CHAPTER 11

ONE-WAY ANALYSIS OF VARIANCE WITH POST HOC COMPARISONS

OBJECTIVES

After completing this chapter, you should

- be able to perform a significance test (ANOVA) involving two or more levels of one independent variable, when the samples are either independent (between-subjects) or dependent (within-subjects or repeated measures).
- be able to perform two different post hoc tests after a significant ANOVA.

CHAPTER REVIEW

The analysis of variance, or ANOVA, is a widely used test for comparing more than two groups. Two reasons for not using the two-sample *t* test are that multiple *t* tests are tedious to compute and that the more tests you do on the same data, the more likely you are to commit a Type I error (reject a true null).

The total variability in some data can be partitioned or divided into the *within-groups variability* and the *between-groups variability.* The variability within each group stems from individual differences and experimental error; the variability between groups comes from individual differences, experimental error, and the treatment effect. The ANOVA test is the ratio of a measure of variability between groups to a measure of the variability within groups. If there is no treatment effect, the computed value of F will be close to 1. However, if there is a treatment effect, the F ratio will be relatively large because of the added source of variability contributing to the between-group differences. One-way between-subjects ANOVA applies to situations in which the data from three or more independent groups are analyzed.

The first step in determining the indices of variability is to compute the sums of squares. The *total sum of squares* is the sum of the squared deviations of each score from the total mean. The *sum of squares within each group* is the sum of the squared deviations of each score in a group from its group mean, with the deviations summed across groups. Finally, the *sum of squares between groups* can be obtained by subtraction: $SS_b = SS_{tot} - SS_w$. Also, SS_b is the square of the deviation between each group mean and the total mean multiplied by the number of subjects in a particular group and summed over groups. It's a good idea to compute SS_b to test the accuracy of your other computations.

After the sums of squares have been determined, an appropriate degrees of freedom is computed for each. For SS_{tot}, or the total sum of squares, $df = N - 1$, where N is the total number of cases sampled. For SS_b, or the sum of squares between groups, $df = K - 1$, where K is the number of groups. df for SS_w, or the sum of squares within groups, is $N - K$.

Both SS_b and SS_w are divided by their respective df to give the average or *mean square.* The ratio of MS_b to MS_w is called the F ratio. A relatively large value of F indicates greater variability between groups than within groups and may indicate sampling from different populations. The computed value of F is compared with values known to cut off deviant portions (5% or 1%) of the distribution of F. If the computed F exceeds critical values from Table C (see Appendix 2), the null hypothesis is rejected, and we conclude that at least one of the samples probably came from a different population. To help summarize the results, as they are computed, values are entered into the analysis of variance summary table shown here.

Summary Table for Between-Subjects ANOVA

Source	*SS*	*df*	*MS*	*F*
Between groups				
Within groups				
Total				

Two tests are presented for further significance testing following a significant F ratio: the Fisher LSD and the Tukey HSD. Both tests are used to make all pairwise comparisons—comparing all groups by looking at one pair at a time. The LSD test is sometimes called a *protected t test* because it follows a significant F test. In the LSD test, the difference between a pair of means is significant if it is greater than LSD, which is computed with a formula; the same is true for the HSD test; that is, a difference between a pair of means is significant if the difference exceeds the computed value of HSD. A table of differences is used to summarize the results of both tests.

The one-way repeated measures ANOVA applies to situations in which the same (or matched) participants are tested on more than two occasions. The first step is to compute the sums of squares. The total and between-groups sums of squares are computed using the same procedures as in one-way between-subjects ANOVA. However, the within-groups sum of squares is divided into two parts: subjects sum of squares (SS_{subj}) and error sum of squares (SS_{error}). SS_{subj} is the squared deviation between the mean score for each subject and the total mean, multiplied by the number of groups and summed over subjects. SS_{error} is the variability remaining after removing SS_b and SS_{subj} from SS_{tot} and can be obtained by subtraction: $SS_{error} = SS_{tot} - SS_b - SS_{subj}$. Computational formulas were given for each of the sums of squares.

As in one-way between-subjects ANOVA, $df_{tot} = N - 1$, and $df_b = K - 1$. Subjects degrees of freedom (df_{subj}) equal the number of subjects minus 1 ($S - 1$), and error degrees of freedom (df_{error}) equal $(K - 1)(S - 1)$.

Both SS_b and SS_{error} are divided by the appropriate df to give MS_b and MS_{error}, respectively. The F ratio is obtained by dividing MS_b by MS_{error}. If the computed F is greater than or equal to the critical values from Table C (Appendix 2), the null hypothesis is rejected. With slight modifications, the LSD and HSD tests can be used for post hoc testing following a significant repeated measures ANOVA.

To summarize the results, values are entered in a summary table, as shown here.

Summary Table for One-Way Repeated Measures ANOVA

Source	*SS*	*df*	*MS*	*F*
Between groups				
Subjects				
Error				
Total				

SYMBOLS

Symbol	*Stands For*
$\overline{X}_{\text{tot}}$	total mean or grand mean (GM)
X_{g}	score within a group
$\overline{X}_g$	mean of a group
SS_{tot}	total sum of squares
SS_{w}	within-groups sum of squares
SS_{b}	between-groups sum of squares
SS_{subj}	subjects sum of squares
SS_{error}	error sum of squares
N_{g}	number of subjects within a group
N	total number of subjects or total number of scores in a repeated measures ANOVA
$\sum_g$	sum over or across groups
MS_{b}	mean square between groups
MS_{w}	mean square within groups
df_{b}	between-groups degrees of freedom
K	number of groups or number of trials in a repeated measures ANOVA
S	number of participants (subjects)
df_{w}	within-groups degrees of freedom
df_{tot}	total degrees of freedom
df_{subj}	subjects degrees of freedom
df_{error}	error degrees of freedom
F	F ratio, ANOVA test
F_{comp}	your computed F ratio
F_{crit}	the critical value of F from Table C
LSD	least significant difference
HSD	honestly significant difference
q	studentized range statistic
LSD_α, HSD_α	LSD and HSD mean difference values required for significance at a particular α level (.05 or .01, usually)

FORMULAS

Before solving any of the formulas introduced, the following values need to be computed for the data: ΣX_{g}, ΣX_{g}^2, N_{g}, ΣX, ΣX^2, and N. ΣX_{g} is the sum of the scores within each group; ΣX_{g}^2 is the sum of the squared scores within each group; N_{g} is the number of observations within each group; ΣX is the sum of all the scores; ΣX^2 is the sum of all the squared scores; and N is the total number of observations. In addition, for one-way repeated measures ANOVA, ΣX_{m}, $(\Sigma X_{\text{m}})^2$, S, and K must be computed. ΣX_{m} is the sum of scores for each participant; $(\Sigma X_{\text{m}})^2$ is the square of the sum of the scores for each participant; S is the number of participants; and K is the number of trials or tests.

