An Alternative Estimator for the Population Mean
In PPS Sampling

Anita¹, Shashi Bahl²

¹Ph.D Scholar of Department of Statistics M.D.U. Rohtak (Haryana), India
²Professor, Department of Statistics M.D.U. Rohtak (Haryana), India

Abstract: In this paper some composite estimators depending on estimated optimum value of a constant based on sample values under PPS sampling scheme have been proposed. The estimators are better in sense (minimum variance) than PPS estimator, ratio estimator or difference estimator under certain conditions.

Keywords: PPS estimator, Ratio estimator & Difference estimator

1. Introduction

It is well know that at large scale surveys use of multiple auxiliary characteristics improves the precision of the estimators. This can be utilized in the construction of the estimator or at the stage of selecting the sample from the population. In this paper two auxiliary characters have been used in different ways as one for the selection of the sample and other for the purpose of estimation to estimate the population mean.

2. Notations

Let yᵢ, x₁ᵢ and x₂ᵢ be two values of characters y under study and values of the two auxiliary characters x₁ and x₂ for the iᵗʰ unit in the population (i=1,2, ..., N) of the population of size N. Let a sample of size n be drawn with PPSWR sampling (based on x₁) and Y, X₁, X₂ be population total of y, x₁, x₂ respectively.

\[ u_i = \frac{x_2_i}{N p_{li}}, \quad v_i = \frac{x_1_i}{X_1} \]

\[ \bar{u}_n = \frac{1}{n} \sum_{i=1}^{n} u_i = \bar{y}_{1pp} \]

\[ \bar{v}_n = \frac{1}{n} \sum_{i=1}^{n} v_i = \bar{x}_{2pp} \]

\[ \sigma_u^2 = \sum_{i=1}^{N} p_{li} (u_i - \bar{y})^2 = \frac{\sigma_u^2}{\bar{y}^2} = C_u \]

\[ \sigma_v^2 = \sum_{i=1}^{N} p_{li} (v_i - \bar{x}_2)^2 = \frac{\sigma_v^2}{\bar{x}_2^2} = C_v \]

\[ \bar{y}_D = \bar{y} + k \left( \bar{x}_2 - \bar{x} \right)^2 \]

\[ \bar{y}_R = \frac{\bar{u}_n}{\bar{v}_n} \bar{x}_2 \]

\[ \sum_{i=1}^{N} p_{li} (u_i - \bar{y})(v_i - \bar{x}_2) \]

\[ \rho_{uv} = \frac{\sigma_u \sigma_v}{\sigma_{uv}} \]
3. Proposed Estimators

For estimating the population mean $\overline{Y}$, when the information on $x_1$, $x_2$ for every unit of the population is available, the proposed estimators of $\overline{Y}$ are

$$\overline{Y}_{D,PPS} = a\overline{Y}_D + (1-a)\overline{y}_{PPS} \quad \text{......... (1)}$$

$$\overline{Y}_{PPS,D} = a\overline{y}_{PPS} + (1-a)\overline{Y}_D \quad \text{......... (2)}$$

$$\overline{Y}_{D,R} = a\overline{Y}_D + (1-a)\overline{y}_R \quad \text{......... (3)}$$

Where $a$ is a constant to be determined so that $\sqrt{\text{var}(\overline{Y}_{D,PPS})}$ and $\sqrt{\text{var}(\overline{Y}_{D,R})}$ is minimum

$$\text{var}(\overline{Y}_{D,PPS}) = a^2\text{var}(\overline{Y}_D) + (1-a)^2\text{var}(\overline{y}_{PPS}) + 2a(1-a)\text{cov}(\overline{Y}_D, \overline{y}_{PPS})$$

$$\text{var}(\overline{Y}_{PPS,D}) = a^2\text{var}(\overline{y}_{PPS}) + (1-a)^2\text{var}(\overline{Y}_D) + 2a(1-a)\text{cov}(\overline{y}_{PPS}, \overline{Y}_D)$$

$$\text{var}(\overline{Y}_{D,R}) = a^2\text{var}(\overline{Y}_D) + (1-a)^2\text{var}(\overline{y}_R) + 2a(1-a)\text{cov}(\overline{Y}_D, \overline{y}_R)$$

