Chapter 13: One-Way Analysis of Variance

Two statisticians, one a newlywed, are sitting in a restaurant eating lunch. The first one says, "How do you like being married?" The second one says, "It's okay, but you lose a degree of freedom!" - Anonymous

Learning Objectives

Upon completion of this chapter, students should know

- When to use a one-way analysis of variance test.
- Why multiple *t* tests cannot be used when there are more than two levels of an independent variable.
- How variation (variance estimates) is partitioned between and within groups.
- The interpretation of the F-ratio.
- How to compute and interpret one-way analysis of variance.
- The purpose of post-hoc tests.
- How to compute and interpret the Tukey HSD

Key Terms

One-way analysis of variance is a test of the null and research hypotheses when the research design involves a comparison of three or more levels of one independent variable.

Mean square is a variance estimate and it is the average of the squared deviation scores used to calculate the variation.

***F* ratio** is the ratio of the mean square between groups estimate and mean square within groups estimate. If the null hypothesis is true, this ratio will be 1.

***F* test** indicates if there is difference between at least two and possibly more of the groups such that the null hypothesis can be rejected. It does not indicate where the difference lies.

Sum of squares is similar to the sum of deviation scores.

Source table is a table that displays the source of variation, sums of squares, degrees of freedom, mean squares, the *F* ratio, and *p* value.

Post-hoc test is a test performed if the F ratio is significant to determine where the significant difference(s) lie. *Post hoc* is Latin for “after the fact.”

Tukey HSD is a post hoc test. The HSD stands for honestly significant difference. It is used to compare sample means when an analysis of variance leads to a significant *F*. It reveals how far apart the sample means must be in order to be significantly different.

Lecture and Demonstration Ideas

Analysis of variance is used when there are more than two levels of a single independent variable. Students may be unclear as to why multiple *t* tests cannot be used to compare more than two means. You may want to discuss the effect this has on Type I errors (Transparency 13-1). Some refer to this as experiment-wise error or alpha-slippage.

When students see the formulas for one-way and two-way ANOVAs, they tend to cringe. One thing I have done in my classes is encourage students to keep the secret – "the math is easy" – instead, let their significant other or sibs see these formulas and maybe ice-tea, sandwiches, or other goodies will be delivered while they work their homework problems.

Memory and Levels of Processing: To introduce students to one-way analysis of variance, use the data collected during one of my class studies on memory (shown on Transparency 13-2). In this study, participants were asked to remember a list of words by either saying the word to themselves, counting the number of consonants, or using the word in a sentence. The words were presented in random order on a computer screen at three-second intervals. It was predicted that participants would remember the most words when the word was more elaboratively processed (sentence) than the other conditions. The solution is shown on Transparency 13-3 and 13-4.

Active-Learning Activities

Magazine Readability Study. To let students practice computing one-way analysis of variance, copy and distribute Handouts 13-A and 13-B. Have students either work independently or in pairs. The solution is shown on Transparency 13-5. Ask students to briefly interpret the results and include the mean and standard deviation in their write-ups.

References:

Shuptrine, F. K., & McVicker, D. D. (1981). Readability levels of magazine ads. *Journal of Advertising Research, 21*, 47-52.

Handout 13-A.

One-Way Analysis of Variance - Formula

$$SS_{bg} = \left[\frac{(\sum X_1)^2}{n_1} + \frac{(\sum X_2)^2}{n_2} + \cdots + \frac{(\sum X_k)^2}{n_k} \right] - \left[\frac{(\sum X_1 + \sum X_1 + \cdots + \sum X_1)^2}{N_{total}} \right]$$

$$SS_{wg} = \left[\sum X_1^2 + \sum X_2^2 + \cdots + \sum X_k^2 \right] - \left[\frac{(\sum X_1)^2}{n_1} + \frac{(\sum X_2)^2}{n_2} + \cdots + \frac{(\sum X_k)^2}{n_k} \right]$$

$$SS_{total} = \left[\sum X_1^2 + \sum X_2^2 + \cdots + \sum X_k^2 \right] - \left[\frac{\left(\sum X_1 + \sum X_1 + \cdots + \sum X_1\right)^2}{N_{total}} \right]$$

$$SS_{total} = SS_{bg} + SS_{wg}$$

$$F = \frac{MS_{bg}}{MS_{wg}} \qquad MS_{bg} = \frac{SS_{bg}}{df_{bg}} \qquad MS_{wg} = \frac{SS_{wg}}{df_{wg}}$$

$$df_{bg} = k - 1 \qquad df_{wg} = (n_1 - 1) + (n_2 - 1) + (n_3 - 1) + \cdots + (n_k - 1)$$

$$HSD = q \cdot \sqrt{\frac{MS_{wg}}{n}}$$

Handout 13-B.

Magazine Educational Level			
Highest Educational Level			
34.00	24.00	32.00	5.00
21.00	39.00	17.00	6.00
37.00	10.00	3.00	6.00
31.00	17.00	10.00	
10.00	18.00	6.00	
Medium Educational Level			
13.00	2.00	29.00	15.00
22.00	9.00	26.00	3.00
25.00	3.00	5.00	8.00
3.00	.00	5.00	
5.00	4.00	24.00	
Lowest Educational Level			
16.00	10.00	9.00	10.00
9.00	11.00	43.00	24.00
10.00	12.00	13.00	15.00
3.00	13.00	14.00	
12.00	1.00	32.00	

Source:

Shuptrine, F. K., & McVicker, D. D. (1981). Readability levels of magazine ads. *Journal of Advertising Research, 21*, 47-52.

Transparency 13-1.

The Effect of Multiple *t* Tests on Type I Error

Alpha set at .05

Number of *t* Tests	α
1	.05
2	.10
3	.15

Transparency 13-2.

