1 Imputation of Missing Data

1.1 Multiple Imputation

Multiple imputation provides a useful strategy for dealing with data sets with missing values. Instead of filling in a single value for each missing value, the multiple imputation procedure replaces each missing value with a set of plausible values that represent the uncertainty about the right value to impute. These multiply imputed data sets are then analyzed by using standard procedures for complete data and then combining the results from all of these analyses. No matter which complete-data analysis is used, the process of combining results from different imputed data sets is essentially the same. This results in valid statistical inferences that properly reflect the uncertainty due to missing values.

1.1.1 Phases of Multiple Imputation

Multiple imputation inference involves three distinct phases:

- The missing data are filled in m times to generate m complete data sets.
- The m complete data sets are analyzed by using standard procedures.
- The results from the m complete data sets are combined for the inference.