Werk

Titel: Regular Estimators for Subclasses of Linear Estimators.

Autor: Prabhu-Ajgaonkar, S.G.

Jahr: 1985

PURL: https://resolver.sub.uni-goettingen.de/purl?358794056_0032|log35

Kontakt/Contact

Digizeitschriften e.V.
SUB Göttingen
Platz der Göttinger Sieben 1
37073 Göttingen

info@digizeitschriften.de

Metrika, Volume 32, 1985, page 265–274

Regular Estimators for Subclasses of Linear Estimators

By *S.G. Prabhu-Ajgaonkar*, Aurangabad[1])

Summary: The concept of regular estimator is due to *Roy/Chakravarti*. For its application they confined to the most general class of linear estimators. The present paper considers some subclasses of linear estimators.

1 Introduction

Consider a population consisting of a finite number N of distinguishable elementary units u_i with associated real numbers (variate values) y_i, $i = 1, 2, \ldots, N$. The mean and the variance of the population are denoted respectively by

$$\mu = \frac{1}{N} \sum_{i=1}^{N} y_i \text{ and } \sigma^2 = \frac{1}{N} \sum_{i=1}^{N} (y_i - \mu)^2. \tag{1.1}$$

A sample of fixed size n is drawn from this population. Let s be a typical sample and p_s be its probability of selection. If S denotes the totality of all samples that could be selected then

$$\sum_{s=1}^{S} p_s = 1 \tag{1.2}$$

Let $n_i(s)$ denote the number of time the i-th population unit occurs in the sample s and let

$$\nu_i(s) = 0\,(1) \text{ if } n_i(s) = 0\,(> 0). \tag{1.3}$$

The number of distinct population units that are included in the sample s is

$$\nu_{(s)} = \sum_{i=1}^{N} \nu_i(s). \tag{1.4}$$

---

[1] *S.G. Prabhu-Ajgaonkar*, Marathwada University, Aurangabad 431004, India.

0026-1335/85/0300265-274$2.50 © 1985 Physica-Verlag, Wien

Roy/Chakravarti called a linear unbiased estimator t of μ, a regular estimator if its variance is of the form

$$V(t) = k\,\sigma^2 = \frac{k(N-1)}{N^2}\sum_{i=1}^{N} y_i^2 - \frac{k}{N^2}\sum_{i\neq j=1}^{N} y_i y_j \tag{1.5}$$

where k is a constant independent of the vector y.

In their discussion of regular estimators, *Roy/Chakravarti* confined to the most general class of linear estimators and demonstrated that there exists a lower bound for the variance of any regular estimator of μ, and that this lower bound is attained when the sampling scheme is balanced.

2 The T_1 Sub-Class

The T_1-subclass of linear estimators were defined symbolically by *Horvitz/Thompson* as follows:

$$T_1 = \sum_{r=1}^{n} \alpha_r y_r \tag{2.1}$$

where α_r $(r = 1, 2, \ldots, n)$ represents a constant used as weight attached to the outcome at the r-th draw $(r = 1, 2, \ldots, n)$. If p_{ir} $(i = 1, 2, \ldots, N; r = 1, 2, \ldots, n)$ denotes the probability of selecting the i-th population unit at the r-th draw, then it is demonstrated by *Prabhu-Ajgaonkar* [1967] that the T_1-subclass is empty when the rank of the probability matrix P, defined as

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ p_{N1} & p_{N2} & & p_{Nn} \end{bmatrix} \tag{2.2}$$

is not the same as that of the augmented matrix p^*, defined as

$$P^* = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} & \frac{1}{N} \\ p_{21} & p_{22} & \cdots & p_{2n} & \frac{1}{N} \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ p_{N1} & p_{N2} & & P_{Nn} & \frac{1}{N} \end{bmatrix} \tag{2.3}$$

Otherwise the T_1-subclass is termed non-empty for the given sampling procedure. It is noted that for the existence of a regular estimator the T_1-subclass should be non-empty.

2.1 Sampling Without Replacement

If sampling is carried out without replacement, then the variance of the T_1-subclass of linear estimators can be expressed as

$$V(T_1) = \sum_{i=1}^{N} y_i^2 \left(\sum_{r=1}^{n} \alpha_r^2 p_{ir} - \frac{1}{N^2} \right) + \sum_{i \neq j=1}^{N} y_i y_j \left(\sum_{r \neq s=1}^{n} \alpha_r \alpha_s p_{irjs} - \frac{1}{N^2} \right) \tag{2.4}$$

where p_{irjs} stands for the joint probability of selecting the i-th and j-th population units at respectively r-th and s-th draws.

