A BOOTSTRAP APPROACH TO PROBABILITY PROPORTIONAL-TO-SIZE SAMPLING

Anders Holmberg, University of Örebro
Dept. of Statistics, University of Örebro, SE-701 82 Örebro, Sweden

KEY WORDS: Bootstrap, Finite Population, Pareto Sampling, Unequal Probability Sampling, Variance Estimation, Gini coefficient

1. Introduction and background

In sampling from finite populations, the original "simple" bootstrap proposed by Efron (1979) does not capture the dependence imposed by without replacement sampling. A number of finite population bootstrap approaches have been proposed. One early approach, BWO (short for Bootstrap WithOut replacement), by Gross (1980), has been further developed and discussed by Bickel & Freedman (1981, 1984), Chao & Lo (1985, 1994), Booth, Butler & Hall (1994), Sitter (1992a, 1992b), and Rao & Katzoff (1996). Among other suggestions we find the Mirror Match method (MM), Sitter (1992a), the Bootstrap With Replacement (BWR), McCarthy & Snowden (1985), and the Rescaling Bootstrap (RB), Rao & Wu (1984, 1988) and Rao, Wu & Yue (1992).

The properties of these bootstrap methods when estimating the variance of a point estimator have been studied for various combinations of sampling design and point estimator. However, few studies exist on bootstrap methods for πps designs, i.e., without replacement probability proportional-to-size designs. Kuk (1989) proposed a bootstrap method for systematic πps sampling, while Rao & Wu (1984, 1988), Rao et al. (1992), and Sitter (1992a, 1992b) studied extensions of the BWO, MM and RB methods to handle the Rao-Hartley-Cochran (RHC) approach for unequal probability sampling. Chao & Lo (1994), finally, give one approach for a design referred to as Murthy's method in Cochran (1977, ch. 9A9).

(In the sequel, we will refer to Gross' approach as the BWO, while BWO$_x$ refers to a modified approach, where the x indicates the relevant authors.)

It has until recently been difficult to find a fixed size πps sampling scheme that has all of the desirable properties mentioned in Särndal, Swensson & Wretman (1992, Section 3.6.2). The Pareto sampling scheme proposed by Rosén (1996, 1997a-b), however, has such properties that most of these difficulties are overcome. Pareto sampling makes it very easy to generate a πps sample, and in a general class of πps sampling schemes Pareto sampling has optimal properties in the sense of producing the smallest asymptotic variance for an estimator of a population total.

In this paper we propose a bootstrap approach for πps samples. In particular, we consider its application to Pareto πps sampling and to the πps scheme proposed by Sunter as desribed in Särndal et al. (1992). A bootstrap variance estimator is suggested, and its properties are studied by Monte Carlo simulations. In these simulations, we consider point estimators of the population mean, the Gini mean difference and the Gini coefficient.

1.1 Notations and definitions

Let $U = \{1, \ldots, k, \ldots, N\}$ denote a finite population of size N. For $k = 1, \ldots, N$, let y_k denote the (unknown) values of the study variable, and let x_k (> 0) denote the (known) size measures. Let s denote a without replacement (fixed size) sample of n elements drawn from U. Furthermore, let π_k and π_{kl} denote first-order and second-order inclusion probabilities.

In πps sampling, we have

$$\pi_k = \frac{n x_k}{t_{xU}}, \tag{1}$$

where $t_{xU} = \sum_U x_k = x_1 + x_2 + \cdots + x_N$. (In the sequel, we assume that $\pi_k \leq 1$ for every $k \in U$.)

The π estimator, $\hat{t}_\pi$, of the population total t_U is given by,[*] $\hat{t}_\pi = \sum_s \frac{y_k}{\pi_k}$ and has the variance, $V(\hat{t}_\pi) = -\frac{1}{2}\sum\sum_{k \neq l \in U} \Delta_{kl} \left(\frac{y_k}{\pi_k} - \frac{y_l}{\pi_l}\right)^2$, where $\Delta_{kl} = (\pi_{kl} - \pi_k \pi_l)$. An unbiased estimator of $V(\hat{t}_\pi)$, provided that every $\pi_{kl} > 0$ ($k \neq l$), is given by the Sen-Yates-Grundy (SYG) variance estimator,

$$\hat{V}(\hat{t}_\pi) = -\frac{1}{2}\sum\sum_{k \neq l \in s} \check{\Delta}_{kl} \left(\frac{y_k}{\pi_k} - \frac{y_l}{\pi_l}\right)^2, \tag{2}$$

where $\check{\Delta}_{kl} = \Delta_{kl}/\pi_{kl}$. Furthermore, let the distribution function of y for the population elements, be defined as $F_U(y) = \#A_y/N$, where $\#$ denotes the number of elements k in the set $A_y = \{k : k \in U,\ and\ y_k \leq y\}$. An estimator for $F_U(y)$ is

[*] Troughout the paper all estimators and parameters lacking the subscript x refers to the y variable.

given by $\hat{F}_s(y) = \sum_{s \cap A_y} \pi_k^{-1} / \sum_s \pi_k^{-1}$, where $s \cap A_y$ is the set of sample elements with values $y_k \leq y$.

