Edpsych/Psych/Stat 587
Spring 2021
C.J. Anderson

R Computer Lab Session 1
Thursday February 11, 2021

The goals of this session are to learn how to

1. Use R commands to compute centered variables.
2. Use the lme4 package, and some basic R commands, to estimate the following models
   - Fixed effects ANOVA
   - ANCOVA
   - Random (and mixed) effects ANOVA
   - Random intercept (HLM/multilevel/random coefficients/mixed effects) models with micro and macro level variables
3. Understand the R output.

General R Programming Tips:

- Do one task at a time and check the results, rather than trying to write all of the R script for all of the models at once.
- After executing commands, always check the console to determine whether an error has occurred.
- R **is case-sensitive** so be aware of this as you name variables and write your script. I also find it helpful to check and make sure all parenthesis are opened and closed appropriately.
- If you want to clear all of the previous commands in the console, you may go to Edit → Clear Console (or CTRL + L on PC)
- You might want to change the working directory to where the data are stored on your machine. To do this using the taskbar, go to File → Working Directory. Alternatively, you can use the `setwd` command.

- *Save* your R script and any output graphics you wrote/created in this lab session. They will be *useful* later in the course and they are **necessary** for completing the first homework assignment.

The data we're using for this computer lab (and accompanying homework) come from the Third International Mathematics and Science Study (TIMSS). The study was funded by the National Science Foundation in conjunction with the National Center for Educational Statistics. The web-site where it downloaded the data was `http://ustimss.msu.edu/`. We are ignoring weights, missing values, etc. The Table below shows the correspondence between the labels given in TIMSS and our data set, as well as a bit more information about them.

Name of Variable in TIMSS	SAS data	Values	Desription
IDSCHOOL	IDSCHOOL	10–263 by 1	School ID
IDSTUD	IDSTUD		Student ID
IDGRADE	GRADE	3 or 4	Student's grade in school
ASSNRSC	SCIENCE	range 103.4 – 185	Science score based on ability estimates using the RASCH IRT model.
ASMNRSC	MATH	range 104.3 – 189	Math scores based on ability estimates using the RASCH model.
ITSEX	GENDER	girl, boy	Student's gender
ACBGST01	GEN_SHORT	none, a little, some, a lot	General school shortages of instructional materials
—	SHORTAGES	0,1,2,3	Re-coded GEN_SHORT as a numerical variable.

Do the following:

1. Install and call these three packages: `lme4`, `lmerTest`, and `texreg`.

2. If you wish to use the `icc` function that I wrote, you can download it from the course web-site under R computer lab 1 or from under lecture notes on "Random Intercept Model".

3. Load the data into R by opening up the data set and R program available on the course web-site (this will act as a template for the lab). The name of the

data set is "lab1". Be sure to check the first few rows of the data set and column names.

4. Use the built-in functionality in R to use compute some means.

$$\begin{aligned} \texttt{ocmath} &= \text{math} - \text{overall mean of math} \\ \texttt{school.mean.math} &= \text{School mean math score} \\ \texttt{group.center.math} &= \texttt{math} - \texttt{school.mean.math} \end{aligned}$$

   Once these means have been created, merged them with the main or master data set.

5. Fit the following models where SCIENCE scores based on the RASCH model is the response variable (i.e., Y_{ij} is SCIENCE score for individual i in school j), schools are the group/macro units, and students are the individual/micro units.

   Using `lmer` fit each of the following models using maximum likelihood estimation:

   (a) Fixed effects ANOVA.

   (b) ANCOVA with `ocmath` as a covariate.

   (c) Random effects ANOVA.

   (d) Random intercept model with no explanatory variables (i.e., empty/null/baseline model).

   (e) Random intercept model with `math` as an explanatory variable.

   (f) Random intercept model with `ocmath` as an explanatory variable.

   (g) Random intercept model with `school.mean.math` as an explanatory variable.

   (h) Random intercept model with `group.center.math` as an explanatory variable.

   (i) Random intercept model with `school.mean.math` and `group.center.math` as explanatory variables.

   (j) Random intercept model with `group.center.math, school.mean.math, gender`, and `grade` as explanatory variables.

   (k) Random intercept model with `group.center.math, school.mean.math, gender, grade` and `gen_short` as explanatory variables.

(l) Random intercept model with `group.center.math, school.mean.math, gender, grade` and `shortages` as explanatory variables.

(m) Fit any model(s) that you think may be interesting, useful, or better than those requested above.