EXERCISE 11

One-factor between subjects ANOVA

Before you start

We suggest that you review the material in Chapter 7 before working through this practical Exercise.

The purpose of a one-factor between subjects ANOVA

In one-factor between subjects ANOVA, the ***F*** **ratio** compares the spread among the treatment means with the (supposedly uniform) spread of the scores within groups about their group means. The purpose of this Exercise is to help clarify the rationale of the *F* ratio by showing how its value is affected by various manipulations of some (or all) of the data. Before proceeding with this Exercise, we ask you to suppose that a one-factor ANOVA has been carried out upon a set of data and yields an *F* value of, say, 7.23. Now suppose we were to multiply every score in the experimental results by a constant, say 10. What would happen to the value of *F*: would it still be 7.23? Or would it increase? Or decrease?

We also invite you to speculate upon the effect that adding a constant (say 10) to all the scores in just one of the groups would have upon *F*: suppose, for example, we were to add 10 to all the scores in the group with the largest mean. Would *F* stay the same, increase or decrease in value? Would the effect be the same if the constant were added to the scores of the group with the smallest mean?

As a first approach to answering these questions, we shall carry out a **one-factor ANOVA** on a set of data. Then we shall see what happens to the value of *F* when the data are transformed as described in the previous paragraphs.

Some data

Table 1. Results of a completely randomised experiment on the effects of different mnemonic systems upon recall of logographic characters

No Mnemonic (10 control subjects)	4	6	4	3	5	7	10	4	9	11
Mnemonic 1 (10 subjects trained in Mnemonic 1)	11	9	16	10	12	17	18	16	8	11
Mnemonic 2 (10 subjects trained in Mnemonic 2)	21	16	15	16	18	11	9	12	19	20

Suppose a researcher is interested in how well non-Chinese-speaking students can learn Chinese characters using different kinds of mnemonic. Independent groups of participants are tested under three conditions: No Mnemonic, Mnemonic 1 and Mnemonic 2. The dependent variable is the number of Chinese characters that are correctly recalled. The data are shown in Table 1.

Opening SPSS

Open SPSS and select the **Type in data** radio button in the opening window. If **Data View** appears first, click the **Variable View** tab to open **Variable View**.

Construction of the SPSS data set

Rearrange the data of Table 1 into a form suitable for analysis by SPSS by following the procedure described in Sections 2.1.5 and 7.3.1. In **Variable View**, name the variables *Case*, *Group*, *Score*, remembering to change the value in the **Decimals** column to 0 each time. The variable *Group* will need appropriate values and value labels specified in the **Values** column. It is also recommended that variable labels should be entered in the **Label** column e.g. *Case Number*, *Training Condition*.

Switch to **Data View** and enter the data. The easiest way of entering the case numbers is to wait until all the other data have been entered. Then access **Compute** and enter *Case* as the **Target Variable** and *$casenum* as the **Numeric Expression**. All the case numbers will automatically appear in the *Case* column of **Data View**.

Save the data set to a file such as ***Ex11 Mnemonics***.

Exploring the data

As always, we recommend a preliminary exploration of the data set before any formal testing is carried out, in case there have been any data entry errors or contraindications for using ANOVA. As in Exercise 8, use the **Means** procedure for descriptive statistics and **Explore** for checks on the distributions of the scores within the groups. (Remember to click the **Plots** radio button to suppress the **Statistics** output. This will save you from being swamped with superfluous statistics.)

The output for **Means** begins with a **Case Processing Summary** table, followed by a table labelled **Report** listing the means, number of cases (N) and standard deviations for the three groups.

- **Examine the output for the Means procedure. Do the means appear to differ? Are the standard deviations similar in value?**

The output for **Explore** begins with a **Case Processing Summary** table, followed by the stem-and-leaf displays for the three groups. The final item shows the side-by-side boxplots.

