Edpsy/Psych/Stat 587
Spring 2019
C.J. Anderson

SAS Computer Lab Session 1
Thursday January 29, 2019

The goals of this session are to learn how to

1. Use PROC MEANS to compute centered variables.
2. Use PROC MIXED to fit the following models
   - Fixed effects ANOVA
   - ANCOVA
   - Random (and mixed) effects ANOVA
   - Random intercept (HLM) models with micro and macro level variables
3. Understand what's what on the computer output.

General SAS Programing Tips:

- Do one task at a time and check the results, rather than trying to write all of the SAS commands for all the models at once.
- After you run a SAS program, always check the log file for error and/or warning messages.
- A very common error that may give you odd error messages is forgetting ";" at the end of a command or to close quotes.
- From time to time you may want to clear your OUTPUT and/or LOG windows. To do this, activate the OUTPUT or LOG window (i.e., put cursor in window and press mouse key) then press "Edit" $\rightarrow$ "Clear all" on the tool bar at the top of the screen.
- If you accidently erase your last SAS program without saving it, you can get it back by going to the program/editor window and press "Run" $\rightarrow$ "Recall last submit".
- Save your final SAS input and output files that 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 that we will be 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 I downloaded the data was http://ustimss.msu.edu/. We are ignoring weights, missing values, etc.

Data were collected from 45 different countries; however, we only will be using data from the US. The variables that we will be using include:

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. Create a SAS data set by running the SAS program: "TIMSS_Lab1_Data2010.sas." This SAS program is available from the course web-site. The name of the data set will be "lab1".

2. Using PROC MEANS and the 'lab1' data set, create the following centered variables:

$$\begin{aligned} \texttt{OCmath} &= \text{MATH} - \text{overall mean of MATH} \\ \texttt{grpMmath} &= \text{School mean math score} \\ \texttt{grpCmath} &= \text{MATH} - \texttt{grpMmath} = \text{school centered math score} \end{aligned}$$

   To compute means overall and for each school, see lecture notes and/or introduction to SAS notes. To create mean centered variables, see lectures notes on SAS.

3. 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 SAS MIXED 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/basline 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 `grpCmath` as an explanatory variable.

   (h) Random intercept model with `grpMmath` as an explanatory variable.

   (i) Random intercept model with `grpCmath` and `grpMmath` as explanatory variables.

   (j) Random intercept model with `grpCmath`, `grpMmath`, `gender`, and `grade` as explanatory variables.

   (k) Random intercept model with `grpCmath`, `grpMmath`, `gender`, `grade`, and `gen_short` as explanatory variables.

   (l) Random intercept model with `grpCmath`, `grpMmath`, `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.