VARIANCE APPROXIMATIONS FOR ESTIMATED PROPORTIONS WITHIN A PROPER SUBSET OF A STRATIFIED RANDOM SAMPLE

Margaret C. Wu, Social Security Administration

Motivated by stochastic modeling and estimation in disability studies--where on numerous occasions proportions of mutually exclusive categories within a proper subset of a stratified random sample have to be estimated--combined, Beale's and modified ratio estimators are derived under the stratified sampling scheme. Improved approximations of the biases and variances of the estimators are developed. The estimators are then compared in terms of bias and efficiency.

1. Introduction and Notation

Let simple random samples of sizes n_h be selected from populations of sizes N_h from strata h=1,2,...,L. We assume that within the sample from stratum h there is a subset of size $n_{h.k}$ possessing characteristic k and that this subset can be further partioned into M mutually exclusive classes of sizes n_{hjk}, for j=1,...,M and h=1,...,L. Although the numbers of individuals selected, $n_1, \ldots, n_L$, will have been fixed in advance, the sizes of individuals possessing the characteristic k namely, $n_{1.k}, \ldots, n_{L.k}$ will vary from sample to sample.

Let $N_{h.k}$ and N_{hjk} denote the unknown population sizes corresponding to $n_{h.k}$ and n_{hjk}, respectively. Define,

$$N_{..k} = \sum_{h=1}^{L} N_{h.k},\ N_{.jk} = \sum_{h=1}^{L} N_{hjk} \text{ and}$$

$$N = \sum_{h=1}^{L} N_h. \tag{1.1}$$

Based on the stratified random sample of size $n = \sum_{h=1}^{L} n_h$, described above, we wish to estimate the proportions of the sizes of those possessing characteristic k and also falling into class j, within the total population to $N_{..k}$, i.e.,

$$P_{j/k} = N_{.jk} / N_{..k}, \text{ for } j=1,\ldots,M \tag{1.2}$$

It is well known [11, p. 22] that,

$$\hat{P}_{stjk} = \sum_{h=1}^{L} N_h (n_{hjk} / n_h) / N \text{ and } \hat{P}_{st.k} = \sum_{h=1}^{L} N_h (n_{h.k} / n_h) / N,$$

are unbiased estimators of $P_{jk} = N_{.jk} / N$ and $P_{.k} = N_{..k} / N$, respectively. Since the variances of $\hat{P}_{stjk}$ and $\hat{P}_{st.k}$ tend to zeros as $n \to \infty$ and $N \to \infty$, $\hat{P}_{stjk}$ and $\hat{P}_{st.k}$ are consistent estimators. Therefore, the estimator

$$r = \hat{P}_{stjk} / \hat{P}_{st.k} \tag{1.3}$$

is a consistent estimator of $P_{j/k}$. Since $\{n_{h.k},\ 1 \le h \le L\}$ vary from sample to sample, r can be treated as a combined ratio estimator.

The estimator r is generally biased. Other ratio estimators had been developed in the literatures to reduce the bias. Four of these estimators, simple ratio, Quenouille's ratio (t_1), Beale's ratio (t_2) and modified ratio (t_3) had been compared under the simple random sampling scheme [12]. It was concluded that the three estimators, t_1, t_2, and t_3 were attractive from theoretical as well as computational point of view. Comparison of the variances, to $O(n^{-2})$, showed that t_2 and t_3 were the most efficient estimators; where n was the size of the simple random sample.

In this paper the combined, Beale's and modified ratio estimators are developed, under the stratified random sampling scheme for proportions. Their biases and variances are derived to the order of $O(n^{-2})$. We assume that $n/n_h \to u_h$ (bounded), as n and n_h becoming large, for all h. Comparisons of the estimators are then made in terms of efficiency and bias. A first approximation to the order of $O(n^{-1})$ of the variance and bias of an estimated domain mean (treated as a combined ratio estimator) based on a stratified random sample, had been derived [1, P. 148;3]. A second approximation to the variance of a simple ratio estimator, by using Taylor's series expansion and retaining terms to $O(n^{-3})$, based on a simple random sample, had also been established [1].

