

Multiple Imputation in Practice
Summer Institute in Survey Research Techniques, 2017

Instructors: Dr. Trivellore Raghunathan and Patricia Berglund
Dates: July 5 - July 9, 2017
Lectures: Wednesday - Sunday, 9 am-noon
Computer Lab: Wednesday - Sunday, 1-4:30 pm
Locations: Lecture - TBD Computer Labs - TBD

Instructor Information:

Patricia Berglund	T.E. Raghunathan
Room 4023- ISR	SRC Director's Office-ISR
pberg@isr.umich.edu	teraghu@umich.edu

Topics and Course Requirements

Course Description

Multiple Imputation is a versatile and general purpose method for analyzing data with some missing values. Under this approach, the missing set of values is replaced by several plausible setsof values to generate completed data sets. Each completed data set is analyzed separately and the resulting inferences, such as point estimates and their covariance matrices and test statistics are combined using simple rules. The broad objective of this course is to provide a practical guide for multiple imputation analysis from simple to complex problems using real and simulated data sets. The course will consist of lectures and practice sessions in the computer lab.

The data sets will be from a range of studies: cross-sectional, retrospective, prospective longitudinal studies; randomized clinical trials, and complex sample surveys. Simulated data sets will be used to illustrate and explore certain methodological aspects. The emphasis will be on the methods of analysis. The lab sessions will use IVEware (iveware.org) as stand-alone software and with the statistical packages SAS, Stata, SPSS and R.

Textbook and Class Reading

Two texts are recommended but not required.

1. Berglund and Heeringa, "Multiple Imputation of Missing Data using SAS", <https://support.sas.com/publishing/authors/berglund.html>
2. Raghunathan, "Missing Data Analysis in Practice", <https://www.crcpress.com/Missing-Data-Analysis-in-Practice/Raghunathan/9781482211924>

Prerequisites

This class requires at least one graduate level course in Statistics or BioStatistics along with basic working knowledge of at least one software such as SAS, SPSS, Stata, or R. Our computing sessions will use UM supplied computers and focus on use of SAS and IVEware. Use of personal laptops with additional software is encouraged, however, we will not be able to provide extensive help with software other than SAS.

Workshop Format**Lectures**

Lectures covering theoretical and statistical background will be held each morning from 9-noon. The lectures and discussion will be at an intermediate statistical level and will assume basic understanding of statistics and computing, as outlined above in the Prerequisites.

Computing Lab

Afternoon computing sessions will include hands-on computing exercises using IVEware for multiple imputation and analysis of imputed data sets. The labs will be held from 1pm-4:30m each afternoon in a campus computing lab. Code to do typical analyses including descriptive analysis, linear, logistic, and survival modeling, and other selected topics will be provided and students will have time to work through exercises during lab sessions.

Accommodations for Students with Disabilities - University of Michigan

If you think you need an accommodation for a disability, please contact Services for Students with Disabilities (SSD) office to help us determine appropriate academic accommodations. SSD (734-763-3000; <http://ssd.umich.edu>) typically recommends accommodations through a Verified Individualized Services and Accommodations (VISA) form. Any information you provide is private and confidential and will be treated as such.

Course Syllabus

Workshop: Multiple Imputation in Practice

Date	Day	Type	Topic / Assignments
7-5	Wednesday	Lecture Lab	Introduction to missing data problems, multiple imputation, analysis of imputed data sets. Theoretical background on MI process. Discussion of descriptive analysis of multiply imputed data, Taylor Series Linearization and Jackknife Repeated Replication techniques for complex sample design data. Overview of software to implement MI with focus on IVEware. Discussion of workshop data sets, examples of use of IVEware to perform imputation of missing data, descriptive analysis of completed data sets including continuous and categorical variables.
7-6	Thursday	Lecture Lab	Lecture on multiple imputation and analysis using linear, logistic and Poisson regression. Short review of lab 1 topics. Imputation of missing data and analysis of imputed data sets using IVEware for linear, logistic, and Poisson regression models.
7-7	Friday	Lecture Lab	Lecture on multiple imputation and analysis using survival analysis and structural equation modeling (SEM). Imputation of missing data and analysis of imputed data sets using IVEware for survival analysis and SEM.
7-8	Saturday	Lecture Lab	Lecture on multiple imputation and analysis of longitudinal data. Imputation of missing data and analysis of longitudinal data using IVEware with SASMOD and PROC GENMOD/PROC MIXED and related techniques.
7-9	Sunday	Lecture Lab	Lecture on multiple imputation using additional techniques available in IVEware. Special topics in MI not previously covered. Discussion of “real world” challenges in imputation and analysis. General review and wrap up. Discussion of “real world” computing challenges in imputation and analysis. Individual consultation with students, if time allows.