2. BWO approaches

In this section we will only consider simple random sampling without replacement (SI).

2.1 The BWO bootstrap

Consider a SI sample $s \subset U$ of size n, with values, y_k, attached to each $k \in s$. An estimator, $\hat{\theta} = \hat{\theta}(s) = \hat{\theta}[(k, y_k) : k \in s]$ is to be used for estimating a finite population parameter, θ, e.g., the population total $t_U = \sum_U y_k$. Let the inverse of the sampling fraction $f^{-1} = \frac{N}{n} = c$ be an integer. The BWO suggested by Gross can be described as follows.

1. Create an artificial resampling population U^* consisting of c copies of each element $k \in s$, i.e., $U^* = \{1^*, \ldots, k^*, \ldots, N^*\}$, where $N^* = nc = N$. All c elements that are copies of element $k \in s$ are assigned the y value y_k. (Hence, $F_{U^*}(y) = \hat{F}_s(y)$, and, e.g., $t_{U^*} = \sum_{U^*} y_{k^*} = c \sum_s y_k = \hat{t}_\pi$.)

2. Draw a SI sample s_1^* of size $n^* = n$ from U^*. We will refer to s_1^* as a bootstrap sample.

3. Compute a bootstrap replicate $\hat{\theta}_1^* = \hat{\theta}(s_1^*) = \hat{\theta}\left[(k^*, y_{k^*}) : k^* \in s_1^*\right]$

4. Repeat step 2 and 3 B times. The Monte Carlo bootstrap variance estimator for $\hat{\theta}$, is then given by
$$\hat{V}_{bwo}(\hat{\theta}) = \left(\tfrac{1}{B-1}\right) \sum_{b=1}^{B} \left(\hat{\theta}_b^* - \bar{\hat{\theta}}^*\right)^2, \qquad (3)$$
where $\bar{\hat{\theta}}^* = \sum_{b=1}^{B} \hat{\theta}_b^* / B$.

Let the subscript E_{boot} and V_{boot} denote expectation and variance, respectively, over all possible bootstrap samples conditional on s. Let $\hat{\theta} = \hat{t}_\pi$, and $\hat{\theta}_b^* = \hat{t}_b^* = \sum_{s_b^*} y_{k^*}/\pi_{k^*} = c \sum_{s_b^*} y_{k^*}$. Since $E_{boot}\hat{V}_{bwo}(\hat{t}_\pi) = V_{boot}(\hat{t}_b^*) = \frac{n-1}{N-1}\frac{N}{n}\hat{V}(\hat{t}_\pi)$, where $\hat{V}(\hat{t}_\pi)$ is the unbiased variance estimator (2), we have $E\hat{V}_{bwo}(\hat{t}_\pi) = \frac{n-1}{N-1}\frac{N}{n}V(\hat{t}_\pi)$. Hence, $\hat{V}_{bwo}(\hat{t}_\pi)$ is biased for $V(\hat{t}_\pi)$ unless it is corrected by the factor $\frac{n}{N}\frac{N-1}{n-1}$.

2.2 Modifications of the BWO

Different approaches have been suggested to modify the BWO scheme to cope with cases where $N = cn + r$, $0 < r < n$.

Bickel & Freedman (1984) suggest the creation of two resampling populations U_A^* and U_B^*, where U_A^* consists of c copies for each $k \in s$, while U_B^* consists of $c+1$ copies for each $k \in s$. One of these resampling populations is selected at random, U_A^* with probability p, U_B^* with probability $1-p$. Denote the outcome of the random choice U^*. Bootstrap samples of size $n^* = n$ are now taken from U^* as in the BWO scheme. By an appropriate choice of p, the variance estimator (3) will be unbiased for $V(\hat{t}_\pi)$ under the designs SI and stratified SI, with the bootstrap procedure applied to each stratum. However, the procedure, (BWO$_{BF}$), is unfeasible for some combinations of N and n, (see McCarthy & Snowden (1985), Sitter (1992b)).

