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. Namboodiri 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

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

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

$$E\left[y \left(\frac{1}{x} - \frac{1}{E(x)} \right) \right] = -\frac{1}{E(x)} \operatorname{cov}\left(\frac{y}{x}, x \right), \tag{2}$$

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

---

* Contribution to the discussion following a paper by L. Kish and N. K. Namboodiri entitled "The ratio bias in sample surveys" presented at the 121st Annual Meeting of the American Statistical Association, New York City, December 26-30, 1961.

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

$$- \frac{1}{E(x)} \rho \sigma_{\frac{y}{x}} \sigma_x , \qquad (3)$$

where ρ is the true correlation between $\frac{y}{x}$ and x, $\sigma_{\frac{y}{x}}$ the square root of the variance of $\frac{y}{x}$, and σ_x is the standard deviation of x, which in the notation of the two authors would reduce to $- \rho\sigma_r Cx$ so that

$$E(r_1 - r_2 - \overline{R_1 - R_2}) = \rho_2 \sigma_{r_2} Cx_2 - \rho_1 \sigma_{r_1} Cx_1 \qquad (4)$$

and, $$\left| E(r_1 - r_2 - \overline{R_1 - R_2}) \right| \leq \left(|\rho_2| \; \sigma_{r_2} Cx_2 + |\rho_1| \; \sigma_{r_1} Cx_1 \right) \qquad (5)$$

Now if Cx_2 is the larger of the two coefficients of variation, then the inequality at (5) becomes

$$\left| E(r_1 - r_2 - \overline{R_1 - R_2}) \right| < Cx_2 (|\rho_2| \; \sigma_{r_2} + |\rho_1| \; \sigma_{r_1}) . \qquad (6)$$

But

$$(\rho_1^2 + \rho_2^2)(\sigma_{r_1}^2 + \sigma_{r_2}^2) \geq (|\rho_1| \; \sigma_{r_1} + |\rho_2| \; \sigma_{r_2})^2$$

by Cauchy's inequality, the sign of equality holding if and only if $|\rho_1| / |\rho_2| = \sigma_{r_1} / \sigma_{r_2}$. Hence

$$\frac{\left| E(r_1 - r_2 - \overline{R_1 - R_2}) \right|}{\sqrt{\sigma_{r_1}^2 + \sigma_{r_2}^2}} < Cx_2 \sqrt{\rho_1^2 + \rho_2^2} < \sqrt{2} \; Cx_2 \qquad (7)$$

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 Cx_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

Koop, J. C. (1951) "A note on the bias of the ratio estimate," Bull. Int. Statist. Inst., 33 (2), 141-146.

Hartley, H. O., and Ross, A. (1954) "Unbiased ratio estimators," Nature, 174 270-271.

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."

INSTITUTE OF STATISTICS

NORTH CAROLINA STATE COLLEGE

(Mimeo Series available for distribution at cost)

