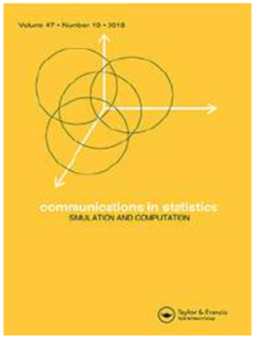

Communications in Statistics - Simulation and Computation

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/lssp20

Modifications on re-scaling bootstrap for adaptive sampling

Sanghamitra Pal & Dipika Patra

To cite this article: Sanghamitra Pal & Dipika Patra (2021): Modifications on re-scaling bootstrap for adaptive sampling, Communications in Statistics - Simulation and Computation, DOI: 10.1080/03610918.2021.1887233

To link to this article: https://doi.org/10.1080/03610918.2021.1887233

Published online: 18 Feb 2021.

COMMUNICATIONS IN STATISTICS - SIMULATION AND COMPUTATION®
https://doi.org/10.1080/03610918.2021.1887233

Modifications on re-scaling bootstrap for adaptive sampling

Sanghamitra Pal and Dipika Patra

Department of Statistics, West Bengal State University, Kolkata, West Bengal, India

ABSTRACT

In complex survey design, the resampling method is often used for assessing the variability and confidence interval of non-linear estimators which is a function of several estimated means or totals. A well-known resampling method, called the re-scaling bootstrap technique was originated by Rao and Wu (1988). This article is an attempt to propose a re-scaling bootstrap technique in estimating the parameters of rare and clustered population for a complex survey design. Adaptive cluster sampling design is a probabilistic approach to reach out to the rare and clustered units. Thompson (1990) introduced this design. Chaudhuri (2000) extended this design for varying probability sampling. In practice, the final sample size of adaptive sampling may be exorbitantly large. From this realistic point of view, Chaudhuri, Bose, and Dihidar (2005) developed the size-constrained adaptive sampling design, in varying probability sampling. We first describe here, how this well-known re-scaling bootstrap technique may be employed for non-linear estimators of the population total, in the case of rare and clustered population. It has been found that there is a need to develop an alternative re-scaling bootstrap procedure to avoid complications in computation to cover Chaudhuri, Bose, and Dihidar (2005). A simulation study has been carried out to demonstrate the proposed methods.

ARTICLE HISTORY
Received 7 July 2020
Accepted 3 February 2021

KEYWORDS
Adaptive sampling; Bootstrap; Rare; Re-scaling; Simulation; Varying probability

AMS SUBJECT CLASSIFICATION
62D05

1. Introduction

Thompson (1990)'s adaptive cluster sampling design considers the initial sample selected by simple random sampling. Later, Chaudhuri (2000) extended this for the varying probability schemes referring as adaptive sampling. The greatest drawback of the adaptive sampling design is that the final sample size may be exorbitantly large in real-life applications. Thompson (1994), Brown (1994), Brown and Manly (1998), Salehi and Seber (1997, 2002), Chaudhuri, Bose, and Ghosh (2004), Chaudhuri, Bose, and Dihidar 2005), and many others suggested alternative way-outs to overcome this drawback. Recently Pal and Patra (2021) developed a prediction approach in adaptive sampling design. In Sec. 2, we briefly review the adaptive sampling design under varying probability schemes.

Chao (2004), Dryver and Chao (2007) suggested ratio estimators under Thompson (1990)'s adaptive cluster sampling design. Variance estimation and confidence interval estimation for such nonlinear estimators are often a complicated task. Rao and Wu (1988)'s re-scaling bootstrap technique works well in estimating the variance of a non-linear statistic for varying probability sampling designs. In this technique, bootstrap resamples are taken from the original sample already at hand and these resampled values are usually re-scaled in such a way that the bootstrap expectation and variance are equal to the original sample-based estimate and its unbiased variance

CONTACT Dipika Patra dipika.patra1988@gmail.com Department of Statistics, West Bengal State University, Kolkata, West Bengal 700126, India.

estimate, in the linear case. Later, modifications on re-scaling bootstrap are recommended in Pal (2009) for non-fixed size design.

Christman and Pontius (2000) investigated the performance of several non-parametric bootstrap methods including Rao and Wu (1988)'s re-scaling bootstrap with Hansen and Hurwitz (1943) estimator under Thompson (1990)'s adaptive cluster sampling. Perez and Pontius (2006) also developed several methods with the Horvitz and Thompson (1952) method of estimation. Their prescribed estimators are biased with extremely low coverage rates. We refer to Mohammadi, Salehi, and Rao (2014) for a detailed study of these shortcomings. Recently Biswas, Rai, and Ahmad (2020) discussed the application of re-scaling bootstrap in rank set sampling.

This article aims to develop a suitable re-scaling bootstrap technique for Chaudhuri (2000)'s general adaptive sampling design and also for Chaudhuri, Bose, and Dihidar (2005)'s size-constrained adaptive sampling design for varying probability sampling. Section 3 is designed for the implementation and necessary modifications on Rao and Wu (1988)'s re-scaling bootstrap method in general adaptive sampling and size-constrained adaptive sampling for varying probability designs considering Horvitz and Thompson (1952) and Rao, Hartley, and Cochran (1962) methods of estimation, respectively. Sections 3.1.1 and 3.2.1 are considered for the Horvitz and Thompson (1952) method of estimation. For, Rao, Hartley, and Cochran (1962) method of estimation, Secs. 3.1.2 and 3.2.2 are considered. Since computation plays a key role in bootstrap, the objective of this current study is to design a suitable algorithm that could avoid complications in computation and also provide satisfactory results. An alternative re-scaling bootstrap technique for size-constrained adaptive sampling is proposed in Sec. 4, for both estimation methods. A simulation study is performed for the variance estimation of a non-linear estimator of the population total, and the performance reports are shown in Sec. 5. The bootstrap confidence intervals are computed by percentile method, and also with the method, assuming normality. The data obtained from https://dolr.gov.in/district-and-category-wise-wastelands-year-2000 has been used for the simulation study to compare proposed methods of rescaling bootstrap procedures.

2. Adaptive sampling design

Consider a finite population $U = \{1, 2, ..., N\}$ of size N and $y = (y_1, y_2, ..., y_N)$ be a study variable with rarity characteristics. The problem is to estimate the population total, $= \sum_{i=1}^{N} y_i$. Adaptive sampling design is applicable for surveying such a rare and clustered population. According to the pioneering work of Thompson (1990), an initial sample of size n is drawn by simple random sampling and the sampled units are observed. If any of them meets the predefined criterion of rarity says, $y_i > c,$ its neighboring units are searched for the further detection of a rarity for the unit i $(i = 1, 2, .n)$. Those neighboring units meeting the rarity condition are included in the sample. The adding process is continued until we get a boundary with edge units. The neighborhood should be well-defined. Those adjacent units which do not meet the criterion of rarity, are called *edge units*. A network $A(i)$ with m_i number of units is found for i^{th} sampled unit $(i = 1, 2, ...n)$. Note that, if i^{th} unit in the initial sample is not rare, its network $A(i)$ consists of that unit, only.

