Sankhyā : The Indian Journal of Statistics
1989, Volume 51, Series B, Pt. 1, pp. 65-83.

ON AVERAGING OVER DISTINCT UNITS IN SAMPLING WITH REPLACEMENT

By BIKAS KUMAR SINHA[1]
Indian Statistical Institute
and
PRANAB KUMAR SEN[2]
University of North Carolina, Chapel Hill

SUMMARY. For an arbitrary convex loss function (including the case of mean absolute deviation), a simple and direct proof of the superiority of averaging over distinct units in sampling with replacement is considered. In this process, the maximal invariance characterization of the group of cyclic permutations is investigated with reference to unbiased estimation of finite population mean and variance. Also, some admissibility results on variance estimation are discussed. The negative factorial moments of the number of distinct units in the sample play an important role in this study, and a detailed treatment of them is provided.

1. Introduction

Asok (1980) gave an elementary proof of the basic fact that in simple random sampling with replacement (SRSWR) the average over distinct units has a smaller variance than the average over the entire sample (including possible repetitions). Various proofs of this well known result were available longtime back (viz., Basu. 1958, Raj and Khamis, 1958, Pathak, 1962 and Chakrabarti, 1965). While Asok's proof is quite interesting and elementary in nature, it can hardly be extended to loss functions other than the quadratic loss. We present here a simple direct proof of the superority of averaging over distinct units in SRSWR with respect to any convex loss fuhction. This, in particular, includes the criterion of mean absolute deviation (MAD). We have exploited Basu's (1958) sufficiency considerations in SRSWR along with a relevant group of cyclic permutations to provide a simple proof of quite general nature. In this respect, we shall see that the group of cyclic permutations is a maximal invariant for inference on the finite population mean, although this characterization may not hold for variance (or higher

AMS (1980) *subject classification* : 62D05, 62C15.

Keywords : Admissibility ; convex loss ; cyclic permutations ; mean ; negative factorial moments ; risk ; sufficiency ; unbiased estimation ; variance.

[1] Currently visiting, University of Illinois at Chicago.

[2] Work of this author was partially supported by the (U.S.) Office of Naval Research, Contract No. N00014-83-K-0387.

order parameter) estimation. Nevertheless, this group is instrumental in providing estimates better than the usual sample variance based on all observations. The admissibility of variance estimation in SRSWR is a natural question in this respect, and we have studied this problem in some detail too. Finally, we present, following a simpler approach, various results on negative factorial moments of $\nu_{N,n}$, the number of distinct units in the SRSWR sample of size n from a poulation of size N. Some applications of these moment formulae are discussed at the end.

2. Inadmissibility of the sample mean

Consider a finite population of N units, serially numbered $1, \ldots, N$ and having variate values $Y_1, \ldots, Y_N$ respectively. The mean of the population is therefore

$$\bar{Y}_N = \left(\sum_{i=1}^{N} Y_i\right)/N. \quad \ldots \quad (2.1)$$

In simple random sampling with replacement (SRSWR), we denote the random variables and indexes by (y_k, r_k), for $k = 1, \ldots, n$. Then $y_k = Y_{r_k}$, for $k \geqslant 1$, and each r_k takes on the values $1, \ldots, N$ with the common probability $1/N$. Then, the *simple mean per unit* estimator of $\bar{Y}_N$ based on SRSWR of size n is given by

$$\bar{y}_n = \left(\sum_{k=1}^{n} y_k\right)/n = \left(\sum_{k=1}^{n} Y_{r_k}\right)/n. \quad \ldots \quad (2.2)$$

Note that in SRSWR, $r_1, \ldots, r_n$ need not be all distinct. Let $\nu_{N,n}$ $(=\nu)$ be the number of distinct units (i.e., the number of the distinct values of $r_1, \ldots, r_n$) in the sample of size n. We may equivalently write

$$(r_1, \ldots, r_n) \equiv \{(i_1, f_1), \ldots, (i_\nu, f_\nu)\}, \quad \ldots \quad (2.3)$$

where $f_1, \ldots, f_\nu$ stand for the frequencies with which the distinct indexes $i_1, \ldots, i_\nu$ occur in the sample. Then, the *mean per distinct unit* is given by

$$\bar{y}_{(\nu_{N,n})} = \left(\sum_{j=1}^{\nu_{N,n}} Y_{i_j}\right) \nu_{N,n}^{-1} = \bar{y}_{(\nu)}, \text{ say}. \quad \ldots \quad (2.4)$$

Let now $\phi = \{\phi(u),\ u \in R\}$ be an arbitrary convex loss function. We like to provide a simpler proof for the following well-known result :

Theorem 1. *For any convex ϕ, for every (N, n),*

$$E\{\phi(\bar{y}_n - \bar{Y}_N)\} \geqslant E\{\phi(\bar{y}_{(\nu)} - \bar{Y}_N)\}, \quad \ldots \quad (2.5)$$

with the equality holding if and only if $n = 1$ or 2. Thus, for every $n > 2$, $\bar{y}_n$ is inadmissible (relative to $\bar{y}_{(\nu)}$) for any convex loss function.

Proof: We may refer to Basu (1958) for some elegant discussion of sufficiency in SRSWR. This may help in understanding the basic motivation of our simple proof of (2.5) based on this sufficiency principle and invariance under suitable group of transformations. In SRSWR, a typical sample (s) may be represented by (2.3), i.e., by $\{(i_1, f_1), \ldots, (i_\nu, f_\nu)\}$. We start with a sample of the type s and permute the units $i_1, \ldots, i_\nu$ in a cyclical manner to generate a set of ν samples of the form

$$\{(i_1, f_j), (i_2, f_{j+1}), \ldots, (i_\nu, f_{\nu+j-1})\} = s_j, \text{ for } j = 1, \ldots, \nu \text{ with } s_1 = s, \quad \ldots \quad (2.6)$$

where $f_{\nu+k} = f_k$, for $k = 1, \ldots, \nu$. If we denote this set of ν samples by π_s, then each s_j (within the set π_s) has the same probability of being drawn, i.e.,

$$P\{s = s_j \mid s \in \pi_s\} = \nu^{-1}, \text{ for each } j = 1, \ldots, \nu. \quad \ldots \quad (2.7)$$

We now write $\bar{y}_n = \bar{y}_{n(s_j)}$ when $s = s_j$, so that

$$\bar{y}_{n(s_j)} = \Big(\sum_{k=1}^{\nu} f_{j+k-1} Y_{i_k}\Big)/n, \quad \text{for } j = 1, \ldots, \nu. \quad \ldots \quad (2.8)$$

Therefore,

$$E\{\bar{y}_{n(s)} \mid s \in \pi_s\} = \nu^{-1} \sum_{j=1}^{\nu} \bar{y}_{n(s_j)} \quad \ldots \quad (2.9)$$

$$= \nu^{-1} \sum_{j=1}^{\nu} n^{-1} \Big(\sum_{k=1}^{\nu} f_{j+k-1} Y_{i_k}\Big) = \nu^{-1} \sum_{k=1}^{\nu} Y_{i_k} = \bar{y}_{(\nu)},$$

as $\sum_{j=1}^{\nu} f_{j+k-1} = n$, for every $k (= 1, \ldots, \nu)$. On the other hand, for every $s_j \in \pi_s$, $\bar{y}_{(\nu(s_j))} = \bar{y}_{(\nu(s))} = \bar{y}_{(\nu)}$, so that exploiting the convexity of $\phi(\,.\,)$, we obtain from (2.9) that

$$\phi(\bar{y}_{(\nu)} - \bar{Y}_N) = \phi\Big(\nu^{-1} \sum_{j=1}^{\nu} (\bar{y}_{n(s_j)} - \bar{Y}_N)\Big) \quad \ldots \quad (2.10)$$

$$\leqslant \nu^{-1} \sum_{j=1}^{\nu} \phi(\bar{y}_{n(s_j)} - \bar{Y}_N) = E[\phi(\bar{y}_{n(s)} - \bar{Y}_N) \mid s \in \pi_s],$$

where for a strictly convex ϕ, the equality in the penultimate step in (2.10) holds if and only if $\nu = 1$ or $f_1 = \ldots = f_\nu$ whenever $\nu > 2$. For $n = 1$, $\nu = 1$, and for $n = 2$, either $\nu = 1$ or $\nu = 2$ and $f_1 = f_2 = 1$, so that in (2.10) we have an equality. On the other hand, for $n \geqslant 3$, $\nu \geqslant 2$ with a non-zero probability, and for $\nu \geqslant 2$, (but $\nu < n$), $f_1, \ldots, f_\nu$ may not be all equal with a positive probability, so that in (2.10), we have a strict inequality in the penultimate step. Since (2.10) holds for all π_s, we immediately obtain on taking expectation over π_s that

$$E[\phi(\bar{y}_{(\nu)}-\bar{Y}_N)] \leqslant E[\phi(\bar{y}_n-\bar{Y}_N)], \quad \text{for every convex } \phi(.), \quad \dots \quad (2.11)$$

where for a strictly convex $\phi(.)$, the equality sign in (2.11) can hold only for $n = 1$ or 2. This completes the proof of the theorem.

