SARA LYNNE STOKES.* An Investigation of the Consequences of Ranked Set Sampling. (Under the direction of NORMAN L. JOHNSON.)

The effects of ranked set sampling on some estimators of location and scale are examined; among these are the MLE's and BLUE's of μ and σ where they are parameters of a random variable X which has c.d.f. of the form \( F\left(\frac{x-\mu}{\sigma}\right) \). Estimators of the correlation coefficient, \( \rho \), of a bivariate normal distribution are introduced, where these estimators are based on a modified form of ranked set sampling in which only extreme order statistics are included in the sample. In both situations, we find that the precision of the estimators is improved by the ranked set sampling process over that from estimators based on the conventional random sample. Two testing and confidence interval procedures for \( \rho \) utilizing ranked set sampling are introduced and compared. Finally, the cost effectiveness of ranked set sampling is briefly discussed.

* This research was supported in part by the Army Research Office, under grant #DAA-G29-74-C-0030.
ACKNOWLEDGEMENTS

I would like to express my appreciation to my advisor, Dr. N.L. Johnson, for his ideas and assistance during the course of this long project. His gentle prodding and encouragement, and his friendship, played important parts in its completion. I also thank the members of my committee, Dr. R. Carroll, Dr. R. Helms, Dr. P.K. Sen and Dr. G. Simons, for their careful reading of the manuscript, and the Department of Statistics for financial support throughout my graduate studies.

I express my gratitude to Ms. June Maxwell for her beautiful and speedy job of typing, and for her ability to take care of details, both of which made the last several weeks proceed much more smoothly.

Special thanks go to my classmates for many reasons, both academic and non-academic; in particular, I would like to thank Chin-Fei Hsu for his generous help in programming, and my friend Carlos Segami for his encouragement from the beginning. And finally, to my family and special friends in Chapel Hill and North Carolina, who have soothed and supported and listened, I express my deepest gratitude.
# TABLE OF CONTENTS

Acknowledgements---------------------------------------------------------- ii

1. INTRODUCTION------------------------------------------------------------- 1

2. ESTIMATION OF LOCATION AND SCALE PARAMETERS FROM RANKED SETS---------- 7
   2.1.1 Efficient Estimators------------------------------------------------ 7
   2.1.2 Fisher Information--------------------------------------------------- 9
   2.2 Best Linear Unbiased Estimation---------------------------------------- 20
   2.3 $s^2$-type Estimators of $\sigma^2$-------------------------------------- 36
   2.4 Judgment Ordering with a Concomitant Variable-------------------------- 46

3. ESTIMATION OF PARAMETERS OF ASSOCIATION-------------------------------- 51
   3.1.1 Estimation of $\rho$ with All Parameters Known---------------------- 51
   3.1.2 Confidence Intervals for $\rho$ with All Parameters Known------------ 57
   3.2 Estimation of $\rho$ with $\mu_y, \sigma_y^2$ Known; $\mu_x, \sigma_x^2$ Unknown 68
   3.3 Estimation of $\rho$ with All Parameters Unknown----------------------- 71

4. HYPOTHESIS TESTING AND RANKED SET SAMPLING------------------------------ 76
   4.1 Tests of Location------------------------------------------------------ 76
   4.2 Tests of Correlation--------------------------------------------------- 82

5. COST CONSIDERATIONS------------------------------------------------------ 98

Appendix------------------------------------------------------------------- 104

Bibliography--------------------------------------------------------------- 111
1. INTRODUCTION

Ranked set sampling was introduced and applied to the problem of estimating pasture yields by McIntyre [16]. It has found applicability in other situations in which observations of a sample are much more easily and inexpensively ordered than quantified. The ranked set sampling process consists of choosing m random samples, each of size m, from the population. The m elements of each sample are ordered among themselves. Then the smallest observation from the first sample is measured accurately, as is the 2nd smallest observation from the 2nd sample. The process continues in this manner until the largest observation from the m-th sample is measured. This entire cycle is repeated n times.

Let $X_{(r:m),i}$ denote the r-th order statistic from the sample of size m, in the i-th cycle. Takahasi and Wakimoto [21] have shown that:

1. $\hat{\mu}_{(m)n} = \frac{1}{mn} \sum_{r=1}^{m} \sum_{i=1}^{n} X_{(r:m),i}$, the arithmetic mean of the mn quantified observations, is an unbiased estimator of the population mean $\mu$.

2. $1 \leq \frac{\text{var}(\overline{X}_{mn})}{\text{var}(\hat{\mu}_{(m)n})} \leq \frac{1}{2(m+1)}$, where $\overline{X}_{mn} = \frac{1}{mn} \sum_{i=1}^{n} X_i$, the conventional estimator of the mean from a random sample of size mn. (Note that both $\overline{X}_{mn}$ and $\hat{\mu}_{(m)n}$ are calculated from mn quantified observations.) The maximum value for $\frac{\text{var}(\overline{X}_{mn})}{\text{var}(\hat{\mu}_{(m)n})}$ of $\frac{1}{2(m+1)}$ is attained only for the rectangular distribution among the class of all continuous distributions with finite variance.

3. If $N = ab = cd$, where a, b, c, d are all positive integers and $a > c$, then
\[ \text{var}(\hat{\mu}_{(c)d}) > \text{var}(\hat{\mu}_{(a)b}). \]

That is, the ranked set sampling estimator based on \( N = nm \) observations is improved by increasing \( m \) and decreasing \( n \) accordingly. Therefore, the best ranked set sampling estimator based on \( N \) quantified observations is \( \hat{\mu}_{(N)1} \). Unfortunately, the ordering process is likely to become difficult, if not impossible, for large \( N \), and the ranked set sampling procedure will become impractical.

Dell and Clutter [6] extend the concept of ranked set sampling to include those situations in which perfectly accurate ordering may not be possible. In this case, they suggest that each of the \( mn \) samples be ordered by eye, and the resulting "judgment order statistics" be used for estimation of the population mean in the same way that the order statistics were used in the original ranked set sampling procedure. That is, let

\[ \hat{\mu}_{[m]n} = \frac{1}{mn} \sum_{r=1}^{m} \sum_{i=1}^{n} X_{[r:m],i} \]

where \( X_{[r:m],i} \) is the \( r \)-th judgment order statistic from a sample of size \( m \) in the \( i \)-th cycle. Then they have shown that for any underlying distribution with the appropriate number of moments, \( \hat{\mu}_{[m]n} \) is an unbiased estimator of \( \mu \) and

\[ \text{var}(\hat{\mu}_{[m]n}) = \frac{1}{mn} \left[ \sigma^2 - \frac{1}{m} \sum_{r=1}^{m} (\hat{\mu}_{[r:m]} - \mu)^2 \right] \leq \frac{\sigma^2}{mn} = \text{var}(\bar{X}_{mn}), \]

where \( \mu_{[r:m]} \) is the mean of the \( r \)-th judgment order statistic from a sample of size \( m \). So

\[ \frac{\text{var}(\bar{X}_{mn})}{\text{var}(\hat{\mu}_{[m]n})} \geq 1, \]

with equality holding only if judgment ordering is no improvement over random ordering.
In addition to McIntyre's original application, the ranked set sampling procedure has been implemented successfully by Halls and Dell [9], who used it for estimating forage yields. Other suggested and actual applications of the ranked set sampling method in which the idea of judgment ordering might be particularly useful and realistic are:

(1) Ordering by observation of aerial photographs for estimating yields of stands of trees.

(2) Ordering by personal interviews for estimating characteristics in human populations.

(3) Ordering by color or appearance for estimating the concentration of a chemical in solution or a pollutant in air.

(4) Ordering by subjective judgment as to the degree of illness of a patient for estimating some measurable characteristic of the progression of his disease.

In order to study the effects of possibly imperfect or judgment ranking on the estimator of the mean, Dell and Clutter propose a model for ranking error. In this model, one assumes that the experimenter ranks the elements from judgment estimates which equal the true value of the observation plus an error component; i.e.,

\[ X_{[r:m]} = X_{(r:m)} + \varepsilon \]

where \( \varepsilon \sim N(0, \sigma_{\varepsilon}^2) \), with \( X_{(r:m)} \) and \( \varepsilon \) independent. They examine the effects of this model on \( \text{var}(\bar{X}_{mn})/\text{var}(\hat{\mu}_{[m]n}) \) by performing computer simulations, varying the underlying distribution of \( X \), the value of \( \sigma_{\varepsilon}^2 \), and the size of \( m \). David and Levine [5] verify the accuracy of some of the computational results by finding the theoretical values of
\[ \frac{\text{var}(\bar{X}_{mn})}{\text{var}(\hat{\mu}_{[m]}n)} \] for an underlying normal distribution.

Dell and Clutter point out that this model for ranking errors may be unrealistic in certain situations. First, judgment errors are likely to be influenced by the size of \( m \), the number to be ranked. Secondly, the independence of the observation and the error component may be a faulty assumption, since the tendency to misrank may be influenced by the specific values of the elements themselves, or specific values of other elements within the sample. This flaw would be difficult to correct, however, since those "specific values" may be peculiar to a given experimental situation.

In Chapter 2, other estimators of the location parameter besides \( \hat{\mu}_{(m)n} \) are examined in order to determine if the benefit derived from ranked set sampling extends to them. In addition, several estimators of dispersion parameters and the effect of ranked set sampling on their variances are studied. In all cases, we show that at least for large enough sample sizes, or for large enough values of \( m \), the ranked set sampling estimator has a smaller variance (or mean square error) than its random counterpart.

In Section 2.4, the concept of judgment ordering is generalized. Suppose we are interested in estimating the mean of \( X \) when not only quantification, but also ordering of the observations of \( X \) is difficult. If the observations of a related variable \( Y \) can be easily ordered, then our goal is to use the information available from the \( Y \)'s to judgment order the random sample of \( X \)'s and therefore to indicate which set of \( X \)'s should be measured. An example of the type of problem where this procedure could prove useful is discussed in [20]. Here an estimate of
cinchona yield in certain plantation areas of India was desired, both in terms of dry bark and quinine contents. The measuring process was an expensive one, however, since it involved killing and stripping the plant, drying and weighing the bark, and chemically analyzing the bark and roots to ascertain quinine yield. But it was also found in the course of the study that the correlation of dry bark yield with volume of bark per plant (a figure easily computed from the height, girth, and bark thickness of a plant) was quite high. So bark volume could have been used as a related variable (Y) for indicating which plants to include in the sample.

In Chapter 3, the ranked set sampling idea is extended to estimation of parameters of association of bivariate data. This technique would be used in situations in which observations of one of the two variables (which we call X) can be ordered easily, while the other, or possibly both, are relatively much more expensive to quantify. Here we have restricted the investigation to the estimation of \( \rho \) where \((X,Y)\) has a bivariate normal distribution. We show that this technique will improve the estimate of \( \rho \) in certain circumstances.

Useful applications of this technique can be found in many fields. Some possibilities are:

1. Estimation of correlation between an external characteristic \(X\) and some internal characteristic (or one that is simply expensive, painful, or inconvenient to measure) of a patient or laboratory animal.

2. Estimation of correlation between socio-economic level or some other observable characteristic, and IQ score, which is tedious and expensive to obtain.

Three cases are considered in this chapter. First we assume that all
parameters except $\rho (\mu_x, \mu_y, \sigma_x^2, \sigma_y^2)$ are known. Secondly, we assume that only $\mu_x$ and $\sigma_x^2$ are known. If $Y$ can be easily measured, it is conceivable that prior information about its mean and variance is available, or at least that they can be quickly and accurately estimated. Finally, we consider the case in which none of the parameters are known.

Chapter 4 describes how ranked set sampling techniques may be put to use in testing of hypotheses. Chapter 5 deals with some considerations of cost in ranked set sampling.

Throughout this work, the emphasis in each analysis is on ranked set sampling when perfect ordering is possible. However, it often happens that the presence of errors in ranking does not negate all its benefits. When that is the case, it will be noted at that point in the discussion.
2. ESTIMATION OF LOCATION AND SCALE PARAMETERS
FROM RANKED SET SAMPLES

2.1.1 Efficient Estimators

Dell and Clutter have considered an estimator of one location parameter, the mean \( \mu \), and have determined in general that it is an unbiased estimator of \( \mu \) (when it exists) and that its variance is smaller than that of the conventional sample mean. In this chapter, other ranked set sampling estimators of location and scale will be examined, with particular attention to their efficiencies with respect to those of comparable estimators based on random samples.

First we will determine the amount of information (in Fisher's sense) available from the ranked set sample about the location and scale parameters. Fisher Information, \( I(\theta) \), is important for several reasons. Among them are the following:

1. Let \( \hat{\theta}_n \), an estimator of \( \theta \), be calculated from a random sample \( X_1, \ldots, X_n \), each \( X_i \) having p.d.f. \( f_\theta(x) \). Then under certain general conditions (see Appendix, Theorem A)

\[
E(\hat{\theta}_n - \theta)^2 \geq \frac{(1 + \frac{db(\theta)}{d\theta})^2}{nI(\theta)}
\]

where \( E\hat{\theta}_n = \theta + b(\theta) \) and

\[
I(\theta) = \int_{-\infty}^{\infty} \left( \frac{\partial \ln f_\theta(x)}{\partial \theta} \right)^2 f_\theta(x) dx.
\]

Consequently, if \( \hat{\theta}_n \) is an unbiased estimator of \( \theta \),

\[
0 \leq \frac{1/n I(\theta)}{\text{var}(\hat{\theta}_n)} = e(\hat{\theta}_n) \leq 1,
\]
and \( e(\hat{\theta}_n) \) will be called the relative efficiency of \( \hat{\theta}_n \). If \( e(\hat{\theta}_n) = 1 \), we say \( \hat{\theta}_n \) is an efficient estimator of \( \theta \); if \( \lim_{n \to \infty} e(\hat{\theta}_n) = 1 \), we say \( \hat{\theta}_n \) is an asymptotically efficient estimator of \( \theta \).

(2) Under fairly general conditions (see Appendix, Theorem B), the maximum likelihood estimator (MLE) of \( \theta \) is an asymptotically efficient estimator.

There are extensions to more general situations of both points (1) and (2) [3]. First consider the case of 2 unknown parameters, \( \theta \) and \( \eta \). Suppose, for simplicity, that \( \hat{\theta}_n \) and \( \hat{\eta}_n \) are unbiased estimators of \( \theta \) and \( \eta \). Now under regularity conditions analogous to those of Theorem A (i.e., the conditions of Theorem A must be satisfied with respect to both parameters \( \theta \) and \( \eta \)), we find that the ellipse

\[
n[E \left( \frac{\partial \ln f}{\partial \theta} \right)^2 (u-\theta)^2 + 2E \left( \frac{\partial \ln f}{\partial \theta} \right) \cdot \left( \frac{\partial \ln f}{\partial \eta} \right) (u-\theta)(v-\eta) + E \left( \frac{\partial \ln f}{\partial \eta} \right)^2 (v-\eta)^2] = 4
\]

is contained within the concentration ellipse of \( (\hat{\theta}_n, \hat{\eta}_n) \), which is

\[
\frac{1}{1-\rho^2} \left( \frac{(u-\theta)^2}{\sigma_1^2} - \frac{2\rho(u-\theta)(v-\eta)}{\sigma_1 \sigma_2} + \frac{(v-\eta)^2}{\sigma_2^2} \right) = 4,
\]

where the covariance matrix of \( (\hat{\theta}_n, \hat{\eta}_n) \) is

\[
\Sigma_n = \begin{bmatrix}
\sigma_1^2 & \rho \sigma_1 \sigma_2 \\
\rho \sigma_1 \sigma_2 & \sigma_2^2
\end{bmatrix}.
\]

If the two ellipses (2.1) and (2.2) coincide, we say that \( \hat{\theta}_n \) and \( \hat{\eta}_n \) are jointly efficient estimators of \( \theta \) and \( \eta \), and

\[
\Sigma_n = \frac{1}{n} I^{-1}(\theta, \eta)
\]
where

\[ I(\theta, \eta) = \begin{bmatrix}
E\left( \frac{\partial \ln f}{\partial \theta} \cdot \frac{\partial \ln f}{\partial \theta} \right) & E\left( \frac{\partial \ln f}{\partial \theta} \cdot \frac{\partial \ln f}{\partial \eta} \right) \\
E\left( \frac{\partial \ln f}{\partial \theta} \cdot \frac{\partial \ln f}{\partial \eta} \right) & E\left( \frac{\partial \ln f}{\partial \eta} \cdot \frac{\partial \ln f}{\partial \eta} \right)
\end{bmatrix} . \]

Furthermore, it is not possible to find a pair of regular unbiased estimators of \((\theta, \eta)\), \(\hat{\theta}_n^*\) and \(\hat{\eta}_n^*\) say, such that \(
\text{var}(\hat{\theta}_n^*) \leq \text{var}(\hat{\theta}_n)\), even if we are unconcerned with \(
\text{var}(\hat{\eta}_n^*)\).

We define the joint efficiency to be

\[ e(\hat{\theta}_n, \hat{\eta}_n) = \left( \frac{1/|I(\theta, \eta)|^{1/2}}{\sum_n} \right) . \]

Then \(0 \leq e(\hat{\theta}_n, \hat{\eta}_n) \leq 1\), and if \(\lim_{n \to \infty} e(\hat{\theta}_n, \hat{\eta}_n) = 1\), we say \((\hat{\theta}_n, \hat{\eta}_n)\) are jointly asymptotically efficient. The MLE's of \(\theta\) and \(\eta\) are jointly asymptotically efficient. These results extend directly to the case of several unknown parameters.

The second important extension is that if \(X_1, \ldots, X_n\) is not a set of iid random variables, but rather has p.d.f.

\[ f(x_1, x_2, \ldots, x_n) \neq f(x_1)f(x_2)\ldots f(x_n) , \]

then results (1) and (2) still hold.

2.1.2 Fisher Information

Let \(X = (X_1, \ldots, X_N)\) be a random sample from a distribution having c.d.f. \(F[(x-\mu)/\sigma]\) and p.d.f. \((1/\sigma)f(x-\mu)/\sigma\). (Note that \(\mu\) is a parameter of location and \(\sigma\) a parameter of scale, although they need not be expectation and standard deviation of \(X\).) Let \(X^* = (X_{1:m}, 1, \ldots, X_{1:m}, n, \ldots, X_{m:m}, 1, \ldots, X_{m:m}, n)\) be a ranked set sample of size \(mn = N\) from the same distribution.

So the likelihood function of \(X\) is
\[ L = \frac{1}{\sigma^N} \prod_{i=1}^{N} f \left( \frac{X_i - \mu}{\sigma} \right) \]

and the likelihood function of \( X^* \) is

\[ L^* = \frac{1}{\sigma^N} \prod_{i=1}^{n} \prod_{r=1}^{m} f \left( \frac{X_{(r:\text{m}),i} - \mu}{\sigma} \right) \]

\[ = \frac{1}{\sigma^N} \prod_{i=1}^{n} \prod_{r=1}^{m} \left( \frac{1}{r-1} \right)^{m-1} f \left( \frac{X_{(r:\text{m}),i} - \mu}{\sigma} \right)^{r-1} \left[ 1 - f \left( \frac{X_{(r:\text{m}),i} - \mu}{\sigma} \right) \right]^{m-r} \]

The Fisher Information of a random sample for such a distribution is easily found.

(1) If the scale parameter \( \sigma \) is known, then

\[ I_N(\mu) = NI(\mu) \quad \text{providing} \quad E \frac{\partial \ln f}{\partial \mu} = 0 \]

\[ = N E \left[ \frac{1}{\sigma} f' \left( \frac{X - \mu}{\sigma} \right) \right]^2 \]

\[ = \frac{N}{\sigma^2} E \left[ \frac{f'(U)}{f(U)} \right]^2 , \]

where \( U = \frac{X - \mu}{\sigma} \).

(2) If the location parameter \( \mu \) is known,

\[ I_N(\sigma) = NI(\sigma) \quad \text{providing that} \quad E \frac{\partial \ln f}{\partial \sigma} = 0 \]

\[ = N E \left[ \frac{1}{\sigma} f \left( \frac{X - \mu}{\sigma} \right) \right]^2 \]

\[ = \frac{N}{\sigma^2} E \left[ U \left( \frac{f'(U)}{f(U)} \right)^2 + 1 \right] \]
\[
= \frac{N}{\sigma^2} \mathbb{E} \left[ \left( \frac{U f'(U)}{f(U)} \right)^2 + 2 \frac{U f'(U)}{f(U)} + 1 \right]
\]

\[
= \frac{N}{\sigma^2} \mathbb{E} \left[ \left( \frac{U f'(U)}{f(U)} \right)^2 - 1 \right]
\]
since
\[
\mathbb{E} \frac{U f'(U)}{f(U)} = -1.
\]

(3) If neither \( \mu \) nor \( \sigma \) is known, then we find the Fisher Information matrix \( I_N(\mu, \sigma) \).

\[
\mathbb{E} \left[ \frac{\partial^2 \ln L}{\partial \mu \partial \sigma} \cdot \frac{\partial \ln L}{\partial \sigma} \right]
\]

\[
= \frac{N}{\sigma^2} \mathbb{E} \left[ \frac{\partial \ln f(X-\mu)}{\partial \mu} \cdot \frac{\partial \ln f(X-\mu)}{\partial \sigma} \right]
\]

if \( \mathbb{E} \frac{\partial \ln f}{\partial \sigma} = \frac{\partial \ln f}{\partial \mu} = 0 \)

\[
= \frac{N}{\sigma^2} \mathbb{E} \left[ \left( \frac{f'(U)}{f(U)} \right)^2 + \frac{f'(U)}{f(U)} \right]
\]

\[
= \frac{N}{\sigma^2} \mathbb{E} \left[ \left( \frac{f'(U)}{f(U)} \right)^2 \right] \quad \text{since} \quad \mathbb{E} \frac{f'(U)}{f(U)} = 0.
\]

Then

\[
I_N(\mu, \sigma) = \frac{N}{\sigma^2} \begin{bmatrix}
\mathbb{E} \left[ \left( \frac{f'(U)}{f(U)} \right)^2 \right] & \mathbb{E} \left[ \frac{f'(U)}{f(U)} \right] \\
\mathbb{E} \left[ \frac{f'(U)}{f(U)} \right] & \mathbb{E} \left[ \left( \frac{f'(U)}{f(U)} \right)^2 - 1 \right]
\end{bmatrix}
\]. \quad (2.3)

Now the same three cases will be considered for a ranked set sample.

(1*) Assume \( \sigma \) known and \( \frac{\partial}{\partial \mu} \int_{-\infty}^{\infty} f(X-\mu)dx = \int_{-\infty}^{\infty} \frac{\partial}{\partial \mu} f(X-\mu)dx \). Then
\[ I_{(m)n}(\mu) = E \left[ \frac{\partial \ln L^*}{\partial \mu} \right]^2 = -E \left[ \frac{\partial^2 \ln L^*}{\partial \mu^2} \right] \]

and

\[ \ln L^* = \text{terms not involving } \mu + \sum_{i=1}^{n} \sum_{r=1}^{m} \ln f \left( \frac{X_{(r:m),i} - \mu}{\sigma} \right) \]

\[ + \sum_{i=1}^{n} \sum_{r=1}^{m} (r-1) \ln F \left( \frac{X_{(r:m),i} - \mu}{\sigma} \right) \]

\[ + \sum_{i=1}^{n} \sum_{r=1}^{m} (m-r) \ln \left[ 1 - F \left( \frac{X_{(r:m),i} - \mu}{\sigma} \right) \right]. \quad (2.4) \]

Thus

\[ \frac{\partial}{\partial \mu} \ln L^* = \sum_{i=1}^{n} \sum_{r=1}^{m} \frac{f' \left( \frac{X_{(r:m),i} - \mu}{\sigma} \right)}{f \left( \frac{X_{(r:m),i} - \mu}{\sigma} \right)} \left( \frac{1}{\sigma} \right) \]

\[ + \sum_{i=1}^{n} \sum_{r=1}^{m} (r-1) \frac{f' \left( \frac{X_{(r:m),i} - \mu}{\sigma} \right)}{F \left( \frac{X_{(r:m),i} - \mu}{\sigma} \right)} \left( \frac{1}{\sigma} \right) \]

\[ + \sum_{i=1}^{n} \sum_{r=1}^{m} (m-r) \frac{f' \left( \frac{X_{(r:m),i} - \mu}{\sigma} \right)}{1 - F \left( \frac{X_{(r:m),i} - \mu}{\sigma} \right)} \left( \frac{1}{\sigma} \right) \]

and

\[ \frac{\partial^2}{\partial \mu^2} \ln L^* = \frac{1}{\sigma^2} \sum_{i} \sum_{r} \left\{ \frac{f''(U_{(r:m),i})}{f(U_{(r:m),i})} - \left[ \frac{f'(U_{(r:m),i})}{f(U_{(r:m),i})} \right]^2 \right\} \]

\[ + \frac{1}{\sigma^2} \sum_{i} \sum_{r} (r-1) \left\{ \frac{f'(U_{(r:m),i})}{F(U_{(r:m),i})} - \left[ \frac{f(U_{(r:m),i})}{F(U_{(r:m),i})} \right]^2 \right\} \]

\[ - \frac{1}{\sigma^2} \sum_{i} \sum_{r} (m-r) \left\{ \frac{f'(U_{(r:m),i})}{1-F(U_{(r:m),i})} + \left[ \frac{f(U_{(r:m),i})}{1-F(U_{(r:m),i})} \right]^2 \right\}. \]
So
\[ -E \left[ \frac{\partial^2 \kappa_0}{\partial \mu^2} \right] \]
\[ = \frac{n}{\sigma^2} \sum_{r=1}^{m} \left\{ \left( -\frac{\kappa_0}{\mu} \right)^2 - \frac{\kappa_0^{(r+1)}}{\mu} \right\} \]
\[ + \frac{n}{\sigma^2} \sum_{r=1}^{m} (r-1)E \left\{ \left( -\frac{\kappa_0}{\mu} \right)^2 - \frac{\kappa_0^{(r-1)}}{\mu} \right\} \]
\[ + \frac{n}{\sigma^2} \sum_{r=1}^{m} (m-r)E \left\{ \left( -\frac{\kappa_0}{\mu} \right)^2 + \frac{\kappa_0^{(r+1)}}{1-F(\mu)} \right\}. \]

(2.5)

Letting \( h(u) = \left[ \frac{f'(u)}{f(u)} \right]^2 - \frac{f''(u)}{f(u)} \), we see that the first term of the right hand side of (2.5) is
\[ \frac{n}{\sigma^2} \sum_{r=1}^{m} \int_{-\infty}^{\infty} h(u) \kappa_0^{(r-1)} f(u) F^{r-1} \left[ 1 - F(u) \right]^n du \]
\[ = \frac{n}{\sigma^2} \int_{-\infty}^{\infty} m h(u) f(u) \sum_{r=1}^{m} \kappa_0^{(r-1)} F^{r-1} \left[ 1 - F(u) \right]_{m-r} du \]
\[ = \frac{nm}{\sigma^2} \int_{-\infty}^{\infty} h(u) f(u) du = \frac{nm}{\sigma^2} E \left[ \left( \frac{f'(u)}{F(u)} \right)^2 - \frac{f''(u)}{f(u)} \right]. \]

Now redefine \( h(u) = \left[ \frac{f(u)}{F(u)} \right]^2 - \frac{f'(u)}{F(u)} \). The second term of the right hand side of (2.5) is
\[ \frac{n}{\sigma^2} \sum_{r=1}^{m} \int_{-\infty}^{\infty} (r-1) h(u) \kappa_0^{(r-1)} f(u) F^{r-1} \left[ 1 - F(u) \right]^{m-r} du \]
\[ = \frac{nm}{\sigma^2} \int_{-\infty}^{\infty} h(u) f(u) \sum_{r=1}^{m} \kappa_0^{(r-1)} F^{r-1} \left[ 1 - F(u) \right]^{m-r} du \]
\[ = \frac{nm}{\sigma^2} \int_{-\infty}^{\infty} h(u) f(u) F(u) du = \frac{nm}{\sigma^2} E \left\{ \frac{F^2(u)}{F(u)} - f'(u) \right\}. \]
Similarly, we can show that the third term of the right hand side of (2.5) is
\[
\frac{nm(m-1)}{\sigma^2} E \left\{ \frac{f^2(U)}{1-F(U)} + f'(U) \right\}.
\]

Then, writing \( f(U) \equiv f \) and \( F(U) \equiv F \), we have
\[
I_\text{(m)n}'(u) = \frac{nm}{\sigma^2} E \left[ \frac{f'(U)}{F(U)} \right]^2 + \frac{nm(m-1)}{\sigma^2} E \frac{f^2(U)}{F(U)[1-F(U)]}
\]
\[
= \frac{N}{\sigma^2} E \left( \frac{f'}{f} \right)^2 + \frac{N(m-1)}{\sigma^2} E \left( \frac{f}{F(1-F)} \right).
\]
Since \( E \frac{f^2}{F(1-F)} \geq 0 \), we see that \( I_\text{(m)n}'(u) \geq I_N(u) \). This tells us that in any regular estimation case, an efficient unbiased estimator of \( \mu \) calculated from a ranked set sample has a smaller variance (by an order of \( m \)) than an efficient unbiased estimator of \( \mu \) calculated from a random sample.

(2*) Assume \( \mu \) known. By the same method as that employed in (1*), we have
\[
I_\text{(m)n}^*(\sigma) = \frac{nm}{\sigma^2} E \left\{ \left[ \frac{Uf'(U)}{f(U)} \right]^2 - 1 \right\} + \frac{nm(m-1)}{\sigma^2} E \left\{ \frac{U^2f^2(U)}{F(U)[1-F(U)]} \right\}
\]
if the p.d.f. of \( U \) has the properties
\[
\lim_{u \to \infty} uf(u) = \lim_{u \to -\infty} uf(u) = \lim_{u \to \infty} u^2f(u) = \lim_{u \to -\infty} u^2f(u) = 0.
\]

Here too, we have
\[
I_\text{(m)n}^*(\sigma) \geq I_N(\sigma).
\]
Therefore, an efficient unbiased estimator of \( \sigma \) based on a ranked set sample is "better" in the sense described in (1*) than its counterpart.
from a random sample.

