Pak. J. Statist.
2002 Vol.18(3) pp 395-414

JACKKNIFING A GENERAL CLASS OF ESTIMATORS – A NEW APPROACH WITH REFERENCE TO VARYING PROBABILITY SAMPLING

Debesh Roy[1]
and
Md. Safiquzzaman[2]
Haldia Government College and Habra High School (H.S)

ABSTRACT

A new method of jackknifing have been suggested to take care of imbalance in the sample, caused either by varying probability of selection or by the varying sizes of the unit being deleted from the sample. This technique of jackknifing is applied to a general class of estimators. Considering a natural population the performance of this jackknife variance estimator has been compared with the other variance estimators.

KEY WORDS

Auxiliary variable, bootstrapping, generalised regression estimator, Horvitz-Thompson estimator, inclusion probability, jackknife, variance estimation.

1. INTRODUCTION

The idea of jackknifing was introduced by Quenouille (1956) in connection with reduction of bias of nonlinear estimators. The possibility of using this technique for the purpose of estimation of variance or mean square error was brought to the light by Tukey (1958). Durbin (1959) perhaps had been the first to use it in the context of finite population. Rao (1965) and Rao and Webster (1966) considered jackknifing classical ratio estimator. Srivastava (1967) defined a variant of the classical ratio estimator, called the Srivastava's modified ratio estimator (SMRE), that reduces in particular situation to classical ratio estimator.

[1] Department of Statistics, Haldia Government College, Haldia, Midnapore (East), West Bengal, INDIA.
[2] Habra High School(H.S.), Habra, North 24 Parganas, West Bengal, INDIA.

Nandi and Aich (1994) observed that a jackknifed version of SMRE induced an improvement and hence was worth considering. Besides classical and modified ratio estimators, many other estimators, which make use of auxiliary information are available in the literature. Some of them which are worth mentioning are generalised regression estimator due to Särndal (1980), asymptocally design unbiased (ADU) linear estimator due to Hajék (1971). Wright (1983) and Särndal and Wright (1984) brought all these estimators under the same umbrella of a general class of estimators, called QR-class of estimators. A slight modification of the QR-class, let it be called modified QR-class (MQR), that includes among others, SMRE and modified forms of estimators due to Särndal (1980), Brewer (1979, 1999) and Hajék (1971) was considered by Roy and Safiquzzaman (2000). On noting that the estimators of the said class are non-linear, they took up the problem of jackknifing the class with the objective of reducing its bias and estimating its variance. Since traditional method of jackknifing does not take into account the imbalance in the sample, caused by varying probability of selection, Roy and Safiquzzaman (2000) faced some problem with lack of stability of the variance estimator by traditional jackknife method.

Modification of standard jackknife procedure was suggested by Hinkley (1977). But his modification was aimed at taking care of the distance of every data point from the centre of the pool of data. He suggested a weighted jackknife in connection with general linear model, the essential idea of which was to scale individual residuals according to the relative importance of corresponding design points. The weighted jackknife technique of Hinkely (1977) was extended to ratio estimation by Chaudhry (1990), who perhaps had been the first to use weighted jackknife in connection with finite population sampling. But this weighted jackknife approach, while being considered for application to finite population sampling with varying probability of selection, takes care of the lack of balance induced only in a limited sense. While the lack of balance, as reflected by the distance of individual data points from a central value, is taken care of by this procedure, it fails to give due consideration to the same as reflected by varying probability of inclusion.

In this paper attempt has made to suggest an alternative weighted jackknife procedure that takes care of the lack of balance induced either by varying probabilities of selection or by varying sizes of the units being selected. Here an alternative weighted jackknife technique has been used to get the jackknife variance estimator of the MQR class. In a particular situation when the MQR estimator reduces to the Generalized Regression estimator (GREG), our weighted jackknife variance estimator reduces to a variance estimator of GREG which is very close to the TAY1 variance estimator proposed by Särndal (1982). To have an idea how our weighted jackknife variance estimator competes with the other variance estimators based on re-sampling techniques, we restored to a bootstrap sampling procedure, named population bootstrap, advocated among others by Davison and Hinkley (1998).

The contents of this article may be divided into five parts. In section **2** we introduce the QR and MQR class of estimators. In section **3** we introduce a new jackknife method and take up the problem of variance estimation by this method. To compare the weighted jackknife variance estimator obtained in section 3 with other variance estimators we introduce in section **4** the variance estimators obtained by linearisation technique and bootstrap re-sampling technique. Lastly in section **5** we consider a natural population of Swedish municipality, named MU284 in the book by Särndal, Swensson and Wretman (1992) to compare the performance of jackknife variance estimators obtained by the new modified technique vis-à-vis the traditional technique to demonstrate that the new method of jackknifing has an edge over the traditional method. The same natural population is used also to compare the variance estimator obtained by bootstrap method and it has been observed that our new method outperforms it by a clear margin.

