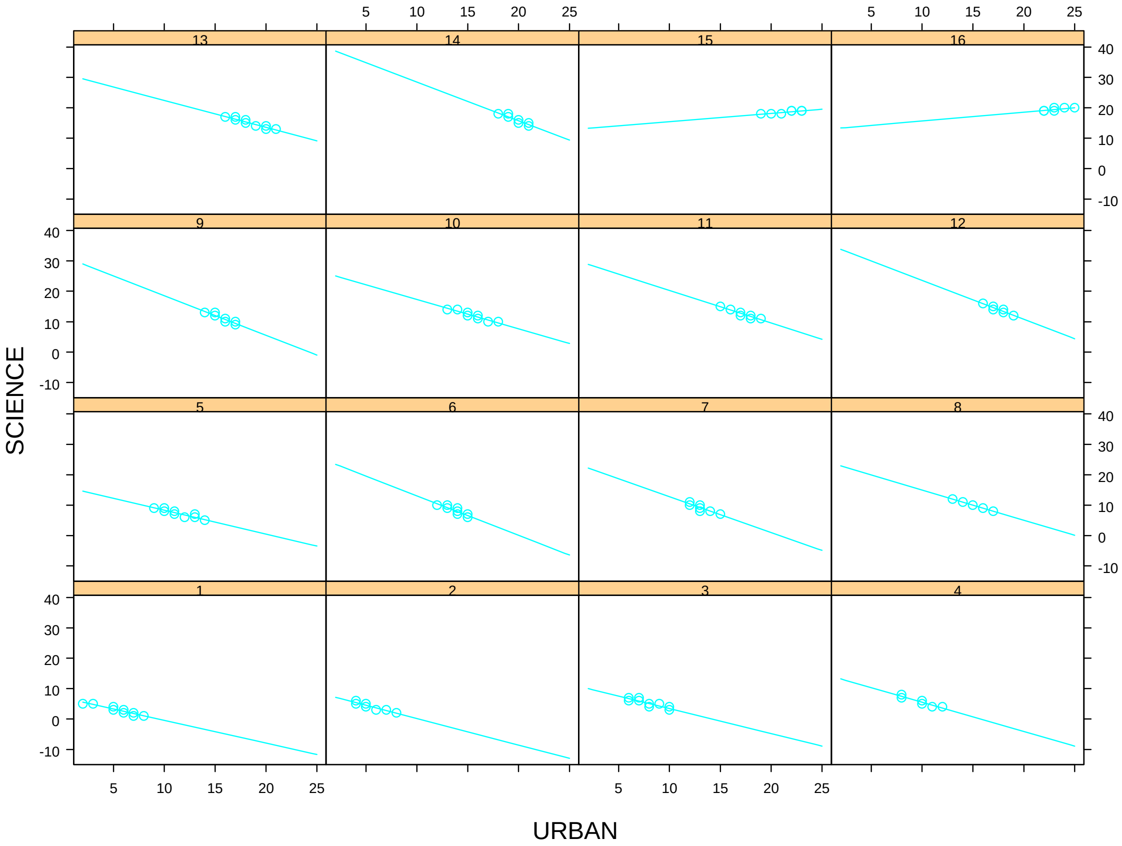
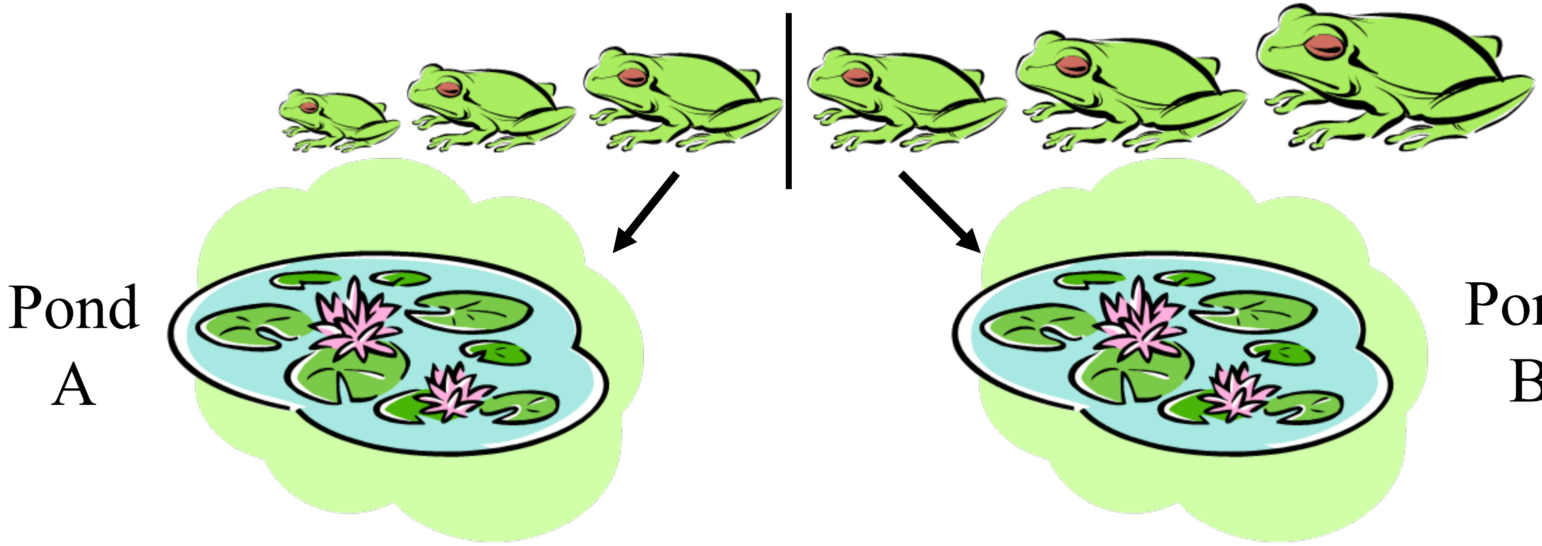
Introduction to Hierarchical Linear Modeling with R



