

Multilevel models

02.28.22

[Click here for PDF of slides](#)

Announcements

- HW 03 due Tue, Mar 01 at 11:59pm
- Raphaël's office hours this week will be Tue, 1:30 - 3:30pm

Learning goals

- Recognize a potential for correlation in a data set
- Identify observational units at varying levels
- Understand issues correlated data may cause in modeling
- Understand how random effects models can be used to take correlation into account
- Use EDA for multilevel data

Correlated observations

Multilevel data

- We can think of correlated data as a multilevel structure
 - Population elements are aggregated into groups
 - There are observational units and measurements at each level
- For now we will focus on data with two levels:
 - **Level one:** Most basic level of observation
 - **Level two:** Groups formed from aggregated level-one observations

Two types of effects

- **Fixed effects:** Effects that are of interest in the study
 - Can think of these as effects whose interpretations would be included in a write up of the study
- **Random effects:** Effects we're not interested in studying but whose variability we want to understand
 - Can think of these as effects whose interpretations would not necessarily be included in a write up of the study

Practice

Radon is a carcinogen – a naturally occurring radioactive gas whose decay products are also radioactive – known to cause lung cancer in high concentrations. The EPA sampled more than 80,000 homes across the U.S. Each house came from a randomly selected county and measurements were made on each level of each home. Uranium measurements at the county level were included to improve the radon estimates.

1. What is the most basic level of observation (Level One)?
2. What are the group units (Level Two, Level Three, etc...)
3. What is the response variable?
4. Describe the within-group variation.
5. What are the fixed effects? What are the random effects?

Ex. 1 from Section 7.10.1 in BMLR



03 : 00

Teratogen and rat pups

Data: Teratogen and rat pups

Today's data are simulated results of an experiment with 24 dams (mother rats) randomly divided into four groups that received different doses of teratogen, a substance that could potentially cause harm to developing fetuses. The four groups are

- High dose (3 mg)
- Medium dose (2 mg)
- Low dose (1 mg)
- No dose (Control)

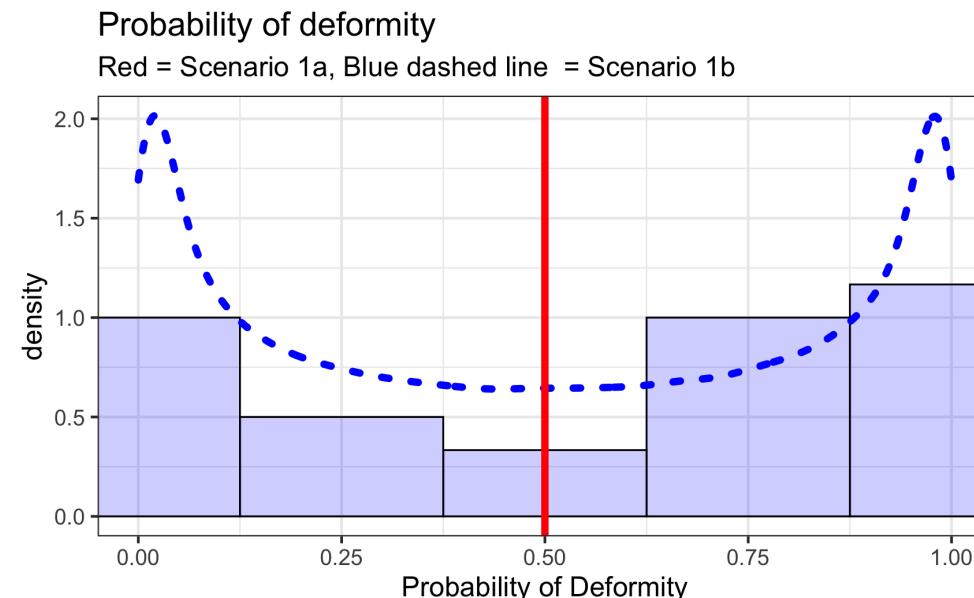
Each dam produced 10 rat pups and the presence of a deformity was noted.