2. NOTATION AND MODIFIED- QR (MQR) CLASS OF ESTIMATORS

Let U be the finite population on which are defined two real variables y and x taking values y_i and $x_i(>0, \text{known})$ with totals Y and X respectively. To estimate Y a sample s of size n is taken with probability p(s). The design p is assumed to admit positive inclusion probabilities π_i for the unit i and $\mathbf{p}_{ij}$ for the pair of units (i, j) of U. By $\sum_U$, $\sum\sum_U$ let us denote sums over i in U and i, j (i<j) in U and by $\sum_S$, $\sum\sum_S$ those in s respectively. Let $Q_i(>0)$, $R_i(\geq 0)$ and **a** be arbitrary constants. On noting that a simple linear estimator of the population total Y can be written as

$$\hat{Y}_S = \frac{1}{n}\sum_S R_i y_i$$

which with $R_i = N,$ reduces to $N\overline{y},$ we define MQR-class of estimators for the finite population total as

$$T_{MQR} = \hat{Y}_S + B_Q\left[\left(\left(\frac{X}{\hat{X}_S}\right)^{\alpha} - 1\right)\hat{X}_S\right]$$

$$= \left(\frac{1}{n}\sum_S R_i y_i\right) + B_Q\left[\left(\left(\frac{X}{\frac{1}{n}\sum_S R_i x_i}\right)^{\alpha} - 1\right)\frac{1}{n}\sum_S R_i x_i\right]$$

where $B_Q = \dfrac{\sum_S Q_i y_i}{\sum_S Q_i x_i}$.

When $\alpha = 1$, MQR-class of estimators reduces to the celebrated QR-class due to Wright (1983). It was shown by Särndal and Wright (1984) that QR-estimator (they called it QR-predictor as it was motivated from a model)

$$T_{QR} = \hat{Y}_s + B_Q\left(X - \hat{X}_s\right)$$

was asymptotically design unbiased (ADU) under the condition

$$1 - \frac{R_i \pi_i}{n} = \lambda Q_i \pi_i \, .$$

They also show that if this sufficient condition for being ADU was satisfied by a QR-estimator, then it was necessarily of the GREG form that can be obtained from T_{QR}, choosing $R_i = \dfrac{n}{\pi_i}$. From GREG form again we get in turn Hajék (1971) form with proper choice of Q_i and also the prediction form of Brewer (1999), called a cosmetically calibrated estimator by him, with proper choice of n.

As for T_{MQR} we see that its asymptotic design expectation, following Brewer (1979), is given by

$$\lim E_p T_{MQR} = \frac{1}{n}\Sigma_U R_i \mathbf{p}_i y_i + \frac{\Sigma_U Q_i \mathbf{p}_i y_i}{\Sigma_U Q_i \mathbf{p}_i x_i}\left[X^{\mathbf{a}}\left(\frac{1}{n}\Sigma_U R_i \mathbf{p}_i x_i\right)^{1-\mathbf{a}} - \frac{1}{n}\Sigma_U R_i \mathbf{p}_i x_i\right]$$

When $R_i = \frac{n}{\mathbf{p}_i}$, $\lim E_p T_{MQR} = Y$ for any choice of Q_i and $\mathbf{a}$. So hence forth we may restrict ourselves to modified GREG form

$$T_{MG} = \Sigma_S \frac{y_i}{\mathbf{p}_i} + \frac{\Sigma_S Q_i y_i}{\Sigma_S Q_i x_i}\left[\left(\frac{X}{\Sigma_S \frac{x_i}{\mathbf{p}_i}}\right)^{\mathbf{a}} - 1\right]\left(\Sigma_S \frac{x_i}{\mathbf{p}_i}\right)$$

whose linearised variance may be put in the form

$$V_{MG} = \Sigma\Sigma_U \mathbf{D}_{ij}\left(\frac{y_i - \mathbf{a}\,\mathbf{b}_Q x_i}{\mathbf{p}_i} - \frac{y_j - \mathbf{a}\,\mathbf{b}_Q x_j}{\mathbf{p}_j}\right)^2$$

where $\mathbf{D}_{ij} = \mathbf{p}_i \mathbf{p}_j - \mathbf{p}_{ij}$ and $\mathbf{b}_Q = \frac{\Sigma_U Q_i \mathbf{p}_i y_i}{\Sigma_U Q_i \mathbf{p}_i x_i}$.

Note that when looked upon as a function of $\mathbf{a}$, variance of T_{MG} may be written as

$$V_{MG}(\mathbf{a}) = V_p\left(\Sigma_S \frac{y_i}{\mathbf{p}_i}\right) - 2\mathbf{a}\,\mathbf{b}_Q C_p\left(\Sigma_S \frac{y_i}{\mathbf{p}_i}, \Sigma_S \frac{x_i}{\mathbf{p}_i}\right) + \mathbf{a}^2 \mathbf{b}_Q^2 V_p\left(\Sigma_S \frac{x_i}{\mathbf{p}_i}\right)$$

where

$$V_p\left(\Sigma_S \frac{u_i}{\mathbf{p}_i}\right) = \Sigma_U \mathbf{D}_{ij}\left(\frac{u_i}{\mathbf{p}_i} - \frac{u_j}{\mathbf{p}_j}\right)^2$$

and

$$C_p\left(\Sigma_s \frac{u_i}{\pi_i}, \Sigma_s \frac{v_i}{\pi_i}\right) = \Sigma\Sigma_U \Delta_{ij}\left(\frac{u_i}{\pi_i} - \frac{u_j}{\pi_j}\right)\left(\frac{v_i}{\pi_i} - \frac{v_j}{\pi_j}\right)$$

Now $V_{MG}^{//}(\alpha) = 2\beta_Q^2 V_p\left(\Sigma_s \frac{x_i}{\pi_i}\right) > 0$. So $V_{MG}(\alpha)$ is a convex function of α having a point of minimum at

$$\alpha = \frac{1}{\beta_Q} \frac{C_p\left(\Sigma_s \frac{y_i}{\pi_i}, \Sigma_s \frac{x_i}{\pi_i}\right)}{V_p\left(\Sigma_s \frac{x_i}{\pi_i}\right)} = \alpha_{opt} \text{ say.}$$

If the regression line of y on x is linear passing through the origin, only then $\alpha_{opt} = 1$. Though the value of the α_{opt} depends on the population values and hence unknown, we can estimate it by

$$\hat{\alpha}_{opt} = \frac{1}{\hat{\beta}_Q} \frac{\hat{C}_p\left(\Sigma_s \frac{y_i}{\pi_i}, \Sigma_s \frac{x_i}{\pi_i}\right)}{\hat{V}_p\left(\Sigma_s \frac{x_i}{\pi_i}\right)}$$

where $\hat{\beta}_Q$, $\hat{C}_p$ and $\hat{V}_p$ are obtained by replacing the population sums in β_Q, C_p and V_p by the respective Horvitz – Thompson estimators.

