AFML-TR-67-105
AF 61(052)-522
MARCH 1967

ESTIMATION OF DISTRIBUTION PARAMETERS BY A COMBINATION OF THE BEST LINEAR ORDER STATISTIC METHOD AND MAXIMUM LIKELIHOOD

WALODDI WEIBULL

ABSTRACT

This report consists of three parts, the first one dealing with the unbiased, minimum-variance estimation of location scale parameters, assuming the shape parameter to be known, the second one presenting formulae for computing the likelihood of a given sample, the third one specifying the estimation procedure. The first part developes general formulae for computing the coefficients of linear estimators, composed of all or part of the elements of a random sample. These formulae are specialized for the cases of exponential distributions and also for estimations, using two of the order statistics only. Formulae for expected values, variances and covariances of standardized Weibull order statistics are deduced and applied to a system of equations, which determines the linear coefficients. For the solution of such systems, a program has been written and applied to a IBM 7090 computer, which delivers the results extremely fast, thus eliminating the need of extensive tables. Tables of expected values and covariance matrices are presented for sample sizes n = 5, 10, 15, 20 and α = 0.1, 0.3, 0.5, 0.7, 0.9, 1.0, useful when no computer is available. The second part presents formulas for computing the likelihood of a given sample for the most general situation, that is, for arbitrarily censored, truncated or grouped samples, and, for the special case of life testing, when the sample may be composed of one subset of items, which have failed after observed time units, a second subset of items, which have accumulated observed time units, without failure, and a third subset of items, which have failed during one or more inspection periods, without knowing their exact life times. The third part defines the procedure of combining the preceding formulas for best estimation, when none of the parameters is known.

TABLE OF CONTENTS

TABLES

1. Introduction

The most useful two general methods for estimating distribution parameters are the Maximum Likelihood Method and the Best Linear Order Statistic Method. They have both their advantages and disadvantages.

The Maximum Likelihood Method is the most efficient one of all general methods, but it requires the solution of a system of rather complicated equations, and, above all, the estimates from small samples are heavily biased.

The Best Linear Order Statistic Method is very easy to use, if necessary tables are available, and it yields unbiased estimates. Its efficiency is less than that of the maximum likelihood method, even if in many cases the loss is quite negligible. It offers the great advantage that the efficiency of the estimates, even from small and censored samples, is easily stated.

The method, now proposed, starts with an unbiased estimation of the scale and location parameters for a properly chosen set of shape parameters by use of the best linear method. For each such set of three parameters, the corresponding likelihood of the sample is computed. Finally, that set which yields the maximum likelihood is determined by interpolation and accepted as the best estimate.

Pertinent formulae will now be deduced and presented below.

2. Best Linear Order Statistic Estimators

2.1 General formulas

From a given set of order statistics, unbiased estimates of the location parameter μ and the scale parameter β with minimum variance can be obtained on the condition that the shape parameter α is known.

The estimator

$$\mu^{*} = \Sigma a_i x_i \tag{1}$$

is a linear combination of all or part of the ordered elements x_i of a sample of size n, where a_i are coefficients to be determined.

From

$$x_i = \mu + \beta z_i \tag{2}$$

where Z is the standardized variate $(\mu = 0,\ \beta = 1)$, it follows that

$$\mu^* = \mu \cdot \Sigma a_i + \beta \cdot \Sigma a_i z_i \tag{3}$$

The estimates will be unbiased on the condition that, for any value of μ and β, the expected value of μ^* is equal to μ itself, that is,

$$E\mu^* = \mu \Sigma a_i + \beta \Sigma(a_i E z_i) = \mu$$

which is satisfied only if

$$\Sigma a_i = 1 \qquad \Sigma(a_i E z_i) = 0 \tag{4}$$

where $E z_i$ is the expected value of z_i and thus independent of the parameters μ and β.

Further, the variance of μ^* is

$$\mathrm{Var}\,\mu^* = \beta^2 . \Sigma a_i a_j \sigma_{ij} \tag{5}$$

where σ_{ij} is the covariance of z_i and z_j including also, for $i = j$, the variance σ_{ii}.

