censoring, which is of primary interest here, T_j is fixed, and number of survivors at these times is random variable. In type II censoring, number of survivors is fixed and T_j is random variable. r_j is independent of life span x.

When failure times x_i $(i = 1, 2, 3, \cdots, n)$ are observed, the likelihood function L for type I multi-censored sample becomes

$$L = C \prod_{i=1}^{n} f(x_i ; \theta) \prod_{j=1}^{k} [1 - F(T_j ; \theta)]^{r_j} , \tag{2}$$

where

C	:	normalizing constant,
θ	:	parameters,
$f(x)$	:	probability density function (p. d. f.), and
$F(x)$	:	cumulative distribution function (c. d. f.).

3. Maximum likelihood estimation based on multi-censored samples

3.1 Location-scale type distribution

Location-scale type distribution which is widely used to represent the statistical interpretation of lifetime is defined as[3)]

$$dF[(x-b)/a] = f[(x-b)/a]\, d[(x-b)/a] , \tag{3}$$

where a and b denote scale and location parameters, respectively. By the transformations

$$y = \frac{x-b}{a}, \qquad g(y) = \ln f[(x-b)/a] ,$$
$$Y = \frac{T-b}{a}, \text{ and } \qquad h(Y) = \ln [1 - F(Y)] , \tag{4}$$

likelihood function, Eq. (2), is written by

$$L = \frac{C}{a^n} \prod_{i=1}^{n} \exp [g(y_i)] \prod_{j=1}^{k} \exp [h(Y_j)] . \tag{5}$$

Taking logarithm of Eq. (5) except constants, we have

$$\ln L = \sum_{i=1}^{n} g(y_i) - n \ln a + \sum_{j=1}^{k} r_j h(Y_j) . \tag{6}$$

Differentiation of Eq. (6) with respect to a and b enables us to obtain likelihood equation such as

$$\sum_{i=1}^{n} g'(y_i) + \sum_{j=1}^{k} r_j h'(Y_j) = 0 , \tag{7-a}$$

and

$$\sum_{i=1}^{n} y_i g'(y_i) + n + \sum_{j=1}^{k} r_j Y_j h'(Y_j) = 0 , \tag{7-b}$$

where prime means the differentiation with respect to y or Y. The solutions of simultaneous equation (7) give ML-estimates of a and b.

3.2 Asymptotic variance-covariance matrix

The asymptotic variance-covariance matrix for the above ML-estimates, $\hat{a}$ and $\hat{b}$, can be derived from Fisher's information matrix[3]. The information matrix of ML-estimates for location and scale parameters

$$I = -E \begin{pmatrix} \dfrac{\partial^2 \ln L}{\partial a^2} & \dfrac{\partial^2 \ln L}{\partial a \partial b} \\ \dfrac{\partial^2 \ln L}{\partial b \partial a} & \dfrac{\partial^2 \ln L}{\partial b^2} \end{pmatrix} \tag{8-a}$$

may, from Eq. (6), become

$$I \simeq -\frac{1}{a^2} \begin{pmatrix} \sum_{i=1}^{n} g_i'' - \sum_{j=1}^{k} r_j h_j'' & \sum_{i=1}^{n} y_i g_i'' + \sum_{j=1}^{k} r_j Y_j h_j'' \\ \sum_{i=1}^{n} y_i g_i'' + \sum_{j=1}^{k} r_j Y_j h_j'' & \sum_{i=1}^{n} y_i^2 g_i'' - n + \sum_{j=1}^{k} r_j h_j'' \end{pmatrix}_{\substack{a=\hat{a} \\ b=\hat{b}}} , \tag{8-b}$$

where

$$g_i'' = g''(y_i)$$

and

$$h_j'' = h''(Y_j) .$$

Hence, inversion of the matrix, Eq. (8-b) gives asymptotic variance-covariance matrix in the form:

$$V = \frac{a^2}{\det I} \begin{pmatrix} \sum_{i=1}^{n} y_i^2 g_i'' - n + \sum_{j=1}^{k} r_j h_j'' & -\sum_{i=1}^{n} y_i g_i'' - \sum_{j=1}^{k} r_j Y_j h_j'' \\ -\sum_{i=1}^{n} y_i g_i'' - \sum_{j=1}^{k} r_j Y_j h_j'' & \sum_{i=1}^{n} g_i'' - \sum_{j=1}^{k} r_j h_j'' \end{pmatrix}_{\substack{a=\hat{a} \\ b=\hat{b}}} . \tag{9}$$