This is also pointed out by Booth et al. (1994), who moreover show that BWO$_{BF}$ can lead to poor results if $c = 1$ (e.g. large sampling fractions). They suggest a method, BWO$_{BBH}$ of 'filling out' U^* to get N elements, as an alternative approach for non-integer f^{-1}. U^* is created by c copies of each element in s plus a random sample (drawn without replacement), of m elements from s, where $m = N - cn = r$.

The modification of the BWO proposed by Sitter (1992b) (BWO$_{\mathrm{Sit}}$), resembles BWO$_{BF}$ in the sense that two resampling populations U_A^* and U_B^* are created. The special characteristics of BWO$_{\mathrm{Sit}}$ can be summarised as follows.

Let $q = f^{-1}\left(1 - \frac{1-f}{n}\right)$ and let $q_A = \lfloor q \rfloor$ and $q_B = \lceil q \rceil$, (where $\lfloor \cdot \rfloor$ denotes the greatest integer equal to or smaller than, and $\lceil \cdot \rceil$ denotes the smallest integer greater than, respectively). U_A^* is now formed to consist of q_A copies of each $k \in s$, while U_B^* is formed to consist of q_B copies of each $k \in s$. The population to resample, U^*, is, for each resample, the result of a random choice between U_A^* (probability p) and U_B^* (probability $1-p$). The bootstrap sample size, n^*, equals $n - 1$ if $U^* = U_A^*$ or n if $U^* = U_B^*$. With a proper choice of p, BWO$_{\mathrm{Sit}}$ yields an unbiased variance estimator. BWO$_{\mathrm{Sit}}$ also overcomes the deficiency for certain combinations of N and n.

The BWO is also discussed by Chao & Lo (1985, 1994). Their generalization of the BWO for noninteger f^{-1}, BWO$_{CL}$, is essentially the same as the one proposed by Booth et al. (1994), (BWO$_{BBH}$), except that the 'filling out' of U^* is done through simple random sampling with replacement.

3. A BWO approach to πps sampling designs

In this section we propose a bootstrap approach to estimate the variance of an estimator $\hat{\theta} = \hat{\theta}(s)$ for a finite population parameter θ under a general fixed size πps sampling design $p(\cdot)$, such that all first-order and second-order inclusion probabilities are strictly positive.

Let $\pi_k = nx_k/t_{xU}$ $(k = 1, ..., N)$ be the first-order inclusion probabilities, let $s \subset U$ be a sample of size n selected by a sample selection scheme obeying the design $p(\cdot)$, and let

$$\pi_k^{-1} = c_k + r_k \quad (k \in s)$$

where $c_k = \lfloor \pi_k^{-1} \rfloor$ and $0 \leq r_k < 1$.

Finally, for $k \in s$, let ε_k be independent Bernoulli random variables with parameters r_k, i.e., $r_k = \Pr(\varepsilon_k = 1) = 1 - \Pr(\varepsilon_k = 0)$.

The bootstrap approach we propose, now proceeds as follows:

1. For $k \in s$, let ε_k be independent realizations of the Bernoulli random variables, and let $d_k = c_k + \varepsilon_k$.

2. Create a resampling population U^* by copying each element $k \in s$ in such a way that element k is copied d_k times, i.e., $U^* = \{1^*, ..., k^*, ..., N^*\}$, where $N^* = \sum_s d_k$. All d_k elements that are copies of element $k \in s$ are assigned the value (y_k, x_k).

3. Draw a bootstrap sample s_1^* of size $n^* = n$ from U^* by applying the same sample selection scheme as for selecting s, which, inter alia, means that $\pi_{k^*} = nx_{k^*}/t_{xU^*}$, where $t_{xU^*} = \sum_{U^*} x_{k^*} = \sum_s d_k x_k$.

4. Compute a bootstrap replicate $\hat{\theta}_1^* = \hat{\theta}(s_1^*)$.

5. Repeat steps 3 and 4 B times. The Monte Carlo bootstrap variance estimator for $\hat{\theta}$ is now given by (3).

We will refer to the above approach as the BWO_{IPA}, where IPA is used to point out that the approach focuses on inclusion probabilities in the creation of the resampling population U^*.

Remark 1. If $r_k = 0$ for all $k \in s$, then U^* simply consists of c_k copies of each $k \in s$. Hence, e.g., $N^* = \sum_s c_k = \sum_s 1/\pi_k = \hat{N}_\pi$, $t_{xU^*} = \sum_s x_k/\pi_k = t_{xU}$ and $F_{U^*}(y) = \hat{F}_s(y)$. If, furthermore, the design is SI, we get $N^* = N$, i.e., the approach is identical to the Gross' BWO.

