

Two-Factor ANOVA in SPSS

Workshop offered by the Mississippi Center for Supercomputing Research and the UM Office of Information Technology

John P. Bentley
Department of Pharmacy Administration
University of Mississippi

October 7, 2009

Outline

- Overview, background, and some terms
- The concept of an interaction
- The null hypotheses in two-way ANOVA
- Just a little statistical theory
- An example and a procedure for interpreting a two-way design: **Using SPSS**
- Revisiting simple effects
- Unbalanced designs: **Using SPSS**
- Higher-way ANOVA: A quick note

Overview and background

- In one-factor (or one-way) ANOVA, we test for differences among two or more means of one independent variable. We might have had many levels, but the bottom line was that we had only one IV.
- But what if I suspect that the difference between an experimental group and a control group is different for males and females? So I might suspect that the treatment works for one gender, but not the other.
- In performing an experiment (e.g., comparing two treatment groups with a control condition), what if I would like to control the effects of a variable (like gender or age group) that is related to the DV? This would not only help me to make sure that such a variable is not confounding the relationship of interest, but can also aid in increasing the power of the test of the comparisons of interest by reducing residual variation (this is called blocking – but it is really a special case of two-way ANOVA).
- Can I do these things? Yes, by incorporating another independent variable into our design and conducting a **two-way ANOVA**.

October 7, 2009

Two-factor ANOVA in SPSS

3

Overview and background

- A research design with two or more categorical independent variables (a.k.a., factors), each studied at two or more levels is called a **factorial design** – when you have 2 IVs, you have a two-way factorial design. The variables can be experimentally manipulated or intrinsic to the subjects (a.k.a., classification variable, subject variable, measured variable, organismic variable, individual-difference variable).
- Factorial designs are those in which the independent variables are completely **crossed**, which refers to the fact that the design includes every possible combination of the levels of the IVs.
- In this seminar, we will consider the case where the factors are **between-subjects factors** – subjects serve in only one of the treatment conditions.

October 7, 2009

Two-factor ANOVA in SPSS

4

Overview and background

- To visualize a two-way factorial design, think of a rectangle:

In this example, FACTOR B has 4 levels, so $G_B = 4$, indexed by j

In this example, FACTOR A has 2 levels, so $G_A = 2$, indexed by i

	B ₁	B ₂	B ₃	B ₄	Total
A ₁	μ_{11}	μ_{12}	μ_{13}	μ_{14}	$\mu_{1.}$
A ₂	μ_{21}	μ_{22}	μ_{23}	μ_{24}	$\mu_{2.}$
Total	$\mu_{.1}$	$\mu_{.2}$	$\mu_{.3}$	$\mu_{.4}$	$\mu_{..}$

This is called a 2 x 4 factorial design. If factor A has G_A levels and factor B has G_B levels, then a two-way factorial contains $G_A * G_B$ combinations (cells). In the current example, there are $2 * 4 = 8$ combinations or cells.

Cell mean $\rightarrow \mu_{ij}$

October 7, 2009

Two-factor ANOVA in SPSS

5

Overview and background

- Row marginal means defined as: $\mu_{i.} = \sum_{j=1}^{G_B} \frac{\mu_{ij}}{G_B}$
- Column marginal means defined as: $\mu_{.j} = \sum_{i=1}^{G_A} \frac{\mu_{ij}}{G_A}$
- Grand mean defined as: $\mu_{..} = \sum_{i=1}^{G_A} \frac{\mu_{i.}}{G_A} = \sum_{j=1}^{G_B} \frac{\mu_{.j}}{G_B} = \frac{1}{G_A G_B} \sum_{i=1}^{G_A} \sum_{j=1}^{G_B} \mu_{ij}$
- We will describe a group of observational units, all receiving the same treatment combination, as **elements** in a cell.

October 7, 2009

Two-factor ANOVA in SPSS

6

Overview and background

- **Balanced design vs. unbalanced design**
 - A balanced design has an equal number of observations in each cell, whereas an unbalanced design does not. For a balanced design, $n_{\text{cell}} = N/(G_A G_B)$, indicating the number of observations in each cell.
 - Unbalanced designs are common in observational studies (as opposed to experimental studies), where the levels of certain factors are determined after, rather than before, the data are collected, or when the proportions of subjects in the levels of the variables of interest are indeed different in the population.
 - The unbalanced design presents special problems in statistical analysis – both in computation and interpretation.

October 7, 2009

Two-factor ANOVA in SPSS

7

Overview and background

- For the remainder of this talk, we will focus on: 1) between-subjects factors, 2) the one DV situation, and 3) replicated designs (i.e., within each combination or cell, there are at least two observations made).
- We could differentiate between, fixed-, random-, and mixed-effects models for two-way ANOVA; our focus in this talk will be on the two-way **fixed-effects** ANOVA model.
- We will start by focusing on balanced designs – there are some slides at the end for unbalanced designs if we have time.
- We will NOT spend any time discussing the assumptions, consequences of assumption violation, or potential remedies.