Where $\text{var}(\overline{Y}_D) = \frac{\overline{y}^2}{n} \left( C_u^2 + k^2C_v^2 - 2k\rho_{uv}C_uC_v \right)$

$$M(\overline{y}_R) = \frac{\overline{y}^2}{n} \left( C_u^2 + C_v^2 - 2\rho_{uv}C_uC_v \right)$$

$$\text{cov}(\overline{Y}, \overline{y}_{PPS}) = \frac{\overline{y}^2}{n} \rho_{uy}C_uC_y$$

$$\text{cov}(\overline{Y}, \overline{y}_{PPS}) = \frac{\overline{Y}^2}{n} \rho_{ux}C_uC_x$$

$$\text{cov}(\overline{Y}, \overline{y}_R) = \frac{\overline{Y}^2}{n} \left( C_u^2 - C_v \right)$$

$$\text{cov}(\overline{X}, \overline{y}_R) = \frac{\overline{Y}^2}{n} \left( C_{uv} - C_v \right)$$

So that

$$\text{var}(\overline{Y}_D, \overline{y}_{PPS}) = \frac{\overline{y}^2}{n} \left[ a^2 \left( C_u^2 + k^2C_v^2 - 2k\rho_{uv}C_uC_v \right) + (1-a)^2 C_u^2 \right.$$ \n
$$+ 2a(1-a)\rho_{uy}C_uC_y - 2a(1-a)k\rho_{ux}C_xC_u \big]$$

$$\text{var}(\overline{y}_{PPS}, \overline{Y}_D) = \frac{\overline{y}^2}{n} \left[ a^2C_u^2 + (1-a)^2 \left( C_u^2 + k^2C_v^2 - 2k\rho_{uv}C_uC_v \right) \right.$$ \n
$$+ 2a(1-a)\rho_{uy}C_uC_y - 2a(1-a)k\rho_{ux}C_xC_u \big]$$

$$\text{var}(\overline{Y}_D, \overline{y}_R) = \frac{\overline{y}^2}{n} \left[ a^2 \left( C_u^2 + k^2C_v^2 - 2k\rho_{uv}C_uC_v \right) \right.$$ \n
$$+ (1-a)^2 \left( C_u^2 + C_v^2 - 2\rho_{uv}C_uC_v \right) +$$ \n
$$2a(1-a)\left( C_u^2 - C_v \right) \right]$$
4. Optimum Values of a

The optimum values of a for which \( V(\bar{y}_{D,PPS}) \), \( V(\bar{y}_{PPS,D}) \) and \( V(\bar{y}_{D,R}) \) are minimized; are respectively

\[
a_{opt} = \frac{C_{uv}}{kC_v^2} a_{opt} = 1 - \frac{C_{uv}}{kC_v^2}
\]

\[
a_{opt} = \frac{C_v^2 - C_{uv}}{(1-k)C_v^2}
\]

For these optimum values of a

\[
V(\bar{y}_D, pps) = \frac{\sigma_u^2}{n}(1 - \rho_{uv}^2)
\]

\[
V(\bar{y}_D, R) = \frac{\sigma_u^2}{n}(1 - \rho_{uv}^2)
\]

5. Empirical Study

To see the performance of the proposed estimator in comparison to other estimators the description of population data are given below:-

Population (Source: Singh and Mangat(1996), pp-220)

N=240 N_1=70 N_2=120 N_3=50
n =24 n_1 =7 n_2=12 n_3=5
\(X_1 = 15.28\) \(X_2 = 17.25\) \(X_3 = 17.8\)
\(\bar{y} = 17.42\) \(C_u = 0.299\) \(C_v = 0.318\)
\(C_{uv}^2 = 0.089401\) \(C_{uv}^2 = 0.101124\)
\(\rho_{uv} = 0.76\) \(\rho_{uv} C_u C_v = 0.072262\)

\[
V(\bar{y}_{D, pps}) = .03778 \quad V(\bar{y}_{D, pps}) = .04242 \quad V(\bar{y}_{D, pps}) = .039091\]
\[
k=1 \quad k=1 \quad k=1
\]

\[
V(\bar{y}_{D, R}) = .03776 \quad V(\bar{y}_{D, R}) = .039090 \quad V(\bar{y}_{D, R}) = .037971\]
\[
k=2 \quad k=2 \quad k=2
\]

Thus the variances of our proposed estimators are less than the usual estimators.

6. Conclusion

When \(a_{opt}\) put in proposed estimator the resulting proposed estimator attains minimum variance. Thus for practical applications proposed estimator may be preferred.

References


Volume 4 Issue 3, March 2015