Hypothesis Testing for the Population Proportion p: Finding the p-value

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

3. Finding the P-value of the Test

So far we've talked about the p-value at the intuitive level: understanding what it is (or what it measures) and how we use it to draw conclusions about the significance of our results. We will now go more deeply into how the p-value is calculated.

It should be mentioned that eventually we will rely on technology to calculate the p-value for us (as well as the test statistic), but in order to make intelligent use of the output, it is important to first **understand** the details, and only then let the computer do the calculations for us. Let's start.

Recall that so far we have said that the p-value is the probability of obtaining data like those observed assuming that H_o is true. Like the test statistic, the p-value is, therefore, a measure of the evidence against H_o. In the case of the **test statistic,** the **larger** it is in magnitude (positive or negative) , the further $\hat{p}$ is from p_0 , the **more evidence we have against H_o.** In the case of the **p-value**, it is the opposite; the **smaller** it is, the more unlikely it is to get data like those observed when H_o is true, the **more evidence it is against H_o**. One can actually draw conclusions in hypothesis testing just using the test statistic, and as we'll see the p-value is, in a sense, just another way of looking at the test statistic. The reason that we actually take the extra step in this course and derive the p-value from the test statistic is that even though in this case (the test about the population proportion) and some other tests, the value of the test statistic has a very clear and intuitive interpretation, there are some tests where its value is not as easy to interpret. On the other hand, the p-value keeps its intuitive appeal across all statistical tests.

How is the p-value calculated?

Intuitively, the p-value is the **probability** of observing **data like those observed** assuming that $H_o$is true. Let's be a bit more formal:

- Since this is a probability question about the **data**, it makes sense that the calculation will involve the data summary, the **test statistic.**
- What do we mean by **"like"** those observed? By "like" we mean **"as extreme or even more extreme."**

Putting it all together, we get that in **general:**

The p-value is the probability of observing a test statistic as extreme as that observed (or even more extreme) assuming that the null hypothesis is true.

Comment

By **"extreme"** we mean extreme **in the direction of the alternative** hypothesis.

Specifically, for the z-test for the population proportion:

1. If the alternative hypothesis is $H_a : p < p_0$ (**less** than), then "extreme" means **small**, and the p-value is:

   The probability of observing a test statistic **as small as that observed or smaller** if the null hypothesis is true.

2. If the alternative hypothesis is $H_a : p > p_0$ (**greater** than), then "extreme" means **large**, and the p-value is:

   The probability of observing a test statistic **as large as that observed or larger** if the null hypothesis is true.

3. if the alternative is $H_a : p \neq p_0$ (**different** from), then "extreme" means extreme in either direction **either small or large (i.e., large in magnitude)**, and the p-value therefore is:

   The probability of observing a test statistic **as large in magnitude as that observed or larger** if the null hypothesis is true.

(Examples: If z = -2.5: p-value = probability of observing a test statistic as small as -2.5 or smaller or as large as 2.5 or larger.

If z = 1.5: p-value = probability of observing a test statistic as large as 1.5 or larger, or as small as -1.5 or smaller.)

OK, that makes sense. But how do we actually calculate it?

Recall the important comment from our discussion about our test statistic,

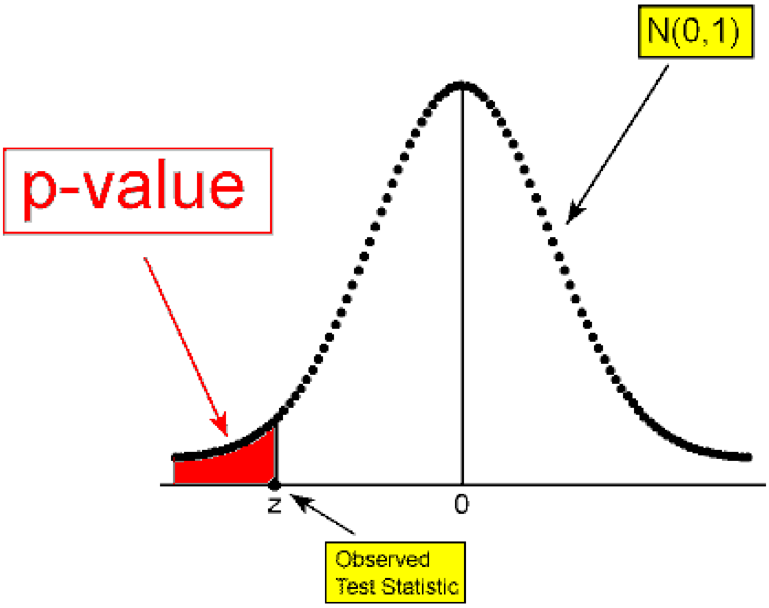
$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