October 7, 2009

Two-factor ANOVA in SPSS

8

Overview and background

- Why do a two-way design? Why not just do two one-way tests?
- The two-way factorial design offers two primary advantages over performing 2 separate one-way ANOVAs:
 - Allows us to look at multiple IVs within the context of a single design.
 - We get to perform multiple tests (the “main effects” of the IVs and the interaction) without having to correct for multiple testing.
 - Whenever a second factor is related to the DV or interacts with the first factor, the MSE for the two-factor ANOVA will be less than for a one-factor ANOVA and the *F*-ratio for the first factor will be larger in the two-factor design than in a one-factor design – i.e., you will have more power with a two-way ANOVA. This is a primary reason for using blocking (i.e., blocked designs).
 - Allows us to assess whether the IVs are interacting with each other.

October 7, 2009

Two-factor ANOVA in SPSS

9

Overview and background

- Main effects, interaction, and simple effects
 - A **main effect** expresses the differences among the means for one independent variable averaged over the levels of the other independent variable.
 - A **simple effect** expresses the differences among the means for one independent variable at a fixed level of the other independent variable.
 - An **interaction** is present when the simple effects of one independent variable are not the same at all levels of the other independent variable.

October 7, 2009

Two-factor ANOVA in SPSS

10

The concept of an interaction

- An interaction occurs when the relationship between the IV and the DV is different at different levels (i.e., values) of the moderator.
- In the case of a two-factor ANOVA, this statement is equivalent to the definition on the previous slide.
- Here are some other equivalent ways of representing an interaction in a two-way ANOVA:
 - An interaction between two factors exists if the mean differences among the levels of factor A are not constant across the levels of factor B (or if the mean differences among the levels of factor B are not constant across the levels of factor A) – interaction as a difference in differences of means.
 - An interaction occurs when some variability among cell means is not fully attributable to the main effects of the two IVs.
 - An interaction is present if the combination of 2 variables has an effect on the DV over and above the effect expected from each separate IV.
 - We will see another way of describing an interaction when we state the two-way fixed-effects ANOVA model and learn about partitioning total variation.

October 7, 2009

Two-factor ANOVA in SPSS

11

The null hypotheses

- There are three primary hypotheses for two-way ANOVA and there are multiple ways to state these null hypotheses.
- Below is one method using our 2x4 design from slide 5 as the example:

→ Main effect of factor A (based on marginal means)

$$H_0 : \mu_{1.} = \mu_{2.} \text{ (no main effect of factor A)}$$

$$\text{vs. } H_A : \mu_{1.} \neq \mu_{2.}$$

→ Main effect of factor B (based on marginal means)

$$H_0 : \mu_{.1} = \mu_{.2} = \mu_{.3} = \mu_{.4} \text{ (no main effect of factor B)}$$

$$\text{vs. } H_A : \mu_{.j} \neq \mu_{.j'}, \text{ for some } j \neq j'$$

October 7, 2009

Two-factor ANOVA in SPSS

12

The null hypotheses

→ Test of interaction between A and B (based on cell means)

You want to test differences in the differences of the means

In general:

$$H_0 : \mu_{ij} - \mu_{i'j} = \mu_{ij'} - \mu_{i'j'}, \text{ for all } i \neq i' \text{ AND } j \neq j'$$

$$\text{vs. } H_A : \mu_{ij} - \mu_{i'j} \neq \mu_{ij'} - \mu_{i'j'}, \text{ for some } i \neq i' \text{ AND } j \neq j'$$

Or equivalently:

$$H_0 : \mu_{ij} - \mu_{ij'} = \mu_{i'j} - \mu_{i'j'}, \text{ for all } i \neq i' \text{ AND } j \neq j'$$

$$\text{vs. } H_A : \mu_{ij} - \mu_{ij'} \neq \mu_{i'j} - \mu_{i'j'}, \text{ for some } i \neq i' \text{ AND } j \neq j'$$

For our 2 x 4 example:

$$H_0 : \begin{cases} \mu_{11} - \mu_{21} = \mu_{12} - \mu_{22} \\ \mu_{11} - \mu_{21} = \mu_{13} - \mu_{23} \\ \mu_{11} - \mu_{21} = \mu_{14} - \mu_{24} \end{cases} \text{ (no interaction between A and B)}$$

October 7, 2009

Two-factor ANOVA in SPSS

13

The null hypotheses

- In words: The factor A effect at B1 equals the factor A effect at B2, B3, and B4.
 - Can also be phrased in terms of the B effect across the levels of A.
- The test of interaction does not test equality of all cell means. There must be no main effect of Factor A, no main effect of Factor B, and no interaction for all cell means to be equal.