2.1. General adaptive sampling for varying probability designs

Chaudhuri (2000) extended the above procedure of adaptive sampling where the initial sample is drawn by varying probability designs, using a suitable size measure variable. Following Chaudhuri (2000) and using Horvitz and Thompson (1952) method of estimation, an unbiased estimator of the population total $Y = \sum_{i=1}^{N} y_i$ is given by

$$t_{HT} = \sum_{i \in s} \frac{t_i}{\pi_i};$$

where $t_i = \frac{1}{m_i} \sum_{j \in A(i)} y_j$, the average value of the study variable belonging to i^{th} unit's network $A(i)$ and π_i be the first order inclusion probability of i^{th} unit.

The fact $T = \sum_{i \in U} t_i = \sum_{i \in U} y_i = Y$, has been used for unbiased estimation of adaptive sampling design.

An unbiased variance estimator of t_{HT}, is formulated as follows:

$$v(t_{HT}) = \sum_{i<j \in s} \sum \frac{(\pi_i \pi_j - \pi_{ij})}{\pi_{ij}} \left(\frac{t_i}{\pi_i} - \frac{t_j}{\pi_j} \right)^2;$$

where π_{ij} be the second-order inclusion probability of *ith* and *jth* units in the sample.

Chaudhuri, Bose, and Ghosh (2004) considered the Rao, Hartley, and Cochran (1962) method of sampling strategy and the estimation procedure in general adaptive sampling.

To choose an initial sample of size n using Rao, Hartley, and Cochran (1962) scheme, the population U of size N is divided at random into n groups, namely $G_1, G_2, ..., G_n$ with sizes $N_1, N_2, ..., N_n$, respectively, such that $\sum_{i=1}^{n} N_i = N$. Then, one unit is drawn from each of the n groups with probability $\frac{p_{kj}}{\sum_{j \in G_k} p_{kj}}$ i.e., $\frac{p_k}{Q_k}$. Here p_{kj} is the probability of selecting one unit (j^{th}) from k^{th} group and $Q_k = \sum_{j \in G_k} p_{kj}$, $k = 1, 2, ..., n$. The adaptive sampling procedure is applied to the initial sample of size n in rare and clustered population.

Hence, an unbiased estimator of the population total in general adaptive sampling design for the Rao, Hartley, and Cochran (1962) method of estimation may be written as,

$$t_{RHC} = \sum_{i=1}^{n} \frac{t_i}{p_i / Q_i}$$

with an unbiased variance estimator as,

$$v(t_{RHC}) = \frac{\sum_{i=1}^{n} N_i^2 - N}{N^2 - \sum_{i=1}^{n} N_i^2} \left(\sum_{i=1}^{n} Q_i \left(\frac{t_i}{p_i} - t_{RHC} \right)^2 \right).$$

2.2. Size-constrained adaptive sampling for varying probability designs

A major disadvantage of adaptive sampling is that the network size may be unmanageably large. Such a population demands spare time and cost. Chaudhuri, Bose, and Dihidar (2005) exhibited a notable way out for varying probability design, called *size-constrained adaptive sampling*. In this design, a subset $B(i)$ is drawn from $A(i)$ by simple random sampling without replacement (SRSWOR) for i^{th} unit in the sample, subject to the condition $\sum_{i \in s} l_i \leq L$; l_i is the cardinality of the subset $B(i)$ and L be the pre-fixed integer value. Then, following Horvitz and Thompson (1952), an unbiased estimator of Y can be written as,

$$t_{HT}^* = \sum_{i \in s} \frac{e_i}{\pi_i} \tag{2.1}$$

writing $e_i = \frac{1}{l_i} \sum_{j \in B(i)} y_j$ as the average y values of the subsampled units of i^{th} unit's network.

The related unbiased estimator of variance is given by

$$v(t_{HT}^*) = \sum_{i \in s} \frac{v_R(e_i)}{\pi_i} + \sum_{i<j \in s} \sum \left(\frac{\pi_i \pi_j - \pi_{ij}}{\pi_{ij}} \right) \left(\frac{e_i}{\pi_i} - \frac{e_j}{\pi_j} \right)^2;$$

$$v_R(e_i) = \left(\frac{1}{l_i} - \frac{1}{m_i}\right)\left(\frac{1}{l_i - 1}\right) \sum_{j \in B(i)} (y_j - e_j)^2 \tag{2.2}$$

Similarly, if the Rao, Hartley, and Cochran (1962) method of estimation is performed, an unbiased estimator of the population total becomes

$$t^*_{RHC} = \sum_{i=1}^{n} \frac{e_i}{p_i / Q_i} \tag{2.3}$$

An unbiased variance estimator is

$$v\left(t^*_{RHC}\right) = \frac{\sum_{i=1}^{n} N_i^2 - N}{N^2 - \sum_{i=1}^{n} N_i^2} \left(\sum_{i=1}^{n} Q_i \left(\frac{e_i}{p_i} - t^*_{RHC}\right)^2\right) + \sum_{i=1}^{n} \left(\frac{1}{l_i} - \frac{1}{m_i}\right)\left(\frac{1}{l_i - 1}\right)\left\{\sum_{j \in B(i)} (y_j - e_j)^2\right\} \frac{Q_i}{p_i}. \tag{2.4}$$

3. Rao and Wu (1988)'s re-scaling bootstrap technique

Efron (1982) developed bootstrap procedures for an independent and identically distributed (iid) sample of fixed size design. An important issue comes up in performing the naive bootstrap with a stratified sampling design. In this design, the ratio of bootstrap variance estimator of the sample mean to the customary unbiased variance estimator of the sample mean does not converge to 1 in probability unless the stratum sample size is very large. Therefore, the bootstrap variance for the sample mean becomes an inconsistent estimator which is described in Rao and Wu (1988). To overcome this issue, Rao and Wu (1988) proposed a re-scaling bootstrap technique.

Let θ be the parameter of interest which is a non-linear function of p population means or totals, say $\theta = f(\theta_1, \theta_2, ..., \theta_p)$ and it is estimated by a non-linear statistic $\hat{\theta} = f\left(\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_p\right)$ which is the same function of p unbiased estimated means or totals. Here, $\hat{\theta}_j$ is a linear estimator of θ_j. Suppose, y_i is the value of the variable of interest for i^{th} unit in the population and the objective is to estimate the finite population total $Y = \sum_{i=1}^{N} y_i$ by a non-linear estimator.

For instance, let us postulate a model, $y_i = \beta x_i + \epsilon_i$ in which β be an unknown constant and x_i be the value of an auxiliary variable x corresponding to i^{th} unit, assuming known $X = \sum_{i=1}^{N} x_i$. Here, ϵ_i's are independent with means 0 and variances σ_i^2, $i = 1, 2, .N$. Then, the generalized regression (greg) estimator t_{greg} (Cassel, Sarndal, and Wretman 1976) may be written as

$$t_{greg} = \sum_s \frac{y_i}{\pi_i} + \left(X - \sum_s \frac{x_i}{\pi_i}\right) \frac{\sum_s y_i x_i w_i}{\sum_s x_i^2 w_i}. \tag{3.1}$$

The usual choices of w_i's are $\frac{1}{\pi_i x_i}$, $\frac{1-\pi_i}{\pi_i x_i}$, $\frac{1}{x_i}$, $\frac{1}{x_i^2}$, $\frac{1}{x_i^g}$ $(0 < g < 2)$, if Horvitz and Thompson (1952) method of estimation is used.

