

Slides Ch. 10

Sherri Verdugo

November 3, 2014

Contents

Chapter 10 One Way ANOVA	3
Sherri Verdugo, M.S.	3
Instructor, CSUF	3
Prologue & Introduction	3
ANOVA is the statistical cousin to the t test.	3
How Analysis of Variance is Used	3
Example of ANOVA Concept Introduction	3
Abortion stance by rural and urban settings	3
Analysis of Variance in Experimental Settings	4
Example 1: random groups	4
ANOVA example 1 continued: DATA	4
F: an intuitive friendly approach	5
Case 1: Day versus Night	5
F: an intuitive friendly approach	5
Case 2: Day versus Night New Scores	5
F: an intuitive friendly approach	6
Case 3: Day versus Night New Scores	6
F: an intuitive friendly approach	6
Case 4: Day versus Night New Scores	6
Terminology: Getting to know ANOVA	7
Terminology: Getting to know ANOVA	7
Terminology: Getting to know ANOVA	8
Summary Recap for Sum of Squares	8
Terminology: Getting to know ANOVA	8
Summary Recap for Mean Squares	8
Terminology: Getting to know ANOVA	8
Summary Recap for Mean Squares	8
Procedure: Calculating One-Way ANOVA	9
Procedure: Calculating One-Way ANOVA	9
Procedure: Calculating One-Way ANOVA	9

One Way ANOVA EXAMPLE:	9
x = Pro-Life Index	9
One Way ANOVA EXAMPLE:	10
x = Pro-Life Index	10
One Way ANOVA EXAMPLE:	10
x = Pro-Life Index	10
Let's calculate F	11
Let's calculate F	11
Let's calculate F	11
Let's calculate F	11
Wrap up our conclusions in a table	12
Comparing t with F	12
Analysis of Variance: Exp. Data	12
Analysis of Variance: Exp. Data	13
Find the squares	13
Find the squares	14
ANOVA STEPS	14
ANOVA STEPS	14
ANOVA STEPS	14
Post Hoc Testing	15
Post Hoc Testing: Steps	15
Post Hoc Testing: Steps	15
From our example:	15
SPSS applications	16
Two-Way ANOVA	16
Input	16
Two-Way ANOVA	16
Output Summary	16
Two Way ANOVA	17
Anova table	17
What about interactions??	17
What about interactions??	17
What about interactions??	18
What about interactions??	18
Two-Way: Interactions and ANOVA	19
Key Concepts	19

Equations of Interest	20
Equations of Interest, Continued	20
Equations of Interest, Continued	20

Chapter 10 One Way ANOVA

Sherri Verdugo, M.S.

Instructor, CSUF

Prologue & Introduction

The Analysis of Variance (ANOVA) can be used for more than one group. We are not limited to just one group. ANOVA is a powerful tool.

ANOVA is the statistical cousin to the t test.

- A technique for comparing sample means.
 - Can be used for more than two means.
 - Friendly to experimental testing.
 - Several sample means are being compared at once.
 - Overview of the process.
 - Can be used to evaluate statistical tests.
-

How Analysis of Variance is Used

Design for non experimental application with two variables:

- * One is interval level of measurement (usually the dependent)
 - * Grouped data of any level of measurement.
 - * ANOVA is non-directional...meaning only non-directional hypothesis apply.
-

Example of ANOVA Concept Introduction

Abortion stance by rural and urban settings

Example of a non-experimental setting

Rural Group 1	Urban Group 2
160	100

Rural Group 1	Urban Group 2
130	110
150	130
140	120
$\Sigma X_1 = 580$	$\Sigma X_2 = 460$
$\bar{X}_1 = 145$	$\bar{X}_2 = 115$

- Means differ
- Hypothesis: $H_0 : \mu_1 = \mu_2$ and $H_A : \mu_1 \neq \mu_2$
- ANOVA is NON-DIRECTIONAL
- Calculate the F ratio (named Fisher, the original development)
- Compare the F to $F_{critical}$ at $\alpha = 0.5$

Analysis of Variance in Experimental Settings

Example 1: random groups

Experiments are the backbone of science. ANOVA is the most common used statistical technique. Participants or subjects are the people being studied in an experiment.

