Total Estimators in Inverse Sampling

L. Greco
Dipartimento di Metodi Quantitativi, Università degli Studi di Siena
greco2@unisi.it

S. Naddeo
Dipartimento di Metodi Quantitativi, Università degli Studi di Siena
naddeo@unisi.it

Abstract. Inverse sampling is an efficient design when the population of interest is subdivided into two groups, one of which contains only few units. With this design it is possible to obtain efficient parameter estimators for the whole population as well as for its subgroups. We consider unequal selection probabilities in sampling with replacement, as well as equal selection probabilities in sampling with and without replacement.

Keywords. sequential sampling, rare populations, unequal selection probabilities, stopping rules.

1. Introduction

Inverse sampling is an efficient design for estimating the parameters of a population in which only few units display the characteristic of interest or when the variable of interest tends to be at or near zero for many of the population units and only one subgroup exhibits values different from zero. This design is also efficient when the aim of the survey is to estimate the parameters of a subgroup formed by few units, in addition to the parameters of the whole population. In many real situations the population may be considered as subdivided into two subgroups, and the subgroup to which a unit belongs is not known until the unit is sampled. In these cases the classical fixed sample-size designs may not give efficient estimates for the population parameters. Instead, in inverse sampling designs the selection of units continues until a prefixed number of units with the characteristic of interest is observed in the sample.

Recently, Christman and Lan (2001) considered inverse sampling in simple random sampling with and without replacement, when all the population units have equal selection probabilities. They provided some unbiased estimators of the population total and their variances, but not the variance estimators.

Salehi and Seber (2001, 2004) considered inverse sampling with equal selection probabilities by using Murthy's method, showing that this method is an application of the Rao-Blackwell theorem. With this approach they obtained the unbiased estimator of the population total using simple random sampling without replacement, with the same results given by Christman and Lan (2001). Salehi and Seber (2001, 2004) also obtained a variance estimator which is unbiased on the basis of Murthy's method.

In this paper we consider inverse sampling when the units have unequal selection probabilities, as happens in environmental surveys when the selection probabilities are evaluated at the moment in which the units are selected. We derive unbiased estimators of the totals of the two subgroups, their variance and the corresponding variance estimators in inverse sampling with replacement (ISWR). Our estimator of the whole population total is equivalent to that given by Christman and Lan (2001) when the population units have equal selection probabilities.

Finally, we obtain similar results in inverse sampling without replacement (ISWOR) when the population units have equal selection probabilities, and we show that the variance estimator reported by Salehi and Seber (2001, 2004) is the unbiased estimator of the variance given by Christman and Lan (2001).

2. Unequal selection probabilities

Let us consider a population of N units divided into two groups of N_1 and N_2 units respectively, with $N_1+N_2=N$, and let y_{1i} (i=1,2,…,N_1) and y_{2i} (i=1,2,…,N_2) be the values of the variable of interest Y in the first and second group. For example, the population can be subdivided into two subgroups according to whether the y-values satisfy some condition C. Moreover, let T_1 and T_2 be the totals of the two groups and $T=T_1+T_2$ the population total, p_{1i} the selection probability of the i-th unit of the first group and p_{2i} the selection probability of the i-th unit of the second group. Finally, let $P=\sum_{i=1}^{N_1} p_{1i}$ be the selection probability of the first group and $1-P=\sum_{i=1}^{N_2} p_{2i}$ the selection probability of the second group.

In inverse sampling designs, the selection of units continues until k units of the first group are observed in the sample, so that the sample size ν is a random variable which can assume the values k, k+1, ... Let us consider the event $\nu=n$ and all the possible samples of size n. Among them the only samples taken into account with inverse sampling are formed by k units coming from the first group and $n-$k units from the second group.

