Hypothesis Testing for the Population Proportion p: One- vs. Two-Sided Alternative

Learning Objective: Carry out hypothesis testing for the population proportion and mean (when appropriate), and draw conclusions in context.

3. One-Sided Alternative vs. Two-Sided Alternative

Recall that earlier we noticed (only visually) that for a given value of the test statistic z, the p-value of the two-sided test is twice as large as the p-value of the one-sided test. We will now further discuss this issue. In particular, we will use our example 2 (marijuana users at a certain college) to gain better intuition about this fact.

For illustration purposes, we are actually going to use example 2* (where out of a **sample of size 400**, 76 were marijuana users). Let's recall example 2*, but this time give two versions of it; the original version, and a slightly changed version, which we'll call example 2**. The differences are highlighted.

Example: 2*

There are rumors that students at a certain liberal arts college are more inclined to use drugs than U.S. college students in general. Suppose that in a simple random sample of 400 students from the college, 76 admitted to marijuana use. Do the data provide enough evidence to conclude that the proportion of marijuana users among the students in the college (p) is **higher** than the national proportion, which is .157? (This number is reported by the Harvard School of Public Health.)

Example: 2**

The dean of students in a certain liberal arts college was interested in whether the proportion of students who use drugs in her college is different than the proportion among U.S. college students in general. Suppose that in a simple random sample of 400 students from the college, 76 admitted to marijuana use. Do the data provide enough evidence to conclude that the proportion of marijuana users among the students in the college (p) **differs** from the national proportion, which is 0.157? (This number is reported by the Harvard School of Public Health.)

Learn By Doing

1/1 point (graded)

Read the two versions carefully. In particular, pay attention to the highlighted parts. What is the **main** difference between the two versions?

- ○ In example 2** we are told that the dean of students initiated the study, and in example 2* we don't know who did.
- ○ In example 2* the alternative is the one-sided H_o: p > 0.157, and the new wording of example 2** suggests that the "other" one-sided alternative H_o: p < 0.157 is appropriate.
- ○ In example 2* the alternative is two-sided, and the new wording in example 2** suggests that a one-sided alternative is appropriate.
- ● In example 2* the alternative is one-sided (H_o: p > 0.157), and the new wording in example 2** suggests that the two-sided alternative is appropriate. ✔

Answer

Correct:

Indeed, in example 2* we suspect from the outset (based on the rumors) that the overall proportion (p) of marijuana smokers at the college is higher than the reported national proportion of 0.157, and therefore the appropriate alternative is H_o: p > 0.157.

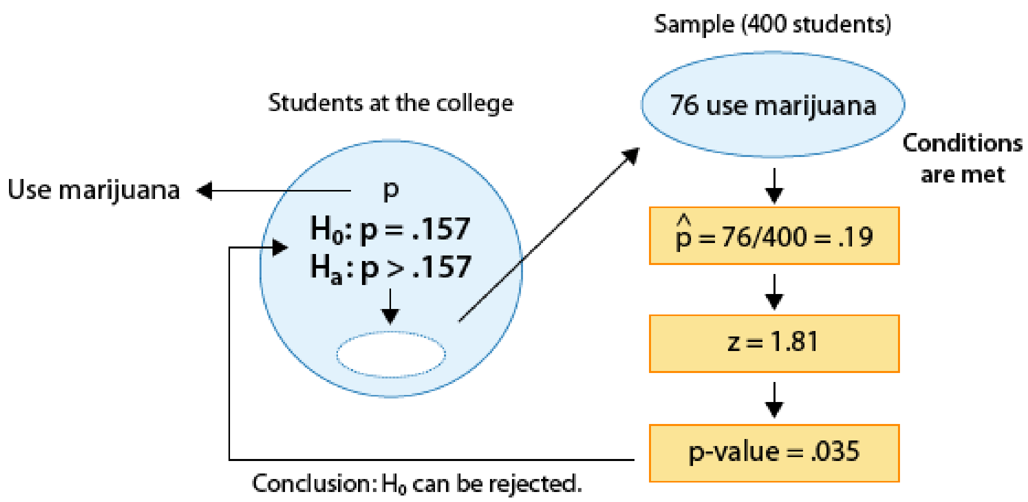
Submit

Indeed, in example 2* we suspect from the outset (based on the rumors) that the overall proportion (p) of marijuana smokers at the college is **higher** than the reported national proportion of 0.157, and therefore the appropriate alternative is H_o: p > 0.157. In example 2**, as a result of the change of wording (which eliminated the part about the rumors), we simply wonder if p is **different** (in either direction) from the reported national proportion of 0.157, and therefore the appropriate alternative is the two-sided test: $H_a : p \neq p_0$. Would switching to the two-sided alternative have an effect on our results?

