On controlled sampling designs

P. Mukhopadhyay [a,*], K. Vijayan [b]

[a] *Indian Statistical Institute, Computer Science Unit, 203 Barrackpore Trunk Road, Calcutta 700 035, India*

[b] *University of Western Australia, Australia*

Received 5 October 1992; revised 26 April 1995

Abstract

Uses of t-designs for obtaining controlled sampling designs to replace certain types of sampling designs for estimating the linear and quadratic functions of population values have been discussed.

AMS classification: 62D05

Keywords: BIB-design; t-design; non-preferred sample

1. Introduction

In practical sample surveys all the possible samples are not generally preferred. The sampling designs which eliminate (or assign very small selection-probabilities to) the non-preferred samples are called controlled sampling designs and have been considered by Chakraborty (1963), Avadhani and Sukhatme (1973), Foody and Hedayat (1977), Srivastava and Saleh (1985), and Rao and Nigam (1990), among others. In this paper we show some usefulness of t-designs to obtain controlled sampling designs.

2. Main results

Let $\mathscr{P}$ be a finite population of units labeled $1,\ldots,i,\ldots,N$. Associated with i are two real numbers (y_i, x_i), values of a main variable 'y' and a closely related auxiliary variable 'x' respectively, $p_i = x_i/X$, $X = \sum_{i=1}^{N} x_i$. Let p be a fixed size $(-n)$ sampling design with $p(s)$ as the probability of selecting a sample (set) s, π_i, π_{ij}, $\pi_{ijk}\ldots$, as the inclusion-probabilities. Assume $p(s) > 0\ \forall s \in \gamma$ where the sample space γ consists of M_0 sample s, $M_i = \binom{N-i}{n-i}$, $i = 0,1,2,\ldots$ Assume that

$$p(s) = \alpha + \sum_{i \in s} \beta_i x_i, \quad s \in \gamma, \tag{1}$$

where α, β_i's are suitable constants.

Consider a controlled sampling design p_1 corresponding to p as follows. The design p_1 is based on a BIB-design with parameters (v,b,r,k,λ) where each variety is identified as a unit and each block as a sample. Thus $v = N$ and $k = n$. Let γ_b denote the collection of all samples so obtained (i.e. the reduced sample space). Assign

$$p_1(s) = \frac{M_0}{b} p(s)\ (0) \text{ if } s \in \gamma_b \text{ (otherwise)}. \tag{2}$$

Further, assume that the above BIB-design is a t-design, i.e. in this design every m-tuple of varieties occur together in a constant number λ_m of blocks $(m \leqslant t)$. It is known that for a t-design

$$\lambda_t \binom{v}{t} = \binom{k}{t} b. \tag{3}$$

Theorem 1. *Consider a sampling design p as given in* (1) *and its controlled version p_1 in* (2). *If a t-design is used to obtain p_1, then*

$$\pi_{i_1 \ldots i_{t-1}}(p) = \pi_{i_1 \ldots i_{t-1}}(p_1). \tag{4}$$

Proof.

$$\begin{aligned} \pi_{i_1 \ldots i_{t-1}}(p) &= \sum{}' p(s) \\ &= \sum{}' \left[\alpha + \sum_{i \in s} \beta_i x_i\right] \\ &= M_{t-1}\{\alpha + \beta_{i_1} x_{i_1} + \cdots + \beta_{i_{t-1}} x_{i_{t-1}}\} \\ &\quad + M_t \sum_{k (\neq i_1,\ldots,i_{t-1})} \beta_k x_k . \end{aligned} \tag{5}$$

where $\sum'$ denotes summation over all $s(\in \gamma)$ containing $(i_1,\ldots,i_{t-1})$. Again,

$$\begin{aligned} \pi_{i_1 \ldots i_{t-1}}(p_1) &= \frac{M_0}{b} \sum{}'' \left[\alpha \sum_{i \in s} \beta_i x_i\right] \\ &= \frac{M_0}{b} \left[\lambda_{t-1}\{\alpha + \beta_{i_1} x_{i_1} + \cdots + \beta_{i_{t-1}} x_{i_{t-1}}\} + \lambda_t \sum_{\substack{k \\ (\neq i_1,\ldots,i_{t-1})}} \beta_k x_k\right] \end{aligned}$$

when $\sum''$ denotes summation over all $s(\in \gamma_b)$ containing $(i_1,\ldots,i_{t-1})$.

