Simple, Marginal, and Interaction Effects in General Linear Models

EPSY 905: Fundamentals of Multivariate Modeling

Online Lecture #3

Today's Class

- Centering and Coding Predictors
- Interpreting Parameters in the Model for the Means
- Main Effects Within Interactions
- GLM Example 1: "Regression" vs. "ANOVA"

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]*

CENTERING AND CODING PREDICTORS

The Two Sides of a Model

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

Our focus today

- **Model for the Means (Predicted Values):**
  - 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)

- **Model for the Variance:**
  - $e_p \sim N(0, \sigma_e^2)$→ 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, **we** 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, **either we or the program** 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

- <u>For more than two groups, need: ***dummy codes = #groups − 1***</u>
  - 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 ☺

- <u>Potential pit-falls:</u>
  - 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)

- <u>Other examples of things people do to categorical predictors:</u>
  - "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?

Categorical Predictors: Manual Coding

- Model: $y_i = \beta_0 + \beta_1 d1_i + \beta_2 d2_i + \beta_3 d3_i + e_i$
  - "Treatgroup" variable: Control=0, Treat1=1, Treat2=2, Treat3=3
  - New variables to be created for the model:
    - *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
- How does the model give us **all possible group differences**? By determining each group's mean, and then the difference…

Control Mean (Reference)	Treatment 1 Mean	Treatment 2 Mean	Treatment 3 Mean
$\boldsymbol{\beta_0}$	$\boldsymbol{\beta_0 + \beta_1 d1_i}$	$\boldsymbol{\beta_0 + \beta_2 d2_i}$	$\boldsymbol{\beta_0 + \beta_3 d3_i}$

- The model for the 4 groups directly provides 3 differences (control vs. each treatment), and indirectly provides another 3 differences (differences between treatments)

Group Differences from Dummy Codes

- Model: $y_i = \beta_0 + \beta_1 d1_i + \beta_2 d2_i + \beta_3 d3_i + e_i$

Control Mean (Reference)	Treatment 1 Mean	Treatment 2 Mean	Treatment 3 Mean
$\boldsymbol{\beta_0}$	$\boldsymbol{\beta_0 + \beta_1 d1_i}$	$\boldsymbol{\beta_0 + \beta_2 d2_i}$	$\boldsymbol{\beta_0 + \beta_3 d3_i}$

	Alt Group		Ref Group	Difference
• Control vs. T1 =	$(\beta_0+\beta_1)$	$-$	(β_0)	$= \beta_1$
• Control vs. T2 =	$(\beta_0+\beta_2)$	$-$	(β_0)	$= \beta_2$
• Control vs. T3 =	$(\beta_0+\beta_3)$	$-$	(β_0)	$= \beta_3$
• T1 vs. T2 =	$(\beta_0+\beta_2)$	$-$	$(\beta_0+\beta_1)$	$= \beta_2 - \beta_1$
• T1 vs. T3 =	$(\beta_0+\beta_3)$	$-$	$(\beta_0+\beta_1)$	$= \beta_3 - \beta_1$
• T2 vs. T3 =	$(\beta_0+\beta_3)$	$-$	$(\beta_0+\beta_2)$	$= \beta_3 - \beta_2$