In ISWR the total probability of all samples is

$$\mathrm{P}^{\mathrm{k}}\left(1-\mathrm{P}\right)^{n-\mathrm{k}},$$

so that the probability of selecting a sample of n elements in a given order is

$$\frac{\prod_{i=1}^{\mathrm{N}_1}\mathrm{p}_{1i}^{m_{1i}}\prod_{i=1}^{\mathrm{N}_2}\mathrm{p}_{2i}^{m_{2i}}}{\mathrm{P}^{k}\left(1-\mathrm{P}\right)^{n-k}}=\prod_{i=1}^{\mathrm{N}_1}\left(\frac{\mathrm{p}_{1i}}{\mathrm{P}}\right)^{m_{1i}}\prod_{i=1}^{\mathrm{N}_2}\left(\frac{\mathrm{p}_{2i}}{1-\mathrm{P}}\right)^{m_{2i}},$$

where m_{ji} counts how many times the i-th unit of j-th group is in the sample (j=1,2; i=1,…, N_j), with $\sum_{i=1}^{\mathrm{N}_1}m_{1i}=\mathrm{k}$ and $\sum_{i=1}^{\mathrm{N}_2}m_{2i}=n-\mathrm{k}$.

For a given n the selection probability of the i-th unit in the first group is $\frac{\mathrm{p}_{1i}}{\mathrm{P}}$, while $\frac{\mathrm{p}_{2\mathrm{i}}}{1-\mathrm{P}}$ is the selection probability of the i-th unit in the second group. Thus, conditionally to $\nu=n$ the selection procedure with inverse sampling is exactly like the selection procedure with stratified sampling, where the two strata correspond to the two subgroups and k and n–k units are respectively selected from the first and from the second stratum. Obviously, the sample results in the two groups are independent.

Let Y_{ji} be the random variable "value of Y" on the i-th sample unit from the j-th group, so that the estimator of the general total may be written as

$$\hat{T}=\hat{T}_1+\hat{T}_2=\frac{\mathrm{P}}{\mathrm{k}}\sum_{i=1}^{\mathrm{k}}\frac{Y_{1i}}{\mathrm{p}_{1i}}+\frac{1-\mathrm{P}}{\nu-\mathrm{k}}\sum_{i=1}^{\nu-\mathrm{k}}\frac{Y_{2i}}{\mathrm{p}_{2i}}. \qquad (1)$$

Usually, the value of P is not known, but an unbiased estimator is

$$\hat{P}=\frac{\mathrm{k}-1}{\nu-1}, \qquad (2)$$

as can be derived from the negative binomial distribution of ν. Hence, if we denote by

$$W_{ji}=\frac{Y_{ji}}{\mathrm{p}_{ji}}$$

the sample i.i.d. random variables, an estimator of the total is

$$\widetilde{T} = \widetilde{T}_1 + \widetilde{T}_2 = \frac{\hat{P}}{\mathrm{k}} \sum_{i=1}^{\mathrm{k}} W_{1i} + \frac{1-\hat{P}}{\nu - \mathrm{k}} \sum_{i=1}^{\nu-\mathrm{k}} W_{2i} = \hat{P}\overline{W_1} + \left(1-\hat{P}\right)\overline{W_2}, \tag{3}$$

where the means of $\overline{W_1}$ and $\overline{W_2}$ are respectively

$$\mathrm{E}\left(\overline{W_1}\right) = \mathrm{W}_1 = \frac{\mathrm{T}_1}{\mathrm{P}}$$

and

$$\mathrm{E}\left(\overline{W_2}\right) = \mathrm{W}_2 = \frac{\mathrm{T}_2}{1-\mathrm{P}}.$$

Since

$$\mathrm{E}_v\left[\mathrm{E}_Y\left(\widetilde{T}\right)\right] = \mathrm{PW}_1 + (1-\mathrm{P})\mathrm{W}_2 = \mathrm{T}_1 + \mathrm{T}_2,$$

estimator $\widetilde{T}$ is unbiased. As regards its variance, we may observe that

$$\mathrm{V}_v\left[\mathrm{E}_Y\left(\widetilde{T}\right)\right] = (\mathrm{W}_1 - \mathrm{W}_2)^2 = \mathrm{V}_v\left(\hat{P}\right),$$