(3*) Assume $\mu$ and $\sigma$ unknown and the conditions of (2*) hold. Then the upper triangle of the matrix

\[
\mathbb{I}_{(m,n)}^*(\mu, \sigma) = \frac{N}{2} \mathbb{I} \left[ \begin{array}{c}
E(\frac{f'}{f})^2 + (m-1)E(\frac{f^2}{F(1-F)}) - E(\frac{f'}{f})^2 \right] \\
E(\frac{f'}{f})^2 - 1] + (m-1)E(\frac{Uf^2}{F(1-F)})
\end{array} \right]
\]

Let $(\hat{\mu}, \hat{\sigma})$ be a pair of jointly efficient unbiased estimators of $(\mu, \sigma)$ calculated from a random sample, and let $(\hat{\mu}^*, \hat{\sigma}^*)$ be jointly efficient unbiased estimators calculated from a ranked set sample in a regular estimation situation. Then

\[
\text{var}(\hat{\mu}^*) = \frac{\sigma^2}{N(m-1)} \left\{ \frac{1}{m-1} E(\frac{f'}{f})^2 - 1 \right\} + \frac{E(\frac{Uf^2}{F(1-F)})}{A^*},
\]

\[
\text{var}(\hat{\sigma}^*) = \frac{\sigma^2}{N(m-1)} \left\{ \frac{1}{m-1} E(\frac{f'}{f})^2 + E(\frac{f^2}{F(1-F)}) \right\}/A^*,
\]

where

\[
A^* = \left\{ \frac{1}{m-1} E(\frac{f'}{f})^2 + E(\frac{f^2}{F(1-F)}) \right\} \left\{ \frac{1}{m-1} E(\frac{f'}{f})^2 \right\} - \left\{ \frac{1}{m-1} E(\frac{f'}{f})^2 U + E(\frac{Uf^2}{F(1-F)}) \right\}^2
\]

and

\[
\text{var}(\hat{\mu}) = \frac{\sigma^2}{N} E(\frac{f'}{f})^2 /A
\]

\[
\text{var}(\hat{\sigma}) = \frac{\sigma^2}{N} E(\frac{f'}{f})^2 /A,
\]

where

\[
A = E(\frac{f'}{f})^2 \left\{ E(\frac{f'}{f})^2 \right\} - \left\{ E(\frac{f'}{f})^2 U \right\}^2
\]

Thus

\[
\lim_{m \to \infty} N(m-1) \text{var}(\hat{\mu}^*) = \sigma^2 E \frac{(Uf^2}{F(1-F)^2} / \left( E \frac{f^2}{F(1-F)} E \frac{(Uf^2}{F(1-F)} - \left[ E \frac{Uf^2}{F(1-F)} \right]^2
\]

\[
\lim_{m \to \infty} N(m-1) \text{var}(\hat{\sigma}^*) = \sigma^2 E \frac{(Uf^2}{F(1-F)^2} / \left( E \frac{f^2}{F(1-F)} E \frac{(Uf^2}{F(1-F)} - \left[ E \frac{Uf^2}{F(1-F)} \right]^2
\]

\[
\lim_{m \to \infty} N(m-1) \text{var}(\hat{\mu}) = \sigma^2 E \frac{(Uf^2}{F(1-F)^2} / \left( E \frac{f^2}{F(1-F)} E \frac{(Uf^2}{F(1-F)} - \left[ E \frac{Uf^2}{F(1-F)} \right]^2
\]

\[
\lim_{m \to \infty} N(m-1) \text{var}(\hat{\sigma}) = \sigma^2 E \frac{(Uf^2}{F(1-F)^2} / \left( E \frac{f^2}{F(1-F)} E \frac{(Uf^2}{F(1-F)} - \left[ E \frac{Uf^2}{F(1-F)} \right]^2
\]
and
\[ \lim_{m \to \infty} N(m-1) \text{var}(\hat{\sigma}^*) = \sigma^2 E \frac{\tilde{f}^2}{F(1-F)} / (E \frac{\tilde{f}^2}{F(1-F)} E \frac{(Uf)^2}{F(1-F)} - [E \frac{Uf^2}{F(1-F)}]^2). \]

But
\[ \lim_{m \to \infty} N(m-1) \text{var}(\hat{\mu}) = \lim_{m \to \infty} N(m-1) \text{var}(\hat{\sigma}) = \infty. \]

Therefore, if \( m \) is large enough,
\[ \text{var}(\hat{\mu}^*) \leq \text{var}(\hat{\mu}) \]
and
\[ \text{var}(\hat{\sigma}^*) \leq \text{var}(\hat{\sigma}). \]

If the random variable \( X \) is symmetric around \( \mu \), then \( E \frac{f}{E} \frac{U}{F(1-F)} = 0 \), and
\[ E \frac{Uf^2}{F(1-F)} = 0, \]
and
\[ I_{(m)n}(\mu, \sigma) = \frac{\sigma^2}{N} \begin{bmatrix} 1/I^*(\mu) & 0 \\ 0 & 1/I^*(\sigma) \end{bmatrix}, \]
\[ I_{N}(\mu, \sigma) = \frac{\sigma^2}{N} \begin{bmatrix} 1/I(\mu) & 0 \\ 0 & 1/I(\sigma) \end{bmatrix}. \]

Hence in that case, \((\hat{\mu}^*, \hat{\sigma}^*)\) has an ellipse of concentration which lies completely within that of \((\hat{\mu}, \hat{\sigma})\) for any \( m \); i.e.,
\[ \text{var}(\hat{\mu}^*) \leq \text{var}(\hat{\mu}) \]
and
\[ \text{var}(\hat{\sigma}^*) \leq \text{var}(\hat{\sigma}) \quad \forall m. \]

Now we will look at the difference in the appearance of the two MLE's of the mean \( \mu \) of a normal distribution. This example is not meant to suggest that the estimation procedure described here is an appropriate one for this situation, but rather is included as an illustration of the modification in the MLE that is caused by the
independence of the order statistics. If the random variable is
known to have a normal distribution, one can obtain more information
about \( \mu \), of course, by selecting the median from each sample of size
\( m \) rather than choosing each order statistic in turn. In fact,
whenever the distribution is known, the estimator might be improved by
selecting a particular order statistic (or a particular set of order
statistics) instead of including \( n \) of each in the ranked set sample.
This altered form of sampling will be called modified ranked set
sampling, and we will see examples of its use later.

Let \( X_1, \ldots, X_N \) be iid random variables and \( X_i \sim N(\mu, \sigma^2) \), \( \sigma^2 \) known.
Then the MLE of \( \mu \) is \( \hat{\mu} = \frac{1}{N} \sum X_i \).

Now let \( X_{(1:m)}, \ldots, X_{(m:m)}, n \), \( n = N \), be a ranked set sample from
the same distribution as above. The MLE \( \hat{\mu}^* \) is found by solving the
equation

\[
\frac{\partial \ln L^*}{\partial \mu} = 0
\]

with \( \ln L^* \) as in (2.4), where now \( f(z) = \phi(z) \), the normal p.d.f. and
\( F(z) = \Phi(z) \), the normal c.d.f. This equation reduces to

\[
\sum_{i=1}^{n} \sum_{r=1}^{m} Z_{(r:m),i} = \sum_{i=1}^{n} \sum_{r=1}^{m} (r-1) g(Z_{(r:m),i}) + \sum_{i=1}^{n} \sum_{r=1}^{m} (m-r) g'(Z_{(r:m),i})
\]

\[
= 0 \tag{2.6}
\]

where

\[
Z_{(r:m),i} = \frac{X_{(r:m),i} - \mu}{\sigma}, \quad g(z) = \frac{\phi(z)}{\Phi(z)}, \quad g'(z) = \frac{\phi(z)}{1-\Phi(z)}.
\]

A solution for this equation may be found computationally [10].

A method for finding an approximate solution to (2.6) is adapted
from Tiku [22]. This solution is based on his approximation

\[
g(z) = \frac{\phi(z)}{\Phi(z)} \approx \alpha + \beta z \tag{2.7}
\]
where
\[
\beta = \frac{g(h) - g(\ell)}{h - \ell}, \quad \alpha = g(h) - h\beta,
\]
and \((\ell, h)\) is an interval containing \(z\). Using this approximation, we find that \(\hat{\mu}^* \approx \tilde{\mu}^*\), where \(\tilde{\mu}^*\) is a solution of the equation

\[
\sum_{i=1}^{n} \sum_{r=1}^{m} Z(r:m),i - \sum_{i=1}^{n} \sum_{r=1}^{m} (r-1)(\alpha_{ri} + \beta_{ri}Z(r:m),i) \\
+ \sum_{i=1}^{n} \sum_{r=1}^{m} (m-r)(\alpha'_{ri} + \beta'_{ri}Z(r:m),i) = 0
\]

where
\[
g(Z(r:m),i) \equiv \alpha_{ri} + \beta_{ri}Z(r:m),i, \quad g'(Z(r:m),i) = g(-Z(r:m),i) \\
= \alpha'_{ri} + \beta'_{ri}Z(r:m),i,
\]

with \(\alpha_{ri}, \beta_{ri}, \alpha'_{ri}, \beta'_{ri}\) as defined by (2.7). Thus

\[
\tilde{\mu}^* = \frac{\sum_{i=1}^{n} \sum_{r=1}^{m} X(r:m),i + \sigma \sum_{i=1}^{n} \sum_{r=1}^{m} [(1-r)\alpha_{ri} + (m-r)\alpha'_{ri}] + \sum_{i=1}^{n} \sum_{r=1}^{m} \left\{ [(1-r)\beta_{ri} + (m-r)\beta'_{ri}] \right\} }{mn + \sum_{i=1}^{n} \sum_{r=1}^{m} [(1-r)\beta_{ri} + (m-r)\beta'_{ri}]} \quad (2.8)
\]

The problem here is that the exact position of each \(Z(r:m),i\) is unknown, so that the interval \((x_{ri}, h_{ri})\) which covers \(Z(r:m),i\) must be estimated \(\forall r, i\).

Let \(\mu(r:m), \sigma^2(r:m)\) be the mean and variance of the \(r\)-th order statistic from a unit normal distribution, let \(z_p\) be such that
\[
\Phi(z_p) = p, \text{ and let } \bar{X} = \frac{1}{mn} \sum_{i=1}^{n} \sum_{r=1}^{m} X(r:m),i.
\]
Then for \(m\) fixed and \(n\) large,
\[
\bar{X} \sim N\left(\mu, \sqrt{\frac{1}{m^2 n} \sum_{r=1}^{m} \sigma^2(r:m)}\right).
\]

Thus for some small \(\alpha\), \((x_{ri}, h_{ri})\) is likely to cover \(Z(r:m),i = (X(r:m),i - \mu)/\sigma\) when
\[ \ell_{ri} = \frac{X^{(r:m),i} - \bar{X}}{\sigma} - z_{1-\alpha/2} \sqrt{\frac{1}{m} \sum_{n=1}^{m} \sigma^2_{r(m)}} \]

and

\[ h_{ri} = \frac{X^{(r:m),i} - \bar{X}}{\sigma} + z_{1-\alpha/2} \sqrt{\frac{1}{m} \sum_{n=1}^{m} \sigma^2_{r(m)}} . \]

Then \( \beta_{ri}', \alpha_{ri} \) may be computed by substituting \( \ell_{ri}', h_{ri} \) into (2.7), and \( \tilde{\mu}^* \) may subsequently be calculated from (2.8).

Alternatively, if \( m \) is large (a condition which is not as likely to occur because of practical restrictions, as we have indicated before, but which produces a nice form for \( \tilde{\mu}^* \)), then we choose

\[ \ell_{ri} = \mu^{(r:m)} - \kappa \sigma^{(r:m)} \]

\[ h_{ri} = \mu^{(r:m)} + \kappa \sigma^{(r:m)}, \quad \kappa \text{ a constant}. \]

This interval is likely to cover \( Z_{r(m),i} \), \( \forall i \), if \( \kappa \) is chosen large enough. Unfortunately, however, unless \( m \) is sufficiently large (causing \( \sigma^{(r:m)} \) to be small), the interval will be too wide to allow \( g(z) \) to be approximated over it accurately by a linear function. For this choice of \( \ell_{ri}', h_{ri} \), we have \( \ell_{ri} = -h_{(m-r+1)i} \) and \( h_{ri} = -\ell_{(m-r+1)i} \), \( \forall i \).

Therefore

\[ \beta_{(m-r+1:m),i}' = \frac{g(-\ell_{(m-r+1:m)}) - g(-h_{(m-r+1:m)})}{-\ell_{(m-r+1:m)} + h_{(m-r+1:m)}} \]

\[ = \frac{g(h_{(r:m)}) - g(\ell_{(r:m)})}{h_{(r:m)} - \ell_{(r:m)}} = \beta_{(r:m),i} \]

and

\[ \alpha_{(m-r+1:m),i}' = g(-\ell_{(m-r+1:m)}) + \ell_{(m-r+1:m)} \beta_{(m-r+1:m)} \]

\[ = g(h_{(r:m)}) - h_{(r:m)} \beta_{(r:m)} = \alpha_{(r:m),i} \quad \forall i , \]
since $g'(z) = g(-z)$. Now suppose that $m$ is even, $m = 2k$. Then

$$\sum_{r=1}^{m} [(1-r)\alpha_{ri} + (m-r)\alpha'_{ri}] = \sum_{r=1}^{k} (1-r)\alpha_{ri} + \sum_{r=k+1}^{m} (1-r)\alpha_{ri}$$

$$+ \sum_{r=1}^{k} (m-r)\alpha'_{ri} + \sum_{r=k+1}^{m} (m-r)\alpha'_{ri}$$

$$= \sum_{r=1}^{k} (1-r)\alpha_{ri} + \sum_{r=1}^{k} (r-m)\alpha_{m-r+1,i} + \sum_{r=1}^{k} (m-r)\alpha_{m-r+1,i}$$

$$+ \sum_{r=1}^{k} (r-1)\alpha'_{ri} = 0.$$ 

Similarly, $\sum_{r=1}^{m} [(1-r)\beta_{ri} + (m-r)\beta'_{ri}] = 0$. Therefore, from (2.8),

$$\hat{\mu}^* = \frac{1}{mn} \sum_{i=1}^{n} \sum_{r=1}^{m} X_{r:m},i + \frac{1}{mn} \sum_{r=1}^{m} \{[(1-r)\beta_{r} + (m-r)\beta'_{r}] \sum_{i=1}^{n} X_{r:m},i\}$$

$$= \bar{X} + \frac{1}{mn} \left[ \sum_{r=1}^{k} (1-r)\beta_{r} + (m-r)\beta_{m-r+1,i} \right] \left( \sum_{i=1}^{n} X_{m-r+1:m},i - X_{r:m},i \right)$$

where

$$\beta_{r} = \frac{g(\mu_{r:m} + \kappa \sigma_{r:m}) - g(\mu_{r:m} - \kappa \sigma_{r:m})}{2\kappa \sigma_{r:m}}.$$ 

If $m$ is odd, $m = 2k+1$, then $\hat{\mu}^*$ is still given by (2.9).

2.2 Best Linear Unbiased Estimation

Let $X$ be a random variable having c.d.f. $F\left(\frac{X-\mu}{\sigma}\right)$ where $F$ is known. (Note again here that $\mu$ is not necessarily the mean; $\sigma$ is not necessarily the standard deviation.) The random variable $U = \frac{X-\mu}{\sigma}$ has a parameter-free distribution, and $(X_{r:m} - \mu)/\sigma$ is distributed as $U_{(r:m)}$, the $r$-th order statistic from a sample of size $m$ of a distribution
having c.d.f. \( F(u) \). The means, variances, and covariances of the set of random variables \( U_{(1:m)}, \ldots, U_{(m:m)} \) can be calculated and will be denoted here as:

\[
\begin{align*}
E U_{(r:m)} &= \alpha_{(r:m)} \\
Var U_{(r:m)} &= \sigma^2_{(r:m)} \\
Cov(U_{(r:m)}, U_{(s:m)}) &= \sigma_{(r:s:m)}.
\end{align*}
\]

Therefore the actual ordered observations \( X_{(1:m)}, \ldots, X_{(m:m)} \) have expectations, variances, and covariances:

\[
\begin{align*}
E X_{(r:m)} &= \mu + \alpha_{(r:m)} \sigma \\
Var X_{(r:m)} &= \sigma^2 \sigma^2_{(r:m)} \\
Cov (X_{(r:m)}, X_{(s:m)}) &= \sigma^2 \sigma_{(r:s:m)}
\end{align*}
\]

Lloyd [15] applied the generalized least squares theorem to these ordered random variables and developed the best linear unbiased estimators (BLUE's) of \( \mu \) and \( \sigma \).

**Generalized Least Squares Theorem.** If \( E \overline{X} = A \theta \) and \( Var \overline{X} = \omega^2 V \) (\( A \) and \( V \) known), then the least squares estimate of \( \theta \) is

\[
\hat{\theta} = (A'V^{-1}A)^{-1}A'V^{-1} \overline{X}
\]

and

\[
Var \hat{\theta} = \omega^2 (A'V^{-1}A)^{-1}.
\]

\( \hat{\theta} \) is also the BLUE.

In Lloyd's application

\[
A = \begin{bmatrix}
1 & \alpha_{(1:N)} \\
1 & \alpha_{(2:N)} \\
\vdots & \vdots \\
1 & \alpha_{(N:N)}
\end{bmatrix} = [1 : \alpha],
\]

the upper triangle of
\[
V = \begin{bmatrix}
\sigma^2_{(1:N)} & \sigma_{(12:N)} & \cdots & \sigma_{(1N:N)} \\
\sigma^2_{(2:N)} & \cdots & \sigma_{(2N:N)} \\
\vdots & \ddots & \ddots \\
\sigma^2_{(N:N)} & \cdots & \sigma^2_{(NN:N)}
\end{bmatrix}
\]

\[\omega^2 = \sigma^2, \text{ and } \underline{\theta} = [\frac{\mu}{\sigma}].\]

Thus the estimator \(\hat{\theta} = [\frac{\hat{\mu}}{\hat{\sigma}}]\) has variance-covariance matrix

\[
\text{var} \hat{\theta} = \frac{\sigma^2}{\Delta} \begin{bmatrix}
\alpha'V^{-1}\alpha & -1'V^{-1}\alpha \\
-1'V^{-1}\alpha & 1'V^{-1}1
\end{bmatrix}
\]

where \(\Delta = (1'V^{-1}1)(\alpha'V^{-1}\alpha) = (1'V^{-1}1)^2\).

For the ranked set sampling case, we again choose \(mn = N\) samples of size \(m\) each, and from each sample, we select one order statistic for quantification. The quantified elements are therefore independent and

\[
\text{var} \underline{X} = \sigma^2V^* \text{ where } V^* = \text{Diag}(\sigma^2_{(1:m)}, \sigma^2_{(2:m)}, \ldots, \sigma^2_{(m:m)}), \ldots,
\]

Then the BLUE of \(\underline{\theta} = [\frac{\mu}{\sigma}]\) is

\[
\hat{\theta}^* = (A'V^*^{-1}A)^{-1}A'V^*^{-1}\underline{X}
\]

and

\[
\text{Var} \hat{\theta}^* = \frac{\sigma^2}{n\Delta^*} \begin{bmatrix}
\sum_{r=1}^{m} \frac{\alpha^2_{(r:m)}}{\sigma^2_{(r:m)}} & -\sum_{r=1}^{m} \frac{\alpha_{(r:m)}}{\sigma^2_{(r:m)}} \\
\sum_{r=1}^{m} \frac{\alpha_{(r:m)}}{\sigma^2_{(r:m)}} & \sum_{r=1}^{m} \frac{\alpha^2_{(r:m)}}{\sigma^2_{(r:m)}}
\end{bmatrix}
\]

where
\[ \Delta^* = \left( \sum_{r=1}^{m} 1/\sigma^2_{(r:m)} \right) \left( \sum_{r=1}^{m} \alpha^2_{(r:m)}/\sigma^2_{(r:m)} \right) - \left( \sum_{r=1}^{m} \alpha_{(r:m)}/\sigma^2_{(r:m)} \right)^2. \]

If the only parameter to be estimated is \( \mu \), and the c.d.f. is of the form \( F(x-\mu) \), then one can show that

\[ \text{Var}(\hat{\mu}) = \sigma^2/\bar{V}^{-1} \]

and

\[ \text{Var}(\hat{\mu}^*) = \sigma^2/\bar{V}^{*^{-1}} = \sigma^2/(n \sum_{r=1}^{m} 1/\sigma^2_{(r:m)}) \]

If the only parameter to be estimated is \( \sigma \), and the c.d.f. is of the form \( F(x/\sigma) \), then one can show that

\[ \text{Var}(\hat{\sigma}) = \sigma^2/\bar{V}^{-1} \]

and

\[ \text{Var}(\hat{\sigma}^*) = \sigma^2/\bar{V}^{*^{-1}} = \sigma^2/(n \sum_{r=1}^{m} \alpha^2_{(r:m)}) \]

If \( X \) is a random variable with a distribution symmetrical around \( \mu \), then \( U \) has a distribution symmetrical around the origin and

\[ \alpha_{(r:m)} = -\alpha_{(m-r+1:m)} , \]

\[ \sigma^2_{(r:m)} = \sigma^2_{(m-r+1:m)} . \]

Then

\[ \text{Var}(\hat{\mu}) = \sigma^2/\bar{V}^{-1} , \quad \text{Var}(\hat{\sigma}) = \sigma^2/\bar{V}^{-1} \]

\[ \text{Cov}(\hat{\mu}, \hat{\sigma}) = 0 ; \]

\[ \text{Var}(\hat{\mu}^*) = \sigma^2/(n \sum_{r=1}^{m} \frac{1}{\sigma^2_{(r:m)}}) , \quad \text{Var}(\hat{\sigma}^*) = \sigma^2/(n \sum_{r=1}^{m} \frac{\alpha^2_{(r:m)}}{\sigma^2_{(r:m)}}) , \]

\[ \text{Cov}(\hat{\mu}^*, \hat{\sigma}^*) = 0 . \] (2.10)

Blom [2] and Jung [12] have suggested alternative linear estimators of \( \mu \) and \( \sigma \) which are more easily computed than \( \hat{\mu} \) and \( \hat{\sigma} \), since they don't
require inversion of an \( N \times N \) matrix. (Blom calls his estimator the unbiased nearly best linear estimator.) Both have shown that their estimators are asymptotically efficient when certain restrictions on the p.d.f. of \( U \) are fulfilled. Therefore, the BLUE of \( \mu \) or \( \sigma \), whose variance can be no larger than that of the nearly best estimator, must also be asymptotically efficient.

Ogawa [18] used the asymptotic distribution of sample quantiles to find large sample BLUE's of \( \mu \) and \( \sigma \) based on selected quantiles from a sample of size \( N \). Their asymptotic distribution is given as follows by Mosteller [171].

**Theorem.** For \( k \) given real numbers

\[ 0 < \lambda_1 < \ldots < \lambda_k < 1, \]

let the \( \lambda_i \)-quantile of the population be \( x_i \); i.e.

\[
\int_{-\infty}^{x_i} g(t) \, dt = \lambda_i, \quad i=1,\ldots,k.
\]

Further, assume that the p.d.f. of the population, \( g(x) \), is differentiable in the neighborhoods of \( x = x_i \) and that \( g_i = g(x_i) \neq 0 \) \( \forall i \). Then the joint distribution of the \( k \) order statistics,

\[ x(n_1:N), \ldots, x(n_k:N) \quad \text{where} \quad n_i = \lfloor N\lambda_i \rfloor + 1, \quad i=1,\ldots,k, \]

tends to a \( k \)-dimensional normal distribution with means \( x_1, \ldots, x_k \) and variance-covariance matrix

\[
\begin{pmatrix}
\frac{1}{N} \frac{\lambda_i (1-\lambda_j)}{g_i g_j}
\end{pmatrix}
\]

as \( N \) tends to infinity.

Note here that the joint distribution of the \( k \) order statistics
\[ X_{(m_1:m)}, \ldots, X_{(m_k:m)} \] from a ranked set sample where \( m_i = [m \lambda_i] + 1 \), \( i = 1, \ldots, k \), \( \lambda_i \)'s as given above, tends to a \( k \)-dimensional normal distribution with means \( x_1, \ldots, x_k \), and variance-covariance matrix
\[
\text{Diag} \left( \frac{\lambda_i (1-\lambda_i)}{mg_i} \right).
\]

The estimators of \( \mu \) or \( \sigma \) for a particular distribution can be improved by an optimal choice of the spacings \( \lambda_1, \ldots, \lambda_k \), for a given \( k \). The optimal spacing for the vector \( \hat{\theta} \) is that which minimizes the generalized variance of \( \hat{\theta} \). Ogawa has found these optimal spacings for some small values of \( k \) for estimating either \( \mu \) or \( \sigma \) (but not both simultaneously except for \( k = 2 \)) when they are parameters of the normal distribution, and for \( \sigma \) from the one-parameter exponential distribution.

Can ranked set sampling be employed to improve unbiased linear estimators of \( \mu \) and \( \sigma \) over those from random samples? Intuition might suggest that this would be true for estimators of \( \mu \), since for most distributions, the BLUE of \( \mu \) is of the form
\[
\hat{\mu} = \sum_{r=1}^{N} \ell_r X_{(r:N)} \quad \text{where} \quad \ell_r > 0, \quad r = 1, \ldots, N.
\]

In such distributions, if \( n = 1 \), then \( m = N \), and
\[
\text{Var}(\hat{\mu}) = \sum_{r=1}^{N} \ell_r^2 \text{Var}(X_{(r:N)}) + 2 \sum_{r<s} \ell_r \ell_s \text{Cov}(X_{(r:N)}, X_{(s:N)})
\]
\[
\geq \sum_{r=1}^{N} c_r^2 \text{Var}(X_{(r:N)})
\]

since \( \text{Cov}(X_{(r:N)}, X_{(s:N)}) \geq 0 \)

when \( X_{(r:N)}, X_{(s:N)} \) are order statistics from a single sample

\[
\geq \sum_{r=1}^{N} c_r^2 \text{Var}(X_{(r:N)})
\]

where \( c_1, \ldots, c_N \) are the coefficients of the BLUE from a ranked set sample

\[ = \text{Var}(\hat{\mu}^*) \].
This shows that ranked set sampling does improve the BLUE of \( \mu \) for all distributions of this type, when \( n = 1 \). However, there are distributions (such as the U-shaped distribution whose density is given on p. 43) for which the coefficients of the BLUE in a random sample are not all positive.

Similarly, intuition might suggest that ranked set sampling would decrease the precision of the estimator of \( \sigma \), since the BLUE's of \( \sigma \) tend to be of the form

\[
\hat{\sigma} = \left[ \frac{N}{2} \right] \sum_{r=1} \ell'_r (X_{N-r+1:N} - X_{r:N}),
\]

with \( \ell'_r > 0 \) \( \forall r \). Then

\[
\text{Var}(\hat{\sigma}) = 2 \left[ \sum_{r=1}^{[N/2]} \ell'_r^2 \text{Var} X_{r:N} + \sum_{r<s<[N/2]} \ell'_r \ell'_s \text{Cov}(X_{r:N}, X_{s:N}) \right] - 2 \left[ \sum_{r=1}^N \ell'_r^2 \text{Cov}(X_{r:N}, X_{N-r+1:N}) \right] + 4 \sum_{r<s<[N/2]} \ell'_r \ell'_s \text{Cov}(X_{r:N}, X_{N-r+1:N})
\]

and when \( n = 1 \), then \( m = N \) and

\[
\text{Var}(\hat{\sigma}^*) = 2 \sum_{r=1}^N c'_r^2 \text{Var} X_{r:N}
\]

where \( c'_1, \ldots, c'_N \) are the coefficients of the BLUE from a ranked set sample.

\( \text{Var}(\hat{\sigma}^*) \) lacks the "benefit" of the negative terms which appear in the expression above for \( \text{Var}(\hat{\sigma}) \).

In fact, one can show that intuition is correct in both cases for \( n = 1 \) and \( m = 2 \) or 3, at least for symmetric distributions. In these cases one can show that

\[
\text{Var}(\hat{\mu}) > \text{Var}(\hat{\mu}^*)
\]

\[
\text{Var}(\hat{\sigma}) < \text{Var}(\hat{\sigma}^*)
\]
Despite these examples, however, we will be able to show that for certain types of distributions, ranked set sampling always improves the BLUE of \( \mu \) and of \( \sigma \), if \( m \) is large enough.

Ranked set sampling as we have discussed it so far has involved choosing for quantification \( n \) each of \( X_{(1:m)}, \ldots, X_{(m:m)} \), from independent samples of size \( m \), until \( mn = N \) observations are quantified. But note that calculation of a BLUE requires knowledge of the distribution of \( U = \frac{X-\mu}{\sigma} \). Having this information may allow us to make a better choice of which order statistic to measure from each of the \( mn \) samples in order to minimize the generalized variance of the BLUE vector. We mentioned this fact in connection with the calculation of MLE's.

