THE COMPARISON OF SOME RATIO ESTIMATORS FOR SMALL SAMPLES

by

Richard Alexander Kuzik

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF SCIENCE

DEPARTMENT OF STATISTICS

UNIVERSITY OF MANITOBA

WINNIPEG, MANITOBA

October 1971

ABSTRACT

One of the main objectives of a sample survey is the estimation of the population mean or total of a character 'y' attached to the units in the population. The ratio method of estimation provides a powerful technique for obtaining increased precision whenever information on an auxiliary character 'x' positively correlated with 'y' is available. Since the classical ratio estimator is biased, considerable attention has been given in recent years to the development of wholly unbiased or approximately unbiased ratio estimators. The relative efficiencies of these ratio estimators have been previously investigated by assuming a linear regression of y on x and gamma distribution for x. In this thesis, various x-populations are employed to investigate relative efficiencies of the estimators empirically, assuming linear regression of y on x. The results obtained indicate that the relative efficiencies are fairly insensitive to the distribution of x-values.

Two classical estimators of variance and a 'jack-knife' variance estimator are also considered and the relative biases and relative stabilities of these variance estimators are investigated empirically.

ACKNOWLEDGEMENT

I am extremely grateful and express my sincere appreciation to Dr. J. N. K. Rao, Department of Statistics, for his most valuable guidance and constructive criticism throughout the period of this study and during the writing of this thesis. I am also thankful to Dr. G. F. Atkinson, Department of Statistics, for his assistance in developing the efficient computer program used in this thesis, and to Dr. K. Mount, Department of Statistics, and Dr. P. J. Dhruvarajan, Department of Economics, who examined this thesis extensively and recommended very helpful suggestions as to style and context.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

1.1 Objective

Sample surveys are conducted to obtain estimates of various unknown population parameters of interest. Here we confine ourselves to estimation of the population mean or total of a character under study 'y'. Ratio estimators incorporate the knowledge of an auxiliary variable 'x', which is positively correlated with the character 'y'.

In such situations, the classical ratio estimator is often used. However, it is biased and, in surveys where small or moderate samples within many strata make it appropriate to use the 'separate' ratio estimators, the bias relative to the standard error could be large. For these types of surveys, modifications to the classical ratio estimator lead to wholly or approximately unbiased ratio estimators. This thesis considers the performance of some modified estimators and investigates the stability of three variance estimators under different population conditions.

1.2 Review of Literature

For simple random sampling of size n, drawn without replacement from a finite population of size N, the classical ratio estimator of the population mean $\bar{Y}$ is given by

$$\bar{y}_r = \frac{\bar{y}}{\bar{x}}\bar{X} = r\bar{X} \tag{1.1}$$

where $\bar{X}$ is the known population mean of the auxiliary variable 'x'; $\bar{y}$ and $\bar{x}$ are the sample means of 'y' and 'x' respectively; and

$$r = \frac{\bar{y}}{\bar{x}} \tag{1.2}$$

is the customary estimator of the ratio

$$R = \frac{\bar{Y}}{\bar{X}} . \tag{1.3}$$

The classical ratio estimator has a bias almost of order $O(n^{-1})$ for large n.

Several modifications to the classical ratio estimator leading to approximately or wholly unbiased estimators have been proposed in the literature. Beale's (1962) approximately unbiased estimator of $\bar{Y}$ is given by

$$\bar{y}_B = r\bar{X} \frac{1 + \left(\frac{1}{n} - \frac{1}{N}\right) \frac{s_{xy}}{\bar{x}\,\bar{y}}}{1 + \left(\frac{1}{n} - \frac{1}{N}\right) \frac{s_x^2}{\bar{x}^2}} \tag{1.4}$$

while Tin's (1965) modification of this estimator is

$$\bar{y}_T = r\bar{X} \left[1 + \left(\frac{1}{n} - \frac{1}{N}\right)\left(\frac{s_{xy}}{\bar{x}\,\bar{y}} - \frac{s_x^2}{\bar{x}^2}\right)\right] \tag{1.5}$$

where

$$s_x^2 = \frac{n(\Sigma x^2)_s - (\Sigma x)_s^2}{n(n-1)} \tag{1.6}$$

$$s_{xy} = \frac{n(\Sigma xy)_s - (\Sigma x)_s(\Sigma y)_s}{n(n-1)} \tag{1.7}$$

where $(\Sigma x)_s$ is the sample total for 'x', i.e.,

$$(\Sigma x)_s = \sum_{i=1}^{n} x_i. \tag{1.8}$$

A wholly unbiased ratio estimator, proposed by Hartley and Ross (1954), is

$$\bar{y}_H = \bar{r}\,\bar{X} + \frac{(N-1)n}{N(n-1)}(\bar{y} - \bar{r}\bar{x}) \tag{1.9}$$

where

$$\bar{r} = n^{-1}(\Sigma[y/x])_s. \tag{1.10}$$

Two estimators based on splitting up the sample at random into 'g'($\geq$ 2) groups, each of size 'm', where n = mg, are due to Quenouille (1956) and Mickey (1959). Mickey's wholly unbiased estimator of $\bar{Y}$ is

$$\bar{y}_{M(g)} = \bar{r}'_g\,\bar{X} + \frac{(N - n + m)g}{N}(\bar{y} - \bar{r}'_g\,\bar{x}) \tag{1.11}$$

where

$$\bar{r}'_g = \sum_{j=1}^{g} \bar{r}'_j/g \tag{1.12}$$

where $\bar{r}'_j$ is the classical ratio estimator computed from the sample after omitting the j^{th} group, i.e.,

$$\bar{r}'_j = \frac{\bar{y}'_j}{\bar{x}'_j} \tag{1.13}$$

where

$$\bar{y}'_j = \frac{n\bar{y} - m\bar{y}_j}{n-m} \; ; \quad \bar{x}'_j = \frac{n\bar{x} - m\bar{x}_j}{n-m} \qquad (1.14)$$

and $\bar{y}_j$, $\bar{x}_j$ are the j^{th} group sample means. Note that $\bar{y}_H = \bar{y}_{M(g)}$ when n = 2.

Quenouille's method of bias reduction leads to the approximately unbiased estimator

$$\bar{y}_{Q(g)} = [\omega_g r - (\omega_g - 1)\bar{r}'_g]\,\bar{X} \qquad (1.15)$$

where

$$\omega_g = g\left[1 - \frac{(n - m)}{N}\right]. \qquad (1.16)$$

The approximately unbiased estimators have bias almost of order $O(n^{-2})$ for large n (terms of order $O(N^{-1})$ do not appear).

Turning to the variance estimators, the classical estimator of the variance of $\bar{y}_r$ is

$$v_1(\bar{y}_r) = \left[\frac{1}{n} - \frac{1}{N}\right](s_y^2 - 2rs_{xy} + r^2 s_x^2). \qquad (1.17)$$

Another variance estimator (Cochran, 1963) is

$$v_2(\bar{y}_r) = \frac{\bar{X}^2}{\bar{x}^2}\left[\frac{1}{n} - \frac{1}{N}\right](s_y^2 - 2rs_{xy} + r^2 s_x^2). \qquad (1.18)$$

Tukey's (1958) 'jack-knife' variance estimator of $\bar{y}_r$ is

$$v_{3g}(\bar{y}_r) = \bar{X}^2\left[1 - \frac{n}{N}\right]\frac{(g - 1)}{g}\sum_{j=1}^{g}(\bar{r}'_j - \bar{r}'_g)^2 \qquad (1.19)$$

For the comparison of the various estimators, several approaches have been used. Tin compared $\bar{y}_r$, $\bar{y}_B$, $\bar{y}_T$, $\bar{y}_H$ and $\bar{y}_{Q(2)}$ for large n without assuming any model. Rao (1965), Rao and Webster (1966) and Rao (1967) compared $\bar{y}_r$, $\bar{y}_T$, $\bar{y}_{Q(g)}$, $\bar{y}_{M(g)}$ and $\bar{y}_H$ under two models assuming infinite populations, viz the regression of y on x is linear with constant error variance and

(1) $x \sim N(\mu, \sigma^2)$, i.e., x is normally distributed with mean μ and variance σ^2;

(2) $x \sim \Gamma(h)$, i.e., x is distributed as a gamma with parameter h.

