

## A Generalised Class of Unbiased Separate Regression Type Estimator under Stratified Random Sampling

Shashi Bhushan<sup>1</sup>, Praveen Kumar Misra<sup>1</sup>

<sup>1</sup>Department of Mathematics and Statistics, Dr. Shakuntala Misra National Rehabilitation University, India

**Abstract:** In this paper a generalized class of regression type estimators using the auxiliary information on population mean and population variance is proposed under stratified random sampling. In order to improve the performance of the proposed class of estimator, the Jack-knifed versions are also proposed. A comparative study of the proposed estimator is made with that of separate ratio estimator, separate product estimator, separate linear regression estimator and the usual stratified sample mean. It is shown that the estimators through proposed allocation always give more efficient estimators in the sense of having smaller mean square error than those obtained through Neyman Allocation.

**Keywords:** Auxiliary information, ratio type Estimator, Bias, Mean Square Error.

### I. Introduction of the Proposed Estimator

Let a population of size 'N' be stratified into 'L' non-overlapping strata, the  $h^{\text{th}}$  stratum size being  $N_h$  ( $h=1,2,\dots,L$ ) and  $\sum_{h=1}^L N_h = N$ . Suppose 'y' be characteristic under study and 'x' be the auxiliary variable. We denote by

$y_{hj}$ : The observation on the  $j^{\text{th}}$  unit of the population for the characteristic 'y' under study ( $j=1,2,\dots,N_h$ ) in the  $h^{\text{th}}$  stratum ( $h=1,2,\dots,L$ )

$x_{hj}$ : The observation on the  $j^{\text{th}}$  unit of the population for the auxiliary characteristic 'x' under study ( $j=1,2,\dots,N_h$ ) in the  $h^{\text{th}}$  stratum ( $h=1,2,\dots,L$ ).

$$\bar{Y}_h = \frac{1}{N_h} \sum_{j=1}^{N_h} y_{hj}, \bar{X}_h = \frac{1}{N_h} \sum_{j=1}^{N_h} x_{hj},$$

$$S_{yh}^2 = \frac{1}{(N_h - 1)} \sum_{j=1}^{N_h} (y_{hj} - \bar{Y}_h)^2, S_{xh}^2 = \frac{1}{(N_h - 1)} \sum_{j=1}^{N_h} (x_{hj} - \bar{X}_h)^2$$

$$\sigma_{xh}^2 = \frac{1}{N_h} \sum_{j=1}^{N_h} (x_{hj} - \bar{X}_h)^2, \sigma_{yh}^2 = \frac{1}{N_h} \sum_{j=1}^{N_h} (y_{hj} - \bar{Y}_h)^2$$

$$S_{xyh} = \frac{1}{(N_h - 1)} \sum_{j=1}^{N_h} (x_{hj} - \bar{X}_h) (y_{hj} - \bar{Y}_h) = \rho_h S_{xh} S_{yh}$$

where  $\rho_h$  is the population correlation coefficient between 'x' and 'y' for the  $h^{\text{th}}$  stratum ( $j=1,2,\dots,N_h$ ).

$$R_h = \frac{\bar{Y}_h}{\bar{X}_h}, C_{yh}^2 = \frac{S_{yh}^2}{\bar{Y}_h^2} = \frac{\mu_{02h}}{\bar{Y}_h^2}, C_{xh}^2 = \frac{S_{xh}^2}{\bar{X}_h^2} = \frac{\mu_{20h}}{\bar{X}_h^2}$$

$\mu_{pqh} = \frac{1}{N_h} \sum_{j=1}^{N_h} (x_{hj} - \bar{X}_h)^p (y_{hj} - \bar{Y}_h)^q$ : the  $(p, q)^{\text{th}}$  population product moment about mean between 'x' and 'y' for the  $h^{\text{th}}$  stratum ( $h=1,2,\dots,L$ ).

$$\beta_{1h} = \frac{\mu_{30h}^2}{\mu_{20h}^2}, \beta_{2h} = \frac{\mu_{40h}^2}{\mu_{20h}^2},$$

$\beta_h = \frac{S_{xyh}}{S_{xh}^2} = \rho_h \frac{S_{yh}}{S_{xh}}$  be the population regression coefficient of y on x for the  $h^{\text{th}}$  stratum ( $h=1,2,\dots,L$ ).

