IS MULTIPLE IMPUTATION SAFE FOR EVERYDAY PRACTICE?

J. Carlin

Royal Children's Hospital, Parkville, Victoria, Australia

† E-mail: jbcarlin@unimelb.edu.au

In recent years the method of multiple imputation has begun to penetrate into widespread data analytic practice as a standard tool for handling missing data, due to the increasing availability of software tools in major statistical packages. At times the use of the method substantially outruns the bounds within which it can be confidently assumed to be trustworthy, and few routine users are likely to appreciate the potential dependence of results on untestable assumptions. In this context, applied research examining the sensitivity of findings in large epidemiological studies with missing data to the method used for imputation is important. In this talk I will briefly review the theory of multiple imputation and the state of development of software tools for creating imputations and for carrying out the analysis of imputed data. In the context of our experience with a large cohort study I will present results of sensitivity analyses and simulation studies examining the dependence of conclusions on imputation methods. In particular I will focus on comparisons between fully model-based imputation assuming multivariate normal distributions and the method of "chained equations" or "switching regressions". The latter method provides great flexibility in model specification and model checking, although its theoretical credentials remain unclear.