The results under model (2) are exact for any sample size n.

Tin showed that $\bar{y}_B$ and $\bar{y}_T$ were better than the other estimators with regard to bias, efficiency and approach to normality.

Using model (1), Rao has shown that

A) g = n is the optimum choice for $\bar{y}_{M(g)}$ and $\bar{y}_{Q(g)}$.

B) the asymptotic variance of $\bar{y}_{M(g)}$ (with g = n) is slightly smaller than that of $\bar{y}_r$, but is slightly larger than the mean square error of $\bar{y}_{Q(g)}$ (with g = n).

Using model (2), Rao and Webster demonstrated that

A) g = n is the optimum choice for $\bar{y}_{M(g)}$ and $\bar{y}_{Q(g)}$.

B) $\bar{y}_{M(g)}$ is considerably better than $\bar{y}_H$ for n > 2 and only slightly better than $\bar{y}_r$ for $n \geq 8$.

C) $\bar{y}_{Q(g)}$ (with g = n) and $\bar{y}_T$ are better than $\bar{y}_r$, $\bar{y}_H$ and $\bar{y}_{M(g)}$.

CHAPTER II

APPROACH TO THE PROBLEM

We assume a relationship between the auxiliary variable 'x' and the character of interest 'y'. However, instead of making a specific distributional assumption on x, a wide variety of live and synthetic x-populations will be employed. The purpose of this approach is to study whether these results will be similar to those obtained under the assumption of a gamma distribution for x (which leads to elegant analytical results). With regard to the estimators used when splitting up the sample into groups, we confine ourselves here to g = n, as previous investigations indicate that g = n is an optimum choice (and also to reduce computer time).

Mickey's and Quenouille's estimators, along with Tukey's variance estimator, for g = n, reduce to

$$\bar{y}_M = \bar{X}\,\bar{r}'_n + \frac{(N - n + 1)}{N}\, n\,(\bar{y} - \bar{r}'_n\,\bar{x}) \tag{2.1}$$

$$\bar{y}_Q = \bar{X}\,[\omega r - (\omega - 1)\,\bar{r}'_n] \tag{2.2}$$

and

$$v_3(\bar{y}_r) = \bar{X}^2 \left(1 - \frac{n}{N}\right) \frac{(n - 1)}{n} \sum_{j=1}^{n} (\bar{r}'_j - \bar{r}'_n)^2 \tag{2.3}$$

where

$$\bar{r}'_n = \frac{(\Sigma r')_s}{n} \tag{2.4}$$

$$\bar{y}_j' = \frac{n\bar{y} - y_j}{n-1} \; ; \quad \bar{x}_j' = \frac{n\bar{x} - x_j}{n-1} \tag{2.5}$$

$$\omega = n \left[1 - \frac{(n - 1)}{N}\right]. \tag{2.6}$$

In developing the above approach, it was originally planned to develop a general relationship between y and x, such as

$$y_i = \alpha + \beta x_i^{p+1} + \gamma x_i^{2(p+1)} + u_i \qquad i = 1, 2,\ldots,N \tag{2.7}$$

for general p, where u_i is the error term such that

$$\varepsilon(u_i^2 | x_i) = \delta x_i^t, \quad \delta > 0; \quad i = 1, 2,\ldots,N \tag{2.8}$$

$$\varepsilon(u_i u_j | x_i, x_j) = 0 \qquad \text{for } i \neq j = 1, 2,\ldots,N \tag{2.9}$$

where ε denotes the conditional expectation under the model for given x.

This approach, however, was found to be too ambitious, even for the classical ratio estimator, as the average mean square error ε(MSE) involved too many parameters. To illustrate this, consider the average MSE of $\bar{y}_r$ under the model

$$\varepsilon(\text{MSE } \bar{y}_r) = \xi \left[\frac{\bar{y}}{\bar{x}} \bar{X} - \bar{Y}\right]^2$$

$$= \xi \left[\frac{\{n\alpha + \beta(\Sigma x^{p+1})_s + \gamma(\Sigma x^{2(p+1)})_s + n\bar{u}\}}{n\bar{x}} \bar{X} - \frac{(N\alpha + \beta\Sigma^N x_i^{p+1} + \gamma\Sigma^N x_i^{2(p+1)} + N\bar{U})}{N}\right]^2 \tag{2.10}$$

where $\bar{u}$ and $\bar{U}$ are the sample mean and the population mean of the errors u_i respectively and $\xi = \varepsilon E$ where E is the expectation over all the $\binom{N}{n}$ possible.

samples for a given finite population. Therefore, (2.10) reduces to

$$\varepsilon(\text{MSE } \bar{y}_r) = E\left[\alpha\left(\frac{\bar{X}}{\bar{x}} - 1\right) + \beta\left(\frac{\bar{X}}{\bar{x}} \frac{(\Sigma x^{p+1})_s}{n} - \frac{\sum^N x_i^{p+1}}{N}\right) + \gamma\left(\frac{\bar{X}}{\bar{x}} \frac{(\Sigma x^{2(p+1)})_s}{n} - \frac{\sum^N x_i^{2(p+1)}}{N}\right)\right]^2 + \delta\, E\left[\frac{\bar{X}^2}{\bar{x}^2} \frac{(\Sigma x^t)_s}{n^2} + \frac{\sum^N x_i^t}{N^2} - 2\,\frac{\bar{X}}{\bar{x}} \frac{(\Sigma x^t)_s}{nN}\right] \tag{2.11}$$

because

$$n^2\, \varepsilon(\bar{u}^2) = \varepsilon[(\Sigma u)_s^2] = \delta(\Sigma x^t)_s \tag{2.12}$$

$$N^2\, \varepsilon(\bar{U}^2) = \varepsilon\left[\sum^N u_i\right]^2 = \delta \sum^N x_i^t \tag{2.13}$$

$$nN\, \varepsilon(\bar{u}\bar{U}) = \varepsilon\left[(\Sigma u)_s \left(\sum^N u_i\right)\right] = \delta(\Sigma x^t)_s \tag{2.14}$$

It is seen that (2.11) involves α^2, β^2, γ^2, $\alpha\beta$, $\alpha\gamma$, $\beta\gamma$ as well as δ, p, t.