and

$$\mathrm{E}_v\left[\mathrm{V}_Y\left(\widetilde{T}\right)\right] = \frac{\sigma_{1w}^2}{\mathrm{k}}\mathrm{E}_v\left(\hat{P}^2\right) + \sigma_{2w}^2\mathrm{E}_v\left[\frac{\left(1-\hat{P}^2\right)}{\nu-\mathrm{k}}\right] =$$
$$= \frac{\sigma_{1w}^2}{\mathrm{k}}\mathrm{E}_v\left(\hat{P}^2\right) + \frac{\sigma_{2w}^2}{\mathrm{k}-1}\mathrm{E}_v\left[\hat{P}\left(1-\hat{P}\right)\right]$$

where

$$\sigma_{jw}^2 = \sum_{i=1}^{N_j} \left(W_j - \mathrm{W}_j\right)^2 \mathrm{p}_{ji} \qquad j=1,2\,,$$

so that

$$\mathrm{V}\left(\tilde{T}\right) = \left(\mathrm{W}_1 - \mathrm{W}_2\right)^2 \mathrm{V}_v\left(\hat{P}\right) + \frac{\sigma_{1w}^2}{\mathrm{k}} \mathrm{E}_v\left(\hat{P}^2\right) + \frac{\sigma_{2w}^2}{\mathrm{k}-1} \mathrm{E}_v\left[\hat{P}\left(1-\hat{P}\right)\right].$$

In order to obtain an unbiased variance estimator, let us note that an unbiased estimator of W_1^2 is

$$\hat{W}_{1q} = \overline{W}_1^2 - \frac{\hat{S}_{1w}^2}{\mathrm{k}},$$

while the corresponding unbiased estimator of W_2^2 is

$$\hat{W}_{2q} = \overline{W}_2^2 - \frac{\hat{S}_{2w}^2}{\nu - \mathrm{k}},$$

where $\hat{S}_{jw}^2$ is the unbiased sample variance of the W_{ji}s in the j-th group.

Moreover, an unbiased estimator of P^2 is

$$\hat{P}_q = \frac{(\mathrm{k}-1)(\mathrm{k}-2)}{(\nu-1)(\nu-2)} \tag{4}$$

so that an unbiased estimator of the variance of $\hat{P}$ is

$$\hat{\mathrm{V}}\left(\hat{P}\right) = \hat{P}^2 - \frac{(\mathrm{k}-1)(\mathrm{k}-2)}{(\nu-1)(\nu-2)} = \frac{\hat{P}\left(1-\hat{P}\right)}{\nu-2}.$$

and finally

$$\hat{\mathrm{V}}\left(\tilde{T}\right) = \left(\overline{W}_1 - \overline{W}_2\right)^2 \frac{\hat{P}\left(1-\hat{P}\right)}{\nu-2} + \frac{\hat{S}_{1w}^2}{\mathrm{k}} \hat{P}_q + \frac{\hat{S}_{2w}^2}{\mathrm{k}-1}\left(\hat{P} - \frac{\mathrm{k}-1}{\mathrm{k}-2}\hat{P}_q\right). \tag{5}$$

As regards the two subgroups, observe that the variances of the two unbiased estimators of the totals $\widetilde{T}_1$ and $\widetilde{T}_2$ are respectively

$$\mathrm{V}\left(\widetilde{T}_1\right) = \mathrm{W}_1^2 \mathrm{V}_v\left(\hat{P}\right) + \frac{\sigma_{1w}^2}{\mathrm{k}} \mathrm{E}_v\left(\hat{P}^2\right)$$

and

$$\mathrm{V}\left(\widetilde{T}_2\right) = \mathrm{W}_2^2 \mathrm{V}_v\left(\hat{P}\right) + \frac{\sigma_{2w}^2}{\mathrm{k}-1} \mathrm{E}_v\left[\hat{P}\left(1-\hat{P}\right)\right]$$

and their unbiased estimators may be obtained from expression (5).

3. Stopping rules

With inverse sampling it is possible to run out of resources (e.g., money or time) prior to selecting k units from the first subgroup. In this case it is still possible to get unbiased estimators of the subgroup totals and their variances.