When the domain of the study is an event of infrequent occurrence, the first approximation does not always have a known accuracy, it is natural to consider improving the approximation by including some higher order terms. Since variance approximations to $O(n^{-1})$ of the combined ratio, Beale's ratio and modified ratio estimators are the same, the improved approximations to $O(n^{-2})$ of the variances are also necessary for the efficiency comparison.

2. The Estimators and Their Biases.

Introduce random variables y_{hi} and x_{hi} for $h=1,\dots,L$, and $i=1,\dots, n_h$ such that

$$y_{hi}=\begin{cases}1 & \text{if the ith individual of the sample from stratum h possesses characteristic k and also falls into class j;}\\ 0 & \text{otherwise;}\end{cases}$$

$$x_{hi}=\begin{cases}1 & \text{if the ith individual of the sample from stratum h possesses characteristic k,}\\ 0 & \text{otherwise.}\end{cases}$$

Thus,

$$r=[(\sum_{h=1}^{L} N_h \bar{y}_h) / N] / [(\sum_{h=1}^{L} N_h \bar{x}_h) / N]$$

$$= \bar{y}_{st} / \bar{x}_{st} \qquad (2.1)$$

is a combined ratio estimator, where

$$\bar{x}_h = \Sigma_{i=1}^{n} x_{hi} / n_h = n_{h.k} / n_h \quad \text{and}$$

$$\bar{y}_h = n_{hjk} / n_h.$$

Also, let Y_{hi}, X_{hi} be the population values corresponding to y_{hi} and x_{hi};

$$\bar{X} = \sum_{h=1}^{L} \Sigma_{i=1}^{N_h} X_{hi} / N=N_{..k} / N,$$

$$\bar{Y} = \sum_{h=1}^{L} \Sigma_{1=1}^{N_h} Y_{hi} / N = N_{.jk} / N,$$

$$\bar{X}_h = \Sigma_{i=1}^{N_h} X_{hi} / N_h, \ \bar{Y}_h = \Sigma_{i=1}^{N_h} Y_{hi} / N_h,$$

$$U_h = N_h / N, \ f_h = \frac{1}{n_h} - \frac{1}{n_h} \quad \text{and}$$

$$g_h = \frac{1}{n_h^2} - \frac{1}{N_h^2} .$$

Therefore, $P_{j/k} = \bar{Y} / \bar{X}$

A first approximation to $O(n^{-1})$ of the bias of r, $B_1(r)$, is

$$B_1(r) = P_{j/k}\{\sum_{h=1}^{L} U_h^2 f_h [\frac{S_{hx}^2}{\bar{X}^2} - \frac{S_{hxy}}{\bar{X}\,\bar{Y}}]\} . \qquad (2.2)$$

The Beale's and the modified ratio estimators, t_2 and t_3, can be expressed as

$$t_2= r [1+ \sum_{h=1}^{L} U_h^2 f_h s_{hxy} / (\bar{x}_{st} \bar{y}_{st})] / (1+ \sum_{h=1}^{L} U_h^2 f_h s_{hx}^2 / \bar{x}_{st}^2) \quad \text{and}$$

$$t_3= r \{1+ \sum_{h=1}^{L} U_h^2 f_h [s_{hxy} / (\bar{x}_{st} \bar{y}_{st}) - (s_{hx}^2 / \bar{x}_{st}^{\,2})]\} . \qquad (2.3)$$