October 7, 2009

Two-factor ANOVA in SPSS

14

TABLE 19.7 General (balanced) two-way ANOVA table²

Source	d.f.	SS	MS	<i>F</i>
				Fixed
Row (main effect)	$r - 1$	SSR	$MSR = \frac{SSR}{r - 1}$	$\frac{MSR}{MSE}$
Column (main effect)	$c - 1$	SSC	$MSC = \frac{SSC}{c - 1}$	$\frac{MSC}{MSE}$
Row \times column (interaction)	$(r - 1)(c - 1)$	SSRC	$MSRC = \frac{SSRC}{(r - 1)(c - 1)}$	$\frac{MSRC}{MSE}$
Error	$rc(n - 1)$	SSE	$MSE = \frac{SSE}{rc(n - 1)}$	
Total	$rcn - 1$	SSY		

© 2007 Thomson Higher Education

From Kleinbaum *et al.* (2008)

Just a little statistical theory

- The test statistic we will employ in ANOVA is the F statistic, which is actually a ratio of variance estimates.
- Under the normality and independence assumptions, such a ratio has the F distribution when H_0 is true and the calculated value of F (the F statistic) can be used to test the null hypotheses of interest.
- We would expect a value close to 1 if H_0 is true, but larger than 1 if H_0 is not true.
- The larger the value of F , the more likely H_0 is to be untrue. We will appeal to the F distribution to help us make our decisions.

An example

- We will use the data set in Table 19.5 (page 521) from Kleinbaum *et al.*
- The DV is forced expiratory volume, a measure of pulmonary function (low FEV indicates possible respiratory dysfunction and high FEV indicates good respiratory function). Observations were taken on $n = 12$ individuals in each of three plants in a given industry where workers are exposed primarily to one of three toxic substances. Thus, we have two factors, each with three levels (3 x 3 design).
- There are a total of 108 observations in the data set, 12 in each of the 9 cells (see next slide).
- We will treat both of the factors as fixed (although you could argue that one or both of these factors could be treated as random).

October 7, 2009

Two-factor ANOVA in SPSS

17

TABLE 19.5 Forced expiratory volume classified by plant and toxic exposure

Plant	Toxic Substance			Row Total
	A	B	C	
1	4.64, 5.92, 5.25, 6.17, 4.20, 5.90, 5.07, 4.13, 4.07, 5.30, 4.37, 3.76 ($T_{11} = 58.78$)	3.21, 3.17, 3.88, 3.50, 2.47, 4.12, 3.51, 3.85, 4.22, 3.07, 3.62, 2.95 ($T_{12} = 41.57$)	3.75, 2.50, 2.65, 2.84, 3.09, 2.90, 2.62, 2.75, 3.10, 1.99, 2.42, 2.37 ($T_{13} = 32.98$)	$R_1 = 133.33$
2	5.12, 6.10, 4.85, 4.72, 5.36, 5.41, 5.31, 4.78, 5.08, 4.97, 5.85, 5.26 ($T_{21} = 62.81$)	3.92, 3.75, 4.01, 4.64, 3.63, 3.46, 4.01, 3.39, 3.78, 3.51, 3.19, 4.04 ($T_{22} = 45.33$)	2.95, 3.21, 3.15, 3.25, 2.30, 2.76, 3.01, 2.31, 2.50, 2.02, 2.64, 2.27 ($T_{23} = 32.37$)	$R_2 = 140.51$
3	4.64, 4.32, 4.13, 5.17, 3.77, 3.85, 4.12, 5.07, 3.25, 3.49, 3.65, 4.10 ($T_{31} = 49.56$)	4.95, 5.22, 5.16, 5.35, 4.35, 4.89, 5.61, 4.98, 5.77, 5.23, 4.86, 5.15 ($T_{32} = 61.52$)	2.95, 2.80, 3.63, 3.85, 2.19, 3.32, 2.68, 3.35, 3.12, 4.11, 2.90, 2.75 ($T_{33} = 37.65$)	$R_3 = 148.73$
Column Total	$C_1 = 171.15$	$C_2 = 148.42$	$C_3 = 103.00$	$G = 422.57$

© 2007 Thomson Higher Education

From Kleinbaum *et al.* (2008)

October 7, 2009

Two-factor ANOVA in SPSS

18

How should I proceed?

- There is no single recipe that outlines the best approach for interpreting a two-way design (however, a useful first step is to plot the cell means).
- I will outline the conventional approach.
- What you do depends on what you find. After plotting, always start with the interaction – look at the test of significance for it first. There are three possible outcomes:

October 7, 2009

Two-factor ANOVA in SPSS

19

How should I proceed?

– No interaction is present

- Without an interaction, the analysis is, in effect, two individual one-way ANOVAs.
- Each factor can be interpreted in isolation, using the principles that apply to the analysis of a one-way design.
- You have a statistical justification for generalizability, meaning that you can generalize the effect of factor A to all the levels of factor B and vice versa. Assume the factors in our study were gender and drug therapy. If there was no treatment-by-gender interaction, the overall treatment effect can be generalized to both sexes. The effects of treatment do not vary by sex.
- You can proceed to look at the main effects of the two factors in the context of a single design.
- There is some debate as to whether you should drop the interaction and refit the model before assessing the main effects. Dropping the interaction is described as “pooling,” because the SS attributed to the interaction are pooled with the residual SS to provide a new estimate of error, based on more degrees of freedom. I would suggest not pooling – use the MSE from the full interaction model when testing the main effects.
- You may decide to perform contrasts of interest, like all pairwise comparisons on the main effects that are significant.
- If neither main effect is significant, you can conclude that there is no evidence of any relationship between the response variable and the predictors (considered either separately or together).

October 7, 2009

Two-factor ANOVA in SPSS

20

How should I proceed?