3. JACKKNIFING THE MODIFIED GREG PREDICTOR AND VARIANCE ESTIMATION BY WEIGHTED JACKKNIFE METHOD

To jackknife $T_{MG} = T_n$ (say) in traditional method, we denote by $T_{n-1}(-j)$ the value of T_n based on a sample of size (n – 1), when the jth pair (x_j, y_j) is deleted from the sample. Then a set of pseudo values are chosen as

$$T_j^* = nT_n - (n-1)T_{n-1}(-j) \quad : j \in s$$

and

$$T_{JK} = \frac{1}{n}\sum_{S} T_i^{*}$$

is proposed as the jackknifed estimator of the population total.

To get the variance estimator we note that T_{JK} is a mean of pseudo values and hence jackknife variance estimator is of the form

$$v_{JK} = \frac{\hat{\mathbf{s}}_T^{\,2}}{n} = \frac{1}{n(n-1)}\sum_{S}\left(T_i^{*} - T_{JK}\right)^2$$

But in varying probability sampling with units having wide diversity in their sizes, the effect of deletion of different units, while constructing pseudo values, appears to be different. In this connection let a unit of the sample, either having a large size measure or having a very high probability of inclusion, be designated as a '*heavy*' unit. In traditional set up, the construction of pseudo values does not take into consideration the lack of balance caused by deletion of a '*heavy*' unit and hence taking an unweighted average of those pseudo values may not be purposeful. So we propose to attach adequate weights to the present set of pseudo values in such a way that it gives a '*cushioning effect*' against the shock suffered due to deletion of a '*heavy*' unit while constructing pseudo values. Following this line of argument let us attach a weight function with each T_j^{*} to get a revised set of weighted pseudo values as

$$d_j = \mathbf{q}_j\left[nT_n - (n-1)T_{n-1}(-j)\right], \quad \left(\mathbf{q}_j \ge 0, \sum_U \mathbf{q}_j = 1\right) \ : j \in s$$

where $\mathbf{q}_j$'s are so chosen that they are proportional either to $\mathbf{p}_j$'s or to the sizes of the data points being deleted. Then we define the weighted jackknife estimator of the population total as

$$T_{JK}(w) = \frac{\sum_S \frac{d_i}{\mathbf{p}_i}}{\sum_S \frac{\mathbf{q}_i}{\mathbf{p}_i}} = \frac{\sum_S \frac{\mathbf{q}_i}{\mathbf{p}_i} T_i^{*}}{\sum_S \frac{\mathbf{q}_i}{\mathbf{p}_i}}$$

Here it is worth noting that $T_{JK}(w)$ is a weighted average of the normal pseudo values T_j^* $(j \in s)$ which reduces to T_{JK} when $\left(\frac{\mathbf{q}_j}{\mathbf{p}_j}\right)$ is constant for all $j \in s$.

To emphasize the role of $\mathbf{q}_j$ as a '*shock absorber*' we may speak in terms of Basu's (1971) elephant example. Let $\mathbf{q}_j$ be proportional to the current weight of the jth elephant in the sample. If '*Jumbo*' the heaviest elephant with small inclusion probability is the jth elephant selected in the sample against all odds, then $\left(\frac{\mathbf{q}_j}{\mathbf{p}_j}\right)$ plays the role of an excellent buffer against the '*erosion*' caused to T_j^* due to deletion of '*Jumbo*'.

Since $T_{JK}(w)$ is a Hajék type estimator, we may take the weighted jackknife variance estimator as

$$v_{JK}(w) = \sum\sum_S \frac{\mathbf{D}_{ij}}{\mathbf{p}_{ij}} \left(\frac{\mathbf{q}_i \mathbf{e}_i}{\mathbf{p}_i} - \frac{\mathbf{q}_j \mathbf{e}_j}{\mathbf{p}_j} \right)^2$$

where $\mathbf{e}_j = T_j^* - T_{JK}(w)$.

Now let us discuss in the following, three different choices of the weight. The first choice takes care of the imbalance caused by varying probabilities of selection, the second one is aimed at compensating the lack of balance caused by varying sizes of the units being deleted and the third is the combination of the two taking care of imbalance caused in either way.

***Choice* 1:** Here $\mathbf{q}_j$'s are chosen such that $\mathbf{q}_j \geq 0$, $\mathbf{q}_j \infty \mathbf{p}_j$ and $\sum_U \mathbf{q}_j = 1$. An obvious choice satisfying these three conditions is

$$\mathbf{q}_j = \frac{\mathbf{p}_j}{n}$$

and for this choice of $\mathbf{q}_j$, the pseudo values are

$$d_j = \frac{\pi_j}{n}\left[nT_n - (n-1)T_{n-1}(-j)\right]$$

For $T_n = \sum_S \frac{y_i}{\pi_i}$ and for a sampling design p(s) for which $\pi_j = np_j \left(p_j \geq 0, \sum_U p_i = 1\right)$ we must have $d_j = y_j$. Thus when $(p(s), T(s)) = (\pi ps, Y_{HT})$, the pseudo values are equal to the actual values of the sample.