For simplicity we assume that $\pi_{k^*} \leq 1$ for $k^* \in U^*$ and every U^*

4. An example - Pareto πps sampling

4.1 Pareto πps sampling

In Rosén (1997b) Pareto πps sampling is introduced, together with the following sample selection scheme: (i) For every element $k \in U$, compute target inclusion probabilities $\lambda_k = nx_k/t_{xU}$. (ii) Generate N independent standard uniform random variables $U_1, U_2, \ldots, U_N$ and form the ranking variables

$$Q_k = \frac{U_k(1-\lambda_k)}{(1-U_k)\lambda_k} \qquad (k = 1, \ldots, N). \qquad (4)$$

(iii) The elements with the n smallest Q_k constitute the sample s.

To estimate the population total t_U, Rosén considers

$$\hat{t}_\lambda = \sum\nolimits_s \frac{y_k}{\lambda_k} = \sum\nolimits_s \frac{y_k}{\pi_k} \cdot \frac{\pi_k}{\lambda_k} \qquad (5)$$

which is very close to the π estimator, since λ_k are very close to π_k $(k = 1, \ldots, N)$.

The variance of $\hat{t}_\lambda$ is, unless n and N are very small, well approximated by,

$$\begin{aligned} AV\left(\hat{t}_\lambda\right) &= \frac{N}{N-1} \sum\nolimits_U \lambda_k(1-\lambda_k) \\ &\quad \times \left(\frac{y_k}{\lambda_k} - \frac{\sum_U y_k(1-\lambda_k)}{\sum_U \lambda_k(1-\lambda_k)}\right)^2 \qquad (6) \\ &= \frac{N}{N-1}\frac{1}{2} \sum\sum_{k \neq l \in U} v_{kl} \left(\frac{y_k}{\lambda_k} - \frac{y_l}{\lambda_l}\right)^2, \end{aligned}$$

where $v_{kl} = \lambda_k(1-\lambda_k)\lambda_l(1-\lambda_l)/(n - \sum_U \lambda_k^2)$. As a variance estimator, Rosén proposes

$$\begin{aligned} \hat{V}\left(\hat{t}_\lambda\right) &= \frac{n}{n-1} \sum\nolimits_s (1-\lambda_k) \qquad (7) \\ &\quad \times \left(\frac{y_k}{\lambda_k} - \frac{\sum_s y_k(1-\lambda_k)/\lambda_k}{\sum_s (1-\lambda_k)}\right)^2 . \end{aligned}$$

which, unless n and N are very small, is approximately unbiased.

In Swensson (1997) approximations to π_{kl} are given by $\lambda_{kl} = \lambda_k\lambda_l(1-\gamma_{kl})$, where $\gamma_{kl} = N(1-\lambda_k)(1-\lambda_l)/(n - \sum_U \lambda_k^2)(N-1)$, $k \neq l \in U$.

4.2 BWO_{IPA} and Pareto πps sampling

To apply BWO_{IPA} to Pareto πps sampling we follow the scheme in section 3, the only difference being that the πs are replaced by λs. We will separate the case when all $r_k = 0$, (i.e. all $\lambda_k^{-1} = c_k$ are integers) from the case when some $r_k \neq 0$

All $r_k = 0$ The remark in section 3, implies that $\lambda_{k^*} = \lambda_k$ for every copy $k^* \in U^*$ linked to element $k \in s$.

Let us consider the point estimator $\hat{t}_\lambda$ of equation (5), and let our objective be to estimate its variance. For the BWO_{IPA} we first observe that the expectation of $\hat{t}_b^*$, conditional on s, is $E_{boot}\hat{t}_b^* = \sum_{U^*} \pi_{k^*} y_{k^*}/\lambda_{k^*} \simeq \sum_{U^*} y_{k^*} = t_{U^*} = \sum_s c_k y_k = \hat{t}_\lambda$. Hence,

$$E\bar{\hat{t}}^* \simeq E\hat{t}_\lambda \simeq E\hat{t}_\pi = t_U. \qquad (8)$$