First let us suppose that \( X \) is a random variable having a 1-parameter distribution with c.d.f. of the form \( F(x-\mu) \); i.e., we assume \( \sigma \) is known. Then the BLUE of \( \mu \) based on \( n \) repetitions of a set of independent order statistics \( X_{(r_1:m)}, \ldots, X_{(r_m:m)} \) (a sampling scheme which we will call modified ranked set sampling) has variance

\[
\sigma^2 / \sum_{i=1}^{m} \frac{1}{\sigma^2(r_i:m)}
\]

where

\[
\begin{align*}
\sigma^2(r_1:m), \ldots, \sigma^2(r_m:m) \\
\sigma^2(r_1:m), \ldots, \sigma^2(r_m:m)
\end{align*}
\]

This variance is minimized by choosing every \( r_i \) so that \( \sigma^2(r_i:m) \) is minimized over the set \( \{r_i : i = 1, \ldots, m\} \). But we know that

\[
\frac{m \sigma^2([\lambda m]+1:m)}{\lambda(1-\lambda)} \rightarrow \frac{\lambda(1-\lambda)}{f^2(u_\lambda)} \text{ as } m \rightarrow \infty
\]
where \( u^\lambda \) is defined by \( \int_{-\infty}^{u^\lambda} f(t) dt = \lambda \). So by minimizing \( \frac{\lambda(1-\lambda)}{f^2(u^\lambda)} \) with respect to \( \lambda \), we can find the quantile to choose from each sample of size \( m \) which will lead, for large \( m \), to the linear estimator with the smallest variance. Let this optimal \( \lambda \) be \( \lambda^* \), and let \( \hat{\mu}^\Delta \) be the BLUE based on the modified ranked set sample. Then

\[
m^2 n \text{Var}(\hat{\mu}^\Delta) + \frac{\sigma^2 \lambda^*(1-\lambda^*)}{f^2(u^\lambda)} \to \text{Var}(\hat{\mu}) \text{ as } m \to \infty.
\]

For distributions for which the Cramèr-Rao inequality holds, we know

\[
m_n \text{Var}(\hat{\mu}) \geq \sigma^2 \int_{-\infty}^{\infty} \left[ \frac{f'(u)}{f(u)} \right]^2 f(u) du.
\]

Therefore

\[
\frac{\text{Var}(\hat{\mu}^\Delta)}{\text{Var}(\hat{\mu})} \to 0 \text{ as } m \to \infty
\]

for these distributions. In fact, any sample \( \lambda \)-quantile \( (0 < \lambda < 1) \) chosen from each sample would lead to the same conclusion, since

\[
\text{Var} \frac{X_{([m\lambda]+1:m)}}{m} = O\left( \frac{1}{m^2} \right) \text{ as } m \to \infty.
\]

This fact allows us to show that the BLUE \( \hat{\mu}^* \) from a conventional ranked set sample has smaller variance than that from a random sample for \( m \) large enough. (We note here again, however, that there would, in practice, be no reason for using the conventional ranked set sample over the modified ranked set sample in any situation in which a BLUE could be calculated.)

\[
\text{Var}(\hat{\mu}^*) = \frac{\sigma^2}{n \sum_{r=1}^{m} \frac{1}{\sigma^2(r:m)}} \leq \frac{\sigma^2}{n \sum_{r=\lceil(\delta/2)m\rceil}^{\lceil(\delta/2)m\rceil} \frac{1}{\sigma^2(r:m)}} \text{ for } \delta > 0
\]
\[
\sigma^2 \sigma^2_{(k:m)}/[(1-\delta)m]-1
\]

where
\[
\sigma^2_{(k:m)} = \max(\sigma^2_{(j:m)} : j = \lfloor \frac{\delta}{2} m \rfloor, ... , \lfloor (1- \frac{\delta}{2})m \rfloor).
\]

But
\[
\limsup_{m \to \infty} \{m \sigma^2_{(j:m)} : j = \lfloor \frac{\delta}{2} m \rfloor, ... , \lfloor (1- \frac{\delta}{2})m \rfloor \}
= \frac{\lambda_* (1- \lambda_*)}{f^2(u_{\lambda_*})}
\]

if \( f(x) > 0 \) for \( u_{\delta/2} < u < u_{1-\delta/2} \), where \( \lambda_* \) is that value which maximizes \( h_1(\lambda) = (\lambda(1-\lambda))/(f^2(\lambda)) \) in the interval \( \delta/2 < \lambda < 1-\delta/2 \).

If there are finitely many values of \( u \) for which \( f(u) = 0 \) when \( u_{\delta/2} < u < u_{1-\delta/2} \), then label these points \( z_1, ... , z_t \), and let \( \gamma_1, ... , \gamma_t \) be such that \( \int_{-\infty}^{z_j} f(u) du = \gamma_j \), \( j = 1, ..., t \). Then
\[
\text{Var}(\hat{\mu}^*) \leq \frac{\sigma^2}{n} \left\{ \sum_{r = \lfloor \frac{\delta}{3} m \rfloor}^{\lfloor \gamma_1 - \frac{\delta}{3t} m \rfloor} \frac{1}{\sigma^2_{(r:m)}} + \sum_{r = \lfloor \gamma_1 + \frac{\delta}{3t} m \rfloor}^{\lfloor \gamma_2 - \frac{\delta}{3t} m \rfloor} \frac{1}{\sigma^2_{(r:m)}} + \ldots + \sum_{r = \lfloor \gamma_t + \frac{\delta}{3t} m \rfloor}^{\lfloor (1- \frac{\delta}{3}) m \rfloor} \frac{1}{\sigma^2_{(r:m)}} \right\}.
\]

Let \( \lambda_* \) be that value of \( \lambda \) which maximizes \( h_1(\lambda) \) over the region
\[
\left( \frac{\delta}{3} < \lambda < \gamma_1 - \frac{\delta}{3t} \right) \cup \cdots \cup \left( \gamma_t + \frac{\delta}{3t} < \lambda < \frac{\delta}{3} \right).
\]

Then for either case,
\[
\lim_{m \to \infty} m^2 n \text{Var}(\hat{\mu}^*) \leq \frac{\sigma^2 \lambda_* (1- \lambda_*)}{(1-\delta)f^2(x_{\lambda_*})}
\]

\[
\implies \frac{\text{Var}(\hat{\mu}^*)}{\text{Var}(\hat{\mu})} \to 0 \text{ as } m \to \infty
\]
if $X$ has a distribution for which the Cramér-Rao inequality holds. Notice that this proves that for $m$ large enough,

$$
\frac{1}{V^{*1/2}} \succeq \frac{1}{V^{-1/2}}
$$

where the upper triangle of the symmetric matrix

$$
V = \begin{bmatrix}
\sigma^2_{(1:m)} & \sigma^2_{(12:m)} & \cdots & \sigma^2_{(1m:m)} \\
\sigma^2_{(2:m)} & \sigma^2_{(22:m)} & \cdots & \sigma^2_{(2m:m)} \\
\vdots & \ddots & \ddots & \vdots \\
\sigma^2_{(m:m)} & & & \\
\end{bmatrix}
$$

and $V^* = \text{Diag}(\sigma^2_{(1:m)}, \ldots, \sigma^2_{(m:m)})$.

Now suppose $X$ has a one-parameter distribution with c.d.f. of the form $F\left(\frac{X}{\theta}\right)$. Here again we would benefit by choosing a modified ranked set sample on which to base the BLUE rather than a conventional ranked set sample. In this case, the optimal quantile to measure from each sample for large $m$ is $X([m\lambda]+1:m)$, where $\lambda$ minimizes

$$
h_2(\lambda) = \frac{\lambda(1-\lambda)}{f^*(u_\lambda)u_\lambda^2}.
$$

Let the optimal $\lambda$ be $\lambda^*_\theta$. Then

$$
\hat{\sigma}^A = \frac{1}{m} \sum_{i=1}^{mn} X([m\lambda^*_\theta]+1:m)/mn\lambda^*_\theta([m\lambda^*_\theta]+1:m)
$$

and

$$
\text{Var}(\hat{\sigma}^A) = \frac{\sigma^2}{mn\lambda^*_\theta([m\lambda^*_\theta]+1:m)}.
$$

Then we have
\[
\begin{align*}
\text{mn} \ Var(\hat{\sigma}^D) & \geq \sigma^2 \int_{-\infty}^{\infty} \left\{ \left[ \frac{uf'(u)}{f(u)} \right]^2 - 1 \right\} f(u) du \\
\text{Therefore} & \quad \frac{\text{Var}(\hat{\sigma}^D)}{\text{Var}(\hat{\sigma}^D)} \rightarrow 0 \quad \text{as } m \rightarrow \infty.
\end{align*}
\]

By an argument identical to that used for showing that \( \text{Var}(\hat{\mu}^*) \) \\
< \text{Var}(\hat{\mu}) \) for large \( m \), we can show
\[
\frac{\text{Var}(\hat{\sigma}^*)}{\text{Var}(\hat{\sigma})} \rightarrow 0 \quad \text{as } m \rightarrow \infty,
\]
and thus \( \alpha'^* \sigma^{-1} \alpha \geq \alpha' \sigma^{-1} \alpha \) for \( m \) sufficiently large.

Now suppose the c.d.f. of \( X \) is of the form \( F\left( \frac{X - \mu}{\sigma} \right) \), \( \mu \) and \( \sigma \) both unknown. Here again one can find, in certain situations, better linear estimators of \( \mu \) and \( \sigma \) than \( \hat{\mu}^* \) and \( \hat{\sigma}^* \) by choosing a modified ranked set sample. However, \( \mu \) and \( \sigma \) cannot both be estimated with linear estimators if the same quantile is chosen from each of the \( mn \) samples, since in that case
\[
A = \begin{bmatrix}
1 & \alpha_{(r:m)} \\
1 & \alpha_{(r:m)} \\
\vdots & \vdots \\
1 & \alpha_{(r:m)}
\end{bmatrix}
\]
has rank 1.

Therefore, at least two different quantiles must be included among those order statistics which make up the modified ranked set sample.
Suppose we include exactly 2 different quantiles in our sample; i.e., suppose we quantify mn/2 r-th and mn/2 s-th order statistics, each from a sample of size m, with r>s; then the generalized variance of \( \hat{\theta}^A \) is

\[
\left| \text{Var} \hat{\theta}^A \right|^{1/2} = \left( \frac{\sigma^2}{mn} \right) \left[ \left( \frac{1}{\sigma^2(r:m)} + \frac{1}{\sigma^2(s:m)} \right) \left( \frac{\alpha^2(r:m)}{\sigma^2(r:m)} + \frac{\alpha^2(s:m)}{\sigma^2(s:m)} \right) \right.
\]

\[
- \left( \frac{\alpha^2(r:m)}{\sigma^2(r:m)} + \frac{\alpha^2(s:m)}{\sigma^2(s:m)} \right) \left] \right)^{1/2}
\]

\[
= \frac{2\sigma^2 \sigma(r:m) \sigma(s:m)}{mn(\alpha(r:m) - \alpha(s:m))}.
\]

If the r-th and s-th order statistics are the \( \lambda_1 \) - and \( \lambda_2 \)-quantiles, \( 0 < \lambda_1 < 1, 0 < \lambda_2 < 1 \), then for large m, \( \left| \text{Var} \hat{\theta}^A \right|^{1/2} \) is minimized by minimizing

\[
\frac{\lambda_1 \lambda_2 (1-\lambda_1) (1-\lambda_2)}{f^2(u_{\lambda_1}) f^2(u_{\lambda_2}) (u_{\lambda_1} - u_{\lambda_2})^2}
\]

wrt \( \lambda_1 \) and \( \lambda_2 \) over the rectangle \( 0 < \lambda_1 < 1, 0 < \lambda_2 < 1 \). From this expression we can see that

\[
\left| \text{Var} \hat{\theta}^A \right|^{1/2} = O\left( \frac{1}{mn} \right) \quad \text{as} \quad m \to \infty
\]

regardless of which pair of quantiles are chosen. We know from (2.3) that

\[
\left| \text{Var} \hat{\theta}^A \right|^{1/2} = O\left( \frac{1}{mn} \right) \quad \text{as} \quad m \to \infty
\]

in any regular estimation situation. Consequently we have

\[
\frac{\left| \text{Var} \hat{\theta}^A \right|^{1/2}}{\left| \text{Var} \hat{\theta} \right|^{1/2}} \to 0 \quad \text{as} \quad m \to \infty.
\]
Suppose $X$ has a distribution symmetric around $\mu$ and we wish to find the optimal pair of quantiles $X_{(\lceil m\lambda_1 \rceil + 1 : m)}$, $X_{(\lceil m\lambda_2 \rceil + 1 : m)}$, where $\lambda_1 = 1 - \lambda_2$, for estimating $\mu$ and $\sigma$. If the quantified sample includes $mn/2$ each of the $\lambda_1$- and $\lambda_2$-quantile, then

$$|\text{Var} \hat{\theta}^{\Delta}|^{1/2} = \sigma^2 \sigma^2 ([m\lambda_1] + 1 : m) / \frac{mn\alpha}{([m\lambda_1] + 1 : m)}$$

since $\text{cov}(\hat{\mu}, \hat{\sigma}^{\Delta}) = 0$, $\sigma(r:m) = -\alpha(m-r+1:m)$, $\sigma^2(r:m) = \sigma(m-r+1:m)$.

These optimal quantiles, $\lambda^*$ and $1-\lambda^*$, are found for large $m$ by minimizing

$$h_2(\lambda) = \left| \frac{\lambda(1-\lambda)}{u_\lambda f^2(u_\lambda)} \right| \text{ for } 0 < \lambda < 1.$$ 

For symmetric distributions, the BLUE of $\theta$ from a conventional ranked set sample has generalized variance (from (2.10))

$$|\text{Var} \hat{\theta}^{*}|^{1/2} = \frac{\sigma^2}{n} \left[ \sum_{r=1}^{m} \frac{1}{\sigma^2(r:m)} \left[ \sum_{r=1}^{m} \frac{\alpha^2(r:m)}{\sigma^2(r:m)} \right] \right]^{1/2}$$

$$\leq \frac{\sigma^2}{n} \left[ \sum_{r=\lceil \delta/2 \rceil m}^{\lceil m/2 \rceil} \frac{1}{\sigma^2(r:m)} \left[ \sum_{r=\lceil \delta/2 \rceil m}^{\lceil m/2 \rceil} \frac{\alpha^2(r:m)}{\sigma^2(r:m)} \right] \right]^{1/2}.$$ 

(2.11)

By using the same argument as for the 1-parameter case, we can show that each of the sums of (2.11) are $O(m^2)$ as $m \to \infty$, which leads us directly to the conclusion that

$$\frac{|\text{Var} \hat{\theta}^{*}|^{1/2}}{|\text{Var} \hat{\theta} |^{1/2}} \to 0 \text{ as } m \to \infty.$$
Suppose $X \sim N(\mu, \sigma)$. Then $h_1(\lambda)$ is minimized by $\lambda^*_\mu = .5$.

This indicates that the best linear unbiased estimator of $\mu$, for $\sigma^2$ known, based on a modified ranked set sample is obtained by choosing the sample median from each sample. If $\mu$ and $\sigma$ are both unknown, but $\mu$ is the only parameter of interest, it will still be estimated most accurately by this linear estimator. Similarly, one can show that $h_2(\lambda)$ is minimized by $\lambda^*_\sigma = .942$. So if $\mu$ is known, $\sigma$ can be best estimated (for large $m$) with a linear estimator by choosing the $([.942m]+1)$-st order statistic from each sample of size $m$. If $\mu$ and $\sigma$ are both unknown, but $\sigma$ is the only parameter of interest, it can be estimated most accurately by choosing \(\frac{mn}{2}([.942m]+1)-\text{st}\) and \(\frac{mn}{2}(m-([.942m]+1))-\text{st}\) order statistics. If neither $\mu$ nor $\sigma$ is known, then the optimal symmetric pair of quantiles for estimating both $\mu$ and $\sigma$ from linear estimators are the 87.3 and the 13.7 sample percentiles.

So far, we have established that modified ranked set BLUE's of $\theta = \begin{bmatrix} \mu \\ \sigma \end{bmatrix}$ can be found whose generalized variance is smaller than that of the BLUE based on a random sample, for any distribution for which the Cramèr-Rao inequality holds. Furthermore, we have shown for this same condition that a BLUE based on the conventional ranked set sample has smaller variance than that of the random sample when its distribution is one-parameter or symmetric. Notice that if we could show that the BLUE based on a ranked set sample is efficient, we would have proved this fact in greater generality. Unfortunately, I have not been able to prove this, although I believe that it is true. Recall that Blom and Jung have shown that BLUE's based on random samples are efficient, for distributions satisfying
certain (fairly restrictive, in Blom's case) conditions. They did this, as we have mentioned, by displaying another asymptotically unbiased linear function of the order statistics which was efficient. To arrive at his results, Blom proved and used a theorem ([2], p. 53) which states that, for distributions for which \( F^{-1}(u) = G(\mu) \) satisfies certain conditions,

\[
\text{EX}_{(r:m)} = G\left(\frac{r}{m+1}\right) + R_{r}
\]

and

\[
|R_{r}| < \frac{M^{*}}{m}, \quad \text{where} \quad M^{*} \text{ does not depend on } r \text{ or } m,
\]

if \( X_{(r:m)} \) belongs to a sequence of random variables of the form \( \{X_{(r_{\nu}, n_{\nu})}\} \) where \( n_{\nu} < n_{\nu+1} \) and \( r_{\nu}/n_{\nu} \rightarrow c \) as \( \nu \rightarrow \infty \), for all but possibly a finite number of points of the interval \([0,1] \) (and another condition for the remaining points) A second theorem ([3], p. 55) states a similar result for covariance:

\[
\text{Cov}(X_{(r:m)}, X_{(s:m)}) = \frac{1}{m+2} \frac{r(m-r+1)}{(m+1)^2} G'(r_{m+1}) G'(s_{m+1}) + R_{rs} \quad (r \leq s)
\]

where \( |R_{rs}| < \frac{M^{*}}{m^2} \), and \( M^{*} \) does not depend on \( r, s, \) or \( m \).

It may be that a proof of my conjecture could be produced by mimicking Blom's method. However, a more direct approach might be possible, although perhaps not without the introduction of additional restrictions, since (for example) we see from Blom's theorem that

\[
\text{Var} \ X_{(r:m)} = \frac{r}{m+1} \left(1 - \frac{r}{m+1}\right) + O\left(\frac{1}{m^2}\right), \quad \text{Var} \left(\hat{\mu}^{*}\right) = \frac{\sigma^2}{n^{1/2}} I^*_{(m)n}(\mu)
\]

and

\[
I^*_{(m)n}(\mu) = nm^{2} \int_{-\infty}^{\infty} \frac{f^2(u)}{F(u)[1-F(u)]} \, du + O(m).
\]
Unfortunately, Blom's theorem holds only for quantiles, and does not assure us that his condition on the variance of $X_{(r:m)}$ holds for all $r$, even in the range $[\delta m] < r < [(1-\delta)m]$, $\delta > 0$.

2.3 $s^2$-type Estimators of $\sigma^2$

So far in this chapter, several methods for estimating scale parameters have been discussed. But for each of the methods, knowledge of the form of the distribution is necessary, as well as the assumption that perfect ordering is present. Dell and Clutter's paper dealt with an estimator of the mean, $\bar{X}$, which was improved (or at least not hurt) by the process of ranked set sampling, regardless of the underlying distribution or of the presence of ranking errors. Here our purpose is to determine how ranked set sampling affects $s^2_N = \frac{1}{N-1} \sum_{i=1}^{N} (X_i - \bar{X})^2$ as an estimator of the variance $\sigma^2$ (when it exists) of any distribution.

Let $X_{[1:m]}, \ldots, X_{[m:m]}$ be a set of judgment ordered random variables, and let $E[X_{(r:m)}] = \mu_{(r:m)}$, $\text{Var} X_{(r:m)} = \sigma^2_{(r:m)}$, $\tau_{(r:m)} = \mu_{(r:m)} - \mu$. Then

$$\sum_{r=1}^{m} E[X_{(r:m)} - a]^k = mE(X-a)^k.$$  \hspace{1cm} (2.12)

As particular cases of this relationship, we see that

$$\sum_{r=1}^{m} \tau_{(r:m)} = 0$$  \hspace{1cm} and \hspace{1cm} $$\sum_{r=1}^{m} \sigma^2_{(r:m)} = m\sigma^2 - \sum_{r=1}^{m} \tau^2_{(r:m)}.$$

A natural variance estimator calculated from the ranked set sample of size $mn = N$, $X_{[1:m], 1}, \ldots, X_{[m:m], n}$ is

$$s^2_{[m]n} = \frac{1}{mn-1} \sum_{i=1}^{n} \sum_{r=1}^{m} (X_{(r:m), i} - \hat{\mu}_{[m]n})^2.$$
Then

\[ E s^2_{[m]} = \frac{n}{mn-1} \left( \sum_{r=1}^{m} E X^2_{r:m} - m \mu^2_{[m]} \right) \]

\[ = \frac{n}{mn-1} \left[ \sum_{r=1}^{m} (\sigma^2_{r:m} + \mu^2_{r:m}) - \frac{m}{m n} \left( \sum_{r=1}^{m} \sigma^2_{r:m} \right) + \mu^2 \right] \]

\[ = \sigma^2 - \frac{1}{m (mn-1)} \sum_{r=1}^{m} \tau^2_{r:m} , \]

so \( s^2_{[m]} \) is a biased estimator of \( \sigma^2 \). However,

\[ E s^2_{[m]} - \sigma^2 = \frac{1}{m (mn-1)} \sum_{r=1}^{m} \tau^2_{r:m} = \frac{\sigma^2}{mn-1} - \frac{m \sigma^2_{r:m}}{m (mn-1)} \leq \sigma^2_{mn-1} . \]

So \( s^2_{[m]} \) is asymptotically unbiased and the bias is no larger than that which results from estimating \( \sigma^2 \) by \( \frac{N-1}{N} s^2_N \), though it cannot be removed in the same way.

Ranked set sampling does not always improve the precision of the estimator \( s^2_N \); i.e., it is not always true that

\[ MSE s^2_{[m]} = \text{Var} s^2_{[m]} + (E s^2_{[m]} - \sigma^2)^2 \leq \text{Var} s^2_N . \]

as was the case for Dell and Clutter's estimator of \( \mu \). However, if \( mn = N \) is sufficiently large, we will show that

\[ MSE s^2_{[m]} \leq \text{Var} s^2_N \]

regardless of the underlying distribution (as long as the appropriate number of moments exist) or of the presence of ranking errors.

To simplify notation in the following calculations, we write

\( X_{r:m} \equiv X_{r i}, \mu_{r:m} \equiv \mu_r, \mu_{[m]} \equiv \mu, s^2_{[m]} \equiv s^2, \sigma^2_{r:m} \equiv \sigma^2_r, \)

\( \tau_{r:m} \equiv \tau_r \). Then
\[(mn-1)s^2 = \sum_{i=1}^{n} \sum_{r=1}^{m} [(X_{ri} - \mu_r) - (\hat{\mu} - \mu) + (\mu_r - \mu)]^2 \]

\[= \sum_{i} \sum_{r} (X_{ri} - \mu_r)^2 + \sum_{i} \sum_{r} (\hat{\mu} - \mu)^2 + \sum_{i} \sum_{r} (\mu_r - \mu)^2 \]

\[-2 \sum_{i} \sum_{r} (X_{ri} - \mu_r)(\hat{\mu} - \mu) + 2 \sum_{i} \sum_{r} (X_{ri} - \mu_r)(\mu_r - \mu) \]

\[-2 \sum_{i} \sum_{r} (\hat{\mu} - \mu)(\mu_r - \mu) \]

Let

\[T = \sum_{i} \sum_{r} (X_{ri} - \mu_r)^2 + mn(\hat{\mu} - \mu)^2 - 2(\hat{\mu} - \mu) \sum_{i} \sum_{r} (X_{ri} - \mu_r) \]

\[+ 2 \sum_{r} \sum_{i} \tau_r(X_{ri} - \mu_r) \]

\[= \sum_{i} \sum_{r} Y_{ri}^2 - \frac{1}{mn} (\sum_{i} \sum_{r} Y_{ri})^2 + 2 \sum_{r} \sum_{i} \tau_r Y_{ri} \]

\[= (\frac{mn-1}{mn}) \sum_{i} \sum_{r} Y_{ri}^2 - \frac{2}{mn} \sum_{r<s} \sum_{i=1}^{m} \sum_{j=1}^{m} Y_{ri} Y_{sj} - \frac{2}{mn} \sum_{r=1}^{m} \sum_{i<j} \sum_{r} Y_{ri} Y_{rj} \]

\[+ 2 \sum_{i} \sum_{r} \tau_r Y_{ri} \]

where \(Y_{ri} = X_{ri} - \mu_r\). Then \(EY_{ri} = 0\), \(\text{Var}Y_{ri} = \sigma_r^2\), and \(\text{cov}(Y_{ri}, Y_{sj}) = 0\) if \(r \neq s\) or \(i \neq j\), and \(\text{var}(mn-1)s^2 = \text{var}T\)

\[= ET^2 - (ET)^2 \]

\[ET = n(\frac{mn-1}{mn}) \sum_{i=1}^{m} \sigma_r^2 \]

\[ET^2 = E\left\{(\frac{mn-1}{mn})^2 \left[\sum_i \sum_r Y_{ri}^4 + 2 \sum_r \sum_{i<j} Y_{ri}^2 Y_{rj}^2 + 2 \sum_r \sum_i Y_{ri}^2 Y_{rj}^2 \right] \right\} \]

\[+ \frac{4}{(mn)^2} \left[\sum_{r<s} \sum_{i} \sum_{j} Y_{ri}^2 Y_{rj}^2 \right] + \frac{4}{(mn)^2} \left[\sum_{r<i<j} \sum_{i} Y_{ri}^2 Y_{rj}^2 \right] \]

\[+ 4 \sum_{i} \sum_{r} \tau_r^2 Y_{ri} + 4(\frac{mn-1}{mn}) \sum_{i} \sum_{r} \tau_r Y_{ri}^3 \text{ terms having expectation 0 } \]
\[ = n \left[ \frac{(mn-1)}{mn} \sum_r \mu_{4r} + 4\sum_r \tau_r \sigma_r^2 + 4\left( \frac{(mn-1)}{mn} \right) \sum_r \tau_r \mu_{3r} \right. \]

\[ \left. + \frac{(m^2 n^2 - 2mn + 3)}{(mn)^2} \left( 2n \sum_{r<s} \sigma_r^2 \sigma_s^2 + (n-1) \sum_r \sigma_r^4 \right) \right] \]

where \( \mu_{kr} = E(X_r - \mu_k)^k \). So

\[ \text{Var}(mn-1)s^2 = n \left[ \frac{(mn-1)}{mn} \sum_r \mu_{4r} + 4\sum_r \tau_r \sigma_r^2 + 4\left( \frac{(mn-1)}{mn} \right) \sum_r \tau_r \mu_{3r} \right. \]

\[ \left. + \frac{4n}{(mn)^2} \sum_{r<s} \sigma_r^2 \sigma_s^2 + \frac{2(n-1) - (mn-1)^2}{(mn)^2} \sum_r \sigma_r^4 \right] . \]

(2.13)

Note as a check that if the ranked set sample were indeed a random sample (i.e., judgment ordering no better than random), then

\( \sigma_r^2 = \sigma^2, \tau_r = 0, \mu_{4r} = \mu_4 = E(X-\mu)^4 \), yielding

\[ \text{Var}(mn-1)s^2 = n \left[ \frac{(mn-1)}{mn} m \mu_4 + \frac{4n}{(mn)^2} \frac{m(m-1)}{2} \sigma^4 + \frac{2(n-1) - (mn-1)^2}{(mn)^2} m \sigma^4 \right] \]

\[ = \frac{(mn-1)^2}{mn} \mu_4 + \frac{(mn-3)(mn-1)}{mn} \sigma^4 \]

\[ = \text{Var}(N-1)s^2_N . \]

(2.14)

Var T will be rewritten in a form which contains only expectations of powers of \((X_r - \mu)\) and \(\tau_r\) by noting:

(a) \[ \sum_r \mu_{4r} = \sum_r E[(X_r - \mu) - (\mu_r - \mu)]^4 \]

\[ = E(X-\mu)^4 - 3\sum_r \tau_r^4 - 4\sum_r E(X_r - \mu) \tau_r^3 + 6 \sum_r E(X_r - \mu)^2 \tau_r^2 \]

(b) \[ \sum_r \tau_r^2 \sigma_r^2 = \sum_r \tau_r^2 E(X_r - \mu)^2 + \sum_r \tau_r^4 \]
(c) \[ \sum_{r} \tau_{r} E(X_{r} - \mu)^{3} r + 2 \sum_{r} \tau_{r}^{4} r - 3 \sum_{r} \tau_{r} E(X_{r} - \mu)^{2} r^{2} \]

(d) \[ \left( \sum_{r} \sigma_{r}^{2} \right)^{2} = \sum_{r} \sigma_{r}^{4} r + 2 \sum_{r<s} \sigma_{r}^{2} \sigma_{s}^{2} \Rightarrow 2 \sum_{r<s} \sigma_{r}^{2} \sigma_{s}^{2} = \left( \sum_{r} \sigma_{r}^{2} \right)^{2} - \sum_{r} \sigma_{r}^{4} \] (2.15)

Thus from (2.13) we see that

\[
\text{Var}(mn-1) s^{*2} = \frac{n}{(mn-1)^2} \left\{ \left( \frac{mn-1}{mn} \right)^2 mnE(X-\mu)^4 + \left( \frac{mn-3}{mn} \right)^2 \sum_{r} \tau_{r}^{4} r \right. \\
+ \frac{4(mn-1)}{(mn)^2} \sum_{r} \tau_{r} E(X_{r} - \mu)^{3} r + \frac{-2(mn-1)^2 + 6}{(mn)^2} \sum_{r} \tau_{r}^{2} E(X_{r} - \mu)^{2} r \\
+ \frac{-2(mn-1)^2 + 2mn-3}{(mn)^2} \sum_{r} \sigma_{r}^{4} r \left. \right. + \frac{2n}{(mn)^2} \left. \left. \left\{ \left( \sum_{r} \sigma_{r}^{2} \right)^{2} \right. \right. \right. \\
\right. \\
Then \[
\text{MSE} (s^{*2}) = \text{Var} (s^{*2}) + (E s^{*2} - \sigma^{2})^2 = \frac{1}{mn} E(X - \mu)^4 \\
+ \frac{(mn-3)(mn+1)}{(mn-1)^2 mn} \sum_{r} \tau_{r}^{4} r + \frac{4}{(mn-1)^2 mn} \sum_{r} \tau_{r} E(X_{r} - \mu)^{3} r \\
+ \frac{-2(mn)^2 + 6}{(mn-1)^2 mn} \sum_{r} \tau_{r}^{2} E(X_{r} - \mu)^{2} r - \frac{mn-2mn+3}{(mn-1)^2 mn} \sum_{r} \sigma_{r}^{4} r \\
+ \frac{2}{mn} \left( \sum_{r} \sigma_{r}^{2} \right)^{2} + \frac{1}{(mn-1)^2 mn} \left( \sum_{r} \tau_{r}^{2} \right)^{2} \] (2.16)

\textbf{Lemma.} Let \( X_{[1:m]} \ldots, X_{[m:m]} \) be a ranked set sample of judgment order statistics from a continuous distribution whose fourth moment exists. Then:

(a) \[ \sum_{r=1}^{m} \tau_{[r:m]}^{2} E(X_{[r:m]} - \mu)^{2} = 0(m) \]

(b) \[ \sum_{r=1}^{m} \tau_{[r:m]}^{4} = 0(m) \text{ and } \sum_{r=1}^{m} \sigma_{[r:m]}^{4} = 0(m) \]
\[ \left( \sum_{r=1}^{m} \tau_{r}^{2} \right)^{2} = 0(m^{2}) \quad \text{and} \quad \left( \sum_{r=1}^{m} \sigma_{r}^{2} \right)^{2} = 0(m^{2}) \]

