Introduction to Multi-Level Modeling

CLASS SESSIONS
Monday, June 11, – Friday, June 15, 2018
Time: 1:30 - 5:30pm
Location: TBD

INSTRUCTOR:
Katherine M. Keyes, PhD MPH
Associate Professor of Epidemiology
Columbia University
722 West 168th Street #724
(212) 304-5652
Kmk2104@columbia.edu

COURSE DESCRIPTION
This course will provide an overview of the theory underlying the use of multi-level models, using a parametric framework, as well as teach the basic application of methods necessary to conduct and interpret multi-level analyses of epidemiologic data.

After reviewing basic statistical and theoretical principles of linear and logistic regression models using only one level of organization, lecture and lab sessions will be focused on the use of random-effects models and generalized estimating equation (GEE) models for the analysis of data with two levels. Hands-on exercises will use data from an investigation of the influence of NYC neighborhoods on obesity, focusing on the application and interpretation of regression models that account for clustered observations and group-level covariates. SAS code will be provided for all in-class exercises and participants will learn how to implement multi-level models using both SAS Proc MIXED and SAS Proc GLIMMIX. R and STATA code will also be provided.

PREREQUISITES
Good understanding of linear and logistic regression analysis and interpretation. Prior exposure to SAS is recommended but not required.

COURSE LEARNING OBJECTIVES
By the end of the course, participants will be able to
• Explain the biostatistical and epidemiologic theories behind using multiple levels of organization to answer population health research questions
• Critically review studies that have employed multi-level analytic techniques
• Compare and contrast random-effects models with GEE models
• Select an analytical strategy appropriate to their research question
- Conduct multi-level analysis and interpret statistical output using SAS

**COURSE READINGS**

All of these articles should be read before the first day of class to help give you a good background of the issues we will be covering during the class:


4. Black JL, Macinko J, Beth Dixon L, Fryer GE. Neighborhoods and obesity in New York City. Health & Place 2010; 16: 489-499. *(Example using the Community Health Survey, which is the source for our practice dataset)*


**COURSE STRUCTURE**

Class meets from 1:30 PM – 5:00 PM Monday – Friday for a week (20 hours total). Each day we will split time between lecture and hands-on exercises.

**COURSE SCHEDULE**

| Session 1 – Introduction to Multi-Level Models: Concepts & Data Structure |
|---|---|
| **6-11-18** | **Learning Objectives:** |
|  | Explain the basic theory behind multi-level models |
|  | Identify the general distinctions between different types of multi-level models |
|  | Introduce basic GIS concepts & methods |
|  | Discuss potential issues that arise when defining and measuring a group level variable |
|  | Begin exploration of the dataset to be used for the in class exercises |

| Session 2 – Introduction to General Linear Mixed Models |
|---|---|
| **6-12-18** | **Learning Objectives:** |
|  | Describe the distinction between OLS and a basic General Linear Mixed Model |
|  | Explain the equations of General Linear Mixed Models |
|  | Present the Intraclass correlation coefficient (ICC) |
|  | Introduce Proc MIXED |
|  | Discuss random intercept models with level-1 and level-2 predictors |
### Session 3 – Graphing Linear Mixed Models & Testing Model Assumptions

**6-13-18**  
**Learning Objectives:**  
- Review random intercept models  
- Introduce random slope models  
- Explore Random intercepts and random slopes with graphs  
- Model Building and Testing assumptions

### Session 4 – Generalized Linear Mixed Models, Cross-level interactions & Sample Size

**6-14-18**  
**Learning Objectives:**  
- Examine cross-level interactions  
- Critically evaluate a mixed model article  
- Explain mixed models with binary outcomes  
- Introduce Proc GLIMMIX  
- Understand sample size & power in a multi-level study

### Session 5 – Sample Size, Generalized Estimating Equations (GEE) & more complex models

**6-15-18**  
**Learning Objectives:**  
- Explain Generalized Estimating Equations  
- Compare GEE to Mixed models  
- Discuss options in Proc GLIMMIX  
- Introduce 3 level and Cross-classified models