Let the variance estimator (3) be denoted $\hat{V}_{mc}(\hat{t}_\lambda)$. Then we have $E_{boot}\hat{V}_{mc}(\hat{t}_\lambda) = V_{boot}(\hat{t}_b^*) \approx AV_{boot}(\hat{t}_b^*)$, where

$$\begin{aligned} AV_{boot}(\hat{t}_b^*) &= \frac{N^*}{N^*-1}\sum\nolimits_{U^*} \lambda_{k^*}(1-\lambda_{k^*}) \\ &\quad \times \left(\frac{y_{k^*}}{\lambda_{k^*}} - \frac{\sum_{U^*} y_{k^*}(1-\lambda_{k^*})}{\sum_{U^*} \lambda_{k^*}(1-\lambda_{k^*})}\right)^2 \\ &= \frac{N^*}{N^*-1}\frac{n-1}{n}\hat{V}\left(\hat{t}_\lambda\right) \qquad (9) \end{aligned}$$

where we have used the fact that $c_k = \lambda_k^{-1}$, and that $\lambda_{k^*} = \lambda_k$ for every copy $k^* \in U^*$ linked to element $k \in s$. Hence, to get an approximately unbiased and consistent variance estimator, the bootstrap variance estimator has to be slightly corrected. Ignoring the factor $N^*/(N^*-1)$ an alternative bootstrap variance estimator is,

$$\hat{V}_{IPA}(\hat{t}_\lambda) = \frac{n}{n-1}\hat{V}_{mc}(\hat{t}_\lambda). \qquad (10)$$

It is not easy to give general results for more complex parameters. However, some indications of what to expect follows from the Monte Carlo simulations of section 5.

Some $r_k \neq 0$ We now have two sources of variation, (I) the generation of U^*, and (II) the subsequent resampling from U^*. Hence $E_{boot}(\cdot) = E_I E_{II}(\cdot)$ and $V_{boot}(\cdot) = E_I V_{II}(\cdot) + V_I E_{II}(\cdot)$. Since t_{xU^*} now typically differs from t_{xU}, λ_{k^*} usually differs (slightly) from λ_k for a copy k^* linked to element $k \in s$. Hence, the evaluation of $E\hat{V}_{mc}(\hat{t}_\lambda)$ will be less straightforward.

However, since

$$\begin{aligned} E_{boot}t_{U^*} &= E_{boot}\sum\nolimits_s (c_k + \varepsilon_k)y_k \\ &= \sum\nolimits_s (c_k + r_k)y_k \\ &= \sum\nolimits_s \lambda_k^{-1}y_k = \hat{t}_\lambda \qquad (11) \end{aligned}$$

we still have $E\bar{\hat{t}}^* \simeq t_U$. Furthermore, $E_{boot}N^* = \sum_s \lambda_k^{-1} \approx \sum_s \pi_k^{-1} = \hat{N}$, and $E_{boot}t_{x_{U^*}} = \sum_s (c_k + r_k)x_k = \sum_s \lambda_k^{-1}x_k = t_{xU}$.

Next, we see that if $\hat{\theta} = \hat{t}_\lambda$, $E_{boot}\hat{V}_{mc}(\hat{t}_\lambda)$ equals

$$V_{boot}\left(\hat{t}_{s_b^*}\right) = E_I V_{II}(\hat{t}_b^*) + V_I E_{II}(\hat{t}_b^*) \qquad (12)$$

where

$$\begin{aligned} V_I E_{II}(\hat{t}_b^*) &= V_{boot}t_{U^*} \\ &= V_{boot}(\sum\nolimits_s c_k y_k + \sum\nolimits_s \varepsilon_k y_k) \\ &= \sum\nolimits_s r_k(1-r_k)y_k. \qquad (13) \end{aligned}$$

Turning to the first term of (12), we see that $V_{II}(\hat{t}_b^*)$ approximately equals

$$\begin{aligned} AV_{II}(\hat{t}_b^*) &= \frac{N^*}{N^*-1}\sum\nolimits_{U^*} \lambda_{k^*}(1-\lambda_{k^*}) \qquad (14) \\ &\quad \times \left(\frac{y_{k^*}}{\lambda_{k^*}} - \frac{\sum_{U^*} y_{k^*}(1-\lambda_{k^*})}{\sum_{U^*} \lambda_{k^*}(1-\lambda_{k^*})}\right)^2 \end{aligned}$$

which is equal to (9). When taking the expectation over all possible bootstrap samples conditioning on s, there is a major difference between the expressions though. Expression (14) cannot be evaluated in the same way. It can be rewritten to a summation over s incorporating the random variable ε_k, hence

$$\begin{aligned} AV_{II}(\hat{t}_b^*) &= \frac{N^*}{N^*-1}\sum\nolimits_s d_k\lambda_k^{'}(1-\lambda_k^{'}) \qquad (15) \\ &\quad \times \left(\frac{y_k}{\lambda_k^{'}} - \frac{\sum_s d_k y_k(1-\lambda_k^{'})}{\sum_s d_k\lambda_k^{'}(1-\lambda_k^{'})}\right)^2, \end{aligned}$$

