Math 21 – Pointers for Section 4.1

In this section we learn to investigate the relationship between two numerical variables, and learn how to determine whether there is significant evidence to conclude that there is indeed a relationship.

For example, is there a relationship between shoe size and height? Probably, because a larger shoe size typically goes with a taller person.

Is there a relationship between shoe size and IQ? Probably not.

Scatter Diagram

First determine which variable affects the other variable. The variable that affects the other is the independent (x) variable or the explanatory variable. The variable that is being affected is the dependent (y) variable or the response variable.

Suppose we are investigating the relationship between study hours and score on an exam. Does "hours of study" change "test score", or does "test score" change "hours of study"? Since the number of hours that a student studies impacts the student's score, we say that "hours of study" is the independent (x) variable and "test score" is the dependent (y) variable.

Once you know which variable is x and which is y, we can put the ordered pairs on an x-y graph. This graph is called a scatter diagram (book) or scatter plot (StatCrunch).

When you create a scatter plot, look for patterns. As you move from left to right, do the points trend upwards, downwards, or neither? Are the points close to forming a line, or are they spread out?

StatCrunch

- Enter all of the ordered pairs in two columns, 1 ordered pair per line.
  Put the x-variable in the first column and the y-variable in the second column.
- Press the Graph button, and select Scatter Plot.
- Select the correct columns for the x-column and y-column.
- Click on Compute!

In general, if the points trend upward from left to right, this suggests a positive association. As x increases, y ***increases***.

In general, if the points trend downward from left to right, this suggests a negative association. As x increases, y ***decreases***.

Pearson's Correlation Coefficient *r*

The correlation coefficient *r* determines the type of relation (positive/negative) as well as the strength of the relation.

r is always between -1 & 1.

Negative *r* indicated a negative relation. Positive *r* indicates a positive relation.

The closer to 0 *r* is, the weaker the relation. The farther from 0 *r* is, the stronger the relation.

DO NOT DO THESE BY HAND. USE STATCRUNCH!

StatCrunch

- Enter all of the ordered pairs in two columns, 1 ordered pair per line.
  Put the x-variable in the first column and the y-variable in the second column.
- Press the Stat button, and select Summary Stats > Correlation.
- Select the two columns.
- Click on Compute!

We typically round *r* to 3 decimal places.

Does a Linear Relation Exist?

To determine whether there is enough evidence to conclude that a linear relation exists, we compare the value of *r* to the critical values found in Table II in Appendix A in the book.

Ignoring the signs, if *r* > critical value we conclude that a linear relation exists. Otherwise we conclude that no linear relation exists.

Three Step Process

1) Draw a scatter plot to get an idea about positive/negative and the strength of the association.

2) Calculate *r*.

3) Compare *r* to the critical value. We can only conclude that a linear relation exists if *r* > critical value.

If *r* is positive, we conclude that a positive linear relation exists.

If *r* is negative, we conclude that a negative relation exists.