A COMPARISON OF ESTIMATORS OF THE GAIN IN EFFICIENCY ACHIEVED BY IMPORTANCE SAMPLING

by

PETER MARK CUMBERBIRCH

B.Sc., University of Victoria, 1971

A THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in the Department

of

Mathematics

We accept this thesis as conforming

to the required standard

[redacted]

UNIVERSITY OF VICTORIA

April 1975

Supervisor: Dr. W. K. Hastings

ABSTRACT

The problem investigated is that of estimating the 'relative gain' in efficiency achieved by importance, or probability proportional to estimated size (ppes) sampling, a variance reduction technique. Monte Carlo sampling experiments are conducted in an exploratory study of several estimation procedures, each of which originates from either a one-sample or a two-sample approach to the problem.

The two-sample approach reduces to the usual variance comparison problem, although the sampled distributions are somewhat more restricted in this case. Assumptions of normality and of identical distributions, except for location and scale, however, are violated; consequently robust estimation procedures are required. Jackknife techniques, both with and without the natural log transformation, as well as slightly modified versions of the Box and Box-Andersen methods of variance comparison, are implemented in the Monte Carlo study. For comparative purposes, the classical F method is also included.

The two-sample Monte Carlo results indicate that the jackknife technique with the natural log transformation, and the Box method with subsample 10, are the leading competitors among the two-sample procedures. The jackknife proves to be the more powerful of the two, but tends to yield significance levels below the nominal level. The Box method, on the other hand, is more conservative and provides significance levels much nearer the nominal level.

For the one-sample approach, P. V. Sukhatme and B. V. Sukhatme have suggested an estimation technique based on a single sample which yields an estimator that allows negative estimates of the variance ratio to occur. As an alternative to the Sukhatme estimator, a non-negative, one-sample estimator is proposed.

The Sukhatme estimator and jackknifed versions of the Sukhatme estimator, the non-negative estimator and the natural log of the non-negative estimator are compared. Results of the Monte Carlo experiments for the one-sample procedures tend to indicate that the non-negative estimator is superior to the Sukhatme estimator.

[redacted]

TABLE OF CONTENTS

CHAPTER		PAGE

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

IMPORTANCE SAMPLING

1.1 Introduction

In sample survey problems, the parameter of interest is often a population total or mean. To estimate this quantity, the statistician is responsible for selecting the most appropriate sampling procedure. Determining the 'best' approach to the problem requires that a variety of factors be considered. Perhaps the most restrictive of these are the limitations imposed by monetary considerations. As a result, the statistician must design a survey which minimizes expenditures while retaining an acceptable degree of precision.

In practice the statistician will often define a cost function for the estimation procedures under consideration. This function attempts to account for the major factors contributing to the overall expense of the survey and in its simplest form is a linear function of the sample size (see Cochran [6, p.82]). The procedure which minimizes the cost while retaining the desired precision is then chosen. To determine which of the estimation procedures minimizes the cost, reliable estimates of the relative efficiency of the competing estimators are essential. That is, estimates of the relative magnitudes of the variances of the estimates are required. These estimates are often obtained by conducting pilot surveys or by studying the results of previous surveys.

This thesis is concerned with the problem of estimating the 'relative efficiency' of estimates of the mean which arise from importance, or probability proportional to estimated size (ppes) sampling, a variance reduction technique. In its discrete form, importance sampling involves selecting the sampling units with probability proportional to some measure of the magnitude of the characteristic of interest, as opposed to the usual method of sampling with equal probability. An unbiased statistic, in the form of a weighted sample mean, is then used to estimate the population mean. If the 'measure of size' is positively correlated with the characteristic of interest, a substantial variance reduction may result. The following example illustrates the probability proportional to estimated size (ppes) sampling method.

Example 1. Consider the problem of estimating the mean (or total) annual dollar volume sales of retail stores in a given area. From previous surveys it is known that the distribution of sales is often skewed. That is, a small percentage of the stores tend to account for a large percentage of the total sales. Simple random sampling is obviously not the most efficient sampling procedure in this case; a properly designed sample would include a high proportion of all stores with sales over a certain size.

Additional information, which can be utilized to improve the sampling procedure, is often available. If, for example, a record of the number of employees could be obtained for each of the stores, then, since positive correlation is likely to exist between the number of employees and the volume of sales, this information could be used as a 'measure' of the volume of sales for each store. Suppose that there is a total of M stores and that the i^{th} store has w_i employees. Then, if W is the total number of employees working in the M stores, the probability proportional to estimated size (ppes) sampling method would require that the probability of selecting an arbitrary store, the j^{th} say, be $P_j = w_j/W$. That is, the stores would be selected with probability proportional to the number of employees working in them. Note that if the sales and the number of employees are positively correlated, the stores with the largest volume of sales would be most likely to enter the sample.

If the ppes sampling procedure led to the selection of stores $1,\ldots,n$, an unbiased estimate of the mean volume would be given by

$$\hat{V} = \sum_{i=1}^{n} h_i T_i/n ,$$

where $h_i = W/Mw_i$, and T_i is the volume of sales for the i^{th} store. An unbiased estimate $\hat{V}_T$ of the total volume of sales is obtained by multiplying $\hat{V}$ by the total number of employees, i.e. $\hat{V}_T = M\hat{V}$.

If the relative efficiency of the importance sampling estimate and the usual estimate obtained by simple random sampling is to be estimated, a one-sample or a two-sample approach may be used. The two-sample method requires that two independent samples be obtained, one by simple random sampling and the other by importance sampling. From these samples, independent estimates of the variance of the competing estimators are obtained. The problem is thus reduced to the usual variance comparison problem, although the distributions are somewhat restricted in this case. These distributions, however, are not necessarily well behaved. In fact, extreme non-normality is not uncommon and the usual assumption of identical distributions, except possibly for location and scale, is unrealistic.

The complexity of the two-sample problem dictates the use of robust estimation procedures, that is, procedures insensitive to departures from the usual assumptions. In an article by Miller [13], Monte Carlo sampling experiments are conducted to study several proposed robust variance comparison techniques. The procedures entering the study require that the sampled distributions differ at most by location and scale, but with slight modifications a number of these methods, which are described in Chapter 3, may be applied to the distributions arising from the importance sampling problem.

A one-sample approach, requiring that estimates of the variance of both estimators of the mean be obtained from a single sample, has been proposed by Sukhatme and Sukhatme [21, pp.50-52]. The ratio of these variance estimators then provides an estimate of the relative efficiency. The Sukhatme estimator, however, permits negative estimates of the ratio to occur. As an alternative to this estimator, we have proposed a non-negative, one-sample estimator which is described in Section 1.4. Note that the one-sample approach requires an estimate of the variance that would have been obtained if the alternative one-sample estimation procedure had been used.

The one-sample procedures have obvious practical advantages over the two-sample methods. Since a single sample is required, both time and expenses are reduced. Alternative sampling procedures may also be investigated, after the fact, thus providing additional information which could be used to aid in the design of future surveys. One should note, of course, that the advantage of the one-sample estimators would be negated if they proved to provide unreliable estimates.

In the remaining sections of this chapter, the importance sampling problem is discussed in more detail, and the Sukhatme estimator and its non-negative alternative are introduced.

1.2 The Importance Sampling Problem

Consider the problem of estimating the integral

$$I = \int f(x)p(x)dx = E_p(f) \quad , \qquad (1.1)$$

where $p(x)$ is a density function (with respect to Lebesgue or counting measure). The usual unbiased estimator of I is $\hat{I}_p = \sum_{i=1}^{n} f(x_i)/n$, where $(x_1,\ldots,x_n)$ is a sample from $p(x)$. The variance of $\hat{I}_p$ is given by

$$\mathrm{Var}_p(\hat{I}_p) = \mathrm{Var}_p(f)/n = [\int f^2(x)p(x)dx - I^2]/n .$$

To estimate I using importance sampling, an estimator $\hat{I}_q$ based on a sample $(y_1,\ldots,y_n)$ from a distribution with density $q(x)$ is used. $\hat{I}_q$ is defined by

$$\hat{I}_q = \sum_{i=1}^{n} f(y_i)p(y_i)/nq(y_i) \tag{1.2}$$

and

$$\mathrm{Var}_q(\hat{I}_q) = \mathrm{Var}_q(fp/q)/n = \{\int [f^2(x)p^2(x)/q(x)]dx - I^2\}/n .$$

If $q(x)$ is chosen to make $f(x)p(x)/q(x)$ approximately constant where most of the probability of $q(x)$ lies, we would expect $\mathrm{Var}_q(\hat{I}_q)$ to be small. In fact, the Schwartz inequality yields

$$\int f^2p^2/q = (\int f^2p^2/q)(\int q) \geq (\int |f|p)^2 ,$$

which is minimized for $q(x) = |f(x)|p(x)/\int|f(x)|p(x)dx$. Thus, the optimum choice of $q(x)$ is

$q_o(x) = |f(x)|p(x)/\int|f(x)|p(x)dx$ and if $f(x) \geq 0$,

$q_o(x) = f(x)p(x)/I$, for which $Var_q(\hat{I}_q) = 0$.

Sampling with probability proportional to estimated size is a special case of importance sampling, where $p(x)$ is the discrete uniform distribution and $q(x)$ is chosen proportional to some measure of the 'size' of $f(x)$. For example, let $(u_1,\ldots,u_N)$ be the sampling units of a finite population and let y_i be the value of a characteristic of interest of the i^{th} sampling unit. Then, with $f(u_i) = y_i$ and $p(u_i) = 1/N$, $q(u_i)$ is chosen proportional to a measure of the size of $f(u_i)$. This measure is often provided by an auxiliary variable, w_i say, that is highly correlated with y_i . The optimum situation, $q_o(u_i) \propto |f(u_i)|p(u_i)$, occurs when w_i and $|y_i|$ are proportional, in which case $Var_q(\hat{I}_q) = 0$, if $y_i \geq 0$.

A numerical example of ppes sampling, applied to an agricultural problem in which a total is estimated, is given in Sukhatme and Sukhatme [21, pp.50-52].

Although a substantial variance reduction may be achieved by importance sampling (or ppes sampling), Hastings [8] has shown that the resulting estimate may suffer from extreme non-normality, even when $q(x)$ is arbitrarily close to the optimum $q_o(x)$. This induced non-normality of the estimate $\hat{I}_q$ not only affects the assessment of error and inferences pertaining to $\hat{I}_q$ but also has an adverse effect on procedures used to estimate the ratio $Var_p(f)/Var_q(fp/q)$.

If the estimates $\hat{I}_p$ and $\hat{I}_q$ are based on the same number of observations, the ratio

$$Var_p(\hat{I}_p)/Var_q(\hat{I}_q) = Var_p(f)/Var_q(fp/q)$$

may be considered as a measure of the relative efficiency of the estimation procedures. This thesis is concerned with the problem of estimating this variance ratio. In sample survey problems, such an estimate would prove to be a valuable aid in the design of future surveys.

Since, in practice, it may not be economically feasible to obtain samples from both $p(x)$ and $q(x)$, an estimate of $Var_p(f)/Var_q(fp/q)$, based on a single sample from either $p(x)$ or $q(x)$, is desirable. With such an estimator, data from previous sample surveys could be utilized to study the feasibility of various density functions $q(x)$ that may be available.

In some situations, however, it may be possible to obtain independent samples from both distributions, in which case, two-sample estimates can be used. This will often be the case, for example, when sampling is done on a computer.

1.3 The Two-sample Variance Ratio Problem

In the introduction we suggested that the two-sample situation reduces to the usual variance comparison problem. To see that this is in fact the case, define $f^*(x) = f(x)p(x)/q(x)$ and let $(x_1,\ldots,x_n)$ and $(y_1,\ldots,y_m)$ be independent samples from $p(x)$ and $q(x)$, respectively. Then $(f(x_1),\ldots,f(x_n))$ and $(f^*(y_1),\ldots,f^*(y_m))$ constitute independent samples from the distributions of the random variables $f(X)$ and $f^*(Y)$, respectively, where X has density $p(x)$ and Y has density $q(x)$. Note that the variance ratio $\mathrm{Var}_p(f)/\mathrm{Var}_q(fp/q)$ is just the ratio of the variances of $f(X)$ and $f^*(Y)$, i.e.

$$\mathrm{Var}_p(f)/\mathrm{Var}_q(fp/q) = \mathrm{Var}[f(X)]/\mathrm{Var}[f^*(Y)] .$$

Thus, by considering $f(x_i)$, $i = 1,\ldots,n$, and $f^*(y_j)$, $j = 1,\ldots,m$, as random observations of the random variables $f(X)$ and $f^*(Y)$, respectively, we find that the problem may be formulated as the usual variance comparison problem.

It is well known that the classical F statistic is of little value when the distributions of $f(X)$ and $f^*(Y)$ stray far from the family of normal distributions (see Box [4]). Considering the findings of Hastings [8], it would be unwise to base tests or confidence intervals on the F distribution.

Several alternative variance comparison techniques that are less sensitive to non-normality have been proposed. One such method that has received a considerable amount of attention in recent years is the 'jackknife method', a procedure based on an estimator introduced by Quenouille [18] for its bias reduction properties and later extended by Tukey [22] to provide approximate confidence intervals and tests. In a Monte Carlo study by Miller [13], the jackknife was found to be a leading competitor among the robust techniques for variance comparisons. Miller's study compared the jackknife's performance with procedures suggested by Box [4], Box-Andersen [5], Moses [15] and Levene [11], as well as the classical procedure based on the F distribution.

The distributions of $f(X)$ and $f^*(Y) = f(Y)p(Y)/q(Y)$, where X has density $p(x)$ and

Y has density $q(x)$, may be radically different in nature, further complicating the estimation procedures. Miller [13] assumed that the sampled distributions were practically identical, differing at most by location and scale parameters. The effect of relaxing the restrictions imposed by Miller, upon the estimation techniques considered in this paper, is discussed in subsequent chapters.

In this thesis a Monte Carlo study is conducted to investigate the behavior of the jackknife method and slight variations of the methods suggested by Box [4] and Box-Andersen [5], when applied to the importance sampling problem. These procedures provide approximate confidence intervals for the ratio $v_p/v_q = \mathrm{Var}_p(f)/\mathrm{Var}_q(f^*)$ that are easily calculated on a computer, whereas the methods suggested by Moses [15] and Levene [11] do not. For comparison, the method based on the F distribution is also used.

A description and discussion of those methods implemented can be found in Chapter 3. The Moses [15] and Levene [11] methods, as well as several others based on ranks, are described in Miller [13].

1.4 The One-sample Variance Ratio Problem

The one-sample approach to the estimation problem requires that an estimate of the variance ratio

v_p/v_q be obtained from a single sample drawn from $q(x)$. That is, if $(x_1,\ldots,x_n)$ is a sample from $q(x)$, the estimator of v_p/v_q is to be of the form $h(x_1,\ldots,x_n)$, where the statistic $h(x_1,\ldots,x_n)$ is a function (measurable) of these sample values.

Sukhatme and Sukhatme [21, pp.50-52] have suggested a procedure for estimating, from a single sample taken from $q(x)$, the 'relative gain' due to ppes sampling over simple random sampling without replacement, when a total is estimated. In this case, $p(x)$ is the discrete uniform distribution and the quantity of interest is $(V_p - V_q)/V_q$, where V_p is the variance of the estimated total that would have been obtained by sampling without replacement from $p(x)$, and V_q is the variance of the estimated total when the sample is taken from $q(x)$. Obviously, since totals are merely multiples of means, the problem is equivalent to estimating $[\mathrm{Var}_p(\hat{I}_p) - \mathrm{Var}_q(\hat{I}_q)]/\mathrm{Var}_q(\hat{I}_q)$, when the sampling from $p(x)$ is without replacement.

In this thesis we assume that the sampling from $p(x)$ and $q(x)$ is with replacement. By introducing the finite population correction in the obvious manner, however, the procedures developed here can be modified to handle the problem of sampling without replacement when $p(x)$ is the discrete uniform distribution. (See Appendix C for an example involving sampling without replacement.)

In the following, we define two estimators of the variance ratio v_p/v_q, both of which are candidates for the jackknife methodology. The first of these estimators, $\hat{\theta}_s$, is a generalized version of the estimator introduced by Sukhatme and Sukhatme [21, pp.50-52]. The second estimator, $\hat{\theta}_w$, is a non-negative estimator that we have proposed as an alternative to the Sukhatme estimator $\hat{\theta}_s$.

(i) The Sukhatme estimator $\hat{\theta}_s$

The procedure followed by Sukhatme and Sukhatme [21, pp.50-52] can be extended to handle the more general importance sampling problem of estimating the ratio v_p/v_q or the 'relative gain', $(v_p - v_q)/v_q$, when $p(x)$ is an arbitrary density function. Following their method, we first obtain unbiased estimates of $v_p = \mathrm{Var}_p(f)$ and $v_q = \mathrm{Var}_q(f^*)$. To this end, let $(x_1,\ldots,x_n)$ be a sample from $q(x)$ and define $f^*(x) = f(x)p(x)/q(x)$, as before. Then $\hat{I}_q$ is an unbiased estimate of I, and s_q^2, defined by

$$s_q^2 = \sum_{i=1}^{n} [f^*(x_i) - \hat{I}_q]^2/(n-1) , \qquad (1.3)$$

is an unbiased estimate of v_q.

Define

$$\hat{v}_{ps} = S_q^2 + \sum_{i=1}^{n} [f^*(x_i)f(x_i) - f^{*2}(x_i)]/n . \quad (1.4)$$

Then $\hat{v}_{ps}$ is an unbiased estimate of v_p . To see this, note that

$$\begin{aligned} E_q(\hat{v}_{ps}) &= E_q(S_q^2) + E_q[f^*(x)f(x) - f^{*2}(x)] \\ &= \int f^{*2}(x)q(x)dx - I^2 + \int f^2(x)p(x)dx - \int f^{*2}(x)q(x)dx \\ &= \int f^2(x)p(x)dx - I^2 \\ &= v_p . \end{aligned}$$

The estimator $\hat{\theta}_s$ of the variance ratio v_p/v_q is then defined by

$$\hat{\theta}_s = \hat{v}_{ps}/S_q^2 , \quad (1.5)$$

and the 'relative gain' may be estimated by $(\hat{v}_{ps} - S_q^2)/S_q^2 = \hat{\theta}_s - 1$.

<u>(ii) The non-negative estimator $\hat{\theta}_w$</u>

Although the estimate $\hat{v}_{ps}$ is unbiased, it has the undesirable property of allowing negative values. That is, this estimator can lead to negative estimates of $\text{Var}_p(f)$. As an alternative to this estimator, we propose the use of an estimate s_w^2 which, although biased, does not permit negative values. Recall that we are concerned with the estimation of the ratio v_p/v_q and that $\hat{\theta}_s$ will not, in general, be unbiased, despite the fact that $\hat{v}_{ps}$ and s_q^2 are unbiased.

To obtain a non-negative estimate of v_p, we first note that

$$\begin{aligned} v_p &= \int [f(x) - I]^2 p(x)dx \\ &= \int [f(x) - I]^2 [p(x)/q(x)] q(x) dx \\ &= E_q\{p(x)[f(x) - I]^2/q(x)\} . \end{aligned}$$

Therefore, $\widehat{\mathrm{var}}_p(f)$, defined by

$$\widehat{\mathrm{var}}_p(f) = \sum_{i=1}^{n} p(x_i)[f(x_i) - I]^2/nq(x_i) , \tag{1.6}$$

where $(x_1,\ldots,x_n)$ is a sample from $q(x)$, is an unbiased estimate of v_p. Obviously the parameter $I = E_p(f)$ will not be known, but if we replace I by its estimate $\hat{I}_q$ in (1.6), we obtain an estimator s_w^2 defined by

$$s_w^2 = \sum_{i=1}^{n} p(x_i)[f(x_i) - \hat{I}_q]^2/nq(x_i) . \tag{1.7}$$

Note that s_w^2 has the form of a weighted sum of squares and is non-negative.

To obtain an expression for the bias of S_w^2, we proceed as follows. Let $p_i = p(x_i)$, $q_i = q(x_i)$, $f_i = f(x_i)$ and $f_i^* = f^*(x_i)$. Then expanding S_w^2, we obtain

$$S_w^2 = [\sum_{i=1}^{n} f_i^2 p_i/q_i - 2(\sum_{i=1}^{n} f_i^*)^2/n + (\sum_{i=1}^{n} f_i^*/n)^2(\sum_{i=1}^{n} p_i/q_i)]/n .$$

But

$$(\sum_{i=1}^{n} f_i^*)^2(\sum_{i=1}^{n} p_i/q_i) = \sum_{i=1}^{n} f_i^{*2} p_i/q_i + \sum_{i\neq j}^{n} f_j^{*2} p_i/q_i + 2\sum_{i\neq j}^{n} f_j^* f_i^* p_i/q_i + \sum_{i\neq j\neq k}^{n} f_i^* f_j^* p_k/q_k ,$$

therefore,

$$S_w^2 = (\sum_{i=1}^{n} f_i^2 p_i/q_i)/n - 2(\sum_{i=1}^{n} f_i^{*2} + \sum_{i\neq j}^{n} f_i^* f_j^*)/n^2 + (\sum_{i=1}^{n} f_i^{*2} p_i/q_i + \sum_{i\neq j}^{n} f_j^{*2} p_i/q_i + 2\sum_{i\neq j}^{n} f_j^* f_i^* p_i/q_i + \sum_{i\neq j\neq k}^{n} f_i^* f_j^* p_k/q_k)/n^3 .$$

Taking expectations with respect to $q(x)$ and simplifying yields

$$E_q(S_w^2) = \int f^2 p - I^2 + (2I\int f^* p - I^2 - \int f f^* p)/n$$
$$+ (2I^2 + \int f^{*2} p - 2I\int f^* p - \int f f^* p)/n^2 ,$$

or

$$E_q(S_w^2) = v_p + [2IE_p(f^*) - I^2 - E_p(ff^*)]/n$$
$$+ [2I^2 + E_p(f^{*2}) - 2IE_p(f^*) - E_p(ff^*)]/n^2 .$$

Thus, the bias of S_w^2 is given by

$$\mathrm{Bias}(S_w^2) = [2IE_p(f^*) - I^2 - E_p(ff^*)]/n$$
$$+ [2I^2 + E_p(f^{*2}) - 2IE_p(f^*) - E_p(ff^*)]/n^2 .$$

Note that in the optimum situation, if $f(x) \geq 0$, we have $q_o(x) = f(x)p(x)/I$ and $f^*(x) = I$, and the bias terms vanish. This would not be the case, however, if $n - 1$ had been used in the place of n in the definition of S_w^2 . In fact, the bias would be equal to $v_p/(n-1)$ in the optimum situation if this alternative form of S_w^2 were used.

The non-negative estimator of v_p/v_q is defined by

$$\hat{\theta}_w = S_w^2/S_q^2 , \tag{1.8}$$

and the 'relative gain' may be estimated by $\hat{\theta}_w - 1$.

The jackknife procedure may be applied to the estimators $\hat{\theta}_s$ and $\hat{\theta}_w$ to obtain approximate confidence intervals and tests for the ratio v_p/v_q .

To evaluate the behavior of the estimators $\hat{\theta}_s$ and $\hat{\theta}_w$, when used in conjunction with the jackknife method, a Monte Carlo study is conducted. An estimate of the variance of $\hat{\theta}_s$, based on a Taylor series expansion, is also used to obtain approximate confidence intervals for v_p/v_q , and this procedure enters the Monte Carlo study as well.

Details of the Monte Carlo study and the estimation techniques are given in subsequent chapters. A general discussion of the jackknife method follows in Chapter 2, and details of its application to the two-sample and one-sample problems may be found in Chapters 3 and 4, respectively.

CHAPTER 2

THE JACKKNIFE

2.1 Introduction

The jackknife estimator was introduced by Quenouille [18] for the purpose of bias reduction. A general method for obtaining approximate confidence intervals was proposed later by Tukey [22], who called his procedure the 'jackknife'.

Subsequent papers have established the jackknife as a robust procedure that is capable of reducing bias and providing approximate confidence intervals and tests.

In this chapter, Quenouille's estimator is defined and the validity of Tukey's conjecture is discussed. Special emphasis is placed on the results of a paper by Arvesen [1], in which the asymptotic normality of Tukey's jackknife is established for a large class of statistics, the U-statistics.

2.2 Quenouille's Estimator

In the following discussion, we adopt the notation of Gray and Schucany [7].

Let $(x_1,\ldots,x_N)$ be a random sample from a c.d.f. F_θ. Divide the sample into n groups of k observations each $(N = nk)$, and let $\hat{\theta} = \hat{\theta}_N(x_1,\ldots,x_N)$

be an estimator of θ based on the sample. Define $\hat{\theta}^i$ to be the same estimator based on the subsample obtained by deleting the i^{th} group of observations. Now let

$$J_i(\hat{\theta}) = n\hat{\theta} - (n-1)\hat{\theta}^i \quad , \quad i = 1,2,\ldots,n \ , \tag{2.1}$$

and

$$J(\hat{\theta}) = \sum_{i=1}^{n} J_i(\hat{\theta})/n \ . \tag{2.2}$$

$J(\hat{\theta})$ is called the jackknife estimator and the estimators $J_i(\hat{\theta})$, $i = 1,\ldots,n$, are called the pseudo-values of the jackknife.

Quenouille introduced these estimators in [18]. He noted that for a large class of statistics of the form $\hat{\theta} = \hat{\theta}_N(x_1,\ldots,x_N)$, the expected value of $\hat{\theta}$ had the form

$$E(\hat{\theta}) = \theta + a_1/N + a_2/N^2 + a_3/N^3 + \ldots \ . \tag{2.3}$$

This class, for example, includes all estimators of the form $\hat{\theta} = f(\bar{x})$ $(\theta = f(\mu_x))$, where $f(\bar{x})$ admits a Taylor series expansion about μ_x , the mean of x .

When (2.3) holds, the pseudo-values completely eliminate a bias term of order $1/N$, as is established below.

$$E(J_i(\hat{\theta})) = nE(\hat{\theta}) - (n - 1)E(\hat{\theta}^i)$$

$$= n(\theta + a_1/N + a_2/N^2 + a_3/N^3 + \ldots)$$

$$-(n - 1)[\theta + a_1/(N - k) + a_2/(N - k)^2$$

$$+ a_3/(N - k)^3 + \ldots]$$

$$= \theta - a_2/N(N - k) - a_3(2N - k)/N^2(N - k)^2 + \ldots .$$

To achieve a minimum loss of efficiency, while retaining the bias reduction properties of the pseudo-values, Quenouille [18] suggested the use of the estimator $J(\hat{\theta})$. He showed that in many instances the standard error is increased by a factor of $o(1/N)$, i.e. $S.E.J(\hat{\theta}) = (S.E.\hat{\theta})[1 + o(1/N)]$, and since in general the S.E. of $\hat{\theta}$ will decrease as $N^{-\frac{1}{2}}$, the reduction in bias is not accompanied by a comparable increase in variability.

The following example is taken from Quenouille [18]. Let $(x_1,\ldots,x_N)$ be a random sample from a normal distribution with mean μ and variance σ^2 , and suppose $\hat{\theta} = \sum_{i=1}^{N} (x_i - \bar{x})^2/N$ is the estimator of σ^2 to be jack-knifed. The bias of the estimator $\hat{\theta}$ is $-\sigma^2/N$. If $k = 1$, it is not difficult to show that the jackknife estimator $J(\hat{\theta})$ is the unique minimum variance unbiased estimate of σ^2 , i.e.

$$J(\hat{\theta}) = \sum_{i=1}^{N} (x_i - \bar{x})^2/(N - 1) .$$

Quenouille [18] also suggested a higher order jackknife estimator to eliminate a bias term of order $1/N^2$ as well. Details of this second order estimator can be found in Miller [14] and Quenouille [18].

2.3 Tukey's Conjecture

Tukey [22] proposed that in most instances the pseudo-values could be treated as n approximately independent identically distributed random variables. Define a statistic T by

$$T = n^{\frac{1}{2}}[J(\hat{\theta}) - \theta]/\{(n - 1)^{-1} \sum_{i=1}^{n} [J_i(\hat{\theta}) - J(\hat{\theta})]^2\}^{\frac{1}{2}} .$$

Tukey then suggested that T has an approximate Student-t distribution with $n - 1$ degrees of freedom.

Miller [12] establishes the 'trustworthiness' of Tukey's method in two situations where the estimators have a linear quality to them: one where the estimator is a linear function of the observations, such as the sample mean, and the other where $\hat{\theta} = f(\bar{x})$ $(\theta = f(\mu))$ and $f(x)$ is a real-valued function with a bounded second derivative in a neighborhood of μ .

In the same paper, Miller illustrates, by counter example, that universal application of the jackknife can be hazardous. He shows that if the estimator $\hat{\theta} = \max(x_1,\ldots,x_N)$, for estimating a truncation point, is jackknifed, the statistic T does not necessarily have an approximate Student-t distribution.

Correlation between the pseudo-values can also be a major source of error in the approximate confidence intervals derived by Tukey's method. The failure of the jackknife in an example by Miller [12] on the preservation of normality is a consequence of this correlation. Miller [12] showed that T does not necessarily approximate a t statistic, even though the pseudo-values have a multivariate normal distribution with intraclass correlation ρ and common variance σ^2 . Gray and Schucany [7, p.173] consider an essentially equivalent form of Miller's example, in which they assume that $(J_1(\hat{\theta}),\ldots,J_n(\hat{\theta}))$ has a multivariate normal distribution with correlation matrix

$$\Sigma = (\rho_{ij})_{n\times n} ,$$

where $\rho_{ij} = 1$ if $i = j$ and $\rho_{ij} = \rho$ if $i \neq j$. The authors show that T is distributed as ct , where t is a Student-t random variable and c is a constant defined by $c = \{[1 + (n - 1)\rho]/(1 - \rho)\}^{\frac{1}{2}}$. Thus, the confidence levels of the approximate confidence intervals may be suspect when the pseudo-values are highly correlated.

Gray and Schucany [7, p.165] suggest that in many situations the correlation among the pseudo-values may be modeled more adequately by $\rho = 1/n$, rather than $\rho = 0$. In this case, $c = [(2n - 1)/(n - 1)]^{\frac{1}{2}}$, and the approximate confidence intervals should be increased in length by a factor of c.

2.4 Jackknifing U-statistics

Arvesen [1] proves that the jackknife technique can be successfully applied to U-statistics and functions of several U-statistics. Under suitable regularity conditions, the results established by Miller [12] can be extended to these cases to provide asymptotic convergence theorems for the Studentized jackknife estimator. Arvesen also demonstrates how the jackknife method can be extended to the two-sample problem and establishes the asymptotic convergence of the estimator in this case as well.

All the estimators that we shall consider as possible candidates for the jackknife technique are U-statistics or functions of them. For this reason, the relevant definitions and theorems of Arvesen [1] will be stated in this section, with slight notational changes to be consistent with the notation introduced earlier.

Definition 2.1.

Let $X_1,\ldots,X_N$ be N independent identically distributed (IID) random variables, and let $k(X_1,\ldots,X_m)$ be an unbiased estimate of some parameter η, where m is the smallest number of observations required to estimate η. Then there exists a symmetric kernel of $k(X_1,\ldots,X_m)$, given by

$$k^*(X_1,\ldots,X_m) = (m!)^{-1} \sum_{P_m} k(X_{a_1},\ldots,X_{a_m}) , \qquad (2.4)$$

where P_m indicates that the sum is over the $m!$ permutations of the subscripts.

The U-statistic for the parameter η is defined by

$$U(X_1,\ldots,X_N) = \binom{N}{m}^{-1} \sum_{C_N} k^*(X_{a_1},\ldots,X_{a_m}) , \qquad (2.5)$$

where C_N indicates that the summation is over all combinations $a_1,\ldots,a_m$ of m integers chosen from $1,\ldots,N$.

The sample variance S^2 is an example of a U-statistic. To see this, note that at least two observations are required to estimate σ^2 and that the function $k(X_1,X_2)$, defined by

$$k(X_1,X_2) = X_1^2 - X_1X_2 ,$$

is an unbiased estimate of σ^2. The symmetric kernel is then

$$k^*(X_1,X_2) = \tfrac{1}{2}(X_1^2 - X_1X_2 + X_2^2 - X_2X_1)$$

$$= \tfrac{1}{2}(X_1 - X_2)^2 .$$

If we now sum $k^*(X_1,X_2)$ over all possible combinations (i,j) of two integers chosen from $1,\ldots,N$, we obtain

$$U(X_1,\ldots,X_N) = \binom{N}{2}^{-1} \sum_{i<j}^{N} \tfrac{1}{2}(X_i - X_j)^2$$

$$= [1/N(N - 1)] \sum_{i<j}^{N} (X_i - X_j)^2 .$$

But

$$\sum_{i<j}^{N} (X_i - X_j)^2 = \tfrac{1}{2} \sum_{j=1}^{N} \sum_{i=1}^{N} (X_i^2 + X_j^2 - 2X_iX_j)$$

$$= N(\sum_{i=1}^{N} X_i^2 - N\bar{X}^2) .$$

Therefore

$$U(X_1,\ldots,X_N) = (\sum_{i=1}^{n} X_i^2 - N\bar{X}^2)/(N - 1)$$

$$= S^2 .$$

Now, let

$$k_c^*(x_1,\ldots,x_c) = E[k^*(X_1,\ldots,X_c\ ,\ X_{c+1},\ldots,X_m) \qquad (2.6)$$

$$|X_1 = x_1,\ldots,X_c = x_c]\ ,$$

$$\zeta_c = \mathrm{Var}[k_c^*(X_1,\ldots,X_c)]\quad ,\quad c = 1,\ldots,m\ , \qquad (2.7)$$

and

$$\zeta_o = 0\ .$$

The following theorems appear in Arvesen [1] and are stated here without proof.

The first group of theorems establish the asymptotic normality of the jackknife statistic when applied to a real-valued function $g:R \to R$ of a U-statistic, thus justifying the use of the normal approximation to obtain approximate confidence intervals and tests. The U-statistics, defined as in (2.5), are functions of independent identically distributed (IID) random variables, in this case.

<u>THEOREM 2.1</u>. Let $X_1,\ldots,X_N$ be IID random variables. If $k^*(X_1,\ldots,X_m)$ is a real-valued symmetric statistic with expectation η and finite second moment $E\{[k^*(X_1,\ldots,X_m)]^2\}$, then as $N\to\infty$, the limiting distribution of $N^{\frac{1}{2}}(U - \eta)$ is normal with mean zero and variance $m^2\zeta_1$.

THEOREM 2.2. Let $X_1,\ldots,X_N$ be IID random variables. If $k^*(X_1,\ldots,X_m)$ is a real-valued symmetric statistic with expectation η and $E|k^*(X_1,\ldots,X_m)| < \infty$, then $U(X_1,\ldots,X_N) \underset{a.s.\ L_1}{\to} \eta$ as $N\to\infty$.

Before proceeding to the next theorem, we adopt the following notation. Let

$$U = U(X_1,\ldots,X_N) \quad , \quad \text{as in (2.5)} \ ,$$

and let

$$U_i = \binom{N-k}{m}^{-1} \sum_{C_{n-1}^i} k^*(X_{b_1},\ldots,X_{b_m}) \quad , \quad i = 1,\ldots,n \ , \tag{2.8}$$

where C_{n-1}^i indicates that the summation is over all combinations $(b_1,\ldots,b_m)$ of m integers chosen from $(1,\ldots,(i-1)k,ik+1,\ldots,N)$, $N = nk$, and $k^*(X_1,\ldots,X_m)$ is as in (2.4).

Let g be a real-valued function, and let

$$\hat\theta = g(U) \quad , \quad \hat\theta^i = g(U_i) \quad , \quad \theta = g(\eta) \ ,$$

$$J_i(\hat\theta) = n\hat\theta - (n-1)\hat\theta^i \ ,$$

$$J(\hat\theta) = \sum_{i=1}^{n} J_i(\hat\theta)/n \ . \tag{2.9}$$

and

$$S_g^2 = (n-1)^{-1} \sum_{i=1}^{n} [J_i(\hat\theta) - J(\hat\theta)]^2 \ .$$

THEOREM 2.3. Let $X_1,...,X_N$ $(N = nk)$ be IID random variables, and let $k^*(X_1,...,X_m)$ be a real-valued symmetric statistic with expectation η, and finite second moment $E\{[k^*(X_1,...,X_m)]^2\}$. Let g be a function defined on the real line, which in a neighborhood of η has a bounded second derivative. Then as $n\to\infty$,

$$n^{\frac{1}{2}}[J(\hat{\theta}) - \theta] \underset{D}{\to} N(0,m^2\zeta_1[g'(\eta)]^2) .$$

THEOREM 2.4. Let $X_1,...,X_N$ $(N = nk)$ be IID random variables, and let $k^*(X_1,...,X_m)$ be a real-valued symmetric statistic with expectation η and finite second moment $E\{[k^*(X_1,...,X_m)]^2\}$. Let g be a function defined on the real line, which in a neighborhood of η has a continuous first derivative. Then, as $n\to\infty$,

$$S_g^2 \underset{P}{\to} m^2\zeta_1[g'(\eta)]^2 .$$

Theorem 7 of Arvesen [1], although a valid result, appears with a faulty proof. This theorem is stated below and can be found in Appendix B with an alternative proof.

THEOREM 2.5. Let $X_1,...,X_N$ $(N = nk)$ be IID random variables, and let $k^*(X_1,...,X_m)$ be a real-valued symmetric statistic with expectation η and finite second moment $E\{[k^*(X_1,...,X_m)]^2\}$. Let g be a function defined on the real line, which in a neighborhood of η has a continuous first derivative. Then, as $k\to\infty$ (n fixed),

$$n^{\frac{1}{2}}[J(\hat{\theta}) - \theta]/S_g \underset{D}{\to} t(n-1) ,$$

where $t(n-1)$ denotes the Student-t distribution with $n-1$ degrees of freedom.

The next group of theorems generalize Theorems 2.3 and 2.4 to establish the asymptotic normality of the jackknife statistic when applied to functions of several U-statistics, where the U-statistics are, in this case, symmetric functions of random vectors.

The advantage of such theorems is obvious. Suppose, for example, that $Z = (X,Y)$ is a random vector and that we wish to estimate $h(\mu_x,\mu_y,\sigma_x^2,\sigma_y^2)$, where $h:R^4 \to R$ is a real-valued function. Under suitable regularity conditions, these theorems indicate that the jackknife can be successfully applied to $h(\bar{X},\bar{Y},S_X^2,S_Y^2)$ to estimate $h(\mu_x,\mu_y,\sigma_x^2,\sigma_y^2)$, where $\bar{X}$ and $\bar{Y}$ are the sample means of X and Y, respectively, and S_X^2 and S_Y^2 are the sample variances of X and Y, respectively.

These generalizations of Theorems 2.3 and 2.4 require the introduction of the following notation. Let $X_1,\ldots,X_N$ be IID random vectors of p components. Let

g be a real-valued function of q arguments and let

$$\theta = g(\eta_1, \ldots, \eta_q) \ .$$

Let $U^1, \ldots, U^q$ be such that

$$U^j = \binom{N}{m_j}^{-1} \sum_{C_N} k^{*j}(X_{a_1}, \ldots, X_{a_{m_j}}) \quad , \quad j = 1, \ldots, q \ ,$$

where C_N is as in (2.5) and k^{*j} is a real-valued symmetric kernel based on m_j observations and is an unbiased estimate of η_j . Let

$$\hat{\theta} = g(U^1, \ldots, U^q)$$

and

$$\hat{\theta}^i = g(U_i^1, \ldots, U_i^q) \quad , \quad i = 1, \ldots, n \ ,$$

where

$$U_i^j = \binom{N-k}{m_j}^{-1} \sum_{C_{n-1}^i} k^{*j}(X_{b_1}, \ldots, X_{b_{m_j}}) \ , \ i=1, \ldots, n \ ; \ j=1, \ldots, q \ ,$$

and C_{n-1}^i and $(b_1, \ldots, b_{m_j})$ are defined as in (2.8).

