



JOHNS HOPKINS  
BLOOMBERG  
SCHOOL of PUBLIC HEALTH

Department of Biostatistics

## BIOSTATISTICS SEMINAR

# Flexible Modeling of Incomplete Multivariate Data for Inference, Imputation and Prediction

**Joseph L. Schafer**

Senior Mathematical Statistician for Analytic Modeling  
Office of the Associate Director for Research and Methodology  
U.S. Census Bureau

### Abstract:

For modeling incomplete data or imputing missing items, the multivariate normal distribution is computationally convenient. But standard formulations of the model are so restrictive that the model has often been abandoned in favor of alternatives (e.g. sequences of conditional univariate regressions) not based on *bonafide* joint model for the variables in question. In this talk, I demonstrate how the multivariate normal model can be extended to handle (a) mixtures of continuous, categorical and semicontinuous responses, (b) flexible regression functions for means and variances with complex combinations of fixed and random effects; (c) enforced conditional independencies among responses where desired, and even (d) skewed and/or heavy-tailed residual errors. I describe Bayesian computational strategies for inference, imputation and out-of-sample prediction, and discuss challenges of creating a user interface to make the software accessible to data analysts. The primary motivating example is an ongoing effort to model unit-level time series from the Census Bureau's monthly business surveys, which feature highly skewed positive values with intermittent zeroes, high rates of missing values, long-term and annual periodic trends that vary by industry and unit. which features highly skewed positive values with intermittent zeroes, high rates of missing values, long-term and annual periodic trends that vary by industry and unit.

Johns Hopkins Bloomberg School of Public Health, Department of Biostatistics  
Monday, September 17, 2018, 12:15:1-15, Room W2008 (Refreshments 12:00pm)

For disability access information or listening devices, please contact the Office of Support Services at 410-955-1197 or on the web [www.jhsph.edu/SupportServices](http://www.jhsph.edu/SupportServices). EO/AA

Dept. of Biostatistics, 615 N. Wolfe Street, Suite E3527 Baltimore, MD 21205