# BALANCED REPEATED REPLICATION

Jun Shao, University of Ottawa
Department of Mathematics, University of Ottawa, Ottawa, K1N 6N5 Canada

KEY WORDS: Variance estimation, stratified multistage sampling, nonsmooth estimates, imputation.

## 1. Introduction

Variance estimation is an important part of sample survey theory. The balanced repeated replication (BRR) method, together with the jackknife and the linearization methods, are the most popular methods used in sample surveys. The purpose of this article is to study asymptotic properties of the BRR method.

We start with an introduction of the adopted sampling design. Nowadays, sample surveys often use a stratified multistage sampling (Kish and Frankel, 1974; Krewski and Rao, 1981). The population under consideration has been stratified into $L$ strata with $N_h$ clusters in the $h$th stratum. Within the $i$th cluster in stratum $h$, there are second, third,..., stage units, and $N_{hi}$ ultimate units. Associated with the $j$th ultimate unit in the $i$th cluster of stratum $h$ is a vector of characteristics $Y_{hij}$, $j = 1, \ldots, N_{hi}$, $i = 1, \ldots, N_h$, $h = 1, \ldots, L$. The finite population distribution function is then given by

$$F = \frac{1}{N} \sum_{h=1}^{L} \sum_{i=1}^{N_h} \sum_{j=1}^{N_{hi}} \delta_{Y_{hij}},$$

where $N = \sum_{h=1}^{L} \sum_{i=1}^{N_h} N_{hi}$ is the total number of ultimate units in the population and $\delta_x$ is the distribution function degenerated at $x$.

For each $h$, $n_h \geq 2$ clusters are selected from stratum $h$ using probability sampling with replacement, independently across the strata. Within the $(h,i)$th first-stage cluster, $n_{hi}$ ultimate units are sampled according to some multistage sampling methods, $i = 1, \ldots, n_h$, $h = 1, \ldots, L$. Let $\{y_{hij}, j = 1, \ldots, n_{hi}, i = 1, \ldots, n_h, h = 1, \ldots, L\}$ be the observations from the sampled ultimate units and $w_{hij}$ be survey weights associated with the $y_{hij}$. We assume that the survey weights are constructed so that

$$H = \frac{1}{N} \sum_{h=1}^{L} \sum_{i=1}^{n_h} \sum_{j=1}^{n_{hi}} w_{hij} \delta_{y_{hij}}$$

is unbiased for the population distribution $F$. However, $H$ may not be a distribution function since $H(\infty)$ is not necessarily equal to one. Furthermore, in many cases $N$ is unknown. Thus, we estimate $N$ by $\hat{N} = \sum_{h=1}^{L} \sum_{i=1}^{n_h} \sum_{j=1}^{n_{hi}} w_{hij}$ and obtain a distribution estimator

$$\hat{F} = \frac{1}{\hat{N}} \sum_{h=1}^{L} \sum_{i=1}^{n_h} \sum_{j=1}^{n_{hi}} w_{hij} \delta_{y_{hij}}.$$

In most survey problems, the parameter of interest $\theta$ is a known functional of $N$ and $F$ and a survey estimate of $\theta$ is obtained by replacing $N$ and $F$ with $\hat{N}$ and $\hat{F}$, respectively. In most cases, the resulting $\hat{\theta}$ can be written as

$$\hat{\theta} = T(\hat{Z}), \quad \hat{Z} = \sum_{h=1}^{L} \sum_{i=1}^{n_h} \sum_{j=1}^{n_{hi}} w_{hij} z_{hij}, \qquad (1.1)$$

with a known functional $T$, where $z_{hij}$ is an appropriately defined vector of data from the $(h,i,j)$th ultimate sample unit. For example, $T = g$, $z_{hij} = (1, y'_{hij})'$ and $\hat{\theta} = g(\hat{Z})$ gives estimates for ratios, correlation coefficients, and regression coefficients; $z_{hij} = (1, \delta_{y_{hij}})'$, $T(\hat{Z}) = (Z_2/Z_1)^{-1}(p)$, where $Z_1$ and $Z_2$ are the first and second components of $\hat{Z}$, and $\hat{\theta} = T(\hat{Z}) = \hat{F}^{-1}(p)$ is the survey estimator of $\theta = F^{-1}(p)$.