Remarks : Note that (2.10) holds for all convex $\phi(.)$, and hence, the case of MAD (for which $\phi(x) = |x|$) is included in the set of all ϕ in (2.5) ; Asok (1980) considered the case of $\phi(x) = x^2$ and his technique may not apply for the MAD. The risk -superiority of $\bar{y}_{(\nu)}$ to $\bar{y}_n$ (in SRSWR) is thus true for any convex loss function, and this reestablishes the inadmissibility of $\bar{y}_n$ in a general setup. Basically, the Basu (1958) sufficiency of $\nu_{N,n}$ and the incorporation of the cyclical permutation group in (2.6) (resulting in the discrete uniform (conditional) distribution in (2.7)) have provided us with the necessary tools for the simple proof in (2.10)—(2.11). To illustrate this point, consider a simple example : $N = 20$, $n = 10$ and $(r_1, \dots, r_{10}) = (7, 2, 3, 7, 5, 2, 7, 15, 2, 3)$. Then, we have $\nu = 5$ and $s_1 = \{(2, 3), (3,2), (5,1), (7,3), (15,1)\}$, $s_2 = \{(2,2), (3,1), (5,3), (7,1), (15,3)\}$, $s_3 = \{(2,1), (3,3), (5,1), (7,3), (15,2)\}$, $s_4 = \{(2,3), (3,1), (5,3), (7,2), (15,1)\}$ and $s_5 = \{(2,1), (3,3), (5,2), (7,1), (15,3)\}$. Thus, here π_s consists of 5 elements $\{s_1, \dots, s_5\}$ and (2.7) asserts an equal conditional probability to each of the s_j.

3. Group of cyclic permutations and maximal invariance

For a given $\nu_n = \nu : 1 \leqslant \nu \leqslant n$, let P_ν denote the permutational uniform distribution in (2.7), and let $\mathcal{C}_\nu$ denote the corresponding sub-sigma field. We investigate below the nature of $(P_\nu, \mathcal{C}_\nu)$; $\nu \in [1, n]$.

(i) Suppose that we are to estimate a first order parameter $\theta_1 = N^{-1} \sum_{i=1}^{N} a(Y_i)$. In SRSWR (N, n), a natural unbiased estimator of θ_1 is $\hat{\theta}_{1,n} = n^{-1} \sum_{i=1}^{n} a(Y_{r_i})$. Note that for every $\nu_n = \nu : 1 \leqslant \nu \leqslant n$,

$$E[\hat{\theta}_{1,n} \mid \mathcal{C}_\nu] = E\left\{n^{-1} \sum_{k=1}^{\nu} f_k a(Y_{i_k}) | \mathcal{C}_\nu\right\} = n^{-1} \sum_{k=1}^{\nu} a(Y_{i_k}) E[f_k | \mathcal{C}_\nu]$$

$$= n^{-1} \sum_{k=1}^{\nu} a(Y_{i_k}) \left\{\nu^{-1} \sum_{l=1}^{\nu} f_l\right\} = \nu^{-1} \sum_{k=1}^{\nu} a(Y_{i_k}) \quad \dots \quad (3.1)$$

$$= \text{distinct unit mean.}$$

Thus, in this case, P_ν preserves the maximal invariance property.

(ii) Consider now a second order parameter $\theta_2 = \left(\sum_{1\leq j\neq k\leq N} \phi(Y_j, Y_k)\right)/N(N-1)$ and its natural estimator $\hat{\theta}_{2,n} = \left(\sum_{1\leq j\neq k\leq n} \phi(Y_{r_j}, Y_{r_k})\right)/n(n-1)$. Let

$$U_{\nu,1} = \nu^{-1} \sum_{k=1}^{\nu} \phi(Y_{i_k}, Y_{i_k}), \qquad \ldots \quad (3.2)$$

$$U_{\nu,2} = \begin{cases} 0, \text{ if } \nu = 1, \\ \left(\sum_{1\leq j\neq k\leq \nu} \phi(Y_{i_j}, Y_{i_k})\right)/\nu(\nu-1), \text{ if } \nu \geqslant 2. \end{cases} \qquad \ldots \quad (3.3)$$

Also, note that by definition,

$$\hat{\theta}_{2,n} = \frac{1}{n(n-1)} \left\{ \sum_{k=1}^{\nu} f_k(f_k-1)\phi(Y_{i_k}, Y_{i_k}) + \sum_{1\leq j\neq k\leq \nu} f_jf_k\phi(Y_{i_j}, Y_{i_k}) \right\} \quad \ldots \quad (3.4)$$

Now, note that $E[f_k|C_\nu] = n\nu^{-1}$, $E[f_k^2|C_\nu] = \nu^{-1} \sum_{i=1}^{\nu} f_i^2$ and for $k \neq q$, $E[f_kf_q| C_\nu] = \left\{n^2 - \sum_{i=1}^{\nu} f_i^2\right\}/\nu(\nu-1)$, so that by (3.2), (3.3) and (3.4), we have

$$\begin{aligned} E[\hat{\theta}_{2,n}|C_\nu] &= n^{-1}(n-1)^{-1} \left\{ \sum_{k=1}^{\nu} \phi(Y_{i_k}, Y_{i_k})E[f_k(f_k-1)|C_\nu] \right. \\ &\quad \left. + \sum_{1\leq k\neq q\leq \nu} \phi(Y_{i_k}, Y_{i_q})E[f_kf_q|C_\nu] \right\} \\ &= U_{\nu,2} - c(\boldsymbol{f}_n, n)(U_{\nu,2} - U_{\nu,1}), \end{aligned} \qquad \ldots \quad (3.5)$$

where

$$c(\boldsymbol{f}_n, n) = n^{-1}(n-1)^{-1}\left(\sum_{k=1}^{\nu} f_k^2 - n\right). \qquad \ldots \quad (3.6)$$

This shows that in general $U_{\nu,2}$ may not be equal to $E(\hat{\theta}_{2,n}|C_\nu)$, for every $\nu \geqslant 2$. Thus, the cyclic group P_ν does not posses the maximal invariance character for second (and higher) order parameters. Nevertheless, it is certainly instrumental in providing an improved estimator $E[\hat{\theta}_{2,n}|C_\nu]$. Therefore, even if $\hat{\theta}_{2,n}$ is a natural unbiased estimator of θ_2, in general, the distinct units U-statistic $U_{\nu,2}$ may not be unbiased for θ_2. If, in particular, $U_{\nu,1} = 0$ (with probability 1), then $\{1-c(\boldsymbol{f}_n, n)\}U_{\nu,2}$ is unbiased for θ_2 with uniformly smaller risk than $\hat{\theta}_{2,n}$. This last observation pertains to the case of the variance of the finite population (σ^2) where $\phi(a, b) = (a-b)^2/2$, so that $U_{\nu,1} = 0$ and $U_{\nu,2} = s_{(\nu)}^2 = (\nu-1)^{-1} \sum_{k=1}^{\nu} (Y_{i_k} - \bar{y}_{(\nu)})^2$, for $\nu \geqslant 2$. Therefore, it follows that the estimator $\{1-c(\boldsymbol{f}_n, n)\}s_{(\nu)}^2$ is unbiased for σ^2 and is uniformly better than $\hat{\theta}_{2,n}$. It must, however, be noted that given ν, $\boldsymbol{f}_n$ has a known distribution, independent of the Y_{i_k} (see Pathak, 1962). Thus, if we set $d(\nu, n) = E[c(f_n, n) \mid \nu]$ and define

