36-309/749
Experimental Design for Behavioral and Social Sciences

Dec 1, 2015
Lecture 11: Mixed Models (HLMs)

Independent Errors Assumption

- *An **error*** is the deviation of an individual observed outcome (DV) from the population mean of all observed outcomes that have the same levels of all of the explanatory variables (IVs). It represents variation unexplained by these IV(s).

- ***Correlation*** is between –1 and +1. [The correlation between two random variables is equal to the square root of R^2 from simple regression.] Zero is uncorrelated. Correlation is "unitless"; the corresponding quantity on the scale of the measurements is ***covariance***.

- The assumption of "independent errors" (which implies uncorrelated errors) comes down to the idea that knowing the error (or its estimate, the residual) for one measurement tells us nothing about the error for another measurement.

Indep. Errors Assumption, cont.

- We often do not have any tools to check this assumption from the data; rather we ***think*** about the likely nature of any correlation (e.g., repeated measures, collusion, or hierarchy).

- In repeated measures ANOVA and some other analyses that we have not studied, errors are ***modeled*** as correlated, so the assumption does not apply to these analyses (actually, it does apply between subjects but not within subjects). In between-subjects ANOVA, regression, logistic regression, and the chi-square test of independence, uncorrelated errors is a strong assumption. Fairly mild violations begin to alter the null sampling distributions of the test statistics resulting in incorrect p-values and confidence intervals and/or poor power. [But estimates of population means and slopes are still unbiased, i.e., on average they are correct over multiple experiments.]

General Linear Mixed Models

- General linear mixed models are best thought of as Normal linear models that flexibly model correlation often in the form of clustering or hierarchy. Complete or partial synonyms include hierarchical linear models (HLM), multilevel modeling, random regression models, and growth curve models.

- The term “mixed” comes from the two model components: fixed effects (retronym) and random effects (see below).

- [Replacing “general” by “generalized” allows non-normal outcomes (e.g., succeed vs. fail) using a more complex method. Replacing “linear” with “non-linear” allows expressions more complicated that $\beta_0+\beta_1 x_1+\ldots+\beta_k x_k$, but this is not commonly needed because the x’s or y’s can be transformed to produce a linear relationship on a different scale.]

HLM: An intuitive hierarchical approach

- Example: Repeated test scores over time are modeled as ***individual*** regression lines. $U_{\bullet i}$'s represent "personal" deviations of intercepts/slopes around population averages.
  - $Y_{iT} = A_{0i} + A_{1i}\,T + \varepsilon_{iT}$, for subject *i* at time T

    $\varepsilon_{iT} \sim N(0, \sigma^2)$ and uncorrelated
  - $A_{0i} = \beta_0 + \beta_{0A}\,(age_i - 40) + \beta_{0M}\,Male_i + \beta_{0I}\,(IQ_i - 100) + U_{0i}$

    $A_{1i} = \beta_1 + \beta_{1I}\,(IQ_i - 100) + U_{1i}$
  - The β's are called ***fixed effects***.
  - ***Random effects***: $U_{0i} \sim N(0, \tau_0^2)$, $U_{1i} \sim N(0, \tau_1^2)$, covariance$(U_{0i}, U_{1i}) = \tau_{01}$

HLM, cont.

- Both categorical and quantitative IVs are allowed. DVs are assumed to be Normally distributed quantitative variables [but discrete quantitative and categorical outcomes are allowed under "generalized linear mixed models"].

- The term "mixed model" derives from a mix of fixed and random effects:
  - Fixed effect factors have levels that would be the same in the next experiment, e.g., types of treatment or gender.

