

6.4: Hypothesis Tests for a Single Population Proportion

In a previous lesson, we were introduced to the four step hypothesis testing process:

- Step 1. Determine hypotheses
- Step 2. Collect sample data
- Step 3. Assess the evidence
- Step 4. State a conclusion in context

We will now take a closer look at each of these steps.

Determine the Hypotheses

In order to test a claim about a population parameter, we create two opposing hypotheses. We call these the null hypothesis, H_0 , and the alternative hypothesis, H_a . Let p represent a given population proportion.

The Null Hypothesis

In every hypothesis test, we assume that the null hypothesis is true. The null hypothesis is always a statement of equality and therefore, should always contain an equal symbol ($=$). When a test involves a single population proportion, the null hypothesis will be

$$H_0 : p = \text{value}$$

Since the value is a proportion, it will be a number between 0 and 1 (inclusive).

The Alternative Hypothesis

The alternative hypothesis is a claim implied by the research question and is an inequality. The alternative hypothesis states that population proportion is greater than ($>$), less than ($<$), or not equal (\neq) to the assumed value in the null hypothesis.

When a test involves a single population proportion, alternative hypothesis will be one of the following:

$$\begin{aligned} H_a : p &> \text{value} \\ H_a : p &< \text{value} \\ H_a : p &\neq \text{value} \end{aligned}$$

Example 1

Research on college completion has shown that about 60% of students who begin college eventually graduate. A publication of higher education claims that the proportion for STEM (science, technology, engineering, math) majors is lower.

Solution: We will let p represent the proportion of all STEM majors who begin college and ultimately graduate. The null hypothesis is $H_0 : p = 0.60$. The alternative hypothesis is $H_a : p < 0.60$. The publication authors have the burden of proof and must produce evidence to support their claim that the proportion of college graduates among STEM majors is lower against the assumption that it is not.

You try!

1. The population proportion is represented by the symbol p . In the following questions, write a sentence in words for what p represents. Then determine the null and alternative hypotheses.

- a. About 67% of registered voters voted in the 2020 presidential election. A student claims that less than 67% of students at our college voted in the 2020 presidential election.

p represents:

H_0 : _____

H_a : _____

- b. In 2013, the US Department of Defense changed a policy that affected women in the armed forces. Under the new rules, women who met physical requirements could be assigned to combat positions. Many people in the US were opposed to the change. About 18% of women were against it. Researchers want to know if the proportion of US men who were against the decision was different⁹.

p represents:

H_0 : _____

H_a : _____

- c. A pro-life advocate believes that a majority (more than 50%) of unplanned pregnancies resulted from no use of contraceptive methods.

p represents:

H_0 : _____

H_a : _____

Collect Sample Data

During a hypothesis test, we work to know if a sample statistic is unusual or not. Therefore, we must think about probabilities from a sampling distribution.

In a previous lesson, we learned about the sampling distribution of sample proportions. The Central Limit Theorem says that a sampling distribution of sample proportions is approximately normal if there are at least 10 expected successes (np) and failures ($n(1 - p)$) in the sample. In the second step of a hypothesis test, we verify that the sampling distribution is approximately normal and we identify or compute any sample statistics.

Example 2

Research on college completion has shown that about 60% of students who begin college eventually graduate. A publication of higher education claims that the proportion for STEM (science, technology, engineering, math) majors is lower. Researchers randomly select 100 STEM majors and determine that 51 eventually graduate.

Solution: Recall, the null hypotheses is

$$H_0 : p = 0.60$$

The number of expected successes in the sample is $np = 100(0.60) = 60$. The number of expected failures in the sample is $n(1 - p) = n - np = 100 - 60 = 40$. Therefore, the sampling distribution of sample proportions is approximately normal.

The sample proportion is

$$\hat{p} = \frac{x}{n} = \frac{51}{100} = 0.51$$

Assess the Evidence

This step is all about probability. Since the sampling distribution is approximately normal (as determined in step 2), we can compute a Z-score and use the standard normal distribution to find probabilities. The sampling distribution of sample proportions has mean

$$\mu_{\hat{p}} = p$$

and standard error

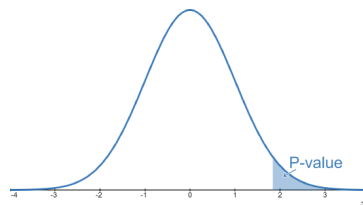
$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

where p is the assumed population proportion, and n is the sample size. The test statistic is

$$z = \frac{x - \mu}{\sigma} \text{ which translates to } z = \frac{\hat{p} - \mu_{\hat{p}}}{\sigma_{\hat{p}}} = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

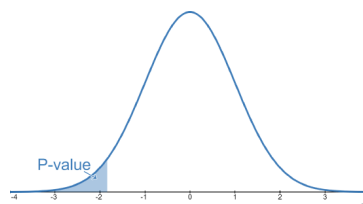
when looking at the sampling distribution of sample proportions.

- When the alternative hypothesis is $H_a : p > \text{value}$, we are conducting a right-tailed test. The P-value is the probability of observing a sample proportion at least as extreme as the one we observed. In this case at least as extreme means “as high or higher”. The P-value is the area to the right of the test statistic (T.S.).



