© Statistical Research and Training Center
مجلّه‌ی پژوهش‌های آماری ایران

J. Statist. Res. Iran **11** (2014): 191–202
دوره‌ی ۱۱، شماره‌ی ۲، پاییز و زمستان ۱۳۹۳، صص ۱۹۱–۲۰۲

On the Simple Inverse Sampling with Replacement

Mohammad Mohammadi

University of Isfahan

Abstract. In this paper we derive some unbiased estimators of the population mean under simple inverse sampling with replacement, using the class of Hansen-Hurwitz and Horvitz-Thompson type estimators and the post-stratification approach. We also compare the efficiency of resulting estimators together with Murthy's estimator. We show that in despite of general belief, the strategy consisting of inverse sampling with Murthy's estimator is highly less efficient when the target population is rare, whereas it can be more efficient when subpopulation means are closed. In fact, for inverse sampling to be highly efficient design one should know the population structure and then use an appropriate estimator.

Keywords. Finite population; Hansen-Hurwitz estimator; Horvitz-Thompson estimator; Murthy's estimator; post-stratification.

MSC 2010: 62D05.

1 Introduction

Simple inverse sampling with replacement (SISWR) introduced firstly by Haldan (1945) to estimate the population proportion. Christman and Lan (2001) provided an unbiased estimator of the population mean under SISWR. Similar result is obtained by Salehi and Seber (2001) by using the Murthy's method (Murthy, 1957). In the context of finite sampling it is known that the inverse sampling is more efficient design for rare populations, where the target population partitioned into two subgroups for which one is small with large y-values and other is large with small (near to zero) y-values.

In this paper, in Section 2 we derive some new unbiased estimators of the population mean under SISWR using the class of Hansen-Hurwitz and Horvitz-Thompson type estimators and the post-stratification idea. Typically, these involved in the size of population subgroups, so should be known. In section 3, we give a theoretical comparison of resulting estimators together with Murthy's estimator. As a main result, the Murthy's estimator is highly less efficient estimator in the rare populations and more efficient estimator when the mean of population subgroups be very closed. In other word, for inverse sampling to be highly efficient design for rare populations one should know the subpopulation sizes and then use an appropriate estimator.

2 Inverse Sampling with Replacement

Suppose that a finite population $U = \{u_1, u_2, \ldots, u_N\}$ of N units is divided into two groups U_C and $U_{\bar{C}}$, with the corresponding sizes M and $N - M$, respectively. With any unit u_k there is an associated value of the variable of interest y_k, for $k = 1, 2, \ldots, N$. Let $\bar{y}_{U_C} = \frac{1}{M}\sum_{U_C} y_k$, $\bar{y}_{U_{\bar{C}}} = \frac{1}{N-M}\sum_{U_{\bar{C}}} y_k$, $\sigma^2_C = \frac{1}{M}\sum_{U_C}(y_k - \bar{y}_{U_C})^2$, and $\sigma^2_{\bar{C}} = \frac{1}{N-M}\sum_{U_{\bar{C}}}(y_k - \bar{y}_{U_{\bar{C}}})^2$. In SISWR the units are selected sequentially, with replacement and with equal probabilities until predetermined number of units, say r, possessing the attribute drown. Let os, os_C and $os_{\bar{C}}$ denote the final samples from U, U_C and $U_{\bar{C}}$, respectively. Obviously, the final sample size, say n_{os}, is a random variable which follows the negative binomial distribution. Salehi and Seber (2001) applied the Murthy's estimator under SISWR and found an unbiased estimator of the population mean as:

$$\hat{\mu}_M = \frac{\hat{M}}{N}\bar{y}_{os_C} + \left(1 - \frac{\hat{M}}{N}\right)\bar{y}_{os_{\bar{C}}}, \tag{1}$$

where $\bar{y}_{os_C}$ and $\bar{y}_{os_{\bar{C}}}$ are sample means from os_C and $os_{\bar{C}}$, and $\hat{M} = \frac{(r-1)N}{n_{os}-1}$ is an unbiased estimator of the subpopulation size M, whether it is known or not. Christman and Lan (2001) showed that the variance of $\hat{\mu}_M$ in (1) is given by:

$$\mathrm{Var}(\hat{\mu}_M) = E\left(\frac{\hat{M}}{N}\right)^2\frac{\sigma^2_C}{r} + E\left\{\frac{(1-\frac{\hat{M}}{N})^2}{n_{os}-r}\right\}\sigma^2_{\bar{C}} + \mathrm{Var}\left(\frac{\hat{M}}{N}\right)(\bar{y}_{U_C} - \bar{y}_{U_{\bar{C}}})^2. \tag{2}$$

