NMARness and approximate population bias of the direct MLE in the analysis of missing data

Yutaka Kano[*]

Osaka University, Japan. *kano@sigmath.es.osaka-u.ac.jp*

Abstract:

The method of direct maximum likelihood (DML), without use of missing-data mechanism, can produce consistent estimators for parameters of interest if the missing-data mechanism is MAR and the model is correctly specified for observable variables. The DML generally causes biased estimators for NMAR missingness and/or misspecified models. Yuan (2007, JMVA) studied unbiasedness of the DMLE when the model is misspecified but the MAR holds.

One way of reducing the bias caused by NMAR missingness is to include auxiliary variables in the model under consideration (Graham 2003 SEM; Graham 2009 Annu. Rev. Psychol.; Collins, Shafer & Kam 2001 Psychol. Methods). It is said that the background of the method is that by inclusion of auxiliary variables the NMAR missing-data mechanism gets close to MAR.

In this talk, we first define NMARness which denotes degree of departure from MAR missingness, and next we define approximate population bias (APB) of the DMLE under NMAR missingness. The APB represents magnitude of the bias of the DMLE for NMAR missingness. Theoretical properties of NMARness and APB are studied. We use those notions to study effects of auxiliary variables on reduction of the bias of the DMLE. It is turned out that one can distinguish among three cases which explain reduction of the bias due to auxiliary variables. To get close to MAR is one of these cases.