Goal: Understand the association between teratogen exposure and the probability a pup is born with a deformity.

Scenario 1: No dose effect

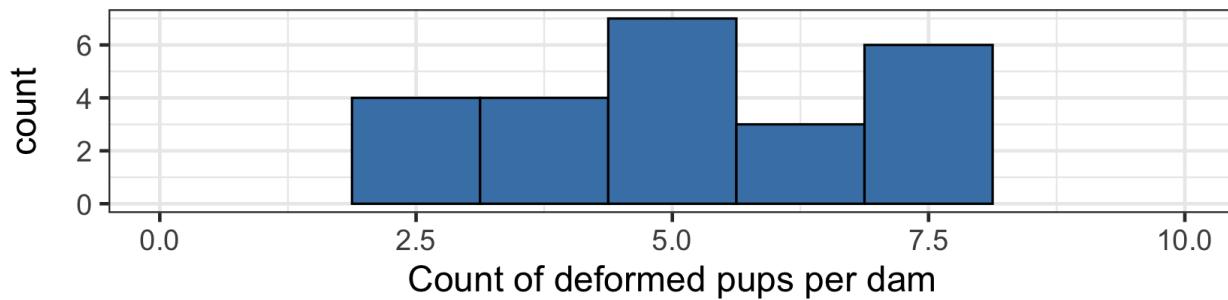
Assume dose has no effect on, p , the probability of a pup born with a deformity.

- **Scenario 1a.:** $p = 0.5$ for each dam
- **Scenario 1b.:** $p \sim Beta(0.5, 0.5)$ (expected value = 0.5)

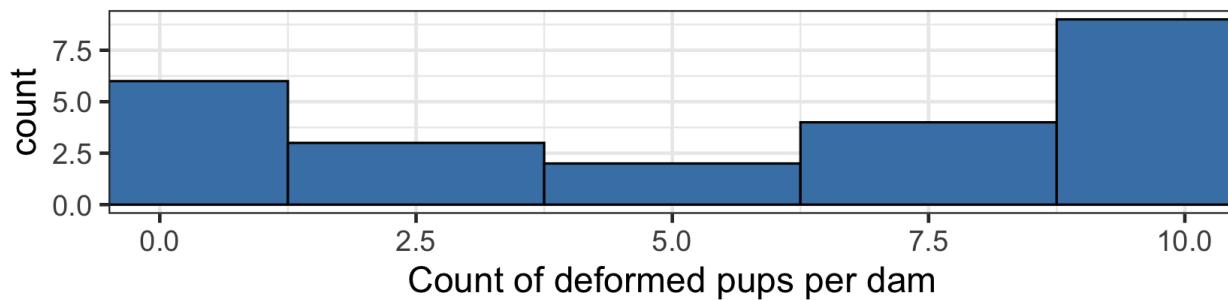


From Figure 7.1 in BMLR

Scenario 1a: Binomial, $p = 0.5$



Scenario 1b: Binomial, $p \sim \text{Beta}(0.5, 0.5)$



mean_1a	sd_1a	mean_1b	sd_1b
5.166667	1.493949	5.666667	4.103727

Scenario 2: Dose effect

Scenario 2: Dose effect

Now we will consider the effect of the dose of teratogen on the probability of a pup born with a deformity. The 24 pups have been randomly divided into four groups:

- High dose (**dose = 3**)
- Medium dose (**dose = 2**)
- Low dose (**dose = 1**)
- No dose (**dose = 0**)

We will assume the true relationship between p and dose is the following:

$$\log\left(\frac{p}{1-p}\right) = -2 + 1.33 \text{ dose}$$

Scenario 2

Scenario 2a.