Dr. J. Kyle Roberts

First Things First

- Robinson (1950) and the problem of contextual effects
- The "Frog-Pond" Theory



A Brief History of Multilevel Models

- Nested ANOVA designs
- Problems with the ANCOVA design
  - “Do schools differ” vs. “Why schools differ?”
  - ANCOVA does not correct for intra-class correlation (ICC)

Strengths of Multilevel Models

- Statistical models that are not hierarchical sometimes ignore structure and report underestimated standard errors
- Multilevel techniques are more efficient than other techniques
- Multilevel techniques assume a general linear model and can perform all types of analyses

Multilevel Examples

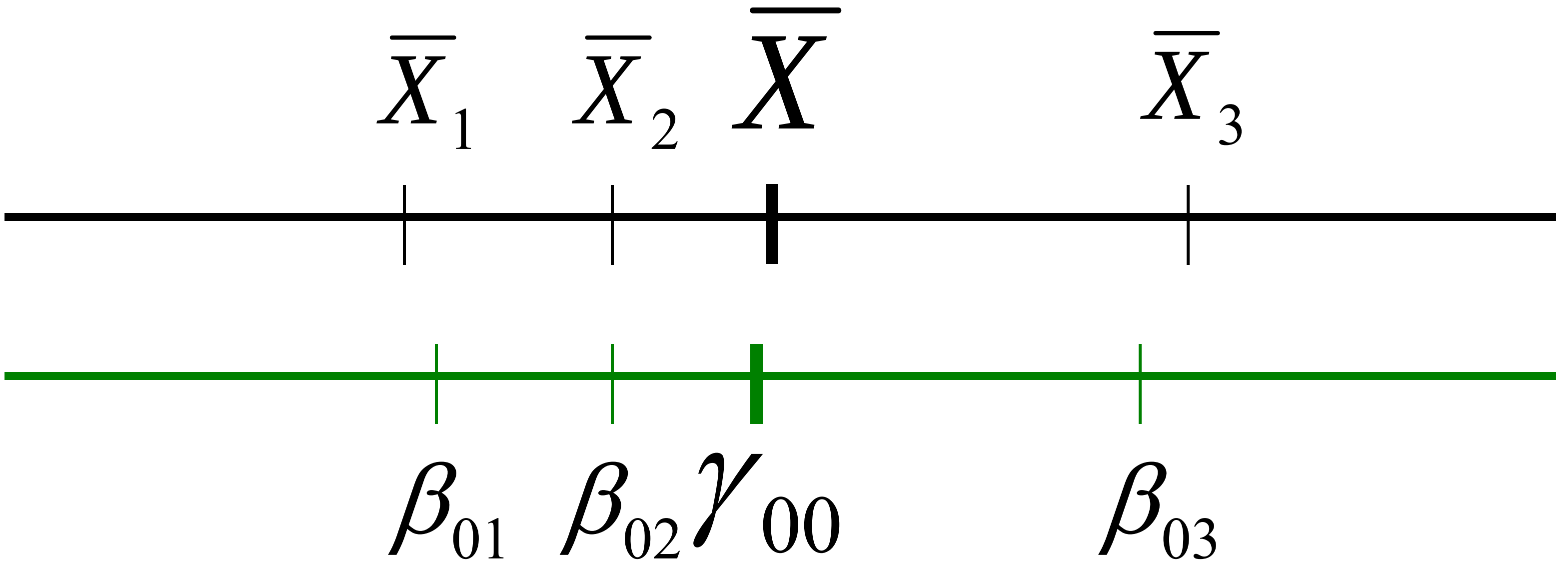
- Students nested within classrooms
- Students nested within schools
- Students nested within classrooms within schools
- Measurement occasions nested within subjects (repeated measures)
- Students cross-classified by school and neighborhood
- Students having multiple membership in schools (longitudinal data)
- Patients within a medical center
- People within households

Do we really need HLM/MLM?