The estimator t_{greg} (Eq. (3.1)) is a non-linear function of four (here, $p = 4$) unbiased estimators of four population totals. It may be written as,

$$\hat{\theta} = f\left(\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \hat{\theta}_4\right) = \hat{\theta}_1 + \left(\theta_2 - \hat{\theta}_2\right) \frac{\hat{\theta}_3}{\hat{\theta}_4}; \tag{3.2}$$

in which $\hat{\theta}_1 = \sum_{i \in s} \frac{y_i}{\pi_i}$, $\hat{\theta}_2 = \sum_{i \in s} \frac{x_i}{\pi_i}$, $\hat{\theta}_3 = \sum_{i \in s} y_i x_i w_i$, $\hat{\theta}_4 = \sum_{i \in s} x_i^2 w_i$ are four unbiased estimators of the four population totals say, $\theta_1 = Y$, $\theta_2 = X = \sum_{i=1}^{N} x_i, \theta_3 = \sum_{i=1}^{N} y_i x_i w_i \pi_i$, $\theta_4 = \sum_{i=1}^{N} x_i^2 w_i \pi_i$, respectively and $f(\theta_1, \theta_2, \theta_3, \theta_4) = Y$, with the assumption of known X .

For estimating the variance of $\hat{\theta}$, Rao and Wu (1988)'s re-scaling bootstrap for varying probability sampling designs may be used here. According to Rao and Wu (1988), a bootstrap sample is drawn in such a way that the bootstrap-based expectation is equal to their original sample-based estimate, and the bootstrap variance is equal to the related usual unbiased variance estimate, in the linear case. This procedure is repeated independently, a large number of times, say B times and found B number of bootstrap estimates for each of the linear estimator. Suppose, $\hat{\theta}^*_{1b}$, $\hat{\theta}^*_{2b}$, $\hat{\theta}^*_{3b}$, and $\hat{\theta}^*_{4b}$ are b^{th} $(b = 1, 2, \ldots, B)$ bootstrap estimates derived from the same bootstrap sample (s^*_b) corresponding to θ_1, θ_2, θ_3, and θ_4 respectively. Therefore, a bootstrap estimate of Y is taken as $f\left(\hat{\theta}^*_1, \hat{\theta}^*_2, \hat{\theta}^*_3, \hat{\theta}^*_4\right)$ writing, $\hat{\theta}^*_1 = \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}^*_{1b}$, $\hat{\theta}^*_2 = \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}^*_{2b}$, $\hat{\theta}^*_3 = \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}^*_{3b}$, and $\hat{\theta}^*_4 = \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}^*_{4b}$. For vivid illustration, we refer to the monograph Chaudhuri (2010, chapter 8), here.

In the next two Secs. 3.1 and 3.2, we briefly describe the re-scaling bootstrap procedure, employing general adaptive sampling and size-constrained adaptive sampling respectively. Each sub-section considers the Horvitz and Thompson (1952) and Rao, Hartley, and Cochran (1962) methods of estimation.

3.1. Proposed re-scaling bootstrap procedure for general adaptive sampling with varying probability designs

We propose an amendment of the re-scaling bootstrap technique for general adaptive sampling, of the non-linear estimator "f" (in Eq. (3.2)) in estimating the population total. Writing $c_i = \frac{1}{m_i}\sum_{j\in A(i)} x_j$, the average of the values of the auxiliary variable x which belong to the same network $A(i)$, the greg estimator with the Horvitz and Thompson (1952) method of estimation for Y can be defined using the Eq. (3.2) as, $f\left(\sum_s \frac{t_i}{\pi_i}, \sum_s \frac{c_i}{\pi_i}, \sum_s t_i c_i w'_i, \sum_s c_i^2 w'_i\right)$. The usual choices of w'_i's are $\frac{1}{\pi_i c_i}$, $\frac{1-\pi_i}{\pi_i c_i}$, $\frac{1}{c_i}$, $\frac{1}{c_i^2}$, $\frac{1}{c_i^g}$; $(0 < g < 2)$.

Similarly, the Rao, Hartley, and Cochran (1962) version of the greg estimator can be written as

$$f\left(\hat{\theta}_{1(RHC)*}, \hat{\theta}_{2(RHC)*}, \hat{\theta}_{3(RHC)*}, \hat{\theta}_{4(RHC)*}\right) = \hat{\theta}_{1(RHC)*} + \left(\theta_2 - \hat{\theta}_{2(RHC)*}\right)\frac{\hat{\theta}_{3(RHC)*}}{\hat{\theta}_{4(RHC)*}}. \tag{3.3}$$

Here, $\hat{\theta}_{1(RHC)*} = \sum_s \frac{t_i}{p_i/Q_i}$, $\hat{\theta}_{2(RHC)*} = \sum_s \frac{c_i}{p_i/Q_i}$, $\hat{\theta}_{3(RHC)*} = \sum_s t_i c_i w''_i$ and $\hat{\theta}_{4(RHC)*} = \sum_s c_i^2 w''_i$. The choices of w''_i's are $\frac{Q_i}{p_i c_i}$, $\frac{Q_i - p_i}{p_i c_i}$, $\frac{1}{c_i}$, $\frac{1}{c_i^2}$, $\frac{1}{c_i^g}$; $(0 < g < 2)$.

3.1.1. Proposed re-scaling bootstrap for Horvitz and Thompson (1952) method of estimation with general adaptive sampling design

For $\sum_s \frac{t_i}{\pi_i}$, the re-scaling bootstrap procedure may be modified as follows:

Step 1. A set of m ordered pairs say, $\left\{\left(\frac{t_{i^*}}{\pi_{i^*}}, \frac{t_{j^*}}{\pi_{j^*}}\right);\ i^*, j^* \in s,\ i^* \neq j^*\right\}$ are drawn using with replacement procedure from the $n(n-1)$ pairs of sampled units with the probability $q_{i^*j^*}$.
Step 2. Then compute, $et_{HT} = t_{HT} + \frac{1}{m}\sum_{i*, j* \in s_1} k_{i*j*}\left(\frac{t_{i*}}{\pi_{i*}} - \frac{t_{j*}}{\pi_{j*}}\right)$;

where k_{ij} is so chosen that $E_*(et_{HT}) = t_{HT}$ and $V_*(et_{HT}) = v(t_{HT})$. Here E_* and V_* denote bootstrap expectation and variance, respectively. This t_{HT} is defined in the Sec. 2.1.

Therefore, $k_{ij}^2 q_{ij} = \frac{\pi_i\pi_j - \pi_{ij}}{2\pi_{ij}} m$.

Note that, a particular choice of (m, q_{ij}, k_{ij}) is $\left(n(n-1), \frac{1}{n(n-1)}, n(n-1)\left(\frac{\pi_i\pi_j - \pi_{ij}}{2\pi_{ij}}\right)^{1/2}\right)$.