T_1 is a regular estimator if

$$V(T_1) = k\,\sigma^2.$$

Equating coefficients of y_i^2 and $y_i y_j$ on both the sides, we obtain the following equations respectively,

$$\frac{k(N-1)}{N^2} = \sum_{r=1}^{n} \alpha_r^2 p_{ir} - \frac{1}{N^2} \qquad i = 1, 2, \ldots, N. \tag{2.5}$$

and

$$-\frac{k}{N^2} = \sum_{r \neq s=1}^{n} \alpha_r \alpha_s p_{irjs} - \frac{1}{N^2} \qquad i \neq j = 1, 2, \ldots, N. \tag{2.6}$$

Let us consider some particular sampling schemes.

2.1.1 Simple Random Sampling Without Replacement

For this sampling scheme, it is noted that

$$p_{ir} = \frac{1}{N}, \quad i = 1, 2, \ldots, N, \; r = 1, 2, \ldots, n.$$

The T_1-subclass is nonempty and the conditions of unbiasedness reduce to

$$\sum_{r=1}^{n} \alpha_r = 1. \tag{2.7}$$

Further either condition (2.5) or (2.6) reduces to the equation

$$\sum_{r=1}^{n} \alpha_r^2 = \frac{k(N-1)+1}{N}. \tag{2.8}$$

The weights α_r of a regular estimator should satisfy conditions (2.7) and (2.8). It is noted that for simple random sampling without replacement there exist regular estimators in the T_1 subclass.

2.1.2 Ikeda-Sen's Sampling Scheme

This sampling scheme formulated by *Ikeda,* a student of *Midzuno* and independently by *Sen* is defined as follows:

On the first draw the population units are selected with varying probabilities and on the subsequent draws, the population units are sampled with equal probabilities without replacement. For this sampling scheme, *Prabhu-Ajgaonkar* [1967] has demonstrated that the T_1-subclass is non-empty and that the conditions of unbiasedness reduce to

$$\alpha_1 = \frac{1}{N}, \sum_{r=2}^{n} \alpha_r = \frac{N-1}{N}. \tag{2.9}$$

The condition (2.5) reduces (for $n > 1$) to

$$\sum_{r=2}^{n} \alpha_r^2 \frac{(1-p_i)}{(N-1)} = \frac{k(N-1)+1}{N^2} - \frac{p_i}{N} \tag{2.10}$$

$i = 1, 2, \ldots, N$, where p_i $(i = 1, 2, \ldots, N)$ denotes the probability of selecting the i-th population unit at the first draw.

It is noted that the above system of linear equations in α_r^2 is not consistent. Thus a regular estimator does not exist for the T_1-subclass, when Ideka-Sen's sampling scheme is employed.

2.2 Sampling With Replacement

If sampling is carried out with replacement, then

$$V(T_1) = \sum_{i=1}^{N} y_i^2 \sum_{r=1}^{n} \alpha_r^2 p_{ir}(1-p_{ir}) - \sum_{i \neq j=1}^{N} y_i y_j \sum_{r=1}^{n} \alpha_r^2 p_{ir} p_{jr} \tag{2.11}$$

T_1 is a regular estimator, if

$$V(T_1) = k\sigma^2.$$

Equating coefficients of y_i^2 and $y_i y_j$ on both sides, we obtain from equations (1.5) and (2.11)

$$\sum_{r=1}^{n} \alpha_r^2 p_{ir}(1 - p_{ir}) = \frac{k(N-1)}{N^2}, \quad i = 1, 2, \ldots, N. \tag{2.12}$$

$$\sum_{r=1}^{n} \alpha_r^2 p_{ir} p_{jr} = \frac{k}{N^2}, \quad i \neq j = 1, 2, \ldots, N. \tag{2.13}$$

We consider two particular sampling schemes.

2.2.1 Simple Random Sampling With Replacement

For this sampling scheme, it is easy to note that

$$p_{ir} = \frac{1}{N}, \quad i = 1, 2, \ldots, N;\ r = 1, 2, \ldots, n,$$

and

$$p_{ir,js} = \frac{1}{N^2}, \quad i, j = 1, 2, \ldots, N$$
$$r \neq s = 1, 2, \ldots, n.$$

The conditions of unbiasedness reduce to

$$\sum_{r=1}^{n} \alpha_r = 1.$$

Further, each of the condition (2.12) and (2.13) reduces to

$$\sum_{r=1}^{n} \alpha_r^2 = k.$$

Thus it is noted that for simple random sampling with replacement there exist regular estimators in the T_1 subclass.

2.2.2 Varying Probability With Replacement

For this sampling scheme, it is noted that

$$p_{ir} = p_i \quad \text{for } i = 1, 2, \ldots, N;\ r = 1, 2, \ldots, n.$$

Prabhu-Ajgaonkar [1967] has pointed out that the T_1 subclass is empty. Therefore there does not exist a regular estimator when this sampling procedure is employed.