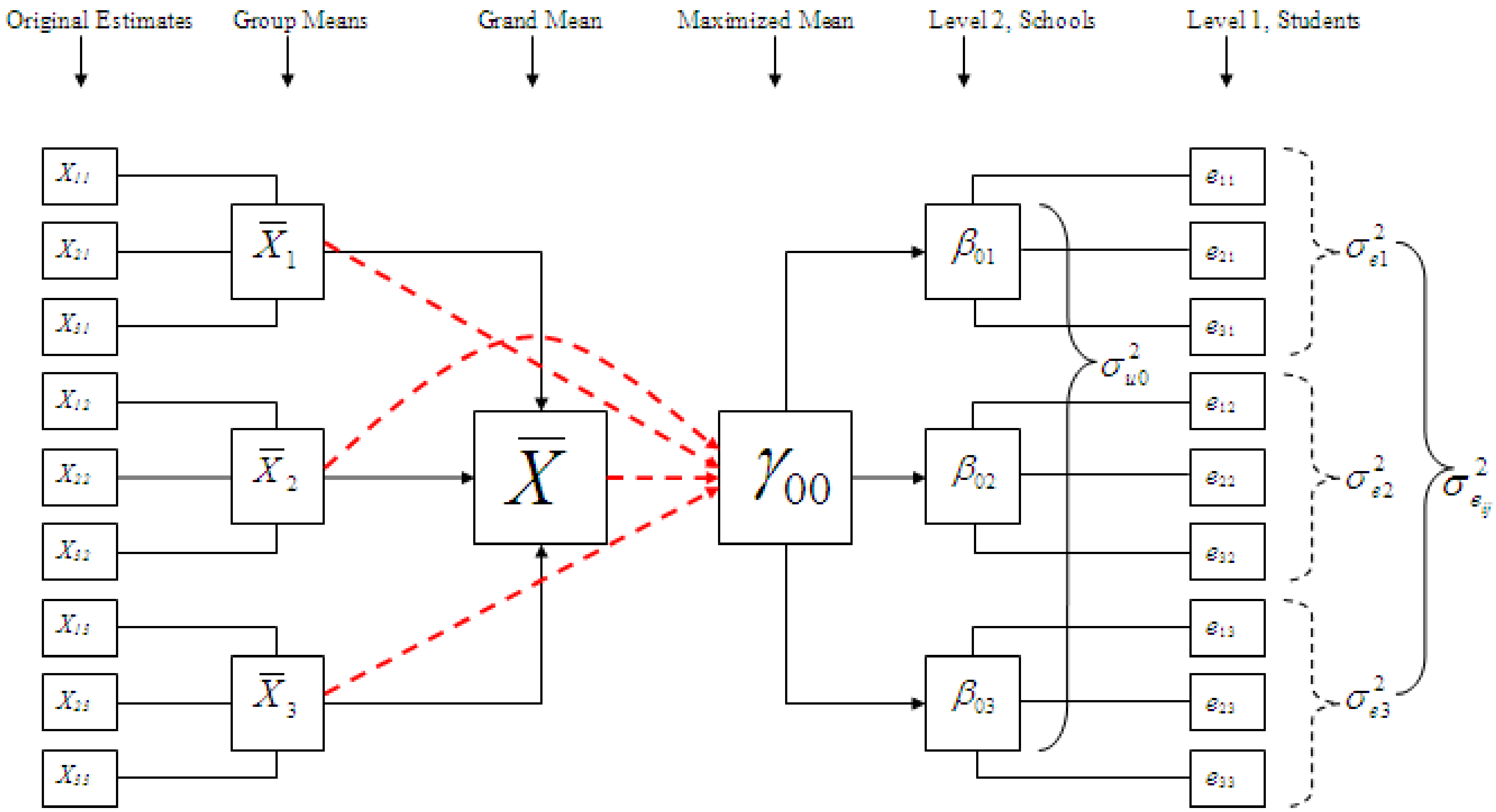
- "All data are multilevel!"
- The problem of independence of observations
- The "inefficiency" of OLS techniques

Differences in HLM and Other Methods

- HLM is based on Maximum Likelihood and Empirical Bayesian estimation techniques
- 1 + 1 = 1.5



Graphical Example of Multilevel ANOVA



Notating the HLM ANOVA

- The full model would be:

$$y_{ij} = \gamma_{00} + u_{0j} + e_{ij}$$

- Level-1 model is:

