Math 1040 StatCrunch Exercise 3

Exercises from Elementary Statistics 12th ed. By Mario F. Triola

The purpose of this exercise is to introduce you to some more features of the StatCrunch software that is available with your textbook and MyStatLab program. You will learn how to create a scatter diagram for bivariate data, find the regression equation that best fits the data, and determine whether or not a significant linear correlation exists.

Part 1: Creating a scatterplot

Using the StatCrunch software, load Data Set 14: Car Measurements. (If you cannot recall how to accomplish this or any of the previously learned tasks, please refer to StatCrunch Exercises 1 and 2.)

First create a scatter diagram by clicking on “Graph” then “Scatter Plot” on the toolbar at the top of the StatCrunch spreadsheet. For the X variable choose the Weight column and for the Y variable choose the Highway column. The weight here is given in pounds and the highway gas mileage is given in miles per gallon. Scroll down until you find the menu that allows you to enter labels and a title. Type **appropriate labels** for your axes. The graph should show what the axes represent and should include the units. Click on “Compute!” to see your results. Notice that the points do not form a perfect line, but they do seem to have an overall linear tendency. This scatter diagram would suggest that there might be a linear correlation between the weight of a car and the gas mileage of that car. **Copy your scatter diagram** into the document file that you will submit (Word is a good choice). Save this document. You will be adding to it!

Part 2: Finding the correlation coefficient and regression equation

Although looking at the plot is a great place to start, we cannot claim we have a linear relationship just by looking at the scatter diagram. We must calculate the linear correlation coefficient and compare it to the appropriate critical value in Table A-6 in the textbook. In StatCrunch click on “Stat” on the toolbar at the top of the page. Select “Regression” and then “Simple Linear”. Again, choose the weight column as the X variable and the highway miles per gallon column for the Y variable. Click on “Compute!” at the bottom of this window. Looking at the information displayed, we can see that the correlation coefficient *r* is −0.7927. The closer *r* is to −1, the stronger the negative linear relationship between our X and Y values. Note that the negative relationship is due to the fact that as the weight increases, the gas mileage decreases and the slope of the regression line is a negative number. Even though *r* is close to −1, we still must do a proper statistical analysis. Looking at Table A-6 on page 732 of the textbook, we see that for a sample size *n* = 20 (which is as close as we can get to the sample size of *n* = 21 in our data set), the critical value is 0.561 with α = 0.01. If the absolute value of the correlation coefficient is greater than the critical value, then a significant linear correlation exists. Since , we conclude that there is a significant linear relationship between the weight of a car and its highway gas mileage at the α = 0.05 level of confidence.

Also note that the StatCrunch output gives the equation of the regression line that best fits the data.

“Highway = 52.354965 – 0.0066952812 Weight”

Another way to write this is *y* = 52.354965 – 0.0066952812*x*. Remember that the coefficient of *x* in this equation is the slope of the line. Because we do have a significant linear correlation (tested by comparing *r* to the critical value) we could use this equation to make a prediction about gas mileage for a car of a given weight. For example, a car that weighs 3000 pounds would have a predicted highway gas mileage of about 32.27 miles per gallon because 52.354965 – 0.0066952812 (3000) ≈ 32.27.

Before closing the regression output window, click on “Options” in the upper left hand corner and choose “Edit”. Scroll down until you “Prediction of Y”. Click the box to make a prediction and let X = 2800. Scroll down again and you will see “Graphs”. Highlight “Plot the fitted line”. This will display the scatter diagram for the data with the equation of the regression line graphed as well. Continue scrolling down and type **appropriate labels** for your axes. Click on “Compute!”. The output shows the same information as before, but now includes the predicted gas mileage for a car weighing 2800 pounds. In the document file that you will submit, **report the linear correlation coefficient r, the linear regression equation, and the predicted gas mileage for a car weighing 2800 pounds**. Go back to the StatCrunch output and click “>” at the bottom of the output window. Now you can see your scatter diagram and line. **Copy this graph** into the document file that you will submit and save your file.

**Warning:** Most of the information in the first StatCrunch output window is beyond the scope of this class. Do not copy and paste this entire output into your file.

Part 3: Practicing your new skills

Delete the car weight data and load Data Set 2: Foot and Height Measurements.

Using the skills you have just learned, determine if there is a linear correlation between foot length (X) and height (Y). Note that both measurements are in centimeters. **Include the following** in the document file that you will submit:

1. The linear correlation coefficient and the linear regression equation.
2. The critical value for your sample size from Table A-6 with α = 0.01. State whether or not a significant linear correlation exists.
3. A copy of the Scatter Plot with fitted regression line. The graph will have **appropriate labels**, including the correct units of measurement.
4. Would the linear regression equation give a good prediction of the height of a person with a foot length of 15.3 centimeters? Explain why or why not.

Now, determine if there is a linear correlation between age in years (X) and foot length in centimeters (Y). **Include the following** in the document file that you will submit:

1. The linear correlation coefficient and the linear regression equation.
2. The critical value for your sample size from Table A-6 with α = 0.01. State whether or not a significant linear correlation exists.
3. A copy of the Scatter Plot with fitted regression line. The graph will have **appropriate labels**, including the correct units of measurement.
4. Would the linear regression equation give a good prediction of the foot length of a person who is 42 years old? Explain why or why not.

Part 4: Wrapping up

Read back through the assignment and double check that you have completed each part. Make sure your document file is neat, orderly, and easy to follow. Make sure your name is on your assignment.

Print your document and turn it in by the posted due date.

Last updated August 2014