- **An interaction is present but is dominated by the main effects**
 - When an interaction is present, the effect of either factor changes with the levels of the other. The simple picture of two main effects is not appropriate, and the two factors cannot be treated completely separately. To investigate the finding, we must consider how the simple effects of one factor differ with the levels of the other.
 - Although the presence of an interaction implies that the simple effects are not the same, we still may not need to interpret each simple effect separately.
 - While you have a departure from parallelism, the levels of one factor are found to be in the same relative order for all levels of the other factor. In other words, at each level of factor B, the difference between the factor A level effects lies in the same direction, but the magnitude of the difference depends on factor B (Kleinbaum *et al.* call this a same-direction interaction).
 - We might approach this problem by trying to understand the main effects first, then see how this pattern is modulated by the other factor.

October 7, 2009

Two-factor ANOVA in SPSS

21

How should I proceed?

- **The interaction dominates the main effects**
 - Here, the strategy of viewing interaction as a modulation of a strong main effect is not useful. Indeed, it can be downright deceptive to look at the marginal (main) effects at all.
 - Kleinbaum *et al.* call this a reverse interaction: the direction of the difference between two cell means for one row (column) is opposite to, or reversed, from, the direction of the difference between the corresponding cell means for some other row (column).
 - The simple effects differ qualitatively. You should look at the simple effects separately ... you would be justified in ignoring the main effects altogether.

October 7, 2009

Two-factor ANOVA in SPSS

22

How should I proceed?

- Simple effects tests
 - Test equality of all cell means within a single row, or equality of all cell means within a single column.
 - For example:

		FACTOR B - DOSAGE			
		0 mg	3 mg	6 mg	
FACTOR A - GENDER	M	μ_{M0}	μ_{M3}	μ_{M6}	$\mu_{M.}$
	F	μ_{F0}	μ_{F3}	μ_{F6}	$\mu_{F.}$
		$\mu_{.0}$	$\mu_{.3}$	$\mu_{.6}$	$\mu_{..}$

October 7, 2009

Two-factor ANOVA in SPSS

23

How should I proceed?

- Test the effect of dose among males:
 - $H_0: \mu_{M0} = \mu_{M3} = \mu_{M6}$
 - Note the similarity to one-way ANOVA
- If this test is significant, further testing is required to determine which levels differ – for example, you may want to do all of the pairwise comparisons (Keppel and Wickens call these simple comparisons). You could implement Tukey's HSD – use the MSE from the two-way ANOVA table.

October 7, 2009

Two-factor ANOVA in SPSS

24

An example

October 7, 2009 Two-factor ANOVA in SPSS 25

An example

Tests of Between-Subjects Effects

Dependent Variable: FEV

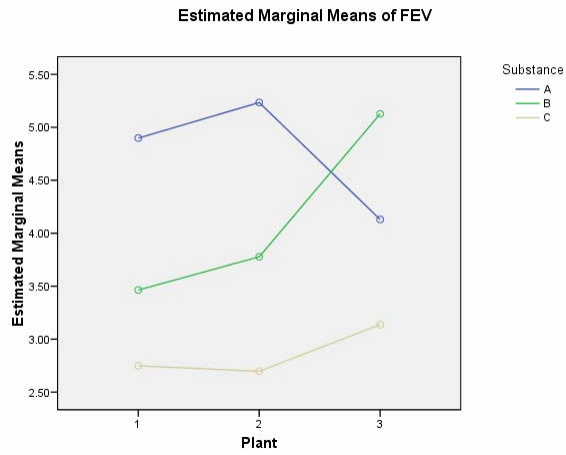
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	94.698 ^a	8	11.837	44.096	.000
Intercept	1653.383	1	1653.383	6159.160	.000
Plant	3.299	2	1.649	6.145	.003
Substance	66.889	2	33.445	124.588	.000
Plant * Substance	24.510	4	6.128	22.826	.000
Error	26.576	99	.268		
Total	1774.658	108			
Corrected Total	121.274	107			

^a. R Squared = .781 (Adjusted R Squared = .763)