$$y_{ij} = \beta_{0j} + e_{ij}$$

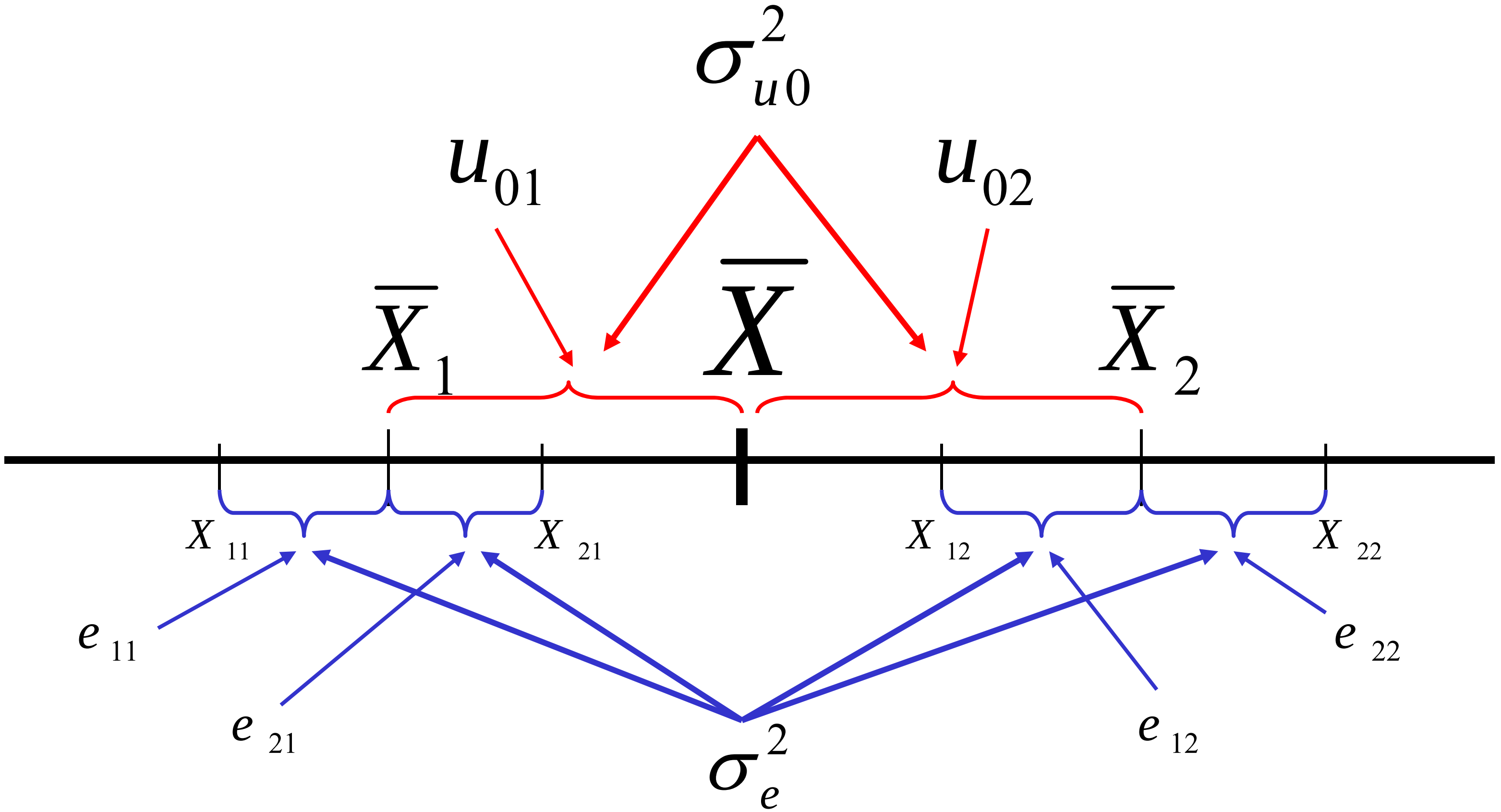
$$\begin{cases} y_{11} = \beta_1 + e_{11} \\ y_{21} = \beta_1 + e_{21} \\ \cdots \\ y_{ij} = \beta_j + e_{ij} \end{cases}$$

- Level-2 model is:

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

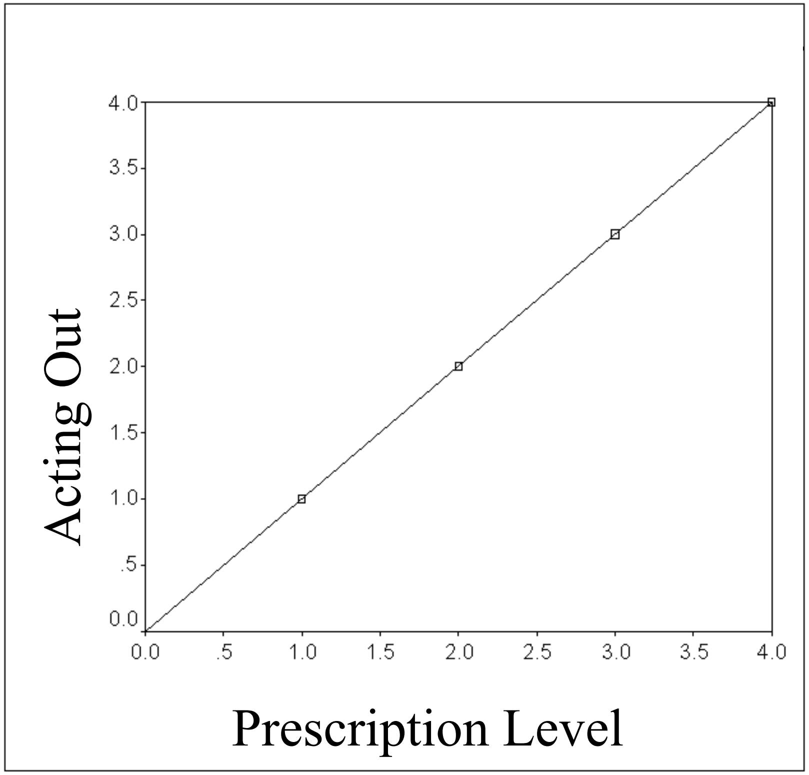
$$\begin{cases} \beta_1 = \gamma_{00} + u_1 \\ \beta_2 = \gamma_{00} + u_2 \\ \cdots \\ \beta_j = \gamma_{00} + u_j \end{cases}$$

