On Inadmissibility of Variance Estimators In Unequal Probability Sampling

by

Shriram H. Biyani
The University of Minnesota
Technical Report No. 339
January 31, 1979

ABSTRACT

On Inadmissibility of Variance Estimators
In Unequal Probability Sampling

Examples are given of sampling designs for which the Yates-Grundy and Ajgaonkar's variance estimators are nonnegative, but inadmissible in the class of nonnegative unbiased quadratic estimators. A non-trivial posterior lower bound can be obtained for any nonnegative definite quadratic function of a finite population. An example is given to show that the Yates-Grundy estimator can take values smaller than this bound, even when the estimator is non-negative.

Key words: Finite population, Variance of Horvitz-Thompson estimator, Yates-Grundy Estimator, Inadmissibility.

1. Introduction

Consider a finite population of N units labelled 1,2,...,N. Let $y_1, y_2, \ldots, y_N$ be the values of some characteristic y for the N units, respectively. Assume the range of $\underline{y} = (y_1, y_2, \ldots, y_N)$ to be R^N.

For a given sampling design, let π_i denote the inclusion probability of a unit i, and π_{ij} denote the joint inclusion probability of units i and j, $1 \le i,j \le N$. In particular, $\pi_{ii} = \pi_i$. The well known Horvitz-Thompson estimator of the population total is

$$e_{HT}(s) = \sum_{i \in s} z_i \text{ , where } z_i = y_i/\pi_i \text{ .} \tag{1.1}$$

Restricting the discussion to fixed sample size designs, the variance of e_{HT} can be expressed as

$$V(e_{HT}) = \tfrac{1}{2} \sum_{i,j \varepsilon s} (\pi_i \pi_j - \pi_{ij}) (z_i - z_j)^2 \text{ .} \tag{1.2}$$

Two unbiased estimators of $V(e_{HT})$, proposed by Horvitz and Thompson (1952) and Yates and Grundy (1953) respectively, are

$$v_{HT}(s) = \sum_{i,j \varepsilon s} (\pi_{ij} - \pi_i \pi_j) \pi_{ij}^{-1} . z_i z_j \text{ , and} \tag{1.3}$$

$$v_{YG}(s) = \tfrac{1}{2} \sum_{i,j \varepsilon s} (\pi_i \pi_j - \pi_{ij}) \pi_{ij}^{-1} . (z_i - z_j)^2 \text{ .} \tag{1.4}$$

The Yates-Grundy estimator is usually preferred to v_{HT} on the grounds that it can be prevented from taking negative values by choosing a design satisfying the condition

$$\pi_{ij} \leq \pi_i \pi_j, \quad 1 \leq i \neq j \leq N. \tag{1.5}$$

No comparable result is available for v_{HT}. In fact, it can be shown (see Remark 1), that v_{HT} can take negative values for all designs for which v_{YG} can, and many more. The Yates-Grundy estimator has two additional desirable properties not shared by v_{HT}. First, when $z_1 = z_2 = \ldots = z_N$, $V(e_{HT})$ vanishes, and as one might expect, v_{YG} is identically zero. Second, v_{YG} is invariant under a shift of origin of the z's, i.e., a transformation of the form $(z_1, z_2, \ldots, z_N) \to (z_1+c, z_2+c, \ldots, z_N+c)$. However, Godambe and Joshi (1965) showed that v_{HT} is admissible, with respect to squared error loss, in the class of unbiased estimators of $V(e_{HT})$. Of course, it is not admissible among all estimators, in general, as it can take negative values. But rather surprisingly, the only available admissibility result for v_{YG} is due to Joshi (1970), for sample size two, in the class of unbiased estimators. Examples 1 and 2 show that for higher sample sizes, v_{YG} can be inadmissible even in the narrower class of non-negative unbiased quadratic estimators. In Example 2, another variance estimator due to Ajgaonkar (1967) is also shown to be inadmissible. In Example 3, v_{YG} is shown to take values smaller than a lower bound for $V(e_{HT})$ obtainable from the sample. This provides an example of inadmissibility of v_{YG} for n=2, in the class of non-negative quadratic estimators. Some results useful in the construction of these examples are given below.