- **Scenario:** Using a random sampling assignment of 30 subjects are assigned to one of three groups. Each group has 10 subjects. Because of the random assignment, any difference in mean satisfaction scores among the three groups will be the result of the time frame for the course (day, night, and Saturday) rather than any other factors.
- Hypothesis for variable selection:

$$H_0 : \mu_{day} = \mu_{night} = \mu_{saturday}$$

$$H_{A1} : \mu_{day} \neq \mu_{night}, H_{A2} : \mu_{day} \neq \mu_{saturday}, \mu_{night} \neq \mu_{saturday}$$

- Question: are the means different from each other to conclude they are not related to sampling error.

ANOVA example 1 continued: DATA

Day	Night	Saturday
5	5	5
5	4	5
5	4	5
5	3	4

Day	Night	Saturday
4	3	4
4	2	4
4	2	3
4	2	3
3	1	3
2	0	2

- Sums for each group: $\Sigma X_D = 41$, $\Sigma X_N = 26$, and $\Sigma X_S = 38$
- Means for each group: $\bar{D} = \frac{41}{10} = 4.1$, $\bar{N} = \frac{26}{10} = 2.6$, and $\bar{S} = \frac{38}{10} = 3.8$
- Question: the means are different... are they different enough to claim that they did not come from sampling error???

F: an intuitive friendly approach

Case 1: Day versus Night

Day	Night
5	5
5	5
5	5
3	3
3	3
3	3
$\Sigma x_D = 24$	$\Sigma x_D = 24$
$\bar{x}_D = 24/6 = 4$	$\bar{x}_D = 24/6 = 4$

- The Grand Mean = $\Sigma G_M = \frac{48}{12} = 4.00$
- We can use the F statistic here.

F: an intuitive friendly approach

Case 2: Day versus Night New Scores

Day	Night
5	5

Day	Night
5	5
5	3
5	3
3	3
3	3
$\Sigma x_D = 26$	$\Sigma x_D = 22$
$\bar{x}_D = 26/6 = 4.33$	$\bar{x}_D = 22/6 = 3.67$

- The Grand Mean = $\Sigma G_M = \frac{48}{12} = 4.00$
- $F = 1.248$
- The classes now differ a bit because of the satisfaction rating.

F: an intuitive friendly approach

Case 3: Day versus Night New Scores

Day	Night
5	5
5	3
5	3
5	3
5	3
3	3
$\Sigma x_D = 28$	$\Sigma x_D = 20$
$\bar{x}_D = 28/6 = 4.67$	$\bar{x}_D = 20/6 = 3.33$

- The Grand Mean = $\Sigma G_M = \frac{48}{12} = 4.00$
- $F = 7.994$
- The classes now differ a bit because of the satisfaction rating.

F: an intuitive friendly approach

Case 4: Day versus Night New Scores

Day	Night
5	3
5	3
5	3
5	3
5	3
5	3
$\Sigma x_D = 30$	$\Sigma x_D = 18$
$\bar{x}_D = 30/6 = 5.00$	$\bar{x}_D = 18/6 = 3.00$

- The Grand Mean = $\Sigma G_M = \frac{48}{12} = 4.00$
- F = mathematically undefined... it is getting larger though and approaching infinity as a limit.

Terminology: Getting to know ANOVA

Some new concepts are needed to deal with the calculations.