(d) \[ \sum_{r=1}^{m} \tau_{r} \mu_{r}^{3} = O(m) . \]

\textbf{Proof:}

(a) \[ \left[ E(X_{r} - \mu) \right]^{2} \leq E(X_{r} - \mu)^{4} \quad \text{and} \]

\[ 0 \leq \tau_{r}^{2} = E(X_{r} - \mu)^{2} - \sigma_{r}^{2} \leq E(X_{r} - \mu)^{2} \]

\[ \Rightarrow \sum_{r} E(X_{r} - \mu)^{2} \tau_{r}^{2} \leq \sum_{r} [E(X_{r} - \mu)^{2}]^{2} \leq \sum_{r} E(X_{r} - \mu)^{4} = mE(X - \mu)^{4} \]

(b) \[ \sum_{r} \tau_{r}^{4} \leq \sum_{r} E(X_{r} - \mu)^{2} \tau_{r}^{2} \leq mE(X - \mu)^{4} \quad \text{and} \]

\[ \sum_{r} \sigma_{r}^{4} \leq \sum_{r} [E(X_{r} - \mu)^{2}]^{2} \leq mE(X - \mu)^{4} \]

(c) \[ \left( \sum_{r} \tau_{r}^{2} \right)^{2} = \left( m\sigma^{2} - \sum_{r} \tau_{r}^{2} \right)^{2} \leq m^{2} \sigma^{4} \quad \text{and} \]

\[ \left( \sum_{r} \sigma_{r}^{2} \right)^{2} = \left( m\sigma^{2} - \sum_{r} \tau_{r}^{2} \right)^{2} \leq m^{2} \sigma^{4} \]

(d) \[ \sum_{r} \tau_{r} E(X_{r} - \mu)^{3} \leq \left( \sum_{r} \tau_{r}^{4} \right)^{1/4} \left\{ \sum_{r} [E(X_{r} - \mu)^{3}]^{4/3} \right\}^{3/4} \text{ by Hölder's inequality} \]

\[ \leq mE(X - \mu)^{4} \left\{ \sum_{r} E(X_{r} - \mu)^{4} \right\}^{3/4} = mE(X - \mu)^{4} = O(m) . \]

\textbf{Theorem.} Let \( X_{[1:m], 1, \ldots, X_{[m:m]}, n} \) be a ranked set sample of judgment order statistics from a continuous distribution whose fourth moment exists. Then \( \exists N^* \) s.t. for \( N = mn > N^* \),

\[ \text{MSE} \ s_{[m:n]}^{2} \leq \text{Var} \ s_{N}^{2} . \]

\textbf{Proof:} From the lemma and (2.16)
\[
\text{MSE } s^*_{(m)N} = \frac{1}{m n} E(X - \mu)^4 + \frac{1}{2 m n} \sum_r \tau_r^4 - \frac{2}{m n} \sum_r \tau_r^2 E(X_r - \mu)^2 - \frac{1}{m n} \sum_r \sigma_r^4 + O\left(\frac{1}{(mn)^2}\right)
\]

\[
= \frac{1}{m n} E(X - \mu)^4 + \frac{1}{2 m n} \sum_r \tau_r^4 - \frac{2}{m n} \sum_r \tau_r^2 E(X_r - \mu)^2 - \frac{1}{m n} \sum_r \left[ E(X_r - \mu)^2 - \tau_r^2 \right]^2
\]

\[
+ O\left(\frac{1}{(mn)^2}\right)
\]

\[
= \frac{1}{m n} E(X - \mu)^4 - \frac{1}{m n} \sum_r \left[ E(X_r - \mu)^2 \right]^2 + O\left(\frac{1}{(mn)^2}\right)
\]

\[
= \frac{1}{m n} E(X - \mu)^4 - \frac{1}{m n} \sum_r \left( \sigma_r^2 + \tau_r^2 \right)^2 + O\left(\frac{1}{(mn)^2}\right),
\]

and from (2.14)

\[
\text{Var } s^2_N = \frac{1}{m n} E(X - \mu)^4 - \frac{1}{m n} \sigma^4 + O\left(\frac{1}{(mn)^2}\right).
\]

But

\[
\sum_r \left( \sigma_r^2 + \tau_r^2 \right) = m \sigma^2 \Rightarrow m \sigma^4 = \left[ \sum_r \left( \sigma_r^2 + \tau_r^2 \right) \right]^2,
\]

so

\[
\frac{1}{m n} \sigma^4 = \frac{1}{3 m n} \left[ \sum_r \left( \sigma_r^2 + \tau_r^2 \right) \right]^2 = \frac{1}{m n} \left[ \sum_r \frac{1}{m} \left( \sigma_r^2 + \tau_r^2 \right) \right]^2
\]

\[
\leq \frac{1}{m n} \left[ \left( \sum_r \frac{1}{m} \right) \sum_r \left( \sigma_r^2 + \tau_r^2 \right) \right]
\]

\[
= \frac{1}{m n} \sum_r \left( \sigma_r^2 + \tau_r^2 \right)^2.
\]

Table 2.1 compares MSE \( s^*_{(m)1}^2 \) and \( \text{var } s^2_m \) for several distributions to see how large \( m \) must be before MSE \( s^*_{(m)n}^2 \leq \text{var } s^2_N \) when \( n = 1 \).
Table 2.3.1

Comparison of MSE $s_{(m)}^2$ and var $s_m^2$. The "cross-over points" are marked by arrows.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>m</th>
<th>MSE $s_{(m)}^2$</th>
<th>var $s_m^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) U-shaped</td>
<td>2</td>
<td>.460</td>
<td>.394 +</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>.117</td>
<td>.143</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>.060</td>
<td>.077</td>
</tr>
<tr>
<td>(2) Uniform</td>
<td>2</td>
<td>.0134</td>
<td>.0097</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>.0046</td>
<td>.0042 +</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>.0024</td>
<td>.0026</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>.0015</td>
<td>.0018</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>.0010</td>
<td>.0014</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>.0076</td>
<td>.0011</td>
</tr>
<tr>
<td>* (3) Normal [1]</td>
<td>2</td>
<td>2.96</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.24</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>.722</td>
<td>.667 +</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>.483</td>
<td>.500</td>
</tr>
<tr>
<td>* (4) Gamma r = 5 [6]</td>
<td>2</td>
<td>91.50</td>
<td>65.00</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>42.95</td>
<td>35.00</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>26.49</td>
<td>24.17 +</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>18.46</td>
<td>18.50</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>13.81</td>
<td>15.00</td>
</tr>
<tr>
<td>Distribution</td>
<td>m</td>
<td>MSE $s^2_{[m]}$</td>
<td>var $s^2_m$</td>
</tr>
<tr>
<td>--------------</td>
<td>---</td>
<td>----------------</td>
<td>-------------</td>
</tr>
<tr>
<td>*(5) Gamma $r=3$ [6]</td>
<td>2</td>
<td>37.11</td>
<td>27.00</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>18.28</td>
<td>15.00</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>11.55</td>
<td>10.50</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>8.18</td>
<td>8.10</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>6.19</td>
<td>6.60</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>4.89</td>
<td>5.57</td>
</tr>
<tr>
<td>*(6) Exponential</td>
<td>2</td>
<td>6.38</td>
<td>5.00</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3.56</td>
<td>3.00</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>2.39</td>
<td>2.17</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1.76</td>
<td>1.70</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>1.37</td>
<td>1.40</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>1.10</td>
<td>1.19</td>
</tr>
<tr>
<td>*(7) Lognormal [7]</td>
<td>2</td>
<td>1337.4</td>
<td>1253.7</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>873.1</td>
<td>828.3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>645.2</td>
<td>621.0</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>510.0</td>
<td>494.9</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>421.0</td>
<td>412.1</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>357.7</td>
<td>353.0</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>310.6</td>
<td>308.8</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>270.7</td>
<td>274.4</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>245.4</td>
<td>246.9</td>
</tr>
</tbody>
</table>

* References for tables of moments of order statistics needed for performing these calculations are indicated.
If \( n > 1 \), an unbiased "\( s^2 \)-type" estimator can be found.

Consider the estimator \( \tilde{s}^2 \), calculated from a ranked set sample:

\[
\tilde{s}^2 = \frac{1}{mn} \left[ \sum_{i=1}^{n} \sum_{r=1}^{m} (X_{[r:m],i} - \hat{\mu}_{[m]})^2 \right] + \frac{1}{(n-1)m} \sum_{i=1}^{n} \sum_{r=1}^{m} (X_{[r:m],i} - \hat{\mu}_{[r:m]})^2
\]

where

\[
\hat{\mu}_{[r:m]} = \frac{1}{n} \sum_{i=1}^{n} X_{[r:m],i}.
\]

Then

\[
E\tilde{s}^2_{[m]n} = \frac{1}{mn} [(mn-1)\sigma^2 - \frac{1}{m} \sum_{r=1}^{m} \tau^2_{[r:m]} + \frac{1}{(n-1)m} \sum_{r=1}^{m} (n-1)\sigma^2_{[r:m]}]
\]

\[
= \sigma^2
\]

and

\[
\text{Var} \tilde{s}^2_{[m]n} = \left( \frac{mn-1}{mn} \right)^2 \text{Var} s^2 + \frac{1}{(mn)^2} \frac{1}{(n-1)^2} \frac{1}{m^2} A - \frac{2}{(mn)^2} \frac{1}{(n-1)m} B
\]

where

\[
A = \text{Var} \left[ \sum_{i} \sum_{r} (X_{ri} - \hat{\mu}_{[r:m]})^2 \right]
\]

and

\[
B = \text{Cov} \left[ \sum_{i} \sum_{r} (X_{ri} - \hat{\mu}_{[r:m]})^4, \sum_{i} \sum_{r} (X_{ri} - \hat{\mu}_{[r:m]})^2 \right].
\]

Using (2.14), one can see that

\[
A = \frac{m}{n} \sum_{r=1}^{m} \text{Var} \left[ \sum_{i=1}^{n} (X_{ri} - \hat{\mu}_{[r:m]})^2 \right] = \frac{(n-1)^2}{n^2} \frac{\sum \mu^4_{4r}}{m^2} - \frac{(n-3)(n-1)}{n^2} \frac{\sum \sigma^4_{r}}{r^2}.
\]

Therefore

\[
\frac{1}{(mn)^2} \frac{1}{(n-1)^2} \frac{1}{m^2} A = \frac{1}{(mn)^2} \frac{\sum \mu^4_{4r}}{m^2 n} - \frac{(n-3)}{(mn)^2 (n-1)m} \frac{\sum \sigma^4_{r}}{r^2}
\]

\[= O(\frac{1}{m}) \text{ by (2.15 a) and the lemma}.
\]

By a method identical to that used for finding \( \text{Var}(mn-1)s^2 \), \( B \) can be obtained.

\[
B = \frac{mn-1}{m^2 n} \sum_{r} \mu_{4r} + \frac{3-mn}{m^2 n} \sum_{r} \sigma^4_{r} + \frac{2}{m} \sum_{r} \frac{\mu}{r^3} \text{ and}
\]
\[
\frac{1}{(mn)^2} \frac{1}{(n-1)m} (mn-1) \sum_{r}^{\infty} \frac{\mu_r}{r^2} + \frac{-mn+3}{(mn)^2m^2(n-1)} \sum_{r}^{\infty} \frac{\sigma_r^4}{r^2} + \frac{2}{(mn)^2(n-1)m^2} \sum_{r}^{\infty} \frac{\tau_r \mu_{3r}}{r}
\]

So \( \text{Var } \tilde{s}^2 = \text{Var } s^2 + O(1/[(mn)^2]) \). Therefore we have shown that for \( mn \) sufficiently large,

\[
\text{Var } \tilde{s}^2_{[m]n} \leq \text{Var } s^2_N
\]

2.4 Judgment Ordering with a Concomitant Variable

In this section, we discuss the use of a concomitant variable to judgment order the \( X \)'s. Suppose the observations of \( X \) cannot be easily ordered, but there is a related variable \( Y \) which is readily observable and which can be easily ordered. We use the related variable \( Y \) to acquire a ranked set sample of \( X \)'s. Again, \( mn \) samples of size \( m \) are chosen; in the first sample of size \( m \), the \( X \) associated with the smallest \( Y \) is measured; in the second sample, the \( X \) associated with the second smallest \( Y \) is measured, etc. As before, this cycle is repeated \( m \) times until \( mn \) \( X \)'s have been quantified.

This procedure is similar in theory to stratified sampling. That is, the "extra information" available from the \( Y \) variable is utilized to divide the population into \( m \) homogeneous strata which are as different as possible from one another.

Obviously, the benefit derived by using this method should depend on how much information \( Y \) yields about \( X \). If \( X \) and \( Y \) are independent, the estimators \( \hat{\mu}_{[m]n} \) and \( \hat{s}^2_{[m]n} \) should be equivalent in precision to \( \bar{X} \) and \( s^2 \). If \( |\rho| = (|\sigma_{xy}|)/(\sigma_x \sigma_y) = 1 \), then \( \hat{\mu}_{[m]n} \) and \( \hat{s}^2_{[m]n} \) should be equivalent in precision to \( \hat{\mu}_{(m)n} \) and \( \hat{s}^2_{(m)n} \), estimators which are calculated from a ranked set sample of actual order statistics of \( X \).
Let \( X_{[1:m]}, 1, \ldots, X_{[m:m]}, n \) be a ranked set sample selected on the basis of an ordered concomitant variable \( Y \). Furthermore, assume that the regression of \( X \) on \( Y \) is linear, and that \( \frac{X - \mu_X}{\sigma_X} \) and \( \frac{Y - \mu_Y}{\sigma_Y} \) both have c.d.f. \( F \). (These properties are characteristic, for example, of both the bivariate normal and bivariate Pareto distributions.) Then

\[
E(X|Y) = \mu_X + \frac{\rho \sigma_X}{\sigma_Y} (Y - \mu_Y)
\]

and

\[
\text{Var}(X|Y) = \sigma_X^2 (1 - \rho^2).
\]

Thus

\[
E X_{[r:m]} = E[E(X|Y_{(r:m)})] = \mu_X + \frac{\rho \sigma_X}{\sigma_Y} (E Y_{(r:m)} - \mu_Y),
\]

\[
= \mu_X + \rho \sigma_X \alpha_{(r:m)},
\]

where \( \alpha_{(r:m)} = E U_{(r:m)} \), \( U = \frac{Y - \mu_Y}{\sigma_Y} \) has c.d.f. \( F \), and

\[
\text{Var} X_{[r:m]} = E[\text{Var}(X|Y_{(r:m)})] + \text{Var}[E(X|Y_{(r:m)})]
\]

\[
= \sigma_X^2 (1 - \rho^2) + \frac{\rho^2 \sigma_Y^2}{\sigma_Y^2} \text{Var} Y_{(r:m)}
\]

\[
= \sigma_X^2 (1 - \rho^2) + \rho^2 \sigma_Y^2 \alpha_{(r:m)}^2
\]

where \( \sigma_{U_{(r:m)}}^2 = \text{Var} U_{(r:m)} \). Then

\[
\text{Var} \left( \frac{\hat{\mu}_{(m)n}}{\hat{\mu}_{[m:n]}^2} \right) = \frac{1}{m n} \left[ \sum_{r=1}^{m} \left( E X_{(r:m)} - \mu_X \right)^2 \right] - \frac{\sigma_X^2}{m n} \sum_{r=1}^{m} \tau_{(r:m)}^2
\]

\[
= \frac{\sigma_X^2}{m n} \sum_{r=1}^{m} \tau_{(r:m)}^2 - \frac{\sigma_Y^2}{m n} \sum_{r=1}^{m} \tau_{(r:m)}^2
\]

\[
= 1 - \frac{1}{m} \sum_{r=1}^{m} \tau_{(r:m)}^2
\]

\[
= 1 - \frac{\rho^2}{m} \sum_{r=1}^{m} \tau_{(r:m)}^2
\]
where \( \tau_{U(r:m)} = EU(r:m) - EU = \alpha(r:m) \). Likewise, we can compare the precision of \( \hat{\mu}_{[m]} \) with \( \bar{X} \) by examining

\[
\frac{\text{Var} \bar{X}}{\text{Var} \hat{\mu}_{[m]}} = \frac{1}{1 - \frac{\rho^2}{m} \sum_{r=1}^{m} \tau_{U(r:m)}^2}
\]

Table 2.4.1 displays some values of \( \text{Var} \hat{\mu}_{(m)}/\text{Var} \hat{\mu}_{[m]} \), a measure of the loss of precision obtained as a result of ranking by a concomitant variable \( Y \) rather than by ordering the \( X \)'s themselves, when \( (X,Y) \) has a bivariate normal distribution.

Table 2.4.2 displays some values of \( \text{Var} \bar{X}/\text{Var} \hat{\mu}_{[m]} \), a measure of the gain in precision obtained as a result of ranking by a concomitant variable \( Y \) over that obtained by random sampling, when \( (X,Y) \) has a bivariate normal distribution.

Similarly, we can examine the effects of ranked set sampling with a concomitant variable on \( s^2_{[m]} \), an estimator of \( \sigma^2_x \).

\[
\text{MSE} s^2_{(m)}/s^2_{[m]} = \frac{1}{mn} E(X-\mu)^4 - \frac{1}{m^2 n} \sum_r [\text{Var} X(r:m) + (EX(r:m) - \mu_x)^2]^2 + O(\frac{1}{mn})
\]

\[
= \frac{E(X-\mu)^4 - \frac{1}{m} \sum_r (\sigma^2_{U(r:m)} + \tau_{U(r:m)}^2)^2}{E(X-\mu)^4 - \frac{1}{m} \sum_r \rho^2 \sigma^2_{U(r:m)} + \rho^2 \tau_{U(r:m)}^2 + (1-\rho^2)^2} + O(\frac{1}{mn})
\]

Likewise, we see that

\[
\frac{\text{Var} s^2_{2mn}}{\text{Var} s^2_{[m]}} = \frac{1}{mn} E(X-\mu)^4 - \frac{1}{m^2 n} \sigma^4_x \sum_r [\rho^2 \sigma^2_{U(r:m)} + \rho^2 \tau_{U(r:m)}^2 + (1-\rho^2)^2] + O(\frac{1}{mn})
\]
Table 2.4.3 displays some values of \( \frac{\text{var } s_{(m)n}^2}{\text{var } s_{[m]n}^2} \), a measure of the loss in precision obtained as a result of ranking by a concomitant variable \( Y \) rather than by ordering the \( X \)'s themselves, when \( (X,Y) \) has a bivariate normal distribution and \( n \) is large. Table 2.4.4 displays some values of \( \frac{\text{var } s_{mn}^2}{\text{var } s_{[m]n}^2} \), a measure of the gain in precision obtained as a result of ranking by a concomitant variable \( Y \) over that obtained by random sampling, when \( (X,Y) \) has a bivariate normal distribution and \( n \) is large.

We see from examining the tables that the advantage of using ranked set sampling for estimating the variance is almost negligible if \( |\rho| < .75 \). Gains in precision of the estimator of \( \hat{\mu} \), however, are substantial for \( |\rho| \) as small as .5. Similarly, the losses in accuracy of estimators which result from ordering the concomitant variable \( Y \) rather than \( X \) itself are considerably more marked in the estimation of mean than in estimation of variance.

<table>
<thead>
<tr>
<th>( m )</th>
<th>( \rho = \pm .25 )</th>
<th>( \rho = \pm .5 )</th>
<th>( \rho = \pm .75 )</th>
<th>( \rho = \pm .9 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>.70</td>
<td>.74</td>
<td>.83</td>
<td>.92</td>
</tr>
<tr>
<td>4</td>
<td>.44</td>
<td>.50</td>
<td>.63</td>
<td>.80</td>
</tr>
<tr>
<td>6</td>
<td>.33</td>
<td>.38</td>
<td>.51</td>
<td>.71</td>
</tr>
<tr>
<td>8</td>
<td>.26</td>
<td>.31</td>
<td>.43</td>
<td>.64</td>
</tr>
<tr>
<td>10</td>
<td>.22</td>
<td>.26</td>
<td>.38</td>
<td>.58</td>
</tr>
</tbody>
</table>
### Table 2.4.2

<table>
<thead>
<tr>
<th>m</th>
<th>$\rho=\pm .25$</th>
<th>$\rho=\pm .5$</th>
<th>$\rho=\pm .75$</th>
<th>$\rho=\pm .9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.02</td>
<td>1.09</td>
<td>1.22</td>
<td>1.35</td>
</tr>
<tr>
<td>4</td>
<td>1.04</td>
<td>1.17</td>
<td>1.48</td>
<td>1.87</td>
</tr>
<tr>
<td>6</td>
<td>1.05</td>
<td>.121</td>
<td>1.63</td>
<td>2.25</td>
</tr>
<tr>
<td>8</td>
<td>1.05</td>
<td>1.23</td>
<td>1.73</td>
<td>2.55</td>
</tr>
<tr>
<td>10</td>
<td>1.05</td>
<td>1.25</td>
<td>1.80</td>
<td>2.79</td>
</tr>
</tbody>
</table>

### Table 2.4.3

<table>
<thead>
<tr>
<th>m</th>
<th>$\rho=\pm .25$</th>
<th>$\rho=\pm .5$</th>
<th>$\rho=\pm .75$</th>
<th>$\rho=\pm .9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>4</td>
<td>.85</td>
<td>.86</td>
<td>.89</td>
<td>.94</td>
</tr>
<tr>
<td>6</td>
<td>.73</td>
<td>.74</td>
<td>.79</td>
<td>.89</td>
</tr>
<tr>
<td>8</td>
<td>.64</td>
<td>.65</td>
<td>.72</td>
<td>.84</td>
</tr>
<tr>
<td>10</td>
<td>.57</td>
<td>.58</td>
<td>.66</td>
<td>.80</td>
</tr>
</tbody>
</table>

### Table 2.4.4

<table>
<thead>
<tr>
<th>m</th>
<th>$\rho=\pm .25$</th>
<th>$\rho=\pm .5$</th>
<th>$\rho=\pm .75$</th>
<th>$\rho=\pm .9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>4</td>
<td>1.00</td>
<td>1.01</td>
<td>1.05</td>
<td>1.11</td>
</tr>
<tr>
<td>6</td>
<td>1.00</td>
<td>1.02</td>
<td>1.10</td>
<td>1.22</td>
</tr>
<tr>
<td>8</td>
<td>1.00</td>
<td>1.02</td>
<td>1.13</td>
<td>1.31</td>
</tr>
<tr>
<td>10</td>
<td>1.00</td>
<td>1.03</td>
<td>1.16</td>
<td>1.40</td>
</tr>
</tbody>
</table>
3. ESTIMATION OF PARAMETERS OF ASSOCIATION

In this chapter, we will examine the process of estimation of the correlation coefficient of a bivariate normal random vector \((X, Y)\). We will show that a modified form of ranked set sampling can improve the estimator of \(\rho\) over that from a random sample of comparable size.

3.1.1 Estimation of \(\rho\) with All Parameters Known

The simplest, but most unrealistic, situation is that in which \(\mu_X, \mu_Y, \sigma_X^2, \text{ and } \sigma_Y^2\) are assumed to be known. Suppose the pairs \((X, Y)\) of a sample of size \(m\) are ordered on the basis of the rank of the \(Y\)'s. Then the pair \((X_{(r:m)}, Y_{(r:m)})\), the \(r\)-th largest \(Y\) and its concomitant \(X\), have joint p.d.f.

\[
f_{(r:m)}(x, y) = f_{(r:m)}(y)f(x | y) = \frac{m!}{(r-1)!(m-r)!} F_{r-1}^{-1}(y)[1-F(y)]^{m-r}f(x, y).
\]

(3.1)

Suppose we choose, in the usual way, a ranked set sample of \(Y\)'s of size \(mn=N\), and measure the pairs consisting of those \(Y\)'s and their concomitant \(X\)'s. (Note the similarity of this procedure and that discussed in Section 2.4.) A maximum likelihood estimator of \(\rho, \hat{\rho}\), can be found by solving the equation

\[
\frac{\partial}{\partial \rho} \ln \left\{ \prod_{r=1}^{m} \prod_{i=1}^{n} f_{(r:m)}(X_{ri}, Y_{ri}) \right\} = 0
\]

where \(f_{(r:m)}(x, y)\) is given in (3.1).
But this is equivalent to solving the equation

$$\frac{\partial}{\partial \rho} \ln \left[ \prod_{r=1}^{m} \prod_{i=1}^{n} f(X_{ri}, Y_{ri}) \right] = 0$$

(3.2)

where $f(x,y)$ is the p.d.f. of a bivariate normal distribution.

This is because $\rho$ does not appear in any factor of $f_{(r:m)}(x,y)$ except $f(x,y)$. Therefore the MLE of $\rho$ from a ranked set sample is calculated in exactly the same way as that from a random sample.

In fact, even if ranking errors do occur in the ordering of the $Y$'s, the MLE of $\rho$ will still be calculated in the same way, since the joint p.d.f. of $Y_{[r:m]}$ (the $r$-th judgment ordered $Y$) and its concomitant $X$ is

$$f_{[r:m]}(x,y) = f_{[r:m]}(y|x) f(x|y)$$

where $f_{[r:m]}(y)$ is the p.d.f. of the $r$-th judgment order statistic and $f(x|y)$ is the usual conditional normal p.d.f. Note here also that regardless of the distribution of the random vector $(X,Y)$, the ranked set sampling MLE of $\rho$ when $\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2$ are known is always calculated in the same way as the MLE of $\rho$ from a random sample.

Equation (3.2) is equivalent to the cubic equation

$$g(\rho) = m n \sigma_Y^2 \rho (1-\rho^2) + \sigma_X \sigma_Y (1+\rho^2) \sum_{r=1}^{m} \sum_{i=1}^{n} (X_{ri} - \mu_X)(Y_{ri} - \mu_Y)$$

$$- \rho \left[ \sigma_Y^2 \sum_{r=1}^{m} \sum_{i=1}^{n} (X_{ri} - \mu_X)^2 + \sigma_X^2 \sum_{r=1}^{m} \sum_{i=1}^{n} (Y_{ri} - \mu_Y)^2 \right] = 0.$$  

There must always be a root of this equation between $-1$ and $+1$. 


since
\[ g(-1) = \sum_{r=1}^{m} \sum_{i=1}^{n} \left[ \sigma_{y} (X_{ri} - \mu_{x}) + \sigma_{x} (Y_{ri} - \mu_{y}) \right]^{2} \geq 0 \]
and
\[ g(+1) = -\sum_{r=1}^{m} \sum_{i=1}^{n} \left[ \sigma_{y} (X_{ri} - \mu_{x}) - \sigma_{x} (Y_{ri} - \mu_{y}) \right]^{2} \leq 0 . \]

Furthermore, note that there is a root \( \geq 0 \) in (-1,1) according to whether \( \sum_{r} \sum_{i} (X_{ri} - \mu_{x}) (Y_{ri} - \mu_{y}) \geq 0 \). By adapting Kendall and Stuart's argument ([13], p. 40), one can show that as \( m \to \infty \) or \( n \to \infty \), the probability of more than one root of the likelihood equation existing will tend to 0. If there are three roots of \( g(\rho) = 0 \), there will be two real turning points satisfying the quadratic equation \( g'(\rho) = 0 \). But this equation has two roots iff
\[ \left[ \frac{1}{mn} \sum_{r} \sum_{i} \left( \frac{X_{ri} - \mu_{x}}{\sigma_{x}} \right)^{2} \left( \frac{Y_{ri} - \mu_{y}}{\sigma_{y}} \right)^{2} \right] - \frac{3}{mn} \left[ \sum_{r} \sum_{i} \left( \frac{X_{ri} - \mu_{x}}{\sigma_{x}} \right)^{2} + \sum_{r} \sum_{i} \left( \frac{Y_{ri} - \mu_{y}}{\sigma_{y}} \right)^{2} - 1 \right] > 0 . \]

Using Chebyshev's inequality, we see that
\[ \Pr \left[ \left| \frac{1}{mn} \sum_{r} \sum_{i} U_{1,ri} U_{2,ri} - \rho^{2} \right| > \epsilon \right] \leq \frac{E(U_{1}U_{2})^{2}}{m \epsilon^{2}} \to 0 \quad \text{as} \quad m \to \infty \]
or as \( n \to \infty \), since
\[ E \frac{1}{mn} \sum_{r} \sum_{i} U_{1,ri} U_{2,ri} = \frac{n}{mn} E \sum_{r} U_{1,r} U_{2,r} = \frac{1}{m} m E U_{1} U_{2} = \rho^{2} \]
and
\[ \text{Var} \left( \frac{1}{mn} \sum_{r} \sum_{i} U_{1,ri} U_{2,ri} \right) = \frac{n}{mn} \frac{1}{2} \sum_{r} \left[ E U_{1}^{2} U_{2}^{2} - (E U_{1} U_{2})^{2} \right] \leq \frac{1}{2} \frac{n}{m} \frac{1}{2} \sum_{r} E U_{1}^{2} U_{2}^{2} = \frac{1}{mn} E(U_{1}U_{2})^{2} \]
where \( (U_{1}, U_{2}) \sim N \left( 0, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right) \).

Similarly,
\[ \Pr \left[ \left| \frac{1}{mn} \sum_{r} \sum_{i} E U_{k,ri}^{2} - 1 \right| > \epsilon \right] \leq \frac{3}{m \epsilon^{2}} \to 0 \quad \text{for} \quad k=1,2 \quad \text{as} \quad m \to \infty \quad \text{or} \quad n \to \infty . \]
since
\[ E \frac{1}{mn} \sum_r \sum_i U_{k,ri}^2 = \frac{1}{m} E \sum_r U_{k,r}^2 = \frac{1}{m} m E U_{k,r}^2 = 1, \quad k=1,2, \]
and
\[ \text{Var}(\frac{1}{mn} \sum_r \sum_i U_{k,ri}^2) = \frac{n}{2m} \sum_r \left[ E U_{k,r}^4 - (E U_{k,r}^2)^2 \right] \leq \frac{1}{2mn} \sum_r E U_{k,r}^4 = \frac{3}{mn}. \]

Therefore, the left hand side of (3.3) converges in probability to
\[ \rho^2 - 3(1+1-1) = \rho^2 - 3 < 0. \]

So for \( m \) or \( n \) sufficiently large, there cannot be three real roots of \( g(\rho) = 0 \), so there is a unique root to the likelihood equation.

