An Upper Limit to the Difference in Bias Between Two Ratio Estimates

by

John Clement Koop

Institute of Statistics
Mimeo Series No. 312
January 1962
DISCUSSION*

J. C. Koop
N. C. State College, Raleigh
North Carolina

In a concluding remark of the abstract to their paper Prof. Kish and Mr. Namboodirí have stated that

"For the comparison of two ratios \((r_1 - r_2)\) the bias ratio is found to have no necessary limit, but a reasonable and empirical limit appears in the greater of the two involved coefficients of variation \(Cx_1\) and \(Cx_2\)."

I think they were in a position to be somewhat more optimistic and not to have erred at all if only they had taken their algebra a little further as follows:

In their notation the identity connecting the expectation of the ratio of the two random variables to the true value and the bias is

\[
E(r) = \frac{E(y)}{E(x)} + E \left[ y \left( \frac{1}{x} - \frac{1}{E(x)} \right) \right],
\]

(1)
given by Koop (1951). Generally \(y\) and \(x\) can represent estimates (simple or otherwise) based on the same set of units in a sample survey. The second term on the right-hand side of the identity, which symbolizes the bias, yields the covariance expression thus

\[
E \left[ y \left( \frac{1}{x} - \frac{1}{E(x)} \right) \right] = -\frac{1}{E(x)} \text{cov} \left( \frac{y}{x}, x \right),
\]

(2)
an expression which was later discovered and used by Hartley and Ross (1954) to correct for the bias of their ratio estimate.

The bias as given by (2) can be further expressed as

\[- \frac{1}{E(x)} \rho \frac{\sigma_y}{\sigma_x} \sigma_y \]

(3)

where \( \rho \) is the true correlation between \( \frac{y}{x} \) and \( x \), \( \sigma_y \) the square root of the variance of \( \frac{y}{x} \), and \( \sigma_x \) is the standard deviation of \( x \), which in the notation of the two authors would reduce to \( -\rho \sigma_x \appa \) so that

\[E(r_{1-r_2} - \frac{R_1-R_2}{1-R_2}) = \rho_2 \sigma_2 \appa \sigma_2 - \rho_1 \sigma_1 \appa \sigma_1\]

(4)

and,

\[|E(r_{1-r_2} - \frac{R_1-R_2}{1-R_2})| \leq (|\rho_2| \appa \sigma_2 + |\rho_1| \appa \sigma_1)\]

(5)

Now if \( \sigma_2 \) is the larger of the two coefficients of variation, then the inequality at (5) becomes

\[|E(r_{1-r_2} - \frac{R_1-R_2}{1-R_2})| \leq \sigma_2 (|\rho_2| \appa \sigma_2 + |\rho_1| \appa \sigma_1) \]

(6)

But

\[(\rho_1^2 + \rho_2^2)(\sigma_1^2 + \sigma_2^2) \geq (|\rho_1| \appa \sigma_1 + |\rho_2| \appa \sigma_2)^2\]

by Cauchy's inequality, the sign of equality holding if and only if

\[|\rho_1| / |\rho_2| = \sigma_1^2 / \sigma_2^2\]

Hence

\[|E(r_{1-r_2} - \frac{R_1-R_2}{1-R_2})| < \sigma_2 \sqrt{\rho_1^2 + \rho_2^2} < \sqrt{2} \sigma_2\]

(7)

showing that an upper limit to the absolute bias difference, standardized by the standard error of \( (r_{1-r_2}) \), exists. Its measure assumes most significance when the individual biases are in opposite directions and when \( \sigma_2 \) is large; in this situation a comparison of differences \( (r_{1-r_2}) \) may not be meaningful.

When \( \rho_1^2 + \rho_2^2 = 1 \), the "empirical limit" which the authors conjectured on the basis of evidence provided by their surveys will be obtained, but only if

standardized as above.

References


An Upper Limit to the Difference in Bias Between Two Ratio Estimates

by

J. C. Koop

Corrigenda

On p. 2, in the 10th line from the bottom of the page, instead of

"by the standard error of \((r_1 - r_2)\)"

read

"by the square root of the sum of the variances of \(r_1\) and \(r_2\)."
290. Schutzenberger, M. P. On the equation \( a^n = b^{m^2} \) in a free group. June, 1961.
292. Bhattacharya, P. K. Some properties of the least square estimator in regression analysis when the 'independent' variables are stochastic. June, 1961.