The best estimate is defined by the condition that

$$\mathrm{Var}\,\mu^* = \text{minimum} \tag{6}$$

with the side conditions (4).

This is a constrained minimum problem, which can be transformed into an unconstrained one by adding to $\mathrm{Var}\,\mu^*$ the two terms $\beta^2 k_1(\Sigma a_i - 1)$ and $\beta^2 k_2 \Sigma(a_i E z_i)$

Hence

$$\Sigma a_i a_j \sigma_{ij} + k_1(\Sigma a_i - 1) + k_2\Sigma(a_i E z_i) = \text{minimum} \tag{7}$$

Setting the derivatives with regard to a_i equal to zero and dividing by 2, we have

$$\Sigma a_i \sigma_{ij} + k_1 + k_2 E z_i = 0 \tag{8}$$

which, together with the conditions (4), provides the necessary $(n+2)$ equations.

The system of equations thus becomes

$$\left.\begin{array}{llllllll} a_1 & + a_2 & + \dots & + a_n & + 0 & + 0 & = 1 \\ a_1 E z_1 & + a_2 E z_2 & + \dots & + a_n E z_n & + 0 & + 0 & = 0 \\ a_1 \sigma_{11} & + a_2 \sigma_{12} & + \dots & + a_n \sigma_{1n} & + k_1 & + k_2 E z_1 & = 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_1 \sigma_{n1} & + a_2 \sigma_{n2} & + \dots & + a_n \sigma_{nn} & + k_1 & + k_2 E z_n & = 0 \end{array}\right\} \tag{9}$$

Multiplying the n last equations by $\beta^2 a_1, \dots \beta^2 a_n$, respectively, and adding, we have, considering equ.(4), the least variance attainable as

$$(\mathrm{Var}\ \mu^*)_{\min} = -k_1 \cdot \beta^2 \tag{10}$$

All or part of the elements of the sample may be used, in the latter alternative by putting equal to zero the coefficients a_i and those σ_{ij} that correspond to the omitted order statistics z_i.

Equ. (10) provides a measure of efficiency of the estimation.

Identically, the coefficients b_i of the unbiased best linear estimator

$$\beta^* = \Sigma\, b_i\, x_i \tag{11}$$

can be obtained, resulting in the system of equations

$$\left.\begin{array}{llllll}
b_1 & + b_2 & + \dots + b_n & + 0 & + 0 & = 0 \\
b_1 Ez_1 & + b_2 Ez_2 & + \dots + b_n Ez_n & + 0 & + 0 & = 1 \\
b_1 \sigma_{11} & + b_2 \sigma_{12} & + \dots + b_n \sigma_{1n} & + k_3 & + k_4 Ez_1 & = 0 \\
\dots & \dots & \dots & \dots & \dots & \dots \\
b_1 \sigma_{n1} & + b_2 \sigma_{n2} & + \dots + b_n \sigma_{nn} & + k_3 & + k_4 Ez_n & = 0
\end{array}\right\} \qquad (12)$$

and the least variance attainable

$$(\mathrm{Var}\,\beta^*)_{min} = -\,k_4 \beta^2 \qquad (13)$$

The general formulae above take simplified forms for the particular cases of an exponential distribution ($\alpha = 1$) and of estimations using two order statistics only, as will be demonstrated below.

The computations of the coefficients a_i, b_i, k_1, k_4 require the numerical values of the expected values Ez_i, the variances σ_{ii}, and the covariances σ_{ij}. Pertinent formulae are presented below.

