### On a Scheme of Sampling of Two Units With Inclusion Probability Proportional to Size

Sarat C. Senapati<sup>1</sup>, L. N. Sahoo<sup>2</sup>, and G. Mishra<sup>2</sup>

<sup>1</sup>Ravenshaw College, Cuttack, India <sup>2</sup>Utkal University, Bhubaneswar, India

Abstract: This paper introduces an unequal probability sampling without replacement scheme with inclusion probability proportional to size. This new scheme possesses some desirable properties with regard to  $\pi_i$  and  $\pi_{ij}$ , and provides a non-negative variance estimator of the Horvitz and Thompson estimator, when the values of the auxiliary variable fulfill some restrictions. On comparing the suggested scheme with some of the existing sampling schemes in respect of efficiency and stability of the variance estimator empirically, it has been observed that the performance of the scheme is satisfactory.

**Zusammenfassung:** Wir stellen hier einen Stichprobenplan vor, der ohne Wiederholung und mit Aufnahmewahrscheinlichkeit proportional zum Umfang arbeitet. Dieser neue Plan weist bzgl.  $\pi_i$  und  $\pi_{ij}$  einige wünschenswerte Eigenschaften auf und liefert einen nicht-negativen Varianzschätzer des Horvitz und Thompson Schätzers, falls die Hilfsvariable einige Restriktionen erfüllt. Vergleicht man empirisch den vorgeschlagenen Plan mit einigen anderen Stichprobenplänen bzgl. Effizienz und Stabilität des Varianzschätzers, so hat sich gezeigt, dass die Güte des Plans zufriedenstellend ist.

Keywords: Joint Inclusion Probability, Unequal Probability Sampling.

### **1** Introduction

Consider a finite population of N units, and let  $y_i$ , i = 1, ..., N, denote the value for the *i*th unit of a certain character y. We are interested in estimating  $Y = \sum_i y_i$ , the total of all y-values based on a sample s of n distinct units selected from the population according to some unequal probability sampling without replacement scheme with  $\pi_i$  as the inclusion probability of the *i*th unit and  $\pi_{ij}$  as the joint inclusion probability of the *i*th and *j*th units. The most commonly used estimator in this connection is the Horvitz and Thompson (1952) (HT) estimator defined by

$$\hat{Y}_{HT} = \sum_{i \in s} \frac{y_i}{\pi_i} \,.$$

From Horvitz and Thompson (1952) we have  $\sum_{i} \pi_{i} = n$ ,  $\sum_{i \neq j} \pi_{ij} = (n-1)\pi_{i}$  and  $\sum_{i} \sum_{j < i} \pi_{ij} = n(n-1)/2$ . Sen (1953) and Yates and Grundy (1953) independently suggested an unbiased estimator of  $\operatorname{var}(\hat{Y}_{HT})$  given by

$$v(\hat{Y}_{HT}) = \sum_{i \in s} \sum_{j < i} \frac{\pi_i \pi_j - \pi_{ij}}{\pi_{ij}} \left(\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j}\right)^2 \,, \tag{1}$$

which holds for a fixed size sampling design. A sufficient condition for (1) to be always non-negative is that  $\pi_{ij} < \pi_i \pi_j, \forall i \neq j$ .

When information is available on an auxiliary character x with value  $x_i$  on unit i, considerable reduction in  $var(\hat{Y}_{HT})$  can be achieved by making  $\pi_i = np_i$ . This is of course only guaranteed if  $y_i/x_i$  is nearly constant for all i, where  $p_i = x_i/X$  is the initial probability to select the *i*th unit and  $X = \sum_i x_i$ . Such a scheme is known as an inclusion probability proportional to size (IPPS or  $\pi ps$ ) sampling scheme. Brewer and Hanif (1983) and Chaudhuri and Vos (1988) have made elaborate discussions on a number of IPPS methods. But a majority of these methods are restricted to n = 2 only. The reasons are not far to seek. The calculation of  $\pi_{ij}$  becomes cumbersome for n > 2 and some procedures seem to be less precise than even probability proportional to size with replacement (PPS-WR) scheme. However, an IPPS sampling scheme with n = 2 is very useful in stratified sampling, where stratification is sufficiently *deep*, i.e. the number of strata and their sizes is such that a sample of two units per stratum meets the requirement on the total sample size. Indeed, the advantages of stratified sampling are exploited to the greatest possible extent if the number of strata is maximum and n = 2 is the minimum sample size to estimate the variance within each stratum (cf Chaudhuri and Vos, 1988, p. 148).

