A FURTHER CONTRIBUTION TO THE THEORY OF SYSTEMATIC STATISTICS

by

Junjiro Ogawa
University of North Carolina

Sponsored by the Office of Naval Research under the contract for research in probability and statistics at Chapel Hill. Reproduction in whole or in part is permitted for any purpose of the United States government. Table 3.1 produced under sponsorship of the Office of Ordnance Research.

Institute of Statistics
Mimeoograph Series No. 168
April, 1957
A FURTHER CONTRIBUTION TO THE THEORY OF SYSTEMATIC STATISTICS\textsuperscript{1}

by

Junjiro Ogawa, Institute of Statistics, University of North Carolina

Introduction. Until 1945 the main interest in the field of statistical estimation seems to have been in the so-called "efficient" estimators. But from the point of view of economy in practical use, it seems reasonable to inquire whether the output of information is worth the input measured in money, man-hours, or otherwise. Thus we may ask whether comparable results could have been obtained by a smaller expenditure. From this standpoint F. Mosteller proposed $\sum_{i=1}^{2} \gamma_i$ in 1946 the use of order statistics for such purposes on the ground that, however large the sample size, the observations can easily be ordered in magnitude with the help of punch-card equipment. He considered the problem of estimation of the mean and standard deviation of a univariate normal population, and that of estimation of the correlation coefficient of a bivariate normal distribution. Then in 1951 J. Ogawa $\sum_{i=4}^{7} \gamma_i$ considered more systematically the problem of estimation of the location and the scale parameter of a population whose density function depends only on these two parameters, and obtained the optimum solutions in some cases for the univariate normal population.

There are many cases in which the samples are by their very nature

\textsuperscript{1}Sponsored by the Office of Naval Research under the contract for research in probability and statistics at Chapel Hill. Reproduction in whole or in part is permitted for any purpose of the United States government. Table 3.1 produced under sponsorship of the Office of Ordnance Research.

\textsuperscript{2}The numbers in square brackets refer to the bibliography listed at the end.
ordered in magnitude as for example in a life test of electric lamps or a fatigue test of a certain material. Usually in such cases the population probability density functions are assumed to be exponential. So at least for the exponential distribution estimation and testing of an hypothesis based upon systematic statistics are of great importance, from the point of view of application.

The first two sections of this paper will be devoted to the general theory of systematic statistics from the population whose density function depends only upon the location and scale parameters for sufficiently large values of sample size. Then in §3 the application of the general theory to the exponential distribution will be given and optimum spacing of the selected sample quantiles and the corresponding best estimators will be tabulated. Finally, in §4, discussion on testing an hypothesis will be presented.

1. **Relative Efficiencies.** The fundamental probability distribution of the large sample theory of systematic statistics from the population whose probability density function depends only on the location parameter m and the scale parameter σ is as follows. The limiting frequency function is

\[ f_{n,k} = \exp \left[ -\frac{n}{2\sigma^2} \left( \sum_{i=1}^{k} \frac{\lambda_{i+1} - \lambda_{i-1}}{\lambda_{i+1} - \lambda_{i}} \right) \sum_{i=1}^{k} \frac{1}{\lambda_{i+1} - \lambda_{i}} \right] \]

\[ = \frac{h(x(n_1), x(n_2), \ldots, x(n_k))}{\sum_{i=1}^{k} \frac{\lambda_{i+1} - \lambda_{i-1}}{\lambda_{i+1} - \lambda_{i}} \sum_{i=1}^{k} \frac{1}{\lambda_{i+1} - \lambda_{i}}} \]

\[ - 2 \sum_{i=1}^{k} \frac{f_{i} f_{i-1}}{\lambda_{i+1} - \lambda_{i}} \left( x(n_i) - m - \sigma u_i \right) \left( x(n_{i-1}) - m - \sigma u_{i-1} \right) \]

\[ , \quad (1.1) \]
where \( f(u) \) denotes the standardized frequency function of the population and \( u_i \) stands for the \( \lambda_i \)-quantile of the population, i.e.,

\[
\lambda_i = \int_{-\infty}^{u_i} f(t) dt, \quad i = 1, 2, \ldots, k
\]  

(1.2)

and

\[
f_i = f(u_i), \quad i = 1, 2, \ldots, k.
\]  

(1.3)

Finally, we put

\[
C_{n,k} = (2\pi \sigma^2)^{-k/2} f_1 f_2 \ldots f_k \prod_{i=1}^{k} \left( \lambda_i - \lambda_{i-1} \right)^{-\frac{k}{2}}, \quad (1.4)
\]

\[
\lambda_0 = 0, \quad \lambda_{k+1} = 1.
\]  

(1.5)

In the first place, we define the relative efficiency of the systematic statistics with respect to parameter to be the ratio of the amount of information in Fisher's sense derived from (1.1) to that derived from the original whole sample.