Formula 11-5. *Computational formula for the total sum of squares*

$$SS_{\text{tot}} = \Sigma X^2 - \frac{(\Sigma X)^2}{N}$$

This equation is identical to the numerator of sample variance, which we said in Chapter 6 was sometimes called the sum of squares or SS.

Formula 11-6. *Computational formula for the within-group sum of squares*

$$SS_{\text{w}} = \sum_{\text{g}} \left[\Sigma X_{\text{g}}^2 - \frac{(\Sigma X_{\text{g}})^2}{N_{\text{g}}} \right]$$

This is just the sum of squares equation computed for each group and then summed across groups.

For three groups, the computational formula for SS_{w} becomes

$$SS_{\text{w}} = \left[\Sigma X_1^2 - \frac{(\Sigma X_1)^2}{N_1} \right] + \left[\Sigma X_2^2 - \frac{(\Sigma X_2)^2}{N_2} \right] + \left[\Sigma X_2^3 - \frac{(\Sigma X_3)^2}{N_3} \right]$$

Formula 11-7. *Computational formula for the between-groups sum of squares*

$$SS_{\text{b}} = \sum_{\text{g}} \left[\frac{(\Sigma X_{\text{g}})^2}{N_{\text{g}}} \right] - \frac{(\Sigma X)^2}{N}$$

For three groups, the computational formula for SS_{b} becomes

$$SS_{\text{b}} = \left[\frac{(\Sigma X_1)^2}{N_1} + \frac{(\Sigma X_2)^2}{N_2} + \frac{(\Sigma X_3)^2}{N_3} \right] - \frac{(\Sigma X)^2}{N}$$

Formulas 11-8, 11-9, and 11-10. *Equations for between-groups degrees of freedom, within-groups degrees of freedom, and total degrees of freedom, respectively*

$df_{\text{b}} = K - 1$

$df_{\text{w}} = N - K$

$df_{\text{tot}} = N - 1$

Formula 11-11. *Equation for the between-groups mean square*

$$MS_{\text{b}} = \frac{SS_{\text{b}}}{df_{\text{b}}}$$

Formula 11-12. *Equation for the within-groups mean square*

$$MS_{\text{w}} = \frac{SS_{\text{w}}}{df_{\text{w}}}$$

Formula 11-13. *Equation for F ratio in one-way between-subjects ANOVA*

$$F = \frac{MS_{\text{b}}}{MS_{\text{w}}}$$

Formula 11-14. *Least significant difference* (*LSD*) *between pairs of means*

$$LSD_{\alpha} = t_{\alpha}\sqrt{MS_{\mathrm{w}}\left(\frac{1}{N_1}+\frac{1}{N_2}\right)}$$

Formula 11-15. *Honestly significant difference* (*HSD*) *between pairs of means*

$$HSD_{\alpha} = q_{\alpha}\sqrt{\frac{MS_{\mathrm{w}}}{N_{\mathrm{g}}}}$$

Formula 11-18. *Computational formula for within-subjects sum of squares in one-way repeated measures ANOVA*

$$SS_{\mathrm{subj}} = \sum_{\mathrm{s}}\left[\frac{(\Sigma X_{\mathrm{m}})^2}{K}\right] - \frac{(\Sigma X)^2}{N}$$

For three subjects, the computational formula for SS_{subj} becomes

$$SS_{\mathrm{subj}} = \left[\frac{(\Sigma X_{\mathrm{s}_1})^2}{K}+\frac{(\Sigma X_{\mathrm{s}_2})^2}{K}+\frac{(\Sigma X_{\mathrm{s}_3})^2}{K}\right] - \frac{(\Sigma X)^2}{N}$$

Formula 11-19. *Computational formula for error sum of squares in one-way repeated measures ANOVA*

$$SS_{\mathrm{error}} = SS_{\mathrm{tot}} - SS_{\mathrm{b}} - SS_{\mathrm{subj}}$$

Formula 11-20. *Computational formula for error degrees of freedom*

$df_{\mathrm{error}} = (K - 1)(S - 1)$

Formula 11-21. *Computational formula for mean square error in one-way repeated measures ANOVA*

$$MS_{\mathrm{error}} = \frac{SS_{\mathrm{error}}}{df_{\mathrm{error}}}$$

Formula 11-22. *Computational formula for F ratio in one-way repeated measures ANOVA*

$$F = \frac{MS_{\mathrm{b}}}{MS_{\mathrm{error}}}$$

The degrees of freedom for the F ratio are the df associated with the numerator ($df_{\mathrm{b}} = K - 1$) and df associated with the denominator [$df_{\mathrm{error}} = (K - 1)(S - 1)$].

TERMS TO DEFINE AND/OR IDENTIFY

ANOVA

one-way ANOVA

additivity

key deviations

total variability

within-groups variability

individual differences

experimental error

between-groups variability

treatment effect

total sum of squares

within-groups sum of squares

between-groups sum of squares

ANOVA summary table

mean square

F ratio

post-ANOVA tests

a posteriori test

post hoc test

a priori test

Fisher LSD test

protected *t* test

pairwise comparisons

Tukey HSD test

studentized range statistic

repeated measures ANOVA

FILL-IN-THE-BLANK ITEMS

Introduction

The *t* test in Chapter 10 was used to compare the results from (1) _________________ samples in order to see whether they were drawn from (2) _________________ populations. One important technique to compare the results from two or more groups is the (3) _________________ _____ _________________. Two reasons not to apply the *t* test to results from more than two groups are that the computations would be (4) _________________ and the more tests you do on the same data, the greater the likelihood of committing a Type (5) _________________ error, rejecting a (6) _________________ null hypothesis.

