Pearson New International Edition

Reading Statistics and Research
Schuyler W. Huck
Sixth Edition

ALWAYS LEARNING™

PEARSON®

Pearson New International Edition

Reading Statistics and Research
Schuyler W. Huck
Sixth Edition

PEARSON®

four things: (1) H_0 was rejected, (2) a statistically significant finding was obtained, (3) a **reliable difference** was observed, or (4) *p* is less than a small decimal value (e.g., $p < .05$). In Excerpts 7.5 through 7.7, we see examples of how researchers sometimes communicate their decision to disbelieve H_0.

EXCERPTS 7.5–7.7 • *Rejecting the Null Hypothesis*

The authors were able to reject the null hypothesis that the program would have no effect on knowledge.

Source: Rethlefsen, M. L., Piorun, M., & Prince, D. (2009). Teaching Web 2.0 technologies using Web 2.0 technologies. *Journal of the Medical Library Association*, *97*(4), 253–259.

ESPN Internet articles included a significantly higher proportion of descriptors about the positive skill level/accomplishments and family roles/personal relationships than CBS SportsLine articles.

Source: Kian, E. T. M., Mondello, M., & Vincent, J. (2009). ESPN—The women's sports network? A content analysis of Internet coverage of March Madness. *Journal of Broadcasting & Electronic Media*, *53*(3), 477–495.

Participants generated more original analogies of time following exposure to dual cultures or a fusion culture (vs. control) ($t = 2.08$, $p < .05$).

Source: Leung, A. K., & Chiu, C. (2010). Multicultural experience, idea receptiveness, and creativity. *Journal of Cross-Cultural Psychology, 41*(5–6), 723–741.

Just as there are different ways for a researcher to tell us that H_0 is considered to be false, there are various mechanisms for expressing the other possible decision concerning the null hypothesis. Instead of saying that a fail-to-reject decision has been reached, the researcher may tell us (1) H_0 was tenable, (2) H_0 was **accepted**, (3) no reliable differences were observed, (4) no significant difference was found, (5) the result was not significant (often abbreviated as *ns* or *NS*), or (6) *p* is greater than a small decimal value (e.g., $p > .05$). Excerpts 7.8 through 7.10 illustrate these different ways of communicating a fail-to-reject decision.

EXCERPTS 7.8–7.10 • *Failing to Reject the Null Hypothesis*

Hence, this null hypothesis was accepted.

Source: Vinodh, S., Sundararaj, G., & Devadasan, S. R. (2010). Measuring organisational agility before and after implementation of TADS. *International Journal of Advanced Manufacturing Technology*, *47*(5–8), 809–818.

EXCERPTS 7.8–7.10 • (*continued*)

The male participants were evenly split, with 51% choosing the true crime book and 49% choosing the war book, $(1, N = 259) = 0.04$, *ns*.

Source: Vicary, A. M., & Fraley, R. C. (2010). Captured by true crime: Why Are women drawn to tales of rape, murder, and serial killers? *Social Psychological and Personality Science*, *1*(1), 81–86.

No significant age variance was found between Jewish and Muslim participants ($t(215) = 1.89$, $p > .05$).

Source: Winstok, Z. (2010). The effect of social and situational factors on the intended response to aggression among adolescents. *Journal of Social Psychology*, *150*(1), 57–76.

It is especially important to be able to decipher the language and notation used by researchers to indicate the decision made concerning H_0, because most researchers neither articulate their null hypotheses nor clearly state that they used the hypothesis testing procedure. Often, the only way to tell that a researcher has used this kind of inferential technique is by noting what happened to the null hypothesis.

Step 2: The Alternative Hypothesis

Near the beginning of the hypothesis testing procedure, the researcher must state an **alternative hypothesis.** Referred to as H_a (or as H_1), the alternative hypothesis takes the same form as the null hypothesis. For example, if the null hypothesis deals with the possible value of Pearson's product–moment correlation in a single population (e.g., H_0: $\rho = .00$), then the alternative hypothesis must also deal with the possible value of Pearson's correlation in a single population. Or, if the null hypothesis deals with the difference between the means of two populations (perhaps indicating that $\mu_1 = \mu_2$), then the alternative hypothesis must also say something about the difference between those populations' means. In general, therefore, H_a and H_0 are identical in that they must (1) deal with the same number of populations, (2) have the same statistical focus, and (3) involve the same variable(s).

The only difference between the null and alternative hypothesis is that the possible value of the population parameter included within H_a always differs from what is specified in H_0. If the null hypothesis is set up to say H_0: $\rho = .00$, then the alternative hypothesis might be set up to say H_a: $\rho \neq .00$, or, if a researcher specifies in Step 1 that H_0: $\mu_1 = \mu_2$, we might find that the alternative hypothesis is set up to say H_a: $\mu_1 \neq \mu_2$.