2.2 The exponential distribution ($\alpha = 1$)

For this particular case, the estimates from a sample of size n with r_1 smallest and r_2 largest elements censored are given by Sarhan & Greenberg [2] [1] as follows

$$\mu^* = C\{[1/C + (n-r_1)\cdot \sum_{i=1}^{r_1+1} (1/(n-i+1)]\cdot x_{r_1+1} - (r_2 \sum_{i=1}^{r_1+1} (1/(n-i+1))\cdot x_{n-r_2} - \sum_{i=1}^{r_1+1} (1/(n-i+1))\cdot \sum x_i\} \qquad (14)$$

and

$$\beta^* = C[\sum_{i=r_1+1}^{n-r_2} x_i - (n-r_1)\cdot x_{r_1+1} + r_2 \cdot x_{n-r_2}] \qquad (15)$$

where $$C = 1/(n-r_1-r_2-1) \qquad (16)$$

The variances of the estimates are

$$\mathrm{Var}(\mu^{*}/\beta) = C[\sum_{i=1}^{r_1+1} (1/(n-i+1))]^2 + \sum_{i=1}^{r_1+1} (1/(n-i+1)^2) \tag{17}$$

and

$$\mathrm{Var}(\beta^{*}/\beta) = C \tag{18}$$

For a complete sample $(r_1 = r_2 = 0)$ thus

$$\mu^{*} = (n x_1 - \bar{x})/(n-1) \tag{19}$$
$$\beta^{*} = n(\bar{x} - x_1)/(n-1)$$

where $\bar{x} = \Sigma x_i/n$ and the variances

$$\mathrm{Var}(\mu^{*}) = 1/n(n-1)$$
$$\mathrm{Var}(\beta^{*}) = 1/(n-1) \tag{20}$$

2.3 Estimations using the best two observations

Applying the preceding formulae to the case of two order statistics z_i and z_j and simplifying the notations by putting

$$Ez_i = E_i \quad \text{and} \quad Ez_j = E_j$$

we have

$$a_i = E_j/(E_j - E_i) \; ; \quad a_j = -E_i/(E_j - E_i) \tag{21}$$

and

$$b_i = -1/(E_j - E_i) \; ; \quad b_j = 1/(E_j - E_i) \tag{22}$$

Consequently

$$\mu^{*} = (x_i E_j - x_j E_i)/(E_j - E_i) \tag{23}$$

with variance

$$\mathrm{Var}(\mu^{*}/\beta) = (\sigma_{ii} E^2{}_j + \sigma_{jj} E_i{}^2 - 2\sigma_{ij} E_i E_j)/E_j - E_i)^2 \tag{24}$$

and

$$\beta^{*} = (x_j - x_i)/(E_j - E_i) \tag{25}$$

with variance

$$\mathrm{Var}(\beta^*/\beta) = (\sigma_{ii} + \sigma_{jj} - 2\sigma_{ij})/(E_j - E_i)^2 \tag{26}$$

The best order numbers i,j are dependent on the shape parameter α and the sample size n. For the exponential distribution $(\alpha = 1)$, the maximum efficiency is, according to Sarhan & Greenberg (l.c.), attained for the following spacings:

Sample size	2...6	7...10	11...15	16...20	21
i	1	1	1	1	1
j	n	n-1	n-2	n-3	n-4

For sample sizes $n = 5$ the same rule holds for $0.1 \leq \alpha \leq 1$, but not for larger sample sizes. For example, the best values are $i = 2$, $j = 10$ for $\alpha = 0.1$. The best pair of order numbers (i,j) can be easily determined for any sample size by use of equs. (24) and (26).

2.4 Expected values, variances and covariances of the order statistics

Corresponding to the distribution function

$$P = 1 - e^{-z^{1/\alpha}} \tag{27}$$

we have, according to formulae derived by Lieblein [2]

the expected value of z_i

$$Ez_i = i.C_i^n.\alpha! \sum_{\mu=0}^{i-1} (-1)^{i-1-\mu}.C_\mu^{i-1}(n-\mu)^{-(1+\alpha)} \tag{28}$$

where

$$C_i^n = n!/i!(n-i)! \quad \text{and} \quad C_\mu^{i-1} = (i-1)!/\mu!(i-1-\mu)! \tag{29}$$

Considering that

$$K_1 = \sum Ez_i/n.\alpha! = 1 \tag{30}$$

the quantity K_1 can be used for checking the accuracy of the computed values of Ez_i.

the variance of z_i

$$\text{Var } z_i = \sigma_{ii} = E(z_i^2) - [E(z_i)]^2 \tag{31}$$

where

$$E(z_i^2) = i.C_i^n \,.\, (2\alpha)! \sum_{\mu=0}^{i-1} (-1)^{i-1-\mu}.C_\mu^{i-1}(n-\mu)^{-(1+2\alpha)} \tag{32}$$

The accuracy of the computed values of $E(z_i^2)$ can be checked by use of the quantity

$$K_2 = \sum_{i=1}^{n} E(z_i^2)/n.\ (2\alpha)! = 1 \tag{33}$$

the covariance of z_i and z_j

$$\text{Cov}(z_i, z_j) = \sigma_{ij}$$

$$= C(1+2\alpha)! \sum_{\mu=0}^{i-1} \sum_{v=0}^{j-i-1} (-1)^{\mu+v}.C_\mu^{i-1}.C_v^{j-i-1}.(tu)^{-(1+\alpha)}.B_p \tag{34}$$

where

$$C = C_j^n.C_{i-1}^j \,.\, (j-1)(j-i-1)$$

$$t = j-i+\mu-v \tag{35}$$

$$u = n-j+v+1$$

$$p = t/(t+u)$$

and B_p is the incomplete Beta function, defined by the series

$$B_p = \frac{p^{\alpha+1}}{\alpha+1} - \frac{\alpha.p^{\alpha+2}}{\alpha+2} + \frac{\alpha(\alpha-1).p^{\alpha+3}}{1.2(\alpha+3)} - \frac{\alpha(\alpha-1)(\alpha-2).p^{\alpha+4}}{1.2.3(\alpha+4)} + \ldots \tag{36}$$

By use of these formulae, programs for the IBM 7090 have been set up by Göran W. Weibull and used for the computation of the expected values and the covariance matrices for

$$n = 5,\ 10,\ 15,\ 20$$

$$\alpha = 0.1,\ 0.3,\ 0.5,\ 0.7,\ 0.9,\ 1.0$$

The values of $\alpha!$ and $(2\alpha)!$ were computed with twelve and those of the Beta function with fifteen significant figures.

The results are presented in the appended tables according to the scheme for each combination of (n,α)

$$
\begin{matrix}
Ez_1 & Ez_2 & Ez_3 & Ez_4 \\
\sigma_{11} & & & \\
\sigma_{21} & \sigma_{22} & & \\
\sigma_{31} & \sigma_{32} & \sigma_{33} & \\
\sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44}
\end{matrix}
$$

where the upper half of the covariances has been omitted, considering that $\sigma_{ij} = \sigma_{ji}$.

The computing time for each set α, $n = 5, 10, 15, 20$, that is, for 50 expected values, 50 variances, and 350 covariances was altogether 132 sec. This time was the same for all values of α except for $\alpha = 1$, where the covariances are equal to corresponding variances, as seen in the tables, and, consequently, the computing time was considerably less than for other values of α.

It is interesting to note that the short computing times eliminate the need of prepared tables, since the expected values and the covariance matrices can be directly provided by the computer for any actually wanted combination (n,α) as well for complete as for censored samples. Also the inflation of the precision in estimating the parameters, due to censoring of some of the elements, is easily and directly provided by the computer.

3. Likelihood Functions of a Sample with Given Parameters

The most general situation with regard to censoring of a sample arises in connection with life testing according to the following scheme, indicated by Hahn & Godfrey [3].

a) The available specimens are randomly assigned to groups to be exposed for differing predetermined testing times. For example, in a given program a small number of units may be put to a 10,000 hour extended life test while a larger sample goes through a 1,000 hour quality control test under the same stress conditions. Testing on any unit would of course be terminated if the unit fails before its planned removal time.

b) Specimens are placed on test at differing times and analysis is required at some random time before the conclusion of the program, at which time the unfailed items differ in exposure time.

c) There is random removal of specimens from test. This might occur when there is a breakdown of test equipment leading to premature test termination for some items. It also is frequently the case in evaluating component information from systems testing when test termination arises on account of system failure due to some component other than the one under study.

The present report gives formulae for the likelihood of a sample observed under the situation described above and for the parameters α, β, μ assumed to be known.

Suppose that a set of items are placed on test (not necessarily at the same time). Among these a first subset of I items are observed to fail after $y_1, y_2, \ldots, y_I$ time units of exposure. The likelihood then becomes

$$L_1 = c_1 . \prod_{i=1}^{I} f(y_i) \tag{37}$$

where f(y) is the density function and the value c is independent of the distribution parameters.

Further, a second subset of J items have accumulated $z_1, z_2, \ldots, z_J$ time units without failure. They are either randomly removed or at various prespecified times. The likelihood of this set is

$$L_2 = c_2 \prod_{j=1}^{J} [1-F(z_j)] \tag{38}$$

Finally, a third subset of d_k items have failed during one or more, say, K inspection periods $(w_k - w_{k-1})$, the number within each period being known but not their exact life-times. The likelihood of this set is

$$L_3 = c_3 \prod_{k=1}^{K} [F(w_k) - F(w_{k-1})]^{d_k} \tag{39}$$

In particular, for $k=1$ we have $w_{k-1}=w_o=\mu$. This subset corresponds to the case of d_1 smallest elements censored with the likelihood

$$L_{31} = c_{31}[F(w_1)]^{d_1} \tag{40}$$

For $k=K$ we have $w_K=\infty$. This set corresponds to the case of d_K largest elements censored with the likelihood

$$L_{3K} = c_{3K}[1-F(w_{k-1})]^{d_K} \tag{41}$$

In the general case when all three types of items are present, the likelihood is obtained by multiplication, that is,

$$L = L_1 . L_2 . L_3 \tag{42}$$

Since the coefficients c are independent of the parameters α, β, μ, they can, without loss of generality, be put equal to 1.

Now, assuming that the distribution functions of all items are identical and

$$F(x) = 1 - e^{-[(x-\mu)/\beta]^m} \qquad (m = 1/\alpha,\ x = y,z,w) \tag{43}$$

and

$$f(x) = \frac{m}{\beta}\left(\frac{x-\mu}{\beta}\right)^{m-1} . e^{-[(x-\mu)/\beta]^m} \tag{44}$$

the logarithms of the likelihoods become

$$\log L_1 = I \log(m/\beta) + (m-1)\Sigma \log[y_i-\mu)/\beta] + \Sigma[(y_i-\mu)/\beta]^m \tag{45}$$

$$\log L_2 = \Sigma[(z_i-\mu)/\beta]^m \tag{46}$$

$$\log L_3 = \Sigma d_k[e^{-[(w_{k-1}-\mu)/\beta]^m} - e^{-[(w_k-\mu)/\beta]^m}] \tag{47}$$

In the general case, the logarithm of the likelihood becomes the sum of the three logarithms (45) - (47).

4. The Estimation Procedure

For a properly chosen set of shape parameters $\alpha = 1/m$, the corresponding values of the parameters β and μ are computed by use of the estimators (1) and (11).

For each set of α, β, μ thus determined, the corresponding value of logL is computed by use of (45 - 47). From these values, the value of α which yields the maximum likelihood is determined by interpolation and the corresponding values of β and μ are finally estimated.

References

1. Sarhan, A. E. and Greenberg, B. G., Contributions to Order Statistics. J. Wiley & Sons, New York, 1962.

2. Lieblein, J., On moments of order statistics from the Weibull distribution, Ann. Math. Statist. 26, 1955, 330-333.

3. Hahn, G. J. and Godfrey, J. T., Estimation of Weibull distribution parameters with differing test times for unfailed items. Presented at 1963 Ann. Convention of the Amer. Stat. Ass., Cleveland, Ohio, Sept 4-7, 1963.