Level of Processing and the Number of Words Correctly Recalled					
Low Condition		**Medium Condition**		**High Condition**	
Count	12	Word	19	Sentence	18
Count	8	Word	17	Sentence	22
Count	11	Word	18	Sentence	21
Count	9	Word	14	Sentence	28
Count	12	Word	18	Sentence	25
Count	9	Word	18	Sentence	22
Count	14	Word	17	Sentence	21
Count	14	Word	11	Sentence	18
Count	5	Word	26	Sentence	20
Count	8	Word	11	Sentence	27
Count	9	Word	14	Sentence	19
Count	6	Word	11	Sentence	18

Transparency 13-3.

Level of Processing and the Number of Words Correctly Recalled		
Low Condition	**Medium Condition**	**High Condition**
$\sum X_1 = 117$	$\sum X_2 = 194$	$\sum X_3 = 259$
$\sum X_1^2 = 1,233$	$\sum X_2^2 = 3,342$	$\sum X_3^2 = 5,721$
$(\sum X_1)^2 = 13,689$	$(\sum X_2)^2 = 37,636$	$(\sum X_3)^2 = 67,081$

Source Table

Source	Sums of Squares	*df*	Mean Square	*F*	*p*
Between	842.167	*2*	421.083	32.404	< .05
Within	428.833	*33*	12.995		
Total	1,271.000	*35*			

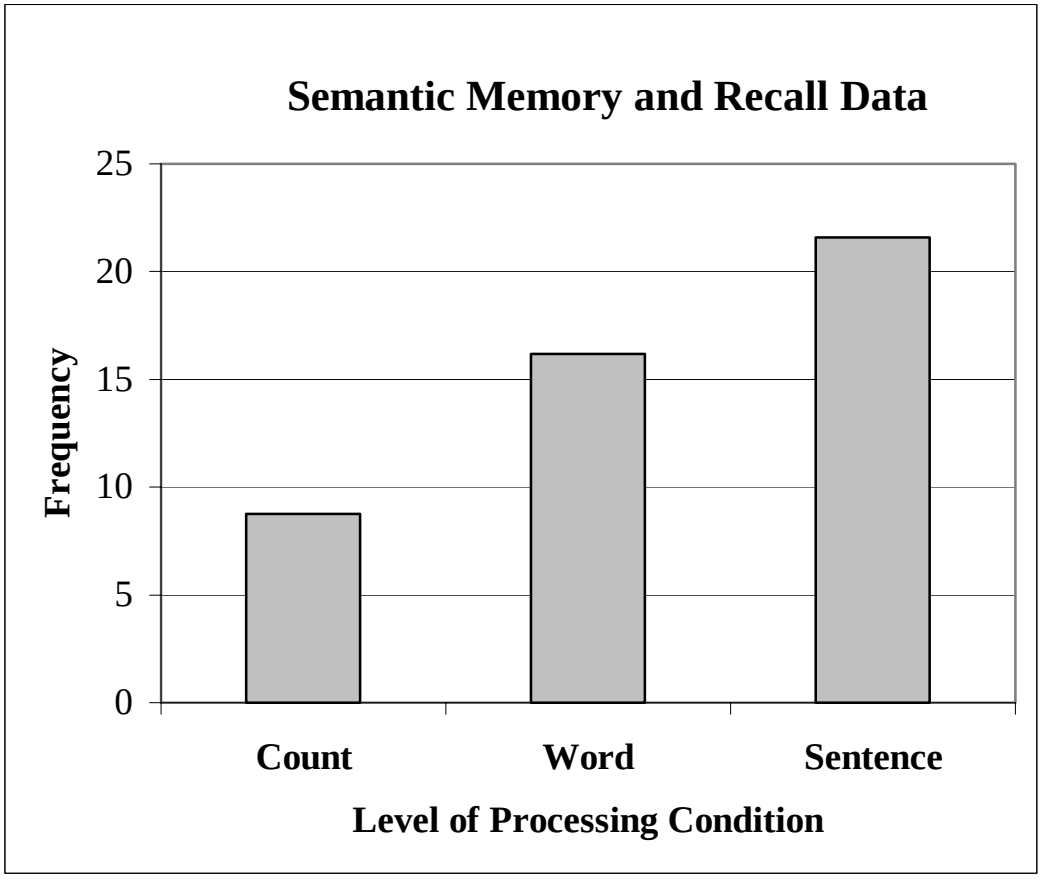
HSD = $3.49 \cdot \sqrt{\frac{12.995}{12}} = 3.632$

Count mean: 9.75 – Word mean: 16.167 = 6.417*

Count mean: 9.75 – Sentence mean: 21.583 = 11.833*

Word mean: 16.167 – Sentence mean: 21.583 = 5.416*

Transparency 13-4



Transparency 13-5.

Magazine Educational Level		
High	**Medium**	**Low**
$\sum X_1 = 326$	$\sum X_2 = 201$	$\sum X_3 = 257$
$\sum X_1^2 = 8,392$	$\sum X_2^2 = 3,863$	$\sum X_3^2 = 5,345$
$(\sum X_1)^2 = 106,276$	$(\sum X_2)^2 = 40,401$	$(\sum X_3)^2 = 66,049$

Source Table

Source	Sums of Squares	*df*	Mean Square	*F*	*p*
Between	435.593	*2*	217.796	.157	> .05
Within	5,781.889	*51*	113.370		
Total	6,217.482	*53*			

Highest Educational Level: Mean = 18.111 S = 12.097

Medium Educational Level: Mean = 11.167 S = 9.757

Lowest Educational Level: Mean = 14.278 S = 9.928