where $\lambda_k^{'}$ are the bootstrap target inclusion probability from U^* connected with element k in s, and d_k is the sample value multiplier for element k. Adding (15) and (13) gives,

$$E_{boot}\hat{V}_{mc}(\hat{t}_\lambda) = E_I AV_{II}(\hat{t}_b^*) + \sum\nolimits_s r_k(1-r_k)y_k, \qquad (16)$$

where an expression for the first term is difficult derive. However, its relation to equation (9) suggests that we can expect a reasonable behaviour of the variance estimator in (10), which also is supported by Monte Carlo simulations. To keep the bootstrap variance estimator as simple as possible for the more complex estimators studied in section 5, we use the same expression i.e. $\hat{V}_{IPA}(\hat{\theta}) = \frac{n}{n-1}\hat{V}_{mc}(\hat{\theta})$.

4.3 Variations of the BWO$_{IPA}$

It is easy to see that there are a number of variations of the BWO$_{IPA}$ for the case of noninteger λ_k^{-1} which more or less mimic some of the BWO$_x$. They are not spelled out here, but are presently under study.

5. Empirical Results

To get some preliminary idea of the performance of the proposed method, a limited Monte Carlo simulation was carried out. The finite population used was MU281 (281 Swedish municipalities) in Särndal, Swensson & Wretman (1992), and the variables considered were P75 (=x), SS82 (=y_1), and REV84 (=y_2) with population correlation coefficients $\rho(x, y_1) = 0.7$ and $\rho(x, y_2) = 0.9$. 5000 independent samples of sizes 5 and 40 were generated according to each of two πps schemes (Pareto and Sunter's.) Three parameters were considered, viz. the mean, $\bar{y}_U$, Gini's mean difference, $G_U = \sum\sum_U |y_k - y_l| / N^2$, and the Gini coefficient, $R_U = G_U / 2\bar{y}_U$ (see table 1).

Table 1: Parameter values

	θ			
y_i	t_U	$\bar{y}_U$	G_U	$100 \times R_U$
y_1	6193	22.039	7.907	17.94
y_2	7.574E5	2694.8	2328.7	43.21

Under Pareto πps sampling the estimators used for the three parameters were $\tilde{y} = (\sum_s y_k/\lambda_k)/\hat{N}$, $\hat{G} = (\sum\sum_s |y_k - y_l| / \lambda_{kl})/\hat{N}^2$, where $\hat{N} = \sum_s \lambda_k^{-1}$, and $\hat{R} = \hat{G}/2\tilde{y}$, respectively. Four variance estimators were considered for $\tilde{y}$: (i) An expression based on the SYG variance estimator where π_k and π_{kl} were replaced by λ_k and λ_{kl}, $\hat{V}_{SYGq}$, (see Swensson (1997)). (ii) An expression based on the variance estimator (7), $\hat{V}_{Ro}$. The two estimators $\hat{V}_{SYGq}$ and $\hat{V}_{Ro}$ were derived through first-order Taylor linearisation of $\tilde{y}$, thus replacing y_k in formulas (2) and (7) by $u_k = (y_k - \tilde{y})/\hat{N}$. (iii) The delete-one jackknife variance estimator, $\hat{V}_J$, and, (iv) the BWO$_{IPA}$ variance estimator, $\hat{V}_{IPA}$, based on $B = 300$ bootstrap replicates in each sample*. For the estimators $\hat{G}$ and $\hat{R}$ only $\hat{V}_J$ and $\hat{V}_{IPA}$ were considered.

Under πps sampling according to the Sunter approach $\hat{V}_{Ro}$ does not apply. The variance estimator for $\tilde{y}$, $\hat{V}_{SYG}$, was obtained from $\hat{V}_{SYGq}$ by replacing λ_k and λ_{kl} by π_k and π_{kl}, respectively. The variance estimators for $\hat{G}$ and $\hat{R}$ were the same as those used for the Pareto approach.

Results from the simulation are presented in table 2.

6. Discussion

As can be expected, a sample size of $n = 5$ is too small for reliable inference based on the non-linear estimators considered here. Hence, in the sequel, the discussion will only concern the sample size $n = 40$.

(a) All three point estimators ($\tilde{y}$, $\hat{G}$, and $\hat{R}$) seem to have negligible bias under both schemes.

(b) The jackknife variance estimator has a large positive bias for each combination of sampling scheme and point estimator. This is partly due to its failure of capturing the dependence due to without replacement sampling, which to some extent might be remedied by applying some finite population correction factor.