By virtue of (3),

$$\frac{M_0}{b}\lambda_{t-1} = M_{t-1}, \quad \frac{M_0}{b}\lambda_t = M_t.$$

Hence the theorem. □

Corollary 1. *It, therefore, follows from Mukhopadhyay* (1972) *that*

$$\pi_{i_1 \ldots i_k}(p) = \pi_{i_1 \ldots i_k}(p_1), \quad k = 1, 2, \ldots, t-1.$$

Theorem 2.

Suppose

$$p(s) = \mu + \sum_{i \in s} \gamma_i x_i^2 + \sum_{i \neq j \in s}\sum \gamma_{ij} x_i x_j,$$

where μ, γ_i, γ_{ij}'*s are suitable constants. Consider* p_1 *as given in* (2). *If a 3-design is used to obtain* p_1, $\pi_i(p) = \pi_i(p_1)$, *and if a 4-design is used,* $\pi_{ij}(p) = \pi_{ij}(p_1)$.

Proof. Similar. □

Example 1. For Midzuno's sampling design, condition (1) is satisfied and the ratio estimator is unbiased for population total Y. The controlled design p_1 preserves the unbiasedness of the ratio estimator.

Example 2. For estimating the population variance $S_y^2 = \sum_{i=1}^{N} (Y_i - \bar{Y})^2/(N-1)$, the ratio estimator $\hat{S}^2_{yR} = (s_y^2/s_x^2)S_x^2$ (when s_y^2, s_x^2 are sample variances) is unbiased under the design $p(s) = s_x^2/M_0S_x^2$, $s \in \gamma$. The same estimator remains unbiased under p_1.

Remarks. For estimating Y, Srivastava (1985, 1988) and Shrivastava and Ouyang (1992a, b) suggested an elegant class of estimators

$$\hat{Y}_{sr1} = r(w)\sum_{i \in w} \frac{Y_i}{\pi_r(w)},$$

where

$$\pi_r(w) = \sum_{w \ni i} g(w), \quad g(w) = p(w)r(w),$$

w is a sample (not necessarily of a fixed size), $r(w)$ is a suitable weight. Suppose we restrict to fixed size sampling designs p. If

$$g(s) = p(s)r(s) = \alpha + \sum_{i \in s} \beta_i x_i,$$

for suitable constants α, β's, then the strategy $(p_1, \hat{Y}_{sr1})$ is also unbiased for Y.

Acknowledgements

The authors thank Prof. J.N. Srivastava for valuable suggestions leading to the substantial improvement of the paper.

References

Avadhani, M.S. and B.V. Sukhatme (1973). Controlled sampling with equal probabilities and without replacement. *Internat. Statist. Rev.* **41**, 175–182.

Chakraborty, M.C. (1963). On the use of incidence matrices of designs in sampling from finite populations. *J. Indian Statist. Assoc.* **1**, 78–85.

Foody, W. and A. Hedayat (1977). On theory and applications of BIB designs and repeated blocks. *Ann. Statist.* **5**, 932–945.

Mukhopadhyay, P. (1972). A sampling design to realise pre-assigned sets of inclusion-probabilities of first two orders. *Calcutta Statist. Assoc. Bull.* **21**, 87–122.

Rao, J.N.K. and A.K. Nigam (1990). Optimal controlled sampling designs. *Biometrika* 77, 807–814.

Srivastava, J.N. (1985). On a general theory of sampling, using experimental design, concepts I: estimation. *Bull. Internat. Statist. Inst.* **51** (2), 10.3.1–10.3.16.

Srivastava, J.N. (1988). On a general theory of sampling, using experimental design concepts. II. Relation with arrays. In: J.N. Srivastava, Ed., *Probability and Statistics*. North-Holland, Amsterdam.

Srivastava, J.N. and Z. Ouyang (1992a). Some properties of a general estimator in finite population sampling. *Sankhya* B **54**, 261–264.

Srivastava, J.N. and Z. Ouyang (1992b). Studies on a general estimator in sampling, using extraneous information through a sample weight function. *J. Statist. Plann. Inference.* **31**, 199–218.

Srivastava, J.N. and F. Saleh (1985). Need of *t*-design in sampling theory. *Utilitas Math.* **25**, 5–17.