***Choice* 2:** Here θ_j's are chosen such that $\theta_j \geq 0$, $\theta_j \propto X_j$ and $\sum_U \theta_j = 1$. An obvious choice satisfying these three conditions is

$$\theta_j = \frac{X_j}{X}$$

***Choice* 3:** Here we take

$$\theta_j = \frac{1}{2}\left(\frac{\pi_j}{n} + \frac{X_j}{X}\right)$$

so that $\theta_j \geq 0$ and $\sum_U \theta_j = 1$.

When $\pi_j \propto X_j$, the above three choices, for obvious reason, are same.

What ever may be the choice of θ_j, if $T_n = T_{MG}$, then assuming

$$\left|\frac{x_j / \pi_j}{\sum_S x_i / \pi_i}\right| < 1, \quad \left|\frac{Q_j x_j}{\sum_S Q_i x_i}\right| < 1 \text{ and } \left|\frac{Q_j y_j}{\sum_S Q_i y_i}\right| < 1$$

$$T_j^* = nT_n - (n-1)T_{n-1}(-j)$$

$$= n\left[\frac{g_{sj} e_j}{\pi_j} + (1-\alpha)\left(\frac{n-1}{n}\right)^{\alpha}\left(\frac{X}{\sum_S \frac{x_i}{\pi_i}}\right)^{\alpha} B_Q \frac{x_j}{\pi_j} + \left\{1 - \left(\frac{n-1}{n}\right)^{\alpha}\right\}\left(\sum_S \frac{x_i}{\pi_i}\right)\left(\frac{X}{\sum_S \frac{x_i}{\pi_i}}\right)^{\alpha}\right]$$

where $e_j = y_j - B_Q x_j$

$$g_{sj} = 1 + \left\{ \left(\frac{n-1}{n} \right)^{\mathbf{a}} \left(\frac{X}{\sum_S \frac{x_i}{\mathbf{p}_i}} \right)^{\mathbf{a}} - 1 \right\} \left(\sum_S \frac{x_i}{\mathbf{p}_i} \right) \frac{\mathbf{p}_j Q_j}{\sum_S Q_i x_i}$$

So the weighted jackknife estimator of the population total is

$$T_{JK}(w) = \frac{\sum_S \frac{d_i}{\mathbf{p}_i}}{\sum_S \frac{\mathbf{q}_i}{\mathbf{p}_i}} = \frac{\sum_S \frac{\mathbf{q}_i}{\mathbf{p}_i} T_i^*}{\sum_S \frac{\mathbf{q}_i}{\mathbf{p}_i}}$$

and the weighted jackknife variance estimator is

$$v_{JK}(w) = \sum \sum_S \frac{\mathbf{D}_{ij}}{\mathbf{p}_{ij}} \left(\frac{d_i - T_{JK}(w)\mathbf{q}_i}{\mathbf{p}_i} - \frac{d_j - T_{JK}(w)\mathbf{q}_j}{\mathbf{p}_j} \right)^2$$

$$= \sum \sum_S \frac{\mathbf{D}_{ij}}{\mathbf{p}_{ij}} \left(\frac{d_i}{\mathbf{p}_i} - \frac{d_j}{\mathbf{p}_j} \right)^2 + T_{JK}^{\ 2}(w) \sum \sum_S \frac{\mathbf{D}_{ij}}{\mathbf{p}_{ij}} \left(\frac{\mathbf{q}_i}{\mathbf{p}_i} - \frac{\mathbf{q}_j}{\mathbf{p}_j} \right)^2$$

$$- 2T_{JK}(w) \sum \sum_S \frac{\mathbf{D}_{ij}}{\mathbf{p}_{ij}} \left(\frac{d_i}{\mathbf{p}_i} - \frac{d_j}{\mathbf{p}_j} \right) \left(\frac{\mathbf{q}_i}{\mathbf{p}_i} - \frac{\mathbf{q}_j}{\mathbf{p}_j} \right)$$

***Note*:** When $\mathbf{a} = 1$ and $Q_i = w_i x_i$; T_{MG} reduces to GREG and in that case for a particular choice of the weight function, viz. $\mathbf{q}_j = \frac{\mathbf{p}_j}{n}$, weighted jackknife variance estimator reduces to

$$v_{JK}(w) = \sum \sum_S \frac{\mathbf{D}_{ij}}{\mathbf{p}_{ij}} \left(\frac{g_{si} e_i}{\mathbf{p}_i} - \frac{g_{sj} e_j}{\mathbf{p}_j} \right)^2$$

where

$$g_{sj} = 1 + \left(\frac{n-1}{n} X - \sum_{S} \frac{x_i}{\mathbf{p}_i} \right) \frac{\mathbf{p}_j w_j x_j}{\sum_{S} w_i x_i^{2}}$$

Note that this weighted jackknife variance estimator of GREG is very close to the variance estimator named TAY1 proposed by Särndal (1982).

