EFFICIENCY OF STRATIFICATION IN SUB-SAMPLING DESIGNS FOR THE RATIO METHOD OF ESTIMATION WITH VARYING PROBABILITIES OF SELECTION

By J. N. K. Rao and H. K. Chawla

Forest Researcn Institute, Dehra Dun and Indian Council of Agricultural Research, New Delhi

Mokashi (1954) has worked out the efficiency of stratification for the ratio method of estimation with equal probabilities of selection in sub-sampling designs. There may be further gain in efficiency if the units are selected with probabilities proportional to some measure highly correlated with the character under study. Following Sukhatme (1954), who has worked out the efficiency of stratification with varying probabilities of selection with replacement for a single variate, formulæ are derived in this paper, to estimate the efficiency of stratification for the ratio method of estimation when sampling is done with varying probabilities of selection and with replacement.

Let

N, be the number of units in the population.

N_t, the number of units in the t-th stratum.

n, the number of units in the sample.

n_t, the number of units in the sample from the t-th stratum.

M_{ti}, the number of sub-units in the i-th unit of the t-th stratum.

m_{ti}, the number of sub-units sampled from i-th unit of the t-th stratum.

k, the number of strata in the population.

Also let,

y_{tij}, be the value of the character y for the j-th sub-unit in the i-th unit of the t-th stratum.

x_{tij}, the value of the character x for the j-th sub-unit in the i-th unit of the t-th stratum.

P_{ti}, the probability of selecting the i-th unit of the t-th stratum.

Define

$$Z_{tij} = \frac{M_{ti} y_{tij}}{M_{t0} P_{ti}} \tag{1}$$

$$v_{tij} = \frac{M_{ti} x_{ti}}{M_{t0} P_{tij}}$$

where

$$\sum_{i=1}^{Nt} M_{ti} = M_{t0}; \quad \sum_{t=1}^{k} M_{t0} = M_0 \text{ and } \frac{M_{t0}}{M_0} = \lambda_t \tag{2}$$

Also let,

$\bar{Z}_{ti(mti)}$, be the sample mean of the i-th unit of the t-th stratum, so that,

$$\bar{Z}_{ti(mti)} = \frac{1}{m_{ti}} \sum_{j=1}^{mti} Z_{tij}$$

$\bar{Z}_{ti.}$, the corresponding population mean of the i-th unit of the t-th stratum, so that

$$\bar{Z}_{ti.} = \frac{1}{M_{ti}} \sum_{j=1}^{Mti} Z_{tij}$$

$\bar{Z}_{ts}$, the sample mean for the t-th stratum, so that

$$Z_{ts} = \frac{1}{n_t} \sum_{i=1}^{nt} \bar{Z}_{ti(mti)}$$

$\bar{Z}_{t..}$, the corresponding population mean, so that

$$\bar{Z}_{t..} = \frac{1}{M_{t0}} \sum_{i=1}^{Nt} \sum_{j=1}^{Mti} Z_{tij}$$

$\bar{Z}_{w}$, the sample mean, so that

$$\bar{Z}_{w} = \sum_{t=1}^{k} \frac{M_{t0} \bar{Z}_{ts}}{M_0} = \sum_{t=1}^{k} \lambda_t \bar{Z}_{ts}$$

$\bar{Z}_{..}$, the corresponding population mean.

The expression for another variate 'v' follows similarly.

The estimate of the ratio of y to x is given by

$$\hat{r} = \frac{\bar{Z}_w}{\bar{v}_w} = \frac{\sum_{t=1}^{k} \lambda_t \bar{Z}_{ts}}{\sum_{t=1}^{k} \lambda_t \bar{v}_{ts}} \tag{3}$$

and the variance of $\bar{r}$ is given by

$$\text{Var}(\bar{r})_s = \frac{1}{\bar{x}_{..}^2}\{V(\bar{Z}_w) - 2R\,\text{Cov}(\bar{Z}_w, v_w) + R^2V(\bar{v}_w)\}. \tag{4}$$

where s stands for stratification

and

$$R = \frac{\bar{y}_{..}}{\bar{x}_{..}} = \frac{\text{population mean of } y}{\text{population mean of } x}$$