Define $J_i(\hat{\theta})$, $i = 1, \ldots, n$, $J(\hat{\theta})$ and S_g^2 as in (2.9). Let

$$g_k = \partial g(t_1, \ldots, t_q)/\partial t_k \Big|_{(\eta_1, \ldots, \eta_q)}$$

and

$$g_{kr} = \partial g(t_1,\ldots,t_q)/\partial t_k \partial t_r \Big|_{(\eta_1,\ldots,\eta_q)} . \tag{2.10}$$

THEOREM 2.6. Let $X_1,\ldots,X_N$ $(N = nk)$ be IID random vectors of p components. Let $k^{*j}(X_1,\ldots,X_{m_j})$ be a real-valued symmetric statistic with expectation η_j, and finite second moment $E\{[k^{*j}(X_1,\ldots,X_{m_j})]^2\}$, $j = 1,\ldots,q$. Let g be a real-valued function defined on R^q, which in a neighborhood of $(\eta_1,\ldots,\eta_q)$ has bounded second partial derivatives. Then, as $n\to\infty$, $n^{\frac{1}{2}}[J(\hat{\theta}) - \theta]$ is asymptotically normally distributed with mean zero and variance

$$\sigma^2 = \sum_{i=1}^{q} \sum_{j=1}^{q} m_i m_j g_i g_j \zeta_1^{i,j} ,$$

where

$$\zeta_c^{i,j} = \mathrm{Cov}[k_c^{*i}(X_1,\ldots,X_c) , k_c^{*j}(X_1,\ldots,X_c)] .$$

THEOREM 2.7. Let $X_1,\ldots,X_N$ $(N = nk)$ be IID random vectors of p components. Let $k^{*j}(X_1,\ldots,X_{m_j})$ be a real-valued symmetric statistic with expectation η_j, and finite second moment $E\{[k^{*j}(X_1,\ldots,X_{m_j})]^2\}$, $j = 1,\ldots,q$. Let g be a real-valued function defined on R^q, which in a neighborhood of $(\eta_1,\ldots,\eta_q)$ has continuous first partial derivatives. Then, as $n\to\infty$,

$$S_g^2 \underset{p}{\to} \left(\sum_{i=1}^{q}\sum_{j=1}^{q} m_i m_j g_i g_j \zeta_1^{i,j}\right) .$$

Consider the problem of estimating a parameter which requires sampling from two distinct distributions. For example, suppose that we wish to estimate $\log(\sigma_x^2/\sigma_y^2)$, where σ_x^2 and σ_y^2 are the variances of random variables X and Y , respectively. If independent samples are obtained from the distributions of each of these random variables, the jackknife procedure could then be applied separately to $\log(S_x^2)$ and to $\log(S_y^2)$ to provide estimates $J(\log(S_x^2))$ of $\log(\sigma_x^2)$ and $J(\log(S_y^2))$ of $\log(\sigma_y^2)$. The results of Theorems 2.3 and 2.4, together with the independence of the two jackknife statistics, could then be utilized to obtain an estimate $J(\log(S_x^2)) - J(\log(S_y^2))$ of $\log(\sigma_x^2/\sigma_y^2)$ and to provide approximate confidence intervals and tests.

This method of jackknifing separately and combining the results has obviously relied heavily on the properties of the log transformation. Another method, requiring equal sample sizes, but which is applicable to a larger class of functions, could also have been used. This procedure requires randomly pairing the observations from the distinct distributions and treating the pairs of observed values as random observations of the random vector (X,Y). Then, with $g: R^2 \to R$ defined by $g(u,v) = \log(u/v)$, Theorem 2.6 and 2.7 justify the use of the statistic $J(\log(S_x^2/S_y^2))$ to estimate $\log(\sigma_x^2/\sigma_y^2)$, where $J(\log(S_x^2/S_y^2))$ is defined as for Theorem 2.6, with $U^1 = S_x^2$ and $U^2 = S_y^2$.

The final group of theorems establish the validity of an alternative way to jackknife the two-sample estimate, which does not require equal sample sizes. This procedure involves computing an estimate of the unknown parameter from both the samples, and then jackknifing by successively deleting each observation (or group of observations if $k > 1$) in the first sample with the second sample intact and then deleting the observations in the second sample with the first intact.

The generalization of the jackknife to the two-sample case proceeds as follows. Let $X_1,\ldots,X_{N_1}$ be IID random vectors of p components, let $Y_1,\ldots,Y_{N_2}$ be IID random vectors of p components and let the X's and the Y's be independent. Let

$$k^*_{c_1 c_2}(x_1,\ldots,x_{c_1},y_1,\ldots,y_{c_2}) = E[k^*(X_1,\ldots,X_{m_1};Y_1,\ldots,Y_{m_2}) \\ |X_1=x_1,\ldots,X_{c_1}=x_{c_1};Y_1=y_1,\ldots,Y_{c_2}=y_{c_2}]$$

and

$$\zeta_{c_1 c_2} = \mathrm{Var}[k^*_{c_1 c_2}(X_1,\ldots,X_{c_1};Y_1,\ldots,Y_{c_2})] \ ,$$

where $k^*(X_1,\ldots,X_{m_1};Y_1,\ldots,Y_{m_2})$ is a real-valued statistic, symmetric in the X's and symmetric in the Y's.

Define

$$U = \left[\binom{N_1}{m_1}\binom{N_2}{m_2}\right]^{-1} \sum_C k^*(X_{a_1},\ldots,X_{a_{m_1}};Y_{b_1},\ldots,Y_{b_{m_2}}) \ , \qquad (2.11)$$

where C indicates that the summation is over all combinations $(a_1,\ldots,a_{m_1})$ from $(1,\ldots,N_1)$ and all combinations $(b_1,\ldots,b_{m_2})$ from $(1,\ldots,N_2)$.

The two-sample jackknife estimator is defined as follows. Let $X_1,\ldots,X_{N_1}$ be N_1 observations from the first population, and let $Y_1,\ldots,Y_{N_2}$ be N_2 observations from the second population, and split the X's into n_1 groups of size k_1 and the Y's into n_2 groups of size k_2 $(N_1 = n_1 k_1 \ , \ N_2 = n_2 k_2)$. Next, let $\hat{\theta}$ be the estimate of θ based on all the observations, and let $\hat{\theta}_{i,\cdot}$ be the estimator obtained after deletion of the i^{th} group of X's , $i = 1,\ldots,n_1$, and let $\hat{\theta}_{\cdot,j}$ be the estimator obtained after deletion of the j^{th} group of Y's , $j = 1,\ldots,n_2$. Then define

$$J_{1,i} = n_1\hat{\theta} - (n_1 - 1)\hat{\theta}_{i,\cdot} \quad , \; i = 1,\ldots,n_1 \; ,$$

$$J_{2,j} = n_2\hat{\theta} - (n_2 - 1)\hat{\theta}_{\cdot,j} \quad , \; j = 1,\ldots,n_2 \; , \tag{2.12}$$

$$J_1 = \sum_{i=1}^{n_1} J_{1,i}/n_1 \quad , \quad J_2 = \sum_{j=1}^{n_2} J_{2,j}/n_2 \; ,$$

and

$$s_g^2 = n_1\{[n_1(n_1-1)]^{-1}\sum_{i=1}^{n_1}(J_{1,i}-J_1)^2+[n_2(n_2-1)]^{-1}\sum_{j=1}^{n_2}(J_{2,j}-J_2)^2\}.$$

The two-sample jackknife estimator is given by

$$J = (\sum_{i=1}^{n_1} J_{1,i} + \sum_{j=1}^{n_2} J_{2,j})/(n_1 + n_2) \; . \tag{2.13}$$

The following theorem, which extends Theorems 2.6 and 2.7 to the two-sample case, is stated for $k_1 = k_2 = 1$.

THEOREM 2.8. Let $X_1,\ldots,X_{n_1}$ be IID random vectors of p components, let $Y_1,\ldots,Y_{n_2}$ be IID random vectors of p components, and let the X's and the Y's be independent. Let $k^{*j}(X_1,\ldots,X_{m1_j} \; ; \; Y_1,\ldots,Y_{m2_j})$ be a real-valued statistic, symmetric in the X's and symmetric in the Y's , with expectation η_j and finite second moment for $j = 1,\ldots,q$. Let U^j be as in (2.11) for $j = 1,\ldots,q$, and let g be a real-valued function on R^q , which in a neighborhood of $(\eta_1,\ldots,\eta_q)$ has bounded second partial derivatives. Let the jackknife estimate J be defined as in (2.13) with $\hat{\theta} = g(U^1,\ldots,U^q)$. Then if $n_1 \le n_2$ and $n_1 \to \infty$ such that $\lim(n_1/n_2)$ exists, $n_1^{\frac{1}{2}}(J - \theta)$ is asymptotically normally distributed with mean zero and variance

$$\sigma^2 = \lim_{n_1 \to \infty} \{ \sum_{i=1}^{q} \sum_{j=1}^{q} g_i g_j [m_i^1 m_j^1 \zeta_{10}^{i,j} + (n_1/n_2) m_i^2 m_j^2 \zeta_{01}^{i,j}]\} ,$$

where

$$\zeta_{c_1 c_2}^{i,j} = \mathrm{Cov}[k^{*i}_{c_1 c_2} (X_1,\ldots,X_{c_1};Y_1,\ldots,Y_{c_2}),$$

$$k^{*j}_{c_1 c_2} (X_1,\ldots,X_{c_1};Y_1,\ldots,Y_{c_2})] ,$$

and the g_i , $i = 1,\ldots,q$ are defined as in (2.10). In addition, $S_g^2 \underset{p}{\to} \sigma^2$.

Before proceeding to the next section, we introduce a final theorem which will prove useful in subsequent chapters. This theorem establishes the preservation of the asymptotic normality of a random variable when transformed by a function with continuous first partial derivatives.

<u>THEOREM 2.9</u>. Let $\{\underset{\sim}{X}_1=(X_{11},\ldots,X_{1s})\ ,\ \underset{\sim}{X}_2=(X_{21},\ldots X_{2s})\ ,\ldots\}$ be a sequence of random vectors of s components such that $\underset{\sim}{X}_n \underset{P}{\to} \underset{\sim}{\mu}$, where $\underset{\sim}{\mu} = (\mu_1,\ldots,\mu_s)$, and let $g:R^s \to R$ be a real-valued function with continuous first partial derivatives in a neighborhood $A(\underset{\sim}{\mu})$ of $\underset{\sim}{\mu}$. Then if

$$\underset{\sim}{Y}_n = n^{\frac{1}{2}}(\underset{\sim}{X}_n - \underset{\sim}{\mu}) \underset{D}{\to} \underset{\sim}{Y} \sim N(\underset{\sim}{0},\Sigma_s) ,$$

$$n^{\frac{1}{2}}[g(\underset{\sim}{X}_n) - g(\underset{\sim}{\mu})] \underset{D}{\to} N(0,\sum_{i=1}^{s}\sum_{j=1}^{s} g_i(\underset{\sim}{\mu})g_j(\underset{\sim}{\mu})\sigma_{ij}) ,$$

where $\sum_s = (\sigma_{ij})_{s\times s}$ and $g_i = \partial_g(t_1,\ldots,t_s)/\partial t_i$, $i = 1,\ldots,s$.

<u>PROOF</u>. On $A(\underset{\sim}{\mu})$ we have

$$n^{\frac{1}{2}}[g(\underset{\sim}{X}_n) - g(\underset{\sim}{\mu})] = \sum_{i=1}^{s} g_i(\underset{\sim}{\zeta}_n)n^{\frac{1}{2}}(X_{ni} - \mu_i)$$

$$= \underset{\sim}{G}_n \cdot \underset{\sim}{Y}_n ,$$

where $\underset{\sim}{\zeta}_n$ indicates that the partial derivatives are evaluated on the line segment between $\underset{\sim}{X}_n$ and $\underset{\sim}{\mu}$, $\underset{\sim}{G}_n = (g_1(\underset{\sim}{\zeta}_n),\ldots,g_s(\underset{\sim}{\zeta}_n))$ and $(\cdot)$ denotes the inner product function.

Now, since $\underset{\sim}{X}_n \underset{P}{\to} \underset{\sim}{\mu}$ and the first partial derivatives of g are continuous on $A(\underset{\sim}{\mu})$, we have

$$g_i(\underset{\sim}{\zeta}_n) \underset{P}{\rightarrow} g_i(\underset{\sim}{\mu}) \quad , \; i = 1,\ldots,s \; .$$

Therefore, $\underset{\sim}{G}_n \underset{P}{\rightarrow} \underset{\sim}{G}$, where $\underset{\sim}{G} = (g_1(\underset{\sim}{\mu}),\ldots,g_s(\underset{\sim}{\mu}))$.

Furthermore, since $\underset{\sim}{Y}_n \underset{D}{\rightarrow} \underset{\sim}{Y} \sim N(\underset{\sim}{0},\Sigma_s)$, Theorem 4.4 of Billingsley [3, p.27] yields

$$(\underset{\sim}{Y}_n,\underset{\sim}{G}_n) \underset{D}{\rightarrow} (\underset{\sim}{Y},\underset{\sim}{G}) \; .$$

To complete the proof we need only note that the inner product function is continuous and apply Corollary 1 of Billingsley [3, p.31], which yields

$$\underset{\sim}{G}_n \cdot \underset{\sim}{Y}_n \underset{D}{\rightarrow} \underset{\sim}{G} \cdot \underset{\sim}{Y} \sim N(0, \sum_{i=1}^{s} \sum_{j=1}^{s} g_i(\underset{\sim}{\mu}) g_j(\underset{\sim}{\mu}) \sigma_{ij}) \; . \qquad \square$$

2.5 The Generalized Jackknife

In a paper by Schucany, Gray and Owen [19], the jackknife method is generalized to handle more general forms of bias than Quenouille's estimator. A comprehensive study of this procedure, together with a variety of applications and examples, can be found in a book by Gray and Schucany [7]. In this book, asymptotic results similar to those developed for the jackknife, which is a special case of the generalized jackknife, are established. Approximate confidence intervals and tests are obtained in a manner analogous to those of the jackknife.

A recent paper by Miller [14] summarizes some of the latest developments in the jackknife technique and indicates several areas where further research is required. In his bibliography, Miller attempts to list all published works on the jackknife methodology, providing an excellent source of references.

CHAPTER 3

THE TWO-SAMPLE VARIANCE RATIO ESTIMATORS

3.1 Introduction

Recall, from Section 1.3, that the two-sample variance ratio problem reduces to that of estimating

$$v_p/v_q = \mathrm{Var}[f(X)]/\mathrm{Var}[f^*(Y)] ,$$

where $f^*(x) = f(x)p(x)/q(x)$, X has density $p(x)$ and Y has density $q(x)$. We proceed by obtaining independent samples $(x_1,\ldots,x_{N_1})$ from $p(x)$ and $(y_1,\ldots,y_{N_2})$ from $q(x)$ so that $(f(x_1),\ldots,f(x_{N_1}))$ and $(f^*(y_1),\ldots,f^*(y_{N_2}))$ constitute independent samples from the distributions of $f(X)$ and $f^*(Y)$, respectively. The problem is thus reduced to the usual variance comparison problem. As we have mentioned earlier, however, the usual assumptions of normality and of identical distributions, except for location and scale, are violated. Consequently, robust estimation procedures are required.

In this chapter we discuss several two-sample estimation techniques and define the two-sample estimators of v_p/v_q that are included in the Monte Carlo study. Three variations of the jackknife method are considered; Arvesen's two-sample jackknife [1], applied to the ratio of the sample variances of $f(X)$ and $f^*(Y)$; Arvesen's two-sample jackknife, applied to the natural logarithm of the ratio of the sample variances; and Miller's jackknife technique [13], which involves jackknifing $\log(S^2)$ separately for each sample and then combining the results.

Arvesen's two-sample jackknife technique with logs differs from Miller's method in that Arvesen's procedure involves computing the natural logarithm of the ratio of the sample variances and then jackknifing by successively deleting each group of observations in the first sample with the second intact and then deleting each group of observations in the second sample with the first intact. On the other hand, Miller's method requires jackknifing the natural logarithm of the sample variances separately and then combining the resulting jackknife statistics.

The other methods discussed are the classical F method of variance comparison; a modified version of a method suggested by Box-Andersen [5]; and the Box procedure [4], which requires dividing the samples into disjoint subgroups, calculating $\log(S^2)$ for each subgroup and then comparing the two sets of values by a t-test for location.

The following definitions will prove helpful in the subsequent discussions. Let X and Y be random variables with densities $p(x)$ and $q(x)$, respectively. Denote the k^{th} central moments of $f(X)$ and $f^*(Y)$ by $\mu_k(f)$ and $\mu_k(f^*)$, respectively, and define

$$\sigma_f^2 = \mu_2(f) \quad , \quad \sigma_{f^*}^2 = \mu_2(f^*) \ ,$$

$$\gamma_1(f) = \mu_3(f)/\sigma_f^3 \quad , \quad \gamma_1(f^*) = \mu_3(f^*)/\sigma_{f^*}^3 \tag{3.1}$$

$$\gamma_2(f) = [\mu_4(f)/\sigma_f^4] - 3 \quad , \quad \gamma_2(f^*) = [\mu_4(f^*)/\sigma_{f^*}^4] - 3 \ .$$

Recall that $\hat{I}_p = \sum_{i=1}^{N_1} f(x_i)/N_1$ and $\hat{I}_q = \sum_{i=1}^{N_2} f^*(y_i)/N_2$ are unbiased estimates of $I = \int f(x)p(x)dx$, where $(x_1,\ldots,x_{N_1})$ and $(y_1,\ldots,y_{N_2})$ are samples from $p(x)$ and $q(x)$, respectively.

3.2 Miller's Two-sample Jackknife Estimator (J_{M2L})

Miller [13] suggests that the log transformation applied to the estimator S^2 before jackknifing can prove beneficial. This transformation tends to stabilize the variance and reduce the asymmetry of the distribution of the estimate.

The use of variance stabilizing transformations on the estimator, in conjunction with the jackknife, is a widely accepted practice among advocates of the jackknife method. In fact, Miller [14] indicates the need for such transformations to prevent distortion of the results. Relatively little research, however, has been done to study the connection between transformations and the jackknife.

We now define J_{M2L}, Miller's estimator applied to the importance sampling problem. Let $(x_1,\ldots,x_{N_1})$ be a sample from $p(x)$ and let $(y_1,\ldots,y_{N_2})$ be a sample from $q(x)$, where $N_1 = n_1k_1$ and $N_2 = n_2k_2$. Define

$$\theta_f = \log(\sigma_f^2) \quad , \quad \theta_{f*} = \log(\sigma_{f*}^2) \; ,$$

and (3.2)

$$\hat{\theta}_f = \log(S_f^2) \quad , \quad \hat{\theta}_{f*} = \log(S_{f*}^2) \; ,$$

where

$$S_f^2 = \sum_{i=1}^{N_1} [f(x_i) - \hat{I}_p]^2/(N_1 - 1)$$

and (3.3)

$$S_{f*}^2 = \sum_{i=1}^{N_2} [f*(y_i) - \hat{I}_q]^2/(N_2 - 1) \; .$$

Divide the sample $(x_1,\ldots,x_{N_1})$ into n_1 subsamples of size k_1 and the sample $(y_1,\ldots,y_{N_2})$ into n_2 subsamples of size k_2. Let $J_i(\hat{\theta}_f)$, $J_j(\hat{\theta}_{f*})$, $J(\hat{\theta}_f)$ and $J(\hat{\theta}_{f*})$ be defined as in (2.1) and (2.2), and let

$$S_1^2 = \sum_{i=1}^{n_1} [J_i(\hat{\theta}_f) - J(\hat{\theta}_f)]^2/(n_1 - 1) ,$$

$$S_2^2 = \sum_{j=1}^{n_2} [J_j(\hat{\theta}_{f*}) - J(\hat{\theta}_{f*})]^2/(n_2 - 1) \tag{3.4}$$

and

$$S_J^2 = [S_1^2 + (n_1/n_2)S_2^2]/n_1 .$$

Define

$$T = \{[J(\hat{\theta}_f) - \theta_f] - [J(\hat{\theta}_{f*}) - \theta_{f*}]\}/S_J . \tag{3.5}$$

Then the statistic T is suggested as a basis for tests and confidence intervals for $\log(\sigma_f^2/\sigma_{f*}^2)$. If $\gamma_2(f) = \gamma_2(f^*)$ and the subsample sizes are equal $(k_1 = k_2)$, then the statistic T is to be treated as a Student-t random variable with $n_1 + n_2 - 2$ degrees of freedom.

It is not difficult to show that the asymptotic variance of the pseudo-values, $J_i(\hat{\theta}_f)$ and $J_j(\hat{\theta}_{f*})$, are $[\gamma_2(f) + 2]/k_1$ and $[\gamma_2(f^*) + 2]/k_2$, respectively. Thus, if $\gamma_2(f) \neq \gamma_2(f^*)$ or $k_1 \neq k_2$, the degrees of freedom are somewhat ambiguous, as for the two-sample t statistic of mean difference with unequal variances.

In practice it is unlikely that $\gamma_2(f) = \gamma_2(f^*)$, and, in fact, the importance sampling problem may lead to radically different values for $\gamma_2(f)$ and $\gamma_2(f^*)$. The effect of $\gamma_2(f)$ and $\gamma_2(f^*)$ upon the distributions of $J(\hat{\theta}_f)$ and $J(\hat{\theta}_{f^*})$ will become apparent in the following. The following theorem is an application of Theorems 2.3 and 2.4.

<u>THEOREM 3.1</u>. Let $X_1,\ldots,X_{N_1}$ $(N_1 = n_1k_1)$ be IID random variables with density $p(x)$, and let $\hat{\theta}_f$, θ_f , S_1^2 and $J(\hat{\theta}_f)$ be defined as in the definition of T (3.5). Let the fourth moment of $f(X_1)$, $E[f(x_1)^4]$, be finite. Then, as $n_1 \to \infty$,

$$n_1^{\frac{1}{2}}[J(\hat{\theta}_f) - \theta_f] \underset{D}{\to} N(0,\sigma_1^2) ,$$

where $\sigma_1^2 = \gamma_2(f) + 2$. In addition, $S_1^2 \underset{P}{\to} \sigma_1^2$.

<u>PROOF</u>. Denote the random variables $f(X_i)$ by F_i , $i = 1,\ldots,N_1$. Then the sample variance S_f^2 , defined as in (3.3), is a U-statistic with symmetric kernel $k^*(F_1,F_2) = (F_1 - F_2)^2/2$, and with k_c^* as in 2.6,

$$k_1^*(f_1) = E[(F_1 - F_2)^2/2 \mid F_1 = f_1]$$

$$= \tfrac{1}{2}[f_1^2 - 2\mu_1(f)f_1 + E(F_2^2)] \ .$$

Now, since $E[f(X_1)^4] < \infty$, ζ_1 exists and is given by

$$\zeta_1 = \mathrm{Var}[k_1^*(F_1)]$$

$$= \tfrac{1}{4}\mathrm{Var}[F_1^2 - 2\mu_1(f)F_1]$$

$$= \tfrac{1}{4}[\mathrm{Var}(F_1^2) + 4\mu_1(f)^2\mu_2(f) - 4\mu_1(f)\mathrm{Cov}(F_1,F_1^2)] \ ,$$

which, upon simplifying, yields

$$\zeta_1 = \tfrac{1}{4}[\mu_4(f) - \mu_2(f)^2] \ .$$

The theorem now follows immediately from Theorems 2.3 and 2.4 with $g(x) = \log(x)$ and $\eta = \sigma_f^2$. □

The next theorem establishes the asymptotic normality of the statistic T defined in (3.5).

<u>THEOREM 3.2</u>. Let $X_1,\ldots,X_{N_1}$ $(N_1 = n_1k_1)$ be N_1 IID random variables with density $p(x)$, let $Y_1,\ldots,Y_{N_2}$ $(N_2 = n_2k_2)$ be N_2 IID random variables with density $q(y)$, and let the X's and the Y's be independent. Let T be defined as in (3.5), and let the fourth moments of $f(X_1)$ and $f^*(Y_1)$ be finite. Then, if $n_1 \to \infty$ such that $\lim(n_1/n_2)$ exists,

$$T \underset{D}{\rightarrow} N(0,1) .$$

PROOF. From Theorem 3.1, we have

$$n_1^{\frac{1}{2}}[J(\hat{\theta}_f) - \theta_f] \underset{D}{\rightarrow} N(0,\sigma_1^2)$$

and

$$S_1^2 \underset{P}{\rightarrow} \sigma_1^2 ,$$

where $\sigma_1^2 = \gamma_2(f) + 2$.

A similar result can be obtained for $J(\hat{\theta}_{f*})$; that is,

$$n_2^{\frac{1}{2}}[J(\hat{\theta}_{f*}) - \theta_{f*}] \underset{D}{\rightarrow} N(0,\sigma_2^2)$$

and

$$S_2^2 \underset{P}{\rightarrow} \sigma_2^2 ,$$

where $\sigma_2^2 = \gamma_2(f*) + 2$.

Now, if $\lim_{n_1 \to \infty}(n_1/n_2) = c$, we have

$$n_1^{\frac{1}{2}}[J(\hat{\theta}_{f*}) - \theta_{f*}] \underset{D}{\rightarrow} N(0,c\sigma_2^2)$$

and

$$(n_1/n_2)S_2^2 \underset{P}{\rightarrow} c\sigma_2^2 .$$

But $J(\hat{\theta}_f)$ and $J(\hat{\theta}_{f*})$ are independent, since the X's and the Y's are independent; thus, as $n_1 \to \infty$,

$$n_1^{\frac{1}{2}}\{[J(\hat{\theta}_f) - \theta_f] - [J(\hat{\theta}_{f*}) - \theta_{f*}]\} \underset{D}{\to} N(0,\sigma_1^2 + c\sigma_2^2)$$

and

$$S_1^2 + (n_1/n_2)S_2^2 \underset{P}{\to} \sigma_1^2 + c\sigma_2^2 \ ,$$

and therefore the statistic T is asymptotically distributed as a normal random variable with mean 0 and variance 1 . □

Theorem 3.2 justifies the use of the normal distribution or the t distribution to obtain approximate confidence intervals for $\log(\sigma_f^2/\sigma_{f*}^2)$. The t distribution will, of course, result in more conservative confidence bounds. To obtain point and interval estimates of the parameter of interest, $v_p/v_q = \sigma_f^2/\sigma_{f*}^2$, we apply the exponential transformation to the estimator of $\log(v_p/v_q)$ and to its associated confidence bounds. This yields an estimator J_{M2L} of the variance ratio v_p/v_q defined by

$$J_{M2L} = \exp(J(\hat{\theta}_f) - J(\hat{\theta}_{f*})) \ . \tag{3.6}$$

In the Monte Carlo study, J_{M2L} is defined as in (3.6) with $N_1 = N_2$ and $k_1 = k_2 = 1$. The confidence intervals for v_p/v_q are obtained as follows.

Let

$$C_u = J(\hat{\theta}_f) - J(\hat{\theta}_{f^*}) + z_{\alpha/2}S_J$$

and

$$C_l = J(\hat{\theta}_f) - J(\hat{\theta}_{f^*}) - z_{\alpha/2}S_J ,$$

where $\text{Prob}(V > z_{\alpha/2}) = \alpha/2$ and $V \sim N(0,1)$. Then (C_l, C_u) is an approximate $100(1 - \alpha)\%$ confidence interval for $\log(v_p/v_q)$, and, since the exponential transformation is a monotonic increasing function, $(\exp(C_l), \exp(C_u))$ is an approximate $100(1 - \alpha)\%$ confidence interval for v_p/v_q.

It is worth noting, at this point, that although the jackknife estimator may eliminate a first order bias term for $\log(v_p/v_q)$, the exponential transformation may reintroduce a first order bias term for v_p/v_q. To see this, consider an estimator $\hat{\theta}$ of some parameter θ, and the estimator $\hat{\beta} = \exp(\hat{\theta})$ of $\exp(\theta)$. Expanding $\exp(\hat{\theta})$ in a Taylor series about θ, we obtain

$$\exp(\hat{\theta}) = \exp(\theta)[1 + (\hat{\theta} - \theta) + (\hat{\theta} - \theta)^2/2 + \ldots]$$

and

$$E(\exp(\hat{\theta})) = \exp(\theta) + \exp(\theta)[\text{bias}(\hat{\theta}) + \tfrac{1}{2}\text{MSE}(\hat{\theta}) + \ldots] .$$

Thus, the bias of $\hat{\beta}$ is given by

$$\text{bias}(\hat{\beta}) = \exp(\theta)[\text{bias}(\hat{\theta}) + \text{MSE}(\hat{\theta}) + \ldots] .$$

Note that even if $\hat{\theta}$ were an unbiased estimate of θ, the bias of $\hat{\beta}$ will generally be of order $1/N$, since $\text{Var}(\hat{\theta})$ is usually of order $1/N$.

3.3 Arvesen's Two-sample Estimator with Logs (J_{A2L})

The J_{A2L} estimator, defined in this section, is based on Arvesen's two-sample jackknife estimator applied to the importance sampling problem. As for Millers' estimate J_{M2L}, the log transformation is used in conjunction with the jackknife.

Let $X_1,\ldots,X_{N_1}$ $(N_1 = n_1 k_1)$ be IID random variables with density $p(x)$, let $Y_1,\ldots,Y_{N_2}$ $(N_2 = n_2 k_2)$ be IID random variables with density $q(y)$, and let the X's and the Y's be independent. Divide the X's into n_1 groups of size k_1 and the Y's into n_2 groups of size k_2. Let

$$F_i = f(X_i) \quad , \quad i = 1,\ldots,N_1 ,$$

$$F_i^* = f^*(Y_i) \quad , \quad i = 1,\ldots,N_2 ,$$

$$\hat{\theta} = \log(S_f^2/S_{f*}^2)$$

and

$$\theta = \log(v_p/v_q) \ ,$$

where S_f^2 and S_{f*}^2 , defined as in the previous section, are the sample variances of $(F_1,\ldots,F_{N_1})$ and $(F_1^*,\ldots,F_{N_2}^*)$, respectively.

Let $\hat{\theta}_{i,\cdot}$ be the estimator $\hat{\theta}$ after deletion of the i^{th} group of F's , and let $\hat{\theta}_{\cdot,j}$ be the estimator $\hat{\theta}$ after deletion of the j^{th} group of F*'s .

Then $J_{1,i}$, $J_{2,j}$, J_1 , J_2 and J , defined as in (2.12) and 2.13), become

$$J_{1,i} = J_i(\hat{\theta}_f) - \log(S_{f*}^2) \ ,$$

$$J_{2,j} = \log(S_f^2) - J_j(\hat{\theta}_{f*}) \ ,$$

$$J_1 = J(\hat{\theta}_f) - \log(S_{f*}^2) \ ,$$

$$J_2 = \log(S_f^2) - J(\hat{\theta}_{f*})$$

and

$$J = \{n_1[J(\hat{\theta}_f) - \log(S_{f*}^2)] + n_2[\log(S_f^2) - J(\hat{\theta}_{f*})]\}/(n_1+n_2) \ ,$$

where $J(\hat{\theta}_f)$, $J(\hat{\theta}_{f*})$, $J_i(\hat{\theta}_f)$ and $J_j(\hat{\theta}_{f*})$ are defined as in the previous section.

The variance estimate S_g^2 , as in (2.12), is given by

$$S_g^2 = S_1^2 + (n_1/n_2)S_2^2 ,$$

where S_1^2 and S_2^2 are defined as in (3.4).

From Theorem 2.8, if the fourth moment of F_1 is finite, the fourth moment of F_1^* is finite, $n_1 \le n_2$ and $n_1 \to \infty$ such that $\lim(n_1/n_2)$ exists, then

$$Z = n_1^{\frac{1}{2}}[J - \log(v_p/v_q)]/S_g \underset{D}{\rightarrow} N(0,1) .$$

Thus, for n_1 large, Z is approximately normally distributed with mean zero and variance 1 .

The estimator J_{A2L} of v_p/v_q is defined by

$$J_{A2L} = \exp(J) . \tag{3.7}$$

As for the J_{M2L} estimator, the statistic J_{A2L} , entering the Monte Carlo study, is defined with $N_1 = N_2$ and $k_1 = k_2 = 1$. Approximate confidence intervals for v_p/v_q are obtained in an analogous manner to those for the estimator J_{M2L} . That is, if

$$C_u = J + z_{\alpha/2}S_g/n_1^{\frac{1}{2}}$$

and

$$C_1 = J - z_{\alpha/2}S_g/n_1^{\frac{1}{2}} ,$$

where $\text{Prob}(V > z_{\alpha/2}) = \alpha/2$ and $V \sim N(0,1)$, then (C_l, C_u) is an approximate $100(1-\alpha)\%$ confidence interval for $\log(v_p/v_q)$. Moreover, since the exponential transformation is a monotonic increasing function, $(\exp(C_l), \exp(C_u))$ is an approximate $100(1-\alpha)\%$ confidence interval for v_p/v_q.

Both Miller's method and Arvesen's method are valid asymptotically, but, as yet, there has been little research done to determine whether one is superior to the other. A comparison of the performance of these two estimators, in the Monte Carlo study, is given in Chapter 6.

3.4 Arvesen's Two-sample Estimator Without Logs (J_{A2})

The estimator J_{A2} is Arvesen's two-sample jackknife estimate applied to $\hat{\theta} = S_f^2/S_{f*}^2$ $(\theta = v_p/v_q)$. That is,

$$J_{A2} = J \; , \qquad (3.8)$$

where J is as in (2.13) with $\hat{\theta} = S_f^2/S_{f*}^2$. This estimate is included in the Monte Carlo study, thus allowing the effect of the logarithmic transformation of the J_{A2L} estimate to be studied.

If the fourth moments of F_1 and F_1^* are finite, $n_1 \leq n_2$ and $n_1 \to \infty$ such that $\lim(n_1/n_2)$

exists, Theorem 2.8 yields

$$Z = n_1^{\frac{1}{2}}(J_{A2} - v_p/v_q)/S_g \underset{D}{\to} N(0,1) \quad ,$$

where S_g^2 is defined as in (2.12).

Thus, for large n_1 , Z is approximately distributed as a normal random variable with mean zero and variance 1 . Confidence intervals for v_p/v_q are then based on this normal approxiamtion in the usual manner. Again, the form of the estimator J_{A2} which enters the Monte Carlo study is as in (3.8) with $N_1 = N_2$ and $k_1 = k_2 = 1$.

3.5 The Box Estimator (B_k)

Let $(x_1,\ldots,x_{N_1})$ and $(y_1,\ldots,y_{N_2})$ be independent samples from $p(x)$ and $q(x)$, respectively. The Box method requires that $(f(x_1),\ldots,f(x_{N_1}))$ and $(f^*(y_1),\ldots,f^*(y_{N_2}))$ each be divided into subgroups of size $k > 1$. $\text{Log}(S^2)$ is then computed for each subgroup and the two sets of values are compared by a two-sample t-test for location. The details of the procedure are given below.

Let $X_1,\ldots,X_{N_1}$ be IID random variables with density $p(x)$, let $Y_1,\ldots,Y_{N_2}$ be IID random variables with density $q(x)$, and let the X's and the Y's be independent. Let

$$F_i = f(X_i) \quad , \quad i = 1,2,\ldots,N_1 \; ,$$

and (3.9)

$$F_i^* = f^*(Y_i) \quad , \quad i = 1,2,\ldots,N_2 \; .$$

Then $F_1,\ldots,F_{N_1}$ are IID random variables and $F_1^*,\ldots,F_{N_2}^*$ are IID random variables and the F's and F*'s are independent. Now divide the F's and the F*'s into subgroups of size $k > 1$ $(N_1 = n_1 k,\; N_2 = n_2 k)$. Define

$$V_i = \log(U_F^i) \quad , \quad i = 1,\ldots,n_1 \; , \tag{3.10}$$

and

$$W_j = \log(U_{F^*}^j) \; , \quad j = 1,\ldots,n_2 \; ,$$

where U_F^i and $U_{F^*}^j$ are the sample variances based on the i^{th} and j^{th} groups of the F's and F*'s , respectively.

The random variables $V_1,\ldots,V_{n_1}$ and $W_1,\ldots,W_{n_2}$ are then treated as independent identically distributed random variables from normal distributions with equal variances and means $\log(\sigma_f^2)$ and $\log(\sigma_{f^*}^2)$, respectively. Define a statistic T by

$$T = \left(\frac{n_1 n_2 (n_1+n_2-2)}{(n_1+n_2)}\right)^{\frac{1}{2}} \frac{[\bar{V}-\bar{W}-\log(\sigma_f^2/\sigma_{f*}^2)]}{[(n_1-1)S_V^2+(n_2-1)S_W^2]^{\frac{1}{2}}} \tag{3.11}$$

where

$$S_V^2 = \sum_{i=1}^{n_1} (V_i - \bar{V})^2/(n_1 - 1)$$

and

$$S_W^2 = \sum_{j=1}^{n_2} (W_j - \bar{W})^2/(n_2 - 1) .$$

T is then assumed to have a Student-t distribution with $n_1 + n_2 - 2$ degrees of freedom and is the basis for confidence interval construction and hypothesis testing.

The theoretical justification of the above t approximation requires that the standardized fourth cumulants of the F's and the F^*'s be equal, i.e. $\gamma_2(f) = \gamma_2(f^*)$. This requirement is necessary to assure the equality of the asymptotic variances of V_i and W_j, as will become apparent in the proof of the following theorem.

<u>THEOREM 3.3</u>. Let $F_1,\ldots,F_{N_1}$ be N_1 $(N_1 = n_1 k)$ independent identically distributed random variables, and let $F_1^*,\ldots,F_{N_2}^*$ be N_2 $(N_2 = n_2 k)$ independent identically distributed random variables, such that the F's and the F^*'s are independent and $\gamma_2(f) = \gamma_2(f^*)$.

Define V_i, $i = 1,\ldots,n_1$, W_j, $j = 1,\ldots,n_2$, and T as in (3.10) and (3.11). Then for n_1 and n_2 fixed, as $N_1 \to \infty$ and $N_2 \to \infty$,

$$T \underset{D}{\to} t(n_1 + n_2 - 1) ,$$

where $t(n_1 + n_2 - 2)$ denotes the Student-t distribution with $n_1 + n_2 - 2$ degrees of freedom.

PROOF. The random variables $V_1,\ldots,V_{n_1}$ are independent and identically distributed since they are based on disjoint subsets of the IID random variables $F_1,\ldots,F_{N_1}$. Similarly, $W_1,\ldots,W_{n_2}$ are IID random variables. Also, the V's and W's are independent because the F's and F*'s are independent.