We therefore confine ourselves to a linear regression of y on x, <u>viz</u>

$$y_i = \alpha + \beta x_i + u_i \qquad i = 1, 2, \ldots, N \tag{2.15}$$

assume that u_i's satisfy (2.8) and (2.9). Then

$$\varepsilon(\text{MSE } \bar{y}_r) = \xi\left[\frac{(\alpha + \beta\bar{x} + \bar{u})}{\bar{x}}\,\bar{X} - (\alpha + \beta\bar{X} + \bar{U})\right]^2$$

$$= \alpha^2 E\left[\frac{(\overset{N}{\Sigma} x_i) - (\Sigma x)_s}{(\Sigma x)_s}\right]^2$$

$$+ \delta E\left[\frac{(\Sigma x^t)_s(\overset{N}{\Sigma} x_i)\{(\overset{N}{\Sigma} x_i)-2(\Sigma x)_s\}+(\overset{N}{\Sigma} x_i^t)(\Sigma x)_s^2}{N^2(\Sigma x)_s^2}\right] \tag{2.16}$$

which involves only α^2, δ and t. Noting that

$$s_{xy} = \beta s_x^2 + s_{xu} \tag{2.17}$$

$$s_y^2 = \beta^2 s_x^2 + 2\beta s_{xu} + s_u^2 \tag{2.18}$$

where

$$s_{xu} = \frac{n(\Sigma xu)_s - (\Sigma x)_s(\Sigma u)_s}{n(n-1)} \tag{2.19}$$

$$s_u^2 = \frac{n(\Sigma u^2)_s - (\Sigma u)_s^2}{n(n-1)} \tag{2.20}$$

we have, for Beale's estimator

$$\varepsilon MSE(\bar{y}_B) = \xi\left[\frac{\{nN\bar{x}\bar{y} + (N-n)(\beta s_x^2 + s_{xu})\}\bar{X}}{nN\bar{x}^2 + s_x^2(N-n)} - \bar{Y}\right]^2$$

$$= \xi\left[\alpha\left(\frac{nN\bar{x}\bar{X}}{nN\bar{x}^2 + s_x^2(N-n)} - 1\right)\right.$$

$$\left. + \bar{u}\left(\frac{nN\bar{x}\bar{X}}{nN\bar{x}^2 + s_x^2(N-n)}\right) - s_{xu}\left(\frac{(N-n)\bar{X}}{nN\bar{x}^2 + s_x^2(N-n)}\right) - \bar{U}\right]^2 \tag{2.21}$$

Similarly, for Tin's estimator

$$\varepsilon MSE(\bar{y}_T) = \xi\left[\alpha\left(\frac{\bar{X}}{\bar{x}}\left\{1 - \frac{(N-n)s_x^2}{nN\bar{x}^2}\right\} - 1\right)\right.$$

$$\left. + \bar{u}\left(\frac{\bar{X}}{\bar{x}}\left\{1 - \frac{(N-n)s_x^2}{nN\bar{x}^2}\right\}\right) + s_{xu}\frac{(N-n)\bar{X}}{nN\bar{x}^2} - \bar{U}\right]^2. \quad (2.22)$$

For the Hartley-Ross unbiased estimator

$$\varepsilon V(\bar{y}_H) = \xi\left[\alpha\left\{\left(\frac{\bar{X}}{n} - \frac{(N-1)\bar{x}}{N(n-1)}\right)(\Sigma x^{-1})_s + \frac{(N-1)n}{N(n-1)} - 1\right\}\right.$$

$$\left. + (\Sigma u x^{-1})_s\left(\frac{\bar{X}}{n} - \frac{(N-1)\bar{x}}{N(n-1)}\right) + \bar{u}\frac{(N-1)n}{N(n-1)} - \bar{U}\right]^2 \quad (2.23)$$

where V denotes the variance over all samples from a given finite population. Finally, for the estimators due to Mickey and Quenouille, we have

$$\varepsilon V(\bar{y}_M) = \xi\left[\alpha\left\{(\Sigma\bar{x}_j'^{-1})_s\left(\frac{\bar{X}}{n} - \frac{(N-n+1)\bar{x}}{N}\right) + \frac{(N-n+1)n}{N} - 1\right\}\right.$$

$$\left. + (\Sigma\bar{u}_j'\bar{x}_j'^{-1})_s\left(\frac{\bar{X}}{n} - \frac{(N-n+1)\bar{x}}{N}\right) + \bar{u}\frac{(N-n+1)n}{N} - \bar{U}\right]^2. \quad (2.24)$$

$$\varepsilon MSE(\bar{y}_Q) = \xi\left[\alpha\left(\frac{n\bar{X}(N-n+1)}{N\bar{x}} - \frac{(N-n)(n-1)\bar{X}}{nN}(\Sigma\bar{x}_j'^{-1})_s - 1\right)\right.$$

$$\left. - (\Sigma\bar{u}_j'\bar{x}_j'^{-1})_s\frac{(N-n)(n-1)\bar{X}}{nN} + \bar{u}\left(\frac{n\bar{X}(N-n+1)}{N\bar{x}}\right) - \bar{U}\right]^2. \quad (2.25)$$

Now, using the formulae given in Appendicies I and II, we get the following average MSE's and average variances:

$$\varepsilon MSE(\bar{y}_B) = \alpha^2 E\left[\frac{(\Sigma x)_s\{n(\Sigma x_i^N) - N(\Sigma x)_s\} - n(N-n)s_x^2}{N(\Sigma x)_s^2 + n(N-n)s_x^2}\right]$$

$$+ \delta E\left[\frac{n^2(N-n)^2(\Sigma x^{t+2})_s(\Sigma x_i^N)^2}{N^2(n-1)^2[N(\Sigma x)_s^2 + n(N-n)s_x^2]^2}\right]$$

$$+ \delta E\left[\frac{2n(N-n)(\Sigma x^{t+1})_s(\Sigma x_i^N)\{[N(n-2)+n](\Sigma x)_s(\Sigma x_i^N)-(n-1)[N(\Sigma x)_s^2+n(N-n)s_x^2]\}}{N^2(n-1)^2[N(\Sigma x)_s^2 + n(N-n)s_x^2]^2}\right]$$

$$+ \delta E\left[\frac{[N(n-2)+n](\Sigma x^t)_s(\Sigma x)_s(\Sigma x_i^N)\{[N(n-2)+n](\Sigma x)_s(\Sigma x_i^N)-2(n-1)[N(\Sigma x)_s^2+n(N-n)s_x^2]\}}{N^2(n-1)^2[N(\Sigma x)_s^2 + n(N-n)s_x^2]^2}\right]$$

$$+ \frac{\delta(\Sigma x_i^{t\,N})}{N^2} . \qquad (2.26)$$

$$\varepsilon MSE(\bar{y}_T) = \alpha^2 E\left[\frac{n(\Sigma x_i^N)[N(\Sigma x)_s^2-n(N-n)s_x^2]-N^2(\Sigma x)_s^3}{N^2(\Sigma x)_s^3}\right]^2$$

$$+ \delta E\left[\frac{n^2(N-n)^2(\Sigma x^{t+2})_s(\Sigma x_i^N)^2}{N^4(n-1)^2(\Sigma x)_s^4}\right] + \delta E\left[\frac{2n(N-n)(\Sigma x^{t+1})_s(\Sigma x_i^N)}{N^2(n-1)(\Sigma x)_s^2}\right.$$

$$\left.\left\{\frac{(\Sigma x_i^N)\{[N(n-2)+n](\Sigma x)_s^2 - n(n-1)(N-n)s_x^2\}}{N^2(n-1)(\Sigma x)_s^3} - \frac{1}{N}\right\}\right]$$

$$+ \delta E\left[\frac{(\Sigma x_i^N)(\Sigma x^t)_s[N(\Sigma x)_s^2 - n(N-n)s_x^2]}{N^2(\Sigma x)_s^3}\right.$$

$$\left\{ \frac{(\Sigma x_i^N)\{(\Sigma x)_s^2[N(n-3) + 2n] - n(n-1)(N-n)s_x^2\}}{N^2(n-1)(\Sigma x)_s^3} - \frac{2}{N} \right\}$$

