Simple, Marginal, and Interaction Effects in General Linear Models: Part 1

PSYC 943 (930): Fundamentals of Multivariate Modeling

Lecture 2: August 24, 2012

Today's Class

- Centering and Coding Predictors
- Interpreting Parameters in the Model for the Means
- Main Effects Within Interactions
- Welcome (Back) to SAS Syntax
- GLM Example 1: "Regression" vs. "ANOVA"

CENTERING AND CODING PREDICTORS

The Two Sides of a Model

$$Y_p = \boxed{\beta_0 + \beta_1 X_p + \beta_2 Z_p + \beta_3 X_p Z_p} + \boxed{e_p}$$

- <u>**Model for the Means (Predicted Values):**</u>
  - Each person's expected (predicted) outcome is a function of his/her values on x and z (and their interaction), each measured once per person
  - **Estimated parameters are called fixed effects** (here, β_0, β_1, β_2, and β_3); although they have a sampling distribution, they are not random variables
  - The number of fixed effects will show up in formulas as ***k*** (so ***k* = *4*** here)

- <u>**Model for the Variance:**</u>
  - $e_p \sim N(0, \sigma_e^2)$ → ONE residual (unexplained) deviation
  - e_p has a mean of 0 with some estimated constant variance σ_e^2, is normally distributed, is unrelated to x and z, and is unrelated across people (across all observations, just people here)
  - **Estimated parameter is the residual variance only** (in the model above)

For now we focus entirely on the fixed effects in the model for the means…

Representing the Effects of Predictor Variables

- From now on, we will think carefully about exactly **<u>how</u>** the **predictor variables** are entered into the **model for the means** (i.e., by which a predicted outcome is created for each person)

- Why don't people always care? Because the scale of predictors:
  - Does NOT affect the amount of outcome variance accounted for (R^2)
  - Does NOT affect the outcomes values predicted by the model for the means (so long as the same predictor fixed effects are included)

- Why should this matter to us?
  - ***Because the Intercept = expected outcome value when X = 0***
  - Can end up with nonsense values for intercept if X = 0 isn't in the data
  - We will almost always need to deliberately **adjust the scale of the predictor variables** so that they have 0 values that could be observed in our data
  - Is much bigger deal in models with random effects (MLM) or GLM once interactions are included (… stay tuned)

Adjusting the Scale of Predictor Variables

- For **continuous** (quantitative) predictors, **<u>we</u>** will make the intercept interpretable by **centering**:
  - **Centering** = subtract a constant from each person's variable value so that **the 0 value** falls within the range of the new centered predictor variable
  - Typical → Center around predictor's mean: $Centered\ X_1 = X_1 - \overline{X_1}$
    - Intercept is then expected outcome for "average X_1 person"
  - Better → Center around meaningful constant *C*: $Centered\ X_1 = X_1 - C$
    - Intercept is then expected outcome for person with that constant (even 0 may be ok)

- For **categorical** (grouping) predictors, **<u>either we or the program</u>** will make the intercept interpretable by **creating a reference group**:
  - **Reference group** is given a 0 value on all predictor variables created from the original grouping variable, such that the intercept is the expected outcome for that reference group specifically
  - Accomplished via "dummy coding" or "reference group coding"
    → Two-group example using *Gender*: 0 = Men, 1 = Women
    (or 0 = Women, 1 = Men)

Adjusting the Scale of Predictor Variables

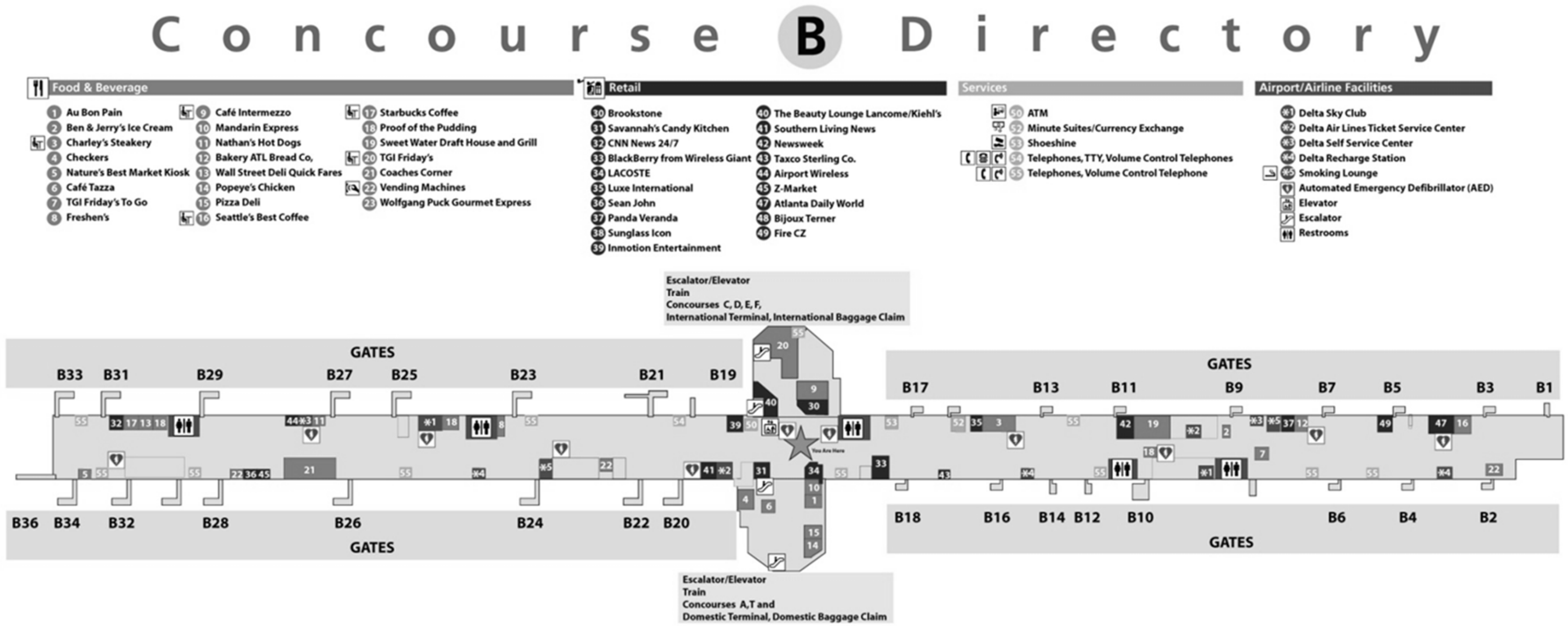
- For more than two groups, need: ***dummy codes = #groups − 1***
  - Four-group example: Control, Treatment1, Treatment2, Treatment3
  - Variables: *d1*= 0, 1, 0, 0 → difference between Control and T1
    d2= 0, 0, 1, 0 → difference between Control and T2
    d3= 0, 0, 0, 1 → difference between Control and T3