Thus, we consider a sampling design in which the selection procedure stops if: (*i*) k units of the first subgroup are observed in the sample; (*ii*) the maximum sample size M is obtained without getting k units of the first subgroup. In the first case, the total estimator is given in expression (3), while in the second case the estimator is the classical estimator of the totals of the two subgroups used in the fixed sample-size designs when M units are randomly selected from the whole population. Obviously, in the last case the number of selected units from the first subgroup is a random variable *X* with a binomial distribution of parameters P and M.

By using this stopping rule, the unbiased estimator of the population total is

$$\widetilde{T}_s = \begin{cases} \widetilde{T}_a = \hat{P}_a \overline{W}_{1a} + \left(1-\hat{P}_a\right)\overline{W}_{2a} & \text{case } (i) \\ \widetilde{T}_b = \hat{P}_b \overline{W}_{1b} + \left(1-\hat{P}_b\right)\overline{W}_{2b} & \text{case } (ii) \end{cases} \qquad (6)$$

where $\widetilde{T}_a$ is used when the k-th unit of the first subgroup is selected at the *n*-th replication (with k≤*n*≤M), $\widetilde{T}_b$ is the estimator which is computed when *n*=M and the number of sampled units of the

first subgroup is x (with $0 \le x < \mathrm{k}$), $\hat{P}_a = \dfrac{\mathrm{k}-1}{\nu-1}$, $\hat{P}_b = \dfrac{X}{\mathrm{M}}$, and finally $\overline{W}_{1a}$, $\overline{W}_{2a}$, $\overline{W}_{1b}$, $\overline{W}_{2b}$ are the sample means in the two cases.

The variance of estimator (6) is

$$\mathrm{V}\left(\widetilde{T}_s\right) = \sum_{n=\mathrm{k}}^{\mathrm{M}} \mathrm{E}_Y\left[\left(\widetilde{T}_a - \mathrm{T}\right)^2\right] f(n;\mathrm{k},\mathrm{P}) + \sum_{x=0}^{\mathrm{k}-1} \mathrm{E}_Y\left[\left(\widetilde{T}_b - \mathrm{T}\right)^2\right] g(x;\mathrm{M},\mathrm{P}),$$

where $f(n;\mathrm{k},\mathrm{P})$ denotes the negative binomial distribution of parameters k and P, while $g(x;\mathrm{M},\mathrm{P})$ is the binomial distribution of parameters M and P. An unbiased variance estimator is

$$\hat{\mathrm{V}}\left(\widetilde{T}_s\right) = \begin{cases} \hat{\mathrm{V}}_a\left(\widetilde{T}_s\right) = \left(\overline{W}_{1a} - \overline{W}_{2a}\right)^2 \left(\hat{P}_a^2 - \hat{P}_{aq}\right) + \dfrac{\hat{S}_{1w}^2}{\mathrm{k}} \hat{P}_{aq} + \dfrac{\hat{S}_{2w}^2}{\nu - \mathrm{k}} \left(1 - 2\hat{P}_a + \hat{P}_{aq}\right) & \text{case } (i) \\ \hat{\mathrm{V}}_b\left(\widetilde{T}_s\right) = \left(\overline{W}_{1b} - \overline{W}_{2b}\right)^2 \left(\hat{P}_b^2 - \hat{P}_{bq}\right) + \dfrac{\hat{S}_{1w}^2}{X} \hat{P}_{bq} + \dfrac{\hat{S}_{2w}^2}{\mathrm{M} - X} \left(1 - 2\hat{P}_b + \hat{P}_{bq}\right) & \text{case } (ii) \end{cases} \tag{7}$$

where $\hat{P}_{aq}$ is given by (4), $\hat{P}_{bq} = \dfrac{X}{\mathrm{M}} \dfrac{X-1}{\mathrm{M}-1}$ and $\hat{\mathrm{V}}_a\left(\widetilde{T}_s\right)$ is equivalent to (5).

Estimators (6) and (7) are unbiased, as is shown in the Appendix.

The unbiased estimators of T_1 and T_2 and their unbiased variance estimators may be easily obtained from the previous paragraph and from (7).