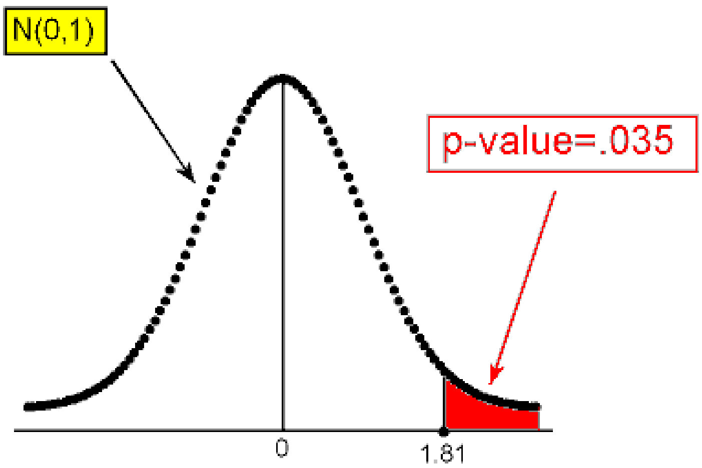
Let's explore that.

Example: 2*

We already carried out the test for this example, and the results are summarized in the following figure:



The following figure reminds you how the p-value was found (using the test statistic):



Example: 2**

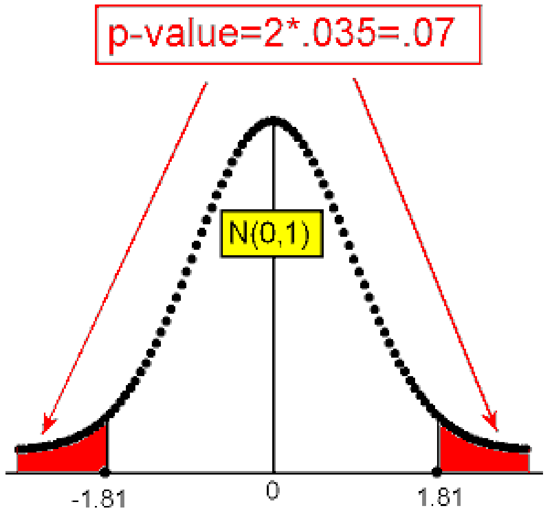
I. Here we are testing:

$$H_0 : p = 1.57$$
$$Ha : p \neq .157$$

II. Since we have the same data as in example 2* (76 marijuana users out of 400), we have the same sample proportion and the same test statistic:

$$\hat{p} = .19$$
$$z = 1.81$$

III. Since the calculation of the p-value depends on the type of alternative we have, here is where things start to be different. Statistical software tells us that the p-value for example 2** is 0.070. Here is a figure that reminds us how the p-value was calculated (based on the test statistic):



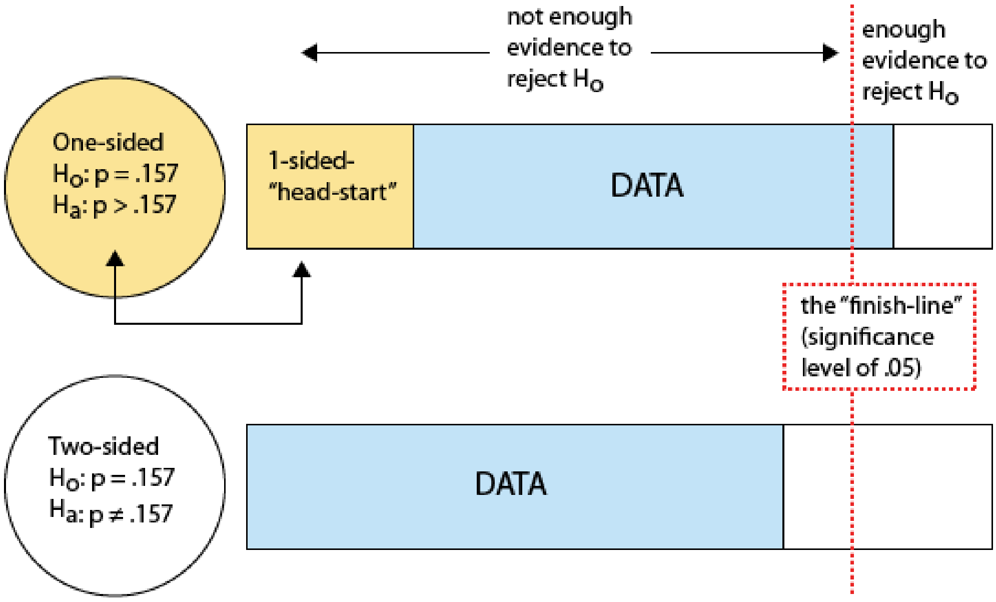
IV. If we use the 0.05 level of significance, the p-value we got is not small enough (0.07 > 0.05), and therefore we cannot reject H_o. In other words, the data do not provide enough evidence to conclude that the proportion of marijuana smokers in the college is different from the national proportion (0.157).

