

## One-Way ANOVA Laboratory

*Purpose:* This first laboratory is intended as a simple introduction to using the computers and to the sorts of analyses that one would conduct on simple independent groups designs.

1. Let's start with some simple computations of basic statistics on three groups.

A developmental psychologist is examining problem-solving ability for grade school children. Random samples of 5-year-old, 6-year-old, and 7-year-old children are obtained with  $n = 3$ . Each child is given a standardized problem-solving task and the psychologist records the number of errors. The data (G&W) are as follows:

	<b>5-YEAR-OLDS</b>	<b>6-YEAR-OLDS</b>	<b>7-YEAR-OLDS</b>	
	5	6	0	
	4	4	1	
	6	2	2	Sum
$\Sigma Y$				
$\bar{Y}$				XXXXX
$\Sigma Y^2$				XXXXX
$SS$				
$s^2$				
$s$				XXXXX
$s_{\bar{Y}}$				XXXXX

Later in the lab, we'll use the bracket terms to compute an ANOVA on these data. For now, however, let's see how we might compute an ANOVA using the summary information.

$MS_{\text{Treatment}}$  ( $MS_A$ ) is the variance of the three means multiplied by the sample size ( $n$ ).

$$MS_A = ns_A^2 = n \frac{SS_{\bar{Y}}}{k-1} =$$

$$SS_{\bar{Y}} = \sum \bar{Y}^2 - \frac{(\sum \bar{Y})^2}{k} =$$

$MS_{\text{Error}}$  ( $MS_{S/A}$ ) is simply the mean of the three group variances, which provides you with an estimate of population variance ( $\hat{\sigma}^2$ ).

$$MS_{S/A} = \frac{s_{5\text{-Year}}^2 + s_{6\text{-Year}}^2 + s_{7\text{-Year}}^2}{3}$$

You are now able to compute the  $F$ -ratio.

$$F = \frac{MS_A}{MS_{S/A}} =$$

2. To further reinforce the conceptual notions that underlie the ANOVA, let's deal with summary data and demonstrate that you can compute an ANOVA from the summary data.

Here are the summary data for an experiment with  $n = 25$ :

	<b>a<sub>1</sub></b>	<b>a<sub>2</sub></b>	<b>a<sub>3</sub></b>
Mean	5	10	15
Variance	150	150	300

Complete the source table below. You'll need to recall that for  $k$  treatment levels (3 in this case),  $MS_{\text{Treatment}}$  is the variance of the  $k$  group means multiplied by the sample size ( $n = 25$ ).

You'll also need to recall that  $MS_{\text{Error}}$  is the mean of the separate group variances (as a way of estimating the population variance through pooling variances—assuming homogeneity of variance).

<b>Source</b>	<b>SS</b>	<b>df</b>	<b>MS</b>	<b>F</b>
A				
S/A				
Total				

Determine if the  $F$ -ratio is significant. (You remember how to use the table of critical  $F$ -values, right?)

What could you conclude from your analyses at this point?

3. Now, suppose that your data had been:

	<b>a<sub>1</sub></b>	<b>a<sub>2</sub></b>	<b>a<sub>3</sub></b>
Mean	5	10	15
Variance	300	300	600

Can you predict the  $F$ -ratio without doing a lot of computation? What would the  $MS_{Treatment}$  be?

How would the  $MS_{Error}$  compare to the first analysis?

Source	SS	df	MS	F
A				
S/A				
Total				

Can you see the implications of the within-group variability on the  $F$ -ratio (and power)? Which of the two outcomes would be more powerful?

4. Now, suppose that the original data had been obtained from a sample size of  $n = 5$ . How would that change affect your ANOVA and interpretation?

	<b>a<sub>1</sub></b>	<b>a<sub>2</sub></b>	<b>a<sub>3</sub></b>
Mean	5	10	15
Variance	150	150	300

Source	SS	df	MS	F
A				
S/A				
Total				

5. To ensure that you're comfortable with computing an ANOVA from raw data using K&W's procedure, do so for the data set from Exercise 1:

	<b>5-YEAR-OLDS</b>	<b>6-YEAR-OLDS</b>	<b>7-YEAR-OLDS</b>	
	5	6	0	
	4	4	1	
	6	2	2	Sum
$\Sigma Y$	15	12	3	30
$\Sigma Y^2$	77	56	5	138

Compute the bracket terms:

$[A] = \frac{\sum A_j^2}{n}$	$[Y] = \sum Y_{ij}^2$	$[T] = \frac{T^2}{an}$

Compute the SS:

$SS_A = [A] - [T]$	
$SS_{S/A} = [Y] - [A]$	
$SS_T = [Y] - [T]$	

Complete the summary table:

<b>Source</b>	<b>SS</b>	<b>df</b>	<b>MS</b>	<b>F</b>
A				
S/A				
Total				

Note that you could determine whether or not the  $F$ -ratio is significant, but you cannot know (at this time) what actually happened in your study. You would need to compute post hoc tests, which we will discuss soon.

Examine the  $SS_{S/A}$  and compare it to the sum of the  $SS$  for each of the three groups.

6. The data from K51 (K51.sav) should be in a folder called PS318 (under Course Files). Locate the data (Ind Groups/One-Way/SPSS Files) and use SPSS to produce an ANOVA using One-Way ANOVA. Note that these data are not yet fully interpretable, because there are more than two groups.

You should also find a data set called K51mod.sav. This data set is the one found in the notes for Chapter 3. Analyze these data to convince yourself of the importance of minimizing the within-group variance.

7. Now, to ensure that you're comfortable with using the computer for data analyses, let's first use the computer to analyze data that you input yourself. Use SPSS to analyze the following data:

A psychologist studying human memory would like to examine the process of forgetting. One group of participants is required to memorize a list of words in the evening just before going to bed. Their recall is tested 10 hours later in the morning. Participants in the second group memorize the same list of words in the morning and then their memories are tested 10 hours later after being awake all day. The psychologist hypothesizes that there will be less forgetting during sleep than during a busy day. The recall scores and some summary information for the two samples of college students are seen below. Interpret the results of this study as completely as you can.

	<b>Asleep Scores</b>	<b>Awake Scores</b>
	15	15
	13	13
	14	14
	14	12
	16	14
	15	13
	16	11
	15	12
	16	13
	15	13
	17	12
	14	14

Note that you can fully interpret the results of this study because you have only two groups. What would you conclude?

8. Although you cannot always influence the variability found within your groups, there are occasions on which you are able to take steps to minimize the within-group variability. (Some strategies involve restricting the nature of sample members (e.g., all females), but such strategies are almost always avoided because of the restriction of generalizability. Other strategies are more reasonable, such as minimizing the noise in the experimental room, giving more explicit instructions, etc.) Suppose that you are dealing with a particular Treatment Effect Size (let's say 5). What would your  $F$ -ratio be if you were to conduct an experiment in which the Individual Differences and Random Variability = 5?

What would your  $F$ -ratio be if you were to conduct an experiment in which the Individual Differences and Random Variability = 100?

In comparing these two experimental settings, can you see why it's so important to keep the Individual Differences and Random Variability as small as possible?

9. The summary information for ANOVA used in APA-style papers actually allows you to fully construct the ANOVA table for one-way independent groups designs. Do so for the following summary:

$$F(2,27) = 10, MSE = 2, p < .05$$

<b>Source</b>	<b><i>SS</i></b>	<b><i>df</i></b>	<b><i>MS</i></b>	<b><i>F</i></b>
<i>A</i>				
<i>S/A</i>				
Total				