Between-Subjects ANOVA

Like the *t* test, one-way ANOVA has two versions: a (7) _________________ ANOVA, which parallels the independent *t*, and a (8) _________________ _________________ ANOVA, which parallels the dependent *t*. The *t* score is a measure of the distance between a group mean and a (9) _________________ mean, or another group (10) _________________ in standard (11) _________________ terms. One of the reasons that variance can be used to determine whether more than two groups differ is the property of (12) _________________. According to the property of additivity, the variance of the sum of independent scores is equal to the (13) _________________ of the variances of the scores. Because variance is additive, we can divide the total variability in a set of scores into its (14) _________________ _________________.

Visualization of ANOVA concepts

The two components of variability in which we're interested are based on within-groups variability and (15) _________________ variability, and the (16) _________________ _________________ for these components are ($X_g - \overline{X}_g$) and ($\overline{X}_g - \overline{X}_{tot}$), respectively. $\overline{X}_{tot}$ is often called the (17) _________________ _________________. To analyze the variances, we are interested in comparing the (18) _________________ variability to the within-groups variability. If the between-groups variability is (19) _________________ relative to within-groups variability, we will probably conclude that the treatments had an effect.

Everyday ANOVA.

The static on your cellular phone or the distracting noise at a party is analogous to (20) _________________ variability; the message or signal you're trying to detect is analogous to (21) _________________ variability.

Three sources of variability in some data are discussed: (22) _________________ variability, between-groups variability, and (23) _________________ variability. The variability within groups is caused by experimental error and (24) _________________ _________________. The variability between groups comes from experimental error, individual differences, and the (25) _________________. The ANOVA test is the ratio of the variability (26) _________________ groups to the variability (27) _________________ groups. If there is no treatment effect, the F ratio will be near (28) _________________, whereas a treatment effect will make the statistic relatively (29) _________________.

Measuring variability: The sum of squares

The (30) _________________ sum of squares is the sum of the squared deviation of each score from the total mean. The sum of squares (31) _________________ groups is the sum of the squared deviations of each group score from a group mean, with the deviations summed across groups. The sum of squares (32) _________________ groups is based on the deviation between each group mean and the total mean.

Computing the sums of squares

Although computation of the sums of squares can be tedious, the "trick" is to first compute the sum of the (33) _________________ in each group, the sum of the (34) _________________ _________________ in each group, the (35) _________________ sum of scores, the total sum of squared scores, the number of subjects per group, and the (36) _________________ number of subjects. The symbols are ΣX_g, (37) _________________, ΣX, (38) _________________, N_g, and N, respectively.

Remember that variance is additive. In one-way between-subjects ANOVA, once SS_{tot} and one of its components (either SS_b or SS_w) have been computed, the other component can be found by (39) _________________, although the value should be computed as a check on the accuracy of your calculations.

The analysis of variance summary table

The ANOVA summary table provides a place for the sums of squares; the (40) _________________ for each of the sums of squares; the (41) _________________ squares, which are computed by dividing each SS by its df; and the (42) _________________ _________________. F is computed by dividing the MS_b by (43) _________________. df_b is equal to $K - 1$, where K is the number of (44) _________________. df_w = (45) _________________. If the computed F ratio is larger than the critical F ratio from Table

(46) _________________, the null hypothesis is (47) _________________. Instead of being symmetrical like the t distributions, the F distributions are (48) _________________ skewed with a peak around (49) _________________. (50) _________________ tests are tests that follow a significant F ratio.

Post Hoc Comparisons

There are many post hoc tests available that avoid the problem of inflation of Type I error by (51) _________________ the critical value needed to reject H_0. Because a posteriori or post hoc tests follow a significant F ratio, they are also called (52) _________________ tests. On the other hand, (53) _________________ _________________ tests are tests designed to look at specific hypotheses before the experiment is performed. When the experimenter cannot predict the patterning of (54) _________________ before the research is performed, post hoc tests are appropriate.

The Fisher LSD

As presented in the text, the LSD test does not require equal (55) _________________ _________________. Also, the LSD test is a (56) _________________ test, which means that we are more likely to be able to reject the null hypothesis with it than with many other post hoc tests available. The LSD test is sometimes called a (57) _________________ t test because it follows a significant (58) _________________ _________________. The significant F ratio tells us that there is at least one (59) _________________ comparison, thus protecting the error rate. With the test, the difference between two sample means is significant if it is greater than (60) _________________, which is found with the following formula: (61) _________________. As before, (62) _________________ is the level of significance, and the value of t is obtained from Table (63) _________________. The results of the Fisher LSD are best summarized in a

(64) _________________ _________________ ___________________.

The Tukey HSD

Although the Tukey HSD test can be used for more complex comparisons, we used it for making all (65) _________________ comparisons when the sample sizes are (66) _________________. Like the

Fisher LSD, the difference between two sample means is significant if it is greater than

(67) _________________, which is found with the following formula: (68) _________________. The

value of q comes from the distribution of the (69) _________________ _________________

_________________, whose critical values are found in Table (70) _________________. HSD stands for

(71) _________________ _________________ _________________.

Repeated Measures ANOVA

Repeated measures ANOVA is appropriate in situations in which the (72) _________________ participants are measured on more than (73) _________________ occasions. In repeated measures ANOVA, each participant serves as his or her own (74) _________________. By using a person as his or her own control, we are able to extract some of the (75) _________________ from our scores.

There are two sources of variability that contribute to SS_w: experimental (76) _________________ and variability in (77) _________________. Thus, $SS_w = SS_{subj}$ + (78) _________________. SS_{error} is used as the (79) _________________ in computing the F ratio in one-way repeated measures ANOVA. Because the property of additivity applies, $SS_{tot} = SS_b + SS_{subj}$ + (80) _________________. As compared to one-way between-subjects ANOVA, the additional step in one-way repeated measures ANOVA is the computation of (81) _________________ sum of squares.