  Done for you in GLM software ☺

- Potential pit-falls:
  - All predictors representing the effect of group (e.g., d1, d2, d3) MUST be in the model at the same time for these specific interpretations to be correct!
  - Model parameters resulting from these dummy codes will not *directly* tell you about differences among non-reference groups (…but stay tuned)

- Other examples of things people do to categorical predictors:
  - "Contrast/effect coding"→ *Gender*: −0.5 = Men, 0.5 = Women (or vice-versa)
  - Test other contrasts among multiple groups → four-group example above:
    Variable: *contrast1*= −1, 0.33, 0.33, 0.34 → Control vs. Any Treatment?

Why the Intercept β_0 *Should* Be Meaningful…

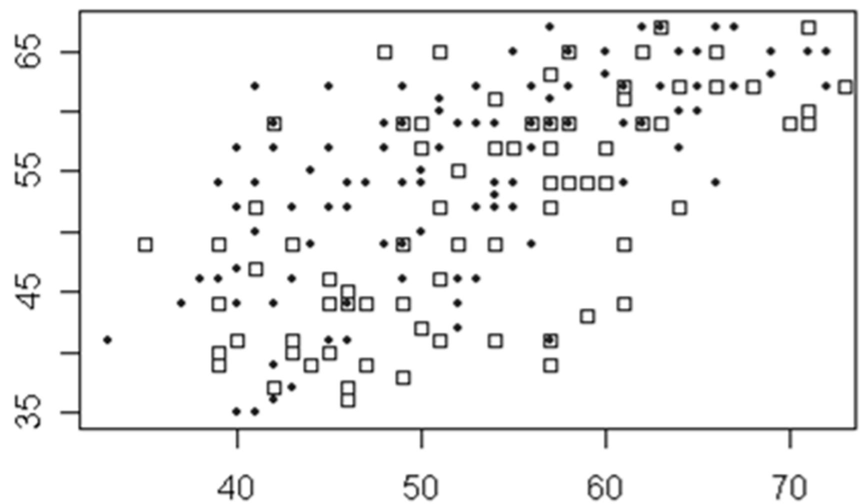


This is a very detailed map…

But what do we need to know to be able to use the map at all?

What the Intercept β_0 Should Mean to You…

The model for the means will describe what happens to the predicted outcome Y
"as X increases" or
"as Z increases"
and so forth…



But you won't know what Y is actually supposed to be unless you know where the predictor variables are starting from!

Therefore, the **intercept** is the "**YOU ARE HERE**" sign in the map of your data… so it should be somewhere in the map*!

* There is no *wrong* way to center (or not), only *weird*…

What the Intercept *WILL* Mean to You…

INTERPRETING PARAMETERS IN THE MODEL FOR THE MEANS

Interpreting Parameters in the Model for the Means

- Last time we saw that each regression slope (or more generally, any estimated **fixed effect**) had 4 relevant pieces of output:
  - **Estimate** = best guess for the fixed effect from our data
  - **Standard Error** = precision of fixed effect estimate (quality of best guess)
  - ***T*-Value** = Estimate / Standard Error → Wald test
  - ***p*-value** = probability that fixed effect estimate is ≠ 0
    - Compare Wald test *T*-value to critical *T*-value at chosen level of significance

- Estimate of β_X for the slope of X in a one-predictor model:

After 1/N-k cancels, is called:

$$\beta_X = \frac{\frac{1}{\mathrm{N-k}}\sum_{p=1}^{N}(X_p-\bar{X})(Y_p-\bar{Y})}{\frac{1}{\mathrm{N-k}}\sum_{p=1}^{N}(X_p-\bar{X})(X_p-\bar{X})} = \frac{covariance\ of\ X\ and\ Y}{covariance\ of\ X\ and\ X} = \frac{sum\ of\ cross{-}products}{sum\ of\ squared\ Xs}$$

After Cov(X) cancels, β_x is in units of y per units of x

- When more than one predictor is included, β_x turns into:
  "unique" covariance of X and Y / "unique" covariance of X and X

Stay tuned for how ☺

Interpreting Parameters in the Model for the Means

- Standard Error (SE) for estimate β_X in a one-predictor model (remember, SE is like the SD of the estimated parameter):

$$SE_{\beta_X} = \sqrt{\frac{\frac{1}{N-k}\sum_{p=1}^{N}\left(Y_p - \hat{Y}_p\right)^2}{\frac{1}{N-k}\sum_{p=1}^{N}\left(X_p - \bar{X}\right)^2 * (N-k)}} = \sqrt{\frac{residual\ variance\ of\ Y}{variance\ of\ X * (N-k)}}$$

SE is also in units of Y / units of X

- When more than one predictor is included, SE turns into:

$$SE_{\beta_X} = \sqrt{\frac{residual\ variance\ of\ Y}{Var(X) * \left(1 - R_X^2\right) * (N-k)}}$$

