LINEAR INVARIANCE AND ADMISSIBILITY IN SAMPLING FINITE POPULATIONS

by

Bikas Kumar Sinha[1] and Sastry G. Pantula
North Carolina State University
Raleigh, North Carolina USA

ABSTRACT

Using a technique of Patel and Dharmadhikari (1977), we come up with certain appropriate classes of linear invariant and sub-invariant admissible unbiased estimators of the mean in finite population sampling. The motivation rests on a study of the variance functions of such estimators. We also deduce explicit forms of linear invariant estimators for two familiar sampling schemes viz., (a) ppswor sampling of size two, (b) Midzuno scheme of size n.

AMS 1970 Subject Classification: Primary 62D05, Secondary 62C25
Key Words and Phrases: Linear invariance, admissible estimators, connected designs, ppswor, Midzuno scheme

---

[1] On leave from the Indian Statistical Institute, Calcutta, India.

1. Introduction

Patel and Dharmadhikari (1977, 1978) developed an interesting and useful technique for constructing a class of linear admissible unbiased estimators of the population total (or mean). In their first paper they suggested a method to construct linear invariant unbiased estimators. They established that for a connected design, a linear invariant unbiased estimator is admissible in the class of linear unbiased estimators. In their second paper, they modified the technique for constructing admissible estimators because explicit forms of the linear invariant estimators were difficult to obtain.

We have two-fold purpose in this article. First, we derive explicit forms of linear invariant estimators of the population mean for two familiar sampling schemes viz., (i) ppswor of sample size two and (ii) Midzuno (1950) scheme of sample size n. We reemphasize the use of linear invariant estimators as a method of constructing admissible estimators. Secondly, we introduce a related concept of 'sub-invariance' as another tool to construct admissible estimators in the class of linear unbiased estimators. We construct admissible linear sub-invariant estimators based on an alternative representation of the admissible estimator of Patel and Dharmadhikari (1977). The motivation lies in a study of the nature of the variance functions of such estimators.

2. Relevant Work of Patel and Dharmadhikari (1977)

Consider a finite population $\mathcal{U} = \{1,2,\ldots,N\}$ of N identifiable units. Let (S,P) denote a sampling design for $\mathcal{U}$ so that $p(s) > 0$ for every $s \in S$

and $\sum_{s \in S} p(s) = 1$. A linear estimator $\ell(s,\bar{Y})$, for the population mean $\bar{Y}$ of a character Y under study, based on a sample s, has the form

$$\ell(s,\bar{Y}) = \sum_{i \in s} \beta(s,i) Y_i \quad . \tag{1}$$

See Raj (1968) for the general notation and definitions. The conditions of unbiasedness and linear invariance are respectively,

$$\sum_{s \ni i} \beta(s,i)\ p(s) = 1/N \ , \quad i = 1,2,\ldots,N \tag{2}$$

and

$$\sum_{i \in s} \beta(s,i) = 1 \ , \ s \in S \ . \tag{3}$$

We assume that the sampling design is connected. That is, any two units (samples) are connected through one or more chains of samples (units) and units (samples). (See Patel and Dharmadhikari (1977) for details.) Let $\pi_i = \sum_{s \ni i} p(s)$ and $\pi_{ij} = \sum_{s \ni \{i,j\}} p(s)$ denote the inclusion probabilities of the unit i and the units i and j, respectively. Following the notation of Patel and Dharmadhikari (1977), let

$$\begin{aligned} c_{ii} &= \pi_i - \sum_{s \ni i} \{p(s)/n(s)\} \ , \ i = 1,2,\ldots, N \\ c_{ij} &= -\sum_{s \ni \{i,j\}} \{p(s)/n(s)\} \qquad i \neq j = 1,2,\ldots, N \end{aligned} \tag{4}$$

where n(s) denotes the size of the sample s . The $N \times N$ matrix $\mathbf{C} = ((c_{ij}))$ is called the $\mathbf{C}$-matrix of the sampling design. Note that the matrix $\mathbf{C}$ is symmetric and that $\mathbf{C}\,\underset{\sim}{1} = \underset{\sim}{0}$ where $\underset{\sim}{1} = (1,1,\ldots,1)'$ and $\underset{\sim}{0} = (0,0,\ldots,0)'$. Since the design is assumed to be connected, the rank of the matrix $\mathbf{C}$ is $N-1$. For the sake of completeness we state the results of Patel and Dharmadhikari (1977) in the following theorem.

Theorem 1. (Patel and Dharmadhikari (1977)) The estimator $\ell(s,\bar{Y})$ with $\beta(s,i) = [1/n(s)] + \lambda_i - \bar{\lambda}_s$, where the λ_i's satisfy

$$C\lambda = d \ , \qquad (5)$$

$\lambda = (\lambda_1,\ldots,\lambda_N)'$, $d_i = 1/N - \sum_{s \ni i} [p(s)/n(s)]$, $i = 1,2,\ldots,N$ and $\bar{\lambda}_s = \sum_{i \in s} \lambda_i/n(s)$, is linear invariant, unbiased for $\bar{Y}$, and minimizes the sum of the variances at the points $Y_{(1)},\ldots,Y_{(N)}$. The system of equations in (5) is consistent and the estimator ℓ is unique and, hence, admissible within the class of all linear unbiased estimators of $\bar{Y}$. Here $Y_{(i)}$ stands for N-vector with 1 in the ith position and 0 elsewhere.