$$+ \left. \frac{(\Sigma x_i^N)(\Sigma x^t)_s(N-n)\{(\Sigma x_i)(N-n) + 2N(n-1)(\Sigma x)_s\}}{N^4(n-1)^2(\Sigma x)_s^2} \right] + \frac{\delta(\Sigma x_i^{t\,N})}{N^2} \qquad (2.27)$$

$$\varepsilon V(\bar{y}_H) = \alpha^2 E\left[\left\{ \frac{(n-1)(\Sigma x_i^N) - (N-1)(\Sigma x)_s}{nN(n-1)} \right\}(\Sigma 1/x)_s + \frac{N-n}{N(n-1)}\right]^2$$

$$+ \delta E\left[(\Sigma x^{t-2})_s \left\{ \frac{(n-1)(\Sigma x_i^N) - (N-1)(\Sigma x)_s}{nN(n-1)} \right\}^2\right]$$

$$+ 2\delta E\left[(\Sigma x^{t-1})_s \left\{ \frac{(n-1)(\Sigma x_i^N) - (N-1)(\Sigma x)_s}{nN(n-1)} \right\}\left\{\frac{N-n}{N(n-1)}\right\}\right]$$

$$+ \delta E\left[\frac{(N-1)(N-2n+1)(\Sigma x^t)_s}{N^2(n-1)^2}\right] + \frac{\delta(\Sigma x_i^{t\,N})}{N^2} \qquad (2.28)$$

$$\varepsilon V(\bar{y}_M) = \alpha^2 E\left[\frac{(n-1)\{(\Sigma x_i^N)-(N-n+1)(\Sigma x)_s\}}{nN}\left\{\Sigma\left[\frac{1}{(\Sigma x)_s - x_i}\right]\right\}_s + \frac{(N-n)(n-1)}{N}\right]^2$$

$$+ \delta E\left[\frac{(\Sigma x^t)_s(N-n+1)(N-n-1)}{N} + \left\{\frac{(\Sigma x_i^N)-(\Sigma x)_s(N-n+1)}{nN}\right\}^2\right.$$

$$\left.\left\{\Sigma\left[\frac{(\Sigma x^t)_s - x_i^t}{[(\Sigma x)_s - x_i]^2}\right] + \sum_{i\neq j}\sum\left[\frac{(\Sigma x^t)_s - x_i^t - x_j^t}{[(\Sigma x)_s - x_i][(\Sigma x)_s - x_j]}\right]\right\}_s\right]$$

$$+ 2\delta E\left[(N-n)\left\{\frac{(\overset{N}{\Sigma} x_i)-(\Sigma x)_s(N-n+1)}{nN^2}\right\}\left\{\Sigma\left[\frac{(\Sigma x^t)_s - x_i^t}{(\Sigma x)_s - x_i}\right]\right\}_s\right] + \frac{\delta \overset{N}{\Sigma} x_i^t}{N^2} \qquad (2.29)$$

$$\varepsilon MSE(\bar{y}_Q) = \alpha^2 E\left[\frac{n^2(N-n+1)(\overset{N}{\Sigma} x_i) - N^2(\Sigma x)_s}{N^2(\Sigma x)_s} - \left\{\frac{(N-n)(n-1)^2(\overset{N}{\Sigma} x_i)}{nN^2}\right\}\right.$$

$$\left.\left\{\Sigma\left[\frac{1}{(\Sigma x)_s - x_i}\right]\right\}_s\right]^2 + \delta E\left[(\Sigma x^t)_s\left\{\frac{n(N-n+1)(\overset{N}{\Sigma} x_i)}{N^2(\Sigma x)_s} - \frac{2}{N}\right\}\left\{\frac{n(N-n+1)(\overset{N}{\Sigma} x_i)}{N^2(\Sigma x)_s}\right\}\right]$$

$$+ \delta\left\{\frac{(N-n)(n-1)(\overset{N}{\Sigma} x_i)}{nN^2}\right\}^2 E\left[\left\{\Sigma\left[\frac{(\Sigma x^t)_s - x_i^t}{[(\Sigma x)_s - x_i]^2}\right] + \underset{i\neq j}{\Sigma\,\Sigma}\left[\frac{(\Sigma x^t)_s - x_i^t - x_j^t}{[(\Sigma x)_s - x_i][(\Sigma x)_s - x_j]}\right]\right\}_s\right]$$

$$- 2\delta E\left[\left\{\frac{(\overset{N}{\Sigma} x_i)(N-n)(n-1)[n(\overset{N}{\Sigma} x_i)(N-n+1) - N(\Sigma x)_s}{(\Sigma x)_s nN^4}\right\}\left\{\Sigma\left[\frac{(\Sigma x^t)_s - x_i^t}{(\Sigma x)_s - x_i}\right]\right\}_s\right]$$

$$+ \delta\,\frac{\overset{N}{\Sigma} x_i^t}{N^2} \qquad (2.30)$$

We now derive the average expectations and average MSE's of the three variance estimators. Under our model, (1.17), (1.18) and (1.19) reduce to

$$v_1(\bar{y}_r) = \frac{(N-n)}{nN}\left[s_x^2\,\frac{(\alpha+\bar{u})^2}{\bar{x}^2} - 2s_{xu}\,\frac{(\alpha+\bar{u})}{\bar{x}} + s_u^2\right] \qquad (2.31)$$

$$v_2(\bar{y}_r) = \frac{\bar{X}^2}{\bar{x}^2}\,\frac{(N-n)}{nN}\left[s_x^2\,\frac{(\alpha+\bar{u})}{\bar{x}^2} - 2s_{xu}\,\frac{(\alpha+\bar{u})}{\bar{x}} + s_u^2\right] \qquad (2.32)$$

$$v_3(\bar{y}_r) = \bar{X}^2 \frac{(N-n)(n-1)}{n^2N}\left[(n-1)\left\{\Sigma\left(\frac{\alpha+\bar{u}_j'}{\bar{x}_j'}\right)^2\right\}_s\right.$$

$$\left.-\left\{\sum_{i\neq j}\Sigma\left(\frac{(\alpha+\bar{u}_i')(\alpha+\bar{u}_j')}{\bar{x}_i'\bar{x}_j'}\right)\right\}_s\right]$$

$$= \bar{X}^2 \frac{(N-n)(n-1)}{n^2N}\left[\alpha^2\left((n-1)\left\{\Sigma\left(\frac{1}{\bar{x}_j'^2}\right)\right\}_s - \left\{\sum_{i\neq j}\Sigma\left(\frac{1}{\bar{x}_i'\bar{x}_j'}\right)\right\}_s\right)\right.$$

$$\left.+ (n-1)\left\{\Sigma\left(\frac{\bar{u}_j'^2}{\bar{x}_j'^2}\right)\right\}_s - \left\{\sum_{i\neq j}\Sigma\left(\frac{\bar{u}_i'\bar{u}_j'}{\bar{x}_i'\bar{x}_j'}\right)\right\}_s\right] + \text{terms involving } \alpha \text{ and } \bar{u}_i' \text{ whose expectation is zero.} \quad (2.33)$$

