ON CERTAIN UNBIASED RATIO ESTIMATORS

T. J. RAO

(Received May 4, 1965; revised Jan. 5, 1966)

Summary

We consider the two ratio type estimators $\frac{\bar{y}}{\bar{x}}\bar{X}$ and $\frac{\bar{X}}{m}\sum\frac{y_i}{x_i}$ for the estimation of $\bar{Y}$ and show how the estimators due to Murthy and Nanjamma, Hartley and Ross, and Nieto de Pascual can be obtained as linear combinations of these two. Some interesting relations among these estimators have been mentioned.

1. Introduction

Consider the two ratio type estimators $\frac{\bar{y}}{\bar{x}}\bar{X}$ and $\frac{\bar{X}}{m}\sum_1^m\frac{y_i}{x_i}$ which are used for the estimation of population mean $\bar{Y}$ of a characteristic $\mathscr{Y}$, when we have auxiliary information on another characteristic $\mathscr{X}$ related to $\mathscr{Y}$ and $\bar{X}$ is known beforehand. In the construction of unbiased ratio estimators, the bias of one of these estimators is estimated and then the estimator is corrected for its bias. Estimators of this type were given by Hartley and Ross [1], Murthy and Nanjamma [2] and Nieto de Pascual [3]. In this paper, we present a general method of deriving these estimators by considering linear combinations of the two typical ratio estimators $\frac{\bar{y}}{\bar{x}}\bar{X}$ and $\frac{\bar{X}}{m}\sum_1^m\frac{y_i}{x_i}$ and determine the coefficients for which the combined estimators are unbiased. For particular values of the coefficients, we show that the estimators given in [1], [2] and [3] are obtained as particular cases. It is also observed that certain interesting relations hold between these coefficients.

2. Main results

Consider a linear combination of the two estimators mentioned above, say

$$(1) \qquad \hat{\bar{Y}}=\theta_1\frac{\bar{y}}{\bar{x}}\bar{X}+\theta_2\frac{\bar{X}}{m}\sum\frac{y_i}{x_i} .$$

For $\hat{\bar{Y}}$ to be unbiased, we have the condition that

$$E(\hat{\bar{Y}})=\theta_1\bar{Y}+\theta_1 B\left(\frac{\bar{y}}{\bar{x}}\bar{X}\right)+\theta_2\bar{Y}+\theta_2 B\left(\frac{\bar{X}}{m}\sum\frac{y_i}{x_i}\right)=\bar{Y},$$

where B stands for the bias of the estimator.

i.e., $\qquad \theta_1+\theta_2=1$

and

$$\theta_1 B\left(\frac{\bar{y}}{\bar{x}}\bar{X}\right)+\theta_2 B\left(\frac{\bar{X}}{m}\sum\frac{y_i}{x_i}\right)=0 ,$$

i.e., $\qquad \theta_2=1-\theta_1$

and

$$\frac{B\left(\frac{\bar{X}}{m}\sum\frac{y_i}{x_i}\right)}{B\left(\frac{\bar{y}}{\bar{x}}\bar{X}\right)}=-\frac{\theta_1}{1-\theta_1} .$$

This ratio is always positive, and hence follows that $\theta_1 \notin (0, 1)$. If we consider the approximate expressions of the biases, we have

$$-\frac{\theta_1}{1-\theta_1}=m$$

and hence

$$\theta_1=\frac{m}{m-1} .$$

The estimator is now given by $\frac{m}{m-1}\frac{\bar{y}}{\bar{x}}\bar{X}-\frac{1}{m-1}\frac{\bar{X}}{m}\sum\frac{y_i}{x_i}$. This is the almost unbiased ratio estimator given by Murthy and Nanjamma [2]. A similar approach by Hartley (unpublished lecture notes, 1954) of combining these two types of ratio estimators has been pointed out by Nieto de Pascual [3], which results in the Murthy-Nanjamma estimator.

Now, consider a combined estimator of the type

$$(2) \qquad \hat{\bar{Y}}'=\theta\frac{\bar{y}}{\bar{x}}\bar{X}+(1-\theta)\frac{\bar{X}}{m}\sum_1^m\frac{y_i}{x_i} ,$$

where θ is a random variable.

Let us further suppose that the scheme is Simple Random Sampling With-Out Replacement (SRSWOR).

We have for the unbiasedness of $\hat{\bar{Y}}'$

$$E\left[\theta\frac{\bar{y}}{\bar{x}}\bar{X}+(1-\theta)\frac{\bar{X}}{m}\sum\frac{y_i}{x_i}\right]=\bar{Y}.$$

i.e.,

$$\text{(A)}\qquad E\left[\theta\left\{\frac{\bar{y}}{\bar{x}}\bar{X}-\frac{\bar{X}}{m}\sum\frac{y_i}{x_i}\right\}\right]=-B\left[\frac{\bar{X}}{m}\sum\frac{y_i}{x_i}\right].$$

Now, we have the following

LEMMA. *For a sample taken by SRSWOR,*

$$B\left[\frac{\bar{X}}{m}\sum\frac{y_i}{x_i}\right]=\frac{m}{m-1}\cdot\frac{N-1}{N}\left[E\left\{\frac{\bar{x}}{m}\sum\frac{y_i}{x_i}-\bar{y}\right\}\right].$$

PROOF.