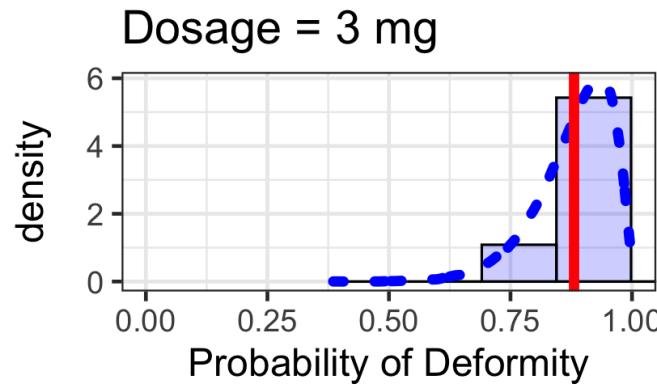
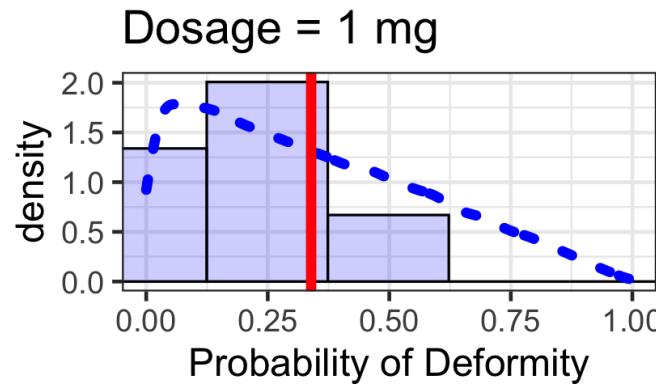
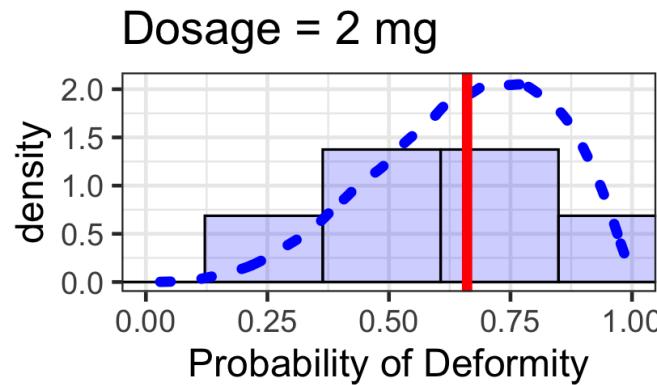
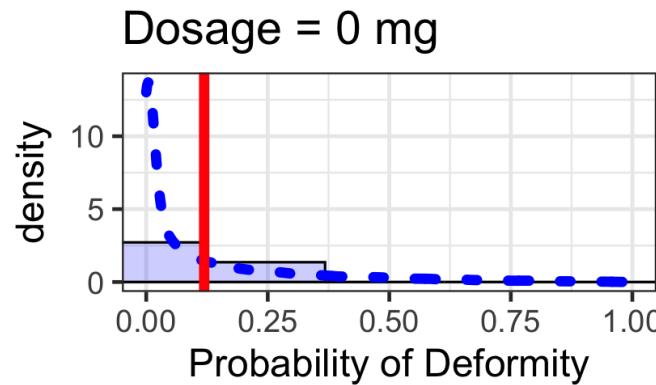
$$p = \frac{e^{-2+1.33 \text{ dose}}}{1 + e^{-2+1.33 \text{ dose}}}$$

Scenario 2b.

$$p \sim Beta\left(\frac{2p}{(1-p)}, 2\right)$$

On average, dams who receive dose x have the same probability of deformed pup as dams with dose x under Scenario 2a.

Distributions under Scenario 2



Replicated from Figure 7.3 in BMLR

Summary statistics under Scenario 2

Summary statistics of Scenario 2 by dose.