In the next section we solve the system of equations (5) for two familiar designs.

3. Explicit Forms of Linear Invariant Estimators Under Two Familiar Schemes

The system (5) seems intractable in general terms. However, for at least two familiar schemes (i) ppswor of sample size 2 and (ii) Midzuno scheme, it is possible to obtain an algebraic expression for λ satisfying (5). The resulting expressions for linear invariant estimators based on these two designs are given below and the details of the calculations are given in the Appendix.

(i) PPSWOR Sampling Scheme of Size Two

For this scheme, $\pi_i = p_i\{1 + \sum_{j \neq i} p_j/(1-p_j)\}$ and $\pi_{ij} = p_i p_j \left(\frac{1}{1-p_i} + \frac{1}{1-p_j}\right)$.

The linear invariance unbiased estimator is given by

$$\ell(s,\bar{Y}) = \tfrac{1}{2}(Y_i + Y_j) + \tfrac{1}{2}(Y_i - Y_j)(\lambda_i - \lambda_j) \tag{6}$$

where $s = \{i,j\}$,

$$\lambda_i - \lambda_j = \frac{p_i - p_j}{N(1+a-ap_i)(1+a-ap_j)} \left[\frac{2(N+1-ax)}{(a+1)x} + \frac{N}{x} - \frac{2}{p_i p_j}\{1+a(1-p_i)(1-p_j)\}\right],$$

$$a = \sum_{i=1}^{N} \{p_i/(1-p_i)\},$$

and

$$x = \sum_{i=1}^{N} (p_i^2/\theta_i).$$

(ii) Midzuno Scheme for Samples of Size n

For this scheme, $\pi_i = \frac{n-1}{N-1} + \frac{N-n}{N-1} p_i$ and $\pi_{ij} = \frac{(n-1)(n-2)}{(N-1)(N-2)} + \frac{(n-1)(N-n)}{(N-1)(N-2)}(p_i + p_j)$.

The linear invariant unbiased estimator $\ell(s,\bar{Y})$ is given by

$$\bar{y}(s) + \sum_{i \in s} (\lambda_i - \bar{\lambda}_s) Y_i \tag{7}$$

where

$$\lambda_i = \frac{x}{a_i} + \ell, \quad a_i = g + fp_i,$$

$$x = \frac{(N-n)(N+1)}{nN(N-1)}, \left\{\frac{f(f+c)}{(1+dy)[f^2 + cg^2 y - cgN + cf - dfc(N-gy)]}\right\},$$

$$f = [N(N-n)(n-1)]/[n(N-1)(N-2)],$$

$$g = [(n-1)(Nn-N-n)]/[n(N-1)(N-2)], \quad b = -(n-1)(n-2)/[n(N-1)(N-2)]$$

$$c = (n-1)(N-n)/[n(N-1)(N-2)]$$

$$z = \sum_{i=1}^{N} \{1/[(Nn-N-n) + N(N-n)p_i]\}, \quad d = b + c$$

$y = n(N-1)(N-2)\, z/(n-1)$ and ℓ is some constant.

4. Linear Invariance and Sub-invariance

In this section we examine the results of Patel and Dharmadhikari (1977) from a different angle. We extend the definition of linear invariance to linear sub-invariance and establish that certain types of linear sub-invariant unbiased estimators are also admissible in the class of linear unbiased estimators of the population mean. The motivation lies in a study of the variance functions of such estimators.

Patel and Dharmadhikari (1977) deduced that a choice of $\beta(s,i)$ as $\beta(s,i) = \lambda_i + a_s$ together with the conditions (2) and (3) results in $\beta(s,i) = \{1/n(s)\} + \lambda_i - \bar{\lambda}_s$ where λ_i's satisfy the system (5). Likewise, it may be seen that such a $\beta(s,i)$ also has an alternative representation like $\beta(s,i) = \{1/N\pi_i\} + w_s - \bar{w}_i$ where $\bar{w}_i = \sum_{s \ni i} w_s\, p(s)/\pi_i$ and the w_s's satisfy an analogous system

$$D\,\underset{\sim}{w} = \underset{\sim}{q}\ . \tag{8}$$

<u>Note that the study of admissibility essentially rests on the reduced forms of sampling designs wherein no two samples are equivalent</u>. Therefore the cardinality M of S is at most 2^N-1 and the matrix equation (8) with

$$\begin{aligned} \underset{\sim}{w} &= (w_1, w_2, \ldots, w_M)' \\ d_{ss} &= n(s)p(s) - p^2(s) \sum_{i \in s} 1/\pi_i \\ d_{ss'} &= -p(s)p(s') \sum_{i \in s \cap s'} 1/\pi_i\ , \quad s \neq s' \end{aligned} \tag{9}$$

and

$$q_s = p(s) - p(s) \sum_{i \in s} 1/N\pi_i$$

represents a finite system.