R_X^2 = X variance accounted for by other predictors, so $1 - R_X^2$ = unique X variance

- So all things being equal, SE is smaller when:
  - More of the outcome variance has been reduced (better model)
    - So fixed effects can become significant later if R^2 is higher then
  - The predictor has less covariance with other predictors (less collinearity)
    - Best case scenario: X is uncorrelated with all other predictors

- If SE is smaller → *T*-value is smaller → *p*-value is smaller

MAIN EFFECTS WITHIN INTERACTIONS

Interactions: $Y_p = \beta_0 + \beta_1 X_p + \beta_2 Z_p + \beta_3 X_p Z_p + e_p$

- **Interaction = Moderation:** the effect of a predictor depends on the value of the interacting predictor
  - Either predictor can be “the moderator” (interpretive distinction only)

- Interactions can always be evaluated for any combination of categorical and continuous predictors, although traditionally…
  - In “ANOVA”: By default, all possible interactions are estimated
    - Software does this for you; oddly enough, nonsignificant interactions usually still are kept in the model (even if only significant interactions are interpreted)
  - In “ANCOVA”: Continuous predictors (“covariates”) do not get to be part of interaction terms → make the “homogeneity of regression assumption”
    - There is no reason to assume this – it is a testable hypothesis!
  - In “Regression”: No default – effects of predictors are as you specify them
    - Requires most thought, but gets annoying because in regression programs you usually have to manually create the interaction as an observed variable:
    - e.g., XZinteraction = centeredX * centeredZ

Done for you in GLM software ☺

Main Effects of Predictors within Interactions in GLM

- Main effects of predictors within interactions should remain in the model regardless of whether or not they are significant
  - An interaction is an over-additive (enhancing) or under-additive (dampening) effect, so *what it is additive to* must be included

- The role of a two-way interaction is to adjust its main effects…
- However, the idea of a "main effect" no longer applies… each main effect is ***conditional*** on the interacting predictor = 0

- e.g., Model of Y = W, X, Z, X*Z:
  - The effect of W is still a "main effect" because it is not part of an interaction
  - The effect of X is now the conditional main effect of X *specifically when Z=0*
  - The effect of Z is now the conditional main effect of Z *specifically when X=0*

- The trick is keeping track of what 0 means for every interacting predictor, which depends on the way each predictor is being represented, as determined by you, or by the software without you!

Interactions: Why 0 Matters

- Y = Student achievement (GPA as percentage grade out of 100)
  X = Parent attitudes about education (measured on 1-5 scale)
  Z = Father's education level (measured in years of education)

- **Model:** $GPA_p = \beta_0 + \beta_1 * Att_p + \beta_2 * Ed_p + \beta_3 * Att_p * Ed_p + e_p$
  $GPA_p = 30 + 2 * Att_p + 1 * Ed_p + 0.5 * Att_p * Ed_p + e_p$

- **Interpret β_0:**
- **Interpret β_1:**
- **Interpret β_2:**
- **Interpret β_3: Attitude as Moderator:**
  Education as Moderator:

- **Predicted GPA** for **attitude of 3** and **Ed of 12**?
  $66 = 30 + 2*(3) + 1*(12) + 0.5*(3)*(12)$

Interactions: Why 0 Matters

- Y = Student achievement (GPA as percentage grade out of 100)
  X = Parent attitudes about education (**still measured on 1-5 scale**)
  Z = Father's education level (**0 = 12 years of education**)

- **Model:** $GPA_p = \beta_0 + \beta_1{*}Att_p + \beta_2{*}Ed_p + \beta_3{*}Att_p{*}Ed_p + e_p$
- **Old Equation:** $GPA_p = 30 + 2{*}Att_p + 1{*}Ed_p - 0 + 0.5{*}Att_p{*}Ed_p - 0 + e_p$
- **New Equation:** $GPA_p = 42 + 8{*}Att_p + 1{*}Ed_p - 12 + 0.5{*}Att_p{*}Ed_p - 12 + e_p$

- **Why did β_0 change?**
- **Why did β_1 change?**
- **Why did β_2 stay the same?**
- **Why did β_3 stay the same?**

- Which fixed effects would have changed if we centered attitudes at 3 but left education uncentered at 0 instead?

Getting the Model to Tell Us What We Want…

- Model equation already says what Y (the intercept) should be…
  Original Model: $\mathbf{GPA_p = \beta_0 + \beta_1{*}Att_p + \beta_2{*}Ed_p + \beta_3{*}Att_p{*}Ed_p + e_p}$
  $$\mathbf{GPA_p = 30 + 2{*}Att_p + 1{*}Ed_p + 0.5{*}Att_p{*}Ed_p + e_p}$$
  - The intercept is always conditional on when predictors = 0

- But the model also tells us any conditional main effect for any combination of values for the model predictors
  - Using intuition: **Main Effect = what it is + what *modifies* it**
  - Using calculus (first derivative of model with respect to each effect):
    $$\textbf{Effect of Attitudes} = \beta_1 + \beta_3{*}Ed_p = 2 + 0.5{*}Ed_p$$
    $$\textbf{Effect of Education} = \beta_2 + \beta_3{*}Att_p = 1 + 0.5{*}Att_p$$
    $$\textbf{Effect of Attitudes*Education} = \beta_3 = 0.5$$
  - Now we can use these new equations to determine what the conditional main effects would be given other predictor values besides true 0…
    …let's do so for a reference point of attitude = 3 and education = 12

Getting the Model to Tell Us What We Want…

Old Equation using uncentered predictors:

$$\mathbf{GPA_p = \beta_0 + \beta_1{*}Att_p + \beta_2{*}Ed_p + \beta_3{*}Att_p{*}Ed_p + e_p}$$

$$\mathbf{GPA_p = 30 + 2{*}Att_p + 1{*}Ed_p + 0.5{*}Att_p{*}Ed_p + e_p}$$

New equation using centered predictors:

$$\mathbf{GPA_p = \_\_\_ + \_\_\_{*}Att_p{-}3 + \_\_\_{*}Ed_p{-}12 + \_\_\_{*}Att_p{-}3{*}Ed_p{-}12 + e_p}$$

- **β_0: expected value of GPA when Att_p=3 and Ed_p=12**
  - **β_0 =**
- **β_1: effect of Attitudes**
  - **β_1 =**
- **β_2: effect of Education**
  - **β_2 =**
- **β_3: two-way interaction of Attitudes and Education:**
  - **β_3:**

Testing the Significance of Model-Implied Fixed Effects

- We now know how to calculate any conditional main effect:
  Effect of interest = what it is + what *modifies* it
  Effect of Attitudes = $\beta_1 + \beta_3$*Ed for example...

