TOPIC 50 THE *t* TEST

Suppose we have a *research hypothesis* that says "homicide investigators who take a short course on the causes of HIV will be less fearful of the disease than investigators who have not taken the course" and test the hypothesis by conducting an experiment in which a random sample of investigators is assigned to take the course and another random sample is designated as the control group that does not take the course.[1] Suppose that at the end of the experiment, the experimental group has a mean of 16.61 on a fear-of-HIV scale and the control group has a mean of 29.67 (where the higher the score, the greater the fear of HIV). These means support the research hypothesis. However, can we be certain that our research hypothesis is correct? If you have been reading the topics on statistics in this book in order from the beginning, you already know the answer is "no" because of the *null hypothesis*, which says that there is no *true* difference between the means. That is, it says the difference was created merely by the chance errors created by random sampling. (These errors are known simply as *sampling errors*.) Put another way, unrepresentative groups may have been assigned to the two conditions quite at random, creating the difference between the two means.

The ***t* test** is often used to test the null hypothesis regarding the observed difference between two means.[2] For the example we are considering, a series of computations, which are beyond the scope of this book, would be performed to obtain a value of t (in this case, it is 5.38) and a value of degrees of freedom (which, in this case, is $df = 179$). The values of t and df are not of any special interest to consumers of research because they are merely substeps in the mathematical procedure used to get the *probability* (p) that the null hypothesis is true. In this particular case, p is less than .05. Thus, in a research report, you may read a statement such as this:

The difference between the means is statistically significant ($t = 5.38$, $df = 179$, $p < .05$).[3]

As you know from Topic 40, the term *statistically significant* indicates that the null hypothesis has been rejected. You should recall that when the probability that the null hypothesis is true is .05 or less (such as .01 or .001), we reject the null hypothesis. (When something is unlikely to be true because it has a low probability of being true, we reject it.)

Having rejected the null hypothesis, we are in a position to assert that our research hypothesis is probably true (assuming no procedural bias was allowed to affect the results, such as testing the control group immediately after a major news story on a famous person with AIDS while testing the experimental group at an earlier time).

What can cause a t test to yield a low probability? Three factors:

1. *Sample size*. The larger the sample, the less likely that an observed difference is due to sampling errors. (You should recall from the topics on sampling that larger samples provide more precise information.) Thus, when a sample is large, we are more likely to reject the null hypothesis than when the sample is small.

2. *The size of the difference between means*. The larger the difference, the less likely the difference is due to sampling errors. Thus, when the difference between the means is large, we are more likely to reject the null hypothesis than when the difference is small.

3. *The amount of variation in the population*. You should recall from Topic 24 that when a population is very heterogeneous (has much variability), there is more potential for sampling error. Thus, when there is little variation (as indicated by the standard deviations of the sample), we are more likely to reject the null hypothesis than when there is much variation.

A special type of t test is also applied to correlation coefficients. Suppose we drew a random sample of 50 students and correlated their hand size with their GPAs and got an r of 0.19 on the Pearson r scale with possible values ranging from 1.00 to

[1] You probably recall that we prefer random sampling because (1) it precludes any bias in the assignment of participants to the groups and (2) we can test for the effect of random errors with significance tests while we cannot test for the effects of bias.

[2] To test the null hypothesis between two *medians*, the *median test* is used. It is a specialized form of the chi square test, whose results you already know how to interpret (see Topics 43 and 44).

[3] Sometimes, researchers leave out the abbreviation *df* and present the result as $t(179) = 5.38$, $p < .05$.

−1.00. The null hypothesis says that the *true* correlation in the population is 0.00. In other words, it says that we obtained 0.19 merely as the result of sampling errors. For this example, the *t* test indicates that $p > .05$. Because the probability that the null hypothesis is true is greater than 5 in 100, we do *not* reject the null hypothesis; we have a statistically insignificant correlation coefficient. (In other words, for $n = 50$, an r of 0.19 is not significantly different from an r of 0.00.) When reporting the results of the *t* test for the significance of a correlation coefficient, it is conventional *not* to mention the value of *t*. Rather, researchers usually indicate only whether or not the correlation is significant at a given probability level.

Exercise on Topic 50

1. What does the null hypothesis say about the difference between two sample means?

2. Are the values of *t* and *df* of any special interest to consumers of research?

3. Suppose you read that $t = 2.000$, $df = 20$, $p > .05$ for the difference between two means. Using conventional standards, should you conclude that the null hypothesis should be rejected?

4. Suppose you read that $t = 2.859$, $df = 40$, $p < .01$ for the difference between two means. Using conventional standards, should you conclude that the null hypothesis should be rejected?

5. Based on the information in Question 4, should you conclude that the difference between the means is statistically significant?

6. When we use a large sample, are we "more" *or* "less" likely to reject the null hypothesis than when we use a small sample?

7. When the size of the difference between means is large, are we "more" *or* "less" likely to reject the null hypothesis than when the size of the difference is small?

8. If we read that for a sample of 92 participants, $r = .41$, $p < .001$, should we reject the null hypothesis?

9. Is the value of *r* in Question 8 statistically significant?

Question for Discussion

10. Of the three factors that lead to a low probability when *t* tests are conducted, which one is most directly under the control of a researcher?

For Students Who Are Planning Research

11. Will you be conducting *t* tests? Explain.