2.1 Unbiased Estimators with Known Subpopulation Sizes

In this subsection we derive the Hansen-Hurwitz and Horvitz-Thompson type estimators for the population mean under SISWR. Furthermore, we give an unbiased version of post-stratified estimator adjusted for probability of no selection from $U_{\bar{C}}$.

The Hansen-Hurwitz Type Estimator

Hansen and Hurwitz (1943) suggested a class of unbiased estimators of the population mean which applied usually for with replacement sampling designs. Let $f(k)$ denote the number of which the kth unit selected in the final sample, for $k = 1, \ldots, N$. So, the Hansen-Hurwitz (HH) type estimator is defined as:

$$\hat{\mu}_{HH} = \frac{1}{N} \sum_{k=1}^{N} \frac{y_k}{E\{f(k)\}} f(k). \tag{3}$$

To derive $\hat{\mu}_{HH}$ and its variance under SISWR we need to have $E\{f(k)\}$, Var$\{f(k)\}$ and Cov$\{f(k), f(l)\}$ for any $(k \neq l)$. It is evident that a SISWR given its size n_{os} is equivalent to a stratified random sampling with replacement with size $(r, n_{os} - r)$ from $(U_C, U_{\bar{C}})$. Hence, for any $\{k_1, \ldots, k_\nu\} \subseteq U_C$ the random vector $(f(k_1), \ldots, f(k_\nu))$ follows multinomial (Mn) distribution with parameters $(r, \frac{1}{M}, \ldots, \frac{1}{M})$. On the other hand, it can be shown that for any $\{k_1, \ldots, k_\omega\} \subseteq U_{\bar{C}}$, the joint distribution of $f(k_1), \ldots, f(k_\omega)$ is a negative multinomial (NMn) with density function

$$P_r\left\{f(k_1) = x_1, \ldots, f(k_\omega) = x_\omega\right\} = \frac{\Gamma(r+t)}{\Gamma(r)\prod_{i=1}^{\omega} x_i!} \left(\frac{M}{M+\omega}\right)^r \left(\frac{1}{M+\omega}\right)^t, \tag{4}$$

where $x_i \geqslant 0$, and $t = \sum_{i=1}^{\omega} x_i$ (see Appendix A). Therefore, $f(k)$ is a negative binomial random variable with parameters $(r, \frac{M}{M+1})$. Hence, for any $k \in U$:

$$E\{f(k)\} = \frac{r}{M}. \tag{5}$$

Also, we have

$$\text{Var}\{f(k)\} = \begin{cases} \frac{r(M-1)}{M^2} & \text{if} \quad k \in U_C, \\ \frac{r(M+1)}{M^2} & \text{if} \quad k \in U_{\bar{C}}, \end{cases} \tag{6}$$

and for any $k \neq l$

$$\mathrm{Cov}\{f(k), f(l)\} = \begin{cases} -\frac{r}{M^2} & \text{if} \quad (k\&l) \in U_C, \\ 0 & \text{if} \quad k \in U_C \& l \in U_{\bar{C}}, \\ \frac{r}{M^2} & \text{if} \quad (k\&l) \in U_{\bar{C}}. \end{cases} \tag{7}$$

Imputing equation (5) in (3) results the HH type estimator of $\bar{y}_U$ as:

$$\hat{\mu}_{HH} = \frac{M}{Nr} \sum_{os} y_k. \tag{8}$$

Using equations (6) and (7), it can be shown that (see Appendix B) the variance of HH estimator in (8) is given by:

$$\mathrm{Var}(\hat{\mu}_{HH}) = \frac{1}{N^2 r} \left\{ M^2 \sigma_C^2 + M(N-M)\sigma_{\bar{C}}^2 + N(N-M)\bar{y}_{U_{\bar{C}}}^2 \right\}. \tag{9}$$