- Variance Estimate = $\hat{\sigma} = \frac{\Sigma(x-\bar{x})^2}{n-1}$
- **Sum of Squares:** The sum of the squared deviations of the values of x from the mean.
- **Mean Square:** The mean squared deviation of a score (a value of x) from the mean of all scores.
 $MS = \frac{SS}{df}$
- **Total Sum of Squares:** The total of the squared deviations of scores about the grand mean.
 $SS_{TOTAL} = \Sigma x^2 - \frac{(\Sigma x)^2}{n}$
- **Between-Groups Sums of Squares [page 328]:** $SS_{Between} = \frac{(\Sigma X_{cat.1})^2}{n_{cat.1}} + \frac{(\Sigma X_{cat.2})^2}{n_{cat.2}} + \dots + \frac{(\Sigma X_{last.cat})^2}{n_{last.cat}} - \frac{(\Sigma X_{total})^2}{n_{total}}$

Terminology: Getting to know ANOVA

- **Between-Groups Sums of Squares:** The portion of the total sum of squares that can be accounted for by the variations of the category means about the grand mean. $SS_{Between} = \frac{(\Sigma X_{cat.1})^2}{n_{cat.1}} + \frac{(\Sigma X_{cat.2})^2}{n_{cat.2}} + \dots + \frac{(\Sigma X_{last.cat})^2}{n_{last.cat}} - \frac{(\Sigma X_{total})^2}{n_{total}}$
- **Within-Groups Sums of Squares (a.k.a. Error Sum of Squares) [page 328]:** The portion of the total sum of squares left unexplained by the variations of the category means about the grand mean.
 $SS_{Within} = SS_{Total} - SS_{Between}$

Terminology: Getting to know ANOVA

Summary Recap for Sum of Squares

- SS_B = the portion of SS_T accounted for by the categories of the independent variable.
 - SS_T = the portion of SS_T not accounted for by the categories of the independent variable.
 - Additive: $SS_T = SS_B + SS_W$
 - We use both SS_T and SS_B to form two separate population variance estimates, these result in the between groups mean squares and between groups degrees of freedom.
-

Terminology: Getting to know ANOVA

Summary Recap for Mean Squares

- **Between-Groups Mean Square:** A variance estimate based on the between-groups sum of squares.
 $MS_{BETWEEN} = \frac{SS_B}{df_B} = \frac{SS_B}{num.categories - 1}$
 - **Degrees of Freedom, between-groups Mean Squares:** $df_{between} = num.categories - 1$
 - **Within-Groups Mean Squares:** A population variance estimate based on what the categories of the independent variable do not explain –the variation of scores within the groups: $MS_{Within} = \frac{SS_{Within}}{df_{Within}} = \frac{SS_{Within}}{(n_{Total} - 1) - (num.categories - 1)}$
 - ****Degrees of Freedom, Within-Groups Mean:** That portion of the total degrees of freedom not accounted for by the number of groups studied. $df_W = (n_{Total} - 1) - (num.categories - 1)$
-

Terminology: Getting to know ANOVA

Summary Recap for Mean Squares

$$df_W = df_{total} - df_{between} = (n - 1) - (num.categories) - 1$$

So what is MS_W ?

- $MS_{Within} = \frac{SS_{Within}}{df_{Within}} = \frac{SS_{Within}}{(n_{Total} - 1) - (num.categories - 1)}$
- The degrees of freedom are additive:
– $df_{total} = df_B + df_W$

Ok, so what is F ?

$$F = \frac{MS_B}{MS_W}$$

- Cool, but what if I get lost in the calculations???
-

Procedure: Calculating One-Way ANOVA

- STEP 1: Calculate the grand mean $\bar{X}_{GRAND} = \frac{\Sigma X_{ALL.SCORERS}}{n_{total.N.in.study}}$
 - STEP 2: Calculate the total sum of squares (sum of squared deviations from the GRAND MEAN), where $SS_{Total} = \Sigma x_{Total}^2 - \frac{(\Sigma x_{Total})^2}{n_{Total}}$
 - STEP 3: Calculate the BETWEEN-group sum of squares (sum of squared deviations of category means from the GRAND mean), where $SS_{Between} = \frac{(\Sigma X_{cat.1})^2}{n_{cat.1}} + \frac{(\Sigma X_{cat.2})^2}{n_{cat.2}} + \dots + \frac{(\Sigma X_{last.cat})^2}{n_{last.cat}} - \frac{(\Sigma X_{total})^2}{n_{total}}$
 - Note: when you see $+\dots+$ it is short hand saying that you CONTINUE repeating the $\frac{(\Sigma X_{cat})}{n_{cat}}$ for ALL OF THE subsequent categories, if any.
-