Let a simple random sample of size  $n_h$  be selected from the  $h^{th}$  stratum without replacement, without any loss of generality, we assume that first  $n_h$  units have been selected in the  $h^{th}$  stratum from  $N_h$  units by SRSWOR. Moreover we assume that  $N_h$  is so large that  $1 - f_h \approx 1$ .

We define

$$\bar{y}_h = \frac{1}{n_h} \sum_{j=1}^{n_h} y_{hj}, \bar{x}_h = \frac{1}{n_h} \sum_{j=1}^{n_h} x_{hj}, s_{xh}^2 = \frac{1}{n_h - 1} \sum_{j=1}^{n_h} (x_{hj} - \bar{x}_h)^2, s_{yh}^2 = \frac{1}{n_h - 1} \sum_{j=1}^{n_h} (y_{hj} - \bar{y}_h)^2$$

$$\hat{\sigma}_{xh}^2 = \frac{1}{n_h} \sum_{j=1}^{n_h} (x_{hj} - \bar{x}_h)^2, \hat{\sigma}_{yh}^2 = \frac{1}{n_h} \sum_{j=1}^{n_h} (y_{hj} - \bar{y}_h)^2, s_{xyh} = \frac{1}{n_h - 1} \sum_{j=1}^{n_h} (x_{hj} - \bar{x}_h)(y_{hj} - \bar{y}_h), b_h = \frac{s_{xyh}}{s_{xh}^2}$$

Assuming that  $\bar{X}_h$  is known  $\forall h = 1, 2, \dots, L$ . The proposed generalized estimator  $\hat{Y}_{gS}$  for estimating the population mean  $\bar{Y}$  of the study variable is given by

$$\hat{Y}_{gS} = \sum_{j=1}^L W_h \left\{ \bar{y}_h g(\bar{w}_h) + b_h (\bar{X}_h - \bar{x}_h) \right\} \tag{1.1}$$

where  $\bar{w}_h = \frac{\hat{\sigma}_{xh}^2}{\sigma_{xh}^2}$  and  $g(\bar{w}_h)$  is such that  $g(\bar{w}_h) = 1$  at  $w_h = 1$ , is a function of  $w_h$  satisfying the following conditions.

1. whatever be the sample chosen  $w_h$  assumes values in the bounded closed interval 'I' of the real line containing the point unity
2. In the interval 'I' the function  $g(\bar{w}_h)$  is continuous and bounded.
3. The first, second and third order derivatives of  $g(\bar{w}_h)$  exist and are continuous.

Strata means  $\bar{X}_h$  and strata variances  $\sigma_{xh}^2$  of the auxiliary variables x are assumed to be known. It should be noted that for  $g(\bar{w}_h) = 1$  the proposed generalized estimator reduces to the separate linear regression estimator given by

$$\bar{y}_{LRS} = \sum_{j=1}^L W_h \left\{ \bar{y}_h + b_h (\bar{X}_h - \bar{x}_h) \right\} \tag{1.2}$$

## II. Bias And Mean Square Error of the Proposed Estimator $\hat{Y}_{gS}$

Expanding  $g(\bar{w}_h)$  about the point  $w_h = 1$  in the third order Taylor's series from (1.1)

$$\hat{Y}_{gS} = \sum_{h=1}^L W_h \left[ \bar{y}_h \left\{ g(1) + (w_h - 1)g'(1) + \frac{(w_h - 1)^2}{2!}g''(1) + \frac{(w_h - 1)^3}{3!}g'''(w_h^*) \right\} + b_h (\bar{X}_h - \bar{x}_h) \right] \tag{2.1}$$

where  $w_h^* = 1 + \theta(w_h - 1); 0 < \theta < 1$  and  $\theta$  may depend on  $w_h$ ;  $g'(1), g''(1), g'''(w_h^*)$  denotes first second and third order partial derivatives of  $g(\bar{w}_h)$  at points  $w=1, 1, w^*$  respectively.