Step 3. The same set of m ordered pairs is considered to get the bootstrap estimates for θ_2, θ_3 and θ_4 whose linear estimators are for $\sum_s \frac{c_i}{\pi_i}$, $\sum_s t_i c_i w_i'$ and $\sum_s c_i^2 w_i'$. Let, ec_{HT}, etc_{HT} and ec_{HT}^* denote the bootstrap estimates for $\sum_s \frac{c_i}{\pi_i}$, $\sum_s t_i c_i w_i'$ and $\sum_s c_i^2 w_i'$, respectively.
Step 4. Finally a bootstrap estimate $f\ (et_{HT}, ec_{HT}, etc_{HT}, ec_{HT}^*)$ is found for the population total Y.
Step 5. Replicate **Step 1** to **Step 4,** independently, for a large number of times say, B times. Let b^{th} bootstrap estimate is represented by $\tilde{t}^{*b}_{HT,greg.AS}$. That is, $\tilde{t}^{*b}_{HT,greg.AS} = f\ (et_{HT}, ec_{HT}, etc_{HT}, ec_{HT}^*)$. Therefore, the B number of bootstrap estimates say, $\tilde{t}^{*1}_{HT,greg.AS}, \tilde{t}^{*2}_{HT,greg.AS}, ..., \tilde{t}^{*B}_{HT,greg.AS}$ are found.
Step 6. The final bootstrap estimate becomes $\tilde{t}^{B}_{HT,greg.AS} = \frac{1}{B}\sum_{b=1}^{B} \tilde{t}^{*b}_{HT,greg.AS}$ and its bootstrap variance estimate can be written as, $\tilde{v}^{HT}_{b.AS} = \frac{1}{B-1}\sum_{b=1}^{B}\left(\tilde{t}^{*b}_{HT,greg.AS} - \tilde{t}^{B}_{HT,greg.AS}\right)^2$.

3.1.2. Proposed re-scaling bootstrap for Rao, Hartley, and Cochran (1962) method of estimation with general adaptive sampling design

Following Rao and Wu (1988, Section 5.1), the re-scaling bootstrap procedure is described below for the linear total $\hat{\theta}_{1(RHC)*} = \sum_s \frac{t_i}{p_i/Q_i}$ (see Eq. (3.3)).

Step 1. A bootstrap sample $\{\frac{t_{i*}}{p_{i*}}\}_{i*=1}^{m}$ of size m is drawn using with replacement procedure from the original sample $\{t_k,\ p_k\}_{k=1}^{n}$, with probabilities Q_k.
Step 2. Then compute, $et_{RHC} = t_{RHC} + \lambda m^{-1/2}\sum_{i^*\in s_1}\left(\frac{t_{i*}}{p_{i*}} - t_{RHC}\right)$. This t_{RHC} is defined in Sec. 2.1.

Here, λ is so chosen that $E_*(et_{RHC}) = t_{RHC}$ and $V_*(et_{RHC}) = v(t_{RHC})$.

That implies, $\lambda = m^{1/2}\left(\frac{\sum_{i=1}^{n} N_i^2 - N}{N^2 - \sum_{i=1}^{n} N_i^2}\right)^{1/2}$.

Step 3. Step 2 is repeated for $\sum_s \frac{c_i}{p_i/Q_i}$, $\sum_s t_i c_i w_i''$, $\sum_s c_i^2 w_i''$ and the bootstrap estimates ec_{RHC}, etc_{RHC}, ec_{RHC}^* are calculated.
Step 4. Proceeding in a similar way, a bootstrap estimate $f\ (et_{RHC}, ec_{RHC}, etc_{RHC}, ec_{RHC}^*)$ is found.
Step 5. Step 1 to **Step 4** are replicated B times to get B number of bootstrap estimates, say, $\tilde{t}^{*1}_{RHC,greg.AS}, \tilde{t}^{*2}_{RHC,greg.AS}, ..., \tilde{t}^{*B}_{RHC,greg.AS}$. Here, $\tilde{t}^{*b}_{RHC,greg.AS}$ denotes b^{th} $(b = 1, 2, ..., B)$ bootstrap estimate.
Step 6. Then, the final bootstrap estimate is $\tilde{t}^{B}_{RHC,greg.AS} = \frac{1}{B}\sum_{b=1}^{B} \tilde{t}^{*b}_{RHC,greg.AS}$ and the bootstrap variance estimate is $\tilde{v}^{RHC}_{b.AS} = \frac{1}{B-1}\sum_{b=1}^{B}(\tilde{t}^{*b}_{RHC,greg.AS} - \tilde{t}^{B}_{RHC,greg.AS})^2$.

3.2. Proposed re-scaling bootstrap for size-constrained adaptive sampling with varying probability designs

The greg estimator for the size-constrained adaptive sampling, using the Horvitz and Thompson (1952) method of estimation can be written as $f\left(\sum_s \frac{e_i}{\pi_i}, \sum_s \frac{c'_i}{\pi_i}, \sum_s e_i c'_i w'_i, \sum_s c'^2_i w'_i\right)$ where $e_i = \frac{1}{l_i}\sum_{j\in B(i)} y_j$ and $c'_i = \frac{1}{l_i}\sum_{j\in B(i)} x_j$ which are the average y and x (auxiliary variable) values of the units in (i), respectively. Here, w_i''s are chosen from $\frac{1}{\pi_i c_i'}$, $\frac{1-\pi_i}{\pi_i c_i'}$, $\frac{1}{c_i'}$, $\frac{1}{c_i'^2}$, $\frac{1}{c_i'^g}$; $(0 < g < 2)$.

Similarly, the Rao, Hartley, and Cochran (1962) version of the greg estimator can be written as $f\left(\sum_s \frac{e_i}{p_i/Q_i}, \sum_s \frac{c_i'}{p_i/Q_i}, \sum_s e_i c_i' w_i'', \sum_s c_i'^2 w_i''\right)$ with the usual choices of w_i'''s as $\frac{Q_i}{p_i c_i'}$, $\frac{Q_i - p_i}{p_i c_i'}$, $\frac{1}{c_i'}$, $\frac{1}{c_i'^2}$, $\frac{1}{c_i'^g}$; $(0 < g < 2)$.

3.2.1. Modifications on re-scaling bootstrap procedure for Horvitz and Thompson (1952) method of estimation with size-constrained adaptive sampling

In Sec. 2.2, we have seen in Eq. (2.2) that the unbiased variance estimator of size-constrained adaptive sampling for the HT method of estimation is written as,

$$v\left(t^*_{HT}\right) = \sum_{i\in s} \frac{v_R(e_i)}{\pi_i} + \sum_{i<j\in s}\sum \left(\frac{\pi_i\pi_j - \pi_{ij}}{\pi_{ij}}\right)\left(\frac{e_i}{\pi_i} - \frac{e_j}{\pi_j}\right)^2 = v_2 + v_1.$$

Following Pal (2009), two independent bootstrap samples are drawn by suitable bootstrap designs which are described below:

Step 1. For $\sum_s \frac{e_i}{\pi_i}$, we have drawn a bootstrap sample (s_1) of m pairs $\left\{\left(\frac{e_{i*}}{\pi_{i*}}, \frac{e_{j*}}{\pi_{j*}}\right); (i*, j*) \in s, i* \neq j*\right\}$ from $n(n-1)$ pairs of the original sample $\{e_i\}_{i=1}^n$, using with replacement procedure with the probability q_{i*j*}.

Step 2. Then calculate, $et_1 = t^*_{HT} + \frac{1}{m}\sum_{i*, j*\in s_1} k_{i*j*}\left(\frac{e_{i*}}{\pi_{i*}} - \frac{e_{j*}}{\pi_{j*}}\right)$.

Here, k_{ij} is so chosen that $E_*(et_1) = t^*_{HT}$ and $V_*(et_1) = v_1$. That implies, $k^2_{ij}q_{ij} = \frac{\pi_i\pi_j - \pi_{ij}}{2\pi_{ij}}m$.