We shall consider the following two cases separately:

Case I. The case in which the scale parameter \( \sigma \) is known and the location parameter \( m \) is unknown. For the sake of convenience, put

\[
S = \sum_{i=1}^{k} \frac{\lambda_{i+1} - \lambda_{i-1}}{(\lambda_{i+1} - \lambda_i)(\lambda_i - \lambda_{i-1})} f_i^2 (x(n_i) - m - \theta u_i)^2
\]

\[
- 2 \sum_{i=2}^{k} \frac{f_{i-1} f_i}{\lambda_i - \lambda_{i-1}} (x(n_i) - m - \theta u_i)(x(n_{i-1}) - m - \theta u_{i-1}), \quad (1.6)
\]
where the right-hand side is the same as the expression in the exponent in the density function \( h(x_{(n_1)}, x_{(n_2)}, \ldots, x_{(n_k)}) \) except for the constant factor \(-n/2\sigma^2\), then we have

\[
\log h(x_{(n_1)}, \ldots, x_{(n_k)}) = -\frac{n}{2\sigma^2} S + \text{term independent of } m. \tag{1.7}
\]

Thus it follows by differentiation with respect to \( m \) that

\[
-\frac{\partial^2 \log h}{\partial m^2} = \frac{n}{2\sigma^2} \frac{\partial^2 S}{\partial m^2} = \frac{n}{\sigma^2} K_1, \tag{1.8}
\]

where

\[
K_1 = \sum_{i=1}^{k+1} \frac{(f_{i}-f_{i-1})^2}{\lambda_i - \lambda_{i-1}} \tag{1.9}
\]

and \( f_{k+1} = 0 \) and \( f_0 \) should be equal to 1 or 0 as the case may be. Hence the amount of information \( I_s(m) \) of the systematic statistics with respect to the location parameter \( m \) is asymptotically equal to

\[
I_s(m) = \mathcal{E} \left( \frac{\partial \log h}{\partial m} \right)^2 = -\mathcal{E} \left( \frac{\partial^2 \log h}{\partial m^2} \right) = \frac{n}{\sigma^2} K_1. \tag{1.10}
\]

The likelihood function of the original whole sample, considered as a random sample of size \( n \), is

\[
L = \sigma^{-n} \prod_{i=1}^{n} f\left( \frac{x_i - m}{\sigma} \right), \tag{1.11}
\]
hence we have

$$\log L = \sum_{i=1}^{n} \log f\left( \frac{x_i - m}{\sigma} \right) - n \log \sigma ,$$  \hspace{1cm} (1.12)$$

and consequently the amount of information \( I_0(m) \) of the original whole sample with respect to the location parameter \( m \) is equal to

$$I_0(m) = \mathcal{E} \left( \frac{\partial \log L}{\partial m} \right)^2 = \frac{1}{\sigma^2} \mathcal{E} \left( \sum_{i=1}^{n} \frac{f'(u_i)}{f(u_i)} \right)^2$$  \hspace{1cm} (1.13)$$

where \( U_i = (X_i - m)/\sigma \), i.e. \( U \) stands for the standardized variate.

If \( f(t) = (2\pi)^{-1/2} \exp \left\{ -\frac{t^2}{2} \right\} \), then

$$I_0(m) = \frac{n}{\sigma^2} \hspace{1cm} (1.14)$$

and if \( f(t) = e^{-t} \) for \( t > 0 \), then

$$I_0(m) = \frac{n^2}{\sigma^2} \hspace{1cm} (1.15)$$

Thus the relative efficiency of the systematic statistics with respect to the location parameter \( m \) is defined as:

$$\eta(m) = \frac{I_s(m)}{I_0(m)} = \frac{nK_1}{\mathcal{E} \left( \sum_{i=1}^{n} \frac{f'(u_i)}{f(u_i)} \right)^2} \hspace{1cm} (1.16)$$

In particular, for the normal distribution we have

$$\eta(m) = K_1 \hspace{1cm} (1.17)$$
and for the exponential distribution we get

\[ \eta_e(m) = \frac{1}{n} \xi_i = \frac{1}{n} \sim 0 \quad , \quad (1.18) \]

In the case of the one-sided exponential distribution defined by

\[ g(x) = \frac{1}{\sigma} e^{-\frac{x-m}{\sigma}} \quad \text{for} \quad x > m, \quad (1.19) \]

\[ = 0 \quad \text{otherwise}, \]

we know already that \( \min_{1 \leq i \leq n} x_i \) i.e. \( x(1) \) is the maximum likelihood estimator, \( x(1) \) is the uniformly minimum variance estimator of \( m \) for all values of \( n \). Thus, the estimator based upon the frequency function (1.1) becomes meaningless.

**Case 2.** The case in which the location parameter \( m \) is known and the scale parameter \( \sigma \) is unknown. In this case

\[ \log h = -k \log \sigma - \frac{n}{2\sigma^2} S + \text{a term independent of } \sigma. \quad (1.20) \]

Therefore we get

\[ \frac{\partial^2 \log h}{2\sigma^2} = \frac{k}{\sigma^2} - \frac{3nS^2}{\sigma^3} - \frac{2n}{3\sigma^2} \frac{\partial S}{\partial \sigma} - \frac{n}{2\sigma^2} \frac{\partial^2 S}{\partial \sigma^2} \quad (1.21) \]

and since it can easily be seen that

\[ \mathcal{L} \left( \frac{n}{\sigma^2} S \right) = k, \quad \mathcal{L} \left( \frac{\partial S}{\partial \sigma} \right) = 0, \quad \frac{\partial^2 S}{\partial \sigma^2} = 2k \quad (1.22) \]
where
\[ K_2 = \sum_{i=1}^{k+1} \frac{(f_{i-1}u_{i-1} - f_iu_i)^2}{\lambda_i \lambda_{i-1}}, \]  

(1.23)

we obtain finally the amount of information of the systematic statistics with respect to the scale parameter \( \sigma \), i.e.,
\[ I_g(\sigma) = \mathcal{E} \left( \frac{\partial \log h}{\partial \sigma} \right)^2 = \mathcal{E} \left( - \frac{\partial^2 \log h}{\partial \sigma^2} \right) = \frac{2k}{\sigma^2} + \frac{n}{\sigma^2} K_2, \]  

