# CHAPTER 24 Appendix

## One-Way Analysis of Variance: Comparing Several Means

One-way analysis of variance (ANOVA) extends the two-sample $t$ test to more than two groups. The null hypothesis is $H_0 : \mu_1 = \mu_2 = \ \ldots \ = \mu_k$ and the alternative hypothesis is $H_a$ : at least one population mean is different the rest.

The first step is to examine the distributions of the groups with graphs and basic statistics to assess the necessary conditions/assumptions:

1. (Approximate) Normality and no outliers. Like the $t$ procedures, ANOVA is fairly robust to failure of this assumption, but you do not want data distributions that are strongly skewed or have clear outliers. Examine boxplots or stemplots of the data for each group to check this assumption (see the appendices for Chapters 1 and 2).
2. All groups have the same population variance (a condition that is generally considered fulfilled if twice the smallest standard deviation is larger than the biggest standard deviation). Check the summary statistics for each group to find the standard deviations (see the Chapter 2 Appendix).

Some software allows you to assess these conditions within the ANOVA platform. With other software, you will have to construct the plots and calculate the sample statistics separately.

**Caution:** Rejecting the null hypothesis in ANOVA does *not* say which population mean(s) are different from the rest—only that they are not all the same. Determining which mean(s) differ depends on *post-hoc* analysis, which is covered in Chapter 26.

**Excel**

ANOVA in Excel requires that the different group (treatment) observations be placed in separate columns. If the data are not formatted in this way, copy and paste so each group has its own column.

Unfortunately, Excel cannot create stemplots or side-by-side boxplots so that you can visually examine the data for symmetry and outliers. If this is your only technology, you will have to create those plots individually for each group (see the Chapter 1 Appendix).

To perform the ANOVA, use **Data ➔ Data Analysis ➔ ANOVA: Single Factor**.

1. Drag the cursor to select the input range (the columns with your data).
2. If you have included the column header (which labels the results), check the box for **Labels in first row**.
3. The default is to put the results in a new worksheet. Click **OK** to obtain the results of the test.

The Excel Video Technology Manual: *One-Way ANOVA* gives more information and an example.

The data analysis and test can all be done through **Analyze ➔ Fit Y by X**.

1. Click to select and enter the response variable (**Y, Response**) into its box and the explanatory (**X, Factor**) variable (which must be modeling type Nominal) into its box. Click **OK**.
2. You are presented with a version of a dotplot of the data for each group. This plot by itself can help you locate any potential outliers. To create side-by-side boxplots of the data, click the red triangle and select **Display Options ➔ Box Plots**. Examine these for outliers and relative symmetry.
3. Click the red triangle and select **Means and Std Dev**. If twice the smallest standard deviation is more than the largest standard deviation, you have met the "equal variability" assumption.
4. Click the red triangle and select **Means/Anova**. This gives the ANOVA table and other summary statistics.

The JMP Video Technology Manual: *One-Way ANOVA* gives more information and an example.

**Minitab**

1. **Stat ➔ ANOVA ➔ One-Way**.
2. Use the drop-down to identify your data layout (response values all in a single column with a separate column of group labels, or each group in its own column).
3. Click to select and enter the column(s) of data and the factor column (if that layout is being used).
4. Click **Graphs**. Check the appropriate box to obtain a boxplot of the data. Click **OK** twice.
5. Examine the boxplots for outliers and relative symmetry. If no problems are seen, minimize or close the graph window to see the ANOVA results.
6. Scroll down in the results window until you see the ANOVA table results. Farther down, summary statistics for each group are displayed. Minitab also provides calculated *individual* 95% confidence intervals for each population mean, based on the pooled standard deviation. If the intervals overlap, you can determine that the population means are not significantly different. If the intervals do not overlap, be careful: "Family-wise" confidence intervals for the difference in means (as discussed in Chapter 26) may still overlap.

The Minitab Video Technology Manual: *One-Way ANOVA* gives more information and an example.

SPSS requires numeric group identifiers for the data. If the grouping variable is categorical, you will must create a numeric version. Use **Transform ➔ Recode into Different Variables**.

1. Select and enter the original variable into the box.
2. Give the new variable a name in the box labeled **Name** and click **Change**.
3. Click **Old and New Values**.
4. For each value of the original variable, type it into the **Old Value** box (exactly as entered) and enter a numeric value into the **New Value** box. Click **Add**. Repeat this step until you have assigned numeric labels to all the groups. Click **Continue** and **OK**.

Perform the ANOVA using **Analyze ➔ Compare Means ➔ One-Way ANOVA**.

1. Enter the response variable in the **Dependent** list and the *numeric* group identifier in the **Factor** list.
2. Click **OK** to obtain the results (SPSS gives only the ANOVA table).

The SPSS Video Technology Manual: *One-Way ANOVA* gives more information and an example.

CRUNCHIT!

1. **Statistics ➔ ANOVA ➔ One-way**
2. Click the appropriate tab for the type of data layout. If you have one column of data and a column for grouping, click **Grouped**. If each sample is in a separate column, click **Columns**.
3. Click to enter the variables.
4. **Calculate** the results. CrunchIt! gives only the ANOVA table.

For more information and an example, see the CrunchIt! Help video *One-Way ANOVA*.

TI-83/-84

1. STAT ▶ ▶ (Tests) ANOVA( to perform the test. This command is the last one on the Tests menu, so pressing ▲ once to find it is easier than scrolling down.
2. For parameters, enter the names of all the lists that contain your data.
3. Press ENTER on **Calculate**. Note that TI calculators do not give the "Total" line for the ANOVA table; instead, its sum of squares and degrees of freedom are simply the sum of the other two entries.

The TI Video Technology Manual: *One-Way ANOVA* gives more information and an example.

For side-by-side boxplots (assuming you have one column of data and a separate column that indicates group), use a command of the following form:

```
> boxplot(mydat$Response ~ mydat$Group)
```

Examine this plot for "approximate Normality" and outliers. If satisfied, compare variability of the groups using this command:

```
> numSummary(mydat$Response, group=mydat$Group)
```

If you want to look at only the standard deviations, you can use this command:

```
> Tapply(mydat$Response, mydat$Group,sd)
```

If the conditions are satisfied, perform the ANOVA:

```
> mod <- aov(mydat$Response ~ mydat$Group)
```

It will appear that this command does not do anything. To see the results, you must enter this command:

```
> summary(mod)
```

The R Video Technology Manual: *One-Way ANOVA and Pairwise Comparisons* gives more information and an example.