The Horvitz-Thompson Type Estimator

Horvitz and Thompson (1952) defined an alternative class of unbiased estimators for the population mean which is constructed based on inclusion probabilities. Let s denote the sample set of selected distinct units and $\pi_k = P_r(k \in s)$ be inclusion probability for kth element. Hence, the Horvitz-Thompson (HT) type estimator for the population mean is:

$$\hat{\mu}_{HT} = \frac{1}{N} \sum_{k \in s} \frac{y_k}{\pi_k}. \tag{10}$$

The variance of $\hat{\mu}_{HT}$ is

$$\mathrm{Var}(\hat{\mu}_{HT}) = \frac{1}{N^2} \sum_{k=1}^{N} \sum_{l=1}^{N} (\pi_{kl} - \pi_k \pi_l) \frac{y_k y_l}{\pi_k \pi_l}, \tag{11}$$

where π_{kl} is the joint inclusion probability of the kth and the lth units, with $\pi_{kk} = \pi_k$. Using the relation between $f(k)$ and π_k, we have:

$$\pi_k = 1 - P_r\{f(k) = 0\} = \begin{cases} 1 - \varphi_{M,1} & \text{if} \quad k \in U_C, \\ 1 - \varphi_{(M+1),1} & \text{if} \quad k \in U_{\bar{C}}, \end{cases} \tag{12}$$

and

$$\pi_{kl}-\pi_k\pi_l = P_r\{f(k)=f(l)=0\} - P_r\{f(k)=0\}P_r\{f(l)=0\}$$
$$= \begin{cases} \varphi_{M,2}-\varphi^2_{M,1} & \text{if} \quad (k,l)\in U_C \\ 0 & \text{if} \quad k\in U_C, l\in U_{\bar{C}} \\ \varphi_{(M+2),2}-\varphi^2_{(M+1),1} & \text{if} \quad (k,l)\in U_{\bar{C}}, \end{cases} \tag{13}$$

where $\varphi_{s,t}=(1-t/s)^r$. Imputing (12) and (13) in (10) and (11), respectively, the HT estimator of $\bar{y}_U$ is found as:

$$\hat{\mu}_{HT}=\frac{1}{N}\left\{\frac{1}{1-\varphi_{M,1}}\sum_{s_C} y_k + \frac{1}{1-\varphi_{(M+1),1}}\sum_{s_{\bar{C}}} y_k\right\}$$

with variance

$$\mathrm{Var}(\hat{\mu}_{HT}) = \frac{1}{N^2}\left[\frac{(\varphi_{M,1}-\varphi_{M,2})\sum_{U_C} y_k^2-(\varphi^2_{M,1}-\varphi_{M,2})T_C^2}{(1-\varphi_{M,1})^2} + \frac{\left\{\varphi_{(M+1),1}-\varphi_{(M+2),2}\right\}\sum_{U_{\bar{C}}} y_k^2-\left\{\varphi^2_{(M+1),1}-\varphi_{(M+2),2}\right\}T_{\bar{C}}^2}{\{1-\varphi_{(M+1),1}\}^2}\right], \tag{14}$$

where $T_C=\sum_{U_C} y_k$ and $T_{\bar{C}}=\sum_{U_{\bar{C}}} y_k$.