When $\hat{\theta} = g(\hat{Z})$, the linearization method produces the following variance estimator (Rao, 1988):

$$v_{\mathrm{L}} = \sum_{h=1}^{L} \frac{1}{n_h} \nabla g(\hat{Z})' s_h^2 \nabla g(\hat{Z}), \tag{1.2}$$

where $s_h^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (z_{hi} - \bar{z}_h)(z_{hi} - \bar{z}_h)'$,

$$z_{hi} = \sum_{j=1}^{n_{hi}} n_h w_{hij} z_{hij}, \quad \text{and} \quad \bar{z}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} z_{hi}. \tag{1.3}$$

The linearization method provides a consistent estimator of the asymptotic variance of $\hat{\theta} = g(\hat{Z})$ (Krewski and Rao, 1981; Bickel and Freedman, 1984). But it requires a separate derivation of the derivatives for each $g$.

When $\hat{\theta}$ is a sample quantile, the linearization is not applicable for deriving a variance estimator. Rao and Wu (1987) obtained a variance estimator by equating Woodruff's interval to a normal theory interval.

We now describe the BRR method which was first proposed by McCarthy (1969) for the case where $n_h \equiv 2$. A set of $R$ replicates (subsets of first-stage sample clusters) is formed in a balanced manner. For the $h$th stratum, let $\mathbf{s}_{rh} \subset \{1, ..., n_h\}$ be the indices of the first-stage sample clusters in the $r$th replicate, where the size of $\mathbf{s}_{rh}$ is $m_h$. When $L$ is large and all $n_h$ are small, a simple and effective choice of $m_h$ is $m_h = 1$ for all $h$. The set $\{\mathbf{s}_{rh} : r = 1, \ldots, R, h = 1, \ldots, L\}$ constitutes a BRR if for fixed $h$ and $h'$, the number of elements in $\{r : i \in \mathbf{s}_{rh}, i' \in \mathbf{s}_{rh}\}$ does not depend on $i$ and $i'$ $(i \neq i')$; and the number of elements in $\{r : i \in \mathbf{s}_{rh}, i' \in \mathbf{s}_{rh'}\}$ does not depend on $i$ and $i'$. A trivial BRR is {all possible subsets of $\{1, \ldots, n_h\}$ of size $m_h$, $h = 1, \ldots, L\}$, in which case $R = \prod_{h=1}^{L} \binom{n_h}{m_h}$.

In the simple case where $n_h = 2$ and $m_h = 1$ for all $h$, the BRR can be constructed using Hadamard matrices and a BRR variance estimator is (McCarthy, 1969)

$$v_{\mathrm{BRR}} = \frac{1}{R} \sum_{r=1}^{R} \left( \hat{\theta}^{(r)} - \frac{1}{R} \sum_{r=1}^{R} \hat{\theta}^{(r)} \right)^2, \tag{1.4}$$

where $\hat{\theta}^{(r)} = T(\hat{Z}^{(r)})$, $\hat{Z}^{(r)} = \sum_{h=1}^{L} \bar{z}_h^{(r)}$, $\bar{z}_h^{(r)} = \frac{1}{m_h} \sum_{i \in \mathbf{s}_{rh}} z_{hi}$, and $z_{hi}$ is as given in (1.3) (also, see (1.1)). In general, if we define the BRR variance estimator using (1.4), then in the case of $\hat{\theta} = g(\hat{Z})$ with linear $g(x) = c'x$, the BRR variance estimator is

$$\frac{1}{R} \sum_{r=1}^{R} \left( \hat{\theta}^{(r)} - \frac{1}{R} \sum_{r=1}^{R} \hat{\theta}^{(r)} \right)^2 = \sum_{h=1}^{L} \frac{n_h - m_h}{m_h n_h} c' s_h^2 c$$

(by the balance property of the BRR), which does not agree with the unbiased and consistent variance estimator $v_{\mathrm{L}}$ in (1.2), unless $m_h = n_h/2$ for all $h$. Hence, some modification has to be made. Wu (1991) considered a rescaling adjustment and derived a BRR variance estimator

$$v_{\mathrm{BRR}} = \frac{1}{R} \sum_{r=1}^{R} \left( \tilde{\theta}^{(r)} - \frac{1}{R} \sum_{r=1}^{R} \tilde{\theta}^{(r)} \right)^2, \tag{1.5}$$

where $\tilde{\theta}^{(r)} = T(\tilde{Z}^{(r)})$, $r = 1, \ldots, R$, and

$$\tilde{Z}^{(r)} = \sum_{h=1}^{L} \left[ \sqrt{\tfrac{m_h}{n_h - m_h}} \bar{z}_h^{(r)} + \left( 1 - \sqrt{\tfrac{m_h}{n_h - m_h}} \right) \bar{z}_h \right]. \tag{1.6}$$

The adjustment in (1.6) ensures that in the case of $\hat{\theta} = c'\hat{Z}$, $v_{\mathrm{BRR}}$ in (1.5) reduces to $v_{\mathrm{L}}$.