3.3 Confidence region of $\hat{a}$ and $\hat{b}$ based on likelihood ratio

In order to determine the confidence region of ML-estimates $\hat{a}$ and $\hat{b}$, consider the likelihood ratio test theory. Let H_0 be the hypothesis

$$H_0 : \theta_r = \theta_{r0} , \tag{10-a}$$

against the alternative one

$$H_1 : \theta_r \neq \theta_{r0} , \tag{10-b}$$

where

θ_r : parameters of underlying distribution,
θ_{r0} : given parameter values in H_0, and
r : number of parameters in H_0 .

Since there is no general UMP (uniformly most powerful) test in this case, it is usually convenient to utilize the likelihood ratio.

From the value defined by

$$\hat{\hat{\theta}}_s = [\theta_s \mid \max_{\theta_s} L(\theta_{r0}, \theta_s)] , \tag{11}$$

likelihood ratio is written as

$$\lambda = \frac{L(\theta_{r0}, \hat{\theta}_s)}{L(\theta)} . \tag{12}$$

The statistic $-2 \log \lambda$ lies in the region $(0, \infty)$ and is asymptotically distributed in χ^2 form with r degrees of freedom[3]. Therefore, $(1-\epsilon)$ confidence set results in

$$\theta_r = [\theta_{r0} \mid -2 \log \lambda \leq \chi_r^2(\epsilon)] , \tag{13}$$

and $(1-\epsilon)$ asymptotic confidence region of a and b must satisfy the relation

$$L(a, b) \leq \exp [- \frac{\chi_r^2(\epsilon)}{2}] \cdot L(\hat{a}, \hat{b}) . \tag{14}$$

4. Application to double exponential distribution

4.1 Estimation procedure

Double exponential distribution which is a type of extreme value (asymptotic smallest value) distributions belongs to location-scale family. The estimation method in section 3 can be applied to this distribution as an example of location-scale type. Its p.d.f. and c.d.f. are

$$f(x) = \frac{1}{a} \exp [(\frac{x-b}{a}) - \exp (\frac{x-b}{a})] , \tag{15-a}$$

and

$$F(x) = 1 - \exp [-\exp (\frac{x-b}{a})] , \tag{15-b}$$

respectively, whose mean and variance are given by

$$\mu = b - a\gamma,$$

and $(\gamma:$ Euler's const., $0.577\cdots)$ (16)

$$\sigma^2 = \frac{\pi^2 a^2}{6}.$$

The transformations corresponding to Eq. (4) are

$$y = \frac{x-b}{a}, \qquad g(y) = y - \exp(y),$$
$$Y = \frac{T-b}{a}, \quad \text{and} \quad h(Y) = -\exp(Y). \tag{17}$$

Likelihood function and its logarithmic form can be easily derived from Eqs. (5) and (6) as

$$L = \frac{C}{a^n}\exp\left[\sum_{i=1}^{n} y_i - \sum_{i=1}^{n}\exp(y_i) - \sum_{j=1}^{k} r_j \exp(Y_j)\right], \tag{18}$$

and

$$ln\,L = ln\,C - n\,ln\,a + \sum_{i=1}^{n} y_i - \sum_{i=1}^{n}\exp(y_i) - \sum_{j=1}^{k} r_j \exp(Y_j). \tag{19}$$

The differentiation of Eq. (17) with respect to y and Y yields

$$g_i' = 1 - \exp(y_i), \qquad g_i'' = -\exp(y_i),$$
$$h_j' = -\exp(Y_j), \quad \text{and} \quad h_j'' = -\exp(Y_j). \tag{20}$$