Procedure: Calculating One-Way ANOVA

- STEP 4: Calculate the within-group sum of squares, where $SS_{Within} = SS_{Total} - SS_{Between}$.
 - STEP 5: Calculate the between and within mean squares:
 - $MS_{Between} = \frac{SS_{Between}}{df_{Between}} = \frac{SS_{Between}}{num.categories-1}$
 - $MS_{Within} = \frac{SS_{Within}}{df_{Within}} = \frac{SS_{Within}}{(n_{Total}-1)-(num.categories-1)}$
 - STEP 6: Calculate the F ratio where $F = \frac{MS_{Between}}{MS_{Within}}$
-

Procedure: Calculating One-Way ANOVA

- STEP 7: Use the table of F values in the textbook to test the F for significance.
 - Numerator $df = df_{between} = num.categories - 1$
 - Denominator $df = df_{within} = (n.total - 1) - (num.categories - 1)$
 - Decisions:
 - If F is significant and the number of categories is greater than TWO, perform a post-hoc procure such as Scheffe's test.
 - If F is significant and this is *non-experimental research*, measure association with the correlation ratio or r (Note: we will cover this next week!)
-

One Way ANOVA EXAMPLE:

x = Pro-Life Index

x_1 =Rural	x_2 =Urban
160	100
130	110

Let's calculate F

- Step One: find the total sum of squares
 - $SS_{Total} = \sum x_{Total}^2 - \frac{(\sum x_{Total})^2}{n_{Total}} = 138,000 - \frac{(1040)^2}{8}$
 - $= 138,000 - \frac{1,081,600}{8} = 138,000 - 135,200 = 2,800$
 - Step Two: find the between-group sum of squares (we already calculated SS_{total})
 - $SS_{Between} = \frac{(\sum X_{cat.1})^2}{n_{cat.1}} + \frac{(\sum X_{cat.2})^2}{n_{cat.2}} - \frac{(\sum X_{total})^2}{n_{total}}$
 - $= \frac{(580)^2}{4} + \frac{(460)^2}{4} - 135,200$
 - $= \frac{336,400}{4} + \frac{211,600}{4} - 135,200$
 - $= 84,100 + 52,900 - 135,200 = 137,000 - 135,200 = 1,800$
-

Let's calculate F

- Step 3: Find the within group sum of squares $SS_{Within} = SS_{Total} - SS_{Between}$
 - $SS = 2,800 - 1,800 = 1,000$
 - Step 4: Calculate the Mean Squares
 - $MS_{Between} = \frac{SS_{Between}}{df_{Between}} = \frac{SS_{Between}}{num.categories-1} = \frac{1,800}{2-1} = \frac{1,800}{1} = 1,800$
 - $MS_{Within} = \frac{SS_{Within}}{df_{Within}} = \frac{SS_{Within}}{(n_{Total}-1)-(num.categories-1)} = \frac{1,000}{(8-1)-(2-1)}$
 - $= \frac{1,000}{(7)-(1)} = \frac{1,000}{6} = 166.67$
-

Let's calculate F

- Step 5: find the F ratio
 - $F = \frac{MS_{Between}}{MS_{Within}} = \frac{1,800}{166.67} = 10.799 = 10.80$
 - Step 6: find the degrees of freedom
 - $df_B = num.categories - 1 = 2 - 1 = 1$
 - $df_W = (n_{Total} - 1) - (num.categories - 1)$
 - $= (8 - 1) - (2 - 1) = 7 - 1 = 6$
-

Let's calculate F

- Step 7: find f critical
 - Link: http://www.socr.ucla.edu/applets.dir/f_table.html
 - Find the section for $\alpha = 0.05$
 - The top row is $df_B = 1$ and we move down the column until we find
 - The row with $df_W = 6$ and the intersection is 5.9874 or 5.99
- Step 8: decisions, decisions...

