How is hypothesis testing used to determine whether a proportion differs significantly from a given value, and what is the process for conducting a two-tailed test?

Authored by

stats writer

June 29, 2024

RECOMMENDED CITATION

stats writer (2024). *How is hypothesis testing used to determine whether a proportion differs significantly from a given value, and what is the process for conducting a two-tailed test?*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=159306

Hypothesis testing is a statistical method used to determine whether a proportion differs significantly from a given value. This method is commonly used in research and data analysis to draw conclusions about a population based on a sample.

The process of conducting a two-tailed test involves formulating a null hypothesis, which states that there is no significant difference between the proportion in the sample and the given value. A alternative hypothesis is also formulated, which states that there is a significant difference between the proportion in the sample and the given value.

A test statistic is then calculated using the sample data, which measures the difference between the observed proportion and the given value. This test statistic is then compared to a critical value, which is determined based on the level of significance chosen by the researcher. If the test statistic falls within the critical value range, the null hypothesis cannot be rejected and it is concluded that there is no significant difference between the proportion and the given value. However, if the test statistic falls outside the critical value range, the null hypothesis is rejected and it is concluded that there is a significant difference between the proportion and the given value.

A two-tailed test is used when the researcher is interested in determining whether the proportion is significantly different from the given value in both directions, either greater or less than. This type of test is appropriate when there is no specific direction or expectation for the difference. Overall, hypothesis testing is a valuable tool for determining the significance of a proportion in relation to a given value and can provide valuable insights for decision making in various fields.

Statistics - Hypothesis Testing a Proportion (Two Tailed)

A population proportion is the share of a population that belongs to a particular category.

Hypothesis tests are used to check a claim about the size of that population proportion.

Hypothesis Testing a Proportion

The following steps are used for a hypothesis test:

Check the conditionsDefine the claimsDecide the significance levelCalculate the test statisticConclusion

For example:

Population: Nobel Prize winners**Category**: Women

And we want to check the claim:

"The share of Nobel Prize winners that are women is **not** 50%"

By taking a sample of 100 randomly selected Nobel Prize winners we could find that:

10 out of 100 Nobel Prize winners in the sample were women

The **sample** proportion is then: (displaystyle frac{10}{100} = 0.1), or 10%.

From this sample data we check the claim with the steps below.

1. Checking the Conditions

The conditions for calculating a confidence interval for a proportion are:

The sample is randomly selectedThere is only two options:
Being in the categoryNot being in the categoryThe sample needs at least:
5 members in the category5 members not in the category

In our example, we randomly selected 10 people that were women.

The rest were not women, so there are 90 in the other category.

The conditions are fulfilled in this case.

Note: It is possible to do a hypothesis test without having 5 of each category. But special adjustments need to be made.

2. Defining the Claims

We need to define a **null hypothesis** ((H_{0})) and an **alternative hypothesis** ((H_{1})) based on the claim we are checking.

The claim was:

"The share of Nobel Prize winners that are women is **not** 50%"

In this case, the **parameter** is the proportion of Nobel Prize winners that are women ((p)).

The null and alternative hypothesis are then:

Null hypothesis: 50% of Nobel Prize winners were women.

Alternative hypothesis: The share of Nobel Prize winners that are women is **not** 50%

Which can be expressed with symbols as:

(H_{0}): (p = 0.50)

(H_{1}): (p neq 0.50)

This is a '**two-tailed**' test, because the alternative hypothesis claims that the proportion is **different** (larger or smaller) than in the null hypothesis.

If the data supports the alternative hypothesis, we **reject** the null hypothesis and **accept** the alternative hypothesis.

3. Deciding the Significance Level

The significance level ((alpha)) is the **uncertainty** we accept when rejecting the null hypothesis in a hypothesis test.

The significance level is a percentage probability of accidentally making the wrong conclusion.

Typical significance levels are:

(alpha = 0.1) (10%)(alpha = 0.05) (5%)(alpha = 0.01) (1%)

A lower significance level means that the evidence in the data needs to be stronger to reject the null hypothesis.

There is no "correct" significance level - it only states the uncertainty of the conclusion.

Note: A 5% significance level means that when we reject a null hypothesis:

We expect to reject a **true** null hypothesis 5 out of 100 times.

4. Calculating the Test Statistic

The test statistic is used to decide the outcome of the hypothesis test.

The test statistic is a standardized value calculated from the sample.

The formula for the test statistic (TS) of a population proportion is:

(displaystyle frac{hat{p} - p}{sqrt{p(1-p)}} cdot sqrt{n})

(hat{p}-p) is the **difference** between the **sample** proportion ((hat{p})) and the claimed **population** proportion ((p)).

(n) is the sample size.