Estimating (Univariate) Linear Models in R

	Alt Group	Ref Group	Difference
1. Control vs. T1 =	$(\beta_0+\beta_1)$	$-\ (\beta_0)$	$=\beta_1$
2. Control vs. T2 =	$(\beta_0+\beta_2)$	$-\ (\beta_0)$	$=\beta_2$
3. Control vs. T3 =	$(\beta_0+\beta_3)$	$-\ (\beta_0)$	$=\beta_3$
4. T1 vs. T2 =	$(\beta_0+\beta_2)$	$-\ (\beta_0+\beta_1)$	$=\beta_2-\beta_1$
5. T1 vs. T3 =	$(\beta_0+\beta_3)$	$-\ (\beta_0+\beta_1)$	$=\beta_3-\beta_1$
6. T2 vs. T3 =	$(\beta_0+\beta_3)$	$-\ (\beta_0+\beta_2)$	$=\beta_3-\beta_2$

```
#R Syntax for Estimating 4-Group Linear Model
#  For Predicting Y in data frame called mydata
library(multcomp)
model01 = lm(y~d1+d2+d3,data=mydata)
summary(model01) #  shows model results

mean1 = matrix(c(1,0,0,0),1); rownames(mean1) = c("Control Mean")
mean2 = matrix(c(1,1,0,0),1); rownames(mean2) = c("T1 Mean")
mean3 = matrix(c(1,0,1,0),1); rownames(mean3) = c("T2 Mean")
mean4 = matrix(c(1,0,0,1),1); rownames(mean4) = c("T3 Mean")
contrast1 = mean2-mean1; rownames(contrast1) = c("Control vs. T1")
contrast2 = mean3-mean1; rownames(contrast2) = c("Control vs. T2")
contrast3 = mean4-mean1; rownames(contrast3) = c("Control vs. T3")
contrast4 = mean3-mean2; rownames(contrast4) = c("T1 vs. T2")
contrast5 = mean4-mean2; rownames(contrast5) = c("T1 vs. T3")
contrast6 = mean4-mean3; rownames(contrast6) = c("T2 vs. T3")

mycontrasts = rbind(mean1,mean2,mean3,mean4,contrast1,contrast2,contrast3,contrast4,
                    contrast5,contrast6)
values = glht(model01,linfct=mycontrasts)
summary(values)
```

Note the order of the equations: the reference group mean *is subtracted from* the alternative group mean.

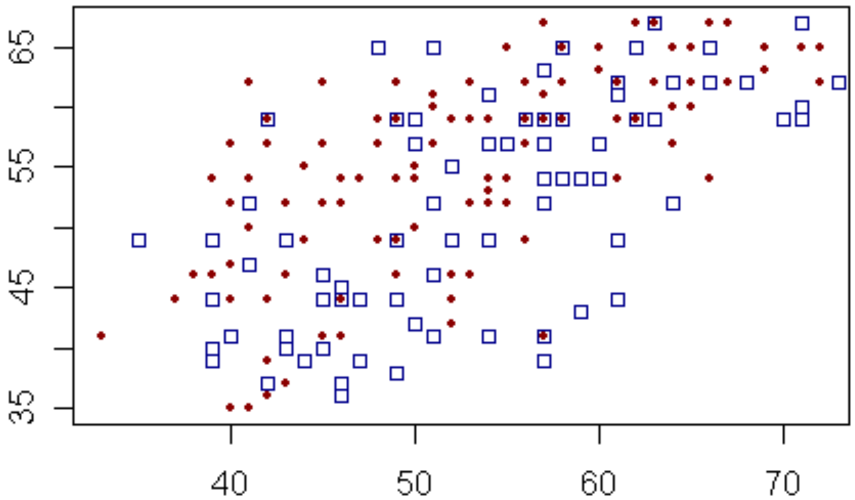
The ~ is the equals sign: to the left goes the DV. To the right go the IVs (a + indicates additive effects of IVs).

The values come from place-holder numbers put in the correct positions for the betas.

The ghlt function is from the multcomp package.

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*…

Continuous Predictors

- For **continuous** (quantitative) predictors, **<u>we</u>** (not R) will make the intercept interpretable by **centering**
  - **Centering** = subtract a constant (e.g., sample mean, other meaningful reference value) from each person's variable value so that **the 0 value** falls within the range of the new centered predictor variable
  - Predicted group means **at** specific levels of continuous predictors can be found using the same procedure (e.g., if X1 SD=5, means at ±1 SD):

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!