The post-stratified Estimator

As another well-known technique, we may use the post-stratification strategy. Since $\bar{y}_U=\frac{M}{N}\bar{y}_{U_C}+(1-\frac{M}{N})\bar{y}_{U_{\bar{C}}}$, an unbiased estimator of the population mean can be achieved by constructing unbiased estimators of the sub-population means $\bar{y}_{U_C}$ and $\bar{y}_{U_{\bar{C}}}$. It can be easily shown that $\bar{y}_{os_C}$ and $\frac{1}{1-(\frac{M}{N})^r}\tilde{y}_{os_{\bar{C}}}$ are unbiased estimators of $\bar{y}_{U_C}$ and $\bar{y}_{U_{\bar{C}}}$, respectively, where $\tilde{y}_{os_{\bar{C}}}=\frac{1}{n_{os}-r}\sum_{os_{\bar{C}}} y_k$ if $n_{os}>r$, and 0 if $n_{os}=r$. Hence, the post-stratified estimator of the population mean is defined as:

$$\hat{\mu}_{p.st}=\frac{M}{N}\bar{y}_{os_C}+\left(1-\frac{M}{N}\right)\frac{1}{1-(\frac{M}{N})^r}\tilde{y}_{os_{\bar{C}}}$$

The variance of $\hat{\mu}_{p.st}$ is

$$\mathrm{Var}(\hat{\mu}_{p.st})=\left(\frac{M}{N}\right)^2\frac{\sigma_C^2}{r}+\delta_{pst}\sigma^2_{\bar{C}}+(1-\frac{M}{N})^2\frac{(\frac{M}{N})^r}{1-(\frac{M}{N})^r}\bar{y}^2_{U_{\bar{C}}}, \tag{15}$$

where

$$\delta_{pst} = \frac{1}{\{1-(\frac{M}{N})^r\}^2} \sum_{x=r+1}^{\infty} \binom{x-1}{r-1} \frac{1}{x-r} \left(\frac{M}{N}\right)^r \left(1-\frac{M}{N}\right)^{x-r+2}. \quad (16)$$

3 Comparison of Efficiencies

The variance of derived estimators as well as the variance of Murthy's estimator in (2) are typically as functions of the subpopulation means and variances. The variance of $\hat{\mu}_{HT}$ given in (14) is more complicated to compare it with others. However, our simulation on the numerous populations show that $\hat{\mu}_{HT}$ is dominated by $\hat{\mu}_{HH}$ and/or $\hat{\mu}_{p.st}$, so we remove it from our comparison. In continue we give a theoretical comparison between $\hat{\mu}_M$, $\hat{\mu}_{HH}$ and $\hat{\mu}_{p.st}$ with respect to subpopulation variances and means.

Subgroup Variance σ^2_C
As the multiplier of σ^2_C, the first part of the variance of both $\hat{\mu}_{HH}$ and $\hat{\mu}_{p.st}$ are similar, and is smaller than those for $\hat{\mu}_M$, since $E(\frac{\hat{M}}{N})^2 > (\frac{M}{N})^2$. This means that as the subpopulation variance σ^2_C increased, the efficiencies of $\hat{\mu}_{HH}$ and $\hat{\mu}_{p.st}$ relative to $\hat{\mu}_M$ are increased.

Subgroup Variance $\sigma^2_{\bar{C}}$
The second factor of the variances of $\hat{\mu}_M$, $\hat{\mu}_{HH}$ and $\hat{\mu}_{p.st}$ are multipliers of $\sigma^2_{\bar{C}}$, say $\delta_M = E\left\{\frac{(1-\frac{\hat{M}}{N})^2}{n_{os}-r}\right\}$, $\delta_{HH} = \frac{M(N-M)}{N^2 r}$ and δ_{pst} which defined in (16), respectively. We give the values of these quantities in Table 1 for $\frac{M}{N} = 0.05, 0.1, 0.2$. As it seen, the δ_{HH} is the smallest for any situation, in favor of $\hat{\mu}_{HH}$. Using the inequality due to Prasad (1982), it can be shown that (see Appendix C) $\delta_M > \delta_{HH}$. On the other hand, in any case we have $\delta_{pst} > \delta_{HH}$, however, as r increases the coefficients δ_M, δ_{HH} and δ_{pst} are closed to each other.