Even though the estimates from a ranked set and from a random sample are calculated in the same way, it is not necessarily true that they have the same asymptotic variance. To simplify calculations of Fisher Information, \( I_N(\rho) \), we assume that \( \mu_x = 0 \), \( \mu_y = 0 \), \( \sigma_x = 1 \), \( \sigma_y = 1 \). In this case
\[ I_N(\rho) = -E \left[ \frac{\partial^2 \ln L}{\partial \rho^2} \right] \]
and
\[ \frac{\partial^2 \ln L}{\partial \rho^2} = \frac{mn}{(1-\rho^2)^2} + \frac{2mnp^2}{(1-\rho^2)^2} + \frac{4\rho}{(1-\rho^2)^2} \sum_r \sum_i x_{ri} y_{ri} \]
\[ - \sum_r \sum_i \left( X_{ri}^2 + 2\rho X_{ri} Y_{ri} + Y_{ri}^2 \right) \left[ \frac{1}{(1-\rho^2)^3} + \frac{4\rho^2}{(1-\rho^2)^3} \right]. \]

If the \( N \) observations are simply a random sample, then
\[ I_N(\rho) = \frac{mn}{(1-\rho^2)^2} - \frac{2mnp^2}{(1-\rho^2)^2} = \frac{mn}{(1-\rho^2)^2} \left[ \frac{1+\rho^2}{1-\rho^2} \right]. \]

(3.4)
This tells us that the MLE \( \hat{\rho}_n \) from a random sample of size \( mn = N \), where \( \mu_x, \mu_y, \sigma_x^2 \), and \( \sigma_y^2 \) are all known, has asymptotic variance \( \frac{((1-\rho^2)^2)}{N(1+\rho^2)} \). Now suppose the \( mn \) observations form a ranked set sample. Notice that

\[
\text{EX}_{[r:m]} Y_{[r:m]} = \int_{-\infty}^{\infty} xy f(r:m)(y) f(x|y) dy \, dx
\]

\[
= \int_{-\infty}^{\infty} y f(r:m)(y) \int_{-\infty}^{\infty} x f(x|y) dx \, dy
\]

\[
= \int_{-\infty}^{\infty} y f(r:m)(y) \rho y \, dy = \rho \text{EY}^2_{[r:m]}
\]

and

\[
\text{EX}^2_{[r:m]} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^2 f(r:m)(y) f(x|y) dy \, dx
\]

\[
= 1 - \rho^2 (1-\text{EY}^2_{[r:m]}).
\]

Then

\[
I^{*}_{(m)n}(\rho) = -E \frac{\partial^2 \ln L}{\partial \rho^2} = \frac{mn}{(1-\rho^2)} \left[ \frac{1+\rho^2}{1-\rho^2} \right].
\]

This tells us that the MLE \( \hat{\rho}^{*}_{(m)n} \) from a ranked set sample of size \( mn = N \), where \( \mu_x, \mu_y, \sigma_x^2 \), and \( \sigma_y^2 \) are known, has asymptotic variance \( \frac{((1-\rho^2)^2)}{mn(1+\rho^2)} \). So far as asymptotic variance is concerned, the process of ranked set sampling does not improve the MLE of \( \rho \).

Suppose now, however, that we modify the ranked set sampling procedure so that instead of choosing \( n \) first order statistics, \( n \) second order statistics, etc., where all \( mn \) are independent, we include in our sample only extreme order statistics. That is, from each group of \( m \) judgment ordered random variables, we choose
for quantification \( Y_{(1:m)} \) and its concomitant \( X \) and include \( N \)

vectors in our sample. (Note that we could just as appropriately

choose all smallest \( Y \)'s, \( Y_{(1:m)} \), or even choose \( Y_{(1:m)} \) and \( Y_{(m:m)} \)

alternately. The reason is that in this case, \( \mu_X, \mu_Y, \sigma^2_X, \) and \( \sigma^2_Y \)

are known, so the probability structure of \( (Y_{(m:m)} - \mu_Y)/\sigma_Y \) is

known and is the same as that of \( -(Y_{(1:m)} - \mu_Y)/\sigma_Y \). Later we will

consider the case where \( \mu_X \) and \( \sigma^2_X \) are unknown. Then we will see

that we are better served to alternate \( Y_{(1:m)} \) and \( Y_{(m:m)} \).

When we sample in this way (which is another form of modified

ranked set sampling), the MLE of \( \rho, \hat{\rho}_{(m)N}^\Delta \), is still a root of

g(\rho) = 0, but the asymptotic variance of the estimator is the

reciprocal of

\[
I_{(m)N}(\rho) = \frac{N}{(1-\rho^2)^2} \left[ EU^2_{(m:m)} + \frac{2\rho^2}{(1-\rho^2)} \right] \geq I_N(\rho), \tag{3.5}
\]

where \( U = (Y - \mu_Y)/\sigma_Y \), since \( EU^2_{(m:m)} \geq 1 \). We note here that the choice

of the extreme order statistics for inclusion in the modified ranked

set sample was made because it maximizes the Fisher Information

over all possible choices of order statistics.

Table 3.1 gives some values of the ratio

\[
e_N(\hat{\rho}, \hat{\rho}_{(m)}^\Delta) = \frac{I_{(m)N}(\rho)}{I_N(\rho)}
\]

for particular values of \( m \) and \( \rho \). These values are a measure of the

amount of improvement an estimator based on a modified ranked set

sample allows over that from a random sample. From Cramér

([3], p. 376) one can see that \( EY^2_{(m:m)} = 0(\ln m) \). Consequently,

\[e_N(\hat{\rho}, \hat{\rho}_{(m)}^\Delta) \to \infty \text{ as } m \to \infty.\]
Table 3.1
Comparison of asymptotic variances of $\hat{\rho}$ and $\hat{\rho}^\Delta$

\[ e_N(\hat{\rho}, \hat{\rho}^\Delta_{(m)}) = I^\Delta_{(m)N}(\rho)/I_N(\rho) \]

<table>
<thead>
<tr>
<th>m</th>
<th>$\rho = 0$</th>
<th>$\rho = \pm 0.5$</th>
<th>$\rho = \pm 0.75$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1.55</td>
<td>1.33</td>
<td>1.15</td>
</tr>
<tr>
<td>6</td>
<td>2.02</td>
<td>1.65</td>
<td>1.29</td>
</tr>
<tr>
<td>8</td>
<td>2.40</td>
<td>1.84</td>
<td>1.39</td>
</tr>
<tr>
<td>10</td>
<td>2.71</td>
<td>2.03</td>
<td>1.48</td>
</tr>
</tbody>
</table>

3.1.2 Confidence Intervals for $\rho$ with All Parameters Known

Suppose a confidence interval for $\rho$ is desired rather than a point estimate. Large sample theory allows us to find an approximate confidence interval based on the asymptotic normality of maximum likelihood estimators. (Theorem B, Appendix). So if the sample size $N$ from which $\hat{\rho}^\Delta_{(m)N}$ of Section 3.1.1 is calculated is "sufficiently large," a 100(1-$\alpha$)% confidence interval for $\rho$ would be

\[
\left[ \hat{\rho}^\Delta_{(m)N} - z_{1-\alpha/2} \sqrt{\frac{1}{N} I_{(m)N}(\hat{\rho}^\Delta)} , \hat{\rho}^\Delta_{(m)N} + z_{1-\alpha/2} \sqrt{\frac{1}{N} I_{(m)N}(\hat{\rho}^\Delta)} \right]
\]

\[
= \left[ \hat{\rho}^\Delta_{(m)N} - z_{1-\alpha/2} \sqrt{\frac{1}{N} \left( \frac{1-\hat{\rho}^2_{(m)N}}{\left(1-\hat{\rho}^\Delta_{(m)N}^2\right)^2} \right)} , \hat{\rho}^\Delta_{(m)N} + z_{1-\alpha/2} \sqrt{\frac{1}{N} \left( \frac{1-\hat{\rho}^2_{(m)N}}{\left(1-\hat{\rho}^\Delta_{(m)N}^2\right)^2} \right)} \right]^{1/2}
\]

\[
\hat{\rho}^\Delta_{(m)N} + z_{1-\alpha/2} \sqrt{\frac{1}{N} \left( \frac{1-\hat{\rho}^2_{(m)N}}{\left(1-\hat{\rho}^\Delta_{(m)N}^2\right)^2} \right)}
\]

(3.6)
where $U \sim N(0,1)$. But what can be done to obtain an exact confidence interval?

We choose a modified ranked set sample as required for $\hat{\rho}_N^\Delta$ above. Note that the random variable $X_{m:m}|Y_{m:m}$ has p.d.f.

$$\frac{f_{m:m}(x,y)}{f_{m:m}(y)}$$

where $f_{m:m}(x,y)$ is given by (3.1) and $f_{m:m}(y)$ is the p.d.f. of $Y_{m:m}$.

$$= \frac{1}{\sqrt{1-\rho^2}} \phi \left( \frac{x-\rho y}{\sqrt{1-\rho^2}} \right)$$

where $\phi(u) = (1/\sqrt{2\pi}) e^{-u^2/2}$,

assuming once again that $\mu_x = \mu_y = 0$, $\sigma^2_x = \sigma^2_y = 1$. Let $T_{m:m},i = X_{m:m},i - \rho Y_{m:m},i$. Then

$$T_{m:m},i | Y_{m:m},i \sim N(0,1-\rho^2).$$

Since this distribution does not depend on $Y_{m:m}$, it follows that

$$T_{m:m},i \sim N(0,1-\rho^2),$$

and since the $T_{m:m},i$'s are independent, we have

$$\frac{1}{N} \sum_{i=1}^{N} T_{m:m},i \sim N(0,N(1-\rho^2)).$$

Thus we can use the statistic

$$T_{(m)N} = \frac{\sum_{i=1}^{N} (X_{m:m},i - \rho Y_{m:m},i)}{\sqrt{N(1-\rho^2)}} \sim N(0,1)$$

(3.7)

to construct a confidence interval for $\rho$. 

\[ \]
Consider the event

\[ A = \{ -z_{1-\alpha/2} \leq T_{(m)N} \leq z_{1-\alpha/2} \} . \]

From (3.7), we know \( \Pr(A) = 1-\alpha \), and \( A \) is equivalent to

\[ \left\{ \left( \frac{1}{\rho} \sum_{i=1}^{N} \frac{Y_{(m):i}}{\rho} - \frac{1}{\rho} \sum_{i=1}^{N} \frac{Y_{(m):i}}{\rho} \right)^2 \leq N(1-\rho^2)z_{1-\alpha/2}^2 \right\}. \]

From this form of the confidence interval, we see that it cannot contain values of \( \rho \) s.t. \( |\rho| > 1 \). This, in turn, may be rewritten as

\[ \left\{ (N\rho_{1-\alpha/2}^2 + N^2\tilde{Y}_{m}^2)\rho^2 - 2N\frac{\tilde{X}_m\tilde{Y}_m}{\rho} + (N^2\tilde{X}_m^2 - N^2_{1-\alpha/2}) \leq 0 \right\} \quad (3.8) \]

where \( \tilde{X}_m = \frac{1}{N} \sum_{i=1}^{N} X_{(m):i} \) and \( \tilde{Y}_m = \frac{1}{N} \sum_{i=1}^{N} Y_{(m):i} \). The left hand side of (3.8) is a parabola in \( \rho \) of the form \( a\rho^2 - 2b\rho + c \). Since

\[ a = N\rho_{1-\alpha/2}^2 + N^2\tilde{Y}_m^2 > 0, \]

we know that the only turning point of this parabola is a minimum, and it occurs at \( \rho = b/a = \frac{N^2\tilde{X}_m\tilde{Y}_m}{(N\rho_{1-\alpha/2}^2 + N^2\tilde{Y}_m^2)} \), with minimum value \( \frac{ac-b^2}{a} \). But if \( \frac{ac-b^2}{a} > 0 \), or equivalently, in our case, \( ac-b^2 > 0 \), there will be no values of \( \rho \) which satisfy (3.8). The probability of this occurring is

\[ \Pr[ac-b^2 > 0] = \Pr[(N^2\tilde{Y}_m^2 + N^2_{1-\alpha/2})(N^2_{1-\alpha/2} - N^2\tilde{X}_m^2 - N^2\tilde{Y}_m^2) - N^2\tilde{X}_m^2 - N^2\tilde{Y}_m^2 > 0] \]

\[ = \Pr[N^2_{1-\alpha/2}(N\tilde{X}_m^2 - N\tilde{Y}_m^2 - \tilde{X}_m^2) > 0] \]

\[ = \Pr[\tilde{X}_m^2 - \tilde{Y}_m^2 > z_{1-\alpha/2}^2/N]. \]

We now show that this probability converges to 0 with increasing \( N \) or \( m \). (Assume that \( \rho^2 < 1 \), since for \( \rho^2 = 1 \), \( \Pr[\tilde{X}_m^2 - \tilde{Y}_m^2 = 0] = 1 \),
so \( \Pr[\overline{x}_m^2 - \overline{y}_m^2 > z_1^{2(1-\alpha)/N}] = 0 \) \( \forall N, m \). First note that the hyperbola
\[
x^2 - y^2 = z_1^{2(1-\alpha)/N}
\]
has as asymptotes the lines \( x = y \) and \( x = -y \). Therefore

\[
\Pr[\overline{x}_m^2 - \overline{y}_m^2 > z_1^{2(1-\alpha)/N}] = \Pr[(\overline{x}_m, \overline{y}_m) \in A_1] + \Pr[(\overline{x}_m, \overline{y}_m) \in A_2]
\]

\[\leq 1 - \Pr[(\overline{x}_m, \overline{y}_m) \in B_1] \]

\[= 1 - \Pr[-\overline{y}_m < \overline{x}_m < \overline{y}_m \text{ and } \overline{y}_m > 0] \]

\[= 1 - \{\Pr[-\overline{y}_m < \overline{x}_m < \overline{y}_m \text{ and } \overline{y}_m > 0 | |\overline{y}_m - \mu_m| < \varepsilon\} \cdot \Pr[|\overline{y}_m - \mu_m| < \varepsilon] \]

\[+ \Pr[-\overline{y}_m < \overline{x}_m < \overline{y}_m \text{ and } \overline{y}_m > 0 | |\overline{y}_m - \mu_m| > \varepsilon\} \cdot \Pr[|\overline{y}_m - \mu_m| > \varepsilon]\}

\[\leq 1 - \Pr[-\overline{y}_m < \overline{x}_m < \overline{y}_m \text{ and } \overline{y}_m > 0 | |\overline{y}_m - \mu_m| < \varepsilon\} \cdot \Pr[|\overline{y}_m - \mu_m| < \varepsilon] \]

\[\leq 1 - \Pr[-\mu_m + \varepsilon < \overline{x}_m < \mu_m - \varepsilon] \cdot \Pr[|\overline{y}_m - \mu_m| < \varepsilon] \text{ where } \varepsilon < \mu_m \]

\[= 1 - \Pr[-\mu_m (1+\rho) + \varepsilon < \overline{x}_m - \rho \mu_m < \mu_m (1-\rho) - \varepsilon] \cdot \Pr[|\overline{y}_m - \mu_m| < \varepsilon] \]

\[ \]
\[
\left\{ \begin{array}{ll}
1 - \Pr(\left| \bar{X}_m - \rho \mu_m \right| < \mu_m (1 + \rho) - \epsilon) \cdot \Pr(\left| \bar{Y}_m - \mu_m \right| < \epsilon) \quad & \text{if } \rho \leq 0 \\
1 - \Pr(\left| \bar{X}_m - \rho \mu_m \right| < \mu_m (1 - \rho) - \epsilon) \cdot \Pr(\left| \bar{Y}_m - \mu_m \right| < \epsilon) \quad & \text{if } \rho > 0
\end{array} \right.
\]

\[
\leq 1 - \left[ 1 - \frac{\text{Var} \bar{X}_m}{(\mu_{(m:m)}(1+\rho)-\epsilon)^2} \right] \cdot \left[ 1 - \frac{\text{Var} \bar{Y}_m}{\epsilon^2} \right]
\]

if \( \rho \leq 0 \)

by Chebyshev's inequality

\[
= 1 - \left[ 1 - \frac{1 + \rho^2 (\sigma^2_{(m:m)} - 1)}{N(\mu_{(m:m)}(1+\rho)-\epsilon)^2} \right] \left[ 1 - \frac{\sigma^2_{(m:m)}}{\epsilon^2} \right] \rightarrow 0 \quad \text{as } m \rightarrow \infty
\]

or as \( N \rightarrow \infty \)

since \([3]\) \( \mu_{(m:m)} = \text{E}U_{(m:m)} = O(\sqrt{n}m) \) and \( \sigma^2_{(m:m)} = \text{Var} U_{(m:m)} = O(\frac{1}{\sqrt{n}m}) \).

If there are values of \( \rho \) which satisfy (3.8) (i.e., if \( ac - b^2 \leq 0 \)), then the 100(1-\( \alpha \))% confidence interval is

\[
\left( \frac{b - \sqrt{b^2 - ac}}{a}, \quad \frac{b + \sqrt{b^2 - ac}}{a} \right)
\]

\[
= \left[ \bar{Y}_m - \frac{z_{1-\alpha/2}}{\sqrt{N} \frac{\bar{Y}_m - X_m^2}{m} + z_{1-\alpha/2}}, \quad \frac{\bar{Y}_m + z_{1-\alpha/2}}{\sqrt{N} \frac{\bar{Y}_m - X_m^2}{m} + z_{1-\alpha/2}} \right].
\]

(3.9)

The interval of (3.9) is an exact 100(1-\( \alpha \))% confidence interval for \( \rho \) for any sample size, but (3.6) gives an interval which is a valid 100(1-\( \alpha \))% confidence interval only when \( N \) is large. Consequently, it is reasonable to compare the lengths of the two confidence intervals only for large \( N \). The length of the confidence interval derived from the properties of the MLE is, from (3.6),

\[
L_{(m)1} = \frac{2 z_{1-\alpha/2} \sqrt{1 - \rho^2}}{\sqrt{N \left[ \mu^2_{(m:m)} + \sigma^2_{(m:m)} + \frac{2 \Delta^2}{(m)} \right]}}
\]
where $\mu_{(m:m)} = E(U_{(m:m)})$, $\sigma^2_{(m:m)} = Var(U_{(m:m)})$, $U \sim N(0,1)$. The length of the exact confidence interval is, from (3.9),

$$L(m) = 2z_{1-\alpha/2} \sqrt{\frac{2}{N} \left( \frac{1}{m} \sum x_i^2 - \frac{1}{m} \bar{x}_m^2 \right) + \frac{2}{m} \left( z_{1-\alpha/2} \right)^2 \left( 1 - \frac{1}{m} \right)^{1/2} + \frac{2}{m+1} \bar{y}_m^2}.$$ 

We know that (under certain regularity conditions which are fulfilled here), a MLE converges a.s. to the parameter it estimates, and that $\frac{1}{m} \sum x_i/N \rightarrow^N EX$ when $X_1, \ldots, X_N$ are iid by SLLN. Therefore

$$\hat{\rho}_{(m)} \rightarrow^N \rho$$

$$\bar{Y}_m \rightarrow^N EY_{(m:m)} = \mu_{(m:m)}$$

$$\bar{X}_m \rightarrow^N EX_{[m:m]} = \rho \mu_{(m:m)}$$

and

$$\sqrt{N} L_1 \rightarrow^N 2z_{1-\alpha/2} \sqrt{1 - \rho^2} \sqrt{\frac{2}{m} \mu^2_{(m:m)} + \frac{2}{m} \sigma^2_{(m:m)} + \frac{2\rho^2}{1 - \rho^2}},$$

$$\sqrt{N} L_2 \rightarrow^N 2z_{1-\alpha/2} \sqrt{1 - \rho^2} \sqrt{\frac{2}{m} \mu^2_{(m:m)}}, \quad \text{all as } N \rightarrow \infty.$$  

Let the ratio of these limits

$$\frac{\mu_{(m:m)}}{\sqrt{\frac{2}{m} \mu^2_{(m:m)} + \frac{2}{m} \sigma^2_{(m:m)} + \frac{2\rho^2}{1 - \rho^2}}}$$

be denoted by $R(m, \rho)$. It is clear that $R(m, \rho) < 1$ for any values of $m$ and $\rho$. But since

$$\mu_{(m:m)} = O(\sqrt{\ln m}) \quad \text{and} \quad \sigma^2_{(m:m)} = O\left(\frac{1}{\ln m}\right),$$

we know that

$$\lim_{m \rightarrow \infty} R(m, \rho) = 1 \quad \text{if} \quad \rho^2 \neq 1.$$
So the limiting length of the exact confidence interval approaches that of the MLE-based confidence interval as \( m \) increases. Some values of \( R(m, \rho) \) are tabulated in Table 3.2. This table shows that \( R(m, \rho) \) approaches 1 fairly rapidly (with \( m \)), especially for small values of \( |\rho| \).

If errors in ranking of the \( Y \)'s occur, this method will still produce an exact \( 100(1-\alpha)\% \) confidence interval since the p.d.f. of \( X_{[m:m]} | Y_{[m:m]} \) will still be

\[
\frac{f_{[m:m]}(x, y)}{f_{[m:m]}(y)} = \frac{f(x | y) f_{[m:m]}(y)}{f_{[m:m]}(y)} = \frac{1}{\sqrt{1-\rho^2}} \phi \left( \frac{x-\rho y}{\sqrt{1-\rho^2}} \right)
\]

where \( f_{[m:m]}(y) \) is the p.d.f. of the largest judgment order statistic. However, if the ordering is no better than random, \( \sqrt{N} L_2 \to \infty \) as \( N \to \infty \) since \( \mu_{[m:m]} = EU = 0 \). In order to use the MLE-based confidence interval for the case where ranking errors in the \( Y \)'s occur, one must first estimate \( \mu_{[m:m]}^2 + \sigma_{[m:m]}^2 \) with an estimator such as

\[
\frac{1}{N} \sum_{i=1}^{N} Y_{[m:m], i}^2 \quad \text{a.s.} \quad \mu_{[m:m]}^2 + \sigma_{[m:m]}^2.
\]

Then the interval will also be a valid confidence interval for large samples, since \( \hat{\rho}_{(m)N}^\Delta \) will still be the MLE of \( \rho \), as noted in Section 3.1.1. In this case, if ordering is no better than random,

\[
\sqrt{N} L_2 \to 2 z_{1-\alpha/2} \sqrt{1-\rho^2} \frac{1-\rho^2}{1+\rho^2},
\]
so this method will still give us a useful confidence interval.

Table 3.2

<table>
<thead>
<tr>
<th>m</th>
<th>( \rho=0 )</th>
<th>( \rho=\pm.25 )</th>
<th>( \rho=\pm.5 )</th>
<th>( \rho=\pm.75 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>.564</td>
<td>.530</td>
<td>.437</td>
<td>.299</td>
</tr>
<tr>
<td>4</td>
<td>.826</td>
<td>.793</td>
<td>.691</td>
<td>.507</td>
</tr>
<tr>
<td>6</td>
<td>.891</td>
<td>.863</td>
<td>.773</td>
<td>.591</td>
</tr>
<tr>
<td>8</td>
<td>.919</td>
<td>.894</td>
<td>.813</td>
<td>.639</td>
</tr>
<tr>
<td>10</td>
<td>.934</td>
<td>.912</td>
<td>.837</td>
<td>.669</td>
</tr>
</tbody>
</table>

There is another approach to finding an exact confidence interval for \( \rho \) when \( \mu_x, \mu_y, \sigma^2_x, \) and \( \sigma^2_y \) are all known. It too, is based on the fact that \( E[X|Y] \) is a linear function of \( Y \) for the bivariate normal distribution. Suppose here again that our data is a modified ranked set sample of size \( N \). Consider the random variable \( \eta_{(m)N} \) where \( \eta_{(m)N} \) is the number of the \( N \) observations which fall in the region \( X_{[m:m]} - \rho Y_{[m:m]} \leq 0 \). (Here again, we assume that \( \mu_x = \mu_y = 0, \sigma^2_x = \sigma^2_y = 1 \).)
Then $n_{(m)N} \sim \text{Bin}(N, \frac{1}{2})$, since

$$
\Pr[X_{[m:m]} - \rho Y_{(m:m)} \leq 0] = \int_{-\infty}^{\infty} f_{(m:m)}(y) \int_{-\infty}^{y} \frac{1}{\sqrt{1-\rho^2}} \phi \left( \frac{x - \rho y}{\sqrt{1-\rho^2}} \right) \, dx \, dy
$$

$$
= \phi(0) = \frac{1}{2}.
$$

(Note that this does not hold true for $\rho = \pm 1$, since in that case

$$
\Pr[X_{[m:m]} - \rho Y_{(m:m)} = 0] = 1.
$$

So by randomizing, one can find a $100(1-\alpha)\%$ confidence interval

$$
[n_* < n_{(m)N} < n^*]
$$

based on the binomial distribution of $n$; or if $N$ is moderately large, the normal approximation to the binomial can be employed to yield

$$
n_* = \frac{N}{2} - z_{1-\alpha/2} \sqrt{N/2}, \quad n^* = \frac{N}{2} + z_{1-\alpha/2} \sqrt{N/2}.
$$

The implementation of this confidence interval procedure can be handled graphically by simply counting the number of observations to the left of the line $x = \rho y$ as its slope changes from $\rho = -1$ to $\rho = +1$. It can also be carried out analytically by calculating for each pair $(X_{[m:m]}, Y_{(m:m)})$ the interval of $\rho$'s satisfying

$$
X_{[m:m]} - \rho Y_{(m:m)} \leq 0.
$$

These intervals are:

$$
\rho \geq \frac{X_{[m:m]}}{Y_{(m:m)}} \quad \text{for} \quad Y_{(m:m)} > 0,
$$

$$
\rho \leq \frac{X_{[m:m]}}{Y_{(m:m)}} \quad \text{for} \quad Y_{(m:m)} < 0,
$$

all $\rho$'s if $Y_{(m:m)} = 0$ and $X_{[m:m]} \leq 0$,

no $\rho$'s if $Y_{(m:m)} = 0$ and $X_{[m:m]} > 0$.  

Then the 100(1-\alpha)\% confidence interval $I_{1-\alpha}$ is the set of $\rho$'s for which $\eta_{(m)N}$, the number of intervals covering $\rho$, lies between $n_*$ and $n^*$.

Now suppose $(X,Y) \sim N_2\left(0, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right)$ and $\rho_1 \neq \rho$. We will show:

(a) $\Pr[\rho_1 \in I_{1-\alpha}] \to 0$ as $m \to \infty$

(b) $\Pr[\rho_1 \in I_{1-\alpha}] \to 0$ as $N \to \infty$.

Proof:

(a) $\Pr[\rho_1 \in I_{1-\alpha}] = \Pr\left[ \frac{N}{2} - z_{1-\alpha/2} \sqrt{N}/2 < \eta_{(m)N} < \frac{N}{2} + z_{1-\alpha/2} \sqrt{N}/2 \right]$.

But

$\eta_{(m)N} \sim N(Np_1, Np_1 q_1)$

where

$p_1 = \int \int_{-\infty}^{\infty} f_{(m:m)}(y) \frac{1}{\sqrt{1-\rho^2}} \phi\left(\frac{x-\rho y}{\sqrt{1-\rho^2}}\right) dx \, dy$

$= \int_{-\infty}^{\infty} f_{(m:m)}(y) \phi\left(\frac{y(\rho_1-\rho)}{\sqrt{1-\rho^2}}\right) dy$.

By Chebyshev's inequality, we know that

$\Pr[|Y_{(m:m)} - \mu_{(m:m)}| > \varepsilon] \leq \frac{\sigma_{(m:m)}^2}{\varepsilon^2}$.

So if $\rho_1 < \rho$

$p_1 = \int_{|Y_{(m:m)} - \mu_{(m:m)}| > \varepsilon} f_{(m:m)}(y) \phi\left(\frac{y(\rho_1-\rho)}{\sqrt{1-\rho^2}}\right) dy$

$+ \int_{|Y_{(m:m)} - \mu_{(m:m)}| \leq \varepsilon} f_{(m:m)}(y) \phi\left(\frac{y(\rho_1-\rho)}{\sqrt{1-\rho^2}}\right) dy$. 
\[ \begin{align*}
&\leq 1 \cdot \Pr[|Y_{m:m} - \mu_{m:m}| \geq \varepsilon] + \phi\left(\frac{(\mu_{m:m} - \varepsilon)(\rho_1 - \rho)}{\sqrt{1 - \rho^2}}\right) \Pr[|Y_{m:m} - \mu_{m:m}| \leq \varepsilon] \\
&\leq \frac{\sigma^2_{m:m}}{\varepsilon^2} + \phi\left(\frac{(\mu_{m:m} - \varepsilon)(\rho_1 - \rho)}{\sqrt{1 - \rho^2}}\right) \to 0 \text{ as } m \to \infty
\end{align*} \]

and if \( \rho_1 > \rho \)

\[ p_1 \geq \int_{|Y_{m:m} - \mu_{m:m}| \leq \varepsilon} \phi\left(\frac{y}{\sqrt{1 - \rho^2}}\right) f_{m:m}(y) \, dy \]

\[ \geq \Pr[|Y_{m:m} - \mu_{m:m}| \leq \varepsilon] \cdot \phi\left(\frac{(\mu_{m:m} - \varepsilon)(\rho_1 - \rho)}{\sqrt{1 - \rho^2}}\right) \]

\[ \geq \left[1 - \frac{\sigma^2_{m:m}}{\varepsilon^2}\right] \cdot \phi\left(\frac{(\mu_{m:m} - \varepsilon)(\rho_1 - \rho)}{\sqrt{1 - \rho^2}}\right) \to 1 \text{ as } m \to \infty. \]

Thus \( \lim_{m \to \infty} \Pr[\rho_1 \varepsilon I_{1-\alpha}] \)

\[ = \lim_{m \to \infty} \Pr \left\{ \frac{\sqrt{N(h_p) - z_{1-\alpha/2}}}{\sqrt{p_1 q_1}} < \frac{n_{m,n} - N p_1}{\sqrt{N p_1 q_1}} < \frac{\sqrt{N(h_p) + z_{1-\alpha/2}}}{\sqrt{p_1 q_1}} \right\} \]

\[ = \left\{ \begin{array}{ll}
\lim_{p_1 \to 0} \left[ \Phi\left(\frac{\sqrt{N(h_p) + z_{1-\alpha/2}}}{\sqrt{p_1 q_1}}\right) - \Phi\left(\frac{\sqrt{N(h_p) - z_{1-\alpha/2}}}{\sqrt{p_1 q_1}}\right) \right] & \text{if } \rho_1 < \rho \\
\lim_{p_1 \to 1} \left[ \Phi\left(\frac{\sqrt{N(h_p) + z_{1-\alpha/2}}}{\sqrt{p_1 q_1}}\right) - \Phi\left(\frac{\sqrt{N(h_p) + z_{1-\alpha/2}}}{\sqrt{p_1 q_1}}\right) \right] & \text{if } \rho_1 > \rho 
\end{array} \right. \]

\[ = 0. \]

(b) Considering \( p_1 = \int_{-\infty}^{\infty} f_{m:m}(y) \phi\left(\frac{y}{\sqrt{1 - \rho^2}}\right) \, dy \) as a function of \( \rho_1 \),

we see that
\[
\frac{\partial p_1}{\partial \rho_1} = \int_{-\infty}^{\infty} f_{(m:m)}(y) \frac{\gamma(\rho_1 - \rho)}{\sqrt{1-\rho^2}} \frac{y}{\sqrt{1-\rho^2}} \, dy
\]

\[
= \int_{0}^{\infty} \frac{\gamma(\rho_1 - \rho)}{\sqrt{1-\rho^2}} \frac{Y}{\sqrt{1-\rho^2}} (f_{(m:m)}(y) - f_{(m:m)}(-y)) \, dy
\]

\[
= \int_{0}^{\infty} \frac{\gamma(\rho_1 - \rho)}{\sqrt{1-\rho^2}} \frac{Y}{\sqrt{1-\rho^2}} \, m \phi(y) [\phi(y)^{m-1} - \phi(-y)^{m-1}] \, dy > 0 \quad \forall \rho_1 \text{ if } m > 1.
\]

Therefore \( p_1 \) is a strictly increasing function of \( \rho_1 \). This tells us that

\[
\rho_1 = \rho \text{ iff } p_1 = \frac{1}{2}.
\]

Thus

\[
\Pr[\rho_1 \in I_{1-\alpha}] = \phi\left(\frac{\sqrt{N}(\frac{1}{2} - p_1) + z_{1-\alpha/2}}{\sqrt{p_1 q_1}}\right) - \phi\left(\frac{\sqrt{N}(\frac{1}{2} - p_1) - z_{1-\alpha/2}}{\sqrt{p_1 q_1}}\right) \to 0
\]

as \( N \to \infty \) if \( m > 1 \).