Start here

October 7, 2009 Two-factor ANOVA in SPSS 26

An example

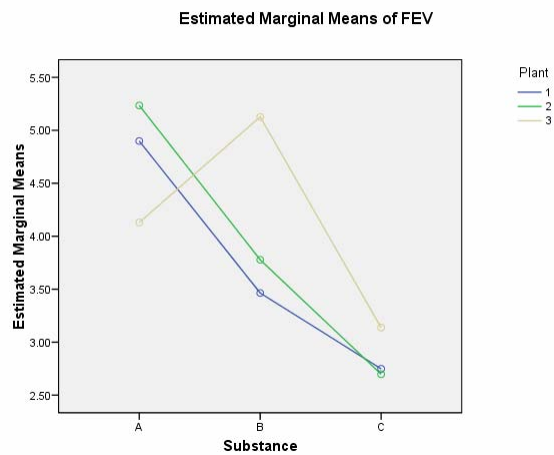


October 7, 2009

Two-factor ANOVA in SPSS

27

An example



October 7, 2009

Two-factor ANOVA in SPSS

28

An example

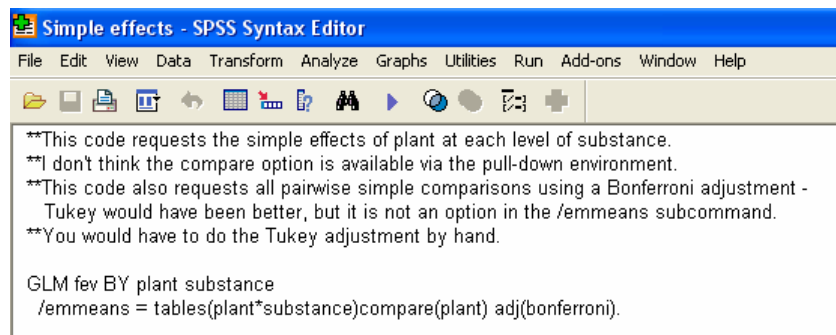
- The interaction is significant and it appears that the interaction dominates the main effects. What should we do next?
- Let's examine the simple effects of plant at each level of substance (I could have just as easily chosen to look at the simple effects of substance at each plant).

October 7, 2009

Two-factor ANOVA in SPSS

29

An example



```
Simple effects - SPSS Syntax Editor
File Edit View Data Transform Analyze Graphs Utilities Run Add-ons Window Help
**This code requests the simple effects of plant at each level of substance.
**I don't think the compare option is available via the pull-down environment.
**This code also requests all pairwise simple comparisons using a Bonferroni adjustment -
  Tukey would have been better, but it is not an option in the /emmeans subcommand.
**You would have to do the Tukey adjustment by hand.

GLM fev BY plant substance
  /emmeans = tables(plant*substance)compare(plant) adj(bonferroni).
```

October 7, 2009

Two-factor ANOVA in SPSS

30

An example

Univariate Tests

Dependent Variable: FEV

Substance		Sum of Squares	df	Mean Square	F	Sig.
A	Contrast	7.689	2	3.845	14.322	.000
	Error	26.576	99	.268		
B	Contrast	18.729	2	9.365	34.885	.000
	Error	26.576	99	.268		
C	Contrast	1.391	2	.695	2.590	.080
	Error	26.576	99	.268		

Each F tests the simple effects of Plant within each level combination of the other effects shown. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

October 7, 2009

Two-factor ANOVA in SPSS

31

An example

Pairwise Comparisons

Dependent Variable: FEV

Substance	(I) Plant	(J) Plant	Mean Difference (I-J)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^a	
						Lower Bound	Upper Bound
A	1	2	-.336	.212	.347	-.851	.179
		3	.768*	.212	.001	.253	1.283
	2	1	.336	.212	.347	-.179	.851
		3	1.104*	.212	.000	.589	1.619
	3	1	-.768*	.212	.001	-1.283	-.253
		2	-1.104*	.212	.000	-1.619	-.589
B	1	2	-.313	.212	.425	-.828	.202
		3	-1.663*	.212	.000	-2.178	-1.147
	2	1	.313	.212	.425	-.202	.828
		3	-1.349*	.212	.000	-1.864	-.834
	3	1	1.663*	.212	.000	1.147	2.178
		2	1.349*	.212	.000	.834	1.864
C	1	2	.051	.212	1.000	-.464	.566
		3	-.389	.212	.206	-.904	.126
	2	1	-.051	.212	1.000	-.566	.464
		3	-.440	.212	.120	-.955	.075
	3	1	.389	.212	.206	-.126	.904
		2	.440	.212	.120	-.075	.955

Based on estimated marginal means

*. The mean difference is significant at the .050 level.

^a. Adjustment for multiple comparisons: Bonferroni.

October 7, 2009

Two-factor ANOVA in SPSS

32

What do we conclude?

- A two-way ANOVA was used to assess the effect of manufacturing plant (3 levels) and exposure to toxic substance (3 levels) on respiratory function measured by forced expiratory volume (FEV). There is evidence of a significant interaction between toxic substance and plant ($F(4,99)=22.83$, $p<0.0001$). Furthermore, the plotting of the results suggests a reverse interaction and thus follow-up analyses using simple effects were conducted in an attempt to understand the nature of the interaction.
- When considering differences in FEV among the plants within each toxic substance, a conflicting picture emerges. For workers exposed to substance C, the plants were not significantly different ($F(2,99)=2.59$, $p=0.08$) and all had rather low FEVs. Statistically significant differences across the plants were observed for both substance A ($F(2,99)=14.32$, $p<0.0001$) and substance B ($F(2,99)=34.89$, $p<0.0001$). Pairwise comparisons among the cell means using a Bonferroni adjustment for multiple comparisons revealed that for substance A, plants 1 and 2 had a significantly higher mean FEV than plant 3 and that for substance B, plants 1 and 2 had significantly lower mean FEV than plant 3. These patterns demonstrate that some plants are associated with better respiratory health than others for one kind of toxic exposure but are worse for other kinds of exposure. These data do not allow us to conclude that one plant was better overall than another, rather, the differences in respiratory health among plants depend on which toxic substance is being considered.

