Journal of Official Statistics, Vol. 23, No. 4, 2007, pp. 453–454

Discussion

Ray Chambers[1]

It is a pleasure to be asked to comment on this article. The issue of how to properly account for the variability of an estimate based partially on imputed data has been around a long time. As the author notes, Rubin's elegant multiple imputation (MI) solution to this problem is technically only valid when the imputation method is "Bayesian proper," i.e., it corresponds to a random draw from the posterior distribution of the missing values. Many imputation methods in common use in National Statistical Institutes (NSIs) are most definitely NOT Bayesian proper. More importantly, the vast majority are single draw imputation methods. As a consequence, there is certainly a perceived need for an approach to variance estimation in these cases that is comparable in simplicity (conceptually at least) to Rubin's Bayesian MI approach.

This article represents an attempt to sketch out how such an approach might work under what might be termed "Bayesian improper" MI. The key to the development in the article is the proposition (1) that represents the estimated variance of the multiple imputation-based estimate $\bar{\theta}^*$ as the sum of the average $\bar{V}^*$ of the estimated "full sample" variances from each imputation and a quantity that is proportional to the between imputation variance B^*, with the constant of proportionality chosen so that this sum is an unbiased estimator of the actual variance of $\bar{\theta}^*$. The article then goes on to show that in a number of simple MI situations this constant depends only on the proportion of values imputed in the data set of interest. In more complicated situations, however, it turns out that this constant depends on the characteristics of the nonrespondents (see Section 3.2) or on the actual variability of the respondents (see Section 5.1.2).

I have two misgivings about the approach set out in this article. The first relates to its practical impact. As I mentioned above, most imputation strategies used in NSIs are single imputation strategies. In such cases the representation (1) for the estimated variance is not defined, since B^* is not defined. Consequently all the results developed in the article become irrelevant. We have no option but to work out the actual variance of the single draw estimator $\hat{\theta}^*$ and how to estimate it. The second relates to the usefulness (under MI) of the results presented in the article. In my experience, surveys are typically highly multivariate, with the amount of missingness varying considerably between different variables. Even under the very straightforward scenarios explored in this article (MCAR or MCAR under stratification) it is clear that implementation of the MI variance "formula" (1) will be variable specific, requiring information (e.g., variable specific response rates) that may not be readily available to a secondary data analyst. This runs contrary to the

[1]University of Wollongong, Centre for Statistical and Survey Methodology, Wollongong NSW 2522, Australia. Email: ray@uow.edu.au

spirit of Rubin's MI variance estimator, which requires no extra knowledge about the nonresponse process beyond access to the multiple imputed data sets. I must also admit to a third (and somewhat carping) criticism, which is that the focus on MCAR in the article makes its results less extensible than might otherwise be the case. It is generally recognised (hoped?) that most nonresponse mechanisms are MAR, and a general development of this case would have been welcomed.

The above criticisms notwithstanding, there is one very positive message in this article. It is not spelled out explicitly, but is implicit in what the author proposes. This corresponds to the statement that if single imputation is a good thing, then multiple imputation is a better thing. It seems obvious that the MI estimator $\bar{\theta}^*$ is preferable to the single imputation estimator $\hat{\theta}^*$ since it has less imputation-induced variability. However, many NSIs persist with single imputation strategies. The argument that this is justified because variance estimation under "Bayesian improper" MI is just too difficult has always seemed rather self-serving to me. This article demonstrates that a simple "MI-like" variance estimation strategy is possible for a number of the imputation methods used in practice, and consequently represents a valuable contribution to official statistics methodology.

Received January 2006