For the case of $\hat{\theta} = \hat{F}^{-1}(p)$, the coefficients in front of $\bar{z}_h^{(r)}$ and $\bar{z}_h$ in formula (1.6) must be nonnegative in order for $\tilde{F}^{(r)}$ (the adjusted estimator of the population distribution function) being a proper distribution function. This requires

$$m_h \leq n_h/2 \quad \text{for all } h. \tag{1.7}$$

It is easy to see that the estimator in (1.4) is a special case of that in (1.5) when $n_h = 2$ and $m_h = 1$ for all $h$. In fact, as long as $m_h = n_h/2$, $\hat{Z}^{(r)}$ and $\hat{F}^{(r)}$ are the same as the adjusted $\tilde{Z}^{(r)}$ and $\tilde{F}^{(r)}$, respectively.

A convenient way of computing $\tilde{\theta}^{(r)}$ is to use the formula for the original estimator $\hat{\theta}$ with the weights $w_{hij}$ changing to

$$w_{hij}^{(r)} = \left( 1 + \sqrt{\tfrac{n_h - m_h}{m_h}} \right) w_{hij} \quad \text{if } i \in \mathbf{s}_{rh}$$

or

$$w_{hij}^{(r)} = \left( 1 - \sqrt{\tfrac{m_h}{n_h - m_h}} \right) w_{hij} \quad \text{if } i \notin \mathbf{s}_{rh}.$$

Condition (1.7) ensures that the new weights $w_{hij}^{(r)} \geq 0$. When $n_h = 2$ and $m_h = 1$ for all $h$, $w_{hij}^{(r)} = 2w_{hij}$ or 0.

Asymptotic properties of the BRR variance estimators are studied in the next section. In particular, the BRR variance estimators are consistent for both smooth estimators $\hat{\theta} = g(\hat{Z})$ and nonsmooth estimators such as sample quantiles. This is an advantage of the BRR over the jackknife or the linearization method, when one prefers to use a single method for both cases of smooth and nonsmooth $\hat{\theta}$.

To compute the BRR variance estimator, it is desired to find a BRR with $R$ as small as possible. In the general case of $n_h \geq 2$, however, the construction of a BRR with a feasible $R$ is much more difficult than in the case of $n_h = 2$ per stratum. In the case of $n_h = p > 2$ clusters per stratum for $p$ prime or power of prime, a BRR can be obtained by using orthogonal arrays of strength two (Gurney and Jewett, 1975), where each balanced sample is obtained by selecting one first-stage sample cluster from each stratum. Gupta and Nigam (1987) and Wu (1991) obtained BRR with $m_h = 1$ in the case of unequal $n_h$, using mixed level orthogonal arrays of strength two to construct balanced replicates. More methods for constructing BRR can be found in Sitter (1993). In Section 3 we consider some approximated BRR which can be easily obtained when an exact BRR is difficult to construct.

## 2. Asymptotic Properties of the BRR

An asymptotic framework is provided by assuming that the finite population under study is a member of a sequence of finite populations indexed by $k = 1, 2, \ldots$ Thus, the quantities $L$, $N$, $N_h$, $N_{hi}$, $Y_{hij}$, $F$, $\theta$, $n_h$, $n_{hi}$, $y_{hij}$, $w_{hij}$, $\hat{F}$, $\hat{\theta}$, $v_{\mathrm{L}}$, $v_{\mathrm{BRR}}$, etc., depend on the population index $k$, but, for simplicity of notation, $k$ will be suppressed in what follows. All limiting process, however, will be understood to be as $k \to \infty$. Note that the parameter of interest $\theta$ is not fixed as $k$ increases; but we always assume that $\{\theta, k = 1, 2, \ldots\}$ is a bounded set.