A more complex design considers an initial sample of a prefixed size, say m. In this case the selection procedure stops if: (*i*) at least *k* units of the first subgroup are selected in the initial sample; (*ii*) exactly k units of the first subgroup are observed in a sample of size *n* (with $\mathrm{m} < n \le \mathrm{M}$); (*iii*) the maximum sample size M is obtained without obtaining k units of the first subgroup.

Obviously, previous results apply in cases (*ii*) and (*iii*), while in case (*i*) the expressions of the totals estimators, of their variances and of the variances estimators may be easily obtained from the previous expressions when the binomial distribution applies, substituting M with m. Also with this stopping rule, the total estimators of the population and of the two subgroups are unbiased, as well as the estimators of the variances of T, T_1 and T_2, as is shown in Appendix.

Finally, it is worth noting that if the procedure selection considers only cases (*i*) and (*ii*), i.e. a maximum sample size M is not fixed, the corresponding estimators are still unbiased.

4. Equal selection probabilities

If the selection probabilities are equal for every population unit so that the selection probability of the *i*-th unit in the *j*-th group (*j*=1,2; *i*=1,2,…,N_j) is 1/N, estimator (3) under ISWR assumes the form

$$\widetilde{T} = \mathrm{N}\left[\hat{P}\overline{Y}_1 + \left(1 - \hat{P}\right)\overline{Y}_2\right] , \tag{8}$$

where $\overline{Y}_1$ and $\overline{Y}_2$ are the sample means in the two subgroups. Its variance is

$$\mathrm{V}\left(\widetilde{T}\right) = \mathrm{N}^2\left\{\left(\mu_1 - \mu_2\right)^2 \mathrm{V}_v\left(\hat{P}\right) + \frac{\sigma_1^2}{\mathrm{k}}\mathrm{E}_v\left(\hat{P}^2\right) + \frac{\sigma_2^2}{\mathrm{k}-1}\mathrm{E}_v\left[\hat{P}\left(1-\hat{P}\right)\right]\right\} \tag{9}$$

where μ_j and σ_j^2 are the mean and the variance of the variable of interest in the *j*-th group (*j*=1,2) respectively. It is worth noting that the estimator of the population total and its variance are equivalent to the expressions given by Christman and Lan (2001).

An unbiased estimator of (9) is

$$\hat{\mathrm{V}}\left(\widetilde{T}\right) = \mathrm{N}^2\left\{\left(\overline{Y}_1 - \overline{Y}_2\right)^2 \frac{\hat{P}\left(1-\hat{P}\right)}{\nu - 2} + \frac{\hat{S}_1^2}{\mathrm{k}}\hat{P}_q + \frac{\hat{S}_2^2}{\mathrm{k}-1}\left[\hat{\mathrm{P}} - \frac{\mathrm{k}-1}{\mathrm{k}-2}\hat{P}_q\right]\right\}$$

where $\hat{S}_j^2$ is the unbiased sample variance in the *j*-th group.

If we consider ISWOR and equal selection probabilities, from the hypergeometric distribution it turns out that the first order inclusion probabilities are the same as in the previous case. In fact, conditionally to $\nu=n$, the number of samples formed by k units from the first subgroup and n–k units from the second subgroup is

$$\binom{\mathrm{N}_1}{\mathrm{k}}\binom{\mathrm{N}_2}{n-\mathrm{k}},$$

while the number of samples containing, for example, the *i*-th unit of the first subgroup is $\binom{\mathrm{N}_1 - 1}{\mathrm{k}-1}\binom{\mathrm{N}_2}{n-\mathrm{k}}$. Thus, its inclusion probability turns out to be $\frac{\mathrm{k}}{\mathrm{N}_1}$. Similarly, the inclusion

probability of any unit coming from the second subgroup is $\frac{n-\mathrm{k}}{\mathrm{N}_2}$. The second order inclusion probability of two units in the first subgroup is $\frac{\mathrm{k}}{\mathrm{N}_1}\frac{\mathrm{k}-1}{\mathrm{N}_1-1}$, the corresponding probability for two units in the second group is $\frac{n-\mathrm{k}}{\mathrm{N}_2}\frac{n-\mathrm{k}-1}{\mathrm{N}_2-1}$, while it is $\frac{\mathrm{k}}{\mathrm{N}_1}\frac{n-\mathrm{k}}{\mathrm{N}_2}$ if the two units come from different subgroups.

