CHAPTER 9

Appendix

One-Way Analysis of Variance with Excel, JMP, Minitab, SPSS, CrunchIt!, R, and TI-83/-84 Calculators

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 from the rest. You have already met the basic ANOVA table that is part of the standard regression output. Here, the terms in the table take on a different meaning.

The first step is to graphically examine the distributions of the groups for (approximate) Normality and outliers (we don't want any of these). We also use descriptive statistics to compare the standard deviations in terms of satisfying the basic assumption that all groups have the same population variance (which is generally considered fulfilled if twice the smallest standard deviation is larger than the largest standard deviation).

Data Analysis and ANOVA

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

For summary statistics on each group, use **Data ➔ Data Analysis ➔ Descriptive Statistics**.

1. Drag the cursor to select the input range (the columns with your data).
2. If you have included the column header (this labels the results), check the box for "Labels in first row."
3. Check the box for "Summary Statistics" (the default is to put the statistics in a new worksheet page). Click **OK**.
4. Compare the largest and smallest standard deviations. If twice the smallest is more than the largest, you have met the "equal variability" assumption.

Unfortunately, Excel cannot make stemplots or side-by-side boxplots to visually examine the data for symmetry and outliers. If this is your only technology, you will have to create those by hand (or try to rely on histograms).

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 (this labels the results), check the box for "Labels in first row."
3. Click **OK** for the results of the test (using the default to place them in a new window).

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) 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**. This superimposes the boxplots on the points graph (which can be unchecked if you wish). 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, 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

Use **Stat → ANOVA → One-Way**.

1. Use the drop-down to tell Minitab what your data layout is (response values all in a single column with a separate column of group labels or each group in its own column).
2. Click to select and enter the column(s) of data and the factor column (if that layout is being used).
3. Click **Graphs**. Check the box to obtain a boxplot of data. You can also ask for a plot of the individual values (like a dotplot). Click **OK** to return to the main dialog and again for the results.
4. Scroll down in the output window to examine the boxplots for outliers and relative symmetry. If no problems are seen, scroll back up to see the ANOVA results.
5. You'll see summary statistics for each group. If twice the smallest standard deviation is more than the largest, you have met the "equal variability" assumption.
6. Scroll back up further for the ANOVA table results.

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

SPSS must have numeric group identifiers. If the grouping variable is categorical, you will have to 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 until you have assigned numeric labels to all the groups. Click **Continue** and **OK**.

For summary statistics and graphs, use **Analyze → Descriptive Statistics → Explore**.

1. Click to select and enter the response variable into the "Dependent" list and either of the two group identifiers into the "Factor" list.
2. If you want side-by-side boxplots instead of individual stem-and-leaf plots, click **Plots** and uncheck the box by Stem-and-leaf. Click **Continue**.
3. Click **OK** for the default of summary statistics, stemplots, and boxplots for each group. If twice the smallest standard deviation is more than the largest, the data meet the "equal variability" condition.

Perform the ANOVA using **Analyze → Compare Means → One-Way ANOVA**. Enter the response variable in the "Dependent" list and the numeric group identifier in the "Factor" list. Click **OK** for the results.

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

CRUNCHIT!

For side-by-side boxplots (presuming you have one column of data and a separate column that indicates group):

1. Click **Graphics → Boxplots**.
2. Click the **Grouped** tab.
3. Click to enter the response (data) variable and the factor (group by) variables.
4. Click **Calculate**.

If these plots look reasonably symmetric with no outliers, the "approximately Normal" condition is satisfied.

Compare standard deviations.

1. Click **Statistics → Descriptive Statistics**.
2. Click the **Grouped** tab.
3. Click to enter the response (data) variable and the factor (group by) variables.
4. Click **Calculate**.

If twice the smallest standard deviation is more than the largest, the data meet the "equal variability" condition.

Compute the ANOVA.

1. Click **Statistics → ANOVA → One-way**.
2. Here there is an option for the type of data layout—one column of data and a column for grouping (Grouped) or each group in a separate column (Columns). Click the appropriate tab.
3. Click to enter the variables into the appropriate boxes.
4. Click **Calculate** for the results.

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

TI-83/-84

TI calculators can display up to three side-by-side boxplots at once. Enter the data for each group into separate lists. Use 2nd Y= (Stat Plots) to define the boxplots for each group using the plot that identifies outliers (if they exist). If no outliers are seen and each distribution is relatively symmetric, the "approximately Normal" condition is satisfied.

Use STAT ▸ ENTER for 1-VarStats on each list to find the standard deviations for each list of data. If twice the smallest standard deviation is more than the largest, the data meet the "equal variability" condition.

Use 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. For parameters, enter the names of all the lists that contain your data, separated by commas. Note that TI calculators do not give the "Total" line for the ANOVA table, but it is the sum of the two sums of squares, and df are simply the sum of the other two entries.

The TI-83/-84 Video Technology Manual: One-Way ANOVA gives more information and an example.

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

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

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

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

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

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

If twice the smallest standard deviation is more than the largest, the data meet the "equal variability" condition. If conditions are satisfied, perform the ANOVA:

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

The command

```
> summary(mod)
```

returns the ANOVA table.