The summary table for one-way repeated measures ANOVA is similar to that for one-way between-subjects ANOVA except that SS_{tot} is divided into (82) _________________ ____________________ instead of two. For subjects, degrees of freedom are found by subtracting one from the number of (83) _________________. Degrees of freedom for error are the product of df_b and (84) _________________ or $(K - 1)(S - 1)$.

Normally, an F ratio is not computed for (85) _________________. In one-way repeated measures ANOVA, F is found by dividing MS_b by (86) _________________.

Troubleshooting Your Computations

Two obvious signs of trouble when computing the sums of squares are a (87) _________________ sign for *SS* and failure of SS_b and SS_w to sum to (88) _________________. The most common error in filling in the summary table is determining incorrectly the (89) _________________ for each *SS*. Remember that $df_{tot} = df_b + df_w =$ (90) _________________ for between-subjects ANOVA, and $df_{tot} = df_b + df_{subj} + df_{error} =$ (91) _________________ for repeated measures ANOVA.

The most common computational error made in calculating either LSD or HSD is to use *N* instead of (92) _________________ in the expression under the radical sign. Another error that is sometimes made is to use a value from the (93) _________________ table rather than the critical (94) _________________ in the formula for LSD or instead of the (95) _________________ value in the HSD formula. Also, be sure to subtract to obtain (96) _________________ differences or to use the (97) _________________ values or your differences in the significance tests. Remember, if your computed value is equal to or (98) _________________ than the critical value, then you reject the null hypothesis for that test.

PROBLEMS

1. From a large introductory psychology class, 32 snake-phobic students were selected and randomly assigned to one of four experimental groups. Group 1 received five sessions of relaxation training; Group 2 received five sessions of imagery training (they were required to imagine each of several feared situations); Group 3 received relaxation training combined with the imagery training; Group 4 participants were told that there would be a few weeks' delay in the beginning of therapy. Three weeks from the beginning of the experiment, each participant was given a behavioral avoidance test to determine how closely he or she would approach a live snake in an aquarium. The response measure is the distance from the snake in feet.

Group 1	Group 2	Group 3	Group 4
10	8	1	10
8	8	2	9
8	7	3	9
9	5	5	8
10	4	7	10
7	4	3	7
6	6	4	8
8	3	5	9

Compute the following:

$\Sigma X_1 =$ $\Sigma X_2 =$ $\Sigma X_3 =$ $\Sigma X_4 =$ $\Sigma X =$

$\Sigma X_1^2 =$ $\Sigma X_2^2 =$ $\Sigma X_3^2 =$ $\Sigma X_4^2 =$ $\Sigma X^2 =$

$N_1 =$ $N_2 =$ $N_3 =$ $N_4 =$ $N =$

Find the sums of squares:
$SS_{\text{tot}} =$

$SS_{\text{w}} =$

Find SS_{b} by subtraction:
$SS_{\text{b}} = SS_{\text{tot}} - SS_{\text{w}} =$

Compute SS_{b}:
$SS_{\text{b}} =$

Complete the ANOVA summary table:

Source	*SS*	*df*	*MS*	*F*
Between groups				
Within groups				
Total				

The computed value of F is_________________. The *df* for the numerator is_________________, and the *df* for the denominator is_________________. The table values required for rejection of H_0 are _________________ at the 5% level and _________________ at the 1% level. What is your decision, and what does it mean in the context of the problem?

2. Use the Fisher LSD test to analyze the data in Problem 1 further.

3. Three different commercial sleeping aids and a placebo are given to four groups of randomly selected young adults. After a suitable period of time for the drugs to take effect, each participant is placed in a room with a bed, and his or her EEG is monitored. The response measured is the length of time before onset of sleep as determined by the EEG. The results are as follows: Group Placebo, $N = 9$, $\Sigma X = 29.7$, $\Sigma X^2 = 105.49$; Group Potion 1, $N = 8$, $\Sigma X = 30.4$, $\Sigma X^2 = 120.22$; Group Potion 2, $N = 9$, $\Sigma X = 32$, $\Sigma X^2 = 121.26$; Group Potion 3, $N = 8$, $\Sigma X = 30.1$, $\Sigma X^2 = 131.51$. Compute the sums of squares and fill in the summary table.

$SS_{\text{tot}} =$

$SS_{\text{w}} =$

$SS_{\text{b}} =$

ANOVA Summary Table

Source	*SS*	*df*	*MS*	*F*
Between groups				
Within groups				
Total				

Is the *F* ratio significant, and what does your conclusion mean in the context of the problem?

4. During WWII, the RAF noticed that a large number of fighter pilots were being killed because they were not dark-adapted during night air raids. An experiment was performed to determine whether different levels of preflight illumination might result in significant differences in time to dark adaptation. Twenty-four pilots were randomly and equally assigned to one of three treatment groups. Group A spent 30 minutes in a brightly lighted room; Group B spent 30 minutes in a dimly lighted room; Group C spent 30 minutes in a brightly lighted room wearing red-tinted goggles. The length of time in minutes for complete dark adaptation was recorded for each pilot. Determine whether the groups differed significantly.

Group A	Group B	Group C
$N = 8$	$N = 8$	$N = 8$
$\Sigma X = 260$	$\Sigma X = 78$	$\Sigma X = 36$
$\Sigma X^2 = 8{,}492$	$\Sigma X^2 = 788$	$\Sigma X^2 = 196$

5. Using the Fisher LSD, do all pairwise comparisons of the groups in Problem 4.

6. We hypothesize that the experience of taking a statistics course will reduce mathematics anxiety. To test this hypothesis, we select nine statistics students and assess their mathematics anxiety on four occasions: on the first day of class, after 3 weeks of class, after 6 weeks of class, and after 9 weeks of class. Perform the appropriate overall test of significance.