(1.24)

or
\[ I_g(\sigma) = \frac{n}{\sigma^2}(K_2 + \frac{2k}{n}) \sim \frac{n}{\sigma^2} K_2 \]  

(1.25)

On the other hand we have
\[ \frac{\partial \log h}{\partial \sigma} = -\frac{n}{\sigma} - \frac{1}{\sigma^2} \sum_{i=1}^{n} \frac{(x_i - m)}{f\left(\frac{x_i - m}{\sigma}\right)} \]  

(1.26)

Hence by making use of the general relation
\[ \mathcal{E} \left( \frac{Uf'(U)}{f(U)} \right) = 1, \]  

(1.27)

we obtain
\[ I_0(\sigma) = \mathcal{E} \left( \frac{\partial \log L}{\partial \sigma} \right)^2 = \mathcal{E} \left( \frac{1}{\sigma^2} \sum_{i=1}^{n} \frac{U_i f'(U_i)}{f(U_i)} \right)^2 \]  

\[ = \frac{n}{\sigma^2} \left\{ \mathcal{E} \left( \frac{Uf'(U)}{f(U)} \right)^2 - 1 \right\}. \]  

(1.28)

In particular, for the normal distribution we have
\[ I_0(\sigma) = \frac{2n}{\sigma^2} \]  

(1.29)

and for the exponential distribution we have

\[ I_0(\sigma) = \frac{n}{\sigma^2} \]  

(1.30)

Thus we obtain the relative efficiency of the systematic statistics with respect to the scale parameter \( \sigma \) as follows:

\[ \eta(\sigma) = \frac{I_0(\sigma)}{I_s(\sigma)} = \frac{\frac{2k}{n} + K_2}{\text{E} \left( \frac{Uf(I)}{F(U)} \right)^2 - 1} \sim \frac{K_2}{\text{E} \left( \frac{Uf(I)}{F(U)} \right)^2 - 1} \]  

(1.31)

In particular, for the normal distribution we have

\[ \eta(\sigma) = \frac{1}{2} K_2 \]  

(1.32)

and for the exponential distribution we have

\[ \eta(\sigma) = K_2 \]  

(1.33)

2. The Best Linear Unbiased Estimators of the Unknown Parameters Based upon the Selected Sample Quantiles for Sufficiently Large Sample Size \( n \).

In the circumstance now under consideration, the basic distribution is given by (1.1), and the unknown parameters are \( m \) and \( \sigma \). Since the distribution given in (1.1) is a \( k \)-dimensional normal distribution, we can apply
the extended Gauss-Markov theorem on least squares $\sqrt{3}$ in order to find the best linear unbiased estimators which are asymptotically efficient estimators. Hence, as far as the large sample problems are concerned there are no other estimators which are more efficient than the best linear unbiased ones.

Case I. The case in which the scale parameter $\sigma$ is known and the location parameter $m$ is unknown.

In this case, we can find the best linear unbiased estimator $\hat{m}$ of the location parameter $m$ by solving the single normal equation

$$\frac{\partial S}{\partial m} \bigg|_{m=\hat{m}} = 0,$$

which turns out to be

$$K_1 \hat{m} = X - K_2 \sigma^2,$$  \hspace{1cm} (2.1)

where

$$X = \sum_{i=1}^{k+1} \frac{(f_i-x_{i-1})(f_i-x_{i-1}^x(n_i-x_{i-1}))}{\lambda_i-\lambda_{i-1}},$$  \hspace{1cm} (2.2)

and

$$K_2 = \sum_{i=1}^{k+1} \frac{(f_i-x_{i-1})(f_i-u_{i-1}-f_i-x_{i-1}u_{i-1})}{\lambda_i-\lambda_{i-1}}.$$  \hspace{1cm} (2.3)

Hence the best linear unbiased estimator $\hat{m}$ of $m$ is

$$\hat{m} = \frac{1}{K_1} X - \frac{K_2}{K_1} \sigma,$$  \hspace{1cm} (2.5)
and it can easily be seen after small calculation that

$$\text{Var}(\hat{m}) = \sigma^2 \frac{1}{n \cdot K_1}.$$  \hspace{1cm} (2.6)

From (2.6), we can see that the best linear unbiased estimator $\hat{m}$ is an "efficient estimator" (so to speak) from the point of view of the relative efficiency given in the preceding section.

If in particular the frequency function $f(t)$ of the population is symmetric with respect to the origin, for example the normal distribution, and the spacing of the selected sample quantiles $x(n_1), x(n_2), \ldots, x(n_k)$ is also symmetric, i.e.,

$$n_i + n_{k-i+1} = n; \ i = 1, 2, \ldots, k$$  \hspace{1cm} (2.7)

or in terms of $\lambda_i$

$$\lambda_i + \lambda_{k-i+1} = 1; \ i = 1, 2, \ldots, k,$$  \hspace{1cm} (2.8)

then we have

$$u_i + u_{k-i+1} = 0; \ i = 1, 2, \ldots, k$$  \hspace{1cm} (2.9)

and

$$f_i = f_{k-i+1}; \ i = 1, 2, \ldots, k.$$  \hspace{1cm} (2.10)

In such a case, it follows clearly that
\[ K_3 = 0. \] (2.11)

Hence we have
\[ \hat{m} = \frac{1}{K_1} X, \] (2.12)

and
\[ \text{Var}(\hat{m}) = \sigma^2 \frac{1}{n} \frac{1}{K_1} \] (2.13)

**Case 2.** The case in which the location parameter \( m \) is known and the scale parameter \( \sigma \) is unknown.

In a quite similar manner, we can find the best linear unbiased estimator \( \hat{\sigma} \) of the scale parameter \( \sigma \) by solving the normal equation
\[ \frac{d\mathcal{S}}{d\sigma} \bigg| \sigma = \hat{\sigma} = 0. \] (2.14)