Let

$$\bar{y}_h - \bar{Y} = e_{0h}, \bar{x}_h - \bar{X} = e_{1h}, s_{xyh} - S_{xyh} = e_{2h}, s_{xh}^2 - S_{xh}^2 = e_{3h}, \hat{\sigma}_{xh}^2 - \sigma_{xh}^2 = e_{4h}$$

$$E(e_{0h}) = E(e_{1h}) = E(e_{2h}) = E(e_{3h}) = E(e_{4h}) = 0; \forall h = 1, 2, \dots, L$$

Now, from (5.2.1), we have

$$\hat{Y}_{gS} = \sum_{j=1}^L W_h \left\{ \bar{Y}_h + e_{0h} + g'(1)\bar{Y}_h \left( \frac{e_{4h}}{\sigma_{xh}^2} + \frac{e_{0h}e_{4h}}{\bar{Y}_h \sigma_{xh}^2} \right) + \frac{e_{xh}^2}{2\sigma_{xh}^4} g''(1)\bar{Y}_h \right\}$$

$$\left\{ \beta_h \left( 1 + \frac{e_{2h}}{S_{xyh}} \right) \left( 1 - \frac{e_{3h}}{S_{xh}^2} + \dots \right) (-e_{1h}) \right\}$$

$$\hat{Y}_{gS} = \sum_{j=1}^L W_h \left\{ (\bar{Y}_h + e_{0h}) \left\{ g(1) + \frac{e_{4h}}{\sigma_{xh}^2} g'(1) + \frac{e_{xh}^2}{2\sigma_{xh}^4} g''(1) + \frac{e_{xh}^3}{6\sigma_{xh}^6} g'''(w^*) \right\} \right\}$$

$$\left\{ + \beta_h \left( 1 + \frac{e_{2h}}{S_{xyh}} \right) \left( 1 + \frac{e_{3h}}{S_{xh}^2} \right)^{-1} (-e_{1h}) \right\}$$

$$\hat{Y}_{gs} = \sum_{j=1}^L W_h \left\{ \bar{Y}_h + e_{0h} + g'(1) \bar{Y}_h \left( \frac{e_{4h}}{\sigma_{xh}^2} + \frac{e_{0h}e_{4h}}{\bar{Y}_h \sigma_{xh}^2} \right) + \frac{e_{xh}^2}{2\sigma_{xh}^4} g''(1) \bar{Y}_h + \beta_h \left( -e_{1h} - \frac{e_{1h}e_{2h}}{S_{xyh}} + \frac{e_{1h}e_{3h}}{S_{xh}^2} + \dots \right) \right\} \quad (2.2)$$

Let the sample size be so large that  $|e_i|$ ,  $i=0,1,2,3,4$ ;  $\forall h=1,2,\dots,L$ ; becomes so small that terms of  $e_i$  s having powers greater than two may be neglected.

$$E(\hat{Y}_{gs}) = \sum_{j=1}^L W_h \left\{ \bar{Y}_h + g'(1) \frac{E(e_{0h}e_{4h})}{\sigma_{xh}^2} + \bar{Y}_h \frac{g''(1)E(e_{xh}^2)}{2\sigma_{xh}^4} + \beta_h \left( \frac{E(e_{1h}e_{3h})}{S_{xh}^2} - \frac{E(e_{1h}e_{2h})}{S_{xyh}} \right) \right\}$$

Using the results given in Sukhatme and Sukhatme (1997) and proved in appendix

$$E(e_{1h}e_{3h}) = \left( \frac{1}{n_h} - \frac{1}{N_h} \right) \mu_{30h}, E(e_{1h}e_{2h}) = \left( \frac{1}{n_h} - \frac{1}{N_h} \right) \mu_{21h}$$

$$E(e_{0h}e_{4h}) = \left( \frac{1}{n_h} - \frac{1}{N_h} \right) \mu_{21h}, E(e_{xh}^2) = \left( \frac{1}{n_h} - \frac{1}{N_h} \right) (\mu_{40h} - \mu_{20h}^2); \forall h=1,2,\dots,L$$

we have,

$$E(\hat{Y}_{gs}) = \sum_{h=1}^L W_h \left[ \bar{Y}_h + \left( \frac{1}{n_h} - \frac{1}{N_h} \right) \left\{ \frac{g'(1)\mu_{21h}}{\sigma_{xh}^2} + \frac{\bar{Y}_h g''(1)}{2\sigma_{xh}^4} (\mu_{40h} - \mu_{20h}^2) + \beta_h \left( \frac{\mu_{30h}}{S_{xh}^2} - \frac{\mu_{21h}}{S_{xyh}} \right) \right\} \right] \quad (2.3)$$