Dosage	Scenario 2a					Scenario 2b				
	Mean p	SD p	Mean Count	SD Count		Mean p	SD p	Mean Count	SD Count	
0	0.119	0	1.333	1.366		0.061	0.069	0.500	0.837	
1	0.339	0	3.167	1.835		0.239	0.208	3.500	2.881	
2	0.661	0	5.833	1.472		0.615	0.195	5.833	1.941	
3	0.881	0	8.833	1.169		0.872	0.079	8.833	1.169	

From Table 7.2 in BMLR



1. In Scenario 2a, dams produced 4.79 deformed pups on average, with standard deviation 3.20. Scenario 2b saw an average of 4.67 with standard deviation 3.58. Why are comparisons by dose more meaningful than these overall comparisons?
2. We will use binomial and quasibinomial regression to model the relationship between dose and probability of pup born with a deformity. What can you say about the center and the width of the confidence intervals under Scenarios 2a and 2b?
 - Which will be similar and why?
 - Which will be different and how?

Scenario 2: Estimated odds ratio

The estimated effect of dose and the 95% CI from the binomial and quasibinomial models are below:

Scenario 2a

	Odds Ratio	95% CI
Binomial	3.536	(2.604, 4.958)
Quasibinomial	3.536	(2.512, 5.186)

Scenario 2b

	Odds Ratio	95% CI
Binomial	4.311	(3.086, 6.271)
Quasibinomial	4.311	(2.735, 7.352)

1. Describe how the quasibinomial analysis of Scenario 2b differs from the binomial analysis of the same simulated data. Do confidence intervals contain the true model parameters? Is this what you expected? Why?
2. Why are differences between quasibinomial and binomial models of Scenario 2a less noticeable than the differences in Scenario 2b?

Summary

- The structure of the data set may imply correlation between observations.
- Correlated observations provide less information than independent observations; we need to account for this reduction in information.
- Failing to account for this reduction could result in underestimating standard error, thus resulting in overstating significance and the precision of the estimates.
- We showed how we can account for this by incorporating the dispersion parameter or a random effect.

Multilevel models

Data: Music performance anxiety

The data [**musicdata.csv**](#) come from the Sadler and Miller (2010) study of the emotional state of musicians before performances. The dataset contains information collected from 37 undergraduate music majors who completed the Positive Affect Negative Affect Schedule (PANAS), an instrument produces a measure of anxiety (negative affect) and a measure of happiness (positive affect). This analysis will focus on negative affect as a measure of performance anxiety.

The primary variables we'll use are

- **na**: negative affect score on PANAS (the response variable)
- **perform_type**: type of performance (Solo, Large Ensemble, Small Ensemble)
- **instrument**: type of instrument (Voice, Orchestral, Piano)

Look at data

id	diary	perform_type	na	gender	instrument
1	1	Solo	11	Female	voice
1	2	Large Ensemble	19	Female	voice
1	3	Large Ensemble	14	Female	voice
43	1	Solo	19	Female	voice
43	2	Solo	13	Female	voice
43	3	Small Ensemble	19	Female	voice

- What are the Level One observations? Level Two observations?
- What are the Level One variables? Level Two variables?

Univariate exploratory data analysis

Level One variables

Two ways to approach univariate EDA (visualizations and summary statistics) for Level One variables:

- Use individual observations (i.e., treat observations as independent)
- Use aggregated values for each Level Two observation

Level Two variables

- Use a data set that contains one row per Level Two observation

Complete Part 1: Univariate EDA in `lecture-15.Rmd`



08 : 00

Bivariate exploratory data analysis

Goals

- Explore general association between the predictor and response variable
- Explore whether subjects at a given level of the predictor tend to have similar mean responses
- Explore whether variation in response differs at different levels of a predictor

There are two ways to visualize these associations:

- One plot of individual observations (i.e., treat observations as independent)
- Separate plots of responses vs. predictor for each Level Two observation (lattice plots)



Fitting the model

Questions we want to answer

The goal is to understand variability in performance anxiety (**na**) based on performance-level and musician-level characteristics. Specifically:

- What is the association between performance type (large ensemble or not) and performance anxiety? Does the association differ based on instrument type (orchestral or not)?