Let $n$ be the number of first-stage sampled clusters. It is assumed that $n \to \infty$ as $k \to \infty$. Also, without loss of generality, we assume that for each $k$, there is a set $\mathcal{H}_k \subset \{1, \ldots, L\}$ (note that $L$ depends on $k$) such that

$$\sup_{h \in \mathcal{H}_k,\, k=1,2\ldots} n_h < \infty; \quad \min_{h \notin \mathcal{H}_k} n_h \to \infty. \qquad (2.1)$$

Note that (2.1) includes the following two common situations in surveys: (1) all the $n_h$ are small (bounded by a constant), in which case $\mathcal{H}_k = \{1, \ldots, L\}$; (2) all the $n_h$ are large, in which case $\mathcal{H}_k = \emptyset$.

It is assumed that no survey weight is disproportionately large, i.e.,

$$\max_{h,i,j} n_{hi} w_{hij} / N = O(n^{-1}). \qquad (2.2)$$

Under this assumption, $\hat{F}$ is consistent for $F$ and is asymptotically normal.

To make the asymptotic treatment simpler, we redefine $\hat{Z}$ and $z_{hi}$ as follows: $\hat{Z}$ and $z_{hi}$ are still as defined in (1.1) and (1.3), respectively, but with $w_{hij}$ replaced by $w_{hij}/N$. This change of notation does not have any effect when $\hat{\theta} = T(\hat{F})$ or $\hat{\theta} = g(\hat{Z})$ is proportional to $g(N\hat{Z})$ (e.g., $g(\hat{Z}) = c'\hat{Z}$). With these redefined $\hat{Z}$ and $z_{hi}$, we assume that

$$0 < c_1 \leq n \mathrm{var}(\hat{Z}) \leq c_2 < \infty \qquad (2.3)$$

for all $n$, where $c_1$ and $c_2$ are some constants. We also assume that the size of the BRR satisfies $R/n^2 \to 0$ and that $0 < \epsilon_0 \leq \frac{m_h}{n_h} \leq \frac{1}{2}$ for all $h$.

When $\hat{\theta} = g(\hat{Z})$, the consistency of $v_{\mathrm{BRR}}$ was established by Krewski and Rao (1981). The following theorem extends the result in Rao and Wu (1985).

**Theorem 1.** *Assume the conditions previously stated and*

$$\sum_{h=1}^{L} \sum_{i=1}^{n_h} E \left\| \frac{z_{hi} - E z_{hi}}{n_h} \right\|^4 = O\left(\frac{1}{n^3}\right). \qquad (2.4)$$

*Suppose further that the function $g$ is twice continuously differentiable with nonzero $\nabla g$ in a compact set containing $\{\mu, k = 1, 2, \ldots\}$, where $\mu = E\hat{Z}$. Then $v_{\mathrm{BRR}}/v_{\mathrm{L}} = 1 + O_p(n^{-1/2})$.*

**Proof.** Under condition (2.3), it suffices to show that $v_{\mathrm{BRR}} - v_{\mathrm{L}} = O_p(n^{-3/2})$. From the second order differentiability of $g$,

$$\tilde{\theta}^{(r)} = \hat{\theta} + l^{(r)} + q^{(r)}$$

with $l^{(r)} = (\tilde{Z}^{(r)} - \hat{Z})'\nabla g(\hat{Z})$ and $q^{(r)} = \frac{1}{2}(\tilde{Z}^{(r)} - \hat{Z})'\nabla^2 g(\xi^{(r)})(\tilde{Z}^{(r)} - \hat{Z})$. Using condition (2.4), the balance property of the replicates and the fact that $\max_{r=1,\ldots,R} \|\tilde{Z}^{(r)} - \mu\| \to_p 0$, we obtain that

$$v_{\mathrm{BRR}} = v_{\mathrm{L}} + \frac{2}{R}\sum_{r=1}^{R} l^{(r)} q^{(r)} + O_p(n^{-2}).$$

Then the result follows from

$$\frac{1}{R}\sum_{r=1}^{R} l^{(r)} q^{(r)} = O_p(n^{-3/2}). \quad \Box$$

In the case where $\hat{\theta} = \hat{F}^{-1}(p)$, a sample quantile, some "differentiability" condition on the population distribution function $F$ is required. Although $F$ is not differentiable for each fixed $k$, we may assume that $F$ is differentiable in the following limiting sense: there exists a sequence of functions $\{f_k(\cdot), k = 1, 2, \ldots\}$ such that $0 < \inf_k f(\theta) \le \sup_k f(\theta) < \infty$ and

$$\lim_{k\to\infty}\left[\frac{F(\theta + O(n^{-1/2})) - F(\theta)}{O(n^{-1/2})} - f(\theta)\right] = 0,$$

Note that the population index $k$ for $F$, $\theta$ and $f(\theta)$ is suppressed.