showing that  $\hat{Y}_{gs}$  is a biased estimator of population mean  $\bar{Y}$  and its bias is given by

$$B(\hat{Y}_{gs}) = E(\hat{Y}_{gs}) - \bar{Y}_h$$

$$B(\hat{Y}_{gs}) = \sum_{h=1}^L W_h \left[ \left( \frac{1}{n_h} - \frac{1}{N_h} \right) \left\{ \frac{g'(1)\mu_{21h}}{\sigma_{xh}^2} + \frac{\bar{Y}_h g''(1)}{2\sigma_{xh}^4} (\mu_{40h} - \mu_{20h}^2) + \beta_h \left( \frac{\mu_{30h}}{S_{xh}^2} - \frac{\mu_{21h}}{S_{xyh}} \right) \right\} \right] \quad (2.4)$$

The mean square error of  $\hat{Y}_{gs}$  is given by

$$MSE(\hat{Y}_{gs}) = E(\hat{Y}_{gs} - \bar{Y})^2 = E \left\{ \sum_{j=1}^L W_h \left( e_{0h} + g'(1) \bar{Y}_h \frac{e_{4h}}{\sigma_{xh}^2} - \beta_h e_{1h} \right)^2 \right\}$$

(Using (5.2.2) to the first order of approximation)

$$MSE(\hat{Y}_{gs}) = \sum_{j=1}^L W_h \left\{ E(e_{0h}^2) + E \left( \frac{g'(1)^2 \bar{Y}_h^2}{\sigma_{xh}^2} e_{4h}^2 \right) + E(\beta_h^2 e_{1h}^2) + 2 \frac{\bar{Y}_h g'(1)}{\sigma_{xh}^2} E(e_{0h}e_{4h}) \right. \\ \left. - 2\beta_h E(e_{0h}e_{1h}) - 2 \frac{g'(1) \bar{Y}_h \beta_h}{\sigma_{xh}^2} E(e_{1h}e_{4h}) \right\}$$

Substituting the following results given in Sukhatme and Sukhatme (1997) and proved in appendix

$$E(e_{0h}^2) = \left( \frac{1}{n_h} - \frac{1}{N_h} \right) S_{yh}^2, E(e_{1h}^2) = \left( \frac{1}{n_h} - \frac{1}{N_h} \right) S_{xh}^2, E(e_{0h}e_{1h}) = \left( \frac{1}{n_h} - \frac{1}{N_h} \right) S_{xyh}$$

$$E(e_{xh}^2) = \left( \frac{1}{n_h} - \frac{1}{N_h} \right) (\mu_{40h} - \mu_{20h}^2), E(e_{0h}e_{4h}) = \left( \frac{1}{n_h} - \frac{1}{N_h} \right) \mu_{21h}$$

$$E(e_{1h}e_{4h}) = \left( \frac{1}{n_h} - \frac{1}{N_h} \right) \mu_{30h}; \forall h=1,2,\dots,L$$

we have

$$MSE(\hat{Y}_{gs}) = \sum_{j=1}^L W_h \left( \frac{1}{n_h} - \frac{1}{N_h} \right) \left\{ S_{yh}^2 + \frac{\{g'(1)\}^2 \bar{Y}_h^2}{\sigma_{xh}^4} (\mu_{40h} - \mu_{20h}^2) + \beta_h^2 S_{xh}^2 \right. \\ \left. + \frac{2g'(1) \bar{Y}_h}{\sigma_{xh}^2} \mu_{21h} - 2\beta_h S_{xyh} - \frac{2g'(1) \bar{Y}_h \beta_h}{\sigma_{xh}^2} \mu_{30h} \right\} \quad (2.5)$$

(2.5) is minimum when

$$g'(1)\bar{Y}_h = \frac{(\beta_h \mu_{20h} - \mu_{21h})}{(\mu_{40h} - \mu_{20h}^2)} \mu_{20h}; \forall h = 1, 2, \dots, L \quad (2.6)$$

and the minimum mean square error of  $\hat{Y}_{gs}$  is given by

$$\begin{aligned} MSE(\hat{Y}_{gs})_{opt} &= \sum_{j=1}^L W_h \left( \frac{1}{n_h} - \frac{1}{N_h} \right) \left\{ (1 - \rho_h^2) S_{yh}^2 - \frac{(\beta_h \mu_{30h} - \mu_{21h})^2}{(\mu_{40h} - \mu_{20h}^2)} \right\} \\ MSE(\hat{Y}_{gs})_{opt} &= \sum_{j=1}^L W_h \left( \frac{1}{n_h} - \frac{1}{N_h} \right) \left\{ (1 - \rho_h^2) S_{yh}^2 - \frac{(\beta_h \mu_{30h} - \mu_{21h})^2}{\mu_{20h}^2 (\beta_{2h} - 1)} \right\} \end{aligned} \quad (2.7)$$