2. Some Useful Results

Theorem (Vijayan 1975): An unbiased polynomial estimator v of $V(e_{HT})$ can be non-negative for all $\underline{y}$, only if it is of the form

$$v(s) = \sum_{i,j \varepsilon s} b_{sij} \phi_{ij} \tag{2.1}$$

where b_{sij} are constants and $\phi_{ij} = (z_i - z_j)^2$.

Remark 1: In his Corollary 1, Vijayan has claimed without proof, that v_{HT} can take negative values for any sampling design. However, this is not true, for most of the well known equal probability designs, including simple random sampling (s.r.s.) and stratified sampling with s.r.s. within strata. For these designs, both v_{HT} and v_{YG} coincide with the usual estimators, which are nonnegative. However, we can see from the following argument, that v_{HT} is nonnegative only if it coincides with v_{YG} (and the latter is nonnegative). If v_{HT} is to be nonnegative, then it must be of the form (2.1). That is, for some constants b_{sij} ,

$$\sum_{i,j \varepsilon s} (\pi_{ij} - \pi_i \pi_j)\, \pi_{ij}^{-1} \cdot z_i z_j = \sum_{i,j \varepsilon s} b_{sij} (z_i - z_j)^2. \tag{2.2}$$

Equating the coefficients of $z_i z_j$ on both sides of (2.2) gives the desired conclusion.

LEMMA 1. Let $c_1, c_2, \ldots, c_k$ be constants such that $\sum_{i=1}^{k} c_i = 0$, and let $\phi_{ij} = (z_i - z_j)^2$. Then

$$\sum_{i,j=1}^{k} c_i c_j \phi_{ij} \leq 0, \text{ with equality if and only if } \sum_{i=1}^{k} c_i z_i = 0.$$

PROOF. $\sum_{i,j=1}^{k} c_i c_j \phi_{ij} = \sum_{i,j=1}^{k} c_i c_j (z_i^2 + z_j^2) - 2 \sum_{i,j=1}^{k} c_i c_j z_i z_j.$

The first sum vanishes by hypothesis, and the second sum is $-2(\sum_{i=1}^{k} c_i z_i)^2$.

COROLLARY 1. $\phi_{12} \leq 2(\phi_{13} + \phi_{23})$.

This is a special case of Lemma 1 with $k=3$, $c_1 = c_2 = 1$, $c_3 = -2$.

3. Inadmissibility Examples for $n>2$.

Example 1:

Let $3 \leq n \leq N/2$. Consider a design p of fixed sample size n, such that

$$\begin{aligned} p(s) &= a, \text{ if s includes any of the units with labels } 1, 2, \ldots, n, \\ &= b \text{ otherwise,} \end{aligned}$$

where $a,b > 0$ are to be suitably chosen subject to

$$\binom{N-n}{n} b + \left[\binom{N}{n} - \binom{N-n}{n}\right] a = 1. \tag{3.1}$$

We have

$$\begin{aligned} \pi_i &= \binom{N-1}{n-1} a, && \text{if } i \leq n, && (3.2) \\ &= \binom{N-n-1}{n-1} b + \left[\binom{N-1}{n-1} - \binom{N-n-1}{n-1}\right] a, && \text{if } i > n. \end{aligned}$$

$$\begin{aligned} \pi_{ij} &= \binom{N-2}{n-2} a, && \text{if } i \text{ or } j \leq n, i \neq j, && (3.3) \\ &= \binom{N-n-2}{n-2} b + \left[\binom{N-2}{n-2} - \binom{N-n-2}{n-2}\right] a, && \text{if } i \text{ and } j > n, i \neq j. \end{aligned}$$

Choose 'a', such that

$$\pi_{ij} - \pi_i\pi_j = 0 \text{ for } 1 \le i \ne j \le n. \quad (3.4)$$

That is, $a = \binom{N-2}{n-2}/\binom{N-1}{n-1}^2$.

From (3.1) we get $\binom{N-n}{n}(b-a) = 1 - \binom{N}{n}a = (N-n)/[n(N-1)] > 0$. It follows that b-a>0, and hence b>0 as required.