Subgroup Means
The third factor in the variance of mentioned estimators is in the basis of subpopulation means. A serious drawback of $\hat{\mu}_{HH}$ and $\hat{\mu}_{p.st}$ is that these have'nt location scale property with resect to the y-values in $U_{\bar{C}}$, so those variances may be affected by the variation of mean in $U_{\bar{C}}$, with minor effect for $\hat{\mu}_{p.st}$. Define α as the difference of $\bar{y}^2_{U_{\bar{C}}}$ multipliers in equations (9) and

(15), i.e $\alpha = \frac{N-M}{Nr} - (1-\frac{M}{N})^2 \frac{(\frac{M}{N})^r}{1-(\frac{M}{N})^r}$. It can be shown easily that

$$\alpha = \left(1 - \frac{M}{N}\right) \frac{\sum_{i=0}^{r-1} \{(\frac{M}{N})^i - (\frac{M}{N})^r\}}{r \sum_{i=0}^{r-1} (\frac{M}{N})^i},$$

so always $\alpha > 0$ (in favor of $\hat{\mu}_{p.st}$). However, as is shown in Figure 1, the difference value α will be negligible, as r increased. So, $\hat{\mu}_{HH}$ is suggested only if the mean of y-values be near to zero or r is chosen very large.
On the other hand, for the variables with near to zero values in $U_{\bar{C}}$ and large values in U_C, the $\hat{\mu}_M$ has highly less efficiency in comparison with $\hat{\mu}_{HH}$ and $\hat{\mu}_{p.st}$. This means that for inverse sampling to be an efficient design for the rare populations we need to know subpopulation size M and use $\hat{\mu}_{HH}$ or $\hat{\mu}_{p.st}$, appropriately.

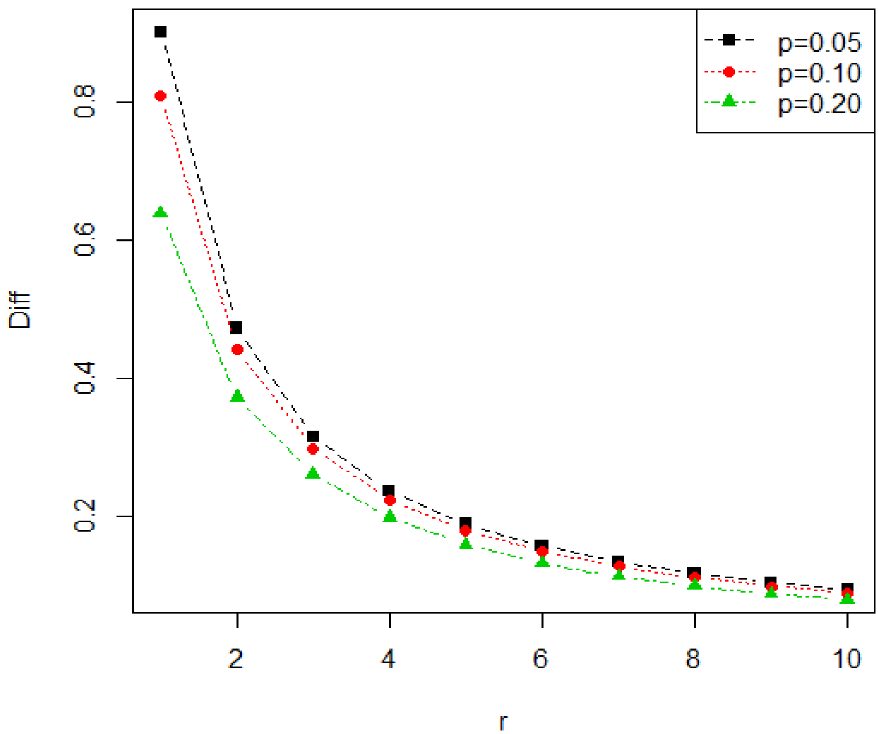

Figure 1. The difference values α against some values of r.