Now, from Theorem 2.9, since the sample variance U_F^i is asymptotically normally distributed with mean σ_f^2 and variance $[\mu_4(f) - \sigma_f^4]/k$, the random variable $V_i = \log(U_F^i)$ is asymptotically normal with mean $\log(\sigma_f^2)$ and variance $[\mu_4(f)/\sigma_f^4 - 1]/k = [\gamma_2(f) + 2]/k$. Similarly, $W_j = \log(U_{F*}^j)$ is asymptotically normal with mean $\log(\sigma_{f*}^2)$ and variance $[\gamma_2(f*) + 2]/k$. Thus, as $k \to \infty$,

$$k^{\frac{1}{2}}[V_i - \log(\sigma_f^2)] \underset{D}{\to} A_i \sim N(0, \gamma_2(f) + 2) , \; i = 1,\ldots,n_1 ,$$

$$k^{\frac{1}{2}}[W_j - \log(\sigma^2_{f*})] \underset{D}{\to} B_j \sim N(0, \gamma_2(f^*) + 2) \quad , \; i = 1,\ldots,n_2 \; ,$$

and the random variables $A_1,\ldots A_{n_1}$ and $B_1,\ldots,B_{n_2}$ retain the properties of independence possessed by the random variables $V_1,\ldots,V_{n_1}$ and $W_1,\ldots,W_{n_2}$.

Define the statistic T^* by

$$T^* = \left(\frac{n_1 n_2 (n_1+n_2-2)}{(n_1+n_2)}\right)^{\frac{1}{2}} \frac{(\bar{A}-\bar{B})}{[(n_1-1)S_A^2+(n_2-1)S_B^2]^{\frac{1}{2}}} \quad ,$$

where

$$S_A^2 = \sum_{i=1}^{n_1} (A_i - \bar{A})^2/(n_1 - 1)$$

and

$$S_B^2 = \sum_{i=1}^{n_2} (B_i - \bar{B})^2/(n_2 - 1) \; .$$

Then, since $\gamma_2(f) = \gamma_2(f^*)$, we have $Var(A_1) = Var(B_1)$, and the statistic T^* has a Student-t distribution with $n_1 + n_2 - 2$ degrees of freedom.

Now, reasoning as in the proof of Theorem 2.5 and applying Corollary 1 of Billingsley [3, p.31], we obtain

$$T \underset{D}{\to} T^* \sim t(n_1 + n_2 - 2) \; . \quad \square$$

If $\gamma_2(f) \neq \gamma_2(f^*)$, then $\sigma_A^2 \neq \sigma_B^2$ (where $\sigma_A^2 = \mathrm{Var}(A_1)$, $\sigma_B^2 = \mathrm{Var}(B_1)$) and, consequently, the statistic T^* will not have a Student-t distribution with $n_1 + n_2 - 2$ degrees of freedom. That is, the statistic T will not be distributed asymptotically as $t(n_1 + n_2 - 2)$.

Define a statistic t^* by

$$t^* = (\bar{A} - \bar{B})/(S_A^2/n_1 + S_B^2/n_2)^{\frac{1}{2}} .$$

Then, Wetherill [23, p.160] suggests that if $\sigma_A^2 \neq \sigma_B^2$, the statistic t^* has an approximate t distribution with 'effective' degrees of freedom f , where

$$1/f = \sigma_A^4/G^2 n_1^2 (n_1 - 1) + \sigma_B^4/G^2 n_2^2 (n_2 - 1) , \tag{3.12}$$

with

$$G = \sigma_A^2/n_1 + \sigma_B^2/n_2 .$$

If σ_A^2 and σ_B^2 are unknown, they may be replaced by their sample estimates S_A^2 and S_B^2 .

The formula (3.12) is obtained by finding the degrees of freedom of a t random variable with approximately the same variance as t^*. In the derivation of (3.12), Wetherill uses $1 + 2/f$ to approximate the variance of the t distribution with f degrees of freedom, instead of its actual variance $1 + 2/(f - 2)$. A derivation of (3.12) follows.

Let $\bar{A} - \bar{B} = X$ and $(S_A^2/n_1 + S_B^2/n_2) = Y$. Then X and Y are independent, since $\bar{A}$ and S_A^2 are independent, $\bar{B}$ and S_B^2 are independent and the A's and B's are independent.

Now

$$\begin{aligned} \mathrm{Var}(t^*) &= \mathrm{Var}(X/Y^{\frac{1}{2}}) \\ &= E[\mathrm{Var}(X/Y^{\frac{1}{2}}|Y)] + \mathrm{Var}[E(X/Y^{\frac{1}{2}}|Y)] , \end{aligned}$$

and, since X and Y are independent,

$$\mathrm{Var}(X/Y^{\frac{1}{2}}|Y) = \mathrm{Var}(X)/Y$$

and

$$\begin{aligned} E(X/Y^{\frac{1}{2}}|Y) &= E(X)/Y^{\frac{1}{2}} \\ &= 0 , \end{aligned}$$

since $E(X) = 0$. Therefore

$$\mathrm{Var}(t^*) = \mathrm{Var}(X)E(1/Y) .$$

To evaluate $E(1/Y)$, we expand $1/Y$ in a Taylor series about $E(Y) = G$. Thus,

$$1/Y = 1/G - (Y - G)/G^2 + (Y - G)^2/G^3 - \ldots$$

and

$$E(1/Y) \doteq 1/G + \mathrm{Var}(Y)/G^3$$

$$= 1/G + [\mathrm{Var}(S_A^2/n_1) + \mathrm{Var}(S_B^2/n_2)]/G^3 .$$

But

$$\mathrm{Var}(S_A^2) = 2\sigma_A^4/(n_1 - 1)$$

and

$$\mathrm{Var}(S_B^2) = 2\sigma_B^4/(n_2 - 1) .$$

Therefore

$$E(1/Y) \doteq 1/G + 2[\sigma_A^4/n_1^2(n_1-1) + \sigma_B^4/n_2^2(n_2-1)]/G^3 .$$

Noting that $\mathrm{Var}(X) = G$, we obtain

$$\mathrm{Var}(t^*) \doteq 1 + 2[\sigma_A^4/n_1^2(n_1-1) + \sigma_B^4/n_2^2(n_2-1)]/G^2 . \quad (3.13)$$

Equating (3.13) to the approximate variance of the t distribution with f degrees of freedom, i.e. to $(1 + 2/f)$, and solving for $1/f$, yields (3.12).

If we were to equate (3.13) to the actual variance of a t random variable with f degrees of freedom, we would obtain 'effective' degrees of freedom f' , where $f' = f + 2$. Thus, Wetherill's procedure is slightly more conservative.

In practice σ_A^2 and σ_B^2 are usually unknown and the assumption that $\sigma_A^2 = \sigma_B^2$ may be unrealistic. If σ_A^2 is very small relative to σ_B^2, most of the error will be due to σ_B^2, and consequently the 'effective' degrees of freedom would be nearer to $n_2 - 1$, as is indicated by (3.12).

Note that the statistic t, defined by

$$t = [\bar{V} - \bar{W} - \log(v_p/v_q)]/(S_V^2/n_1 + S_W^2/n_2)^{\frac{1}{2}} ,$$

is asymptotically distributed as t^*. Thus, Wetherill's method may be used to obtain approximate confidence intervals and tests for $\log(v_p/v_q)$ when $\gamma_2(f) \neq \gamma_2(f^*)$. Point and intervals estimates of v_p/v_q can then be obtained as for J_{M2L} and J_{A2L}.

The problem of determining an optimum subgroup size k has not, as yet, been solved. Recall that the statistics T and t will be approximately distributed as t random variables if V_i and W_j are approximately normally distributed, that is, if the subgroup size k is large. Increasing k, however, results in loss of degrees of freedom of the t statistic. Furthermore, $\bar{V} - \bar{W}$, although consistent, is a biased estimator of $\log(v_p/v_q)$ and small k may result in a large bias accompanied by a short confidence interval. To complicate matters further, if $\gamma_2(f) \neq \gamma_2(f^*)$, we must use

S_V^2 and S_W^2 in place of σ_A^2 and σ_B^2, in (3.12), to estimate the 'effective' degrees of freedom f. Thus, the value obtained for f may not be reliable for small k.

In view of the fact that $\gamma_2(f)$ and $\gamma_2(f^*)$ may differ greatly in magnitude, we shall consider a statistic T_z which is distributed asymptotically as a t random variable, but requires that the sample sizes and the group sizes be equal, i.e. $n_1 = n_2$ and $N_1 = N_2$.

We begin by randomly pairing the V_i's and the W_i's and defining

$$Z_i = V_i - W_i \quad , \quad i = 1,\ldots,n \; ,$$

$$S_z^2 = \sum_{i=1}^{n} (Z_i - \bar{Z})^2/(n - 1)$$

and

$$T_z = n^{\frac{1}{2}}[\bar{Z} - \log(v_p/v_q)]/S_z \; .$$

Then $Z_1,\ldots,Z_n$ are IID random variables and as $k\to\infty$, with n fixed,

$$T_z \underset{D}{\to} t(n - 1) \; .$$

The proof of this result is similar to that for Theorem 3.3.

Now define

$$B_k = \exp(\bar{Z}) \ , \qquad (3.14)$$

where k indicates the group size. Then B_k is a consistent estimator of v_p/v_q, and confidence intervals for v_p/v_q are obtained as follows. Let

$$C_u = \exp(\bar{Z} + t_{\alpha/2} S_z / n^{\frac{1}{2}})$$

and

$$C_l = \exp(\bar{Z} - t_{\alpha/2} S_z / n^{\frac{1}{2}}) \ ,$$

where $\mathrm{Prob}(X > t_{\alpha/2}) = \alpha/2$ and $X \sim t(n - 1)$. Then (C_l, C_u) is an approximate $100(1 - \alpha)\%$ confidence interval for v_p/v_q .

In the Monte Carlo study, the estimators B_5 and B_{10} are implemented with $n = 30$.

3.6 The Box-Andersen Estimator (F_{BA})

The Box-Andersen test [5] for comparing variances involves adjusting the degrees of freedom of the classical F or beta test in order to reduce the effects of non-normality. The adjustment is obtained by equating the first two moments of the beta distribution to the first two moments of the beta statistic under the permutation distribution. For example, if $(X_1, \ldots, X_n)$ and

$(Y_1, \ldots, Y_m)$ are independent samples, the test is obtained by comparing the F ratio S_x^2/S_y^2 with the critical points of an F distribution with $d(n-1), d(m-1)$ degrees of freedom. The adjustment d is given by

$$d = (1 + \tfrac{1}{2}\hat{\gamma}_2)^{-1} ,$$

where $\hat{\gamma}_2$ is an estimate of $\gamma_2(x)$, $\gamma_2(x)$ being defined in a manner analogous to $\gamma_2(f)$.

This procedure requires that the random variables $(X_1 - \mu_x)/\sigma_x$ and $(Y_1 - \mu_y)/\sigma_y$ be identically distributed, where $\mu_x = E(X_1)$, $\mu_y = E(Y_1)$, $\sigma_x^2 = \mathrm{Var}(X_1)$ and $\sigma_y^2 = \mathrm{Var}(Y_1)$. With this requirement satisfied, $\gamma_2(x) = \gamma_2(y)$ and a pooled estimate of γ_2 can be obtained. Thus, if $r = \sigma_x^2/\sigma_y^2$, the estimator $\hat{\gamma}_2$, suggested by Shorack [20], is defined by

$$\hat{\gamma}_2 + 3 = (m+n)\left[r^4 \sum_{i=1}^{n} (X_i - \bar{X})^4 + \sum_{j=1}^{m} (Y_j - \bar{Y})^4\right]$$

$$\Big/\left[r^2 \sum_{i=1}^{n} (X_i - \bar{X})^2 + \sum_{j=1}^{m} (Y_j - \bar{Y})^2\right]^2 . \qquad (3.14)$$

Miller [13] defines $\hat{\gamma}_2$ as in (3.14) with $r = 1$, which is the value obtained under the null hypothesis of equal variances. Approximate confidence intervals for r can be constructed for this procedure, but the calculations are somewhat tedious.

In view of the fact that it is unrealistic to assume that the random variables $(X_1 - \mu_x)/\sigma_x$ and $(Y_1 - \mu_y)/\sigma_y$ are identically distributed, we consider an alternative approach, which is outlined in Plackett [7, pp.83-84].

Adopting notation similar to Plackett's, let $S_x^2(\gamma_2(x);\ n-1)$ and $S_y^2(\gamma_2(y);\ m-1)$ be the sample variances of the X's and Y's, respectively, where $\gamma_2(x)$ and $\gamma_2(y)$ are the standardized fourth cumulants of X and Y, respectively, as in (3.1). Then

$$E[S_x^2(\gamma_2(x);n-1)] = \sigma_x^2 ,$$

$$E[S_y^2(\gamma_2(y);m-1)] = \sigma_y^2 ,$$

$$\text{Var}[S_x^2(\gamma_2(x);n-1)] = 2\sigma_x^4[1+(n-1)\gamma_2(x)/2n]/(n-1)$$

and

$$\text{Var}[S_y^2(\gamma_2(y);m-1)] = 2\sigma_y^4[1+(m-1)\gamma_2(y)/2m]/(m-1) .$$

Since $S_x^2(\gamma_2(x);n-1)$ and $S_y^2(\gamma_2(y);m-1)$ are asymptotically normal, they have the same limiting distributions as $S_x^2(0;\delta_x(n-1))$ and $S_y^2(0;\delta_y(m-1))$, respectively, where

$$\delta_x = [1 + (n-1)\gamma_2(x)/2n]^{-1}$$

and (3.15)

$$\delta_y = [1 + (m-1)\gamma_2(y)/2m]^{-1} .$$

Therefore, the statistic F defined by

$$F = S_x^2(\gamma_2(x);n-1)/S_y^2(\gamma_2(y);m-1) ,$$

has the same limiting distribution as

$$F' = S_x^2(0;\delta_x(n-1))/S_y^2(0;\delta_y(m-1)) .$$

Thus, for finite samples we infer that rF is approximately distributed as an F random variable with $\delta_x(n-1),\delta_y(m-1)$ degrees of freedom. Estimates d_x and d_y of δ_x and δ_y , respectively, are obtained by replacing $\gamma_2(x)$ and $\gamma_2(y)$, in (3.15), by sample estimates $g_2(x)$ and $g_2(y)$, respectively.

To apply this procedure to the importance sampling problem we proceed as follows. Let $F_1,\ldots,F_{N_1}$ and $F_1^*,\ldots,F_{N_2}^*$ be defined as in (3.9). Define

$$M_4(f) = \sum_{i=1}^{N_1} (F_i - \bar{F})^4/(N_1 - 1) ,$$

$$M_4(f^*) = \sum_{i=1}^{N_2} (F_i^* - \bar{F}^*)^4/(N_2 - 1) ,$$

$$g_2(f) = M_4(f)/S_f^4 - 3$$

and

$$g_2(f^*) = M_4(f^*)/S_{f^*}^4 - 3 ,$$

where S_f^2 and S_{f*}^2 are defined as in (3.3). Then $g_2(f)$ and $g_2(f^*)$ are consistent estimators of $\gamma_2(f)$ and $\gamma_2(f^*)$, respectively.

Now define

$$d_f = [1 + (N_1 - 1)g_2(f)/2N_1]^{-1}$$

and

$$d_{f*} = [1 + (N_2 - 1)g_2(f^*)/2N_2]^{-1} .$$

The estimator F_{BA} is defined by

$$F_{BA} = S_f^2/S_{f*}^2 , \qquad (3.16)$$

the classical F statistic. Confidence intervals and tests, however, are based on the F distribution with $d_{f*}(N_2 - 1)$, $d_f(N_1 - 1)$ degrees of freedom. That is, we assume that

$$(v_p/v_q)/F_{BA} \sim F(d_{f*}(N_2 - 1) , d_f(N_1 - 1)) .$$

Let

$$C_u = f_u F_{BA}$$

and

$$C_l = f_l F_{BA} ,$$

where f_u and f_l are such that $\text{Prob}(F' > f_u) = \alpha/2$ and $\text{Prob}(F' < f_l) = \alpha/2$, $F' \sim F(d_{f*}(N_2-1), d_f(N_1-1))$. Then (C_l, C_u) is an approximate $100(1-\alpha)\%$ confidence interval for v_p/v_q.

The interval estimate of v_p/v_q thus accounts for differences in $\gamma_2(f)$ and $\gamma_2(f^*)$. Note that large values of γ_2 could result in very small values for d_f or d_{f*}. This situation, however, would suggest that large samples were required to compensate for the high degree of non-normality indicated by γ_2 or its sample estimate.

3.7 The Classical F Statistic (F)

It is well known that the classical F statistic is extremely non-robust with respect to departures from normality (see Box [4]). For comparative purposes, this classical procedure of variance comparison is included in the Monte Carlo study. In the following we define the F statistic and its associated interval estimate, applied to the problem of estimating v_p/v_q.

Let $F_1, \ldots, F_{N_1}$ and $F_1^*, \ldots, F_{N_2}^*$ be defined as in (3.9). Then define the estimate F of v_p/v_q by

$$F = S_f^2 / S_{f*}^2 , \tag{3.17}$$

where S_f^2 and S_{f*}^2 are as in (3.3).

Let

$$C_l = f_l F$$

and

$$C_u = f_u F ,$$

where $\mathrm{Prob}(F' < f_l) = \alpha/2$ and $\mathrm{Prob}(F' > f_u) = \alpha/2$, $F' \sim F(N_2 - 1, N_1 - 1)$. Then (C_l, C_u) is assumed to be a $100(1 - \alpha)\%$ confidence interval for v_p/v_q.

CHAPTER 4

THE ONE-SAMPLE ESTIMATORS

4.1 Introduction

Recall, from Chapter 1, that the one-sample approach to the problem requires that an estimate of the variance ratio, v_p/v_q , be obtained from a single sample, $(x_1,\ldots,x_n)$, drawn from $q(x)$. We introduced two such estimators in Section 1.4; the Sukhatme estimator, $\hat{\theta}_S$, and the non-negative estimator, $\hat{\theta}_W$.

The Sukhatme estimator is defined as in (1.5), by

$$\hat{\theta}_S = 1 + \{\sum_{i=1}^{n} [f^*(x_i)f(x_i) - f^{*2}(x_i)]/n\}/S_q^2 ,$$

and the non-negative estimator is defined as in (1.8), by

$$\hat{\theta}_W = S_W^2/S_q^2 ,$$

where

$$S_W^2 = \sum_{i=1}^{n} p(x_i)[f(x_i) - \hat{I}_q]^2/nq(x_i) .$$

In this chapter we define the one-sample estimators J_{S1} , T_{S1} , J_{W1} and J_{WIL} and their associated interval estimates. For convenience we shall use the same symbols when referring to the interval estimates as are used to denote the point estimates.

The four methods of obtaining approximate confidence intervals for v_p/v_q are; the jackknife procedure applied to the Sukhatme estimator (J_{S1}) ; the jackknife procedure applied to the non-negative estimator (J_{W1}) ; the jackknife technique applied to the natural logarithm of the non-negative estimator (J_{WIL}) ; and a method based on the asymptotic normality of the Sukhatme estimator and a consistent estimate of its variance (T_{S1}) .

Since the Sukhatme estimator $\hat{\theta}_s$ may yield negative estimates of the variance ratio, we cannot jackknife the log of this estimator as we can for the non-negative estimator $\hat{\theta}_W$. It is possible, however, to obtain an estimate of the variance of $\hat{\theta}_s$ by estimating the variance of the first order terms of a Taylor series expansion of $\hat{\theta}_s$. As we shall see in Section 4.5, the Sukhatme estimator is asymptotically normally distributed, and its variance estimator converges in probability to the asymptotic variance of $\hat{\theta}_s$. These results then justify the use of a normal approximation to obtain approximate confidence intervals for v_p/v_q . The resulting interval estimate is denoted by T_{S1} , and for convenience this symbol is also used for the point estimate, i.e. $T_{S1} = \hat{\theta}_s$.

4.2 The Jackknifed Sukhatme Estimator (J_{S1})

The jackknife technique can be successfully applied to the Sukhatme estimator $\hat{\theta}_s$. Keeping notation consistent with that of Section 2.4, we proceed as follows.

Let $X_1,\ldots,X_N$ $(N = nk)$ be IID random variables with density $q(x)$. Define

$$F_i = f(X_i) \ , \quad i = 1,\ldots,N \ ,$$
$$F_i^* = f^*(X_i) \ , \quad i = 1,\ldots,N \ , \tag{4.1}$$

and the random vectors $Z_1,\ldots,Z_N$ by

$$Z_i = (F_i, F_i^*) \ , \quad i = 1,\ldots,N \ .$$

Let

$$U^1 = \sum_{i=1}^{N} (F_i^* - \bar{F}^*)^2/(N-1) \ , \tag{4.2}$$

and

$$U^2 = \sum_{i=1}^{N} (F_i F_i^* - F_i^{*2})/N \ .$$

Then U^1 and U^2 are U-statistics with symmetric kernels

$$k^{*1}(Z_1, Z_2) = (F_1^* - F_2^*)^2/2$$

and

$$k^{*2}(Z_1) = F_1F_1^* - F_1^{*2} ,$$

respectively. Note that $U^1 = S_q^2$, where S_q^2 is defined as in (1.3), and that $\hat{\theta}_s$ is given by

$$\hat{\theta}_s = (U^1 + U^2)/U^1 .$$

Now, divide the random vectors $Z_1,\ldots,Z_N$ into n subgroups of size k , let $g:R^2 \to R$ be defined by

$$g(t_1,t_2) = (t_1 + t_2)/t_1 ,$$

and let

$$\eta_1 = E[k^{*1}(Z_1,Z_2)] = v_q$$

and

$$\eta_2 = E[k^{*2}(Z_1)] = v_p - v_q .$$

Define the estimator J_{S1} by

$$J_{S1} = J(\hat{\theta}_s) , \tag{4.3}$$

where $J(\hat{\theta}_s)$ is defined as in (2.2) with $\hat{\theta}_s = g(U^1,U^2)$.

Then, if $E\{[k^{*1}(Z_1,Z_2)]^2\} < \infty$, $E\{[k^{*2}(Z_1)]^2\} < \infty$ and $v_q > 0$, the hypotheses of Theorems 2.6 and 2.7 are satisfied. Thus, as $n\to\infty$,

$$n^{\frac{1}{2}}(J_{S1} - v_p/v_q)/S_g \underset{D}{\rightarrow} N(0,1) ,$$

where S_g^2 is defined as in (2.9).

Let

$$C_u = J_{S1} + z_{\alpha/2}S_g/n^{\frac{1}{2}}$$

and

$$C_l = J_{S1} - z_{\alpha/2}S_g/n^{\frac{1}{2}} ,$$

where $\text{Prob}(Y > z_{\alpha/2}) = \alpha/2$ and $Y \sim N(0,1)$. Then approximate $100(1 - \alpha)\%$ confidence intervals for v_p/v_q are given by (C_l, C_u) .

4.3 The Jackknifed $\hat{\theta}_W$ Estimator Without Logs (J_{W1})

The J_{W1} estimator is the jackknife statistic obtained by jackknifing $\hat{\theta} = \hat{\theta}_W$, with $\hat{\theta}_W$ defined as in (1.8). As for the J_{S1} estimator, Theorems 2.6 and 2.7 guarantee the asymptotic normality of the Studentized jackknife when the regularity conditions are satisfied. To define J_{W1} formally, we begin by expressing $\hat{\theta}_W$ as a function of U-statistics. To this end let $F_1,\ldots,F_N$ and $F_1^*,\ldots,F_N^*$ be defined as in (4.1), and define the random variables $C_1,\ldots,C_N$ by

$$C_i = p(X_i)/q(X_i) \quad , \quad i = 1,\ldots,N , \tag{4.4}$$

where $X_1, \ldots, X_N$ are IID random variables with density $q(x)$.

Now define the IID random vectors $Z_1, \ldots, Z_N$ by

$$Z_i = (F_i, F_i^*, C_i) \quad , \quad i = 1, \ldots, N .$$

Let

$$U^1 = \sum_{i=1}^{N} (F_i^* - \bar{F}^*)^2/(N - 1) \ ,$$

$$U^2 = \sum_{i=1}^{N} C_i F_i^2/N \ ,$$

$$U^3 = \sum_{i=1}^{N} F_i^*/N$$

and

$$U^4 = \sum_{i=1}^{N} C_i/N \ .$$

Then U^1 , U^2 , U^3 and U^4 are U-statistics with symmetric kernels

$$k^{*1}(Z_1, Z_2) = (F_1^* - F_2^*)^2/2 \ ,$$

$$k^{*2}(Z_1) = C_1 F_1^2 \ ,$$

$$k^{*3}(Z_1) = F_1^*$$

and

$$k^{*4}(Z_1) = C_1 \ ,$$

respectively.

Define a function $g:R^4 \to R$ by

$$g(t_1,t_2,t_3,t_4) = (t_2 - 2t_3^2 + t_3^2 t_4)/t_1 .$$

The estimator $\hat{\theta}_W$ is then given by

$$\hat{\theta}_W = g(U^1,U^2,U^3,U^4) .$$

Now divide the random vectors $Z_1,\ldots,Z_N$ into n subgroups of size k $(N = nk)$ and let

$$\begin{aligned} \eta_1 &= E[k^{*1}(Z_1,Z_2)] = v_q , \\ \eta_2 &= E[k^{*2}(Z_1)] = v_p + I^2 , \\ \eta_3 &= E[k^{*3}(Z_1)] = I \end{aligned} \tag{4.5}$$

and

$$\eta_4 = E[k^{*4}(Z_1)] = 1 ,$$

where $I = \int f(x)p(x)dx$, as in (1.1) . (Note that $g(\eta_1,\eta_2,\eta_3,\eta_4) = v_p/v_q$.)

The estimator J_{W1} is then defined by

$$J_{W1} = J(\hat{\theta}_W) , \tag{4.6}$$

where $J(\hat{\theta}_W)$ is defined as in (2.2).

To satisfy the hypotheses of Theorems 2.6 and 2.7, we require the second moments of k^{*1} , k^{*2} , k^{*3}

and k^{*4} to be finite and $v_q > 0$. Having met these requirements, we then have

$$n^{\frac{1}{2}}(J_{W1} - v_p/v_q)/S_g \underset{D}{\rightarrow} N(0,1) ,$$

where S_g^2 is defined as in (2.9).

Approximate $100(1-\alpha)\%$ confidence intervals are then given by $(J_{W1}-z_{\alpha/2}S_g/n^{\frac{1}{2}}, J_{W1}+z_{\alpha/2}S_g/n^{\frac{1}{2}})$, where $z_{\alpha/2}$ is such that $\text{Prob}(V > z_{\alpha/2}) = \alpha/2$ and $V \sim N(0,1)$.

4.4 The Jackknifed $\hat{\theta}_W$ Estimator With Logs (J_{WIL})

The form of the estimator $\hat{\theta}_W$ suggests that the logarithmic transformation applied to $\hat{\theta}_W$ before jackknifing could prove advantageous. This transformation has proven its worth when applied to the variance components problem, as is indicated by the results of Arvesen and Schmitz [2]. In fact, the jackknife without the log transformation was found to behave poorly with respect to power.

To define J_{WIL}, let U^1, U^2, U^3 and U^4 be defined as in the previous section and let $g:R^4 \rightarrow R$ be defined by

$$g(t_1,t_2,t_3,t_4) = \log((t_2 - 2t_3^2 + t_3^2t_4)/t_1) .$$

Then

$$g(\eta_1,\eta_2,\eta_3,\eta_4) = \log(v_p/v_q) ,$$

where η_1, η_2, η_3 and η_4 are defined as in (4.5).

Now define

$$J_{W1L} = \exp(J(\hat{\theta})) , \tag{4.7}$$

where $\hat{\theta} = g(U^1,U^2,U^3,U^4)$ and $J(\hat{\theta})$ is as in (2.2).

The hypotheses of Theorems 2.6 and 2.7 will be satisfied if the second moments of the kernels of U^1, U^2, U^3 and U^4 are finite and $v_p,v_q > 0$. When these conditions are fulfilled, we have

$$n^{\frac{1}{2}}[J(\hat{\theta}) - \log(v_p/v_q)]/S_g \xrightarrow[D]{} N(0,1) , \tag{4.8}$$

where S_g^2 is as in (2.9), with $\hat{\theta} = g(U^1,U^2,U^3,U^4)$.

From (4.8) we may obtain approximate confidence intervals for $\log(v_p/v_q)$ based on the normal distribution. Then, by applying the exponential transformation to the confidence bounds, approximate confidence intervals for the parameter v_p/v_q may be obtained.

4.5 The Sukhatme Estimator (T_{S1})

Recall that all the estimators that we have considered as candidates for the jackknife method have been of the form $\hat{\theta} = g(\underset{\sim}{U})$ with $\theta = g(\underset{\sim}{\eta})$, where $\underset{\sim}{U} = (U^1,\ldots,U^s)$ is a vector of U-statistics and $\underset{\sim}{\eta} = (E(U^1),\ldots,E(U^s))$. When $\hat{\theta}$ is of this form the convergence theorems of Chapter 2 establish the asymptotic normality of $N^{1/2}[J(\hat{\theta}) - \theta]$. The asymptotic normality of the jackknife estimator, however, should not be surprising since the statistic $N^{1/2}g(\underset{\sim}{U})$ is itself asymptotically normally distributed. In fact, Theorem 7.1 of Hoeffding [9] establishes the convergence in distribution of $N^{1/2}(\underset{\sim}{U} - \underset{\sim}{\eta})$ to the multivariate normal distribution with mean $\underset{\sim}{0}$ and covariance matrix $(m_i m_j \zeta_1^{i,j})_{s\times s}$ and, from Theorem 2.9 of Chapter 2, we obtain the convergence in distribution of $N^{1/2}[g(\underset{\sim}{U})-g(\underset{\sim}{\eta})]$ to the univariate normal distribution with mean 0 and variance

$$\sigma^2 = \sum_{i=1}^{s} \sum_{j=1}^{s} g_i g_j m_i m_j \zeta_1^{i,j} , \qquad (4.9)$$

as in Theorem 2.7. Thus, $J(g(\underset{\sim}{U}))$ and $g(\underset{\sim}{U})$ are asymptotically equivalent.

In the following, we derive an estimate S_T^2 of the variance of the Sukhatme estimator. This variance estimate will then be used in conjunction with Sukhatmes' estimator to obtain approximate confidence intervals for v_p/v_q , based on the normal distribution.

Since we shall use the same symbol when referring to both the point and the interval estimate of v_p/v_q , we denote the Sukhatme estimator by

$$T_{S1} = \hat{\theta}_s \ . \tag{4.10}$$

This notation will remind the reader that T_{S1} is the one-sample Sukhatme estimator with associated variance estimator S_T^2 .

To obtain the estimate S_T^2 of $\mathrm{Var}(N^{\frac{1}{2}}T_{S1})$, we first note that, from Hoeffding [9], we have

$$\lim_{N\to\infty} N\mathrm{Cov}(U^i, U^j) = m_i m_j \zeta_1^{i,j} ,$$

where m_i, m_j and $\zeta_1^{i,j}$ are defined as in Theorem 2.7. Therefore, σ^2 , as in (4.9), may also be written as

$$\sigma^2 = \lim_{N\to\infty} N\mathrm{Var}[\sum_{i=1}^{s} g_i(U^i - \eta_i)] ,$$

where g_i , $i = 1,\ldots,s$, are the first partial derivatives of g evaluated at $\underset{\sim}{\eta}$. This suggests that we may estimate σ^2 by obtaining a consistent estimator of the variance of the first order terms of the Taylor series expansion of $g(\underset{\sim}{U})$ about $\underset{\sim}{U} = \underset{\sim}{\eta}$.

Now, let U^1 and U^2 be defined as in (4.2) and let $Y_i = F_i^*$, $i = 1,\ldots,N$, and $W_i = F_i F_i^* - F_i^{*2}$, $i = 1,\ldots,N$, so that $U^1 = S_Y^2$, $U^2 = \bar{W}$ and $T_{S1} = 1 + \bar{W}/S_Y^2$. Noting that $\mathrm{Var}(T_{S1}) = \mathrm{Var}(\bar{W}/S_Y^2)$, we proceed as follows.

Expand $\bar{W}/S_Y^2$ about $E(S_Y^2) = \eta_1$ and $E(\bar{W}) = \eta_2$, to obtain

$$\bar{W}/S_y^2 \doteq \eta_2/\eta_1 + (\bar{W} - \eta_2)/\eta_1 - \eta_2(S_y^2 - \eta_1)/\eta_1^2$$

and

$$\lim_{N\to\infty} N\mathrm{Var}(\bar{W}/S_y^2) = \lim_{N\to\infty} N\mathrm{Var}(\bar{W})/\eta_1^2 + \lim_{N\to\infty} N\eta_2^2\mathrm{Var}(S_y^2)/\eta_1^4 \quad (4.11)$$

$$- 2\lim_{N\to\infty} N\eta_2\mathrm{Cov}(\bar{W},S_y^2)/\eta_1^3 .$$

But

$$\lim_{N\to\infty} N\mathrm{Var}(\bar{W})/\eta_1^2 = \mathrm{Var}(W_1)/\eta_1^2 , \quad (4.12)$$

$$\lim_{N\to\infty} N\eta_2^2\mathrm{Var}(S_y^2)/\eta_1^4 = \eta_2^2[\mu_4(Y_1) - \eta_1^2]/\eta_1^4 \quad (4.13)$$

and

$$\lim_{N\to\infty} N\eta_2\mathrm{Cov}(\bar{W},S_y^2)/\eta_1^3 = \eta_2[\mathrm{Cov}(W_1,Y_1^2) - 2I\mathrm{Cov}(W_1,Y_1)]/\eta_1^3 . \quad (4.14)$$

We now define S_T^2 by substituting (4.12), (4.13) and (4.14) in (4.11) and replacing all unknown parameters by their sample estimates; i.e., define

$$S_T^2 = \{S_W^2 + \bar{W}^2[M_4(Y) - S_y^4]/S_y^4 - 2\bar{W}(S_{wy^2} - 2\bar{Y}S_{wy})/S_y^2\}/S_y^4 ,$$

where

$$S_W^2 = \sum_{i=1}^{N} (W_i - \bar{W})^2/(N - 1) ,$$

$$M_4(Y) = \sum_{i=1}^{N} (Y_i - \bar{Y})^4/(N - 1) , \quad (4.15)$$

$$S_{wy} = \sum_{i=1}^{N} (W_i - \bar{W})(Y_i - \bar{Y})/(N - 1)$$

and

$$S_{wy^2} = \sum_{i=1}^{N} (W_i - \bar{W})(Y_i^2 - \sum_{i=1}^{N} Y_i^2/N)/(N - 1) .$$

Since all the sample estimates in (4.15) are consistent (assuming that the second moments of U^1 and U^2 exist), we have

$$S_T^2 \underset{p}{\rightarrow} \sigma_T^2 ,$$

where

$$\sigma_T^2 = \text{Var}(W_1)/\eta_1^2 + \eta_2^2[\mu_4(Y) - \eta_1^2]/\eta_1^4$$
$$- 2\eta_2[\text{Cov}(W_1,Y_1^2) - 2I\text{Cov}(W_1,Y_1)]/\eta_1^3 .$$

It is easily verified that $\sigma_T^2 = \sigma^2$, where σ^2 is defined as in (4.9) with $g(t_1,t_2) = t_2/t_1$.

Since $N^{\frac{1}{2}}T_{S1}$ is asymptotically normally distributed with mean v_p/v_q and variance σ_T^2 , we have

$$N^{\frac{1}{2}}(T_{S1} - v_p/v_q)/S_T \underset{D}{\rightarrow} N(0,1) .$$

This result justifies the use of the normal approximation to obtain approximate confidence intervals for v_p/v_q . That is, if we define

$$C_l = T_{S1} - z_{\alpha/2} S_T / N^{\frac{1}{2}}$$

and

$$C_u = T_{S1} + z_{\alpha/2} S_T / N^{\frac{1}{2}} ,$$

where $z_{\alpha/2}$ is such that $\text{Prob}(V > z_{\alpha/2}) = \alpha/2$ and $V \sim N(0,1)$, then (C_l, C_u) is an approximate $100(1 - \alpha)\%$ confidence interval for v_p/v_q.

The performance of the T_{S1} estimator and its associated interval estimator is compared with that of the other one-sample estimators in Chapter 6.

CHAPTER 5

THE MONTE CARLO STUDY

5.1 Introduction

The Monte Carlo study was run on an IBM 370 computer, Model 145, and the random number generator was taken from IBM's scientific subroutine package.

The density functions, $p(x)$ and $q(x)$, were chosen to belong to either the family of Pareto distributions or to the family of Gamma distributions. Details of the sampling procedures used to obtain samples from the Pareto distribution and the Gamma distribution are given in Sections 5.2 and 5.3, respectively. Also found in these sections are the derivations of the expressions used to calculate the moments of $f(X)$ and $f^*(Y)$, where X has density $p(x)$ and Y has density $q(x)$. The measures of skewness and kurtosis (γ_1 and γ_2) of $f(X)$ and $f^*(Y)$ are easily obtained from these expressions. In Section 5.4 we discuss the rationale behind the choice of the functions $f(x)$, $p(x)$ and $q(x)$.

A total of 18 Monte Carlo sampling experiments were conducted. For each of these experiments and for each estimator under consideration, 400 independent observations (the Monte Carlo samples) were obtained and their 400 associated confidence intervals were constructed.

The comparisons of Chapter 6 are based on these observed values of the point and interval estimators.

For the two-sample estimators, $M = 400$ (the Monte Carlo sample size) independent samples of size $n = 30$ were generated from $p(x)$ and $q(x)$. The two-sample estimators, as well as their approximate 90% confidence intervals, were then calculated for each pair of samples. The one-sample estimators and their approximate 90% confidence intervals were obtained for each of $M = 400$ samples of size $n = 30$ generated from $q(x)$. Summary statistics, such as the mean of the 400 sample values, the estimated coefficient of variation of the sample mean, the mean confidence interval length and the coefficient of variation of the mean confidence interval length, were then obtained for each of the variance ratio estimators.

If $\hat{\theta}$ represents an arbitrary estimator of v_p/v_q, the $M = 400$ observations, $(\hat{\theta}_1, \ldots, \hat{\theta}_M)$ say, constitute a random sample from the distribution of $\hat{\theta}$. Therefore, if the true ratio v_p/v_q were known, we could then obtain estimates of the bias and its contribution to the mean square error, based on these sample values. In practice, however, we found that, in the majority of cases, the coefficient of variation of the sample mean

$\bar{\hat{\theta}}$ was too large, resulting in unreliable estimates of these quantities. An obvious solution to this problem would be to increase the Monte Carlo sample size, but, to reduce the coefficient of variation of $\bar{\hat{\theta}}$ to an acceptable level, the Monte Carlo sample size would have had to been increased considerably, making the survey prohibitively expensive. For example, since the coefficient of variation decreases as $1/M^{\frac{1}{2}}$, a Monte Carlo sample size of 1600 would have been required to decrease the coefficient of variation by a factor of ½.

To study the behavior of the interval estimates, coverage percentages were calculated for various multiples of the true variance ratio, $R = v_p/v_q$. That is, for each estimator, we calculated the fraction of the $M = 400$ confidence intervals that contained $c \times R$, where twenty values for c were chosen ranging from 0 to 100. Note that when $c = 1$, the coverage of $c \times R$ is an estimate of the true confidence level.