- But if we want to test whether that new effect is ≠ 0, we also need its **standard error** (**SE** needed to get Wald test *T*-value → *p*-value)

- Even if the conditional main effect is not *directly* given by the model, its estimate and SE are still *implied* by the model

- **3 options** to get the new conditional main effect estimate and SE (in order of least to most annoying):

1. **Ask the software to give it to you** using your original model (e.g., ESTIMATE in SAS, TEST in SPSS, NEW in Mplus)

Testing the Significance of Model-Implied Fixed Effects

2. **Re-center your predictors** to the interacting value of interest (e.g., make attitudes=3 the new 0 for attitudes) and **re-estimate** your model; repeat as needed for each value of interest

3. **Hand calculations** (what the program is doing for you in option #1)

For example: **Effect of Attitudes = $\beta_1 + \beta_3$*Ed**

- SE^2 = sampling variance of estimate → e.g., $Var(\beta_1) = SE_{\beta 1}^2$
- $\mathbf{SE_{\beta 1}^2 = Var(\beta_1) + Var(\beta_3)*Ed + 2Cov(\beta_1,\beta_3)*Ed}$

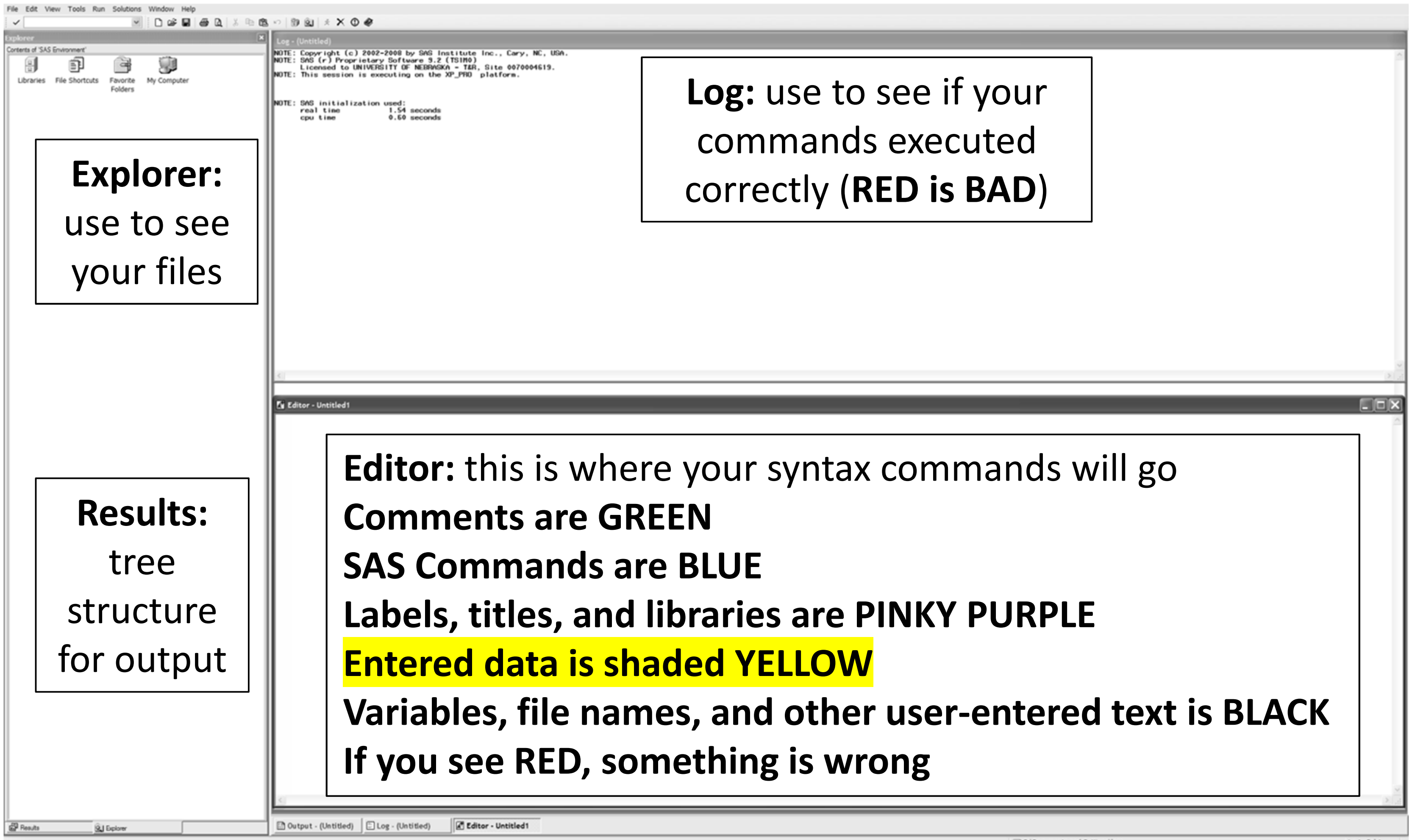
Stay tuned for why ☺

  - Values come from "asymptotic (sampling) covariance matrix"
  - Variance of a sum of terms always includes covariance among them
  - Here, this is because what each main effect estimate could be is related to what the other main effect estimates could be
  - Note that if a main effect is unconditional, its $SE^2 = Var(\beta)$ only

WELCOME (BACK) TO SAS SYNTAX

Welcome to SAS! (or Welcome Back to SAS!)