Student	First Day	3 Weeks	6 Weeks	9 Weeks
A	14	12	9	8
B	8	7	5	3
C	6	7	4	2
D	9	10	8	7
E	15	12	10	9
F	12	10	8	9
G	9	8	7	6
H	7	6	5	3
I	10	9	7	7

7. Use the Fisher LSD test to perform all pairwise comparisons for the data in Problem 6.

8. A manufacturing company is concerned about the effect of fatigue on the speed with which its workers can assemble pocket calculators. For 10 workers, the average time (in seconds) it takes to assemble a pocket calculator is measured at the beginning, in the middle, and at the end of the shift. Does performance change across periods of the workers' shift?

Worker	Beginning	Middle	End
A	20	21	23
B	28	30	31
C	22	23	24
D	19	19	22
E	24	26	28
F	26	27	29
G	19	18	19
H	24	25	27
I	20	21	22
J	19	21	22

9. Use the Tukey HSD to make all pairwise comparisons for the data in Problem 8.

10. In a study of dark adaptation, eight participants seated in an almost totally dark room were asked to determine visually the presence or absence of an object. All participants were given 10 trials after 1 minute of adaptation, after 15 minutes, and after 30 minutes. At each testing, the number of correct detections out of 10 trials was recorded. Perform an overall significance test, and tell what your conclusion means in the context of the problem.

Participant	1 Minute	15 Minutes	30 Minutes
A	2	6	6
B	0	2	4
C	4	7	9
D	3	5	6
E	6	8	10
F	0	2	4
G	2	5	1
H	3	5	8

11. Use the Fisher LSD test to make all pairwise comparisons and summarize your results.

12. A study was done to see whether the source of dietary fat affects visual discrimination. Rats were placed on one of four diets for 2 months: Diet 1 had 5% corn oil; Diet 2 was the same as Diet 1 with the addition of 20% safflower oil; Diet 3 was Diet 1 with 20% added coconut oil; Diet 4 was Diet 1 with 20% added olive oil. All the rats were trained on a simple visual discrimination task, and their errors before achieving a certain criterion were recorded. Test the data to see whether the different diets affected learning of the task.

Diet 1	Diet 2	Diet 3	Diet 4
13	10	7	14
20	20	17	19
31	34	11	27
18	27	23	31
11	7	14	15
11	27	13	21
11	10	26	14
12	12	23	
12	32	4	
12	11		

USING SPSS—EXAMPLES AND EXERCISES

SPSS has several techniques for performing analysis of variance (ANOVA). The ONEWAY procedure is one such method, and it will perform a variety of post hoc tests. For the repeated measures ANOVA, we will need to use the SPSS GLM (General Linear Model)–Repeated Measures procedure to obtain the analysis. The SPSS GLM procedures will perform analyses for many different types of ANOVA designs. Unfortunately—given our desire to keep this as simple as possible—the SPSS GLM procedures are some of the "fancier" SPSS techniques and provide extensive output that is well beyond the level of the textbook.

Example—Independent Groups ANOVA: As an example of an independent groups ANOVA, we will work Problem 1 using SPSS. We will illustrate how to perform the ANOVA, how to do the LSD and HSD post hoc tests, how to graph the means, and how to provide an Error Bar chart of the groups. The steps are as follows:

1. Start SPSS and enter the data. The data entry is an extension of the set-up we used for the two-sample independent-groups *t* test. Name the two variables **group** and **distance.** The group variable will have a 1 entered for each distance score from Group 1, a 2 for each distance score from Group 2, and so on.
2. Select *Analyze>Compare Means>One-Way ANOVA*.
3. Move **distance** to the Dependent List box because it is the dependent variable. Move **group** to the Factor box.
4. Select the Post Hoc box and choose LSD and Tukey (HSD) in the Post Hoc Multiple Comparisons box. Then select *Continue*.
5. Select the Options box; click *Descriptive* and *Means plot* (this will give you a line graph of the group means), then *Continue>OK*. The results should appear in the output Viewer window.
6. As an extra illustration for this exercise, we will create an Error Bar chart for the groups, which shows a plot of the confidence intervals for each group. This type of graph is helpful in understanding and emphasizing that there is within-group variability that is not shown in graphs of group means as point estimates.
7. Select *Graphs>Error Bar>Simple>Summaries for groups of cases>Define*.
8. Move **distance** into the Variable box and move **group** into the Category Axis box. Other settings in this dialog box should indicate a 95% confidence interval for means. Next click *OK,* and the graph should appear in the output Viewer window.

Notes on Reading the Output

1. The ANOVA output box gives the source table. The "Sig." after the F value is the exact probability value for the obtained F ratio. For example, $p = .000$ means that p is 0 when rounded to three decimal places. Because p is never exactly 0, it is better to express this probability as $p < .001$.
2. The Multiple Comparisons box is highly redundant. It does not give a test statistic value for each comparison or a minimum difference required for significance between two groups. Instead the box indicates the significant comparisons by an asterisk beside the Mean Difference and the exact p value given in the Sig. column. For example, the Tukey HSD results indicate that Group 1 versus Group 4 and Group 2 versus Group 3 are the only comparisons that are not statistically different. The more powerful LSD test indicates that only Groups 1 and 4 are not statistically different.
3. The Means Plots and the Graph showing confidence intervals provide pictures of the results. Because their confidence intervals overlap considerably, we would expect Groups 1 and 4 and Groups 2 and 3 not to be statistically different. Of course, this is what we found with the HSD test.

```
ONEWAY
  distance BY group
  /STATISTICS DESCRIPTIVES
  /PLOT MEANS
  /MISSING ANALYSIS
  /POSTHOC = TUKEY LSD ALPHA(.05).
```

Oneway

Descriptives

DISTANCE

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
1.00	8	8.2500	1.3887	.4910	7.0890	9.4110	6.00	10.00
2.00	8	5.6250	1.9226	.6797	4.0177	7.2323	3.00	8.00
3.00	8	3.7500	1.9086	.6748	2.1543	5.3457	1.00	7.00
4.00	8	8.7500	1.0351	.3660	7.8846	9.6154	7.00	10.00
Total	32	6.5938	2.5635	.4532	5.6695	7.5180	1.00	10.00