What is the problem with using the following model to draw conclusions?

term	estimate	std.error	statistic	p.value
(Intercept)	15.721	0.359	43.778	0.000
orchestra	1.789	0.552	3.243	0.001
large_ensemble	-0.277	0.791	-0.350	0.727
orchestra:large_ensemble	-1.709	1.062	-1.609	0.108

Other modeling approaches

- 1 Condense each musician's set of responses into a single outcome (e.g., mean max, last observation, etc.) and fit a linear model on these condensed observations
 - Leaves few observations (37) to fit the model
 - Ignoring a lot of information in the multiple observations for each musician
- 2 Fit a separate model for each musician understand the association between performance type (Level One models). Then fit a system of Level Two models to predict the fitted coefficients in the Level One model for each subject based on instrument type (Level Two model).

Let's look at approach #2

Two-level model

We'll start with the Level One model to understand the association between performance type and performance anxiety for the i^{th} musician.

$$na_{ij} = a_i + b_i \text{LargeEnsemble}_{ij} + \epsilon_i, \quad \epsilon_{ij} \sim N(0, \sigma^2)$$

Why is it more meaningful to use performance type for the Level One model than instrument?

For now, estimate a_i and b_i using least-squares regression.

Example Level One model

Below is partial data for observation #22

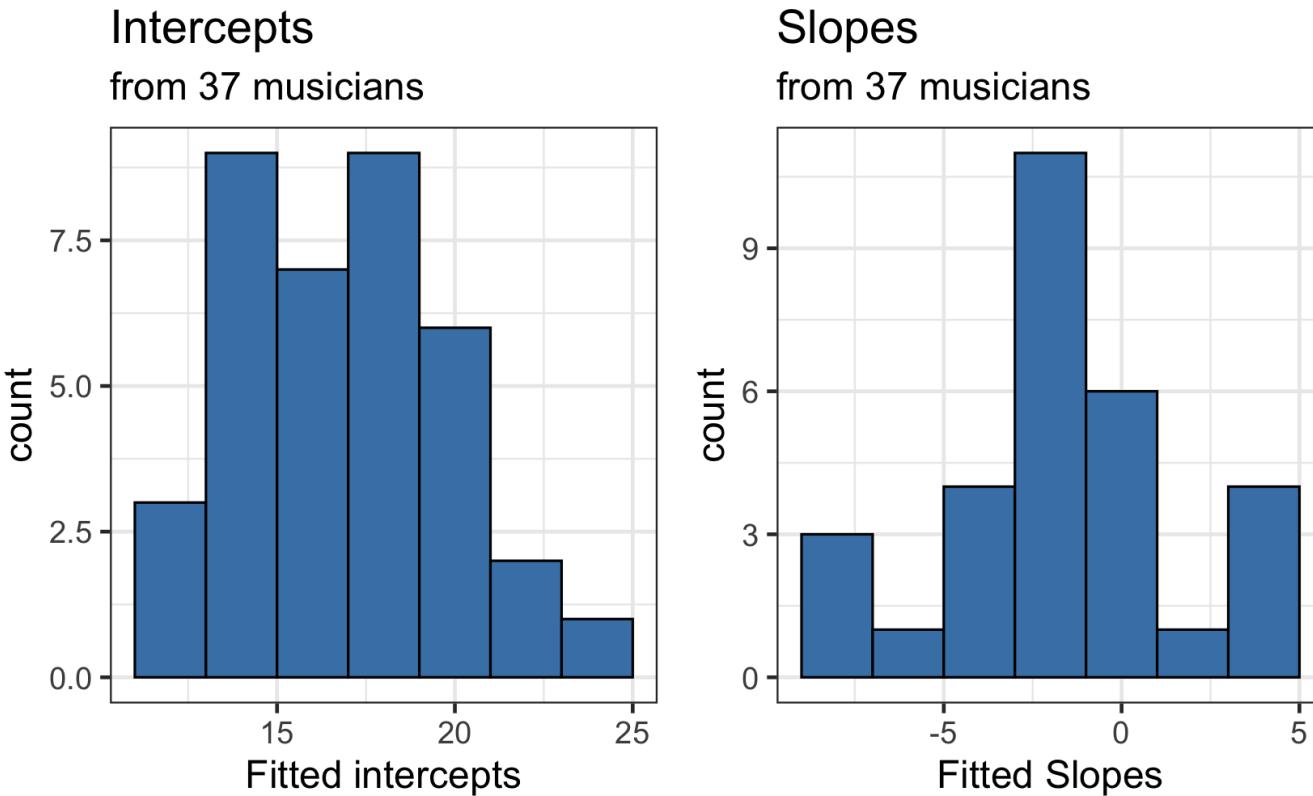
id	diary	perform_type	instrument	na
22	1	Solo	orchestral instrument	24
22	2	Large Ensemble	orchestral instrument	21
22	3	Large Ensemble	orchestral instrument	14
22	13	Large Ensemble	orchestral instrument	12
22	14	Large Ensemble	orchestral instrument	19
22	15	Solo	orchestral instrument	25