SAS opens with 5 different windows by default



Explorer: use to see your files

Log: use to see if your commands executed correctly (**RED is BAD**)

Results: tree structure for output

Editor: this is where your syntax commands will go

Comments are GREEN

SAS Commands are BLUE

Labels, titles, and libraries are PINKY PURPLE

Entered data is shaded YELLOW

Variables, file names, and other user-entered text is BLACK

If you see RED, something is wrong

Output: (not shown here) where your model results will be

Things to Know about SAS Syntax

- It's awesome, and can be used to automate nearly any task!
- All SAS commands in end a semi-colon
- **Only code in pinky-purple is line-, case-, or space-sensitive**
- Homework data should always be imported into the 'work' library
  - Is temporary directory
  - So if something goes wrong, you can easily re-create the data file
  - You do not need to save your SAS dataset, just the syntax file (.sas)
- Use the colors to help you – if syntax is the wrong color (or if it is **red**), that means something is wrong
  - Missing quotes? Missing semi-colon? Missing parentheses?
  - Are all variable and dataset names spelled correctly?
  - Did nothing happen? You are missing "**RUN;**" to execute the command
  - Always check the log AND the data to see if something worked correctly

Two Types of SAS Commands

- PROC : stands for "procedure"
  - e.g., PROC IMPORT, PROC REG, PROC GLM, PROC MIXED
  - Used to do something (import non-SAS data, run a model)
  - Each will be explained specifically when relevant

- DATA : is used to do something to existing SAS dataset, has 3 necessary commands:

```
DATA place.NameOfDatafileToBeCreated;

     SET place.NameOfDatafileCreatedFrom;

     *** variable transformations go here;

RUN;
```

Change the black text as needed to refer to your data…

This is where you'd write your code…

Today's Example: GLM as "Regression" vs. "ANOVA"

- Study examining effect of new instruction method (where New: 0=Old, 1=New) on test performance (% correct) in college freshmen vs. seniors (where Senior: 0=Freshmen, 1=Senior), *n* = 25 per group
- $\boldsymbol{Test_p = \beta_0 + \beta_1 Senior_p + \beta_2 New_p + \beta_3 Senior_p New_p + e_p}$

Test Mean *(SD)*, $\left[SE = \frac{SD}{\sqrt{n}}\right]$	**Freshmen**	**Seniors**	**Marginal (Mean)**
Old Method	80.20 *(2.60),[0.52]*	82.36 *(2.92),[0.59]*	81.28 *(2.95),[0.42]*
New Method	87.96 *(2.24),[0.45]*	87.08 *(2.90),[0.58]*	87.52 *(2.60),[0.37]*
Marginal (Mean)	84.08 *(4.60),[0.65]*	84.72 *(3.74),[0.53]*	84.40 *(4.18),[0.42]*

Importing and Describing Data for Example #1

```
* Location for files to be saved - CHANGE PATH;
%LET examples=F:\Example Data\943; LIBNAME examples "&examples.";
* Read in data to work (temporary) library;
DATA work.example1; SET examples.example1; RUN;
* Send all results to an excel file;
ODS HTML FILE="&examples.\Example1.xls" STYLE=MINIMAL;
TITLE "Descriptive Statistics for Test Score by Group";
PROC MEANS MEAN STDDEV STDERR MIN MAX DATA=work.example1;
  CLASS Senior New;  * By CLASS variable unique combo;
  VAR Test;          * Do for VAR variables;
RUN; TITLE;
* Models would all go here…

* Close excel file so I can use it;
ODS HTML CLOSE;
```

Next, we estimate our models…

Analysis Variable : Test Test: Test Score Outcome							
Senior: Year (0=Freshman, 1=Senior)	**New: Instruction (0=Old, 1=New)**	**N Obs**	**Mean**	**Std Dev**	**Std Error**	**Minimum**	**Maximum**
0	**0**	**25**	80.20	2.60	0.52	75	86
	1	**25**	87.96	2.24	0.45	83	93
1	**0**	**25**	82.36	2.93	0.59	76	89
	1	**25**	87.08	2.90	0.58	81	92

SAS PROCs are hierarchical…

Hierarchy for General Models:

- PROC MIXED
- **PROC GLM**
- PROC Reg
- PROC ANOVA
- PROC T-Test

- We'll begin with PROC GLM, which can estimate every model below it using least squares.

- We'll eventually move to PROC MIXED, which can estimate every model below it using maximum likelihood (stay tuned).

- Although PROC Reg is subsumed by PROC GLM, only PROC Reg directly provides:
  - Standardized beta weights
  - R^2 change tests

We'll see code next time to do these in GLM if needed!

- Unfortunately, PROC Reg is annoying:
  - Must code categorical predictors manually
  - Must create interaction terms manually
  - No cell means; limited specific estimates possible

GLM EXAMPLE 1: “REGRESSION” VS. “ANOVA”

GLM via Dummy-Coding in "Regression"

```
TITLE "GLM via Dummy-Coded Regression";
PROC GLM DATA=work.example1;
* Model y = predictor effects;
  MODEL Test = Senior New Senior*New / SOLUTION;
* Get predicted test score per group;
  ESTIMATE "Intercept for Freshmen-Old" Intercept 1 Senior 0 New 0 Senior*New 0;
  ESTIMATE "Intercept for Freshmen-New" Intercept 1 Senior 0 New 1 Senior*New 0;
  ESTIMATE "Intercept for Senior-Old"   Intercept 1 Senior 1 New 0 Senior*New 0;
  ESTIMATE "Intercept for Senior-New"   Intercept 1 Senior 1 New 1 Senior*New 1;
RUN; QUIT; TITLE;
```

