ON THE BIAS AND MEAN SQUARE ERROR
OF THE RATIO OF TWO BIASED ESTIMATORS*

BY

Hermann Wiedenhofer, Jacob E. Bearman and

Frank B. Martin

Technical Report No. 192

December 1972

University of Minnesota
Minneapolis, Minnesota

*Research supported in part by Project 2830, Agricultural Experimental Station, University of Minnesota and be PHS (Grant/Contract) T01 HS-00053 from the National Center for Health Services Research and Development.

1. Summary and Introduction.

In the current literature of finite sampling theory, expressions for the bias and M.S.E. for the ratio of two unbiased estimators are well developed. Similar theoretical values for the ratio of two biased estimators, not available in the current literature, are obtained here.

It is well known that $\hat{R}$, the ratio of two unbiased estimators, is a biased estimator of the corresponding population ratio, R. It will be shown that the bias and M.S.E. of $\hat{R}$, when $\hat{R}$ is the ratio of two biased estimators has additional terms. An exact expression and a first and second approximation are given for the bias of $\hat{R}$. Further, an expression for the M.S.E. of $\hat{R}$ is given, utilizing the first approximation above.

The notation and procedures used by Cochran [1] and Raj [2] for the ratio estimation are gnerally followed.

2. Definitions.

Let R be a population ratio such that

$$R = \frac{Y}{X} \quad , \tag{2.1}$$

where X and Y are means or totals of population characteristics.

Let $\hat{R}$ be an estimator of R ,

$$\hat{R} = \frac{y}{x} \quad , \tag{2.2}$$

where x and y are biased estimators of X and Y respectively. Then,

$$E(y) = Y + B(y) \tag{2.3}$$

$$E(x) = X + B(x) \tag{2.4}$$

$$E(\hat{R}) = R + B(\hat{R}) \tag{2.5}$$

3. The Exact Bias of $\hat{R}$

Theorem 1

$$B(\hat{R}) = \frac{R[\frac{B(y)}{Y} - \frac{B(x)}{X}] - \frac{1}{X} \operatorname{Cov}(\hat{R},x)}{1 + \frac{B(x)}{X}} \quad . \tag{3.1}$$

Proof:

Using the usual definition for $\operatorname{Cov}(\hat{R},x)$ and expression (2.2), it follows that

$$\frac{1}{X} \operatorname{Cov}(\hat{R},x) = \frac{1}{X} [E(y) - E(\hat{R})\, E(x)] \quad , \tag{3.2}$$

and by substitution from (2.3) , (2.4) and (2.5) ,

$$\frac{1}{X} \operatorname{Cov}(\hat{R},x) = \frac{Y}{X} + \frac{B(y)}{X} - R[1 + \frac{B(x)}{X}] - B(\hat{R})\, [1 + \frac{B(x)}{X}]. \tag{3.3}$$

Solving (3.3) for $B(\hat{R})$ yields

$$B(\hat{R}) = \frac{R[\frac{B(y)}{Y} - \frac{B(x)}{X}] - \frac{1}{X} \operatorname{Cov}(\hat{R},x)}{1 + \frac{B(x)}{X}} \quad . \tag{3.4}$$

Q. E. D.

$\frac{B(x)}{X}$ and $\frac{B(y)}{Y}$ are the relative bias of x and y respectively, similar to the usual definition of $\frac{B(\hat{R})}{R}$ in Cochran [1]. Since the exact expression (3.4) for the bias of $\hat{R}$ is not always very useful, approximate expressions are obtained as $B_1(\hat{R})$ based on the first term, and $B_2(\hat{R})$ based on the first two terms in expression (4.1) of the Taylor's series expansion of $\frac{1}{x}$ around X.