Understanding Errors



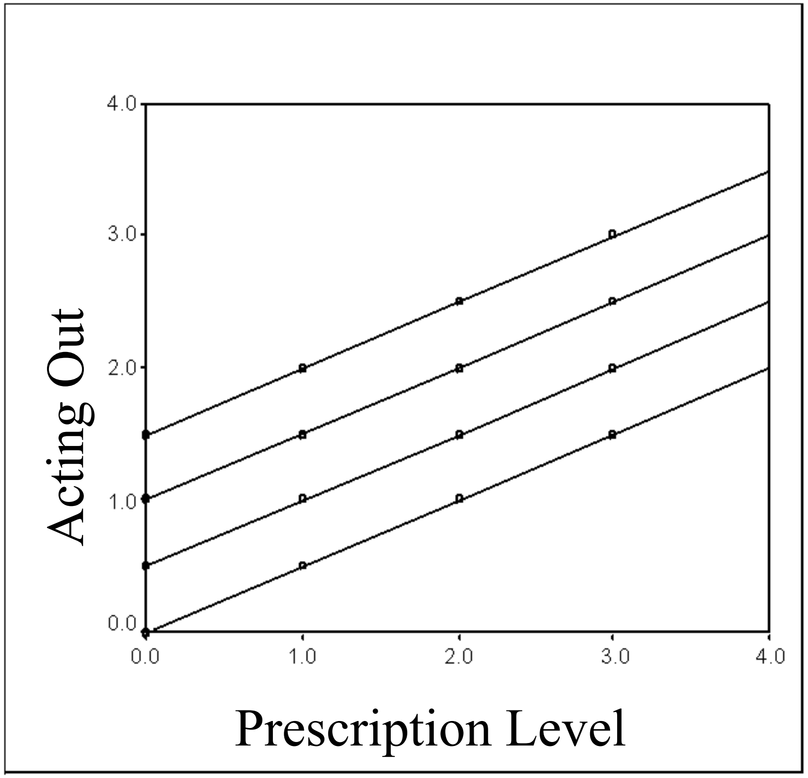
Fixed vs. Random Coefficients

- Fixed Slopes and Intercepts



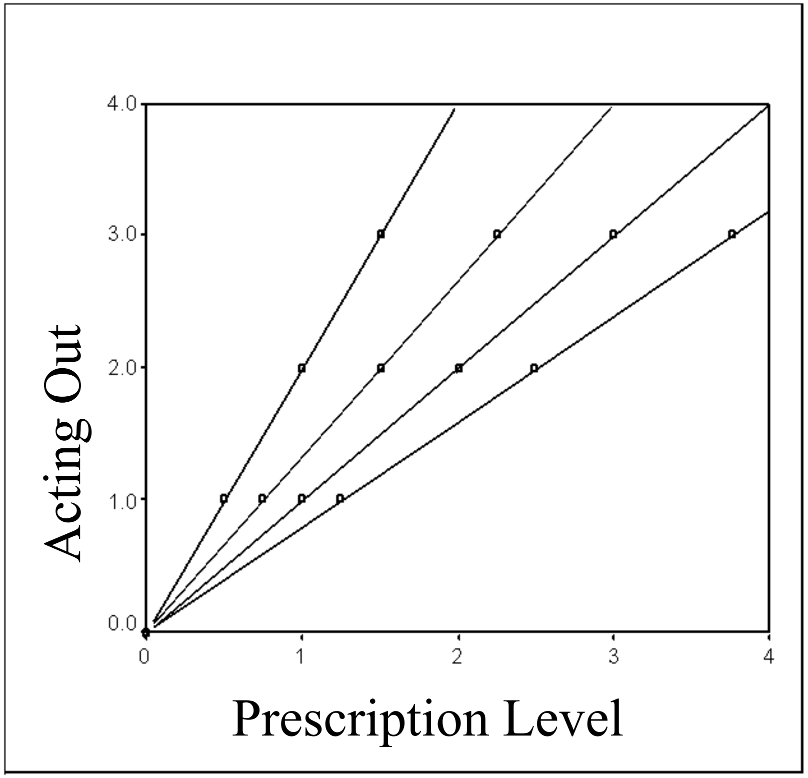
Fixed vs. Random Coefficients

- Fixed Slopes and Intercepts
- Random Intercepts and Fixed Slopes



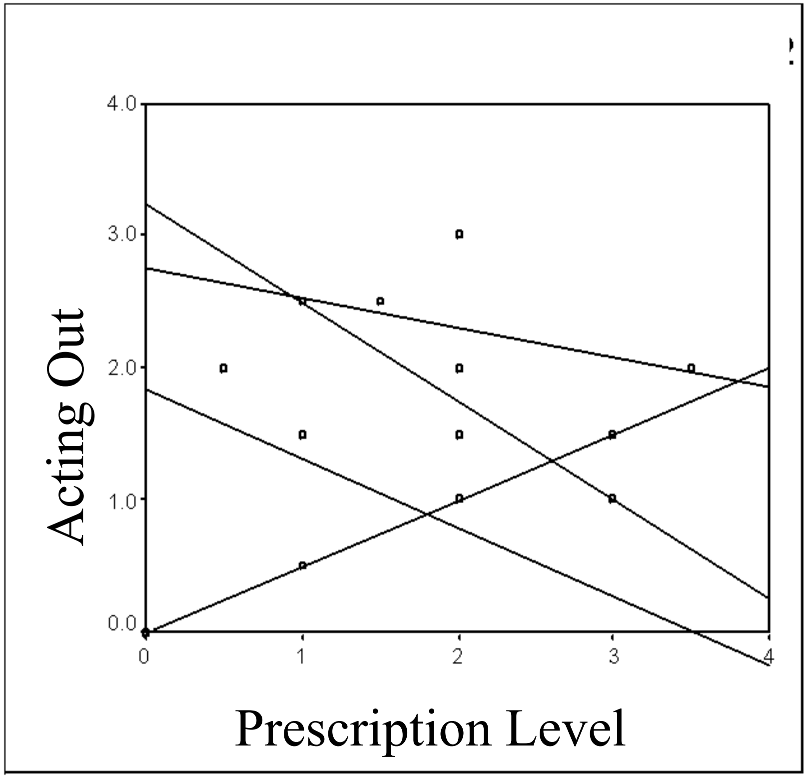
Fixed vs. Random Coefficients

- Fixed Slopes and Intercepts
- Random Intercepts and Fixed Slopes
- Fixed Intercepts and Random Slopes



Fixed vs. Random Coefficients

- Fixed Slopes and Intercepts
- Random Intercepts and Fixed Slopes
- Fixed Intercepts and Random Slopes
- Random Slopes and Intercepts



Let's Give This A Shot!!!

- An example where we use a child's level of "urbanicity" (a SES composite) to predict their science achievement
- Start with Multilevel ANOVA (also called the "null model")

$$science_{ij} = \gamma_{00} + u_{0j} + r_{ij}$$

Grand mean (γ_{00}) | Group deviation (u_{0j}) | Individual diff. (r_{ij})

Intraclass Correlation

- The proportion of total variance that is *between* the groups of the regression equation
- "The degree to which individuals share common experiences due to closeness in space and/or time" Kreft & de Leeuw, 1998.
- a.k.a – ICC is the proportion of group-level variance to the total variance
- LARGE ICC DOES NOT EQUAL LARGE DIFFERENCES BETWEEN MLM AND OLS (Roberts, 2002)
- Formula for ICC: $\rho = \dfrac{\sigma_{u0}^2}{\sigma_{u0}^2 + \sigma_{e0}^2}$

Statistical Significance???

- Chi-square vs. degrees of freedom in determining model fit
- The problem with the df
- Can also compute statistical significance of variance components (only available in some packages)

The Multilevel Model – Adding a Level-1 Predictor

- Consider the following 1-level regression equation:
  - $y = a + bx + e$
    - y = response variable
    - a = intercept
    - b = coefficient of the response variable (slope)
    - x = response variable
    - e = residual or error due to measurement