ESTIMATE requests **predicted outcomes from model for the means**:

$$\widehat{Test_p} = \beta_0 + \beta_1 Senior_p + \beta_2 New_p + \beta_3 Senior_p New_p$$

- Freshmen-Old: $Test_p = \beta_0 + \beta_1 0 + \beta_2 0 + \beta_3 0 * 0$
- Freshmen-New: $Test_p = \beta_0 + \beta_1 0 + \beta_2 1 + \beta_3 0 * 1$
- Senior-Old: $Test_p = \beta_0 + \beta_1 1 + \beta_2 0 + \beta_3 1 * 0$
- Senior-New: $Test_p = \beta_0 + \beta_1 1 + \beta_2 1 + \beta_3 1 * 1$

Dummy-Coded "Regression": Results

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	1041.44	347.15	48.26	<.0001
Error	96	690.56	7.19		
Corrected Total	99	1732.00			

F-Test of $R^2 > 0$

R-Square
0.601293

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Senior	1	58.32	58.32	**8.11**	0.0054
New	1	752.72	752.72	**104.64**	<.0001
Senior*New	1	57.76	57.76	**8.03**	0.0056

These **"omnibus" F-tests** tell us if each effect is significant. Because each effect df=1 and **because it's using our coding**, the results match the **fixed effects table below** ($F = T^2$).

This table was created by the **ESTIMATE** commands to get per-group intercepts (i.e., predicted outcomes).

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept for Freshmen-Old	80.20	0.54	149.51	<.0001
Intercept for Freshmen-New	87.96	0.54	163.98	<.0001
Intercept for Senior-Old	82.36	0.54	153.54	<.0001
Intercept for Senior-New	87.08	0.54	162.34	<.0001

This **fixed effects** table uses our coding. However, not all possible conditional main effects are provided...

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	80.20	0.54	149.51	<.0001
Senior	2.16	0.76	2.85	0.0054
New	7.76	0.76	10.23	<.0001
Senior*New	-3.04	1.07	-2.83	0.0056

Dummy-Coded "Regression": Mapping Results to Data

ESTIMATE commands table

Parameter	Estimate	Standard Error
Intercept for Freshmen-Old	80.20	0.54
Intercept for Freshmen-New	87.96	0.54
Intercept for Senior-Old	82.36	0.54
Intercept for Senior-New	87.08	0.54

FIXED EFFECTS table

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept (β0)	80.20	0.54	149.51	<.0001
Senior (β1)	2.16	0.76	2.85	0.0054
New (β2)	7.76	0.76	10.23	<.0001
Senior*New (β3)	-3.04	1.07	-2.83	0.0056

Test Mean [SE]	**Freshmen**	**Seniors**	**Marginal**
Old Method	β_0 80.20 *[0.52]*	β_1 82.36 *[0.59]*	81.28 *[0.42]*
New Method	β_2 87.96 *[0.45]*	β_3 87.08 *[0.58]*	87.52 *[0.37]*
Marginal	84.08 *[0.65]*	84.72 *[0.53]*	84.40 *[0.42]*

Dummy-Coded “Regression”: *Model-Implied* Main Effects

```
TITLE "GLM via Dummy-Coded Regression";
PROC GLM DATA=work.example1;
* Model y = predictor effects;
  MODEL Test = Senior New Senior*New / SOLUTION;
* Get all possible conditional main effects;
  ESTIMATE "Senior Effect: Old"    Senior 1 Senior*New 0;
  ESTIMATE "Senior Effect: New"    Senior 1 Senior*New 1;
  ESTIMATE "New Effect: Freshmen"  New 1    Senior*New 0;
  ESTIMATE "New Effect: Seniors"   New 1    Senior*New 1;
RUN; QUIT; TITLE;
```

ESTIMATE requests **conditional main effects from model for the means**:

Model for the Means: $\widehat{Test}_p = \beta_0 + \beta_1 Senior_p + \beta_2 New_p + \beta_3 Senior_p New_p$

Main Effect = what it is + what *modifies* it

- Senior Effect for Old Method: $\beta_1 + \beta_3 * 0$
- Senior Effect for New Method: $\beta_1 + \beta_3 * 1$
- New Method Effect for Freshmen: $\beta_2 + \beta_3 * 0$
- New Method Effect for Seniors: $\beta_2 + \beta_3 * 1$

Dummy-Coded "Regression": *Model-Implied* Main Effects

ESTIMATE commands table

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Senior Effect: Old	2.16	0.76	2.85	0.0054
Senior Effect: New	-0.88	0.76	-1.16	0.2489
New Effect: Freshmen	7.76	0.76	10.23	<.0001
New Effect: Seniors	4.72	0.76	6.22	<.0001

FIXED EFFECTS table

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept (β0)	80.20	0.54	149.51	<.0001
Senior (β1)	2.16	0.76	2.85	0.0054
New (β2)	7.76	0.76	10.23	<.0001
Senior*New (β3)	-3.04	1.07	-2.83	0.0056

Effect of Senior for New: $\beta_1 + \beta_3(New_p)$; Effect of New for Seniors: $\beta_2 + \beta_3(Senior_p)$

Test Mean [SE]	**Freshmen**	**Seniors**	**Marginal**
Old Method	β_0 80.20 *[0.52]*	β_1 82.36 *[0.59]*	81.28 *[0.42]*
New Method	β_2 87.96 *[0.45]*	β_3 $\beta_2 + \beta_3$ 87.08 *[0.58]*	87.52 *[0.37]*
Marginal	84.08 *[0.65]*	$\beta_1 + \beta_3$ 84.72 *[0.53]*	84.40 *[0.42]*

GLM via "ANOVA" instead

- So far we've used "regression" to analyze our 2x2 design:
  - We manually dummy-coded the predictors
  - SAS treats them as "continuous" predictors, so it uses our variables as is