### III. Optimum Allocation With The Proposed Class

Consider the cost function  $C = C_0 + \sum_{h=1}^L c_h n_h$ , where  $C_0$  is the fixed cost and  $c_h$  be the cost of drawing per unit sample within  $h^{th}$  stratum respectively, we have

$$V(\hat{Y}_{gs})_{min} = \sum_{j=1}^L \frac{W_h^2}{n_h} \left\{ (1 - \rho_h^2) S_{yh}^2 - \frac{(\beta_h \mu_{30h} - \mu_{21h})^2}{(\mu_{40h} - \mu_{20h}^2)} \right\} \quad (3.1)$$

we wish to choose  $n_h$  such that  $V(\hat{Y}_{gs})_{min}$  is further least for the fixed cost. To achieve this objective, we apply the Lagrange's method of multipliers for maxima and minima. Accordingly, we define

$$\phi = V(\hat{Y}_{gs})_{min} + \lambda \left( \sum_{h=1}^L c_h n_h - C + C_0 \right) \quad (3.2)$$

where  $\lambda$  is a constant, known as Lagrange's multiplier.

Differentiating (5.3.2) with respect to  $n_h$  and then equating it to zero, we get

$$\begin{aligned} -\frac{W_h^2}{n_h^2} \left( \frac{1}{n_h} - \frac{1}{N_h} \right) \left\{ (1 - \rho_h^2) S_{yh}^2 - \frac{(\beta_h \mu_{30h} - \mu_{21h})^2}{\mu_{20h}^2 (\beta_{2h} - 1)} \right\} + \lambda c_h &= 0 \\ \text{or } n_h = \frac{1}{\sqrt{\lambda}} \frac{W_h}{\sqrt{c_h}} \left\{ (1 - \rho_h^2) S_{yh}^2 - \frac{(\beta_h \mu_{30h} - \mu_{21h})^2}{\mu_{20h}^2 (\beta_{2h} - 1)} \right\}^{\frac{1}{2}}; \forall h = 1, 2, \dots, L \end{aligned} \quad (3.3)$$

Summing over all strata we have

$$n = \frac{1}{\sqrt{\lambda}} \sum_{j=1}^L \frac{W_h}{\sqrt{c_h}} \left\{ (1 - \rho_h^2) S_{yh}^2 - \frac{(\beta_h \mu_{30h} - \mu_{21h})^2}{\mu_{20h}^2 (\beta_{2h} - 1)} \right\}^{\frac{1}{2}} \quad (3.4)$$

Taking ratio of (5.3.3) and (5.3.4) we obtain

$$\begin{aligned} n_h = n \frac{\frac{1}{\sqrt{\lambda}} \frac{W_h}{\sqrt{c_h}} \left\{ (1 - \rho_h^2) S_{yh}^2 - \frac{(\beta_h \mu_{30h} - \mu_{21h})^2}{\mu_{20h}^2 (\beta_{2h} - 1)} \right\}^{\frac{1}{2}}}{\frac{1}{\sqrt{\lambda}} \sum_{j=1}^L \frac{W_h}{\sqrt{c_h}} \left\{ (1 - \rho_h^2) S_{yh}^2 - \frac{(\beta_h \mu_{30h} - \mu_{21h})^2}{\mu_{20h}^2 (\beta_{2h} - 1)} \right\}^{\frac{1}{2}}} \forall h = 1, 2, \dots, L \end{aligned} \quad (3.5)$$