4. VARIANCE ESTIMATION BY OTHER METHODS

To ascertain how the jackknife variance estimator obtained in the earlier section competes with the other rival estimators we consider a linearised variance estimator given by

$$v_T = \sum\sum_s \frac{\mathbf{D}_{ij}}{\mathbf{p}_{ij}}\left(\frac{y_i - \mathbf{a}\hat{\mathbf{b}}_Q x_i}{\mathbf{p}_i} - \frac{y_j - \mathbf{a}\hat{\mathbf{b}}_Q x_j}{\mathbf{p}_j}\right)^2$$

In a further comparative study we consider another variance estimator based on bootstrap re-sampling technique. To estimate the variance of T_{MG} by bootstrap re-sampling technique, a sample of size n is drawn following Midzuno (1952) scheme of sampling. Then assuming for the moment that k = N/n is an integer, k copies of these sampled units are taken to generate an artificial population U^* of size N. From this population generated artificially, M samples, called bootstrap samples, of size n each are drawn by Midzuno (1952) scheme of sampling. Denoting by $T_{MG}(r)$, the modified GREG estimator based on rth bootstrap sample, the bootstrap variance estimator v_B is calculated as

$$v_B = \frac{1}{M-1}\sum_{r=1}^{M}\left(T_{MG}(r) - \frac{1}{M}\sum_{r=1}^{M} T_{MG}(r)\right)^2$$

If N/n is not an integer, we write $N = kn + l$, where $0 < l < n$, and U^* is constructed by taking k copies of n sampled units and then adding to them a sample of size *l* taken from the original sample.

5. A SIMULATION STUDY

We consider a natural population of Swedish Municipality, named MU284 in the book by Särndal, Swensson and Wretman (SSW) (1992). Sweden is divided into 284 municipality having considerable variation in size and other characteristics. The data on a few variables selected by SSW include among others RMT85, the revenue from municipal taxation (in millions of kronor) in 1985 and P85, the population (in thousands) in the same year for all the 284 municipality. We take RMT85 as the study variable Y and P85 as the auxiliary variable X. From population of size N = 284 we select R (= 2000) samples of size n (= 20) each following Midzuno (1952) scheme of sampling. On the basis of each of these 2000 samples we calculate for different values of α, 2000 variance estimators by

(i) Linearisation technique (v_T),
(ii) Standard jackknife method (v_J),
(iii) Bootstrap Method (v_B).
(iv) Weighted jackknife method $(v_{J\mathbf{p}})$, with $\mathbf{q}_j = \frac{\mathbf{p}_j}{n}$ as weight,

The other two choices of $\mathbf{q}_j$ has not been considered because in the present study, with Midzuno scheme of sampling, the inclusion probability $\mathbf{p}_j$, for the present choice of N and n, is not far from being proportional to X_j.

Denoting by v(r) the variance estimator for the rth (r = 1, 2, ……, 2000) sample of a group of 2000 samples, we calculate

$$v_T = \frac{1}{2000}\sum_{r=1}^{2000} v_T(r);\, v_J = \frac{1}{2000}\sum_{r=1}^{2000} v_J(r);\, v_B = \frac{1}{2000}\sum_{r=1}^{2000} v_B(r);\, v_{J\mathbf{p}} = \frac{1}{2000}\sum_{r=1}^{2000} v_{J\mathbf{p}}(r)$$

and compare their performance for different values of $\mathbf{a}$, on observing their closeness to the actual value of the variance given by V_{MG}.

But before going into the task of actual comparison let us observe that with $Q_j = w_j x_j$ where $w_j = x_j^{-g}$ $(0 \le g \le 2)$, T_{MG} depends on two parameters $\mathbf{a}$ and g. So to strike upon an optimum choice of $(\mathbf{a}, g)$ we first observe that in MU284 if RMT85 and P85 are taken as y_j and x_j, then the variance function is likely to be proportion to $x_j^{\mathbf{g}}$, where $\mathbf{g}$ takes value between 1 and 2. If $\mathbf{g}$ is known then the optimum choice of g would be $\mathbf{g}$, but even 284 values of y_j on x_j may not be sufficient to estimate $\mathbf{g}$. So we take a set of trial values for g and α, such that g = 1.0, 1.1, 1.2,………., 2.0 and α = 0.0, 0.1, 0.2, 0.3, ………., 2.0 and for each such trial value we compute the value of V_{MG}, shown in the following table.

Table 5.1

Showing the values of V_{MG} for different values of g and α

$N = 284,\ n = 20,\ Q_j = w_j x_j$ and $w_j = x_j^{-g}$

α \ g	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
0.	8325.2	8325.2	8325.2	8325.2	8325.2	8325.2	8325.2	8325.2	8325.2	8325.2	8325.
0	7125.8	7204.8	7232.3	7257.8	7280.5	7301.5	7319.5	7335.2	7348.8	7360.4	2
0.	6122.0	6175.1	6225.5	6272.3	6314.9	6352.9	6386.3	6415.4	6440.5	6462.1	7370.
1	5163.5	5235.6	5304.4	5368.5	5426.9	5479.1	5525.2	5565.4	5600.1	5630.0	4
0.	4300.4	4386.6	4469.2	4546.4	4617.0	4680.3	4736.3	4785.2	4827.6	4864.1	6480.
2	3532.6	3628.1	3719.9	3806.1	3885.3	3956.6	4019.9	4075.1	4123.1	4164.6	6
0.	2860.4	2960.0	3056.0	3147.7	3231.8	3307.8	3375.4	3434.8	3486.6	3531.3	5655.
3	2283.4	2382.4	2479.1	2571.0	2656.4	2734.0	2803.4	2864.6	2918.1	2964.4	6
0.	1801.9	1895.2	1987.4	2076.1	2159.1	2235.1	2303.5	2364.2	2417.4	2463.8	4895.
4	1415.9	1498.6	1581.8	1663.0	1739.9	1811.2	1875.9	1933.8	1984.8	2029.4	4
0.	1125.2	1192.3	1262.0	1331.6	1399.0	1462.4	1520.7	1573.3	1620.2	1661.5	4200.
5	929.9	976.5	1028.0	1082.0	1136.1	1188.5	1237.6	1282.8	1323.5	1359.3	2
0.	830.0	851.2	880.0	914.2	951.4	989.5	1026.8	1062.1	1094.7	1124.3	3569.
6	825.5	816.3	817.8	823.2	844.8	865.6	888.3	911.4	934.0	955.2	9
0.	916.3	871.8	841.6	828.9	826.4	816.6	822.0	830.6	841.2	852.3	3004.
7	1102.6	1017.8	951.2	901.4	866.1	842.5	827.9	819.8	816.3	815.	5
0.	1384.3	1254.3	1146.7	1060.7	994.0	943.5	906.1	878.9	859.4	7	2504.
8	1761.4	1581.2	1428.1	1301.7	1200.0	1119.4	1056.6	1008.0	970.5	845.5	4
0.	2233.9	1998.5	1795.3	1624.6	1484.1	1370.3	1279.3	1206.9	1149.4	941.6	2068.
9	2801.8	2506.3	2248.5	2029.2	1846.3	1696.2	1574.3	1475.8	1396.5	1103.9	3
1.	3465.1	3104.6	2787.6	2515.6	2286.7	2097.1	1941.5	1814.7	1711.5	1332.6	1697.
0										1627.5	5
1.											1391.
1											7
1.											1150.
2											7
1.											974.7
3											863.5
1.											817.2
4											835.7
1.											919.2