A particular choice of (m, q_{ij}, k_{ij}) is $\left(n(n-1), \frac{1}{n(n-1)}, n(n-1)\left(\frac{\pi_i\pi_j - \pi_{ij}}{2\pi_{ij}}\right)^{1/2}\right)$.

Step 3. Another bootstrap sample (s_2) is drawn independently with a success probability r'_i for i^{th} unit, following Hajek (1964)'s Poisson sampling scheme and $et_2 = \sum_{i*\in s_2} \frac{(v_R(e_{i*}))^{1/2}}{r'_{i*}} - \sum_{i\in s} v_R(e_i)$ is calculated, considering $\frac{1}{r'_i} = \frac{\pi_i + 1}{\pi_i}$.

Step 4. So, the bootstrap estimate correspond to the linear estimator $\sum_s \frac{e_i}{\pi_i}$ becomes, $\tilde{t}^*_{HT} = et_1 + et_2$ which is derived from the bootstrap sample $s* = (s_1, s_2)$. Clearly, $E_*\left(\tilde{t}^*_{HT}\right) = t^*_{HT}$ and $V_*\left(\tilde{t}^*_{HT}\right) = v(t^*_{HT})$.

Step 5. Next, the same set $s^* = (s_1, s_2)$ is considered to get the bootstrap estimate for $\sum_s \frac{c'_i}{\pi_i}, \sum_s e_i c'_i w'_i$ and $\sum_s c'^2_i w'_i$. Thereafter a bootstrap estimate of the population total is derived as described in the previous section.

Step 6. Step 1 to ***Step 5*** are replicated independently, a large number of times, say B times, and the bootstrap estimates $\tilde{t}^{*1}_{HT,greg}, \tilde{t}^{*2}_{HT,greg}, ..., \tilde{t}^{*B}_{HT,greg}$ are found. Herein, $\tilde{t}^{*b}_{HT,greg}$ denotes b^{th} bootstrap estimate of the greg estimator.

Step 7. Therefore, the final bootstrap estimate is $\tilde{t}^B_{HT,greg} = \frac{1}{B}\sum_{b=1}^B \tilde{t}^{*b}_{HT,greg}$ and its bootstrap variance estimate is given by, $\tilde{v}_b{}^{HT} = \frac{1}{B-1}\sum_{b=1}^B \left(\tilde{t}^{*b}_{HT,greg} - \tilde{t}^B_{HT,greg}\right)^2$.

3.2.2. Modifications on re-scaling bootstrap procedure for Rao, Hartley, and Cochran (1962) method of estimation with size-constrained adaptive sampling

The unbiased variance estimator of size-constrained adaptive sampling with Rao, Hartley, and Cochran (1962) method of estimation is (as written in Eq. 2.4),

$$v\left(t^*_{RHC}\right) = \frac{\sum_{i=1}^n N_i^2 - N}{N^2 - \sum_{i=1}^n N_i^2}\left(\sum_{i=1}^n Q_i\left(\frac{e_i}{p_i} - t^*_{RHC}\right)^2\right) + \sum_{i=1}^n \left(\frac{1}{l_i} - \frac{1}{m_i}\right)\left(\frac{1}{l_i - 1}\right)\left\{\sum_{j\in B(i)} (y_j - e_j)^2\right\}\frac{Q_i}{p_i}$$
$$= v_1 + v_2$$

Similar to the Horvitz and Thompson (1952) method of estimation defined in Sec. 3.2.1, two independent bootstrap samples are drawn from the original sample s for $\sum_s \frac{e_i}{p_i/Q_i}$.

The steps are as follows:

Step 1. We have drawn a with replacement sample $s_1 : \{\frac{e_{i*}}{p_{i*}}\}_{i*=1}^{m}$ from the original sample with probability Q_k and then compute, $es_1 = t^*_{RHC} + \lambda m^{-1/2} \sum_{i^* \in s_1} \left(\frac{e_{i^*}}{p_{i^*}} - t^*_{RHC}\right)$; t^*_{RHC} is defined in Eq. (2.3). Here, λ is so chosen that $E_*(es_1) = t^*_{RHC}$ and $V_*(es_1) = v_1$.

Therefore, $\lambda = m^{1/2}\left(\frac{\sum_{i=1}^{n} N_i^2 - N}{N^2 - \sum_{i=1}^{n} N_i^2}\right)^{1/2}$.

Step 2. Another sample s_2 is drawn independently, by Poisson Sampling Scheme (Hajek 1964), with the success probability r_i and then compute, $es_2 = \sum_{i^* \in s_2} \frac{(v_R(e_{i*}))^{1/2}}{r_{i*}} - \sum_{i \in s} v_R(e_i)$.

Here, r_i is so chosen that $E_*(es_2) = 0$ and $V_*(es_2) = v_2$, yielding $r_i = \frac{p_i}{Q_i + p_i}$.

Step 3. Finally, the bootstrap estimate for $\sum_s \frac{e_i}{p_i / Q_i}$ becomes $\tilde{t}^*_{RHC} = es_1 + es_2$ with the basic conditions $E_*\left(\tilde{t}^*_{RHC}\right) = t^*_{RHC}$ and $V_*\left(\tilde{t}^*_{RHC}\right) = v(t^*_{RHC})$.

Note that, a variance of the estimator $\tilde{t}^*_{RHC(HT)}$ can be derived as follows,

$$\begin{aligned} V\left(\tilde{t}^*_{RHC(HT)}\right) &= E_{AS} V_B\left(\tilde{t}^*_{RHC(HT)}\right) + V_{AS} E_B\left(\tilde{t}^*_{RHC(HT)}\right) \\ &= E_{AS}\left(v\left(t^*_{RHC(HT)}\right) + V_{AS}\left(t^*_{RHC(HT)}\right) = 2V\left(t^*_{RHC(HT)}\right);\right. \end{aligned}$$

E_{AS}, V_{AS} denote the expectation and the variance operators for general adaptive sampling design while E_B, V_B for bootstrap expectation and bootstrap variance, respectively.

Therefore, its unbiased variance estimator becomes $2v\left(t^*_{RHC(HT)}\right)$ for the RHC (HT) method of estimation. $v(t^*_{HT})$ and $v(t^*_{RHC})$ are defined in Eq. (2.2) and Eq. (2.4), respectively.

Step 4. Similarly, we get bootstrap estimates for θ_2, θ_3 and θ_4 whose linear estimators are $\sum_s \frac{c_i'}{p_i / Q_i}$, $\sum_s e_i c_i' w_i''$ and $\sum_s c_i'^2 w_i''$. Thereafter, one may get the final bootstrap estimate as narrated above.

Step 5. ***Step 1*** to ***Step 4*** are replicated B times to find B number of estimates say, $\tilde{t}^{*1}_{RHC,greg}, \tilde{t}^{*2}_{RHC,greg}, \ldots, \tilde{t}^{*B}_{RHC,greg}$. Next, we consider the average of these estimates to reach out to the final bootstrap estimate, say $\tilde{t}^{B}_{RHC,greg}$.

Therefore, the bootstrap variance estimate is given by, $\tilde{v}_b{}^{RHC} = \frac{1}{B-1}\sum_{b=1}^{B}\left(\tilde{t}^{*b}_{RHC,greg} - \tilde{t}^{B}_{RHC,greg}\right)^2$.