which said that when the null hypothesis is true (i.e., when $p = p_0$), the possible values of our test statistic (because it is a z-score) follow a standard normal (N(0,1), denoted by Z) distribution. Therefore, the p-value calculations (which assume that H_o is true) are simply standard normal distribution calculations for the 3 possible alternative hypotheses.

Less Than

The probability of observing a test statistic as **small as that observed or smaller**, assuming that the values of the test statistic follow a standard normal distribution. We will now represent this probability in symbols and also using the normal distribution.

- $Ha: p < p_0 \Rightarrow p-value = P(Z \le z):$

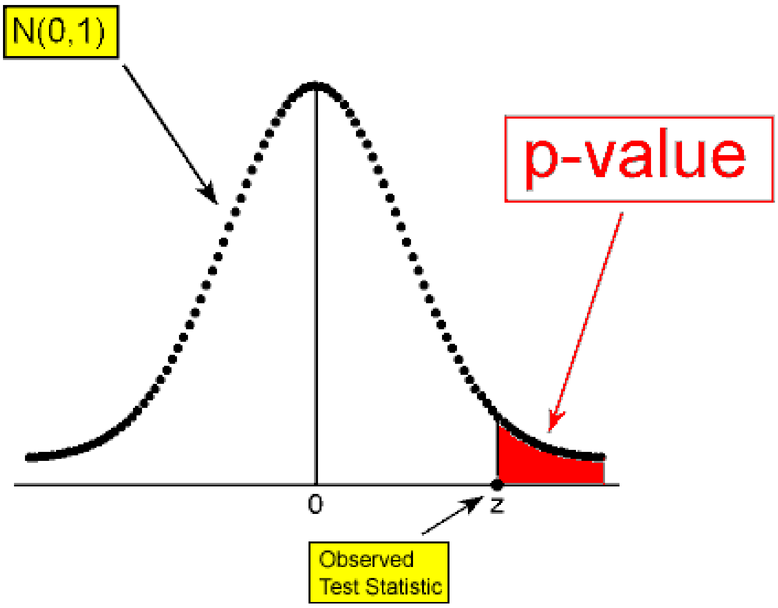


Looking at the shaded region, you can see why this is often referred to as a **left-tailed** test. We shaded to the left of the test statistic, since less than is to the left.

Greater Than

The probability of observing a test statistic as **large as that observed or larger**, assuming that the values of the test statistic follow a standard normal distribution. Again, we will represent this probability in symbols and using the normal distribution.

- $Ha: p > p_0 \Rightarrow p-value = P(Z \ge z):$

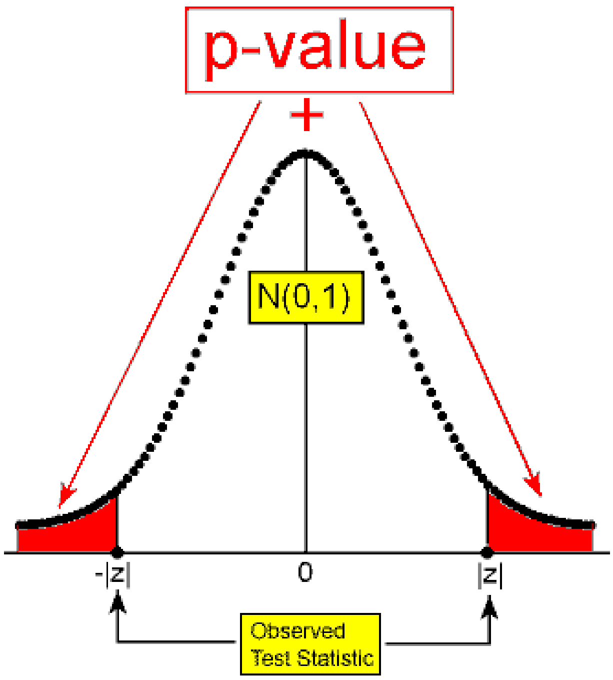


Looking at the shaded region, you can see why this is often referred to as a **right-tailed** test. We shaded to the right of the test statistic, since greater than is to the right.

