DISCUSSION

Stanley Lemeshow, University of Massachusetts

The first paper presented in this session by Maurer, Jones and Bryant addresses a very important issue in deciding among competing methods of variance estimation - namely the issue of cost. This is particularly important in large-scale sample surveys, where literally hundreds upon hundreds of such calculations are performed and costs can rapidly become prohibitive.

The authors chose to compare two of the available methods for variance estimation - balanced repeated replication and linearization. There are others available - specifically the jackknife procedure - but, to my knowledge, this method is not currently used in any of the ongoing large-scale surveys and there undoubtedly was no computer program readily available for including this procedure in the present paper.

There is some evidence that the jackknife deserves more serious attention than it has received in the past but I'll discuss that in more detail a bit later on.

The basic idea of the paper is good: use the competing methods to do the same thing with the same data and see which does so most efficiently. Unfortunately, the results of this investigation are not so much a commentary on the methods of variance estimation as upon the computer programs available for computation. The authors recognize this for, among other things, they point out that:

-Neither program is written in its most efficient form.

-Different programmers wrote the two programs and used different languages (PL1 for the BRR program, FORTRAN for the LIN program).

-The LIN program is not set-up to handle a large table in a single pass efficiently. Instead, each cell of the table is treated as a separate problem.

Of interest are the findings "within" a particular method. Specifically, for LIN: CPU time is effected only minimally by the number of strata; the CPU time is dramatically increased by the number of cells in the table since the Taylor expansion algorithm is re-derived for each cell; the CPU time increases as the number of input records increases. For BRR: CPU time is increased by increasing the number of strata; CPU time is uneffected by the number of cells in the table; CPU time increases as the number of records increases.

Of course, if these are the only two computer programs available for computing BRR or Linearization estimates, and if one assumes that the same variance estimates are obtainable from each, the comparison dictates which program to use under a variety of circumstances. The authors wisely do not attempt to judge the two methods since this research does not add to our knowledge concerning their relative cost efficiencies.

Now, I'd like to spend a little time discussing the jackknife method. My reference to the jackknife differs from those who use the jackknife for the classical purpose of reducing bias and variability of an estimator under simple random sampling. Instead, I refer to its use as a method of estimating variances in complex, large scale surveys for which precise expressions for variances of desired estimates cannot be derived. Its potential use is exactly the same as BRR or Linearization.

The major advantage that both the jackknife and BRR have over linearization is that irrespective of the type of estimator (i.e., ratio estimate, slope, correlation coefficient, percentile) for which a standard error is desired, the process is the same. With the linearization method however, the Taylor Series expansion must be derived for the specific type of estimator. This, of course, is not really so objectionable since, once derived, the process is quite repetitive and easily programmed. However, in some instances, the forms of these expansions are not at all obvious and, in others, may not be possible to obtain.

With both BRR and the jackknife (JRR), any estimator of interest is calculated from the entire sample as well as from a number of subsets of the original sample. It is the sampling variability of the estimates from the subsets which allows an estimate of the standard error to be determined.

Past research has shown that JRR estimates do just as well as BRR and Linearized ones in controlled situations. An objection raised about the jackknife (as recently as the present session in the paper by Donald Sears) is that it is extremely time consuming and expensive. This objection should be studied carefully to see if it is, in fact, true. In my research, I have not found this to be the case although I haven't studied it systematically or extensively. Intuitively, it would not seem to be much more expensive than BRR since, by clever programming, much time may be saved in the calculation process.

For example, if there are 16 strata from each of which 2 PSU's have been selected, 16 half-sample estimates must be made as compared to 32 jackknife estimates. However, by employing the fact that each jackknife estimate is based on all but one of the PSU's, calculations may be made within each PSU and new jackknife estimates obtained by subtracting the results of one out and adding in the results of another. This would be far more difficult to do within the process of obtaining half-sample estimates.

The potential for the jackknife method has been documented in the past. In particular:

-JRR variance estimates are easily obtained when there are more than two PSU's per stratum. While it is possible to obtain a balanced set of "third samples", it is not necessarily an easy task to do so for 4 or more PSU's per stratum.

-There is no reason why the jackknife couldn't be used, with the appropriate weighting factors, for situations where there are unequal numbers of PSU's in each stratum.

-The jackknife method produces variances of regression coefficients with far more stability and far less bias than does the BHS method. This is particularly true when the number of individuals in each PSU is small - as may well be the case when variance estimates are desired for specific domains of interest. When there are many individuals in each PSU, both methods seem to perform well.

-When neither the numerator or denominator random variable of a combined ratio estimate are distributed normally, the BRR method produces standard errors which are highly biased and quite variable. The jackknife estimates are considerably better - particularly when there are few observations in each PSU (as with small domains of interest).

In summary, this is an area in which there are many unanswered questions and much research can certainly be done. I do not believe the three methods BRR, JRR and Linearization are all equally good - although I don't doubt the equality of their abilities in the linear case or with ratio estimates. The question of efficiency in those cases would be the deciding factor. I do hope there will be more research done in this area and great flexibility and open-mindedness shown by the large survey organizations for implementing the conclusions of that research.