$$U^*_{\nu,2} = U_{\nu,2}\{1-d(\nu, n)\}, \quad \dots \quad (3.7)$$

then we have immediately, for $U_{\nu.1} \equiv 0$,

$$E[U^*_{\nu,2}-\theta_2]^2 \leqslant E[E[\hat{\theta}_{2,n} \mid C_\nu)-\theta_2]^2 \leqslant E[\hat{\theta}_{2,n}-\theta_2]^2 \quad \dots \quad (3.8)$$

and a similar inequality holds for any other convex loss function (including the MAD). Finally, by direct enumeration of the $d(\nu, n)$, $\nu \geqslant 2$, we obtain that

$$U^*_{\nu,2} = c_{\nu,n}s^2_{(\nu)} \text{ where } c_{\nu,n} = 1-(\Delta^\nu O^{n-1})/(\Delta^\nu O^n), \text{ for } \nu \geqslant 2, \quad \dots \quad (3.9)$$

and this agrees with Pathak (1962).

4. Admissible estimators of variance

Motivated by (3.8) and (3.9), we raise the following question : How good is the estimator $U^*_{\nu,2}$ for σ^2 ? It is admissible in an appropriate sub-class of estimators of σ^2 ? In the following study, we (partly) resolve this issue.

We confine ourselves to the class Q of homogeneous, quadratic, unbiased estimators of σ^2 of the form

$$\hat{\sigma}^2 = c_\nu s^2_{(\nu)} \text{ if } \nu \geqslant 2, \text{ and } = 0, \text{ for } \nu = 1. \quad \dots \quad (4.1)$$

Note that in SRSWR (N, n), we have

$$P\{\nu_n = k\} = \binom{N}{k}\Delta^k 0^n/N^n, \text{ for } k = 1, n \text{ (and 0, otherwise)}, \quad \dots \quad (4.2)$$

so that the c_ν satisfy the restraint :

$$\sum_{k=1}^{n} c_k P\{\nu_n = k\} = (N-1)/N. \quad \dots \quad (4.3)$$

Clearly, (3.9) satisfies (4.3). The choice of simpler a $c_\nu = (N-1)N^{-1}(1-N^{-n+1})^{-1}$, for $\nu \geqslant 2$, also satisfies (4.3) ; in this case, c_ν does not depend on ν (only on N, n), and (4.1) reduces to $N^{-1}(N-1)(1-N^{-n+1})^{-1}s^2_{(\nu)}$, for every $\nu \geqslant 2$ (and 0, for $\nu = 1$).

We want to consider admissible unbiased variance estimators of the form (4.1) under a quadratic loss. Towards this, we compute first the variance (i.e., risk under the quadratic loss) of $\hat{\sigma}^2$ in (4.1). Clearly, denoting by E_2 and V_2, the conditional expectation and variance, given $\nu_n = \nu$, and by E_1 and V_1, the expectation and variance on ν, we have

$$V(\hat{\sigma}^2) = V_1[E_2(\hat{\sigma}^2)]+E_1[V_2(\hat{\sigma}^2)] = T_1+T_2, \text{ say} \quad \dots \quad (4.4)$$

where

$$T_1 = S^4V_1(c^2_\nu) = S^4\Big[\sum_{k=2}^{n} c^2_k P\{\nu_n = k\}-((N-1)/N)^2\Big]; \; S^2 = \frac{\sum_{i=1}^{N}(Y_i-\overline{Y}_N)^2}{N-1} \quad \dots \quad (4.5)$$

and

$$V_2(\hat{\sigma}^2) = \begin{cases} c_\nu^2 \, V(s_{(\nu)}^2 \mid \text{SRSWR}(N, \nu)) & \nu \geqslant 2, \\ 0, & \text{otherwise}, \end{cases} \qquad \ldots \quad (4.6)$$

Note that s_k^2 is a U-statistic (of degree 2), and hence, by reference of Nandi and Sen (1963), we obtain that for every $k \geqslant 2$,

$$V(s_k^2 \mid \text{SRSWR}(N, k)) = \frac{4(N-k)(k-2)}{k(k-1)(N-3)} \zeta_1 + \frac{2}{k(k-1)} \zeta_2 \left[1 - \frac{(k-2)(k-3)}{(N-2)(N-3)} \right] \qquad \ldots \quad (4.7)$$

where

$$\xi_2 = \binom{N}{2}^{-1} \sum_{1 \leqslant i < j \leqslant N} [\tfrac{1}{2}(Y_i - Y_j)^2 - S^2]^2 = S^{*4} - S^4 ; \qquad \ldots \quad (4.8)$$

$$S^{*4} = \{ \sum_{1 \leqslant i < j \leqslant N} (Y_i - Y_j)^4 \} / \{2N(N-1)\}, \qquad \ldots \quad (4.9)$$

$$\xi_1 = [\Sigma_N^* \{\tfrac{1}{2}(Y_i - Y_j)^2 - S^2\} \{\tfrac{1}{2}(Y_i - Y_k)^2 - S^2\}] / [N(N-1)(N-2)], \quad \ldots \quad (4.10)$$

and the summation Σ_N^* extends over all distinct (i, j, k) from $\{1, \ldots, N\}$.

By virtue of (4.6), (4.7), (4.8), (4.9) and (4.10), we obtain that

$$T_2 = \zeta_1 \{ \sum_{k \geqslant 2} 4(N-k)(k-2) c_k^2 P\{\nu_n = k\} / [k(k-1)(N-3)] \}$$

$$+ 2\zeta_2 \left\{ \sum_{k \geqslant 2} \frac{c_k^2 P\{\nu_n = k\}}{k(k-1)} \left[1 - \frac{(k-2)(k-3)}{(N-2)(N-3)} \right] \right\}. \qquad \ldots \quad (4.11)$$

Thus, we may combine (4.5) and (4.11) to provide an expression for $V(\hat{\sigma}^2)$. Alternatively, we may also set

$$\begin{aligned} V(\hat{\sigma}^2) &= E(\hat{\sigma}^2)^2 - \sigma^4 = E_1[E_2(\hat{\sigma}^4)] - \sigma^4 \\ &= \sum_{k \geqslant 2} P\{\nu_n = k\} \, E(c_k^2 s_{(\nu)}^4 \mid \nu = k) - \sigma^4 \\ &= \sum_{k \geqslant 2} P\{\nu_n = k\} \, c_k^2 E(s_{(\nu)}^4 \mid \nu = k) - \sigma^4. \end{aligned} \qquad \ldots \quad (4.12)$$

To suggest an admissible variance estimator, we now proceed as follows. We choose a particular point $\boldsymbol{Y}_0 \in R^N$ and compute ζ_1 and ζ_2 based on $\boldsymbol{Y}_0$. Then $V(\hat{\sigma}^2)$ becomes a quadratic function in the c_k, $k \geqslant 2$. Next, we minimize this quadratic expression subject to the linear restraint in (4.3). If the resulting solution, say, $\{c_{k0} ; 2 \leqslant k \leqslant n\}$, is unique, then

$$\hat{\sigma}_0^2 = c_{k0} s_{(\nu)}^2 \text{ for } \nu_n = k \geqslant 2, \text{ and } = 0 \text{ for } \nu_n = 1, \qquad \ldots \quad (4.13)$$

turns out to be admissible in the relevant class of competing unbiased estimators of σ^2. Suppose now that at $\boldsymbol{Y}_0$, $E(s_{(\nu)}^4 \mid \nu = k)$ is given by ξ_k^0, for

$k \geqslant 2$. Then, using the (Cauchy-Schwarz) inequality that $(\Sigma p_k c_k^2 \xi_k^0)(\Sigma p_k/\xi_k^0) \geqslant (\Sigma c_k p_k)^2 = [(N-1)/N]^2$, we obtain that $[\sum_{k \geqslant 2} c_k^2 P\{\nu_n = k\}\xi_k^0] \geqslant (N-1)^2 N^{-2}[\sum_{k \geqslant 2} P\{\nu_n = k\}/\xi_k^0]^{-1}$, and the strict equality sign holds iff $c_k \propto (\xi_k^0)^{-1}$, for every $k \geqslant 2$, i.e.,

$$c_{k0} = (\xi_k^0)^{-1} / \left\{ \sum_{k=2}^{n} P\{\nu_n = k\} (\xi_k^0)^{-1}, \right\}, \text{ for } k = 2, \ldots, n. \quad \ldots \quad (4.14)$$

This yields a choice of the c_k corresponding to an admissible estimator.