When cost of drawing per unit sample is same in each stratum, (5.3.5) reduces to:

$$\begin{aligned} n_h = n \frac{W_h \left\{ (1 - \rho_h^2) S_{yh}^2 - \frac{(\beta_h \mu_{30h} - \mu_{21h})^2}{\mu_{20h}^2 (\beta_{2h} - 1)} \right\}^{\frac{1}{2}}}{\sum_{j=1}^L W_h \left\{ (1 - \rho_h^2) S_{yh}^2 - \frac{(\beta_h \mu_{30h} - \mu_{21h})^2}{\mu_{20h}^2 (\beta_{2h} - 1)} \right\}^{\frac{1}{2}}} \forall h = 1, 2, \dots, L \end{aligned} \quad (3.6)$$

Substituting the value from (5.3.6) in (5.3.1) we have

$$V\left(\hat{Y}_{gS}\right)_{\min.opt} = \frac{1}{n} \sum_{h=1}^L \left[ W_h \left\{ (1-\rho_h^2) S_{yh}^2 - \frac{(\beta_h \mu_{30h} - \mu_{21h})^2}{\mu_{20h}^2 (\beta_{2h} - 1)} \right\}^{\frac{1}{2}} \right]^2 = V_{opt}(say) \quad (3.7)$$

#### IV. Concluding Remarks

The mean square error of the separate linear regression estimator is given by

$$MSE(\bar{y}_{LRS}) = \sum_{h=1}^L W_h^2 \left( \frac{1}{n_h} - \frac{1}{N_h} \right) (1-\rho_h^2) S_{yh}^2 \quad (4.1)$$

Also the minimum mean square error of the proposed generalized regression type estimator  $\hat{Y}_{\theta S}$  is given by

$$MSE\left(\hat{Y}_{\theta S}\right)_{\min} = \sum_{h=1}^L W_h^2 \left( \frac{1}{n_h} - \frac{1}{N_h} \right) \left\{ (1-\rho_h^2) S_{yh}^2 - \frac{(\beta_h \mu_{30h} - \mu_{21h})^2}{\mu_{20h}^2 (\beta_{2h} - 1)} \right\} \quad (4.2)$$

Therefore the proposed generalized class of estimators  $\hat{Y}_{\theta S}$  may be preferred to the separate linear regression estimator, separate ratio estimator, separate product estimator and the usual stratified sample mean in the sense of smaller mean square error. Further the parameter involved  $\theta_h$  may be estimated by the corresponding sample value in order to get a class of estimators depending upon estimated optimum value.

The variance of stratified sample mean  $\bar{y}_{st}$  under Neyman allocation  $n_h = n \frac{W_h S_{yh}}{\sum_{h=1}^L W_h S_{yh}}$

$$\text{Is given by } V(\bar{y}_{st})_{Ney} = \frac{1}{n} \left( \sum_{h=1}^L W_h S_{yh} \right)^2 \quad (\text{ignoring f.p.c}) \quad (4.3)$$

It is evident that  $V_{opt}$  is always smaller than  $V(\bar{y}_{st})_{Ney}$  except for the case when  $\rho_h=0$  and  $\beta_h \mu_{30h} = \mu_{21h}$  simultaneously.

#### References

- [1]. Jack – knife technique of Gray and Schucany and Sukhatme et al.
- [2]. W.G.Cochran, *Sampling Techniques*, third ed. John Wiley &sons, New York, 1977.
- [3]. R.S.Jessen, *Statistical Survey Techniques*, John Wiley &sons, New York, 1978.
- [4]. D.Singh, F.S.Chaudhary, *Theory and Analysis of sample Survey Design*, New Age Publication, New Delhi, India 1986.
- [5]. S.K.Srivastava, *An estimator using auxiliary information in sample surveys*, Calcutta Statistical Association Bulletin 16(1967) 121-132.
- [6]. Bhushan. S. (2013) "Improved sampling strategies in finite population, Scholars Press, Germany.
- [7]. Bahli, S. And Tuteja, R.K. (1991): Ratio and Product type exponential estimator, *Information and Optimization sciences*, Vol.XII, I, 159-163.
- [8]. Gray H. L. and Schucany W.R. (1972). *The Generalized Jack-knife Statistic*, Marcel Dekker, New York.
- [9]. Naik, V.D. and Gupta, P.C. (1996): A note on estimation of mean with known population proportion of an auxiliary character. *Jour. Ind. Soc. Agr. Stat.*, 48(2),151-158.
- [10]. Sukhatme, P.V. and Sukhatme, B.V. (1970): *Sampling theory of surveys with applications*. Iowa State University Press, Ames, U.S.A.