Note that this confidence interval is still valid if errors in ranking occur. However, unless the ranking is somewhat better than random, we cannot be assured that \( \lim_{N \to \infty} \Pr[\rho_1 \in I_{1-\alpha}] = 0 \).

3.2 Estimation of \( \rho \) with \( \mu_y, \sigma_y^2 \) Known; \( \mu_x, \sigma_x^2 \) Unknown.

Now suppose we are interested in the estimation of \( \rho \) when only \( \mu_y \) and \( \sigma_y^2 \) are known, with \( \mu_x, \sigma_x^2 \) unknown. The MLE's of \( \mu_x, \sigma_x^2 \), and \( \rho \) from a random sample of size \( N \) are [11]

\[
\begin{align*}
\bar{\mu}_x &= \bar{X} - \left(\frac{r}{s_y}\right)(\bar{Y} - \mu_y) \\
\tilde{s}_x^2 &= s_x^2(1-r^2+s_y^2/s_y^2) \\
\tilde{\rho} &= (r \sigma_y/s_y)(1-r^2+s_y^2/s_y^2)^{-\frac{1}{4}}
\end{align*}
\]
where \( s_x^2 = \frac{1}{N} \sum_{i=1}^{N} (X_i - \bar{X})^2 \), \( s_y^2 = \frac{1}{N} \sum_{i=1}^{N} (Y_i - \bar{Y})^2 \), \( r_{xys} = \frac{1}{N} \sum_{i=1}^{N} (X_i - \bar{X})(Y_i - \bar{Y}) \).

If we choose as our bivariate sample a ranked set sample of \( Y \)'s and their concomitant \( X \)'s, then the MLE's for \( \rho \), \( \sigma_x^2 \), and \( \mu_x \) will be calculated in the same way. This is true for any modified ranked set sample as described in Section 3.1, or, in fact, for any ranked set sampling scheme where the \( Y \)'s are judgment order statistics.

The reason for this, of course, is that the p.d.f. of \( (X_{[r:m]}, Y_{[r:m]}) \) is

\[
f_{[r:m]}(x,y) = f_{[r:m]}(y)f(x|y)
\]

where \( f_{[r:m]}(y) \) is not a function of \( \rho \), \( \mu_x \), or \( \sigma_x^2 \).

In order to find the asymptotic variance of the MLE of \( \rho \), we must determine the Fisher Information matrix \( I_N(\mu_x, \sigma_x^2, \rho) \). After calculations similar to those of the previous section, we find that, once again, a random sample and a ranked set sample produce identical Fisher Information matrices, which are both equivalent to

\[
I_N(\mu_x, \sigma_x^2, \rho) = \frac{N}{(1-\rho^2)} \begin{bmatrix}
\frac{1}{\sigma_x^2} & 0 & 0 \\
0 & \frac{2-\rho^2}{\sigma_x^2} & -\frac{\rho}{\sigma_x} \\
0 & -\frac{\rho}{\sigma_x} & \frac{1+\rho^2}{1-\rho^2}
\end{bmatrix}.
\]

Thus

\[
I^{-1}_N(\mu_x, \sigma_x^2, \rho) = \frac{(1-\rho^2)}{N} \begin{bmatrix}
\frac{\sigma^2}{\sigma_x^2} & 0 & 0 \\
0 & \frac{\sigma_x^2(1+\rho^2)}{2} & \frac{\rho \sigma_x(1-\rho^2)}{2} \\
0 & \frac{\rho \sigma_x(1-\rho^2)}{2} & \frac{(1-\rho^2)(2-\rho^2)}{2}
\end{bmatrix}.
\]
Now suppose the observations in our sample are chosen according to a modified ranked set sampling scheme; i.e., choose for quantification \(N/2\) each of \(Y_{(m:m)}\) and \(Y_{(1:m)}\), along with their concomitant \(X\)'s. Then the upper triangle of the Fisher Information matrix for this sample will be

\[
\frac{1}{\Delta} \frac{1}{(m)N} \chi^2(x, \sigma^2, \rho) = \frac{N}{(1-\rho^2)} \begin{bmatrix}
\frac{1}{\sigma^2_x} & 0 & 0 \\
2(1-\rho^2) + \rho^2 \frac{\sigma^2_x}{2 \text{EU}} & \frac{\rho \sigma_x (2-\text{EU})}{\text{EU}^2} & \frac{\rho \sigma_x (2-\text{EU}) (1-\rho^2)}{\text{EU}^2} \\
\frac{2(1-\rho^2) + \rho^2 \sigma_x^2}{\text{EU}^2} & \frac{\rho \sigma_x (2-\text{EU})}{\text{EU}^2} & \frac{\rho \sigma_x (2-\text{EU}) (1-\rho^2)}{\text{EU}^2}
\end{bmatrix}
\]

Thus

\[
\frac{1}{\Delta} \frac{1}{(m)N} \chi^2(x, \sigma^2, \rho)
\]

\[
= \frac{(1-\rho^2)}{N} \begin{bmatrix}
\sigma^2_x & 0 & 0 \\
0 & \frac{\sigma^2_x [(1-\rho^2) \text{EU} + 2 \rho^2]}{2 \text{EU}^2} & \frac{\rho \sigma_x (2-\text{EU})}{\text{EU}^2} & \frac{\rho \sigma_x (2-\text{EU}) (1-\rho^2)}{\text{EU}^2} \\
0 & 0 & \frac{\rho \sigma_x (2-\text{EU}) (1-\rho^2)}{\text{EU}^2} & \frac{(1-\rho^2) [2(1-\rho^2) + \rho^2 \text{EU}^2]}{2 \text{EU}^2}
\end{bmatrix}
\]

(3.12)

where \(U \sim N(0,1)\). Since

\[
\frac{(1-\rho^2)(2-\rho^2)}{2N} = \frac{(1-\rho^2)}{2 \text{EU}^2} \left[ \frac{2 \text{EU}^2}{(m:m)} - \rho^2 \text{EU}^2 \right]
\]
\[
\tilde{\rho}_N = \frac{(1-\rho^2)^2}{2\text{NEU}^2_{(m:m)}} \left[ 2\text{EU}^2_{(m:m)}(1-\rho^2) + \rho^2\text{EU}^2_{(m:m)} \right]
\]

we know that the asymptotic variance of \(\tilde{\rho}_N\) is always larger than that of the MLE from the modified ranked set sample, \(\tilde{\rho}_{(m)N}^\Delta\). Note also from \(T_{(m)N}^{\Delta-1}(\mu_x, \sigma_x, \rho)\) that the asymptotic variance of \(N_{(m)N}^{\Delta} \tilde{\rho}_{(m)N}^\Delta\) as \(m \to \infty\).

Here again, we can use the fact that \(\tilde{\rho}_{(m)N}^\Delta\) is asymptotically normally distributed to find an approximate 100(1-\(\alpha\))% confidence interval for \(\rho\). Unfortunately, the methods of Section 3.1 cannot be used to find exact confidence intervals, since estimates of \(\mu_x\) and \(\sigma_x^2\) have to be introduced. If \(\mu_x\) and \(\sigma_x^2\) are replaced by their maximum likelihood estimators in \(T_{(m)N}\) or \(\eta_{(m)N}\), those statistics could still be used to find approximate confidence intervals for \(\rho\) when \(N\) is large. They are asymptotically normally distributed and will have the same asymptotic properties as when \(\mu_x, \sigma_x^2\) were known because of the consistency of the MLE's. However, I have not demonstrated that a confidence interval procedure based on \(T_{(m)N}\) or \(\eta_{(m)N}\) is in any way preferable to that based on \(\tilde{\rho}_{(m)N}^\Delta\), although no investigation of the speed with which they reach their asymptotic distribution has been undertaken.

3.3 Estimation of \(\rho\) with All Parameters Unknown.

Now suppose that \(\mu_x, \mu_y, \sigma_x^2, \) and \(\sigma_y^2\) are all unknown. The MLE's of all five parameters are well known if the sample \(X_1, \ldots, X_N\) is a random sample. But suppose our sample is a modified ranked set
sample of the type which includes \( N/2 \) largest and \( N/2 \) smallest \( Y \) observations and their concomitant \( X \)'s; i.e., our bivariate sample is \((X_{[1:m]}, 1, Y_{[1:m]}, 1), \ldots, (X_{[1:m]}, N/2, Y_{(1:m)}, N/2), (X_{[m:m]}, 1, Y_{(m:m)}, 1), \ldots, (X_{[m:m]}, N/2, Y_{(m:m)}, N/2)\). Then (writing \( Y_{mi} \) for \( Y_{(m:m), i} \), \( X_{1i} \) for \( X_{[1:m], i} \), etc.) the new likelihood function is

\[
L = \prod_{i=1}^{N/2} m^{-1}(Y_{mi}) f(X_{mi}, Y_{mi}) \prod_{i=1}^{N/2} m[1-F(Y_{mi})]^{-1} f(X_{mi}, Y_{mi})
\]

where \( f \) is a joint normal p.d.f. with parameters \( \mu_x, \mu_y, \sigma_x^2, \sigma_y^2 \) and \( \rho \); and \( F \) is a normal c.d.f. with parameters \( \mu_y \) and \( \sigma_y^2 \). Then the five likelihood equations are:

\[
\begin{align*}
\frac{\partial \ln L}{\partial \mu_x} &= \frac{N}{\sigma_x(1-\rho^2)} \left\{ \frac{\bar{X} - \mu_x}{\sigma_x} - \rho \frac{\bar{Y} - \mu_y}{\sigma_y} \right\} = 0 \\
\frac{\partial \ln L}{\partial \mu_y} &= \frac{N}{\sigma_y(1-\rho^2)} \left\{ \frac{\bar{Y} - \mu_y}{\sigma_y} - \rho \frac{\bar{X} - \mu_x}{\sigma_x} \right\} \\
&\quad + \frac{(m-1)}{\sigma_y} \sum_{i=1}^{N/2} \left\{ \frac{f(Y_{li} - \mu_y)}{\sigma_y} \frac{f(Y_{mi} - \mu_y)}{F(Y_{mi} - \mu_y)} - \frac{f(Y_{li} - \mu_y)}{1-F(Y_{li} - \mu_y)} \right\} = 0
\end{align*}
\]

where \( \bar{X} = \frac{1}{N} \sum_{i=1}^{N/2} (X_{1i} + X_{mi}) \) and \( \bar{Y} = \frac{1}{N} \sum_{i=1}^{N/2} (Y_{li} + Y_{mi}) \).

\[
\begin{align*}
\frac{\partial \ln L}{\partial (\sigma_x^2)} &= -\frac{1}{2\sigma_x^2(1-\rho^2)} \left\{ N(1-\rho^2) - \sum_{i=1}^{N/2} \left[ (X_{li} - \mu_x)^2 + (X_{mi} - \mu_x)^2 \right] \right\} \\
&\quad + \rho \frac{\sum_{i=1}^{N/2} [(X_{li} - \mu_x)(Y_{li} - \mu_y) + (X_{mi} - \mu_x)(Y_{mi} - \mu_y)]}{\sigma_x \sigma_y} = 0 \\
\frac{\partial \ln L}{\partial (\sigma_y^2)} &= -\frac{1}{2\sigma_y^2(1-\rho^2)} \left\{ N(1-\rho^2) - \sum_{i=1}^{N/2} \left[ (Y_{li} - \mu_y)^2 + (Y_{mi} - \mu_y)^2 \right] \right\}
\end{align*}
\]
\[
\begin{align*}
&\frac{\gamma_{N/2}}{\sigma_x \sigma_y} \sum_{i=1}^{N/2} [(X_{li} - \mu_x)(Y_{li} - \mu_y) + (X_{mi} - \mu_x)(Y_{mi} - \mu_y)] \\
&\quad + \rho \frac{\gamma_{N/2}}{\sigma_x \sigma_y} \sum_{i=1}^{N/2} \left[ \frac{f((Y_{mi} - \mu_y)/\sigma_y)}{F((Y_{mi} - \mu_y)/\sigma_y)} \right] \left( \frac{Y_{mi} - \mu_y}{\sigma_y} \right)^2 \\
&\quad - \frac{f((Y_{li} - \mu_y)/\sigma_y)}{1 - F((Y_{li} - \mu_y)/\sigma_y)} \left( \frac{Y_{li} - \mu_y}{\sigma_y} \right)^2 = 0
\end{align*}
\]

\[
\frac{\partial \ln L}{\partial \rho} = \frac{1}{(1-\rho^2)} \left( N\rho - \frac{1}{(1-\rho^2)} \left[ \rho \frac{\gamma_{N/2}}{\sigma_x} \sum_{i=1}^{N/2} [(X_{li} - \mu_x)^2 + (Y_{mi} - \mu_y)^2] \\
+ \frac{\gamma_{N/2}}{\sigma_y} \sum_{i=1}^{N/2} (Y_{li} - \mu_y)^2 + (Y_{mi} - \mu_y)^2 \right] \\
- (1+\rho^2) \frac{\gamma_{N/2}}{\sigma_x \sigma_y} \sum_{i=1}^{N/2} [(X_{li} - \mu_x)(Y_{li} - \mu_y) + (X_{mi} - \mu_x)(Y_{mi} - \mu_y)] \right) \\
= 0.
\]

An exact solution to these equations would be difficult to find numerically. One finds, using Tiku's approximation which was introduced in Section 2.1.3, that an approximate solution to (3.13) is (see Appendix, p.107).

\[
\hat{\mu}_x = \bar{X}, \quad \hat{\mu}_y = \bar{Y}.
\]

But this approximation does not produce an explicit solution to the remaining three equations.

One intuitive approach to estimating all parameters with only the data at hand is to first estimate \( \mu_y \) and \( \sigma_y \) with their BLUE's. These estimates are, for our modified ranked set sample of \( Y \)'s,
\[ \hat{\mu}_y = \bar{y} \quad \text{and} \quad \hat{\sigma}_y = \frac{\sum_{i=1}^{N/2} (Y_{mi} - Y_{1i})}{\alpha_m N} \]

where \( \alpha_m = \text{EU}_{(m:m)}, \quad U \sim N(0,1) \). Then replace \( \mu_y \) and \( \sigma_y \) in (3.10) with their BLUE's. The remaining estimators will then be

\[ \hat{\mu}_x = \bar{x} \]
\[ \hat{\sigma}_x^2 = s_x^2 (1 - r^2 + r^2 \hat{\sigma}_y^2 / s_y^2) \]
\[ \hat{\rho} = \left( r \hat{\sigma}_y / s_y \right) (1 - r^2 + r^2 \hat{\sigma}_y^2 / s_y^2)^{-1/2} \]

where

\[ s_x^2 = \frac{\sum_{i=1}^{N/2} [(X_{1i} - \bar{x})^2 + (X_{mi} - \bar{x})^2]}{N} \]
\[ s_y^2 = \frac{\sum_{i=1}^{N/2} [(Y_{1i} - \bar{y})^2 + (Y_{mi} - \bar{y})^2]}{N} \]
\[ s_{x,y}^2 = \frac{\sum_{i=1}^{N/2} [(X_{1i} - \bar{x})(Y_{1i} - \bar{y}) + (X_{mi} - \bar{x})(Y_{mi} - \bar{y})]}{N} \]

A simple Monte Carlo study was undertaken in order to gain some information about the behavior of the estimator \( \hat{\rho} \). The mean and variance of \( \hat{\rho} \) were calculated on the basis of 200 samples, which were chosen by the modified ranked set sampling procedure described in this section. The results of this experiment are given in Table A.1 of the Appendix (p. 109) for \( m = 5, 7, 10, \rho = 0, .5, .75, \) and \( N = 10, 20, 30 \). For comparison purposes, we give in Table A.2 the ratio of the Cramér-Rao lower bound of the estimators of \( \rho \) from a random sample \( ((1 - \rho^2)^2 / N) \) to the variance of \( \hat{\rho} \). For large \( N \), \( (1 - \rho^2)^2 / N \) is the approximate variance of \( r \), the Pearson product-
moment correlation coefficient. Examination of this table shows that in all cases, these ratios are larger than 1, indicating that ranked set sampling substantially improves the precision of the estimator of \( \rho \) over that from a random sample for an \( m \) even as small as 5.

A drawback of the estimator \( \hat{\rho} \) is that it is not valid when errors in ranking occur, since \( \hat{\sigma}_y \) will no longer be an unbiased estimator of \( \sigma_y \) and \( \hat{\mu}_y \) will not be an unbiased estimator of \( \mu_y \) unless the ranking errors one makes are "symmetrical" (i.e., unless \( E(Y_{[m:m]} - \mu_y) = -E(Y_{[1:m]} - \mu_y) \)), which is unlikely. If errors in the ranking of the \( Y \)'s does occur, then a two stage estimation process could be employed; i.e., independent estimates of \( \mu_y \) and \( \sigma_y \) can be introduced into (3.9). Presumably, this process will be relatively inexpensive, since, according to the framework of the problem, the observations of \( Y \) are more accessible than those of \( X \).
4. HYPOTHESIS TESTING AND RANKED SET SAMPLING

In this chapter, test statistics calculated from ranked set samples are compared with appropriate conventional tests, and, in other cases, with each other.

4.1 Tests of Location

Since the first step in the ranked set sampling procedure is an ordering process, it is natural in some instances to compare a test based on a ranked set statistic with a nonparametric rank test. Let \( X_{(1:m)}, \ldots, X_{(m:m)}, n_x \) and \( Y_{(1:m)}, \ldots, Y_{(m:m)}, n_y \) be ranked set samples from each of two populations, with \( n_x + n_y = n \). Suppose we wish to test the hypothesis that the two populations are identical against the alternative that they are of the same form, but differ in location; i.e.,

\[
H_0: F_X(x) = F_Y(x) \quad \text{for all } x
\]

against

\[
H_1: F_X(x) = F_Y(x-\theta) \quad \text{for all } x \text{ and some } \theta \neq 0.
\]

Consider the test statistic

\[
T(m)_{n_x, n_y} = \frac{\hat{\mu}_y(m)n_x - \hat{\mu}_x(m)n_y}{\sqrt{(n_x-1) \sum_{r=1}^{m} s^2_y(r:m) + (n_y-1) \sum_{r=1}^{m} s^2_x(r:m) \left( \frac{n_x + n_y - 2}{n_x + n_y} \right)}}
\]

where \( \hat{\mu}_y = (1/m_n) \sum_{i=1}^{n_y} \sum_{r=1}^{m} Y_{(r:m),i} \),

\[
s^2_y(r:m) = \frac{\sum_{i=1}^{n_y} (Y_{(r:m),i} - \hat{\mu}_y)^2}{n_y - 1}\]
\[ \hat{\mu}_y(r) = \left(1/n_y\right) \sum_{i=1}^{n_y} y(r:m)_i, \] and likewise for \( X \). For large values of \( n_x \) and \( n_y \),

\[
\frac{(n_x - 1) \sum_{r=1}^{m} s_x^2(r:m) + (n_y - 1) \sum_{r=1}^{m} s_y^2(r:m)}{n_x + n_y - 2} + \sum_{r=1}^{m} \sigma^2(r:m)
\]

as \( n_x, n_y \to \infty \).

Therefore, when \( n_x \) and \( n_y \) are large, \( T(m)_{n_x, n_y} \) is approximately normally distributed with

\[
\mu_{T(m)_{n_x, n_y}}(\theta) = \frac{\theta}{\sqrt{\sum_r \sigma^2(r:m)}}, \quad \sigma^2_{T(m)_{n_x, n_y}}(\theta) = \frac{1}{m^2} \left( \frac{1}{n_x} + \frac{1}{n_y} \right)
\]

for samples from distributions satisfying certain weak conditions. A test based on \( T(m) \) will be compared with the Mann-Whitney test, a linear rank test for detecting location differences.

Recall that because of the ranked set sampling process, there were originally \( mn_x \) samples of \( X \) observations and \( mn_y \) samples of \( Y \) observations, each of size \( m \). But \( T(m)_{n_x, n_y} \) is calculated only from the quantified observations of which there are \( mn_x \) \( X \)'s and \( mn_y \) \( Y \)'s, a total of \( mn \). Computation of the Mann-Whitney statistic, however, requires only that the order of the observations within a combined sample be known, and no exact measurements are needed. So if the data is the type where ranked set sampling is appropriate, i.e., where quantifying is much more difficult than ordering, a Mann-Whitney test would seem a reasonable choice. Notice, however, that it is not always possible to eye-order the combined sample, since the sample of \( X \)'s may not become available at the same time or at the same place as the sample of \( Y \)'s, thus eliminating the opportunity to directly com-
pare the members of the two samples.

Two situations will be considered. In the first, \( T_{(m)n_x, n_y} \)
is compared with the Mann-Whitney statistic \( U_{(m)n_x, n_y} \), which isbased on one Mann-Whitney test applied to all the observations.

That is, assume the ranking of all \( m^2 n \) observations to be established,and calculate \( U_{(m)n_x, n_y} \) from this information. This is usually anunrealistic approach, of course, because our concern all along hasbeen to keep the number to be ordered in each sample reasonably smallso that the ordering may be done accurately. In the second situation, this "advantage" given to \( U_{(m)} \) is partially eliminated. Herelet \( n_x = n_y = N/2 \). Then suppose we combine one sample of \( X \)'s withone sample of \( Y \)'s, both of size \( m \), and that these \( 2m \) observationsare ranked among themselves and a Mann-Whitney test performed. Thisprocess is repeated \( mn/2 \) times (since there are \( mn/2 \) pairs of samples),and the resulting \( mn/2 \) test statistics are combined to produce a"Mann-Whitney type" statistic

\[
U^*_{(m)n} = \sum_{i=1}^{mn/2} U_{mi},
\]

where \( U_{mi} \) is the Mann-Whitney statistic from the \( i \)-th combined samples.

For large values of \( n \), both \( U_{(m)n_x, n_y} \) (because it is a linear rankstatistic) and \( U^*_{(m)n} \) (by the CLT) have approximately normal distributions, and

\[
\begin{align*}
U_{(m)n_x, n_y} (0) &= (m^2 n_x) (m^2 n_y) Pr[Y > X] = m n_x n_y \int_{-\infty}^{\infty} F(z+\theta) f(z) dz \\
\sigma^2_{U_{(m)n_x, n_y}} (0) &= m^2 n_x n_y (m^2 (n_x + n_y) + 1)/12 = \frac{m n_x n_y (mn+1)}{12}
\end{align*}
\]

(4.2)
\[
\begin{align*}
\mu_{U^*_{(m)n}}(\theta) &= \frac{mn}{2} m^2 \Pr[Y > X] \\
\sigma^2_{U^*_{(m)n}}(\theta) &= \frac{mn}{2} m^2 \frac{(2m+1)}{12} = \frac{m^3 n (2m+1)}{24} 
\end{align*}
\]

(4.3)

Theorem C of the Appendix enables us to calculate in a simple manner the Pitman efficiency of \( T \) wrt \( U \) and \( U^* \). Following the notation of the theorem, each of our statistics is based on \( mn \) observations. (Note that we get \( m^2 n - mn = mn(m-1) \) observations "free" in the nonparametric tests.) From (4.1), we have

\[
\frac{d}{d\theta} \mu_{T_{(m)n_x,n_y}}(0) = \frac{1}{\sqrt{\sum_{r=1}^{m} \sigma^2_{r:m}}} 
\]

and

\[
\sigma_{T_{(m)n_x,n_y}}(0) = \sqrt{\frac{1}{m} \left( \frac{1}{n_x} + \frac{1}{n_y} \right)} = \sqrt{\frac{n}{2 \, n_x \, n_y}} 
\]

From (4.2) and (4.3),

\[
\frac{d}{d\theta} \mu_{U_{(m)n_x,n_y}}(0) = m^4 n_x n_y \int_{-\infty}^{\infty} t^2(z) dz 
\]

and

\[
\sigma_{U_{(m)n_x,n_y}}(0) = m^2 \sqrt{\frac{n_x n_y (mn+1)}{12}} 
\]

\[
\frac{d}{d\theta} \mu_{U^*_{(m)n}}(0) = \frac{m^3 n}{2} \int_{-\infty}^{\infty} t^2(z) dz 
\]

and

\[
\sigma_{U^*_{(m)n}}(0) = \sqrt{\frac{m^3 n (2m+1)}{24}} 
\]
Thus if \( \frac{n_x}{n} \to p \) and \( \frac{n_y}{n} \to (1-p) \) as \( n \to \infty \),

\[
c_t^{(m)} = \lim_{n \to \infty} \frac{1}{\sqrt{mn} \sum_{r=1}^{m} \sigma_r^2(r:m)} \cdot m \sqrt{n_x n_y / n}
\]

\[
= \sqrt{\frac{mp(1-p)}{\sum_{r} \sigma_r^2(r:m)}}
\]

\[
c_u^{(m)} = \lim_{n \to \infty} \frac{\sqrt{\frac{12}{m} n_x n_y \int_{-\infty}^{\infty} f^2(z)dz}}{m \sqrt{mn} n_x n_y (mn+1) / n}
\]

\[
= m \sqrt{\frac{12p(1-p)}{\int_{-\infty}^{\infty} f^2(z)dz}}
\]

\[
c_u^{*}\n_{(m)} = \lim_{n \to \infty} \frac{\sqrt{\frac{24}{m^2} n_x n_y \int_{-\infty}^{\infty} f^2(z)dz}}{2\sqrt{(mn)} n_x n_y (2m+1) / n}
\]

\[
= \sqrt{\frac{6m^2}{2m+1}} \int_{-\infty}^{\infty} f^2(z)dz
\]

Therefore, the Pitman efficiency of \( U^{(m)} \) relative to \( T^{(m)} \),

\[
e_{U^{(m)}, T^{(m)}} = \frac{12m^2 p(1-p) \int_{-\infty}^{\infty} f^2(z)dz}{mp(1-p)} \sum_{r=1}^{m} \sigma_r^2(r:m)
\]

\[
= 12m^2 \left[ \int_{-\infty}^{\infty} f^2(z)dz \right] \left( \sum_{r=1}^{m} \sigma_r^2(r:m) / \sigma^2 \right)
\]

\[
= m^2 \text{var} \hat{\mu}_{(m)n} / \text{var} \hat{X}_{mn}
\]

where \( u, t \) are the conventional Mann-Whitney and \( t \)-tests. Similarly,

\[
e_{U^{*}\n_{(m)}, T^{(m)}} = \left( \frac{c_{U^{*}\n_{(m)}}}{c_{T^{(m)}}} \right)^2
\]
\[
\frac{6m^2}{2m+1} \left[ \frac{\int f^2(z)dz}{\sum_{r=1}^{m} \sigma_r^2(r;m)} \right]^{2m} \left[ \frac{\var \mu_{(m)n}}{\var \bar{X}_{mn}} \right]^{m/4}
\]

We noted in Chapter 1 that \(\frac{\var \mu_{(m)n}}{\var \bar{X}_{mn}} \geq \frac{2}{(m+1)}\), with equality holding only for the uniform distribution. So for the uniform distribution

\[
e_{U_{(m)}}, T_{(m)} = \frac{2m^2}{(m+1)} = O(m),
\]

\[
e_{U^*_{(m)}}, T_{(m)} = \frac{4m^2}{(m+1)(2m+1)} + 2 \quad \text{as } m \to \infty.
\]

For many other distributions, \(\frac{\var \mu_{(m)n}}{\var \bar{X}_{mn}}\) is close to its lower bound of \(2/(m+1)\). For these distributions then,

\[
e_{U_{(m)}}, T_{(m)} \approx \frac{2m^2}{m+1} e_{u,t},
\]

\[
e_{U^*_{(m)}}, T_{(m)} \approx \frac{4m^2}{(2m+1)(m+1)} e_{u,t}.
\]

From these results, we see that \(U_{(m)}\) is better as far as Pitman efficiency is concerned. However, as we pointed out earlier, the accurate ranking of all \(mn^2\) observations will quite possibly be very difficult. Table 4.1 displays \(e_{U^*_{(m)}}, T_{(m)}\) for \(m = 2, \ldots, 5\) for the normal and uniform distributions. Here again, the nonparametric test would be preferred over that based on the statistic \(T_{(m)}\) on the basis of Pitman efficiency. We must keep in mind, however, that:
(1) direct comparison of the two samples may not be feasible

(2) \( T_m \) requires the ordering of \( m \) observations while \( U^*_{(m)} \) requires ordering of \( 2m \)

(3) if ranking errors occur, \( T_m \) will still produce a valid size \( \alpha \) test for large samples, while \( U^*_{(m)} \) (or \( U_{(m)} \)) will not.