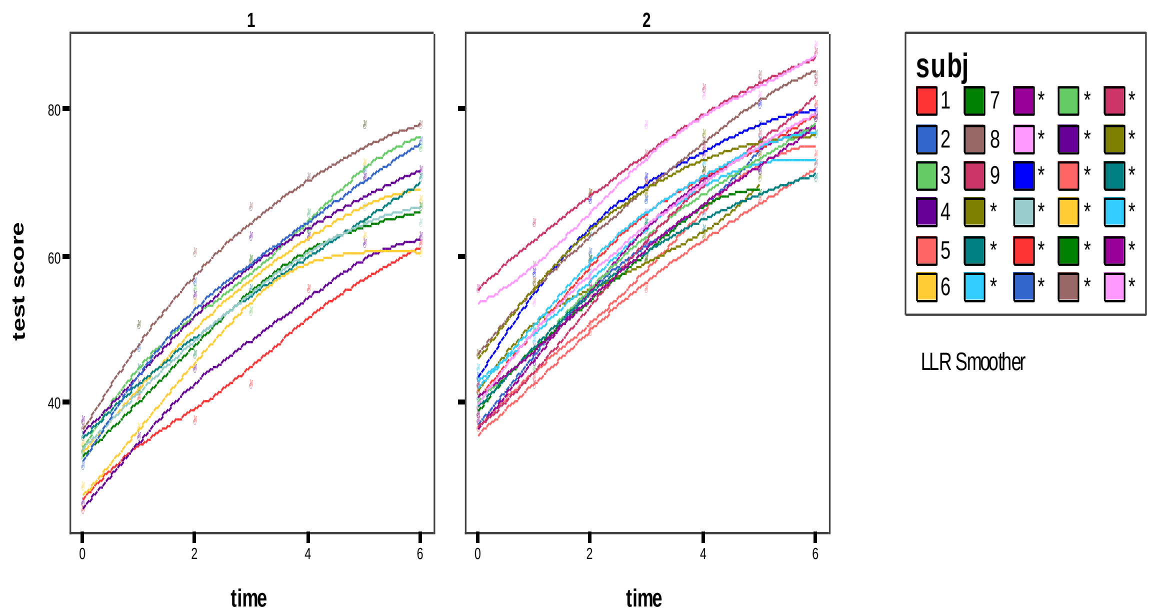
    Parameter count: 1 for quantitative IVs or k-1 for factors (categ. IVs)
  - Random effect factors have levels (often "subjects") that would be different in the next experiment, and are assumed to come from a Normal distribution with mean zero. Random intercepts and slopes are "one per upper level item". Random effects "induce" correlation. Parameter count: 1 (a variance describing the "spread" of the intercept or a slope across all *n* subjects)

HLM, cont.

- Mixed models are generally an improvement over repeated measures analyses:
  - In addition to spherical (univariate) and unstructured (multivariate) correlation, several other useful correlation structures are allowed.
  - Random intercepts and slopes are often scientifically meaningful.
  - Unexplained variability is appropriately assigned to upper vs. lower levels of the hierarchy.
  - Unequally and/ inconsistently spaced data are handled correctly.
  - Missing data are allowed and correctly handled (if missingness is not related to the value of the outcome).

HLM Example

- Random regression or growth curve modeling for a learning task. Subjects are repeatedly tested for performance on a learned task. Subjects are randomized to one of 2 manuals giving instructions for how to learn the task quickly.
  - EDA:



HLM Example, cont.

- **Model 1: Fixed quadratic model (i.e., just regression/ANCOVA)**

Model Dimension(a)

		Number of Levels	Number of Parameters
Fixed Effects	Intercept	1	1
	Man	2	1
	Time	1	1
	Time2	1	1
Residual			1
Total		5	5

Information Criteria

Akaike's Information Criterion (AIC)	1202.060
Schwarz's Bayesian Criterion (BIC)	1205.301

Information Criteria absolute values are not interpretable. A BIC-to-BIC or AIC-to-AIC difference of more than about 2 is clear evidence that the model with the lower IC is better.

Example, cont.

Estimates of Fixed Effects(b)

Parameter	Estimate	Std. Error	df	T	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	41.37	.94	189.0	44.61	.000	39.54	43.20
[man=1]	-8.68	.82	189	-10.61	.000	-10.29	-7.06
[man=2]	0(a)	0	.	.	.	.	.
Time	10.15	.70	189.0	14.52	.000	8.77	11.53
time2	-.68	.11	189.0	-5.99	.000	-.90	-.46

a. This parameter is set to zero because it is redundant.

b. Dependent Variable: test score.

