

Missing data analysis in surveys

Universitat Autònoma de Barcelona

22, 24 and 26, November, 2010

Gonzalo Rivero ✉ gonzalo.rivero@nyu.edu
New York University

Aim: Familiarize the participants with the fundamental concepts of multiple imputation applied to social surveys.

Software: The workshop is designed as an introduction to the commands `ice` and `mi` in **Stata**. I will make occasional references to **R** and **Amelia II**.

Programme

Session 1 *Introduction to incomplete datasets*

- Motivation and definition of *missing value*.
- Structure of the missing data. The ignorability assumption.
- Management and description of incomplete data using **Stata**.
- Single imputation. Limitations.

Session 2 *Multiple imputation (I)*

- Basic intuitions of multiple imputation.
- Combination of estimates.
- Number of imputations.
- Multivariate normal method. Implementation in **Stata**.

Session 3 *Multiple imputation (II)*

- Chained equations. Implementation in **Stata**.
- Specification of the imputation model.
- Transformation of variables.
- Diagnoses and sensitivity analysis.
- Practical case.

Bibliography

- Allison, P. D. (2002). *Missing Data*. Sage University.
- de Leeuw, E. (2001). Reducing missing data in surveys: an overview of methods. *Quality and Quantity* 35, 147–160.
- Graham, J. W., A. E. Olchowski, and T. D. Gilreath (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science* 8, 206–213.
- Horton, N. and S. Lipsitz (2001). Multiple imputation in practice: Comparison of software packages for regression models with missing variables. *American Statistician* 55, 244–254.

- Kenward, M. G. and J. Carpenter (2007). Multiple imputation: Current perspectives. *Statistical Methods in Medical Research* 16(3), 199–218.
- King, G., J. Honaker, A. Joseph, and K. Scheve (2001). Analyzing incomplete political science data: An alternative algorithm for multiple imputation. *American Political Science Review* 95, 49–69.
- Little, R. and D. Rubin (1987). *Statistical Analysis with Missing Data*. Wiley New York.
- Longford, N. T. (2005). *Missing Data and Small-Area Estimation: Modern Analytical Equipment for the Survey Statistician*. Springer.
- Rubin, D. (1987). *Multiple Imputation for Non-response in Surveys*. Wiley.
- Schafer, J. (1997). *Analysis of Incomplete Multivariate Data*. Chapman & Hall/CRC.
- Stata Corp (2009). *Multiple Imputation Reference Manual*. Stata Press.
- von Hippel, P. (2009). How to impute interactions, squares, and other transformed variables. *Sociological Methodology* 39, 265–291.