**Lab Assignment 09**

**I. Objectives**: To test hypotheses regarding the means of two or more populations using ANOVA, and to examine the effects of incorporating controls using ANCOVA.

**II. Datasets**: “anorexia.csv”

**III. Packages**: reshape

**IV. Preparation**

1) Open RStudio by double-clicking the icon or selecting RStudio from the Windows Start menu.

2) Clear any data in memory: > rm(list=ls())

3) Download dataset “anorexia.csv” and place it in your working directory

4) Download R script “POLS 6480 Lab 09.R” and place it in your working directory.

5) Open the R script by typing *Ctrl*+*O* or by clicking on File in the upper-left corner, using the dropdown menu, and navigating to the script in your working directory.

**V. Instructions for Lab 09**

**A. Effectiveness of Treatments for Eating Disorders**

The dataset you will use is from an experiment on treating eating disorders. Subjects’ weights were measured before and after two different treatments for anorexia: family therapy and cognitive behavioral therapy. The study also included a control group, so we can examine the three conditions simultaneously using ANOVA.

1. Load the anorexia dataset, typing the following line, changing the directory if needed:

> experiment <- read.csv("C:/anorexia.csv")

Start by making separate data frames for each of the three conditions. Lines 3–5 separate the data according to the variable experiment$therapy. To examine the subjects’ weights after therapy, you can run lines 7–8 to graph horizontal box-and-whisker plots.

2. To warm up, the first test we will carry out is a comparison of family therapy to the cognitive behavioral therapy. Use the built-in R command for carrying out a difference-of-means test:

> T = t.test(treatment.f$after, treatment.b$after)

We are naming the test result T, because we are going to compare some of the output to the ANOVA output. To see what results are saved after performing a difference of means test, you can type names(T). In order to examine the test results, you can type the following:

> T$statistic

> T$p.value

These commands will show you the *t*-statistic and *p*-value, respectively. What is the difference-of-means *t* statistic? Does family therapy lead to a significantly different change in weight than cognitive behavioral therapy? (That is, do you reject the null hypothesis?)

Recall that R has three default settings: R presents the test statistic for a two-tailed test (which you can change by typing alt = "greater"), R uses a 95% confidence level when computing the confidence interval, and R uses the unequal variances test (which you can change by typing var.equal = TRUE). Suppose you changed your hypothesis to fit a directional test; does family therapy lead to a significantly greater increase in weight than cognitive behavioral therapy?

3. Now let us turn to Analysis of Variance (ANOVA). The key to ANOVA is to understand that it compares two variances. The between-groups variance is based on the differences between the group means and the overall mean. The within-groups variance is based on the differences between observations in each group and the group mean. (These two variances sum to equal the total variance, which is based on all observations’ differences from the overall mean.)

Suppose we have a few groups of observations that we are comparing. If the groups are tightly clustered, and those clusters are separated from each other by wide distances, then the between-groups component of the total variance will be proportionately large while the within-groups component of the total variance will be proportionately small. In such a situation, knowing what group an observation belongs to will be very informative. The F-test is a measure of how informative group membership is in a given situation.

In the R script, lines 10–11 calculate the mean post-treatment weights, variances of post-treatment weights, and sample sizes for the family therapy group and the cognitive behavioral therapy group. (We are ignoring the control group for the moment.) Line 12 generates the overall mean post-treatment weight from these two groups, using the weighted average of groups’ means.

Lines 13–14 calculate the between-groups sums-of-squares and the within-groups sum of squares, respectively. When calculating the between-groups sum of squares, you are creating weighted (by group size) sums of the squared differences between group means and the overall mean. The “numerator degrees of freedom” equals the number of groups, minus 1. When calculating the within-groups sum of squares, you are creating weighted (by group size, minus 1) sums of the variances of each group. The “denominator degrees of freedom” equals the number of observations minus the number of groups. (You need to remember these d.f. for the F test.)

The test you are about to conduct is called a one-way ANOVA because there is only one dimension along which the groups differ (i.e., which treatment the subjects received). The numerator for the F statistic, calculated in line 15, is the between-groups mean-squared error, which equals the between-groups sum of squares, divided by the number of groups minus one (which equals the numerator d.f). The denominator for the F statistic, calculated in line 16, is the within-groups mean squared error, which equals the within-groups sum of squares divided by the number of observations minus the number of groups (which equals the denominator d.f.).