Estimates of Covariance Parameters

Parameter	Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Residual	29.59	3.04	9.72	.000	24.19	36.20

Example, cont.

Estimated Marginal Means

manual version(b)

				95% Confidence Interval	
manual version	Mean	Std. Error	Df	Lower Bound	Upper Bound
1	32.696(a)	1.009	189.000	30.705	34.686
2	41.373(a)	.928	189.000	39.544	43.203

a Covariates appearing in the model are evaluated at the following values: time = 0, time2 = 0.
b Dependent Variable: test score.

Note: Calculation of p-values and CI's depend on the model assumptions being true.

Learning Example: Add R.I.

- Model 2: Random Intercept model (each subject has his or her own intercept)

Model Dimension

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	Man	2		1	
	Time	1		1	
	time2	1		1	
Random Effects	Intercept	1	Variance Components	1	subj
Residual				1	
Total		6		6	

R.I. Model, cont.

Note: “Covariance Structure” only applies when there are multiple random effects

Information Criteria(a) [vs. prior model BIC=1205.3]

Akaike's Information Criterion (AIC)	1066.127
Schwarz's Bayesian Criterion (BIC)	1072.610

Estimates of Fixed Effects

						95% Confidence Interval	
Parameter	Estimate	Std. Error	Df	t	Sig.	Lower Bound	Upper Bound
Intercept	41.45	1.17	38.2	35.32	.000	39.07	43.82
[man=1]	-8.82	1.79	28.1	-4.91	.000	-12.48	-5.14
[man=2]	0(a)	0	.	.	.	.	.
Time	10.04	.40	161.4	25.14	.000	9.25	10.83
time2	-.67	.06	161.4	-10.34	.000	-.79	-.54

a. This parameter is set to zero because it is redundant.

R.I. Model, cont.

Estimates of Covariance Parameters

Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Residual		9.56	1.06	8.97	.000	7.68	11.89
Intercept [subject = subj]	Variance	20.87	5.98	3.49	.000	11.90	36.58

Estimated Marginal Means

1. manual version(b)

manual version	Mean	Std. Error	Df	95% Confidence Interval	
				Lower Bound	Upper Bound
1	32.64(a)	1.49	33.58	29.60	35.67
2	41.45(a)	1.17	38.20	39.07	43.82

a. Covariates appearing in the model are evaluated at the following values: time = 0, time2 = 0.

Learning Example, cont.

- Conclusions: Manual 2 is better by 8.8 points (95% CI [5.1,12.5]) at each time. Subject-to-subject variability is about twice as large as within-subject variability. (The random intercept is justified, so the CIs for the "manual" means and the mean difference from the fixed model are wrong.) For each manual, across subjects, 95% of intercepts vary roughly $\pm 2\sqrt{20.9}$. Consider other models to check random slope, random curvature, interaction and/or serial correlation.

- Equation hierarchy "collapsed":

  $Y_{iT} = (\beta_0 + U_{0i} + \beta_M M) + \beta_T T + \beta_{T2} T^2 + \varepsilon_{iT}$

  where *i* is subject number, *T* is time, β_0 is the average intercept, U_i is a random per-subject intercept deviation, β_M is the effect of manual, M is a manual indicator, T is time, ε_{iT} is residual Normally distributed error, and U_{0i} is distributed Normally with mean 0 and variance τ_0^2 estimated at 20.9.

Example 2: NCES

- Hierarchical or multi-level model from the NCES “High School and Beyond” study. We will look at the relationship between the outcome math achievement test score (Math ACH) and the explanatory variable SES, taking into account the correlation of subjects within the same school.
- Model 1: Random Intercept Only

Model Dimensions(a)

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
Random Effects	Intercept	1	Identity	1	School
Residual				1	
Total		2		3	

a. Dependent Variable: Math ACH test score

NCES: R.I. model, cont.

