

AFML-TR-67-197

**ESTIMATION OF PARAMETERS FROM LARGE
SAMPLES ARBITRARILY CENSORED
OR TRUNCATED**

WALODDI WEIBULL

*LA ROSIAZ s/LAUSANNE
SWITZERLAND*

TECHNICAL REPORT AFML-TR-67-197

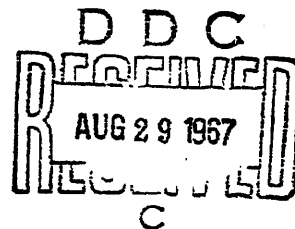
JUNE 1967

Distribution of this document is unlimited. It may be released to the
Clearinghouse, Department of Commerce, for sale to the general public.

AIR FORCE MATERIALS LABORATORY
RESEARCH AND TECHNOLOGY DIVISION
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO



Reproduced by the
CLEARINGHOUSE
for Federal Scientific & Technical
Information Springfield Va. 22151



AD 657318

**ESTIMATION OF PARAMETERS FROM LARGE
SAMPLES ARBITRARILY CENSORED
OR TRUNCATED**

WALODDI WEIBULL

*LA ROSIAZ s/LAUSANNE
SWITZERLAND*

Distribution of this document is unlimited. It may be released to the Clearinghouse, Department of Commerce, for sale to the general public.

FOREWORD

This report was prepared by Prof. Dr. Waloddi Weibull, La Rosiaz, Lausanne, Switzerland under USAF Contract No. AF 61(052)-943. This contract was initiated under Project No. 7351, "Metallic Materials", Task No. 735106, "Behavior of Metals". The contract was administered by the European Office, Office of Aerospace Research. The work was monitored by the Directorate of Materials and Processes, Aeronautical Systems Division, under the direction of Mr. W. J. Trapp.

This report covers work performed during the period May 1966 to May 1967.

The manuscript was released by the author 31 May 1967 for publication as a technical report.

This technical report has been reviewed and is approved.



W. J. TRAPP, Chief
Strength and Dynamics Branch
Metals and Ceramics Division
Air Force Materials Laboratory

ABSTRACT

Approximation formulas for the expected values, variances and coveriances of the order statistics y_i , which provide a very good approximation for sample sizes equal to or larger than $N = 20$, are developed. Their applications to graphical analysis and parameter estimations by use of desk computing machines and digital computers are demonstrated. Tables which simplify the computing procedure are presented.

TABLE OF CONTENTS

	Page
1. Introduction	1
2. Approximate expected values $E(y_i)$	2
3. Approximate covariances $Cov(y_i, y_j)$	4
4. Practical application of the formulas	6
4.1 Practical analysis	6
4.2 Estimations by use of desk machines	7
4.3 Estimations by use of digital computers	12
References	14
Tables	15 - 22
Illustration	23

LIST OF TABLES

	Page
1. Values of $c = i - N \cdot \bar{P}_i$ as function of $(i-1)/N$.	15
2. Values of $K_i = 0.56146 \cdot e^{\Sigma(i-1)^{-1}}$.	15
3. Accuracy of approximate values of \bar{P}_i	16
4. Values of the factors g_i and h_i for $N=20$ and 100 .	17
5. Accuracy of the approximate covariance matrix for $N=20$.	18
6. Accuracy of the approximate variances for $N=20$ and 100 .	19
7. Values of g_i, h_i, w_i and v_i as functions of \bar{P}_i .	20
8. Values of $y = \log[-\ln(1-P)]$ for P from $.001$ to $.999$.	21 - 22

Estimation of Parameters from Large Samples Arbitrarily Censored
or Truncated

1. Introduction

Large samples are usually presented as grouped data, that is, only the number of observations within a set of intervals are given and not the exact value of each element of the sample. Even if all values are known, it is a practical necessity to pick out a set of the ordered elements or to group the data. Very little information will be lost, if the selection of the order numbers is properly done, or if the number of intervals is sufficiently large.

The specific problem of estimation from large samples thus consists in finding methods applicable to arbitrarily censored samples. This requirement excludes the method of moments. Theoretically, the method of maximum-likelihood would be possible to apply also for this situation, but it involves heavy and tedious computations and the estimators are subjected to unknown biases. The asymptotic efficiency of complete samples is as good as could be, but for finite, censored samples the efficiency cannot be stated.

Remains to be examined the linear-order-statistic estimators. These estimators have many attractive properties. If properly constructed, they yield unbiased estimates with good efficiency, and they are extremely simple to handle. Furthermore, the variances of the estimates are easily computed for any type of censoring. The requisite for these advantages is, however, that the expected values and the covariance matrices of the order statistics are known. This is already the case for the y -estimators for sample sizes up to $N=20$, but for larger samples the covariance matrices become unmanageably large and acceptable accuracies require sophisticated computational means.

In order to eliminate this drawback a study of possible approximation formulas has been performed, resulting in formulas for the expected values and covariance matrices of the order statistics y_i , which are very simple and have a remarkable accuracy, as will now be demonstrated.

2. Approximate expected values $E(y_i)$

Let $e_i = E(y_i)$ be the expected value of y_i and \bar{P}_i the corresponding percentage point, defined by

$$\log[-\ln(1-\bar{P}_i)] = e_i \quad (1)$$

Putting

$$\bar{P}_i = (i-c)/N \quad (2)$$

we have

$$c = i - N\bar{P}_i \quad (3)$$

where c is a correction term. After having computed the values of c for various sample sizes it was found, as indicated in Sci.Rep.No.4, that c is a practically unique function of $(i-1)/N$ independent of the sample size, at least for $(i-1)/N \leq 80\%$, as demonstrated in Fig.1, where values of c have been plotted for $N=20, 50$ and 100 . The drawn curve may act as a master curve, by use of which very accurate plotting positions can be obtained. The values of c can also be read from Table 1 for given values of $(i-1)/N$.