(c) Comparing $\hat{V}_{IPA}(\tilde{y})$ with its competitors, the Monte Carlo results indicate that it behaves at least as well with respect to relative bias, while its variability seems to be at about the same level.

(d) For the more complex point estimators $\hat{G}$ and $\hat{R}$, $\hat{V}_{IPA}$ consistently underestimates their variation. Different bias-reducing factors are presently being studied.

7. Conclusion

The suggested bootstrap approach, BWO$_{IPA}$, for without replacement probability-proportional to size sampling is relatively simple to implement. By focusing on the inclusion probabilities for the construction of a finite resampling population, its degree of generality is high. The results from the Monte Carlo simulation indicate that the proposed variance estimator might be useful. However, more extensive studies are needed. Such studies are presently in progress. Furthermore, different variations of the approach are also under study.

REFERENCES

Bickel, P. & Freedman, D. (1981), 'Some asymptotic theory for the bootstrap.', *The Annals of Statistics* **9**, 1196–1217.

Bickel, P. & Freedman, D. (1984), 'Asymptotic normality and the bootstrap in stratified sampling.', *The Annals of Statistics* **12**, 470–482.

Booth, J. G., Butler, T. W. & Hall, P. (1994), 'Bootstrap methods for finite populations.', *Journal of the American Statistical Association* **89**, 1282–1289.

Chao, M.-T. & Lo, S.-H. (1985), 'A bootstrap method for finite population.', *Sankya Ser A.* **47**, 399–405.

Chao, M.-T. & Lo, S.-H. (1994), 'Maximum likelihood summary and the bootstrap method in structured finite populations.', *Statistica Sinica* **4**, 389–406.

Cochran, W. G. (1977), *Sampling Techniques*, 3:rd edn, Wiley, New York.

Efron, B. (1979), 'Bootstrap methods: another look at the jackknife.', *Annals of Statistics* **7**, 1–26.

Gross, S. (1980), Median estimation in sample surveys., *in* 'Proceedings of the Section on Survey Research Methods.', American Statistical Association., pp. 181–184.

Kuk, A. Y. C. (1989), 'Double bootstrap estimation of variance under systematic sampling with probability proportional to size.', *Journal of*

*Clearly, the design of the Monte Carlo study does not permit the calculation of reliable bootstrap percentile-t confidence intervals, and the application of the normal approximation is crude in several respects. Hence, no effort is made to study coverage rates of confidence intervals in the present limited Monte Carlo study.

Table 2: Expected values, variances and relative biases for point estimators, $(\hat{\theta})$, and variance estimators, $(\hat{V})$, estimated from $S = 5000$ Monte Carlo runs, (sample sizes $n = 5$ and $n = 40$, variables y_1 and y_2).