It can be easily verified that (1.5) holds, and hence v_{YG} is non-negative. In particular, for $i \le n$ and $j > n$,

$$(\pi_i\pi_j - \pi_{ij})/\pi_{ij} = (n-1)^{-1} > 0. \quad (3.5)$$

Now consider another unbiased quadratic estimator v_h defined as follows (with h to be suitably chosen).

Let $s_1 = \{1,2,\ldots,n\}$, $s_2 = \{1,2,\ldots,n-1, n+1\}$, and

$$\begin{aligned} v_h(s) &= v_{YG}(s) + h\phi_{12}, && \text{if } s=s_1, \\ &= v_{YG}(s) - h\phi_{12}, && \text{if } s=s_2, \\ &= v_{YG}(s) && \text{otherwise.} \end{aligned}$$

For any h, the unbiasedness of v_h can be easily seen from $E(v_h - v_{YG}) = 0$.

$$\begin{aligned} V(v_{YG}) - V(v_h) &= \sum_s p(s)\left[v_{YG}^2(s) - v_h^2(s)\right] \\ &= \sum_{s=s_1,s_2} a\left[v_{YG}(s) - v_h(s)\right]\left[v_{YG}(s) + v_h(s)\right] \end{aligned}$$

$$= -ah\phi_{12}\left[2v_{YG}(s_1) + h\phi_{12}\right] + ah\phi_{12}\left[2v_{YG}(s_2) - h\phi_{12}\right]$$

$$= 2ah\phi_{12}\left[-h\phi_{12} - v_{YG}(s_1) + v_{YG}(s_2)\right].$$

By (3.4) $v_{YG}(s_1) = 0$, and

$$v_{YG}(s_2) = \sum_{i=1}^{n-1} (\pi_i\pi_{n+1} - \pi_{i,n+1})\pi_{i,n+1}^{-1} \cdot \phi_{i,n+1}$$

$$= (n-1)^{-1}\sum_{i=1}^{n-1} \phi_{i,n+1} \qquad \text{by (3.5)}$$

$$\geq (n-1)^{-1} (\phi_{1,n+1} + \phi_{2,n+1})$$

$$\geq \phi_{12}/[2(n-1)] \qquad \text{by Corollary 1.}$$

Hence,

$$V(v_{YG}) - V(v_h) \geq 2ah\phi_{12}[-h\phi_{12} + \phi_{12}/\{2(n-1)\}]$$

$$\geq 0, \qquad \text{if } 0 < h < 1/[2(n-1)].$$

Strict inequality holds whenever $\phi_{12} \neq 0$, i.e., $z_1 \neq z_2$.

Using Corollary 1, it is easy to see that for the above choice of h, v_h is nonnegative. This shows that v_{YG} is inadmissible among nonnegative unbiased quadratic estimators of $V(e_{HT})$, for the design in the example.

The basic idea used in the example is simple. For the sample s_1, the Yates- Grundy estimator has the value zero, which "common-sense" tells us to be too low. Hence, the alternative estimator increases the value of v_{YG} for the sample s_1 and compensates for it elsewhere, to retain unbiasedness. A similar idea is used in the next example.

Example 2:

Let $N = 5$, $n = 4$, and

$$p(s) = 1/2, \text{ if } s = \{1,2,3,4\}$$
$$= 1/8 \text{ otherwise.}$$

The following unbiased estimator of $V(e_{HT})$ was proposed by Ajgaonkar (1967).

$$v_A(s) = \tfrac{1}{2}\sum_{i,j\in s}(\pi_i\pi_j - \pi_{ij})\phi_{ij}\Big/\left[\binom{N-2}{n-2}p(s)\right]. \tag{3.6}$$

For the above design, (1.5) holds, and hence v_A and v_{YG} are both nonnegative. However, both can be seen to be inadmissible in the class of nonnegative unbiased quadratic estimators.

Two estimators v_1, v_2 dominating v_A and v_{YG} respectively, are given by

$$\begin{aligned} v_1(s) - v_A(s) = v_2(s) - v_{YG}(s) &= h\phi_{12}, \quad \text{if } s = \{1,2,3,4\} \\ &= -2h\phi_{12}, \quad \text{if } s = \{1,2,3,5\} \text{ or } \\ &\qquad\qquad\qquad \{1,2,4,5\} \\ &= 0 \qquad \text{otherwise,} \end{aligned}$$

and $h>0$ to be suitably chosen.