This method is quite convenient for graphical examination of the sample data, but formulas are preferable for computational work. Such formulas were indicated in Sci.Rep.No.4. A straight line was substituted for the master curve with acceptable accuracy for $N=50$. (it should be noted that the required accuracy of c decreases with increasing N .)

A more refined approximation formula can be obtained in the following way: From the exact formulas for $E(y_i)$ it can be proved that

$$E(y_{i+1}) - E(y_i) \rightarrow \log e/i \quad \text{if } N \rightarrow \infty \quad (4)$$

Since for any sample size the exact value of $E(y_1)$ is

$$E(y_1) = C_1 - \log N \quad (5)$$

where

$$C_1 = -0.250,682 \quad (6)$$

it follows that

$$E(y_i) = C_1 - \log N + \log e \cdot \Sigma(i-1)^{-1} \quad (i \geq 2) \quad (7)$$

The corresponding value of \bar{P}_i is for small \bar{P}_i

$$E(y_i) = \log[-\ln(1-\bar{P}_i)] \approx \log \bar{P}_i \quad (8)$$

Hence, from (7) and (8),

$$N \cdot \bar{P}_i = 0.561,460 \cdot e^{\Sigma(i-1)^{-1}} \quad (9)$$

This expression yields a good approximation for small values of i/N . Its applicability can be remarkably extended by introducing a correction term $0.23i/N^2$.

With the notations

$$K_1 = 0.56146 \quad ; \quad K_i = 0.56146 \cdot e^{\Sigma(i-1)^{-1}} \quad (i \geq 2) \quad (10)$$

we thus arrive at the formula

$$\bar{P}_i = K_i/N - 0.23i/N^2 \quad (11)$$

Values of K_i are presented in Table 2. The quantity K_i is a function¹ of i only. It has the convenient property that, with increasing i ,

$$K_i \longrightarrow (i-0.5)$$

Hence, for $i \geq 10$, say, we may for any sample size put

$$\bar{P}_i = (i-0.5)/N - 0.23(i-1)/N^2 \quad (12)$$

The applicability of equ.(11) is demonstrated in Table 3, from which it can be seen that it yields a very close approximation of \bar{P}_i for $N=25$ and $(i-1)/N \leq 60\%$.

For $(i-1)/N \geq 60\%$ it should be replaced by

$$\bar{P}_i = (i-0.64)/N \quad (13)$$

3. Approximate covariances $\text{Cov}(y_i, y_j)$

In a report by Lieblein [1] it is quoted from Mosteller [2] that any order statistic X_i becomes, if $N \rightarrow \infty$ and i/N is not too near 0 or 1, normally distributed with mean, variance and covariance

$$\begin{aligned} E(y_i) &= t_{i/N} \\ V(y_i) &= \frac{i/N(1 - i/N)}{N[f(t_{i/N})]^2} \\ \text{Cov}(y_i, y_j) &= \frac{i/N(1 - j/N)}{N \cdot f(t_{i/N}) \cdot f(t_{j/N})} \end{aligned} \tag{14}$$

In particular, for the variate y the quantity $t_{i/N}$ becomes

$$t_{i/N} = \log\left[-\ln\left(1 - \frac{i}{N}\right)\right]$$

from which it can be concluded that the eqs.(14) are based on the assumption that $\bar{P}_i = i/N$, which is correct only as $N \rightarrow \infty$.

The approximations of equ.(14) can, for small and moderate sample sizes, be considerably improved by substituting the exact or even the approximate values \bar{P}_i for i/N .

Hence, defining, for convenience, the quantities g_i and h_i by

$$\begin{aligned} g_i &= \bar{P}_i / (1 - \bar{P}_i) \cdot \log(1 - \bar{P}_i) \\ h_i &= 1/10 \cdot \log(1 - \bar{P}_i) \end{aligned} \tag{15}$$

we have

$$\begin{aligned} V(y_i) &= \sigma_{ii} = 0.35574 \cdot g_i \cdot h_i / N \\ \text{Cov}(y_i, y_j) &= \sigma_{ij} = 0.35574 \cdot g_i \cdot h_j / N \end{aligned} \tag{16}$$

Values of g_i and h_i for $N=20$ and 100 are presented in Table 4.

It is a great convenience that the covariances can be represented by a product of two factors, one depending on i and the other on j only, because in that way the covariance matrix can be specified by $2N$ distinct values instead of by $N(N+1)/2$, which for $N=25$ reduces the number of values from 325 to 50.

The very good approximation of equ.(16) down to as small a sample as $N=20$ is demonstrated in Table 5, in which the quotients of approximate and exact values of variances and covariances are presented. The accuracy generally improves with increasing N , as demonstrated in Table 6, except for small or large order numbers, as will be illustrated by the following considerations.