It is quite clear that a large class of admissible solution $\{c_{k0}\}$ may be generated by appropriate choice of $\boldsymbol{Y}_0 \in R^N$. The interest is therefore to see how some of the proposed estimators match with this spectrum. In the following table, we consider some special cases.

TABLE 1 : $\{c_{k0}\}$ CORRESPONDING TO SOME TYPICAL CHOICE OF $\boldsymbol{Y}_0 \in R^N$

	$\boldsymbol{Y}_0$	c_{k0} $(k = 2, \ldots, n)$ (proportional to)
(i)	$(1, \boldsymbol{0}_{N-1})$	k
(ii)	$(1, -1, \boldsymbol{0}_{N-2})$	$k(k-1)/[N(k-1) + k(k+1)]$
(iii)	$(1, 1, \boldsymbol{0}_{N-2})$	$k(k-1)/[N(k-1) + k^2 - 7k + 8]$
(iv)	$(\boldsymbol{1}_M, \boldsymbol{0}_M) : N = 2M$	$k(k-1)/[k(k-1) + 2]$
(v)	$(\boldsymbol{0}_M, \boldsymbol{1}_{M+1}) : N = 2M + 1$	same as in (iv) for large M (N)
(vi)	$(\boldsymbol{0}_{[N\alpha]}, \boldsymbol{1}_{N-[N\alpha]}), 0 < \alpha < 1$	$\dfrac{k(k-1)}{[(k-2)(k-3) + (k-1)/\alpha(1-\alpha)]}$ (large N)
(vii)	$(\boldsymbol{0}_M, \boldsymbol{1}_M, \boldsymbol{2}_M) : N = 3M$	For $n = 3$, $c_2 : c_3 = 2 : 3$, which corresponds to Pathak's choice in (3.9) ; for $n \geqslant 4$, the solution does not lead to (3.9).
(viii)	$(\boldsymbol{0}_M, \boldsymbol{1}_M, \boldsymbol{2}_M, \boldsymbol{3}_M) : N = 4M$	For $n = 4$, $c_2 : c_3 : c_4 = 35.3 : 23.5 : 20.2$ which very closely approximates Pathak's choice in (3.9). However, for $n \geqslant 5$, the solution may not be very close to (3.9).

Some of the relevant computations have been relegated to the Appendix. Following the choice of $\boldsymbol{Y}_0$ in (iv) and (v) in Table 1, we may conclude that the choice of $c_k \propto k$, $k \geqslant 2$, may easily be made, while the choice of $c_k = c(> 0)$, $k \geqslant 2$, is nearly admissible for relatively large populations. We also observe that for $n \geqslant 4$, the simple choice in (viii) may not lead to the Pathak's choice in (3.9). We therefore pose this question : Is it possible to obtain Pathak's choice in (3.9) by allowing $\boldsymbol{Y}_0$ to have n distinct elements, say, $a_0, a_1, \ldots, a_{n-1}$ (with possibly equal or unequal frequencies $m_0, \ldots, m_{n-1}$, where $\sum_{j=0}^{n-1} m_j = N$), for $n \geqslant 4$? For $n \leqslant 5$, we have an affirmative answer, while, for $n \geqslant 6$, we show that the present method of characterization of admissibility does not work out for the Pathak solution in (3.9). We relegate the details of this finding in the Appendix. However, in passing we may comment that for $n \geqslant 6$, the inability of the present method does not necessarily mean that the Pathak solution in (3.9) is inadmissible. However, even if that solution is admissible, a different method of establishing that is needed. We therefore pose that as an open problem.

5. Negative (factorial) moments of $\nu_{N,n}$

We note that for every (N, n), in SRSWR,

$$E[\bar{y}_{(\nu_{N,n})} - \bar{Y}_N]^2 = \sigma^2\{N/(N-1)\}\{E\nu_{N,n}^{-1} - N^{-1}\}, \quad \ldots \quad (5.1)$$

where σ^2 stands for the population variance of $Y_1, \ldots, Y_N$. In fact, higher order moments of $\bar{y}_{(\nu)}$ involve the negative factorial moments of $\nu_{N,n}$ which we may consider as follows. For every $p, q \geqslant 1$, let

$$p^{[q]} = p \ldots (p-q+1) \text{ and } p^{-(q)} = \{p \ldots (p+q-1)\}^{-1}; \quad \ldots \quad (5.2)$$

conventionally, we let $p^{[0]} = p^{-(0)} = 1$. Our objective is to provide a convenient expression for the negative factorial moment :

$$\mu_r^*(n\,; N) = E[\nu_{N,n}^{-(r)}], \text{ for every } r \geqslant 1, (N, n). \quad \ldots \quad (5.3)$$

For every (N, n), let

$$p_{N,k}^{(n)} = P\{\nu_{N,n} = k\},\ k \geqslant 1. \text{ (Note that } p_{N,k}^{(n)} = 0 \text{ for } k > n). \quad \ldots \quad (5.4)$$

Note that for every $k \geqslant 1$, $n \geqslant 1$ and $N \geqslant 1$,

$$p_{N,k}^{(n)} = N^{-n}N^{[k]}c(n, k), \quad \ldots \quad (5.5)$$

where the positive numbers $c(n, k)$ depend only on (n, k), but not on N ; we may refer to Basu (1958) for some details. Thus,

$$N^{-1}(N-k)p_{N,k}^{(n)} = N^{-n}(N-1)^{[k]}c(n, k) = \{(N-1)/N\}^n p_{N-1,k}^{(n)}, \quad \ldots \quad (5.6)$$

for every $k \geqslant 1$, $N \geqslant 1$, $n \geqslant 1$. Hence, for every real a (> -1),

$$\begin{aligned} E\{(\nu_{N,n}+a)^{-1}\} - (N+a)^{-1} &= (N+a)^{-1}E\{(N-\nu_{N,n})/(\nu_{N,n}+a)\} \\ &= N(N+a)^{-1}\sum_{k=1}^{N}\{(N-k)/N\}(k+a)^{-1}p_{N,k}^{(n)} \quad \ldots \quad (5.7) \\ &= N(N+a)^{-1}\sum_{k=1}^{n}(k+a)^{-1}p_{N-1,k}^{(n)}\left(\frac{N-1}{N}\right)^n \\ &= \frac{N}{N+a}\{(N-1)/N\}^n E((\nu_{N-1,n}+a)^{-1}). \end{aligned}$$

Using this chain relation recursively and noting that $E((\nu_{1,n}+a)^{-1}) = (1+a)^{-1}$, we obtain that

$$E\{(\nu_{N,n}+a)^{-1}\} = N^{-n} \sum_{k=1}^{N} N^{[k-1]} (N-k+1)^n/(N+a)^{[k]}. \quad \ldots \quad (5.8)$$

In particular, we have

$$E(\nu_{N,n}^{-1}) = N^{-n} \sum_{k=1}^{N} (N-k+1)^{n-1} = N^{-n} \sum_{j=1}^{N} j^{n-1}. \quad \ldots \quad (5.9)$$

From (5.8), (5.9) and the identity that $x^{-(2)} = x^{-1}-(x+1)^{-1}$, we obtain that

$$\mu_2^*(n\,;N) = N^{-n} \sum_{j=1}^{N} j^{n-1}(1-j/(N+1)). \quad \ldots \quad (5.10)$$