Observe that $D\underset{\sim}{1} = \underset{\sim}{0}$ and $\underset{\sim}{q}'\underset{\sim}{1} = 0$ so that the system (6) is consistent and the solution is unique (up to $w_s - \bar{w}_i$). This reminds one of similar studies in block-design set-up. It can be shown that rank (D) = M-1 iff rank (C) = N-1 . In practice, therefore, given a connected sampling design, one might first check the aspect of simplicity in solving the systems (5) and (8). If M is smaller than N, one might prefer the system (6) to (5).

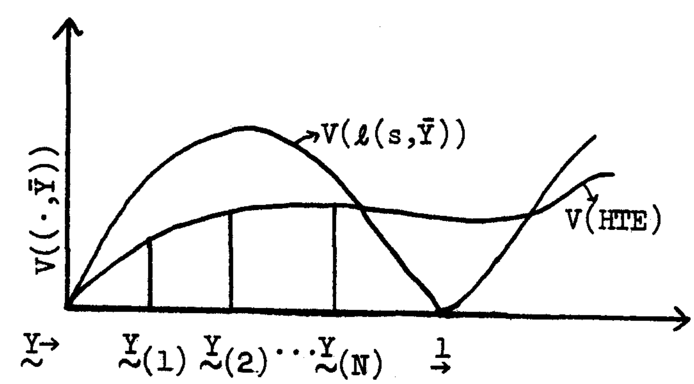
The two alternative representations of the admissible linear invariant unbiased estimator $\ell(s,\bar{Y})$ e.g.,

$$\ell(s,\bar{Y}) = \bar{y}(s) + \sum_{i\in s} (\lambda_i - \bar{\lambda}_s)Y_i \ , \ \bar{y}(s) = \sum_{i\in s} Y_i/n(s)$$

and

$$\ell(s,\bar{Y}) = \frac{1}{N}\sum_{i\in s} Y_i/\pi_i + \sum_{i\in s} (w_s - \bar{w}_i)Y_i$$

have interesting interpretations. In the first case, it is the sample mean adjusted for unbiasedness by an error function and in the second case, it is the Horwitz-Thompson Estimator (abbreviated henceforward as HTE) adjusted for linear invariance by another error function. This latter representation allows us to explore further the behavior of variance functions of such estimators and investigate the question of availability of similar other admissible estimators. Towards this, let us now examine the variance curves of the two admissible estimators: HTE and $\ell(s,\bar{Y})$.



The curve of $V[\ell(s,\bar{Y})]$ hits the X-axis at $\underset{\sim}{1}$ and stays above the curve of V(HTE) at the points $\underset{\sim}{Y}_{(i)}$, $i=1,2,\ldots,N$. Naturally, no estimator is better than the other and both are admissible. The following two questions naturally arise.

Q1: Do there exist other admissible variance curves that intersect the curve of V(HTE) at one or more of the points $\underset{\sim}{Y}_{(1)}, \underset{\sim}{Y}_{(2)}, \ldots, \underset{\sim}{Y}_{(N)}$ and yet hit the X-axis at a point?

Q2. Do there exist other admissible variance curves that hit the X-axis at more than one point? (It must be understood that we are referring to points that are not proportional.)

We establish that such admissible curves do exist. We define systems analogous to (8) and construct admissible estimators for the population mean based on such admissible curves. We first introduce the concept of sub-invariance.

Definition 4.1: An unbiased estimator is said to be type I - linear sub-invariant of order m if for some choice of $(i_1,i_2,\ldots,i_m) \subset (1,2,\ldots,N)$, the variance curve of the estimator attains the minimum value at the points $\underset{\sim}{Y}_{(i_1)}, \underset{\sim}{Y}_{(i_2)}, \ldots, \underset{\sim}{Y}_{(i_m)}$ and $\underset{\sim}{1} - \sum_{j=1}^{m} \underset{\sim}{Y}_{(i_j)}$, simultaneously. For $m=0$, it is a linear invariant estimator.

We first establish the existance of such estimators.

Lemma 4.1: For every m, $0 \leq m \leq N$, there exists at least one type I - linear sub-invariant estimator of order m .

Proof: For $m=0$, we know that a linear invariant estimator exists. For $m=N-1$ or N, HTE attains the minimum variance at $\underset{\sim}{Y}_{(1)}, \underset{\sim}{Y}_{(2)}, \ldots, \underset{\sim}{Y}_{(N)}$. For $1 \leq m \leq N-2$, let us set $i_j=j$, $j=1,2,\ldots,m$ for the simplicity of notation. Define $\ell^{(m)}(s,\bar{Y}) = \sum_{i \in s} \beta^{(m)}(s,i) Y_i$

with

$$\beta^{(m)}(s,i) = 1/N\pi_i \ , \ i \in s, \ i=1,2,\ldots, m$$

$$= 1/N\pi_i + w_s - \bar{w}_i \ , \ i \in s \ , \ i=m+1,\ldots,N \ , \qquad (10)$$

where

$$\bar{w}_i = \sum_{s \ni i} w_s p_{(s)}/\pi_i \ , \ i=m+1,\ldots,N \ .$$

Clearly, w_s's are relevant only for the samples $s \not\subset \{1,2,\ldots,m\}$ and the estimating equations are given by

