

MISSING DATA: beyond star trek, the next generation

Abstract

Missing data is ubiquitous. Arguably every dataset from Biostatistics to Banking suffers from missingness, censorship, rounding, coarsening, survivorship bias, non-response bias, or a failure to simultaneously observe all relevant values. This workshop will outline whether and how one should address these incomplete data problems. We begin with an evaluation of the complete case analysis. In this common practice, cases with any missing variables are simply deleted, which may lead to inconsistent estimates. We discuss various ways of dealing with this inconsistency. We examine techniques which adjust for missing data, including reweighting the observed data, and imputing the missing observations under a frequentist and Bayesian paradigm. We will address questions such as the following:

- When do we need to construct a model for missingness?
- Does it matter how we impute missing values? For example, can we just use the last value observed?
- To what extent does our analysis depend on the assumptions made about the missing data mechanism? How sensitive are the results to the model assumptions?
- What are the ramifications for design? When a cheap surrogate variable for an expensive covariate is easily obtained, how can we use it in designing two-stage studies?
- Do we need to adjust for survivorship biases?
- What software is available for the analysis of missing data, and what assumptions are made?

Biography



Dr. Don L. McLeish is Adjunct Professor of Statistics and Actuarial Science at the University of Waterloo. He has taught at universities in Canada and internationally since 1973. He has acted as President of the S.S.C. and Director of the Center for Advanced Studies in Finance at the University of Waterloo. In 2007 he was awarded the Statistical Society of Canada Gold Medal for contributions to the theory of martingales, estimating functions, Monte Carlo Methods and quantitative finance. He is the author of three books, one on Monte Carlo methods in finance and a host of refereed publications; and has supervised of over 50 graduate students. He has an abiding interest in the proper treatment of missing or censored data in statistics, biostatistics and finance. He has co- taught a graduate course at the University of Waterloo on the treatment of incomplete data together with

Dr. Richard Cook, and co-organized a very successful Fields Institute 2004 workshop on missing data.



Dr. Cynthia A. Struthers is Associate Professor of Statistics and Actuarial Science at the University of Waterloo where she is currently teaching fellow in the Faculty of Mathematics at the University of Waterloo. She combines expertise in teaching methodology and biostatistics and has coauthored several papers on Missing data with D. McLeish. Her interests are in the communication of statistical ideas, and she has taught undergraduate and graduate courses in probability and statistics at all levels, including a graduate course at the University of Waterloo on the treatment of incomplete data.

Schedule on May 25, 2015

08:30-09:15	Registration & Breakfast
09:15-09:30	Welcoming Remarks <i>Wendy Lou, Janet McDougall, Xin Gao, Ruth Croxford</i>
	Introduction <i>Georges Monette</i>
09:30-10:45	INCOMPLETE DATA: A (fairly) complete review of basic practice Don McLeish and Cynthia Struthers
10:45-11:00	Coffee Break
11:00-12:30	COMPLETING THE INCOMPLETE: Estimating functions, likelihood, Imputation and Bayes Don McLeish and Cynthia Struthers
12:30-14:00	Networking Lunch Career-Advice Panel Ruth Croxford, Arthur Tabachneck, Marguerite Ennis, Jacky Bai, Peter Mercurio
13:00-14:00	Poster Presentations
14:00-15:15	Example: LONGITUDINAL DATA AND COMPUTATION Don McLeish and Cynthia Struthers
15:15-15:30	Coffee Break
15:30-16:30	SURVIVORSHIP AND MISSING DATA IN FINANCE Don McLeish
16:30-17:00	SENSITIVITY AND SOFTWARE Don McLeish and Cynthia Struthers
17:00-17:15	Awards Presentation & Closing Remarks Wendy Lou, Paul Corey, Georges Monette
17:15-18:00	SORA Annual General Meeting Xin Gao, Hugh McCague
18:00-20:00	Networking Dinner (optional, pay for your own)