Therefore

$$\xi v_1(\bar{y}_r) = \frac{(N-n)}{nN(n-1)}\left\{n\alpha^2 E\left[\frac{n(\Sigma x^2)_s-(\Sigma x)_s^2}{(\Sigma x)_s^2}\right]\right.$$

$$\left.+ \delta E\left[(\Sigma x^t)_s\left\{\frac{(\Sigma x^2)_s+(\Sigma x)_s^2}{(\Sigma x)_s^2}\right\}\right] - 2\delta E\left[\frac{(\Sigma x^{t+1})_s}{(\Sigma x)_s}\right]\right\} \quad (2.34)$$

$$\xi v_2(\bar{y}_r) = \frac{\bar{X}^2}{\bar{x}^2}\,\xi v_1(\bar{y}_r) \quad (2.35)$$

$$\xi v_3(\bar{y}_r) = \bar{X}^2 \frac{(N-n)(n-1)}{n^2N}\left\{(n-1)^2\alpha^2 E\left[(n-1)\left\{\Sigma\left(\frac{1}{[(\Sigma x)_s-x_i]^2}\right)\right\}_s\right.\right.$$

$$\left.-\left\{\sum_{i\neq j}\Sigma\left(\frac{1}{[(\Sigma x)_s-x_i][(\Sigma x)_s-x_j]}\right)\right\}_s\right] + \delta E\left[\frac{(n-1)}{n}\left\{\Sigma\left(\frac{(\Sigma x^t)_s-x_i^t}{[(\Sigma x)_s-x_i]^2}\right)\right\}_s\right.$$

$$-\left\{\sum_{i\neq j}\sum\left(\frac{(\Sigma x^t)_s-x_i^t-x_j^t}{[(\Sigma x)_s-x_i][(\Sigma x)_s-x_j]}\right)\right\}_s\Bigg]\Bigg\}. \quad (2.36)$$

Turning to the evaluations of the average MSE's of the variance estimators, we further assume that the u_i are independently and normally distributed so that

$$\varepsilon(u_i^4|x_i) = 3\delta^2 x_i^{2t} \quad (2.37)$$

$$\varepsilon(u_i^2 u_j^2|x_i,\ x_j) = \delta^2 x_i^t x_j^t \qquad i \neq j = 1,\ 2, \ldots, N. \quad (2.38)$$

Now

$$v_1^2(\bar{y}_r) = \frac{(N-n)^2}{n^2N^2}\left[s_x^4\ \frac{(\alpha+\bar{u})^4}{\bar{x}^4} + 4s_{xu}^2\ \frac{(\alpha+\bar{u})^2}{\bar{x}^2} + s_u^4\right.$$

$$\left. - 4s_x^2 s_{xu}\ \frac{(\alpha+\bar{u})^3}{\bar{x}^3} + 2s_x^2 s_u^2\ \frac{(\alpha+\bar{u})^2}{\bar{x}^2} - 4s_u^2 s_{xu}\ \frac{(\alpha+\bar{u})}{\bar{x}}\right]$$

$$= \frac{(N-n)^2}{n^2N^2}\left[\frac{s_x^4}{\bar{x}^4}\ (\alpha^4+6\alpha^2\bar{u}^2+\bar{u}^4) + \frac{4s_{xu}^2}{\bar{x}^2}\ (\alpha^2+\bar{u}^2) + s_u^4 - \frac{4s_x^2 s_{xu}}{\bar{x}^3}\ (3\alpha^2\bar{u}+\bar{u}^3)\right.$$

$$\left. + \frac{2s_x^2s_u^2}{\bar{x}^2}\ (\alpha+\bar{u}^2) - 4\ \frac{s_{xu}s_u^2}{\bar{x}}\ \bar{u}\right] + \text{terms whose average expectation is zero.} \quad (2.39)$$

$$v_3^2(\bar{y}_r) = \frac{\bar{X}^4(N-n)^2(n-1)^2}{n^4N^2}\{\alpha^4[(n-1)(\Sigma\bar{x}_j'^{-2})_s - (\sum_{i\neq j}\Sigma\bar{x}_i'^{-1}\bar{x}_j'^{-1})_s]^2$$

$$+ (n-1)^2(\Sigma\bar{u}_j'^2\bar{x}_j'^{-2})_s^2 + (\sum_{i\neq j}\Sigma\bar{u}_i'\bar{u}_j'\bar{x}_i'^{-1}\bar{x}_j'^{-1})_s^2 + \alpha^2[4(n-1)^2(\Sigma\bar{u}_j'\bar{x}_j'^{-2})_s^2$$

$$+ (\sum_{i\neq j}\Sigma\{\bar{u}_i'+\bar{u}_j'\}\bar{x}_i'^{-1}\bar{x}_j'^{-1})_s^2 - 4(n-1)(\Sigma\bar{u}_j'\bar{x}_j'^{-2})_s(\sum_{i\neq j}\Sigma\{\bar{u}_i'+\bar{u}_j'\}\bar{x}_i'^{-1}\bar{x}_j'^{-1})_s]$$

$$- 2(n-1)(\Sigma \bar{u}_j'^{2} \bar{x}_j'^{-2})_s \Big(\sum_{i\neq j} \Sigma \bar{u}_i'\bar{u}_j'\bar{x}_i'^{-1}\bar{x}_j'^{-1}\Big)_s + 2\alpha^2\Big[(n-1)^2(\Sigma\bar{x}_j'^{-2})_s(\Sigma\bar{u}_j'^{2}\bar{x}_j'^{-2})_s$$

$$- (n-1)(\Sigma\bar{x}_j'^{-2})_s\Big(\sum_{i\neq j}\Sigma \bar{u}_i'\bar{u}_j'\bar{x}_i'^{-1}\bar{x}_j'^{-1}\Big)_s - (n-1)\Big(\sum_{i\neq j}\Sigma\bar{x}_i'^{-1}\bar{x}_j'^{-1}\Big)_s(\Sigma\bar{u}_j'^{2}\bar{x}_j'^{-2})_s$$

$$+ \Big(\sum_{i\neq j}\Sigma \bar{x}_i'^{-1}\bar{x}_j'^{-1}\Big)_s\Big(\sum_{i\neq j}\Sigma\bar{u}_i'\bar{u}_j'\bar{x}_i'^{-1}\bar{x}_j'^{-1}\Big)_s\Big]\Big\} + \text{terms whose average expectation is zero.} \qquad (2.40)$$

Using the formulae given in Appendices I and II, we get, after considerable simplification,

$$\xi v_1^2(\bar{y}_r) = \frac{(N-n)^2}{n^2N^2}\Bigg\{\alpha^4 E\left[\frac{n^4 s_x^4}{(\Sigma x)_s^4}\right] + 2\alpha^2\delta E\Bigg[\frac{2n^2(\Sigma x^{t+2})_s}{(n-1)^2(\Sigma x)_s^2}$$

$$- \frac{2n(\Sigma x^{t+1})_s}{(n-1)(\Sigma x)_s}\left(\frac{2}{n-1} + \frac{3ns_x^2}{(\Sigma x)_s^2}\right) + (\Sigma x^t)_s\left(\frac{2}{(n-1)^2} + \frac{ns_x^2}{(\Sigma x)_s^2}\left\{\frac{3ns_x^2}{(\Sigma x)_s^2} + \frac{6}{n-1} + 1\right\}\right)\Bigg]$$

$$+ \delta^2 E\Bigg[\frac{4(\Sigma x^t)_s(\Sigma x^{t+2})_s}{(n-1)^2(\Sigma x)_s^2} - \frac{8(\Sigma x^{2t+1})_s}{(n-1)^2(\Sigma x)_s} + (\Sigma x^t)_s^2\left(\frac{3s_x^4}{(\Sigma x)_s^4} + \frac{n^2+2n+3}{n^2(n-1)^2}\right.$$

$$\left. + \frac{2(n+3)s_x^2}{n(n-1)(\Sigma x)_s^2}\right) + \frac{2(\Sigma x^{2t})_s}{n-1}\left(\frac{n+2}{n(n-1)} + \frac{2s_x^2}{(\Sigma x)_s^2}\right) + \frac{4(\Sigma x^{t+1})_s}{(n-1)(\Sigma x)_s}$$