Let $\sigma^2 = \sigma_k^2$ be the asymptotic variance of $[F(\theta) - \hat{F}(\theta)]/f(\theta)$. It can be shown (Francisco and Fuller, 1991; Shao, 1994) that $(\hat{\theta} - \theta)/\sigma$ is asymptotically $N(0, 1)$.

It is known in this case that the jackknife variance estimator is inconsistent. The BRR variance estimator is still consistent. The following result was shown in Shao and Rao (1993).

**Theorem 2.** *Assume the conditions previously stated and that*

$$R = o\Big(\exp\{c\sum_{h=1}^{L} m_h\}\Big) \qquad (2.5)$$

*for a constant* $c > 0$. *Then* $v_{\mathrm{BRR}}/\sigma^2 \to_p 1$.

For $\hat{\theta} = \hat{F}[\frac{1}{2}\hat{F}^{-1}(\frac{1}{2})]$, the sample low income proportion, Shao and Rao (1993) showed that the BRR variance estimator is also consistent.

## 3. Approximations

Despite of the existence of several elegant methods of forming balanced replicates, the enumeration of the balanced replicates may require a separate software or involve nontrivial mathematical developments. A feasible BRR may not always exist for arbitrary $n_h$.

One easy but inefficient alternative is to randomly divide the $n_h$ clusters in stratum $h$ into two groups and then apply the BRR method by treating the two groups as two "clusters". This method, called the grouped BRR, requires $R$ repeated computations of the point estimator with $L \le R \le L + 4$. When $L$ is large, it provides a consistent variance estimator, although the efficiency of the grouped BRR variance estimator can be quite low if $L$ is much smaller than $n$ (see, e.g., Krewski, 1978). When $L$ is small and $n_h$ are large, however, the grouped BRR variance estimator may be inconsistent (Rao and Shao, 1993).

To increase the efficiency of the grouped BRR variance estimator, we may independently repeat the grouping $G$ times and take the average of the $G$ grouped BRR variance estimators. The resulting variance estimator is called the repeatedly grouped BRR and denoted by $v_{\mathrm{RGBRR}}$ (Rao and Shao, 1993). It requires $GR$ repeated computations of the point estimator. $v_{\mathrm{RGBRR}}$ is quite efficient if $G$ and $R$ are chosen so that $GR$ is comparable with $n$. Since $R$ has the same order as $L$, one must choose a large $G$ if $L$ is small. On the other hand, one may simply set $G = 1$ when $L$ is large.

**Theorem 3.** *Assume the conditions in Theorem 1 (or Theorem 2) and that*

$$\max_{h=1,\ldots,L} \tfrac{n_h}{nG} \to 0 \quad \text{and} \quad \tfrac{GR}{n^2} \to 0. \qquad (3.1)$$

*Then* $v_{\mathrm{RGBRR}}/\sigma^2 \to_p 1$, *where* $\sigma^2$ *is the asymptotic variance of* $\hat{\theta} = g(\hat{Z})$ *(or* $\hat{\theta} = \hat{F}^{-1}(p)$*).*

Shao and Wu (1992) introduced a random subsampling method. Let $\mathbf{s}_h$ denote a subset of

$\{1,\ldots,n_h\}$ of size $m_h$ and $\mathsf{S}$ be the collection of all possible elements of the form $(\mathsf{s}_1,\ldots,\mathsf{s}_L)$. Suppose that $\mathsf{S}=\mathsf{S}_1\cup\mathsf{S}_2\cup\cdots\cup\mathsf{S}_U$, where $\mathsf{S}_j$ are disjoint subsets of $\mathsf{S}$ of $\ell$ elements. Let $\{\mathsf{S}_j^*, j=1,\ldots,u\}$ be a simple random sample of size $u$ (with or without replacement) from $\{\mathsf{S}_j, j=1,\ldots,U\}$ ($u$ is usually much smaller than $U$). Let $(\mathsf{s}_1^{(r)},\ldots,\mathsf{s}_L^{(r)})$ be the $r$th element in the collection $\mathsf{S}_1^*\cup\mathsf{S}_2^*\cup\cdots\cup\mathsf{S}_u^*$, $r=1,\ldots,R=u\ell$. Then the random subsampling BRR variance estimator, denoted by $v_{\mathrm{RSBRR}}$, is defined by (1.5) with $\tilde{\theta}^{(r)}$ being the estimator based on the first-stage sampled clusters indexed by $(\mathsf{s}_1^{(r)},\ldots,\mathsf{s}_L^{(r)})$. If $\ell=1$, then this method amounts to taking a simple random sample of size $R=u$ from $\mathsf{S}$. Another special case for $n_h\equiv n_0$ and $m_h\equiv 1$ for all $h$ is to define $\mathsf{S}_j$ to be a collection of $n_0$ mutually exclusive subsamples, each of which contains $L$ first-stage sampled clusters with one from each stratum. Then each $\mathsf{S}_j$ amounts to grouping the $n_0L$ clusters into $n_0$ exclusive subsamples; there are $n_0^{L-1}$ such groupings to make up $\mathsf{S}$ with $U=n_0^{L-1}$. The random subsampling BRR in this special case is also called the repeated random-group method and is studied in Kovar, Rao and Wu (1988).