5 1. 6 1. 7 1. 8 1. 9 2. 0											1067. 6 1280. 8 1559. 0

From the above table it is obvious that optimum choice of $(g, \mathbf{a})$, that minimize V_{MG} is $(g_0, \mathbf{a}_0) = (1.9, 1.5)$ and the value of V_{MG} at $(g_0, \mathbf{a}_0)$ is 815.7. So in this case one should be tempted to use

$$T_{MG} = \Sigma_s \frac{y_i}{\mathbf{p}_i} + \frac{\Sigma_s x_i^{-1.9} x_i y_i}{\Sigma_s x_i^{-1.9} x_i^2} \left[\left(\frac{X}{\Sigma_s \frac{x_i}{\mathbf{p}_i}} \right)^{1.5} - 1 \right] \Sigma_s \frac{x_i}{\mathbf{p}_i}$$

as an estimator of Y.

But this way of getting optimum choice for $(g, \mathbf{a})$ will not work in practice as all the population observations will not be known and hence V_{MG} will not be an observable quantity. So for different trial values of g (g = 1.0, 1.1, ……., 2.0) we shall find the optimum choice of $\mathbf{a}$ given by

$$\hat{\mathbf{a}}_{opt}(g) = \frac{1}{\hat{\mathbf{b}}_Q} \frac{\hat{C}_p\left(\Sigma_s \frac{x_i}{\mathbf{p}_i}, \Sigma_s \frac{y_i}{\mathbf{p}_i} \right)}{\hat{V}_p\left(\Sigma_s \frac{x_i}{\mathbf{p}_i} \right)}$$

$$= \left(\frac{\Sigma_s x_i^{-g} x_i y_i}{\Sigma_s x_i^{-g} x_i^2} \right) \frac{\Sigma \Sigma_s \frac{\mathbf{D}_{ij}}{\mathbf{p}_{ij}} \left(\frac{x_i}{\mathbf{p}_i} - \frac{x_j}{\mathbf{p}_j} \right)\left(\frac{y_i}{\mathbf{p}_i} - \frac{y_j}{\mathbf{p}_j} \right)}{\Sigma \Sigma_s \frac{\mathbf{D}_{ij}}{\mathbf{p}_{ij}} \left(\frac{x_i}{\mathbf{p}_i} - \frac{x_j}{\mathbf{p}_j} \right)^2}$$

and the choice of $\left(g, \hat{\mathbf{a}}_{opt}(g)\right)$, which minimise the linearised variance estimator $v_T = v_T\left(g, \hat{\mathbf{a}}_{opt}(g)\right)$ is taken as the optimum choice for $(g, \mathbf{a})$. Necessary calculation in this context are shown in the following table.

Table 5.2

Showing the values of $\hat{\alpha}_{opt}(g)$ and $v_T\left(g, \hat{\alpha}_{opt}(g)\right)$ for different values of g

g	$\hat{\mathbf{a}}_{opt}(g)$	$v_T\left(g, \hat{\mathbf{a}}_{opt}(g)\right)$
1.0	1.246120	849.125
1.1	1.358421	839.705
1.2	1.362451	799.524
1.3	1.379845	794.954
1.4	1.381002	839.843
1.5	1.394142	844.615
1.6	1.410013	853.624
1.7	1.429727	825.261
1.8	1.523112	809.415
1.9	1.553279	822.623
2.0	1.568430	835.624

From the above table it is obvious that $v_T\left(g, \hat{\mathbf{a}}_{opt}(g)\right)$ is minimum when $\left(g, \hat{\mathbf{a}}_{opt}(g)\right) = (1.8,\ 1.523112)$ which is very close to the actual choice $(g_0, \mathbf{a}_0) = (1.9,\ 1.50)$, obtained table 5.1.

However in the following numerical study we use the choice $(g, \mathbf{a}) = (1.9,\ 1.50)$ and proceed further with the numerical investigation.