It is surprising that for un-rare populations with near and large values of $\bar{y}_{U_C}$ and $\bar{y}_{U_{\bar{C}}}$, the Murthy's estimator may be even more efficient than $\hat{\mu}_{p.st}$, special for small r-values. In this case, it is dangerous to use $\hat{\mu}_{HH}$, so we compare only $\hat{\mu}_M$ and $\hat{\mu}_{p.st}$. To see the behavior of $\hat{\mu}_M$ and $\hat{\mu}_{p.st}$ with respect to the $\bar{y}_{U_{\bar{C}}}$, we give a small simulation. Define following model for the population:

$$F_C = N(100, 10); \qquad F_{\bar{C}} = N(\bar{y}_{U_{\bar{C}}}, 10),$$

where $F(\cdot)$ denotes the underlying distribution for generating population y-values. The population size is $N = 100$ with $M = 10$, and the r is chosen as $r = 3$. The variance of $\hat{\mu}_M$ and $\hat{\mu}_{p.st}$ are plotted versus of $\bar{y}_{U_{\bar{C}}}$ in the Figure 2. When $\bar{y}_{U_{\bar{C}}}$ is small, $\hat{\mu}_{p.st}$ is more efficient. However, as $\bar{y}_{U_{\bar{C}}}$ increases from 70 to 100, the $\hat{\mu}_M$ is more efficient than $\hat{\mu}_{p.st}$.

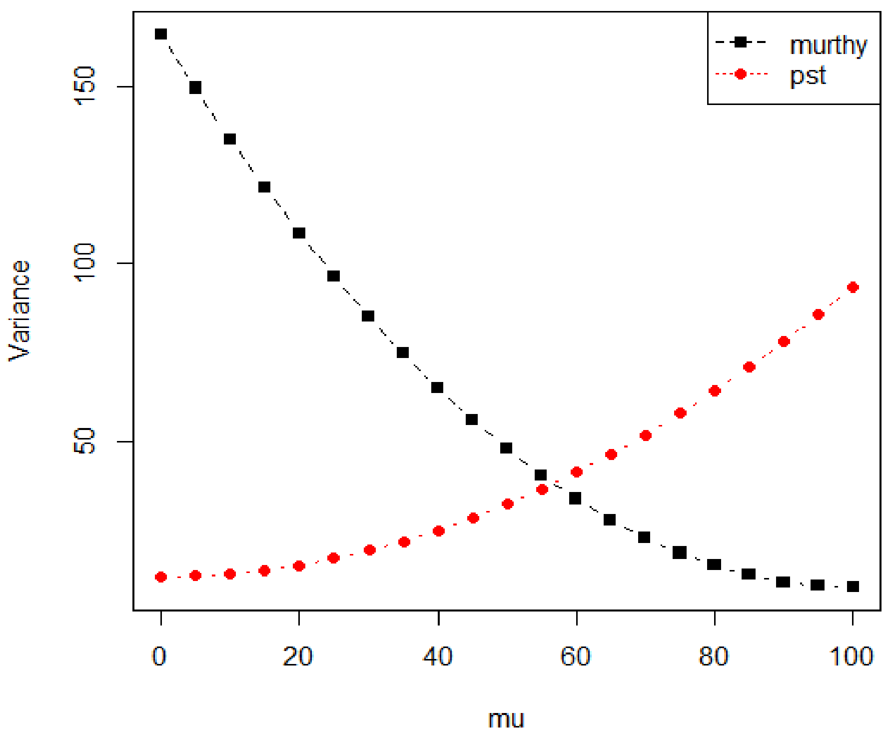

Figure 2. The variance of $\hat{mu}_M$ and $\hat{\mu}_{p.st}$ versus the values of subpopulation mean $\bar{y}_{U_{\bar{C}}}$.

Table 1. The values of variance coefficients, δ_M, δ_{HH}, and δ_{pst} for the σ_C^2 for the unbiased estimators under SISWR design.

$\frac{M}{N}$	r	δ_M	δ_{HH}	δ_{pst}
0.05	2	0.0422	0.0237	0.0499
	3	0.0228	0.0158	0.0250
	4	0.0154	0.0119	0.0163
	5	0.0117	0.0095	0.0121
	6	0.0094	0.0079	0.0096
	8	0.0067	0.0059	0.0068
	10	0.0052	0.0047	0.0053
	12	0.0043	0.0040	0.0043
0.10	2	0.0745	0.0450	0.0934
	3	0.0417	0.0300	0.0493
	5	0.0218	0.0180	0.0234
	8	0.0127	0.0112	0.0131
	10	0.0099	0.0090	0.0101
	12	0.0081	0.0075	0.0083
	15	0.0064	0.0060	0.0065
	18	0.0053	0.0050	0.0053
0.20	2	0.1196	0.0800	0.1558
	3	0.0701	0.0533	0.0916
	6	0.0306	0.0267	0.0344
	10	0.0173	0.0160	0.0184
	15	0.0113	0.0107	0.0117
	20	0.0083	0.0080	0.0085
	25	0.0066	0.0064	0.0067
	30	0.0055	0.0053	0.0056