Information Criteria

The information criteria are displayed in smaller-is-better forms.

Akaike's Information Criterion (AIC)	47120.8
Schwarz's Bayesian Criterion (BIC)	47134.6

a Dependent Variable: Math ACH test score.

Estimates of Fixed Effects(a)

						95% Confidence Interval	
Parameter	Estimate	Std. Error	Df	t	Sig.	Lower Bound	Upper Bound
Intercept	12.64	.24	156.6	51.7	.000	12.15	13.12

a Dependent Variable: Math ACH test score.

Estimates of Covariance Parameters

						95% Confidence Interval	
Parameter		Estimate	Std. Error	Wald Z	Sig.	Lower Bound	Upper Bound
Residual		39.15	.66	59.26	.000	37.87	40.46
Intercept [subject = school]	Variance	8.61	1.08	7.98	.000	6.74	11.01

Estimated marginal means table: mean=12.64, SE(mean)=0.24

NCES: Model 2

- Add mean school SES as a fixed effect school-level variable

Model Dimensions

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	Meanses	1		1	
Random Effects	Intercept	1	Identity	1	school
Residual				1	
Total		3		4	

Information Criteria [previous BIC=47135]

Akaike's Information Criterion (AIC)	46965.3
Schwarz's Bayesian Criterion (BIC)	46979.0

Estimate of Fixed Effects

						95% Confidence Interval	
Parameter	Estimate	Std. Error	Df	t	Sig.	Lower Bound	Upper Bound
Intercept	12.65	.15	153.7	84.74	.000	12.35	12.94
meanses	5.86	.36	153.4	16.22	.000	5.15	6.58

Note: meanses is a “z-score” so mean=0, sd=1

NCES: Model 3

- Add student (relative) SES as a fixed effect

Model Dimensions

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	meanses	1		1	
	Cses	1		1	
Random Effects	Intercept	1	Identity	1	School
Residual				1	
Total		4		5	

Information Criteria [previous BIC=46979]

Akaike's Information Criterion (AIC)	46572.6
Schwarz's Bayesian Criterion (BIC)	46586.3

NCES: Model 3, cont.

(Adding student SES as a fixed effect)

Estimates of Fixed Effects

Parameter	Estimate	Std. Error	Df	T	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	12.66	.15	153.6	84.76	.000	12.37	12.96
meanses	5.87	.36	153.3	16.22	.000	5.15	6.58
Cses	2.19	.11	7021.5	20.16	.000	1.98	2.40

Estimates of Covariance Parameters

Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Residual		37.02	.62	59.2	.000	35.81	38.26
Intercept [subject = school]	Variance	2.69	.40	6.6	.000	2.00	3.62

NCES: Model 4

- Add student SES random effect

Model Dimensions

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	meanses	1		1	
	cses	1		1	
Random Effects	Intercept + cses	2	Diagonal	2	school
Residual				1	
Total		5		6	

Information Criteria [previous BIC=46586]

Akaike's Information Criterion (AIC)	46564.8
Schwarz's Bayesian Criterion (BIC)	46585.4

NCES: Model 4, cont.

(Adding student SES random effect)

Estimates of Covariance Parameters

Parameter		Estimate	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Residual		36.708	.626	.000	35.502	37.956
Intercept + cses [subject = school]	Var: Intercept	2.699	.405	.000	2.011	3.622
	Var: cses	.695	.281	.013	.315	1.535

- Conclusions: Your score is better if you are in a high SES school, but other school-level factors remain to be discovered. It helps to have above average SES for your school, but there is little evidence that the size of this effect varies much from school to school. The largest component of variation is residual (after correcting for SES) student-to-student variation.

Mixed Models Summary

- Mixed (Hierarchical) models are used when the independence error assumption is violated
- Handles missing data and unequally spaced data better than “repeated measures”
- More flexibly models correlation structure
- Can model effects at different levels
- There are complex model selection issues