- More commonly, a factorial design like this would use an ANOVA approach to the GLM
  - It is the *same model* accomplished with less code
  - However – it will give us different (seemingly conflictory) information…

```
TITLE "GLM via ANOVA Instead (uses CLASS and LSMEANS)";
PROC GLM DATA=work.example1;
* CLASS statement denotes predictors as "categorical";
  CLASS Senior New;
* Model y = predictor effects like before;
  MODEL Test = Senior New Senior*New / SOLUTION;
* Get predicted test score per group, all differences across groups;
  LSMEANS Senior*New / PDIFF=ALL;
RUN; QUIT; TITLE;
```

"ANOVA": Results (duplicate test of R^2 omitted)

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Senior	1	10.24	10.24	**1.42**	0.2358
New	1	973.44	973.44	**135.33**	<.0001
Senior*New	1	57.76	57.76	**8.03**	0.0056

Unlike the dummy-coded regression, the **omnibus *F*-tests do NOT match** the fixed effect *t*-test results below ($F \neq T^2$), except for the interaction (within rounding error).

Parameter	Estimate		Standard Error	t Value	Pr > |t|
Intercept	87.08	B	0.54	162.34	<.0001
Senior 0	0.88	B	0.76	1.16	0.2489
Senior 1	0	B	.	.	.
New 0	-4.72	B	0.76	-6.22	<.0001
New 1	0	B	.	.	.
Senior*New 0 0	-3.04	B	1.07	-2.83	0.0056
Senior*New 0 1	0	B	.	.	.
Senior*New 1 0	0	B	.	.	.
Senior*New 1 1	0	B	.	.	.

$1.16^2 \approx 1.35$

$-6.22^2 \approx 38.69$

$-2.83^2 \approx 8.01$

To explain the dots, SAS will say this to you, but it's not a problem…

```
The X'X matrix has been
found to be singular, and
a generalized inverse was
used to solve the normal
equations. Terms whose
estimates are followed by
the letter 'B' are not
uniquely estimable.
```

LSMEANS-Created Tables

Senior	New	Test LSMEAN	LSMEAN Number
0	**0**	80.20	1
0	**1**	87.96	2
1	**0**	82.36	3
1	**1**	87.08	4

Least Squares Means for effect Senior*New
Pr > |t| for H0: LSMean(i)=LSMean(j)
Dependent Variable: Test

i/j	1	2	3	4
1		<.0001	0.0272	<.0001
2	<.0001		<.0001	0.6534
3	0.0272	<.0001		<.0001
4	<.0001	0.6534	<.0001	

This table shows the *p*-values for all cell differences. No SE or *t*-values are provided for these differences.

So do results match across "regression" and "ANOVA"?

Dummy-Coded Regression Omnibus *F*-Tests

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Senior	1	58.32	58.32	**8.11**	0.0054
New	1	752.72	752.72	**104.64**	<.0001
Senior*New	1	57.76	57.76	**8.03**	0.0056

Parameter: Fixed Effects	Estimate	t Value	Pr > |t|	Calc. F value
Intercept (β0)	80.20	149.51	<.0001	**22353.24**
Senior (β1)	2.16	2.85	0.0054	**8.12**
New (β2)	7.76	10.23	<.0001	**104.65**
Senior*New (β3)	-3.04	-2.83	0.0056	**8.01**

Estimated Main Effects	Estimate	t Value	Pr > |t|	Calc. F value
Senior Effect: Old (β1)	2.16	2.85	0.0054	**8.12**
Senior Effect: New (β1+β3)	-0.88	-1.16	0.2489	**1.35**
New Effect: Freshmen (β2)	7.76	10.23	<.0001	**104.65**
New Effect: Seniors (β2+β3)	4.72	6.22	<.0001	**38.69**

ANOVA Omnibus *F*-Tests

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Senior	1	10.24	10.24	**1.42**	0.2358
New	1	973.44	973.44	**135.33**	<.0001
Senior*New	1	57.76	57.76	**8.03**	0.0056

Omnibus F-Tests Above: ???? No Match!!

Below: From the *t*-tests and the dots, we can see that SAS reversed the 0/1 coding of each predictor to make **1 the reference** instead of 0.

Parameter	Estimate	t Value	Pr > |t|	Calc. F value
Intercept	87.08	162.34	<.0001	
Senior 0	0.88	1.16	0.2489	**1.35**
Senior 1	0	.	.	
New 0	-4.72	-6.22	<.0001	**38.69**
New 1	0	.	.	
Senior*New 0 0	-3.04	-2.83	0.0056	**8.01**
Senior*New 0 1	0	.	.	
Senior*New 1 0	0	.	.	
Senior*New 1 1	0	.	.	

Um, what?

- When using the CLASS statement in SAS PROC GLM (or equivalently, the BY statement in SPSS GLM):
  - This is an "ANOVA" approach in which SAS codes your categorical predictors
  - By default, the group highest numbered/last alphabetically is the reference
  - So a 0/1 variable effectively becomes 1/0 in the model
  - Can change default by sorting, but is easier just to recode the predictor (e.g., code Senior=1, Freshmen=2, so that Senior is still the "0" reference)

- That explains why the tables with the main effects (estimates, SE, *t*-values, and *p*-values) did not match across regression vs. ANOVA:
  - Regression: SAS reported the main effects we told it to (0 = reference)
  - ANOVA: SAS reported the *other* model-implied main effects (1 = reference)
  - This isn't really a problem so long as you can keep track of what "0" is!

- However, this does NOT explain why the omnibus *F*-tests for the main effects don't match across regression and ANOVA!

Why the omnibus *F*-tests don't match…

- When a predictor is NOT part of an interaction, its main effect is "**unconditional**" → it is the main effect for **anyone** in the sample
  - The main effect is "controlling for" the other predictors, but is not specific to any other predictor value – the lack of interaction says its main effect would have been the same for any value of the other predictors

- When a predictor IS part of an interaction, its main effect is "**conditional**" → it is the main effect **specifically for the interacting predictor(s) = 0**
  - The main effect is "controlling for" the other predictors, AND specifically for predictor(s) = 0 for any predictor it interacts with
  - The interaction implies that the main effect would be different at some other value of the interacting predictor(s) besides 0, so it matters what 0 is!