4. An alternative approach of re-scaling bootstrap technique for size-constrained adaptive sampling: Unit-wise bootstrap

An alternative re-scaling bootstrap technique for the size-constrained adaptive sampling has been suggested here which is termed as *unit-wise bootstrap*. For every observed sampled unit $\in s$, $B(i)$ is a subset of $A(i)$. The unit-wise bootstrap is performed for every subsampled network $B(i)$ of i^{th} unit. The procedure is elaborated in Secs. 4.1 and 4.2 for Horvitz and Thompson (1952) and Rao, Hartley, and Cochran (1962) methods of estimation, respectively.

4.1. Unit-wise bootstrap for Horvitz and Thompson (1952) method of estimation

The procedure is described as follows:

Step 1. From $B(i)$, a bootstrap sample of size l_i^* is selected by SRSWOR.
Step 2. Then,

$\tilde{y}_{jWOR} = e_i + \left(\frac{l_i^*}{l_i-1}\right)^{\frac{1}{2}}\left(1-\frac{l_i}{m_i}\right)^{\frac{1}{2}}(y_j^* - e_i)$ is computed, denoting y_j^* as the value of the selected bootstrap unit, $i = 1, 2, .n$ and $j = 1, 2, \ldots, \; l_i^*$.

Also, compute $e_{iWOR} = \frac{1}{l_i^*}\sum_{j=1}^{l_i^*} \tilde{y}_{jWOR} \; \forall i.$

Undoubtedly, $E_*(e_{iWOR}) = E(e_i)$ and $V_*(e_{iWOR}) = v(e_i)$ which can be written as,

$$V_*(e_{jWOR}) = V_*\left(\frac{1}{l_i^*}\sum_{j=1}^{l_i^*} \tilde{y}_{jWOR}\right)$$

$$= \frac{1}{l_i^{*2}}.l_i^* \frac{l_i^*}{l_i-1}\left(1-\frac{l_i}{m_i}\right)\sum_{j=1}^{l_i}(y_j-e_i)^2\frac{1}{l_i}$$

$$= \frac{1}{l_i-1}\left(\frac{1}{l_i}-\frac{1}{m_i}\right)\sum_{j=1}^{l_i}(y_j-e_i)^2.$$

The choice of l_i^* depends on l_i. Following Rao and Wu (1988), if $l_i > 3$ the choice of l_i^* may be obtained matching bootstrap third-order moment with the unbiased sample-based third-order moment estimate implying $l_i^* = \frac{(l_i-2)^2}{(l_i-1)} \simeq l_i - 3$. We use $l_i^* = l_i - 1$ if $1 < l_i \leq 3$.

Step 3. Now following Horvitz and Thompson (1952) method of estimation, the bootstrap estimate for $\sum_s \frac{e_i}{\pi_i}$ may be defined as $\tilde{t}_{HT}^{p*} = \sum_{i=1}^{n}\frac{e_{iWOR}}{\pi_i}$.

Here $E_P, E_R, \; E_B$ denote the expectation operators with respect to design, subsampling, and bootstrapping, respectively while V_P, V_R, V_B denote the variance operators, accordingly. The final variance of this estimator can be written as, $V\left(\tilde{t}_{HT}^{p*}\right) = E_pE_RV_B\left(\tilde{t}_{HT}^{p*}\right) + E_pV_RE_B\left(\tilde{t}_{HT}^{p*}\right) + V_pE_RE_B\left(\tilde{t}_{HT}^{p*}\right)$; that is,

$$V\left(\tilde{t}_{HT}^{p*}\right) = 2\sum_{i=1}^{N}\frac{V_R(e_i)}{\pi_i} + \sum_{i<j}\sum_{=1}^{N}(\pi_i\pi_j - \pi_{ij})\left(\frac{t_i}{\pi_i}-\frac{t_j}{\pi_j}\right)^2$$

An unbiased estimator of the $V\left(\tilde{t}_{HT}^{p*}\right)$ may be written as,

$$v\left(\tilde{t}_{HT}^{p*}\right) = 2\sum_{i\in s}\frac{v_R(e_i)}{\pi_i^2} + \sum_{i<j}\sum_{\in s}\left(\frac{\pi_i\pi_j-\pi_{ij}}{\pi_{ij}}\right)\left(\frac{e_i}{\pi_i}-\frac{e_j}{\pi_j}\right)^2$$

Step 4. For the other three linear estimators $\sum_s \frac{c_i'}{\pi_i}$, $\sum_s e_ic_i'w_i'$ and $\sum_s c_i'^2w_i'$, the bootstrap estimates can be derived as described above. The final bootstrap estimate for the greg estimator will be the same non-linear function *"f"* (defined in Sec. 3) of the related bootstrap estimates of the linear totals.
Step 5. The whole procedure is replicated a large number of times say, B times. Denoting $\tilde{t}_{HT.greg}^{pb}$ as b^{th} bootstrap estimate, the final bootstrap estimate may be defined as,

$$t^P_{HT,greg} = \frac{1}{B}\sum\nolimits_{b=1}^{B} \tilde{t}^{pb}_{HT,greg}.$$

Therefore, the related bootstrap variance estimate is,

$$\tilde{v}^P_{b,HT} = \frac{1}{B-1}\sum_{b=1}^{B}(\tilde{t}^{pb}_{HT,greg} - t^P_{HT,greg})^2.$$

4.2. Unit-wise bootstrap for Rao, Hartley, and Cochran (1962) method of estimation

As described in Sec. 4.1, we draw a bootstrap sample from $B(i)\ \forall\ i = 1,2,\ldots,n$. Here, ***Step 1*** and ***Step 2*** will be the same as described in the previous section. In ***Step 3***, the bootstrap estimate for $\sum_s \frac{e_i}{p_i/Q_i}$, using Rao, Hartley, and Cochran (1962) method of estimation may be defined as,

$$\tilde{t}^{p*}_{RHC} = \sum_{i=1}^{n}\frac{e_{iWOR}Q_i}{p_i}.$$

The related variance may be defined as,

$$V\left(\tilde{t}^{p*}_{RHC}\right) = E_pE_RV_B\left(\tilde{t}^{p*}_{RHC}\right) + E_pV_RE_B\left(\tilde{t}^{p*}_{RHC}\right) + V_pE_RE_B\left(\tilde{t}^{p*}_{RHC}\right);$$

It reduces to

$$V\left(\tilde{t}^{p*}_{RHC}\right) = 2\sum_{i=1}^{N}\frac{Q_iV_R(e_i)}{p_i} + \frac{\sum_{i=1}^{n}N_i^2 - N}{N^2 - \sum_{i=1}^{n}N_i^2}\left(\sum_{i=1}^{n}Q_i\left(\frac{t_i}{p_i} - t_{RHC}\right)^2\right);$$

and the related unbiased variance estimator is

$$v\left(\tilde{t}^{p*}_{RHC}\right) = 2\sum_{i=1}^{n}\frac{Q_i^2v_R(e_i)}{p_i^2} + \frac{\sum_{i=1}^{n}N_i^2 - N}{N^2 - \sum_{i=1}^{n}N_i^2}\left(\sum_{i=1}^{n}Q_i\left(\frac{e_i}{p_i} - t^*_{RHC}\right)^2\right).$$