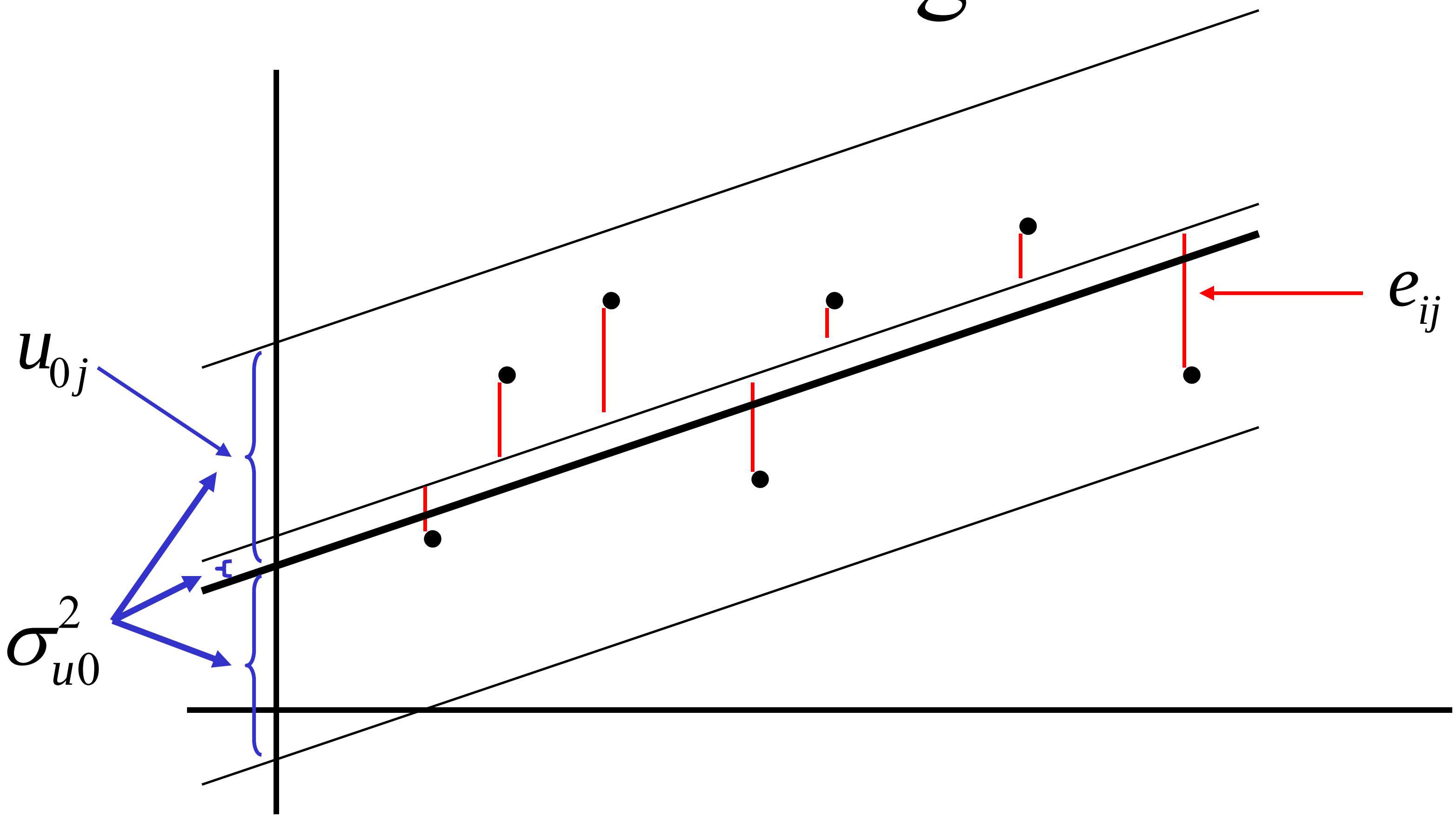
The Multilevel Model (2)

- The fixed coefficients multilevel model is a slight variation on the OLS regression equation:
  - $y_{ij} = a + bx_{ij} + u_j + e_{ij}$
    - Where "i" defines level-1, "j" defines level-2, u_j is the level-2 residual and e_{ij} is the level-1 residual
- Using slightly different annotation we can transform the above equation to:
  - $y_{ij} = \gamma_{00} + \gamma_{10}x_{ij} + u_{0j} + e_{ij}$
    - Where γ_{00} now defines the constant/intercept "a" and γ_{10} defines the slope

The Multilevel Model (3)

- From the previous model:
  - $y_{ij} = \gamma_{00} + \gamma_{10}x_{ij} + u_{0j} + e_{ij}$
- We can then transform this model to:
  - $y_{ij} = \beta_{0j} + \beta_1 x_{1ij} + e_{ij}$ ← Level-1 Model
  - $\beta_{0j} = \gamma_{00} + u_{0j}$ ← Level-2 Model
  - $\beta_{1j} = \gamma_{10}$ ←
  - With variances $u_{0j} = \sigma^2_{u0}$ $e_{ij} = \sigma^2_{e_{ij}}$