Practically, these results show that, conditionally to $\nu=n$, the selection procedure with inverse sampling is exactly like the selection procedure with stratified sampling, where k and n–k units are respectively selected without replacement in the first stratum and, independently, in the second stratum.

Since in this case ν has a negative hypergeometric distribution, an unbiased estimator of P is still given by (2) and the estimator of the population total is given by expression (8). Its variance is

$$\mathrm{V}\left(\widetilde{T}\right)=\mathrm{N}^2\left\{\left(\mu_1-\mu_2\right)^2\mathrm{V}_\nu\left(\hat{P}\right)+\frac{\mathrm{S}_1^2}{\mathrm{k}}\left(1-\frac{\mathrm{k}}{\mathrm{N}_1}\right)\mathrm{E}_\nu\left(\hat{P}^2\right)+\frac{\mathrm{S}_2^2}{\mathrm{k}-1}\mathrm{E}_\nu\left[\hat{P}\left(1-\hat{P}\right)\left(1-\frac{\nu-\mathrm{k}}{\mathrm{N}_2}\right)\right]\right\},$$

where

$$\mathrm{S}_j^2=\frac{1}{\mathrm{N}_j-1}\sum_{i=1}^{\mathrm{N}_j}\left(\mathrm{y}_{ji}-\mu_j\right)^2, \qquad j=1,2,$$

and it is similar to the variance of $\widetilde{T}$ in ISWR with finite population correction factors. Also in this case the estimator of the total and its variance correspond to the expressions given by Christman and Lan (2001).

To obtain an unbiased variance estimator, let us note that an unbiased estimator of μ_1^2 is

$$\hat{\mu}_{1q}=\overline{Y}_1^2-\left(1-\frac{\mathrm{k}}{\mathrm{N}_1}\right)\frac{\hat{S}_1^2}{\mathrm{k}},$$

while the corresponding unbiased estimator of μ_2^2 is

$$\hat{\mu}_{2q}=\overline{Y}_2^2-\left(1-\frac{\nu-\mathrm{k}}{\mathrm{N}_2}\right)\frac{\hat{S}_2^2}{\nu-\mathrm{k}}.$$

An unbiased estimator of P^2 is

$$\hat{P}_q = \frac{k-1}{\nu-1}\frac{k-2}{\nu-2} + \frac{1}{N}\frac{k-1}{\nu-1}\left(1-\frac{k-2}{\nu-2}\right),$$

so that we have

$$\hat{V}\left(\hat{P}\right) = \left(\frac{k-1}{\nu-1}\right)^2 - \hat{P}_q = \frac{\hat{P}\left(1-\hat{P}\right)}{\nu-2}\left(1-\frac{\nu-1}{N}\right).$$

Thus, it is easy to obtain the unbiased estimator (10) of the variance of $\tilde{T}$

$$\hat{V}\left(\tilde{T}\right) = N^2\left[\left(\bar{Y}_1-\bar{Y}_2\right)^2\hat{V}\left(\hat{P}\right) + \frac{\hat{S}_1^2}{k}\left(\hat{P}_q - \frac{k}{N}\hat{P}\right) + \frac{\hat{S}_2^2}{\nu-k}\left(\hat{P}_q - \frac{\nu-k}{N}\left(1-\hat{P}\right)\right)\right],$$

which is equivalent to the expression given by Salehi and Seber (2004).

The estimators of the two subgroup totals and of their variances in ISWOR may be easily obtained from the results given in paragraph 2. Finally, Salehi and Seber (2004) give the unbiased estimators of the population total and its variance for stopping rules similar to those considered in this paper.

References

Christman, M.C. and Lan, F. (2001). Inverse Adaptive Cluster Sampling. *Biometrics* **57**, 1096-1105.