Step 4 and ***Step 5*** will be the same as described in Sec. 4.1, with the necessary changes required for the RHC method of estimation. Denoting $\tilde{t}^{pb}_{RHC.greg}$ as b^{th} bootstrap estimate for the greg estimator, our final bootstrap estimate may be written as, $t^P_{RHC,greg} = \frac{1}{B}\sum_{b=1}^{B}\tilde{t}^{pb}_{RHC,greg}$ with related bootstrap variance estimate as,

$$\tilde{v}^P_{b,RHC} = \frac{1}{B-1}\sum_{b=1}^{B}(\tilde{t}^{pb}_{RHC,greg} - t^P_{RHC,greg})^2.$$

5. Simulation study

A numerical illustration is performed to examine the performances of the proposed re-scaling bootstrap methods discussed in Secs. 3 and 4. We have used here the data from "The National Wasteland Identification Project", available in the link https://dolr.gov.in/district-and-category-wise-wastelands-year-2000. District-wise wastelands due to salinity, degradation, water-logging, Marshy land are given there. We concentrate on the Southern region of India including six states (Andhra Pradesh, Kerala, Karnataka, Tamil Nadu, Goa, and Maharashtra) in estimating its salinity wasteland areas (in km^2). So, the salinity wasteland area (in km^2) is considered as the study

Table 1. Overall comparisons of the proposed bootstrap method.

	Horvitz Thompson			Rao Hartley Cochran		
Re-scaling Bootstrap	Relative Bias	SMSE	RRMSE	Relative Bias	SMSE	RRMSE
General Adaptive Sampling	−0.01118	12698.1	0.105243	0.087345	66263.56	0.240415
Size-constrained Adaptive Sampling	0.048065	21625.12	0.137342	0.064481	63933.024	0.236149
Size-constrained Adaptive Sampling (Alternative method)	0.104782	31077.28	0.164644	0.086732	65701.84	0.239394

Table 2. Overall comparisons of the proposed bootstrap by ACV, ACP, and AL.

	Horvitz Thompson			Rao Hartley Cochran		
Re-scaling Bootstrap	Average length	ACV (%)	ACP (%)	Average length	ACV (%)	ACP (%)
General Adaptive Sampling	8.9225 (9.0224)	0.4426	39.097 (39.098)	899.0191 (922.4029)	21.3718	88.235 (86.275)
Size-constrained Adaptive Sampling	22.4916 (23.0232)	1.7215	35.9677 (36.76)	878.2599 (933.1224)	22.3058	88.235 (92.157)
Size-constrained Adaptive Sampling (Alternative method)	265.6508 (298.3653)	6.4223	60.78 (62.745)	1542.466 (1782.6)	23.0034	96.078 (100)

variable (y). These six states comprise 86 districts (N) in total, 23 of which have such wasteland covering an area of 1070.72 $\text{km}^2 (Y)$. The wastelands due to upland and district-wise geographical area are considered as an auxiliary (x) and a size measure variables (z) respectively. The auxiliary variable(x) is highly correlated with the study variable (y). The initial sample is drawn by varying probability sampling scheme using the size measure variable (z). The total area under upland (X) is 31341.27 km^2.

First, an initial sample of size 19 (n) is drawn by Lahiri (1951), Midzuno (1951), and Sen (1953) sampling scheme. Considering the presence of salinity wasteland as a rare unit, the adaptive sampling procedure is used to enhance more rare units in the initial sample. For each of the initially sampled unit i, its network $A(i)$ is found. The performances of the bootstrap procedures are measured using the following criteria and shown in Tables 1 and 2. The criteria are (a) Average coefficient of variation (ACV), (b) Average coverage percentage (ACP), (c) Relative Bias (RB), (d) Simulated mean square error (SMSE), and e) Relative root mean square error (RRMSE).

For this, a large number of replicated samples (say R = 10,000) have been drawn to judge the efficacy of our proposed procedures. The RB, SMSE, RRMSE may be written as,

$$RB(\hat{t}_{greg}) = \frac{1}{R}\sum_{j=1}^{R}\frac{(\hat{t}_{greg}^{(j)} - \tau)}{\tau}, \quad SMSE = \frac{1}{R}\sum_{j=1}^{R}(\hat{t}_{greg}^{(j)} - \tau)^2 \text{ and}$$

$$RRMSE = \frac{\sqrt{\frac{1}{R}\sum_{j=1}^{R}(\hat{t}_{greg}^{(j)} - \tau)^2}}{\tau}$$

where $\hat{t}_{greg}^{(j)}$ is the estimate of the population parameter (τ) obtained for j^{th} replicate, using any bootstrap method discussed in Secs. 3 and 4. ACV is the average over $R = 10{,}000$ replicates of estimated CV's, i.e., $100.\frac{1}{R}\sum_R \frac{\sqrt{Estimated\ \ MSE(estimate)}}{estimate}$. Usually, ACV less than 10% indicates an excellent estimator of t_{greg} and at most 30% is acceptable. The 95% confidence intervals (CI) are calculated using the percentile method and with normal approximations. The percentage number of times the CI covers Y is called ACP.

To get the confidence interval (CI) based on the percentile method, lower and upper 2.5% points are taken from the histogram constructed with the values $\tilde{t}^{1}_{*,greg}, \tilde{t}^{2}_{*,greg}, \ldots \tilde{t}^{B}_{*,greg}$; $B = 10{,}000$

Table 3. Performance of proposed bootstrap in RHC method of estimation using percentile (normal) method.

Simulation Number 500					
Re-scaling Bootstrap	Bootstrap Estimate	Sqrt. MSE	Lower 2.5% point	Upper 2.5% point	Length
General Adaptive Sampling	999.2539	269.08	482.5308 (471.8572)	1510.034 (1526.651)	1027.503 (1054.793)
Size-constrained Adaptive Sampling	1077.181	207.815	697.0772 (669.8636)	1456.3006 (1484.498)	759.2234 (814.6349)
Size-constrained Adaptive Sampling (Alternative method)	1137.164	382.3698	478.2055 (387.7191)	1607.724 (1886.609)	1129.519 (1498.89)

Table 4. Performance of proposed bootstrap with HT method of estimation using percentile (normal) method.

Simulation Number 500					
Re-scaling Bootstrap	Bootstrap Estimate	Sqrt. MSE	Lower 2.5% point	Upper 2.5% point	Length
General Adaptive Sampling	1074.191	2.2942	1069.986 (1069.694)	1078.78 (1078.688)	8.794 (8.9932)
Size-constrained Adaptive sampling	1061.982	7.2429	1049.252 (1047.786)	1079.939 (1076.178)	30.687 (28.392)
Size-constrained Adaptive Sampling (Alternative method)	1165.013	70.94305	1037.048 (1025.965)	1267.424 (1304.061)	230.376 (278.096)

for the entire sample, drawn from the population, and before doing this, make sure that the process is simulated a large number of times. For each simulation, the CI $(l_{0.025}(r), u_{0.975}(r))$ is evaluated. Tables 3 and 4 show performance reports of a particular simulation.

In Table 1, it is seen that the HT method of estimation is more efficient than the RHC method of estimation, in terms of RB, SMSE, RRMSE. Also, the performances based on ACV and AL (in Table 2) indicate that the HT method of estimation is more efficient than the RHC method of estimation. But the comparison based on ACP contradicts the above consideration.