The objective of this paper is to develop an IPPS sampling scheme for n = 2 which possesses desirable properties and also performs well when compared to popular probability sampling schemes for a number of natural populations with different characteristics.

#### 2 Description of the Suggested Sampling Scheme

Let us assume that  $x_1, \ldots, x_N$  are known and  $x_i > 0 \forall i$ . Consider a set of revised probabilities  $\{P_1, \ldots, P_N\}$ , where  $P_i$  is defined by

$$P_i = \frac{(1 - g_i)(2p_i^2 - \lambda)}{p_i(1 - 2g_i)}, \qquad i = 1, \dots, N,$$

such that  $g_i = \bar{p}_h / N p_i$ ,  $\bar{p}_h$  being the simple harmonic mean of  $p_1, \ldots, p_N$ , and  $\lambda$  is a known constant. In actual practice, we choose  $\lambda$  such that  $\sum_i P_i = 1$ . Thus, we now have

$$\lambda = \sum_{i} \frac{p_i}{1 - 2g_i} \left/ \sum_{i} \frac{1 - g_i}{p_i (1 - 2g_i)} \right.$$
(2)

It must be noted here that the computation of the revised probabilities is restricted only to situations for which  $p_i^2 \ge \lambda/2$  and  $g_i \le 1/2 \quad \forall i$ , because otherwise (1) would give negative results (see (6) below). This means that the proposed sampling scheme cannot be used when at least one of the  $p_i$ 's is small compared to the others.

Our sampling scheme for n = 2 is defined as follows:

- Draw the first unit, say *i*, with revised probability  $P_i$  and without replacement,  $\lambda$  being given by (2).
- Draw the second unit, say j, from the remaining (N-1) units with conditional probability  $P_{j|i} = g_j/(1-g_i)$ .

# **3** Inclusion Probabilities

By definition,

$$\pi_{i} = P_{i} + \sum_{j \neq i} P_{j} \frac{g_{i}}{1 - g_{j}}$$

$$= \frac{(1 - g_{i})(2p_{i}^{2} - \lambda)}{p_{i}(1 - 2g_{i})} + \sum_{j \neq i} \frac{g_{i}(2p_{j}^{2} - \lambda)}{p_{j}(1 - 2g_{j})}$$

$$= 2p_{i} - \frac{\lambda}{p_{i}} + g_{i} \sum_{j} \frac{2p_{j}^{2} - \lambda}{p_{j}(1 - 2g_{j})}.$$
(3)

From (2) we have

$$\sum_{i} \frac{p_i}{1 - 2g_i} - \frac{\lambda}{2} \sum_{i} \frac{1 + (1 - 2g_i)}{p_i(1 - 2g_i)} = 0,$$

i.e.,

$$\sum_{i} \frac{2p_i^2 - \lambda}{p_i(1 - 2g_i)} - \lambda \sum_{i} \frac{1}{p_i} = 0.$$

Noting that

$$g_i = \frac{1}{p_i} \left/ \sum_j \frac{1}{p_j} \right|,$$

we obtain

$$g_i \sum_{j} \frac{2p_j^2 - \lambda}{p_j(1 - 2g_j)} - \frac{\lambda}{p_i} = 0.$$
 (4)

Hence, from (3) together with (4) we get

 $\pi_i = 2p_i \,.$ 

Again, by definition

$$\pi_{ij} = P_i P_{j|i} + P_j P_{i|j} \tag{5}$$

$$= \frac{g_j(2p_i^2 - \lambda)}{p_i(1 - 2g_i)} + \frac{g_i(2p_j^2 - \lambda)}{p_j(1 - 2g_j)}.$$
(6)

# **4 Properties**

The important properties of the proposed sampling scheme are as follows:

(i) 
$$\sum_{i} \pi_{i} = 2 \sum_{i} p_{i} = 2$$
,

(ii)

$$\begin{split} \sum_{j \neq i} \pi_{ij} &= \frac{(2p_i^2 - \lambda)}{p_i(1 - 2g_i)} \sum_{j \neq i} g_j + g_i \sum_{j \neq i} \frac{2p_j^2 - \lambda}{p_j(1 - 2g_j)} \\ &= \frac{(2p_i^2 - \lambda)(1 - g_i)}{p_i(1 - 2g_i)} - \frac{g_i(2p_i^2 - \lambda)}{p_i(1 - 2g_i)} + g_i \sum_j \frac{2p_j^2 - \lambda}{p_j(1 - 2g_j)} \\ &= 2p_i - \frac{\lambda}{p_i} + g_i \sum_j \frac{2p_j^2 - \lambda}{p_j(1 - 2g_j)} \\ &= 2p_i \qquad (\text{using (4)}) \\ &= \pi_i \,, \end{split}$$

(iii)  $\pi_i \pi_j - \pi_{ij} > 0$ ,  $i \neq j = 1, \dots, N$ .

Following Konijn (1973, p. 253) we now have

$$\pi_{i}\pi_{j} - \pi_{ij} = \left(\pi_{ij} + \sum_{k \neq i \neq j} \pi_{ik}\right) \left(\pi_{ij} + \sum_{k \neq i \neq j} \pi_{jk}\right) - \pi_{ij}$$
$$= \pi_{ij} \left(1 - \sum_{l \neq i \neq j} \sum_{l < k} \pi_{lk}\right) + \sum_{k \neq i \neq j} \pi_{ik} \sum_{k \neq i \neq j} \pi_{jk} - \pi_{ij}$$
$$= \sum_{k \neq i \neq j} \pi_{ik} \sum_{k \neq i \neq j} \pi_{jk} - \pi_{ij} \sum_{l \neq i \neq j} \sum_{l < k} \pi_{lk},$$
(7)

$$\sum_{k \neq i \neq j} \pi_{ik} \sum_{k \neq i \neq j} \pi_{jk} = \left[ \frac{2p_i^2 - \lambda}{p_i(1 - 2g_i)} \sum_{k \neq i \neq j} g_k + g_i \sum_{k \neq i \neq j} \frac{2p_k^2 - \lambda}{p_k(1 - 2g_k)} \right] \times \left[ \frac{2p_j^2 - \lambda}{p_j(1 - 2g_j)} \sum_{k \neq i \neq j} g_k + g_j \sum_{k \neq i \neq j} \frac{2p_k^2 - \lambda}{p_k(1 - 2g_k)} \right] \\ = \frac{(2p_i^2 - \lambda)(2p_j^2 - \lambda)}{p_i p_j(1 - 2g_i)(1 - 2g_j)} \left( \sum_{k \neq i \neq j} g_k \right)^2 \\ + g_i g_j \left[ \sum_{k \neq i \neq j} \frac{2p_k^2 - \lambda}{p_k(1 - 2g_k)} \right]^2 + \pi_{ij} \sum_{k \neq i \neq j} g_k \sum_{k \neq i \neq j} \frac{2p_k^2 - \lambda}{p_k(1 - 2g_k)}, \quad (8)$$

$$\pi_{ij} \sum_{l \neq i \neq j} \sum_{l < k} \pi_{lk} = \pi_{ij} \sum_{l \neq i \neq j} \sum_{l < k} \left[ \frac{g_k (2p_l^2 - \lambda)}{p_l (1 - 2g_l)} + \frac{g_l (2p_k^2 - \lambda)}{p_k (1 - 2g_k)} \right] \\ = \pi_{ij} \sum_{k \neq i \neq j} \left[ \left( \sum_{l \neq i \neq j} g_l - g_k \right) \frac{2p_k^2 - \lambda}{p_k (1 - 2g_k)} \right].$$
(9)

Hence, from (7), (8) and (9) we have

$$\pi_i \pi_j - \pi_{ij} = \frac{(2p_i^2 - \lambda)(2p_j^2 - \lambda)}{p_i p_j (1 - 2g_i)(1 - 2g_j)} \left(\sum_{k \neq i \neq j} g_k\right)^2 + g_i g_j \left[\sum_{k \neq i \neq j} \frac{2p_k^2 - \lambda}{p_k (1 - 2g_k)}\right]^2 + \pi_{ij} \sum_{k \neq i \neq j} \frac{g_k (2p_k^2 - \lambda)}{p_k (1 - 2g_k)} \\> 0.$$

Thus, an unbiased and positive estimator of the variance of the HT estimator can always be obtained under the suggested sampling scheme.

### 5 An Example

In order to examine how the suggested sampling scheme operates in a survey situation and possesses  $\pi$ ps properties, we consider a small artificial population of 5 units given in Mukhopadhyay (1998, p. 220). The following table gives the values of  $y_i$ ,  $x_i$ ,  $p_i$ , and  $P_i$ .

| $y_i$ | 123    | 192    | 212    | 267    | 290    |
|-------|--------|--------|--------|--------|--------|
| $x_i$ | 17     | 18     | 20     | 22     | 23     |
| $p_i$ | 0.17   | 0.18   | 0.20   | 0.22   | 0.23   |
| $P_i$ | 0.0763 | 0.1240 | 0.2072 | 0.2795 | 0.3129 |

In this simple example the auxiliary variable x is not proportional to the study variable y and therefore we cannot speak of a probability proportional to size sampling design. But, it is of course a curiosity to note that in this case the proposed scheme works as all  $P_i > 0$ implying that the restrictive conditions for the existence of the scheme, i.e.  $p_i^2 \ge \lambda/2$  and  $g_i \leq 1/2 \ \forall i \text{ are satisfied. For } n = 2 \text{ we compute } \pi_1 = 0.34, \ \pi_2 = 0.36, \ \pi_3 = 0.40,$  $\pi_4 = 0.44, \pi_5 = 0.46, \pi_{12} = 0.0587, \pi_{13} = 0.0796, \pi_{14} = 0.0970, \pi_{15} = 0.1047, \pi_{23} = 0.0000, \pi_{15} = 0.000$  $0.0880, \pi_{24} = 0.1032, \pi_{25} = 0.1101, \pi_{34} = 0.1136, \pi_{35} = 0.1188$ , and  $\pi_{45} = 0.1263$ , and verify that  $\pi_i \pi_j - \pi_{ij} > 0 \ \forall i \neq j$ . Consequently, the Sen-Yates-Grundy variance estimator of the HT estimator under the suggested scheme is non-negative. The variances of  $\hat{Y}_{HT}$ under the proposed scheme, the mean per unit estimator  $\hat{Y} = (N/n) \sum_{i \in s} y_i$  under simple random sampling without replacement scheme and the conventional estimator  $\hat{Y}_{PPS}$  =  $(1/n) \sum_{i \in s} y_i/p_i$  based on the PPSWR scheme are calculated to be 12318, 32465, and 16575, respectively. These calculations clearly show that the variance reductions of the suggested scheme compared to simple random sampling without replacement scheme and **PPSWR** scheme in this example are remarkable although  $p_i$  is not chosen proportional to the size of  $y_i$ . Hence, from these findings we may expect that the scheme can be safely programmed in some situations.

### 6 Performance of the Scheme

To compare the performance of the proposed sampling scheme with some other well known sampling schemes, we consider two different performance measures, viz.,

- 1. *Relative efficiency:* Here, the relative efficiency of a sampling scheme is defined as the variance ratio of the scheme and the PPSWR scheme.
- 2. Stability of the variance estimator: We accept the Hanurav (1967) criterion  $\phi = \min(\pi_{ij}/\pi_i\pi_j) > \beta$ ,  $\forall i \neq j$ , for  $\beta$  sufficiently away from zero, to study stability of the variance estimator of a sampling scheme.

The following sampling plans are taken into consideration:

- A Conventional estimator under PPSWR sampling scheme
- **B** HT estimator under the sampling scheme of Brewer (1963)
- C HT estimator under the sampling scheme of P. Singh (1978)
- **D** HT estimator under the sampling scheme of Deshpande and Prabhu Ajgaonkar (1982)
- **E** Ordered estimator of Raj (1956)
- **F** Estimator of Rao, Hartley, and Cochran (1962)
- **G** Unordered estimator of Murthy (1957)
- **H** HT estimator under the suggested sampling scheme.

Three IPPS sampling procedures B, C, and D are considered for comparison with our suggested procedure H with respect to efficiency and stability of the variance estimator. To examine the efficiency of the HT estimator based on the suggested sampling scheme over other estimators based on probability proportional to size without replacement scheme, we also include the three well known estimators due to Raj, Rao-Hartely-Cochran, and Murthy in our comparison. Since a theoretical comparison is impracticable, we resort to an empirical study of the above procedures for a collection of 18 small natural populations, because a sample of size two only has been considered.

Table 1 describes source, size (N), nature of y and x, and the correlation coefficient  $\rho$  between y and x. Numerical values of the relative efficiency of the comparable sampling plans (in %), and the stability standard  $\phi$  of the variance estimators of the plans are presented in Tables 2 and 3, respectively. Relative efficiency of a sampling scheme is computed by using the exact variance formula for the full population. On the other hand, the  $\phi$ -value of a scheme is decided after computing the quantity  $\pi_{ij}/\pi_i\pi_j$  for all C(N, n) possible samples of n = 2 drawn from a population. Findings in these tables indicate that the suggested scheme is more efficient than other schemes for all populations, whereas its variance estimator is more stable than others for 17 populations (except population 3).

| Pop. | Source  | N  | y   | x  |      |
|------|---|----|---|--|------|
| 1    | Cochran (1977, p. 203)                                | 10 | actual weight of pea-<br>ches                         | estimated weight of peaches                            | 0.97 |
| 2    | D. Singh and Chaud-<br>hary (1986, p. 155)            | 17 | no. of milch animals in survey                        | no. of milch animals in census                         | 0.72 |
| 3    | Konijn (1973, p. 49)                                  | 16 | expenses on food                                      | total expenses   | 0.95 |
| 4    | Cochran (1977, p. 325)                                | 10 | no. of persons  | no. of rooms   | 0.65 |
| 5    | R. Singh and Singh<br>Mangat (1996, p. 193,<br>1-14)  | 14 | milk yield after in-<br>troduction of the new<br>feed | milk yield before in-<br>troduction of the new<br>feed | 0.98 |
| 6    | R. Singh and Singh<br>Mangat (1996, p. 193,<br>15-27) | 14 | milk yield after in-<br>troduction of the new<br>feed | milk yield before in-<br>troduction of the new<br>feed | 0.99 |
| 7    | R. Singh and Singh<br>Mangat (1996, p. 255)           | 15 | tax evaded  | passengers   | 0.59 |
| 8    | Mukhopadhyay (1998, p. 114)                           | 9  | census population for 1961                            | census population for 1951                             | 0.93 |
| 9    | Mukhopadhyay (1998, p. 193, 1-10)                     | 10 | quantity of raw materi-<br>als                        | no. of laborers  | 0.91 |
| 10   | Mukhopadhyay (1998, p. 193, 11-20)                    | 10 | quantity of raw materi-<br>als                        | no. of laborers  | 0.56 |
| 11   | Sukhatme and Sukhat-<br>me (1970, p. 166, 1-<br>10)   | 10 | no. of banana bunches                                 | no. of banana pits                                     | 0.64 |
| 12   | Sukhatme and Sukhat-<br>me (1970, p. 166, 11-<br>20)  | 10 | no. of banana bunches                                 | no. of banana pits                                     | 0.84 |
| 13   | R. Singh and Singh<br>Mangat (1996, p. 173,<br>1-13)  | 13 | area harvested with combine                           | area under paddy                                       | 0.94 |
| 14   | R. Singh and Singh<br>Mangat (1996, p. 173,<br>14-26) | 13 | area harvested with combine                           | area under paddy                                       | 0.95 |
| 15   | Yates (1953, p. 169)                                  | 17 | area under wheat                                      | total acreage of crops<br>and grass                    | 0.51 |
| 16   | R. Singh and Singh<br>Mangat (1996, p. 202,<br>1-11)  | 11 | leaf area   | leaf weight  | 0.95 |
| 17   | R. Singh and Singh<br>Mangat (1996, p. 202,<br>12-22) | 11 | leaf area   | leaf weight  | 0.92 |
| 18   | R. Singh and Singh<br>Mangat (1996, p. 202,<br>23-33) | 11 | leaf area   | leaf weight  | 0.94 |

Tabelle 1: Description of populations

|      | Sampling Plan |        |        |        |        |        |        |        |
|------|---------------|--------|--------|--------|--------|--------|--------|--------|
| Pop. | А             | В      | С      | D      | Ē      | F      | G      | Н      |
| 1    | 100.00        | 112.10 | 112.09 | 111.98 | 111.12 | 112.49 | 112.53 | 114.62 |
| 2    | 100.00        | 106.73 | 106.72 | 106.72 | 106.27 | 106.24 | 106.69 | 106.77 |
| 3    | 100.00        | 107.28 | 107.26 | 107.25 | 106.74 | 107.14 | 107.23 | 108.41 |
| 4    | 100.00        | 111.64 | 111.66 | 111.56 | 110.90 | 112.50 | 112.24 | 112.66 |
| 5    | 100.00        | 108.75 | 108.64 | 108.68 | 107.95 | 108.33 | 108.64 | 113.55 |
| 6    | 100.00        | 107.92 | 107.46 | 107.87 | 108.11 | 108.33 | 108.84 | 120.33 |
| 7    | 100.00        | 107.51 | 107.51 | 107.47 | 107.14 | 107.14 | 107.69 | 107.91 |
| 8    | 100.00        | 113.85 | 113.82 | 113.71 | 112.53 | 112.49 | 114.34 | 114.76 |
| 9    | 100.00        | 112.10 | 112.15 | 112.00 | 111.31 | 112.50 | 112.79 | 113.25 |
| 10   | 100.00        | 112.94 | 112.56 | 112.77 | 111.50 | 112.00 | 113.01 | 119.05 |
| 11   | 100.00        | 113.72 | 113.76 | 113.48 | 111.66 | 112.49 | 113.22 | 114.00 |
| 12   | 100.00        | 112.30 | 112.38 | 112.15 | 111.25 | 112.50 | 112.71 | 112.99 |
| 13   | 100.00        | 107.95 | 107.79 | 107.88 | 108.26 | 108.83 | 109.05 | 113.97 |
| 14   | 100.00        | 109.15 | 108.88 | 109.05 | 108.59 | 108.33 | 109.42 | 112.30 |
| 15   | 100.00        | 106.60 | 106.61 | 106.58 | 106.38 | 106.24 | 106.82 | 106.93 |
| 16   | 100.00        | 111.14 | 111.03 | 111.02 | 110.05 | 109.59 | 111.18 | 115.32 |
| 17   | 100.00        | 112.09 | 111.94 | 111.91 | 110.42 | 110.00 | 111.63 | 117.33 |
| 18   | 100.00        | 111.86 | 111.69 | 111.69 | 110.34 | 109.99 | 111.53 | 114.91 |

Tabelle 2: Relative efficiency of different sampling plans

Tabelle 3: Stability standard of different sampling schemes

|      | Sampling Scheme |         |         |         |  |  |
|------|-----------------|---------|---------|---------|--|--|
| Pop. | В               | С       | D       | Н       |  |  |
| 1    | 0.54401         | 0.54393 | 0.54430 | 0.54740 |  |  |
| 2    | 0.53120         | 0.53118 | 0.53109 | 0.53142 |  |  |
| 3    | 0.52579         | 0.52627 | 0.52585 | 0.49623 |  |  |
| 4    | 0.54758         | 0.54754 | 0.54763 | 0.55777 |  |  |
| 5    | 0.53178         | 0.53166 | 0.53183 | 0.55757 |  |  |
| 6    | 0.53446         | 0.53372 | 0.53463 | 0.54241 |  |  |
| 7    | 0.54015         | 0.53991 | 0.53988 | 0.54510 |  |  |
| 8    | 0.55105         | 0.55105 | 0.55151 | 0.55625 |  |  |
| 9    | 0.52881         | 0.53241 | 0.53037 | 0.53510 |  |  |
| 10   | 0.56454         | 0.55505 | 0.56186 | 0.57013 |  |  |
| 11   | 0.54218         | 0.54109 | 0.54226 | 0.55003 |  |  |
| 12   | 0.53559         | 0.53793 | 0.53689 | 0.54030 |  |  |
| 13   | 0.51989         | 0.52104 | 0.52056 | 0.52779 |  |  |
| 14   | 0.52362         | 0.52511 | 0.52422 | 0.59560 |  |  |
| 15   | 0.53271         | 0.53247 | 0.53266 | 0.53538 |  |  |
| 16   | 0.54258         | 0.54286 | 0.54261 | 0.54930 |  |  |
| 17   | 0.55030         | 0.54982 | 0.54993 | 0.56442 |  |  |
| 18   | 0.54467         | 0.54179 | 0.54418 | 0.54640 |  |  |