The results of the 18 sampling experiments are summarized in Tables I through XVIII. The main body of the tables records the coverages of $c \times R$, which, for notational convenience, are represented as being out of a total possible of 1000. The first column to the right of the coverages records the Monte Carlo sample mean of the estimators, divided by $R = v_p/v_q$; the second column records the estimated coefficient of variation of the statistic in column 1; the third column records the average length of the confidence intervals, divided by R ; and the fourth column records the estimated coefficient of variation of the statistic in column 3. Each table also indicates the true variance ratio R ; the function $f(x)$; the measures γ_1 and γ_2 of skewness and kurtosis of $f(X)$ and $f^*(Y)$; the density functions $p(x)$ and $q(x)$; and the position of $R = 1$, which is indicated by an arrow. The function $f^*(y)$ can easily be obtained from the tables as well, since $f^*(y) = f(y)p(y)/q(y)$.

5.2 The Pareto Distribution

Recorded in Tables I to VIII are the results of the Monte Carlo experiments when $p(x)$ and $q(x)$ are both density functions belonging to the Pareto family of distributions. The function $f(x)$ is of the form $f(x) = cx^k$ in each case.

In the following, we indicate how samples are obtained from the Pareto distribution and derive expressions for the skewness and the kurtosis of $f(X)$ and $f^*(Y)$, where $X \sim \text{Pareto}(a)$ and $Y \sim \text{Pareto}(b)$. An expression for the actual value of the ratio v_p/v_q is also obtained.

The density function $h(x)$ of the Pareto distribution with parameter a $(a > 0)$ is given by

$$h(x) = \begin{cases} a/x^{a+1} & \text{if } 1<x<\infty , \\ 0 & \text{otherwise.} \end{cases} \tag{5.1}$$

The cumulative distribution function $F_x(c)$ is then

$$F_x(c) = \begin{cases} 0 & \text{if } c < 1 , \\ 1 - 1/c^a & \text{if } c \geq 1 . \end{cases}$$

To obtain a sample from $h(x)$, defined as in (5.1), we use the random number generator to generate N independent Uniform(0,1) observations $u_1,\ldots,u_N$. Then $x_1,\ldots,x_N$, defined by

$$x_i = 1/(1 - u_i)^{1/a} \quad , \quad i = 1,\ldots,N ,$$

constitute a sample from the Pareto distribution with parameter a. This result is an immediate consequence of Theorem 8A of Parzen [16, p.314] (the Probability Integral Transformation), since

$$F_x^{-1}(u) = 1/(1 - u)^{1/a} .$$

Let $p(x)$ and $q(x)$ be the density functions of the Pareto distribution with parameters a and b, respectively. Define the function $f: R \to R$ by

$$f(x) = cx^k , \tag{5.2}$$

where c and k are real-valued constants. Let X and Y be random variables with density functions $p(x)$ and $q(x)$, respectively. Then if $a - rk > 0$, the r^{th} moment of $f(X)$ is given by

$$\begin{aligned} \mu_r'(f) &= E_p(c^r X^{rk}) \\ &= \int_1^\infty ac^r x^{rk}/x^{a+1} dx \\ &= ac^r/(a - rk) . \end{aligned} \tag{5.3}$$

Therefore, the central moments of $f(X)$, denoted by $\mu_r(f)$, are

$$\begin{aligned} \mu_r(f) &= c^r E_p\{[X^k + a/(k - a)]^r\} \\ &= c^r E_p\{\sum_{i=0}^{r} \binom{r}{i} [a/(k - a)]^{r-i} x^{ik}\} \\ &= c^r \sum_{i=0}^{r} \binom{r}{i} [a/(k - a)]^{r-i} a/(a - ik) . \end{aligned} \tag{5.4}$$

The value of the integral I is given by

$$\begin{aligned} I &= E_p(f) \\ &= ac/(a - k) , \end{aligned}$$

and the variance v_p is

$$v_p = \mu_2(f)$$
$$= c^2[a/(a - 2k) - a^2/(a - k)^2]$$
$$= ac^2k^2/(a - 2k)(a - k)^2 .$$

With $f(x)$ as in (5.2), the function $f^*(y)$, defined by $f^*(y) = f(y)p(y)/q(y)$, becomes

$$f^*(y) = cy^k ay^{b+1}/by^{a+1}$$
$$= acy^{k+b-a}/b .$$

Now, if $a - k > (r - 1)/rb$, the raw and central moments of $f^*(Y)$ are obtained by replacing k by $k + b - a$, c by ac/b and a by b in (5.3) and (5.4). Thus

$$\mu'_r(f^*) = a^r c^r/b^{r-1}[r(a - k) - b(r - 1)]$$

and (5.5)

$$\mu_r(f^*) = (ac/b)^r \sum_{i=0}^{r} \binom{r}{i} [b/(k - a)]^{r-i} b/[b - i(k+b-a)] ,$$

where $\mu'_r(f^*)$ and $\mu_r(f^*)$ are the r^{th} raw and central moments, respectively, of $f^*(Y)$.

The variance ratio $R = v_p/v_q$ is given by

$$R = \mu_2(f)/\mu_2(f^*) ,$$

which, upon simplifying, yields

$$R = k^2 b[2(a-k)-b]/a(a-2k)\{(a-k)^2-b[2(a-k)-b]\} .$$

The measures of skewness, $\gamma_1(f)$ and $\gamma_1(f^*)$ of $f(X)$ and $f^*(Y)$, respectively, and the measures of kurtosis, $\gamma_2(f)$ and $\gamma_2(f^*)$ of $f(X)$ and $f^*(Y)$, respectively, are easily obtainable from (5.4) and (5.5), where $\gamma_i(f)$ and $\gamma_i(f^*)$, $i = 1,2$, are defined in Section 3.1.

5.3 The Gamma Distribution

The results of the Monte Carlo experiments when $p(x)$ and $q(x)$ are density functions belonging to the family of Gamma distributions are recorded in Tables IX through XVIII. For all but one of the cases, the function $f(x)$ is of the form cx^k. The exceptional case is $f(x) = x^4\exp(x/9)/60$.

The density function of the Gamma distribution with parameters a and b $(a,b > 0)$ is given by

$$g(x) = \begin{cases} x^{a-1}\exp(-x/b)/b^a\Gamma(a) & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{5.6}$$

We denote by Gamma(a,b) the Gamma distribution with parameters a and b.

For the cases we studied, the parameter a in (5.6) was chosen to be an integer. This choice of a simplifies the sampling procedure and saves computer time.

To obtain a sample from Gamma(a,b) when a is integer-valued, we proceed as follows. Generate Na independent observations from the Uniform(0,1) distribution using the random number generator, where N is the desired sample size. Divide the Na observations into N groups of a observations each and define

$$A_i = -b\times\log(U_{1i}\times\ldots\times U_{ai}) \quad , \quad i = 1,\ldots,N \ ,$$

where $U_{1i},\ldots,U_{ai}$ is the i^{th} group of Uniform(0,1) values and log is the natural logarithmic transformation. Then $A_1,\ldots,A_N$ constitutes a sample from Gamma(a,b). The following claim will establish the truth of the above.

Claim. Let $U_1,\ldots,U_a$ be independent Uniform(0,1) random variables and let $b > 0$ be a real-valued constant. Then the random variable A defined by

$$A = -b\times\log(U_1\times\ldots\times U_a) \ ,$$

has a Gamma(a,b) distribution.

<u>Proof</u>. To prove the claim we use the moment generating function technique. The moment generating function of A is

$$M_A(t) = E[\exp(At)]$$
$$= \int_0^1 (U_1 \times \ldots \times U_a)^{-bt} dU_1 \ldots dU_a \text{ , by independence,}$$
$$= (1 - bt)^{-a} \quad , \quad \text{for} \quad t < 1/b \ .$$

But this is the moment generating function of $\text{Gamma}(a,b)$. The claim follows immediately from the uniqueness of the moment generating function. □

In the following, we derive expressions for the moments of $f(X)$ and $f^*(Y)$, where $X \sim \text{Gamma}(a_1,b_1)$ and $Y \sim \text{Gamma}(a_2,b_2)$. The functions $f(x)$, $p(x)$ and $q(x)$ are defined to include all cases entering the Monte Carlo study.

Let $p(x)$ be the density of the Gamma distribution with parameters a_1 and b_1 $(\text{Gamma}(a_1,b_1))$, let $q(x)$ be the density function of the Gamma distribution with parameters a_2 and b_2 $(\text{Gamma}(a_2,b_2))$ and let

$$f(x) = cx^k \exp(dx) \ ,$$

where c , d and k are real-valued constants. Then $f^*(x) = f(x)p(x)/q(x)$ is given by

$$f^*(x) = Cx^a \exp(bx) ,$$

where

$$C = cb_2^{a_2}\Gamma(a_2)/b_1^{a_1}\Gamma(a_1) ,$$

$$a = k + a_1 - a_2$$

and

$$b = (b_1b_2d - b_2 + b_1)/b_1b_2 .$$

Let X and Y be random variables with densities $p(x)$ and $q(x)$, respectively. The r^{th} raw moment of $f^*(Y)$ is then given by

$$\begin{aligned}\mu_r'(f^*) &= E\{[f^*(Y)]^r\} \\ &= C^r E_q[X^{ar}\exp(rbX)] \\ &= C^r \int_0^\infty x^{ar+a_2-1} \exp(-x(1/b_2 - br))dx ,\end{aligned}$$

which, for $ar + a_2 > 0$ and $1 - rbb_2 > 0$, yields

$$\mu_r'(f^*) = [C^r/b_2^{a_2}\Gamma(a_2)]\Gamma(ar + a_2)[b_2/(1 - rbb_2)]^{(ar+a_2)} . \quad (5.7)$$

The r^{th} raw moment of $f(X)$ is obtained in an analogous manner and is given by

$$\mu_r'(f) = [c^r/b_1^{a_1}\Gamma(a_1)]\Gamma(kr + a_1)[b_1/(1 - rdb_1)]^{kr+a_1} , \quad (5.8)$$

providing $rk + a_1 > 0$ and $rdb_1 < 1$.

The central moments of $f(X)$ and $f^*(Y)$ and their measures of skewness and kurtosis can easily be obtained from (5.7) and (5.8), as for the Pareto distributions.

5.4 Selection of the Sampling Experiments

The main objective of the Monte Carlo study is to examine the behavior of the variance ratio estimators under varying degrees of non-normality of the random variables $f(X)$ and $f^*(Y)$, where X has density $p(x)$ and Y has density $q(x)$. In practice, the random variables $f(X)$ and $f^*(Y)$ may differ radically in the degree of non-normality that each displays. This situation is of particular interest since the usual methods of variance comparison in the two-sample case requires the assumption of identical distributions of $f(X)$ and $f^*(Y)$, except for location and scale. We are also interested in studying the effect of extreme non-normality of $f^*(Y)$ upon the one-sample estimators. Recall that Hastings [8] warns that importance sampling may result in extreme non-normality of $f^*(Y)$, even when $q(x)$ is arbitrarily close to the optimum $q_o(x)$.

Since the Monte Carlo study is to serve as an exploratory investigation involving a large number of estimators, we have had to limit the number of cases studied. For this reason, and because in practice it is likely that a knowledgeable choice of $q(x)$ will lead to a variance reduction, we have chosen $f(x)$, $p(x)$ and $q(x)$ to yield $R \geq 1$ in the majority of cases - in fact, $R \geq 1$ for all but one of the sampling experiments.

The Pareto and Gamma distributions were selected for several reasons, not the least important of which is that the sampling procedures are easily programmed on a computer. These distributions also allow the use of simple functional forms, such as $f(x) = x^k$, yet still provide a large class of distributions of $f(X)$ and $f^*(Y)$. Furthermore, as we have seen in Sections 5.2 and 5.3, the raw and central moments of $f(X)$ and $f^*(Y)$ are obtainable by straightforward calculations and, from the expressions for the central moments of $f(X)$ and $f^*(Y)$, we may then calculate the measures of skewness, $\gamma_1 = \mu_3/\mu_2^{3/2}$, and kurtosis, $\gamma_2 = \mu_4/\mu_2^2 - 3$. These parameters, γ_1 and γ_2 , may then be used to indicate the degree of non-normality displayed by $f(X)$ and $f^*(Y)$. Recall that for the normal distribution $\gamma_1 = \gamma_2 = 0$.

Although the calculations of $\gamma_1(f)$, $\gamma_2(f)$, $\gamma_1(f^*)$ and $\gamma_2(f^*)$, defined as in (3.1), are routine computations once the functions $f(x)$, $p(x)$ and $q(x)$ have been specified, it is difficult to obtain specific values of γ_1 and γ_2 for $f(X)$ and $f^*(Y)$. For example, suppose $p(x)$ and $q(x)$ are density functions of the Pareto distribution with parameters a and b, respectively, and suppose $f(x) = x^k$. Then, to guarantee convergence of the variance ratio estimators, we require $E\{[f(X)]^4\} < \infty$ and $E\{[f^*(Y)]^4\} < \infty$. If, in addition, we require $R \geq 1$, then the parameters a, b and k are subject to

$$a - 4k > 0 ,$$

$$a - k > 3/4b$$

and

$$k^2 b[2(a-k)-b]/a(a-2k)\{(a-k)^2-b[2(a-k)-b]\} \geq 1 .$$

With these restrictions on the parameters a, b and k, it is apparent that it may, in fact, be impossible to obtain predetermined values of $\gamma_1(f)$, $\gamma_2(f)$, $\gamma_1(f^*)$ and $\gamma_2(f^*)$.

The method employed to select $f(x)$, $p(x)$ and $q(x)$ was essentially that of trial and error. That is, the values of R , $\gamma_1(f)$, $\gamma_2(f)$, $\gamma_1(f^*)$ and $\gamma_2(f^*)$ were calculated for a large number of functions $f(x)$, $p(x)$ and $q(x)$ and the cases to be studied were selected from these. Table XXIX records the values of R , $\gamma_1(f)$, $\gamma_2(f)$, $\gamma_1(f^*)$ and $\gamma_2(f^*)$ for each of the Monte Carlo sampling experiments. To aid in the discussions of the results of the sampling experiments, a distributional classification scheme is introduced in Section 6.2. The column on the right of Table XXIX records the distributional classification for each of the sampling experiments. If further details of the distributions of $f(X)$ or $f^*(Y)$ are required, one may refer to the original tables, i.e. to Tables I through XVIII.

CHAPTER 6

THE MONTE CARLO RESULTS

6.1 Introduction

In this chapter, we describe the results of the Monte Carlo sampling experiments and discuss the implications of these results with regard to the importance sampling problem. It is important that the performance of the estimators be analyzed in the context of the estimation problem under consideration. Recall that we wish to estimate the 'relative efficiency' (v_p/v_q) of two competing techniques of estimating a population mean (or total). Based on an estimate of the relative efficiency, we hope to determine which of the estimation procedures is preferable. Obviously the worst possible error is to choose the least efficient method when this procedure leads to an estimate of the population mean which displays extreme non-normality. A somewhat less serious error is one which leads to the use of the less efficient technique but results in an estimator of the mean which is distributed more closely to a normal random variable than the alternative statistic. In fact, in some instances, unless a substantial gain in efficiency is obtained, it may advantageous to choose the less efficient method.

A detailed description of each of the two-sample and one-sample estimation techniques was given in Chapters 3 and 4, respectively. The class of two-sample procedures consists of (i) the classical F method (F) ; (ii) the modified Box-Andersen method (F_{BA}) , which involves adjusting the degrees of freedom of the usual F statistic; (iii) the Box methods (B_5 and B_{10}), which require dividing the samples into disjoint subgroups, calculating $\log(S^2)$ for each subgroup and comparing these two sets of values by the t-test for location; (iv) Arvesen's two-sample jackknife procedure (J_{A2}) applied to the ratio of the sample variances; (v) Arvesen's two-sample jackknife (J_{A2L}) applied to the natural log of the ratio of the sample variances; and (vi) Miller's jackknife method (J_{M2L}) , which involves jackknifing the natural log of each of the sample variances separately and then combining the results.

The one-sample methods include (i) the method based on the asymptotic normality of the Sukhatme statistic and a consistent estimator of its variance (T_{S1}) ; (ii) the jackknife method applied to Sukhatmes' estimator (J_{S1}) ; (iii) the jackknife procedure applied to the non-negative estimator (J_{W1}) ; and (iv) the jackknife version of the natural log of the non-negative estimator (J_{W1L}).

Due to the difference in nature of the one-sample and the two-sample procedures, we have discussed the results pertaining to these two classes of estimators separately. A discussion of the results of the sampling experiments relating to the two-sample estimators may be found in Sections 6.5 and 6.6 while those for the one-sample estimators are discussed in Sections 6.3 and 6.4. For the reader interested only in a summary of the findings, Sections 6.3.2, 6.3.3, 6.5.2 and 6.5.3, which contain a detailed examination of the Monte Carlo results, may be omitted.

Since a large number of estimators are studied, we have had to limit the depth of the survey. The Monte Carlo sampling experiments, however, serve as a preliminary investigation from which we are able to eliminate the obviously poor estimation procedures and suggest directions of further study regarding the more promising techniques. Some of the obvious limitations of the survey are those imposed by the small Monte Carlo sample size ($M = 400$), the limited number of distributions sampled, the fixed sample sizes ($n = 30$) and the single confidence level of 90%.

6.2 Preliminary Comments

6.2.1 Discussion of the Tables

The results of the 18 Monte Carlo sampling experiments are recorded in Tables I through XVIII. Each of these tables, which is described in Chapter 5, contains the results for both the one-sample and the two-sample estimators for specific functions $f(x)$, $p(x)$ and $q(x)$. That is, each table summarizes the results of a particular sampling experiment. For convenience we identify the sampling experiments by the number of their associated table.

To study the effect of non-normality of $f(X)$ and $f^*(Y)$, we have roughly classified their distributions as displaying low, moderate or high degrees of non-normality, denoted by L, M and H, respectively. With this classification scheme we can identify, for each sampling experiment, the distributional properties of $f(X)$ and $f^*(Y)$ by a pair of letters, the first denoting the distributional classification of $f(X)$ and the second denoting the distributional classification of $f^*(Y)$. For example, L - M indicates that $f(X)$ displays low non-normality and $f^*(Y)$ displays moderate non-normality. When referring to the distributional classification of a single random variable,

$f^*(Y)$ say, we shall use the notation $f^* \varepsilon L$, $f^* \varepsilon M$ or $f^* \varepsilon H$, with a similar notation for $f(X)$. Table XXIX lists the Monte Carlo sampling experiments, their classification and the values of $\gamma_1(f)$, $\gamma_2(f)$, $\gamma_1(f^*)$, $\gamma_2(f^*)$ and the variance ratio, $R = v_p/v_q$.

To aid in assessing the behavior of the one-sample estimators, we have reproduced the relevant portions of Tables I through XVIII and arranged them according to the distributional properties of $f^*(Y)$ in Tables XIX through XXII. An attempt has been made to arrange the results in order of increasing non-normality of $f^*(Y)$, beginning with the first entry in Table XIX and ending with the last entry in Table XXII. It must be emphasized, however, that this ordering is approximate, being based solely on the values of $\gamma_1(f^*)$ and $\gamma_2(f^*)$. The table from which each of the reproduced portions was taken is indicated by a Roman numeral and, as for the complete tables, we have indicated with an arrow the position of $v_p/v_q = 1$, thus permitting the approximate power of the test $H_o: v_p = v_q$ vs. $H_a: v_p \neq v_q$ to be readily obtained.

For the two-sample estimators we have also reproduced the relevant portions of Tables I to XVIII. These results, which are recorded in Tables XXIII to

XXVIII, have been grouped according to the distributional classifications of both $f(X)$ and $f^*(Y)$. For example, Table XXIII records the results for the L-L classification and Table XXVII contains the results for the L-H classification. As for the one-sample estimators, we have indicated, by a Roman numeral, the original table from which each of the reproduced portions was taken and have also indicated the position of $v_p/v_q = 1$. If additional information regarding the distributional properties of $f(X)$ and $f^*(Y)$ is required, one may refer to either the original table or to Table XXIX, which provides a summary of this information for all distributions sampled.

Tables XXX, XXXI and XXXII record the coverages of R, $R/5$ and $5 \times R$, respectively, for both the one-sample and the two-sample estimators and for all 18 Monte Carlo sampling experiments. Table XXX is particularly useful for comparing the estimated confidence levels. In addition to these tables, a table of maximum coverages has also been constructed. This table (Table XXXIII) records the maximum coverage and the point at which the maximum occurs for each estimator and for all 18 sampling experiments.

6.2.2 Precision of the Estimates of the Probabilities of Containment

The coverages of $c \times R$ are merely estimates of the probability that the point $c \times R$ is contained in an arbitrary confidence interval constructed by the estimator under consideration. Thus, the coverage of R is an estimate of the true confidence level. Obviously a statistical procedure is required to compare coverages obtained by the various estimation procedures.

Consider the problem of estimating the probability of containment, p_o say, of a point x_o. If 400 confidence intervals are constructed independently, the total number of confidence intervals containing x_o has a Binomial distribution with parameters $N = 400$ and $p = p_o$. For our purposes, since $N = 400$ is large, the normal approximation to the Binomial distribution is adequate for constructing approximate confidence intervals and tests for p_o. That is, we assume that the statistic $(\hat{p}_o - p_o)/[\hat{p}_o(1-\hat{p}_o)/N]^{\frac{1}{2}}$ has a normal distribution with mean 0 and variance 1, where $\hat{p}_o$ is the fraction of the confidence intervals containing x_o, i.e. the coverage of x_o. For information regarding approximations to the Binomial distribution we refer to Johnson and Kotz [10, pp.61-67].

For each of the Monte Carlo sampling experiments, the majority of the estimators were based on the same sample. That is, $M = 400$ samples were obtained from $p(x)$ and $q(x)$ and for each pair of samples, $(y_1,\ldots,y_n)$ from $q(x)$ and $(x_1,\ldots,x_n)$ from $p(x)$, the one-sample estimators were based on $(y_1,\ldots,y_n)$ and the two-sample estimators were based on both $(x_1,\ldots,x_n)$ and $(y_1,\ldots,y_n)$. As a result, one would expect the coverages of $c \times R$, for the various estimators, to be positively correlated. Furthermore, since the variance of the difference of two positively correlated estimators is less than the sum of their variances, one would expect tests and confidence intervals for the difference to be conservative if they were developed under the assumption of independence.

Consider the problem of comparing the coverages of a point x_o for two competing estimators. Let $\hat{p}_1$ be the coverage of x_o for the first estimator (i.e. the fraction of the 400 confidence intervals which contain x_o) and let $\hat{p}_2$ be the coverage of x_o for the second estimator. Then $E(\hat{p}_1) = p_1$ and $E(\hat{p}_2) = p_2$, where p_1 and p_2 are the probabilities of containment of x_o for the first and second estimators, respectively. To obtain approximate confidence intervals for the difference, $p_1 - p_2$, using the normal approximation and assuming that $\hat{p}_1$ and $\hat{p}_2$ are uncorrelated, we treat

$$[(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)]/[\widehat{\text{var}}(\hat{p}_1 - \hat{p}_2)]^{\frac{1}{2}}$$

as a normal random variable with mean 0 and variance 1, where

$$\widehat{\text{var}}(\hat{p}_1 - \hat{p}_2) = [\hat{p}_1(1 - \hat{p}_1) + \hat{p}_2(1 - \hat{p}_2)]/N .$$

Therefore, the approximate confidence bounds obtained by this method are

$$\hat{p}_1 - \hat{p}_2 \pm z_o[\widehat{\text{var}}(\hat{p}_1 - \hat{p}_2)]^{\frac{1}{2}} ,$$

where the appropriate value of z_o is obtained from tables of the normal distribution.

Note that $\widehat{\text{var}}(\hat{p}_1 - \hat{p}_2)$ achieves its maximum value of $1/2N$ when $\hat{p}_1 = \hat{p}_2 = \frac{1}{2}$. Thus, the maximum confidence interval length obtainable is $z_o(2/N)^{\frac{1}{2}}$, which, for a 90% confidence level is less than .115 and for a 95% confidence level is approximately .14 (N = 400) .

The maximum confidence interval length obtainable for a single parameter, p_1 say, is approximately .08 for a 90% confidence level and approximately .10 for a 95% confidence level.

The conservative confidence bounds, obtained by assuming that the maximum variance is achieved, are adequate for the majority of comparisons in the following discussions. In Sections 6.3 and 6.5, all comparisons of probabilities of containment and all tests regarding bias and mean confidence interval lengths are made at the $\alpha = .05$ level of significance. It will be apparent, from the context, whether a two-sided or a one-sided test has been used.

6.3 The One-sample Estimators

6.3.1 General Properties

Recall that the Sukhatme estimator $\hat{\theta}_s$ may yield negative estimates of the variance ratio v_p/v_q. This property of the Sukhatme estimator led to the proposal of the non-negative estimator $\hat{\theta}_w$.

The form of $\hat{\theta}_w$ (defined in 1.8) suggests that the log transformation, applied to this estimator before jackknifing, could prove beneficial. The results of the Monte Carlo sampling experiments, however, indicate that the jackknifed version of $\hat{\theta}_w$ without logs (J_{W1}) is superior to the jackknifed version with logs (J_{W1L}). In fact, the J_{W1} estimator proved generally superior to the other one-sample estimators, J_{S1} and T_{S1}, as well.

For each of the one-sample estimators, we now relate some of the general results of the Monte Carlo study.

(i) <u>The jackknifed Sukhatme estimator (J_{S1})</u>

In the majority of the cases run, the J_{S1} estimator was found to be highly variable. As a result, confidence intervals based on this statistic were generally long and uninformative. In fact, the coverage of zero was over 90% for 7 of the 18 sampling experiments. Although the J_{S1} interval estimates did not, in general, provide reasonable lower confidence bounds for R, its estimated confidence levels were usually close to the nominal 90% level. It is interesting to note that if the estimators were compared solely by their estimated confidence levels, the J_{S1} estimator would appear superior.

(ii) <u>The Sukhatme estimator (T_{S1})</u>

The Sukhatme estimator T_{S1} was the poorest of the one-sample estimators. It behaved similarly to J_{S1} with respect to coverages of $c \times R$ when $f^* \varepsilon L$; but was adversely affected by non-normality of the random variable $f^*(Y)$. Large values of $\gamma_2(f^*)$ were associated with low confidence levels and positive bias. The T_{S1} estimator was also inclined to yield long confidence intervals which provided poor lower confidence bounds.

(iii) <u>The jackknifed non-negative estimator (J_{W1})</u>

For the majority of the 18 sampling experiments, the J_{W1} estimator provided much shorter confidence intervals than did J_{S1}. These short confidence intervals, however, were often shifted slightly to the left of R, resulting in confidence levels somewhat lower than desired. The effect of the non-normality of $f^*(Y)$ was to increase the length of the confidence intervals, which was accompanied by a corresponding increase in the confidence levels. In fact, the confidence levels did not differ significantly from the nominal level for the majority of cases for which $f^* \in M$ or H. This estimator also provided rather conservative lower confidence bounds in many instances, although these conservative bounds resulted only when $f^*(Y)$ displayed relatively high non-normality.

(iv) <u>The jackknifed version of the natural log of the non-negative estimator (J_{W1L})</u>

Containment of zero is not a problem associated with the J_{W1L} interval estimator. Confidence levels of this estimator, however, were, in general, well below the nominal 90% level. Although the J_{W1L} estimator behaved similarly to J_{W1} when $f^* \in L$, large $\gamma_2(f^*)$

values were found to be accompanied by considerable positive bias.

6.3.2 Comparison of Coverages for Selected Values of c×R

(i) Estimated confidence levels (Table XXX)

Restricting attention to the one-sample estimators J_{W1L} , J_{W1} , J_{S1} and T_{S1} , we find, from Table XXX, that the estimated confidence levels for J_{S1} are, in general, much nearer the nominal level of 90% than those for the other one-sample estimators. The non-normality of $f^*(Y)$ has no apparent effect on the confidence levels for the J_{S1} estimator. With the exception of runs XIV (M-H) and XVIII (H-L) , the coverages of R for J_{S1} are between 85% and 95%.

Note the similarity of the coverages of R for the J_{S1} and T_{S1} estimators when $f^* \varepsilon L$. Non-normality of $f^*(Y)$, however, appears to have an adverse effect on the coverages of R for T_{S1} . When extreme non-normality of $f^*(Y)$ is present, the true confidence levels for T_{S1} are lower than those for J_{S1} ; the maximum coverage of R (77%) for this category is well below the nominal 90% level.

With the exception of XV, the confidence levels for J_{W1} are significantly lower than the nominal level when $f^* \varepsilon L$. The coverage of R , however, does appear to be improved when $f^* \varepsilon$ M or H. In fact, the coverages are similar to those for J_{S1} in these cases.

The confidence levels for the J_{W1L} interval estimate are too low in general. Although the coverages of R for J_{W1L} are similar to those for J_{W1} when $f^* \varepsilon$ L, we find that they are lower than those for J_{W1} when $f^* \varepsilon$ M or H.

(ii) Coverages of 5×R (Table XXXII)

Inspection of Table XXXII indicates that the coverages of 5×R are generally higher for J_{S1} than for either J_{W1L} or J_{W1}. The estimator T_{S1} behaves similarly to J_{S1} with respect to coverage of 5×R, although the coverages for T_{S1} are, in general, somewhat lower than those for J_{S1} when $f^* \varepsilon$ M or H. There is no striking difference in the coverages of 5×R for J_{W1} and J_{W1L}. Both these estimators result in low coverage of 5×R when $f^* \varepsilon$ L.

(iii) Coverages of R/5 (Table XXXI)

Of the 18 sampling experiments, 14 result in coverages of R/5 greater than 70% for the J_{S1} estimator. The coverages for T_{S1} are similar to those for J_{S1} when $f^* \varepsilon$ L, but are lower than those for J_{S1} when $f^* \varepsilon$ M or H. Despite this fact, over half of the sampling experiments result in coverages of R/5 greater than 70% for T_{S1} .

When $f^* \varepsilon L$, the coverages of this point (R/5) are, in general, substantially lower for J_{W1} than those for J_{S1} and T_{S1} . High coverages of R/5 for J_{W1} are associated with extreme non-normality of $f^*(Y)$. When $f^* \varepsilon L$, the coverages of R/5 are essentially zero for the J_{W1L} estimator and, as for the other estimators, the maximum coverages of R/5 occur when $f^* \varepsilon M$ or H . This estimator provides by far the lowest coverages of R/5 , in general.

(iv) Maximum coverages (Table XXXIII)

In Table XXXIII, the numbers in parentheses are the approximate points at which the maximum coverages occur, and the numbers recorded directly above these parenthesized entries are the coverages of these points, i.e. the maximum coverages. From the examination of the coverages of R/5 , R and 5×R , we have seen that in many cases there is little variation among the coverages of these points - this is especially true for the J_{S1} and T_{S1} estimators. Because of this fact, the points of maximum coverage are not well defined in some cases.

We have found that when $f^* \varepsilon L$, the coverages of R , R/5 and 5×R are, in general, lower for the

J_{W1} and J_{W1L} estimators than for the J_{S1} and T_{S1} estimators. Since the confidence levels for J_{W1} and J_{W1L} are well below the nominal level for the majority of sampling experiments of this category, one might suspect that the estimates of the standard errors required to construct the confidence intervals for these estimators are underestimating the actual standard errors involved. The generally high maximum coverages recorded for J_{W1} and J_{W1L} when $f^* \varepsilon L$, however, tend to indicate that the poor confidence levels are a result of bias, rather than a fault of the estimates of the standard errors. It is interesting to note that for the sampling experiments of this classification, although the maximum coverages for J_{W1} and J_{W1L} occur to the left of R, these estimators generally yield lower coverages of $R/5$ than do either J_{S1} or T_{S1}. Before leaving Table XXXIII, we draw attention to the low maximum coverages associated with extreme non-normality of $f^*(Y)$ for T_{S1} and J_{W1L}.

6.3.3 The Effect of Non-normality

In the following we compare the performances of the one-sample estimators under varying degrees of non-normality of the random variable $f^*(Y)$. Recall that Tables XIX, XX, XXI and XXII, which were created for this purpose, consist of the relevant portions of Tables I through XVIII, arranged according to the distributional properties of $f^*(Y)$.

The four columns immediately to the right of the main body of the tables record the values of four statistics calculated from the Monte Carlo sample values of the various estimators. The first column records the mean of these sample values, divided by R ; the second records the estimated coefficient of variation of the statistic in column 1; the third records the mean confidence interval length, divided by R ; and the fourth records the coefficient of variation of the statistic in column 3. The column on the far right records the sampling experiment and the distributional classification for each data set.

Since the discussion of the coverages of $c \times R$, for selected values of c , has revealed the T_{S1} interval estimate as being of little value, future discussions of this estimator shall be brief.

(i) <u>The case $f^* \varepsilon L$</u> (Tables XIX and XX)

Examining Tables XIX and XX, we find that the results of the sampling experiments I, II, IV, VII, VIII and XVI are similar in nature. In each of these cases, the estimators J_{S1} and T_{S1} yield confidence intervals much longer than do J_{W1} and J_{W1L}. These long confidence intervals appear to be caused by high variance of the Sukhatme estimators. The resulting conservative confidence bounds for R are uninformative. Consider, for example, run number IV for which 0 and $20 \times R$ belong to approximately 90% of the confidence intervals constructed by J_{S1}. In this case the mean confidence interval length is about 50 times as long as for the J_{W1} and J_{W1L} estimators.

The most striking difference between the coverages for the Sukhatme estimators, J_{S1} and T_{S1}, and those for J_{W1} and J_{W1L}, in the six sampling experiments listed in the previous paragraph, is the rate at which these coverages decrease as one moves in either direction from the point at which the maximums occur. The coverages for both J_{W1} and J_{W1L} drop rapidly as one moves away from the points of maximum coverage, while the coverages for the Sukhatme estimators decrease almost imperceptively in comparison.

Before discussing the results of the three remaining sampling experiments of this category, note the similarity of the coverages of $c \times R$ for J_{W1} and J_{W1L} when $f^* \varepsilon L$. J_{W1L}, however, does tend to yield coverages slightly lower than J_{W1} for small values of c.

The results of the remaining cases are discussed individually. All the estimators performed well in XV. The confidence intervals were short and the estimated confidence levels were within 4% of the nominal level for each of the estimators. No significant bias was indicated for any of the estimators, but the points of maximum coverage were slightly less than R.

The confidence intervals for J_{W1} and J_{W1L} are very short in III. This, together with a negative bias, results in poor coverage of R for these estimators. The Sukhatme estimators again provide fairly conservative lower confidence bounds for R.

For run number XVIII, all the estimators behave poorly. The confidence levels are well below the nominal level for each of the estimators in this case and the maximum coverages, which occur at $.7 \times R$, are also less than 90%. Perhaps the large value of R (=218) is responsible for these poor performances; further investigation is necessary to evaluate the effect of this parameter.

Figures 1 and 2 are normal probability plots of J_{W1} and $\log(J_{W1L})$, respectively, for run number XVIII. Although both graphs indicate that the distributions are skewed to the right, this is much less pronounced for $\log(J_{W1L})$. Recall that $J_{W1L} = \exp(J(\log(\hat{\theta}_w))$; thus $\log(J_{W1L})$ is just the jackknife estimator applied to $\log(\hat{\theta}_w)$. Therefore, in this case, the log transformation, used in conjunction with the jackknife, appears to have produced an estimator whose distribution more closely resembles that of a normal random variable.

(ii) <u>The case $f^* \varepsilon M$</u> (Table XXI)

For the sampling experiments of this category, $\gamma_2(f^*)$ ranges in value from approximately 71 for run number V to approximately 600 for run number XVII. Thus, within the moderate non-normality category, there is considerable variation in the degree of non-normality exhibited by $f^*(Y)$.

The jackknifed Sukhatme estimator J_{S1} provides conservative confidence bounds for all cases represented in Table XXI. Although the coverages of R are near the nominal level, we find that zero is contained over 90% of the time for all the sampling experiments of this category. The long confidence intervals for J_{S1} are indicative of high variance of this estimator.

The non-normality of the random variable $f^*(Y)$ seems to have caused an increase in the variance of the J_{W1} estimator. This increased variability, however, is compensated for by an increase in the length of the confidence intervals. As a result, the coverages of R for J_{W1} are close to the nominal level for each of the sampling experiments in this category. Note that although the confidence levels have improved, the coverages of zero have increased considerably. Recall that when $f^* \varepsilon L$, the coverages for J_{W1} were similar to those for J_{W1L}. This is obviously not the case for the sampling experiments of Table XXI. In fact, the J_{W1} estimator appears to behave more like J_{S1} than J_{W1L}, with respect to coverages of $c \times R$, when $f^* \varepsilon M$. With the exception of run number XIII, a significant negative bias is indicated for the J_{W1} estimator.

The J_{W1L} estimator performed quite well for V and IX, although the confidence levels were somewhat lower than desired. Despite the fact that the non-normality of $f^*(Y)$ appears to have increased the confidence interval length considerably, the maximum coverages appear, in general, to have decreased. This could be caused by the jackknife variance S_g^2 underestimating the variance of $J(\log(\hat{\theta}_w))$. The decrease in coverage was also accompanied by a significant positive bias.

(iii) <u>The case $f^* \varepsilon H$</u> (Table XXII)

For the sampling experiments of this final category, $\gamma_2(f^*)$ is of the order 4×10^4 and greater.

It would be unlikely for the estimators to perform well in these cases and, in general, this is true.

The jackknifed Sukhatme estimator J_{S1} behaves much as it did when $f^* \varepsilon M$. That is, its confidence intervals are generally too long and in some cases contain zero more often than they contain R. Estimated confidence levels for J_{S1}, however, are near the nominal level for all runs except run number XIV.

In the three most extreme cases, X, XI and XIV, J_{W1} and J_{S1} display almost identical behavior with respect to coverages of $c \times R$. In the other two cases, VI and XII, however, the J_{W1} estimator yields shorter confidence intervals than J_{S1} and, although their maximum coverages are comparable, J_{W1} results in lower coverages of zero than does J_{S1}. Again, the major disadvantage of these estimators lies in the tendency of their confidence intervals to contain zero.

The J_{W1L} estimator works suprisingly well for run number VI, for which the value of $\gamma_2(f^*)$ is 1.5×10^6. In this instance, the confidence level for the J_{W1L} estimator is approximately 81% and the maximum coverage of approximately 90% occurs at $.9 \times R$, despite the fact that a significant positive bias is indicated. The mean confidence interval length of $2.7 \times R$ is smaller than those for the other one-sample estimators, in this case.

The J_{W1L} estimator performs poorly in all other cases. Its confidence levels are typically too low (as low as 55%) and its confidence intervals are shifted to the right, indicating a tendency to overestimate R. The positive bias of J_{W1L}, which is indicated for all sampling experiments of this category, is highest for runs X and XI.

6.4 Discussion of the One-sample Estimators

Perhaps the most obvious disadvantage of the estimators J_{S1}, T_{S1} and J_{W1} lies in their tendency to provide negative lower confidence bounds. This problem, however, is not unique to the importance sampling problem. Suppose, for example, that the mean μ of a normal random variable X is to be estimated, where the variance σ^2 is known and $\mu > 0$. If the sample mean $\bar{x}$ is used to estimate μ, then

$$\text{Prob}(\bar{x}-z(1-\alpha/2)\sigma/n^{\frac{1}{2}}<0) = \text{Prob}(n^{\frac{1}{2}}(\bar{x}-\mu)/\sigma<z(1-\alpha/2)-1/cv(\bar{x})),$$

where $cv(\bar{x}) = \sigma/n^{\frac{1}{2}}\mu$ is the coefficient of variation of $\bar{x}$ and $z(1-\alpha/2)$ is such that $\text{Prob}(Z<z(1-\alpha/2)) = 1-\alpha/2$, $Z \sim N(0,1)$. Therefore, if $cv(\bar{x}) > 1/z(1-\alpha/2)$, the lower confidence bound will be less than zero with probability greater than ½. Moreover, since J_{S1}, T_{S1} and J_{W1} are asymptotically normally distributed, one would expect large values of the coefficients of variation of these estimators to result in high coverages of zero. Furthermore, since the coverages will, in general, decrease as $1/n^{\frac{1}{2}}$, a substantial increase in sample size may be necessary to achieve a reasonably small probability of containment of zero.