In (2.2) and (2.3),

$$S_{hx}^2 = \Sigma_{i=1}^{N_h} (X_{hi} - \bar{X}_h) / (N_h - 1),$$

$$S_{hxy} = \Sigma_{i=1}^{N_h} (X_{hi} - \bar{X}_h)(Y_{hi} - \bar{Y}) / (N_h - 1), \qquad (2.4)$$

$$s_{hx}^2 = \Sigma_{i=1}^{n_h} (x_{hi} - \bar{x}_h)^2 / (n_h - 1) \quad \text{and}$$

$$s_{hxy} = \Sigma_{i=1}^{n_h} (x_{hi} - \bar{x}_h)(y_{hi} - \bar{y}_h) / (n_h - 1).$$

We assume that when the sample size n is sufficiently large and hence n_h for $1\le h\le L$, are large

$$|(\bar{x}_h - \bar{X}) / \bar{X}|<1 \text{ and hence}$$

$$|(\bar{x}_{st} - \bar{X}) / \bar{X}|<1 . \qquad (2.5)$$

Condition (2.5) is satisfied when

$$n_h> (N-n_h)(\bar{X} - \bar{X}_h') / \bar{X} \quad \text{if } \bar{X} > \bar{X}_h' \text{ and}$$

$$n_h> (N-n_h)(\bar{X}_h' - \bar{X}) / \bar{X} \quad \text{if } \bar{X} < \bar{X}_h'$$

where $\bar{X}_h' = (N\bar{X} - n_h \bar{x}_h) / (N-n_h)$.

For notational simplicity define,

$$\bar{e}_x = \bar{x}_{st} - \bar{X},\ \bar{e}_y = \bar{y}_{st} - \bar{Y},$$

$$\bar{e}_{hx} = \bar{x}_h - \bar{X}_h,\ \bar{e}_{hy} = \bar{y}_h - \bar{Y},$$

$$N_h\ U_{hab} = \sum_{i=1}^{N_h} (Y_{hi} - Y_h)^a\ (X_{hi} - X_h)^b \quad \text{and}$$

$$C_{hab} = U_{hab} / Y^a X^b\ ,$$

for constants a and b. Thus, r can be expressed as

$$r = P_{j/k}\ [1 + \bar{e}_y][1 + \bar{e}_x]^{-1}$$

$$= P_{j/k}\ [1 + \bar{e}_y][\sum_{i=0}^{\infty} (-1)^i\ \bar{e}_x{}^i]\ .$$

The necessary expected values, some of which due to Kendall and Stuart [9, Chap, 13] and [12] are listed below:

$$E\ \bar{e}_{hx}^3 = (g_h - 3f_h / N_h)\ U_{h03}$$

$$E\ \bar{e}_{hx}^2\ \bar{e}_{hy} = (g_h - 3f_h / N_h)\ U_{h12} \tag{2.7}$$

$$E\ \bar{e}_{hx}^4 = 3f_h^2\ U_{h02}^2 + 0(n^{-3})$$

$$E\ \bar{e}_{hx}^2\ \bar{e}_{hy}^2 = f_h^2\ (U_{h20}\ U_{h02} + 2U_{h11}^2)$$

$$E\ \bar{e}_{hx}^3\ \bar{e}_{hy} = 3f_h^2\ U_{h02}\ U_{h11} + 0(n^{-3})$$

$$E\ \bar{e}_{hx}\ (s_{hx}^2 - S_{hx}^2) = f_h\ U_{h03}$$

$$E\ \bar{e}_{hy}\ (s_{hx}^2 - S_{hx}^2) = F_h\ U_{h12}$$

$$E\ \bar{e}_{hx}\ (s_{hxy} - S_{hxy}) = f_h\ U_{h12}$$

$$E\ \bar{e}_{hy}\ (s_{hxy} - S_{hxy}) = f_h\ U_{h21}\ \ ;$$

$$E\ \bar{e}_x^4 = N^{-4} \{ \sum_{h=1}^{L} N_h^4\ E\ \bar{e}_{hx}^4 + 3 \sum_{\substack{h,h'=1 \\ h\neq h'}}^{L} N_h^2\ N_{h'}^2 E\ \bar{e}_{hx}^2\ E\ \bar{e}_{h'x}^2 \}\ , \tag{2.8}$$