<table>
<thead>
<tr>
<th>( m )</th>
<th>Normal</th>
<th>Uniform</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.04</td>
<td>1.07</td>
</tr>
<tr>
<td>3</td>
<td>1.28</td>
<td>1.29</td>
</tr>
<tr>
<td>4</td>
<td>1.45</td>
<td>1.42</td>
</tr>
<tr>
<td>5</td>
<td>1.57</td>
<td>1.51</td>
</tr>
</tbody>
</table>

4.2 Tests of Correlation

Suppose that \((X,Y)\) has a bivariate normal distribution with \( \text{corr}(X,Y) = \rho \). In Chapter 3, two procedures for producing confidence intervals for \( \rho \) were developed. In this section, these procedures will be used to construct tests of

\[ H_0: \ \rho = \rho_0 \]

against

\[ H_1: \ \rho \neq \rho_0 \]

First suppose \( \mu_x, \mu_y, \sigma_x^2, \sigma_y^2 \) are known. For the following discussions we assume that \( \mu_x = \mu_y = 0, \sigma_x^2 = \sigma_y^2 = 1 \). Recall from Chapter 3 that

\[
T_{(m)N} = \frac{\sum_{i=1}^{n} (X_{[m:m],i} - \rho_0 Y_{(m:m),i})}{\sqrt{N(1-\rho_0^2)}} \sim N(0,1)
\]

if \( H_0 \) is true. The power function of this test is
\[ \beta_T^{(m)}(\rho) = \Pr_{\rho} \{ \text{the test rejects } H_0 \} \]

\[ = \Pr_{\rho} \left\{ \left| \sum_{i=1}^{N} (X_{m:m},i - \rho_0 Y_{m:m},i) \right| > z_{1-\alpha/2} \sqrt{N(1-\rho_0^2)} \right\} \]

\[ = 1 - \Pr_{\rho} \left\{ (\rho_0 - \rho) \sum_{i=1}^{N} Y_{m:m},i - z_{1-\alpha/2} \sqrt{N(1-\rho_0^2)} \right\} \]

\[ \leq \sum_{i=1}^{N} (X_{m:m},i - \rho Y_{m:m},i) \]

\[ \leq (\rho_0 - \rho) \sum_{i=1}^{N} Y_{m:m},i + z_{1-\alpha/2} \sqrt{N(1-\rho_0^2)} \right\} \]

\[ = \begin{cases} 
1 - \mathcal{E} \left\{ \left( \frac{(\rho_0 - \rho) \sum_{i=1}^{N} Y_{m:m},i + z_{1-\alpha/2} \sqrt{N(1-\rho_0^2)}}{\sqrt{N(1-\rho_0^2)}} \right) \right\} \\
- \mathcal{E} \left\{ \left( \frac{(\rho_0 - \rho) \sum_{i=1}^{N} Y_{m:m},i - z_{1-\alpha/2} \sqrt{N(1-\rho_0^2)}}{\sqrt{N(1-\rho_0^2)}} \right) \right\} 
\end{cases} \]

if \( |p| < 1 \),

\[ \Pr\{ |(\rho - \rho_0) \sum_{i=1}^{N} Y_{m:m},i| \leq z_{1-\alpha/2} \sqrt{N(1-\rho_0^2)} \} \quad \text{if } |p| = 1 \]

where \( \Phi \) is the c.d.f. of the unit normal distribution. This function is a difficult one to evaluate, so we introduce an approximation which is valid when \( N \) is large enough that \( \sqrt{N(1-\rho_0^2)} \) \( T_N = \sum_{i=1}^{N} (X_{m:m},i - \rho_0 Y_{m:m},i) \) will have an approximately normal distribution when \( \text{corr}(X,Y) = \rho \neq \rho_0 \). We know that

\[ ET_{(m)}^N = \frac{(N/\sqrt{N(1-\rho_0^2)})(\rho \mu_{(m:m)} - \rho_0 \mu_{(m:m)})}{\rho_0^2} \]

\[ = \sqrt{N(\rho - \rho_0)} \mu_{(m:m)} / \sqrt{1-\rho_0^2} \]
and
\[ \text{var } T_{(m)N} = \text{E}[\text{var}(T_{(m)N}|Y_{(m):i}, i=1,\ldots,N)] \\
+ \text{var}[\text{E}(T_{(m)N}|Y_{(m):i}, i=1,\ldots,N)] \]
\[ = \frac{1}{N(1-\rho_0^2)} [N(1-\rho^2) + N(\rho-\rho_0)^2\sigma^2_{(m:m)}] \]
\[ = \frac{[1-\rho^2 + (\rho-\rho_0)^2\sigma^2_{(m:m)}]}{(1-\rho_0^2)} \]

where \( \mu_{(m:m)} = \text{E}U_{(m:m)}, \sigma^2_{(m:m)} = \text{var } U_{(m:m)}, U \sim N(0,1) \).

Thus, if we approximate
\[ \Pr\left[ \frac{1}{N} \sum_{i=1}^{N} (X_{[m:m],i} - \rho_0 Y_{(m:m),i}) \right] \geq z_{1-\alpha/2} \]
by
\[ \Pr[|S_N| \geq z_{1-\alpha/2}] \]
where
\[ S_N \sim N \left( \frac{\sqrt{N}(\rho-\rho_0)\mu_{(m:m)}}{\sqrt{1-\rho_0^2}}, \frac{1-\rho^2 + (\rho-\rho_0)^2\sigma^2_{(m:m)}}{(1-\rho_0^2)} \right) , \]

we have
\[ \beta_T^{(m)}(\rho) \approx \beta_I^{(m)}(\rho) \]
\[ = 1 - \Pr \left[ \frac{\sqrt{N}(\rho-\rho)^\mu_{(m:m)} - z_{1-\alpha/2} \sqrt{1-\rho_0^2}}{\sqrt{1-\rho^2 + (\rho-\rho_0)^2\sigma^2_{(m:m)}}} \right] \leq \frac{S_N - \mu_{S_N}}{\sigma_{S_N}} \]
\[ \leq \frac{\sqrt{N}(\rho-\rho)^\mu_{(m:m)} + z_{1-\alpha/2} \sqrt{1-\rho_0^2}}{\sqrt{1-\rho^2 + (\rho-\rho_0)^2\sigma^2_{(m:m)}}} \]
\[
1 - \phi \left( \frac{\sqrt{N}(\rho_0 - \rho)u_{(m:m)} + z_{1-\alpha/2}\sqrt{1-\rho_0^2}}{\sqrt{1-\rho^2} + (\rho - \rho_0)^2\sigma^2_{(m:m)}} \right)
\]
\[
- \phi \left( \frac{\sqrt{N}(\rho_0 - \rho)u_{(m:m)} - z_{1-\alpha/2}\sqrt{1-\rho_0^2}}{\sqrt{1-\rho^2} + (\rho - \rho_0)^2\sigma^2_{(m:m)}} \right). \]

Now we calculate the power function of the test statistic \( \eta_{(m)} \) from Chapter 3 where \( \eta_{(m)}N \) = number of observations \( (X_{[m:m]}, Y_{(m:m)}) \) in a sample of size \( N \) such that \( X_{[m:m]} \leq \rho_0 Y_{(m:m)} \). If \( H_0 \) is true, \( \eta_{(m)}N \sim Bin(N,k) \), and if \( N \) is large enough, \( \eta_{(m)}N \sim N(N/2,N/4) \). The power function of this test is

\[
\beta_{\eta_{(m)}}(\rho) = \Pr_{\rho} \text{[the test rejects } H_0] = 1 - \Pr_{\rho} \left[ \frac{N}{2} - \frac{\sqrt{N} z_{1-\alpha/2}}{2} \leq \eta_{(m)} \leq \frac{N}{2} + \frac{\sqrt{N} z_{1-\alpha/2}}{2} \right]
\]

where

\[
p = \Pr_{\rho} [X_{[m:m]} \leq \rho_0 Y_{(m:m)}] = \int_{-\infty}^{\infty} f_{(m:m)}(y) \phi \left( \frac{(\rho_0 - \rho)y}{\sqrt{1-\rho^2}} \right) dy
\]

\[
= E_{Y_{(m:m)}} \left[ \phi \left( \frac{(\rho_0 - \rho)y}{\sqrt{1-\rho^2}} \right) \right].
\]

Thus

\[
\beta_{\eta_{(m)}}(\rho) = 1 - \left[ \phi \left( \frac{\sqrt{N}(k-p) + \frac{1}{2}z_{1-\alpha/2}}{\sqrt{p(1-p)}} \right) - \phi \left( \frac{\sqrt{N}(k-p) - \frac{1}{2}z_{1-\alpha/2}}{\sqrt{p(1-p)}} \right) \right].
\]

To test the one-sided hypothesis

\[
H_0: \rho \geq \rho_0
\]

against \( H_1: \rho < \rho_0 \).
we may still use either test statistic $T_{(m)}$ or $\eta_{(m)}$. We would reject $H_0$ if $T_{(m)N}$ is too small or if $\eta_{(m)}$ is too large; i.e., reject at level $\alpha$ if:

$$T_{(m)N} < -z_{1-\alpha}$$

or if

$$\eta_{(m)N} > \frac{N}{2} + \sqrt{N} z_{1-\alpha}/2,$$

where $\Phi(z_{1-\alpha}) = 1 - \alpha$.

Since once again all conditions of Theorem C are fulfilled, the Pitman efficiency of $T_{(m)}$ with respect to $\eta_{(m)}$, $e_{T_{(m)},\eta_{(m)}}$, can be calculated by observing that

$$\frac{d}{d\rho} \mu_{T_{(m)}}(\rho_0) = \sqrt{N} \frac{\mu_{(m:m)}/\sqrt{1-\rho^2}}{0},$$

$$\sigma_{T_{(m)}}(\rho_0) = 1,$$

$$\frac{d}{d\rho} \mu_{\eta_{(m)}}(\rho_0) = n \frac{d}{d\rho} \left[ \int_{-\infty}^{\infty} f_{(m:m)}(y) \Phi \left( \frac{(\rho_0 - \rho)y}{\sqrt{1-\rho^2}} \right) dy \right]_{\rho=\rho_0}$$

$$= N \left\{ \int_{-\infty}^{\infty} f_{(m:m)}(y) \left[ \Phi \left( \frac{(\rho_0 - \rho)y}{\sqrt{1-\rho^2}} \right) \left( \rho y (\rho_0 - \rho_0) - \frac{y}{\sqrt{1-\rho^2}} \right) \right] dy \right\}_{\rho=\rho_0}$$

$$= -N \frac{\mu_{(m:m)}}{\sqrt{1-\rho^2}} \phi(0),$$

$$\sigma_{\eta_{(m)}}(\rho_0) = \sqrt{N}/2.$$

Thus

$$c_{T_{(m)}} = \lim_{N \to \infty} \frac{\sqrt{N} \mu_{(m:m)}}{\sqrt{N}(1-\rho^2)} = \frac{\mu_{(m:m)}}{\sqrt{1-\rho^2}}$$

and

$$c_{\eta_{(m)}} = \lim_{N \to \infty} \frac{-N\phi(0)\mu_{(m:m)}}{N(1-\rho^2)/2} = \frac{-2\phi(0)\mu_{(m:m)}}{\sqrt{1-\rho^2}}.$$
Therefore
\[ e^2_{T(m)} \eta_{(m)} = \left( \frac{c_{T(m)}}{c_{\eta_{(m)}}} \right)^2 \]
\[ = \frac{1}{4 \Phi^2(0)(1 - \rho_0^2)} \]
\[ = \frac{1.57}{1 - \rho_0^2} . \]

There are, of course, infinitely many test statistics similar to \( \eta_{(m)} \) which could be used to test the hypothesis \( \rho = \rho_0 \). For example, let \( n^{k}_{(m)N} \) be the number of observations in a sample of size \( N \) such that
\[ X^{[m:m]}_{(m:m)} - \rho_0 Y^{[m:m]}_{(m:m)} \leq k \cdot \sqrt{1 - \rho_0^2} \]

We know that if \( H_0 \) is true
\[ n^{k}_{(m)N} \sim \text{Bin}(N, \Phi(k)) . \]

So we construct a size \( \alpha \) test by rejecting \( H_0 \) if
\[ \left| \frac{n_{(m)k} - N\Phi(k)}{(N\Phi(k)(1 - \Phi(k)))^{1/2}} \right| > z_{1-\alpha/2} . \]

The power function of this test is
\[ \beta_{n_{(m)}}(\rho) = 1 - \phi \left( \frac{\sqrt{N}(\Phi(k) - p) + z_{1-\alpha/2}(\Phi(k)(1 - \Phi(k)))^{1/2}}{\sqrt{pq}} \right) \]
\[ - \phi \left( \frac{\sqrt{N}(\Phi(k) - p) - z_{1-\alpha/2}(\Phi(k)(1 - \Phi(k)))^{1/2}}{\sqrt{pq}} \right) \]

where
\[ p = E_{Y^{[m:m]}} \left[ \frac{(\rho_0 - \rho)Y + k\sqrt{1 - \rho_0^2}}{\sqrt{1 - \rho_0^2}} \right] \]
The form of this function makes it difficult to compare with the power function of $\eta_{(m)}$ or to find an optimal value of $k$ for a particular $p$. Suppose, however, that we restrict our definition of optimality to be that value of $k$ which maximizes the efficacy of $\eta^k_{(m)}$ for a given $\rho_0$ and $m$. Then

$$\frac{d}{d\rho} \mu_{(m)}(\rho_0) = \frac{d}{d\rho} N E_{Y_{(m:m)}} \left[ \phi \left( \frac{(\rho_0 - \rho)Y + k\sqrt{1-\rho_0^2}}{\sqrt{1-\rho^2}} \right) \right]$$

and

$$\sigma_{k_{(m)}}(\rho_0) = \sqrt{N\phi(k)(1-\phi(k))}.$$  

So

$$c_{k_{(m)}} = \frac{\phi^2(k)[(\mu_{(m:m)}/\sqrt{1-\rho_0^2}) - (kp_0/(1-\rho_0^2))]^2}{\phi(k)(1-\phi(k))}.$$  

$c_{k_{(m)}}$ can be maximized for any values of $m$ and $\rho_0$. Table 4.2 shows the optimal value of $k$ for various values of $\rho_0$ and $m$ (which we write as $k$), and the Pitman efficiency of $T_{(m)}$ with respect to $\eta^K_{(m)}$, $e_{T_{(m)}}, \eta^K_{(m)}$. 

Note that for $\rho_0 = 0$, the optimal value of $k$ regardless of $m$ is $k = 0$, yielding, as we have seen, $e_{T_{(m)}}, \eta^K_{(m)} = 1.57$. We see from this table that inasmuch as Pitman efficiency is concerned, $T_{(m)}$ is a better test statistic than $\eta^K_{(m)}$ when $|\rho_0|$ is small. As $|\rho_0|$ increases, however, the situation reverses. Note also that for small values of $m$, $\eta^K_{(m)}$'s advantage is more marked.

Graphs of the approximate power functions of $T_{(m)N}$ (from (4.4)) and the power functions of $\eta_{(m)N}$ (from (4.5)), with $N = 25$, for testing
Table 4.2
Optimal Values of $k$ ($\kappa$) and $e_t^{\kappa}(m)$, $\eta^{\kappa}(m)$

<table>
<thead>
<tr>
<th>$\rho_0$</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$\kappa$</td>
<td>1.57</td>
<td>0</td>
</tr>
<tr>
<td>.1</td>
<td>-.45</td>
<td>1.45</td>
<td>-.31</td>
</tr>
<tr>
<td>.3</td>
<td>-.96</td>
<td>.939</td>
<td>-.77</td>
</tr>
<tr>
<td>.5</td>
<td>-1.19</td>
<td>.544</td>
<td>-1.04</td>
</tr>
<tr>
<td>.7</td>
<td>-1.33</td>
<td>.280</td>
<td>-1.23</td>
</tr>
<tr>
<td>.9</td>
<td>-1.45</td>
<td>.0881</td>
<td>-1.40</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\rho_0$</th>
<th>5</th>
<th>6</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$\kappa$</td>
<td>1.57</td>
<td>0</td>
</tr>
<tr>
<td>.1</td>
<td>-.23</td>
<td>1.54</td>
<td>-.20</td>
</tr>
<tr>
<td>.3</td>
<td>-.62</td>
<td>1.33</td>
<td>-.56</td>
</tr>
<tr>
<td>.5</td>
<td>-.91</td>
<td>1.01</td>
<td>-.84</td>
</tr>
<tr>
<td>.7</td>
<td>-1.12</td>
<td>.665</td>
<td>-1.07</td>
</tr>
<tr>
<td>.9</td>
<td>-1.34</td>
<td>.272</td>
<td>-1.30</td>
</tr>
</tbody>
</table>

Note: For negative $\rho$, the signs of $\kappa$ are reversed.
$H_0: \rho = 0$

against $H_0: \rho \neq 0$

at level $\alpha = .05$ are illustrated in Figures 4.1 and 4.2 respectively. The four solid lines of each figure are the power functions for $m = 2, 5, 7, 10$. The dotted line of Figure 4.2 shows, for purposes of comparison, the position of $\beta_{T_{(10)}}(\rho)$. Values of the power functions of $T_{(m)}(\rho)$ and $\eta_{(m)}(\rho)$ ($N = 25, m = 5, 7, 10$) for some particular values of $\rho$ are given in Table 4.3, along with values of the power function of $r$, Pearson's product moment correlation coefficient, the conventional test statistic from a random sample. (Note that $r$ is used as a test statistic when $\mu$ and $\sigma^2$ are unknown, while $T_{(m)}$ and $\eta_{(m)}$ require knowledge of these parameters, although they can be estimated from the data, as we have seen. Hence, judging $r$ by the standards of $\eta_{(m)}$ and $T_{(m)}$ is not entirely appropriate, but is given here, with that in mind, for purposes of a rough comparison.) The values of the power function of $r$ are interpolated according to David's method from her tables [4].

Graphs of the approximate power functions of $T_{(m)N}$ and the power functions $\eta_{(m)N}^{K}, N = 25$, for testing

$$H_0: \rho \geq .5$$

against $H_1: \rho < .5$

and for testing

$$H_0: \rho \geq .9$$

against $H_1: \rho < .9$

all at level $\alpha = .05$, are shown in Figures 4.3, 4.4, 4.5 and 4.6, respectively. (Note that the optimal value of $k$ for each particular $m$ and $\rho_0$ is used for $\eta_{(m)N}^{K}$.) Tables 4.4 and 4.5 give some particular
Power Function of $T_{(m)25}$ for $\rho_0 = 0$

Figure 4.1

Power Function of $\eta_{(m)25}^K$ for $\rho_0 = 0$

Figure 4.2
Power Function of $T_{(m)25}$ for $\rho_0 = .5$

Figure 4.3

Power Function of $n_{(m)25}$ for $\rho_0 = .5$

Figure 4.4
Power Function of $T_{(m)25}$ for $\rho_0 = .9$

Figure 4.5

Power Function of $\eta_{(m)25}^K = \rho_0 = .9$

Figure 4.6
Table 4.3

Power Functions of $T$, $\eta$, and $r$ for $H_0: \rho = 0$ vs. $H_1: \rho \neq 0$
when $N = 25$, $\alpha = .05$

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$\pm .1$</th>
<th>$\pm .3$</th>
<th>$\pm .5$</th>
<th>$\pm .7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{T(5)} (\rho)$</td>
<td>.089</td>
<td>.413</td>
<td>.846</td>
<td>.996</td>
</tr>
<tr>
<td>$\beta_{T(7)} (\rho)$</td>
<td>.103</td>
<td>.528</td>
<td>.939</td>
<td>.999+</td>
</tr>
<tr>
<td>$\beta_{T(10)} (\rho)$</td>
<td>.119</td>
<td>.640</td>
<td>.980</td>
<td>.999+</td>
</tr>
<tr>
<td>$\beta_{\eta(5)} (\rho)$</td>
<td>.074</td>
<td>.279</td>
<td>.667</td>
<td>.963</td>
</tr>
<tr>
<td>$\beta_{\eta(7)} (\rho)$</td>
<td>.082</td>
<td>.358</td>
<td>.804</td>
<td>.996</td>
</tr>
<tr>
<td>$\beta_{\eta(10)} (\rho)$</td>
<td>.092</td>
<td>.443</td>
<td>.902</td>
<td>.999+</td>
</tr>
<tr>
<td>$\beta_{T} (\rho)$</td>
<td>.084</td>
<td>.329</td>
<td>.763</td>
<td>.987</td>
</tr>
</tbody>
</table>

Table 4.4

Power Functions of $T$, $\eta$, and $r$ for $H_0: \rho \geq .5$ vs. $H_1: \rho < .5$
when $N = 25$, $\alpha = .05$

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$.4$</th>
<th>$.2$</th>
<th>$0$</th>
<th>$-.2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{T(5)} (\rho)$</td>
<td>.179</td>
<td>.625</td>
<td>.920</td>
<td>.993</td>
</tr>
<tr>
<td>$\beta_{T(7)} (\rho)$</td>
<td>.208</td>
<td>.727</td>
<td>.969</td>
<td>.999</td>
</tr>
<tr>
<td>$\beta_{T(10)} (\rho)$</td>
<td>.238</td>
<td>.813</td>
<td>.990</td>
<td>.999+</td>
</tr>
<tr>
<td>$\beta_{\eta(5)} (\rho)$</td>
<td>.188</td>
<td>.600</td>
<td>.878</td>
<td>.977</td>
</tr>
<tr>
<td>$\beta_{\eta(7)} (\rho)$</td>
<td>.204</td>
<td>.665</td>
<td>.930</td>
<td>.994</td>
</tr>
<tr>
<td>$\beta_{\eta(10)} (\rho)$</td>
<td>.224</td>
<td>.733</td>
<td>.966</td>
<td>.999</td>
</tr>
<tr>
<td>$\beta_{T} (\rho)$</td>
<td>.149</td>
<td>.507</td>
<td>.841</td>
<td>.975</td>
</tr>
</tbody>
</table>
Table 4.5
Power Functions of $T$, $\eta$, and $r$ for $H_0: \rho \geq .9$ vs. $H_1: \rho < .9$
when $N = 25, \alpha = .05$

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$.8$</th>
<th>$.7$</th>
<th>$.6$</th>
<th>$.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_T$ ($\rho$) $T(5)$</td>
<td>.411</td>
<td>.730</td>
<td>.894</td>
<td>.962</td>
</tr>
<tr>
<td>$\beta_T$ ($\rho$) $T(7)$</td>
<td>.473</td>
<td>.809</td>
<td>.945</td>
<td>.986</td>
</tr>
<tr>
<td>$\beta_T$ ($\rho$) $T(10)$</td>
<td>.535</td>
<td>.872</td>
<td>.974</td>
<td>.996</td>
</tr>
<tr>
<td>$\beta_\eta$ ($\rho$) $\eta(5)$</td>
<td>.663</td>
<td>.919</td>
<td>.981</td>
<td>.994</td>
</tr>
<tr>
<td>$\beta_\eta$ ($\rho$) $\eta(7)$</td>
<td>.694</td>
<td>.940</td>
<td>.990</td>
<td>.997</td>
</tr>
<tr>
<td>$\beta_\eta$ ($\rho$) $\eta(10)$</td>
<td>.725</td>
<td>.957</td>
<td>.995</td>
<td>.999</td>
</tr>
<tr>
<td>$\beta_r$ ($\rho$)</td>
<td>.566</td>
<td>.898</td>
<td>.981</td>
<td>.990</td>
</tr>
</tbody>
</table>

values of these functions, along with those of the power functions of $r$ for these tests.

We see from the graphs and tables that the power of $T_{(m)N}$ appears to be larger for small values of $|\rho_0|$ and smaller for large values of $|\rho_0|$ than that of $\eta_{(m)N}^K$, over all values of $\rho$. This conclusion is the same as that suggested by Pitman efficiency for the power of the two test statistics close to the null hypothesis; i.e., that $T_{(m)}$ is a better test statistic for $|\rho_0|$ small and $\eta_{(m)}^K$ is better for $|\rho_0|$ large.

The use of any $\eta_{(m)}^k$ besides $\eta_{(m)}$ (i.e., letting $k = 0$) for the construction of a confidence interval is questionable, however. For some small values of $m$ and an unfortunate choice of $k$, there will exist at least 2 values of $\rho$ which satisfy the equation
\[ p(\rho) = E_Y \left[ \phi \left( \frac{(\rho_0 - \rho) Y + k\sqrt{1-\rho_0^2}}{\sqrt{1-\rho_0^2}} \right) \right] = \phi(k), \]

One solution to this equation is \( \rho_0 \), the true value of the parameter.

If \( k \) is such that

\[ \frac{\rho_0}{\sqrt{1-\rho_0^2}} = \frac{\mu_{(m:m)}}{k}, \]

then there will exist another value \( \rho \) which satisfies \( p(\rho) = \phi(k) \).

This is because, for such a \( k \),

\[ \frac{d}{d\rho} p(\rho_0) = \phi(k) \left[ \frac{k \rho_0}{(1-\rho_0^2)} - \frac{\mu_{(m:m)}}{\sqrt{1-\rho_0^2}} \right] = 0 \]

and

\[ \frac{d^2}{d\rho^2} p(\rho_0) = \phi(k) \left[ \frac{\mu_{(m:m)}}{\rho_0(1-\rho_0^2)^{3/2}(1+\rho_0)} \right] \neq 0. \]

This tells us that the function \( p(\rho) \) has either a maximum or a minimum at \( \rho = \rho_0 \). But

\[ \lim_{\rho \to 1} p(\rho) = 0 \]
\[ \lim_{\rho \to 1} p(\rho) = 1. \]

Since \( p(\rho) \) is continuous over the interval \( -1 < \rho < 1 \), we know that \( p(\rho) \) must intersect the line \( y = \phi(k) \) at least once more, say at \( \rho = \rho_1 \).
So the test statistic $\eta^{k}_{(m)}$ cannot distinguish between $\rho_0$ and $\rho_1$, since

$$
\eta^{k}_{(m)} \sim \text{Bin}(N, \phi(k)) \text{ if } \rho = \rho_0 \text{ or if } \rho = \rho_1.
$$

This problem does not develop when we use the test statistic $\eta^{(m)}$ (i.e., $k=0$) though, because in that case

$$
\frac{d}{d\rho} P(\rho_0) = \phi(0) \left[ \frac{-\mu_{(m:m)}}{\sqrt{1-\rho^2_0}} \right] < 0 \text{ for } m > 1.
$$
5. COST CONSIDERATIONS

The benefit derived from ranked set sampling depends, of course, not only on the amount of increase in precision of the estimator, but also on the relative cost of the two methods of sampling. Dell and Clutter define:

\( c_s \) - the cost of stratification associated with each quantified observation. This is the cost of drawing \( m-1 \) elements and judgment ordering the \( m \) sample elements.

\( c_q \) - the cost of drawing and quantifying 1 element (without classification).

Then the cost of a random sample of size \( mn = N \) is \( Nc_q \), and of a ranked set sample is \( mn(c_q + cs) \). They then consider an expression for the relative efficiency (RE) of the two procedures, given as the ratio of the variance of the estimate from a random sample and that from a ranked set sample, when the total cost \( C \) is held constant for the two procedures. They show that

\[
RE(\bar{X}, \hat{\mu}_{[m]} n) = \frac{c_q \sigma^2}{1} \frac{1}{m} \sum_{r=1}^{m} \left( c_q + c_s \right) \sigma^2_{[r:m]} \\
= (1 - \frac{1}{m\sigma^2} \sum_{r=1}^{m} \tau^2_{[r:m]})^{-1} \frac{c_q}{c_q + c_s}
\]

where \( \tau_{[r:m]} = \bar{\mu}_{[r:m]} - \mu \). So

\[
RE(\bar{X}, \hat{\mu}_{[m]} n) \geq 1 \text{ if } \frac{1}{m\sigma^2} \sum_{r=1}^{m} \tau^2_{[r:m]} \geq \frac{c_s}{c_q + c_s}.
\]

Since the cost of stratifying a random sample involves choosing
m-1 observations and ordering m, I have modified Dell and Clutter's cost function to reflect the fact that $c_s$ should be dependent on m. Let $c_s = (m-1)c^*$. Then the cost of choosing a ranked set sample of size mn is $mn(c_s + c_q) = mn((m-1)c^* + c_q)$. In that case,

$$\text{RE}(\hat{X_n}, \mu_{[m]n}) = \left(1 - \frac{1}{2\sigma^2} \sum_{r=1}^{m} \tau^2_{[r:m]} \right)^{-1} \frac{c_q}{(m-1)c^* + c_q}$$

and

$$\text{RE}(\hat{X_n}, \mu_{[m]n}) \geq 1 \text{ if } \frac{c_s}{c_q} \leq \frac{1}{m-1} \left[ \frac{\sum_{r=1}^{m} \tau^2_{[r:m]} \sigma^2_{[r:m]} + \tau^2_{[r:m]}^2}{\frac{\mu_4 - \sigma^4}{\mu_4 - \sigma^4}} \right]$$

Consider the "s^2-type" estimators of variance discussed in Section 2.3. We know that

$$\text{Var}\ s^2_n = \frac{1}{mn} \mu_4 - \frac{1}{mn} \sigma^4 + O\left(\frac{1}{mn^2}\right), \text{ where } \mu_4 = \text{E}(X-\mu)^4$$

and $s^2_n$ is the sample variance from a random sample and

$$\text{MSE}\ s^2_{[m]n} = \frac{1}{mn} \mu_4 - \frac{1}{m^2 n} \sum_{r=1}^{m} \left(\sigma^2_{[r:m]} + \tau^2_{[r:m]}\right)^2 + O\left(\frac{1}{mn^2}\right).$$

Thus for large samples

$$\text{RE}(s^2_n, s^2_{[m]n}) = \frac{c_q}{c^* + (m-1)c^*} \frac{(\mu_4 - \sigma^4)}{(\mu_4 - \sigma^4)} \frac{1}{m} \sum_{r=1}^{m} \left(\sigma^2_{[r:m]} + \tau^2_{[r:m]}\right)^2$$

= \left(\frac{c_q}{c^* + (m-1)c^*} \frac{(\mu_4 - \sigma^4)}{(\mu_4 - \sigma^4)} \right) \frac{1}{m} \sum_{r=1}^{m} \left(\sigma^2_{[r:m]} + \tau^2_{[r:m]}\right)^2.$$

Therefore

$$\text{RE}(s^2_n, s^2_{[m]n}) \geq 1 \text{ if } \frac{c_s}{c_q} \leq \frac{1}{m-1} \left[ \frac{\sigma^4 - \frac{1}{m} \sum_{r=1}^{m} \left(\sigma^2_{[r:m]} + \tau^2_{[r:m]}\right)^2}{\mu_4 - \sigma^4} \right]$$
and if \( n \) is sufficiently large.