Although all three of the estimators J_{S1}, T_{S1} and J_{W1} often result in high coverages of zero, the J_{S1} estimator appears to be the worst offender. As we have seen in the discussion of the results, the estimated probability of containment of zero for J_{S1} was often over 90% - even when $f^* \epsilon L$. An obvious consequence of this tendency to provide negative lower confidence bounds is that a substantial variance reduction may go undetected. This is, in fact, the case for sampling experiment number VII, for which the coverage of zero is approximately 95%, despite the fact that the variance ratio R is approximately equal to 25. Although the jackknifed Sukhatme estimator does tend to yield conservative interval estimates, its true confidence levels, which were, in general, the closest to the nominal level, appear unaffected by non-normality of $f^*(Y)$.

The Monte Carlo results indicate that T_{S1} is the least valuable of the one-sample estimators.

When $f^* \varepsilon H$, the poor confidence levels, accompanied by a significant positive bias, suggest that the use of this estimator could be hazardous. That is, a significant variance reduction may be incorrectly indicated, resulting in the decision to use the importance sampling estimator $\hat{I}_q$ instead of the alternative estimator $\hat{I}_p$ when $f^*(Y)$ displays extreme non-normality. As we have mentioned in the introduction, this is the most serious type of error.

The J_{W1} estimator is perhaps the best of the one-sample estimators. This estimator behaves similarly to the J_{S1} estimator when $f^* \varepsilon H$, resulting in conservative tests and confidence intervals, yet appears to be much less variable than the J_{S1} estimator when $f^* \varepsilon L$. Although the confidence intervals for J_{W1} are very short when $f^* \varepsilon L$, this estimator has a slight negative bias which, combined with these short confidence intervals, results in confidence levels somewhat lower than desired. This characteristic does not appear too serious, however, since the maximum coverages which are, in general, over 90%, occur between $.8{\times}R$ and R for the majority of cases and the coverages of $c{\times}R$ drop off rapidly as c decreases. Thus, J_{W1} appears far more powerful than J_{S1} , for testing $H_o: v_p = v_q$ vs.

$H_a: v_p \neq v_q$ when $v_p > v_q$, despite the fact that J_{W1} is negatively biased.

In an attempt to explain the difference in the performances of J_{S1} and J_{W1}, we have derived expressions for their asymptotic variances, that is, the asymptotic variances σ_s^2 of $N^{\frac{1}{2}}J_{S1}$ and σ_w^2 of $N^{\frac{1}{2}}J_{W1}$. These derivations are found in Appendix A and the results are stated here:

$$\begin{aligned}\sigma_s^2 = & (\gamma_2(f^*)+2)(v_p^2/v_q^2) + 4I\mu_3(f^*)v_p/v_q^3 + 4I^2/v_q \\ & + \mathrm{Var}(F_1F_1^*)/v_q^2 - 2v_p\mathrm{Cov}(F_1^{*2},F_1F_1^*)/v_q^3 \\ & + 4I(v_p-v_q)\mathrm{Cov}(F_1^*,F_1F_1^*)/v_q^3\end{aligned}$$

and

$$\begin{aligned}\sigma_w^2 = \sigma_s^2 + \{ & I^4 v_q \mathrm{Var}(C_1) - 2I^2 v_p \mathrm{Cov}(F_1^{*2},C_1) \\ & + (v_q-v_p)[2I^2 v_q - 4I^3\mathrm{Cov}(F_1^*,C_1)]\}/v_q^3 ,\end{aligned}$$

where $C_1 = p(X_1)/q(X_1)$ is defined as in (4.4).

From the above, we find that σ_s^2 and σ_w^2 are increasing functions of $\gamma_2(f^*)$, which could explain the long confidence intervals that occur when $f^* \varepsilon H$.

To obtain further insight into the problem, we have derived an expression for the difference

$\sigma_s^2 - \sigma_w^2$, when $q(x)$ is 'near' the optimum $q_o(x)$. This was accomplished by assuming

$$q(x) = \{f(x)p(x)/[1+cg(x)]\}/\int\{f(x)p(x)/[1+cg(x)]\} ,$$

where the constant c is small so that $q(x)$ is close to the optimum $q_o(x)$. Assuming this form of $q(x)$, we find that

$$\sigma_s^2 - \sigma_w^2 = (c^2/v_q^3)\Big(2I^4 v_p \mathrm{Var}_p(g) + 2I^2 v_p[\mathrm{Cov}_p(f,g)]^2 - I^4\{IE_p(fg^2) - [E_p(fg)]^2\}[E_p(1/f)I - 1]\Big) + \mathcal{O}(c^3) , \tag{6.1}$$

where $\lim_{c\to 0} \mathcal{O}(c^3)/c^3$ equals a constant. It is easily shown, using the Schwartz inequality, that

$$IE_p(fg^2) - [E_p(fg)]^2 \geq 0$$

and

$$IE_p(1/f) - 1 \geq 0 .$$

The apparent arbitrary nature of $g(x)$ adds to the complexity of the expression (6.1). Once f , p and q are specified, however, it may be possible to obtain a rough approximation of the difference in variances.

The jackknifed version of the natural log of the non-negative estimator J_{W1L} does not yield negative lower confidence bounds. We have seen, however, that this estimator behaves similarly to J_{W1} with respect to coverages of $c \times R$ when $f^*(Y)$ displays low non-normality. Although no significant bias was indicated for the sampling experiments in this category, the maximum coverages occurred slightly to the left of R. The disadvantage of using this estimator lies in its tendency to overestimate R and provide poor confidence levels when extreme non-normality of $f^*(Y)$ is displayed. Thus, as for the T_{S1} estimator, extreme non-normality of $f^*(Y)$ could result in tests based on J_{W1L} incorrectly indicating a variance reduction, with the obvious consequences.

6.5 The Two-sample Estimators

We begin this section by relating some of the basic results of the survey pertaining to the two-sample estimators. Recall that the two-sample estimators consist of the classical F statistic (F), the modified Box-Andersen statistic (F_{BA}), the Box statistics with subsample sizes 5 and 10 (B_5 and B_{10}), Arvesen's two-sample jackknifed variance ratio estimator (J_{A2}), Arvesen's two-sample jackknife version of the log of the ratio of the sample variances (J_{A2L}) and Miller's jackknife estimator with the log transformation (J_{M2L}).

The J_{A2} estimator permits negative confidence bounds and is unique among the two-sample estimators in this regard.

6.5.1 General Properties

Of the seven two-sample estimators, four were found to be superior to the others. These four consist of the jackknife estimators with the log transformation, J_{A2L} and J_{M2L}, and the Box estimators, B_5 and B_{10}.

With the exception of J_{A2}, all the two-sample interval estimators were affected similarly by non-normality of $f(X)$ and $f^*(Y)$. That is, large values of $\gamma_2(f)$ resulted in the confidence intervals being shifted to the left of R while large $\gamma_2(f^*)$ values resulted in a shift in the opposite direction. The J_{A2} interval estimator appeared to be shifted to the left of R in every case, although this shift was more pronounced for large $\gamma_2(f)$ values.

We now comment on the performance of each of the two-sample estimators individually.

(i) The classical F statistic (F)

The classical F statistic performed poorly, as expected. Non-normality of $f(X)$ and $f^*(Y)$ had a disastrous effect upon this estimator. Even in the cases where $f(X)$ and $f^*(Y)$ were not distributed radically different from normal random variables, the true confidence levels for this estimator were far too low.

(ii) The modified Box-Andersen estimator (F_{BA})

The modified Box-Andersen estimator was definitely an improvement over the classical F statistic. This estimator, however, also resulted in confidence levels that were well below the nominal level.

(iii) Arvesen's jackknife estimator without the log transformation (J_{A2})

Without the log transformation, the jackknife estimator J_{A2} proved to be of little value. It appeared to be much more sensitive to the non-normality of $f(X)$ and $f^*(Y)$ than did either J_{A2L} or J_{M2L}. The estimated confidence levels for this statistic ranged from 22% to 96% and, although it seemed to provide fairly good upper confidence bounds, the lower bounds were often negative.

(iv) The jackknifed estimators with the log transformation (J_{A2L} and J_{M2L})

The J_{A2L} and J_{M2L} interval estimators behaved almost identically with respect to coverages of $c \times R$. Of the four leading competitors, J_{A2L} and J_{M2L}

appeared to be the most powerful for testing $H_o: v_p = v_q$ vs. $H_a: v_p \neq v_q$. These estimators, however, had a tendency to provide confidence levels somewhat lower than desired. Typical confidence levels were between 70% and 80%.

(v) The Box estimator with subsample size 10 (B_{10})

The B_{10} interval estimator was the most conservative of the two-sample interval estimators. Confidence levels for the B_{10} estimator were generally 'close' to the nominal 90% level but the coverages of $c \times R$ did not decrease as rapidly as those for J_{A2L} and J_{M2L} as $|c - 1|$ increased.

(vi) The Box estimator with subsample size 5 (B_5)

The B_5 estimator proved less reliable than B_{10} and appeared more sensitive to non-normality of $f(X)$ and $f^*(Y)$. Although this estimator often led to confidence levels close to the nominal level, this was not a general rule. In fact, on occasion its true confidence level was far below those for J_{A2L} and J_{M2L} . The B_5 estimator provided shorter confidence intervals than B_{10} , yet was often more biased than this estimator.

6.5.2 Comparison of Coverages for Selected Values of c×R

As for the one-sample estimators, we compare the coverages of R , R/5 and 5×R and the maximum coverages obtained for the two-sample estimators.

(i) Estimated confidence levels (Table XXX)

The confidence levels for F and F_{BA} are obviously far too low. Although F_{BA} provides higher confidence levels than F , it is apparent that neither estimator warrants further consideration.

The J_{A2} estimator yields confidence levels that are lower than those for J_{A2L} and J_{M2L} for the M-L and H-L classifications, but higher than those for J_{A2L} and J_{M2L} for the L-M and L-H classifications. From Tables XXX and XXIX, we find that when $\gamma_2(f) < \gamma_2(f^*)$, the coverages of R are, in general, close to the nominal level for J_{A2} .

The jackknife estimators with the log transformation, J_{A2L} and J_{M2L} , yield confidence levels between 75% and 85% for the majority of cases. The lowest confidence levels occur for the L-H and H-L classifications.

The Box estimator with subsample size $k = 10$, B_{10} , provides the best confidence levels, in general. For 15 of the 18 sampling experiments, this estimator yields estimated confidence levels between 80% and 91%. The lowest confidence levels (approximately 54%) occur for runs XI and XVIII, for which the distributional classifications are L-H and H-L, respectively.

The B_5 estimator yields coverages of R which range in value from approximately 9% to approximately 91%, but for the majority of the sampling experiments, the confidence levels are greater than 70%. As for the other non-negative two-sample estimators, the lowest coverages occur for the L-H and H-L classifications.

(ii) Coverages of R/5 (Table XXXI)

With exception of sampling experiment number XVIII, the J_{A2} estimator yields coverages of R/5 which are greater than or equal to 70%. In fact, the coverages of R/5 are higher than those of R for all sampling experiments for which $\gamma_2(f) > \gamma_2(f^*)$.

The J_{A2L} and J_{M2L} estimators yield coverages of R/5 which are, in general, between 20% and 50%. The highest coverages occur when $f \in M$ or H; the maximum coverages of approximately 70% occur for runs XVII (H-M) and XVIII (H-L).

The B_{10} estimator tends to yield coverages of R/5 greater than those for J_{A2L}, J_{M2L} and B_5, but less than those for J_{A2}. The highest values (near 80%) occur when $f \varepsilon M$ or H, and the lowest values (approximately 30% to 50%) occur when $f \varepsilon L$.

The coverages of R/5 for B_5 range from as low as 1% to approximately 80%. Again, the lowest coverages occur when $f \varepsilon L$ and the highest coverages occur when $f \varepsilon M$ or H.

(iii) Coverages of 5×R (Table XXXII)

The coverages of 5×R for J_{A2} are generally lower than those for the Box estimators, B_5 and B_{10}, and the jackknife estimators, J_{A2L} and J_{M2L}. The highest coverages occur when $f^* \varepsilon M$ or H. With the exception of run number XI, all coverages of 5×R are less than 50%, the lowest values occuring when $f^* \varepsilon L$.

The other two-sample estimators appear to yield coverages of 5×R of roughly the same magnitude as they did for R/5, but the highest coverages occur when $f^* \varepsilon M$ or H, in this case.

(iv) Maximum coverages (Table XXXIII)

The maximum coverages for J_{A2} are close to 90% for all runs. Note, however, that the points at which the maximum coverages occur are, with the exception of run number IX, to the left of R. Of the 18 sampling experiments, 14 result in the point of maximum coverage being less than or equal to $.5 \times R$ for the J_{A2} estimator.

For all of the other two-sample estimators, the points at which the maximum coverages occur are less than R when $\gamma_2(f) > \gamma_2(f^*)$ and greater than R when $\gamma_2(f) < \gamma_2(f^*)$. That is, the confidence intervals appear to be shifted to the left of R when $\gamma_2(f) > \gamma_2(f^*)$ and shifted to the right of R when $\gamma_2(f) < \gamma_2(f^*)$.

The Box estimators yield almost identical maximum coverages and points of maximum coverage. Furthermore, the maximum coverages for these estimators are, in the majority of cases, approximately equal to 90%. The coverages of $R/5$ and $5 \times R$, however, are lower for B_5 than for B_{10}, indicating that the confidence intervals for B_5 are inclined to be shorter than those for B_{10}. Thus, we would expect bias to have a greater effect upon the confidence levels for the B_5 estimator, which would explain the tendency for B_5 to provide lower confidence levels than B_{10}.

Note that the maximum coverages for J_{A2L} and J_{M2L} are, for the majority of cases, between 75% and 85%. This suggests that the jackknife variance estimators, S_g^2 , may be underestimating the actual variance of the jackknife statistics. The resulting 'short' confidence intervals, together with a slight shift to either side of R , would account for their confidence levels being somewhat lower than desired.

6.5.3 The Effect of Non-normality

In the following, we discuss the performance of the two-sample estimators under the various distributional classifications. It is apparent from the discussion of the coverages of $R/5$, R and $5\times R$, that the F , F_{BA} and J_{A2} interval estimators are of little value. F and F_{BA} consistently provided confidence levels well below the nominal level, while J_{A2} proved incapable of providing realistic lower confidence bounds. As a result, the following discussion shall focus on the four main competitors, J_{A2L} , J_{M2L} , B_5 and B_{10} .

Tables XXIII through XXVII shall be used to aid in the discussion. Recall that these tables are comprised of the relevant portions of Tables I through XVIII, arranged according to the distributional classifications of $f(X)$ and $f^*(Y)$. Table XXIX, which records the values of R , $\gamma_1(f)$, $\gamma_2(f)$, $\gamma_1(f^*)$ and $\gamma_2(f^*)$ and the classification for each of the sampling experiments, will also prove useful.

Before examining the results of the Monte Carlo sampling experiments in detail, we make some preliminary comments. Let $\hat{\beta}$ be a random variable with a symmetric distribution such that $E(\hat{\beta}) = \beta$, and let I be a symmetric confidence interval for β . Then, if $c > 0$ is a real-valued constant, we have

$$
\begin{aligned}
\text{Prob}(c\times\exp(\beta) \in \exp(I)) &= \text{Prob}(\beta + \log(c) \in I) \\
&= \text{Prob}(\beta - \log(c) \in I) \\
&= \text{Prob}(\exp(\beta)/c \in \exp(I)) .
\end{aligned}
$$

Now, since the estimators J_{A2L} , J_{M2L} , B_5 and B_{10} are of the form $\exp(\hat{\beta})$, where $\hat{\beta}$ has a limiting symmetric distribution, we would expect the coverages of $c\times R$ and R/c to be approximately the same.

(i) <u>Classification L-L</u> (Table XXIII)

As one might expect, all the estimators display their best performances in this case. The jackknife estimators J_{A2L} and J_{M2L} , which behave almost identically with respect to coverages of $c\times R$, both yield coverages of $c\times R$ and R/c that are, as we would hope, approximately the same, thus reflecting the symmetry of the distributions before the exponential transformation. Their estimated confidence levels of approximately

83% and 84%, however, are below the nominal level.

The B_{10} estimator appears to be more conservative than J_{A2L} and J_{M2L}. Its estimated confidence levels are within 4% of the nominal level for all three sampling experiments of this category. The mean confidence interval length for B_{10} is higher than those for J_{A2L} and J_{M2L}, and the coverages of $c{\times}R$ do not decrease as rapidly as those for J_{A2L} and J_{M2L} as c increases or decreases from $c = 1$.

The B_5 estimator also yields mean confidence interval lengths that are shorter than those for B_{10}. These shorter confidence intervals, combined with a positive bias, appear to be the reason for the lower confidence level for run number VII. The confidence levels for the other two sampling experiments of this category, however, do not differ significantly from the nominal level. Note that B_5 appears to yield coverages of $c{\times}R$ and R/c that are less symmetric than those for J_{A2L}, J_{M2L} and B_{10}.

For run number VII, all the estimators, with the exception of J_{A2}, provide tests of $H_o{:}v_p = v_q$ vs. $H_a{:}v_p \neq v_q$ which lead to rejection of H_o, with probability close to 1. The power of the test based on J_{A2} is less than 50%, despite the fact that a considerable variance reduction is achieved ($R = 25$).

(ii) Classification M-L (Table XXIV)

In this category, the random variable $f(X)$ is said to display moderate non-normality. One should not be mislead by this classification. Referring to Table XXIX, we find that $\gamma_2(f)$ ranges in value from approximately 50 for run number XV to approximately 290 for run number XVI.

For the sampling experiments of this category, all the interval estimates appear shifted to left of R. The points of maximum coverage are less than R for each of the estimators and the coverages of $R/5$ are greater than those of $5 \times R$ in every case.

The jackknife estimators J_{A2L} and J_{M2L} again behave similarly with respect to coverages of $c \times R$, and both provide confidence levels lower than desired (74% to 80%). Despite the apparent negative shift of the confidence intervals, a significant positive bias is indicated for J_{A2L} and J_{M2L} for runs I and XVI. One possible explanation is that the distributions of these estimators are positively skewed and that a small number of the Monte Carlo sample values are the major contributors to the sample means. Note the large values obtained for the sample means and for the average confidence interval lengths for both of these estimators in sampling experiment number XV. Although the J_{M2L} estimator yields much larger values for both of these statistics, in this case, no significant difference is indicated. (The coefficients of variation of these quantities are very large.)

The negative shift of the confidence intervals has had an adverse effect on the confidence levels for all the estimators. The Box estimator with subsample size 10 appears to be affected the least in this regard. As for the L-L classification, the B_{10} estimator yields coverages of $c \times R$ which are, in general, higher than those for B_5, J_{A2L} and J_{M2L}. This is particularly noticeable for values of c between .1 and .5.

The B_5 estimator yields shorter confidence intervals than B_{10}, and the coverages decrease more rapidly than those for B_{10} as one moves in either direction from the points of maximum coverage. Furthermore, since the points of maximum coverage, which occur to the left of R, are approximately the same for both B_5 and B_{10}, the confidence levels for B_5 are generally lower than those for B_{10}.

Before leaving Table XXIV, we comment on the power of the test $H_o : v_p = v_q$ vs. $H_a : v_p \neq v_q$, for run number XV. As one might expect, J_{A2} is the least powerful, although the Box methods are not much better, with probability of rejecting H_o approximately .5 for both B_5 and B_{10}. The power of the tests based on J_{A2L} and J_{M2L} is roughly 90%.

(iii) Classification H-L (Table XXV)

The extreme non-normality of $f(X)$, which is characteristic of this category, tends to exaggerate the trends mentioned in the M-L classification.

None of the estimators performed well for sampling experiment number XVIII. Note the extremely large values of the Monte Carlo sample means of the J_{A2L} and J_{M2L} estimators for this sampling experiment. Examination of the Monte Carlo sample values of these estimators revealed the existence of extreme values of approximately $(5.5\times10^6)\times R$ for J_{M2L} and $(1.2\times10^5)\times R$ for J_{A2L}.

The B_{10} estimator again appears superior with respect to confidence levels, but yields coverages of $c\times R$ higher than those of J_{A2L} and J_{M2L} for small values of c.

The power of the test $H_o:v_p = v_q$ vs. $H_a:v_p \neq v_q$, for run number III, is high for all estimators except J_{A2}, but is much lower for run number XVIII. For sampling experiment number XVIII, the Box estimators appear to lag far behind J_{A2L} and J_{M2L} in power, although additional coverages between 0 and $.01\times R$ would be necessary to verify this.

(iv) Classification L-M (Table XXVI)

With the exception of J_{A2}, the non-normality of $f^*(Y)$ appears to have caused the confidence intervals to be shifted to the right of R. The differences in coverage of $c\times R$ and R/c are roughly of the same magnitude as in the M-L case, but are of opposite sign. Confidence levels for the L-M classification are also approximately the same as those of the M-L category.

A significant positive bias is indicated for all estimators of this category, thus reflecting a tendency to overestimate the variance ratio R.

Note that for run number XIII the power of the test $H_o:V_p = v_q$ vs. $H_a:v_p \neq v_q$ is significantly higher for B_5 than for B_{10}, J_{A2L} and J_{M2L}. The B_{10} estimator again appears to lag behind J_{A2L} and J_{M2L} in power. J_{A2} is obviously the least powerful, with probability of rejecting H_o approximately 10%.

(v) Classification L-H (Table XXVII)

As one might expect, the increased non-normality of $f^*(Y)$ tends to exaggerate the effects mentioned previously in the discussion of the L-M classification. That is, the confidence levels were generally low and the points of maximum coverage were shifted to the right for each of the non-negative two-sample estimators.

Although a significant positive bias is indicated for all the non-negative estimators, we do not find such extreme values for J_{A2L} and J_{M2L} as occurred in the H-L category.

Note the large values of the Monte Carlo sample means of the Box estimators for run number XI. These values indicate a very large positive bias, which is reflected by the points of maximum coverage that occur at approximately $100 \times R$.

(vi) <u>Classifications M-H and H-M</u> (Table XXVIII)

The four main competitors, J_{A2L} , J_{M2L} , B_5 and B_{10} , work surprisingly well in these cases of extreme non-normality. The B_5 estimator's performance, however, is generally superior. Its coverages of $c \times R$ are comparable to the coverages of $c \times R$, for J_{A2L} and J_{M2L} , for values of c less than .5 and greater than .5, yet its confidence levels, of approximately 90%, are higher than those of J_{A2L} and J_{M2L} in each case. Although the B_5 estimator is positively biased, the points of maximum coverage occur near R (between $.8 \times R$ and $1.25 \times R$) .

The B_{10} estimator also yields confidence levels that are not significantly different than the nominal level. This estimator, however, results in higher coverage of $c \times R$ than do the other non-negative two-sample estimators.

Examining the results of run number XIV, we find that J_{A2L}, J_{M2L} and B_5 provide tests of $H_o: v_p = v_q$ vs. $H_a: v_p \neq v_q$ of approximately the same power. The B_{10} estimator again proves less powerful than its competitors.

6.6 Discussion of the Two-sample Estimators

Although the two-sample formulation of the importance sampling problem reduces to the form of the usual variance comparison problem, we have seen that the nature of the distributions of $f(X)$ and $f^*(Y)$ result in violation of the usual assumptions of identical distributions, except for location and scale. The often radically different values of $\gamma_2(f)$ and $\gamma_2(f^*)$ are indications of the departure from these assumptions. Recall that, with the exception of the J_{A2} estimator, when $\gamma_2(f) > \gamma_2(f^*)$ the two-sample interval estimates appear shifted to the left of R and when $\gamma_2(f) < \gamma_2(f^*)$ they appear shifted in the opposite direction, thus indicating a tendency to overestimate R

when $\gamma_2(f^*)$ is the larger value and to underestimate R when $\gamma_2(f)$ is the larger value. This property of these estimators is obviously undesirable since tests based on an estimator with this property could likely lead to the use of the less efficient estimate of I if its standardized fourth cumulant were sufficiently larger than that of the alternative estimator. In this case, the distributional properties of the estimate chosen could have disastrous effects upon inferences and tests based on it.

In light of the very poor performances of the interval estimates F, F_{BA} and J_{A2}, we feel that they do not warrant further consideration as possible estimators of the 'relative efficiency' in the importance sampling problem.

Of the remaining two-sample estimators, J_{M2L} and J_{A2L} appear to be the most powerful but generally yield confidence levels below the nominal level. Considering the degree of non-normality displayed by $f(X)$ and $f^*(Y)$, however, these estimators performed reasonably well.

In Miller's review paper [14], he warns that the jackknife is not a device for correcting for outliers. In view of this, one should be suspicious if extremely large or small variance reductions are indicated. Inspection of the sample values $f(x_i)$ and $f^*(y_i)$ and of the pseudo-values of the jackknife estimators would be advisable in this case.

Recall that for the sampling experiments which resulted in extreme values of J_{M2L} and J_{A2L}, the sample means were much larger for the J_{M2L} estimator – over 44 times as large for sampling experiment number XVIII. Does this indicate that J_{M2L} is more sensitive to outliers than J_{A2L}? This deserves investigation. Consider a sample $(x_1,\ldots,x_n)$ for which x_1 is an extreme value. For our purposes, we assume that x_1 is the only extreme value. Then the sample variance, U say, is approximately equal to x_1^2/n and the jackknife estimator

$$J(\log(U)) = n\times\log(U) - \sum_{i=1}^{n} (n-1)\log(U_i)/n$$

$$\doteqdot n\times\log(x_1^2/n) - (n-1)^2\log(x_1^2/(n-1))/n$$

$$\doteqdot 2\log(x_1^2) .$$

Now, suppose F_1 is the only extreme value of $(F_1,\ldots,F_n)$ and $(F_1^*,\ldots,F_n^*)$. Then

$$\log(J_{M2L}) \doteqdot 2\log(F_1^2) .$$

(Recall that $J_{M2L} = \exp(J(\hat{\theta}_f) - J(\hat{\theta}_{f*}))$, where $\hat{\theta}_f = \log(S_f^2)$ and $\hat{\theta}_{f*} = \log(S_{f*}^2)$.)

For Arvesen's estimator J_{A2L} , defined in Section 3.3, we have

$$\log(J_{A2L}) = [J(\hat{\theta}_f) - J(\hat{\theta}_{f*}) + \log(S_f^2/S_{f*}^2)]/2$$

$$\doteqdot \log(F_1^2) + \tfrac{1}{2}\log(F_1^2/n)$$

$$\doteqdot 1.5\log(F_1^2) \ .$$

Therefore,

$$J_{A2L} \doteqdot \exp(1.5\log(F_1^2)) = F_1^3$$

and

$$J_{M2L} \doteqdot \exp(2\log(F_1^2)) = F_1^4 \ .$$

On the other hand, if F_1^* is extreme,

$$J_{A2L} \doteqdot \exp(-1.5\log(F_1^{*2})) = F_1^{*-3}$$

and

$$J_{M2L} \doteqdot \exp(-2\log(F_1^{*2})) = F_1^{*-4} \ .$$

From the above, it is apparent that Arvesen's method is, in fact, affected to a lesser degree than J_{M2L} by extreme values of $f(X)$ and $f^*(Y)$. Since these estimators behave similarly with respect to confidence levels and power, the J_{A2L} estimator appears preferable to J_{M2L}. Note that jackknifing has actually increased the effects of the outliers upon the estimator.

The major drawback of the Box method is that its bias is largest for small subsample sizes, yet increasing the subsample size results in loss of degrees of freedom of the associated t-statistic, which in turn results in long confidence intervals and conservative tests. If, on the other hand, the subsample size were to remain fixed while the sample size was increased, the bias would eventually become the major contributor to the mean square error of the estimator, with the obvious effect of decreasing the confidence level. This, in fact, appears to be the reason for the failure of the B_5 estimator since, as we have seen, the maximum coverage for B_5 was close to 90% in each of the runs and the low confidence levels resulted when the bias was large.

Since the magnitude of the bias depends on the distributional properties of $f(X)$ and $f^*(Y)$, an optimum subsample size may not be obtainable in general. Further research in this area is necessary. If, however, extreme non-normality of either $f(X)$ or or $f^*(Y)$ is suspected, it would be advisable to choose as large a subsample size as possible while retaining an acceptable number of degrees of freedom.

6.7 Summary

Although this investigation has not been conclusive, it has served to reveal the complexity of the importance sampling problem, to expose those estimation techniques most sensitive to non-normality, to provide a promising alternative to Sukhatmes' one-sample estimator and to indicate directions of further study.

We have seen that the estimators F, F_{BA}, J_{A2} and T_{S1} are of little value as interval estimates of the variance ratio R. They were all found to be sensitive to non-normality of $f(X)$ or $f^*(Y)$. The estimators F and F_{BA} consistently provided poor confidence levels, with the lowest values occurring when extreme non-normality of either $f(X)$ or $f^*(Y)$ was indicated. Both the estimators J_{A2} and T_{S1} were inclined to yield negative lower confidence bounds and poor confidence levels. The confidence levels for J_{A2} were lowest when $f(X)$ displayed extreme non-normality, and the lowest values for T_{S1} occurred when $f^*(Y)$ displayed extreme non-normality.

The two-sample estimators J_{A2L} , J_{M2L} , B_5 and B_{10} performed reasonably well, considering the degree of non-normality exhibited by the random variables f(X) and f*(Y) . The Monte Carlo results, however, indicated a tendency of these estimators to overestimate R when $\gamma_2(f) < \gamma_2(f^*)$ and to underestimate R when $\gamma_2(f) > \gamma_2(f^*)$, an obviously undesirable property. To evaluate the effect of this property upon tests of hypothesis, a larger number of distributions should be sampled to provide a wider range of values of the parameters $\gamma_1(f)$, $\gamma_2(f)$, $\gamma_1(f^*)$, $\gamma_2(f^*)$ and R . It would also be advisable to increase the Monte Carlo sample size in order to improve the precision of the estimators involved. Other possible changes to consider are varying the sample sizes (and subsample sizes of the Box estimators) and changing the significance level α .

In the discussion of the two-sample estimators in the previous section, we touched on some of the problems associated with choosing the most appropriate subsample size for the Box estimator. At this point, we wish merely to re-emphasize the need for further research in this area and remind the reader that B_k appears to become more conservative as k increases.

As point estimates, the two-sample estimators, with the exception of J_{A2} , have proven to be far more biased and more variable than the one-sample estimators. Consequently, these point estimators appear less reliable than the one-sample alternatives. A larger Monte Carlo sample size would have allowed a more detailed investigation of the bias and its contribution to the mean square error.

The one-sample approach to the importance sampling problem has received little attention in the past. As a result, very little was known about the reliability of the Sukhatme estimator prior to this investigation. The results of this study, however, do suggest that an alternative estimator is required in view of the fact that the Sukhatme estimator (or the jackknifed version of it) tends to yield negative lower confidence bounds for R. The jackknifed versions of the non-negative estimator $(\hat{\theta}_W)$ appear to be viable alternatives, although further investigation is obviously necessary to determine which is superior.

In this thesis we have studied the performance of the one-sample estimators under varying degrees of non-normality of $f^*(Y)$, without regard to the distributional properties of $f(Y)$ or the relationship between $f^*(Y)$ and $f(Y)$, where Y has density $q(x)$. Obviously the distributions of J_{S1} and J_{W1} are affected by the distributional properties of both these random variables, as is indicated by the expressions for the asymptotic variances of J_{S1} and J_{W1} (see Appendix A).

Although we may find that neither of the estimators J_{S1} or J_{W1} is universally superior to the other, it may be possible to determine conditions under which one is preferable to the other. Obviously this is another area requiring further research. In Appendix A we derived an expression for the difference in the asymptotic variances of J_{W1} and J_{S1} when $q(x)$ was 'arbitrarily close' to the optimum density $q_o(x)$. The resulting expression, however, was of such a general nature that little insight into the problem was achieved. Perhaps an alternative analytic approach, such as assuming a different form of $q(x)$, might prove more fruitful. It is possible, however, that the complexity of the problem will dictate the use of a numerical approach to the problem or require further restrictions on the functions $f(x)$, $p(x)$ and $q(x)$.

TABLE I

Data For Sampling Experiment I

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (C×R)[a]																				MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH
	C=0	.01	0.1	0.2	0.3	0.5	0.7	0.8	0.9	▼ 1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100				
J_{W1L}	0	0	0	0	0	002	247	785	972	780	477	175	100	040	015	0	0	0	0	0	.997	.009	.636	.039
J'_{W1L}	015	015	030	037	077	165	570	885	995	780	445	112	060	022	005	0	0	0	0	0	.968	.008	.673	.031
J_{S1}	755	760	772	787	807	857	900	915	930	945	960	982	980	957	785	265	055	0	0	0	1.07	.060	5.07	.019
T_{S1}	755	760	772	782	813	855	890	907	917	922	947	940	907	732	510	205	045	0	0	0	.842	.089	4.63	.015
J_{A2L}	0	0	095	430	630	817	835	822	787	770	672	560	505	447	395	330	277	160	102	027	1.57	.126	58.4	.425
J_{M2L}	0	0	057	357	595	790	817	805	782	765	680	570	525	457	420	347	292	182	112	040	2.46	.204	149	.467
J_{A2}	430	447	727	895	935	900	792	727	672	612	515	385	340	255	192	132	037	022	010	0	1.10	.075	2.62	.102
F	0	0	017	182	380	552	575	547	515	475	432	337	280	210	165	092	060	012	007	0	1.15	.075	1.52	.075
F_{BA}	0	0	032	245	470	657	712	707	697	667	617	492	452	395	340	272	210	120	055	012	1.15	.075	7.36	.208
B_5	0	005	370	707	865	917	850	807	782	730	630	465	422	327	262	182	112	022	002	0	.607	.038	2.40	.057
B_{10}	0	020	467	752	867	922	912	892	870	850	802	695	655	592	532	438	370	212	102	015	.767	.039	9.47	.128

$\gamma_1(f) = 4.6$ $\gamma_1(f^*) = -0.9$ $R = 1.0$ SAMPLED DISTRIBUTIONS: $p(X) = \text{Pareto}(30)$

$\gamma_2(f) = 70.8$ $\gamma_2(f^*) = 0.1$ $f(X) = X^6$ $M = 400$ $n = 30$[b] $q(X) = \text{Pareto}(18)$

[a]Coverages are represented as being out of a total of 1000; the nominal confidence level is 90%.
▼ Indicates the position of R=1.
[b]Samples from p(x) and q(x) are of equal size.

TABLE II

Data For Sampling Experiment II

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (CxR)[a]																				MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH
	C=0	.01	0.1	0.2	0.3	0.5	0.7	0.8	0.9	▽ 1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100				
J_{W1L}	0	0	0	0	0	002	165	797	945	740	485	232	135	055	027	005	0	0	0	0	1.02	.011	.677	.042
J_{W1}	010	012	017	042	047	102	607	915	972	747	445	152	077	032	012	0	0	0	0	0	.994	.010	.703	.041
J_{S1}	712	720	742	770	795	870	905	932	942	952	980	975	910	620	325	077	015	0	0	0	1.03	.047	3.60	.014
T_{S1}	707	710	740	757	782	847	887	902	922	932	938	865	765	505	285	072	015	0	0	0	.944	.058	3.40	.022
J_{A2L}	0	0	115	438	677	777	765	732	700	667	602	515	472	410	355	315	255	155	092	027	1.51	.132	57.9	.506
J_{M2L}	0	0	092	390	620	735	750	722	695	667	605	520	485	435	375	332	285	180	110	037	2.37	.221	148	.566
J_{A2}	342	352	680	887	920	815	697	627	590	552	467	332	290	220	177	127	080	032	007	0	1.06	.079	2.44	.108
F	0	0	037	262	430	520	512	502	457	435	392	270	220	175	142	077	063	017	007	0	1.10	.077	1.45	.077
F_{BA}	0	0	052	297	530	647	660	635	617	587	547	460	430	355	315	255	200	110	055	012	1.10	.077	6.71	.207
B_5	0	002	477	762	865	870	762	730	677	617	535	380	313	235	177	115	090	022	005	0	.505	.041	2.03	.067
B_{10}	0	025	575	807	872	890	865	840	817	802	740	632	582	527	470	365	305	155	102	015	.666	.046	13.8	.323

$\gamma_1(f) = 7.0$ $\gamma_1(f^*) = -0.5$ $R = 1.0$ SAMPLED DISTRIBUTIONS: $p(X) = $ Pareto(30)

$\gamma_2(f) = 1.5\times 10^6$ $\gamma_2(f^*) = -0.6$ $f(X) = X^{7.499}$ $M = 400$ $n = 30$[b] $q(X) = $ Pareto(15.002)

[a]Coverages are represented as being out of a total of 1000; the nominal confidence level is 90%.
▽Indicates the position of R=1.
[b]Samples from p(x) and q(x) are of equal size.

TABLE III

Data For Sampling Experiment III

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (C×R)[a]																				MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH
	C=0	.01▼	0.1	0.2	0.3	0.5	0.7	0.8	0.9	1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100				
J_{W1L}	0	0	0	0	0	0	007	047	767	563	147	032	017	012	010	002	002	0	0	0	.984	.009	.246	.101
J_{W1}	0	0	0	0	0	0	010	055	782	537	130	017	012	005	002	0	0	0	0	0	.977	.007	.207	.055
J_{S1}	313	317	375	442	512	642	787	830	837	942	992	632	265	087	050	025	012	0	0	0	1.02	.027	1.92	.034
T_{S1}	302	305	372	430	512	650	795	842	892	940	920	207	060	005	0	0	0	0	0	0	.816	.037	1.48	.025
J_{A2L}	0	020	270	522	712	845	847	852	835	822	757	662	627	540	462	372	310	202	115	047	1.50	.085	28.2	.276
J_{M2L}	0	025	260	487	667	825	827	832	825	800	740	660	612	547	472	380	317	217	135	052	2.00	.121	56.2	.343
J_{A2}	832	855	922	962	967	915	837	807	755	707	632	445	382	313	255	195	150	037	010	0	1.07	.069	3.44	.067
F	0	0	067	215	337	500	482	490	467	442	405	295	277	240	185	127	080	025	002	0	1.22	.063	1.61	.063
F_{BA}	0	007	157	362	517	642	730	727	707	682	635	542	487	410	360	292	245	145	080	010	1.22	.063	7.07	.154
B_5	0	002	365	670	820	907	915	917	895	870	797	660	612	522	447	347	277	115	040	0	.878	.052	4.51	.067
B_{10}	0	090	550	762	860	905	910	897	890	890	872	800	767	722	672	597	522	360	235	022	.979	.055	25.2	.250

$\gamma_1(f) = 6.9$ $\gamma_1(f^*) = 2.5$ $R = 25.07$ SAMPLED DISTRIBUTIONS: $p(X) = \text{Pareto}(20.2)$

$\gamma_2(f) = 2.0 \times 10^6$ $\gamma_2(f^*) = 10.5$ $f(X) = X^5$ $M = 400$ $n = 30$[b] $q(X) = \text{Pareto}(16.255)$

[a]Coverages are represented as being out of a total of 1000; the nominal confidence level is 90%.
▼Indicates the position of R=1.
[b]Samples from p(x) and q(x) are of equal size.