First we arrange the 2000 samples in increasing order of magnitude of an ancillary $\left(\frac{\sum_S x_i / \mathbf{p}_i}{X}\right)$. The ordered set of 2000 samples are then grouped into 20 subsets of 100 samples each. On the basis of 100 samples in each of 20 groups, we calculate the 100 variance estimators by four different methods. Then to study the conditional performance of the four variance estimators we calculate for each of the 20 groups the values of

$$v_T = \frac{1}{100}\sum_{r=1}^{100} v_T(r),\ v_J = \frac{1}{100}\sum_{r=1}^{100} v_J(r),\ v_B = \frac{1}{100}\sum_{r=1}^{100} v_B(r),\ v_{J\mathbf{p}} = \frac{1}{100}\sum_{r=1}^{100} v_{J\mathbf{p}}(r)$$

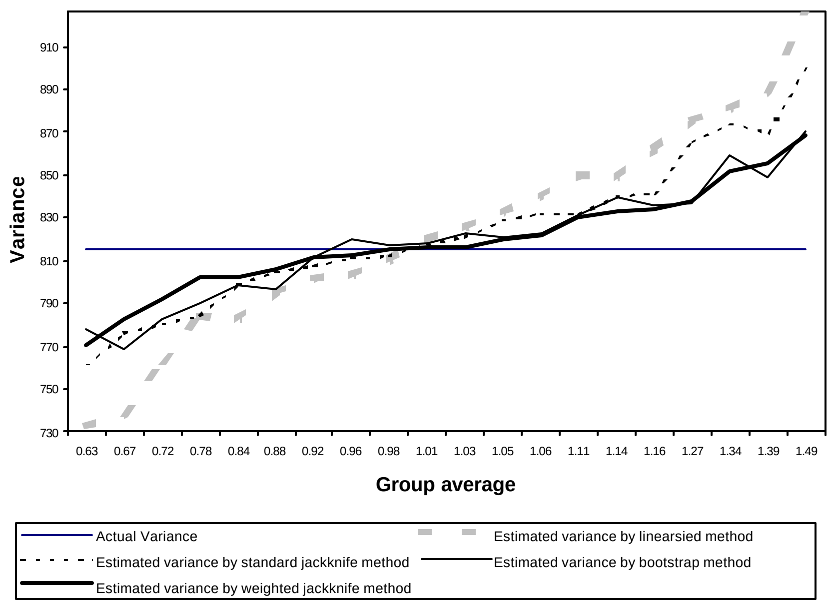
and compare their performance within each group on observing their closeness to the value of V_{MG}, which is 815.7 for $\mathbf{a} = 1.5$. The relevant results in this context are given in the following table.

Table 5.3

Showing the values of four variance estimators, averaged over 100 samples, when $V_{MG} = 815.7,\ N = 284,\ n = 20,\ \mathbf{a} = 1.5,\quad Q_j = w_j x_j,\ w_j = x_j^{g},\ g = 1.9$

$\left(\frac{\sum_S x_i / \mathbf{p}_i}{X}\right)$	v_T	v_J	v_B	$v_{J\mathbf{p}}$	Estimator closest to the actual parameter
0.63	732.5	760.3	778.2	770.2	v_B
0.67	738.4	776.2	768.8	782.6	$v_{J\mathbf{p}}$
0.72	762.8	780.2	782.4	791.6	$v_{J\mathbf{p}}$
0.78	784.2	784.1	790.1	802.1	$v_{J\mathbf{p}}$
0.84	782.5	798.6	798.1	801.9	$v_{J\mathbf{p}}$
0.88	795.3	804.5	797.0	806.4	$v_{J\mathbf{p}}$
0.92	801.6	807.5	811.1	811.2	$v_{J\mathbf{p}}$
0.96	803.5	810.9	819.7	812.6	$v_{J\mathbf{p}}$
0.98	810.5	811.8	817.2	815.0	$v_{J\mathbf{p}}$
1.01	820.1	817.6	817.9	816.4	$v_{J\mathbf{p}}$
1.03	825.9	821.0	822.3	816.5	$v_{J\mathbf{p}}$
1.05	832.6	828.8	820.8	819.7	$v_{J\mathbf{p}}$
1.06	840.0	831.8	822.6	821.7	$v_{J\mathbf{p}}$
1.11	849.7	831.2	831.0	830.5	$v_{J\mathbf{p}}$
1.14	849.2	839.8	839.2	832.8	$v_{J\mathbf{p}}$
1.16	862.2	841.5	836.1	833.5	$v_{J\mathbf{p}}$
1.27	875.9	865.5	837.1	838.0	v_B
1.34	881.0	874.0	859.1	851.4	$v_{J\mathbf{p}}$
1.39	889.4	869.6	849.2	855.3	v_B
1.49	925.2	899.2	870.1	868.5	$v_{J\mathbf{p}}$

Fig:1-Comparison of variance estimates by different methods with actual variance



Remarks: Form the last column of the above table and Figure 1 the weighted jackknife variance estimator seems to have an edge over its rival estimators. The improvement appears to be significant especially for those samples for which the average value of the ancillary $\left(\frac{\sum_S x_i / \mathbf{p}_i}{X}\right)$ is far from 1.

Next we compare the performance of four variance estimators on the basis of Actual Coverage Percentage (ACP); which gives the percentage of cases in which the Confidence Interval (CI)

$$T_{MG} \mp \mathbf{t}_{\mathbf{b}/2}\sqrt{v}$$

covers the actual value of the finite population total Y. The closer it is to $100\beta\%$, the better is the CI.

For the four different estimators of the variance, we construct six different sets of confidence intervals(CI) and observe the conditional coverage percentage in each of 20 groups of 100 samples. To justify the use of normal score $t_{\mathbf{b}/2}$ in the construction of CI, we calculate the values of

$$b_1 = \frac{1}{R}\sum_{r=1}^{R}\left(\frac{T_{MG}(r) - \overline{T}_{MG}}{s_n}\right)^3 \text{ and } b_2 = \frac{1}{R}\sum_{r=1}^{R}\left(\frac{T_{MG}(r) - \overline{T}_{MG}}{s_n}\right)^4$$

where

$$\overline{T}_{MG} = \frac{1}{R}\sum_{r=1}^{R} T_{MG}(r) \text{ and } s_n = \sqrt{\frac{1}{R}\sum_{r=1}^{R}\left(T_{MG}(r) - \overline{T}_{MG}\right)^2}$$

It is noted that for R = 2000 samples $\left(\sqrt{b_1}, b_2\right) = (0.062, 2.981)$. Relevant results in this context are given in the following table.