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

Comparing Group Means

A significant ANOVA only tells us at least one population mean is different from the rest. We'd like to know which one(s) differ.

Excel cannot compute these intervals and tests.

JMP does not compute contrasts.

It has five different methods to compute simultaneous confidence intervals (the LSD—labeled t in JMP, Tukey's HSD, and three others), but none of these are the Bonferroni method described in your text. To access any of these, click the red triangle at the top of the ANOVA results, select "Compare Means," and select a method. Output will be added to the bottom of the ANOVA results.

Minitab does not compute contrasts.

It has four different methods to compute simultaneous confidence intervals (the LSD—labeled Fisher in Minitab, Tukey's HSD, and two others), but none of these are the Bonferroni method described in your text. To access any of these, click the **Comparisons** button in the main ANOVA dialog and check the box for the comparison type you desire. Output will be added to the bottom of the ANOVA results.

To create a contrast, click **Contrasts** at the upper right of the One-Way ANOVA dialog. Enter the coefficients for the contrast (specify a coefficient for each group) and click **Add** after each. To add another contrast, click **Next**. When finished adding contrasts, click **Continue** to return to the main dialog.

To compute simultaneous confidence intervals, click **Post Hoc** in the main ANOVA dialog. Check the box for the type of intervals you would like. There are several methods to choose from; the Bonferroni method described in your text is on the left-hand column. Click **Continue** to return to the main dialog.

CrunchIt! cannot compute these intervals and tests.

TI calculators cannot compute these intervals and tests.

For Bonferroni multiple comparisons, you can use the command

```
> pairwise.t.test(Response,Group,p.adj='bonferroni')
```

This gives a matrix of *P*-values for testing equality of means. Another option is

```
> TukeyHSD(mod)
```

where mod was used to store the model in the aov command. This gives pairwise comparisons of all possible combinations of group means with confidence intervals for the differences as well as *P*-values. By default, the family-wise confidence level is 95%.

Other Topics

Your text also discusses Levene's test for equal variances and the power of an ANOVA test. Not all technologies can perform these functions.

Excel can compute an *F* test for equal variances, but this test is very sensitive to departures from normality. It cannot compute Levene's test, nor can it find the power of the ANOVA test for equality of means.

To calculate the modified Levene's test:

1. **Click Analyze → Fit Y by X**.
2. Select the response variable into Y, Columns" and the explanatory (factor) variable into "X, Factor."
3. Click **OK**.
4. Use the red arrow to select **Unequal Variances**. You will have results for the original Levene's test and the modified Levene's test (labeled Brown-Forsythe).

To calculate the power of the test:

1. From the red arrow, click **Power**. If you simply want to know the power of your test, click to place a check next to "Solve for Power" and click **Done**.
2. If you want to calculate power in advance of sample collection, use **DOE → Sample Size and Power → k Sample Means**. Enter the alpha level you plan to use, an estimated pooled standard deviation, and up to 10 possible sample means. Enter the total sample size or desired power and click **Continue**. JMP returns the total sample size, assuming each sample is the same size.

To test for equal variance:

1. Click **Stat → ANOVA → Test for Equal Variances**.
2. Use the drop-down to select your data layout (one column of data with another column of factor levels or separate columns for each sample).
3. Select and enter the data column(s) and factor column into the appropriate boxes.
4. Click **OK**.
5. Minitab first gives you the standard deviations of each group and confidence intervals for the population standard deviation of these. That is followed by the results of the hypothesis test. Finally, the results are displayed graphically.

To calculate the power of the ANOVA (assuming equal sample sizes for each group):

1. Click **Stat →Power and Sample Size → One-Way ANOVA**.
2. Enter the number of factor levels; any two of sample size, value of the maximum difference in means, or power; and an estimated pooled standard deviation.
3. Click **OK**.
4. Minitab gives you a numerical value for the power for your inputs and a graphical power curve for your set of inputs. This allows you to see how manipulating one assumption will affect the power.

To calculate the test of equal variances, on the main One-Way ANOVA dialog, click **Options...**, then check the box next to Homogeneity of Variance. Click **Continue** to return to the main dialog. The Brown-Forsythe statistic is labeled "Based on medians" in the output.

SPSS cannot do "what if" power calculations, but it can calculate the power of your particular samples in the ANOVA.

1. Click. **Analyze → General Linear Model → Univariate**.
2. Click to select the response variable. Select the factor variable into the "Fixed Factor" box.
3. Click **Options**.

4. Check the box next to "Observed Power."
5. Click **Continue** to return to the main dialog, then click **OK** for the results.

CrunchIt! cannot compute power or Levene's test.

TI-83/-84

TI calculators cannot compute power or Levene's test.

Calculating the test for equal variances requires a package. This example uses the simplest package I have found, "lawstat."

```
> levene.test(Response,Group)
```

Power calculations are a bit more complex. The simplest possible command is

```
➢ pwr.anova.test(k=numlevels, n=sampsize,f=effectsize,
  sig.level=alpha,power=value)
```

where three of the four possible parameters (sample size, effect size, significance level, and power) are specified and the fourth is calculated.