

Applied Linear Mixed Models in Behavioral Research

Course Information and Requirements, Spring 2007

January 30, 2007

Instructor: Yuelin Li, PhD.

e-mail: liy12@mskcc.org
Office hours: Mon (2 – 3) and Wed (1 – 2)

Class Meeting Time: Friday at 10 – 11:15 in the library, 641 Lexington Ave.

Course Objectives

We meet approximately every other week to discuss a few statistical issues related to designing a clinical trial in behavioral research. A randomized psychotherapy clinical trial may involve several hospitals (or “sites”). At each site there are several participating psychotherapists. Each therapist sees a few patients. The patients’ psychological wellbeing is monitored at baseline (before the psychotherapy) and after the therapy. Often several follow-up telephone assessments are made and outcome data collected. Within each site the patients are randomized with equal chance into either the intervention group or the standard care group. To make the two treatment groups comparable, patients may be *stratified* by age, sex, and type of cancer. This a simple “longitudinal study with a nested design”—patients are *nested* within therapists and therapists are nested within sites. Study designs like this are also called *hierarchical*, *multilevel*, or *clustered* designs.

Most behavioral scientists know these designs by the name Hierarchical Linear Models (HLM) and also as a statistical computer package of the same name (Raudenbush and Bryk (2002)). More generally, HLM is known as Linear Mixed Effects Models (LME, Laird and Ware (1982)) or simply Mixed Models.

We will begin with the basics and move gradually into more advanced topics. We will carry out most analyses with the `nlme()` library (Pinheiro & Bates, 2000) in an open-source statistical package called R. The basics will be based on the hierarchical models in HLM (e.g., the 2-level models in Raudenbush & Bryk, 2002, 2nd Ed). Soon you will realize that the hierarchical models in HLM is a special case of the general LME model. We will also use R to visualize the longitudinal outcomes and run Monte Carlo simulations to estimate the statistical power of a study with a nested design. We hope to move into more advanced topics such as Ecological Momentary Data Analysis (EMA) (Schwartz & Stone, 1998) and General Growth Mixture Modeling (Muthén et al. 2002). When we reach GMM, we will use Mplus to supplement R.

The goal is to help you absorb the critical abstract theoretical concepts without expending too much energy mastering the mechanics of statistical software packages. Do not worry about how much you have forgotten about introductory statistics. We will make adjustments as we go. It is important for you to ask for clarifications. Questions are always welcome.

Textbooks

- Required:
 1. Pinheiro and Bates (2000): covers R and `nlme()`.
 2. Brown and Prescott (2006): applied LME in medicine using SAS PROC NLME.
- Recommended:
 1. Raudenbush and Bryk (2002): applied HLM in behavioral sciences.
 2. Venables and Ripley (2002): general R.
 3. Fitzmaurice, Laird, and Ware (2004): a good reference on applied LME in medicine.

Readings

Readings from the text and journal articles are assigned for each class. Since the lectures will assume you know the information in the readings, it is important that these assignments be read before class. You are expected to be responsible for the contents of the required text.

Course Structure

	Date	Topic	Content	Reading materials
1	2/2/07	R Basics	- How to work with R	"An Introduction to R" on r-project.org; Baron & Li
		HLM Basics	- Review of longitudinal data analysis in a hierarchical nested design	Maxwell and Delaney (1990); Raudenbush and Bryk (2002): Ch1; Brown and Prescott (2006): Ch1 & 2
2	3/2/07	LME Basics	- "Fixed" and "random" effects explained using <code>xyp1ot()</code> in R - Hierarchical models in HLM is a special case of LME - <code>groupedData()</code> in R	Pinheiro and Bates (2000): Ch1 & 4 Fox (2002); http://idecide.mskcc.org/stats/multi/hlm.html Pinheiro and Bates (2000): Ch 3
3	3/16/07	LME 2	- Advanced topics such as within group correlation - Group-Randomized Trials - Correlation structure for modeling dependence	Brown and Prescott (2006): Ch5 & 6 Murray, Varnell, and Blitstein (2004); Murray (1998, Ch1 & 2) Pinheiro and Bates (2000): Ch 5
4	4/6/07	GLMM	- Mixed models with binary outcomes	Feng, Diehr, Peterson, and McLerran (2001); Neuhaus (2001)
5	4/13/07	EMA	- EMA model and data preparation - Repeated measures data	Schwartz and Stone (1998); Stone and Shiffman (1994); Shiffman (2005) Brown and Prescott (2006): Ch5 & 6
6	4/27/07	GEE	- Another method to analyze categorical data	http://www.gseis.ucla.edu/courses/ed231c/notes2/gee.html ; Hall et al. (2001); Lee, Herzog, Meade, Webb, and Brandon (2007); Shiffman et al. (2002)
7	5/11/07	Power	- How to calculate statistical power for LME	Brown and Prescott (2006): Sections 5.5.2; 6.7; Raudenbush and Liu (2000); Horton, Brown, and Quian (2004)
8	5/25/07	GMM	- Growth Mixture Models and Latent Class Analysis	Muthén et al. (2002); Muthén and Shedden (1999); McCulloch, Lin, Slate, and Turnbull (2002); Muthén et al. (2002)