Thus, likelihood equation is reduced to the following simple simultaneous equation:

$$na + \sum_{i=1}^{n}(x_i - b) - \Sigma^*(x_i - b)\exp\left(\frac{x_i - b}{a}\right) = 0, \tag{21-a}$$

and

$$-\frac{a}{n} + \frac{1}{a}\,\Sigma^*\exp\left(\frac{x_i - b}{a}\right) = 0, \tag{21-b}$$

where Σ^* signifies summation over the entire N observations, and

$$\Sigma^*(x_i - b)^\delta \exp\left(\frac{x_i - b}{a}\right) = \sum_{i=1}^{n}(x_i - b)^\delta \exp\left(\frac{x_i - b}{a}\right) + \sum_{j=1}^{k} r_j (T_j - b)^\delta \exp\left(\frac{T_j - b}{a}\right), \quad (\delta = 0, 1, 2) \tag{22}$$

and

$$\Sigma^*\exp\left(\frac{x_i - b}{a}\right) = \sum_{i=1}^{n}\exp\left(\frac{x_i - b}{a}\right) + \sum_{j=1}^{k} r_j \exp\left(\frac{T_j - b}{a}\right).$$

From Eqs. (9) and (20), the asymptotic variance-covariance matrix of ML-estimates is written in the form:

$$V = (I)^{-1}_{\substack{a=\hat{a}\\ b=\hat{b}}} \simeq \frac{n}{(\det I)_{\substack{a=\hat{a}\\ b=\hat{b}}}} \begin{bmatrix} \hat{a}^{-2} & \dfrac{\bar{x}-\hat{b}}{\hat{a}} \\ \dfrac{\bar{x}-\hat{b}}{\hat{a}} & \hat{a}^2 + \dfrac{1}{n}\Sigma^*(x_i-\hat{b})^2 \exp(\dfrac{x_i-\hat{b}}{\hat{a}}) \end{bmatrix} . \tag{23}$$

Further the confidence region by use of likelihood ratio can be obtained by the same procedure as described in section 3.3.

4.2 Numerical calculation method of likelihood equation

It has been shown in the previous section that maximum likelihood estimators are given by Eq. (21). After a little reduction, this simultaneous equation can be rewritten as the simpler form:

$$a = \frac{\Sigma^* x_i \exp(\frac{x_i}{a})}{\Sigma^* \exp(\frac{x_i}{a})} - \bar{x} , \tag{24-a}$$

and

$$b = a [ln\, \Sigma^* \exp(\frac{x_i}{a}) - ln\, n] , \tag{24-b}$$

where

$$\Sigma^* x_i \exp(\frac{x_i}{a}) = \sum_{i=1}^{n} x_i \exp(\frac{x_i}{a}) + \sum_{j=1}^{k} r_j T_j \exp(\frac{T_j}{a}) , \tag{25-a}$$

and

$$\Sigma^* \exp(\frac{x_i}{a}) = \sum_{i=1}^{n} \exp(\frac{x_i}{a}) + \sum_{j=1}^{k} r_j \exp(\frac{T_j}{a}) . \tag{25-b}$$

Because Eq. (24-a) is free of parameter b, this estimation procedure may be initiated from solving Eq. (24-a) with respect to a. Since Eq. (24-a) is non-linear in a, however, it can not be explicitly solved but requires numerical methods to calculate the estimates. For this purpose, the following procedures of evaluating Eq. (24-a) can be considered:

(1) Construct a recurrence equation for a_m

$$a_{m+1} = \frac{\Sigma^* x_i \exp(\frac{x_i}{a_m})}{\Sigma^* \exp(\frac{x_i}{a_m})} - \bar{x} \qquad (m = 0, 1, 2, \cdots) \tag{26}$$

which is derived from Eq. (24-a) and gives the ML-estimate $\hat{a}$ as a convergent solution of a_m. Since successive solution of Eq. (26) shows usually oscillating behavior, substitute a modified approximation such as

$$a'_{m+1} = \frac{a_{m+1} + a_m}{2} \tag{27}$$

into a_m in the right-hand side of Eq. (26).

(2) Apply Newton-Raphson's method to

$$f(a) = a - \frac{\Sigma^* x_i \exp(\frac{x_i}{a})}{\Sigma^* \exp(\frac{x_i}{a})} + \bar{x} = 0, \tag{28}$$

then we have $\hat{a}$.