ANOVA

DISTANCE

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	131.344	3	43.781	16.938	.000
Within Groups	72.375	28	2.585		
Total	203.719	31			

Post Hoc Tests

Multiple Comparisons

Dependent Variable: DISTANCE

	(I) GROUP	(J) GROUP	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Tukey HSD	1.00	2.00	2.6250*	.804	.014	.4302	4.8198
		3.00	4.5000*	.804	.000	2.3052	6.6948
		4.00	-.5000	.804	.924	-2.6948	1.6948
	2.00	1.00	-2.6250*	.804	.014	-4.8198	-.4302
		3.00	1.8750	.804	.115	-.3198	4.0698
		4.00	-3.1250*	.804	.003	-5.3198	-.9302
	3.00	1.00	-4.5000*	.804	.000	-6.6948	-2.3052
		2.00	-1.8750	.804	.115	-4.0698	.3198
		4.00	-5.0000*	.804	.000	-7.1948	-2.8052
	4.00	1.00	.5000	.804	.924	-1.6948	2.6948
		2.00	3.1250*	.804	.003	.9302	5.3198
		3.00	5.0000*	.804	.000	2.8052	7.1948
LSD	1.00	2.00	2.6250*	.804	.003	.9783	4.2717
		3.00	4.5000*	.804	.000	2.8533	6.1467
		4.00	-.5000	.804	.539	-2.1467	1.1467
	2.00	1.00	-2.6250*	.804	.003	-4.2717	-.9783
		3.00	1.8750*	.804	.027	.2283	3.5217
		4.00	-3.1250*	.804	.001	-4.7717	-1.4783
	3.00	1.00	-4.5000*	.804	.000	-6.1467	-2.8533
		2.00	-1.8750*	.804	.027	-3.5217	-.2283
		4.00	-5.0000*	.804	.000	-6.6467	-3.3533
	4.00	1.00	.5000	.804	.539	-1.1467	2.1467
		2.00	3.1250*	.804	.001	1.4783	4.7717
		3.00	5.0000*	.804	.000	3.3533	6.6467

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

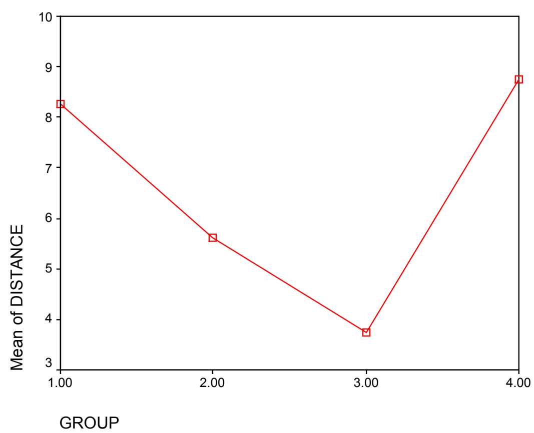
DISTANCE

	GROUP	N	Subset for alpha = .05	
			1	2
Tukey HSD[a]	3.00	8	3.7500	
	2.00	8	5.6250	
	1.00	8		8.2500
	4.00	8		8.7500
	Sig.		.115	.924

Means for groups in homogeneous subsets are displayed.

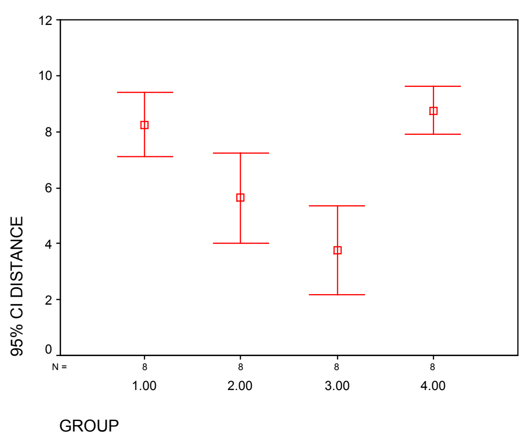
a. Uses Harmonic Mean Sample Size = 8.000.

Means Plots




```
GRAPH
  /ERRORBAR( CI 95 )=distance BY group
  /MISSING=REPORT.
```

Graph

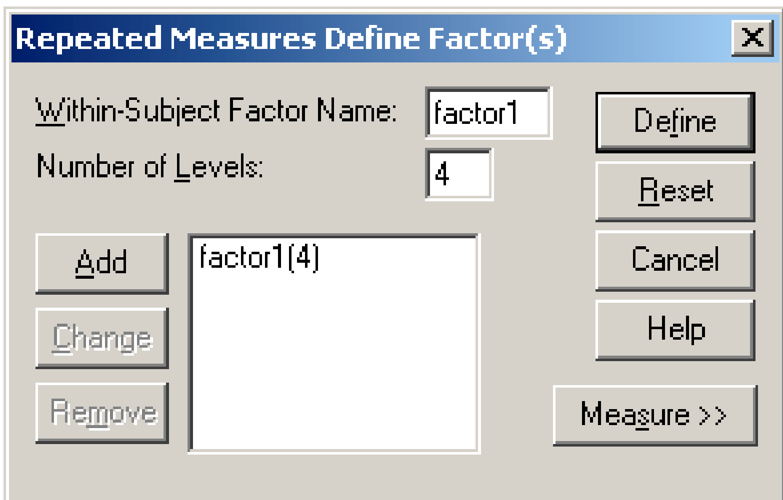


Example—Repeated Measures ANOVA: Our example of the repeated measures ANOVA will use the SPSS GLM–Repeated Measures procedure. This technique will perform the analysis for the one-way repeated measures exercises in the text and in this study guide in addition to analyzing more extensive designs. The procedure does not do post hoc tests for this completely within-subjects design, so we will not be concerned with this part of our computer solution. The supplemental text on using SPSS suggested in Appendix 4 explains how to do such tests. (In short, the post hoc tests are performed by computing sequential pairwise dependent *t* tests and testing for significance by using the α level obtained by dividing .05 by the number of tests performed.) Alternatively, you can compute the post hoc tests by hand, using information from the output and the procedures described in the text. We will solve Problem 6 as an example. Here are the steps to follow:

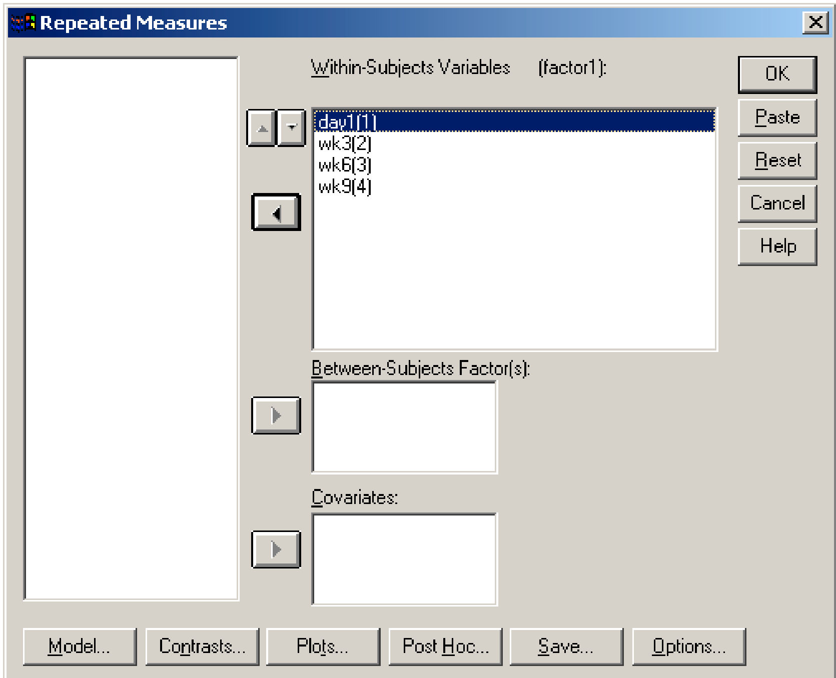
1. Start SPSS and name the variables **day1**, **wk3**, **wk6**, and **wk9**. Enter the data for each of these variables. Note that this data entry arrangement is an extension of the arrangement used for the paired-samples *t* test. Each participant's data are given on one row.

2. Select *Analyze>General Linear Model>GLM-Repeated Measures*.

3. In the dialog box, enter the number of levels (4 for the four times of measurement), and click *Add*. The Define Factor(s) dialog box should appear as follows. Then select *Define*.

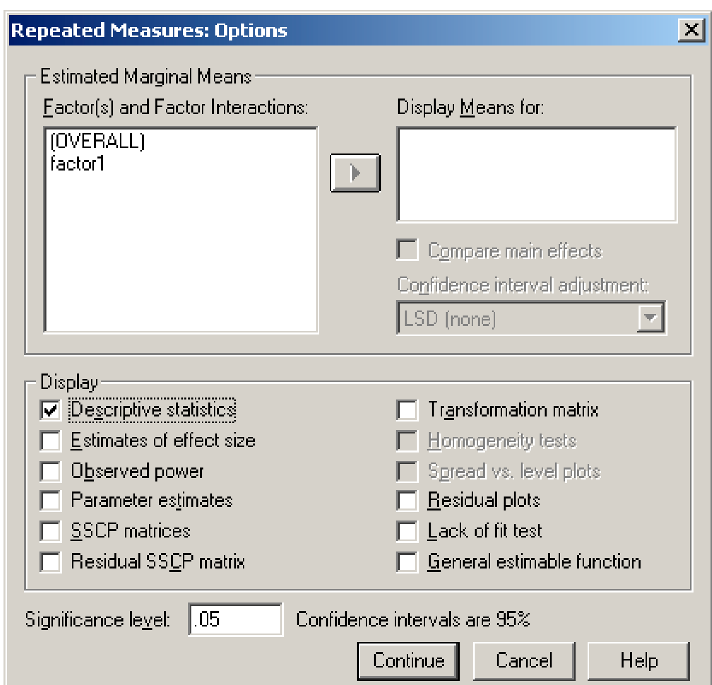


4. In the GLM-Repeated Measures dialog box, highlight each of the variables and move them into the Within-Subjects Variables box in order—that is, **day1** is first and **wk9** is fourth. The dialog box should appear as follows:



5. Although it is not required for the analysis, we will also get a plot of the means by selecting the Plots box, highlighting “factor 1” and moving it to the Horizontal Axis box, then clicking *Add>Continue*.

6. We want descriptive statistics for our groups, so select *Options*, then click on *Descriptive Statistics* in the Options dialog box, which should appear as follows. Click on *Continue*.



7. You should now be back to the GLM-Repeated Measures dialog box. Click *OK*, and the results should appear in the output Viewer window.

Notes on Reading the Output

1. As you learn to use statistical software, one skill that you will need to develop is the ability to ignore parts of the output that are superfluous for what you are trying to do. You will also need to learn to focus on the important and necessary parts of the output for your particular problem. In fact, both of these skills are necessary for you to extract the information from the output that you need for solving the present exercise.
2. In the following output, we have included only the portions that are needed for the present exercise. Your task is to ignore other parts of the output that are produced by the process we have described.
3. The Descriptive Statistics box gives exactly that information.
4. Locate the box labeled Tests of Within-Subjects Effects. Also locate the box labeled Tests of Between-Subjects Effects. The following figure shows the information needed from these two boxes to construct the source table needed for this exercise.