This comes to be
\[ K_2 \hat{\sigma} = Y - K_3 m, \] (2.15)

where
\[ Y = \sum_{i=1}^{k+1} \frac{(f_{i1} - f_{i-1} u_{i-1}) (f_{i1} x(n_1) - f_{i-1} x(n_{i-1}))}{\lambda_i - \lambda_{i-1}}. \] (2.16)

Hence we get
\[ \hat{\sigma} = \frac{1}{K_2} Y - m \frac{K_3}{K_2}. \] (2.17)
and

\[ \text{Var}(\hat{\sigma}) = \frac{\sigma^2}{n} \cdot \frac{1}{K_2} \]  \hspace{1cm} (2.18)

In the particular case in which \( f(t) \) and the spacing of the selected sample quantiles are symmetric we get as before

\[ \hat{\sigma} = \frac{1}{K_2} Y \]  \hspace{1cm} (2.19)

and the variance of \( \hat{\sigma} \) is the same as before.

3. Determination of the Optimum Spacings for the Estimation of the Scale Parameter of the One-Parameter Exponential Distribution Based on the Selected Sample Quantiles for Sufficiently Large Sample Size \( n \).

In \( \int_4^7 \), the author developed the general theory of estimation of the location and the scale parameters based upon the selected sample quantiles for sufficiently large sample size \( n \) and determined the optimum spacings in the case of the normal population.

Owing to its importance in practical applications, optimum spacings are calculated herein for the estimation of the scale parameter \( \sigma \) of the exponential distribution whose density function is given as follows:

\[ g(x) = \begin{cases} 
\frac{1}{\sigma} e^{-\frac{x}{\sigma}}, & \text{if } x > 0 \\
0, & \text{otherwise}
\end{cases} \]  \hspace{1cm} (3.1)
Suppose we are given an ordered sample of size $n$, where $n$ is sufficiently large. In other words, the sample size $n$ is large enough so that the conclusions drawn making use of the limit distribution are valid with enough accuracy. For given $k$ real numbers such that

$$0 < \lambda_1 < \lambda_2 < \ldots < \lambda_k < 1,$$

we select $k$ sample $\lambda_i$-quantiles for $i = 1, 2, \ldots, k$, i.e., $k$ order statistics

$$x^{(n_1)}, x^{(n_2)}, \ldots, x^{(n_k)},$$

where

$$n_i = \lceil n\lambda_i \rceil + 1, \quad i = 1, \ldots, k$$

and the symbol $\lceil n\lambda_i \rceil$ stands for the greatest integer not exceeding $n\lambda_i$.

Furthermore, let the $\lambda_i$-quantile of the standardized exponential distribution, whose density is given by

$$f(x) = \begin{cases} 
  e^{-x}, & \text{if } x > 0 \\
  0, & \text{otherwise}
\end{cases}$$

be $u_i$, and that of (3.1) be $x_i$, then it is clear that

$$x_i = u_i \sigma, \quad i = 1, 2, \ldots, k.$$ (3.5)

The limiting joint distribution of the $k$ sample quantiles $x^{(n_1)}, \ldots, x^{(n_k)}$ has the density function
\[ c_{n,k} \exp \left[ - \frac{n}{2\sigma^2} \left\{ \sum_{i=1}^{k} \frac{\lambda_{i+1} - \lambda_{i-1}}{(\lambda_{i+1} - \lambda_1)(\lambda_i - \lambda_{i-1})} f_i^2(x(n_i) - u_1 \sigma)^2 \right. \right. \\
\left. \left. - 2 \sum_{i=2}^{k} \frac{f_if_{i-1}}{\lambda_i - \lambda_{i-1}} (x(n_i) - u_1 \sigma)(x(n_{i-1}) - u_{i-1} \sigma) \right\} \right], \quad (3.6) \]

where

\[ f_i = e^{-u_1}, \quad i = 1, 2, \ldots, k \quad (3.7) \]

\[ \lambda_i = 1 - e^{-u_1}, \quad i = 1, 2, \ldots, k \quad (3.7) \]

Hence we get

\[ \log h = - \frac{k}{2} \log \sigma^2 - \frac{n}{2\sigma^2} S + \text{term which is independent of } \sigma \quad (3.8) \]

where

\[ S = \sum_{i=1}^{k} \frac{\lambda_{i+1} - \lambda_{i-1}}{(\lambda_{i+1} - \lambda_1)(\lambda_i - \lambda_{i-1})} f_i^2(x(n_i) - u_1 \sigma)^2 \\
- 2 \sum_{i=2}^{k} \frac{f_if_{i-1}}{\lambda_i - \lambda_{i-1}} (x(n_i) - u_1 \sigma)(x(n_{i-1}) - u_{i-1} \sigma), \quad (3.9) \]

Thus we get the amount of information of the systematic statistics with respect to the scale parameter \( \sigma \) in this case.
\[ I_s(c) = E\left(-\frac{\partial^2 \log h}{\partial c^2}\right) = \frac{2k}{\sigma^2} + \frac{n}{\sigma^2} K_2, \] (3.10)

where

\[ K_2 = \sum_{i=1}^{k+1} \left( \frac{(u_i - u_i - u_{i-1})^2}{\lambda_i - \lambda_{i-1}} \right) = \sum_{i=1}^{k+1} \frac{\left( e^{-u_i} - e^{-u_i - u_{i-1}} \right)^2}{e^{-u_i - u_{i-1}} - e^{-u_i}}. \] (3.11)

