Generalized Linear Latent and Mixed Models (GLLAMMs) for Complex Survey, Biometric, and Educational Data

Course Syllabus

Instructor: Stephen Schilling - (734-355-9801) (schillsg@umich.edu) , Institute for Social Research – Survey Research Center, School of Medicine, School of Education, The University of Michigan

Recommended Text: Rabe-Hesketh, S. & Skrondal, A. (2012). *Multilevel and Longitudinal Modeling using STATA, Volumes 1 & 2,* College Station TX: Sage: StataCorp LP.

Course Description: The last 50 years have seen development and use multilevel and mixed models, latent and structural equation models, generalized linear models, generalized linear mixed models, item response theory (IRT) models, and longitudinal models across a wide variety of disciplines. Statisticians often note the overlap between these methods but lacked a unifying approach towards estimation, testing, and application. Recently Rabe-Hesketh and colleagues have developed a unified approach, Generalized Linear Latent and Mixed Models. GLLAMMs allow estimation of multilevel models for binary, ordinal, and count data that include structural equation relationships among latent variables underlying observed data. Diverse applications include 1) multilevel models of complex survey data with multistage sampling, unequal sampling probabilities, and stratification; 2) explanatory IRT Models: 2) modeling endogenous switching and sample selection models for binary, count, and ordinal data; 4) generalized linear models with covariate measurement error; 5) biometrical modeling of twin and family data; 6) multivariate methods for meta-analysis of genetic association studies; 7) multilevel models for discrete choice and rankings; 8) conjoint choice models of consumer preference; 9) differential item functioning and test bias; and 10) enhancing the validity and cross cultural comparability of survey research

This class will consist of daily morning lecture and afternoon lab. Lectures begin with an introduction to GLAMMS, focusing on their structure and estimation and then move on to specific applications, concentrating on modeling of complex survey data, biometric data, and educational data. Worked examples will be performed in the labs using the GLLAMM software available for free with use with STATA. We will also demonstrate Bayesian estimation of GLLAMM models using STAN within R.

Pre-requisites: One or more courses in statistical methods that include regression analysis and/or structural equation models. Familiarity with STATA or R would be helpful, but is not essential.

Tentative Course Schedule

A Brief Introduction to Multilevel and IRT Models

Lecture 1: Monday, June 25, 9:00 – 10:25. Multilevel Model Basics – Model derivation and assumptions: RHS Chapters 1-4
Lecture 2: Monday, June 25, 10:35 – 12:00. IRT Models – The 1 and 2 parameter IRT model. Class Handout
Lab 1: Monday, June 25, 1:00 – 4:00. Introduction to STATA and GLLAMM software

GLLAMM Model Basics

Lecture 3: Tuesday, June 26, 9:00 – 10:25. Multilevel Models for Binary Data and IRT Models as Multilevel models. Assessment of Model/Data Fit: RHS Chapter 10.
Lecture 4: Tuesday, June 26, 10:25 – 12:00. GLLAMM Models: Estimation and Testing RHS Chapter 10 section 11. Rabe-Hesketh, S., Skrondal, A. and Pickles, A. (2005).
Lab 2: Tuesday, June 26, 1:00 – 4:00. Fitting Multilevel IRT Models using GLLAMM Software: Class Handout.

GLLAMM Models Extensions: Multilevel Ordinal Models and Multilevel IRT Models

Lecture 5: Wednesday, June 27, 9:00 – 10:25. Multilevel Ordinal Models: RHS Chapter 11 and Class Handout
Lecture 6: Wednesday, June 27, 10:35 – 12:00. Multilevel IRT Models: Class Handout
Lab 3: Wednesday, June 27, 1:00 – 4:00. Fitting Ordinal and Multilevel IRT Models: Class Handout

Multidimensional and Latent Class Models

Lecture 7: Thursday Jun 28, 9:00 – 10:25. Structural Equation Models and Multidimensional IRT Models: RHS Chapter 11 and Class Handout
Lecture 8: Thursday Jun 28, 10:35 – 12:00. Latent Class Models: Class Handout
Lab 4: Thursday Jun 28, 1:00 – 4:00. Bayesian Methods – An Introduction to STAN.

Models for Count and Survival Data

Lecture 9: Friday, Jun 29, 9:00 – 10:25. Multilevel Count Data: RHS Chapter 13.
Lecture 10: Friday, Jun 29, 10:35 – 12:00. Survival Models: RHS Chapters 14-15.
Lab 4: Friday, June 29, 1:00 – 4:00. Fitting Multilevel Models in Stan: Class Handout

Reference Books

Skrondal, A. and Rabe-Hesketh, S. (2004). Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models. Boca Raton, FL: Chapman & Hall/CRC

Embretson, S. E., & Reise, S. P. (2000). *Item Response Theory for Psychologists.* Psychology press.

Reference Papers

Rabe-Hesketh, S., Skrondal, A. and Pickles, A. (2005). Maximum likelihood estimation of limited and discrete dependent variable models with nested random effects. *Journal of Econometrics* **128** (2), 301-323. Local

Rabe-Hesketh, S., Skrondal, A. and Pickles, A. (2004). Generalized multilevel structural equation modelling. *Psychometrika* **69** (2), 167-190.Local

Rabe-Hesketh, S., Skrondal, A. and Pickles, A. (2002). Reliable estimation of generalized linear mixed models using adaptive quadrature. *The Stata Journal* **2** (1), 1-21