UCONN

NEAG SCHOOL OF EDUCATION

Educational Research Basics by Del Siegle

Statistical Significance

Is there a relationship (difference) or isn't there a relationship (difference)?

Once you have learned the correlation coefficient (r) for your sample, you need to determine what the likelihood is that the r value you found occurred by chance. In other words, does the relationship you found in your sample really exist in the population or were your results a fluke? **— OR —** In the case of a *t* test, did the difference between the two means in your sample occurred by chance and not really exist in your population.

Prior to collecting data, researchers predetermine an alpha level, which is how willing they are to be wrong when they state that there is a relationship (in the case of correlation research) or difference (in the case of a *t* test) between the two variables they measured. A common alpha level for educational research is .05.

Let's assume that you recollected your data with 100 different samples from the same population and calculated r (or t) each time (you would not normally do this). You would assume that the value for r (or t) would not be the same for each of the 100 samples that you collected. When you set your alpha level to .05, you are saying that you are willing to be wrong (say there was a relationship in your sample when there was not one in your population 5 times out of 100). In other words, a maximum of 5 of those 100 samples might show a relationship (r <> 0) when there really was no relationship in the population (r = 0). Another way of looking at it is at least 95 times out of a 100 the relationship (difference in the case of a *t*-test) you found with your sample probably also exists in the populations from which you drew your sample (although it might be stronger or weaker). On the other hand, the relationship (difference in the case of a *t*-test) you found in the sample might occur by chance (r = 0 in the population, but you found r <> 0 in the sample) 5 times out of a 100 times. For example, if 100

times you repeatedly drew samples of 27 pairs of scores from a population where the correlation was exactly 0, by chance five of those times your sample would get a correlation of .381 or higher (even though the correlation coefficient in population from which the samples were drawn was zero.S

Use the following for Correlation...

For this example, we have set the alpha level (likelihood of being incorrect when we say the relationship we found in our sample reflects a relationship in the population) at .05. In order to determine if the r value we found with our sample meets that requirement, we will use a critical value table for Pearson's Correlation Coefficient. To use the table, you need two pieces of information, how many subjects you had and the correlation coefficient r for your study.

First you must determine something called degrees of freedom (df). For a correlation study, the degrees of freedom is equal to 2 less than the number of subjects you had. If you collected data from 27 pairs, the degrees of freedom would be 25. Use the critical value table to find the intersection of alpha .05 (see the columns) and 25 degrees of freedom (see rows). The value found at the intersection (.381) is the minimum correlation coefficient r that you would need to confidently state 95 times out of a hundred that the relationship you found with your 27 subjects exists in the population from which they were drawn.

If the absolute value of your correlation coefficient is above .381, you reject your null hypothesis (there is no relationship) and accept the alternative hypothesis: There is a statistically significant relationship between arm span and height, r (25) = .87, $p < .05$.

If the absolute value of your correlation coefficient were less than .381, you would fail to reject your null hypotheses: There is not a statistically significant relationship between arm span and height, r (25) = .12, $p > .05$.

Note: The number in parentheses following the r is the degrees of freedom and the number following the equal sign is your correlation coefficient r. $p<.05$ means your correlation coefficient exceeded the critical value found on the table and you are 95% confident that a relationship exists. $p > .05$ means that your correlation coefficient was less than the critical value on the table and you cannot be 95% confident that a relationship exists.

When using the critical value table, use the absolute value of your r (in other words...ignore the negative sign of your r if you have a negative relationship). The sign tells the DIRECTION of the relationship...not the STRENGTH. Since we are wondering if there is a strong enough relationship to be statistically significant, we

are only concerned about strength when using the table. An r = -.85 has the same strength as r = .85. You
do need to report the direction in your answer and must place the negative sign in front of the r value.

If your degree of freedom is not on the correlation table, go to the next lowest degree of freedom (df) that is.

Del Siegle, Ph.D.
Neag School of Education – University of Connecticut
del.siegle@uconn.edu

Last revised 10/14/2009

© University of Connecticut Disclaimers, Privacy & Copyright Accessibility Webmaster Login
A-Z Index