$$E\ \bar{e}_y\ \bar{e}_x^3 = N^{-4} \{ \sum_{h=1}^{L} N_h^4\ E\ \bar{e}_{hy}\ \bar{e}_{hx}^3 + 3 \sum_{\substack{h,h'=1 \\ h\neq h'}}^{L} N_h^2\ N_{h'}^2 E\ \bar{e}_{hx}\ \bar{e}_{hy}\ E\ \bar{e}_{h'x}^2 \} \quad \text{and}$$

$$E\ \bar{e}_y^2\ \bar{e}_x^2 = N^{-4} \{ \sum_{h=1}^{L} N_h^4\ E\ \bar{e}_{hy}^2\ \bar{e}_{hx}^2 + \sum_{\substack{h,h'=1 \\ h\neq h'}}^{L} N_h^2\ N_{h'}^2 [2\ E\ \bar{e}_{hx}\ \bar{e}_{hy}\ E\ \bar{e}_{h'x}\ \bar{e}_{h'y} + E\ \bar{e}_{hx}^2\ E\ \bar{e}_{h'y}^2\ ]\}\ .$$

Using (2.6), (2.7) and (2.8) the expectations of r, t_2 and t_3 to $0(n^{\ 2})$, under assumption (2.5), are obtained as:

$$\begin{aligned} E(r) = P_{j/k}\{1 + \sum_{h=1}^{L} [&U_h^2\ f_h\ (C_{h02} - C_{h11}) \\ &+ U_h^3\ (g_h - 3f_h / N_n)(C_{h12} - C_{h03}) \\ &+ 3\ U_h^4\ f_h^2\ C_{h02}\ (C_{h02} - C_{h11}) \\ &+ 3 \sum_{\substack{h,h'=1 \\ h\neq h'}}^{L} U_h^2\ U_{h'}^2 f_h\ f_{h'} C_{h02} (C_{h'02} - C_{h'11}) \} \end{aligned} \tag{2.9}$$

$$\begin{aligned} E(t_2) = P_{j/k}\{1 - \sum_{h=1}^{L} U_h^3 (2g_h - 3f_h/N_h)(C_{h12} - C_{h03}) \\ + 2U_h^4\ f_h\ C_{h02}\ (C_{h02} - C_{h11})] \\ -2 \sum_{\substack{h,h'=1 \\ h\neq h'}}^{L} U_h^2 U_{h'}^2 f_h\ f_{h'} C_{h02} (C_{h'02} - C_{h'11}) \end{aligned}$$

and

$$\begin{aligned} E(t_3) = P_{j/k}\{1 - \sum_{h=1}^{L} [U_h^3 (2g_h - 3f_h/N_h)(C_{h12} - C_{h03}) \\ + 3U_h^4\ f_h^2\ C_{h02}\ (C_{h02} - C_{h11})] \\ - 3 \sum_{\substack{h,h'=1 \\ h\neq h'}}^{L} U_h^2\ U_{h'}^2 f_h\ f_{h'} C_{h02}\ (C_{h'02} - C_{h'11})\}. \end{aligned}$$

For the case of proportions, let

$$P_{h.k} = N_{h.k} / N_h,\ P_{hjk} = N_{hjk} / N_h,$$

$$P_{hj/k} = N_{hjk} / N_{h.k} \quad \text{and}$$

$$W_h = N_h / N_{..k}\ .$$

we have,

$$\begin{aligned} U_{h20} &= P_{hjk}\ (1 - P_{hjk}), \\ U_{h02} &= P_{h.k}\ (1 - P_{h.k}), \\ U_{h11} &= P_{hjk}\ (1 - P_{h.k}), \\ U_{h12} &= P_{hjk} - 3P_{hjk}\ P_{h.k} + 2P_{hjk}\ P_{h.k}^2, \\ U_{h21} &= P_{hjk} - 2P_{hjk}^2 + 2P_{hjk}^2\ P_{h.k} \\ &\quad - P_{hjk}\ P_{h.k} \quad \text{and} \\ U_{h03} &= P_{h.k} - 3P_{h.k}^2 + 2P_{h.k}^3\ . \end{aligned} \tag{2.10}$$