The one-way F statistic is calculated in line 17; what is the F statistic?

We keep track of all the calculations made while performing an F test using an ANOVA Table. For a one-way ANOVA it will contain three rows (between-groups, within-groups, and total variances) and four main columns (sum of squared error, degrees of freedom, mean squared error, and the F statistic itself). Fill in the ANOVA table using the calculations above (all the key terms are defined in the **Environment** tab.)

Source SSE d.f. MSE F

Between

Within

Total

4. The 95% critical value for the F test is found using either the F distribution table or the code in line 18. When finding the critical value, it is necessary to use two degrees of freedom: the numerator d.f. equals 1 (= 2 groups minus 1) and the denominator d.f. equals 44 (= 29 + 17 minus 2 groups). What is the F critical value for 1 and 44 degrees of freedom?

When the F statistic is larger than the F critical value, this indicates that the between-groups difference is larger than we would have expected by chance alone, under the assumption that the two groups actually were drawn from a single population, at a given confidence level. Do we reject or retain the null hypothesis of no difference between groups at the 95% confidence level?

The p-value for the F statistic can be found using the code in line 19; what does *p* equal?

5. There are two different shortcut ways of calculating the F statistic; the first, shown in lines 22–23, use R’s built-in aov command. Check your F statistic, *p* value, and ANOVA table against the results shown by R. (First, run line 21 to create the correct data frame for this analysis.)

The second shortcut, shown in lines 25–26, uses R’s built-in lm command, which you will use extensively in lab 10, lab 11, and POLS 6481. This method estimates a linear regression model, and then derives the F statistic from its results. Check your F statistic, *p* value, and ANOVA table against the results shown by R.

6. Before we move on, it is important to note the relationship between the one-way ANOVA F statistic and the difference-of-means *t* statistic. You already calculated F and the F critical value above. Take the square root of F ( = \_\_\_\_\_ ) and the square root of the critical value ( = \_\_\_\_\_ ). Now use the tools from lab 8 to calculate the *t* statistic ( = \_\_\_\_\_ ) under the assumption of equal variances, and find the *t* critical value with 45 degrees of freedom ( = \_\_\_\_ ). What do you observe? If you retain the null hypothesis (of no difference between group means) using the F test, then you should also retain the null hypothesis using the *t* test.

7. Looking back at the box-and-whisker plots, perhaps it is not surprising that the F and *t* tests yield a not-quite-significant result; the two groups’ post-treatment weights do overlap somewhat. Since the experiment was designed as a panel study, we have additional data on the two groups’ pre-treatment weights. How can we incorporate that information into the analysis?

The method called Analysis of Covariance (ANCOVA) controls for the linear relationship between pre-treatment and post-treatment weight. It turns out that pre- and post-treatment weights are correlated at *r* = 50. Run line 30 to specify and estimate the model; notice that all you add to what you ran in line 25 is + before! Run line 31 to show the result of the F test.

Looking at the first line of the ANOVA Table, you should see that the between-groups sum of squared error and mean squared error are the same as before – but the F statistic is not!

Looking at the second line of the ANOVA Table, you will see that the pre-treatment weight accounts for a great deal of the variance. Note that if you add the sum of squares for the ‘before’ variable and the sum of squares for the ‘residual’, your sum will equal the sum of squares for the ‘residual’ after running line 26. Thus, looking at the third line of the ANOVA Table, you will see that much of the residual variance has been accounted for by knowing the pre-treatment weight. The within-groups mean squared error is 32% smaller than it was previously, consequently the F statistic for therapy is 32% larger.

Consider the *p* value; after controlling for pre-treatment weight, do you still retain the null hypothesis of no difference between treatments?

Earlier I focused on how the difference-of-means test yields the same information as a one-way ANOVA. Hopefully you can see how the ANCOVA is an improvement over the simple *t* test.

8. So far, we have been examining whether there is a difference between the two treatments, not whether treatment has an effect at all. Below we are going to answer this question by comparing the post-treatment weights of the two treated groups and the control group, but we can also compare pre- and post-treatment weights of the treated groups.

In order to make the before/after comparison, we will need to reshape the dataset. Run line 33 to install a package named reshape, and run line 34 to install the package. This package allows us to take data that are in columns and stack those columns vertically, or unstack columns if we so desire. Run line 35 to “melt” the data and create a new data frame named long. Run line 36 to create a new variable in the long data frame named prepost, which distinguishes between whether the observation was taken in the pre-treatment or the post-treatment phase.

