Discussion on

“Dissecting Multiple Imputation from a Multi-phase Inference Perspective: What Happens When God’s, Imputer’s and Analyst’s Models Are Uncongenial?”

Yi-Hau Chen

Institute of Statistical Science, Academia Sinica

yhchen@stat.sinica.edu.tw

The authors make an important contribution to the discussion of the variance estimation of Rubin’s multiple imputation (MI) inference (Rubin (1987)). In particular, assuming the imputer’s model is correctly specified while the analyst’s may not be, the “uncongeniality” considered in the paper, the authors identify sufficient conditions for the validity of the MI inference in terms of the relative efficiency between the imputer’s and the analyst’s observed-data estimators. Although it has been well known in practice that imputation should be based on a sufficiently saturated model, the results in Xie and Meng (2016), especially Theorems 6 and 7, do provide substantial new insights into how the MI inference works in general.

Using the notation in the paper, the two components of Rubin’s MI variance estimator $T_\infty, \bar{U}_\infty$ and $B_\infty$, are respectively consistent estimators for the variances of $\hat{\theta}^A_{\text{com}}$ and $\bar{\theta}_{\infty} - \hat{\theta}^A_{\text{com}}$, where $\hat{\theta}^A_{\text{com}}$ and $\bar{\theta}_{\infty}$ are, respectively, the analyst’s complete-data and the MI estimates for the analyst’s model parameter, regardless of whether the imputer’s and the analysts’s models are congenial or not. The sufficient and necessary condition for $T_\infty$ consistently estimating