$$\left(\frac{2(\Sigma x^{t+1})_s - (\Sigma x)_s(\Sigma x^t)_s}{(n-1)(\Sigma x)_s} - \frac{3(\Sigma x^t)_s}{n(n-1)} - \frac{3s_x^2(\Sigma x^t)_s}{(\Sigma x)_s^2}\right)\Bigg]\Bigg\}. \qquad (2.41)$$

$$\xi v_2^2(\bar{y}_r) = \frac{\bar{X}^4}{\bar{x}^4}\,\xi v_1^2(\bar{y}_r). \qquad (2.42)$$

$$\xi(v_3^2) = \frac{\bar{X}^4 (N-n)^2 (n-1)^2}{n^4 N^2} \Bigg\{ (n-1)^4 \alpha^4 E \Bigg[(n-1) \Bigg\{ \Sigma \Bigg(\frac{1}{[(\Sigma x)_s - x_j]^2} \Bigg) \Bigg\}_s$$

$$- \Bigg\{ \sum_{i \neq j} \sum \Bigg(\frac{1}{[(\Sigma x)_s - x_i][(\Sigma x) - x_j]} \Bigg) \Bigg\}_s \Bigg]^2 + \delta^2 E \Bigg[3(n-1)^2 \Bigg\{ \Sigma \Bigg(\frac{(\Sigma x^t)_s^2 - 2x_j^t (\Sigma x^t)_s + x_j^{2t}}{[(\Sigma x^t)_s - x_j]^4} \Bigg) \Bigg\}_s$$

$$+ [2+(n-1)^2] \Bigg\{ \sum_{i \neq j} \sum \Bigg(\frac{(\Sigma x^t)_s [3(\Sigma x^t)_s - 5x_i^t - 5x_j^t] + 2[x_i^{2t} + x_j^{2t}] + 5x_i^t x_j^t}{[(\Sigma x)_s - x_i^t]^2 [(\Sigma x)_s - x_j^t]^2} \Bigg) \Bigg\}_s$$

$$- 12(n-1) \Bigg\{ \sum_{i \neq j} \sum \Bigg(\frac{(\Sigma x^t)_s [(\Sigma x^t)_s - 2x_i^t - x_j^t] + x_i^{2t} + x_i^t x_j^t}{[(\Sigma x)_s - x_i]^3 [(\Sigma x)_s - x_j]} \Bigg) \Bigg\}_s$$

$$- 2(n-3) \Bigg\{ \sum_{i \neq j \neq k} \sum \sum \Bigg(\frac{(\Sigma x^t)_s [3(\Sigma x^t)_s - 5x_i^t - 3x_j^t - 3x_k^t] + 2[x_i^{2t} + x_j^t x_k^t] + 3x_i^t [x_j^t + x_k^t]}{[(\Sigma x)_s - x_i]^2 [(\Sigma x)_s - x_j][(\Sigma x)_s - x_k]} \Bigg) \Bigg\}_s$$

$$+ \Bigg\{ \sum_{i \neq j \neq k \neq \ell} \sum \sum \sum \Bigg(\frac{3(\Sigma x^t)_s [(\Sigma x^t)_s - x_i^t - x_j^t - x_k^t - x_\ell^t] + 2x_i^t [x_j^t + x_k^t + x_\ell^t] + 2x_j^t [x_k^t + x_\ell^t] + 2x_k^t x_\ell^t}{[(\Sigma x)_s - x_i][(\Sigma x)_s - x_j][(\Sigma x)_s - x_k][(\Sigma x)_s - x_\ell]} \Bigg) \Bigg\}_s \Bigg]$$

$$+ \alpha^2 \delta E \Bigg[6(n-1)^4 \Bigg\{ \Sigma \Bigg(\frac{(\Sigma x^t)_s - x_j^t}{[(\Sigma x)_s - x_j]^4} \Bigg) \Bigg\}_s + 4[(n-1)^2 + (n-1)^4]$$

$$\Bigg\{ \sum_{i \neq j} \sum \Bigg(\frac{(\Sigma x^t)_s - x_i^t - x_j^t}{[(\Sigma x)_s - x_i]^2 [(\Sigma x)_s - x_j]^2} \Bigg) \Bigg\}_s + 2(n-1)^2 \Bigg\{ \sum_{i \neq j} \sum \Bigg(\frac{4(\Sigma x^t)_s - 3[x_i^t + x_j^t]}{[(\Sigma x)_s - x_i][(\Sigma x)_s - x_j]^2} \Bigg) \Bigg\}_s$$

$$+ 2(n-1)^4 \Bigg\{ \sum_{i \neq j} \sum \Bigg(\frac{(\Sigma x^t)_s - x_j^t}{[(\Sigma x)_s - x_i]^2 [(\Sigma x)_s - x_j]^2} \Bigg) \Bigg\}_s$$

$$- 8(n-1)^3 \left\{ \sum_{i \neq j} \sum \left[\frac{2(\Sigma x^t)_s - 2x_i^t - x_j^t}{[(\Sigma x)_s - x_i]^3 [(\Sigma x)_s - x_j]} \right] \right\}_s - 4(n-1)^3$$

$$\left\{ \sum_{i \neq j} \sum \left[\frac{(\Sigma x^t)_s - x_i^t - x_j^t}{[(\Sigma x)_s) - x_i]^3 [(\Sigma x)_s - x_j]} \right] \right\}_s - 4(n-1)^3 \left\{ \sum_{i \neq j} \sum \left[\frac{(\Sigma x^t)_s - x_j^t}{[(\Sigma x)_s - x_i][(\Sigma x)_s - x_j]^3} \right] \right\}_s$$

$$+ 4(n-1)^2 \left\{ \sum_{i \neq j \neq k} \sum \sum \left[\frac{4(\Sigma x^t)_s - 3x_i^t - 2[x_j^t + x_k^t]}{[(\Sigma x)_s - x_i]^2 [(\Sigma x)_s - x_j][(\Sigma x)_s - x_k]} \right] \right\}_s$$

$$- 4(n-1)^3 \left\{ \sum_{i \neq j \neq k} \sum \sum \left[\frac{2(\Sigma x^t)_s - 2x_i^t - x_j^t - x_k^t}{[(\Sigma x)_s - x_i]^2 [(\Sigma x)_s - x_j][(\Sigma x)_s - x_k]} \right] \right\}_s$$

$$- 2(n-1)^3 \left\{ \sum_{i \neq j \neq k} \sum \sum \left[\frac{(\Sigma x^t)_s - x_i^t - x_j^t}{[(\Sigma x)_s - x_i][(\Sigma x)_s - x_j][(\Sigma x)_s - x_k]^2} \right] \right\}_s$$

$$- 2(n-1)^3 \left\{ \sum_{i \neq j \neq k} \sum \sum \left[\frac{(\Sigma x^t)_s - x_j^t}{[(\Sigma x)_s - x_i][(\Sigma x)_s - x_j]^2 [(\Sigma x)_s - x_k]} \right] \right\}_s$$

$$+ 8(n-1)^2 \left\{ \sum_{i \neq j \neq k} \sum \sum \left[\frac{(\Sigma x^t)_s - x_i^t - x_j^t}{[(\Sigma x)_s - x_i]^2 [(\Sigma x)_s - x_j][(\Sigma x)_s - x_k]} \right] \right\}_s$$

$$+ 2(n-1)^2 \left\{ \sum_{i \neq j \neq k \neq \ell} \sum \sum \sum \left[\frac{2(\Sigma x^t)_s - x_i^t - x_j^t - x_k^t - x_\ell^t}{[(\Sigma x)_s - x_i][(\Sigma x)_s - x_j][(\Sigma x)_s - x_k][(\Sigma x)_s - x_\ell]} \right] \right\}_s$$

$$\left. \left. + 2(n-1)^2 \left\{ \sum_{i \neq j \neq k \neq \ell} \sum \sum \sum \left[\frac{(\Sigma x^t)_s - x_i^t - x_j^t}{[(\Sigma x)_s - x_i][(\Sigma x)_s - x_j][(\Sigma x)_s - x_k][(\Sigma x)_s - x_\ell]} \right] \right\}_s \right] \right\} \quad (2.42)$$

The identities given in Appendix III were used in the evaluation of the multiple-fold summations involved in $\xi(v_3^2)$.

