

DISCUSSION: MISCELLANEOUS STATISTICAL ANALYSIS, SOCIAL STATISTICS  
CONTRIBUTED PAPERS SESSION XXV

Gary M. Shapiro, U. S. Bureau of the Census

I will discuss only two of the papers given, those by Sukhatme-Chand and by Kondo-Schnack.

I. SUKHATME-CHAND PAPER

The Sukhatme-Chand paper is a good, professional paper. I do question, however, how much practical value it has. Are there really many instances in which the results can be applied? The paper is primarily concerned with the following situation: Initially, a large sample is taken and one variable is observed; then a subsample is taken and a second variable is observed on this subsample; then a second subsample (a subset of the first subsample) is taken and the variable of real interest is observed. Do situations like this really occur?

In my personal experience, the auxiliary variables have usually not been estimated from samples, but rather have been subject to zero variance. This situation is considered in the paper as a special case and Theorem 5.1 in particular is useful. However, the formulae given for the bias and mean square error for this special case are trivial. Also, the cost discussion and example are not applicable to this situation.

II. KONDO-SCHNACK PAPER

One major advantage of the Keyfitz (or Taylor series linearized) approach [3], [4], [5] over the replication approach to variance estimation has been the inability of the replication variance approach to estimate the components of variance separately. Thus, a general elimination of this inability would significantly improve the value of replication variance estimation. In Casady [2] and Bean [1] (referenced in the Kondo-Schnack papers), it has been shown that components estimation is possible in the case of linear estimates. Linear estimates are of little interest, though, as much simpler methods of variance estimation than either replication or Keyfitz are acceptable. It appears to me, however, that the same procedure will work as well for nonlinear estimates as for linear. I believe Casady's results can be easily generalized for nonlinear estimates.

The Kondo-Schnack paper fully accomplishes its rather modest goal of illustrating the application of the previously developed theory of Casady [2] and Bean [1]. The replication procedure for estimating components is somewhat inconvenient in that completely separate sets of replications are needed for the within and total variance estimates, but this doesn't appear to have caused the authors any major problems. The empirical results are acceptable; some negative estimates of between PSU variance are obtained, but these would also be likely to occur if Keyfitz variance estimates had been made.

I have two minor criticisms of the paper. First, a reader of the paper is left with the impression that the National Center for Health Statistics (NCHS) designed and conducted the Health Interview Survey. In fact, the Bureau of the Census, under contract to NCHS, conducted the survey and was primarily responsible for the design. This is a common problem: Papers written by staff members of organizations who sponsor surveys quite frequently fail to properly acknowledge the organization which actually designed and conducted a survey. The second criticism relates to some incorrect numbers given in the paper. Health Interview Survey was redesigned in 1973, with 376 sample PSU's instead of the previous 357 sample PSU's, and with clusters averaging four households instead of the previous six.

REFERENCES

- [1] Bean, Judy A. "Unbiased BRR Estimator of Within Component for Two-Stage Stratified Design," Internal University of Iowa memorandum, 1977.
- [2] Casady, Robert J. "The Estimation of Variance Components Using Balanced Repeated Replication." Proceedings of the Social Statistics Section, American Statistical Association, 1975, pp. 352-357.
- [3] Keyfitz, Nathan. "Estimates of Sampling Variance Where Two Units are Selected from Each Stratum." Journal of the American Statistical Association, Vol. 52, (Dec. 1957), pp. 503-510.
- [4] Tepping, Benjamin J. "Variance Estimation in Complex Surveys." Proceedings of the Social Statistics Section, American Statistical Association, 1968, pp. 11-18.
- [5] Woodruff, Ralph H. "A Simple Method of Approximating the Variance of a Complicated Estimate." Journal of the American Statistical Association, Vol. 66, No. 334, (June 1971), pp. 411-414.