Consequently we get the relative efficiency of the systematic statistics with respect to \( \sigma \) as

\[ \eta(c) = K_2. \] (3.12)

The best linear unbiased estimator \( \hat{\sigma} \) of \( \sigma \) is

\[ \hat{\sigma} = \frac{1}{K_2} Y, \] (3.13)

where

\[ Y = \sum_{i=1}^{k+1} \frac{(e^{-u_i} - e^{-u_i - u_{i-1}})(e^{-u_i x(n_i)} - e^{-u_i - u_{i-1} x(n_{i-1})})}{e^{-u_i - u_{i-1}} - e^{-u_i}} \] (3.14)

and this can be written as

\[ Y = \sum_{i=1}^{k} a_i x(n_i) \] (3.15)

where
\[
a_i = e^{u_i} \left\{ \frac{e^{-u_i} - e^{u_i-1}}{e^{-u_i-1} - e^{-u_i}} - \frac{e^{-u_i+1} - e^{u_i}}{e^{-u_i} - e^{-u_i+1}} \right\}, \quad (3.16)
\]

The coefficients \(a_i\) for the optimum spacings are calculated in Table 3.1.

From (3.13) and (3.15) we have

\[
\hat{\sigma} = \sum_{i=1}^{k} a_i' x_i(n_i)
\]

where

\[
a_i' = \frac{a_i}{K_2}, \quad i = 1, 2, \ldots, k, \quad (3.17)
\]

Although it may appear to be more convenient to tabulate \(a_i'\) than \(a_i\), for \(k = 10\), for instance, there are ten rounding errors in using (3.17), while the calculation of \(\hat{\sigma}\) by the formula

\[
\hat{\sigma} = \sum_{i=1}^{k} \frac{a_i x_i(n_i)}{K_2}
\]

has the error obtained by only one rounding process.

The required optimum spacing is the set \((\lambda_1, \lambda_2, \ldots, \lambda_k)\), or equivalently the set of values \((u_1, u_2, \ldots, u_k)\) which gives the maximum value of the relative efficiency \(K_2\). We can obtain the optimum spacings step by step as follows:
(i) The case where $k = 1$: in this case it is easily seen that

$$K_2 = \frac{u_1^2}{u_1 e^{-1}}$$

which has only one maximum $K_2 = 0.6471$ at $u_1 = 1.59$, and the corresponding probability is $\lambda_1 = 1 - e^{-1.59} = 0.7961$.

(ii) The case where $k = 2$: put $u_2 = u_1 + x$, then we have

$$K_2 = e^{-u_1} \left( \frac{u_1^2}{u_1 e^{-1}} + u_1^2 + \frac{x^2}{e^{x-1}} \right)$$

and this function of $x$ and $u_1$ is maximized at the point

$$u_1 = 1.02, \quad x = 1.59$$

i.e.,

$$u_1 = 1.02, \quad u_2 = 2.61$$

and the maximum value of $K_2$ is

$$K_2 = 0.8203$$

(iii) The case where $k = 3$: put

$$u_2 = u_1 + x, \quad u_3 = u_2 + y = u_1 + x + y,$$

then it follows after some calculations that
\[ K_2 = e^{-u_1} \left[ \frac{u_1^2}{u_1 - 1} + u_1^2 + e^{-x} \left\{ \frac{x^2}{e^x - 1} + x^2 + \frac{y^2}{e^y - 1} \right\} \right], \quad (3.22) \]

It is easily seen that this function of \(x, y\) and \(u_1\) is maximized at the point

\[ u_1 = 0.75, \quad x = 1.02, \quad y = 1.59. \]

Hence we get the optimum spacing in this case as follows:

\[ u_1 = 0.75, \quad u_2 = 1.77, \quad u_3 = 3.36, \]

and the corresponding maximum value of \(K_2\) is

\[ K_2 = 0.8905. \quad (3.23) \]

In a similar way, the following results in Table 3.1 have been calculated. In Table 3.1, \(u_1\) stands for the \(\lambda_1\)-quantile of the standardized population. The bottom row of the table which representing the relative efficiencies \(K_2\) has been graphed as Fig. 1 to show the increasing rate of relative efficiencies against the number of sample quantiles which has been selected. It will be seen from Fig. 1 that after \(k = 10\), the gain in relative efficiency is not appreciable.
Table 5.1 Optimum Spacings for Estimates of Relative Efficiencies and the Coefficients of Best Estimates

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.59</td>
<td>1.02</td>
<td>0.75</td>
<td>0.61</td>
<td>0.50</td>
<td>0.45</td>
<td>0.37</td>
<td>0.33</td>
<td>0.30</td>
<td>0.27</td>
<td>0.25</td>
<td>0.23</td>
<td>0.21</td>
<td>0.20</td>
<td>0.19</td>
</tr>
<tr>
<td>2</td>
<td>2.61</td>
<td>1.77</td>
<td>1.36</td>
<td>1.11</td>
<td>0.93</td>
<td>0.80</td>
<td>0.70</td>
<td>0.63</td>
<td>0.57</td>
<td>0.52</td>
<td>0.48</td>
<td>0.44</td>
<td>0.41</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3.36</td>
<td>2.38</td>
<td>1.86</td>
<td>1.54</td>
<td>1.30</td>
<td>1.13</td>
<td>1.00</td>
<td>0.90</td>
<td>0.82</td>
<td>0.75</td>
<td>0.69</td>
<td>0.64</td>
<td>0.60</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

...
Fig. 1. The curve of relative efficiencies against the number of selected sample quantiles to be used.
Example: As an illustration of the estimation procedure, we shall calculate the estimate of the standard deviation of the time intervals in days between explosions in mines, involving more than 10 men killed from 6th December 1875 to 29th May 1951. The data are from Maguire, Pearson and Wynn \( \sqrt{1.7} \). It should be noted that this example may not be the best to point out the value of this method because the sample size \( n = 109 \) is not so large that the labor necessary for the classical estimation is almost comparable to that necessary for our estimate. This example should be regarded as an illustration of the calculating procedure itself. There are cases, which are important in applications such as in life testing of the electric lamps and in fatigue tests of certain material, where the samples are ordered in magnitudes in natural order. In such a case the procedure here may be of great help in getting quickly the estimate of \( \sigma \), especially for a large bulk of data.