265. Eicker, Friedhelm. Consistency of parameter-estimates in a linear time-series model. October, 1960.
266. Eicker, Friedhelm. A necessary and sufficient condition for consistency of the LS estimates in linear regression. October, 1960.
267. Smith, W. L. On some general renewal theorems for nonidentically distributed variables. October, 1960.
268. Duncan, D. B. Bayes rules for a common multiple comparisons problem and related Student-t problems. November, 1960.
269. Bose, R. C. Theorems in the additive theory of numbers. November, 1960.
270. Cooper, Dale and D. D. Mason. Available soil moisture as a stochastic process. December, 1960.
271. Eicker, Friedhelm. Central limit theorem and consistency in linear regression. December, 1960.
272. Rigney, Jackson A. The cooperative organization in wildlife statistics. Presented at the 14th Annual Meeting, Southeastern Association of Game and Fish Commissioners, Biloxi, Mississippi, October 23-26, 1960. Published in Mimeo Series, January, 1961.
273. Schutzenberger, M. P. On the definition of a certain class of automata. January, 1961.
274. Roy, S. N. and J. N. Shrizastaza. Inference on treatment effects and design of experiments in relation to such inferences. January, 1961.
275. Ray-Chaudhuri, D. K. An algorithm for a minimum cover of an abstract complex. February, 1961.
276. Lehman, E. H., Jr. and R. L. Anderson. Estimation of the scale parameter in the Weibull distribution using samples censored by time and by number of failures. March, 1961.
277. Hotelling, Harold. The behavior of some standard statistical tests under non-standard conditions. February, 1961.
278. Foata, Dominique. On the construction of Bose-Chaudhuri matrices with help of Abelian group characters. February, 1961.
279. Eicker, Friedhelm. Central limit theorem for sums over sets of random variables. February, 1961.
280. Bland, R. P. A minimum average risk solution for the problem of choosing the largest mean. March, 1961.
281. Williams, J. S., S. N. Roy and C. C. Cockerham. An evaluation of the worth of some selected indices. May, 1961.
282. Roy, S. N. and R. Gnanadesikan. Equality of two dispersion matrices against alternatives of intermediate specificity. April, 1961.
283. Schutzenberger, M. P. On the recurrence of patterns. April, 1961.
284. Bose, R. C. and I. M. Chakravarti. A coding problem arising in the transmission of numerical data. April, 1961.
285. Patel, M. S. Investigations on factorial designs. May, 1961.
286. Bishir, J. W. Two problems in the theory of stochastic branching processes. May, 1961.
287. Konsler, T. R. A quantitative analysis of the growth and regrowth of a forage crop. May, 1961.
288. Zaki, R. M. and R. L. Anderson. Applications of linear programming techniques to some problems of production planning over time. May, 1961.
289. Schutzenberger, M. P. A remark on finite transducers. June, 1961.
290. Schutzenberger, M. P. On the equation $a^{2+n} = b^{2+m}c^{2+p}$ in a free group. June, 1961.
291. Schutzenberger, M. P. On a special class of recurrent events. June, 1961.
292. Bhattacharya, P. K. Some properties of the least square estimator in regression analysis when the 'independent' variables are stochastic. June, 1961.
293. Murthy, V. K. On the general renewal process. June, 1961.
294. Ray-Chaudhuri, D. K. Application of geometry of quadrics of constructing PBIB designs. June, 1961.
295. Bose, R. C. Ternary error correcting codes and fractionally replicated designs. May, 1961.
296. Koop, J. C. Contributions to the general theory of sampling finite populations without replacement and with unequal probabilities. September, 1961.
297. Foradori, G. T. Some non-response sampling theory for two stage designs. Ph.D. Thesis. November, 1961.
298. Mallios, W. S. Some aspects of linear regression systems. Ph.D. Thesis. November, 1961.
299. Taeuber, R. C. On sampling with replacement: an axiomatic approach. Ph.D. Thesis. November, 1961.
300. Gross, A. J. On the construction of burst error correcting codes. August, 1961.
301. Srivastava, J. N. Contribution to the construction and analysis of designs. August, 1961.
302. Hoeffding, Wassily. The strong laws of large numbers for u-statistics. August, 1961.
303. Roy, S. N. Some recent results in normal multivariate confidence bounds. August, 1961.
304. Roy, S. N. Some remarks on normal multivariate analysis of variance. August, 1961.
305. Smith, W. L. A necessary and sufficient condition for the convergence of the renewal density. August, 1961.
306. Smith, W. L. A note on characteristic functions which vanish identically in an interval. September, 1961.
307. Fukushima, Kozo. A comparison of sequential tests for the Poisson parameter. September, 1961.
308. Hall, W. J. Some sequential analogs of Stein's two-stage test. September, 1961.
309. Bhattacharya, P. K. Use of concomitant measurements in the design and analysis of experiments. November, 1961.