With this new data frame, we can perform a simple analysis and a complex analysis.

The simple analysis is a comparison of post-treatment weight to pre-treatment weight, for both treatment groups. Run line 37 to estimate the one-way ANOVA model where the only variable is prepost, which you created in line 36; notice that we are using the long dataframe that you created in line 35. The results from the model reg.new show that treatment had a highly significant effect on average weight. If you run line 38, R will tell you the coefficient for the variable prepost, which tells you the difference between the pre-treatment and post-treatment mean weights for subjects in both treatment groups.

The complex analysis accounts for mean differences between the two treatment groups by adding two variables: therapy, which distinguishes between family and behavioral therapy type, and the interaction of therapy\*prepost. The first of these variables will control for whether subjects assigned to one treatment group had significantly different average weight than subjects assigned to the other treatment group before the experiment. The second of these variables will control for whether subjects assigned to one treatment group had significantly different average weight than subjects assigned to the other treatment group after the experiment. Run line 39 to estimate the two-way ANOVA model and view its results.

You should see that the effect of treatment still attains a statistically significant F. The pre-treatment difference between groups is significant at the *p* < .10 level only. If you are comfortable looking at regression tables, you can type summary(reg.twoway). Or, if you run line 40, R will tell you the coefficients only, indicating the differences between mean weights across treatments and periods. If you calculate the difference between mean post-treatment weights (m.f-m.b), the value equals the sum of the coefficients for therapyf and prepost:therapy !

In fact, before moving on, understanding both the data and two-way ANOVA will be facilitated by performing the following tasks:

1. find the mean weights before and after treatment for the family therapy and behavioral therapy groups;
2. hand-draw two axes, and label the vertical axis every unit between 80 and 90;
3. for the family therapy group, plot the mean weight before treatment on the left side and plot the mean weight after treatment on the right side, and draw a line connecting them;
4. for the behavioral therapy group, plot the mean weight before treatment on the left side and plot the mean weight after treatment on the right side, and draw a line connecting them.
5. Now examine the slopes of the lines and the vertical gaps between them.

The vertical gap between pre-treatment weights should be about 0.54, which equals the coefficient for therapyf. The slope of the line for the behavioral therapy group should be about 3, which equals the coefficient for prepost. The vertical gap between post-treatment weights should be about 4.8, which is equal to the sum of coefficients for therapy and prepost:therapyf. Finally, the slope of the line for the family therapy group should be more than 7.26, which is equal to the sum of coefficients for prepost and prepost:therapyf.

Drawing these lines cannot tell you whether differences between group means are statistically significant, but the process can help you understand how the group means relate to each other.

9. While it is useful to know that subjects in the treatment groups had higher weight after treatment than before treatment, we might run into the so-called Hawthorne Effect: subjects who know that they are being watched have an incentive to behave in ‘socially desirable’ ways whether they are part of a treatment or in the control group.

Run line 42 to estimate the model with post-treatment weight as a function of which treatment (or control) the subject was assigned to; note that the data frame that we are using has changed again. Examine the results of the ANOVA table. Does therapy have a significant effect on post-treatment weight?

R reports the F statistic and the *p* value, but does not report the F critical value. You will need the degrees of freedom reported by R (2 and 69, respectively) to find the correct critical value. For the 95% confidence level, you can type qf(.95, 2, 69); for the 99% confidence level, you could type qf(.99, 2, 69); does the F statistic exceed either or both of these values?

Notice that the degrees of freedom consumed by therapy equals 2, because there are three groups in this analysis. ANOVA will tell you the differences between group means; in this case it is choosing behavioral therapy as the default group because *b* comes before *c* and *f* in the alphabet. If you run line 43, then R will tell you the mean post-treatment weight for subjects in the behavioral therapy group ((Intercept) = 85.7), the difference in mean post-treatment weights between the behavioral therapy group and subjects in the control group (therapyc = –4.6), and the difference in mean post-treatment weights between the behavioral therapy group and the family therapy group (therapyf = 4.8).

10. On your own, run an ANCOVA model in which you control for pre-treatment weight. Knowing before-treatment weight accounts for some of the variance that previously was left in the residual; removing that variance from the residual should improve the F statistic for therapies.

11. To clear the **Environment**, type rm(list=ls()) or click on the broom icon.

To clear the **Console** window, type Ctrl-*l*