Excerpt 7.11 contains an alternative hypothesis, labeled H_1, as well as the null hypothesis with which it was paired. Notice that both H_0 and H_1 deal with the same two populations and have the same statistical focus (a percentage). The null

EXCERPT 7.11 • *The Alternative Hypothesis*

The null and alternative hypotheses are as follows:

$$H_0: p_1 - p_2 = 0$$

and

$$H_1: p_1 - p_2 \neq 0$$

where p_1 is the population proportion of administrators who select a certain outcome, and p_2 is the population proportion of school social workers who also select that outcome (for example, school social work services lead to increased attendance).

Source: Bye, L., Shepard, M., Patridge, J., & Alvarez, M. (2009). School social work outcomes: Perspectives of school social worker and school administrators. *Children & Schools*, *31*(2), 97–108.

hypothesis states that the two populations—administrators and social workers—are identical in the proportion of the population choosing a particular outcome. The alternative hypothesis states the two populations are not identical.

As indicated in the previous section, the hypothesis testing procedure terminates (in Step 6) with a decision to either reject or fail to reject the null hypothesis. In the event that H_0 is rejected, H_a represents the state of affairs that the researcher considers to be probable. In other words, H_0 and H_a always represent two opposing statements as to the possible value of the parameter in the population(s) of interest. If, in Step 6, H_0 is rejected, then belief shifts *from* H_0 *to* H_a. Stated differently, if a reject decision is made at the end of the hypothesis testing procedure, the researcher will reject H_0 *in favor of* H_a.

Although researchers have flexibility in the way they set up alternative hypotheses, they normally will set up H_a either in a **directional** fashion or in a **nondirectional** fashion.[2] To clarify the distinction between these options for the alternative hypothesis, let's imagine that a researcher conducts a study to compare men and women in terms of intelligence. Further suppose that the statistical focus of this hypothetical study is on the mean, with the null hypothesis asserting that H_0: $\mu_{men} = \mu_{women}$. Now, if the alternative hypothesis is set up in a nondirectional fashion, the researcher simply states H_a: $\mu_{men} \neq \mu_{women}$. If, however, the alternative hypothesis is stated in a directional fashion, the researcher specifies a direction in H_a. This could be done by asserting H_a: $\mu_{men} > \mu_{women}$ *or* by asserting H_a: $\mu_{men} < \mu_{women}$.

The directional/nondirectional nature of H_a is highly important within the hypothesis testing procedure. The researcher must know whether H_a was set up in

[2]A directional H_a is occasionally referred to as a *one-sided* H_a; likewise, a nondirectional H_a is sometimes referred to as a *two-sided* H_a.

a directional or nondirectional manner in order to decide whether to reject (or to fail to reject) the null hypothesis. No decision can be made about H_0 unless the directional/nondirectional character of H_a is clarified.

In most empirical studies, the alternative hypothesis is set up in a nondirectional fashion. Thus, if I were to guess what H_a says in studies containing the null hypotheses presented as shown on the left, I would bet that the researchers had set up their alternative hypotheses as indicated on the right.

Possible H_0	***Corresponding nondirectional H_a***
$H_0: \mu = 100$	$H_a: \mu \neq 100$
$H_0: \rho = +.20$	$H_a: \rho \neq +.20$
$H_0: \sigma^2 = 4$	$H_a: \sigma^2 \neq 4$
$H_0: \mu_1 - \mu_2 = 0$	$H_a: \mu_1 - \mu_2 \neq 0$

Researchers typically set up H_a in a nondirectional fashion because they do not know whether the pinpoint number in H_0 is too large or too small. By specifying a nondirectional H_a, the researcher permits the data to point one way or the other in the event that H_0 is rejected. Hence, in our hypothetical study comparing men and women in terms of intelligence, a nondirectional alternative hypothesis allows us to argue that μ_{women} is probably higher than μ_{men} (in the event that we reject the H_0 because $M_{\text{women}} > M_{\text{men}}$); or, such an alternative hypothesis allows us to argue that μ_{men} is probably higher than μ_{women} (if we reject H_0 because $M_{men} > M_{women}$).

Occasionally, a researcher believes so strongly (based on theoretical consideration or previous research) that the true state of affairs falls on one side of H_0 's pinpoint number that H_a is set up in a directional fashion. So long as the researcher makes this decision prior to looking at the data, such a decision is fully legitimate. It is, however, totally inappropriate for the researcher to look at the data first and then subsequently decide to set up H_a in a directional manner. Although a decision to reject or fail to reject H_0 could still be made after first examining the data and then articulating a directional H_a, such a sequence of events would sabotage the fundamental logic and practice of hypothesis testing. Simply stated, decisions concerning how to state H_a (and how to state H_0) must be made without peeking at any data.

When the alternative hypothesis is set up in a nondirectional fashion, researchers sometimes use the phrase **two-tailed test** to describe their specific application of the hypothesis testing procedure. In contrast, directional H_as lead to what researchers sometimes refer to as **one-tailed tests.** Inasmuch as researchers rarely specify the alternative hypothesis in their technical write-ups, the terms *one-tailed* and *two-tailed* help us to know exactly how H_a was set up. For example, consider Excerpts 7.12 and 7.13. Here, we see how researchers sometimes use the term *two-tailed* or *one-tailed* to communicate their decisions to set up H_a in a nondirectional or directional fashion.

