# Stat 414 - Day 21
# Random Slopes cont.

---

**Last Time**

- We fit a "random coefficient model" $y_{ij} = \beta_{0j} + \beta_{1j} x_{1ij} + \epsilon_{ij}$ where $\beta_{0j} = \beta_{00} + u_{0j}$ with $u_{0j} \sim N(0, \tau_0^2)$ and $\beta_{1j} = \beta_{10} + u_{1j}$ with $u_{1j} \sim N(0, \tau_1^2)$.
- This adds a variance component for the slopes ($\tau_1^2$) as well as a covariance between the random slopes and random intercepts ($\tau_{01}$)

---

**Example 1:** Continuing with our beach data. Richness varies by Beach, so including Beach in the model (as fixed or random effects) should give us more accurate standard errors. But we see some patterns (related to the beaches) in the residuals and this tells us that we might be able to improve the fit between the model and the data. In this case, we tried a random slopes model, and the variation in the slopes ($\tau_1^2$) was statistically significant.

```
Random effects:
 Groups   Name        Variance Std.Dev. Corr
 Beach    (Intercept) 10.949   3.309
          NAP          2.502   1.582    -1.00
 Residual              7.174   2.678
Number of obs: 45, groups:  Beach, 9
```

```
model1: Richness ~ NAP + (1 | Beach)
model2: Richness ~ NAP + (1 + NAP | Beach)
       npar    AIC    BIC  logLik deviance Chisq Df Pr(>Chisq)
model1    4 249.83 257.06 -120.92   241.83
model2    6 246.66 257.50 -117.33   234.66 7.173  2    0.02769 *
```

(a) Between what two values do we expect 95% of the slopes to fall?

But we notice this gives us a second new parameter as well.

(b) Interpret the strong, negative correlation ($corr(u_{0j}, u_{1j}) = \hat{\tau}_{01}$ = -0.99) between the slopes and intercepts in context: beaches with larger intercepts (meaning what?) tend to have what kinds of slopes (meaning what?).

Only one beach has Exposure = 8, so we are going to combine that with Exposure 10 make this a binary variable (the rest are exposure 11).

```
rikzdata$ExposureCat = (rikzdata$Exposure > 10)
```

(c) Now consider adding Exposure, a Level 2 variable, to the model. What do you expect to change in the model?

Explore how exposure might be related to the intercepts and/or the slopes.

(d) Is Exposure "positively" or "negatively" related to the intercepts? How about the slopes? What are the implications to the model?

Fit the model including Exposure

```
model3 = lmer(Richness ~ NAP + ExposureCat + (1 + NAP | Beach), data  = rikzdata)
```

(e) Interpret the slope coefficient of Exposure in this model.

(f) Is the Exposure variable significant? Do you see a substantial improvement in the fit of the model? How do the variance components change/what has been the main impact?

Visualize the reduction in the Level 2 variability:

(g) Why didn't including the Exposure variable explain much variation in the slopes?

(h) To expand our model to allow for Exposure to explain variation in slopes, write the Level 1 and Level 2 equations, including Exposure in both Level 2 equations.

(i) Now make the composite equation, what happens? (*Hint*: What is the expression for the intercepts and what is the expression for the slopes?)

So what do we tell R? How many parameters will this add to the model?

```
model4 = lmer(Richness ~ NAP*ExposureCat + (1 + NAP | Beach), data  = rikzdata)
```

(i) How many parameters did we add to the model? What is the estimate for that parameter? Is it statistically significant? How are you deciding?

(j) Does Model 4 appear to be a better fitting model? How are you deciding? (*Hint*: What measures of "model fit" do we have that we haven't done much with recently?)

**Notes:**

- With a level 2 variable, we are thinking of the level 1 intercepts and slopes as "outcomes" and then running a regression model to explain that variation

- "In cases where the explanation of the random effects works extremely well, one may end up with models with no random effects at level two… random intercepts, slope have zero variance.. Omitted.. The resulting model may be analyzed just as well with OLS regression analysis… within group dependence has been fully explained by the available explanatory variables/interactions (no more dependence in the residuals)."