The variance of $\bar{Z}_w$ is given by Sukhatme (1954) as

$$\text{Var}(\bar{Z}_w) = \sum_{t=1}^{k} \lambda_t^2 \left\{\frac{\sigma^2_{tb(Zt)}}{n_t} + \frac{1}{n_t}\sum_{i=1}^{Nt} P_{ti} \times \left(\frac{1}{m_{ti}} - \frac{1}{M_{ti}}\right) S^2_{ti(Zt)}\right\} \tag{5}$$

where

$$\sigma^2_{tb(Zt)} = \sum_{i=1}^{Nt} P_{ti}(\bar{Z}_{ti.} - \bar{Z}_{t..})^2$$

$$S^2_{ti(Z_t)} = \frac{1}{M_{ti}-1}\sum_{j=1}^{Mti}(Z_{tij} - \bar{Z}_{ti.})^2$$

Expressions for $V(\bar{v}_w)$ and $\text{Cov}(\bar{Z}_w, \bar{v}_w)$ follow similarly.

Also

$$\text{Var}(\bar{r})_{U.S.} = \frac{1}{\bar{x}_{..}^2}\{V(\bar{Z}_s) - 2R\,\text{Cov}(\bar{Z}_s\bar{v}_s) + R^2V(\bar{v}_s)\} \tag{6}$$

where $U.S.$ stands for unstratified.

The variance of $\bar{Z}_s$ is given as

$$\text{Var}(\bar{Z}_s) = \frac{\sigma^2_{b(Z)}}{n} + \frac{1}{n}\sum_{l=1}^{N} P_l\left(\frac{1}{m_l} - \frac{1}{M_l}\right) S^2_{l(Z)} \tag{7}$$

where

$$\sigma^2_{b(Z)} = \sum_{l=1}^{N} P_l(\bar{Z}_{l.} - \bar{Z}_{..})^2$$

$$S^2_{l(Z)} = \frac{1}{M_l - 1}\sum_{j=1}^{Ml}(Z_{lj} - \bar{Z}_{l.})^2 \qquad \begin{matrix} j = 1, 2, \ldots M_l \\ l = 1, 2, \cdots N \end{matrix}$$

and P_l is the probability of selecting the l-th unit. Expressions for $V(\bar{v}_s)$ and Cov $(\bar{Z}_s, \bar{v}_s)$ follow similarly

$$\therefore \quad \{\text{Var}(\hat{r})_{U.S.} - \text{Var}(\hat{r})_S\}$$

$$= \frac{1}{\bar{x}_{..}^2}\Bigg\{\left(\frac{\sigma^2_{b(Z)}}{n} - 2R\frac{\sigma_{b(Zv)}}{n} + R^2\frac{\sigma^2_{b(v)}}{n}\right)$$

$$- \sum_{t=1}^{k} \frac{\lambda_t^2}{n_t}\left(\sigma^2_{tb(Z_t)} - 2R\sigma_{tb(Z_tv_t)} + R^2\sigma^2_{tb(v_t)}\right)$$

$$+ \frac{1}{n}\sum_{1}^{N} P_l\left(\frac{1}{m_l} - \frac{1}{M_l}\right)\left(S^2_{l(Z)} - 2RS_{l(Zv)} + R^2S^2_{l(v)}\right)$$

$$- \sum_{t=1}^{k} \frac{\lambda_t^2}{n_t}\sum_{1}^{N_t} P_{ti}\left(\frac{1}{m_{ti}} - \frac{1}{M_{ti}}\right)\left(S^2_{ti(Z_t)} - 2RS_{ti(Z_tv_t)}\right.$$

$$\left. + R^2S^2_{ti(v_t)}\right)\Bigg\} \tag{8}$$

The estimate of $V(\bar{Z}_s)$ and $V(\bar{v}_s)$ without stratification have already been derived by Sukhatme (1954). To estimate (8) we have to estimate Cov $(\bar{Z}_s, \bar{v}_s)$ without stratification in terms of the units included in the sample with stratification.

Now,

$$\text{Cov}(\bar{Z}_s, \bar{v}_s)_{U.S.} = \frac{\sigma_{b(Zv)}}{n} + \frac{1}{n}\sum_{l=1}^{N} P_l\left(\frac{1}{m_l} - \frac{1}{M_l}\right) S_{l(Zv)} \tag{9}$$

where

$$\sigma_{b(Z,v)} = \sum_{l=1}^{N} P_l(\bar{Z}_{l.} - \bar{Z}_{..})(\bar{v}_{l.} - \bar{v}_{..}) \tag{10}$$

and

$$S_{l(Z,v)} = \frac{1}{M_l - 1}\sum_{j=1}^{M_l}(Z_{lj} - \bar{Z}_{l.})(v_{lj} - \bar{v}_{l.}) \tag{11}$$

Let, the l-th unit in the population correspond to the i-th unit of the t-th stratum.

Then,

$$P_{ti} = \frac{P_l}{P_{t.}}$$

where

$$P_{t.} = \sum_{l=1}^{N_t} P_l$$

and

$$Z_{tj} = \frac{M_{ti}}{M_0} \frac{y_{tij}}{P_{t.} P_{ti}} = \frac{Z_{tij}}{P_{t.}} \frac{M_{t0}}{M_0} = \lambda_t \frac{Z_{tij}}{P_{t.}}$$

similarly

$$v_{tj} = \frac{M_{ti}}{M_0} \frac{x_{tij}}{P_{t.} P_{ti}} = \frac{v_{tij}}{P_{t.}} \frac{M_{t0}}{M_0} = \lambda_t \frac{v_{tij}}{P_{t.}}$$

$$\therefore \sigma_{b(Z,v)} = \sum_{t=1}^{k} \sum_{i=1}^{N_t} P_{t.} P_{ti} \left(\frac{\lambda_t}{P_{t.}} \bar{Z}_{ti.} - \bar{Z}_{..} \right) \left(\frac{\lambda_t}{P_{t.}} \bar{v}_{ti.} - \bar{v}_{..} \right)$$

$$= \sum_{t=1}^{k} P_{t.} \sum_{i=1}^{N_t} P_{ti} \left(\frac{\lambda_t}{P_{t.}} \bar{Z}_{ti.} - \frac{\lambda_t}{P_{t.}} \bar{Z}_{t..} + \frac{\lambda_t}{P_{t.}} \bar{Z}_{t..} - \bar{Z}_{..} \right)$$

$$\times \left(\frac{\lambda_t}{P_{t.}} \bar{v}_{ti.} - \frac{\lambda_t}{P_{t.}} \bar{v}_{t..} + \frac{\lambda_t}{P_{t.}} \bar{v}_{t..} - \bar{v}_{..} \right)$$

$$= \sum_{t=1}^{k} P_{t.} \sum_{i=1}^{N_t} P_{ti} \left\{ \frac{\lambda_t^2}{P_{t.}^2} (\bar{Z}_{ti.} - \bar{Z}_{t..}) (v_{ti.} - \bar{v}_{t..}) \right.$$

$$\left. + \left(\frac{\lambda_t}{P_{t.}} \bar{Z}_{t..} - \bar{Z} \right) \left(\frac{\lambda_t}{P_{t.}} \bar{v}_{t..} - \bar{v}_{.} \right) \right\}$$

$$= \sum_{t=1}^{k} \frac{\lambda_t^2}{P_{t.}} \sigma_{tb(Z_t v_t)} + \sum_{t=1}^{k} \frac{\lambda_t^2}{P_{t.}} \bar{Z}_{t..} \bar{v}_{t..} - \bar{Z}_{..} \bar{v}_{..} \qquad (12)$$

To get an estimate of $\sigma_{b(Z,v)}$ we have to estimate the second and third terms in (12).

Now, we have

$$\text{Cov} (\bar{Z}_{ts}, \bar{v}_{ts}) = E (\bar{Z}_{ts} \cdot \bar{v}_{ts}) - \bar{Z}_{t..} \bar{v}_t$$

$$= \frac{\sigma_{tb(Z_t v_t)}}{n_t} + \frac{1}{n_t} \sum_{i=1}^{N_t} P_{ti} \left(\frac{1}{m_{ti}} - \frac{1}{M_{ti}} \right) S_{ti(Z_t v_t)}$$

$$\therefore \quad \text{Est. } \bar{Z}_{t..}\,\bar{v}_{t..} = \bar{Z}_{ts}\,\bar{v}_{ts} - \frac{\hat{\sigma}_{tb\,(Z_t v_t)}}{n_t}$$

$$- \frac{1}{n_t^2} \sum_{i=1}^{n_t} \left(\frac{1}{m_{ti}} - \frac{1}{M_{ti}} \right) s_{ti\,(Z_t v_t)} \tag{13}$$

Where

$$\hat{\sigma}_{tb\,(Z_t v_t)} = s_{tb\,(Z_t v_t)} - \frac{1}{n_t} \sum_{i=1}^{n_t} \left(\frac{1}{m_{ti}} - \frac{1}{M_{ti}} \right) s_{ti\,(Z_t v_t)} \tag{14}$$

where

$$s_{tb\,(Z_t v_t)} = \frac{1}{n_t - 1} \sum_{i=1}^{n_t} (\bar{Z}_{ti\,(m_{ti})} - \bar{Z}_{ts})(\bar{v}_{ti\,(m_{ti})} - \bar{v}_{ts})$$

and

$$s_{ti\,(Z_t v_t)} = \frac{1}{m_{ti} - 1} \sum_{j=1}^{m_{ti}} (Z_{tij} - \bar{Z}_{ti\,(m_{ti})})(v_{tij} - \bar{v}_{ti\,(m_{ti})})$$

Similarly

$$\text{Cov}\,(\bar{Z}_w, \bar{v}_w) = \sum_{t=1}^{k} \lambda_t^2 \left\{ \frac{\sigma_{tb\,(Z_t v_t)}}{n_t} \right.$$

$$\left. + \frac{1}{n_t} \sum_{i=1}^{N_t} P_{ti} \left(\frac{1}{m_{ti}} - \frac{1}{M_{ti}} \right) S_{ti\,(Z_t v_t)} \right\}$$

$$= E\,(\bar{Z}_w \cdot \bar{v}_w) - Z_{..}\bar{v}_{..}$$

$$\text{Est. } \bar{Z}_{..}\bar{v}_{..} = \bar{Z}_w \cdot \bar{v}_w - \sum_{t=1}^{k} \lambda_t^2 \left\{ \frac{\hat{\sigma}_{tb\,(Z_t v_t)}}{n_t} \right.$$

$$\left. + \frac{1}{n_t^2} \sum_{i=1}^{n_t} \left(\frac{1}{m_{ti}} - \frac{1}{M_{ti}} \right) s_{ti\,(Z_t v_t)} \right\} \tag{15}$$

Also

$$P_i S_{I\,(Zv)} = \frac{\lambda_t^2}{P_{t.}} P_{ti} S_{ti\,(Z_t v_t)} \tag{16}$$

Hence from (12) and (16)

$$\text{Cov}\ (\bar{Z}_s, \bar{v}_s)_{U.S.} = \sum_{t=1}^{k} \frac{\lambda_t^2}{P_{t.}} \frac{\sigma_{tb\,(Z_t v_t)}}{n} + \frac{1}{n}\left(\sum_{t=1}^{k} \frac{\lambda_t^2}{P_{t.}} \bar{Z}_{t..}\bar{v}_{t..} - \bar{Z}_{..}\,\bar{v}_{..}\right)$$

$$+ \frac{1}{n} \sum_{t=1}^{k} \frac{\lambda_t^2}{P_{t.}} \sum_{i=1}^{Nt} P_{ti} \left(\frac{1}{m_{ti}} - \frac{1}{M_{ti}}\right) S_{ti\,(Z_t v_t)} \quad (17)$$

Hence substituting the estimated value from (13), (15) and (16) and writing estimate of

$\sigma_{tb\,(Z_t v_t)}$ as $\hat{\sigma}_{tb\,(Z_t v_t)}$

we get

$$\text{Est. Cov}\ (\bar{Z}_s \bar{v}_s)_{U.S.} = \sum_{t=1}^{k} \frac{\lambda_t^2}{P_{t.}} \frac{\hat{\sigma}_{tb\,(Z_t v_t)}}{n}$$

$$+ \frac{1}{n}\left\{\sum_{t=1}^{k} \frac{\lambda_t^2}{P_{t.}} \bar{Z}_{ts}\,\bar{v}_{ts} - \bar{Z}_w\,\bar{v}_w\right.$$

$$- \sum_{t=1}^{k} \frac{\lambda_t^2}{n_t} \left(\frac{1}{P_{t.}} - 1\right) \hat{\sigma}_{tb\,(Z_t v_t)}$$

$$- \sum_{t=1}^{k} \frac{\lambda_t^2}{n_t^2} \left(\frac{1}{P_{t.}} - 1\right)$$

$$\left.\times \sum^{nt} \left(\frac{1}{m_{ti}} - \frac{1}{M_{ti}}\right) s_{ti\,(Z_t v_t)}\right\}$$

$$+ \frac{1}{n} \sum_{t=1}^{k} \frac{\lambda_t^2}{P_{t.}} \frac{1}{n_t} \sum_{i=1}^{nt} \left(\frac{1}{m_{ti}} - \frac{1}{M_{ti}}\right) s_{ti\,(Z_t v_t)}. \quad (18)$$

which after substitution for $\hat{\sigma}_{tb\,(Z_t v_t)}$ from (14) reduces to

$$\text{Est. Cov}\ (\bar{Z}_s, \bar{v}_s)_{U.S.} = \frac{1}{n}\left(\sum_{t=1}^{k} \frac{\lambda_t^2}{P_{t.}} \bar{Z}_{ts}\,\bar{v}_{ts} - \bar{Z}_w \bar{v}_w\right)$$

$$+ \sum_{t=1}^{k} \frac{\lambda_t^2}{n_t} \left(\frac{n_t}{nP_{t.}} - \frac{1}{nP_{t.}} + \frac{1}{n}\right) s_{tb\,(Z_t v_t)} \quad (19)$$

Also

$$\text{Cov}\,(\bar{Z}_w, \bar{v}_w)_S = \sum_{t=1}^{k} \lambda_t^2 \left\{ \frac{\sigma_{tb\,(Z_t v_t)}}{n_t} \right.$$

$$\left. + \frac{1}{n_t} \sum_{i=1}^{N_t} P_{ti} \left(\frac{1}{m_{ti}} - \frac{1}{M_{ti}} \right) S_{ti\,(Z_t v_t)} \right\} \tag{20}$$

and its estimated value is given by

$$\text{Est. Cov}\,(\bar{Z}_w, \bar{v}_w)_S = \sum_{t=1}^{k} \frac{\lambda_t^2}{n_t} s_{tb\,(Z_t v_t)} \tag{21}$$

Hence from (19) and (21)

$$\text{Est.}\,\{\text{Cov}\,(\bar{Z}_s, \bar{v}_s)_{U.S.} - \text{Cov}\,(\bar{Z}_w, \bar{v}_w)_S\}$$

$$= \frac{1}{n} \left(\sum_{t=1}^{k} \frac{\lambda_t^2}{P_{t.}} \bar{Z}_{ts}\, \bar{v}_{ts} - \bar{Z}_w \bar{v}_w \right)$$

$$+ \sum_{t=1}^{k} \frac{\lambda_t^2}{n_t} \left(\frac{n_t}{nP_{t.}} - 1 - \frac{1}{nP_{t.}} + \frac{1}{n} \right) s_{tb\,(Z_t v_t)} \tag{22}$$

From Sukhatme (1954)

$$\text{Est.}\,\{\text{Var}\,(\bar{Z}_s)_{U.S.} - \text{Var}\,(\bar{Z}_w)_s\}$$

$$= \frac{1}{n} \left(\sum_{t=1}^{k} \frac{\lambda_t^2}{P_{t.}} \bar{Z}_{ts}^2 - \bar{Z}_w^2 \right)$$

$$+ \sum_{t=1}^{k} \frac{\lambda_t^2}{n_t} \left(\frac{n_t}{nP_{t.}} - 1 - \frac{1}{nP_{t.}} + \frac{1}{n} \right) s^2_{tb\,(Z_t)} \tag{23}$$

Using similar expressions for the variate 'v'.

A consistent estimate of $\{\text{Var}\,(\bar{r})_{U.S.} - \text{Var}\,(\bar{r})_S\}$ is given by

$$\text{Est.}\,\{\text{Var}\,(\bar{r})_{U.S.} - \text{Var}\,(\bar{r})_S\}$$

$$= \frac{1}{\bar{v}_w^2} \left[\frac{1}{n} \left\{ \left(\sum_{t=1}^{k} \frac{\lambda_t^2}{P_{t.}} \bar{Z}_{ts}^2 - \bar{Z}_w^2 \right) \right. \right.$$

$$- 2\bar{r} \left(\sum_{t=1}^{k} \frac{\lambda_t^2}{P_{t.}} \bar{Z}_{ts}\, \bar{v}_{ts} - \bar{Z}_w\, \bar{v}_w \right) + \bar{r}^2 \left(\sum_{t=1}^{k} \frac{\lambda_t^2}{P_{t.}} \bar{v}_{ts}^2 - \bar{v}_w^2 \right)$$

$$+\sum^{k}\frac{\lambda_t^2}{n_t}\left(\frac{n_t}{nP_{t.}}-1-\frac{1}{nP_{t.}}+\frac{1}{n}\right)$$

$$\times\left(s^2_{tb\,(Z_t)}-2\bar{r}s_{tb\,(Z_tv_t)}+\bar{r}^2s^2_{tb\,(v_t)}\right)\Big\}\Big] \tag{24}$$

If

$$P_l=\frac{M_l}{M_0},$$

then

$$P_{ti}=\frac{P_l}{\sum\limits_{l=1}^{Nt}P_l}=\frac{M_l}{\sum\limits_{l=1}^{Nt}M_l}=\frac{M_{ti}}{M_{t0}}$$

It follows immediately that

$$Z_{lj}=y_{lj}=y_{tij}=Z_{tij}$$

Also

$$P_{t.}=\sum_{l=1}^{Nt}P_l=\sum_{l=1}^{Nt}\frac{M_l}{M_0}=\frac{M_{t0}}{M_0}=\lambda_t$$

Then (24) reduces to

Est. $\{\text{Var}\,(\bar{r})_{U.S.}-\text{Var}\,(\bar{r})_{S.}\}$

$$=\frac{1}{\bar{x}_{..}^2}\left[\frac{1}{n}\left\{\left(\sum_{t=1}^{k}\lambda_t\bar{y}_{ts}^2-\bar{y}_w^2\right)-2\bar{r}\left(\sum_{t=1}^{k}\lambda_t\bar{y}_{ts}\bar{x}_{ts}-\bar{y}_w\bar{x}_w\right)\right.\right.$$

$$+\bar{r}^2\left(\sum_{t=1}^{k}\lambda_t\bar{x}_{ts}^2-\bar{x}_w^2\right)+\sum_{t=1}^{k}\frac{\lambda_t^2}{n_t}\left(\frac{n_t}{nP_{t.}}-1-\frac{1}{nP_{t.}}+\frac{1}{n}\right)$$

$$\times\left(s^2_{tb\,(y)}-2\bar{r}s_{tb\,(yx)}+\bar{r}^2s^2_{tb\,(x)}\right)\Big\}\Big] \tag{25}$$

Example

A sample survey for estimating the live-stock population was conducted during the year 1954 by Indian Council of Agricultural Research in the District of Poona, Bombay State. All the villages in the district were classified into three strata (1) villages having $\leqslant 75$ occupied houses, (2) villages having more than 75 but $\leqslant 200$ occupied houses, and finally, (3) villages having more than 200 occupied houses. From each stratum a sample of villages was selected with probability proportional to the number of occupied houses in the village with replacement.

Each selected village was divided into clusters of 5 consecutive serially ordered houses. From the clusters so formed a simple random sample of 4 clusters was drawn (without replacement). Since the clusters were not of uniform sizes, so it is assumed for the purpose of illustration that a certain number of houses were selected at random from each selected village. The geographical area (expressed in acres) of the village was regarded as a supplementary variate.

Table A shows the number of villages and the number of occupied houses in the population, the number of villages in the sample, the estimated live stock number per house and the values for $s^2_{tb(y)}$, $s^2_{tb(x)}$ and $s_{tb(yx)}$ for each stratum.

TABLE A

Live-Stock Sample Survey, Poona, 1954

Stratum	N_t	M_{to}	n_t	$\bar{y}_{ts}$	$\bar{x}_{ts}$	$s^2_{tb(y)}$	$s^2_{tb(x)}$	$s_{tb(xy)}$
1 ..	628	26070	10	5·631	0·0698	1·6977	0·0057	0·0426
2 ..	563	69781	20	4·407	0·0264	3·8045	0·0001	0·0120
3 ..	281	103429	30	3·245	0·0234	2·4450	0·0001	0·0039
Sum ..	1472	199280	60	..	..	..	..	..

The variance of the estimate of the ratio in the district is given by

$$\text{Est. Var }(\bar{r})_S = \frac{1}{\bar{x}_{..}^2}\left\{\sum_{t=1}^{k}\frac{\lambda_t^2}{n_t}\left(s^2_{tb(y)} - 2\bar{r}s_{tb(xy)} + \bar{r}^2 s^2_{tb(x)}\right)\right\}$$

The weights λ_t are computed in the first column of Table B.

TABLE B

Computation for Estimating Gain in Precision due to Stratification

Stratum	λ_t	$\lambda_t\bar{y}_{ts}$	$\lambda_t\bar{y}_{ts}^2$	$\lambda_t\bar{x}_{ts}$	$\lambda_t\bar{x}_{ts}^2$	$\lambda_t\bar{y}_{ts}\bar{x}_{ts}$
1 ..	0·13082	0·7366	4·1478	0·00913	0·00064	0·05141
2 ..	0·35017	1·5432	6·8009	0·00924	0·00024	0·04071
3 ..	0·51901	1·6842	5·4652	0·01216	0·00028	0·03946
Sum ..	..	3·9640	16·4139	0·03053	0·00116	0·13158

Substituting from the table we have,

Est. Var $(\bar{r})_S = 170\cdot9224$.

The difference between the variance of the ratio of an unstratified sample and that of a stratified sample is estimated (consistent estimate) by equation (25) as $43\cdot1552$. The necessary computations are made in Table B.

Thus the relative increase in variance if the sample were not stratified would be

$$\frac{43\cdot1552}{170\cdot9224} = 0\cdot2525$$

or $25\cdot2\%$.

Summary

It is generally felt that in a sub-sampling design, where the primary units vary largely in their sizes, the selection of primary units with varying probabilities leads to a more efficient estimate than sampling with equal probabilities. So, following Sukhatme (1954) who has worked out the efficiency of stratification with varying probabilities of selection with replacement for a single variate, formulæ are derived in this paper, to estimate the efficiency of stratification for the ratio method of estimation when sampling is done with varying probabilities of selection and with replacement.

Acknowledgment

We are grateful to Dr. G. R. Seth and Dr. K. R. Nair for their guidance in the preparation of this paper.

References

1. Sukhatme, P. V. .. *Sampling Theory of Surveys with Application.* Ind. Soc. Agric. Stats., New Delhi and Iowa State College Press, Ames, Iowa, U.S.A.
2. Mokashi, V. K. .. "Efficiency of stratification in sub-sampling design for the Ratio Method of Estimation," *Ind. Soc. Agric. Statis.*, 1954.
3. I.C.A.R., New Delhi .. *Report of the Scheme for Improvement of Livestock Statistics, Bombay State,* 1954.