October 7, 2009

Two-factor ANOVA in SPSS

33

Revisiting simple effects

- **There are differences of opinions about when and whether simple effects should be tested.**
 - Simple effects are not a breakdown of the AxB interaction into orthogonal parts. Simple effects contain a portion of the AxB interaction and a portion of the main effect under examination.
 - By testing the simple effects of B at each separate level of A, the main effects of A, in essence, were removed, but no correction was made for the influence of the main effects of B – these main effects are still present.
 - We have this relationship: $SS_B + SS_{AxB} = \sum SS_{B \text{ at } a_i}$

The sum of the simple effects of factor B equals the total of the sums of squares associated with factor B and the AxB interaction.

October 7, 2009

Two-factor ANOVA in SPSS

34

Revisiting simple effects

- Recall: $SS_B + SS_{AxB} = \sum SS_{B \text{ at } a_i}$
- Let factor B = plant and let factor A = toxic substance
 - $SS_B = 3.299$
 - $SS_{AxB} = 24.510$
 - $SS_{B \text{ at } SubA} = 7.689$
 - $SS_{B \text{ at } SubB} = 18.729$
 - $SS_{B \text{ at } SubC} = 1.391$
 - $3.299 + 24.510 = 7.689 + 18.729 + 1.391 = 27.809$

October 7, 2009

Two-factor ANOVA in SPSS

35

Revisiting simple effects

- **There are differences of opinions about when and whether simple effects should be tested.**
 - What this means, then, is that a significant simple effect does not represent significant interaction effects at the particular level of the other IV but, rather, the variability of the interaction and main effects **combined**.
 - Thus, examination of simple effects does not provide the researcher with an assessment of the pristine effects of interaction.
 - Myers and Well (2003) state, "Several individuals (e.g., Marascuilo & Levin, 1970; Rosnow & Rosenthal, 1989, 1991, 1995) have argued that when the interaction is significant, tests should be performed not on the simple effects of each factor but on interaction effects ... In contrast, Tukey (1991) gave equal 'first class' status to comparisons of simple effects and 'cross comparisons' of interaction effects."

October 7, 2009

Two-factor ANOVA in SPSS

36

Revisiting simple effects

- **Type I error control**
 - There is general consensus that the tests of the interaction and the two main effects are planned tests and do not require error correction.
 - However, there is no consensual standard for the level of familywise error to use with simple effects. Some researchers argue that a significant interaction justifies the use of uncorrected tests for the simple effects.
 - Others control the error rate over the family of simple effects. Let's revisit our 2 x 4 design from slide 5. Assume we set FWE at 0.10. The two simple effects of B at A1 and B at A2 would be evaluated at a per-test rate of $0.10/2=0.05$. The four simple effects of A at B1, A at B2, A at B3, and A at B4 would be evaluated at a per-test rate of $0.10/4=0.025$.
 - Error control is more important when a significant simple effect is followed by tests of simple comparisons.
 - In our example, we used uncorrected tests for simple effects but we corrected for tests of simple comparisons.

October 7, 2009

Two-factor ANOVA in SPSS

37

Unbalanced designs

- There are two good reasons for preferring balanced designs:
 - The possible consequences of heterogeneity of variance will be minimized.
 - The design will be orthogonal.

October 7, 2009

Two-factor ANOVA in SPSS

38

Unbalanced designs

For the unbalanced design:

$$\begin{aligned}
 SST &= \sum_{i=1}^r \sum_{j=1}^c \sum_{k=1}^{n_{ij}} (Y_{ijk} - \bar{Y}_{...})^2 & SSE &= \sum_{i=1}^r \sum_{j=1}^c \sum_{k=1}^{n_{ij}} (Y_{ijk} - \bar{Y}_{ij\cdot})^2 \\
 SSR &= \sum_{i=1}^r \sum_{j=1}^c \sum_{k=1}^{n_{ij}} (\bar{Y}_{i..} - \bar{Y}_{...})^2 & SSC &= \sum_{i=1}^r \sum_{j=1}^c \sum_{k=1}^{n_{ij}} (\bar{Y}_{\cdot j\cdot} - \bar{Y}_{...})^2 \\
 SSRC &= \sum_{i=1}^r \sum_{j=1}^c \sum_{k=1}^{n_{ij}} (\bar{Y}_{ij\cdot} - \bar{Y}_{i..} - \bar{Y}_{\cdot j\cdot} + \bar{Y}_{...})^2
 \end{aligned}$$

- When you have a balanced design, the collection of sums of squares are said to be orthogonal (i.e., $SST = SSR$ (rows) + SSC (columns) + $SSRC$ (row x column interaction) + SSE). Thus, the terms on the right-hand side must partition the total sum of squares into non-overlapping sources of variation.

October 7, 2009

Two-factor ANOVA in SPSS

39

Unbalanced designs

- When you have an unbalanced design, the sums of squares no longer represent completely separate (i.e., orthogonal) sources of variation. With unequal cell numbers: $SST \neq SSR$ (rows) + SSC (columns) + $SSRC$ (row x column interaction) + SSE . Thus, the traditional sums of squares approach cannot be used – the order in which the effects are tested (i.e., enter the model) matters.
- A variety of methods have been developed to compute sums of squares for each effect.
- Before exploring these, let's briefly explore the effects of unequal cell sizes when we collapse across cells to look at marginal means.