$$B\left[\frac{\bar{X}}{m}\sum\frac{y_i}{x_i}\right]=-\operatorname{cov}\left(\frac{y}{x},x\right)$$

$$=-\operatorname{cov}\left(\frac{1}{m}\sum\frac{y_i}{x_i},\bar{x}\right)\frac{Nm}{N-m}\cdot\frac{N-1}{N}$$

$$=\frac{-m(N-1)}{N-m}\left\{E\left(\frac{\bar{x}}{m}\sum\frac{y_i}{x_i}\right)-\bar{Y}-B\left(\frac{\bar{X}}{m}\sum\frac{y_i}{x_i}\right)\right\}.$$

On simplification it follows that

$$B\left[\frac{\bar{X}}{m}\sum\frac{y_i}{x_i}\right]=\frac{m}{m-1}\cdot\frac{N-1}{N}\left[E\left\{\frac{\bar{x}}{m}\sum\frac{y_i}{x_i}-\bar{y}\right\}\right].$$

From (A), using the lemma, we have

$$E\left[\theta\frac{\bar{X}}{\bar{x}}\left\{\bar{y}-\frac{\bar{x}}{m}\sum\frac{y_i}{x_i}\right\}\right]=E\left[\frac{m}{m-1}\cdot\frac{N-1}{N}\left\{\bar{y}-\frac{\bar{x}}{m}\sum\frac{y_i}{x_i}\right\}\right]$$

for which $\theta=\frac{m}{m-1}\frac{N-1}{N}\frac{\bar{x}}{\bar{X}}$ is a solution and substituting this value of θ in the combined estimator, we get the Hartley-Ross estimator [1].

When θ is a constant, it follows from (A) that

$$\theta=\frac{B\left(\frac{\bar{X}}{m}\sum\frac{y_i}{x_i}\right)}{B\left(\frac{\bar{X}}{m}\sum\frac{y_i}{x_i}\right)-B\left(\frac{\bar{y}}{\bar{x}}\bar{X}\right)}\simeq+\frac{m}{m-1}$$

and this gives the Murthy-Nanjamma estimator. Further, it is interesting to note that while $\theta > 1$ in the case of Murthy-Nanjamma estimator, $E(\theta) > 1$ in the case of Hartley-Ross estimator. Next, we consider a combined estimator of the type

$$\hat{\bar{Y}}'' = (1-\theta)\frac{\bar{y}}{\bar{x}}\bar{X} + \theta\frac{\bar{X}}{m}\sum\frac{y_i}{x_i}\,, \tag{3}$$

where θ is a random variable. Then for $\hat{\bar{Y}}''$ to be unbiased, we have

$$E(\hat{\bar{Y}}'') = E\left[\theta\left\{\frac{\bar{X}}{m}\sum\frac{y_i}{x_i} - \frac{\bar{y}}{\bar{x}}\bar{X}\right\} + \frac{\bar{y}}{\bar{x}}\bar{X}\right] = \bar{Y}\,,$$

i.e.,

$$E\left[\theta\left\{\frac{\bar{X}}{m}\sum\frac{y_i}{x_i} - \frac{\bar{y}}{\bar{x}}\bar{X}\right\}\right] = -B\left[\frac{\bar{y}}{\bar{x}}\bar{X}\right]. \tag{A$'$}$$

We have

$$B\left[\frac{\bar{y}}{\bar{x}}\bar{X}\right] = -\operatorname{cov}\left(\frac{\bar{y}}{\bar{x}},\, \bar{x}\right) \simeq -\frac{1}{m}\operatorname{cov}\left(\frac{y}{x};\, x\right).$$

So, by the lemma it follows that

$$E\left[\theta\left\{\frac{\bar{X}}{m}\sum\frac{y_i}{x_i} - \frac{\bar{y}}{\bar{x}}\bar{X}\right\}\right] = -\frac{1}{m-1}\cdot\frac{N-1}{N}\left[E\left\{\frac{\bar{x}}{m}\sum\frac{y_i}{x_i} - \bar{y}\right\}\right]$$

$$= E\left[-\frac{N-1}{N(m-1)}\cdot\frac{\bar{x}}{\bar{X}}\left\{\frac{\bar{X}}{m}\sum\frac{y_i}{x_i} - \frac{\bar{y}}{\bar{x}}\bar{X}\right\}\right].$$

A solution of this equation gives $\theta = -\frac{N-1}{N(m-1)}\,\frac{\bar{x}}{\bar{X}}$, which on substitution gives the ratio estimator given by Nieto de Pascual [3].

If θ is a constant, it follows from (A$'$) that

$$\theta = -\frac{B\left[\frac{\bar{y}}{\bar{x}}\bar{X}\right]}{B\left[\frac{\bar{X}}{m}\sum\frac{y_i}{x_i}\right] - B\left[\frac{\bar{y}}{\bar{x}}\bar{X}\right]} \simeq -\frac{1}{m-1}$$

and substituting this value in the combined estimator we have once again the Murthy-Nanjamma estimator. It is interesting to note that while $\theta < 0$ in the case of Murthy-Nanjamma estimator, $E(\theta) < 0$ in the case of Nieto's estimator.

INDIAN STATISTICAL INSTITUTE, CALCUTTA

REFERENCES

[1] H. O. Hartley and A. Ross, "Unbiased ratio estimators," *Nature*, 174 (1954), 270-271.
[2] M. N. Murthy and N. S. Nanjamma, "Almost unbiased ratio estimates based on interpenetrating subsample estimates," *Sankhyā*, 21 (1959), 381-392.
[3] J. Nieto de Pascual, "Unbiased ratio estimators in stratified sampling," *J. Amer. Statist. Ass.*, 56 (1961), 70-87.