Next, using the identity that $x^{-(q)} = ((q-1)\,!\,)^{-1} \sum_{j=1}^{q} (-1)^{j-1} \binom{q-1}{j-1} (x+j-1)^{-1}$, for all $q \geqslant 1$, we obtain from (5.8) and (5.10) that for every $r \geqslant 1$,

$$\mu_r^*(n\,;N) = N^{-n}((r-1)\,!\,)^{-1} \sum_{j=1}^{r} (-1)^{j-1} \binom{r-1}{j-1} (N+j-1)^{-[j-1]} \times$$

$$\sum_{s=1}^{N} s^{n-1} (s+j-2)^{[j-1]}. \quad \ldots \quad (5.11)$$

Finally, using the fact that $x^{-2} = \sum_{q \geqslant 2} (q-2)\,!x^{-(q)}$, we have from (5.9) and (5.11) that

$$\text{var}(\nu_{N,n}^{-1}) = \sum_{q \geqslant 2} (q-1)^{-1}N^{-n} \sum_{j=1}^{q} \binom{q-1}{j-1}(N+j-1)^{-[j-1]} \times$$

$$\sum_{s=1}^{N} s^{n-1}(s+j-2)^{[j-1]} - N^{-2n}\left(\sum_{j=1}^{N} j^{n-1}\right)^2. \quad \ldots \quad (5.12)$$

Thus, the simple recursion relation in (5.6) provides a very convenient way for the computation of the negative (factorial) moments of $\nu_{N,n}$. A direct evaluation of these moments would have been much more involved. Only the expression in (5.9) has been known so far (vide Pathak (1961), Chakrabarti (1965) and Lanke (1975), for example).

We also want to evaluate the negative factorial moments of $(N-\nu_{N,n})$ [as these may also arise in some applications to be considered later on]. For this purpose, we rewrite (5.6) as

$$p_{N,k}^{(n)} = ((N+1)/N)^n\{(N-k+1)/(N+1)\}p_{N+1,k}^{(n)},\ k \geqslant 1,\ N \geqslant 1,\ n \geqslant 1. \quad \ldots \quad (5.13)$$

Further, we may express $(N-k)^{-1}$ as $\sum_{r=1}^{\infty} (r-1)!\,(N-k+1)^{-(r)}$. Using this together with (5.13), we obtain that

$$E\{(N-\nu_{N,n})^{-1}\} = \sum_{k=1}^{n} (N-k)^{-1} p_{N,k}^{(n)} = \sum_{k=1}^{n} p_{N,k}^{(n)} \Big\{ \sum_{r=1}^{\infty} (r-1)!\,(N-k+1)^{-(r)} \Big\}$$

$$= \sum_{r=1}^{\infty} (r-1)! \,\Big\{ \sum_{k=1}^{n} (N-k+1)^{-(r)} p_{N,k}^{(n)} \Big\}$$

$$= \sum_{r=1}^{\infty} (r-1)! \,\Big\{ \sum_{k=1}^{n} (N-k+1)^{-(r)} \{(N+r)/N\}^{n} (N-k+r)^{(r)} \times$$

$$(N+1)^{-(r)}\, p_{N+r,k}^{(n)} \Big\}$$

$$= \sum_{r=1}^{\infty} (r-1)!\,\{(N+r)/N\}^{n} (N+1)^{-(r)}. \qquad \ldots \quad (5.14)$$

Similarly, using the identity that $(N-x)^{-2} = \sum_{q=2}^{\infty} (q-2)!\,(N-x)^{-(q)}$

$$E\{(N-\nu_{N,n})^{-2}\} = \sum_{k=1}^{n} (N-k)^{-2} p_{N,k}^{(n)} = \sum_{k=1}^{n} \Big\{ \sum_{q=2}^{\infty} (q-2)!\,(N-k)^{-(q)}\, p_{N,k}^{(n)} \Big\}$$

$$= \sum_{q=2}^{\infty} \Big\{ \sum_{k=1}^{n} (N-k)^{-(q)}\, p_{N,k}^{(n)} \Big\} (q-2)! \qquad \ldots \quad (5.15)$$

$$= \sum_{q=2}^{\infty} (q-2)! \,\Big\{ \sum_{k=1}^{n} p_{N,k}^{(n)} \{(q-1)!\}^{-1} \times$$

$$\sum_{j=1}^{q} (-1)^{j-1} \binom{q-1}{j-1} (N-k+j-1)^{-1} \Big\}$$

$$= \sum_{q=2}^{\infty} (q-2)! \sum_{j=1}^{q} (-1)^{j-1} \Big\{ \binom{q-1}{j-1} / (q-1)! \Big\} \times$$

$$\sum_{k=1}^{n} p_{N,k}^{(n)} (N-k+j-1)^{-1}$$

$$= \sum_{q=2}^{\infty} (q-1)^{-1} \sum_{j=1}^{q} (-1)^{j-1} \binom{q-1}{j-1} \times$$

$$\sum_{r=1}^{\infty} (r-1)!\;\; \{(N+r+j-1)/N\}^{n} (N+1)^{-(r)}.$$

Therefore, $\operatorname{var}\{(N-\nu_{N,n})^{-1}\}$ can be computed from (5.14) and (5.15).

Finally, we consider the incomplete negative moments $E\{(\nu_{N,n}-s+1)^{-1} | \nu_{N,n} \geqslant s\}$.

Let us write

$$\phi(n, N, s) = \sum_{k \geqslant s} (k-s+1)^{-1} p_{N,k}^{(n)}, \text{ for } s \geqslant 1, n \geqslant 1 \text{ and } N \geqslant 1, \quad \ldots \quad (5.16)$$

so that

$$E\{(\nu_{N,n}-s+1)^{-1} | \nu_{N,n} \geqslant s\} = \phi(n, N, s)\{\sum_{k \geqslant s} p_{N,k}^{(n)}\}^{-1}. \quad \ldots \quad (5.17)$$

Consider the expression

$$\sum_{k \geqslant s} \{(k-s+1)^{-1}-(N-s+1)^{-1}\} p_{N,k}^{(n)} = \sum_{k \geqslant s} (N-k)p_{N,k}^{(n)}/\{(N-s+1)(k-s+1)\}$$

$$= N(N-s+1)^{-1} \sum_{k \geqslant s} N^{-1}(N-k)p_{N,k}^{(n)}(k-s+1)^{-1}$$

$$= N(N-s+1)^{-1} \sum_{k \geqslant s} \{(N-1)/N\}^n p_{N-1,k}^{(n)}(k-s+1)^{-1} \text{ [by (5.6)]}$$

$$= \{(N-1)/N\}^n N(N-s+1)^{-1} \sum_{k \geqslant s} (k-s+1)^{-1} p_{N-1,k}^{(n)}. \quad \ldots \quad (5.18)$$

Thus, writing $\Lambda(n, N, s) = (N-s+1)^{-1}\{\sum_{k \geqslant s} p_{N,k}^{(n)}\}$, we have from (5.16) and (5.18)

$$\phi(n, N, s) = \Lambda(n, N, s)+\{(N-1)/N\}^n N(N-s+1)^{-1}\phi(n, N-1, s)$$

$$= \Lambda(n, N, s)+\{(N-1)/N\}^n N(N-s+1)^{-1}\Big\{\Lambda(n, N-1, s)$$

$$+\left(\frac{N-2}{N-1}\right)^n \frac{(N-1)}{(N-s)}\phi(n, N-2, s)\Big\}$$

$$= \sum_{j=0}^{N-s-1} \{(N-j)/N\}^n N^{(j)}\{(N-s+1)^{(j)}\}^{-1} \Lambda(n, N-j, s)$$

$$+(s/N)^n\phi(n, s, s)\, N^{(N-s)}\{N-s+1)\,!\}^{-1}, \quad \ldots \quad (5.19)$$

where for $n \geqslant \nu_{N,n} \geqslant s$, by (5.16)

$$\phi(n, s, s) = p_{s,s}^{(n)} = s^{-n}(\Delta^s 0^n), \text{ as } p_{s,k}^{(n)} = 0, \ \forall\, k > s\,; \quad \ldots \quad (5.20)$$

$$\Lambda(n, m, s) = (m-s+1)^{-1} \sum_{k \geqslant s} p_{m,k}^{(n)},\ m \geqslant s, \text{ are known functions.} \quad \ldots \quad (5.21)$$

