

1/3/2022

Hierarchical Linear Models (Multilevel Level Models)
Edpsy/Psych/Stat 587
Spring 2023

Instructor: Carolyn Anderson
Office: 236C Education Bldg.
office hours: Monday 1:00 – 3:00pm
e-mail: cja@illinois.edu

TAs: Lizzy Wu
Office: rm 230 Education Bldg.
office hours: Tuesday 1:00—3:00pm
e-mail: lixinwu2@illinois.edu

Selim Havan
Office: rm 230 Education Bldg.
Office hours: Thursday 10:00—12
e-mail: shavan2@illinois.edu

Lecture: Mon/Wed 10:00-11:50 am. This is an in-person course.

Office Hours

We will use google sheets for you to sign up for specific time slots with Carolyn , Lizz, and/or Selim. If you are registered for the course, you will be given permission to edit the office hours sheet. I will be adding your email addresses (i.e., netid@illinois.edu), and once added you may need to request access to edit. General policies on office hours:

- You must sign up for office hours.
- On the sign up sheet, indicate whether you want in person or zoom.
- Do not sign up for more than 30 minutes.
- If you sign up for a time slot, but decide not to keep it, please remove your name in case someone else would like to use the time.
- Sign up at least 2 hours prior to office hours. We may not see late sign ups.
- Do not add hours to those listed on sign up sheet.

Prerequisites: Edpsy 581 and 582 or Psych 406 and 407 are required. Edpsy/Soc 584 or Psych 594 useful but not required.

Course web-site: <https://cja.education.illinois.edu/ed-psy-587> Besides course information and materials, there are links to multilevel web-sites. I will up-date this through out the semester. Please notify me (Carolyn) if any of the links are not working. I will up-date this through out the semester.

Questions and feedback:

You can send e-mail to Edpsy587@gmail.com, cja@illinois.edu, lixinwu2@illinois.edu, or shavan2@illinois.edu. If you use one of the latter two, please include the “Edpsy 587” in the subject header.

Course objectives: Independence of observations is a key assumption in many statistical inferential procedures, but if this assumption is not met, then inference is not valid. Data in education, psychology, medicine, public health, sociology and other applied sciences are often clustered or show a multilevel or hierarchical structure where observations within groups are dependent or correlated. For example, in educational research, researchers frequently select a sample of schools from which a sample of classes is selected from which a sample of students is selected. Measurements of attributes or characteristics of schools, classes (teachers) and students may be available. Longitudinal and repeated measures are also hierarchical or multilevel data. Most standard statistical models and tests critically rely on the assumption of independent observations. When data are clustered or show a multilevel structure, observations are typically correlated, which violates the standard independence assumption and invalidates conclusions based on standard statistical methods.

This course provides an introduction to hierarchical or more generally multilevel models that take into account dependencies between observations. Students will learn the basic ideas and theory of hierarchical linear models, as well as have many opportunities to apply the methods to real data from studies in education, psychology and social sciences. Topics that will be covered include an introduction to multilevel analyses, random intercept and slope models, 2 and 3 level models, hypothesis testing, model assessment, longitudinal data, and generalized hierarchical models for dichotomous response/dependent variables.

Texts:

Required: Snijders, T. & Bosker, R. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*, 2nd Edition. Thousand Oaks, CA: Sage.

Additional readings are available online either on course web-site or from the Uofl online collection.

Computing: About 4 lectures will be held using software where you will be doing lab exercises. For these “lectures”, you will need a computer.

Only R will be covered in labs with more emphasis will be placed on R in lecture; however, you can use STATA or SAS if you desire. Note that both R and SAS are free. There is both SAS and R scripts on the course web-site.

For those who wish to use SAS® (version 9.4), it is available free from [SAS.com](https://www.sas.com) under their education program. You will not get SAS graph, but all other procedures we will use in the course are included. If you wish to purchase SAS, you can get a license and media from webstore.illinois.edu. The fee for a full year is \$70 last time I checked, and you can borrow the media to install the program. For current prices and options see <https://webstore.illinois.edu/Shop/search.aspx?keyword=SAS>.

R can be downloaded for free from <https://cran.r-project.org/mirrors.html> and for most information see <https://www.r-project.org>. You should download the most recent version and you will need to reinstall all you packages (you can get a list by using `library()` and then save this). Rstudio can be very useful. You should have a basic familiarity with R.

You may use other programs, but you will be responsible for learning how to use these programs to do multilevel analysis and interpret the results. SPSS is not recommended or accepted.

Evaluation: Students will be evaluated on the basis of homework assignments, mini-projects (computer lab work), class participation (e.g., attending class is a requirement), and a final take home exam **or** project. 60% of grade is based on homework and 40% on final/project. Homework is due on the stated due date unless you get prior permission from the instructor. The project or final is **due Friday May 5th**. Late projects/finals will be docked 10 points (out of 100) per day late.

Homework and final/project should be submitted to edpsy587@gmail.com.