4. Approximations for the Bias of $\hat{R}$.

$$\hat{R} - R = \frac{y-Rx}{X}\left[1 - \frac{(x-X)}{X} + \frac{(x-X)^2}{X^2} - \ldots\right] . \tag{4.1}$$

If $\frac{x-X}{X}$ in absolute value is rather small, then a satisfactory approximation can be obtained by considering the expectation of the R.H.S. of (4.2)

$$\hat{R} - R \doteq \frac{1}{X}(y-Rx) \tag{4.2}$$

By adding and subtracting $RB(x)$, substituting RX for Y , and by the definition of $E(x)$, the expression $y-Rx$ in (4.2) becomes

$$y - Rx = y-E(y) - R[x-E(x)] - RB(x) + B(y) . \tag{4.3}$$

Substituting back into (4.2) , yields

$$\hat{R} - R \doteq \frac{1}{X}\{ y-E(x) - R[x-E(x)] - RB(x) + B(y) \} . \tag{4.4}$$

Taking expectations of both sides of (4.4), the first approximation to the bias of $\hat{R}$ is

$$E(\hat{R}-R) \doteq B_1(\hat{R}) = R\left[\frac{B(y)}{Y} - \frac{B(x)}{X}\right] . \tag{4.5}$$

Using the first two terms in expression (4.1), $\hat{R}-R$ is better approximated by

$$\hat{R} - R \doteq (y-Rx)\left(\frac{1}{X} - \frac{x-X}{X^2}\right) . \tag{4.6}$$

By (4.3) and the fact that $x - X = [\ x-E(x)\ ] + B(x)$, the R.H.S. of (4.6) may be written as

$$\begin{aligned}(y-Rx)\left(\frac{1}{X} - \frac{x-X}{X^2}\right) = \frac{1}{X}\{\ & y-E(y) - R[\ x-E(x)\ ] \\ & - RB(x) + B(y)\} - \frac{1}{X^2}\{[x-E(x)]\ [y-E(y)] \\ & - R[x-E(x)]^2 + [x-E(x)]\ [B(y) - RB(x)]\} \\ & - \frac{B(x)}{X^2}\ \{y - E(y) - R[x-E(x)] - RB(x) + B(y)\} .\end{aligned} \tag{4.7}$$

Taking expectations of the R.H.S. of (4.7) ,

$$B_2(\hat{R}) = R[\frac{B(y)}{Y} - \frac{B(x)}{X}] \ [1 - \frac{B(x)}{X}]$$
$$+ \frac{1}{X^2} [RV(x) - Cov(x,y)] \ . \qquad (4.8)$$

To obtain a form parallel to that commonly occurring in textbooks, the following coefficients of variation and covariation are introduced.

$$CV(x) = \frac{[V(x)]^{\frac{1}{2}}}{X} \ . \qquad (4.9)$$

$$CC(x,y) = \frac{Cov(x,y)}{XY} \ . \qquad (4.10)$$

By substitution and manipulation, $B_2(\hat{R})$ can be written as

$$B_2(\hat{R}) = R\,\{[CV(x)]^2 - CC(x,y) + [\frac{B(y)}{Y} - \frac{B(x)}{X}] \ [1 - \frac{B(x)}{X}]\} \ . \qquad (4.11)$$

Note that if x and y are both unbiased or have the same relative bias, then (4.11) reduces to the commonly current textbook expression for the bias of the ratio estimator.

5. <u>The First Approximation to the Mean Square Error of $\hat{R}$.</u>

Squaring both sides of (4.4) and taking expectations on both sides,

$$M.S.E.(\hat{R}) = E(\hat{R}-R)^2 \doteq \frac{1}{X^2} \{V(y) + R^2 V(x) - 2R\,Cov(x,y) + [B(y) - R\,B(x)]^2\}. \qquad (5.1)$$

Factoring out R^2 from the R.H.S. of (5.1) ,

$$MSE(\hat{R}) \doteq R^2 \left\{[CV(y)]^2 + [CV(x)]^2 - 2CC\,(x,y) + \left[\frac{B(y)}{Y} - \frac{B(x)}{X}\right]^2\right\} \qquad (5.2)$$

Note that if x and y are both unbiased, then expression (5.2) reduces to the commonly current textbook expression for the approximation of the MSE of the ratio estimator.

In [3] a second approximation to the MSE is obtained, however, it is too cumbersome to be presented here.

6. References.

1. Cochran, William G., Sampling Techniques, 2nd ed., New York: John Wiley and Sons, Inc., 1963.

2. Raj, Des, Sampling Theory, New York: McGraw-Hill Book Co. Inc., 1968.

3. Wiedenhofer, Hermann, "Linear and Ratio Estimation Methods in Multistage Sampling, with Special Reference to a System of Voluntary Animal Disease Reporting," Ph.D. dissertation, University of Minnesota, 1972.