- **Do the boxplots suggest any anomalies in the distributions of the data in any of the three groups? Write a statement assessing the suitability of the data for ANOVA.**

Procedure for the one-way ANOVA in GLM

The procedure for the one-way ANOVA is described in detail in Sections 7.5.2 & 7.5.3. Remember to click the **Post Hoc…** button, select **Tukey** and click **Continue** to return to the **Univariate** dialog box. This is because if the ANOVA F-ratio is significant, you will want to know which pairs of means differ significantly.

Click **OK** to run the ANOVA and the multiple comparisons procedure.

Output for the one-way ANOVA

Examine the ANOVA Summary Table.

- **Write down the value of *F* and its associated p-value. Is *F* significant? What are the implications of this result for the experimental hypothesis? What was the null hypothesis? What does the falsity of the null hypothesis imply?**

Look at the table of **Multiple Comparisons**.

- **Note which pairs of levels are significantly different and which are not.**

RE-ANALYSIS OF TRANSFORMED DATA SETS

In this section, we return to the question of the effects of transforming the data upon the ANOVA statistics.

1) Multiplying every score by a constant

We recommend that whenever you have occasion to transform the values of a variable in an original SPSS data set, you should construct a new target variable, rather than change (perhaps irreversibly) the original data. Use the **Compute** procedure (Section 4.4.2) to multiply each value in the data set by a factor of 10. Follow the instructions in that section, choosing, for the target variable, a mnemonic name such as *AllByTen*. Now change the **One-way ANOVA** dialog box so that the dependent variable is *AllByTen* instead of *Score* and click **OK** to run the analysis.

- **Write down the value of *F* and its associated p-value. Is *F* significant? What are the implications of this result for the experimental hypothesis?**

In the output, you will see that both the between groups and within groups variance estimates have increased by a factor of 100. It is easy to show algebraically that when each of a set of scores is multiplied by a constant, the new variance is the old variance times the square of the constant. Since, however, the factors of 100 in the numerator and denominator of the *F* ratio cancel out, the value of the *F* ratio remains unchanged.

2) Adding a constant to the scores in only one group

This time, we want a dependent variable that contains, for two of the three groups, the original scores. In the third (Mnemonic 2) group, however, every score must be increased by 10. To preserve the original scores of the third group safely as well, we shall use two computing operations:

(1) We shall copy all the scores to a new variable;

(2) We shall add a constant of 10 to the scores of the third group only.

First use **Compute** to copy the values in *Score* to a new target variable *NewScore.* Use **Compute** again to add *10* to the numbers in this new variable **only when the grouping variable has the value *3***. To do this, type *NewScore* into the **Target Variable** box and *NewScore +10* in the **Numeric Expression** box. Click **If** to open the **Compute Variable: If Cases** dialog box. Transfer the grouping variable name *Group* into the box and add the expression *=3*. Click **Continue** and **OK** to run the procedure. In **Data View**, check that the values in *NewScore* for the third group have changed but the rest have their original values. Now re-run the **one-way ANOVA**, using *NewScore* as the dependent variable.

- **Write down the value of *F* and its associated p-value. Is *F* significant? What are the implications of this result for the experimental hypothesis?**

You will see that the effect of adding a constant of 10 to all scores in the Mnemonic 2 group has no effect at all upon the within groups variance estimate. Adding the same constant to all the scores in a set has no effect upon the spread of the scores – it merely shifts the mean. The between groups mean square, however, computed from the values of the treatment means alone, has increased considerably. The within groups mean square, on the other hand, is the average of the variance estimates of the scores within groups and is quite independent of the spread among the group means. Consequently, it is quite possible to change the value of the former without affecting that of the latter and vice versa. The effect of increasing the mean of the third group is to increase the spread of the three treatment means. This increases the value of $MS_{between}$ while leaving MS_{within} unaltered. The result is an increase in *F*.

Finishing the session

Close down SPSS and any other windows before logging out of the network.