3 The T_2-Subclass

Horvitz/Thompson formulated symbolically the T_2-subclass estimators as follows:

$$T_2 = \sum_i^n \beta_i y_i \tag{3.1}$$

where β_i $(i = 1, 2, \ldots, N)$ is a constant used as weight associated with the i-th population unit whenever it is included in a sample. *Horvitz/Thompson* showed that there exists only one unbiased estimator for the T_2-subclass, which is defined as

$$t_2 = \frac{1}{N} \sum_i^n \frac{y_i}{\pi_i} \tag{3.2}$$

where π_i $(i = 1, 2, \ldots, N)$ represents the probability of including the i-th population unit in a sample. Further its variance is

$$V(t_2) = \sum_{i=1}^{N} \frac{(1-\pi_i)}{\pi_i} \frac{y_i^2}{N^2} + \sum_{i \neq j=1}^{N} \frac{(\pi_{ij} - \pi_i \pi_j)}{\pi_i \pi_j} \frac{y_i y_j}{N^2} \tag{3.3}$$

where π_{ij} $(i \neq j = 1, 2, \ldots, N)$ denotes the joint probability of selecting the i-th and j-th population units into a sample, t_2 is a regular estimator, if

$$V(t_2) = k\,\sigma^2.$$

From equations (1.5) and (3.3), we obtain after equating coefficient of y_i^2

$$\pi_i = \frac{1}{1 + k\,(N-1)} \qquad i = 1, 2, \ldots, N.$$

Since π_i's vary, the above equation is not possible.

For sampling with varying probabilities, there does not exist a regular estimator for the T_2-subclass.

4 The T_3-Subclass

The T_3-subclass was formulated symbolically by Horvitz and Thompson (1952) as follows:

$$T_3 = \nu_s \left(\sum_{i\,\sigma\,s} y_i \right) \tag{4.1}$$

where ν_s ($s = 1, 2, \ldots, S$) is a constant used as weight attached to s-th sample whenever it is selected.

The conditions of unbiasedness are

$$\sum_{s \supset i} \nu_s p_s = \frac{1}{N}, \quad i = 1, 2, \ldots, N, \tag{4.2}$$

where the summation is over the samples that contain the i-th population unit. The variance of T_3 is given by

$$V(T_3) = \sum_{i=1}^{N} y_i^2 \left(\sum_{s \supset i} \nu_s^2 p_s - \frac{1}{N^2} \right) \tag{4.3}$$

$$+ \sum_{i \neq j = 1}^{N} y_i y_j \left(\sum_{s \supset i,j} \nu_s^2 p_s - \frac{1}{N^2} \right).$$

T_3 is a regular estimator if

$$V(T_3) = k \sigma^2.$$

From equations (1.5) and (4.3), we obtain after equating coefficients of y_i^2 and $y_i y_j$,

$$\sum_{s \supset i} \nu_s^2 p_s = \frac{k(N-1)+1}{N^2}, \quad i = 1, 2, \ldots, N \tag{4.4}$$

$$\sum_{s \supset i,j} \nu_s^2 p_s = \frac{1-k}{N^2}, \quad i \neq j = 1, 2, \ldots, N. \tag{4.5}$$

For the population vector y, when

$$y_i = 1, \quad i = 1, 2, \ldots, N$$

it is noted that Var $(T_3) = 0$, when T_3 is regular.

This shows that a regular estimator is constant, i.e.,

$$T_3 = \nu_s n$$

$$= \text{constant}.$$

However, this contradicts the conditions of unbiasedness at equation (4.2). Hence, a regular estimator for the T_3-class does not exist for varying probability sampling.

5 The T_4-Subclass

The T_4-subclass was defined by *Koop* and independently by *Prabhu-Ajgaonkar* as follows:

$$T_4 = \sum_{i=1}^{N} \left(\sum_{r=1}^{n} \bar{\beta}_{ri} \right) y_i \tag{5.1}$$

where $\bar{\beta}_{ri}$ $(r = 1, 2, \ldots, n, i = 1, 2, \ldots, N)$ is a random variable assuming the value β_{ri} whenever the i-th population unit is selected at the r-th draw, otherwise it is zero.