Course Project: Students may do a course project in lieu of a final exam. The final is basically a project but I provide the data and questions, I will give you 5 options. The instructor must pre-approve course projects and I strongly suggest getting approval earlier rather than later. The range of possible projects is very broad and could include papers describing analysis of data as part of a student's research or collaborative research, studying a topic not covered in the course (and applying it to data), or do a comparative study (what happens when you use ordinary regression versus HLM, or compare software packages). I assume that you have IRB approval. If you cannot complete the project by the end of the semester, you should do the final.

Collaboration: Students may work on assignments/project/final on their own OR work with a partner. If you choose to work on an assignment or project/final with another student, you should include both of your names on the assignment or project/final. You can change partners at any point in the semester. I expect

that both individuals to contribute equally and will received the same grade on the assignment.

Illness: If you are sick, do not come to lecture, the instructor's office hours or the TAs office hours. Face covering is recommended but not required.

The policy on mask wearing and other measures to mitigate the spread of COVID-19 will follow the university policy, which could change. See <https://covid19.illinois.edu/on-campus/on-campus-instructors>.

Fair Use/Plagiarism Policy: Please see go to the following link for policy on academic integrity: <http://education.illinois.edu/edpsy/about/academic-integrity> The definition as spelled out in this document is "The definition of plagiarism is straightforward: Presenting someone else's words, materials, manner of expression, or ideas as your own. This means that even if another person agrees to let you present his or her content as if it were yours, it is still plagiarism. Plagiarism does **not** require intent: it can be intentional or unintentional."

In this class almost every semester students turn in work that is not their own. This includes such things as

- One student does all the computer work and shares it with other students
- One student does all the typing of summary tables and shares the file.
- Students give answers to questions that are nearly identical (i.e., word for word) or the answers are nearly identical to answer keys from past offerings of the course.
- Students ask each other for help on final and/or homework – theses should be only your work. The only **exception** is if you work with a partner and acknowledge this.

These are all examples of plagiarism and will results in at a minimum a reduction in course grade and at a maximum a failing grade. If you work with a partner, both students' names need to be on their joint work.

I take this very seriously.

Emergencies: Review <http://police.illinois.edu/emergency-preparedness/>

In an emergency in this building, we'll have three choices: **RUN** (get out), **HIDE** (find a safe place to stay inside), or **FIGHT** (with anything available to increase our odds for survival).

First, take a few minutes this week and learn the different ways to leave this building (exits are to the North, South and two to the West). If there's ever a fire alarm or something like that, you'll know how to get out, and you'll be able to help others get out too.

Second, if there's severe weather and leaving isn't a good option, go to a low level, in the Education building the east side of the basement (away from windows).

If there's a security threat, such as an active shooter, **RUN** out of the building if we can do it safely or **HIDE** by finding a safe place where the threat cannot see us. We will lock or barricade the door and we will be as quiet as possible, which includes placing our cell phones on silent. We will not leave our area of safety until we receive an Illini-Alert that advises us it is safe to do so. If we cannot run out of the building safely or we cannot find a place to hide, we must be prepared to fight with anything we have available in order to survive.

Remember, RUN away or HIDE if you can, FIGHT if you have no other option.

Finally, if you sign up for emergency text messages at emergency.illinois.edu, you'll receive information from the police and administration during these types of situations.

If you have any questions, go to police.illinois.edu, or call [217-333-1216](tel:217-333-1216).

Tentative Course Schedule

Topic	Reading*
Introduction: Multilevel data	S&B: ch 1, ch2
Review of multiple regression and mixed effects ANOVA & multilevel data. (i.e., statistical treatment of clustered data)	S&B: ch 3
Random intercept model	S&B: ch 4
Random intercept and slope	S&B: ch 5 Singer, J.D. (1998). Using SAS PROC MIXED to fit multilevel models, Hierarchical Models and Individual Growth Models. <i>JEBS</i> , 23, 323-355. https://www.ida.liu.se/~732G34/info/singer.pdf or Illinois library online e-collection.
Estimation methods & problems Tips & Strategies for Mixed Modelling with SAS/STAT Procedures	S&B: p 60-61, 89-90 http://support.sas.com/resources/papers/proceedings12/332-2012.pdf
Bayesian Estimation of multilevel models	online document
Inference for fixed effects	S&B: ch 6
Inference for random effects	S&B: ch 6
Testing assumptions & model evaluation	S&B: ch 9 Leckie, et al. (2014). Modeling heterogeneous variance-covariance components in two-level models. <i>JEBS</i> , 39, 307-332. You can get this online from the Illinois Library e-collection.
Exploratory data analysis	GLMM Draft Chapter 7.1 on LMM (on course web-site)
Longitudinal data	S&B: ch 15

* S&B = Snijders & Bosker

1/3/2022

Discrete dependent variables (multilevel logistic regression) S&B: ch 17.1, 17.2 & 17.3

3--level models (in class and in context of logistic regression) S&B: p 67-71, 90-93

Dealing with skewed data and cross random effects (time permitting)