Substituting (2.10) into (2.9) to obtain,

$$E(r) = P_{j/k} + \sum_{h=1}^{L} W_h^2 P_{h.k} (1-P_{h.k})(P_{j/k} - P_{hj/k}) [f_h + W_h (g_h - 3f_h / N_h)(2P_{h.k} - 1)] + 3 \sum_{h,h'=1}^{L} W_h^2 W_{h'}^2 f_h f_{h'} P_{h.k} (1-P_{h.k}) P_{h'.k} (1-P_{h'.k})(P_{j/k} - P_{h'j/k}) ,$$

$$E(t_2)=P_{j/k} - \sum_{h=1}^{L} W_h^3(2g_h - 3f_h / N_h) P_{h.k} (1-P_{h.k})(2P_{h.k} - 1)(P_{j/k} - P_{hj/k}) - 2 \sum_{h,h'=1}^{L} W_h^2 W_{h'}^2 f_h f_{h'} P_{h.k} (1-P_{h'.k}) P_{h'.k} (1-P_{h'.k})(P_{j/k} - P_{h'j/k}) \tag{2.11}$$

and

$$E(t_3) = P_{j/k} - \sum_{h=1}^{L} W_h^3 (2g_h - 3f_h/N_h) P_{h.k} (1-P_{h.k})(2P_{h.k} - 1)(P_{j/k} - P_{hj/k}) - 3 \sum_{h,h'=1}^{L} W_h^2 W_{h'}^2 f_h f_{h'} P_{h.k} (1-P_{h.k}) P_{h'.k} (1-P_{h'.k})(P_{j/k} - P_{h'j/k}).$$

From (2.11) it can be seen that t_2 and t_3 are both less biased than r. In most cases, t_2 appears to be slightly less biased than t_3.

3. Variance Approximations

The variances of r, t_2 and t_3 are, by defination

$$V(r) = Er^2 - (Er)^2 , \quad V(t_2) = Et_2 - (Et_2)^2 \text{ and } \quad V(t_3) = Et_3 - (Et_3)^2 . \tag{3.1}$$

Squaring both sides of equation (2.6) and using the expansion,

$$(1 + \bar{e}_x/\bar{X})^{-2} = \sum_{i=0}^{\infty} (-1)^i(i+1)(\bar{e}_x/\bar{X})^i \tag{3.2}$$

the variances of r, t_2 and t_3 to $0(n^{-2})$ are obtained as,

$$V(r) = P_{j/k}^2 \{ \sum_{h=1}^{L} U_h^2 f_h(C_{h20}-2C_{h11}+C_{h02}) - 2U_h^3(g_h-3f_h/N_h)(C_{h21}-2C_{h12}+C_{h03}) + U_h^4 f_h^2 (8C_{h02}^2 - 16C_{h02} C_{h11} + 3C_{h20} C_{h02} + 5C_{h11}^2)] + \sum_{\substack{h,h'=1 \\ h\neq h}}^{L} U_h U_{h'} f_h f_{h'}(8C_{h02}C_{h'02} - 16C_{h02}C_{h'11} + 3C_{h02}C_{h'20} + 5C_{h11}C_{h'11}\} \tag{3.3}$$

and

$$V(t_2) = V(t_3) = P_{j/k}^2\{ \sum_{h=1}^{L} [U_h^2 f_h (C_{h20} - 2C_{h11} + C_{h02}) + 2U_h^3 (f_n/N_h) (C_{h03} - 2C_{h12} + C_{h21} + U_h^4 f_h^2 (2C_{h02}^2 - 4C_{h02} C_{h11} + C_{h20} C_{h02} + C_{h11}^2)] + \sum_{\substack{h,h'=1 \\ h=h}}^{L} U_h U_{h'} f_h f_{h'}(2C_{h02} C_{h'02} - 4C_{h02} C_{h'11} + C_{h02} C_{h'20} + C_{h11}C_{h'11}\}. \tag{3.4}$$