After some simple algebra, we get

$$V(v_1) - V(v_A) = \tfrac{1}{2}h\phi_{12}\Big[3h\phi_{12} - 1/48(3\phi_{12}+\phi_{13}+\phi_{23}+\phi_{14}+\phi_{24}-\phi_{34}) - 1/6(2\phi_{15}+2\phi_{25}+\phi_{35}+\phi_{45})\Big], \text{ and}$$

$$V(v_2) - V(v_{YG}) = \tfrac{1}{2}h\phi_{12}\Big[3h\phi_{12} + 1/48(\phi_{13}+\phi_{23}+\phi_{14}+\phi_{24}+2\phi_{34}) - 1/6(2\phi_{15}+2\phi_{25}+\phi_{35}+\phi_{45})\Big].$$

It is easy to see that for h=1/72,

$$V(v_1) - V(v_A) \le V(v_2) - V(v_{YG})$$
$$\le 1/144\phi_{12} \cdot 1/24 \left[\phi_{12}+\phi_{13}+\phi_{23}+\phi_{14}+\phi_{24}+\phi_{34} -4(\phi_{15}+\phi_{25}+\phi_{35}+\phi_{45})\right]$$
$$\le 0,$$

since the quantity inside the square brackets is nonpositive by Lemma 1, with k=5, $c_1=c_2=c_3=c_4=1$ and $c_5=-4$. Strict inequality holds almost everywhere.

4. A Lower Bound For Nonnegative Definite Quadratic Forms.

The following result from matrix algebra can be used to obtain a lower bound for any nonnegative definite quadratic function of a finite population vector $\underline{y}$, based on a sample of observations.

Let $\underline{y}$ be a Nx1 vector and A be an NxN symmetric nonnegative definite matrix. If $\underline{y}' = (\underline{y}_1', \underline{y}_2')$ and A is correspondingly partitioned as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \text{ then}$$

$$y'A\,y \ge \underline{y}_1'(A_{11} - A_{12}A_{22}^*A_{21})\underline{y}_1 + \underline{y}_{2.1}'A_{22}\underline{y}_{2.1}, \qquad (4.1)$$

where A_{22}^* is symmetric, and satisfies $A_{22}A_{22}^*A_{22} = A_{22}$ and $A_{22}^*A_{22}A_{22}^* = A_{22}^*$. The result can be proved by verifying the identity

$$\underline{y}'A\underline{y} = \underline{y}_1'(A_{11} - A_{12}A_{22}^*A_{21})\underline{y}_1 + \underline{y}_{2.1}'A_{22}\underline{y}_{2.1},$$

where $\underline{y}_{2.1} = A_{22}^*(A_{22}\underline{y}_2 + A_{21}\underline{y}_1)$.

It may be noted that the bound given by (4.1) is the best possible, based on y_1, because it is attained when $\underline{y}_2 = -A_{22}^{*}A_{21}\underline{y}_1$.

Remark 2: If $\underline{y}_1$ denotes the vector of the sample observations, then (4.1) provides a lower bound for $\underline{y}'A\underline{y}$, known after the sample is drawn. Thus it seems reasonable to expect that any estimator of $\underline{y}'A\underline{y}$ based on $\underline{y}_1$ should be not only nonnegative, but also not smaller than the bound given by (4.1). Any estimator which fails to satisfy this would be inadmissible, not only with respect to "squared error loss", but with respect to any loss function which increases with the distance between the estimate and the "true value".

The following example shows that the Yates-Grundy estimator can take values smaller than the lower bound obtainable from the sample.

Example 3:

Let N=3, n=2 and consider the design p, such that $p(\{1,3\}) = p(\{2,3\}) = .2$ and $p(\{1,2\}) = .6$. We have $\pi_1 = \pi_2 = .8$, $\pi_3 = .4$, $\pi_{12} = .6$, $\pi_{13} = \pi_{23} = .2$.