TABLE IV

Data For Sampling Experiment IV

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (C×R)[a]																				MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH
	C=0	.01	0.1	0.2	0.3	0.5	0.7	0.8	0.9	∇ 1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100				
J_{W1L}	0	0	0	012	020	090	390	757	982	780	507	267	172	095	063	032	017	0	0	0	1.01	.012	.890	.056
J_{W1}	137	140	160	182	237	342	590	925	997	790	490	202	127	067	037	015	002	0	0	0	.933	.013	1.04	.063
J_{S1}	895	895	897	897	897	905	907	912	917	917	925	935	935	940	945	965	975	982	902	007	2.53	.255	52.5	.034
T_{S1}	897	897	900	900	905	907	912	915	915	915	920	922	925	930	938	947	950	922	310	0	2.61	.278	40.8	.023
J_{A2L}	0	002	035	255	432	675	795	822	842	832	813	697	642	540	452	340	255	072	032	005	1.32	.048	5.44	.173
J_{M2L}	0	005	090	257	422	640	770	810	830	827	810	695	635	542	460	342	260	087	035	005	1.41	.058	6.57	.226
J_{A2}	597	617	727	832	897	925	902	882	855	805	710	535	480	367	265	152	092	017	0	0	1.14	.043	2.81	.049
F	0	0	012	097	220	442	542	560	595	605	565	465	405	332	235	132	075	012	0	0	1.25	.041	1.66	.041
F_{BA}	0	0	060	215	350	577	662	715	747	747	707	602	547	465	392	262	180	052	017	0	1.25	.041	3.25	.080
B_5	0	0	140	407	570	792	882	912	907	907	890	795	740	642	552	420	292	087	012	0	1.14	.035	4.45	.057
B_{10}	0	010	305	540	682	837	885	902	905	910	900	835	797	747	685	592	515	287	130	005	1.19	.040	10.2	.026

$\gamma_1(f) = 2.2$ $\gamma_1(f^*) = -1.9$ $R = 1.0$ SAMPLED DISTRIBUTIONS: $p(X) = \text{Pareto}(30)$

$\gamma_2(f) = 7.9$ $\gamma_2(f^*) = 4.5$ $f(X) = X$ $M = 400$ $n = 30$[b] $q(X) = \text{Pareto}(28)$

[a]Coverages are represented as being out of a total of 1000; the nominal confidence level is 90%.
∇Indicates the position of R=1.
[b]Samples from p(x) and q(x) are of equal size.

TABLE V

Data For Sampling Experiment V

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (CxR)[a]																				MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH
	C=0	.01	0.1	0.2	0.3	0.5	0.7	0.8	0.9	▽ 1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100				
J_{W1L}	0	0	022	055	122	410	677	922	925	815	650	440	387	297	232	160	112	015	0	0	1.14	.037	1.93	.053
J_{W1}	490	492	527	587	627	697	940	992	980	867	675	417	340	242	170	107	070	010	0	0	.881	.048	2.56	.053
J_{S1}	957	957	955	952	950	942	930	917	905	890	880	847	830	777	742	652	575	340	077	0	.981	.249	15.5	.041
T_{S1}	862	862	870	872	872	865	875	870	860	855	832	787	762	710	672	587	502	277	063	0	3.27	.069	9.20	.032
J_{A2L}	0	020	137	235	332	525	617	662	720	750	787	800	802	720	622	505	380	085	005	0	1.76	.043	4.33	.037
J_{M2L}	0	035	147	265	355	532	620	677	720	750	785	782	770	677	577	450	330	060	005	0	1.70	.045	3.92	.036
J_{A2}	652	657	705	760	815	857	870	890	897	895	865	780	717	597	530	352	260	030	0	0	1.49	.052	4.71	.040
F	0	0	005	045	122	292	430	472	472	515	575	527	505	465	405	280	205	030	0	0	1.83	.041	2.43	.041
F_{BA}	0	012	120	210	320	452	590	617	647	667	707	690	657	592	510	392	245	032	0	0	1.74	.037	3.05	.033
B_5	0	0	022	112	232	422	557	630	680	730	807	920	917	910	887	813	707	370	107	005	2.56	.034	9.85	.058
B_{10}	0	015	212	370	495	655	767	802	822	850	892	920	922	900	892	815	752	467	262	020	2.06	.033	21.1	.208

$\gamma_1(f) = -0.9$ $\gamma_1(f^*) = 4.6$ $R = 1.0$ SAMPLED DISTRIBUTIONS: $p(X) =$ Pareto(18)

$\gamma_2(f) = 0.1$ $\gamma_2(f^*) = 70.8$ $f(X) = (5/3)X^{-6}$ $M = 400$ $n = 30$[b] $q(X) =$ Pareto(30)

[a]Coverages are represented as being out of a total of 1000; the nominal confidence level is 90%.
▽Indicates the position of R=1.
[b]Samples from p(x) and q(x) are of equal size.

TABLE VI

Data For Sampling Experiment VI

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (C×R)[a]																				MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH
	C=0	.01	0.1	0.2	0.3	0.5	0.7	0.8	0.9	▽ 1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100				
J_{W1L}	0	005	042	100	160	400	695	882	895	813	697	570	517	412	340	225	170	047	010	0	1.30	.045	2.71	.063
J_{W1}	575	585	627	665	697	775	970	997	980	887	737	563	465	345	260	160	110	017	007	0	.864	.072	3.76	.068
J_{S1}	950	950	952	942	938	912	885	880	877	872	862	825	807	752	707	655	570	265	077	0	1.21	.174	13.6	.044
T_{S1}	800	795	800	802	807	805	782	775	780	772	757	720	710	660	617	585	507	250	063	0	3.24	.064	7.95	.036
J_{A2L}	0	027	155	255	347	472	563	602	622	667	732	785	777	755	695	595	438	115	007	0	1.95	.042	4.79	.037
J_{M2L}	0	037	180	285	357	485	572	605	630	667	722	760	735	702	657	547	390	092	005	0	1.86	.044	4.27	.035
J_{A2}	657	660	700	745	780	825	872	872	870	872	867	813	775	685	602	440	322	050	002	0	1.65	.051	5.47	.041
F	0	0	017	063	130	220	342	392	410	435	502	542	520	500	465	350	262	037	002	0	2.06	.041	2.73	.041
F_{BA}	0	012	110	200	300	430	502	547	572	587	635	657	647	620	567	430	297	052	005	0	2.06	.041	3.47	.036
B_5	0	0	020	090	142	313	480	535	595	617	730	822	870	890	880	832	762	477	197	002	3.42	.039	12.4	.052
B_{10}	0	015	155	305	420	582	695	740	760	802	840	877	890	900	877	855	807	575	317	025	2.61	.037	22.4	.087

$\gamma_1(f) = -0.5$ $\gamma_1(f^*) = 7.0$ $R = 1.0$ SAMPLED DISTRIBUTIONS: $p(X) = \text{Pareto}(15.002)$

$\gamma_2(f) = -0.6$ $\gamma_2(f^*) = 1.5\times10^6$ $f(X) = (30/15.002)X^{-7.499}$ $M = 400$ $n = 30$[b] $q(X) = \text{Pareto}(30)$

[a]Coverages are represented as being out of a total of 1000; the nominal confidence level is 90%.
▽Indicates the position of R=1.
[b]Samples from p(x) and q(x) are of equal size.

TABLE VII

Data For Sampling Experiment VII

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (CxR)[a]																				MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH
	C=0	.01 ▼	0.1	0.2	0.3	0.5	0.7	0.8	0.9	1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100				
J_{W1L}	0	0	0	010	012	080	520	732	880	780	515	282	202	087	052	025	010	0	0	0	1.02	.015	.877	.047
J_{W1}	110	120	137	172	232	402	697	813	917	785	502	230	155	067	037	007	002	0	0	0	.964	.015	1.02	.050
J_{S1}	952	952	962	962	962	967	947	938	932	912	865	750	705	612	522	327	212	025	002	0	.877	.076	5.55	.036
T_{S1}	782	792	820	847	865	887	877	877	870	852	807	705	645	570	467	282	190	020	0	0	1.60	.144	3.56	.029
J_{A2L}	0	0	070	192	290	502	705	755	807	827	857	813	750	622	480	257	147	015	0	0	1.30	.030	2.71	.030
J_{N2L}	0	002	087	207	302	527	715	757	807	827	852	792	720	577	430	227	107	007	0	0	1.25	.031	2.53	.029
J_{A2}	542	547	612	700	770	852	910	917	930	920	880	722	645	465	317	140	063	005	0	0	1.16	.034	3.02	.034
F	0	0	0	022	112	317	520	595	657	692	715	637	575	432	287	140	063	002	0	0	1.36	.029	1.80	.029
F_{BA}	0	0	045	150	277	522	680	750	772	802	775	675	605	455	332	152	077	007	0	0	1.34	.033	2.14	.029
B_5	0	0	012	072	202	490	637	705	767	800	880	870	862	827	785	630	542	197	042	0	2.03	.035	6.48	.054
B_{10}	0	0	112	305	445	657	785	822	847	862	867	852	837	777	732	642	532	332	130	010	1.65	.032	10.3	.088

$\gamma_1(f) = -0.5$ $\gamma_1(f^*) = 2.5$ $R = 24.99$ SAMPLED DISTRIBUTIONS: $P(X) =$ Pareto(18)

$\gamma_2(f) = -0.7$ $\gamma_2(f^*) = 11.0$ $f(X) = 5/(3X^{9.89})$ $M = 400$ $n = 30$[b] $Q(X) =$ Pareto(30)

[a]Coverages are represented as being out of a total of 1000; the nominal confidence level is 90%.
▼Indicates the position of R=1.
[b]Samples from p(x) and q(x) are of equal size.

TABLE VIII

Data For Sampling Experiment VIII

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (C×R)[a]																				MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH
	C=0	.01	0.1	0.2	0.3	0.5	0.7	0.8	0.9	▼1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100				
J_{W1L}	0	0	002	020	032	132	472	797	980	792	517	297	220	152	097	050	032	002	0	0	1.02	.015	1.06	.057
J_{W1}	180	185	212	262	313	415	682	967	1000	800	510	237	180	092	057	030	012	0	0	0	.916	.018	1.29	.064
J_{S1}	862	862	862	860	857	855	852	852	850	850	845	842	837	830	813	785	772	682	477	042	-1.04	.281	58.5	.036
T_{S1}	855	855	852	852	850	847	845	837	837	835	835	827	827	817	815	782	755	655	445	035	5.30	.190	43.2	.029
J_{A2L}	0	005	100	272	417	672	765	800	810	840	832	750	700	582	500	360	247	060	017	002	1.29	.040	4.11	.105
J_{M2L}	0	010	105	280	415	665	762	792	807	837	820	735	672	582	492	332	250	060	020	002	1.32	.042	4.44	.148
J_{A2}	630	637	722	797	862	917	905	892	882	867	787	582	515	397	290	167	102	005	0	0	1.12	.043	2.94	.038
F	0	0	012	072	175	442	585	597	612	617	610	467	415	315	260	132	085	007	0	0	1.28	.038	1.69	.038
F_{BA}	0	0	052	180	335	547	675	707	732	747	715	607	577	477	385	280	215	060	012	0	1.38	.043	3.09	.065
B_5	0	0	087	292	495	740	845	890	895	907	912	832	792	717	637	487	407	140	035	0	1.39	.038	5.59	.061
B_{10}	0	005	287	515	652	797	897	900	910	910	902	865	837	772	717	622	540	305	145	010	1.36	.040	12.6	.111

$\gamma_1(f) = -1.9$ $\gamma_1(f^*) = 2.2$ $R = 1.0$ SAMPLED DISTRIBUTIONS: $p(X) = \text{Pareto}(28)$

$\gamma_2(f) = 4.5$ $\gamma_2(f^*) = 7.9$ $f(X) = 15/(14X)$ $M = 400$ $n = 30$[b] $q(X) = \text{Pareto}(30)$

[a]Coverages are represented as being out of a total of 1000; the nominal confidence level is 90%.
▼Indicates the position of R=1.
[b]Samples from p(x) and q(x) are of equal size.

TABLE IX

Data For Sampling Experiment IX

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (C×R)[a]																				MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH
	C=0	.01	0.1	0.2	0.3	0.5	0.7	0.8	0.9	▽ 1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100				
J_{W1L}	0	0	027	082	142	380	785	857	842	807	727	592	515	402	320	192	100	005	0	0	1.20	.029	2.06	.039
J_{W1}	502	517	572	635	690	890	990	992	950	897	770	540	442	327	235	100	042	0	0	0	.953	.037	2.62	.040
J_{S1}	902	902	905	917	925	927	915	912	912	907	877	805	765	680	575	442	290	045	0	0	1.15	.076	6.00	.031
T_{S1}	707	710	737	765	785	805	832	832	842	842	810	745	712	632	545	405	252	035	0	0	1.88	.047	4.07	.023
J_{A2L}	0	010	075	185	322	520	680	722	757	800	840	813	765	675	537	335	222	040	010	0	1.48	.038	3.50	.049
J_{M2L}	0	012	082	205	322	522	680	722	755	795	827	797	745	635	500	317	220	035	012	0	1.47	.040	3.46	.057
J_{A2}	577	587	675	742	785	867	912	902	897	900	862	737	637	475	342	190	112	015	0	0	1.32	.041	3.21	.034
F	0	0	010	045	115	322	492	565	617	635	647	580	547	430	315	190	110	017	0	0	1.50	.036	1.98	.036
F_{BA}	0	005	042	155	247	472	630	657	702	732	777	710	685	547	457	307	155	017	005	0	1.47	.033	2.71	.041
B_5	0	0	017	120	280	552	732	785	830	857	917	925	905	845	742	595	452	135	012	0	1.68	.029	5.19	.038
B_{10}	0	005	177	377	522	692	785	825	857	870	887	875	870	837	797	672	585	330	130	005	1.59	.035	10.9	.070

$\gamma_1(f) = 1.1$ $\gamma_1(f^*) = 3.5$ R = 1.0 SAMPLED DISTRIBUTIONS: p(X) = Gamma(3,1)

$\gamma_2(f) = 10.0$ $\gamma_2(f^*) = 287$ f(X) = X M = 400 n = 30[b] q(X) = Gamma(5,1)

[a]Coverages are represented as being out of a total of 1000; the nominal confidence level is 90%.
▽Indicates the position of R=1.
[b]Samples from p(x) and q(x) are of equal size.

TABLE X

Data For Sampling Experiment X

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (C×R)[a]																				MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH
	C=0	.01	0.1	0.2	0.3	0.5	0.7	0.8	0.9	▼1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100				
J_{W1L}	0	100	247	307	362	472	515	522	535	552	577	637	662	670	690	712	702	602	297	002	4.98	.066	15.5	.045
J_{W1}	880	880	895	907	910	922	922	915	915	910	902	907	895	872	845	807	745	537	212	0	.657	.800	26.5	.056
J_{S1}	865	865	872	880	890	905	907	907	907	910	907	900	895	867	847	817	752	512	225	005	1.69	.351	27.8	.052
T_{S1}	460	460	490	512	525	547	565	567	577	597	605	627	645	635	615	597	557	370	160	005	7.11	.071	10.7	.050
J_{A2L}	0	145	375	537	637	730	755	767	787	780	817	817	820	800	770	727	688	550	417	210	15.3	.350	3.9×10^4	.922
J_{M2L}	0	150	342	482	590	682	725	730	742	752	767	765	757	737	720	680	657	530	405	222	66.6	.663	3.2×10^5	.962
J_{A2}	967	967	970	970	960	940	925	917	897	875	840	767	737	675	640	552	492	325	175	045	4.40	.210	31.0	.121
F	0	010	085	127	150	207	262	295	290	300	325	327	315	292	265	240	215	140	075	027	6.61	.149	8.75	.149
F_{BA}	0	060	257	362	410	475	517	535	547	557	567	577	575	592	580	525	490	410	295	100	8.30	.255	119	.340
B_5	0	040	447	620	730	822	865	872	887	897	900	870	857	825	815	752	695	530	325	057	2.38	.077	33.1	.148
B_{10}	0	262	617	715	782	840	865	880	882	892	897	902	877	895	872	847	817	700	572	317	3.73	.129	995	.398

$\gamma_1(f) = 13.6$ $\gamma_1(f^*) = 66.3$ R= 1.0 SAMPLED DISTRIBUTIONS: p(X)= Gamma(13,1)

$\gamma_2(f) = 597$ $\gamma_2(f^*) = 3.6\times10^6$ $f(X) \doteq X^6$ M = 400 n = 30[b] q(X)= Gamma(25,1)

[a]Coverages are represented as being out of a total of 1000; the nominal confidence level is 90%.
▼Indicates the position of R=1.
[b]Samples from p(x) and q(x) are of equal size.

TABLE XI

Data For Sampling Experiment XI

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (C×R)[a]																			▼ 218	MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH
	C=0	.01	0.1	0.2	0.3	0.5	0.7	0.8	0.9	1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100				
J_{W1L}	0	070	205	292	340	470	537	555	582	600	607	622	622	620	610	627	615	637	637	245	14.7	.091	97.0	.162
J_{W1}	942	942	942	942	942	940	930	927	925	922	917	897	892	877	860	842	815	705	560	085	-1.72	1.01	84.1	.058
J_{S1}	930	930	932	932	932	932	932	932	932	932	932	932	932	925	922	917	917	877	770	255	-.759	4.61	162	.061
T_{S1}	715	715	717	717	717	720	722	722	727	727	732	732	742	747	752	762	772	810	827	145	30.5	.072	68.2	.020
J_{A2L}	0	142	252	310	337	380	425	435	455	465	512	563	577	597	632	657	688	742	730	465	39.4	.146	309	.100
J_{M2L}	0	170	285	337	357	425	452	477	487	505	535	580	597	605	630	672	685	717	665	355	36.2	.158	174	.101
J_{A2}	960	960	960	960	960	960	960	957	955	957	952	955	950	942	945	940	927	862	742	365	3.24	2.13	344	.096
F	0	0	027	047	050	070	065	077	087	090	102	130	150	188	192	210	235	265	255	135	44.0	.134	58.2	.132
F_{BA}	0	035	127	200	210	240	292	305	313	322	347	400	405	440	460	510	517	512	465	202	44.0	.134	96.6	.127
B_5	0	0	0	010	015	032	063	070	075	085	117	175	192	215	250	322	367	542	722	907	159	.069	2170	.109
B_{10}	0	050	210	270	335	412	465	487	510	535	557	622	637	682	697	742	767	867	922	900	78.1	.087	19900	.213

$\gamma_1(f) = 0.8$ $\gamma_1(f^*) = 310$ $R = .0046$ SAMPLED DISTRIBUTIONS: $p(X) = \text{Gamma}(6,1)$

$\gamma_2(f) = 1.0$ $\gamma_2(f^*) = 1.2\times10^5$ $f(X) \doteq X$ $N = 400$ $n = 30$[b] $q(X) = \text{Gamma}(3,0.9)$

[a]Coverages are represented as being out of a total of 1000; the nominal confidence level is 90%.
▼Indicates the position of R=1.
[b]Samples from p(x) and q(x) are of equal size.

TABLE XII

Data For Sampling Experiment XII

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (C×R)[a]																				MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH
	C=0	.01	0.1▽	0.2	0.3	0.5	0.7	0.8	0.9	1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100				
J_{W1L}	0	040	125	200	252	367	460	512	592	635	725	727	642	527	400	222	117	0	0	0	1.48	.033	2.18	.030
J_{W1}	460	465	492	530	560	617	685	727	750	772	820	735	630	472	335	157	082	0	0	0	1.07	.065	3.46	.043
J_{S1}	690	690	727	762	792	820	850	865	880	870	837	772	720	620	547	355	260	050	007	0	1.06	.099	5.99	.051
T_{S1}	167	172	180	272	357	490	557	592	605	605	607	585	557	515	422	277	200	036	005	0	2.09	.044	2.77	.042
J_{A2L}	0	082	250	352	438	590	662	712	732	762	787	815	817	792	757	657	577	335	142	012	2.58	.076	12.3	.091
J_{M2L}	0	100	257	362	447	592	675	705	720	740	770	780	775	735	695	600	520	285	120	017	2.59	.081	11.8	.110
J_{A2}	872	875	907	922	930	942	942	940	927	922	907	817	795	682	617	500	430	167	052	0	1.81	.093	9.90	.065
F	0	005	050	095	140	245	350	367	392	415	435	455	420	375	365	305	255	092	032	0	2.62	.072	3.47	.072
F_{BA}	0	035	188	262	342	462	563	587	605	622	637	637	615	567	510	462	365	185	085	002	2.62	.072	7.18	.096
B_5	0	0	067	170	310	512	620	662	690	732	787	862	887	890	885	862	840	640	345	047	4.03	.052	25.8	.074
B_{10}	0	077	352	527	615	740	825	852	860	877	902	910	902	902	880	845	807	688	515	150	2.84	.064	67.7	.159

$\gamma_1(f) = 1.9$ $\quad$ $\gamma_1(f^*) = 8.1$ $\quad$ $R = 8.38$ $\quad$ SAMPLED DISTRIBUTIONS: $p(X) = \text{Gamma}(8,1)$

$\gamma_2(f) = 6.0$ $\quad$ $\gamma_2(f^*) = 4.0\times10^4$ $\quad$ $f(X) = X^2$ $\quad$ $M = 400$ $\quad$ $n = 30$[b] $\quad$ $q(X) = \text{Gamma}(13, 0.7501)$

[a]Coverages are represented as being out of a total of 1000; the nominal confidence level is 90%.
▽Indicates the position of R=1.
[b]Samples from p(x) and q(x) are of equal size.

TABLE XIII

Data For Sampling Experiment XIII

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (C×R)[a]																				MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH
	C=0	.01	0.1	▼0.2	0.3	0.5	0.7	0.8	0.9	1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100				
J_{W1L}	0	030	140	222	317	532	620	667	705	722	745	730	710	672	615	517	402	117	010	0	1.78	.044	4.58	.042
J_{W1}	720	732	777	825	852	915	932	927	912	897	850	775	725	640	552	405	302	060	002	0	.896	.119	6.61	.051
J_{S1}	955	955	960	962	960	952	945	947	945	940	905	842	797	725	662	552	450	167	037	0	.743	.225	10.8	.055
T_{S1}	725	725	752	775	790	802	815	813	820	825	790	730	700	620	575	462	375	140	020	0	2.60	.051	5.23	.040
J_{A2L}	0	035	207	355	472	585	682	720	752	772	805	820	822	777	730	645	570	280	132	017	2.32	.069	25.5	.569
J_{M2L}	0	045	227	365	472	605	685	727	747	762	802	785	782	742	705	617	507	247	102	020	2.45	.103	52.0	.790
J_{A2}	875	875	900	925	932	947	940	930	932	920	892	797	747	685	605	472	382	152	040	0	1.61	.081	8.57	.059
F	0	0	020	090	142	287	375	410	438	435	438	442	445	400	372	313	250	075	020	0	2.30	.059	3.05	.059
F_{BA}	0	010	135	265	402	517	587	620	632	650	660	645	635	600	565	457	357	155	063	005	2.30	.059	6.69	.107
B_5	0	0	037	145	282	480	580	647	677	720	800	880	907	915	915	877	837	590	325	032	3.98	.055	22.9	.076
B_{10}	0	030	305	462	560	700	772	805	817	827	862	907	902	900	892	867	817	627	442	142	3.07	.055	108	.318

$\gamma_1(f) = 1.9$ $\gamma_1(f^*) = 7.1$ $R = 6.71$ SAMPLED DISTRIBUTIONS: $p(X) = \text{Gamma}(8,1)$

$\gamma_2(f) = 6.0$ $\gamma_2(f^*) = 337$ $f(X) = X^2$ $M = 400$ $n = 30$[b] $q(X) = \text{Gamma}(13,0.8)$

[a]Coverages are represented as being out of a total of 1000; the nominal confidence level is 90%.
▼Indicates the position of R=1.
[b]Samples from p(x) and q(x) are of equal size.

TABLE XIV

Data For Sampling Experiment XIV

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (C×R)[a]																			MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH
	C=0	.01	▼0.1	0.2	0.3	0.5	0.7	0.8	0.9	1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100				
J_{W1L}	0	017	080	135	177	207	450	517	582	652	710	670	587	430	307	135	077	005	002	0	1.44	.031	1.94	.052
J_{W1}	377	380	410	425	470	552	647	705	752	795	802	655	555	382	240	105	040	0	0	0	1.17	.045	2.62	.039
J_{S1}	340	340	370	395	427	517	632	680	707	735	767	662	567	400	247	080	042	0	0	0	1.21	.043	2.49	.044
T_{S1}	002	002	002	002	005	037	200	277	350	417	540	542	457	287	180	055	032	0	0	0	1.64	.024	1.11	.030
J_{A2L}	0	067	287	462	577	707	765	787	790	802	800	782	770	735	690	612	540	362	210	077	4.27	.230	2010	.886
J_{M2L}	0	080	292	462	557	670	745	757	772	775	782	747	737	697	662	582	515	345	207	085	9.03	.447	9820	.940
J_{A2}	912	917	952	962	960	940	915	892	865	847	790	717	667	590	502	415	347	165	077	007	1.93	.124	12.4	.152
F	0	002	070	137	207	300	375	382	395	397	397	367	357	315	265	222	192	120	032	005	2.86	.122	3.79	.122
F_{BA}	0	030	188	347	432	530	580	595	612	620	630	620	587	557	500	442	392	235	142	035	2.86	.122	22.5	.276
B_5	0	007	255	465	602	752	842	857	870	872	912	887	860	855	830	750	692	447	255	010	2.35	.059	18.0	.086
B_{10}	0	105	467	652	742	845	877	885	890	897	890	887	870	827	805	772	737	607	457	162	2.48	.081	496	.756

$\gamma_1(f) = 4.8$ $\gamma_1(f^*) = 9.9$ $R = 11.34$ SAMPLED DISTRIBUTIONS: $p(X) = \text{Gamma}(12,1)$

$\gamma_2(f) = 49.0$ $\gamma_2(f^*) = 6.0\times10^6$ $f(X) = X^4$ $M = 400$ $n = 30$[b] $q(X) = \text{Gamma}(20, 0.751)$

[a]Coverages are represented as being out of a total of 1000; the nominal confidence level is 90%.

▼Indicates the position of R=1.

[b]Samples from p(x) and q(x) are of equal size.

TABLE XV

Data For Sampling Experiment XV

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (C×R)[a]																				MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH
	C=0	.01	▼0.1	0.2	0.3	0.5	0.7	0.8	0.9	1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100				
J_{W1L}	0	0	0	0	0	063	582	760	897	897	702	247	140	052	017	0	0	0	0	0	1.03	.012	.856	.023
J_{W1}	005	005	010	012	027	292	815	930	940	887	630	175	087	015	005	0	0	0	0	0	.999	.012	.850	.021
J_{S1}	0	002	005	007	027	345	825	927	955	938	712	202	063	005	0	0	0	0	0	0	1.01	.011	.889	.013
T_{S1}	0	0	0	0	012	205	688	857	925	917	720	217	065	005	0	0	0	0	0	0	1.05	.012	.834	.012
J_{A2L}	0	0	125	447	688	800	807	797	772	737	642	535	507	438	375	300	272	165	122	050	2.56	.452	3150	.988
J_{M2L}	0	0	075	382	612	770	790	782	770	732	665	557	517	452	400	325	280	188	140	055	9.87	.784	20900	.996
J_{A2}	337	362	645	865	907	862	747	685	622	565	462	325	277	227	177	112	090	025	007	002	1.17	.152	2.84	.212
F	0	0	032	225	427	567	550	517	495	455	390	262	247	185	142	080	063	020	002	002	1.20	.151	1.59	.151
F_{BA}	0	0	050	300	502	685	700	695	688	665	592	487	447	372	322	265	210	120	055	017	1.20	.151	9.97	.348
B_5	0	005	515	802	877	857	782	717	652	605	512	360	320	220	177	110	067	010	002	0	.467	.043	1.83	.059
B_{10}	0	047	540	762	860	890	860	817	802	760	685	612	580	520	450	375	317	177	080	015	.653	.044	8.55	.141

$\gamma_1(f) = 4.8$ $\gamma_1(f^*) = -0.5$ $R = 11.34$ SAMPLED DISTRIBUTIONS: $p(X) = \text{Gamma}(12,1)$

$\gamma_2(f) = 49.0$ $\gamma_2(f^*) = -0.9$ $f(X) = X^4$ $M = 400$ $n = 30$[b] $q(X) = \text{Gamma}(12,1.496016)$

[a]Coverages are represented as being out of a total of 1000; the nominal confidence level is 90%.
▼Indicates the position of R=1.
[b]Samples from p(x) and q(x) are of equal size.

TABLE XVI

Data For Sampling Experiment XVI

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (C×R)[a]																				MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH
	C=0	.01	0.1	0.2	0.3	0.5	0.7	0.8	0.9	▼ 1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100				
J_{W1L}	0	0	005	015	032	160	787	892	875	777	680	460	380	237	140	057	025	0	0	0	1.08	.019	1.33	.039
J_{W1}	177	180	235	305	407	705	967	990	935	830	675	385	280	142	080	025	010	0	0	0	.972	.020	1.55	.043
J_{S1}	845	847	867	890	910	920	917	912	917	910	860	752	662	515	352	172	072	002	0	0	1.00	.050	3.58	.020
T_{S1}	680	688	722	752	792	857	875	872	875	877	832	725	637	495	345	170	070	002	0	0	1.29	.041	2.94	.014
J_{A2L}	0	0	040	222	462	765	857	840	837	800	722	597	520	442	355	262	185	075	047	010	1.18	.054	5.29	.158
J_{M2L}	0	0	035	220	435	745	845	827	822	795	722	597	522	438	367	277	205	082	055	012	1.31	.065	6.69	.193
J_{A2}	410	430	607	790	887	930	860	835	775	720	600	402	350	240	172	102	057	022	0	0	1.01	.047	2.18	.059
F	0	0	017	110	255	547	652	647	655	635	565	382	322	195	152	072	045	010	0	0	1.08	.046	1.43	.046
F_{BA}	0	0	025	165	337	655	760	780	767	737	662	522	477	365	295	202	142	057	027	0	1.08	.046	3.12	.098
B_5	0	0	135	452	692	905	930	917	880	857	785	642	552	415	317	202	120	017	0	0	.818	.032	2.65	.045
B_{10}	0	005	330	585	740	870	880	887	875	870	825	730	692	627	565	462	377	177	063	005	.912	.040	9.63	.291

$\gamma_1(f) = 3.5$ $\gamma_1(f^*) = 1.1$ $R = 1.0$ SAMPLED DISTRIBUTIONS: $p(X) = \text{Gamma}(5,1)$

$\gamma_2(f) = 287$ $\gamma_2(f^*) = 10.0$ $f(X) = 12X^{-1}$ $M = 400$ $n = 30$[b] $q(X) = \text{Gamma}(3,1)$

[a]Coverages are represented as being out of a total of 1000; the nominal confidence level is 90%.
▼Indicates the position of R=1.
[b]Samples from p(x) and q(x) are of equal size.

TABLE XVII

Data For Sampling Experiment XVII

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (C×R)[a]																				MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH
	C=0	.01	0.1	0.2	0.3	0.5	0.7	0.8	0.9	▽ 1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100				
J_{W1L}	0	112	262	385	480	563	605	617	625	622	637	680	675	692	688	660	635	445	210	0	3.55	.073	12.1	.050
J_{W1}	922	922	932	940	942	935	932	922	912	907	895	862	837	800	750	682	632	340	140	0	.098	3.74	20.7	.053
J_{S1}	925	925	927	930	932	932	935	930	930	930	920	892	875	852	810	752	657	430	151	002	-.083	5.44	24.2	.059
T_{S1}	537	537	570	597	602	645	662	652	660	652	640	630	632	627	607	565	490	277	097	002	4.98	.069	8.23	.055
J_{A2L}	0	210	550	688	757	820	822	817	797	780	767	737	730	692	657	575	537	375	287	145	56.8	.655	2.0×10^6	.924
J_{M2L}	0	222	530	657	710	757	767	767	757	752	730	697	682	660	615	537	482	342	272	150	630	.756	2.6×10^7	.944
J_{A2}	977	980	980	950	932	830	790	777	760	742	705	617	577	502	450	377	315	185	102	035	4.42	.445	40.8	.437
F	0	027	140	215	270	315	337	325	335	300	295	235	207	177	157	137	127	085	032	010	7.44	.407	9.85	.407
F_{BA}	0	120	382	507	550	582	565	570	580	585	560	515	487	447	402	350	332	225	160	077	7.44	.407	120	.483
B_5	0	057	530	695	790	857	890	900	900	897	872	850	822	792	760	672	620	447	242	040	1.53	.063	21.0	.117
B_{10}	0	317	700	817	857	897	900	897	897	892	880	852	840	832	800	747	715	617	485	262	1.78	.083	1180	.663

$\gamma_1(f) = 66.3$ $\gamma_1(f^*) = 13.6$ $R = 1.0$ SAMPLED DISTRIBUTIONS: $p(X) = \text{Gamma}(25,1)$

$\gamma_2(f) = 3.6 \times 10^6$ $\gamma_2(f^*) = 597$ $f(X) = ((24!)/(12!))X^{-6}$ $M = 400$ $n = 30$[b] $q(X) = \text{Gamma}(13,1)$

[a]Coverages are represented as being out of a total of 1000; the nominal confidence level is 90%.
▽Indicates the position of R=1.
[b]Samples from p(x) and q(x) are of equal size.

TABLE XVIII

Data For Sampling Experiment XVIII

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (C×R)[a]																				MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH
	C=0	.01	0.1	0.2	0.3	0.5	0.7	0.8	0.9	1.0 ▼	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100				
J_{W1L}	0	0	0	002	097	575	777	765	717	662	520	337	287	230	190	132	097	047	020	002	1.07	.047	2.82	.183
J_{W1}	0	0	002	032	257	762	800	722	637	575	400	222	170	087	072	032	015	002	0	0	.920	.032	1.02	.052
J_{S1}	0	0	005	077	287	745	800	750	672	607	417	222	175	117	072	030	020	002	0	0	.949	.034	1.08	.056
T_{S1}	0	0	0	032	210	675	762	705	632	557	372	172	132	070	032	015	005	0	0	0	.872	.029	.851	.041
J_{A2L}	0	442	760	692	652	602	557	535	512	505	475	432	422	402	375	332	307	245	210	142	327	.937	5.2×10^{8}	.993
J_{M2L}	0	345	712	675	657	605	577	557	540	527	495	460	445	427	405	377	360	282	237	170	14600	.974	2.4×10^{10}	.996
J_{A2}	865	955	685	505	432	332	272	242	227	217	185	162	155	135	125	107	092	047	027	010	7.97	.784	30.1	.777
F	0	125	277	255	222	150	100	095	087	092	100	095	085	080	072	052	037	017	012	005	8.52	.779	11.3	.779
F_{BA}	0	205	497	527	477	442	400	380	370	367	340	305	272	242	210	195	188	137	100	027	8.52	.779	160	.793
B_5	0	907	542	367	267	192	135	117	107	085	070	040	032	025	022	010	010	0	0	0	.029	.101	.384	.113
B_{10}	0	900	867	767	722	637	587	557	550	535	487	440	412	380	345	305	270	210	145	050	.085	.106	54.1	.440

$\gamma_1(f) = 310$ $\gamma_1(f^*) = 0.8$ $R = 217.71$ SAMPLED DISTRIBUTIONS: $p(X) = \text{Gamma}(3,0.9)$

$\gamma_2(f) = 1.2\times10^5$ $\gamma_2(f^*) = 1.0$ $f(X) = (1/60)X^4\exp(X/9)$ $M = 400$ $n = 30$[b] $q(X) = \text{Gamma}(6,1)$

[a]Coverages are represented as being out of a total of 1000; the nominal confidence level is 90%.

▼Indicates the position of R=1.

[b]Samples from p(x) and q(x) are of equal size.

TABLE XIX

The One-Sample Results for $f^* \in L$[a]; Sampling Experiments I, II, XV and XVIII

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (C×R)																				MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH	EXP. # & D.CLASS.
	C=0	.01	0.1	0.2	0.3	0.5	0.7	0.8	0.9	1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100					
J_{W1L}	0	0	0	0	0	002	247	785	972	720	477	175	100	040	015	0	0	0	0	0	.997	.009	.636	.039	
J_{W1}	015	015	030	037	077	165	570	885	995	780	445	112	060	022	005	0	0	0	0	0	.968	.008	.673	.031	I
J_{S1}	755	760	772	787	807	857	900	915	930	945	960	982	980	957	785	265	055	0	0	0	1.07	.060	5.07	.019	(H-L)
T_{S1}	755	760	772	782	813	855	890	907	917	922	947	940	907	732	510	205	045	0	0	0	.842	.039	4.63	.015	
J_{W1L}	0	0	0	0	0	002	165	797	945	740	485	232	135	055	027	005	0	0	0	0	1.02	.011	.677	.042	
J_{W1}	010	012	017	042	047	102	607	915	972	747	445	152	077	032	012	0	0	0	0	0	.994	.010	.703	.041	II
J_{S1}	712	720	742	770	795	870	905	932	942	952	980	975	910	620	325	077	015	0	0	0	1.03	.047	3.60	.014	(H-L)
T_{S1}	707	710	740	757	782	847	887	902	922	932	938	865	765	505	285	072	015	0	0	0	.944	.058	3.40	.022	
J_{W1L}	0	0	0	0	0	063	582	760	897	897	702	247	140	052	017	0	0	0	0	0	1.03	.012	.856	.023	
J_{W1}	005	005	010	012	027	292	815	930	940	887	630	175	087	015	005	0	0	0	0	0	.999	.012	.850	.021	XV
J_{S1}	0	002	005	007	027	345	825	927	955	938	712	202	063	005	0	0	0	0	0	0	1.01	.011	.889	.013	(H-L)
T_{S1}	0	0	0	0	012	205	628	857	925	917	720	217	065	005	0	0	0	0	0	0	1.05	.012	.834	.012	
J_{W1L}	0	0	0	002	097	575	777	765	717	662	520	337	287	230	190	132	097	047	020	002	1.07	.047	2.82	.183	
J_{W1}	0	0	002	032	257	762	800	722	637	575	400	222	170	037	072	032	015	002	0	0	.920	.032	1.02	.052	XVIII
J_{S1}	0	0	005	077	287	745	800	750	672	607	417	222	175	117	072	030	020	002	0	0	.949	.034	1.08	.056	(H-L)
T_{S1}	0	0	0	032	210	675	762	705	632	557	372	172	132	070	032	015	005	0	0	0	.872	.029	.851	.041	

[a]These results were obtained from Tables I, II, XV and XVIII.