Salehi, M. M. and Seber, G.A.F. (2001). A new proof of Murthy's estimator which applies to sequential sampling. *Australian and New Zealand Journal of Statistics* **43**, 281-286.

Salehi, M. M. and Seber, G.A.F. (2004). A General Inverse Sampling Scheme and its Application to Adaptive Cluster Samplig. *Australian and New Zealand Journal of Statistics* **46**, 483-494.

Appendix

It is worth pointing out the connection between the negative binomial and the binomial distributions, from which the probability that more than M replications are necessary to get *k* units from a given subgroup is equal to the probability that less than *k* units from the same subgroup are selected in M replications. Thus,

$$\sum_{n=\mathrm{M}+1}^{\infty} f(n;\mathrm{k},\mathrm{P}) = \sum_{x=0}^{\mathrm{k}-1} g(x;\mathrm{M},\mathrm{P}),$$

where $f(n;\mathrm{k},\mathrm{P})$ denotes the negative binomial distribution of parameters k and P, while $g(x;\mathrm{M},\mathrm{P})$ is the binomial distribution of parameters M and P, from which it turns out

$$\sum_{n=\mathrm{k}}^{\mathrm{M}} f(n;\mathrm{k},\mathrm{P}) + \sum_{x=0}^{\mathrm{k}-1} g(x;\mathrm{M},\mathrm{P}) = 1. \qquad \text{(A1)}$$

From the previous equality it is apparent that estimator (6) is unbiased. In fact

$$\begin{aligned}
\mathrm{E}(\widetilde{T}_s) &= \sum_{n=\mathrm{k}}^{\mathrm{M}} \mathrm{E}_{\mathrm{y}}(\widetilde{T}_a) f(n;\mathrm{k},\mathrm{P}) + \sum_{x=0}^{\mathrm{k}-1} \mathrm{E}_{\mathrm{y}}(\widetilde{T}_b) g(x;\mathrm{M},\mathrm{P}) = \\
&= \sum_{n=\mathrm{k}}^{\mathrm{M}} [\hat{p}_a(\mathrm{W}_1 - \mathrm{W}_2) + \mathrm{W}_2] f(n;\mathrm{k},\mathrm{P}) + \sum_{x=0}^{\mathrm{k}-1} [\hat{p}_b(\mathrm{W}_1 - \mathrm{W}_2) + \mathrm{W}_2] g(x;\mathrm{M},\mathrm{P}) = \\
&= \mathrm{P}(\mathrm{W}_1 - \mathrm{W}_2) + \mathrm{W}_2 = \mathrm{T}_1 + \mathrm{T}_2 ,
\end{aligned}$$

where $\hat{p}_a = \dfrac{\mathrm{k}-1}{n-1}$, and $\hat{p}_b = \dfrac{x}{\mathrm{M}}$, as results from (A1) and the following equality

$$\sum_{n=\mathrm{k}}^{\mathrm{M}} \hat{p}_a f(n;\mathrm{k},\mathrm{P}) + \sum_{x=0}^{\mathrm{k}-1} \hat{p}_b g(x;\mathrm{M},\mathrm{P}) = \mathrm{P}. \qquad \text{(A2)}$$

As regards the variance estimator (7), observe that

$$\mathrm{V}\left(\widetilde{T}_s\right)=\sum_{n=\mathrm{k}}^{\mathrm{M}}\mathrm{E}_{\mathrm{y}}\left[\left(\widetilde{T}_a-\mathrm{T}\right)^2\right]f\left(n;\mathrm{k},\mathrm{P}\right)+\sum_{x=0}^{\mathrm{k}-1}\mathrm{E}_{\mathrm{y}}\left[\left(\widetilde{T}_b-\mathrm{T}\right)^2\right]g\left(x;\mathrm{M},\mathrm{P}\right)=$$