October 7, 2009

Two-factor ANOVA in SPSS

40

Unbalanced designs

		Gender	
		Female	Male
Education	College degree	$n_{11}=8$ $\bar{X}_{11}=25$	$n_{12}=3$ $\bar{X}_{12}=27$
	No college degree	$n_{21}=4$ $\bar{X}_{21}=17$	$n_{22}=7$ $\bar{X}_{22}=20$

- Is there a difference between the salaries of men and women?

October 7, 2009

Two-factor ANOVA in SPSS

41

Unbalanced designs

- Approach 1 (group-based averages, a.k.a., estimated marginal means, adjusted means, or least-squares means):

$$\text{Women: } \left(\frac{25+17}{2} \right) = 21.0$$

$$\text{Men: } \left(\frac{27+20}{2} \right) = 23.5$$

- Approach 2 (score-based averages):

$$\text{Women: } \left(\frac{8*25 + 4*17}{12} \right) = 22.3$$

$$\text{Men: } \left(\frac{3*27 + 7*20}{10} \right) = 22.1$$

October 7, 2009

Two-factor ANOVA in SPSS

42

Unbalanced designs

- Neither approach is intrinsically better than the other - they give the individuals and the groups different importance. The first approach gives equal weight to each group mean; the second approach weights each cell mean by its cell size (each individual is given equal weight).
- If group size is irrelevant to the conclusion of the study, you should use the first approach.
- This is usually the case when group sizes are different for random reasons (random dropouts or incomplete records) or are the result of properties of the subject population that are irrelevant (like having more women than men sign up for a study).

October 7, 2009

Two-factor ANOVA in SPSS

43

Unbalanced designs

- However, sometimes the sampling proportions are part of the hypotheses (like in some areas of sociology and political science). Here, the second approach should be used.
- How does the type of means relate to the analysis of variance with an unbalanced design?
- If we are interested in testing null hypotheses that treat all the groups equivalently, we should use the first approach and also the procedure known as Type III sums of squares to perform the tests.
- Type III SS observes the unique contribution of each factor above and beyond the other effects included in the model - the cell means are weighted equally - the cell means are unweighted.

October 7, 2009

Two-factor ANOVA in SPSS

44

Unbalanced designs

- If you would like each cell mean to be weighted by its cell size when computing marginal means, you should use Type I sums of squares. Perhaps your design intentionally has unequal cell sizes and you want your analysis to reflect this feature. In this approach, the order the factors are considered makes a difference in how the SS are computed.
- Considering there are also Type II and Type IV sums of squares, and you didn't even discuss them, this is just too much. **What should I do?**
- Keppel and Wickens (2004) make the following recommendation (which is supported by many statisticians): "Unless you wish your conclusions to depend on the size of the samples, the analysis of variance with unequal sample sizes should use Type III sums of squares. You should interpret them using marginal averages based on the cell means (i.e., the estimated, adjusted, or least-squares means)."

October 7, 2009

Two-factor ANOVA in SPSS

45

Unbalanced designs: An example

- We will use the data set in Table 20.3 (page 559) of Kleinbaum *et al.* The DV is satisfaction with medical care and the two IVs are patient's worry (2 levels) and a measure of affective communication (grouped into 3 levels).
- There are a total of 50 observations in the data set, with cell sizes ranging from 4 to 12 (see next slide).

October 7, 2009

Two-factor ANOVA in SPSS

46

TABLE 20.2 Satisfaction with medical care (Y), classified by patient worry and affective communication between patient and physician

Affective Communication	Worry		Row Marginals
	Negative	Positive	
High	2, 5, 8, 6, 2, 4, 3, 10 ($n_{11} = 8$) ($\bar{Y}_{11\cdot} = 5$)	7, 5, 8, 6, 3, 5, 6, 4, 5, 6, 8, 9 ($n_{12} = 12$) ($\bar{Y}_{12\cdot} = 6$)	$n_{1\cdot} = 20$ $\bar{Y}_{1\cdot\cdot} = 5.60$
Medium	4, 6, 3, 3 ($n_{21} = 4$) ($\bar{Y}_{21\cdot} = 4$)	7, 7, 8, 6, 4, 9, 8, 7 ($n_{22} = 8$) ($\bar{Y}_{22\cdot} = 7$)	$n_{2\cdot} = 12$ $\bar{Y}_{2\cdot\cdot} = 6.00$
Low	8, 7, 5, 9, 9, 10, 8, 6, 8, 10 ($n_{31} = 10$) ($\bar{Y}_{31\cdot} = 8$)	5, 8, 6, 6, 9, 7, 7, 8 ($n_{32} = 8$) ($\bar{Y}_{32\cdot} = 7$)	$n_{3\cdot} = 18$ $\bar{Y}_{3\cdot\cdot} = 7.56$
Column Marginals	$n_{\cdot 1} = 22$ $\bar{Y}_{\cdot 1\cdot} = 6.18$	$n_{\cdot 2} = 28$ $\bar{Y}_{\cdot 2\cdot} = 6.57$	$n_{\cdot\cdot} = 50$ $\bar{Y}_{\cdot\cdot\cdot} = 6.40$

© 2007 Thomson Higher Education

From Kleinbaum *et al.* (2008)

October 7, 2009

Two-factor ANOVA in SPSS

47

Unbalanced designs: An example

The screenshot displays the SPSS Univariate dialog box and its sub-dialog, Univariate: Model. In the Univariate dialog, the dependent variable is 'Satisfaction', and the fixed factors are 'Worry' and 'Aff_Comm'. The Univariate: Model dialog shows 'Full factorial' selected, with 'Worry(F)' and 'Aff_Comm(F)' listed as factors. The 'Sum of squares' dropdown menu is open, showing 'Type III' selected, which is highlighted by a blue box and an arrow pointing to a callout box that says 'This is the default'.