$$D^{(m)} \underset{\sim}{w} = \underset{\sim}{q}^{(m)} \qquad (11)$$

where

$$d_{ss}^{(m)} = n^*(s)\ p(s) - p^2(s) \sum_{i \in s - \underset{\sim}{u}_m} 1/\pi_i$$

$$d_{ss'}^{(m)} = -p(s)\ p(s') \sum_{i \in s \cap s' - \underset{\sim}{u}_m} 1/\pi_i \ , \ s \neq s'$$

$$q_s^{(m)} = (N-m)p(s)/N - p(s) \sum_{i \in s - \underset{\sim}{u}_m} (1/N\pi_i) \ ,$$

$$n^*(s) = n(s) - n(s \cap \underset{\sim}{u}_m),$$

and

$$\underset{\sim}{u}_m = \{1,2,\ldots,m\} \ .$$

It can be shown that $D^{(m)}\underset{\sim}{1} = \underset{\sim}{0}$ and $\underset{\sim}{1}'\underset{\sim}{q}^{(m)} = 0$. Therefore, the system of equations in (11) is consistent. It is easy to verify that $\ell^{(m)}(s,\bar{Y})$ given in (10) is a type I - linear invariant estimator of order m . Further, since the sampling scheme is connected, there always exists, for every m, $1 \leq m \leq N-2$, $\{j_1,j_2,\ldots,j_m\}$ such that at least one s is not a proper subset of $\{j_1,j_2,\ldots,j_m\}$. We choose such $(j_1,j_2,\ldots,j_m)$ as $(i_1,i_2,\ldots,i_m)$. Also, for this choice, $D^{(m)}$ is such that the rank of $D^{(m)}$ is equal to the number of rows in $D^{(m)}$ minus one. Hence the system (11) yields unique estimator $\ell^{(m)}(s,\bar{Y})$. ▯

We are now in a position to answer Q1 by establishing admissibility of $\ell^{(m)}(s,\bar{Y})$ given in (10) for at least one m, $1 \leq m \leq N-2$.

Theorem 4.1: For every m, $0 \leq m \leq N$, the estimator $\ell^{(m)}(s,\bar{Y})$ is admissible in the class of linear unbiased estimators.

Proof. For $m=0$, N-1 or N, the proof is immediate. For any other value of m, take $(i_1, i_2, \ldots, i_m) = (1,2,\ldots,m) = \underset{\sim}{u}_m$. Now it can be shown that the estimator $\ell^{(m)}(s,\bar{Y})$ determined through (10) and (11) is the unique unbiased estimator minimizing the sum of the variances at the points $\underset{\sim}{Y}_{(m+1)}, \underset{\sim}{Y}_{(m+2)}, \ldots, \underset{\sim}{Y}_{(N)}$ and attaining minimum variance at each of the points $\underset{\sim}{Y}_{(1)}, \ldots, \underset{\sim}{Y}_{(m)}$. Therefore, $\ell^{(m)}(s,\bar{Y})$ is admissible. ▯

Now we introduce another type of sub-invariance that is related to Q2.

Definition 4.2: An unbiased estimator is said to be a type II linear sub-invariant estimator of order m if for some choice of $\{i_1, i_2, \ldots, i_m\} \subset \{1,2,\ldots,N\}$, the variance of the estimator is zero at the points $\underset{\sim}{1} - \underset{\sim}{Y}_{(i_1)}, \underset{\sim}{1} - \underset{\sim}{Y}_{(i_2)}, \ldots, \underset{\sim}{1} - \underset{\sim}{Y}_{(i_m)}$, simultaneously. For $m=0$, it is a linear invariant estimator.

The existence of type II sub-invariant estimator is established in the following lemma.

Lemma 4.2: Whenever an m^{th} order inclusion probability is strictly positive, there exists at least one type II sub-invariant estimator of order m .

Proof. For simplicity of the notation let $\pi_{12\ldots m} > 0$. Set then $(i_1, \ldots, i_m) = (1,2,\ldots,m) = \underset{\sim}{u}_m$. Now construct the estimator

$$\ell_{(m)}(s,\bar{Y}) = \sum_{i \in s} \beta_{(m)}(s,i) Y_i \qquad (12)$$

using

$$\beta_{(m)}(s,i) = \frac{1}{N\pi_{12\ldots m}} \qquad \text{for } i=1,2,\ldots,m, \quad s \supseteq \underset{\sim}{u}_m$$

$$= 0 \qquad \text{for } i=1,2,\ldots,m, \quad s \not\supseteq \underset{\sim}{u}_m$$

$$= \frac{1}{N\pi_i} + w_s - \bar{w}_i \qquad \text{for } i \in s, \ i=m+1,\ldots,N .$$