The Table 1 of B. A. Maguire, E. S. Pearson and A. H. A. Wynn is reproduced here with observations arranged in order of magnitude as follows:
### Table 3.2. Time intervals in days between explosions in mines, involving more than 10 men killed from 6 Dec. 1875 to 29 May 1951.

(B. A. Maguire, E. S. Pearson and A. H. A. Wynn)

<table>
<thead>
<tr>
<th>Order</th>
<th>Observation</th>
<th>Order</th>
<th>Observation</th>
<th>Order</th>
<th>Observation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>40</td>
<td>113</td>
<td>91</td>
<td>354</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>47</td>
<td>114</td>
<td>92</td>
<td>361</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>48</td>
<td>120</td>
<td>93</td>
<td>364</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>49</td>
<td>120</td>
<td>94</td>
<td>369</td>
</tr>
<tr>
<td>5</td>
<td>11</td>
<td>50</td>
<td>123</td>
<td>95</td>
<td>378</td>
</tr>
<tr>
<td>6</td>
<td>13</td>
<td>51</td>
<td>124</td>
<td>96</td>
<td>390</td>
</tr>
<tr>
<td>7</td>
<td>15</td>
<td>52</td>
<td>129</td>
<td>97</td>
<td>457</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>53</td>
<td>131</td>
<td>98</td>
<td>467</td>
</tr>
<tr>
<td>9</td>
<td>17</td>
<td>54</td>
<td>137</td>
<td>99</td>
<td>498</td>
</tr>
<tr>
<td>10</td>
<td>18</td>
<td>55</td>
<td>145</td>
<td>100</td>
<td>517</td>
</tr>
<tr>
<td>11</td>
<td>19</td>
<td>56</td>
<td>151</td>
<td>101</td>
<td>566</td>
</tr>
<tr>
<td>12</td>
<td>19</td>
<td>57</td>
<td>156</td>
<td>102</td>
<td>644</td>
</tr>
<tr>
<td>13</td>
<td>20</td>
<td>58</td>
<td>171</td>
<td>103</td>
<td>745</td>
</tr>
<tr>
<td>14</td>
<td>20</td>
<td>59</td>
<td>176</td>
<td>104</td>
<td>871</td>
</tr>
<tr>
<td>15</td>
<td>22</td>
<td>60</td>
<td>182</td>
<td>105</td>
<td>1205</td>
</tr>
<tr>
<td>16</td>
<td>23</td>
<td>61</td>
<td>188</td>
<td>106</td>
<td>1312</td>
</tr>
<tr>
<td>17</td>
<td>28</td>
<td>62</td>
<td>189</td>
<td>107</td>
<td>1357</td>
</tr>
<tr>
<td>18</td>
<td>29</td>
<td>63</td>
<td>195</td>
<td>108</td>
<td>1613</td>
</tr>
<tr>
<td>19</td>
<td>31</td>
<td>64</td>
<td>203</td>
<td>109</td>
<td>1650</td>
</tr>
<tr>
<td>20</td>
<td>32</td>
<td>65</td>
<td>208</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>36</td>
<td>66</td>
<td>215</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>37</td>
<td>67</td>
<td>217</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>47</td>
<td>68</td>
<td>217</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>48</td>
<td>69</td>
<td>217</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>49</td>
<td>70</td>
<td>224</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>50</td>
<td>71</td>
<td>228</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>54</td>
<td>72</td>
<td>233</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>54</td>
<td>73</td>
<td>255</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>55</td>
<td>74</td>
<td>271</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>58</td>
<td>75</td>
<td>275</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>59</td>
<td>76</td>
<td>275</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>59</td>
<td>77</td>
<td>275</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>61</td>
<td>78</td>
<td>286</td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>61</td>
<td>79</td>
<td>286</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>66</td>
<td>80</td>
<td>312</td>
<td></td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>72</td>
<td>81</td>
<td>312</td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>72</td>
<td>82</td>
<td>315</td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>75</td>
<td>83</td>
<td>326</td>
<td></td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>78</td>
<td>84</td>
<td>326</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>78</td>
<td>85</td>
<td>329</td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>81</td>
<td>86</td>
<td>330</td>
<td></td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>93</td>
<td>87</td>
<td>336</td>
<td></td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>96</td>
<td>88</td>
<td>338</td>
<td></td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>99</td>
<td>89</td>
<td>345</td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>108</td>
<td>90</td>
<td>348</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
For \( k = 10 \), we can see from Table 3.1 that
\[
\begin{align*}
n_1 &= \sqrt{n \lambda_1} + 1 = \sqrt{109 \times .23662} + 1 = 26 \\
n_2 &= \sqrt{n \lambda_2} + 1 = \sqrt{109 \times .43447} + 1 = 48 \\
n_3 &= \sqrt{n \lambda_3} + 1 = \sqrt{109 \times .59343} + 1 = 65 \\
n_4 &= \sqrt{n \lambda_4} + 1 = \sqrt{109 \times .71917} + 1 = 79 \\
n_5 &= \sqrt{n \lambda_5} + 1 = \sqrt{109 \times .81732} + 1 = 90 \\
n_6 &= \sqrt{n \lambda_6} + 1 = \sqrt{109 \times .88920} + 1 = 97 \\
n_7 &= \sqrt{n \lambda_7} + 1 = \sqrt{109 \times .93980} + 1 = 103 \\
n_8 &= \sqrt{n \lambda_8} + 1 = \sqrt{109 \times .97156} + 1 = 106 \\
n_9 &= \sqrt{n \lambda_9} + 1 = \sqrt{109 \times .98931} + 1 = 108 \\
n_{10} &= \sqrt{n \lambda_{10}} + 1 = \sqrt{109 \times .99791} + 1 = 109
\end{align*}
\]
Thus, we used the ordered observations as follows:

<table>
<thead>
<tr>
<th>Order ((n_1))</th>
<th>Observations ((x_{(n_1)})</th>
<th>Coefficients ((a_1))</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>50</td>
<td>.21646</td>
</tr>
<tr>
<td>48</td>
<td>120</td>
<td>.17729</td>
</tr>
<tr>
<td>65</td>
<td>208</td>
<td>.14135</td>
</tr>
<tr>
<td>79</td>
<td>291</td>
<td>.11122</td>
</tr>
<tr>
<td>90</td>
<td>348</td>
<td>.08396</td>
</tr>
<tr>
<td>97</td>
<td>459</td>
<td>.06037</td>
</tr>
<tr>
<td>103</td>
<td>745</td>
<td>.04000</td>
</tr>
<tr>
<td>106</td>
<td>1312</td>
<td>.02407</td>
</tr>
<tr>
<td>108</td>
<td>1613</td>
<td>.01218</td>
</tr>
<tr>
<td>109</td>
<td>1630</td>
<td>.00418</td>
</tr>
</tbody>
</table>
Thus we calculate the expression
\[ \sum_{i=1}^{10} a_i x(n_i) = 238.50937, \]
and consequently we have
\[ \hat{\sigma} = \frac{1}{K_2} \sum_{i=1}^{10} a_i x(n_i) = 242.59466. \]

If we compare the estimate \( \sigma = 241 \) which was calculated using all of the sample by the classical method with the above, there is quite good agreement.

4. Testing Statistical Hypothesis. In order to test the hypothesis

\[ H_0: \sigma = \sigma_0, \quad (4.1) \]

if we start with the frequency function given by (1.1), then this can be done as follows: if the null-hypothesis \( H_0 \) is true, then

\[ \frac{n}{\sigma_0^2} \left\{ \sum_{i=1}^{k} \frac{\lambda_{i+1} - \lambda_{i-1}}{(\lambda_{i+1} - \lambda_1)(\lambda_1 - \lambda_{i-1})} f_i^2(x(n_i) - u_i \sigma_0)^2 \right\} \]

\[ -2 \sum_{i=2}^{k} \frac{f_i f_{i-1}}{\lambda_1 - \lambda_{i-1}} (x(n_i) - u_i \sigma_0)(x(n_{i-1}) - u_{i-1} \sigma_0) \]
is distributed according to the $\chi^2$-distribution with degrees of freedom $k$. This comes out as

$$\frac{n}{\sigma^2} \sum_{i=1}^{k+1} \frac{(f_i^X(n_i)-f_{i-1}^X(n_{i-1}))^2}{\lambda_i - \lambda_{i-1}} + K_2(\sigma^2 - 2\sigma^2_{0})$$  \quad (4.3)

The minimum value $S_0$ of $S$ under the variation of $\sigma$ is given by

$$S_0 = \sum_{i=1}^{k+1} \frac{(f_i^X(n_i)-f_{i-1}^X(n_{i-1}))^2}{\lambda_i - \lambda_{i-1}} - K_2\sigma^2$$  \quad (4.4)

and $\frac{n}{\sigma^2} S_0$ is distributed according to the $\chi^2$-distribution with degrees of freedom $k-1$, provided $H_0$ is true.

Hence

$$\frac{n}{\sigma^2} K_2(\sigma^2 - \sigma_0^2)$$  \quad (4.5)

is independent of $S_0$ and is distributed according to the $\chi^2$-distribution with degree of freedom 1. Thus the statistic

$$t = \sqrt{\frac{K_2(\sigma^2 - \sigma_0^2)}{S_0}}$$  \quad (4.6)

follows the Student's t-distribution with degrees of freedom $k-1$ if the
null-hypothesis $H_0$ is true.

If the null-hypothesis $H_0$ is not true, and an alternative hypothesis $H: \sigma = \sigma_0(\neq \sigma_0)$ is true, then the distribution of $t$ in (4.6) follows the non-central $t$ distribution with non-centrality parameter

$$\delta = \sqrt{\frac{\sigma}{\sigma_0}} (1 - \frac{\sigma_0}{\sigma})$$

(4.7)

and the power function of the $t$-test is an increasing function of the absolute value of $\delta$. Hence it is reasonable to select sample quantiles the spacing of which makes $k_2$ maximum, i.e. optimum spacing for estimation purpose is also optimum for testing purpose. Thus we can use Table 3.1 also for testing purposes.