October 7, 2009

Two-factor ANOVA in SPSS

This is the default

48

Will give you score-based averages

Will give you group-based averages

October 7, 2009 Two-factor ANOVA in SPSS 49

Unbalanced designs: An example

Tests of Between-Subjects Effects

Dependent Variable: Satisfaction

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	72.000 ^a	5	14.400	4.224	.003
Intercept	1693.608	1	1693.608	496.792	.000
Worry	11.134	1	11.134	3.266	.078
Aff_Comm	44.577	2	22.289	6.538	.003
Worry * Aff_Comm	27.383	2	13.692	4.016	.025
Error	150.000	44	3.409		
Total	2270.000	50			
Corrected Total	222.000	49			

a. R Squared = .324 (Adjusted R Squared = .248)

Descriptive Statistics

Dependent Variable: Satisfaction

Worry	Aff. Comm	Mean	Std. Deviation	N
Negative	High	5.00	2.878	8
	Medium	4.00	1.414	4
	Low	8.00	1.633	10
	Total	6.18	2.684	22
Positive	High	6.00	1.758	12
	Medium	7.00	1.512	8
	Low	7.00	1.309	8
	Total	6.57	1.597	28
Total	High	5.60	2.257	20
	Medium	6.00	2.045	12
	Low	7.56	1.542	18
	Total	6.40	2.129	50

1. Worry

Dependent Variable: Satisfaction

Worry	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Negative	5.667	.424	4.812	6.522
Positive	6.667	.355	5.951	7.383

2. Aff. Comm

Dependent Variable: Satisfaction

Aff. Comm	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
High	5.500	.421	4.651	6.349
Medium	5.500	.565	4.361	6.639
Low	7.500	.438	6.617	8.383

October 7, 2009 Two-factor ANOVA in SPSS 51

Unbalanced designs

- Before we leave the topic of unbalanced designs, you should also ask why you have unequal n 's in the first place.
- If both factors are experimentally manipulated, and you attempted to recruit equal numbers, but ended up with a very unbalanced design, this suggests the loss of data is systematic.
- Perhaps the conditions associated with one of the treatment cells were more aversive than the others, leading to a greater loss of participants.
- Using Type III sums of squares could lead to very biased results in this case.
- There is no single, clearly best approach to deal with missing data due to systematic factors.

Higher-way ANOVA

- Two-way ANOVA can be generalized to any number of factors (i.e., IVs).
- Adding one factor leads to a three-way ANOVA and the concept of a three-way interaction (a.k.a. second-order interaction or triple interaction).
- The three-way interaction is distinctly more difficult to understand than the two-way interaction.
- Three variables are said to interact when the *interaction* of two of the variables is not the same at all levels of the *third* variable.
- Thus, the AxB interaction at level c_1 (called a simple interaction) is different than the AxB interaction at level c_2 (or the AxC interaction is different at different levels of B or the BxC interaction is different at different levels of A – if any one set reflects a three-way interaction, so do all of the others).

October 7, 2009

Two-factor ANOVA in SPSS

53

Higher-way ANOVA

- When no three-way interaction is present, the simple interactions are the same, and they also agree with the two-way interaction in the two-way configuration of marginal means.
- It is important to plot the interaction. Sometimes you are interested in examining one variable as the moderator of the interaction – but sometimes you don't know this. In the latter case, you may find that one plot (AxB at c_1 , AxC at b_1 , BxC at a_1) reveals the three-way interaction more effectively than the others.
- In general, you don't see very many designs in the social sciences greater than the three-factor design. This is not only because of the complexity of interpreting a four-way interaction (i.e., the interaction of any three variables is a function of the level of a fourth variable – say what?), but also because it becomes increasingly difficult to ensure sufficient numbers of observations in all of the cells (it is possible to design experimental studies in which only a fraction of the total number of possible cells need be used, but these designs do not permit testing of all possible effects).

October 7, 2009

Two-factor ANOVA in SPSS

54

Higher-way ANOVA

- From Keppel and Wickens (2004): "...higher-order effects are not easy to understand. Anyone – the present authors emphatically included! – who has attempted to understand a three-way interaction (or three two-way interactions) has had to struggle to find a satisfactory way to think about it. There is no single set of rules that ensures a clear explanation. The overall significance tests help you to decide when you can simplify the picture, but they are limited guides. You should begin by looking at the two-way tables created by holding one factor fixed and treating them as if they were actual two-factor designs. What you do next depends on what these tables (and the tests of their effects) show. You may need to look at simple effects, simple interactions, or some combination of them."