		$n = 5$						$n = 40$					
	$\hat{\theta}$ $[y_1, y_2]$	$\widehat{E(\hat{\theta})}$		$S^2(\hat{\theta})$		$\widehat{\% \, bias}^a$		$\widehat{E(\hat{\theta})}$		$S^2(\hat{\theta})$		$\widehat{\% \, bias}^a$	
P	$\tilde{y}$	23.03	3111	14.14	1.59e6	4.5	15.4	22.12	2731	1.62	130825	0.4	1.3
A	$\hat{G}$	7.529	2418	7.28	8.09e5	-4.8	3.8	7.814	2333	0.72	82662	-1.2	0.2
R	$\hat{R}(\%)$	16.83	41.04	44.58	160.85	-6.2	-5.0	17.74	42.92	5.26	14.07	-1.1	-0.7
E	$\hat{V}_{SYGq}(\tilde{y})$	10.11	1.05e6	69.33	8.24e11	-28.5	-33.8	1.49	123527	0.38	3.47e9	-8.6	-5.6
T	$\hat{V}_{Ro}(\tilde{y})$	10.15	1.06e6	71.29	8.56e11	-28.2	-33.2	1.49	123945	0.39	3.63e9	-8.6	-5.6
O	$\hat{V}_J(\tilde{y})$	18.98	2.19e6	623.9	9.89e12	34.2	37.7	1.84	161569	0.67	5.59e9	13.5	23.5
	$\hat{V}_{IPA}(\tilde{y})$	15.19	1.87e6	193.5	3.20e12	7.4	17.6	1.54	129024	0.40	3.63e9	-5.2	-1.3
	$\hat{V}_J(\hat{G})$	10.56	1.08e6	113.1	9.41e11	44.9	33.5	0.96	118125	0.20	4.66e9	33.1	42.9
	$\hat{V}_{IPA}(\hat{G})$	8.35	1.01e6	30.22	3.94e11	14.6	25.5	0.69	78604	0.06	1.89e9	-3.4	-4.9
	$\hat{V}_J(\hat{R})$	58.79	246.05	5263	81195	31.9	53.0	6.29	17.64	18.54	89.47	19.6	25.3
	$\hat{V}_{IPA}(\hat{R})$	40.86	180.67	1238	16040	-8.3	12.3	4.64	13.00	7.11	42.03	-11.7	-7.6
S	$\tilde{y}$	23.11	3134	15.02	1.67e6	4.9	16.3	22.16	2731	1.415	120622	0.5	1.3
U	$\hat{G}$	7.423	2415	7.15	8.22e5	-6.2	3.7	7.850	2337	0.724	82722	-0.7	0.4
N	$\hat{R}(\%)$	16.52	40.78	42.67	161.77	-7.9	-5.6	17.78	42.96	5.11	13.54	-0.9	-0.6
T	$\hat{V}_{SYG}(\tilde{y})$	9.74	1.04e6	65.09	8.09e11	-35.1	-37.5	1.39	114731	0.21	4.53e9	-1.9	-4.8
E	$\hat{V}_J(\tilde{y})$	18.10	2.14e6	557.1	8.92e12	20.5	28.7	1.76	158843	0.38	6.69e9	24.4	31.7
R	$\hat{V}_{IPA}(\tilde{y})$	14.43	1.82e6	168.0	2.86e12	-3.9	9.1	1.43	117834	0.27	3.79e9	0.8	-2.3
	$\hat{V}_J(\hat{G})$	10.32	1.10e6	109.6	1.08e12	44.3	34.2	0.94	122116	0.13	7.78e9	29.7	47.6
	$\hat{V}_{IPA}(\hat{G})$	8.05	1.00e6	28.89	4.00e11	12.5	22.2	0.68	73737	0.05	2.22e9	-6.1	-10.8
	$\hat{V}_J(\hat{R})$	56.69	239.11	5231	77730	32.8	47.8	5.99	16.59	10.30	45.16	17.2	22.4
	$\hat{V}_{IPA}(\hat{R})$	39.10	174.51	1182	15044	-8.3	7.9	4.63	12.18	5.67	26.78	-9.4	-10.1

[a] For the point estimators, $\widehat{\% \, bias} = 100(\widehat{E(\hat{\theta})} - \theta)/\theta$, for the variance estimators, $\widehat{\% \, bias} = 100(\widehat{E\hat{V}(\hat{\theta})} - S^2(\hat{\theta}))/S^2(\hat{\theta})$

statistical computation and simulation **31**, 73–82.

McCarthy, P. J. & Snowden, C. B. (1985), The bootstrap and finite population sampling., *in* 'Vital and Health Statistics', number 95 *in* '2', U.S. government printing office., Public health service publication Washington DC, pp. 1–23.

Rao, J. N. K. & Wu, C. F. J. (1984), Bootstrap inference for sample surveys., *in* 'ASA proceedings of the section of survey research methods.', American Statistical Association, pp. 106–112.

Rao, J. N. K. & Wu, C. F. J. (1988), 'Resampling inference with complex survey data.', *Journal of the American Statistical Association* **83**, 231–241.

Rao, J. N. K., Wu, C. F. J. & Yue, K. (1992), 'Some recent work on resampling methods for complex surveys.', *Survey Methodology* **18**, 209-217.

Rao, Poduri, S. R. S. & Katzoff, M. J. (1996), 'Bootstrap for finite populations.', *Communications in Statistics-Simulation and Computation* **25**(4), 979–994.

Rosén, B. (1996), On sampling with probability proportional to size., Technical Report 1, R D Report Statistics Sweden.

Rosén, B. (1997a), 'Asymptotic theory for order sampling', *Journal of Statistical Planning and Inference* **62**, 135–158.

Rosén, B. (1997b), 'On sampling with probability proportional to size', *Journal of Statistical Planning and Inference* **62**, 159–191.

Särndal, C.-E., Swensson, B. & Wretman, J. (1992), *Model Assisted Survey Sampling.*, Springer-Verlag.

Sitter, R. (1992a), 'A resampling procedure for complex survey data.', *Journal of the American Statistical Association* **87**, 755–765.

Sitter, R. (1992b), 'Comparing three bootstrap methods for survey data.', *The Canadian Journal of Statistics* **20**, 135–154.

Swensson, B. (1997), 'On Pareto πps sampling', Working Paper Series (2), ESA, University of Örebro. To appear in Theory of Stochastic Processes.