Conditions under which the Ratio Estimator is a Best Linear Unbiased Estimator

by Jee Euen Sook

Kwangwoon University, Seoul, Korea

A well-known result in regression theory indicates the type of population under which the ratio estimate may be called the best among a wide class of estimates. The result was first proved for infinite populations. Brewer and Royall extended the result to finite populations. The result holds if two conditions are satisfied.

1. The relation between $y_i$ and $x_i$ is a straight line through the origin.
2. The variance of $y_i$, about this line is proportional to $x_i$.

A "best linear unbiased estimator" is defined as follows. Consider all estimators $\hat{Y}$ of $Y$ that are linear functions of the sample values $y_i$, that is, that are of the form

$$l_1y_1 + l_2y_2 + \cdots + l_ny_n$$

where the $l_i$'s do not depend on the $y_i$, although they may be functions of the $x_i$. The choice of $l_i$'s is restricted to those that give unbiased estimation of $Y$. The estimator with the smallest variance is called the best linear unbiased estimator (BLUE).

Formally, Brewer and Royall assume that the $N$ population values $(y_i, x_i)$ are a random sample from a superpopulation in which

$$y_i = \beta x_i + e_i$$

where the $e_i$ are independent of the $x_i$ and $x_i > 0$. In arrays in which $x_i$ is fixed, $e_i$ has mean 0 and variance $\lambda x_i$. The $x_i$ ($i=1, 2, \ldots, N$) are known.

In the randomization theory, the finite population total $Y$ has been regarded as a fixed quantity. Under model (1), on the other hand, $Y = \beta X + \sum_{i=1}^{N} e_i$ is a random variable. In defining an unbiased estimator under this model, Brewer and Royall use a concept of unbiasedness which differs from that in randomization theory. They regard an estimator $\hat{Y}$ as unbiased if $E(\hat{Y}) = E(Y)$ in repeated selections of the finite population and sample under the model. Such an estimator might be called model-unbiased.

**Theorem.** Under model (1) the ratio estimator $\hat{Y}_R = \frac{X\bar{y}}{\bar{x}}$ is a best linear unbiased estimator for any sample, random or not, selected solely according to the values of the $x_i$.

**Proof.** Since $E(e_i/x_i) = 0$ in repeated sampling, it follows from (1) that

$$Y = \beta X + \sum_{i=1}^{N} e_i; \quad E(Y) = \beta X$$

Furthermore, with the model (1) any linear estimator $\hat{Y}$ is of the form

$$\hat{Y} = \sum_{i=1}^{N} l_i y_i = \beta \sum_{i=1}^{N} l_i x_i + \sum_{i=1}^{N} l_i e_i$$
If we keep the \( n \) sample values \( x_i \) fixed in repeated sampling under the model (1),

\[
E(\hat{Y}) = \beta \sum_{i=1}^{n} l_i x_i; \quad V(\hat{Y}) = \lambda \sum_{i=1}^{n} l_i^2
\]

(4)

From (2) and (3), \( \hat{Y} \) is clearly model-unbiased if \( \sum_{i=1}^{n} l_i x_i = X \). Minimizing \( V(\hat{Y}) \) under this condition by a Lagrange multiplier gives

\[
2l_i x_i = c x_i; \quad l_i = \text{constant} = X/n \hat{x}
\]

(5)

The constant must have the value \( X/n \hat{x} \) in order to satisfy the model-unbiased condition \( \sum_{i=1}^{n} l_i x_i = X \).

Hence the BLUE estimator \( \hat{Y} \) is \( n \hat{y} X/n \hat{x} = X \hat{y}/\hat{x} = \hat{Y}_R \), the usual ratio estimator. This completes the proof.

Furthermore, from (2) and (3), with \( l_i = X/n \hat{x} \),

\[
\hat{Y}_R - Y = \sum_{i=1}^{n} l_i e_i = (X/n \hat{x}) (\sum_{i=1}^{n} e_i) - \sum_{i=1}^{n} e_i
\]

(6)

\[
= \frac{(X-n \hat{x})}{n \hat{x}} \sum_{i=1}^{n} e_i - \sum_{i=1}^{n} e_i
\]

(7)

where \( \sum_{i=1}^{n} \) denotes the sum over the \( (N-n) \) population values that are not in the sample. Hence

\[
V(\hat{Y}_R) = \frac{\lambda(X-n \hat{x})(n \hat{x})}{(n \hat{x})^2} + \lambda(X-n \hat{x}) = \frac{\lambda(X-n \hat{x})X}{n \hat{x}}
\]

(8)

A model-unbiased estimator of \( \lambda \) from this sample is easily shown to be

\[
\lambda = \frac{1}{\sum_{i=1}^{n} x_i^2} \sum_{i=1}^{n} (y_i - \hat{R} x_i)^2 / (n-1)
\]

(9)

where \( \hat{R} = \hat{y}/\hat{x} \), as usual. This value may be substituted in (8) to give a model-unbiased sample estimate of \( V(\hat{Y}_R) \).

The practical relevance of these results is that they suggest the conditions under which the ratio estimator is superior not only to \( \hat{y} \) but is the best of a whole class of estimators. When we are trying to decide what kind of estimate to use, a graph in which the sample values of \( y_i \) are plotted against those of \( x_i \) is helpful. If this graph shows a straight line relation passing through the origin and if the variance of the points \( y_i \) about the line seems roughly proportional to \( x_i \), the ratio estimator will be hard to beat.

Sometimes the variance of the \( y_i \) in arrays in which \( x_i \) is fixed is not proportional to \( x_i \). If this residual variance is of the form \( \lambda v(x_i) \), \( v(x_i) \) is known, Brewer and Royall showed that the BLUE estimator becomes

\[
\hat{Y} = X \frac{\sum_{i=1}^{n} w_i y_i x_i}{\sum_{i=1}^{n} w_i x_i^2}
\]

(10)

where \( w_i = 1/v(x_i) \). In a population sample of Greece, Jessen et al. (1947) judged that the residual variance increased roughly as \( x_i^2 \). This suggests a weighted regression with \( w_i = 1/x_i^2 \), which gives

\[
\hat{Y} = \frac{X}{n} \sum_{i=1}^{n} \left( \frac{y_i}{x_i} \right) = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{y_i}{x_i} \right)
\]

(11)

For a given population and given \( n \), \( V(\hat{Y}_R) \) in (8) is clearly minimized, given every \( x_i > 0 \), when the sample consists of the \( n \) largest \( x_i \) in the population. In [16] small natural populations of the type to which ratio estimates have been applied, Royall (1970) found for samples having \( n=2 \) to
that selection of the \( n \) largest \( x_i \) usually increased the accuracy of \( \hat{Y}_b \).

In summary, the Brewer-Royall results show that the assumption of a certain type of model leads to an unbiased ratio estimator and formulas for \( V(\hat{Y}_b) \) and practice in cases where examination of the \( y, x \) pairs from the available data suggests that the model is reasonably correct. The variance formulas (8) and (9) appear to be sensitive to inaccuracy in the model, although this issue needs further study.

Further work by Royall and Herson (1973) discusses the type of sample distribution needed with respect to the \( x_i \) in order that \( \hat{Y}_b \) remains unbiased when there is a polynomial regression of \( y_i \) on \( x_i \).

**References**