Stat 462 Lab 10 solutions

March 24, 2014

```
load( "reading.RData" )
library( lme4 )
```

1. Which of these variables would you consider to be fixed effects, and which would you consider to be random effects? Minority, Sex, and SES are all probably best treated as fixed effects. School is probably a random effect because we are not interested in these schools *per se*, but rather in any given school whether included in the dataset or not.

2. What is a situation in which we would treat School as a fixed effect? If we were interested in these particular schools (for example, if these were all of the schools in the country), then we might treat them as a fixed effect.

3. What is a situation in which we would treat School as a random effect? If these schools were a random sample of all possible schools, then we would treat them as a random effect because we would want to generalize to other students at other schools not in the sample.

4. Write down the model that has Score as the response and includes all covariates except SESschool. Treat School as a random effect. There are many ways to write down the same model. Here is one:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \alpha_{j[i]} + \epsilon_i$$

where

- y_i is the reading score for student i
- β_0 is the overall intercept
- β_1 is the coefficient for the minority effect, x_{1i} is the minority code for student i
- β_2 is the coefficient for the sex effect, x_{2i} is the sex code for student i
- β_3 is the coefficient for the SES slope, x_{3i} is the SES for student i
- $\alpha_{j[i]}$ is the random effect of school j corresponding to student i
- ϵ_i is the random effect of student i
- $\alpha_j \sim N(0, \sigma_\alpha^2)$ for all schools j
- $\epsilon_i \sim N(0, \sigma^2)$ for all students i

5. Analyze these data using the model you specified in (4.). Are there statistically significant effects of Minority, Sex, and SES? Compute an estimate of the proportion of variation attributable to differences between schools given the covariates. Using lmer to fit this mixed-effects model, we get

```
> read.fit1 <- lmer( Score ~ Minority + Sex + SES + (1|School), data = reading )
> ( read.sum1 <- summary(read.fit1) )
Linear mixed model fit by REML ['lmerMod']
Formula: Score ~ Minority + Sex + SES + (1 | School)
   Data: reading

REML criterion at convergence: 46394.4

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.2427 -0.7216  0.0342  0.7620  2.8631

Random effects:
 Groups   Name        Variance Std.Dev.
 School   (Intercept)  3.674   1.917
 Residual             35.909   5.992
Number of obs: 7185, groups: School, 160

Fixed effects:
            Estimate Std. Error t value
(Intercept)  14.1145     0.1970   71.64
MinorityYes  -2.9615     0.2058  -14.39
SexM         -1.2298     0.1627   -7.56
SES           2.0894     0.1057   19.77

Correlation of Fixed Effects:
            (Intr) MnrtyY SexM
MinorityYes -0.292
SexM        -0.434  0.014
SES         -0.078  0.195  0.058
```

Minority, Sex, and SES are all significant. The proportion of variation attributable to differences between schools is $3.674/(3.674 + 35.909) = 9.3\%$.

6. Write down the model that has Score as the response and includes all covariates except SESschool. Include separate intercepts and slopes for each school, treating the schools as fixed effects. Here is one description of this model

$$y_i = \alpha_{j[i]} + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{3j[i]} x_{3i} + \epsilon_i$$

where

- y_i is the reading score for student i
- $\alpha_{j[i]}$ is the intercept for school j associated with student i
- β_1 is the coefficient for the minority effect, x_{1i} is the minority code for student i
- β_2 is the coefficient for the sex effect, x_{2i} is the sex code for student i

- $\beta_{3j[i]}$ is the coefficient for the SES slope for school j associated with student i, and x_{3i} is the SES for student i
- ϵ_i is the random effect of student i
- $\epsilon_i \sim N(0, \sigma^2)$ for all students i

```
> read.fit2 <- lm( Score ~ Minority + Sex + SES:School + School , data = reading )
> read.preds <- predict( read.fit2 , data.frame( Minority = "Yes", Sex = "F"
+                   , SES = 1, School = unique( reading$School ) ) )
> mean( read.preds )
[1] 12.99457
```

This prediction only applies to these schools due to the fixed effects treatment. Also, this prediction assumes that it is equally likely that a student will go to any school in the dataset.

8. Write down the model that has Score as the response and includes all covariates except SESschool. Include random intercepts and slopes for each school. Here is one description of this model:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \alpha_{j[i]} + \alpha_{3j[i]} x_{3i} + \epsilon_i$$

where

- y_i is the reading score for student i
- β_0 is the overall intercept
- β_1 is the coefficient for the minority effect, x_{1i} is the minority code for student i
- β_2 is the coefficient for the sex effect, x_{2i} is the sex code for student i
- β_3 is the coefficient for the SES slope, x_{3i} is the SES for student i
- $\alpha_{j[i]}$ is the random effect of school j corresponding to student i
- $\alpha_{3j[i]}$ is the random slope for school j corresponding to student i
- ϵ_i is the random effect of student i
- $\alpha_j \sim N(0, \sigma_\alpha^2)$ for all schools j
- $\alpha_{3j} \sim N(0, \sigma_{\alpha 3}^2)$ for all schools j
- $\epsilon_i \sim N(0, \sigma^2)$ for all students i

9. Analyze these data using the model you specified in (8.). Are there statistically significant effects of Minority, Sex, and SES? Predict the expected Score for a female student with Minority status and SES = 1. For which school does this prediction apply? Using lmer() to fit this model, we get

```
> read.fit3 <- lmer( Score ~ Minority + Sex + SES + (1+SES|School), data = reading )
> ( read.sum3 <- summary(read.fit3) )
Linear mixed model fit by REML ['lmerMod']
Formula: Score ~ Minority + Sex + SES + (1 + SES | School)
   Data: reading

REML criterion at convergence: 46389.1

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.2025 -0.7196  0.0373  0.7638  2.8871

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 School   (Intercept)  3.6600  1.9131
          SES          0.2599  0.5098   -0.43
 Residual             35.7878  5.9823
Number of obs: 7185, groups: School, 160

Fixed effects:
            Estimate Std. Error t value
(Intercept)  14.1463     0.1969   71.83
MinorityYes  -2.9984     0.2067  -14.50
SexM         -1.2177     0.1625   -7.50
SES           2.0957     0.1136   18.44

Correlation of Fixed Effects:
            (Intr) MnrtyY SexM
MinorityYes -0.289
SexM        -0.433  0.012
SES         -0.196  0.177  0.052
> fixef(read.fit3)%*%c(1,1,0,1)
         [,1]
[1,] 13.24361
```

This prediction applies to a “future” school chosen at random from the larger population of schools, assuming that the schools in the dataset were sampled at random from this larger population as well.