For the case of proportions, substituting (2.10) into (3.3) and (3.4) to yield,

$$V(r) = \sum_{h=1}^{L} W_h^2 P_{h.k} \{(1-P_{h.k})(P_{j/k}-P_{hj/k})^2 [f_h + 2W_h (g_h - 3f_h/N_h)(2P_{h.k}-1) + 8W_h^2 f_h^2 P_{h.k}(1-P_{h.k})] + P_{hj/k} (1-P_{hj/k}) [f_h - 2W_h (g_h - 3f_h/N_h)(1-P_{h.k}) + 3W_h^2 f_h^2 P_{h.k} (1-P_{h.k})]\} + \sum_{\substack{h,h'=1 \\ h\neq h'}}^{L} W_h^2 W_{h'}^2 f_h f_{h'}\{P_{h.k} P_{h'.k} (1-P_{h.k}) (1-P_{h'.k})(P_{j/k}-P_{hj/k})^2 + 3P_{h'j/k}(1-P_{h'j/k}) + 5(1-P_{h'.k})P_{h'j/k}(P_{hj/k} - P_{h'j/k})]\} \tag{3.5}$$

and

$$V(t_2) = V(t_3)= \sum_{h=1}^{L} W_h^2 P_{h.k} \{(1-P_{h.k}) (P_{j/k}-P_{hj/k})^2 [f_h+2W_h(f_h/N_h)(1-2P_{h.k}) + 2W_h^2 f_h^2 P_{h.k}(1-P_{h.k})] + P_{hj/k}(1-P_{hj/k}) [f_h + 2W_h (f_h/N_h)(1-P_{h.k})$$

$$+W_h^2 f_h^2 P_{h.k}(1-P_{h.k})]\}$$

$$+\sum_{\substack{h,h'=1\\h\neq h'}}^{L} W_h^2 W_{h'}^2 f_h f_{h'} \{P_{h.k}(1-P_{h.k}) P_{h'.k}$$

$$[2(1-P_{h'.k})(P_{j/k}-P_{h'j/k})^2$$

$$+P_{h'j/k}(1-P_{h'j/k}) + P_{h'j/k}(1-P_{h'.k})$$

$$(P_{hj/k}-P_{h'j/k})]\} . \qquad (3.6)$$

On comparing the variances in (3.5) and (3.6) we have,

$$V(r)-V(t_2) = 2\{\sum_{h=1}^{L}\sum_{h'=1}^{L} W_h^2 W_{h'}^2 f_h f_{h'} P_{h.k}$$

$$(1-P_{h.k})P_{h'.k}[3(1-P_{h'.k})(P_{j/k}-P_{h'j/k})^2$$

$$+P_{h'j/k}(1-P_{h'j/k})]$$

$$-\sum_{h=1}^{L} W_h^3 f_h^2 P_{h.k}(1-P_{h.k})$$

$$[(1-2P_{h.k})(P_{j/k}-P_{hj/k})^2$$

$$+P_{hj/k}(1-P_{hj/k})]$$

$$-2\sum_{h'>h}\sum_{h=1}^{L-1} W_h^2 W_{h'}^2 f_h f_{h'} P_{h.k}(1-P_{h.k})$$

$$P_{h'.k}(1-P_{h'.k})(P_{hj/k}-P_{h'j/k})^2\} .$$

The estimators t_2 and t_3 appear to be more efficient than r in many cases, when the population parameters, $P_{h.k} > \frac{1}{2}$ for $h=1,\ldots,L$.