Not Equal To

The probability of observing a test statistic which is as large as in **magnitude** as that observed or larger, assuming that the values of the test statistic follow a standard normal distribution.

- $Ha: p \neq p_0 \Rightarrow p-value = P(Z \leq -|z|) + P(Z \geq |z|) = 2P(Z \geq |z|)$:



This is often referred to as a **two-tailed** test, since we shaded in both directions.

As noted earlier, before the widespread use of statistical software, it was common to use 'critical values' instead of p-values to assess the evidence provided by the data. Even though the critical values approach is not used in this course, students might find it insightful. Thus, the interested students are encouraged to review the critical value method on the next page. If your instructor clearly states that you are required to have knowledge of the critical value method, you should definitely review the information.

Learn By Doing

1/1 point (graded)

Which of the following p-values will give the strongest evidence against H_o?

- p = 0.31
- p = 0.14
- p = 0.02 ✔

Answer

Correct:

A small p-value (like 0.02) indicates that the sample result is not likely to occur in random sampling from a population in which H_o is true. So a small p-value provides strong evidence against H_o.

Submit

Learn By Doing

1/1 point (graded)

If we are testing an alternative hypothesis of H_a: $p \neq p_o$, which of the following test statistics will give the smallest p-value?

- z = -0.5
- z = 1.1
- z = -2 ✔

Answer

Correct:
If $z = -2$, the data's $\hat{p}$ is 2 standard deviations below p_o. So it is very **unlikely** that $\hat{p}$s from random sampling will be located more standard deviations away from p_o than the observed data. Hence the small p-value.

Submit

Learn By Doing

1/1 point (graded)
Let's return to the scenario where we are studying the population of part-time college students. We know that in 2008, 60% of this population was female. We are curious if the proportion has decreased this year. We test the hypotheses: H_o: $p = 0.60$ and H_a: $p < 0.60$, where p is the proportion of part-time college students that are female this year.

Which of the following $\hat{p}$ values will give the smallest p-value?

- ○ $\hat{p} = 14/25 = 0.56$
- ○ $\hat{p} = 12/25 = 0.48$
- ● $\hat{p} = 10/25 = 0.40$ ✔

Answer
Correct:
For the alternative hypothesis H_a: $p < 0.60$, we are asking, "what is the probability of observing a test statistic smaller than that given by the data?" Of the options given, the probability is the smallest for $\hat{p} = 0.40$, since it is the furthest from $p = 0.60$.

Submit

Learn By Doing

1/1 point (graded)
From the three figures above, it is (at least visually) clear that for a given value of the test statistic z, the p-value of the two-sided test (equal vs. not equal) is __________ the p-value of any of the one-sided tests.

- ○ exactly half as large as

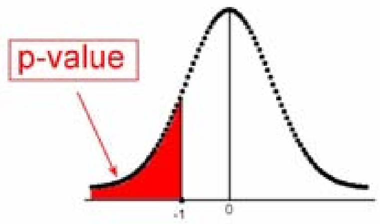
- ○ equal to
- ● exactly twice as large as ✔

Answer
Correct:
Indeed, for any given p_o, the p-value of the two-sided test: H_a: $p \neq p_o$ is twice as large as the p-value of either one of the one-sided tests.

Submit

Did I Get This

1/1 point (graded)
Which pair(s) of hypotheses and z statistic for the population proportion (p) match the figure?



- ○ H_o: $p = 0.56$, H_a: $p > 0.56$, $z = 1$
- ○ H_o: $p = 0.56$, H_a: $p < 0.56$, $z = 1$
- ○ H_o: $p = 0.56$, H_a: $p > 0.56$, $z = -1$
- ● H_o: $p = 0.56$, H_a: $p < 0.56$, $z = -1$ ✔
- ○ H_o: $p = 0.56$, H_a: $p \neq 0.56$, $z = 1$ or H_o: $p = 0.56$, H_a: $p \neq 0.56$, $z = -1$

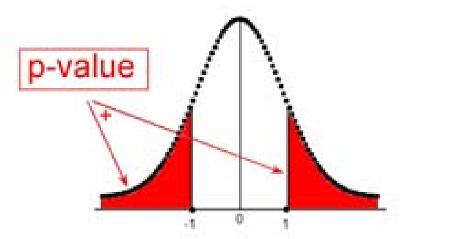
Answer
Correct:
In the figure the test statistic is $z = -1$. Also, the area to the **left** is shaded in to represent $p < 0.56$.