Another estimate $\hat{b}$ is also easily obtained by substitution of $\hat{a}$ into Eq. (24-b). The moment estimators for a and b are

$$a = \frac{\sqrt{6}\, s'}{\pi}, \tag{29-a}$$

and

$$b = \bar{x}' + a\gamma, \tag{29-b}$$

where

$$\bar{x}' = \bar{x} + \frac{\sum_{j=1}^{k} r_j T_j}{\sum_{j=1}^{k} r_j}, \tag{30-a}$$

and

$$s' = \left[\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x}')^2 + \sum_{j=1}^{k} r_j (T_j - \bar{x}')^2 / \sum_{j=1}^{k} r_j \right]^{\frac{1}{2}}. \tag{30-b}$$

The estimate $\hat{a}$ in Eq. (29-a) may be used as an initial value in Eqs. (26) and (28).

5. An example in CFRP fatigue life testing

Let double exponential distribution mentioned in section 4 be assumed as a lifetime distribution of CFRP, then the estimation procedures in section 4 give ML-estimates, variance-covariance matrix and confidence region.

Table 1 shows observed data[4], and each observed value presents the number of cycles to fatigue failure. Total number of specimens exposed to life test is 59 ($N = 59$), in which r_1 and r_2 are number of specimens censored at time T_1 and T_2 ($T_1 = 1.44$ and $T_2 = 3.31$), respectively. Throughout the life test, number of failure specimens is 18 ($n = 18$).

Table 2 stands for the calculation results of estimates of scale and location parameters, (a, b). In this table, $(\hat{a}_0, \hat{b}_0)$ are the initial values used in numerical calculation which are obtained by the moment method. $(\hat{a}_1, \hat{b}_1)$ and $(\hat{a}_2, \hat{b}_2)$ are ML-estimates obtained as solutions of the likelihood equation Eq. (24). $(\hat{a}_1, \hat{b}_1)$ calculated by a recurrence equation Eq. (26) coincide with $(\hat{a}_2, \hat{b}_2)$ which are the results by (2) in

Table 1. Uni-CFRP fatigue test data at stress level 90 kg/mm^2 (pulsating load type)

i	x_i	
1	0.30 ($\times 10^4$)	
2	0.44	
3	0.63	
4	0.70	
5	0.93	
6	1.02	
7	1.03	
8	1.28	
9	1.34	
	---------	T_1
10	1.66	
11	1.76	
12	1.77	
13	1.80	
14	2.22	
15	2.83	
	---------	T_2
16	8.83	
17	10.94	
18	14.50	

* Censored sample
$T_1 = 1.44$
$r_1 = 22$
$T_2 = 3.31$
$r_2 = 19$
($N = 59$, $n = 18$, $k = 2$)

Table 2. ML-estimates

$\hat{a}_0 = 4.31798$	$\hat{a}_1 = 3.79651$	$\hat{a}_2 = 3.79651$
$\hat{b}_0 = 7.79750$	$\hat{b}_1 = 8.58636$	$\hat{b}_2 = 8.58636$

section 4.2. However, procedure (1) prefers to (2) because calculation of estimates by the former converges faster than by the latter. Furthermore containing the evaluation of derivative, calculation in (2) is more complicated than that in (1). Consequently, the processing time required for the former calculation becomes less than the latter. It is also seen from the table that moment estimates $(\hat{a}_0, \hat{b}_0)$ which are taken for the initial values in numerical calculation are fairly different from ML-estimates.

The asymptotic variance-covariance matrix of ML-estimates $(\hat{a}_1, \hat{b}_1)$ is given in Table 3. The confidence region by likelihood ratio is illustrated in Fig. 1. This figure clarifies that the ML-estimates are within the narrower confidence region compared to that from moment estimates.

Table 4 represents the comparison between observed probability and theoretical one.