Now suppose \((X,Y) \sim N(\mu, \begin{pmatrix} \sigma_x^2 & \rho \sigma_x \sigma_y \\ \rho \sigma_x \sigma_y & \sigma_y^2 \end{pmatrix})\). If \( \mu_x, \mu_y \), \( \sigma_x^2, \sigma_y^2 \) are known, we saw in Section 3.1 that we could calculate an MLE of \( \rho \) from either a modified ranked set or a random sample. Then from (3.4) and (3.5), we see that for large samples

\[
RE(\hat{\rho}_N, \hat{\sigma}^N) = \frac{c_q(1-\rho^2)}{(m-1)c_s^*+c_q}(1-\rho^2) \frac{1}{E(U_{[m:m]}^2)}
\]

\[
= \frac{c_q}{(m-1)c_s^*+c_q} \left(1-\frac{\rho^2}{1+\rho^2}\right) EU_{[m:m]}^2
\]

(5.1)

where \( U \sim N(0,1) \). Thus for large \( N \),

\[
RE(\hat{\rho}_N, \hat{\sigma}^N) \geq 1 \text{ if } \frac{c_q}{c_s^*+c_q} \leq \frac{1}{m-1} \frac{1-\rho^2}{1+\rho^2} \frac{EU_{[m:m]}^2}{2}.\]

If only \( \mu_y \) and \( \sigma_y^2 \) are known, then the MLE's can still be computed. (See Section 3.2.) From (3.11) and (3.12), we see that for large values of \( N \),

\[
RE(\tilde{\rho}_N, \tilde{\sigma}^N) = \frac{c_q(2-\rho^2)}{2} \frac{2(1-\rho^2)+\rho^2 EU_{[m:m]}^2}{2EU_{[m:m]}^2}
\]

\[
= \frac{c_q}{(m-1)c_s^*+c_q} \frac{EU_{[m:m]}^2(2-\rho^2)}{2(1-\rho^2)+\rho^2 EU_{[m:m]}^2}
\]

(5.2)

Thus for large \( N \),

\[
RE(\tilde{\rho}_N, \tilde{\sigma}^N) \geq 1 \text{ if } \frac{c_q}{c_s^*+c_q} \leq \frac{1}{m-1} \frac{2(1-\rho^2)(EU_{[m:m]}^2)-1}{2(1-\rho^2)+\rho^2 EU_{[m:m]}^2}.
\]

If the form of the distribution is known and if perfect ranking
is possible, one can determine if ranked set sampling is beneficial as far as relative efficiencies for equal cost are concerned. Tables 5.1 and 5.2 display $\text{RE}(\hat{\mu}_{\text{(m)n}})$ and $\text{RE}(\hat{s}_{\text{(m)n}})$ respectively for several values of $m$ and $c_{s*}/c_q$ for the normal and uniform distributions. Tables 5.3 and 5.4 display $\text{RE}(\hat{\rho}_{\text{N},\rho_{\text{(m)n}}})$ (all parameters known) and $\text{RE}(\tilde{\rho}_{\text{N},\rho_{\text{(m)n}}})$ (only $\mu, \sigma^2$ known).

One can observe from the tables that $\text{RE}(\hat{\rho}_{\text{N},\rho_{\text{(m)n}}})$ and $\text{RE}(\tilde{\rho}_{\text{N},\rho_{\text{(m)n}}})$ begin to decline with $m$ for $\rho = \pm 5$ and $\rho = \pm 75$ respectively when $c_{s*}/c_q = .1$. This is true not only for these particular values of $\rho$ and $c_{s*}/c_q$, but for all $\rho$, $c_{s*}/c_q \neq 0$, since, from (5.1),

$$\text{RE}(\hat{\rho}_{\text{N},\rho_{\text{(m)n}}}) = O(\ln m) \to 0 \text{ as } m \to \infty,$$

and from (5.2)

$$\text{RE}(\tilde{\rho}_{\text{N},\rho_{\text{(m)n}}}) = O(\frac{1}{m}) \to 0 \text{ as } m \to \infty,$$

since

$$E\text{U}_{\text{(m:n)}}^2 = O(\ln m).$$

However, if the major expense in the stratification step is the ranking process itself, rather than choosing the (m-1) observations, it is quite likely that $c_s$ is not a linear function in $m$ of $c_{s*}$, but rather increases more slowly with $m$. If this is the case, the relative efficiencies may not decrease to 0. In fact, if $c_s/c_{s*} = o(\ln m)$, then $\text{RE}(\hat{\rho}_{\text{N},\rho_{\text{(m)n}}}) \to \infty$ and $\text{RE}(\tilde{\rho}_{\text{N},\rho_{\text{(m)n}}}) \to \infty$ as $m \to \infty$.

If the form of the distribution is not known, or if the samples are judgment ordered, then a rough idea of the benefit derived from ranked set sampling may be obtained in some cases by estimating the relative efficiencies. For example, let $A = \sum_{r=1}^{m} (\bar{X}_r - \mu_{[r:m]})^2$ where

$$\bar{X}_r = \frac{1}{n} \sum_{i=1}^{n} X_{[r:m],i} \text{ and let } B = \sum_{i=1}^{n} \sum_{T=1}^{m} (X_{ri} - \bar{X}_r)^2.$$

Then
EA = \frac{m-1}{nm} \frac{\sum_{r=1}^{m} (\sigma^2_{[r:m]} + \tau^2_{[r:m]})}{n} \quad \text{and} \quad EB = (n-1) \frac{\sum_{r=1}^{m} \sigma^2_{[r:m]}}{m}.

So A - \frac{(m-1)}{nm(n-1)} B \text{ is an unbiased estimator of } \frac{\sum_{r=1}^{m} \tau^2_{[r:m]}}{n}. \text{ If } \sigma^2 \text{ is not known, it too must be estimated, and the two estimates can be used to get an idea if } \text{RE}(\bar{X}_{N}, \mu_{[m]n}) \geq 1. \text{ Similarly, if } \hat{\rho}^\Delta_{(m)N} \text{ and } \tilde{\rho}^\Delta_{(m)N} \text{ are based on judgment order statistics, then } \text{EU}^2_{(m:m)} \text{ in (5.1) and (5.2) is replaced by } \frac{(\sigma^2_{y[r:m]} + \tau^2_{y[r:m]})/\sigma^2_{y}}{n} \text{ in the expression for relative efficiencies. Then } \frac{1}{N-1} \sum_{i=1}^{N} (Y_{[m:m],i} - \mu_{y})^2 \text{ is an unbiased estimator of } \sigma^2_{[m:m]} + \tau^2_{[m:m]}, \text{ allowing us to estimate the relative efficiencies.}

### Table 5.1

\text{RE}(\bar{X}_{N}, \mu_{(m)n})

<table>
<thead>
<tr>
<th>m</th>
<th>\text{c}_{s*}/c_q = .01</th>
<th>\text{c}_{s*}/c_q = .1</th>
<th>\text{c}_{s*}/c_q = .01</th>
<th>\text{c}_{s*}/c_q = .1</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.45</td>
<td>1.33</td>
<td>1.49</td>
<td>1.37</td>
</tr>
<tr>
<td>3</td>
<td>1.88</td>
<td>1.59</td>
<td>1.96</td>
<td>1.67</td>
</tr>
<tr>
<td>4</td>
<td>2.28</td>
<td>1.81</td>
<td>2.43</td>
<td>1.93</td>
</tr>
<tr>
<td>5</td>
<td>2.66</td>
<td>1.98</td>
<td>2.89</td>
<td>2.14</td>
</tr>
<tr>
<td>7</td>
<td>3.40</td>
<td>2.25</td>
<td>3.79</td>
<td>2.50</td>
</tr>
<tr>
<td>10</td>
<td>4.40</td>
<td>2.52</td>
<td>5.05</td>
<td>2.89</td>
</tr>
</tbody>
</table>

### Table 5.2

\text{RE}(s^2_{N}, s^2_{(m)n})

<table>
<thead>
<tr>
<th>m</th>
<th>\text{c}_{s*}/c_q = .01</th>
<th>\text{c}_{s*}/c_q = .1</th>
<th>\text{c}_{s*}/c_q = .01</th>
<th>\text{c}_{s*}/c_q = .1</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>.90</td>
<td>.91</td>
<td>.99</td>
<td>.91</td>
</tr>
<tr>
<td>3</td>
<td>1.06</td>
<td>.90</td>
<td>1.09</td>
<td>.93</td>
</tr>
<tr>
<td>4</td>
<td>1.15</td>
<td>.91</td>
<td>1.21</td>
<td>.96</td>
</tr>
<tr>
<td>5</td>
<td>1.23</td>
<td>.91</td>
<td>1.35</td>
<td>1.00</td>
</tr>
<tr>
<td>7</td>
<td>1.37</td>
<td>.91</td>
<td>1.62</td>
<td>1.07</td>
</tr>
<tr>
<td>10</td>
<td>1.62</td>
<td>.93</td>
<td>2.02</td>
<td>1.16</td>
</tr>
</tbody>
</table>
Table 5.3
RE(\hat{\rho}_N, \hat{\rho}_N^{\Delta})

<table>
<thead>
<tr>
<th>m</th>
<th>(c_{s^*}/c_q = .01)</th>
<th>(c_{s^*}/c_q = .1)</th>
<th>(c_{s^*}/c_q = .01)</th>
<th>(c_{s^*}/c_q = .1)</th>
<th>(c_{s^*}/c_q = .01)</th>
<th>(c_{s^*}/c_q = .1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>.99</td>
<td>.91</td>
<td>.99</td>
<td>.91</td>
<td>.99</td>
<td>.91</td>
</tr>
<tr>
<td>3</td>
<td>1.25</td>
<td>1.07</td>
<td>1.15</td>
<td>.97</td>
<td>1.06</td>
<td>.90</td>
</tr>
<tr>
<td>4</td>
<td>1.51</td>
<td>1.19</td>
<td>1.29</td>
<td>1.02</td>
<td>1.12</td>
<td>.89</td>
</tr>
<tr>
<td>5</td>
<td>1.73</td>
<td>1.29</td>
<td>1.42</td>
<td>1.06</td>
<td>1.17</td>
<td>.87</td>
</tr>
<tr>
<td>7</td>
<td>2.09</td>
<td>1.39</td>
<td>1.63</td>
<td>1.08</td>
<td>1.26</td>
<td>.84</td>
</tr>
<tr>
<td>10</td>
<td>2.49</td>
<td>1.43</td>
<td>1.86</td>
<td>1.07</td>
<td>1.36</td>
<td>.78</td>
</tr>
</tbody>
</table>

Table 5.4
RE(\tilde{\rho}_N, \tilde{\rho}_N^{\Delta})

<table>
<thead>
<tr>
<th>m</th>
<th>(c_{s^*}/c_q = .01)</th>
<th>(c_{s^*}/c_q = .1)</th>
<th>(c_{s^*}/c_q = .01)</th>
<th>(c_{s^*}/c_q = .1)</th>
<th>(c_{s^*}/c_q = .01)</th>
<th>(c_{s^*}/c_q = .1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>.99</td>
<td>.91</td>
<td>.99</td>
<td>.91</td>
<td>.99</td>
<td>.91</td>
</tr>
<tr>
<td>3</td>
<td>1.25</td>
<td>1.07</td>
<td>1.20</td>
<td>1.02</td>
<td>1.13</td>
<td>.96</td>
</tr>
<tr>
<td>4</td>
<td>1.51</td>
<td>1.19</td>
<td>1.40</td>
<td>1.11</td>
<td>1.24</td>
<td>.98</td>
</tr>
<tr>
<td>5</td>
<td>1.73</td>
<td>1.29</td>
<td>1.55</td>
<td>1.15</td>
<td>1.32</td>
<td>.98</td>
</tr>
<tr>
<td>7</td>
<td>2.09</td>
<td>1.39</td>
<td>1.78</td>
<td>1.18</td>
<td>1.42</td>
<td>.94</td>
</tr>
<tr>
<td>10</td>
<td>2.49</td>
<td>1.43</td>
<td>2.00</td>
<td>1.15</td>
<td>1.49</td>
<td>.85</td>
</tr>
</tbody>
</table>
Theorem A [3]. Let \( X_1, \ldots, X_n \) be a random sample of iid random variables, each with p.d.f. \( f_\theta(x), \theta \in \Theta \). Consider a transformation from the set of variables \( X_1, \ldots, X_n \) to a new set of variables \( \xi_1, \ldots, \xi_{n-1}, \hat{\theta} \). Let \( J \) be the Jacobian of the transformation. Let \( g, h \) be probability density functions s.t.

\[
\prod_{i=1}^{n} f_\theta(X_i) |J| = g_\theta(\hat{\theta}) h(\xi_1, \ldots, \xi_{n-1} | \hat{\theta}, \theta).
\]

Suppose that for almost all values of \( X, \hat{\theta}, \xi_1, \ldots, \xi_{n-1} \), the partial derivatives

\[
\left| \frac{\partial f}{\partial \theta} \right| < F_0(x), \quad \left| \frac{\partial g}{\partial \theta} \right| < G_0(\hat{\theta}), \quad \left| \frac{\partial h}{\partial \theta} \right| < H_0(\xi_1, \ldots, \xi_{n-1}|\hat{\theta})
\]

where \( F_0, G_0, \hat{\theta} G_0, \) and \( H_0 \) are integrable over their respective spaces.

Then

\[
E(\hat{\theta} - \theta)^2 \geq \frac{\left( 1 + \frac{d b(\theta)}{d \theta} \right)^2}{n \int_{-\infty}^{\infty} \left( \frac{\partial}{\partial \theta} \ln f_\theta(x) \right)^2 f_\theta(x) dx}
\]

where \( b(\theta) = E(\hat{\theta} - \theta) \). The sign of equality holds here, for every \( \theta \in \Theta \), when and only when the following two conditions are satisfied whenever \( g_\theta(\hat{\theta}) > 0 \):

A) \( h(\xi_1, \ldots, \xi_{n-1} | \hat{\theta}, \theta) \) is independent of \( \theta \).

B) We have \( \frac{\partial}{\partial \theta} \ln g = k(\hat{\theta} - \theta) \), where \( k \) is independent of \( \hat{\theta} \), but may depend on \( \theta \).

For a rigorous proof of this last statement, see Wijsman [23].
Theorem B [3]. Let \( X_1, \ldots, X_n \) be a sample of size \( n \) where \( f_\theta(x) \) is the p.d.f. of \( X_1 \). Suppose that the following conditions are satisfied:

1. For almost all \( x \), the derivatives \( \frac{\partial \ln f}{\partial \theta}, \frac{\partial^2 \ln f}{\partial \theta^2}, \frac{\partial^3 \ln f}{\partial \theta^3} \) exist \( \forall \theta \) belonging to a non-degenerate interval \( \Theta \).

2. For every \( \theta \in \Theta \), we have

\[
\left| \frac{\partial}{\partial \theta} \right| < F_1(x), \quad \left| \frac{\partial^2}{\partial \theta^2} \right| < F_2(x), \quad \text{and} \quad \left| \frac{\partial^3}{\partial \theta^3} \right| < H(x),
\]

the functions \( F_1 \) and \( F_2 \) being integrable over \((-\infty, \infty)\),

while \( \int_{-\infty}^{\infty} H(x)f_\theta(x)dx < M \), where \( M \) is independent of \( \theta \).

(A.1)

3. For every \( \theta \in \Theta \),

\[
\int_{-\infty}^{\infty} \left( \frac{\partial \ln f}{\partial \theta} \right)^2 f \, dx \text{ is finite and positive.}
\]

Then there exists a solution to the likelihood equation, and it converges in probability to the true value of \( \theta \) as \( n \to \infty \). This solution is an asymptotically normal and asymptotically efficient estimator of \( \theta \).

LeCam [14] has replaced the conditions of (A.1) with a set of less stringent conditions which do not involve or imply the existence of derivatives higher than the first.

Theorem C ([13], p. 276 and 285). Let \( T_1 = \{ T_{1N} \} \) and \( T_2 = \{ T_{2N} \} \) be two sequences of test statistics for testing \( H_0: \theta = \theta_0 \) against \( H_1: \theta \neq \theta_0 \), each based on \( N \) observations, where \( N \to \infty \). Let

\[
\mu(T_{1N}|\theta) = ET_{1N} \quad \text{when the parameter is } \theta,
\]

\[
\sigma(T_{1N}|\theta) = \text{standard deviation of } T_{1N} \quad \text{when the parameter is } \theta,
\]

for \( i = 1, 2 \).
Suppose that for all \( k > 0, \ i = 1, 2 \):

1. \( \lim_{{N \to \infty}} \frac{\frac{\partial}{\partial \theta} \mu(T_{iN} | \theta = \theta_0 + k/\sqrt{N})}{\sqrt{N} \sigma(T_{iN} | \theta = \theta_0)} = 1 \)

2. \( \lim_{{N \to \infty}} \frac{\frac{\partial}{\partial \theta} \mu(T_{iN} | \theta = \theta_0 + k/\sqrt{N})}{\sqrt{N} \sigma(T_{iN} | \theta = \theta_0)} = c_i \)

3. \( \lim_{{N \to \infty}} \frac{\sigma(T_{iN} | \theta = \theta_0 + k/\sqrt{N})}{\sigma(T_{iN} | \theta = \theta_0)} = 1 \)

4. \( \lim_{{N \to \infty}} \frac{\Pr[\frac{T_{iN} - \mu(T_{iN} | \theta = \theta_0 + k/\sqrt{N})}{\sigma(T_{iN} | \theta = \theta_0 + k/\sqrt{N})} \leq z | \theta = \theta_0 + k/\sqrt{N}] = G(z) \)

for a fixed continuous distribution function \( G \).

Then \( e_{T_1, T_2} = \left( \frac{c_1}{c_2} \right)^2 \).

1. Show that under appropriate regularity conditions

\[
I_1^*(\sigma) = \frac{N}{\sigma^2} E(f')^2 + \frac{N(m-1)}{\sigma^2} E\left( \frac{f^2}{f(1-F)} \right).
\]

**Proof:**

\[
I_1^*(\sigma) = E\left( \frac{\partial \ln L^*}{\partial \theta} \right)^2
\]

where

\[
\ln L^* = -nm \ln \sigma + \sum_{i=1}^{n} \sum_{r=1}^{m} \ln f\left( \frac{X(r:m)}{\sigma} \right) - \mu
\]

\[
+ \sum_{i=1}^{n} \sum_{r=1}^{m} (r-1) \ln F\left( \frac{X(r:m) - \mu}{\sigma} \right)
\]

\[
+ \sum_{i=1}^{n} \sum_{r=1}^{m} (m-r) \ln \left[ 1 - F\left( \frac{X(r:m) - \mu}{\sigma} \right) \right]
\]

+ terms not involving \( \sigma \).
So

\[
\frac{\partial}{\partial \sigma} \ln L^* = -\frac{nm}{\sigma} + \sum_{i} \frac{f'(U_{ri}) (U_{ri})}{f(U_{ri})} \sum_{r} \frac{f(U_{ri}) (U_{ri})}{F(U_{ri})} + \sum_{i} \frac{f(U_{ri}) (U_{ri})}{1-F(U_{ri})} \sum_{r} (r-1) \frac{f(U_{ri}) (U_{ri})}{F(U_{ri})} 
\]

Suppose \( L^* \) is such that

\[
E\left( \frac{\partial \ln L^*}{\partial \sigma} \right)^2 = -E\left( \frac{\partial^2 \ln L^*}{\partial \sigma^2} \right).
\]

Then we have

\[
\frac{\partial^2 \ln L^*}{\partial \sigma^2} = \frac{nm}{\sigma^2} \sum_{i} \frac{f''(U_{ri}) U_{ri}^2}{f(U_{ri})} - \frac{1}{\sigma^2} \sum_{i} \left[ \frac{f'(U_{ri}) U_{ri}}{f(U_{ri})} \right]^2 
\]

\[
+ \frac{2}{\sigma^2} \sum_{i} \left[ \frac{f'(U_{ri}) U_{ri}}{f(U_{ri})} \right] 
\]

\[
+ \frac{2}{\sigma^2} \sum_{i} \left[ \frac{f(U_{ri}) U_{ri}}{1-F(U_{ri})} \right] 
\]

\[
- \frac{2}{\sigma^2} \sum_{i} \left[ \frac{f(U_{ri}) U_{ri}}{1-F(U_{ri})} \left[ \frac{U_{ri}^2 f'(U_{ri})}{1-F(U_{ri})} \right] \right] 
\]

\[
+ \frac{1}{\sigma^2} \sum_{i} \left[ \frac{U_{ri}^2 f'(U_{ri})}{1-F(U_{ri})} \right] \right] \right] 
\]

So

\[
E\left( \frac{\partial^2 \ln L^*}{\partial \sigma^2} \right) = \frac{nm}{\sigma^2} + \frac{nm}{\sigma^2} E\left( f'U \right)^2 - \frac{nm}{\sigma^2} E\left( \frac{f''U^2}{f} \right) - \frac{nm(m-1)}{\sigma^2} E\left( U^2 f' \right) + \frac{nm(m-1)}{\sigma^2} E\left( \frac{U^2 f^2}{F} \right) - \frac{2nm(m-1)}{\sigma^2} E\left( U f \right) + \frac{2nm(m-1)}{\sigma^2} E\left( U^2 f \right) + \frac{nm(m-1)}{\sigma^2} E\left( U^2 f \right) + \frac{nm(m-1)}{\sigma^2} E\left( \frac{U^2 f^2}{1-F} \right)
\]
\[
= \frac{nm}{\sigma^2} \left( 1 + E\left( f' \frac{f''}{f} u^2 \right) - E\left( f'' \frac{u^2}{f^2} \right) \right) + \frac{nm(m-1)}{\sigma^2} E\left[ \frac{u^2}{f(1-F)} \right]. \tag{A.2}
\]

If we assume that two moments exist, we can show that \( E\left( \frac{f''}{2} u^2 \right) = 2 \).

\[
E\left( \frac{f''}{f} u^2 \right) = \int_{-\infty}^{\infty} \frac{f''(u)}{f(u)} u^2 f(u) du = \int_{-\infty}^{\infty} u^2 f''(u) du
\]

\[
= u^2 f'(u) \bigg|_{-\infty}^{+\infty} - 2 \int_{-\infty}^{\infty} f'(u) u du. \tag{A.3}
\]

Consider the first term of (A.3)

\[
\int_{-\infty}^{\infty} u^2 f'(u) du = u^2 f(u) \bigg|_{-\infty}^{+\infty} - 2 \int_{-\infty}^{\infty} uf(u) du < \infty
\]

since first and second moments exist. Therefore,

\[
\lim_{u \to \infty} u^2 f'(u) = \lim_{u \to -\infty} u^2 f'(u) = 0.
\]

Now consider the second term of (A.3)

\[
\int_{-\infty}^{\infty} f'(u) du = uf(u) \bigg|_{-\infty}^{+\infty} - \int_{-\infty}^{\infty} f(u) du = 0 - 1
\]

since the first moment exists. Therefore, we have \( E\left( \frac{f''}{f} u^2 \right) = -2(-1) = 2 \) and from (A.2)

\[
I_N' = \frac{nm}{\sigma^2} \left[ E\left( f' \frac{f''}{f} u^2 \right) \right] + \frac{nm(m-1)}{\sigma^2} E\left[ \frac{u^2}{f(1-F)} \right].
\]

2. We now show that Tiku's approximation allows an explicit approximate solution to equations (3.15) to be found when \( m \) is large enough. Recall that this approach is based on Tiku's linear approximation of the function

\[
f(z) = \frac{\phi(z)}{\Phi(z)}
\]

where \( \phi, \Phi \) are the p.d.f. and c.d.f. of the unit normal distribution;
i.e., we have

\[ g(z) \approx \alpha + \beta z \quad \text{where} \quad \beta = \frac{g(h) - g(\ell)}{h - \ell}, \quad \alpha = g(h) - h\beta, \]

and \((\ell, h)\) is an interval containing \(z\). Note that \((Y_{mi} - \mu_y)/\sigma_y\) and \(-(Y_{i1} - \mu_y)/\sigma_y\) have the same distribution; that of \(Z_{(m:m)}\), where \(Z \sim N(0,1)\). Note also that \((\phi(z))/(1 - \Phi(z)) = g(-z)\).

Now if \(m\) is fairly large, then \(\text{Var } Z_{(m:m)}\) is fairly small (for \(m=10\), say, \(\text{Var } Z_{(10:10)} = .344\)). Then \((EZ_m - \kappa \sqrt{\text{Var } Z_m}, EZ_m + \kappa \sqrt{\text{Var } Z_m})\) will have a high probability of covering \((Y_{m} - \mu_y)/\sigma_y\) and \(-(Y_{i1} - \mu_y)/\sigma_y\) if \(\kappa\) is large enough (but still small enough to keep the linear approximation valid). Then we see that (3.13) simplifies to:

\[
\frac{\bar{x} - \mu_x}{\sigma_x} = \frac{\bar{y} - \mu_y}{\sigma_y} = \rho
\]

and

\[
\frac{\bar{y} - \mu_y}{\sigma_y} = \rho \frac{\bar{x} - \mu_x}{\sigma_x} + \frac{(m-1)(1-\rho^2)}{N} \left\{ \frac{N}{2} \sum_{i=1}^{N/2} \left[ (\alpha_m - \beta_m) \cdot \frac{Y_{i1} - \mu_y}{\sigma_y} - (\alpha_m + \beta_m) \cdot \frac{Y_{mi} - \mu_y}{\sigma_y} \right] \right\}
\]

\[
= \rho \frac{\bar{x} - \mu_x}{\sigma_x} + \frac{(m-1)(1-\rho^2)}{N\sigma_y} \beta_m \sum_{i=1}^{N/2} \left[ 2\mu_y - (Y_{i1} + Y_{mi}) \right].
\]

(A.4)

(A.5)

After substituting (A.4) into (A.5), we have

\[
\bar{y}(1-(m-1)\beta_m) = \mu_y(1-(m-1)\beta_m),
\]

yielding the only solution to (A.4) and (A.5) as

\[
\hat{\mu}_y = \bar{y}, \quad \hat{\mu}_x = \bar{x}.
\]
Table A.1

Means and Variances of $\hat{\rho}$

<table>
<thead>
<tr>
<th>N</th>
<th>$E \hat{\rho}$</th>
<th>$Var \hat{\rho}$</th>
<th>$E \hat{\rho}$</th>
<th>$Var \hat{\rho}$</th>
<th>$E \hat{\rho}$</th>
<th>$Var \hat{\rho}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td>$\rho = 0$</td>
<td></td>
<td>$\rho = 0.5$</td>
<td></td>
<td>$\rho = 0.75$</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-0.024</td>
<td>0.056</td>
<td>-0.006</td>
<td>0.027</td>
<td>-0.007</td>
<td>0.019</td>
</tr>
<tr>
<td>7</td>
<td>-0.003</td>
<td>0.053</td>
<td>-0.004</td>
<td>0.025</td>
<td>0.003</td>
<td>0.015</td>
</tr>
<tr>
<td>10</td>
<td>-0.004</td>
<td>0.049</td>
<td>-0.001</td>
<td>0.019</td>
<td>0.005</td>
<td>0.013</td>
</tr>
<tr>
<td>5</td>
<td>0.497</td>
<td>0.034</td>
<td>0.492</td>
<td>0.016</td>
<td>0.498</td>
<td>0.011</td>
</tr>
<tr>
<td>7</td>
<td>0.513</td>
<td>0.031</td>
<td>0.511</td>
<td>0.016</td>
<td>0.509</td>
<td>0.010</td>
</tr>
<tr>
<td>10</td>
<td>0.520</td>
<td>0.027</td>
<td>0.509</td>
<td>0.012</td>
<td>0.508</td>
<td>0.008</td>
</tr>
<tr>
<td>5</td>
<td>0.750</td>
<td>0.014</td>
<td>0.742</td>
<td>0.006</td>
<td>0.750</td>
<td>0.004</td>
</tr>
<tr>
<td>7</td>
<td>0.760</td>
<td>0.013</td>
<td>0.760</td>
<td>0.006</td>
<td>0.755</td>
<td>0.004</td>
</tr>
<tr>
<td>10</td>
<td>0.765</td>
<td>0.011</td>
<td>0.756</td>
<td>0.005</td>
<td>0.754</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Table A.2

Ratio of Cramer-Rao lower bound of $Var \hat{\rho}$ from a random sample to $Var \hat{\rho}$

<table>
<thead>
<tr>
<th>N</th>
<th>10</th>
<th>20</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td>$\rho = 0$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.78</td>
<td>1.84</td>
<td>1.76</td>
</tr>
<tr>
<td>7</td>
<td>1.89</td>
<td>2.02</td>
<td>2.20</td>
</tr>
<tr>
<td>10</td>
<td>2.06</td>
<td>2.61</td>
<td>2.55</td>
</tr>
<tr>
<td>$\rho = 0.5$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.63</td>
<td>1.80</td>
<td>1.66</td>
</tr>
<tr>
<td>7</td>
<td>1.80</td>
<td>1.81</td>
<td>1.86</td>
</tr>
<tr>
<td>10</td>
<td>2.06</td>
<td>2.43</td>
<td>2.25</td>
</tr>
</tbody>
</table>
Table A.2 - Cont'd

<table>
<thead>
<tr>
<th>N</th>
<th>10</th>
<th>20</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ρ = ±.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.35</td>
<td>1.52</td>
<td>1.53</td>
</tr>
<tr>
<td>7</td>
<td>1.47</td>
<td>1.65</td>
<td>1.60</td>
</tr>
<tr>
<td>10</td>
<td>1.68</td>
<td>1.96</td>
<td>1.84</td>
</tr>
</tbody>
</table>
BIBLIOGRAPHY