What happened here?

It should be pretty clear what happened here numerically. The p-value of the one-sided test (example 2*) is 0.035, suggesting the results are significant at the 0.05 significant level. However, the p-value of the two sided-test (example 2**) is twice the p-value of the one-sided test, and is therefore 2 * 0.035 = 0.07, suggesting that the results are not significant at the 0.05 significance level.

Here is a more conceptual explanation:

The idea is that in Example 2*, we began our hypothesis test with a piece of information (in the form of a rumor) about unknown population proportion p, which gave us a sort of head-start towards the goal of rejecting the null hypothesis. We found that the evidence that the data provided were then enough to cross the finish line and reject H_o. In Example 2**, we had no prior information to go on, and the data alone were not enough evidence to cross the finish line and reject H_o. The following figure illustrates this idea:



We can summarize and say that in general it is harder to reject H_o against a two-sided H_a because the p-value is twice as large. Intuitively, a one-sided alternative gives us a head-start, and on top of that we have the evidence provided by the data. When our alternative is the two-sided test, we get no head-start and all we have are the data, and therefore it is harder to cross the finish line and reject H_o.

Scenario: Online Credit Card Fraud

Consider the following two hypothesis testing scenarios for the population proportion (p) and corresponding studies:

I. The UCLA Internet Report (February 2003) estimated that roughly 8.7% of Internet users are extremely concerned about credit card fraud when buying online. A study was designed in order to examine whether that proportion has changed since.

II. The UCLA Internet Report (February 2003) estimated that roughly 8.7% of Internet users are extremely concerned about credit card fraud when buying online. In light of the increasing problem of spyware, a study was designed in order to examine whether that proportion has increased since.

Did I Get This

1/1 point (graded)

Assume that the two studies (I and II) used the same data, and that the p-value of the test corresponding to study I was found to be 0.024. The p-value of the test corresponding to study II:

- ○ must also be 0.024, since both studies used the same data.

- [ ] must be 0.048.
- [x] must be 0.012. ✔
- [ ] cannot be determined based on the information provided by the problem.

Answer

Correct:

Indeed, in study I we are testing H_o: p = 0.087 and H_a: p ≠ 0.087, whereas in study II, we are testing: H_o: p = 0.087 and H_a: p > 0.087. Since the p-value of the two-sided test is always twice the p-value of the one-sided test, the p-value of the test associated with study II must be 0.024 / 2 = 0.012.

Submit

Did I Get This

1/1 point (graded)

Consider again the two scenarios and studies about online credit card fraud (but ignore the information about the p-value of the test corresponding to study I in the question above). Which of the following (if any) is **NOT** possible (at the 0.05 significance level)?

- [ ] H_o is rejected in both tests.
- [ ] H_o cannot be rejected in both tests.
- [ ] H_o cannot be rejected in test I, but is rejected in test II.
- [x] H_o is rejected in test I, but cannot be rejected in test II. ✔

Answer

Correct:

Recall that the p-value of the two-sided test is twice the p-value of a one-sided test (based on the same data). If H_o can be rejected in test I (at the 0.05 significance level), it means that the p-value is below 0.05. This implies that since the p-value of test II must be, then, below 0.05/2 = 0.025, H_o will also be rejected in that test. There is no way, then, that H_o can be rejected in the two-sided test and not rejected in the one-sided test.

Submit