References

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

Haldane, J.B.S. (1945). On a Method of Estimating Frequencies. *Biometrika*, **33**, 222-225.

Hansen, M.H. and Hurwitz, W.N. (1943). On the Theory of Sampling from Finite Populations. *The Annals of Mathematical Statistics*, **14**, 333-362.

Horvitz, D.G. and Thompson, D.J. (1952). A Generalization of Sampling Without Replacement from a Finite Universe. *The Journal of the American Statistical Association*, **47**, 663-685.

Murthy, M.N. (1957). Ordered and Unordered Estimators in Sampling Without Replacement. *Sankhyã* **18**, 379-390.

Prasad, G. (1982). Sharper Variance Upper Bound for Unbiased Estimation in Inverse Sampling. *Trabajos De Estadistica Y De Investigation Operative*, **33**, 130-132.

Salehi, M.M. and Seber, A.F. (2001). A New Proof of Murthy's Estimator which Applies to Sequential Sampling. *Aust. N. Z. J. Stat*, **44**, 63-74.

Appendix A

Joint distribution of observed frequencies in $U_{\bar{C}}$ under SISWR

Consider the subset $\{k_1, \ldots, k_\omega\} \subseteq U_{\bar{C}}$. Since $f(k_1), \ldots, f(k_\omega)$ given the sample size n_{os} follows multinomial distribution with parameters $(n_{os} - r, \frac{1}{N-M}, \ldots, \frac{1}{N-M})$, we have:

$$
\begin{aligned}
&P_r\{f(k_1) = x_1, \ldots, f(k_\omega) = x_\omega\} \\
&= \sum_{n_{os}=r+t}^{\infty} P_r\{f(k_1) = x_1, \ldots, f(k_\omega) = x_\omega \big| n_{os}\} P_r(n_{os}) \\
&= \sum_{n_{os}=r+t}^{\infty} \left\{ \binom{n_{os} - r}{x_1, \ldots, x_\omega} \left(\frac{1}{N-M}\right)^t \left(1 - \frac{\omega}{N-M}\right)^{n_{os}-r-t} \binom{n_{os}-1}{r-1} \right. \\
&\quad \left. \times \left(\frac{M}{N}\right)^r \left(1 - \frac{M}{N}\right)^{n_{os}-r} \right\} \\
&= \frac{\Gamma(r+t)}{\Gamma(r) \prod_{i=1}^{\omega} x_i!} \left(\frac{1}{N-M}\right)^t \left(\frac{M}{N}\right)^r \left(1 - \frac{M}{N}\right)^t \\
&\quad \times \sum_{n_{os}=r+t}^{\infty} \binom{n_{os}-1}{r+t-1} \left(1 - \frac{M+\omega}{N}\right)^{n_{os}-r-t} \\
&= \frac{\Gamma(r+t)}{\Gamma(r) \prod_{i=1}^{\omega} x_i!} \left(\frac{M}{M+\omega}\right)^r \left(\frac{1}{M+\omega}\right)^t, \qquad \left(t = \sum_{i=1}^{\omega} x_i,\ x_i \geqslant 0\right)
\end{aligned}
$$

which is a negative multinomial (NMn) distribution with parameters (r, M).