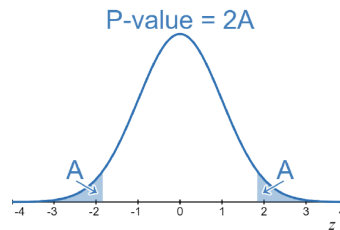
Images are created with the graphing calculator, used with permission from Desmos Studio PBC.

- When the alternative hypothesis is $H_a : p < \text{value}$, we are conducting a left-tailed test. The P-value is the probability of observing a sample proportion at least as extreme as the one we observed. In this case at least as extreme means “as low or lower”. The P-value is the area to the left of the test statistic (T.S.).



Images are created with the graphing calculator, used with permission from Desmos Studio PBC.

- When the alternative hypothesis is $H_a : p \neq \text{value}$, we are conducting a two-tailed test, and the P-value is twice the area of either the tail to the right of a positive test statistic (T.S.), or the tail to the left of a negative test statistic (T.S.).



Images are created with the graphing calculator, used with permission from Desmos Studio PBC.

Example 3

Research on college completion has shown that about 60% of students who begin college eventually graduate. A publication of higher education claims that the proportion for STEM (science, technology, engineering, math) majors is lower. Researchers randomly select 100 STEM majors and determine that 51 eventually graduate.

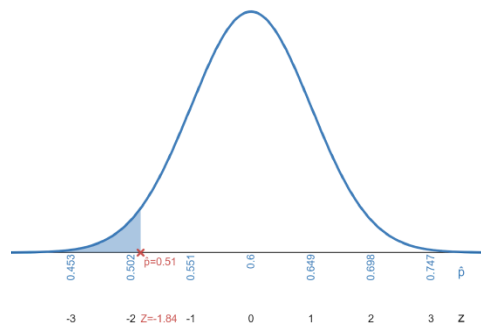
Solution: Recall, the null hypotheses is

$$H_0 : p = 0.60, \hat{p} = \frac{x}{n} = \frac{51}{100} = 0.51$$

The sampling distribution of sample proportions is approximately normal. The mean of the sampling distribution of sample proportions is $\mu_{\hat{p}} = p = 0.60$ and the standard error is

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{(0.60)(1-0.60)}{100}} \approx 0.049$$

The sampling distribution is shown below. The major tick marks have been labeled with values of \hat{p} and the corresponding Z-scores.



Images are created with the graphing calculator, used with permission from Desmos Studio PBC.

We compute the Z-score for the sample statistic,

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{0.51 - 0.60}{\sqrt{\frac{(0.60)(1-0.60)}{100}}} \approx -1.84$$

The sample statistic is -1.84 standard errors below the assumed population proportion. We perform a left-tailed test because the alternative hypothesis, $H_a : p < 0.60$, contains a less than inequality. We can now find the P-value, which is the probability of seeing a sample proportion as low or lower than 0.51, by finding the probability from the standard normal distribution. Go to <https://www.desmos.com/calculator> and type in normaldist(), click the zoom fit button, check the CDF box, and put -1.84 as the maximum. We see that $P(\hat{p} \leq 0.51) = P(z \leq -1.84) \approx 0.0329$.

State a Conclusion

Hypothesis tests are all about making decisions. We use the P-value to make a decision about the null and alternative hypotheses.

We compare our P-value to a level of significance. The level of significance, denoted (the Greek letter “alpha”), is how unlikely a sample statistic needs to be to convince us about a claim. It is also the level of risk we accept in being wrong.

We have only two possible conclusions:

- If the $P\text{-value} \leq \alpha$, we reject the null hypothesis and support the alternative hypothesis.
- If the $P\text{-value} > \alpha$, we fail to reject the null hypothesis and cannot support the alternative hypothesis.
 - This does not make the null hypothesis true—we cannot prove the null hypothesis because sample data cannot reveal the true value of the population proportion.

Example 4

Research on college completion has shown that about 60% of students who begin college eventually graduate. A publication of higher education claims that the proportion for STEM (science, technology, engineering, math) majors is lower. Researchers randomly select 100 STEM majors and determine that 51 eventually graduate. Test the claim at a 5% level of significance.

Solution: Recall, the P-value is about 0.0329. The level of significance is 5% which is 0.05 as a decimal. $0.0329 < 0.05$ so we reject the null hypothesis and support the alternative hypothesis.

The sample data support the claim that the proportion of all STEM majors who eventually graduate is less than 60%.

You try!

A pro-life advocate believes that a majority of unplanned pregnancies resulted from no use of contraceptive methods. She randomly surveyed 125 people who had unplanned pregnancies and found that 64 did not use a contraceptive method in the month they became pregnant. Test the claim at a 5% level of significance.

Step 1. Determine the hypotheses

a. p represents:

b. H_0 : _____

c. H_a : _____

d. Right-, left-, or two-tailed test? Explain how you know.

Step 2. Collect sample data

a. Explain why the sampling distribution of sample proportions is approximately normal.

b. Calculate the sample proportion, \hat{p}

Step 3 Assess the evidence

a. Compute the Z-score that corresponds to the sample proportion.

