Missing Data Analysis

Presenter: Craig Enders, PhD, Associate Professor, Psychology, Arizona State University

Purpose of the workshop, including specific learning objectives

There have been substantial methodological advances in the area of missing data analyses during the last 25 years. Two missing data techniques, maximum likelihood (ML) and multiple imputation (MI) are currently considered “state of the art” in the methodological literature. These techniques have a strong theoretical foundation, and are “robust”, in the sense that they make less strict assumptions about the cause of the missing data. These procedures have a strong theoretical foundation, and are also supported by a large body of empirical work. Methodological studies have demonstrated that ML and MI are less prone to bias, and are more powerful than “traditional” missing data techniques (e.g., listwise deletion of cases, replacing missing values with the mean) that are ubiquitous in published research studies.

The primary goal of this course is to provide participants with the skills necessary to understand and appropriately implement ML or MI in their own research studies. In achieving this goal, it will be important to present technical information in a way that is easily accessible to researchers with limited statistical expertise. The session will provide a mixture of theoretical information, and computer applications. All of the lecture material will be provided in presentation slide format, and extensive computer code will be provided in the handouts, along with computer outputs. The session will emphasize the application of these techniques, so participants should leave the training session with new skills and the background knowledge that allows them to appropriately apply these new methods to their own research. The specific goals of the session are as follows: (a) provide a conceptual overview of missing data theory and assumptions, (b) briefly discuss “traditional” techniques, and demonstrate their shortcomings, (c) discuss ML estimation and how it applies to missing data, (d) discuss MI, (e) demonstrate the use of ML and MI, and discuss the relevant computer output from these procedures.

Target workshop audience

The targeted workshop audience is individuals at any professional stage who are working with, or contemplating working with, data sets with missing values. Participants are assumed to have
familiarity with multiple regression (e.g., a graduate-level course in regression, and experience applying regression models to data). Primarily, the presentation will be using Mplus; however, participants will not need extensive prior experience in Mplus. A brief primer on Mplus will be included in the slides.

**Materials to be provided to attendees**

Attendees will be provided with handouts of all presentations, including computer syntax and output.

**Presenter**

Craig Enders, Ph.D., is an Associate Professor in the Quantitative Psychology concentration in the Department of Psychology at Arizona State University, where he teaches graduate-level courses in missing data analyses, multilevel modeling, and longitudinal modeling. The majority of his research focuses on analytic issues related to missing data analyses and multilevel modeling. His book, *Applied Missing Data Analysis*, was published with Guilford Press in 2010.

**Workshop Outline**

- Orientation to Workshop

  The workshop topics will roughly divide into three sessions.

Session 1

1. Overview of missing data theory and assumptions
2. Discussion of “traditional” missing data techniques

Session 2

3. Discussion of ML estimation
4. Implementation of ML estimation in software packages (review of computer code and output)

Session 3

5. Discussion of MI.
6. Implementation of MI (review of computer code and output).

**Schedule**

8:30 am – 10 am
10:00 am – 10:15 am break
10:15 am – 12:15 pm
12:15 pm – 1:15 pm lunch
1:15 pm – 3:00 pm Session
3:00 pm – 3:15 break