The coefficients $\beta_{(m)}(s,i)$ are well defined because $\pi_{12\ldots m}$ is assumed to be positive. Once again the w_s's are relevant only for the samples $s \not\supseteq \underset{\sim}{u}_m$. In order that $\ell_{(m)}(s,\bar{Y})$ is type II sub-invariant, it is necessary and sufficient that w_s's satisfy,

$$D_{(m)} \underset{\sim}{w} = \underset{\sim}{q}_{(m)} \tag{13}$$

where $D_{(m)}$ has the same form as $D^{(m)}$ in (11) and

$$q_{(m)s} = (1-\tfrac{1}{N})\, p(s) - \frac{(m-1)p(s)}{N\,\pi_{12\ldots m}}\, \delta_{s:\underset{\sim}{u}_m} - \sum_{i \in s-\underset{\sim}{u}_m} (1/N\,\pi_i)$$

with

$$\delta_{s:\underset{\sim}{u}_m} = \begin{cases} 1 & \text{if } s \supseteq \underset{\sim}{u}_m , \\ 0 & \text{otherwise} . \end{cases}$$

Consistency and uniqueness are again easily verifiable. □

<u>Remark 4.1</u>: In general, m can extend up to $\max_{s \in S} \{n(s)\}$. The case of $m=N$ corresponds to the census. We now state the result on admissibility of $\ell_{(m)}(s,\bar{Y})$.

<u>Theorem 4.2</u>: The estimator $\ell_{(m)}(s,\bar{Y})$, whenever available, is admissible within the class of linear unbiased estimators of the population mean. The proof is similar to that of Theorem 4.1 and hence is omitted.

Remark 4.2: The system (11) can be translated into the corresponding system $C^{(m)}\underset{\sim}{\lambda} = \underset{\sim}{d}^{(m)}$ where

$$c_{ii}^{(m)} = \pi_i - \sum_{s \ni i} p(s)/n^*(s) \ , \ c_{ij}^{(m)} = -\sum_{s \ni i \cap j} p(s)/n^*(s)$$

$$d_i^{(m)} = \frac{1}{N} - \frac{N-m}{N} \sum_{s \ni i} p(s)/n^*(s)$$

for $m+1 \leq i \neq j \leq N$.

Similarly, the system (13) has the equivalent counterpart $C_{(m)} \lambda = \underset{\sim}{d}_{(m)}$ with $C_{(m)} = C^{(m)}$ and $d_{(m)i} = \frac{1}{N} - \frac{N-1}{N} \sum_{s \ni i} p(s)/n^*(s)$ $+ \frac{m-1}{N\pi_{12\ldots m}} \sum_{s \ni i} \delta_{s:\underset{\sim}{u}_m} p(s)/n^*(s)$.

We conclude this section with the following examples based on SRSWOR(N,n) sampling procedure. Choose any $m \leq n$.

$$\text{(i)} \quad \ell^{(m)}(s,\bar{Y}) = \left(1 - \frac{n^*(s)}{n}\right)\bar{y}(s \cap \underset{\sim}{u}_m) + \left(1 - \frac{m}{N}\right) \bar{y}(s \cap \underset{\sim}{u}_m')$$

$$\text{(ii)} \quad \ell_{(m)}(s,\bar{Y}) = \frac{(N-1)^{(m-1)}}{n^{(m)}} \bar{y}(\underset{\sim}{u}_m)\delta_{s:\underset{\sim}{u}_m} + \left\{1 - \frac{1}{N} - \frac{(N-1)^{(m-1)}}{n^{(m)}} \delta_{s:\underset{\sim}{u}_m}\right\} \bar{y}(s \cap \underset{\sim}{u}_m')$$

where $\bar{y}(\cdot)$ = mean of the elements in the set $(\cdot)$

$$n^{(m)} = n(n-1) \ldots (n-m+1), \ \underset{\sim}{u}_m' = u - \underset{\sim}{u}_m$$

and $n^*(s)$ and $\delta_{s:\underset{\sim}{u}_m}$ are as defined before. It should be noted that if $m = n$ and $s = \underset{\sim}{u}_n$, then $\bar{y}(s \cap \underset{\sim}{u}_m') = 1$. If $m = 0$ we get $\ell^{(0)}(s,\bar{Y}) = \ell_{(0)}(s,\bar{Y}) = \bar{y}(s)$.

Appendix

In this appendix we derive explicit expressions for linear invariant estimators for two familiar sampling schemes. Observe that the system

$$C\,\underset{\sim}{t} = \underset{\sim}{b}$$

with

$$C\,\underset{\sim}{1} = \underset{\sim}{0}$$

$$\underset{\sim}{b}'\underset{\sim}{1} = \underset{\sim}{0}$$

where C is any symmetric matrix, is equivalent to

$$(\underset{\sim}{C} + \underset{\sim}{\alpha}\underset{\sim}{\alpha}')\underset{\sim}{t} = \underset{\sim}{b}$$

for any $\underset{\sim}{\alpha}$ satisfying $\underset{\sim}{\alpha}'\underset{\sim}{1} \neq 0$. Now a suitable choice of $\underset{\sim}{\alpha}$ can be made such that (i) rank of $C + \underset{\sim}{\alpha}\underset{\sim}{\alpha}' = N$ and (ii) explicit algebraic expressions for the elements of $(C + \underset{\sim}{\alpha}\underset{\sim}{\alpha}')^{-1}$ are available. Moreover, in applications, we are concerned with contrasts involving the t_i's so that in any solution to $\underset{\sim}{t}$, terms proportional to $\underset{\sim}{1}$ can be ignored.

<u>PPSWOR (N,n=2) Sampling Scheme</u>:

Write $\delta_i = p_i(1-p_i)^{-1}$, $a = \sum_{i=1}^{N} \delta_i$ and $\theta_i = ap_i + \delta_i$. Then (5) reads as

$$H\underset{\sim}{\lambda} = \underset{\sim}{h} \qquad \text{(A.1)}$$

where

$$H = M - (\underset{\sim}{p}\underset{\sim}{\delta}' + \underset{\sim}{\delta}\underset{\sim}{p}') ,$$

$$M = \text{diag}(\theta_1, \theta_2, \ldots, \theta_N) ,$$

$$\underset{\sim}{h} = \frac{2}{N}\underset{\sim}{1} - \underset{\sim}{\theta} + 2(\underset{\sim}{\delta} - \underset{\sim}{p}) ,$$

$$\underset{\sim}{1} = (1,1,\ldots,1)' ,$$

$$\underset{\sim}{\theta} = (\theta_1, \theta_2, \ldots, \theta_N)' ,$$

$$\underset{\sim}{\delta} = (\delta_1, \delta_2, \ldots, \delta_N)'$$

and

$$\underset{\sim}{p} = (p_1, p_2, \ldots, p_N)' .$$

We take $\boldsymbol{\alpha} = (\mathbf{p}+\boldsymbol{\delta})$ in the above formulation. Then

$$\boldsymbol{\lambda} = (H + \boldsymbol{\alpha}\boldsymbol{\alpha}')^{-1}\mathbf{h}$$

$$= [A^{-1} - A^{-1}\boldsymbol{\delta}\boldsymbol{\delta}' A^{-1}(1+\boldsymbol{\delta}' A^{-1}\boldsymbol{\delta})^{-1}]\mathbf{h}$$

where

$$A^{-1} = M^{-1} - M^{-1}\mathbf{p}\,\mathbf{p}'M^{-1}(1+\mathbf{p}'M^{-1}\mathbf{p})^{-1}.$$

Let $\quad x = \mathbf{p}'M^{-1}\mathbf{p} = \sum_{i=1}^{N} (p_i^2/\theta_i)$. Then

$$\mathbf{p}'M^{-1}\boldsymbol{\delta} = 1 - ax$$

and

$$\boldsymbol{\delta}'M^{-1}\boldsymbol{\delta} = a^2 x .$$

This yields

$$\boldsymbol{\lambda} = M^{-1}\mathbf{h} - M^{-1}\mathbf{p}(\mathbf{p}'M^{-1}\mathbf{h})(1+x)^{-1}$$

$$- \{M^{-1}\boldsymbol{\delta} - M^{-1}\mathbf{p}(1-ax)(1+x)^{-1}\}\,\boldsymbol{\delta}' A^{-1}\mathbf{h}\,[x(a+1)^2]^{-1}(1+x) .$$

Note that

$$\boldsymbol{\delta}'A^{-1}\mathbf{h}\,(1+x) = \boldsymbol{\delta}' M^{-1}\mathbf{h}(1+x) - \mathbf{p}'M^{-1}\mathbf{h}(1-ax)$$

$$= (\boldsymbol{\delta}' - \mathbf{p}')\,M^{-1}\mathbf{h} + x(\boldsymbol{\delta} + a\mathbf{p})'M^{-1}\mathbf{h}$$

$$= (\boldsymbol{\delta}' - \mathbf{p}')\,M^{-1}\mathbf{h}$$

since $\mathbf{1}'\mathbf{h} = 0$ and $M\,\mathbf{1} = \boldsymbol{\theta} = \boldsymbol{\delta} + a\mathbf{p}$.

Therefore,

$$\boldsymbol{\lambda} = M^{-1}\mathbf{h} - M^{-1}\mathbf{p}(\mathbf{p}'M^{-1}\mathbf{h})(1+x)^{-1}$$

$$- \{M^{-1}(\boldsymbol{\delta}-\mathbf{p}) + x\,M^{-1}\boldsymbol{\theta}\}(\boldsymbol{\delta}' - \mathbf{p}')M^{-1}\mathbf{h}[x(x+1)(a+1)^2]^{-1} .$$

Since $\underset{\sim}{h} = \frac{2}{N}\underset{\sim}{1} - \underset{\sim}{\theta} + 2(\underset{\sim}{\delta}-\underset{\sim}{p})$ and $M^{-1}\underset{\sim}{\theta} = \underset{\sim}{1}$ we get

$$\underset{\sim}{\lambda} = \frac{2}{N} M^{-1}\underset{\sim}{1} + M^{-1}(\underset{\sim}{\delta} - \underset{\sim}{p})(2-z) - y\, M^{-1}\underset{\sim}{p} + \ell\, \underset{\sim}{1}$$

where $y = \underset{\sim}{p}'M^{-1}\underset{\sim}{h}(1+x)^{-1}$,

$$z = (\underset{\sim}{\delta}' - \underset{\sim}{p}')M^{-1}\underset{\sim}{h}\,[x(x+1)(a+1)^2]^{-1}, \text{ and}$$

$$\ell = -(zx+1)\ .$$

Since $\underset{\sim}{\delta} - \underset{\sim}{p} = \underset{\sim}{\theta} - (a+1)\underset{\sim}{p}$ we get

$$\underset{\sim}{\lambda} = \frac{2}{N} M^{-1}\underset{\sim}{1} - [(a+1)(2-z) + y]\, M^{-1}\underset{\sim}{p} + (\ell + 2-z)\,\underset{\sim}{1}\,.$$

Therefore,

$$\lambda_i - \lambda_j = \frac{2}{N}\left(\frac{1}{\theta_i} - \frac{1}{\theta_j}\right) - [(a+1)(2-z)+y]\left[\frac{p_i}{\theta_i} - \frac{p_j}{\theta_j}\right]$$

$$= \frac{(p_i - p_j)}{N(a+1-ap_i)(a+1-ap_j)p_ip_j}\,[Np_ip_j\{(a+1)(2-z)+y\} - 2\{a(1-p_i)(1-p_j)+1\}]\ .$$

Now,

$$\underset{\sim}{p}'M^{-1}\underset{\sim}{h} = \frac{2}{N}\,\underset{\sim}{p}'M^{-1}\underset{\sim}{1} - \underset{\sim}{p}'M^{-1}\,\underset{\sim}{\theta} + 2\underset{\sim}{p}'M^{-1}(\underset{\sim}{\delta} - \underset{\sim}{p})$$

$$= \frac{2}{N}\,\underset{\sim}{p}'M^{-1}\underset{\sim}{1} - 1 - 2x(a+1)$$

and

$$\underset{\sim}{\delta}'M^{-1}\underset{\sim}{h} = (\underset{\sim}{\theta} - a\,\underset{\sim}{p})'M^{-1}\underset{\sim}{h}$$

$$= \underset{\sim}{1}'\underset{\sim}{h} - a\,\underset{\sim}{p}'M^{-1}\underset{\sim}{h}$$

$$= -a\,\underset{\sim}{p}'M^{-1}\underset{\sim}{h}\,.$$

Therefore,

$$y = -x(a+1)z\ ,$$

and

$$\lambda_i - \lambda_j = \frac{(p_i - p_j)}{xN(a+1-ap_i)(a+1-ap_j)p_ip_j}[2\, p_ip_j\, \underset{\sim}{p}'M^{-1}\underset{\sim}{1} + N\, p_ip_j - 2x\{a(1-p_i)(1-p_j)+1\}] .$$

Observe that

$$\sum_{i=1}^{N} \frac{p_i}{\theta_i} + \sum_{i=1}^{N} \frac{p_i\delta_i}{\theta_i} = \underset{\sim}{p}'M^{-1}\underset{\sim}{1} + (\underset{\sim}{\delta} - \underset{\sim}{p})'M^{-1}\underset{\sim}{1}$$

$$= \underset{\sim}{\delta}'\, M^{-1}\underset{\sim}{1}$$

and

$$\underset{\sim}{p}'M^{-1}\underset{\sim}{1} = \underset{\sim}{\delta}'M^{-1}\underset{\sim}{1} + \underset{\sim}{p}'M^{-1}\underset{\sim}{\delta}$$

$$= (\underset{\sim}{\theta} - a\underset{\sim}{p})'M^{-1}\underset{\sim}{1} + (1-ax) .$$

Therefore

$$\underset{\sim}{p}'M^{-1}\underset{\sim}{1} = [N + (1-ax)](1+a)^{-1} ,$$

and finally,

$$\lambda_i - \lambda_j = \frac{(p_i - p_j)}{N(a+1-ap_i)(a+1-ap_j)}\, [\frac{2(N+1-ax)}{(a+1)x} + \frac{N}{x} - \frac{2}{p_ip_j}\{a(1-p_i)(1-p_j)+1\}] .$$

<u>Midzuno Scheme</u>:

For Midzuno scheme (5) reads as

$$R\underset{\sim}{\lambda} = \underset{\sim}{r} \tag{A.2}$$

where

$$R = P + b\,\underset{\sim}{1}\,\underset{\sim}{1}' - c(\underset{\sim}{p}\,\underset{\sim}{1}' + \underset{\sim}{1}\,\underset{\sim}{p}') ,$$

$$P = \text{diag}\,(a_1,a_2,\ldots,a_p) ,$$

$$b = -(n-1)(n-2)/[n(N-1)(N-2)]$$

$$c = (n-1)(N-n)/[n(N-1)(N-2)]$$

$$a_i = (n-1)\{(Nn-N-n) + N(N-n)p_i\}/[n(N-1)(N-2)] = g + fp_i \text{ (say)}$$

and

$$\underset{\sim}{r} = \frac{(N-n)}{n(N-1)} \{\frac{1}{N}\underset{\sim}{1} - \underset{\sim}{p}\} .$$

We take $\underset{\sim}{\alpha} = (\underset{\sim}{1} + \underset{\sim}{p})$ in the above formulation.

Then

$$\underset{\sim}{\lambda} = (R + \underset{\sim}{\alpha}\underset{\sim}{\alpha}')^{-1} \underset{\sim}{r}$$

$$= [R + (c+b)\underset{\sim}{1}\,\underset{\sim}{1}' + c\,\underset{\sim}{p}\underset{\sim}{p}']^{-1}\underset{\sim}{r}$$

$$= \left[B^{-1} - \frac{d\,B^{-1}\underset{\sim}{1}\underset{\sim}{1}'B^{-1}}{(1+d\underset{\sim}{1}'B^{-1}\underset{\sim}{1})}\right] \frac{(N-n)}{n(N-1)} \{\frac{1}{N}\underset{\sim}{1} - \underset{\sim}{p}\}$$

where $B = R + c\underset{\sim}{p}\underset{\sim}{p}'$, $d = (c+b)$ and

$$B^{-1} = P^{-1} - \frac{c \cdot P^{-1}\underset{\sim}{p}\underset{\sim}{p}'P^{-1}}{(1 + c\,\underset{\sim}{p}'P^{-1}\underset{\sim}{p})} .$$

Next note that

$$(R + \underset{\sim}{\alpha}\underset{\sim}{\alpha}')\underset{\sim}{1} = c(N+1)(\underset{\sim}{1} + \underset{\sim}{p})$$

and hence

$$(R + \underset{\sim}{\alpha}\underset{\sim}{\alpha}')^{-1}\underset{\sim}{p} = \frac{1}{c(N+1)}\underset{\sim}{1} - (R + \underset{\sim}{\alpha}\underset{\sim}{\alpha}')^{-1}\underset{\sim}{1} .$$

Further $P^{-1}\underset{\sim}{p} = \frac{1}{f}P^{-1}\underset{\sim}{a} - \frac{g}{f}P^{-1}\underset{\sim}{1} = \frac{1}{f}\underset{\sim}{1} - \frac{g}{f}P^{-1}\underset{\sim}{1}$.

This yields, after routine algebra, for some ℓ_1 and ℓ_2 ,

$$\underset{\sim}{\lambda} = (R + \underset{\sim}{\alpha}\underset{\sim}{\alpha}')^{-1}\underset{\sim}{1} \cdot \frac{(N-n)(N+1)}{nN(N-1)} + \ell_1\underset{\sim}{1}$$

$$= \frac{(N-n)(N+1)}{nN(N-1)} \cdot \frac{1}{(1+d\underset{\sim}{1}'B^{-1}\underset{\sim}{1})} \left(1 + \frac{c\underset{\sim}{p}'P^{-1}\underset{\sim}{1}}{(1+c\underset{\sim}{p}'P^{-1}\underset{\sim}{p})} \cdot \frac{g}{f}\right)P^{-1}\underset{\sim}{1} + \ell_2\,\underset{\sim}{1} .$$

Therefore, $\lambda_i = \frac{x}{a_i} + \ell_2$ where

$$x = \frac{(N-n)(N+1)}{nN(N-1)} \cdot \frac{1}{1+d\,\underset{\sim}{1}'B^{-1}\underset{\sim}{1}} \left[1 + \frac{c\underset{\sim}{p}'P^{-1}\underset{\sim}{1}}{(1+c\underset{\sim}{p}'\underset{\sim}{P}^{-1}\underset{\sim}{p})} \cdot \frac{g}{f} \right] .$$

Now,

$$\underset{\sim}{p}'P^{-1}\,\underset{\sim}{1} = \sum_{i=1}^{N} (p_i/a_i) = \frac{N-gy}{f}$$

$$\text{where } y = \sum_{i=1}^{N} (1/a_i) = \frac{n(N-1)(N-2)}{(n-1)} \sum_{i=1}^{N} \left[\left\{ \frac{1}{(Nn-N-n) + N(N-n)p_i} \right\} \right] .$$

Also,

$$\underset{\sim}{p}'P^{-1}\underset{\sim}{p} = \frac{1}{f^2} \sum_{i=1}^{N} \{(\theta_i - g)^2/\theta_i\}$$

$$= \frac{1}{f^2} [g^2 y - gN + f] ,$$

$$\underset{\sim}{1}'B^{-1}\underset{\sim}{1} = y - \frac{c \cdot (N-gy)}{f[1 + \frac{c}{f^2}(g^2 y - gN + f)]}$$

and

$$1 + \frac{gc\underset{\sim}{p}'P^{-1}\underset{\sim}{1}}{f + cf\underset{\sim}{p}'P^{-1}\underset{\sim}{p}} = \frac{f^2 + cf}{f^2 + c[g^2 y - gN + f]} .$$

Therefore,

$$x = \frac{(N-n)(N+1)}{nN(N-1)} \cdot \frac{f(f+c)}{(1+dy)[f^2 + cg^2 y - cgN + cf - dfc(N-gy)]} .$$

References

Midzuno, H. (1950). An outline of the theory of sampling systems. _Am. Inst. Statist. Math._ **1**, 149-156.

Patel, H. C. and Dharmadhikari, S. W. (1977). On linear invariant unbiased estimators in survey sampling. _Sankhyā: Series C_, **39**, 21-27.

Patel, H. C. and Dharmadhikari, S. W. (1978). Admissibility of Murthy's and Midzuno's estimators within the class of linear unbiased estimators of finite population totals. _Sankhyā: Series C_, **40**, 21-28.

Raj, Des (1968). _Sampling Theory_. McGraw-Hill, New York.