The average bias of $v_i(\bar{y}_r)$ $(i = 1, 2, 3)$ as an estimator of $V(\bar{y}_r)$ is given by

$$\varepsilon B(v_i) = \xi v_i(\bar{y}_r) - \varepsilon V(\bar{y}_r). \tag{2.43}$$

The average MSE of $v_i(\bar{y}_r)$ is given by

$$\varepsilon MSE(v_i) = \xi[v_i(\bar{y}_r) - V(\bar{y}_r)]^2. \tag{2.44}$$

However, the evaluation of $\varepsilon MSE(v_i)$ is difficult. We therefore employ the alternative criterion

$$\begin{aligned} \varepsilon MSE^*(v_i) &= \xi[v_i(\bar{y}_r) - \varepsilon V(\bar{y}_r)]^2 \\ &= \xi[v_i(\bar{y}_r)]^2 - 2[\xi v_i(\bar{y}_r)][\varepsilon V(\bar{y}_r)] + [\varepsilon V(\bar{y}_r)]^2 \end{aligned} \tag{2.45}$$

which is readily evaluable by using the formulae for $\xi[v_i(\bar{y}_r)]^2$ and $\xi v_i(\bar{y}_r)$ derived above.

CHAPTER III

EMPIRICAL STUDY

In this thesis, we examine only six populations and three sample sizes ($n = 4$, 6 and 8). A more extensive study, however, will appear in a Technical Report of the Department of Statistics, University of Manitoba. Moreover, the conclusions presented here are similar to those derived from the more extensive study. Table 1 shows the source, nature of the character x, the population size N and the coefficient of variation of x, C_x. The populations are arranged according to increasing order of C_x, where N ranges from 49 to 270 and C_x ranges from 0.42 to 1.01.

To compute the coefficients of α^2 and δ, in the formulae for average MSE's or average variances of the estimators given in Chapter II, we need to draw all the $\binom{N}{n}$ possible samples of x-values. However, $\binom{N}{n}$ is very large for the population given here and, due to limitations on available computer time, we adopted the following procedure for computing the coefficients: From a given population of size N 2,000 independent samples, each of a given size n, were drawn at random without replacement and

$$\binom{N}{n}^{-1} \sum_{s=1}^{\binom{N}{n}} h(s) \text{ is approximated by } \sum_{s=1}^{2000} h(s)/2000,$$ where h(s) is a

function of the x-values in the sample s. Some preliminary calculations indicate that this approximation is very satisfactory for the comparison of the estimators or variance estimators.

TABLE 1

Description of the Population (Arranged in Increasing Order of C_x)

Pop. No.	Source	Nature of the Character	N	C_x
1	Murthy (1967) p. 131	Length of Timber	176	0.42
2	Biometrika (1959) Vol 46 p. 178	Gamma (h = 2)	100	0.59
3	Sukhatme (1970) p. 256	No. of Villages in a Circle	89	0.61
4	Biometrika (1959) Vol 46 p. 178	Log Normal	100	0.75
5	Kish (1965) p. 625	No. of Dwellings	270	0.99
6	Cochran (1963) p. 156	Size of Cities in U.S. in 1920	49	1.01

Comparison of the estimators

We denote the coefficients of α^2 and δ in the average MSE of the estimators by A_r, A_B, A_T, A_H, A_M, A_Q and by D_r, D_B, D_T, D_H, D_M, D_Q respectively, where the subscript in A or D indicates the estimator. Tables 2, 3 and 4 give the values of the ratios $E_{1\alpha} = A_H/A_M$, $E_{2\alpha} = A_r/A_T$, $E_{3\alpha} = A_B/A_T$, $E_{4\alpha} = A_M/A_T$ and $E_{5\alpha} = A_Q/A_T$ and of the ratios for $E_{1\delta},\ldots,E_{5\delta}$, where $E_{i\delta}$ is obtained from $E_{i\alpha}$ by replacing A by D ($i = 1,\ldots,5$). The following conclusions may be drawn from these tables: (1) Mickey's unbiased estimator is superior to the Hartley-Ross estimator as $E_{1\alpha}$ is considerably greater than 1 (especially when C_x is large), $E_{1\delta} > 1$ for $t = 0, 1$ ($E_{1\delta}$ is considerable for $t = 0$); for $t = 2$ we have $0.96 \leq E_{1\delta} \leq 1.00$ for the first four populations (with moderate C_x) and $0.91 \leq E_{1\delta} \leq 0.96$ for the populations 5 and 6 (with larger C_x). (2) As expected, the efficiencies of Beale's and Tin's are virtually equal, expecially when C_x is not large. (3) Tin's estimator is significantly better than Mickey's as $E_{4\alpha} > 1$ and, for all t, $E_{4\delta} \geq 1$. (4) For $t = 1$ and 2, Tin's estimator is better than Quenouille's as $E_{5\delta} > 1$ and $0.99 \leq E_{5\alpha} \leq 2.61$; for $t = 0$, $0.97 \leq E_{5\delta} \leq 1.67$. (5) Tin's estimator is slightly more efficient than the classical estimator as $E_{2\alpha} > 1$ and, for $t = 0$, $E_{2\delta} > 1$. However, this efficiency is reversed for $t = 1$ and 2, since, for $t = 1$, $0.95 \leq E_{2\delta} \leq 1.00$, and, for $t = 2$, $0.82 \leq E_{2\delta} \leq 0.98$. (6) As n increases, all the E-values (excepting E_1) tend to 1.

TABLE 2

Values of: $E_{1\alpha} = A_H/A_M$, $E_{2\alpha} = A_r/A_T$, $E_{3\alpha} = A_B/A_T$, $E_{4\alpha} = A_M/A_T$, $E_{5\alpha} = A_Q/A_T$

$E_{1\delta} = D_H/D_M$, $E_{2\delta} = D_r/D_T$, $E_{3\delta} = D_B/A_T$, $E_{4\delta} = D_M/D_T$, $E_{5\delta} = D_Q/D_T$