The conditions of unbiasedness give

$$\sum_{r=1}^{n} E(\bar{\beta}_{ri}) = \frac{1}{N}, \quad i = 1, 2, \ldots, N. \tag{5.2}$$

The variance of T_4 is noted to be

$$V(T_4) = \sum_{i=1}^{N} y_i^2 \left[E \left(\sum_{r=1}^{n} \bar{\beta}_{ri} \right)^2 - \frac{1}{N^2} \right]$$

$$+ \sum_{i \neq j=1}^{N} y_i y_j \left[E \left(\sum_{r=1}^{n} \bar{\beta}_{ri} \right) \left(\sum_{r=1}^{n} \bar{\beta}_{rj} \right) - \frac{1}{N^2} \right]. \tag{5.3}$$

T_4 is a regular estimator, if

$$V(T_4) = k\,\sigma^2.$$

From equations (1.5) and (5.3), we obtain after equating coefficients of y_i^2 and $y_i y_j$,

$$E \left(\sum_{r=1}^{n} \bar{\beta}_{ri} \right)^2 = \frac{k(N-1)+1}{N^2}, \quad i = 1, 2, \ldots, N. \tag{5.4}$$

$$E \left[\left(\sum_{r=1}^{n} \bar{\beta}_{ri} \right) \left(\sum_{r=1}^{n} \bar{\beta}_{rj} \right) \right] = \frac{1-k}{N^2}, \quad i \neq j = 1, 2, \ldots, N \tag{5.5}$$

Consider a quantity $\emptyset$ given gy

$$\emptyset = \sum_{i=1}^{N} \sum_{r=1}^{n} \beta'_{ri} \tag{5.6}$$

where the β'_{ri} are weights of a regular estimator. It is noted that $E(\emptyset) = 1$ and $V(\emptyset) = 0$. This shows that $\emptyset$ is constant. Therefore

$$\emptyset = 1, \quad \text{for all } S. \tag{5.7}$$

Consider a quantity

$$z = \sum_{i=1}^{N} \left[\left(\sum_{r=1}^{n} \beta'_{ri} \right) - \frac{\nu_i(s)}{\nu(s)} \right]^2$$

$$= \sum_{i=1}^{N} \left(\sum_{r=1}^{n} \beta'_{ri} \right)^2 - \frac{1}{\nu(s)}, \tag{5.8}$$

because of (5.7).

Using (5.4) and (5.5) we get

$$E(z) = \frac{(N-1)k}{N} + \frac{1}{N} - E\left[\frac{1}{\nu(s)}\right]. \tag{5.9}$$

Since $E(z)$ is non-negative, we obtain a lower bound for k given by

$$k \geqslant \frac{N}{N-1} E\left[\frac{1}{\nu(s)}\right] - \frac{1}{(N-1)}. \tag{5.10}$$

The regular estimators belonging to subclass T_4 defined at equation (5.1) and satisfying condition (5.2) have

$$V(T_4) \geqslant k\,\sigma^2$$

where k is given by equation (5.10). This lower bound is attained if and only if $E(z) = 0$. This is possible if

$$\left(\sum_{r=1}^{n} \beta'_{ri} \right) = \frac{\nu_i(s)}{\nu(s)} \quad \text{for every } S. \tag{5.11}$$

Equation (5.11) defines regular estimator having the minimum variance among the class of regular estimators belonging to the T_4 subclass. Since a regular estimator is unbiased, we have

$$E\left(\sum_{r=1}^{n} \beta'_{ri} \right) = E\left[\frac{\nu_i(s)}{\nu(s)}\right]$$

$$= \frac{1}{N}, \quad s = 1, 2, \ldots, S. \tag{5.12}$$

Roy/Chakravarti have termed sampling scheme balanced when the condition (5.12) holds. Further they derived this result by considering the most general class of linear estimators. However, it is interesting to note that the same result follows if we restrict to the T_4-subclass of linear estimators.

6 Acknowledgement

The author is thankful to the referee for going through the paper and making many useful suggestions.

References

Horvitz, D.G., and *D.J. Thompson*: A generalization of sampling without replacement from a finite universe. J. Amer. Statist. Assoc., **47**, 1952, 663–685.

Koop, J.C.: On the axioms of sample formation and their bearing on the construction of linear estimators in sampling theory for finite universes. Metrika 7, 81–114 and 165–204.

Midzuno, H.: On the sampling system with probability proportional to sum of sizes. Ann. Inst. Statist. Math., **3**, 1952, 99–107.

Prabhu-Ajgaonkar, S.G.: On Horvitz and Thompson's T_1 class of linear estimators. Ann. Math. Statist. **38**, 1967, 1882–1886.

–: A contribution to the theory of unique estimators for Horvitz and Thompson's classes of linear estimators. Metrika **15**, 1970, 40–47.

Roy, J., and *I.M. Chakravarti*: Estimating the mean of a finite population. Ann. Math. Statist. **31**, 1980, 392–398.

Sen, A.R.: Recent advances in sampling with varying probabilities. Bull. Calcutta Stat. Assoc., **5**, 1953, 1–15.

Received March 3, 1982
(Revised version January 29, 1983)