CHAPTER 28 Appendix

Multiple and Logistic Regression

Multiple regression in most cases uses the same regression dialogs as simple linear regression with additional predictor variables. See Chapters 3, 4 and 23 for more details. The multiple regression setting allows you to calculate four types of models:

1. Two parallel lines. Here you will have a column of 0s and 1s denoting group membership in addition to a quantitative x variable. The coefficient of the grouping variable is the offset in the intercept for the group coded as 1.
2. Two nonparallel lines. Here you will also have a column that multiplies the x variable by the column of 0s and 1s. The coefficient of the multiplied column denotes the offset in the slope for the group coded as 1 in the grouping variable.
3. A curve, such as a quadratic or cubic model with one quantitative predictor.
4. A multidimensional surface. You will have at least two quantitative x variables.

Logistic regression has a binary response variable. The idea is to model the probability of a "success" using quantitative predictor variables.

Multiple Regression

Excel

Performing a multiple regression works just like simple regression; all the predictor variables must be in side-by-side columns. Specify the range of predictors as o block—for example, b1:h23.

1. **Analyze ➔ Fit Model**
2. Select and enter the response (y) variable and enter the predictor variables into the box labeled **Construct Model Effects**.
3. Create interaction variables by selecting two or more variables in the list and clicking **Cross**.
4. For powers of variables (transforms to allow for quadratic, cubic, or other terms), right-click on the variable name and select **Transform**. Then select the desired transformation.
5. Click **Run**.
6. If confidence intervals for parameters are desired, click the red triangle in the output next to **Response** and select **Regression Reports ➔ Show all Confidence Intervals**.
7. Confidence intervals and prediction intervals are obtained by clicking the red triangle and selecting **Save Columns ➔ Mean** (or **Indiv**) **Confidence Interval.** Two new columns will be added to the data set corresponding to the upper and lower limits of the interval.

Note: If the variable values for which you wish to make predictions do not appear in the original data set, create a dummy observation with those values and a missing value for the response variable.

For a video that shows how to use JMP here with an example, see the JMP Video Technology Manual, *Multiple Regression: Fitting and Inference*.

Minitab

1. **Stat ➔ Regression ➔ Regression ➔ Fit Regression Model**
2. Enter the list of predictors in the (**Continuous**) predictor box. For the two line models (either parallel or nonparallel), enter the grouping variable in **Categorical Predictors** box.
3. For interactions or powers of variables, click **Model** after entering the basic predictors. Selecting two variables in the dialog and clicking **Add** next to **Interactions Through Order 2** will add the basic interaction term. Clicking **Add** next to **Terms Through Order** (the default is 2) will add the square of a predictor into a model.
4. If you want to use the model to make predictions, use **Stat ➔ Regression ➔ Regression ➔ Predict** after fitting the model. Input the values of the predictors (in the order specified in the model) and click **OK**.

Residuals plots are obtained just as they were in Chapter 23 (use the **Graphs** button in the main regression dialog). Prediction and confidence intervals for responses are still created through **Stat ➔ Regression ➔ Regression ➔ Predict**.

For a video that shows how to use JMP here with an example, see the JMP Video Technology Manual, *Multiple Regression: Fitting and Inference*.

1. Use **Analyze ➔ Regression ➔ Linear** for a model that is linear in all predictors.
2. Use **Analyze ➔ Regression ➔ Curve Estimation** for a model that uses only one predictor variable, but has powers of that variable (e.g., quadratic or cubic regression).
3. To add quadratic or cubic terms to a general linear model, use **Transform ➔ Compute Variable** to create squares or cubes of individual predictors.

Residuals plots are still created using **Plots** in the dialog. Prediction and confidence intervals are created using the **Save** dialog option.

For a video that shows how to use SPSS with an example, see the JMP Video Technology Manual, *Multiple Regression: Fitting and Inference*.

CRUNCH IT!

1. **Statistics ➔ Regression ➔ Multiple Linear**
2. Use the drop-down to select the response variable, and check the boxes to select the predictor variables.
3. Use the second drop-down to select numeric results or one of two residuals plots.

If you want to add an interaction or power term to the model, use **Insert ➔ Evaluate Formula** to create the new variable before trying the regression.

CrunchIt! cannot create confidence or prediction intervals for multiple regression responses, but it can provide confidence intervals for the coefficients if you input a confidence level.

For more information (and an example), see the CrunchIt! Help Video, *Multiple Linear Regression*.

TI-83/-84

These TI calculators cannot perform multiple regression.

The basic fitting command is of the form

```
> model<-lm(y ~ x1+x1+x1:x2)
> summary(model)
```

The **x1:x2** portion of the statement is an interaction term. You can use **x1:x1** to create a quadratic term in your model.

For more information, and an example, see the R Video Technology Manual videos, *Multiple Regression: Fitting Models* and *Multiple Regression: Inference*.

Logistic Regression

Excel

Excel cannot do logistic regression.

The response variable for logistic regression must have modeling type Nominal or Ordinal. To assign the roles properly, right-click the column name of the response variable and select **Value Ordering**. The category that indicates a success should come first in this ordering. Click **Reverse** at the bottom right, if needed.

1. **Analyze ➔ Fit Y by X**
2. Select the binary variable and click **Y, Response**.
3. Select the *x* variable and click **X, Factor**.
4. **OK**

For more than one predictor, use **Analyze ➔ Fit Model**.

Minitab

1. **Stat ➔ Regression ➔ Binary Logistic Regression ➔ Fit Binary Logistic Model**
2. Click to select and enter the binary response variable into the box labeled **Response**.
3. If necessary, change the value of the response variable that corresponds to a success.
4. Click to select and enter the continuous predictor(s) into their box.
5. If desired, use **Graphs** to define residuals plots and **Model** to refine the predictors (add interaction terms, for example).
6. **OK**

1. **Analyze ➔ Regression ➔ Binary Logistic**
2. Click to select and enter the binary **Dependent** variable into its box.
3. Click to enter the predictor(s) into the **Covariates** box.
4. **OK**

CrunchIt! can do logistic regression with only one predictor, but not more than one.

1. **Statistics ➔ Regression ➔ Logistic**
2. Use the drop-downs to select the binary response and predictor variables.
3. Use the drop-down to define the binary value that designates a success.
4. **Calculate**

TI-83/-84

TI calculators cannot do logistic regression.

Use the following commands to calculate and display a logistic regression. Note that additional predictors can be added in the model statement.

```
> model <- glm(Response ~ predictor, family=binomial)
> summary(model)
```