Level One model

```
music %>%
  filter(id == 22) %>%
  lm(na ~ large_ensemble, data = .) %>%
  tidy() %>%
  kable(digits = 3)
```

term	estimate	std.error	statistic	p.value
(Intercept)	24.500	1.96	12.503	0.000
large_ensemble	-7.833	2.53	-3.097	0.009

Repeat for all 37 musicians. See Part 3: Level One Models in **lecture-15.Rmd**.



Recreated from BMLR Figure 8.9

Now let's consider if there is an association between the estimated slopes, estimated intercepts, and the type of instrument

Level Two Model

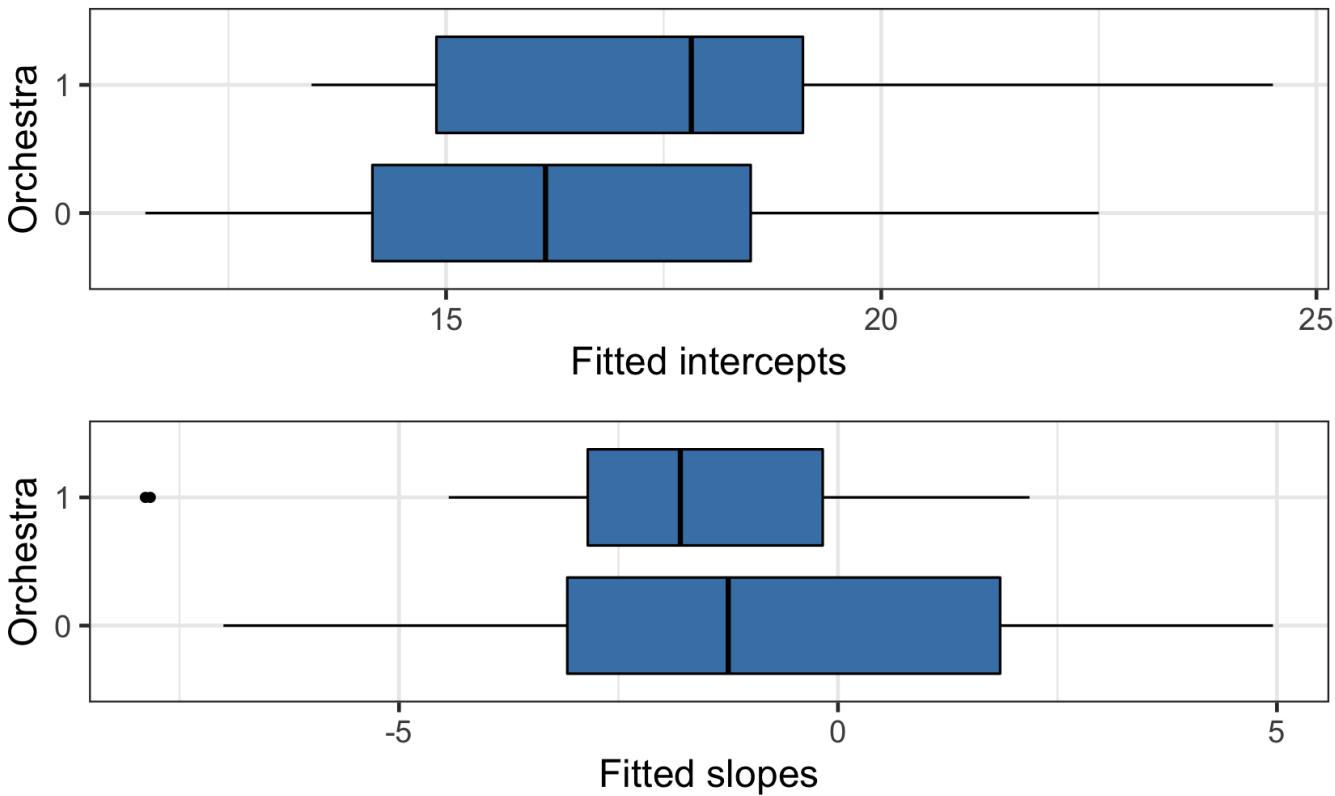
The slope and intercept for the i^{th} musician can be modeled as