- $F_{obtained} = 10.80 > F_{critical} = 5.99$ we reject H_0
- Step 9: why why why???
- Because we need to state our probability of falsely rejecting a true null hypothesis at $\alpha = 0.05$
- Go back to the table and look up $F_{critical}$ with $\alpha = 0.01$. This is 13.74.
- Result: since $.01 < p < .05$, we can state that our probability of falsely rejecting the true null hypothesis as $p < .05$

Wrap up our conclusions in a table

Source	SS	df	MS	F	P
TOTAL	2,800				
Between	1,800	1	1,800	10.80	<.05
Within	1,000	6	166.67		

- Typically, we do not report the df_{total} or MS_{total} because we don't really need it to calculate F .

Comparing t with F

- Yep, I know a few statistical tests, why can't I use what I'm comfortable with.
 - Sigh, if only it WAS that easy. We are methodical. Remember, that is our purpose in this class.
- So, what, I can use either t or F ... does it matter?
 - Of course, for two reasons:
 - 1. You can make a directional assumption in a t test's H_1 and reduced our error by 1/2.
 - 2. ANOVA **assumes** EQUAL POPULATION VARIANCE. This is not always the case! But, alas it is not covered too much in this class.
- Can I assume the ANOVA assumptions???
 - In this class, yes, but you must use the word "robust" to justify your answer!
 - Outside of this class and in the real world... ONE NEVER ASSUMES ANYTHING... EVER EVER EVER EVER!
- Robust: accurate even when underlying assumptions (such as equal population variances) are violated.

Analysis of Variance: Exp. Data

Day	Night	Saturday
5	5	5

Day	Night	Saturday
5	4	5
5	4	5
5	3	4
4	3	4
4	2	4
4	2	3
4	2	3
3	1	3
2	0	2
$n_1 = 10$	$n_2 = 10$	$n_3 = 10$
$\Sigma X_1 = 41$	$\Sigma X_2 = 26$	$\Sigma X_3 = 38$
$\bar{x}_1 = 4.1$	$\bar{x}_2 = 2.6$	$\bar{x}_3 = 3.8$

Analysis of Variance: Exp. Data

- $\Sigma X_{Total} = \Sigma X_1 + \Sigma X_2 + \Sigma X_3$
- $\Sigma X_{Total} = 41 + 26 + 38 = 105$
- Grand Mean = $\frac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3}{n.cases} = \frac{10.5}{3} = 3.5$

Find the squares

Day^2	$Night^2$	$Saturday^2$
$5^2 = 25$	$5^2 = 25$	$5^2 = 25$
$5^2 = 25$	$4^2 = 16$	$5^2 = 25$
$5^2 = 25$	$4^2 = 16$	$5^2 = 25$
$5^2 = 25$	$3^2 = 9$	$4^2 = 16$
$4^2 = 16$	$3^2 = 9$	$4^2 = 16$
$4^2 = 16$	$2^2 = 4$	$4^2 = 16$
$4^2 = 16$	$2^2 = 4$	$3^2 = 9$
$4^2 = 16$	$2^2 = 4$	$3^2 = 9$
$3^2 = 9$	$1^2 = 1$	$3^2 = 9$
$2^2 = 4$	$0^2 = 0$	$2^2 = 4$

Find the squares

- $\Sigma X_1^2 = 177$ and $\Sigma X_2^2 = 88$ and $\Sigma X_3^2 = 154$
 - $\Sigma X_{total} = \Sigma X_1^2 + \Sigma X_2^2 + \Sigma X_3^2$
 - $\Sigma X_{total} = 177 + 88 + 154 = 419$
-

