

Stat 414 – Day 27 Case Study

Last Time: Looking for unusual observations

- Outliers vs. Influential observations
- Don't just flag but investigate!
- Difference between dropping a level 2 unit and a level 1 unit

Part 1: Stage fright can be a serious problem for performers, and understanding the personality underpinnings of performance anxiety is an important step in determining how to minimize its impact. Sadler and Miller (2010) studied the emotional state of musicians before performances and factors which may affect their emotional state. Data were collected on 37 undergraduate music majors over the course of an academic year. Students completed diaries prior to performances, including the [Positive Affect Negative Affect Schedule \(PANAS\)](#) before each performance. The *negative affect* measure of this instrument is used as a measure of performance anxiety. Factors include type of performance (solo, large ensemble, small ensemble), audience (orchestral vs. keyboard/vocalist), age, gender, instrument, and years studying. (This dataset involves for categorical explanatory variables, not sure why it takes me a little longer to understand the model, but keep in mind that “slopes” correspond to “group differences.”)

Initial Data Exploration

1. Is there much variation in negative affect (na) across performances? Is there much variation in na across musicians?
2. Confirm the types of variables/number of categories and start thinking about how you will incorporate these into a model.
3. Does negative affect seem to vary by performance type? What about instrument? Do you think either variable will be a useful variable to include?
4. Does the relationship between negative affect and number of previous performances appear to differ across musicians? What does that suggest including in the model?

For the graph of negative affect vs. NEM

(c) What interesting pattern do you notice about this graph? Why is it expected? What's the difference between the two plots? Is one better than the other?

Some of the questions we might want to explore: Which characteristics of individual performances are most associated with performance anxiety? Which characteristics of student musicians are most associated with performance anxiety? Are any of these associations statistically significant? Does the significance remain after controlling for other covariates? But of course, need to account for lack of independence in performances by the same musician.

(d) Identify Level 1 and Level 2. Identify some variables at each level.

(e) Start with a random intercepts model to assess the variation in performance anxiety ("na") among the musicians. What is the ICC?

(f) Fit a model to predict performance anxiety based on the type of performance (large ensemble or not) with random intercepts and random slopes. Explain in plain language what it means for this model to have "random intercepts." What does it mean for this model to have "random slopes" for that variable?

(g) What does $\hat{\sigma}$ represent here in this new model? What do the two Level 2 variance components represent?

(h) The graph shows the fitted equations from the multilevel model for each performer. How do you think the graph will differ if we fit a separate line for each performer?

(i) Interpret your model output: Do the signs of the coefficients of the fixed effects make sense in context? What do you learn about the effect of large ensemble performances on anxiety? How much of the performance-to-performance variation is explained by the type of performance? How did the intercept variance change? Does this surprise you?

(j) Which is larger, the variation in the intercepts or in the slopes? What does that tell you in context?

(k) Interpret the slope/intercept correlation in this context. Are the effects “fanning in” or “fanning out”? Why does this relationship make sense in context.

(l) Write out the (new) model (by level and then composite) that uses the type of performance (large ensemble or not) with random intercepts and slopes that depend on type of instrument (orchestral or not).

Fit the model for (l): Make a binary variable for orchestra. Include orchestra, largeperformance, and their interaction as fixed effects, and then random intercepts and random slopes for ensemble, nested in performer ID).

(m) Interpret the interaction between performance type and orchestra type in context.

(n) How much variability in the intercepts does including type of instrument explain? How much variability in the slopes? How did the estimate of within group variation change?

(o) Summarize what you learn about the effect of type of instrument on the intercepts and the slopes.

(p) Maybe with the interaction between performance type and instrument type we no longer need the random slopes... Investigate this. Document how you did so (both the model equations and the R code).