Short Assignments

There will be short, written assignments. Each will consist of short-answer questions and other exercises that should require about 1 hour to complete if you understand the material. These will deal primarily with material

developed in the lectures and the readings. These exercises are intended primarily to help you solidify your understanding of the material.

Grades

You will not be graded. But I will try to provide feedback on the short assignments as a learning aide.

References

- Brown, H., & Prescott, R. (2006). *Applied mixed models in medicine*. New York: John Wiley & Sons.
- Feng, Z., Diehr, P., Peterson, A., & McLerran, D. (2001). Selected statistical issues in group randomized trials. *Annu Rev Public Health, 22*, 167-87.
- Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2004). *Applied longitudinal analysis*. Hoboken, NJ: John Wiley & Sons.
- Fox, J. (2002). *Linear mixed models*.
<http://cran.r-project.org/doc/contrib/Fox-Companion/appendix-mixed-models.pdf>; last assessed: January, 2007.
- Hall, S. M., Delucchi, K. L., Velicer, W. F., Kahler, C. W., Ranger-Moore, J., Hedeker, D., et al. (2001). Statistical analysis of randomized trials in tobacco treatment: longitudinal designs with dichotomous outcome. *Nicotine Tob Res, 3*(3), 193-202.
- Horton, J., N, Brown, E. R., & Quian, L. (2004). Use of r as a toolbox for mathematical statistics exploration. *The American Statistician, 58*, 343-357.
- Laird, N. M., & Ware, J. H. (1982). Random effects models for longitudinal data. *Biometrics, 38*, 963-974.
- Lee, J. H., Herzog, T. A., Meade, C. D., Webb, M. S., & Brandon, T. H. (2007). The use of gee for analyzing longitudinal binomial data: A primer using data from a tobacco intervention. *Addictive Behaviors, 32*, 1871-93.
- Maxwell, S., & Delaney, H. (1990). *Designing experiments and analyzing data*. Belmont, CA: Wadsworth Inc.
- McCulloch, C. E., Lin, H., Slate, E. H., & Turnbull, B. W. (2002). Discovering subpopulation structure with latent class mixed models. *Stat Med, 21*(3), 417-29.
- Murray, D. M. (1998). *Design and analysis of group-randomized trials*. New York: Oxford University Press.
- Murray, D. M., Varnell, S. P., & Blitstein, J. L. (2004). Design and analysis of group-randomized trials: a review of recent methodological developments. *Am J Public Health, 94*(3), 423-32.
- Muthén, B., Brown, C. H., Masyn, K., Jo, B., Khoo, S. T., Yang, C. C., et al. (2002). General growth mixture modeling for randomized preventive interventions. *Biostatistics, 3*(4), 459-75.
- Muthén, B., Brown, C. H., Masyn, K., Jo, B., Khoo, S. T., Yang, C. C., et al. (2002). General growth mixture modeling for randomized preventive interventions. *Biostatistics, 3*(4), 459-75.
- Muthén, B., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the em algorithm. *Biometrics, 55*(2), 463-9.
- Neuhaus, J. M. (2001). Assessing change with longitudinal and clustered binary data. *Annu Rev Public Health, 22*, 115-28.
- Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in s and s-plus*. New York: Springer-Verlag.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models* (2 ed.). Newbury Park, CA: Sage Publications, Inc.
- Raudenbush, S. W., & Liu, X. (2000). Statistical power and optimal design for multisite randomized trials. *Psychological Methods, 5*, 199-213.
- Schwartz, J. E., & Stone, A. A. (1998). Strategies for analyzing ecological momentary assessment data. *Health Psychol, 17*, 6-16.
- Shiffman, S. (2005). Dynamic influences on smoking relapse process. *J Pers, 73*(6), 1715-48.
- Shiffman, S., Gwaltney, C. J., Balabanis, M. H., Liu, K. S., Paty, J. A., Kassel, J. D., et al. (2002). Immediate antecedents of cigarette smoking: an analysis from ecological momentary assessment. *J Abnorm Psychol, 111*(4), 531-45.
- Stone, A. A., & Shiffman, S. (1994). Ecological momentary assessment (ema) in behavioral medicine. *Annals of Behavioral Medicine, 16*, 199-202.
- Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with s* (4th ed.). New York: Springer.