**Theorem 4.** *Assume the conditions in Theorem 3 with condition* (3.1) *replaced by* $R\to\infty$ *and* $R/n^2\to 0$. *Then* $v_{\mathrm{RSBRR}}/\sigma^2\to_p 1$.

## 4. BRR for Imputed Data

Most surveys have missing observations and use imputation for missing data. That is, if an observation $y_{hij}$ is missing, we replace it by a value $\eta_{hij}$ obtained under a given imputation rule. We consider the simple case where $y_{hij}$ is univariate and the units respond independently with the same probability. A commonly employed imputation rule, called the hot deck imputation (Kalton, 1981; Sedransk, 1985), imputes the missing values with a random sample from the respondents, where each respondent $y_{hij}$ is selected with (imputation) probability proportional to $w_{hij}$ (Rao and Shao, 1992). Let $a_{hij}$ be 1 if $y_{hij}$ is a respondent and $a_{hij}=0$ otherwise. Then the imputed estimate of $Y$ is

$$\hat{Y}_I=\sum w_{hij}[a_{hij}y_{hij}+(1-a_{hij})\eta_{hij}],$$

where $\sum$ is over all indices $(h,i,j)$ in the sample. Let $E_I$ be the imputation expectation. Then $E_I(\eta_{hij})=\sum w_{hij}a_{hij}y_{hij}/\sum w_{hij}a_{hij}$ and

$$\hat{Y}_I=\frac{\sum w_{hij}\sum w_{hij}a_{hij}y_{hij}}{\sum w_{hij}a_{hij}} + \sum w_{hij}(1-a_{hij})[\eta_{hij}-E_I(\eta_{hij})]. \quad (4.1)$$

Let $A$ and $B$ be the first and the second terms on the right side of (4.1). Then $(\hat{Y}_I-Y)/\sigma$ is asymptotically $N(0,1)$, where $\sigma^2$ is the sum of the asymptotic variances of $A$ and $B$.

Since $v_{\mathrm{BRR}}=v_{\mathrm{L}}$ in the linear case and $v_{\mathrm{L}}$ is inconsistent when the data are imputed (Rao and Shao, 1992), we obtain a modified BRR variance estimator for $\hat{Y}_I$ as follows. Let

$$\tilde{Y}_I^{(r)}=\sum w_{hij}^{(r)}[a_{hij}y_{hij}+(1-a_{hij})\eta_{hij}^{(r)}],$$

where $\eta_{hij}^{(r)}$ is an adjusted value of $\eta_{hij}$ given by

$$\eta_{hij}^{(r)}=\eta_{hij}+\frac{\sum w_{hij}^{(r)}a_{hij}y_{hij}}{\sum w_{hij}^{(r)}a_{hij}}-\frac{\sum w_{hij}a_{hij}y_{hij}}{\sum w_{hij}a_{hij}}.$$

The modified BRR variance estimator for $\hat{Y}_I$ is then

$$\tilde{v}_{\mathrm{BRR}}=\frac{1}{R}\sum_{r=1}^{R}\left(\tilde{Y}_I^{(r)}-\frac{1}{R}\sum_{r=1}^{R}\tilde{Y}_I^{(r)}\right)^2.$$