$$=\sum_{n=\mathrm{k}}^{\mathrm{M}}\left[\left(\mathrm{W}_1-\mathrm{W}_2\right)^2\left(\hat{P}_a-\mathrm{P}\right)^2+\frac{\sigma_{1w}^2}{\mathrm{k}}\hat{P}_a^2+\frac{\sigma_{2w}^2}{n-\mathrm{k}}\left(1-\hat{P}_a\right)^2\right]f\left(n;\mathrm{k},\mathrm{P}\right)+$$

$$+\sum_{x=0}^{\mathrm{k}-1}\left[\left(\mathrm{W}_1-\mathrm{W}_2\right)^2\left(\hat{P}_b-\mathrm{P}\right)^2+\frac{\sigma_{1w}^2}{\mathrm{M}}\hat{P}_{\mathrm{b}}+\frac{\sigma_{2t}^2}{\mathrm{M}}\left(1-\hat{P}_b\right)\right]g\left(x;\mathrm{M},\mathrm{P}\right)$$

and

$$\mathrm{E}\left[\hat{\mathrm{V}}\left(\widetilde{T}_s\right)\right]=\sum_{n=\mathrm{k}}^{\mathrm{M}}\mathrm{E}_{\mathrm{y}}\left[\hat{\mathrm{V}}_a\left(\widetilde{T}_s\right)\right]f\left(n;\mathrm{k},\mathrm{P}\right)+\sum_{x=0}^{\mathrm{k}-1}\mathrm{E}_{\mathrm{y}}\left[\hat{\mathrm{V}}_b\left(\widetilde{T}_b\right)\right]g\left(x;\mathrm{M},\mathrm{P}\right)=$$

$$=\sum_{n=\mathrm{k}}^{\mathrm{M}}\left(\mathrm{W}_1-\mathrm{W}_2\right)^2\left(\hat{P}_a^2-\hat{P}_{aq}\right)+\frac{\sigma_{1w}^2}{\mathrm{k}}\hat{P}_a^2+\frac{\sigma_{2w}^2}{n-\mathrm{k}}\left(1-\hat{P}_a\right)^2 f\left(n;\mathrm{k},\mathrm{P}\right)+$$

$$+\sum_{x=0}^{\mathrm{k}-1}\left[\left(\mathrm{W}_1-\mathrm{W}_2\right)^2\left(\hat{P}_b^2-\hat{P}_{bq}\right)+\frac{\sigma_{1w}^2}{\mathrm{M}}\hat{P}_{\mathrm{b}}+\frac{\sigma_{2t}^2}{\mathrm{M}}\left(1-\hat{P}_b\right)\right]g\left(x;\mathrm{M},\mathrm{P}\right).$$

From (A1), (A2) and

$$\sum_{n=\mathrm{k}}^{\mathrm{M}}\hat{p}_{aq}f\left(n;\mathrm{k},\mathrm{P}\right)+\sum_{x=0}^{\mathrm{k}-1}\hat{p}_{bq}g\left(x;\mathrm{M},\mathrm{P}\right)=\mathrm{P}^2,\qquad\text{(A3)}$$

where $\hat{p}_{aq}=\frac{\mathrm{k}-1}{n-1}\frac{\mathrm{k}-2}{n-2}$ and $\hat{p}_{bq}=\frac{x}{\mathrm{M}}\frac{x-1}{\mathrm{M}-1}$, it turns out that $\mathrm{E}\left[\hat{\mathrm{V}}\left(\widetilde{T}_s\right)\right]=\mathrm{V}\left(\widetilde{T}_s\right)$.

When at least m and at most M replications are performed, the equality analogous to (A1) is

$$\sum_{x=\mathrm{k}}^{\mathrm{m}}g\left(x;\mathrm{m},\mathrm{P}\right)+\sum_{n=\mathrm{m}+1}^{\mathrm{M}}f\left(n;\mathrm{k},\mathrm{P}\right)+\sum_{x=0}^{\mathrm{k}-1}g\left(x;\mathrm{M},\mathrm{P}\right)=1.\qquad\text{(A4)}$$

From (A4), equalities analogous to (A2) and (A3) may be easily obtained and these equalities are sufficient to show that the totals estimators are unbiased, as well as the estimators of their variances.