Table 5.4

Showing the ACP's for four different variance estimators when $V_{MG} = 815.7$, $N = 284$, $n = 20$, $\mathbf{b} = 0.05$, $\left(\sqrt{b_1}, b_2\right) = (0.062, 2.981)$, $Q_j = w_j x_j$, $w_j = x_j^{-g}$, $g = 1.9$, $\mathbf{a} = 1.5$

$\left(\frac{\sum_S x_i / \mathbf{p}_i}{X}\right)$	ACP of the CI corresponding to the estimators			
	v_T	v_J	v_B	$v_{J\mathbf{p}}$
0.63	84	87	90	88
0.67	84	88	88	90
0.72	86	89	89	91
0.78	88	90	91	93
0.84	88	92	91	93
0.88	90	93	94	94
0.92	93	93	95	95
0.96	94	95	95	95
0.98	94	95	95	95
1.01	95	95	95	95
1.03	96	95	95	95
1.05	96	96	95	95
1.06	97	96	95	95
1.11	97	96	97	96

1.14	97	97	97	96
1.16	98	97	97	96
1.27	98	98	97	97
1.34	99	99	98	97
1.39	100	98	97	97
1.49	100	99	98	97

Remark: From the above table it appears that with respect to ACP the performance of the weighted jackknife variance estimator is very good, if not the best. However the samples for which the average value of the ancillary is very close to 1, all the variance estimators seem to be equally efficient in yielding confidence intervals with a given confidence co-efficient.

ACKNOWLEDGEMENT

We wish to extend our thanks to an anonymous referee and an associate editor for their constructive criticisms and useful comments on the paper leading to substantial improvement over an earlier draft.

REFERENCES

1. Basu, D. (1971) An essay on the logical foundations of survey sampling. *Foundation of Statistical Inference.* Edited by Godambe, V. P. and Sprott, D. A. Hotl, Rinehart, Winston, Toronto. Part I, 203 – 242.
2. Brewer, K.R.W. (1979) A class of robust sampling designs for large scale surveys. *J. Amer. Statist. Assoc.*, 74, 911-915.
3. Brewer, K.R.W. (1999) Cosmetic calibration with unequal probability sampling. *Survey Methodology*, 25, 205 – 212.
4. Chaudhuri, A. & Stenger, H. (1992) *Survey Sampling, Theory and Methods*. Marcel Dekker, Inc.
5. Chaudhry, N.A. (1990) The weighted jackknife for ratio estimation. *Comm. In Statist. Theory and Methods*, 19(9), 3283 – 3313.
6. Davison, A.C. & Hinkley, D.V. (1998) *Bootstrap Method and their application*. Cambridge University Press.
7. Durbin, J. (1959) A note on the application of Quenouille's method of bias reduction to the estimation of ratios. *Biometrika*, 46, 477-480.
8. Hajék, J. (1971) Comment on a paper by Basu, D. *Foundation of Statistical Inference*, 203-242.

9. Hinkley, J.D.V. (1977) Jackknifing in unbalance situations. *Technometrics*. 19(3), 258 – 292.
10. Horvitz, D.G. and Thompson D.J. (1952) A generalization of sampling without replacement from a finite universe. *J. Amer. Statist. Assoc.*, 47, 203 – 242.
11. Midzuno, H. (1952) On the system with probabilities proportionate to sum of sizes. *Ann. Institute of Statist. Math*. 3, 99 – 107.
12. Nandi, S.B. & Aich, A.B. (1994) Jackknifing modified ratio estimator under a model. *Sankhyâ: The Indian J. Statist.*, 56 B1, 18 – 25.
13. Quenouille, M.H. (1956) Notes on bias in estimation. *Biometrika*, 43, 353-360.
14. Rao, J.N.K. (1965) A note on estimation of ratios by Quenouille's method. *Biometrika*, 52, 647-649.
15. Rao, J.N.K & Webster, J.T. (1966) On two methods of bias reduction in the estimation of ratios. *Biometrika*, 53, 571-577.
16. Roy, D. & Safiquzzaman, M. (2000) Jackknifing a general class of estimator. *Brazilian J. Probability and Statist.;* 14, 47 – 60.
17. Särndal, C.E.(1980) On π-inverse weighting versus best linear weighting in probability sampling. *Biometrika*, 67, 639-650.
18. Särndal, C. E. (1982) Implication of survey design for generalized regression Estimation of linear function. J. Statist. Planning and Inference, 7, 155 – 170.
19. Särndal, C.E., Swensson, B. & Wretman, J. (1992) *Model Assisted Survey Sampling*, Springer-Verlag.
20. Särndal, C. E. & Wright, R.L. (1984) Cosmetic form of estimators in survey sampling. *Scandinavian J. Statist*., 11, 146-156.
21. Srivastava, S. K. (1967) An estimator using auxiliary information in sample survey. *Calcutta Statistical Association Bulletin*, 16, 121 – 132.
22. Tukey, J. (1958) Bias and confidence in not quite large sample. *Ann. Math. Statist*., 29, 614.
23. Wright, R.L. (1983) Finite population sampling with multivariate auxiliary information. *J. Amer. Statist. Assoc.*, 78, 879-884.