Appendix B

Deriving the variance of $\hat{\mu}_{HH}$
Using the probability function (4), we have $\mathrm{Var}(f(k)) = \frac{r(M+1)}{M^2}$ and $\mathrm{Cov}\{f(k), f(l)\} = \frac{r}{M^2}$ (for $k \neq l \in U_{\bar{C}}$). On the other hand, we may rewritten $\hat{\mu}_{HH}$ as $\frac{M}{Nr}\sum_{k=1}^{N} f(k) y_k$. Hence,

$$\begin{aligned}
\mathrm{Var}(\hat{\mu}_{HH}) &= \left(\frac{M}{Nr}\right)^2 \left[\sum_{k=1}^{N} y_k^2 \mathrm{Var}\{f(k)\} + \sum_{k\neq l}\sum y_k y_l \mathrm{Cov}\{f(k), f(l)\}\right] \\
&= \left(\frac{M}{Nr}\right)^2 \Bigg\{ \sum_{k\in U_C} \frac{r(M-1)}{M^2} y_k^2 + \sum_{k\in U_{\bar{C}}} \frac{r(M+1)}{M^2} y_k^2 \\
&\quad + \sum_{(k\neq l)}\sum_{\in U_C} \left(\frac{-r}{M^2}\right) y_k y_l + \sum_{(k\neq l)}\sum_{\in U_{\bar{C}}} \left(\frac{r}{M^2}\right) y_k y_l \Bigg\} \\
&= \frac{1}{N^2 r}\Bigg\{ (M-1)\sum_{k\in U_C} y_k^2 - \sum_{(k\neq l)}\sum_{\in U_C} y_k y_l + (M+1)\sum_{k\in U_{\bar{C}}} y_k^2 \\
&\quad + \sum_{(k\neq l)}\sum_{\in U_{\bar{C}}} y_k y_l \Bigg\} \\
&= \frac{1}{N^2 r}\Bigg\{ M\sum_{k\in U_C} y_k^2 - M^2 \bar{y}_{U_C}^2 + M\sum_{k\in U_{\bar{C}}} y_k^2 + (N-M)^2 \bar{y}_{U_{\bar{C}}}^2 \Bigg\}.
\end{aligned}$$

By some simplifications we get the variance of $\hat{\mu}_{HH}$ as:

$$\mathrm{Var}(\hat{\mu}_{HH}) = \frac{1}{N^2 r}\left\{ M^2 \sigma_C^2 + M(N-M)\sigma_{\bar{C}}^2 + N(N-M)\bar{y}_{U_{\bar{C}}}^2 \right\}.$$

Appendix C

Proof of $\delta_M > \delta_{HH}$

Prasad (1982) found a sharper upper bound for the variance of $\frac{r-1}{n_{os}-1}$ (as the unbiased estimator of the population proportion $\frac{M}{N}$) under SISWR as

$$\text{Var}\left(\frac{r-1}{n_{os}-1}\right) \leqslant \frac{(r-2)(\frac{M}{N})^2(1-\frac{M}{N})}{(r-1)^2\frac{M}{N}+(r-2)^2(1-\frac{M}{N})}$$

On the other hand we have $\delta_M = E\left(\frac{(1-\frac{\hat{M}}{N})^2}{n_{os}-r}\right) = \frac{M}{N(r-1)} - (r-1)E(\frac{1}{n_{os}-1})^2$. Using the above inequality we have:

$$\delta_M \geqslant \frac{M}{N(r-1)} - \frac{(\frac{M}{N})^2 + \frac{(r-2)(\frac{M}{N})^2(1-\frac{M}{N})}{(r-1)^2\frac{M}{N}+(r-2)^2(1-\frac{M}{N})}}{r-1}$$

Since $\delta_{HH} = \frac{M(N-M)}{rN^2}$, by some simplifications we have:

$$\delta_M - \delta_{HH} \geqslant \frac{M(N-M)}{N^2}\left[\frac{(r-2)^2-(r-1)(r-3)\frac{M}{N}}{r(r-1)\left\{(r-1)^2\frac{M}{N}+(r-2)^2(1-\frac{M}{N})\right\}}\right]$$

It is evident that the numerator of biggest bracket in above inequality is positive, so $\delta_M > \delta_{HH}$.

Mohammad Mohammadi
Department of Statistics,
University of Isfahan,
Isfahan, Iran.
email: *m.mohammadi@sci.ui.ac.ir*