

GLOSSARY

Error of Non-observation:

Inability to obtain data on all questionnaire items from all persons (Missing Data)

Unit Nonresponse:

A whole unit fails to provide data.

Item Nonresponse:

Data on particular items are unavailable for analysis, but the unit participates.

Missing Completely At Random (MCAR):

Data on a variable are missing **completely** at random if the missingness is unrelated to the (unknown) value of that variable, **and** is unrelated to the values of other variables. That is, the missing values are a random sample of all values and not related to any observed or unobserved variable.

Missing At Random (MAR):

Data are missing at random if the missingness is possibly related to the observed data in the data set, but, conditional on these data, **not** to the (unknown) value of the variable itself. In other words, the missing values are a random sample of all values within classes defined by observed values (i.e., conditional on the observed data the missingness is the result of a random process).

Not Missing At Random (NMAR):

The missingness on a variable is related to the unobserved (missing) value of that variable.

Ignorable:

If the data are MCAR or MAR **and** if the proper statistical model is used, the missingness is said to be ignorable with respect to a particular type of inference (e.g., likelihood-based or Bayesian). In this case, the observed data likelihood does not depend on the missingness probabilities.

Nonignorable:

When the missingness is dependent on an unobserved variable (NMAR) and the cause of missingness is unknown, and therefore cannot be included in the analysis, the missingness is said to be non-ignorable. In this case, a model for the missingness must be postulated and included in the analysis to prevent bias.

Single imputation: Each missing value in a data set is filled in only once with one value, yielding a completed data set.

Multiple imputation: Each missing value is replaced by a set of m ($m > 1$) values in turn, resulting in m data sets. These are analyzed and the results are combined in a single final estimate and its associated variation. This variation reflects both sampling variance and the imprecision with which the missing data can be predicted from the observed data.