If $N \rightarrow \infty$, then

$$\bar{P}_1 \rightarrow 0 ; -g_1 \rightarrow 2.30258 ; -h_1 \rightarrow 0.230,258/N$$

Observing that

$$N \cdot \bar{P}_1 \rightarrow 0.561,460$$

we have

$$-h_1 \rightarrow 0.410,106 ; g_1 h_1 \rightarrow 0.944,301 ; V(y_1) \rightarrow 0.335,926$$

Since the exact value of $V(y_1) = 0.310,254$ it follows that

$$V(y_1)_{\text{appr.}} \rightarrow 1.0827 \cdot V(y_1)_{\text{exact}}$$

The fact that the approximations of $V(y_i)$ for small order numbers do not improve with increasing N is, however, amply compensated by the small-order properties of the "weight" w_i which has to be attributed to the observations.

Defining the weight w_i of the i :th observation, as being proportional to its variance, by

$$w_i = 1/g_i h_i = 10(1 - \bar{P}_i) \cdot \log^2(1 - \bar{P}_i) / \bar{P}_i \quad (17)$$

it follows that w_i takes, for any sample size, a maximum

$$w_{\max} = 1.22145 \text{ for } \bar{P}_i = 0.795$$

and $V(y_i)$ takes a minimum

$$V(y_i)_{\min} = 0.29124/N$$

For large N the values of $V(y_i)$ tend to those given in the following table, from which the w_i/w_{\max} ratio have been computed as presented in the table

i	$V(y_i)$	$N \cdot w_i / w_{\max}$
1	0.310,254	0.94
2	0.121,644	2.39
3	0.074,561	3.91
4	0.053,591	5.43

From the table it is seen that for $N=100$ the weights of the small-order observations are only a few per cents of those of the middle part of the sample. The fact that the weights vary as $1:N$ within the sample makes it quite understandable that a neglect of weighting the observations (as frequently done) is strongly rejectable.

Values of g_i , h_i and w_i , as being functions of \bar{P}_i only, are presented in Table 7.

4. Practical application of the formulas

The use of the formulas depends on the computation tools at disposal. Three stages can be distinguished: paper and pencil, desk computing machine, digital computer. In the first case examination of the sample data is limited to graphical methods.

4.1 Graphical analysis

It is strongly recommended to start analysis of the observations x_i by plotting the values $\log x_i$ against the expected values e_i , in order to check whether the sample data

belong to a simple or a complex distribution, that is, one which is composed of two or more components, indicated by discontinuities in the plotted curve. There is no sense in estimating parameters from a sample belonging to a complex population before the components have been separated. If the location parameter $\mu \neq 0$, the curve can be rectified by plotting $\log(x_i - \mu)$ against e_i , where μ is that value, which provides a straight line, found by trial and error.

The values of e_i are determined by computing the values $(i-1)/N$ and reading the corresponding values of c in Table 1 or by use of the master curve in Fig.1. The correct plotting positions then are

$$\bar{P}_i = (i - c) / N \quad (2)$$

which can be directly used, if a probability paper is available. Otherwise, the corresponding value e_i can be read in Table 8 or computed by use of

$$e_i = \log \ln [N / (N - i + c)] \quad (1)$$

Then a straight line is fitted by eye to the data points. Its slope is a - not very accurate - estimate of the shape parameter $m = 1/\alpha$ and its intersection with the $\log(x - \mu)$ -axis an estimate of the parameter $(\log \beta)$.

Considering that the data points with small order numbers have very large variances compared to the middle part points, not too much consideration should be taken to them, when fitting the line.

4.2 Estimations by use of desk machines

The percentage points \bar{P}_i are computed by use of

$$\bar{P}_i = (i - 0.5) / N - 0.23(i - 1) / N^2 \quad (i - 1) / N \leq 60\% \quad (12)$$

$$\bar{P}_i = (i - 0.64) / N \quad (i - 1) / N \geq 60\% \quad (13)$$

For moderate sample sizes ($N = 20 \div 100$) and $i \leq 10$ the formula

$$\bar{P}_i = K_i / N - 0.23 \cdot i / N^2 \quad (11)$$

should be used.

The values e_i can then be read in Table 8 and g_i , h_i and w_i in Table 7 or computed by use of

$$e_i = \log[-\ln(1 - \bar{P}_i)] \quad (1)$$

$$g_i = \bar{P}_i / (1 - \bar{P}_i) \cdot \log(1 - \bar{P}_i) \quad (15)$$

$$h_i = 1/10 \log(1 - \bar{P}_i) \quad (15)$$

and then

$$\sigma_{ii} = 0.35574 \cdot g_i \cdot h_i / N ; \quad \sigma_{ij} = 0.35574 \cdot g_i \cdot h_j / N \quad (j \geq i) \quad (16)$$

The shape parameter α and the scale parameter $(\log \beta)$ are estimated by use of the linear estimators

$$\begin{aligned} \hat{\alpha} &= \sum a_i \cdot \log(x_i - \mu) \\ (\log \hat{\beta}) &= \sum b_i \cdot \log(x_i - \mu) \end{aligned} \quad (19)$$

and the coefficients a_i , b_i from the system of equations

$$\sum a_i = 0 ; \quad \sum a_i \cdot e_i = 1 ; \quad a_i = -w_i (k_1 + k_2 \cdot e_i) \quad (20)$$

$$\sum b_i = 1 ; \quad \sum b_i \cdot e_i = 0 ; \quad b_i = -w_i (k_3 + k_4 \cdot e_i) \quad (21)$$

These equations define "the first approximation" of the minimum-variance estimate in so far as all $\sigma_{ij} = 0$. Introducing the exact values of σ_{ij} makes the computations too heavy for a desk machine and will require a digital computer, as demonstrated in the next section. This approximation is quite acceptable when the data are grouped, because the mutual dependence of the order statistics decrease with the difference in order numbers. The "second approximation" is defined by putting all variances $\sigma_{ii} = 1$, that is, applying the least-squares method to the unweighted obser-

vations. This method should be rejected because of its poor efficiency.

From equ.(20) it follows that

$$a_i = \frac{w_i e_i \cdot \Sigma w_i \cdot \Sigma w_i e_i}{\Sigma w_i \cdot \Sigma w_i e_i^2 - (\Sigma w_i e_i)^2} \quad (22)$$

and from equ.(21) that

$$b_i = \frac{w_i \cdot \Sigma w_i e_i^2 - w_i e_i \cdot \Sigma w_i e_i}{\Sigma w_i \cdot \Sigma w_i e_i^2 - (\Sigma w_i e_i)^2} \quad (23)$$

It should be observed that, since w_i and e_i are uniquely determined by \bar{P}_i , this rule holds also for a_i and b_i .

Even if the covariances have been neglected in the preceding formulas, they are required for computing the variances (efficiencies) of the estimates, without surpassing the capacity of a desk machine.

Observing that

$$V[\log(x_i - \mu)] = \alpha^2 \cdot \sigma_{ii} \quad (24)$$

it follows that

$$\begin{aligned} V[\hat{\alpha}/\alpha] &= \Sigma a_i a_j \sigma_{ij} \\ V[(\hat{\log \beta})/\alpha] &= \Sigma b_i b_j \sigma_{ij} \end{aligned} \quad (25)$$

or from equ.(16)

$$\begin{aligned} 2.811 \cdot V[\hat{\alpha}/\alpha] &= \Sigma a_i g_i \cdot a_j h_j / N \\ 2.811 \cdot V[(\hat{\log \beta})/\alpha] &= \Sigma b_i g_i \cdot b_j h_j / N \end{aligned} \quad (26)$$

Observing that j must never be less than 1 the sum-

nation is performed in the following way:

$$\begin{aligned}
 2.811.N.V[\hat{\alpha}/\alpha] &= a_1 g_1 [a_1 h_1 + 2(a_2 h_2 + a_3 h_3 + \dots)] + \\
 & a_2 g_2 [a_2 h_2 + 2(a_3 h_3 + a_4 h_4 + \dots)] + \\
 & \text{etc.}
 \end{aligned}
 \tag{27}$$

and likewise for $V[(\log \hat{\beta})/\alpha]$

The formulas above can be used for any sample size and for any selection of the elements. They will be applied to the very simple case that the parameters are estimated from two observations only, having the order number i and j .

The coefficients a_i, b_i are then determined by the unbiased conditions, which implies that neither the variances nor the covariances enter into the equations and thus, in this particular case, both the first and the second approximation are identical with the exact solution.

Hence,

$$a_i + a_j = 0 \quad ; \quad a_i e_i + a_j e_j = 1$$

or

$$a_i = 1/(e_i - e_j) \quad a_j = 1/(e_i - e_j)$$

and

$$2.811.N.V(\hat{\alpha}/\alpha) = a_i g_i \cdot a_i h_i + 2a_i g_i \cdot a_j h_j + a_j g_j \cdot a_j h_j$$

If now $N \rightarrow \infty$, then $\bar{P}_i \rightarrow i/N$ and the formulas become identical with a percentile estimator, proposed and examined by Dubey[3]. He found that the minimum variance of the estimate was obtained by choosing $i/N = 0.1673$ and $j/N = 0.9737$, approximated by the 17th and 97th sample percentile. The Dubey formulas are exact for large sample sizes only.

Introducing $N = 20$; $i = 3$; $j = 19$, in the preceding formula we obtain

$$V(\hat{\alpha}/\alpha) = 1.04832/N$$

and for $N=100$; $i=17$; $j=97$ we obtain

$$V(\hat{\alpha}/\alpha) = 0.92372/N$$

The Dubey value, being $0.91627/N$, is very close to the exact value already for $N=100$.

The formulas in this section satisfy the requirement to obtain estimates with (approximately) minimum variances. The estimation procedure can, however, be regarded from another point-of-view.

It frequently occurs that life-time distributions are presented, not as numbers of failure, i , but as number of survivals, n , at given time points $T=t_i$. If N is the number of items existing at $T=0$, then

$$1 - P_i = 1 - i/N = (N - i)/N = n/N \quad (28)$$

Hence,

$$y_i = \log[-\ln(1 - P_i)] = \log \ln(N/n) = n \cdot \log t - n \cdot \log \beta \quad (29)$$

If now a table of n is given, it may, in some cases, be of greater interest to represent this table by an algebraic expression as true as possible than to have the most precise parameters. That is, the problem is more a fitting problem than an estimation problem. The conditions then become: to fit a curve to the data points with the least possible deviations between observed and calculated values, not of y , but of n .

From equ.(29) it follows that

$$dn = n \cdot \ln(N/n) \cdot dy / \log e \quad (30)$$

The least sum of squared deviations Δn will be obtained by introducing the weights (omitting some constants)

$$v_i = n^2 \cdot \log^2(N/n) = (1 - P_i)^2 \cdot \log^2(1 - P_i) \quad (31)$$

It is of interest to compare the two types of weight v_i and w_i . Both of them tend to 0 when $P_i \rightarrow 0$ or 1,

and both have a maximum

$$v_{\max} = 0.025,526 \quad \text{at } P_i = 0.632$$

$$w_{\max} = 1.22145 \quad \text{at } P_i = 0.795$$

The values of v_i are presented in Table 7, multiplied by 47.852 in order to have the same maximum values as w_i . From the table it can be concluded that the differences are not very large within the interval $P_i = 0.20 \div 0.80$. Outside this interval the values of v_i become very small compared to w_i . Accepting the rule to suppress all elements with a weight less than 2%, it follows that only elements corresponding to a P_i within the intervals

$$2\% \leq P_i \leq 99.98\% \quad \text{for the weight } w_i$$

$$5\% \leq P_i \leq 98\% \quad \text{for the weight } v_i$$

should be included in the computations.

In the preceding formulas, the location parameter μ is assumed to be known. This is frequently the case, for instance, with regard to strength of brittle materials, fatigue life of ball bearings and reliability of relays and electronic devices.

If, on the other hand, the parameter μ is unknown, it can be estimated by methods indicated in Sci.Rep.No.12 of Contract AF 61(052)-522 or in Sci.Rep.No.4 of this contract.

4.3 Computations by use of digital computers

If a digital computer is available, also the covariances can be included in the computations, resulting in improved efficiencies. Since the values of e_i , σ_{ii} and σ_{ij} have been determined, as indicated in Section 4.2, the σ_{ij} coefficients a_i , k_1 and k_2 can be computed by use of the system of equations

$$\begin{bmatrix} 1 & 1 & 1 & \dots & 0 & 0 \\ e_1 & e_2 & e_3 & \dots & 0 & 0 \\ \sigma_{11} & \sigma_{12} & \sigma_{13} & \dots & 1 & e_1 \\ \sigma_{21} & \sigma_{22} & \sigma_{23} & \dots & 1 & e_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \sigma_{n1} & \sigma_{n2} & \sigma_{n3} & \dots & 1 & e_n \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ \dots \\ k_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \dots \\ 0 \end{bmatrix} \quad (32)$$

The covariance matrix can be put in a more convenient form by multiplying each of the n last equations by $2.811.N$, resulting in the matrix

$$\begin{matrix} \varepsilon_1^{h_1} & \varepsilon_1^{h_2} & \varepsilon_1^{h_3} & \varepsilon_1^{h_4} \\ \varepsilon_2^{h_2} & \varepsilon_2^{h_2} & \varepsilon_2^{h_3} & \varepsilon_2^{h_4} \\ \varepsilon_3^{h_3} & \varepsilon_2^{h_3} & \varepsilon_3^{h_3} & \varepsilon_3^{h_4} \\ \varepsilon_4^{h_4} & \varepsilon_2^{h_4} & \varepsilon_3^{h_4} & \varepsilon_4^{h_4} \end{matrix} \quad (33)$$

considering that j must not be less than i and that $\sigma_{21} = \sigma_{12} = 0.35574 \varepsilon_i \cdot h_j / N$. The coefficients k_1 and k_2 are replaced by k'_1 and k'_2 and

$$k'_i = 2.811.N.k_i \quad (i=1,2) \quad (34)$$

Hence,

$$V(\hat{\alpha}/\alpha) = -0.35574 k'_1 / N \quad (35)$$

If the i :th element is excluded from the computation, the corresponding row and column shall be suppressed, that is, we have to put $a_i = \varepsilon_i = h_i = 0$.

References

1. Lieblein, J., "A new method of analysing extreme-value data.", NACA TN 3053, 1954.
2. Mosteller, F., "On some useful "inefficient" statistics.", Ann.Math.Statistics, 17, 1946, 377-408.
3. Dabey, S.D., "Some percentile estimators for Weibull parameters.", Technometrics, 2, 1967, 119-129.

Table 1. Values of $c = i - N\bar{P}_i$ as Function of $(i-1)/N$

$\frac{i-1}{N}$	c	$\frac{i-1}{N}$	c	$\frac{i-1}{N}$	c
.00	.440	.20	.547	.80	.650
.01	.478	.25	.559	.85	.651
.02	.490	.30	.570	.90	.648
.03	.497	.35	.581	.91	.647
.04	.502	.40	.591	.92	.645
.05	.506	.45	.601	.93	.643
.06	.509	.50	.611	.94	.641
.07	.513	.55	.620	.95	.639
.08	.516	.60	.628	.96	.636
.09	.519	.65	.635	.97	.633
.10	.521	.70	.642	.98	.632
.15	.535	.75	.647	.99	.645

Table 2. Values of $K_i = 0.56146 \cdot e^{\sum(i-1)^{-1}}$ ($i \geq 2$)

i	K_i	i	K_i	i	K_i
1	0.5615	16	15.5027	35	34.5013
2	1.5260	17	16.5025	40	39.5011
3	2.5163	18	17.5023	45	44.5010
4	3.5118	19	18.5022	50	49.5009
5	4.5092	20	19.5021	55	54.5008
6	5.5072	21	20.5020	60	59.5007
7	6.5064	22	21.5020	70	69.5007
8	7.5055	23	22.5019	80	79.5006
9	8.5049	24	23.5018	90	89.5006
10	9.5044	25	24.5017	100	99.5005
11	10.5040	26	25.5017	-	-
12	11.5036	27	26.5016	-	-
13	12.5033	28	27.5016	-	-
14	13.5031	29	28.5015	-	-
15	14.5029	30	29.5014	-	-

$$\bar{P}_i = K_i/N - 0.231/N^2$$

$K_i \longrightarrow (i - 0.5)$ with increasing i

Table 3. Accuracy of Approximate Values of \bar{P}_i -

N = 25			N = 100			N = 1000		
i	\bar{P}_i	Error	i	\bar{P}_i	Error	i	\bar{P}_i	Error
1	.02209	-.00012	1	.00559	-.00001	1	.000561	.000000
2	.06030	-.00011	2	.01521	-.00001	10	.009502	.000000
3	.08855	-.00008	3	.02509	-.00001	20	.019498	.000000
4	.13900	-.00007	4	.03503	.00000	30	.029494	-.000001
5	.17853	-.00006	5	.04498	.00000	40	.039492	-.000002
10	.37650	-.00013	10	.09481	.00000	50	.049489	-.000004
11	.41629	-.00018	20	.19456	.00000	100	.099477	-.000021
12	.45596	-.00023	30	.29432	.00000	200	.199454	-.000079
13	.49565	-.00030	40	.39409	-.00002	300	.299431	-.000157
14	.53535	-.00038	50	.49386	-.00005	400	.399408	-.000210
15	.57506	-.00046	60	.59363	-.00011	500	.499365	-.000203
						600	.599362	-.000044
16	.61440	-.00040	70	.69360	.00000	700	.699360	.000041
17	.65440	-.00015	80	.79360	.00009	800	.799360	.000054
18	.69440	.00007	90	.89360	.00009	900	.899360	.000026
19	.73440	.00026	95	.94360	.00001	1000	.999360	-.000019
20	.77440	.00042	96	.95360	-.00001	-	-	-
21	.81440	.00054	97	.96360	-.00004	-	-	-
22	.85440	.00062	98	.97360	-.00007	-	-	-
23	.89440	.00065	99	.98360	-.00008	-	-	-
24	.93440	.00066	100	.99360	.00005	-	-	-
25	.97440	.00101	-	-	-	-	-	-

The formula $\bar{P}_i = K_i/N - 0.23i/N^2$ used for $\begin{cases} N = 20 & i = 1 \div 15 \\ N = 100 & i = 1 \div 60 \\ N = 1000 & i = 1 \div 600 \end{cases}$

The formula $\bar{P}_i = (i - 0.64)/N$ used for $\begin{cases} N = 20 & i = 16 \div 20 \\ N = 100 & i = 70 \div 100 \\ N = 1000 & i = 700 \div 1000 \end{cases}$

Table 4. Values of the Factors g_i and h_i for $N = 20$ and 100

N = 20			N = 100		
i	$-g_i$	$-h_i$	i	$-g_i$	$-h_i$
1	2.3351	8.2026	1	2.3091	41.0012
2	2.3951	2.9405	2	2.3203	15.0130
3	2.4623	1.7414	3	2.3321	9.0579
4	2.5365	1.2090	4	2.3441	6.4571
5	2.6186	.9139	5	2.3564	5.0030
6	2.7102	.71389	10	2.4212	2.31159
7	2.8129	.59375	20	2.5707	1.06420
8	2.9291	.49684	30	2.7549	.66053
9	3.0619	.42225	40	2.9892	.45954
10	3.2157	.36287	50	3.2996	.33811
11	3.3960	.31430	60	3.7359	.25562
12	3.6112	.27367	70	4.4066	.19466
13	3.8736	.23897	80	5.6092	.14596
14	4.2023	.20878	90	8.6261	.10281
15	4.6289	.18203	95	13.3964	.08008
16	5.2098	.15785	96	15.4145	.07499
17	6.0584	.73548	97	18.4127	.06947
18	7.4444	.11409	98	23.4114	.06331
19	10.2227	.09251	99	33.7242	.05595
20	19.5862	.06776	100	70.3234	.04565

Table 5. Accuracy of the Approximate Covariance Matrix for N = 20

j \ i	1	2	4	6	8	10	12	14	16	18	20
1	1.098										
2	1.022	1.029									
3	1.025	1.009									
4	1.008	1.026	1.016								
5	1.005	1.007	1.003								
6	1.004	1.005	1.024	1.014							
7	1.003	1.004	1.007	1.003							
8	1.002	1.002	1.002	1.018	1.014						
9	1.001	1.001	1.004	1.007	1.008						
10	1.001	1.000	1.002	1.004	1.014	1.014					
11	1.000	0.999	1.001	1.003	1.008	1.009					
12	0.998	0.998	1.000	1.001	1.004	1.011	1.014				
13	0.997	0.998	0.998	1.000	1.002	1.006	0.998				
14	0.996	0.996	0.994	0.998	1.000	1.003	1.020	1.015			
15	0.995	0.994	0.995	0.996	0.998	1.001	1.005	1.000			
16	0.993	0.993	0.995	0.995	0.996	0.998	1.001	1.008	1.018		
17	0.991	1.000	0.992	0.992	0.993	0.995	0.997	1.002	1.008		
18	0.988	0.990	0.988	0.988	0.989	0.990	0.992	0.996	1.007	1.026	
19	0.996	0.981	0.981	0.982	0.982	0.983	0.985	0.988	0.995	1.008	
20	0.967	0.964	0.965	0.966	0.966	0.967	0.968	0.969	0.974	0.994	1.106

Table 6. Accuracy of the Approximate
Variances for N=20 and 100

i	N = 20	i	N = 100
1	1.098	1	1.086
2	1.029	2	1.018
3	1.019	3	1.008
4	1.016	4	1.005
5	1.015	5	1.003
6	1.015	10	1.002
7	1.014	20	1.001
8	1.014	30	1.002
9	1.014	40	1.002
10	1.014	50	1.003
11	1.014	60	1.003
12	1.014	70	1.003
13	1.015	80	1.004
14	1.015	90	1.005
15	1.016	95	1.010
16	1.018	96	1.012
17	1.020	97	1.015
18	1.026	98	1.022
19	1.040	99	1.037
20	1.106	100	1.113

Table 7. Values of g_i , h_i , w_i and v_i as Functions of \bar{P}_i -

\bar{P}_i	$-g_i$	$-h_i$	w_i	v_i
.005	2.30833	45.93688	.0094	.0000
.006	2.30954	38.26140	.0113	.0006
.008	2.31187	28.66725	.0151	.0009
.010	2.31420	22.91055	.0189	.0035
.020	2.32599	11.39744	.0377	.0139
.05	2.36268	4.48906	.0943	.0214
.08	2.40130	2.76150	.1508	.0531
.10	2.42826	2.18543	.1884	.0812
.20	2.57971	1.03188	.3757	.2876
.30	2.76672	.64557	.5599	.5625
.40	3.00505	.45076	.7383	.8478
.50	3.32193	.33219	.9062	1.0841
.60	3.76941	.25129	1.0557	1.2124
.70	4.46247	.19125	1.1717	1.1775
.80	5.72271	.14307	1.2214	.9351
.90	9.00000	.10000	1.1111	.4785
.95	14.60381	.07686	.8909	.2025
.97	21.23170	.06567	.7172	.0999
.98	28.84100	.05889	.5887	.0552
.99	49.50000	.05000	.4040	.0191
.992	59.13463	.04769	.3546	.0135
.995	86.48300	.04346	.2661	.0108
.997	131.72770	.03964	.1915	.0027
.999	333.00000	.03333	.0900	.0004
.9999	2,499.75000	.02500	.0160	.0000

Table 8.

VALUES OF $V=LOGI-LN(I-P)$ FOR P FROM .001 TO .999

P	0	1	2	3	4	DIFF	5	6	7	8	9	DIFF
.00	-2.9999	-2.9998	-2.9997	-2.9996	-2.9995	.0001	-2.9994	-2.9993	-2.9992	-2.9991	-2.9990	.0001
.01	-1.9978	-1.9977	-1.9976	-1.9975	-1.9974	.0001	-1.9973	-1.9972	-1.9971	-1.9970	-1.9969	.0001
.02	-1.6946	-1.6945	-1.6944	-1.6943	-1.6942	.0001	-1.6941	-1.6940	-1.6939	-1.6938	-1.6937	.0001
.03	-1.5018	-1.5017	-1.5016	-1.5015	-1.5014	.0001	-1.5013	-1.5012	-1.5011	-1.5010	-1.5009	.0001
.04	-1.3675	-1.3674	-1.3673	-1.3672	-1.3671	.0001	-1.3670	-1.3669	-1.3668	-1.3667	-1.3666	.0001
.05	-1.2899	-1.2898	-1.2897	-1.2896	-1.2895	.0001	-1.2894	-1.2893	-1.2892	-1.2891	-1.2890	.0001
.06	-1.2085	-1.2084	-1.2083	-1.2082	-1.2081	.0001	-1.2080	-1.2079	-1.2078	-1.2077	-1.2076	.0001
.07	-1.1392	-1.1391	-1.1390	-1.1389	-1.1388	.0001	-1.1387	-1.1386	-1.1385	-1.1384	-1.1383	.0001
.08	-1.0789	-1.0788	-1.0787	-1.0786	-1.0785	.0001	-1.0784	-1.0783	-1.0782	-1.0781	-1.0780	.0001
.09	-1.0254	-1.0253	-1.0252	-1.0251	-1.0250	.0001	-1.0249	-1.0248	-1.0247	-1.0246	-1.0245	.0001
.10	-0.9773	-0.9772	-0.9771	-0.9770	-0.9769	.0001	-0.9768	-0.9767	-0.9766	-0.9765	-0.9764	.0001
.11	-0.9335	-0.9334	-0.9333	-0.9332	-0.9331	.0001	-0.9330	-0.9329	-0.9328	-0.9327	-0.9326	.0001
.12	-0.8934	-0.8933	-0.8932	-0.8931	-0.8930	.0001	-0.8929	-0.8928	-0.8927	-0.8926	-0.8925	.0001
.13	-0.8562	-0.8561	-0.8560	-0.8559	-0.8558	.0001	-0.8557	-0.8556	-0.8555	-0.8554	-0.8553	.0001
.14	-0.8215	-0.8214	-0.8213	-0.8212	-0.8211	.0001	-0.8210	-0.8209	-0.8208	-0.8207	-0.8206	.0001
.15	-0.7891	-0.7890	-0.7889	-0.7888	-0.7887	.0001	-0.7886	-0.7885	-0.7884	-0.7883	-0.7882	.0001
.16	-0.7586	-0.7585	-0.7584	-0.7583	-0.7582	.0001	-0.7581	-0.7580	-0.7579	-0.7578	-0.7577	.0001
.17	-0.7297	-0.7296	-0.7295	-0.7294	-0.7293	.0001	-0.7292	-0.7291	-0.7290	-0.7289	-0.7288	.0001
.18	-0.7023	-0.7022	-0.7021	-0.7020	-0.7019	.0001	-0.7018	-0.7017	-0.7016	-0.7015	-0.7014	.0001
.19	-0.6763	-0.6762	-0.6761	-0.6760	-0.6759	.0001	-0.6758	-0.6757	-0.6756	-0.6755	-0.6754	.0001
.20	-0.6514	-0.6513	-0.6512	-0.6511	-0.6510	.0001	-0.6509	-0.6508	-0.6507	-0.6506	-0.6505	.0001
.21	-0.6276	-0.6275	-0.6274	-0.6273	-0.6272	.0001	-0.6271	-0.6270	-0.6269	-0.6268	-0.6267	.0001
.22	-0.6047	-0.6046	-0.6045	-0.6044	-0.6043	.0001	-0.6042	-0.6041	-0.6040	-0.6039	-0.6038	.0001
.23	-0.5828	-0.5827	-0.5826	-0.5825	-0.5824	.0001	-0.5823	-0.5822	-0.5821	-0.5820	-0.5819	.0001
.24	-0.5616	-0.5615	-0.5614	-0.5613	-0.5612	.0001	-0.5611	-0.5610	-0.5609	-0.5608	-0.5607	.0001
.25	-0.5411	-0.5410	-0.5409	-0.5408	-0.5407	.0001	-0.5406	-0.5405	-0.5404	-0.5403	-0.5402	.0001
.26	-0.5213	-0.5212	-0.5211	-0.5210	-0.5209	.0001	-0.5208	-0.5207	-0.5206	-0.5205	-0.5204	.0001
.27	-0.5021	-0.5020	-0.5019	-0.5018	-0.5017	.0001	-0.5016	-0.5015	-0.5014	-0.5013	-0.5012	.0001
.28	-0.4835	-0.4834	-0.4833	-0.4832	-0.4831	.0001	-0.4830	-0.4829	-0.4828	-0.4827	-0.4826	.0001
.29	-0.4654	-0.4653	-0.4652	-0.4651	-0.4650	.0001	-0.4649	-0.4648	-0.4647	-0.4646	-0.4645	.0001
.30	-0.4477	-0.4476	-0.4475	-0.4474	-0.4473	.0001	-0.4472	-0.4471	-0.4470	-0.4469	-0.4468	.0001
.31	-0.4305	-0.4304	-0.4303	-0.4302	-0.4301	.0001	-0.4300	-0.4299	-0.4298	-0.4297	-0.4296	.0001
.32	-0.4138	-0.4137	-0.4136	-0.4135	-0.4134	.0001	-0.4133	-0.4132	-0.4131	-0.4130	-0.4129	.0001
.33	-0.3974	-0.3973	-0.3972	-0.3971	-0.3970	.0001	-0.3969	-0.3968	-0.3967	-0.3966	-0.3965	.0001
.34	-0.3814	-0.3813	-0.3812	-0.3811	-0.3810	.0001	-0.3809	-0.3808	-0.3807	-0.3806	-0.3805	.0001
.35	-0.3657	-0.3656	-0.3655	-0.3654	-0.3653	.0001	-0.3652	-0.3651	-0.3650	-0.3649	-0.3648	.0001
.36	-0.3504	-0.3503	-0.3502	-0.3501	-0.3500	.0001	-0.3499	-0.3498	-0.3497	-0.3496	-0.3495	.0001
.37	-0.3353	-0.3352	-0.3351	-0.3350	-0.3349	.0001	-0.3348	-0.3347	-0.3346	-0.3345	-0.3344	.0001
.38	-0.3205	-0.3204	-0.3203	-0.3202	-0.3201	.0001	-0.3200	-0.3199	-0.3198	-0.3197	-0.3196	.0001
.39	-0.3060	-0.3059	-0.3058	-0.3057	-0.3056	.0001	-0.3055	-0.3054	-0.3053	-0.3052	-0.3051	.0001
.40	-0.2917	-0.2916	-0.2915	-0.2914	-0.2913	.0001	-0.2912	-0.2911	-0.2910	-0.2909	-0.2908	.0001
.41	-0.2777	-0.2776	-0.2775	-0.2774	-0.2773	.0001	-0.2772	-0.2771	-0.2770	-0.2769	-0.2768	.0001
.42	-0.2638	-0.2637	-0.2636	-0.2635	-0.2634	.0001	-0.2633	-0.2632	-0.2631	-0.2630	-0.2629	.0001
.43	-0.2502	-0.2501	-0.2500	-0.2499	-0.2498	.0001	-0.2497	-0.2496	-0.2495	-0.2494	-0.2493	.0001
.44	-0.2367	-0.2366	-0.2365	-0.2364	-0.2363	.0001	-0.2362	-0.2361	-0.2360	-0.2359	-0.2358	.0001
.45	-0.2234	-0.2233	-0.2232	-0.2231	-0.2230	.0001	-0.2229	-0.2228	-0.2227	-0.2226	-0.2225	.0001
.46	-0.2103	-0.2102	-0.2101	-0.2100	-0.2099	.0001	-0.2098	-0.2097	-0.2096	-0.2095	-0.2094	.0001
.47	-0.1973	-0.1972	-0.1971	-0.1970	-0.1969	.0001	-0.1968	-0.1967	-0.1966	-0.1965	-0.1964	.0001
.48	-0.1845	-0.1844	-0.1843	-0.1842	-0.1841	.0001	-0.1840	-0.1839	-0.1838	-0.1837	-0.1836	.0001
.49	-0.1718	-0.1717	-0.1716	-0.1715	-0.1714	.0001	-0.1713	-0.1712	-0.1711	-0.1710	-0.1709	.0001