- To understand why the omnibus F-tests didn't match, we need to consider yet another way to create a "0" reference group…

2 Kinds of "Conditional" Main Effects

- **"Simple" conditional main effects**
  - Specifically for a "0" value in the interacting predictor, where the meaning of "0" is usually chosen deliberately with the goal of inferring about a particular kind of person (or group of persons)
  - e.g., the "simple" main effect of Education *for Attitudes = 3*
    the "simple" main effect of Attitudes *for Education = 12 years*
  - e.g., the "simple" effect of Old vs. New Instruction *for Seniors*
    the "simple" effect of Freshman vs. Senior *for New Instruction*

- **"Marginal" (omnibus) conditional main effects**
  - What is done for you without asking in ANOVA! The fixed effects solution is not given by default (and not often examined at all); the omnibus *F*-tests are almost always used to interpret "main effects" instead
  - Tries to produce the "average" main effect in the sample, marginalizing over other predictors
  - Consequently, a "0" person may not even be logically possible…

Making Regression Replicate ANOVA Omnibus *F*-Tests

```
* Centering group variables at "mean" to mimic ANOVA;
DATA work.example1; SET work.example1;
  SeniorC = Senior -.5;  NewC = New - .5;
LABEL SeniorC = "SeniorC :0=Second Semester Sophomore?"
      NewC =   "NewC: 0=Halfway New Instruction?";
RUN;
TITLE "GLM via Regression to Mimic ANOVA";
PROC GLM DATA=work.example1;
  MODEL Test = SeniorC NewC SeniorC*NewC / SOLUTION;
RUN; QUIT; TITLE;
```

Regression (ANOVA-Mimic) Omnibus *F*-Tests

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Senior	1	10.24	10.24	**1.42**	0.2358
New	1	973.44	973.44	**135.33**	<.0001
Senior*New	1	57.76	57.76	**8.03**	0.0056

Previous ANOVA Omnibus *F*-Tests

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Senior	1	10.24	10.24	**1.42**	0.2358
New	1	973.44	973.44	**135.33**	<.0001
Senior*New	1	57.76	57.76	**8.03**	0.0056

The same thing would happen in unbalanced data (i.e., with unequal group sizes), so long as groups were still coded as ±0.5 in the regression…

Making Regression Replicate ANOVA Omnibus *F*-Tests

Regression (ANOVA-Mimic) Fixed Effects

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept (new β0)	84.40	0.27	314.69	<.0001
SeniorC (new β1)	0.64	0.54	1.19	0.2358
NewC (new β2)	6.24	0.54	11.63	<.0001
SeniorC*NewC (β3)	-3.04	1.07	-2.83	0.0056

Dummy-Coded Regression Fixed Effects

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept (β0)	80.20	0.54	149.51	<.0001
Senior (β1)	2.16	0.76	2.85	0.0054
New (β2)	7.76	0.76	10.23	<.0001
Senior*New (β3)	-3.04	1.07	-2.83	0.0056

New β_1 is senior main effect for halfway new method ("marginal" conditional)
New β_2 is method main effect for second-semester sophomores ("marginal" conditional)

Test Mean [SE]	**Freshmen**	**Seniors**	**Marginal**
Old Method	β_0 80.20 *[0.52]*	β_1 82.36 *[0.59]*	81.28 *[0.42]*
	β_2	$\boldsymbol{\beta_3}$ / $\beta_2 + \beta_3$	$\boldsymbol{\beta_2}$
New Method	87.96 *[0.45]*	87.08 *[0.58]*	87.52 *[0.37]*
		$\beta_1 + \beta_3$	
Marginal	84.08 *[0.65]*	$\boldsymbol{\beta_1}$ 84.72 *[0.53]*	84.40 *[0.42]* $\boldsymbol{\beta_0}$

SUMMARY

Purpose of Today's Lecture…

- To examine exactly what we can learn from our model output
  - Meaning of estimated fixed effects; how to get model-implied fixed effects
  - Interpretation of omnibus significance tests

- To understand why results from named GLM variants may differ:
  - Regression/ANOVA/ANCOVA are all the same GLM
    - Linear model for the means + and a normally-distributed residual error term
    - You can fit main effects and interactions among any kind of predictors; whether they should be there is always a testable hypothesis in a GLM

- When variants of the GLM provide different results, it's because:
  - Your predictor variables are being recoded (if using CLASS/BY statements)
  - Simple conditional main effects and marginal conditional main effects do not mean the same thing (so they will not agree when in an interaction)
  - By default your software picks your model for the means for you:
    - "Regression" = whatever you tell it, exactly how you tell it
    - "ANOVA" = marginal main effects + all interactions for categorical predictors
    - "ANCOVA" = marginal main effects + all interactions for categorical predictors; continuous predictors only get to have main effects

SAS vs. SPSS for General Linear Models

- Analyses using least squares (i.e., any GLM) can be estimated equivalently in SAS PROC GLM or SPSS GLM ("univariate")…
- However… see below for a significant limitation

How do I tell it…	SAS GLM	SPSS GLM
What my DV is	First word after MODEL	First word after UNIANOVA
I have continuous predictors (or to leave them alone!!)	Assumed by default	WITH option
I have categorical predictors (and to dummy-code them for me)	CLASS statement	BY option
What fixed effects I want	After = on MODEL statement	After = on /DESIGN statement
To show me my fixed effects solution (Est, SE, t-value, p-value)	After / on MODEL statement	/PRINT = PARAMETER
To give me means per group	LSMEANS statement	/EMMEANS statement
To estimate model-implied effects	ESTIMATE statement	**NO CAN DO.**