$$a_i = \alpha_0 + \alpha_1 \text{Orchestra}_i + u_i$$
$$b_i = \beta_0 + \beta_1 \text{Orchestra}_i + v_i$$

Note the response variable in the Level Two models are not observed outcomes but the (fitted) slope and intercept from each musician

See Part 4: Level Two Models in **lecture-15.Rmd**.

Estimated coefficients by instrument



Level Two model

Model for intercepts

term	estimate	std.error	statistic	p.value
(Intercept)	16.283	0.671	24.249	0.000
orchestra	1.411	0.991	1.424	0.163

Model for slopes

term	estimate	std.error	statistic	p.value
(Intercept)	-0.771	0.851	-0.906	0.373
orchestra	-1.406	1.203	-1.168	0.253

Writing out the models

Level One

$$\hat{n}a_{ij} = \hat{a}_i + \hat{b}_i \text{ LargeEnsemble}_{ij}$$

for each musician.

Level Two

$$\hat{a}_i = 16.283 + 1.441 \text{ Orchestra}_i$$

$$\hat{b}_i = -0.771 - 1.406 \text{ Orchestra}_i$$

Composite model

$$\hat{na}_i = 16.283 + 1.441 \text{ Orchestra}_i - 0.771 \text{ LargeEnsemble}_{ij} \\ - 1.406 \text{ Orchestra : LargeEnsemble}_{ij}$$

(Note that we also have the error terms ϵ_{ij}, u_i, v_i that we will discuss next class.)

- What is the predicted average performance anxiety before solos and small ensemble performances for vocalists and keyboardists? For those who play orchestral instruments?
- What is the predicted average performance anxiety before large ensemble performances for those who play orchestral instruments?

Disadvantages to this approach

- ⚠ Weighs each musician the same regardless of number of diary entries
- ⚠ Drops subjects who have missing values for slope (7 individuals who didn't play a large ensemble performance)
- ⚠ Does not share strength effectively across individuals

Plot the R^2 values calculated in Part 3: Level One Model of **lecture-15.Rmd**.

We will use a unified approach that utilizes likelihood-based methods to address some of these drawbacks.

Acknowledgements

The content in the slides is from

- [BMLR: Chapter 7 - Correlated data](#)
- [BMLR: Chapter 8 - Introduction to Multilevel Models](#)
- Sadler, Michael E., and Christopher J. Miller. 2010. "Performance Anxiety: A Longitudinal Study of the Roles of Personality and Experience in Musicians." Social Psychological and Personality Science 1 (3): 280–87.
<http://dx.doi.org/10.1177/1948550610370492>.