ANOVA STEPS

- Step 1: $SS_{Total} = \Sigma x_{Total}^2 - \frac{(\Sigma x_{Total})^2}{n_{Total}}$
 - $SS_{Total} = 419 - \frac{(105^2)}{30} = 419 - \frac{11,025}{30}$
 - $= 419 - 367.50 = 51.50$
 - Step 2: $SS_B = \frac{(\Sigma X_{cat.1})^2}{n_{cat.1}} + \frac{(\Sigma X_{cat.2})^2}{n_{cat.2}} + \frac{(\Sigma X_{cat.3})^2}{n_{cat.3}} - \frac{(\Sigma X_{total})^2}{n_{total}}$
 - $\frac{(41)^2}{10} + \frac{(26)^2}{10} + \frac{(38)^2}{10} - 367.50$
 - $\frac{1681}{10} + \frac{676}{10} + \frac{1444}{10} - 367.50$
 - $168.10 + 67.60 + 144.40 - 367.50 = 12.60$
-

ANOVA STEPS

- Step 3: $SS_W = SS_T - SS_B = 51.50 - 12.60 = 38.90$
 - Step 4: Calculate the MS Between and Within
 - $MS_B = \frac{SS_B}{df_B} = \frac{SS_B}{num.categories-1}$
 - $MS_B = \frac{12.60}{3-1} = \frac{12.60}{2} = 6.30$
 - $MS_W = \frac{SS_W}{df_W} = \frac{SS_W}{(n_{Total}-1)-(num.categories-1)}$
 - $MS_W = \frac{38.90}{(30-1)-(3-1)} = \frac{38.90}{29-2} = \frac{38.90}{27} = 1.44$
-

ANOVA STEPS

- Step 5: Calculate F
 - $F = \frac{MS_B}{MS_W} = \frac{6.30}{1.44} = 4.375$
 - Step 6: Consult the F table http://www.socr.ucla.edu/applets.dir/f_table.html
 - For $F_{critical}$ at $df = 2$ and $df = 27$
 - $F_{critical}\alpha = 0.05 = 3.35 < F_{obtained} = 4.375$ REJECT H_0
 - $F_{critical}\alpha = 0.01 = 5.49 > 4.375$ with $p < 0.05$
-

Post Hoc Testing

- How do we detect subtle differences between the satisfaction and time of day scores? Isn't the previous procedure only covering one inequality????
- We do this by setting up our inequalities:
 - $\mu_D \neq \mu_N \neq \mu_S$
 - $\mu_D \neq \mu_N = \mu_S$
 - $\mu_D = \mu_N \neq \mu_S$
 - $\mu_D = \mu_S \neq \mu_N$

Post Hoc Testing: Steps

- Summarize the ANOVA source table from the previous problem.

Source	SS	df	MS	F	P
TOTAL	51.50				
Between	12.60	2	6.30	4.375	<.05
Within	38.90	27	1.44		

- For any two categories, i and j , the following formula generates Scheffe's Critical value.

$$(\bar{X}_i - \bar{X}_j) = \pm \sqrt{(df_B)(F_{critical})(MS_W)(\frac{1}{n_i} + \frac{1}{n_j})}$$

Post Hoc Testing: Steps

From our example:

- $(\bar{X}_i - \bar{X}_j) = \pm \sqrt{(2)(3.35)(1.44)(\frac{1}{10} + \frac{1}{10})}$
- $\pm \sqrt{1.9296} = \pm 1.389$
- So how does this work:

Null	$absolute \bar{x}_i - \bar{x}_j$	Scheffe's Crit.	Conclusions
$\mu_D = \mu_N$	$absolute 4.1 - 2.6 = 1.5$	> 1.389	Reject H_0
$\mu_D = \mu_S$	$absolute 4.1 - 3.8 = 0.3$	< 1.389	Cannot reject H_0
$\mu_N = \mu_S$	$absolute 2.6 - 3.8 = 1.2$	< 1.389	Cannot reject H_0

SPSS applications

- To complete ANOVA in SPSS: “statistics -> Compare Means -> One-Way ANOVA”
 - See files for SPSS
-