Model-Implied Simple Main Effects

- **Original:** $GPA_p = \beta_0 + (\beta_1{*}Att_p) + (\beta_2{*}Ed_p) + (\beta_3{*}Att_p{*}Ed_p) + e_p$

  $GPA_p = 30 + (1{*}Att_p) + (2{*}Ed_p) + (0.5{*}Att_p{*}Ed_p) + e_p$
- Given any values of the predictor variables, the model equation provides predictions for:
  - Value of outcome (model-implied intercept for non-zero predictor values)
  - Any conditional (simple) main effects implied by an interaction term
  - **Simple Main Effect = what it is + what *modifies* it**

- Step 1: **Identify** all terms in model involving the predictor of interest
  - e.g., Effect of Attitudes comes from: $\beta_1{*}Att_p + \beta_3{*}Att_p{*}Ed_p$
- Step 2: **Factor out** common predictor variable
  - Start with $[\beta_1{*}Att_p + \beta_3{*}Att_p{*}Ed_i] \rightarrow [Att_p\,(\beta_1 + \beta_3{*}Ed_p)] \rightarrow Att_p$ **(new** β_1**)**
  - Value given by ( ) is then the model-implied coefficient for the predictor
- Step 3: **ESTIMATEs** calculate model-implied simple effect and SE
  - Let's try it for **a new reference point of attitude = 3 and education = 12**

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: Expected GPA for 0 attitude and 0 years of education**
- **Interpret β_1: Increase in GPA per unit attitude for 0 years of education**
- **Interpret β_2: Increase in GPA per year education for 0 attitude**
- **Interpret β_3: Attitude as Moderator: Effect of education (slope) increases by .5 for each additional unit of attitude (more positive)**

  Education as Moderator: Effect of attitude (slope) increases by .5 for each additional year of education (more positive)

- **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? 0 = 12 years of education**
- **Why did β_1 change? Conditional on Education = 12 (new zero)**
- **Why did β_2 stay the same? Attitude is the same**
- **Why did β_3 stay the same? Nothing beyond to modify two-way interaction (effect is unconditional)**

- 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: $GPA_p = \beta_0 + \beta_1 * Att_p + \beta_2 * Ed_p + \beta_3 * Att_p * Ed_p + e_p$

  $$GPA_p = 30 + 2 * Att_p + 1 * Ed_p + 0.5 * Att_p * Ed_p + e_p$$

  - 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:

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

$GPA_p = 30 + 2 * Att_p + 1 * Ed_p + 0.5 * Att_p * Ed_p + e_p$

New equation using centered predictors:

$GPA_p = 66 + 8 * (Att_p - 3) + 2.5 * (Ed_p - 12) + .5 * (Att_p - 3) * (Ed_p - 12) + e_p$

- **β_0: expected value of GPA when $Att_p = 3$ and $Ed_p = 12$**

  $\beta_0 = 66$

- **β_1: effect of Attitudes**

  $\beta_1 = 2 + 0.5 * Ed_p = 2 + 0.5 * 12 = 8$

- **β_2: effect of Education**

  $\beta_2 = 1 + 0.5 * Att_p = 1 + .5 * 3 = 2.5$

- **β_3: two-way interaction of Attitudes and Education:**

  $\beta_3 = 0.5$

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., glht in R, 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., $Var(\beta_1) = SE_{\beta 1}^2$
- $SE_{\beta 1}^2 = Var(\beta_1) + Var(\beta_3) * Ed + 2Cov(\beta_1, \beta_3) * Ed$

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

GLM EXAMPLE 1:
"REGRESSION" VS. "ANOVA"

GLM via Dummy-Coding in "Regression"

```
#MODEL #1 -- Using 0/1 coding instead of factors
model1 = lm(Test~Senior+New+Senior*New,data=data01)
summary(model1)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  80.2000     0.5364 149.513  < 2e-16 ***
Senior        2.1600     0.7586   2.847  0.00539 **
New           7.7600     0.7586  10.229  < 2e-16 ***
Senior:New   -3.0400     1.0728  -2.834  0.00561 **

#MODEL #1 – ANOVA Table
anova(model1)

Analysis of Variance Table

Response: Test
           Df Sum Sq Mean Sq  F value    Pr(>F)
Senior      1  10.24   10.24   1.4235  0.235762
New         1 973.44  973.44 135.3253 < 2.2e-16 ***
Senior:New  1  57.76   57.76   8.0297  0.005609 **
Residuals  96 690.56    7.19
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Note: these ANOVA table is displaying marginal tests for the main effects. Marginal tests are for the main effect only and are not conditional on any interacting variables.

Getting Each of the Means as a Contrast