Understanding Errors



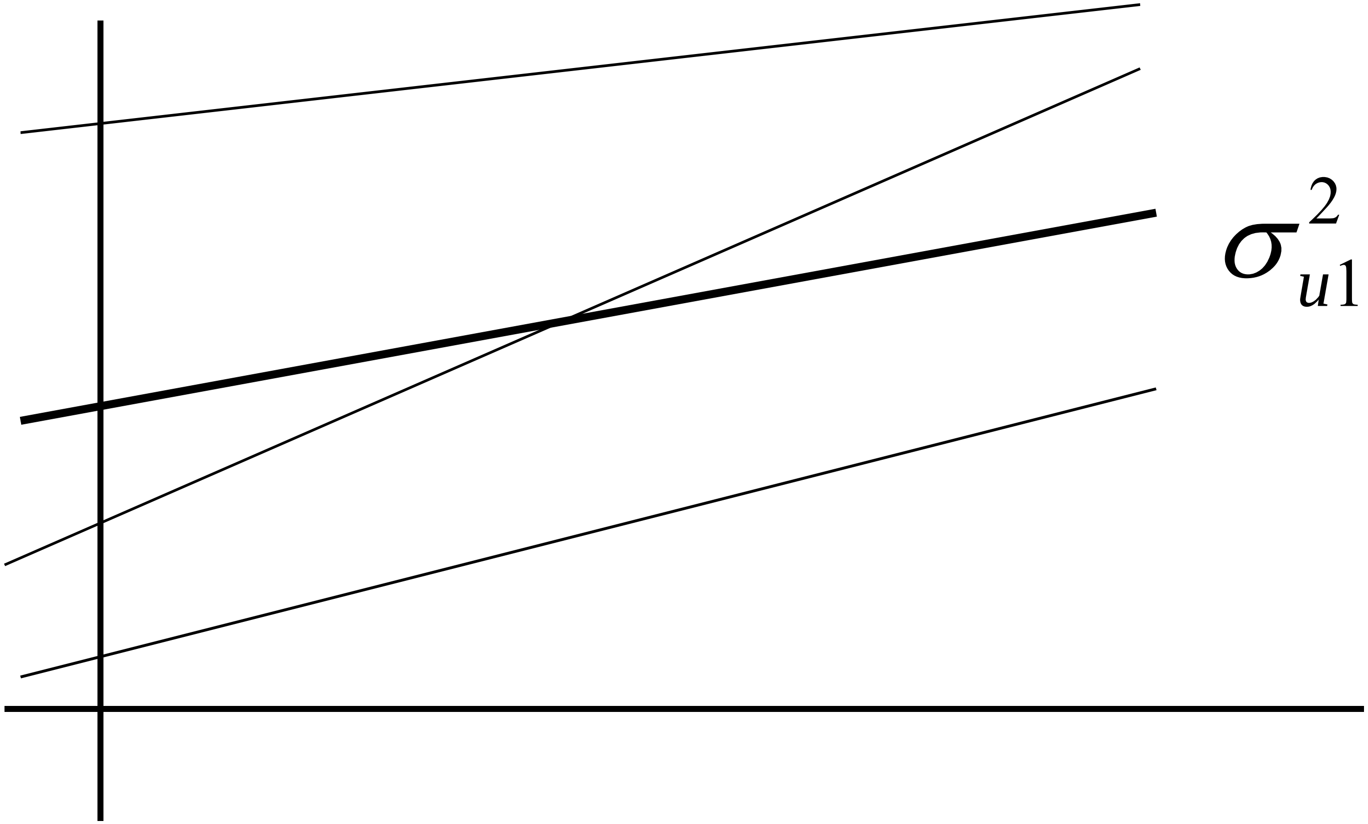
Adding a Random Slope Component

- Suppose that we have good reason to assume that it is inappropriate to "force" the same slope for "urbanicity" on each school
  - Level-1 Model → $y_{ij} = \beta_{0j}x_0 + \beta_{1j}x_{1ij} + r_{ij}$
  - Level-2 Model → $\beta_{0j} = \gamma_{00} + u_{0j}$

    $\beta_{1j} = \gamma_{10} + u_{1j}$

Complete Model → $science_{ij} = \gamma_{00} + u_{0j} + (\gamma_{10} + u_{1j})urban + r_{ij}$

Understanding Errors Again



σ_{u1}^{2} ????

Model Fit Indices

- Chi-square $-2*\ell$
- Akaike Information Criteria

$$AIC = -2*\ell + 2K$$

- Bayesian Informaiton Criteria

$$BIC = -2*\ell + K*Ln(N)$$

To Center or Not to Center

- In regression, the intercept is interpreted as the expected value of the outcome variable, when all explanatory variables have the value zero
- However, zero may not even be an option in our data (e.g., Gender)
- This will be especially important when looking at cross-level interactions
- General rule of thumb: If you are estimating cross-level interactions, you should grand mean center the explanatory variables.

An Introduction to R

R as a Big Calculator

- Language was originally developed by AT&T Bell Labs in the 1980's
- Eventually acquired by MathSoft who incorporated more of the functionality of large math processors
- The commands window is like a big calculator

```
> 2+2
[1] 4
```

```
> 3*5+4
[1] 19
```

Object Oriented Language

- A Big Plus for R is that it utilizes object oriented language.

```
> x<-2+4
> x
[1] 6

> y<-3+5
> y
[1] 8

> x+y
[1] 14
```

```
> x<-1:10
> x
 [1]  1  2  3  4  5  6
7  8  9 10
> mean(x)
[1] 5.5

> 2*x
 [1]  2  4  6  8 10 12 14 16
18 20
> x^2
 [1]   1   4   9  16  25  36
49  64  81 100
```

Utilizing Functions in R

- R has many "built in" functions (c.f., "Language Reference" in the "Help" menu)
- Functions are commands that contain "arguments"
- `seq` function has 4 arguments
  - `seq(from, to, by, length.out, along.with)`

```
> ?seq
> seq(from=1, to=100, by=10)
 [1]  1 11 21 31 41 51 61 71
81 91
```

```
> seq(1, 100, 10)
> seq(1, by=10,
length=4)
[1]  1 11 21 31
```

Making Functions in R

```
> squared<-function(x){x^2}
> squared(5)
[1] 25

> inverse<-function(x){1/x}
> num<-c(1,2,3,4,5)
> inverse(num)
[1] 1.0000000 0.5000000 0.3333333
0.2500000 0.2000000
```

Reading a Dataset

```
> example<-read.table(file, header=T)
> example<-read.table("c:/aera/example.txt", header=T)
> head(example)
```

Other Software Packages for HLM Analysis

- Good Reviews at http://www.cmm.bristol.ac.uk/
  - MLwiN
  - SAS – PROC MIXED, PROC NLMIXED, GLIMMIX
  - S-PLUS – lme, nlme, glme
  - M*plus*
  - SPSS – Version 12 and later
  - STATA