Substituting in (1.2), we get

$$V(e_{HT}) = .04(Z_1-Z_2)^2 + .12(Z_1-Z_3)^2 + .12(Z_2-Z_3)^2 = \underline{Z}'\,A\,\underline{Z}\,,$$

where $$\underline{Z} = \begin{bmatrix} Z_1 \\ Z_2 \\ Z_3 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 0.16 & -.04 & -.12 \\ -.04 & 0.16 & -.12 \\ -.12 & -.12 & 0.24 \end{bmatrix}.$$

The lower bound for $V(e_{HT})$ based on the sample $\{1,2\}$ is,

$$[Z_1\ Z_2]\left\{\begin{bmatrix} .16 & -.04 \\ -.04 & .16 \end{bmatrix} - \begin{bmatrix} -.12 \\ -.12 \end{bmatrix}(0.24)^{-1}[-.12, -12]\right\}\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}$$

$$= [Z_1\ Z_2]\begin{bmatrix} .10 & -.10 \\ -.10 & .10 \end{bmatrix}\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = (Z_1 - Z_2)^2/10\,.$$

But, for $s =' \{1,2\}$, $v_{YG}(s) = .04(Z_1 - Z_2)^2/.6 = (Z_1 - Z_2)^2/15$, which is smaller than the above bound, except when $Z_1 = Z_2$. The estimator which equals $(Z_1 - Z_2)^2/10$ for $s =' \{1,2\}$, and agrees with v_{YG} for other samples is uniformly better than v_{YG}. This shows the inadmissibility of v_{YG} for the above design, in the class of nonnegative quadratic estimators.

5. Remarks

Comparing expressions (1.2) and (1.4), we can see that the Yates-Grundy estimator is obtained by dividing the general term in the expression for $V(e_{HT})$ by the corresponding inclusion probability, and restricting the sum to the sample instead of the population. A similar process is used in constructing the Horvitz-Thompson estimator e_{HT} of the population total. Although this is a neat way of constructing an unbiased estimator, the result may not always be a good estimator. We can rationalize e_{HT} with a superpopulation model under which the expected value of y_i is proportional to π_i. For example, Godmabe and Thompson (1973) have shown that e_{HT} is optimal, in some sense, when $y_1/\pi_1, y_2/\pi_2, \ldots, y_N/\pi_N$ are the realized values of exchangeable random variables. However, such a rationalization may not always be possible, and this is the case with v_{YG}. It may be tempting to consider a model under which $E\left[(\pi_i\pi_j - \pi_{ij})\pi_{ij}^{-1}\phi_{ij}\right]$ is constant over all $i \neq j$. But such a model may not be consistent, because the ϕ_{ij}'s are subject to constraints such as given by Lemma 1. Thus the factors π_{ij}^{-1} in case of v_{YG} and $[p(s)]^{-1}$ in case of v_A are no more than arbitrary inflation factors, introduced to satisfy an algebraic relation (unbiasedness). Examples 1 and 2 show that better unbiased estimators might exist for $n>2$. But for $n=2$, it is known (Lanke 1974) that v_{YG} is the unique nonnegative unbiased estimator of $V(e_{HT})$. However, in example 3, we see that v_{YG} is not a reasonable estimator. This suggests that we must not restrict ourselves to estimators which are unbiased with respect to the sampling design.

REFERENCES

1. Ajgaonkar, S.G.P. (1967), "Unbiased estimator of variance of Narain, Horvitz and Thompson estimator." Sankhya Ser. A, 29, 55-60

2. Godambe, V.P. and Joshi, V.M. (1965), "Admissibility and Bayes estimation in sampling finite populations - I." Annals of Mathematical Statistics, 36, 1707-1722.

3. Horvitz, D.G. and Thompson, D.J. (1952), "A generalization of sampling without replacement from a finite universe." Journal of American Statistical ASsociation, 47, 663-685.

4. Joshi, V.M. (1970), "Note on the admissibility of Sen-Yates-Grundy estimator." Sankhya Ser. A, 32, 431-438.

5. Lanke, J. (1974), "On nonnegative variance estimation in sample surveys." Sankhya Ser. C, 36, 33-42.

6. Vijayan, K. (1975), "On estimating the variance in unequal probability sampling." Journal of American Statistical Association, 70, 713-716.

7. Yates, F. and Grundy, P.M. (1953), "Selection without replacement from within strata with probability proportional to size." Jounral of Royal Statistical Society Ser. B, 15, 235-261.