TOPIC 40 INTRODUCTION TO THE NULL HYPOTHESIS

Suppose we drew random samples of engineers and psychologists, administered a self-report measure of sociability, and computed the mean (the most commonly used average) for each group. Furthermore, suppose the mean for engineers is 65.00 and the mean for psychologists is 70.00. Where did the five-point difference come from? There are three possible explanations:

1. Perhaps the population of psychologists is truly more sociable than the population of engineers, and our samples correctly identified the difference. (In fact, our *research hypothesis* may have been that psychologists are more sociable than engineers, which now appears to be supported by the data we collected.)

2. Perhaps there was a bias in procedures. By using random sampling, we have ruled out sampling bias, but other procedures, such as measurement, may be biased. For instance, maybe the psychologists were contacted during December, when many social events take place, and the engineers were contacted during a gloomy February. The only way to rule out bias as an explanation for the difference between the two means is to take *physical steps* to prevent it. In this case, we would want to make sure that the sociability of both groups was measured in the same way at the same time.

3. Perhaps the populations of psychologists and engineers are the same but the samples are unrepresentative of their populations because of random sampling errors. For example, the random draw may have given us a sample of psychologists who are more sociable, on the average, than their population purely by chance (at random).

The third explanation has a name: the **null hypothesis**. The general form in which it is stated varies from researcher to researcher. Here are three versions, all of which are consistent with each other:

Version A of the null hypothesis:

The observed difference was created by sampling error. (Note that the term *sampling error* refers only to *random errors*, not errors created by a bias. There is no bias in sampling in our example because we sampled at random.)

Version B of the null hypothesis:

There is no true difference between the two groups. (The term *true difference* refers to the difference we would find in a census of the two populations. That is, the *true difference* is the difference we would find if there were no sampling errors.)

Version C of the null hypothesis:

The true difference between the two groups is zero.

Significance tests determine the probability that the null hypothesis is true. (We will be considering the use of specific significance tests in Topics 43–44 and 50–52.) Suppose for our example we conduct a significance test and find that the probability that the null hypothesis is a correct hypothesis is less than 5 in 100. This would be stated as $p < .05$, where p stands for the word *probability*. Of course, if the chances that something is true are less than 5 in 100, it is likely that it is *not* true. Thus, we would *reject the null hypothesis*, leaving us with only the first two explanations we started with as viable explanations for the difference.

There is no rule of nature that dictates at what probability level the null hypothesis should be rejected. However, conventional wisdom suggests that .05 or less (such as .01 or .001) is reasonable because there are low probabilities that it is true. Of course, researchers should state in their reports the probability level they used to determine whether to reject the null hypothesis.

Note that when we fail to reject the null hypothesis because the probability is greater than .05, we do just that: We "fail to reject" the null hypothesis, and it remains on our list of possible explanations. We *cannot* "accept" the null hypothesis as the only explanation for a difference based on inferential statistics because there are three possible explanations for a difference, and failing to reject one possible explanation (the null hypothesis) does *not* mean we are accepting it as the only explanation (because there are two explanations left on the list).

An alternative way to say we have rejected the null hypothesis is to state that the difference is **statistically significant**. Thus, if we state that a difference is statistically significant at the .05 level (meaning .05 or less), it is equivalent to stating that the null hypothesis has been rejected at that level.

When you read research reported in academic journals, you will find that the null hypothesis is seldom stated by researchers, who assume that readers know the sole purpose of a significance test is to test a null hypothesis. Instead, researchers report which differences were tested for significance, which significance test they used, and which differences were found to be statistically significant. It is more common to find null hypotheses stated in theses and dissertations because committee members may want to make sure the students they are supervising understand the reason they have conducted a significance test.

We will consider specific significance tests later in this part of the book.

Exercise on Topic 40

1. How many explanations were there for the difference between psychologists and engineers in the example in this topic?

2. What does the null hypothesis say about sampling error?

3. Does the term *sampling error* refer to "random errors" *or* to "bias"?

4. The null hypothesis says the true difference equals what numerical value?

5. Significance tests are designed to determine the probabilities regarding the truth of what hypothesis?

6. The expression $p < .05$ stands for what words?

7. Do we reject the null hypothesis when the probability of its truth is "high" *or* when the probability is "low"?

8. What do we do if the probability is greater than .05?

9. What is an alternative way of saying we have rejected the null hypothesis?

10. Are you more likely to find a null hypothesis stated in a "journal article" *or* in a "dissertation"?

Question for Discussion

11. All individuals use probabilities in everyday activities to make decisions. For instance, before crossing a street with traffic, we estimate the odds that we will get across the street safely before beginning to cross it. Briefly describe one other specific use of probabilities in everyday decision-making.

For Students Who Are Planning Research

12. Will you be testing a null hypothesis in your research? Explain.