**Theorem 5.** *Assume the conditions previously stated. Then* $\tilde{v}_{\mathrm{BRR}}/\sigma^2\to_p 1$.

**Proof.** It can be shown that

$$\tilde{v}_{\mathrm{BRR}}=\frac{1}{R}\sum_{r=1}^{R}\left(\tilde{Y}_I^{(r)}-\hat{Y}_I\right)^2+o\left(\frac{1}{n}\right)$$
$$=\frac{1}{R}\sum_{r=1}^{R}A_r^2+\frac{1}{R}\sum_{r=1}^{R}B_r^2+o\left(\frac{1}{n}\right), \quad (4.2)$$

where

$$A_r=\frac{\sum w_{hij}^{(r)}\sum w_{hij}^{(r)}a_{hij}y_{hij}}{\sum w_{hij}^{(r)}a_{hij}}-A,$$

$$B_r=\sum w_{hij}^{(r)}(1-a_{hij})[\eta_{hij}-E_I(\eta_{hij})]-B,$$

and $A$ and $B$ are defined by (4.1). The result follows from the fact that the first and the second terms on the right side of (4.2) are standard BRR variance estimators for $A$ and $B$, respectively.

## Reference


Bickel, P. J. and Freedman, D. A. (1984). Asymptotic normality and the bootstrap in stratified sampling, *Ann. Statist.*, **12**, 470–482.

Francisco, C. A. and Fuller, W. A. (1991). Quantile estimation with a complex survey design, *Ann. Statist.*, **19**, 454–469.

Gupta, V. K. and Nigam, A. K. (1987). Mixed orthogonal arrays for variance estimation with unequal numbers of primary selections per stratum, *Biometrika*, **74**, 735–742.

Gurney, M. and Jewett, R. S. (1975). Constructing orthogonal replications for standard errors, *J. Amer. Statist. Assoc.*, **70**, 819–821.

Kalton, G. (1981). *Compensating for Missing Data*, ISR research report series. Ann Arbor: Survey Research Center, University of Michigan.

Kish, L. and Frankel, M. R. (1974). Inference from complex samples (with discussion), *J. R. Statist. Soc. B*, **36**, 1–37.

Kovar, J. G., Rao, J. N. K. and Wu, C. F. J. (1988). Bootstrap and other methods to measure errors in survey estimates, *Canadian J. Statist.*, **16**, Supplement, 25–45.

Krewski, D. (1978). On the stability of some replication variance estimators in the linear case, *J. Statist. Plan. Inf.*, **2**, 45–51.

Krewski, D. and Rao, J. N. K. (1981). Inference from stratified samples: properties of the linearization, jackknife and balanced repeated replication methods, *Ann. Statist.*, **9**, 1010–1019.

McCarthy, P. J. (1969). Pseudo-replication: half samples, *Rev. Internat. Statist. Inst.*, **37**, 239–264.

Rao, J. N. K. (1988). Variance estimation in sample surveys, *Handbook of Statistics*, P. K. Krishnaiah and C. R. Rao eds., vol. **6**, 427–447.

Rao, J. N. K. and Shao, J. (1992). Jackknife variance estimation with survey data under hot deck imputation, *Biometrika*, **79**, 811–822.

Rao, J. N. K. and Shao, J. (1993). On balanced half-sample variance estimation in stratified sampling, preprint.

Rao, J. N. K. and Wu, C. F. J. (1985). Inference from stratified samples: Second-order analysis of three methods for nonlinear statistics, *J. Amer. Statist. Assoc.*, **80**, 620–630.

Rao, J. N. K. and Wu, C. F. J. (1987). Methods for standard errors and confidence intervals from sample survey data: Some recent work, *Bull. Internat. Statist. Inst.*, *Proceedings of the 46th session*, **3**, 5–19.

Sedransk, J. (1985). The objective and practice of imputation, *Proceedings of First Annual Research Conference*, Bureau of the Census, Washington D.C., 445–452.

Shao, J. (1994). L-statistics in complex surveys, *Ann. Statist.*, to appear.

Shao, J. and Rao, J. N. K. (1993). Standard errors for low income proportions estimated from stratified multistage samples, preprint.

Shao J. and Wu, C. F. J. (1992). Asymptotic properties of the balanced repeated replication method for sample quantiles, *Ann. Statist.*, **20**, 1571–1593.

Sitter, R. R. (1993). Balanced repeated replications based on orthogonal multi-arrays, *Biometrika*, **80**, 211–221.

Wu, C. F. J. (1991). Balanced repeated replications based on mixed orthogonal arrays, *Biometrika*, **78**, 181–188.