$n = 4$

Coefficient		Pop. No.	E_1	E_2	E_3	E_4	E_5	Pop. No.	E_1	E_2	E_3	E_4	E_5
α^2		1	1.96	1.35	1.02	1.22	0.99	4	2.30	1.22	1.00	1.48	1.11
δ	$t = 0$	1	1.13	1.10	1.01	1.02	0.99	4	1.79	1.19	1.02	1.06	0.99
δ	$t = 1$	1	1.01	0.99	1.00	1.00	1.01	4	1.10	0.97	0.99	1.06	1.05
δ	$t = 2$	1	0.99	0.92	0.99	1.01	1.02	4	0.96	0.84	0.97	1.16	1.12
α^2		2	2.88	1.28	1.01	1.30	1.02	5	6.09	1.75	1.09	2.58	2.61
δ	$t = 0$	2	1.59	1.16	1.01	1.02	0.99	5	7.85	1.44	1.06	1.51	1.67
δ	$t = 1$	2	1.04	0.98	0.99	1.01	1.02	5	1.52	0.95	0.98	1.12	1.23
δ	$t = 2$	2	0.98	0.88	0.98	1.05	1.05	5	0.91	0.82	0.95	1.27	1.27
α^2		3	1.55	1.16	1.00	1.25	1.05	6	9.06	1.09	0.98	1.72	1.23
δ	$t = 0$	3	1.14	1.14	1.01	1.01	0.99	6	11.58	1.16	1.01	1.12	1.01
δ	$t = 1$	3	1.02	0.98	0.99	1.02	1.02	6	1.30	0.97	0.99	1.14	1.07
δ	$t = 2$	3	0.97	0.87	0.98	1.06	1.06	6	0.95	0.82	0.96	1.37	1.18

TABLE 3

Values of: $E_{1\alpha} = A_H/A_M$, $E_{2\alpha} = A_r/A_T$, $E_{3\alpha} = A_B/A_T$, $E_{4\alpha} = A_M/A_T$, $E_{5\alpha} = A_Q/A_T$

$E_{1\delta} = D_H/D_M$, $E_{2\delta} = D_r/D_T$, $E_{3\delta} = D_B/D_T$, $E_{4\delta} = D_M/D_T$, $E_{5\delta} = D_Q/D_T$

$n = 8$

Coefficient		Pop. No.	E_1	E_2	E_3	E_4	E_5	Pop. No.	E_1	E_2	E_3	E_4	E_5
α^2			1.99	1.14	1.00	1.06	1.00		2.90	1.12	1.00	1.15	1.01
δ	t = 0	1	1.06	1.05	1.00	1.00	1.00	4	1.50	1.13	1.01	1.00	0.98
	t = 1		1.01	1.00	1.00	1.00	1.00		1.06	0.99	1.00	1.01	1.01
	t = 2		1.00	0.97	1.00	1.00	1.00		0.97	0.89	0.99	1.04	1.05
α^2			2.95	1.15	1.00	1.10	0.99		13.88	1.43	1.03	1.44	1.14
δ	t = 0	2	1.28	1.08	1.00	1.00	0.99	5	8.39	1.26	1.03	1.05	0.97
	t = 1		1.03	0.99	1.00	1.00	1.00		1.37	0.98	1.00	1.02	1.03
	t = 2		0.99	0.94	1.00	1.01	1.01		0.93	0.86	0.98	1.08	1.09
α^2			1.87	1.12	1.00	1.09	0.99		8.77	1.02	0.99	1.23	1.08
δ	t = 0	3	1.09	1.08	1.00	1.00	0.99	6	8.20	1.14	1.01	0.99	0.97
	t = 1		1.02	0.99	1.00	1.00	1.00		1.19	0.98	1.00	1.02	1.02
	t = 2		0.98	0.93	1.00	1.01	1.01		0.95	0.85	0.98	1.11	1.09

TABLE 4

Values of: $E_{1\alpha} = A_H/A_M$, $E_{2\alpha} = A_r/A_T$, $E_{3\alpha} = A_B/A_T$, $E_{4\alpha} = A_M/A_T$, $E_{5\alpha} = A_Q/A_T$,

$E_{1\delta} = A_H \; A_M$, $E_{2\delta} = A_r \; A_T$, $E_{3\delta} = A_B \; A_T$, $E_{4\delta} = A_M \; A_T$, $E_{5\delta} = A_Q \; A_T$

$n = 12$

Coefficient	Pop. No.	E_1	E_2	E_3	E_4	E_5	Pop. No.	E_1	E_2	E_3	E_4	E_5
α^2	1	1.94	1.09	1.00	1.04	1.00	4	3.33	1.08	1.00	1.09	1.00
δ, t = 0	1	1.04	1.03	1.00	1.00	1.00	4	1.34	1.09	1.00	1.00	0.99
δ, t = 1	1	1.01	1.00	1.00	1.00	1.00	4	1.05	0.99	1.00	1.00	1.00
δ, t = 2	1	1.00	0.98	1.00	1.00	1.00	4	0.98	0.92	1.00	1.02	1.02
α^2	2	2.88	1.08	1.00	1.06	1.00	5	17.14	1.29	1.02	1.19	0.97
δ, t = 0	2	1.18	1.05	1.00	1.00	1.00	5	6.20	1.17	1.01	1.01	0.97
δ, t = 1	2	1.02	1.00	1.00	1.00	1.00	5	1.26	0.99	1.00	1.00	1.01
δ, t = 2	2	1.00	0.96	1.00	1.00	1.00	5	0.96	0.90	0.99	1.03	1.04
α^2	3	1.98	1.08	1.00	1.05	1.00	6	7.74	1.05	1.00	1.11	1.01
δ, t = 0	3	1.06	1.05	1.00	1.00	1.00	6	5.59	1.10	1.00	0.99	0.98
δ, t = 1	3	1.01	1.00	1.00	1.00	1.00	6	1.13	0.99	1.00	1.01	1.01
δ, t = 2	3	0.99	0.95	1.00	1.00	1.01	6	0.96	0.89	0.99	1.04	1.04

The above conclusions are remarkably similar to the analytical results obtained by Rao and Rao (1971) under the assumption of a gamma distribution for x and infinite population size N.

Comparison of the variance estimators

Tables 5, 6 and 7 give the values of the coefficients of α^2 and δ in the biases of v_1, v_2 and v_3 as estimators of $V(\bar{y}_r)$. We denote the coefficients of α^2 by $-C_{1\alpha}$, $-C_{2\alpha}$ and $C_{3\alpha}$ and of δ by $-C_{1\delta}$, $-C_{2\delta}$ and $C_{3\delta}$ for v_1, v_2 and v_3 respectively. We draw the following conclusions from these tables: (1) v_1 consistently underestimates $V(\bar{y}_r)$ whereas v_3 consistently overestimates $V(\bar{y}_r)$ for all t and v_2 for $t = 0$ only. (2) $C_{1\alpha} > C_{2\alpha}$ and $C_{1\delta} > C_{2\delta}$ for $t = 1$ and $C_{1\delta} + C_{2\delta} > 0$ for $t = 0$ so that $|B(v_1)| > |B(v_2)|$ for $t = 0$ and 1. For $t = 2$, $C_{1\delta} < C_{2\delta}$ so that the comparison of $|B(v_1)|$ with $|B(v_2)|$ depends on the value of α^2/δ unless $\alpha \doteq 0$ (i.e., regression approximately through the origin). (3) $C_{3\alpha} > C_{2\alpha}$ and $C_{3\delta} > -C_{2\delta}$ for $t = 0$, $C_{3\delta} > C_{2\delta}$ for $t = 1$ so that $|B(v_2)| < |B(v_3)|$ for $t = 0$ and 1; for $t = 2$, $C_{3\delta} < C_{2\delta}$ so that the comparison depends on α^2/δ. (4) For $n \geq 8$, $C_{3\delta} < C_{1\delta}$ for all t and $C_{3\alpha} < C_{1\alpha}$ or $C_{3\alpha} \doteq C_{1\alpha}$ (excepting for $n = 8$ and population 5) so that v_3 may be preferable to v_1 with regard to absolute bias.

We now turn to the mean square errors of the variance estimators v_1, v_2 and v_3. Here we confine ourselves to examining only the δ^2 term in the m.s.e. (denoted by F_1, F_2 and F_3) given in Tables 8, 9 and 10 (assuming $\alpha \doteq 0$). It is clear from these tables that