```
mean1 = matrix(c(1,0,0,0),1); rownames(mean1)="Freshman-Old"
mean2 = matrix(c(1,0,1,0),1); rownames(mean2)="Freshman-New"
mean3 = matrix(c(1,1,0,0),1); rownames(mean3)="Senior-Old"
mean4 = matrix(c(1,1,1,1),1); rownames(mean4)="Senior-New"

meansvec = rbind(mean1,mean2,mean3,mean4)
means = glht(model1,linfct=meansvec)

summary(means)
```

```
         Simultaneous Tests for General Linear Hypotheses

Fit: lm(formula = Test ~ Senior + New + Senior * New, data =
data01)

Linear Hypotheses:
                   Estimate Std. Error t value Pr(>|t|)
Freshman-Old == 0   80.2000     0.5364   149.5   <2e-16 ***
Freshman-New == 0   87.9600     0.5364   164.0   <2e-16 ***
Senior-Old == 0     82.3600     0.5364   153.5   <2e-16 ***
Senior-New == 0     87.0800     0.5364   162.3   <2e-16 ***
```

glht 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: $Test_p = \beta_0 + \beta_1 0 + \beta_2 1 + \beta_3 0 * 0$
- 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": Mapping Results to Data

glht 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

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

```
effect1 = matrix(c(0,1,0,0),1); rownames(effect1) = "Senior Effect: Old"
effect2 = matrix(c(0,1,0,1),1); rownames(effect2) = "Senior Effect: New"
effect3 = matrix(c(0,0,1,0),1); rownames(effect3) = "New Effect: Freshmen"
effect4 = matrix(c(0,0,1,1),1); rownames(effect4) = "New Effect: Seniors"

effectsvec = rbind(effect1,effect2,effect3,effect4)
effects = glht(model1,linfct=effectsvec)
summary(effects)
```

```
         Simultaneous Tests for General Linear Hypotheses

Fit: lm(formula = Test ~ Senior + New + Senior * New, data = data01)

Linear Hypotheses:
                          Estimate Std. Error t value Pr(>|t|)
Senior Effect: Old == 0     2.1600     0.7586   2.847   0.0194 *
Senior Effect: New == 0    -0.8800     0.7586  -1.160   0.5939
New Effect: Freshmen == 0   7.7600     0.7586  10.229   <0.001 ***
New Effect: Seniors == 0    4.7200     0.7586   6.222   <0.001 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- single-step method)
```

glht 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

glht 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]* $\beta_1 + \beta_3$	87.52 *[0.37]*
Marginal	84.08 *[0.65]*	84.72 *[0.53]*	84.40 *[0.42]*

GLM via "ANOVA" instead – in R with Factors

- 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

```
#MODEL #2 -- Using factors (R coded)
model2 = lm(Test~SeniorF+NewF+SeniorF*NewF,data=data01)
summary(model2)
anova(model2)

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)     80.2000     0.5364 149.513  < 2e-16 ***
SeniorF1         2.1600     0.7586   2.847  0.00539 **
NewF1            7.7600     0.7586  10.229  < 2e-16 ***
SeniorF1:NewF1  -3.0400     1.0728  -2.834  0.00561 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Analysis of Variance Table

Response: Test
             Df Sum Sq Mean Sq  F value    Pr(>F)
SeniorF       1  10.24   10.24   1.4235  0.235762
NewF          1 973.44  973.44 135.3253 < 2.2e-16 ***
SeniorF:NewF  1  57.76   57.76   8.0297  0.005609 **
Residuals    96 690.56    7.19
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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*
  - **These are given in the summary() function output of R**

- **"Marginal" (omnibus) 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…
  - **These are given in the anova() function output of R**

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")...

How do I tell it...	R
What my DV is	The first command in lm(Y~X): Before the ~
I have continuous predictors (or to leave them alone!!)	Assumed by default (can tell if you use class(data$variable)) function and find predictors are numeric
I have categorical predictors (and to dummy-code them for me)	class(data$variable) function says factor
What fixed effects I want	glht() function from multcomp package
To show me my fixed effects solution (Est, SE, t-value, p-value)	summary() function applied to lm() object
To give me means per group	glht() function or use factor type
To estimate model-implied effects	glht() function