TABLE XX

The One-Sample Results for $f^* \in L$[a]; Sampling Experiments III, IV, VII, VIII and XVI

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (C×R) C=0	.01	0.1	0.2	0.3	0.5	0.7	0.8	0.9	1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100	MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH	EXP. # & D.CLASS.
J_{W1L}	0	0	0	012	020	090	390	757	982	780	507	267	172	095	063	032	017	0	0	0	1.01	.012	.890	.056	IV (L-L)
J_{W1}	137	140	160	182	237	342	590	925	997	790	490	202	127	067	037	015	002	0	0	0	.933	.013	1.04	.063	
J_{S1}	895	895	897	897	897	905	907	912	917	917	925	935	935	940	945	965	975	982	902	007	2.53	.255	52.5	.034	
T_{S1}	897	897	900	900	905	907	912	915	915	915	920	922	925	930	933	947	950	922	310	0	-2.61	.278	40.8	.023	
J_{W1L}	0	0	005	015	032	160	787	892	875	777	680	460	380	237	140	057	025	0	0	0	1.08	.019	1.33	.039	XVI (M-L)
J_{W1}	177	180	235	305	407	705	967	990	935	830	675	385	290	142	080	025	010	0	0	0	.972	.020	1.55	.043	
J_{S1}	845	847	867	890	910	920	917	912	917	910	860	752	662	515	352	172	072	002	0	0	1.00	.050	3.58	.020	
T_{S1}	680	632	722	752	792	857	875	872	875	877	832	725	637	495	345	170	070	002	0	0	1.29	.041	2.94	.014	
J_{W1L}	0	0	002	020	032	132	472	797	980	792	517	297	220	152	097	050	032	002	0	0	1.02	.015	1.06	.057	VIII (L-L)
J_{W1}	180	185	212	262	313	415	632	967	1000	800	510	237	180	092	057	030	012	0	0	0	.916	.018	1.29	.064	
J_{S1}	862	862	862	860	857	855	852	852	850	850	845	842	837	830	813	785	772	682	477	042	-1.04	.821	58.5	.026	
T_{S1}	855	855	852	852	850	847	845	837	837	835	835	827	827	817	815	782	755	655	445	035	5.30	.190	43.2	.029	
J_{W1L}	0	0	0	0	0	0	007	047	767	563	147	032	017	012	010	002	002	0	0	0	.984	.009	.246	.101	III (H-L)
J_{W1}	0	0	0	0	0	0	010	055	782	537	130	017	012	005	002	0	0	0	0	0	.977	.007	.207	.055	
J_{S1}	313	317	375	442	512	642	787	830	837	942	992	632	265	087	050	025	012	0	0	0	1.02	.027	1.92	.034	
T_{S1}	302	305	372	430	512	650	795	842	892	940	920	207	060	005	0	0	0	0	0	0	.816	.037	1.48	.025	
J_{W1L}	0	0	0	010	012	080	520	732	880	780	515	282	202	087	052	025	010	0	0	0	1.02	.015	.877	.047	VII (L-L)
J_{W1}	110	120	137	172	232	402	697	813	917	785	502	230	155	067	037	007	002	0	0	0	.964	.015	1.02	.050	
J_{S1}	952	952	962	962	962	967	947	938	932	912	865	750	705	612	522	327	212	025	002	0	.877	.076	5.55	.036	
T_{S1}	782	792	820	847	865	887	877	877	870	852	807	705	645	570	467	282	190	020	0	0	1.60	.144	3.56	.029	

[a]These results were obtained from Tables III, IV, VII, VIII and XVI.

TABLE XXI

The One-Sample Results for $f \in M$; Sampling Experiments V, IX, XIII and XVII[a]

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (C×R)																				MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN. LENGTH	EXP. # & D.CLASS.
	C=0	.01	0.1	0.2	0.3	0.5	0.7	0.8	0.9	1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100					
J_{W1L}	0	0	022	055	122	410	677	922	925	815	650	440	387	297	232	160	112	015	0	0	1.14	.037	1.93	.053	V (L-M)
J_{W1}	490	492	527	587	627	697	940	992	930	867	675	417	340	242	170	107	070	010	0	0	.881	.048	2.56	.053	
J_{S1}	957	957	955	952	950	942	930	917	905	890	880	847	830	777	742	652	575	340	077	0	.931	.249	15.5	.041	
T_{S1}	862	862	870	872	872	865	875	870	860	855	832	787	762	710	672	587	502	277	063	0	3.27	.069	9.20	.032	
J_{W1L}	0	0	027	082	142	320	785	857	842	807	727	592	515	402	320	192	100	005	0	0	1.20	.029	2.06	.039	IX (L-M)
J_{W1}	502	517	572	635	690	890	990	992	950	897	770	540	442	327	235	100	042	0	0	0	.953	.037	2.62	.040	
J_{S1}	902	902	905	917	925	927	915	912	912	907	877	805	765	680	575	442	290	045	0	0	1.15	.076	6.00	.031	
T_{S1}	707	710	737	765	735	805	832	832	842	842	810	745	712	632	545	405	252	035	0	0	1.88	.047	4.07	.023	
J_{W1L}	0	030	140	222	317	532	620	667	705	722	745	730	710	672	615	517	402	117	010	0	1.78	.044	4.58	.042	XIII (L-M)
J_{W1}	720	732	777	825	852	915	932	927	912	897	850	775	725	640	552	405	302	060	002	0	.896	.119	6.61	.051	
J_{S1}	955	955	960	962	960	952	945	947	945	940	905	842	797	725	662	552	450	167	037	0	.743	.225	10.8	.055	
T_{S1}	725	725	752	775	790	802	815	813	820	825	790	730	700	620	575	462	375	140	020	0	2.60	.051	5.23	.040	
J_{W1L}	0	112	262	385	480	563	605	617	625	622	637	600	675	692	688	660	635	445	210	0	3.55	.073	12.1	.050	XVII (H-M)
J_{W1}	922	922	932	940	942	935	932	922	912	907	895	862	837	800	750	682	632	340	140	0	.098	3.74	20.7	.053	
J_{S1}	925	925	927	930	932	932	935	930	930	930	920	892	875	852	810	752	657	430	151	002	-.083	5.44	24.2	.059	
T_{S1}	537	537	570	597	602	645	662	652	660	652	640	630	632	627	607	565	490	277	097	002	4.98	.069	8.23	.055	

[a]These results were obtained from Tables V, IX, XIII and XVII.

TABLE XXII

The One-Sample Results for $f \in H$[a]; Sampling Experiments VI, X, XI, XII and XIV

ESTIMATOR	C=0	.01	0.1	0.2	0.3	0.5	0.7	0.8	0.9	1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100	MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH	EXP. # & D.CLASS.
	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (C×R)																								
J_{W1L}	0	040	125	200	252	367	460	512	592	635	725	727	642	527	400	222	117	0	0	0	1.48	.033	2.18	.030	XII (L-H)
J_{W1}	460	465	492	530	560	617	635	727	750	772	820	735	630	472	335	157	082	0	0	0	1.07	.065	3.46	.043	
J_{S1}	690	690	727	762	792	820	850	865	860	870	837	772	720	620	547	355	260	050	007	0	1.06	.099	5.99	.051	
T_{S1}	167	172	180	272	357	490	557	592	605	605	607	585	557	515	422	277	200	036	005	0	2.09	.044	2.77	.042	
J_{W1L}	0	005	042	100	160	400	695	882	895	813	697	570	517	412	340	225	170	047	010	0	1.30	.045	2.71	.063	VI (L-H)
J_{W1}	575	585	627	665	697	775	970	997	980	887	737	563	465	345	260	160	110	017	007	0	.864	.072	3.76	.063	
J_{S1}	950	950	952	942	938	912	835	880	877	872	862	825	807	752	707	655	570	265	077	0	1.21	.174	13.6	.044	
T_{S1}	800	795	800	802	807	805	782	775	780	772	757	720	710	660	617	585	507	250	063	0	3.24	.064	7.95	.036	
J_{W1L}	0	017	080	135	177	207	450	517	582	652	710	670	587	430	307	135	077	005	002	0	1.44	.031	1.94	.052	XIV (M-H)
J_{W1}	377	380	410	425	470	552	647	705	752	795	802	655	555	382	240	105	040	0	0	0	1.17	.045	2.62	.039	
J_{S1}	340	340	370	395	427	517	632	680	707	735	767	662	567	400	247	080	042	0	0	0	1.21	.043	2.49	.044	
T_{S1}	002	002	002	002	005	037	200	277	350	417	540	542	457	287	180	055	032	0	0	0	1.64	.024	1.11	.030	
J_{W1L}	0	100	247	307	362	472	515	522	535	552	577	637	662	670	690	712	702	602	297	002	4.98	.066	15.5	.045	X (M-H)
J_{W1}	880	880	895	907	910	922	922	915	915	910	902	907	895	872	845	807	745	537	212	0	.657	.800	26.5	.056	
J_{S1}	865	865	872	880	890	905	907	907	907	910	907	900	895	867	847	817	752	512	225	005	1.69	.351	27.8	.052	
T_{S1}	460	460	490	512	525	547	565	567	577	597	605	627	645	635	615	597	557	370	160	005	7.11	.071	10.7	.050	
J_{W1L}	0	070	205	292	340	470	537	555	582	600	607	622	622	620	610	627	615	637	637	245	14.7	.091	97.0	.162	XI (L-H)
J_{W1}	942	942	942	942	942	940	930	927	925	922	917	897	892	877	860	842	815	705	560	085	-1.72	1.01	84.1	.058	
J_{S1}	930	930	932	932	932	932	932	932	932	932	932	932	932	925	922	917	917	877	770	255	-.759	4.61	162	.061	
T_{S1}	715	715	717	717	717	720	722	722	727	727	732	732	742	747	752	762	772	810	827	145	30.5	.072	68.2	.020	

[a] These results were obtained from Tables VI, X, XI, XII and XIV.

TABLE XXIII

The Two-Sample Results for the L-L Classification[a]; Sampling Experiments IV, VII and VIII

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (C×R) C=0	.01	0.1	0.2	0.3	0.5	0.7	0.8	0.9	1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100	MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH	EXP. # & D.CLASS.
J_{A2L}	0	002	035	255	432	675	795	822	842	832	813	697	642	540	452	340	255	072	032	005	1.32	.048	5.44	.173	
J_{M2L}	0	005	070	257	422	640	770	810	830	827	810	695	635	542	460	342	260	087	035	005	1.41	.053	6.57	.226	
J_{A2}	597	617	727	832	897	925	902	882	855	805	710	535	480	367	265	152	092	017	0	0	1.14	.043	2.81	.049	IV (L-L)
F	0	0	012	097	220	442	542	560	595	605	565	465	405	332	235	132	075	012	0	0	1.25	.041	1.66	.041	
F_{BA}	0	0	060	215	350	577	662	715	747	747	707	602	547	465	392	262	180	052	017	0	1.25	.041	3.25	.080	
B_5	0	0	140	407	570	792	882	912	907	907	890	795	740	642	552	420	292	037	012	0	1.14	.035	4.45	.057	
B_{10}	0	010	305	540	632	837	835	902	905	910	900	835	797	747	685	592	515	287	130	005	1.19	.040	10.2	.026	
J_{A2L}	0	005	100	272	417	672	765	800	810	840	832	750	700	582	500	360	247	060	017	002	1.29	.040	4.11	.105	
J_{M2L}	0	020	105	280	415	665	762	792	807	837	820	735	672	582	492	332	250	060	020	002	1.32	.042	4.44	.148	
J_{A2}	630	637	722	797	862	917	905	892	882	867	787	582	515	397	290	167	102	005	0	0	1.12	.043	2.94	.038	VIII (L-L)
F	0	0	012	072	175	442	585	597	612	617	610	467	415	315	260	132	085	007	0	0	1.28	.038	1.69	.038	
F_{BA}	0	0	052	180	335	547	675	707	732	747	715	607	577	477	385	280	215	060	012	0	1.38	.043	3.09	.065	
B_5	0	0	087	292	495	740	845	890	895	907	912	832	792	717	637	487	407	140	035	0	1.39	.038	5.59	.061	
B_{10}	0	005	287	515	652	797	897	900	910	910	902	865	837	772	717	622	540	305	145	010	1.36	.040	12.6	.111	
J_{A2L}	0	0	070	192	290	502	705	755	807	827	857	813	750	622	480	257	147	015	0	0	1.30	.030	2.71	.030	
J_{M2L}	0	002	087	207	302	527	715	757	807	827	852	792	720	577	430	227	107	007	0	0	1.25	.031	2.53	.029	
J_{A2}	542	547	612	700	770	852	910	917	930	920	880	722	645	465	317	140	063	005	0	0	1.16	.034	3.02	.034	VII (L-L)
F	0	0	0	022	112	317	520	595	657	692	715	637	575	432	287	140	063	002	0	0	1.36	.029	1.80	.029	
F_{BA}	0	0	045	150	277	522	680	750	772	802	775	675	605	455	332	152	077	007	0	0	1.34	.033	2.14	.029	
B_5	0	0	012	072	202	490	637	705	767	800	880	870	862	827	785	630	542	197	042	0	2.03	.035	6.48	.054	
B_{10}	0	0	112	305	445	657	785	822	847	862	867	852	837	777	732	642	532	332	130	010	1.65	.032	10.3	.088	

[a]These results were obtained from Tables IV, VII and VIII.

TABLE XXIV

The Two-Sample Results for the M-L Classification[a]; Sampling Experiments I, XV and XVI

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (C×R) C=0	.01	0.1	0.2	0.3	0.5	0.7	0.8	0.9	1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100	MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH	EXP. # & D.CLASS.
J_{A2L}	0	0	125	447	688	800	807	797	772	737	642	535	507	438	375	300	272	165	122	050	2.56	.452	3150	.978	XV (M-L)
J_{M2L}	0	0	075	382	612	770	790	782	770	732	665	557	517	452	400	325	280	188	140	055	9.87	.784	20900	.996	
J_{A2}	337	362	645	855	907	862	747	685	622	565	462	325	277	227	177	112	090	025	007	002	1.17	.152	2.84	.212	
F	0	0	032	225	427	587	550	517	495	455	390	262	247	185	142	080	063	020	002	002	1.20	.151	1.59	.151	
F_{BA}	0	0	050	300	502	685	700	695	638	665	592	487	447	372	322	265	210	120	055	017	1.20	.151	9.97	.318	
B_5	0	005	515	802	877	857	782	717	652	605	512	360	320	220	177	110	067	010	002	0	.467	.043	1.83	.059	
B_{10}	0	047	540	762	850	890	860	817	802	760	685	612	580	520	450	375	317	177	080	015	.653	.044	8.55	.142	
J_{A2L}	0	0	095	430	630	817	835	822	787	770	672	560	505	447	395	330	277	160	102	027	1.57	.126	58.4	.425	I (M-L)
J_{M2L}	0	0	057	357	595	790	817	805	782	765	680	570	525	457	420	347	292	182	112	040	2.46	.204	149	.467	
J_{A2}	430	447	727	895	935	900	792	727	672	612	515	385	340	255	192	132	087	022	010	0	1.10	.075	2.62	.102	
F	0	0	017	182	380	552	575	547	515	475	432	337	280	210	165	092	060	012	007	0	1.15	.075	1.52	.075	
F_{BA}	0	0	032	245	470	657	712	707	697	667	617	492	452	395	340	272	210	120	055	012	1.15	.075	7.36	.203	
B_5	0	005	370	707	865	917	850	807	782	730	630	465	422	327	262	182	112	022	002	0	.607	.038	2.40	.057	
B_{10}	0	020	467	752	867	922	912	892	870	850	802	695	655	592	532	438	370	212	102	015	.767	.039	9.47	.123	
J_{A2L}	0	0	040	222	462	765	857	840	837	800	722	597	520	442	355	262	185	075	047	010	1.18	.054	5.29	.158	XVI (M-L)
J_{M2L}	0	0	035	220	435	745	845	827	822	795	722	597	522	436	367	277	205	082	055	012	1.31	.065	6.69	.193	
J_{A2}	410	430	607	790	887	930	860	835	775	720	600	402	350	240	172	102	057	022	0	0	1.01	.047	2.18	.059	
F	0	0	017	110	255	547	652	647	655	635	565	382	322	195	152	072	045	010	0	0	1.08	.046	1.43	.046	
F_{BA}	0	0	025	165	337	655	760	730	767	737	662	522	477	365	295	202	142	057	027	0	1.08	.046	3.12	.098	
B_5	0	0	135	452	692	905	930	917	880	857	725	642	552	415	317	202	120	017	0	0	.818	.032	2.65	.045	
B_{10}	0	005	330	585	740	870	880	867	875	870	825	730	692	627	565	462	377	177	063	005	.912	.040	9.63	.291	

[a]These results were obtained from Tables I, XV and XVI.

TABLE XXV

The Two-Sample Results for the H-L Classification; Sampling Experiments II, III and XVIII

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (C×R) C=0	.01	0.1	0.2	0.3	0.5	0.7	0.8	0.9	1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100	MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH	EXP. # & D.CLASS.
J_{A2L}	0	020	270	522	712	845	847	852	835	822	757	662	627	540	462	372	310	202	115	047	1.50	.035	28.2	.276	
J_{M2L}	0	025	260	487	667	825	827	832	825	800	740	660	612	547	472	330	317	217	135	052	2.00	.121	56.2	.343	
J_{A2}	332	855	922	962	957	915	837	807	755	707	632	445	382	313	255	195	150	037	010	0	1.07	.069	3.44	.067	III
F	0	0	067	215	337	500	482	490	467	442	405	295	277	240	185	127	080	025	002	0	1.22	.063	1.61	.063	(H-L)
F_{BA}	0	007	157	362	517	642	730	727	707	682	635	542	487	410	360	292	245	145	020	010	1.22	.063	7.07	.154	
B_5	0	002	365	670	820	907	915	917	895	870	797	660	612	522	447	347	277	115	040	0	.878	.052	4.51	.067	
B_{10}	0	090	550	762	850	905	910	897	890	890	872	800	767	722	672	597	522	360	235	022	.979	.055	25.2	.250	
J_{A2L}	0	442	760	692	652	602	557	535	512	505	475	432	422	402	375	332	307	245	210	142	327	.937	$5.2 \cdot 10^3$	.993	
J_{M2L}	0	345	712	675	657	605	577	557	540	527	495	460	445	427	405	377	360	282	237	170	14600	.974	$2.4 \cdot 10^{10}$	.995	
J_{A2}	865	955	685	505	432	332	272	242	227	217	185	162	155	135	125	107	092	047	027	010	7.97	.784	30.1	.777	XVIII
F	0	125	277	255	222	150	100	095	087	092	100	095	085	080	072	052	037	017	012	005	8.52	.779	11.3	.779	(H-L)
F_{BA}	0	205	497	527	477	442	400	380	370	367	340	305	272	242	210	195	188	137	100	027	8.52	.779	160	.793	
B_5	0	907	542	367	267	192	135	117	107	085	070	040	032	025	022	010	010	0	0	0	.029	.101	.304	.113	
B_{10}	0	900	867	767	722	637	587	557	550	535	487	440	412	380	345	305	270	210	145	050	.085	.106	54.1	.440	
J_{A2L}	0	0	115	438	677	777	765	732	700	667	602	515	472	410	355	315	255	155	092	027	1.51	.132	57.9	.506	
J_{M2L}	0	0	092	390	620	735	750	722	695	667	605	520	485	435	375	332	285	180	110	037	2.37	.221	148	.566	
J_{A2}	342	352	680	837	920	815	697	627	590	552	467	332	290	220	177	127	080	032	007	0	1.06	.079	2.44	.103	II
F	0	0	037	262	430	520	512	502	457	435	392	270	220	175	142	077	063	017	007	0	1.10	.077	1.45	.077	(H-L)
F_{BA}	0	0	052	297	530	647	660	635	617	587	547	460	430	355	315	255	200	110	055	012	1.10	.077	6.71	.207	
B_5	0	002	477	762	865	870	762	730	677	617	535	380	313	235	177	115	090	022	005	0	.505	.041	2.03	.067	
B_{10}	0	025	575	807	872	890	865	840	817	802	740	632	582	527	470	365	305	155	102	015	.666	.046	13.8	.323	

[a] These results were obtained from Tables II, III and XVIII.

TABLE XXVI

The Two-Sample Results for the L-M Classification[a]; Sampling Experiments V, IX and XIII

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (C×R)																				MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH	EXP. # & D.CLASS.
	C=0	.01	0.1	0.2	0.3	0.5	0.7	0.8	0.9	1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100					
J_{A2L}	0	020	137	235	332	525	617	662	720	750	787	800	802	720	622	505	380	085	005	0	1.76	.043	4.33	.037	V (L-M)
J_{M2L}	0	035	147	265	355	532	620	677	720	750	785	782	770	677	577	450	330	060	005	0	1.70	.045	3.92	.036	
J_{A2}	652	657	705	760	815	857	870	890	897	895	865	780	717	597	530	352	260	030	0	0	1.49	.052	4.71	.040	
F	0	0	005	045	122	292	430	472	472	515	575	527	505	465	405	280	205	030	0	0	1.83	.041	2.43	.041	
F_{BA}	0	012	120	210	320	452	590	617	647	667	707	690	657	592	510	392	245	032	0	0	1.74	.037	3.05	.033	
B_5	0	0	022	112	232	422	557	630	680	730	807	920	917	910	887	813	707	370	107	005	2.56	.034	9.85	.058	
B_{10}	0	015	212	370	495	655	767	802	822	850	892	920	922	900	892	815	752	467	262	020	2.06	.033	21.1	.208	
J_{A2L}	0	010	075	185	322	520	680	722	757	800	840	813	765	675	537	335	222	040	010	0	1.48	.038	3.50	.049	IX (L-M)
J_{M2L}	0	012	082	205	322	522	630	722	755	795	827	797	745	635	500	317	220	035	012	0	1.47	.040	3.46	.057	
J_{A2}	577	587	675	742	785	867	912	902	897	900	862	737	637	475	342	190	112	015	0	0	1.32	.041	3.21	.034	
F	0	0	010	045	115	322	492	565	617	635	647	580	547	430	315	190	110	017	0	0	1.50	.036	1.98	.036	
F_{BA}	0	005	042	155	247	472	630	657	702	732	777	710	685	547	457	307	155	017	005	0	1.47	.033	2.71	.041	
B_5	0	0	017	120	280	552	732	785	830	857	917	925	905	845	742	595	452	135	012	0	1.68	.029	5.19	.038	
B_{10}	0	005	177	377	522	692	785	825	857	870	887	875	870	837	797	672	585	330	130	005	1.59	.035	10.9	.070	
J_{A2L}	0	035	207	355	472	585	682	720	752	772	805	820	822	777	730	645	570	280	132	017	2.32	.069	25.5	.569	XIII (L-M)
J_{M2L}	0	045	227	365	472	605	685	727	747	762	802	785	782	742	705	617	507	247	102	020	2.45	.103	52.0	.790	
J_{A2}	875	875	900	925	932	947	940	930	932	920	892	797	747	685	605	472	382	152	040	0	1.61	.031	8.57	.059	
F	0	0	020	090	142	287	375	410	438	435	438	442	445	400	372	313	250	075	020	0	2.30	.059	3.05	.059	
F_{BA}	0	010	135	265	402	517	587	620	632	650	660	645	635	600	565	457	357	155	063	005	2.30	.059	6.69	.107	
B_5	0	0	037	145	282	480	580	647	677	720	800	830	907	915	915	877	837	590	325	032	3.98	.055	22.9	.076	
B_{10}	0	030	305	462	560	700	772	805	817	827	862	907	902	900	892	867	817	627	442	142	3.07	.055	108	.318	

[a]These results were obtained from Tables V, IX and XIII.

TABLE XXVII

The Two-Sample Results for the L-H Classification[a]; Sampling Experiments VI, XI and XII

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (C×R) C=0	.01	0.1	0.2	0.3	0.5	0.7	0.8	0.9	1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100	MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH	EXP. # & D.CLASS.
$J_{\lambda 2L}$	0	032	250	352	438	590	662	712	732	762	787	815	817	792	757	657	577	335	142	012	2.58	.076	12.3	.091	XII (L-H)
$J_{\mu 2L}$	0	100	257	362	447	592	675	705	720	740	770	780	775	735	695	600	520	285	120	017	2.59	.081	11.3	.110	
$J_{\lambda 2}$	872	875	907	922	930	942	942	940	927	922	907	817	795	682	617	500	430	167	052	0	1.81	.093	9.90	.065	
F	0	005	050	095	140	245	350	367	392	415	435	455	420	375	365	305	255	092	032	0	2.62	.072	3.47	.072	
F_{BA}	0	035	188	262	342	462	563	587	605	622	637	637	615	567	510	462	365	185	035	002	2.62	.072	7.18	.096	
B_5	0	0	067	170	310	512	620	662	690	732	787	862	887	890	835	862	840	640	345	047	4.03	.052	25.8	.074	
B_{10}	0	077	352	527	615	740	825	852	860	877	902	910	902	902	880	845	807	638	515	150	2.84	.064	67.7	.159	
$J_{\lambda 2L}$	0	142	252	310	337	380	425	435	455	465	512	563	577	597	632	657	688	742	730	465	39.4	.146	309	.100	XI (L-H)
$J_{\mu 2L}$	0	170	235	337	357	425	452	477	487	505	535	580	597	605	630	672	635	717	665	355	36.2	.158	174	.101	
$J_{\lambda 2}$	960	960	960	960	960	960	960	957	955	957	952	955	950	942	945	940	927	862	742	365	3.24	2.13	344	.096	
F	0	0	027	047	050	070	065	077	087	090	102	130	150	188	192	210	235	265	255	135	44.0	.134	58.2	.132	
F_{BA}	0	035	127	200	210	240	292	305	313	322	347	400	405	440	460	510	517	512	465	202	44.0	.134	96.6	.127	
B_5	0	0	0	010	015	032	063	070	075	085	117	175	192	215	250	322	367	542	722	907	159	.069	2170	.109	
B_{10}	0	050	210	270	335	412	465	487	510	535	557	622	637	682	697	742	767	867	922	900	78.1	.087	19900	.213	
$J_{\lambda 2L}$	0	027	155	255	347	472	563	602	622	667	732	785	777	755	695	595	435	115	007	0	1.95	.042	4.79	.027	VI (L-H)
$J_{\mu 2L}$	0	037	180	285	357	485	572	605	630	667	722	760	735	702	657	547	390	092	005	0	1.86	.044	4.27	.035	
$J_{\lambda 2}$	657	660	700	745	780	825	872	872	870	872	867	813	775	685	602	440	322	050	002	0	1.65	.051	5.47	.041	
F	0	0	017	063	130	220	342	392	410	435	502	542	520	500	465	350	262	037	002	0	2.06	.041	2.73	.041	
F_{BA}	0	012	110	200	300	430	502	547	572	587	635	657	647	620	567	430	297	052	005	0	2.06	.041	3.47	.036	
B_5	0	0	020	090	142	313	480	535	595	617	730	822	870	890	820	832	762	477	197	002	3.42	.039	12.4	.052	
B_{10}	0	015	155	305	420	582	695	740	760	802	840	877	890	900	877	855	807	575	317	025	2.61	.037	22.4	.087	

[a]These results were obtained from Tables VI, XI and XII.

TABLE XXVIII

The Two-Sample Results for the H-M and M-H Classifications[a]; Sampling Experiments X, XIV and XVII

ESTIMATOR	COVERAGE OF MULTIPLES OF THE VARIANCE RATIO (c×R)																				MEAN EST. ÷ R	CV. of MEAN EST.	MEAN LENGTH ÷ R	CV. of MEAN LENGTH	EXP. # & D.CLASS.
	c=0	.01	0.1	0.2	0.3	0.5	0.7	0.8	0.9	1.0	1.25	1.75	2.0	2.5	3.0	4.0	5.0	10	20	100					
$J_{\lambda 2L}$	0	210	550	680	757	820	822	817	797	780	767	737	730	692	657	575	537	375	237	145	56.8	.655	2.0×10^6	.924	XVII (H-M)
J_{M2L}	0	222	530	657	710	757	767	767	757	752	730	697	682	660	615	537	482	342	272	150	630	.756	2.6×10^7	.944	
$J_{\lambda 2}$	977	960	960	950	932	830	790	777	760	742	705	617	577	502	450	377	315	185	102	035	4.42	.445	40.8	.437	
F	0	027	140	215	270	315	337	325	335	300	295	235	207	177	157	137	127	085	032	010	7.44	.407	9.85	.407	
F_{BA}	0	120	382	507	550	592	565	570	580	585	560	515	487	447	402	350	332	225	160	077	7.44	.407	120	.483	
B_5	0	057	530	695	790	857	890	900	900	897	872	850	822	792	760	672	620	447	242	040	1.53	.063	21.0	.117	
B_{10}	0	317	700	817	857	897	900	897	897	892	880	852	840	832	800	747	715	617	485	262	1.78	.083	1180	.663	
$J_{\lambda 2L}$	0	145	375	537	637	730	755	767	787	780	817	817	820	800	770	727	688	550	417	210	15.3	.350	3.9×10^4	.922	X (M-H)
J_{M2L}	0	150	342	482	590	682	725	730	742	752	767	765	757	737	720	680	657	530	405	222	66.6	.663	3.2×10^5	.962	
$J_{\lambda 2}$	967	967	970	970	960	940	925	917	897	875	840	767	737	675	640	552	492	325	175	045	4.40	.210	31.0	.121	
F	0	010	085	127	150	207	262	295	290	300	325	327	315	292	265	240	215	140	075	027	6.61	.149	8.75	.149	
F_{BA}	0	060	257	362	410	475	517	535	547	557	567	577	575	592	580	525	490	410	295	100	8.30	.255	119	.340	
B_5	0	040	447	620	730	822	865	872	887	897	900	870	857	825	815	752	695	530	325	057	2.38	.077	33.1	.148	
B_{10}	0	262	617	715	782	840	865	880	882	892	897	902	877	895	872	847	817	700	572	317	3.73	.129	995	.392	
$J_{\lambda 2L}$	0	067	287	462	577	707	765	787	790	802	800	782	770	735	690	612	540	362	210	077	4.27	.230	2010	.836	XIV (M-H)
J_{M2L}	0	080	292	462	557	670	745	757	772	775	782	747	737	697	662	582	515	345	207	085	9.03	.447	9820	.940	
$J_{\lambda 2}$	912	917	952	962	960	940	915	892	865	847	790	717	667	590	502	415	347	165	077	007	1.93	.124	12.4	.152	
F	0	002	070	137	207	300	375	382	395	397	397	367	357	315	265	222	192	120	032	005	2.86	.122	3.79	.122	
F_{BA}	0	030	183	347	432	530	580	595	612	620	630	620	587	557	500	442	392	235	142	035	2.86	.122	22.5	.275	
B_5	0	007	255	465	602	752	842	857	870	872	912	837	860	855	830	750	692	447	255	010	2.35	.059	18.0	.086	
B_{10}	0	105	467	652	742	845	877	885	890	897	890	887	870	827	805	772	737	607	457	162	2.48	.081	496	.756	

[a]These results were obtained from Tables X, XIV and XVII.

TABLE XXIX

Summary of the Distributional Properties of f(X) and $f^*(Y)$

EXP.#	R	$\gamma_1(f)$	$\gamma_2(f)$	$\gamma_1(f^*)$	$\gamma_2(f^*)$	D.CLASS
I	1.0	4.6	70.8	-0.9	0.1	M-L
II	1.0	7.0	1.5×10^6	-0.5	-0.6	H-L
III	25.07	6.9	2.0×10^6	2.5	10.5	H-L
IV	1.0	2.2	7.9	-1.9	4.5	L-L
V	1.0	-0.9	0.1	4.6	70.8	L-M
VI	1.0	-0.5	-0.6	7.0	1.5×10^6	L-H
VII	24.99	-0.5	-0.7	2.5	11.0	L-L
VIII	1.0	-1.9	4.5	2.2	7.9	L-L
IX	1.0	1.1	10.0	3.5	287	L-M
X	1.0	13.6	597	66.3	3.6×10^6	M-H
XI	.0046	0.8	1.0	310	1.2×10^5	L-H
XII	8.38	1.9	6.0	8.1	4.0×10^4	L-H
XIII	6.71	1.9	6.0	7.1	337	L-M
XIV	11.34	4.8	49.0	9.9	6.0×10^6	M-H
XV	11.34	4.8	49.0	-0.5	-0.9	M-L
XVI	1.0	3.5	287	1.1	10.0	M-L
XVII	1.0	66.3	3.6×10^6	13.6	597	H-M
XVIII	217.71	310	1.2×10^5	0.8	1.0	H-L

TABLE XXX

Coverages of R[a]

EST. \ EXP. #	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII
J_{W1L}	780	740	563	780	815	813	780	792	807	552	600	635	722	652	897	777	622	662
J_{W1}	780	747	537	790	867	887	785	800	897	910	922	772	897	795	887	830	907	575
J_{S1}	945	952	942	917	890	872	912	850	907	910	932	870	940	735	938	910	930	607
T_{S1}	922	932	940	915	855	772	852	835	842	597	727	605	825	417	917	877	652	557
J_{A2L}	770	667	822	832	750	667	827	840	800	780	465	762	772	802	737	800	780	505
J_{M2L}	765	667	800	827	750	667	827	837	795	752	505	740	762	775	732	795	752	527
J_{A2}	612	552	707	805	895	872	920	867	900	875	957	922	920	847	565	720	742	217
F	475	435	442	605	515	435	692	617	635	300	090	415	435	397	455	635	300	092
F_{BA}	667	587	682	747	667	587	802	747	732	557	322	622	650	620	665	737	585	367
B_5	730	617	870	907	730	617	800	907	857	897	085	732	720	872	605	857	897	085
B_{10}	850	802	890	910	850	802	862	910	870	892	535	877	827	897	760	870	892	535
D. CLASS.	M-L	H-L	H-L	L-L	L-M	L-H	L-L	L-L	L-M	M-H	L-H	L-H	L-M	M-H	M-L	M-L	H-M	H-L

[a]These entries were obtained from Tables I through XVIII.

TABLE XXXI

Coverages of R/5[a]

EST. \ EXP. #	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII
J_{W1L}	0	0	0	012	055	100	010	020	082	307	292	200	222	135	0	015	385	002
J_{W1}	037	042	0	182	587	665	172	262	635	907	942	530	825	425	012	305	940	032
J_{S1}	787	770	442	897	952	942	962	860	917	880	932	762	962	395	007	890	930	077
T_{S1}	782	757	430	900	872	802	847	852	765	512	717	272	775	002	0	752	597	032
J_{A2L}	430	438	522	255	235	255	192	272	185	537	310	352	355	462	447	222	688	692
J_{M2L}	357	390	487	257	265	285	207	280	205	482	337	362	365	462	382	220	657	675
J_{A2}	895	887	962	832	760	745	700	797	742	970	960	922	925	962	865	790	950	505
F	132	262	215	097	045	063	022	072	045	127	047	095	090	137	225	110	215	255
F_{BA}	245	297	362	215	210	200	150	180	155	362	200	262	265	347	300	165	507	527
B_5	707	762	670	407	112	090	072	292	120	620	010	170	145	465	802	452	695	367
B_{10}	752	807	762	540	370	305	305	515	377	715	270	527	462	652	762	585	817	767
D. CLASS.	M-L	H-L	H-L	L-L	L-M	L-H	L-L	L-L	L-M	M-H	L-H	L-H	L-M	M-H	M-L	M-L	H-M	H-L

[a]These entries were obtained from Tables I through XVIII.

TABLE XXXII

Coverages of 5×R[a]

EST. \ EXP. #	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII
J_{W1L}	0	0	002	017	112	170	010	032	100	702	615	117	402	077	0	025	635	097
J_{W1}	0	0	0	002	070	110	002	012	042	745	815	082	302	040	0	010	632	015
J_{S1}	055	015	012	975	575	570	212	772	290	752	917	260	450	042	0	072	657	020
T_{S1}	045	015	0	950	502	507	190	755	252	557	772	200	375	032	0	070	490	005
J_{A2L}	277	255	310	255	380	438	147	247	222	638	638	577	570	540	272	185	537	307
J_{M2L}	292	285	317	260	330	390	107	250	220	657	685	520	507	515	280	205	482	360
J_{A2}	087	080	150	092	260	322	063	102	112	492	927	430	382	347	090	057	315	092
F	060	063	080	075	205	262	063	085	110	215	235	255	250	192	063	045	127	037
F_{BA}	210	200	245	180	245	297	077	215	155	490	517	365	357	392	210	142	332	188
B_5	112	090	277	292	707	762	542	407	452	695	367	840	837	692	067	120	620	010
B_{10}	370	305	522	515	752	807	532	540	585	817	767	807	817	737	317	377	715	270
D.CLASS.	M-L	H-L	H-L	L-L	L-M	L-H	L-L	L-L	L-M	M-H	L-H	L-H	L-M	M-H	M-L	M-L	H-M	H-L

[a]These entries were obtained from Tables I through XVIII.

TABLE XXXIII

Maximum Coverages[a]

EST. \ EXP.#	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII
J_{W1L}	972 (0.9)	945 (0.9)	767 (0.9)	982 (0.9)	925 (0.9)	895 (0.9)	880 (0.9)	930 (0.9)	857 (0.8)	712 (4.0)	637 (15)	727 (1.75)	745 (1.25)	710 (1.25)	897 (0.95)	892 (0.8)	692 (2.5)	777 (0.7)
J_{W1}	995 (0.9)	972 (0.9)	782 (0.9)	997 (0.9)	992 (0.8)	997 (0.8)	917 (0.9)	1000 (0.9)	992 (0.8)	922 (0.6)	942 (0.3)[b]	820 (1.25)	932 (0.7)	802 (1.25)	940 (0.9)	990 (0.8)	942 (0.3)	800 (0.7)
J_{S1}	932 (1.75)	930 (1.25)	992 (1.25)	982 (10)	957 (0.01)[b]	952 (0.1)	967 (0.5)	962 (0.01)[b]	927 (0.5)	910 (1.0)	932 (0.95)[c]	880 (0.9)	962 (0.2)	767 (1.25)	955 (0.9)	920 (0.5)	935 (0.7)	800 (0.7)
$\tilde{J}_{S1}$	947 (1.25)	938 (1.25)	940 (1.0)	950 (5)	875 (0.7)	807 (0.3)	837 (0.5)	855 (0.01)[b]	842 (0.95)	645 (2.0)	827 (20)	607 (1.25)	825 (1.0)	542 (1.75)	925 (0.9)	877 (1.0)	662 (0.7)	762 (0.7)
J_{A2L}	835 (0.7)	777 (0.5)	852 (0.8)	842 (0.9)	802 (2.0)	735 (1.75)	857 (1.25)	840 (1.0)	840 (1.25)	820 (2.0)	742 (10)	817 (2.0)	822 (2.0)	802 (1.0)	807 (0.7)	857 (0.7)	822 (0.7)	760 (0.1)
J_{M2L}	817 (0.7)	750 (0.7)	832 (0.8)	830 (0.9)	785 (1.25)	760 (1.75)	852 (1.25)	837 (1.0)	827 (1.25)	767 (1.25)	717 (10)	780 (1.75)	802 (1.25)	782 (1.25)	782 (0.8)	845 (0.7)	767 (0.75)	712 (0.1)
J_{A2}	935 (0.3)	920 (0.3)	967 (0.3)	925 (0.5)	897 (0.9)	872 (0.75)	930 (0.9)	917 (0.5)	900 (1.0)	970 (0.15)	960 (0.7)[b]	942 (0.6)	947 (0.5)	962 (0.2)	907 (0.3)	930 (0.5)	930 (0.01)	955 (0.01)
P	575 (0.7)	520 (0.5)	500 (0.5)	605 (1.0)	575 (1.25)	542 (1.75)	715 (1.25)	617 (1.0)	647 (1.25)	327 (1.75)	265 (10)	455 (1.75)	445 (2.0)	397 (1.13)	587 (0.5)	655 (0.9)	337 (0.7)	277 (0.1)
P_{BA}	712 (0.7)	660 (0.7)	730 (0.7)	747 (0.95)	707 (1.25)	657 (1.75)	802 (1.0)	747 (1.0)	777 (1.25)	592 (2.5)	517 (5)	637 (1.5)	660 (1.25)	630 (1.25)	700 (0.7)	780 (0.8)	585 (1.0)	527 (0.2)
B_5	917 (0.5)	870 (0.5)	917 (0.8)	912 (0.8)	920 (1.75)	890 (2.5)	880 (1.25)	912 (1.25)	925 (1.75)	900 (1.25)	907 (100)	890 (2.5)	915 (2.75)	912 (1.25)	877 (0.3)	930 (0.7)	900 (0.25)	907 (0.01)
B_{10}	922 (0.5)	890 (0.5)	910 (0.7)	910 (1.0)	922 (2.0)	900 (2.5)	867 (1.25)	910 (0.95)	837 (1.25)	902 (1.75)	922 (20)	910 (1.75)	907 (1.75)	897 (1.0)	890 (0.5)	887 (0.8)	900 (0.7)	900 (0.01)
D.CLASS.	M-L	H-L	H-L	L-L	L-M	L-H	L-L	L-L	L-M	M-H	L-H	L-H	L-M	M-H	M-L	H-L	H-M	H-L

[a]This table was constructed from Tables I through XVIII; the entries in parentheses are the points at which the maximum coverages occur.