In particular, for $s = 2$, we have

$$E\{(\nu_{N,n}-1)^{-1} | \nu_{N,n} \geqslant 2\} = \phi(n, N, 2)\{\sum_{k \geqslant 2} p_{N,k}^{(n)}\}^{-1} = \phi(n, N, 2)\,(1-N^{-n+1})^{-1}, \quad \ldots \quad (5.22)$$

where

$$\phi(n, N, 2) = \sum_{j=0}^{N-3} \left(\frac{N-j}{N}\right)^{n-1}(N-j-1)^{-1}\{1-(N-j)^{-n+1}\}+(2/N)^{n-1}\{1-2^{-n+1}\}. \quad \ldots \quad (5.23)$$

Further, along the same line, we have

$$E[\{\nu_{N,n}^{(2)}\}^{-1} \mid \nu_{N,n} \geqslant 2] = E[\nu_{N,n}^{-1}(\nu_{N,n}-1)^{-1} \mid \nu_{N,n} \geqslant 2]$$

$$= E[(\nu_{N,n}-1)^{-1} \mid \nu_{N,n} \geqslant 2] - E[\nu_{N,n}^{-1} \mid \nu_{N,n} \geqslant 2], \quad \ldots \quad (5.24)$$

where

$$E[\nu_{N,n}^{-1} \mid \nu_{N,n} \geqslant 2] = \{E[\nu_{N,n}^{-1}] - p_{N,1}^{(n)}\}/\{1-p_{N,1}^{(n)}\}, \quad \ldots \quad (5.25)$$

so that using (5.9) along with (5.24) and (5.25), we obtain that

$$E[\{\nu_{N,n}^{(2)}\}^{-1} \mid \nu_{N,n} \geqslant 2] = (1-N^{-n+1})^{-1}[\phi(n, N, 2) - N^{-n} \sum_{j=1}^{N} j^{n-1} + N^{-n+1}]. \quad \ldots \quad (5.26)$$

These expressions will be useful for the specific applications in the next section.

6. Some applications

(I) Pathak (1962) suggested the following admissible unbiased estimator of the population mean $\bar{Y}$ in SRSWR(N, n) :

$$\hat{\bar{Y}}_\nu = a_\nu \bar{y}_{(\nu)}, \text{ where } a_\nu = \{(N-\nu)^{-1} - N^{-1}\} / \sum_{k=1}^{n} p_{N,k}^{(n)}\{(N-k)^{-1} - N^{-1}\},$$

$$1 \leqslant \nu \leqslant n \, ; \quad \ldots \quad (6.1)$$

where $\bar{y}_{(\nu)}$ is defined by (2.4) and $\nu = \nu_{N,n}$. We want to provide an expression for the variance of this estimator. Note that $a_\nu = \lambda(\nu^{-1} - N^{-1})^{-1}$ for $\nu = 1, \ldots, n$, where λ is a positive constant. As such, using the same formula as in (4.4), we have

$$V(\hat{\bar{Y}}_\nu) = V_1(E_2\{\hat{\bar{Y}}_\nu\}) + E_1[V_2(\hat{\bar{Y}}_\nu)], \quad \ldots \quad (6.2)$$

with the notations V_1, V_2 and E_1, E_2 explained there. First, we note that

$$E_1[V_2(\hat{\bar{Y}}_\nu)] = E_1[a_\nu^2(\nu^{-1} - N^{-1})S^2] = S^2\lambda^2 E_1\{(\nu^{-1} - N^{-1})^{-1}\}$$

$$= S^2\{E_1(\nu^{-1} - N^{-1})^{-1}\}^{-1} \text{ (on simplification)}$$

$$= N^2 S^2 \{E_1[(N-\nu)^{-1} - N^{-1}]\}^{-1}$$

$$= N^2 S^2 [E_1\{(N-\nu)^{-1}\} - N^{-1}]^{-1}$$

$$= N^2 S^2 \left[\sum_{r=1}^{\infty} (r-1)\,!\,\{(N+r)/N\}^n (N+1)^{-(r)} - N^{-1} \right]^{-1}, \quad \ldots \quad (6.3)$$

where, in the last step, we have made use of (5.14). Next, we note that

$$V_1(E_2(\hat{\bar{Y}}_\nu)) = V_1(a_\nu \bar{Y}) = \bar{Y}^2 V_1(a_\nu) = \bar{Y}^2 \lambda^2 \, V_1\{(\nu^{-1} - N^{-1})^{-1}\}$$

$$= \lambda^2 \bar{Y}^2 \, V_1\{N^2[(N-\nu)^{-1} - N^{-1}]\} = \lambda^2 \bar{Y}^2 N^4 \, V_1\{(N-\nu)^{-1}\}, \quad \ldots \quad (6.4)$$

where by (6.1) $(N^2\lambda)^{-1} = \sum_{k=1}^{n} p_{N,k}^{(n)}\{(N-k)^{-1}-N^{-1}\} = E_1\{(N-\nu)^{-1}-N^{-1}\}$, so that by (6.3) and (6.4), we have

$$V(\hat{\bar{Y}}_\nu) = N^2S^2\Big[\sum_{r=1}^{\infty}(r-1)\,!\,\{(N+r)/N\}^n(N+1)^{-(r)}-N^{-1}\Big]^{-1}$$

$$+\bar{Y}^2[E_1\{(N-\nu)^{-1}\}-N^{-1}]^{-2}V_1\{(N-\nu)^{-1}\}, \quad \ldots \quad (6.5)$$

so that (5.14) and (5.15) may be incorporated in a complete write-up of (6.5). This explains the utility of the negative moments in (5.14) and (5.15).

(II) Consider next some admissible unbiased variance estimators of the form

$$\hat{\sigma}_\nu^2 = c_\nu s_{(\nu)}^2 \text{ where } c \propto \nu, \text{ for } \nu = 2, \ldots, n. \quad \ldots \quad (6.6)$$

We want to compute Var $(\hat{\sigma}_\nu^2)$. This can be done systematically by using (4.5), (4.11) and (5.22), after noting that $c_k \propto k$, for $2 \leqslant k \leqslant n$.

(III) In the same lines, consider the Pathak estimator of the variance.

$$\hat{\sigma}_\nu^2 = (1-\{\Delta^\nu 0^{n-1}\}/\{\Delta^\nu 0^n\})s_{(\nu)}^2,\ \nu \geqslant 2 \ \ (\text{and} = 0, \nu = 1). \quad \ldots \quad (6.7)$$

The expression in (4.4) applies to this situation where we need to take for the c_k, $c_{k,n} = 1-\Delta^k 0^{n-1}/\Delta^k 0^n$, for $k = 2, \ldots, n$. Therefore, looking at (4.5), we observe that our first task is to evaluate

$$\sum_{k=2}^{n} c_{k,n}^2\, p_{N,k}^{(n)} = T_1(n, N), \text{ say}. \quad \ldots \quad (6.8)$$

Also, looking at (4.11), we observe that here we need to evaluate

$$\sum_{k=2}^{n} (N-k)\,(k-2)k^{-1}\,(k-1)^{-1}\,c_{k,n}^2\,p_{N,k}^{(n)} = T_2(n, N), \quad \ldots \quad (6.9)$$

$$\sum_{k=2}^{n} k^{-1}\,(k-1)^{-1}\,c_{k,n}^2\,p_{N,k}^{(n)} = T_3(n, N), \quad \ldots \quad (6.10)$$

$$\sum_{k=2}^{n} (k-2)\,(k-3)k^{-1}\,(k-1)^{-1}\,c_{k,n}^2\,p_{N,k}^{(n)} = T_4(n, N). \quad \ldots \quad (6.11)$$

Since $p_{N,k}^{(n)} = \binom{N}{k}\Delta^k 0^n.\ N^{-n}$, $k = 1, \ldots, n$, we may simplify $T_1\,(n, N)$ as

$$(1-N^{-n+1})-2N^{-1}+N^{-n}\sum_{k\geqslant 2}(\Delta^k 0^{n-1})^2\binom{N}{k}(\Delta^k 0^n)^{-1}$$

$$=(1-N^{-n+1})-2N^{-1}+\Big(\sum_{k\geqslant 2}\{p_{N,k}^{(n-1)}/p_{N,k}^{(n)}\}^2\,p_{N,k}^{(n)}\Big)N^{-2}. \quad \ldots \quad (6.12)$$

Similarly, writing $(N-k)(k-2)/k(k-1) = 2N/k-(N-1)/(k-1)-1$, we may simplify $T_2(n, N)$ as

$$\sum_{k\geqslant 2}\{2N/k-(N-1)/(k-1)-1\}p_{N,k}^{(n)}-2N^{-1}\sum_{k\geqslant 2}\{2N/k-(N-1)/(k-1)-1\}p_{N,k}^{(n-1)}$$
$$+N^{-2}\Big(\sum_{k\geqslant 2}\{2N/k-(N-1)/(k-1)-1\}\{p_{N,k}^{(n-1)}/p_{N,k}^{(n)}\}^2 p_{N,k}^{(n)}, \qquad \dots \quad (6.13)$$

where, for the evaluation of the first two terms in (6.13), the results in Section 5 can readily be incorporated. The last term is, however, trifle harder.

