Title:
Proper Handling of Missing Data: An Overview and Brief Tutorial

Abstract/summary:
Whether missing data are present or not, the goal of a statistical model is to obtain valid and efficient estimates of population parameters. It is often the case that some of the data being used to fit the statistical model are missing. The extent to which these data are not missing completely at random can substantially compromise one's ability to obtain the desired valid and efficient parameter estimates. Statistic procedures to properly account for missing data are well-studied and readily available, yet many scientists and practitioners continue to employ inappropriate techniques (e.g., listwise or pairwise deletion). Oftentimes, these inappropriate techniques lead to invalid and inefficient estimates. In this workshop, I will give an overview of the missing data problem and present the assumptions that are made when missing data are present. I will discuss the two modern approaches to missing data - Bayesian multiple imputation and full information maximum likelihood. Last, I will demonstrate how these techniques are carried out using readily available statistical software.

Bio:
Kimberly Henry is an Associate Professor at Colorado State University. She holds a joint appointment in the Psychology Department and the Colorado School of Public Health. She received her Ph.D. in Biobehavioral Health from the Pennsylvania State University in 2002. As a behavioral scientist oriented toward prevention, her work focuses on the psychological and social factors that produce or mitigate the health risk behaviors of adolescents and young adults. Her goals are to develop and test theoretical models in order to understand the complex interactions of risk, promotive, and protective factors that influence these risk behaviors and, ultimately, to create and test methods for prevention. Over the past ten years, her research studies have focused on four interrelated topics: 1.) school engagement as a promotive and protective factor in positive youth development, 2.) prevention of child maltreatment and exposure to family violence, 3.) development, implementation, and evaluation of community-based interventions, and 4.) innovative application of statistical models in prevention science. Dr. Henry has published over 50 peer reviewed papers and chapters and she has been continuously funded by federal agencies (National Institutes of Health and the Centers for Disease Control and Prevention) for the past 10 years.

2-3 learning objectives for the session
1. Know the assumptions that one makes when missing data are present
2. Be familiar with the two modern methods to properly handle missing data - Bayesian multiple imputation and full information maximum likelihood
3. Discover how multiple imputation and full information maximum likelihood are implemented in practice