TESTING A SINGLE CORRELATION COEFFICIENT

THE HYPOTHESIS AND NATURE OF THE EFFECT

The hypothesis tested is that linear relationship exists between two variables, x and y, as seen in the correlation coefficient (r). The NULL HYPOTHESIS, however, states that no linear relationship exists between the two variables. As in all hypothesis tests, the goal is to reject the NULL HYPOTHESIS and accept the ALTERNATIVE HYPOTHESIS. In other words, to decide that an effect, in this case a relationship, exists.

Suppose a study was performed which examined the relationship between attitude toward boxing and life-satisfaction. The attitude toward boxing measured by the following statement on a questionnaire is:

"1. I enjoy watching a good boxing match."

Life-satisfaction is measured with the following statement:

"2. I am pretty much satisfied with my life."

Both items were measured with the following scale.

1=Strongly Disagree 2=Disagree 3=No Opinion 4=Agree 5=Strongly Agree

The questionnaire was given to N=33 people. The obtained correlation coefficient between these two variables was r = -.30. Because the correlation is negative we conclude that the people who said they enjoyed watching a boxing match were less satisfied with their lives. The corollary, that individuals who said they were satisfied with their lives did not say they enjoyed watching boxing, is also true. On the basis of this evidence the researcher argues that there was a relationship between the two variables.

Before he or she could decide that there was a relationship, however, a hypothesis test had to be performed to negate, or at least make improbable, the hypothesis that the results were due to chance. The ever-present devil's advocate argues that there really is no relationship between the two variables; the obtained correlation was due to chance. In this case chance caused the results. The researcher just happened to select 33 people who had a negative correlation. If another sample were taken, the correlation was just as likely to be positive and just as large. Furthermore, if a sample of infinite size (population) was taken and the correlation coefficient computed, the true correlation coefficient would be 0.0. In order to answer this argument, an hypothesis test is needed.

THE MODEL OF NO EFFECTS

THE SAMPLING DISTRIBUTION OF A CORRELATION COEFFICIENT

In a thought experiment, the study is repeated an infinite number of times using the same two questions and a different sample of 33 individuals each time, assuming the null hypothesis is true. Computing the correlation coefficient each time results in a sampling distribution of the correlation coefficient. The correlation coefficients can be graphed in a relative frequency distribution which might be similar to the following:
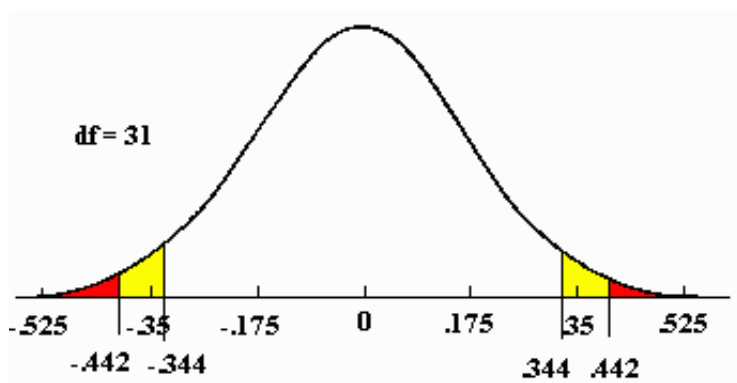
Note that this distribution looks like a normal distribution. It could not be normal, however, because the scores are limited to the range of -1.0 and 1.0.

Because the sampling distribution of the correlation coefficient has a unique shape, program is used to find values which cut off a given proportion of area. This program is included in this book. to use this program, one must first find the degrees of freedom using the following formula:

$$df = N - 2$$

After computing the degrees of freedom, click on the "Find" button:

These are the same values appearing on the sampling distribution of the correlation coefficient presented above. The values appearing in the row corresponding to the degrees of freedom are areas (probabilities) falling below the tail(s) of the distribution. In the above example, .95 area falls between correlations of -.344 and .344 and .99 area between -.442 and .442.

COMPARING THE OBTAINED STATISTIC WITH THE MODEL

The obtained correlation coefficient is now compared with what would be expected given the model created under the Null Hypothesis. In the above example, the value of -.30 falls inside the critical values of -.344 and .344 which were found in the table. Because the obtained correlation coefficient did not fall in the tails of the distribution under the Null Hypothesis, the model and the corresponding null hypothesis must be retained. The model of no effects could explain the results. The correlation coefficient is not significant at the .05 level.

If the obtained correlation coefficient had been -.55, however, the decision would have been different. In this case the obtained correlation is unlikely given the model, because it falls in the tails of the sampling distribution. The model and corresponding null hypothesis are rejected as unlikely and the alternative hypothesis, that of a real effect, accepted. In this case the obtained correlation is said to be significant at the .05 level.