The performances of re-scaling bootstrap methods, proposed in Secs. 3.2 and 4 can be compared using the second and third row of Tables 1 and 2. In terms of RRMSE and ACV, no such significant differences are observed for these two proposed methods, mentioned in Secs. 3.2.2 and 4.2, for the RHC method of estimation. But the ACP value is improved significantly for the unit-wise bootstrap method, irrespective of the choices of the methods of estimation.

6. Concluding remarks

In this study, the modifications of re-scaling bootstrap methods for general adaptive sampling and size-constrained adaptive sampling are proposed for Horvitz and Thompson (1952) and Rao, Hartley, and Cochran (1962) methods of estimation. Through the simulation study, it is observed that the coverage probabilities for the Horvitz and Thompson (1952) method of estimation are extremely low. In terms of RB, SMSE, RRMSE, AL, and ACV, the HT method of estimation is better than the RHC method of estimation. Also, we have derived the unit-wise bootstrap technique for size-constrained adaptive sampling (in Sec. 4) which gives us satisfactory results in terms of ACP and ACV values for both methods of estimation.

So, it may be concluded that the efficacy of re-scaling bootstrap procedure in general adaptive sampling and size-constrained adaptive sampling for the HT method of estimation is more efficient than the RHC method of estimation in terms of ACV, AL, RB, RRMSE, and SMSE. But the performance in terms of ACP value is unsatisfactory for the HT method of estimation. It is noteworthy that Perez and Pontius (2006) also worked on bootstrap procedures for Thompson's adaptive sampling design and pointed out the same issue regarding the coverage rate for the HT method of estimation. But the proposed unit-wise bootstrap technique for the size-constrained adaptive sampling procedure improves the ACP value significantly, for both HT and RHC methods of estimation. Also, the implementation of the proposed unit-wise bootstrap procedure is easier than the methods as described in Sec. 3.2.

Acknowledgments

The authors are grateful to the editor and the referees for their valuable comments and suggestions.

ORCID

Sanghamitra Pal http://orcid.org/0000-0001-9106-4402
Dipika Patra http://orcid.org/0000-0003-4318-1123

References

Biswas, A., A. Rai, and T. Ahmad. 2020. Rescaling bootstrap technique for variance estimation for ranked set samples in finite population. *Communications in Statistics-Simulation and Computation*. 49 (10):2704–18. doi:10.1080/03610918.2018.1527349.

Brown, J. A. 1994. The application of adaptive cluster sampling to ecological studies. *Statistics in Ecology and Environmental Monitoring*, 2:86–97.

Brown, J. A., and B. J. Manly. 1998. Restricted adaptive cluster sampling. *Environmental and Ecological Statistics* 5 (1):49–63. doi:10.1023/A:1009607403647.

Cassel, C. M., C. E. Sarndal, and J. H. Wretman. 1976. Some results on generalized difference estimation and generalized regression estimation for finite population. *Biometrika* 63 (3):615–20. doi:10.1093/biomet/63.3.615.

Chao, C. T. 2004. Ratio estimation on adaptive cluster sampling. *Journal of Chinese Statistical Association* 42 (3): 307–27.

Chaudhuri, A. 2000. Network and adaptive sampling for unequal probabilities. *Calcutta Statistical Association Bulletin* 50 (3-4):237–53. doi:10.1177/0008068320000310.

Chaudhuri, A. 2010. *Essentials of survey sampling*. New Delhi: Prentice Hall of India.

Chaudhuri, A., M. Bose, and K. Dihidar. 2005. Sample size- restrictive adaptive sampling: An application in estimating localized elements. *Journal of Statistical Planning and Inference* 134 (1):254–67. doi:10.1016/j.jspi.2004.02.016.

Chaudhuri, A., M. Bose, and J. K. Ghosh. 2004. An application of adaptive sampling to estimate highly localized population segments. *Journal of Statistical Planning and Inference* 121 (2):175–89. doi:10.1016/S0378-3758(03)00117-4.

Christman, M. C., and J. S. Pontius. 2000. Bootstrap confidence intervals for adaptive cluster sampling. *Biometrics* 56 (2):503–10. doi:10.1111/j.0006-341x.2000.00503.x.

Dryver, A. L., and C. T. Chao. 2007. Ratio estimators in adaptive cluster sampling. *Environmetrics* 18 (6):607–20. doi:10.1002/env.838.

Efron, B. 1982. *The jackknife, the bootstrap and other resampling plans*. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Hajek, J. 1964. Asymptotic theory of rejective sampling with varying probabilities from a finite population. *The Annals of Mathematical Statistics* 35 (4):1491–523. doi:10.1214/aoms/1177700375.

Hansen, M. M., and W. N. Hurwitz. 1943. On the theory of sampling from finite populations. *The Annals of Mathematical Statistics* 14 (4):333–62. doi:10.1214/aoms/1177731356.

Horvitz, D. G., and D. J. Thompson. 1952. A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association* 47 (260):663–85. doi:10.1080/01621459.1952.10483446.

Lahiri, D. B. 1951. A method of sample selection providing unbiased ratio estimates. *Bulletin of International Statistical Institute* 3 (2):133–40.

Midzuno, H. 1951. On the sampling system with probability proportional to the sum of the sizes. *Annals of the Institute of Statistical Mathematics* 3 (1):99–107. doi:10.1007/BF02949779.

Mohammadi, M., M. Salehi, and J. N. Rao. 2014. Bootstrap confidence intervals for adaptive cluster sampling design based on Horvitz-Thompson type estimators. *Environmental and Ecological Statistics* 21 (2):351–71. doi:10.1007/s10651-013-0258-z.

Pal, S. 2009. Rao and Wu's rescaling bootstrap modified to achieve extended coverage. *Journal of Statistical Planning and Inference* 139 (10):3552–58. doi:10.1016/j.jspi.2009.04.006.

Pal, S., and D. Patra. 2021. A prediction approach in adaptive sampling. *Metron*. Advance online publication. doi: 10.1007/s40300-020-00195-1.

Perez, D. T., and J. S. Pontius. 2006. Conventional bootstrap and normal confidence interval estimation under adaptive cluster sampling. *Journal of Statistical Computation and Simulation* 76 (9):755–64. doi:10.1080/10629360500109002.

Rao, J. N., H. O. Hartley, and W. C. Cochran. 1962. On a simple procedure of unequal probability sampling without replacement. *Journal of Royal Statistical Society B* 24:482–91.

Rao, J. N., and C. F. Wu. 1988. Resampling inference with complex survey data. *Journal of the American Statistical Association* 83 (401):231–41. doi:10.1080/01621459.1988.10478591.

Salehi, M. M., and G. A. Seber. 1997. Adaptive cluster sampling with networks selected without replacement. *Biometrika* 84:209–19.

Salehi, M. M., and G. A. Seber. 2002. Unbiased estimators for restricted adaptive cluster sampling. *Australian & New Zealand Journal of Statistics* 44 (1):63–74. doi:10.1111/1467-842X.00208.

Sen, A. R. 1953. On the estimator of the variance in sampling with varying probabilities. *Journal of Indian Society of Agricultural Statistics* 5:119–27.

Thompson, S. K. 1990. Adaptive cluster sampling. *Journal of the American Statistical Association* 85 (412):1050–59. doi:10.1080/01621459.1990.10474975.

Thompson, S. K. 1994. Factors influencing the efficiency of adaptive cluster sampling. Technical Report number 94-0301. Pennsylvania State University, Department of Statistics, State College, PA.