4. Conclusion

The computation formulas for estimators t_2 and t_3 are not as simple as that for r. However, in view of the comparisons, estimators t_2 and t_3 not only are less biased than r, but in many situations are also more efficient than r. Furthermore, since the sample values s_{hx}^2 and s_{hxy} are required in estimating the variance of r, even to the order of $O(n^{-1})$, the computation formulas (2.3) for t_3 and t_2 cannot be considered as tidious for the case when they are more efficient. Let $\hat{V}(r)$ and $\hat{V}(t_2)$ be the estimated variances of r and t_2. Also, define,

$$\hat{P}_{h.k} = n_{h.k}/n_h,\quad \hat{P}_{hjk} = n_{hjk}/n_h \quad \text{and}$$

$$\hat{P}_{hj/k} = n_{hjk}/n_{h.k} .$$

When the parameter values are not available, the following equation could be used in comparing the estimated variances of the estimators.

$$\hat{V}(r)-\hat{V}(t_2)=2\{\sum_{h=1}^{L}\sum_{h'=1}^{L} W_h^2 W_{h'}^2 f_h f_{h'}\hat{P}_{h.k}$$

$$(1-\hat{P}_{h.k})\hat{P}_{h'.k}[3(1-\hat{P}_{h.k})(r-\hat{P}_{h'j/k})^2$$

$$+\hat{P}_{h'j/k}(1-\hat{P}_{h'j/k})]$$

$$-\sum_{h=1}^{L} W_h^3 f_h^2 \hat{P}_{h.k}(1-\hat{P}_{h.k})$$

$$[(1-2\hat{P}_{h.k})(r-\hat{P}_{hj/k})^2+\hat{P}_{hj/k}(1-\hat{P}_{hj/k})]$$

$$-2\sum_{h'>h}\sum_{h=1}^{L-1} W_h^2 W_{h'}^2 f_h f_{h'}\hat{P}_{h.k}$$

$$(1-\hat{P}_{h.k})\hat{P}_{h'.k}(1-\hat{P}_{h'.k})(\hat{P}_{hj/k}-\hat{P}_{h'j/k})^2\}.$$

REFERENCES

1. Cochran, W. G. (1963) Sampling Techniques John Wiley & Sons, Inc. New York.

2. Durbin, J. (1967) Design of Multi-stage Surveys for the estimation of Sampling errors. Applied Statistics 16; 152-164.

3. Durbin, J. (1958) Sampling theory for estimates based on fewer individuals than the number selected. Bull. Int. Stat. Inst. 36, 3, 113-119.

4. Fieller, E. C. (1951) The Biological Standardization of Insulin J. Roy. Stat. Soc. Supplement VII.

5. Hanson, Hurwitz & Madow (1953) Sample Survey Methods and Theory John Wiley & Sons, Inc. New York.

6. Hartely, H. O. and Ross, A. (1954) Unbiased ratio estimate. Nature, 174, 270-271.

7. Kish, L. (1965) Survey Sampling. John Wiley & Sons, Inc. New York.

8. Kish, L., Nambovdiri, N. K. and Pillai, P.K. (1962) The ratio bias in surveys, J. Am. Statist. A. 56 (300); 863-876.

9. Kendall, M. G. and Stuart, A., The Advance Theory of Statistics, Vol. 1. London, Charles Griffin and Company, 1958.

10. Lauh E. & Williams (1963) Some small sample results for the variance of ratio estimators. Proceedings of the Social Statistics Section of the American Statist. Assoc.

11. Sukhatme, P. V. & Sukhatme, B. V. (1970) Sampling Theory of Surveys with Applications. Iowa State University Press, Ames, Iowa.

12. Tin, M. (1965) Comparison of Some ratio estimators J. Am. Statist. A. 60 (309); 294-307.