Note that $k^{-1}(k-1)^{-1} = (k-1)^{-1}-k^{-1}$, and $(k-2)(k-3)/k(k-1) = 1+2/(k-1)-6/k$. Hence, the treatment of (6.10) and (6.11) would be very similar to (6.13). Thus, our basic problem reduces to finding an expression for

$$N^{-2}\sum_{k\geqslant 2}(\alpha+\beta k^{-1}+\gamma(k-1)^{-1})\{p_{N,k}^{(n-1)}/p_{N,k}^{(n)}\}^2 p_{N,k}^{(n)}, N \geqslant 1; n \geqslant 2, \qquad \dots \quad (6.14)$$

for arbitrary (α, β, γ). We write $f(k) = \alpha+\beta k^{-1}+\gamma(k-1)^{-1}$, $k = 2, \dots, n$, and also let $b_{n,k} = (\Delta^k 0^{n-1})^2/(\Delta^k 0^n))$, $k=2, \dots, n$. Moreover, for every $m \geqslant 1$, we set

$$A_h^{(0)}(m, n) = m^{-n}\sum_{k\geqslant 2} b_{n,k}\binom{m}{k-h}, \ h = 0, 1, \dots, \qquad \dots \quad (6.15)$$

$$A_h^{(1)}(m, n) = m^{-n}\sum_{k\geqslant 2} k^{-1}b_{n,k}\binom{m}{k-h}, \ h = 0, 1, \dots, \qquad \dots \quad (6.16)$$

$$A_h^{(2)}(m, n) = m^{-n}\sum_{k\geqslant 2} (k-1)^{-1}b_{n,k}\binom{m}{k-h}, \ h = 0, 1, \dots \qquad \dots \quad (6.17)$$

Note then that (6.14) can be written as

$$\alpha A_0^{(0)}(N, n)+\beta A_0^{(1)}(N, n)+\gamma A_0^{(2)}(N, n) = A_0(N, n), \text{ say.} \qquad \dots \quad (6.18)$$

Writing $d_{n,k} = f(k)b_{n,k}$, $k \geqslant 2$, and using the identity that $\binom{N}{k} = \binom{n}{k} + \sum_{j=n}^{N-1}\binom{j}{k-1}$ we obtain from (6.18) that

$$A_0(N, n) = N^{-n}\sum_{k\geqslant 2} d_{n,k}\binom{N}{k} = N^{-n}\sum_{k\geqslant 2} d_{n,k}\left\{\binom{n}{k}+\sum_{j=n}^{N-1}\binom{j}{k-1}\right\}$$
$$= (n/N)^n A_0(n, n)+\sum_{j=n}^{N-1}(j/N)^n\left\{j^{-n}\sum_{k\geqslant 2} d_{n,k}\binom{j}{k-1}\right\}$$
$$= (n/N)^n A_0(n, n)+\sum_{j=n}^{N-1}(j/N)^n A_1(j, n)$$
$$= (n/N)^n A_0(n, n)+\sum_{j=n}^{N-1}(j/N)^n\left\{(n/j)^n A_1(n, n)+\sum_{l=n}^{j-1}(l/j)^n A_2(l, n)\right\}$$
$$= \dots$$
$$= (n/N)^n\left\{\sum_{h=0}^{n}\binom{N-n}{h}A_h(n, n)\right\}, \qquad \dots \quad (6.19)$$

where the $A_h(m, n)$ are defined as in (6.18) with the $A_0^{(r)}$, $r = 0, 1, 2$ being replaced by $A_h^{(r)}$, $r = 0, 1, 2$, respectively, for $h \geqslant 0$. Thus, it suffices to consider suitable expressions for the basic formulae in (6.15), (6.16) and (6.17), for $m = n$. In this development, we have tacitly assumed that $N \geqslant n$; some other adjustments may be necessary for the other case when $n > N$. Since the $A_h^{(r)}(n, n)$ ($r = 0, 1, 2$; $h \geqslant 0$) are independent of N, they can be systematically derived for any given n and for each $h : 0 \leqslant h \leqslant n$. While tabulating these entries for various n is a possibility, we omit these details (for our primary emphasis on the theory only).

Appendix

The entries in Table 1 in Section 4 are based on some intricate computations which are presented here. Consider the point $\mathbf{Y}_0 \in R^N$ given by

$$\mathbf{Y}_0 = (a_0 \mathbf{1}_{m_0}, \ldots, a_{n-1}\mathbf{1}_{m_{n-1}}),\ a_0, \ldots, a_{n-1} \text{ real} \quad \ldots \quad \text{(A.1)}$$

$$m_j \geqslant 0, \text{ for } j = 0, 1, \ldots, n-1, \text{ and } \sum_{j=0}^{n-1} m_j = N.$$

In SRSWR(N, n), suppose that $\nu_{N,n} = k (\leqslant n)$, and let $f_0, \ldots, f_{n-1}$ be respectively the number of units in the sample with distinct units with values $a_0, \ldots, a_{n-1}$. Note that $f_0 + \ldots + f_{n-1} = k$, and $\mathbf{f} = (f_0, \ldots, f_{n-1})$ has the Hypergeometric law

$$\binom{m_0}{f_0}\binom{m_1}{f_1} \cdots \binom{m_{n-1}}{f_{n-1}} \Big/ \binom{N}{k}, \quad \text{for } f_j \geqslant 0, \quad j=0, \ldots, n-1. \quad \ldots \quad \text{(A.2)}$$

This automatically leads us to

$$E\left[f_0^{(r_0)} \cdots f_{n-1}^{(r_{n-1})} \mid \nu_{N,n} = k\right] = k^{(r)} m_0^{(r_0)} \cdots m_{n-1}^{(r_{n-1})} / N^{(r)}, \quad \ldots \quad \text{(A.3)}$$

where $r = r_0 + \ldots + r_{n-1}$ ($\geqslant 0$) and the r_j are all nonnegative integers. Note that the right hand side of (A.3) is a polynomial in k of degree r. Further, the usual joint moments $E[f_0^{r_0} \cdots f_{n-1}^{r_{n-1}} \mid \nu_{N,n} = k]$ can be expressed as linear combinations of the factorial moments (of same and lower orders), and hence, these will be polynomials in k of degree r. Thus, schematically, we may write

$$E\left[f_0^{r_0} \cdots f_{n-1}^{r_{n-1}} \mid \nu_{N,n} = k\right] = \sum_{j=0}^{r-1} \alpha_j(m_0, \ldots, m_{n-1}) k^{r-j} / N^{(r)}, \quad \ldots \quad \text{(A.4)}$$

where $r = r_0 + \ldots + r_{n-1}$ ($\geqslant 1$) and the $\alpha_j(m_0, \ldots, m_{n-1})$ depends on $(r_0, \ldots, r_{n-1})$ as well as the m_s, $0 \leqslant s \leqslant n-1$ (there is no contribution of m_j if $r_j = 0$, $0 \leqslant j \leqslant n-1$).