Submit

Did I Get This

1/1 point (graded)

Which pair(s) of hypotheses and z statistic for the population proportion (p) match the figure?



- H_o: p = 0.56, H_a: p > 0.56, z = 1
- H_o: p = 0.56, H_a: p < 0.56, z = 1
- H_o: p = 0.56, H_a: p > 0.56, z = -1
- H_o: p = 0.56, H_a: p < 0.56, z = -1
- H_o: p = 0.56, H_a: p ≠ 0.56, z = 1 or H_o: p = 0.56, H_a: p ≠ 0.56, z = -1 ✔

Answer

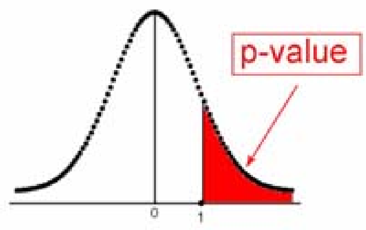
Correct: In the figure the test statistic is z = 1 and -1. Also, both tails are shaded in to represent p ≠ 0.56.

Submit

Did I Get This

1/1 point (graded)

Which pair(s) of hypotheses and z statistic for the population proportion (p) match the figure?



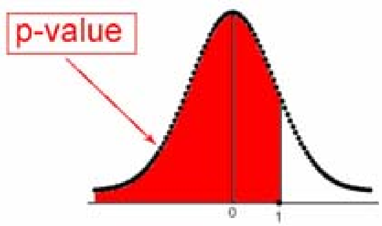
- H_o: p = 0.56, H_a: p > 0.56, z = 1 ✔
- H_o: p = 0.56, H_a: p < 0.56, z = 1
- H_o: p = 0.56, H_a: p > 0.56, z = -1
- H_o: p = 0.56, H_a: p < 0.56, z = -1
- H_o: p = 0.56, H_a: p ≠ 0.56, z = 1 or H_o: p = 0.56, H_a: p ≠ 0.56, z = -1

Answer
Correct:
In the figure the test statistic is z = 1. Also, the area to the **right** is shaded in to represent p > 0.56.

Submit

Did I Get This

1/1 point (graded)
Which pair(s) of hypotheses and z statistic for the population proportion (p) match the figure?



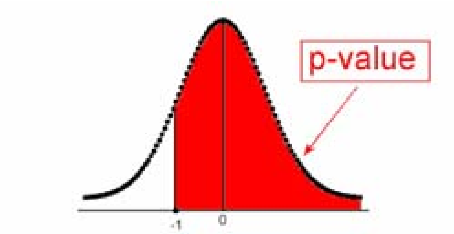
- H_o: $p = 0.56$, H_a: $p > 0.56$, $z = 1$
- H_o: $p = 0.56$, H_a: $p < 0.56$, $z = 1$ ✔
- H_o: $p = 0.56$, H_a: $p > 0.56$, $z = -1$
- H_o: $p = 0.56$, H_a: $p < 0.56$, $z = -1$
- H_o: $p = 0.56$, H_a: $p \neq 0.56$, $z = 1$ or H_o: $p = 0.56$, H_a: $p \neq 0.56$, $z = -1$

Answer
Correct:
In the figure the test statistic is $z = 1$. Also, the area to the **left** is shaded in to represent $p < 0.56$.

Submit

Did I Get This

1/1 point (graded)
Which pair(s) of hypotheses and z statistic for the population proportion (p) match the figure?



- H_o: $p = 0.56$, H_a: $p > 0.56$, $z = 1$
- H_o: $p = 0.56$, H_a: $p < 0.56$, $z = 1$
- H_o: $p = 0.56$, H_a: $p > 0.56$, $z = -1$ ✔
- H_o: $p = 0.56$, H_a: $p < 0.56$, $z = -1$
- H_o: $p = 0.56$, H_a: $p \neq 0.56$, $z = 1$ or H_o: $p = 0.56$, H_a: $p \neq 0.56$, $z = -1$

Answer
Correct:
In the figure the test statistic is z = -1. Also, the area to the **right** is shaded in to represent p > 0.56.

Submit