

Roderick Little, University of Michigan

Partially Missing At Random and Ignorable Inferences for Parameter Subsets with Missing Data

Abstract : For likelihood-based inferences from data with missing values, models are generally needed for both the data and the missing-data mechanism. Modeling the mechanism can be difficult, because its parameters are often poorly identified and results are often sensitive to model misspecification. Rubin (1976) showed that sufficient conditions for ignoring the missing data mechanism are that (a) the missing data are missing at random (MAR), and (b) the parameter of the data model and the missing-data mechanism are distinct. These conditions are not always necessary, and they relate to the complete set of parameters in the model. We propose definitions of partial MAR and ignorability for a subvector of parameters of particular substantive interest. We apply these definitions to regression with missing covariates, where an ignorable likelihood method is applied to a subsample of observations that are complete on a set of variables, but possibly incomplete on others. This method gives valid estimates for mechanisms where both complete-case analysis and full ignorable likelihood methods are biased.