Next, we note that given $\nu_{N,n} = k$ and $\boldsymbol{f}$, $s^2_{(\nu)} = s^2_k = (k-1)^{-1} \sum_{j=0}^{n-1} f_j \, [a_j - k^{-1} \sum_{s=0}^{n-1} f_s \, a_s]^2$, for $k \geqslant 2$, and $s^2_k = 0$, for $k = 1$. Further, under (A.1),

$$S^2 = (N-1)^{-1} \left[\sum_{j=0}^{n-1} m_j a_j^2 - N^{-1} \left(\sum_{j=0}^{n-1} m_j \, a_j \right)^2 \right]. \qquad \ldots \quad (A.5)$$

Looking back at (4.14), our principal task is to evaluate $E[s^4_{(\nu)} \,|\, \nu = k]$, for $k = 2, \ldots, n$. In this context, we note that under $\nu_{N,n} = k (\geqslant 2)$ and given $\boldsymbol{f}$,

$$s_k^4 = (k-1)^{-2} [\sum_j a_j^4 f_j^2 + \sum_{i \neq j} a_i^2 a_j^2 f_i f_j + k^{-2} (\sum_j f_j \, a_j)^4 \qquad \ldots \quad (A.6)$$

$$-2k^{-1} \{ \sum_j a_j^4 f_j^2 + \sum_{i \neq j} a_i^2 \, a_j^2 f_i f_j^2 + 2 \sum_{i \neq j} a_i^3 \, a_j \, f_i^2 f_j + \sum_{i \neq j \neq s} a_i^2 \, a_j \, a_s \, f_i \, f_j f_s \}].$$

From (A.4) and (A.6), we readily obtain that for $k \geqslant 2$,

$$E[s^4_{(\nu)} \,|\, \nu = k] = (k-1)^{-2} \, [A_1 k^2 + A_2 k + A_3 + A_4 k^{-1}], \qquad \ldots \quad (A.7)$$

where the coefficients $A_1, \ldots, A_4$ depend on $a_0, \ldots, a_{n-1}$, $m_0, \ldots, m_{n-1}$ and (N, n). For the particular cases treated in Table 1, explicit expressions for (A. 6) lead us to the computations of the c_{k_0} reported there. To have a deeper look at the Pathak solution in (3.9), we now look at (A.7) and (4.14). Let us denote by

$$\phi(k) = \{k(k-1)^2\} \, \{\Delta^k 0^n / [\Delta^k 0^n - \Delta^k 0^{n-1}]\}, \text{ for } k = 2, \ldots, n. \qquad \ldots \quad (A.8)$$

Also, let $\boldsymbol{A} = (A_1, \ldots, A_4)'$ and let

$$\begin{pmatrix} 2^3 & 2^2 & 2 & 1 \\ \cdots & \cdots & \cdots & \cdots \\ n^3 & n^2 & n & 1 \end{pmatrix} = \boldsymbol{W} \text{ of order } (n-1) \times 4. \qquad \ldots \quad (A.9)$$

Then the scheme in (A.1) leads to the admissibility of (3.9), through (4.14), if

$$\boldsymbol{WA} = \gamma \boldsymbol{\phi} = \gamma(\phi(2), \ldots, \phi(n-1))', \text{ for some } \gamma \neq 0. \qquad \ldots \quad (A.10)$$

For $n \leqslant 5$, $\boldsymbol{W}$ has full (row–) rank $(n-1) (\leqslant 4)$, and hence, (A.10) holds. Thus, for $n \leqslant 5$, the admissibility of (3.9) can be established by using (4.14). For $n \geqslant 6$, we have some basic difficulties in verifying (4.14) with (A.10). In order that (A.10) holds, we must have $\phi(k) = d_{n1} + d_{n2} k + d_{n3} k^2 + d_{n4} k^3$, for $k = 2, \ldots, n$, where the coefficients d_{n1}, d_{n2}, d_{n3} and d_{n4} do not depend on $k (= 2, \ldots, n-1)$. However, if we look at (A.8), the first factor $k(k-1)^2 = k^3 - 2k^2 + k$ satisfies the above system, but the second factor (i.e., $\Delta^k 0^n / [\Delta^k 0^n -$

$\Delta^k 0^{n-1}\} = e_{n,k}$, say) is not a constant, for all $k\ (= 2, \ldots, n)$. In order that (A.10) holds, for an arbitrary $n (\geqslant 6)$, we require that

$$e_{n,k} = a_{n0} + k^{-1} a_{n1}, \text{ for } k = 2, \ldots, n, \qquad \ldots \quad \text{(A.11)}$$

where a_{n0} and a_{n1} are real numbers. Note that (A.11) ensures that $k(k-1)^2 e_{n,k} = d_{n1} + d_{n2}k + d_{n3}k^2 + d_{n4}k^3$, for every $k = 2, \ldots, n$. To examine (A.11), we write

$$b_{q,k} = \Delta^k 0^q - \Delta^k 0^{q-1}, \text{ for } k, q > 1 \ \ (b_{q,k} = 0 \text{ if } q < k). \qquad \ldots \quad \text{(A.12)}$$

Then we have $e_{n,k} = (b_{k,k} + \ldots + b_{n,k}) / b_{n,k}$, for $k = 2, \ldots, n$, so that

$$e_{n,k} = 1 + b_{n,k}^{-1}\, b_{n-1,k}\, e_{n-1,k}, \text{ for } k = 2, \ldots, n. \qquad \ldots \quad \text{(A.13)}$$

Note that $b_{n-1,n} = 0$ for every $n \geqslant 2$, so that $e_{n,n} = 1$, for every $n \geqslant 2$. Since the $b_{q,k}$ are all positive numbers (for $q \geqslant k$), we immediately obtain from (A.13) that $e_{n,n-1} > 1$, for every $n \geqslant 3$. Thus, $e_{n,n-1} > e_{n,n}$, for every $n \geqslant 3$. Also, using (A.13) in a chain-rule, we obtain that

$$e_{n,2} = 2^{-1} - 2^{-n} \text{ and } e_{n,3} = [3(3^{n-1} - 2^n) + 2]/[2.3^{n-1} - 3.2^{n-1}), \qquad \ldots \quad \text{(A.14)}$$

for every $n \geqslant 3$. For $n \geqslant 6$, we have $n-1$ elements $e_{n,2}, \ldots, e_{n,n}$, and looking at (A.14), we observe that $e_{n,2}$, $e_{n,3}$, $e_{n,n-1}$ and $e_{n,n}\ (=1)$ fail to satisfy (simultaneously) (A.11). Thus, (A.10) does not hold for $n \geqslant 6$. This can also be verified numerically with the specific case of $n = 6$. Here, $(e_{6,2}, \ldots, e_{6,6}) =$ (31/16, 72/13, 117/11, 120/7, 1), and for (A.11) to hold, we require that $k(k+1)[e_{n,k} - e_{n,k+1}]$ is a constant, for every $k = 2, \ldots, n-1$; in this case, we have

k	$k(k+1)\,[e_{n,k} - e_{n,k+1}]$
2	-21.6
3	-61.175
4	-130.130
5	$+484.285$

This clearly shows that (A.10) does not hold for $n = 6$. Thus, the current method of establishing admissibility does not work for the Pathak estimator, for $n \geqslant 6$.

Acknowledgements. The authors are grateful to the referee for his helpful comments on the paper.

References

Asok, C (1980): A note on the comparison between sample mean and mean based on distinct units in sampling with replacement. *Amer. Statist.*, **34**, 158.

Basu, D. (1958): On sampling with and without replacement. *Sankhyā*, Ser. A **20**, 287-294.

Chakrabarti, M. C. (1965): Query and solution. *Jour. Indian Statist. Assoc*, **3**, 228-232.

Lanke, J. (1975): Some contributions to the theory of survey sampling. *Ph.D. Thesis, Univ. of Lund*, Sweden.

Nandi, H. K. and Sen, P. K. (1963): On the properties of U-statistics when the observations are not independent. Part Two: Unbiased estimation of the parameters of a finite population. *Calcutta Statist. Assoc. Bull.*, **12**, 124-148.

Pathak, P. K. (1961): On the evaluation of moments of distinct units in a sample. *Sankhyā*, Ser. A. 23, 415-420.

——— (1962): On simple random sampling with replacement. *Sankhyā*, Ser. A **24**, 287-302.

Raj, D. and Khamis, S. H. (1958): Some remarks on sampling with replacement. *Ann. Math. Statist.*, **29**, 550-557.

Paper received: July, 1987.

Revised: March, 1988.