Two-Way ANOVA

Input

Day	Night	Saturday
5	5	5
5	4	5
5	4	5
5	3	4
4	3	4
4	2	4
4	2	3
4	2	3
3	1	3
2	0	2
$n_1 = 10$	$n_2 = 10$	$n_3 = 10$

Two-Way ANOVA

Output Summary

Groups	Count	Sum	Average	Variance
Column 1	10	41	4.1	0.988889
Column 2	10	26	2.6	2.266667
Column 3	10	38	3.8	1.066667

- Column 1 = Day
 - Column 2 = Night
 - Column 3 = Saturday
-

Two Way ANOVA

Anova table

- Summarize the ANOVA source table from the previous problem.

Source	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P</i>	F crit
Between	12.6	2	6.30	4.375	0.022642	3.354131
Within	38.9	27	1.440741			
TOTAL	51.5	29				

What about interactions??

We look at the means and that indicates that an *interaction* might be occurring. We run this on SPSS because the calculations can become daunting!

- Case One: Advertising Strategists

Day	Night	Saturday
5	5	5
5	4	5
5	4	5
5	3	4
4	3	4
$\bar{x} = 4.8$	$\bar{x} = 3.8$	$\bar{x} = 4.6$

- Not much of a difference here...relatively similar.

What about interactions??

We look at the means and that indicates that an *interaction* might be occurring. We run this on SPSS because the calculations can become daunting!

- Case Two: Media Strategists

Day	Night	Saturday
4	2	4

Day	Night	Saturday
4	 2	 3
4	 2	 3
3	 1	 3
2	 0	 2
$\bar{x} = 3.4$	$\ \bar{x} = 1.4$	$\ \bar{x} = 3.0$

- Something is different here.

What about interactions??

We look at the means and that indicates that an *interaction* might be occurring. We run this on SPSS because the calculations can become daunting!

- Case Three: Advertising Strategists

Day	Night	Saturday
5	 5	 5
4	 4	 4
4	 4	 4
3	 3	 3
2	 2	 2
$\bar{x} = 3.6$	$\ \bar{x} = 3.6$	$\ \bar{x} = 3.6$

- Not much of a difference here...they are all equal.

What about interactions??

We look at the means and that indicates that an *interaction* might be occurring. We run this on SPSS because the calculations can become daunting!

- Case Four: Media Strategists

Day	Night	Saturday
5	 3	 5
5	 2	 5

Day	Night	Saturday
5	 2	 4
4	 1	 3
4	 0	 3
$\bar{x} = 4.6$	 $\bar{x} = 1.6$	 $\bar{x} = 4.0$

- They are different... especially the Night. We would need to investigate this further.

Two-Way: Interactions and ANOVA

- “Statistics”, “General Linear Model”, and then “GLM - General Factorial.”
- Let’s do this together.