References

- Aerts, M, Geys, H, Moldenberghs, G., & Ryan, L.M. (2002). *Topics in Modelling of Clustered Data*. NY: Chapman & Hall/CRC.
- Aitkin, M, Anderson, D, & Hinde, J. (1981). Statistical modelling of data on teaching styles. *Journal of the Royal Statistical Society, A*, 144, 419-461.
- Aitkin, M. & Longford, N. (1986). Statistical modeling issues in school effectiveness studies. *Journal of the Royal Statistical Society, A*, 149, 1-43. (with discussion).
- Anderson, C.J, Johnson, T. & Verkuilen, J.V. (in preparation). *Applied Generalized Linear Mixed Models: Discrete and Continuous Data*.
- Arminger, G., Clogg, C.C., & Sobel, M.E. (1995). *Handbook of Statistical modeling for the social and behavioral sciences*. NY: Plenum Press.
- Cox, D.R., & Solomon, P.J. (2002). *Components of Variance*. Boca Raton, FL: Chapman & Hall.
- De Boeck, P. & Wilson, M. (2004). *Explanatory Item Response Models: A Generalized Linear and Nonlinear Approach*. NY: Springer.
- De Leeuw, J., & Meijer, E.. (2008). *Handbook of Multilevel Analysis*. Springer.
- Diggle, S.L., Heagerty, P., Liang, K.Y, & Zeger, S.L. (2002). *Analysis of longitudinal data, 2nd Edition*. Oxford, Great Britian: Oxford Science Publications.
- Gelman, A., & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. NY: Cambridge.
- Goldstein, H. (1995). *Multilevel statistical models, 2nd Edition*. London: Arnold.
- Goldstein, H. (1999). *Multilevel statistical models*, Internet version (<http://www.ioe.ac.uk/multilevel>).
- Heck, R.H., & Thomas, S.L. (2000). *An introduction of multilevel modeling techniques*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Hedeker, D. & Gibbons, R.D. (2006). *Longitudinal Data Analysis*. Wiley: Hoboken, NJ.
- Hox, J.J. (1995). *Applied multilevel analysis*. Amsterdam: TT-Publikaties. (out of print but via internet, check <http://www.ioe.ac.uk/multilevel>).

- Hox, J.J. (2002). *Multilevel Analysis: Techniques and Applications*. Mahwah: N.J.: Erlbaum.
- Hox, J.J. & J.K. Roberts (2011). *Handbook of Advanced Multilevel Analysis*. NY: Routledge (Taylor & Francis).
- Kreft, I., de Leeuw, J., & Aiken, L.S. (1995). The effect of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research*, 30, 1-12.
- Kreft, I. & de Leeuw, J (1998). *Introducing multilevel modeling*. Thousand Oaks: Sage.
- Little, R.C., Milliken, G.A., Stroup, W.W., & Wolfinger, R.D. (1996). *SAS system for mixed models*. Cary, NC: SAS Institute.
- Little, T.D., Schnabel, K.U, & Baumeter, J. (2000). *Modeling longitudinal and multilevel data*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Longford, N.T. (1993). *Random coefficients models*. NY: Oxford Science Publications.
- Longford, N.T. (1995). Random coefficient models. In G. Arminger, C.C. Clogg and M.E. Sobel (eds) *Handbook of Statistical modeling for the social and behavioral sciences*. pp 519-577. NY: Plenum Press.
- Luke, D.A. (2004). *Multilevel Modeling*. Thousand Oaks: Sage Publications.
- McCulloch, C.E., & Searle, S.R. (2001). *Generalized, Linear, and Mixed Models*. NY: Wiley.
- Molenburgs, G. & Verbeke, G. (2005). *Models for Discrete Longitudinal Data*. NY: Springer.
- Raudenbush, S.W. & Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods*, 2nd Edition. Thousand Oaks, CA: Sage.
- Raudenbush, S.W., Bryk, A.S., Cheong, Y.F, & Congdon, R (2004). *HLM6: Hierarchical linear and nonlinear modeling*. Lincolnwood, IL: Scientific software international.
- Recchia, A. (2010). R-squared measures for two-level hierarchical linear models using SAS. *Journal of Statistical Software*, 32, 1-9. <http://www.jstatsoft.org/v32/c02> (paper, MACRO, and example).

1/3/2022

Singer, J.D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics*, 23, 323-355.

(can download from her web-page at

<https://www.ida.liu.se/~732G34/info/singer.pdf>).

Singer, J.D. & Willett, J.B. (2003). *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. NY: Oxford University Press.

Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized Latent Variable Modeling: Multilevel, Longitudinal and Structural Equation Models*. NY: Chapman & Hall.

Snijders, T. & Bosker, R. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*, 2nd Edition. Thousand Oaks, CA: Sage.

Verbeke, G, & Moldenberghs, G. (editors). (1997). *Linear mixed models in practice*. NY: Springer-Verlag.

Verbeke, G, & Moldenberghs, G. (2000). *Linear mixed models for longitudinal data*. NY: Springer-Verlag.

There are many more references and numerous web-sites devoted to multilevel models, HLM, etc.