EXCERPTS 7.12–7.13 • *Two-Tailed and One-Tailed Tests*

All tests of significance were two-tailed.

Source: Miller, K. (2010). Using a computer-based risk assessment tool to identify risk for chemotherapy-induced febrile neutropenia. *Clinical Journal of Oncology Nursing*, *14*(1), 87–91.

To investigate what variables might be important predictors of company support for fathers taking leave, [Pearson] correlations were calculated. . . . One-tailed tests of significance were used.

Source: Haas, L., & Hwang, P. C. (2010). Is fatherhood becoming more visible at work? Trends in corporate support for fathers taking parental leave in Sweden. *Fathering: A Journal of Theory, Research, & Practice about Men as Fathers*, *7*(3), 303–321.

Step 4: Collection and Analysis of Sample Data

So far, we have covered Steps 1, 2, and 6 of the hypothesis testing procedure. In the first two steps, the researcher states the null and alternative hypotheses. In Step 6, the researcher either (1) rejects H_0 in favor of H_a or (2) fails to reject H_0. We now turn our attention to the principal "stepping stone" used to move from the beginning points of the hypothesis testing procedure to the final decision.

Inasmuch as the hypothesis testing procedure is, by its very nature, an empirical strategy, it should come as no surprise that the researcher's ultimate decision to reject or to retain H_0 is based on the collection and analysis of sample data. No crystal ball is used, no Ouija board is relied on, and no eloquent argumentation is permitted. Once H_0 and H_a are fixed, only scientific evidence is allowed to affect the disposition of H_0.

The fundamental logic of the hypothesis testing procedure can now be laid bare because the connections between H_0, the data, and the final decision are as straightforward as what exists between the speed of a car, a traffic light at a busy intersection, and a lawful driver's decision as the car approaches the intersection. Just as the driver's decision to stop or to pass through the intersection is made after observing the color of the traffic light, the researcher's decision to reject or to retain H_0 is made after observing the sample data. To carry this analogy one step further, the researcher looks at the data and asks, "Is the empirical evidence inconsistent with what one would expect if H_0 were true?" If the answer to this question is yes, then the researcher has a green light and rejects H_0. However, if the data turn out to be consistent with H_0, then the data set serves as a red light telling the researcher not to discard H_0.

Because the logic of hypothesis testing is so important, let us briefly consider a hypothetical example. Suppose a valid intelligence test is given to a random sample of 100 males and a random sample of 100 females attending the same university. If the null hypothesis was first set up to say H_0: $\mu_{male} = \mu_{female}$ and if the data reveal that the two sample means (of IQ scores) differ by only two-tenths of a point, the sample data would be consistent with what we expect to happen when two samples are selected from populations having identical means. Clearly, the notion of sampling error could fully explain why the two *M*s might differ by two-tenths of an IQ point even if $\mu_{male} = \mu_{female}$. In this situation, there is no empirical justification for making the data-based claim that males at our hypothetical university have a different IQ, on average, than do their female classmates.

Now, let's consider what would happen if the difference between the two sample means turns out to be equal to 20 IQ points. If the empirical evidence turns out this way, we have a situation where the data are inconsistent with what one would expect if H_0 were true. Although the concept of sampling error strongly suggests that neither sample mean will turn out exactly equal to its population parameter, the difference of 20 IQ points between M_{males} and $M_{females}$ is quite improbable if, in fact, μ_{males} and $\mu_{females}$ are equal. With results such as this, the researcher would reject the arbitrarily selected null hypothesis.

To drive home the point I am trying to make about the way the sample data influence the researcher's decision concerning H_0, let's shift our attention to a real study that had Pearson's correlation as its statistical focus. In Excerpt 7.14, the hypothesis testing procedure was used to evaluate three bivariate correlations based on data that came from 90 men who had surgery after going to an infertility clinic. Each man was measured in terms of the number of left and right spermatic arteries as well the number of left and right lymphatic channels. Then, the left-right data were correlated for each of the two kinds of arteries and for the channels.

EXCERPT 7.14 • *Rejecting H_0 When the Sample Data Are Inconsistent with H_0*

An analysis of the relationship between the right and left spermatic cord anatomy in the bilateral varicocelectomy cases ($n = 90$) revealed a significant correlation between the number of right and left internal spermatic arteries ($r = 0.42$, $P < .05$). However, we did not identify a significant correlation between the number of right and left external spermatic arteries ($r = 0.13$, $P > .05$) or the number of right and left lymphatic channels ($r = 0.19$, $P > .05$).

Source: Libman, J. L., Segal, R., Baazeem, A., Boman, J., & Zini, A. (2010). Microanatomy of the left and right spermatic cords at subinguinal microsurgical varicocelectomy: Comparative study of primary and redo repairs. *Urology*, *75*(6), 1324–1327.

In the study associated with Excerpt 7.14, the hypothesis testing procedure was used separately to evaluate each of the three sample *r*s. In each case, the null hypothesis stated H_0: $\rho = 0.00$. The sample data, once analyzed, yielded correlations