## 7 Conclusions

On the basis of the analytical and empirical results derived in this work, we may conclude that the suggested sampling procedure is in no way inferior to some standard sampling procedures. But no general conclusion can be drawn from the empirical study as the conclusion is based on the results for 18 populations only and the gain in efficiency of the suggested scheme compared to other leading alternatives is in fact rather small. However, our empirical investigation gives an indication that the suggested sampling scheme (if it exists) compares well with other popularized schemes in terms of efficiency as well as the stability of the estimated variance.

#### Acknowledgement

The authors are grateful to the referees for providing many helpful comments and suggestions which have improved the presentation of the paper.

### Literatur

- Brewer, K. R. W. (1963). Ratio estimation in finite populations: Some results deducible from the assumption of an underlying stochastic process. *Australian Journal of Statistics*, *5*, 93-105.
- Brewer, K. R. W., and Hanif, M. (1983). *Sampling with Unequal Probabilities*. Springer-Verlag.
- Chaudhuri, A., and Vos, J. W. E. (1988). Unified Theory and Strategies of Survey Sampling. North Holland.
- Cochran, W. G. (1977). Sampling Techniques (3rd ed.). John Wiley and Sons.
- Deshpande, M. N., and Prabhu Ajgaonkar, S. G. (1982). An IPPS (inclusion probability proportional to size) sampling scheme. *Statistica Neerlandica*, *36*, 209-212.
- Hanurav, T. V. (1967). Optimum utilization of auxiliary information:  $\pi$ ps sampling of two units from a stratum. *Journal of the Royal Statistical Society, B*, 29, 374-391.
- Horvitz, D. G., and Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47, 663-685.
- Konijn, H. S. (1973). *Statistical Theory of Sample Survey Design and Analysis*. North Holland.
- Mukhopadhyay, P. (1998). *Theory and Methods of Survey Sampling*. New Delhi: Prentice-Hall of India.
- Murthy, M. N. (1957). Ordered and unordered estimators in sampling without replacement. *Sankhyâ*, *18*, 379-390.
- Raj, D. (1956). Some estimators in sampling with varying probabilities without replacement. *Journal of the American Statistical Association*, *51*, 269-284.
- Rao, J. N. K., Hartley, H. O., and Cochran, W. G. (1962). A simple procedure of unequal probability sampling without replacement. *Journal of the Royal Statistical Society*, *B*, 24, 482-491.

- Sen, A. R. (1953). On the estimator of the variance in sampling with varying probabilities. Journal of the Indian Society of Agricultural Statistics, 5, 119-127.
- Singh, D., and Chaudhary, F. S. (1986). *Theory and Analysis of Sample Survey Designs*. Wiley Eastern Limited.
- Singh, P. (1978). The selection of samples of two units with inclusion probability proportional to size. *Biometrika*, 65, 450-454.
- Singh, R., and Singh Mangat, N. (1996). *Elements of Survey Sampling*. The Netherlands: Kluwer Academic Publishers.
- Sukhatme, P. V., and Sukhatme, B. V. (1970). Sampling Theory of Surveys with Applications. Calcutta: Asia Publishing House.
- Yates, F. (1953). Sampling Methods for Censuses and Surveys. London: Charles Griffin & Company.
- Yates, F., and Grundy, P. M. (1953). Selection without replacement from within strata with probability proportional to size. *Journal of the Royal Statistical Society, B*, *15*, 235-261.

Author's addresses:

S. C. Senapati Department of Statistics Ravenshaw College Cuttack 753003 India E-mail: scsenapati2002@rediffmail.com

L. N. Sahoo and G. Mishra Department of Statistics Utkal University Bhubaneswar 751004 India