[b]The coverages from zero to this point are the same.

[c]This value is misleading, the coverages of .1×R to 2×R are the same (932).

TABLE XXXIV

The Application of the One-Sample Estimators J_{S1}, J_{W1} and J_{W1L} to the Sukhatme Example[a]

ESTIMATOR	ESTIMATED VARIANCE RATIO	CONFIDENCE BOUNDS LOWER	CONFIDENCE BOUNDS UPPER	CONFIDENCE INTERVAL LENGTH
$K_n J_{S1}$	15.67	3.08 (0.67)	28.26 (30.67)	25.18 (30.00)
$K_n J_{W1}$	21.15	8.01 (5.49)	34.29 (36.81)	26.28 (31.32)
$K_n J_{W1L}$	22.67	12.23 (10.87)	42.01 (47.28)	29.78 (36.41)

[a]Results for 90% and 95% confidence levels are recorded; the entries in parentheses refer to the 95% confidence level.

Figure 1

A Normal Probability Plot of J_{W1} for Sampling Experiment XVIII

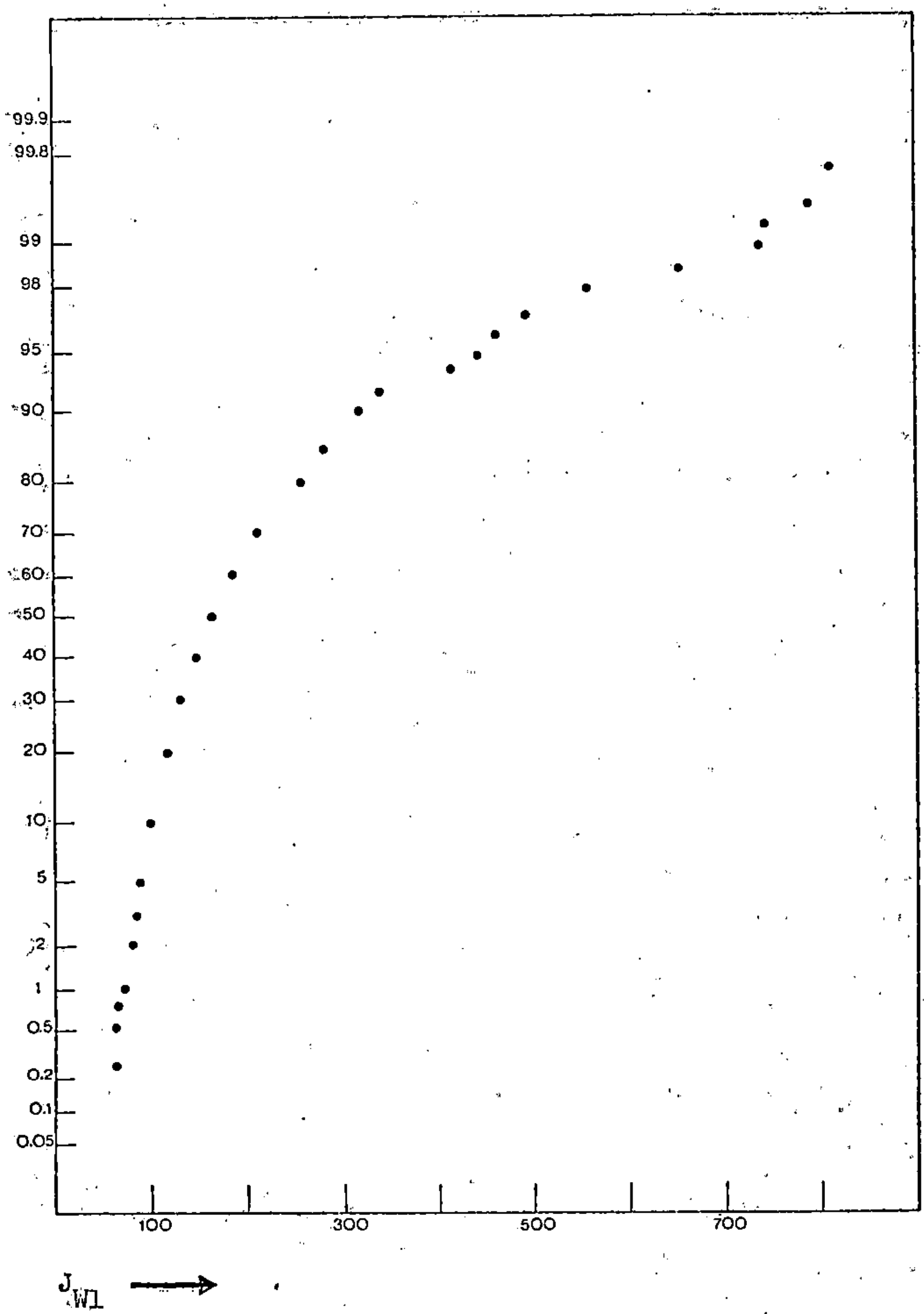
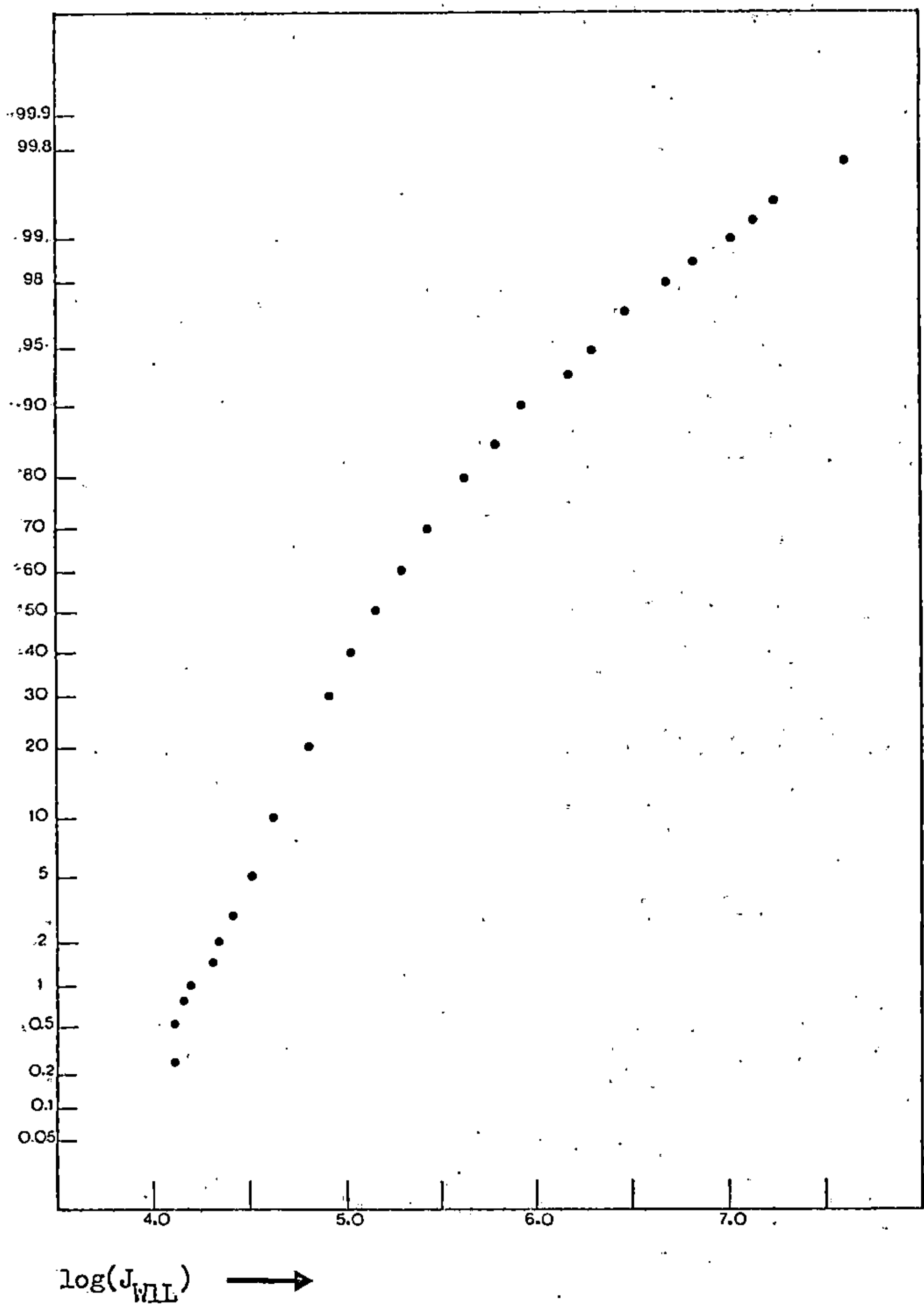


Figure 2

A Normal Probability Plot of $\log(J_{WIL})$ for Sampling Experiment XVIII



REFERENCES

[1] ARVESEN, J. N. (1969). Jackknifing U-statistics. Ann. Math. Statist., V. 40, 2076-2100.

[2] ARVESEN, J. N. & SCHMITZ, T. H. (1970). Robust procedures for variance component problems using the jackknife. Biometrics, V. 26, 677-686.

[3] BILLINGSLEY, P. (1968). Convergence of Probability Measures. Wiley, New York.

[4] BOX, G. E. P. (1953). Non-normality and test on variances. Biometrika, V. 40, 318-335.

[5] BOX, G. E. P. & ANDERSEN, S. L. (1955). Permutation theory in the derivation of robust criteria and the study of departures from assumption. J. Roy. Statist. Sec. Ser. B, V. 17, 1-26.

[6] COCHRAN, W. G. (1963). Sampling Techniques. Second edition, Wiley, New York.

[7] GRAY, H. L. & SCHUCANY, W. R. (1972). The Generalized Jackknife Statistic. Marcel Dekker, New York.

[8] HASTINGS, W. K. (1974). Variance reduction and nonnormality. Biometrika, V. 61, 143-149.

[9] HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. Ann. Math. Statist., V. 19, 293-325.

[10] JOHNSON, N. L. & KOTZ, S. (1969). Discrete Distributions. Houghton Mifflin, Boston.

[11] LEVENE, H. (1960). Robust tests for equality of variances. Contributions to Probability and Statistics (I. Olkin, et al, editors), 278-292.

[12] MILLER, R. G., Jr. (1964). A trustworthy jackknife. Ann. Math. Statist., V. 35, 1594-1605.

[13] MILLER, R. G., Jr. (1968). Jackknifing variances. Ann. Math. Statist., V. 39, 567-582.

[14] MILLER, R. G., Jr. (1974). The jackknife – a review. Biometrika, V. 61, 1-15.

[15] MOSES, L. E. (1963). Rank tests of dispersion. Ann. Math. Statist., V. 34, 973-983.

[16] PARZEN, E. (1960). Modern Probability Theory and Its Applications. Wiley, New York.

[17] PLACKETT, R. L. (1960). Principles of Regression Analysis. Oxford University Press, London.

[18] QUENOUILLE, M. H. (1956). Notes on bias in estimation. Biometrika, V. 43, 353-360.

[19] SCHUCANY, W. R., GRAY, H. L. & OWEN, D. B. (1971). On bias reduction in estimation. J. Am. Statist. Assoc., V. 66, 524-533.

[20] SHORACK, G. R. (1969). Testing and estimating ratios of scale parameters. J. Am. Statist. Assoc., V. 64, 999-1013.

[21] SUKHATME, P. V. & SUKHATME, B. V. (1970). Sampling Theory of Surveys With Applications. Second, revised edition, Iowa State University Press, Ames.

[22] TUKEY, J. W. (1958). Bias and confidence in not-quite large samples. Ann. Math. Statist., V. 29, 614.

[23] WETHERILL, G. B. (1967). Elementary Statistical Methods. Methuen, London.

APPENDIX A

A COMPARISON OF THE ASYMPTOTIC VARIANCES OF J_{S1} AND J_{W1}

A.1 The Asymptotic Variance of J_{S1}

In the following, we adopt the notation and definitions of Section 4.2. Recall that if the hypothesis of Theorem 2.6 is satisfied, the asymptotic variance of $N^{\frac{1}{2}}J_{S1}$ is

$$\sigma_s^2 = \sum_{i=1}^{2} \sum_{j=1}^{2} m_i m_j g_i g_j \zeta_1^{i,j} \ . \tag{A.1}$$

To evaluate (A.1), let k_1^{*1} and k_1^{*2} be defined as in (2.6). Then

$$k_1^{*2} = E[(F_1F_1^*-F_1^{*2}) | (F_1,F_1^*) = (f_1,f_1^*)]$$

$$= f_1f_1^* - f_1^{*2} \ .$$

and

$$k_1^{*1} = E[(F_1^*-F_2^*)^2/2 | (F_1,F_1^*) = (f_1,f_1^*)]$$

$$= \tfrac{1}{2}[f_1^{*2} - 2If_1^* + E(F_2^{*2})] \ .$$

Thus

$$\zeta_1^{1,1} = \mathrm{Var}[k_1^{*1}(Z_1)]$$

$$= \tfrac{1}{4}[\mathrm{Var}(F_1^{*2}) + 4I^2\mathrm{Var}(F_1^*)$$

$$- 4I\mathrm{Cov}(F_1^*,F_1^{*2})] \ ,$$

$$\zeta_1^{2,2} = \mathrm{Var}[k_1^{*2}(Z_1)]$$

$$= \mathrm{Var}(F_1F_1^*) + \mathrm{Var}(F_1^{*2}) - 2\mathrm{Cov}(F_1F_1^*, F_1^{*2})$$

and

$$\zeta_1^{1,2} = \zeta_1^{2,1} = \mathrm{Cov}[k_1^{*1}(Z_1), k_1^{*2}(Z_1)]$$

$$= \mathrm{Cov}[\tfrac{1}{2}(F_1^{*2} - 2IF_1^*), F_1F_1^* - F_1^{*2}] \ .$$

Now

$$g_1 = \frac{\partial g}{\partial t_1}\bigg|_{(\eta_1,\eta_2)}$$

$$= (v_q - v_p)/v_q^2$$

and

$$g_2 = \frac{\partial g}{\partial t_2}\bigg|_{(\eta_1,\eta_2)}$$

$$= 1/v_q \ .$$

Recalling that $m_1 = 2$ and $m_2 = 1$, we can now evaluate (A.1). Thus

$$\sigma_s^2 = (v_q - v_p)^2[\mathrm{Var}(F_1^{*2}) + 4I^2\mathrm{Var}(F_1^*) - 4I\mathrm{Cov}(F_1^*, F_1^{*2})]/v_q^4$$

$$+ [\mathrm{Var}(F_1F_1^*) + \mathrm{Var}(F_1^{*2}) - 2\mathrm{Cov}(F_1F_1^*, F_1^{*2})]/v_q^2$$

$$+ 4(v_q - v_p)[\tfrac{1}{2}\mathrm{Cov}(F_1^{*2}, F_1F_1^*) - I\mathrm{Cov}(F_1^*, F_1F_1^*) - \tfrac{1}{2}\mathrm{Var}(F_1^{*2})$$

$$+ I\mathrm{Cov}(F_1^*, F_1^{*2})]/v_q^3 \ .$$

Upon simplifying, the above becomes

$$\sigma_s^2 = (\gamma_2(f^*)+2)(v_p^2/v_q^2) + 4I\mu_3(f^*)v_p/v_q^3 + 4I^2/v_q$$
$$+ \mathrm{Var}(F_1F_1^*)/v_q^2 - 2v_p \mathrm{Cov}(F_1^{*2},F_1F_1^*)/v_q^3$$
$$+ 4I(v_p-v_q)\mathrm{Cov}(F_1^*,F_1F_1^*)/v_q^3 .$$

A.2 The Asymptotic Variance of J_{W1}

Proceeding as in the previous section, we adopt the definitions and notation of Section 4.3 and let g_1, g_2, g_3 and g_4 be defined as in (2.10).

From Theorem 2.6, the asymptotic variance of $N^{\frac{1}{2}}J_{W1}$ is then

$$\sigma_w^2 = \sum_{i=1}^{4} \sum_{j=1}^{4} m_i m_j g_i g_j \zeta_1^{i,j} . \qquad \text{(A.2)}$$

It is easily verified that

$$m_1 = 2 ,$$

$$m_i = 1 , \ i = 2,3,4 ,$$

$$g_1 = -v_p/v_q^2 ,$$

$$g_2 = 1/v_q ,$$

$$g_3 = -2I/v_q$$

and

$$g_4 = I^2/v_q .$$

Now

$$\zeta_1^{1,1} = \mathrm{Var}[k_1^{*1}(Z_1)]$$

$$= \tfrac{1}{4}[\mathrm{Var}(F_1^{*2}) + 4I^2\mathrm{Var}(F_1^*) - 4I\mathrm{Cov}(F_1^*,F_1^{*2})] \ ,$$

$$\zeta_1^{2,2} = \mathrm{Var}[k_1^{*2}(Z_1)]$$

$$= \mathrm{Var}(C_1F_1^2) \ ,$$

$$\zeta_1^{3,3} = \mathrm{Var}[k_1^{*3}(Z_1)]$$

$$= \mathrm{Var}(F_1^*)$$

$$= v_q \ ,$$

$$\zeta_1^{4,4} = \mathrm{Var}[k_1^{*4}(Z_1)]$$

$$= \mathrm{Var}(C_1) \ ,$$

$$\zeta_1^{1,2} = \mathrm{Cov}[k_1^{*1}(Z_1),k_1^{*2}(Z_1)]$$

$$= \mathrm{Cov}[\tfrac{1}{2}(F_1^{*2}-2IF_1^*),C_1F_1^2] \ ,$$

$$\zeta_1^{1,3} = \mathrm{Cov}[k_1^{*1}(Z_1),k_1^{*3}(Z_1)]$$

$$= \mathrm{Cov}[\tfrac{1}{2}(F_1^{*2}-2IF_1^*),F_1^*] \ ,$$

$$\zeta_1^{1,4} = \mathrm{Cov}[k_1^{*1}(Z_1),k_1^{*4}(Z_1)]$$

$$= \mathrm{Cov}[\tfrac{1}{2}(F_1^{*2}-2IF_1^*),C_1] \ ,$$

$$\zeta_1^{2,3} = \mathrm{Cov}[k_1^{*2}(Z_1), k_1^{*3}(Z_1)]$$

$$= \mathrm{Cov}(C_1F_1^2, F_1^*) \ ,$$

$$\zeta_1^{2,4} = \mathrm{Cov}[k_1^{*2}(Z_1), k_1^{*4}(Z_1)]$$

$$= \mathrm{Cov}(C_1F_1^2, C_1)$$

and

$$\zeta_1^{3,4} = \mathrm{Cov}[k_1^{*3}(Z_1), k_1^{*4}(Z_1)]$$

$$= \mathrm{Cov}(F_1^*, C_1) \ .$$

Noting that

$$C_1F_1^2 = p(X_1)f^2(X_1)/q(X_1)$$

$$= F_1^*F_1 \ ,$$

we can write

$$\zeta_1^{2,3} = \mathrm{Cov}(F_1F_1^*, F_1^*)$$

and

$$\zeta_1^{2,2} = \mathrm{Var}(F_1F_1^*) \ .$$

Substituting the above quantities into (A.2) and simplifying, yields

$$\sigma_w^2 = \sigma_s^2 + I^4 Var(C_1)/v_q^2 - 2I^2 v_p Cov(F_1^{*2}, C_1)/v_q^3$$
$$+ 2I^2 Cov(F_1 F_1^*, C_1)/v_q^2 + 4I^3 (v_p - v_q) Cov(F_1^*, C_1)/v_q^3 .$$

A.3 The Comparison

Consider the difference of the asymptotic variances

$$\sigma_s^2 - \sigma_w^2 = [2I^2 v_p Cov(F_1^{*2}, C_1) - I^4 v_q Var(C_1) \tag{A.3}$$
$$- 2I^2 v_q Cov(F_1 F_1^*, C_1) - 4I^3 (v_p - v_q) Cov(F_1^*, C_1)]/v_q^3 .$$

Note that

$$Cov(F_1 F_1^*, C_1) = \int f^2 p^2/q - (\int p)(\int f^2 p)$$
$$= v_q - v_p .$$

Thus (A.3) becomes

$$\sigma_s^2 - \sigma_w^2 = \{2I^2 v_p Cov(F_1^{*2}, C_1) - I^4 v_q Var(C_1) \tag{A.4}$$
$$+ (v_p - v_q)[2I^2 v_q - 4I^3 Cov(F_1^*, C_1)]\}/v_q^3 .$$

Now, suppose that $f(x) \geq 0$ and that

$$q(x) = \{f(x)p(x)/[1+cg(x)]\}/\int\{f(x)p(x)/[1+cg(x)]\}dx ,$$

where the constant c is small so that $q(x)$ is close to the optimum $q_o(x)$. Then

$$\mathrm{Cov}(F_1^*, C_1) = \mathrm{Cov}_q(fp/q, p/q)$$

$$= \int fp^2/q - I .$$

But

$$\int fp^2/q = [\int fp^2(1+cg)/fp][\int fp(1-cg+c^2g^2- \dots)]$$

$$= I - c\mathrm{Cov}_p(f,g) + c^2[E_p(fg^2) - E_p(g)E_p(fg)] + \mathcal{O}(c^3) \ ,$$

where $\lim_{c \to 0} \mathcal{O}(c^3)/c^3$ equals a constant. Therefore

$$\mathrm{Cov}(F_1^*, C_1) \doteq -c\mathrm{Cov}_p(f,g) + c^2[E_p(fg^2) - E_p(g)E_p(fg)] .$$

To obtain an expression for v_q , we note that

$$\int f^2p^2/q = [\int fp(1+cg)][\int fp(1-cg+c^2g^2- \dots)]$$

$$= I^2 + c^2\{IE_p(fg^2) - [E_p(fg)]^2\} + \mathcal{O}(c^3) \ ,$$

thus

$$v_q \doteq c^2\{IE_p(fg^2) - [E_p(fg)]^2\} .$$

To evaluate $\mathrm{Cov}(F_1^{*2}, C_1)$, first note that

$$\mathrm{Cov}(F_1^{*2}, C_1) = \mathrm{Cov}_q(f^2p^2/q^2, p/q)$$

$$= \int f^2p^3/q^2 - \int f^2p^2/q$$

and

$$\int f^2p^3/q^2 \doteq [\int f^2p^3(1+cg)^2/f^2p^2][\int fp(1-cg+c^2g^2)]^2$$

$$= I^2 - 2cICov_p(f,g)$$

$$+c^2\{2IE_p(fg^2)+[E_p(fg)]^2-4IE_p(g)E_p(fg)+I^2E_p(g^2)\}+\mathcal{O}(c^3) .$$

Thus

$$Cov(F_1^{*2},C_1) \doteq -2cICov_p(f,g)+c^2\{IE_p(fg^2)+2[E_p(fg)]^2\}$$

$$+c^2[I^2E_p(g^2)-4IE_p(g)E_p(fg)] .$$

Finally

$$Var(C_1) = Var_q(p/q)$$

$$= \int p^2/q - 1$$

and

$$\int p^2/q \doteq [\int p^2(1+cg)/fp][\int fp(1-cg+c^2g^2)]$$

$$= E_p(1/f)I+c[IE_p(g/f)-E_p(1/f)E_p(fg)]$$

$$+c^2[E_p(1/f)E_p(fg^2)-E_p(g/f)E_p(fg)]+\mathcal{O}(c^3) .$$

Therefore

$$Var(C_1) \doteq E_p(1/f)I - 1 + c[IE_p(g/f) - E_p(1/f)E_p(fg)]$$

$$+ c^2[E_p(1/f)E_p(fg^2) - E_p(g/f)E_p(fg)] .$$

Substituting the above quantities into the numerator of (A.4) and simplifying, we obtain

$$\sigma_s^2 - \sigma_w^2 = (c^2/v_q^3)(2I^4 v_p \mathrm{Var}_p(g) + 2I^2 v_p [\mathrm{Cov}_p(f,g)]^2$$
$$-I^4\{IE_p(fg^2) - [E_p(fg)]^2\}[E_p(1/f)I-1]) + \mathcal{O}(c^3) \ .$$

APPENDIX B

THE PROOF OF THEOREM 2.5

Before proving Theorem 2.5, we introduce two lemmas which will aid in the proof.

Lemma 1. Let $X_1,\dots,X_N$ $(N = nk)$ be IID random variables. Let U_n be a U-statistic with symmetric kernel $k^*(X_{a_1},\dots,X_{a_m})$ and let U_s be a U-statistic with the same symmetric kernel, based on a subset $(X_{1'},\dots,X_{S'})$ of $(X_1,\dots,X_N)$, where $m \le S \le N$. Then

$$\mathrm{Cov}(U_n,U_s) = \mathrm{Var}(U_n) .$$

Proof.

$\mathrm{Cov}(U_n,U_s) =$

$$\binom{N}{m}^{-1}\binom{S}{m}^{-1}\sum_{C_n}\sum_{C_s}\mathrm{Cov}[k^*(X_{a_1},\dots,X_{a_m}) , k^*(X_{b_1},\dots,X_{b_m})] ,$$

where C_n denotes the sum over all combinations $(a_1,\dots,a_m)$ of m integers chosen from $I_N = (1,\dots,N)$ and C_s denotes the sum over all combinations $(b_1,\dots,b_m)$ of m integers chosen from $I_S = (1',\dots,S')$.

Hoeffding [9] has shown that

$$\mathrm{Cov}[k^*(X_{a_1},\dots,X_{a_m}),k^*(X_{b_1},\dots,X_{b_m})] = \zeta_c ,$$

where ζ_c is defined as in (2.7) and c is the number of integers common to $(a_1,\ldots,a_m)$ and $(b_1,\ldots,b_m)$. Therefore,

$$\mathrm{Cov}(U_n,U_s) = \binom{N}{m}^{-1}\binom{S}{m}^{-1} \sum_{c=1}^{m} K_c \zeta_c ,$$

where K_c is equal to the number of ways we can choose a set $(b_1,\ldots,b_m)$ of integers from I_S and a set of integers $(a_1,\ldots,a_m)$ from I_N such that the sets have exactly c integers in common.

Now, we can choose m integers from I_S in $\binom{S}{m}$ ways and we can choose the c integers to be common to both sets in $\binom{m}{c}$ ways, thus completely specifying $(b_1,\ldots,b_m)$ and c of the a's. The remaining a's can be chosen in $\binom{N-m}{m-c}$ ways. Therefore,

$$K_c = \binom{S}{m}\binom{m}{c}\binom{N-m}{m-c}$$

and

$$\mathrm{Cov}(U_n,U_s) = \binom{N}{m}^{-1} \sum_{c=1}^{m} \binom{m}{c}\binom{N-m}{m-c}\zeta_c \qquad \text{(B.1)}$$

$$= \mathrm{Var}(U_n) \text{ , from Hoeffding [9].} \quad \square$$

Lemma 2. Let $X_1,\ldots,X_N$ $(N = nk)$ be IID random variables and let U_n and U_s be defined as in Lemma 1, with $S = sk$. Then if n and s are fixed,

$$\lim_{N\to\infty} NVar(U_n) = m^2\zeta_1$$

and

$$\lim_{N\to\infty} NVar(U_s) = (n/s)m^2\zeta_1 .$$

<u>Proof</u>. Expanding (B.1), we obtain

$$Var(U_n) = (m^2/N)\frac{(1-m/N)\ldots(1-(2m+2)/N)}{(1-1/N)\ldots(1-(m+1)/N)}\zeta_1$$

$$+ \text{(terms of order } 1/N^2) . \qquad (B.2)$$

Thus, $\lim_{N\to\infty} NVar(U_n) = m^2\zeta_1$.

To obtain $Var(U_s)$, we need only replace N by $S = sN/n$ in (B.2), which yields

$$\lim_{N\to\infty} NVar(U_s) = (n/s)m^2\zeta_1 . \qquad \square$$

We now restate and prove Theorem 2.5 of Chapter 2.

<u>THEOREM 2.5</u>. Let $X_1,\ldots,X_N$ $(N = nk)$ be N IID random variables and let $k^*(X_1,\ldots,X_m)$ be a real-valued symmetric statistic with expectation η and finite second moment $E[k^*(X_1,\ldots,X_m)]^2$. Let g be a function defined on the real line which, in a neighborhood of η , has a continuous first derivative. Then, as $N\to\infty$, with n fixed,

$$n^{\frac{1}{2}}[J(\hat{\theta}) - \theta]/S_g \underset{D}{\rightarrow} t(n-1) \ ,$$

where $t(n-1)$ denotes the Student-t distribution with $n-1$ degrees of freedom.

PROOF. Without loss of generality let $n = 0$. Then the pseudo-values

$$J_i = ng(U) - (n-1)g(U_i) = g(V_i) + d_i \ , \quad i = 1,\ldots,n \ ,$$

where

$$d_i = ng(U) - (n-1)g(U_i) - g(V_i) \ ,$$

and V_i is a U-statistic with symmetric kernel $k^*(X_1,\ldots,X_m)$, based on the i^{th} group of X's.

Claim: As $N\rightarrow\infty$, $N^{\frac{1}{2}}d_i \underset{P}{\rightarrow} 0$, $i = 1,\ldots,n$.

To prove the claim, we proceed as follows. Let $i \in (1,\ldots,n)$ be arbitrary and let Y_N be a random variable defined by

$$Y_N = N^{\frac{1}{2}}g'(0)[nU - (n-1)U_i - V_i] \ .$$

Then $E(Y_N) = 0$ for all N and the variance of Y_N is given by

$$\text{Var}(Y_N) = [g'(0)]^2 N[n^2\text{Var}(U) + (n-1)^2\text{Var}(U_i) + \text{Var}(V_i)$$
$$- 2n(n-1)\text{Cov}(U,U_i) - 2n\text{Cov}(U,V_i)] \ .$$

Note that $\text{Cov}(U_i,V_i) = 0$, since U_i and V_i are based on disjoint subsets of $(X_1,\ldots,X_N)$. From Lemmas 1 and 2, we have

$$\lim_{N\to\infty} N\text{Var}(U) = \lim_{N\to\infty} N\text{Cov}(U,U_i) = \lim_{N\to\infty} N\text{Cov}(U,V_i) = m^2\zeta_1 \ ,$$

$$\lim_{N\to\infty} N\text{Var}(U_i) = [n/(n-1)]m^2\zeta_1$$

and

$$\lim_{N\to\infty} N\text{Var}(V_i) = nm^2\zeta_1 \ .$$

Therefore

$$\lim_{N\to\infty} \text{Var}(Y_N) = m^2\zeta_1[g'(0)]^2[n^2+n(n-1)+n-2n(n-1)-2n]$$
$$= 0 \ .$$

Hence, $Y_N \underset{P}{\to} 0$, by Tchebycheff's theorem.

Let A be an interval containing 0 such that g' is continuous on A . It can be shown (see the proof of Theorem 5, Arvesen [1]) that $\text{Prob}(U,U_i,V_i \in A \text{ simultaneously}) \to 1$. Note that for $U,U_i,V_i \in A$, we have

$$N^{\frac{1}{2}} d_i = Y_N - Z_N ,$$

where

$$Z_N = [g'(0)-g'(\xi_1)]nN^{\frac{1}{2}}U - [g'(0)-g'(\xi_2)](n-1)N^{\frac{1}{2}}U_i$$
$$- [g'(0)-g'(\xi_3)]N^{\frac{1}{2}}V_i ,$$

and ξ_1 lies between U and 0 , ξ_2 lies between U_i and 0 , and ξ_3 lies between V_i and 0 .

But $Z_N \underset{P}{\rightarrow} 0$, since $g'(\xi_i) \underset{P}{\rightarrow} g'(0)$, $i = 1,2,3$, and $N^{\frac{1}{2}}V_i$, $nN^{\frac{1}{2}}U$ and $(n-1)N^{\frac{1}{2}}U_i$ are asymptotically normally distributed. Hence, $N^{\frac{1}{2}}d_i \underset{P}{\rightarrow} 0$, and the claim holds.

Now, since $N^{\frac{1}{2}}V_i$ is asymptotically normally distributed, we have

$$N^{\frac{1}{2}}[g(V_i) - g(0)] \underset{D}{\rightarrow} B_i \quad , \quad i = 1,\ldots,n ,$$

where the B_i's are independent normal random variables. Therefore, as $N\rightarrow\infty$, the random vector $\underline{J} = (N^{\frac{1}{2}}[J_1 - g(0)],\ldots,N^{\frac{1}{2}}[J_n - g(0)])$ converges in distribution to the random vector $\underline{B} = (B_1,\ldots,B_n)$.

The theorem now follows immediately from Corollary 1 of Billingsley [3, p.31], which states that $h(\underline{J}) \underset{D}{\rightarrow} h(\underline{B})$ if h is measurable and

$\text{Prob}(\underline{B} \in \text{(the set of discontinuities of } h)) = 0$. □

APPENDIX C

SUKHATMES' EXAMPLE

In the following, we apply the estimators J_{S1}, J_{W1} and J_{W1L} to an agricultural sampling problem which is discussed in Sukhatme & Sukhatme [21, pp.50-52].

The sampled population $(u_1,\ldots,u_N)$ consists of $N = 892$ villages. For each village (sampling unit) the characteristic of interest, $f(u_i) = y_i$, is the area of land under rice. A record of the total cultivated area, w_i, is available for each village, thus providing a convenient auxiliary variable.

To estimate the total area under rice by sampling with probability proportional to cultivated area (ppes sampling), define

$$q(u_i) = w_i / \sum_{j=1}^{N} w_j \quad , \quad i = 1,\ldots,N \; ,$$

and

$$f^*(u_i) = f(u_i)/Nq(u_i) \quad , \quad i = 1,\ldots,N \; .$$

Then, if $(u_1,\ldots,u_n)$ is a sample from $q(u_i)$, $N\hat{I}_q$ is an unbiased estimate of the total area under rice, where $\hat{I}_q$ is the average of the $f^*(u_i)$ values. Furthermore, if S_q^2 is defined by

$$s_q^2 = \sum_{i=1}^{n} [f^*(u_i) - \hat{I}_q]^2/(n-1) ,$$

$N^2 s_q^2/n$ provides an unbiased estimate of $\text{Var}(N\hat{I}_q) = N^2 v_q/n$.

Now, let $p(u_i) = 1/N$, $i = 1,\ldots,N$, be the density function of the discrete uniform distribution. Then, if the villages were selected with equal probability and with replacement, the variance of the estimated total area under rice would be

$$\text{Var}(N\bar{y}_n) = N^2 v_p/n ,$$

where

$$v_p = \sum_{i=1}^{N} [f(u_i)]^2 p(u_i) - [\sum_{i=1}^{N} f(u_i)p(u_i)]^2$$

$$= \sum_{i=1}^{N} (y_i - \bar{Y})^2/N .$$

Therefore, in this case,

$$\text{Var}(N\bar{y}_n)/\text{Var}(N\hat{I}_q) = v_p/v_q . \quad \text{(C.1)}$$

If, however, the villages were selected with equal probability but without replacement, the variance of the estimated total would be

$$\text{Var}(N\bar{y}_n) = N^2 K_n v_p/n , \quad \text{(C.2)}$$

and the ratio (C.1) would become

$$\mathrm{Var}(N\bar{y}_n)/\mathrm{Var}(N\hat{I}_q) = K_n v_p/v_q \ , \qquad (C.3)$$

where

$$K_n = (N-n)/(N-1) \ .$$

Denote the ratio (C.3) by V_R .

Sukhatme and Sukhatme [21, p.51] have proposed estimating the $\mathrm{Var}(N\bar{y}_n)$ in (C.2) by

$$\hat{V}_S = N^2 K_n \hat{v}_{ps}/n \ ,$$

where $\hat{v}_{ps}$ is defined as in (1.4). (This estimator appears in an alternative form in [21] but is easily shown to be equivalent to $\hat{V}_S$.) Substituting $\hat{V}_S$ for $\mathrm{Var}(N\bar{y}_n)$ and $N^2 s_q^2/n$ for $\mathrm{Var}(N\hat{I}_q)$ in (C.3), they obtain an estiamte $\hat{V}_R$ of the ratio V_R given by

$$\hat{V}_R = K_n \hat{v}_{ps}/s_q^2$$

$$= K_n \hat{\theta}_s \ ,$$

where $\hat{\theta}_s$ is defined as in (1.5). Recall that the estimator $\hat{\theta}_s$ may take on negative values.

An alternative estimator of the ratio V_R can be obtained by replacing v_p/v_q by $\hat{\theta}_w$ in (C.3), where $\hat{\theta}_w$ is defined as in (1.8). This yields a non-negative estimator $\hat{V}_w$ defined by

$$\hat{V}_w = K_n \hat{\theta}_w .$$

The jackknife method may be applied to both $\hat{V}_R$ and $\hat{V}_w$. The resulting estimators, however, will not necessarily equal zero when $n = N$. This problem may be overcome by jackknifing $\hat{\theta}_w$ and $\hat{\theta}_s$ before multiplying by the finite population correction, K_n. Also, since $\hat{\theta}_w$ is non-negative, the logarithmic transformation may be used in conjunction with jackknife in this case.

The estimators $K_n J_{S1}$, $K_n J_{W1}$ and $K_n J_{W1L}$ were applied to the data presented in Table 2.1 of Sukhatme and Sukhatme [21, p.51]. The results are recorded in Table XXXIV.

Note that the value of the estimate $K_n J_{S1}$ is considerably lower than the values of the estimators $K_n J_{W1}$ and $K_n J_{W1L}$. All three procedures for testing $H_o: v_p = v_q$ vs. $H_a: v_p \neq v_q$ lead to rejection of H_o at the $\alpha = 0.1$ level of significance, but when $\alpha = .05$, the test based on J_{S1} does not lead to rejection of H_o. For the one-sided test, $H_o: v_p = v_q$ vs. $H_a: v_p > v_q$ is rejected at the $\alpha = .05$ level of significance in each case.

VITA

Surname: CUMBERBIRCH Given Names: PETER MARK

Place of Birth: VICTORIA, B. C. Date of Birth: FEBRUARY 25, 1948

Educational Institutions Attended, with Dates of Entering and Leaving:

UNIVERSITY OF VICTORIA, VICTORIA 1966/67 and 1968 to 1971

Degrees, Diplomas, Etc., Awarded, with Dates and Names of Institutions:

B.Sc.	1971	UNIVERSITY OF VICTORIA, VICTORIA

Honors and Awards:

University of Victoria Graduate Scholarship, 1972/73 and 1973/74

Publications: