EPSY 584 - Hierarchical Linear Models

4 credit hours (CRN: 43707)

Semester: Fall 2020 Time: Tuesdays 5:00 - 8:00pm Classroom: Blackboard Collaborate Lab room: ETMSW 2027 (1040 W. Harrison St.)	Professor: George Karabatsos (home page) Phone: 312-413-1816 E-mail: georgek@uic.edu (most reachable) Office Hours: Tuesday 2-4pm (EPASW 1034)

Course Description:

This course introduces students to Hierarchical Linear Models, which are mixture models that are widely applied in education, psychology, medicine, and other fields. Sometimes, they are referred to as random-effects models. But more generally, a hierarchical model not only treats the dependent variable observations as random, but also treats model parameters (e.g., regression coefficients, error variance parameter) as random variables, that follow some distribution (e.g., normal, or gamma, etc.). In a Bayesian statistical framework, all model parameters are treated as random variables, arising from a prior distribution.

A Hierarchical Linear Model (HLM) can be viewed as having a nested structure, in that the model allows regression coefficients to vary from one context to another. For example, in educational research, a HLM is often used to analyze data about student math achievement. Here, students are nested within schools, and the model permits the investigation of the relationship between student socioeconomic status and math achievement, by school, and allows the investigation of school-level factors that affect this relationship. To give another example, for longitudinal data analysis, the HLM provides an approach to learn the growth curve of each individual subject, and also provide a way to identify the subject-level predictor variables that significantly predict changes in the subjects' growth curves.

The HLM provides a single flexible framework for statistical modeling that applies to many important tasks of data analysis, including:

(1) ***analysis of variance* (*ANOVA*)**, ***analysis of covariance* (*ANCOVA*)**,
(2) ***random-coefficients regression analysis,***
(3) ***categorical data analysis (rating data, or classification involving unordered categories)***,
(4) ***longitudinal (repeated-measures) analysis***,
(5) ***meta-analysis***,
(6) ***causal inference in nonrandomized studies,***
(7) ***spatial regression,***
(8) ***the analysis of censored data (e.g., for survival analysis),***
(9) ***psychometric analysis with predictor variables (e.g., Rasch models, random item IRT)***.

This course will present various Hierarchical Linear Models from both a Bayesian and a frequentist perspective of statistical inference. Moreover, this course will also present nonparametric Hierarchical Linear Models, which provide a way to circumvent the empirically parametric assumptions of "off-the-shelf" versions of Hierarchical Linear Models. These assumptions include the normality of error distribution, the normality of random effects, and the (inverse-)link function being defined by the standard logistic distribution. All Hierarchical Linear Models discussed in the course will be illustrated on data sets arising from education, psychology, medicine, and other fields.
Through in-class exercises (which count as credit towards three open-notes exams), students will learn how to perform data analysis with Hierarchical Linear Models using the HLM software, and also using the free R software with package nlme.
Also, I will provide free menu-driven software that I have developed, which can be used to perform data analysis using Bayesian hierarchical models.
They include 2-level and 3-level mixture models, with mixing distribution modeled either normally, or modeled nonparametrically as an infinite mixture of distributions.
The software, along with the user's manual, can be downloaded from the webpage: http://tigger.uic.edu/~georgek/HomePage/BayesSoftware.html

Students will also give a 25 minute presentation about a practical implementation of a hierarchical linear model on data. Students are encouraged to present an analysis of their own data set. Also, we may use the Bayesian Regression and SPSS software for data management, and for the imputation of missing data. Students can use other software, such as SAS and Stata.

If you need to find a data set, rich data sets can be found online, on world-wide assessments of educational progress.
Data sets include the PIRLS data set (literacy assessment), the TIMSS data set (for math assessment),
and the PISA data sets (for either Math, science, or literacy assessments), among others:
PIRLS or TIMSS data: http://timssandpirls.bc.edu/
PISA data: http://www.oecd.org/pisa/pisaproducts/
Open Psychology Data https://openpsychometrics.org/_rawdata/
Illinois School Report Card data https://www.isbe.net/ilreportcarddata
EPA data https://www.epa.gov/outdoor-air-quality-data
Machine Learning Data https://archive.ics.uci.edu/ml/datasets.html
Machine Learning Data http://mldata.org/

Course Prerequisites: A previous course covering multiple regression, or equivalents.

Textbook:
Raudenbush, S., & Bryk, A.S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks, CA: Sage. ISBN 0-7619-1904-X.

Suggested (optional) Readings:
Optional Article readings:
De Iorio, M., Müller, P., Rosner, G.L., and MacEachern. S.N. (2004). An ANOVA Model for Dependent Random Measures. *Journal of the American Statistical Association*, *99*, 205-215.
Imbens, G.W., and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, *142*, 615-635.
Karabatsos, G. (2020). Fast search and estimation of Bayesian nonparametric mixture models using a Classification Annealing EM Algorithm. *Journal of Computational and Graphical Statistics*.
Karabatsos, G., and Walker, S.G. (2012). Adaptive-Modal Bayesian Nonparametric Regression. *Electronic Journal of Statistics*, *6*, 2038-2068.
Karabatsos, G., and Walker, S.G. (2012). A Bayesian nonparametric causal model. *Journal of Statistical Planning and Inference*, *142*, 925-934.
Karabatsos, G., and Walker, S.G. (2012). Bayesian nonparametric mixed random utility models. *Computational Statistics & Data Analysis*, *56,* 1714-1722.
Karabatsos, G., Talbott, E., and Walker, S.G. (2015). A Bayesian nonparametric meta-analysis model. *Research Synthesis Methods*, *6,* 28-44.
Karabatsos, G. (2017, in press). Marginal Maximum Likelihood Estimation Methods for the Tuning Parameters of Ridge, Power Ridge, and Generalized Ridge Regression. *Communication in Statistics: Simulation and Computation*.
Kleinman K., & Ibrahim J.G. (1998) A semi-parametric Bayesian approach to the random effects model. *Biometrics*, *54*, 921-938.

Kleinman, K.P. & Ibrahim, J.G. (1998). A semiparametric Bayesian approach to generalized linear mixed models. *Statistics In Medicine*, *17*, 2579-2596.
Robins, J.M., Hernan, M.A., & Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*, *11*, 550-560.
Savitz, N.V., & Raudenbush, S.W. (2009). Exploiting spatial dependence to improve measurement of neighborhood social processes. *Sociological Methodology*, *39*, 151-183.
Stuart, E.A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science*, *25*, 1-21.

Optional book readings:
Cooper, H., Hedges, L.V., & Valentine, J.C. (2009). *The Handbook of Research Synthesis and Meta Analysis*. New York: Russell Sage.
Fox, J. (2008). *A Mathematical Primer for Social Statistics*. Thousand Oaks, CA: Sage.
Gelman, A., and Hill, J. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge.
Raudenbush, S., Bryk, A., Cheong, Y.-F., & Congdon, R. (2004). *HLM 6: Hierarchical Linear and Nonlinear modeling*. Lincolnwood, IL: Scientific Software International. Available as pdf in the course.
Rosenthal, R. (1994). *Parametric measures of effect size*. In H. Cooper & L. V. Hedges (Eds.), *The Handbook of Research Rynthesis* (pp. 231-244). New York: Russell Sage.
Ward, M.D. & Gleditsch, K.S. (2008). *Spatial Regression Models*. Los Angeles: Sage.
Also, I will provide notes on various related topics.

Journal Publications of students who have taken this HLM course in the past:

Arenson, E., and Karabatsos, G. (2018). A Bayesian beta-mixture model for nonparametric IRT (BBM-IRT). *Journal of Modern Applied Statistical Methods*, 1.
Brow, M. (2018). Significant predictors of mathematical literacy for top-tiered countries/economies, Canada, and the United States on PISA 2012: Case for the sparse regression model. *British Journal of Educational Psychology*.
Fujimoto, K., and Karabatsos, G. (2014). Dependent Dirichlet Process Rating Model (DDP-RM). *Applied Psychological Measurement*, *38*, 217-228.
Kaminski-Öztürk, N., and Karabatsos, G. (2017). A Bayesian robust IRT outlier detection model. *Applied Psychological Measurement*, *41*, 195-208.
Muckle, T., and Karabatsos, G. (2009). Hierarchical generalized linear models for the analysis of judge ratings. *Journal of Educational Measurement*, *46*, 198-219.
Talbott, E., Zurheide, J. L., Karabatsos, G., & Kumm, S. (2020). Similarity in Teacher Ratings of the Externalizing Behavior of Twins: A Meta-Analysis. *Behavioral Disorders*.
Tang, X., Karabatsos, G., & Chen, H. (2020). Detecting local dependence: A Threshold-Autoregressive Item Response Theory (TAR-IRT) approach for polytomous items. *Applied Measurement in Education*.

COURSE SCHEDULE

Date	Topic	Read
Aug25	**Assignments/tasks. Introduction and motivation for Hierarchical Linear Models.**	
Sep1	**Hierarchical Linear Models:** HLMs for continuous outcomes	Ch. 1-2
Sep8	**HLM for one-way random-effects ANOVA, ANCOVA, ordinary linear regression, Random Coefficients regression,** **HLM for Means-as-outcomes model, model for non-random varying slopes, and Full HLM**	Ch. 3-5
15	**Hierarchical Linear Models:** Estimation and model fit assessment. Frequentist and Bayesian approaches.	Ch.9, 13-14
22	**Bayesian Hierarchical Linear Models:** HLMs for continuous outcomes	Ch. 13
Sep29	**HLM for Repeated Measures and longitudinal analysis. 3-level HLM.** ***Exam #1 is due.***	Ch. 6,8
Oct6	**Imputation of missing data**. **Power analysis.**	
13	**Generalized HLM for binary, binomial, and counts as outcomes;** **Spatial regression.** **Analysis of Censored data.**	Ch. 10-12
20	**Generalized HLM and nonparametric HLM,** **for discrete outcomes (dependent) variables.**	Ch. 10
27	**HLM for Meta-Analysis.**	Ch. 7
Nov3	**Election day (no class)**	
10	**HLM for Causal Analysis in nonrandomized studies: Regression discontinuity methods.** ***Exam #2 Is Due*** **(along with statement of data set for Final Presentation)**	
17	**HLM and Semiparametric HLM for Psychometric Analysis**	
24	Student Presentations	
Dec1	Student Presentations	
Dec 7 -11	***FINAL EXAM WEEK: Final Exam due before Wednesday Midnight***	

Grading Policy:
The three take-home exams, together, are worth 60% of the final grade, and the Final Presentation is worth 30% of the final grade.
Class participation is worth 10%, including class attendance, and honest class participation. Points will be deducted for lack of attendance or honest class participation.
Each exam involves addressing data analysis tasks using HLM methods described in the course. The second exam will require a written one-paragraph description of the data set that will be used for the Final Presentation.
Final grades will be given out according to the following scale:

A	90% - 100%
B	79% - 89%
C	68% - 78%
D	57% - 67%
F	56% - Lower

Borderline grades will be decided on the basis of class participation.
Each assignment submitted late receive a 20% grade reduction for each week after the original due date.
Students will spend substantial amounts of time reading, and on the computer. It is assumed that students will exert individual initiative in solving computing/analysis problems as they arise.
(Standard policy: There are no exceptions to the above grading scale, and no extra credit work will be accepted. Incomplete grades will be considered only for students with extenuating circumstances. Poor performance on assignments will not be considered in a request for an incomplete).

Data Analyses Presentation:
The presentation, which should be 25 minutes in length (about 15 Power-Point slides; no more), will deal with an application of a parametric or semiparametric HLM on a real data set. The presentation should (at least) include:

INTRODUCTION
Describe the substantive research questions or open problems that your study will address (10 points).

METHODS
-- Describe sample characteristics. (5 points)
-- Fully describe the HLM model you will use to answer your research questions
(using words and mathematical notation), and include a discussion of the assumptions of your model. (10 points)
-- Describe the parameters will you interpret to answer your research questions. (10 points)
-- Use the appropriate coding for all the predictor variables in your model. (10 points)
-- Use the model that is appropriate for the type of dependent variable you will analyze. (15 points)

RESULTS - Accurately describe all the relevant results of your HLM model, including all significant and non-significant effects at Level 1 and Level 2 of the model. (25 points)

DISCUSSION - What are the implications of the results of your study, and potential directions for future research?
(10 points).

Disability Services:
UIC strives to ensure the accessibility of programs, classes, and services to students with disabilities. Reasonable accommodations can be arranged for students with various types of disabilities, such as documented learning disabilities, vision, or hearing impairments, and emotional or physical disabilities. If you need accommodations for this class, please let your instructor know your needs and he/she will help you obtain the assistance you need in conjunction with the Office of Disability Services (1190 SSB, 413-2183).