Key Concepts

- **One-Way Analysis of Variance (ANOVA) [page 318]:** A technique for comparing sample means that can be used to compare more than two means. Analysis of variance with one dependent variable and one independent variable.
- **F/The F Ratio: [page 319]:** The statistic generated by the analysis of variance procedure.
- **Participants or Subjects in an Experiment [page 320]:** The people being studied in an experiment.
- **Grand Mean [page 323]:** The mean of all scores.
- **Sum of Squares [page 327]:** The sum of the squared deviations of the values of x from the mean.
- **Mean Square [page 327]:** The mean squared deviation of a score (a value of x) from the mean of all scores. $MS = \frac{SS}{df}$
- **Total Sum of Squares [page 328]:** The total of the squared deviations of scores about the grand mean. $SS_{TOTAL} = \sum x^2 - \frac{(\sum x)^2}{n}$
- **Between-Groups Sums of Squares [page 328]:** The portion of the total sum of squares that can be accounted for by the variations of the category means about the grand mean. $SS_{Between} = \frac{(\sum X_{cat.1})^2}{n_{cat.1}} + \frac{(\sum X_{cat.2})^2}{n_{cat.2}} + \dots + \frac{(\sum X_{last.cat})^2}{n_{last.cat}} - \frac{(\sum X_{total})^2}{n_{total}}$
- **Within-Groups Sums of Squares (a.k.a. Error Sum of Squares) [page 328]:** The portion of the total sum of squares left unexplained by the variations of the category means about the grand mean. $SS_{Within} = SS_{Total} - SS_{Between}$
- **Between-Groups Mean Square [page 329]:** A variance estimate based on the between-groups sum of squares. $MS_{BETWEEN} = \frac{SS_B}{df_B} = \frac{SS_B}{num.categories - 1}$
- **Degrees of Freedom, between-groups Mean Squares [page 329]:** $df_{between} = num.categories - 1$
- **Within-Groups Mean Squares [page 329]:** A population variance estimate based on what the categories of the independent variable do not explain –the variation of scores within the groups: $MS_{Within} = \frac{SS_{Within}}{df_{Within}} = \frac{SS_{Within}}{(n_{Total} - 1) - (num.categories - 1)}$
- ****Degrees of Freedom, Within-Groups Mean [page 329]:** That portion of the total degrees of freedom not accounted for by the number of groups studied. $df_W = (n_{Total} - 1) - (num.categories - 1)$

- **ANOVA source table [page 337]:** A table summarizing the results of the main steps in the ANOVA procedure.
 - **Robustness (of a test of significance) [page 338]:** accurate even when underlying assumptions (such as equal population variances) are violated.
 - **Post hoc procedures/ post hoc test of multiple comparisons [page 341]:** A follow on procedure that is used once a null hypothesis has been rejected.
 - **Scheffe's test [page 341]:** A test that finds the critical difference between any two sample means that is necessary to reject the null hypothesis that their corresponding population means are equal.

$$(\bar{X}_i - \bar{X}_j) = \pm \sqrt{(df_B)(F_{critical})(MS_W)(\frac{1}{n_i} + \frac{1}{n_j})}$$
 - **Scheffe's critical value [page 342]:** The value in this test needed to reject the null hypothesis.
 - **Two-way analysis of variance/ two way ANOVA [page 350]:** Analysis of variance that includes a second independent variable.
 - **Interaction Effects [page 351]:** Situations where the relationship between the dependent variable and one of the independent variables is a function of the levels of the other independent variable.
-

Equations of Interest

- Grand Mean = $\frac{\Sigma X}{N}$ it is the mean of all scores divided by N.
 - Variance Estimate = $\hat{\sigma} = \frac{\Sigma(x-\bar{x})^2}{n-1}$
-

Equations of Interest, Continued

- Analysis of Variance Components:
 - $SS_{Total} = \Sigma x_{Total}^2 - \frac{(\Sigma x_{Total})^2}{n_{Total}}$
 - $SS_{Between} = \frac{(\Sigma X_{cat.1})^2}{n_{cat.1}} + \frac{(\Sigma X_{cat.2})^2}{n_{cat.2}} + \dots + \frac{(\Sigma X_{last.cat})^2}{n_{last.cat}} - \frac{(\Sigma X_{total})^2}{n_{total}}$
 - $SS_{Within} = SS_{Total} - SS_{Between}$
 - $MS_{Between} = \frac{SS_{Between}}{df_{Between}} = \frac{SS_{Between}}{num.categories-1}$
 - $MS_{Within} = \frac{SS_{Within}}{df_{Within}} = \frac{SS_{Within}}{(n_{Total}-1)-(num.categories-1)}$
 - $F = \frac{MS_{Between}}{MS_{Within}}$
 - $df_B = num.categories - 1$
 - $df_W = (n_{Total} - 1) - (num.categories - 1)$
-

Equations of Interest, Continued

- Scheffe's Test: For any two categories, i and j , the following formula generates Scheffe's critical value.
 - $(\bar{X}_i - \bar{X}_j) = \pm \sqrt{(df_B)(F_{critical})(MS_W)(\frac{1}{n_i} + \frac{1}{n_j})}$