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
FACTOR1	Sphericity Assumed	90.000	3	30.000	41.143	.000
	Greenhouse-Geisser	90.000	2.264	39.747	41.143	.000
	Huynh-Feldt	90.000	3.000	30.000	41.143	.000
	Lower-bound	90.000	1.000	90.000	41.143	.000
Error(FACTOR1)	Sphericity Assumed	17.500	24	.729		
	Greenhouse-Geisser	17.500	18.115	.966		
	Huynh-Feldt	17.500	24.000	.729		
	Lower-bound	17.500	8.000	2.187		

Desired Source Table

Source	SS	df	MS	F
Between	90.0	3	30.0	$\frac{30.0}{0.729} = 41.14$
Subjects	186.5	8		
Error	17.5	24	0.729	
Total				

Tests of Between-Subjects Effects

Measure: MEASURE_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	2304.000	1	2304.000	98.831	.000
Error	186.500	8	23.313		

Following is the solutions output for the example based on Problem 6:

```
GLM
  day1 wk3 wk6 wk9
  /WSFACTOR = factor1 4 Polynomial
  /METHOD = SSTYPE(3)
  /PLOT = PROFILE( factor1 )
  /PRINT = DESCRIPTIVE
  /CRITERIA = ALPHA(.05)
  /WSDESIGN = factor1 .
```

General Linear Model

Within-Subjects Factors

Measure: MEASURE_1

FACTOR1	Dependent Variable
1	DAY1
2	WK3
3	WK6
4	WK9

Descriptive Statistics

	Mean	Std. Deviation	N
DAY1	10.0000	3.0822	9
WK3	9.0000	2.1794	9
WK6	7.0000	2.0000	9
WK9	6.0000	2.6926	9

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
FACTOR1	Sphericity Assumed	90.000	3	30.000	41.143	.000
	Greenhouse-Geisser	90.000	2.264	39.747	41.143	.000
	Huynh-Feldt	90.000	3.000	30.000	41.143	.000
	Lower-bound	90.000	1.000	90.000	41.143	.000
Error(FACTOR1)	Sphericity Assumed	17.500	24	.729		
	Greenhouse-Geisser	17.500	18.115	.966		
	Huynh-Feldt	17.500	24.000	.729		
	Lower-bound	17.500	8.000	2.187		

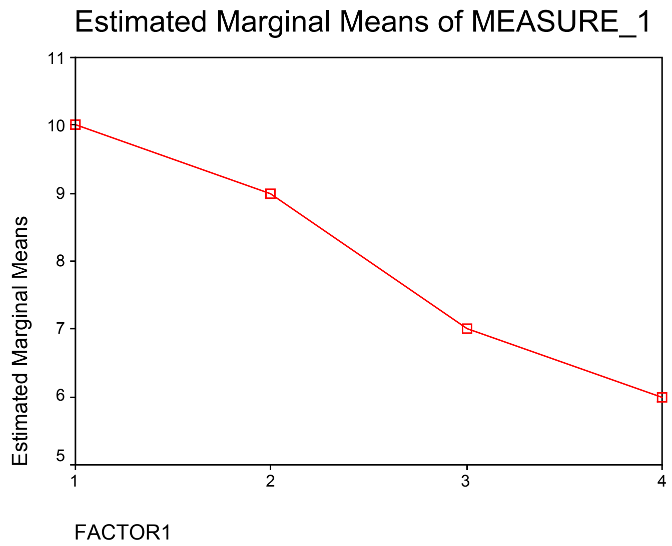
Tests of Between-Subjects Effects

Measure: MEASURE_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	2304.000	1	2304.000	98.831	.000
Error	186.500	8	23.313		

Profile Plots



Exercises Using SPSS

1. Work Self-Test Exercise 4 using SPSS. Obtain LSD post hoc test results and an Error Bar graph of the 95% confidence intervals for each group. Write a complete conclusion for the ANOVA and LSD test.
2. Work Problem 8 using SPSS with the GLM–Repeated Measures procedure. Obtain a plot of the means and use the output to construct the desired source table.

CHECKING YOUR PROGRESS: A SELF-TEST

1. If there is no treatment effect, the *F* ratio should be close to which of the following?
   a. 0
   b. 1
   c. 10
   d. ∞

2. True or False: A significant *F* ratio reveals which of the possible between-group comparisons is significant.

3. Match the following:

_______ df_{error}

_______ df_{tot}

_______ df_{subj}

_______ df_{b}

_______ df_{w}

_______ MS_{w}

_______ F (between subjects)

_______ MS_{b}

_______ MS_{error}

_______ F (repeated measures)

a. $\dfrac{MS_{\text{b}}}{MS_{\text{w}}}$

b. $N - 1$

c. $\dfrac{SS_{\text{w}}}{df_{\text{w}}}$

d. $\dfrac{SS_{\text{b}}}{df_{\text{b}}}$

e. $\dfrac{SS_{\text{b}}}{df_{\text{w}}}$

f. $K - 1$, where K is the number of groups

g. $K - N$

h. $N - K$

i. $\dfrac{MS_{\text{b}}}{MS_{\text{error}}}$

j. $(K - 1)(S - 1)$

k. $\dfrac{SS_{\text{error}}}{df_{\text{error}}}$

l. $S - 1$

4. At the end of the study described earlier in Problem 12 of this chapter, blood samples from each animal were analyzed for total cholesterol and HDL (high-density lipoprotein) cholesterol. The results are reported below in total cholesterol/HDL ratios; lower ratios are better, according to current health guidelines. Compute the F ratio and test it for significance.

Diet 1	Diet 2	Diet 3	Diet 4
2.7	1.5	2.5	2.2
2.2	1.8	2.4	2.3
2.1	1.7	2.2	1.6
2.0	2.0	1.6	2.6
1.6	1.9	1.7	2.2
2.2	1.5	2.2	2.8
2.8	1.6	2.3	2.7
2.0	1.7	2.0	
2.6	1.7	2.2	
2.6	1.8		

5. A child psychologist is interested in the course of development of object conservation in infants. The psychologist studies seven infants over a 6-month period. The infants are given 20 test trials at the ages of 9 months, 12 months, and 15 months. On each trial, an object is shown to the child and then is covered by a cloth. The child shows conservation if he or she looks for the object or becomes distressed when it is covered. The number of trials, out of 20, on which the child shows conservation is recorded. Perform the appropriate analysis; if significant, do all pairwise comparisons with the Fisher LSD test. Tell what your answers mean in the context of the problem.

Child	9 Months	12 Months	15 Months
A	0	3	17
B	2	4	17
C	3	6	16
D	1	2	14
E	0	1	19
F	4	9	2
G	4	3	20