Finally, it should be noted that the confidence interval of $\sigma$ with confidence coefficient $100(1-\alpha)$ per cent it given by

$$\hat{\sigma} - t_{k-1}(100\alpha) \sqrt{\frac{3}{(k-1)K_2}} < \sigma < \hat{\sigma} + t_{k-1}(100\alpha) \sqrt{\frac{3}{(k-1)K_2}}$$

(4.8)

where $t_{k-1}(100\alpha)$ stands for the $100\alpha$ per cent point of the $t$-distribution with degrees of freedom $k-1$.

Example: We shall calculate the $95\%$ confidence interval for $\sigma$ in the example of the preceding section, as an illustration of the procedure explained in this section. We consider case $k = 10$. Then we have
\[ S_0 = \sum_{i=1}^{11} \frac{(f_i x(n_i) - f_{i-1} x(n_{i-1}))^2}{\lambda_i - \lambda_{i-1}} - K_2 \lambda^2 \]

\[ = \sum_{i=1}^{11} \frac{(f_i x(n_i) - f_{i-1} x(n_{i-1}))^2}{\lambda_i - \lambda_{i-1}} - \frac{10}{K_2} \left( \sum a_i x(n_i) \right)^2 \]

(4.9)

and calculation will be executed as shown in the following table. The 95\% confidence interval calculated below seems to be somewhat wide, which may be due to the smallness of the sample size.
<table>
<thead>
<tr>
<th>$a_0$</th>
<th>$\lambda_0$</th>
<th>$\lambda_1 - \lambda_0$</th>
<th>$f_0$</th>
<th>$x(n_1)$</th>
<th>$f_1^x(n_1)$</th>
<th>38.16900</th>
<th>$\frac{f_1^x(n_1)}{\lambda_1}$</th>
<th>16.13095</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>$\lambda_1$</td>
<td>$\lambda_2 - \lambda_1$</td>
<td>$f_1$</td>
<td>$x(n_2)$</td>
<td>$f_2^x(n_1)$</td>
<td>38.16900</td>
<td>$\frac{f_2^x(n_2)}{\lambda_2 - \lambda_1}$</td>
<td>15.00864</td>
</tr>
<tr>
<td>$a_2$</td>
<td>$\lambda_2$</td>
<td>$\lambda_3 - \lambda_2$</td>
<td>$f_2$</td>
<td>$x(n_3)$</td>
<td>$f_3^x(n_2)$</td>
<td>38.16900</td>
<td>$\frac{f_3^x(n_3)}{\lambda_3 - \lambda_2}$</td>
<td>10.50765</td>
</tr>
<tr>
<td>$a_3$</td>
<td>$\lambda_3$</td>
<td>$\lambda_4 - \lambda_3$</td>
<td>$f_3$</td>
<td>$x(n_4)$</td>
<td>$f_4^x(n_3)$</td>
<td>38.16900</td>
<td>$\frac{f_4^x(n_4)}{\lambda_4 - \lambda_3}$</td>
<td>-2.25064</td>
</tr>
<tr>
<td>$a_4$</td>
<td>$\lambda_4$</td>
<td>$\lambda_5 - \lambda_4$</td>
<td>$f_4$</td>
<td>$x(n_5)$</td>
<td>$f_5^x(n_4)$</td>
<td>38.16900</td>
<td>$\frac{f_5^x(n_5)}{\lambda_5 - \lambda_4}$</td>
<td>-184.09973</td>
</tr>
<tr>
<td>$a_5$</td>
<td>$\lambda_5$</td>
<td>$\lambda_6 - \lambda_5$</td>
<td>$f_5$</td>
<td>$x(n_6)$</td>
<td>$f_6^x(n_5)$</td>
<td>38.16900</td>
<td>$\frac{f_6^x(n_6)}{\lambda_6 - \lambda_5}$</td>
<td>-184.09973</td>
</tr>
</tbody>
</table>
\[ S_0 = \sum_{i=1}^{11} \frac{(f_i^{x(n_i)} - f_{i-1}^{x(n_{i-1})})^2}{\lambda_i - \lambda_{i-1}} \cdot \frac{10}{K_2} \left( \sum a_i^{x(n_i)} \right)^2 \]

\[ = 2600.72348 \]

\[ t_{9(5)} \cdot \sqrt{\frac{S_0}{(k-1)K_2}} = 2.262 \times \sqrt{\frac{2600.72348}{9 \times 0.98316}} = 38.77971 \]
REFERENCES

\[\text{1. Maguire, B. A., Pearson, E. S. and Wynn, A. H. A., "The time}
\text{intervals between industrial accidents," }\text{Biometrika, Vol. 39}
\text{ (1957). pp. 168-180.}\]

\[\text{2. Mosteller, F., "On some useful 'inefficient' statistics," }\text{Annals}

\[\text{3. Neyman, J and David, F. N., "Extension of the Markoff's theorem}
\text{on least squares," }\text{Statistical Research Memoirs, Vol. 1(1936).}
\text{pp. 105-116.}\]

\[\text{4. Ogawa, J., "Contributions to the theory of systematic statistics,}

\[\text{5. Ogawa, J., Determination of optimum spacings for the estimation}
\text{of the scale parameter of the exponential distribution based on}
\text{sample quantiles. Typewritten manuscript at the Department}
\text{of Biostatistics, School of Public Health, University of North}
\text{Carolina, Chapel Hill, North Carolina.}\]
ACKNOWLEDGEMENT

The author wishes to express his thanks to Professor B. G. Greenberg and Dr. A. E. Sarhan of the Department of Biostatistics, School of Public Health, University of North Carolina for their kind encouragements, essential suggestions and discussions given to him while this paper was prepared. The author's thanks are also due to Professor Harold Hotelling and Dr. David B. Duncan of the University of North Carolina for their valuable comments on this paper.