Table 3. Variance-covariance matrix

V_{11} = 0.00349	V_{12} = −0.07253
V_{21} = −0.07253	V_{22} = 2.33922

Table 4. Comparison between observed probability and theoretical probability (N = 59, n = 18, k = 2)

i	x_i	Obs. prob.	Theor. (i)	Theor. (ii)
1	0.30 (× 10^4)	0.016667	0.106614	0.145251
2	0.44	0.033333	0.110390	0.166698
3	0.63	0.050000	0.115714	0.197848
4	0.70	0.066667	0.117735	0.209851
5	0.93	0.083333	0.124611	0.250952
6	1.02	0.100000	0.127401	0.267612
7	1.03	0.116667	0.127715	0.269480
8	1.28	0.133333	0.135790	0.317008
9	1.34	0.150000	0.137797	0.328592
10	1.66	0.179762	0.148967	0.390773
11	1.76	0.209524	0.152623	0.410157
12	1.77	0.239286	0.152993	0.412090
13	1.80	0.269047	0.154108	0.417883
14	2.22	0.298809	0.170508	0.497221
15	2.83	0.328571	0.197102	0.602893
16	8.83	0.546825	0.655713	0.979051
17	10.94	0.765079	0.844151	0.993089
18	14.50	0.983333	0.991330	0.998946

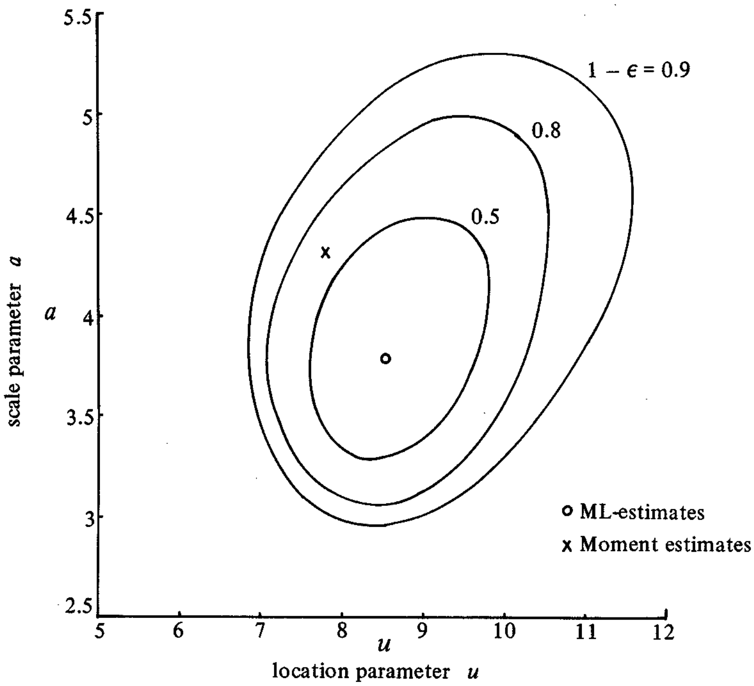

Fig. 1. Confidence region of estimates

The observed probability is calculated by the median rank method in [4], while there are two selections of parameters for calculation of theoretical probability: (i) ML-estimates $\hat{a}_1$ and $\hat{b}_1$ obtained in consideration of multi-censored sample, and (ii) ML-estimates calculated only from failure times x_i $(i = 1, 2, 3, \cdots, n)$. As seen in Table 4, the theoretical probability (i) agrees better with the observed probability than (ii).

6. Conclusion

An established procedure of ML-estimator from multi-censored samples is successfully proposed. Furthermore a calculation method of variance-covariance matrix and a determination procedure of confidence region for ML-estiamtes are also derived from the likelihood theory. A feature of this present method lies in its potentiality to propose the general formulae of estimating location and scale parameters from multi-censored samples. A numerical example of CFRP fatigue test data as a multi-censored sample suggests that present method is applicable to analysis of life test for other materials.

Reference

1) Cohen, A. C., "Multi-Censored Sampling in the Three Parameter Weibull Distribution", Technometrics, Vol. 17, No. 3. pp. 347–351 (1975).
2) Wingo, D. R., "Solution of the Three-Parameter Weibull equations Constrained Modified Quasilinearization (Progressively Censored Samples)", IEEE Trans. on Reliability, Vol. R-22, No. 2, pp. 96–102 (1972).
3) Kendall, M. G. and A. Stuart, The Advanced Theory of Statistics, Vol. 2, Griffin, London, 1969.
4) Nakayasu, H., Z. Maekawa, T. Fujii and K. Mizukawa, "Statistical Study on the Properties of Fatigue and Static Strength of Uni-Directional CFRP Subjected to Tensile Load", Proc. of the 19th Japan Congress on Materials Research, pp. 195–200 (1975).