In our example:

The claimed ((H_{0})) population proportion ((p)) was (0.50)

The sample proportion ((hat{p})) was 10 out of 100, or: (displaystyle frac{10}{100} = 0.10)

The sample size ((n)) was (100)

So the test statistic (TS) is then:

(displaystyle frac{0.1-0.5}{sqrt{0.5(1-0.5)}} cdot sqrt{100} = frac{-0.4}{sqrt{0.5(0.5)}} cdot sqrt{100} = frac{-0.4}{sqrt{0.25}} cdot sqrt{100} = frac{-0.4}{0.5} cdot 10 = underline{-8})

You can also calculate the test statistic using programming language functions:

Example

With Python use the scipy and math libraries to calculate the test statistic for a proportion.

```
import scipy.stats as stats

import math

# Specify the number of occurrences (x), the sample size (n), and the proportion claimed in the null-hypothesis (p)

x = 10

n = 100

p = 0.5

# Calculate the sample proportion

p_hat = x/n

# Calculate and print the test statistic

print((p_hat-p)/(math.sqrt((p*(1-p))/(n))))
```

Example

With R use the built-in math functions to calculate the test statistic for a proportion.

```
# Specify the sample occurrences (x), the sample size (n), and the null-hypothesis claim (p)

x <- 10

n <- 100

p <- 0.5

# Calculate the sample proportion

p_hat = x/n

# Calculate and output the test statistic

(p_hat-p)/(sqrt((p*(1-p))/(n)))
```

5. Concluding

There are two main approaches for making the conclusion of a hypothesis test:

The **critical value** approach compares the test statistic with the critical value of the significance level.The **P-value** approach compares the P-value of the test statistic and with the significance level.

Note: The two approaches are only different in how they present the conclusion.

The Critical Value Approach

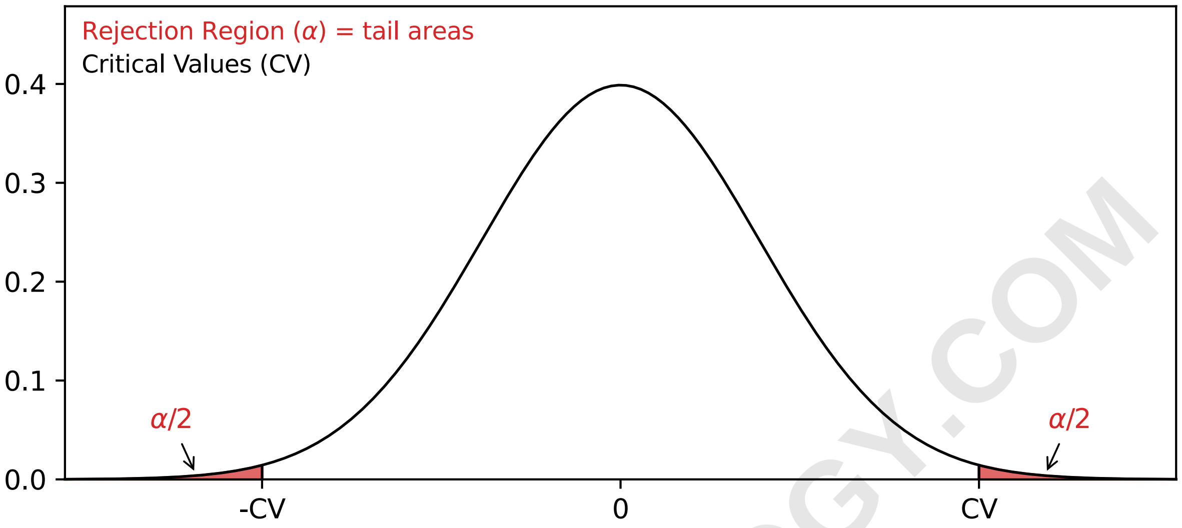
For the critical value approach we need to find the **critical value** (CV) of the significance level ((alpha)).

For a population proportion test, the critical value (CV) is a **Z-value** from a standard normal distribution.

This critical Z-value (CV) defines the **rejection region** for the test.

The rejection region is an area of probability in the tails of the standard normal distribution.

Because the claim is that the population proportion is **different** from 50%, the rejection region is split into both the left and right tail:



The size of the rejection region is decided by the significance level ((alpha)).

Choosing a significance level ((alpha)) of 0.01, or 1%, we can find the critical Z-value from a Z-table, or with a programming language function:

Note: Because this is a two-tailed test the tail area ((alpha)) needs to be split in half (divided by 2).

Example

With Python use the Scipy Stats library `norm.ppf()` function find the Z-value for an (alpha)/2 = 0.005 in the left tail.

import scipy.stats as stats

print(stats.norm.ppf(0.005))

Example

With R use the built-in `qnorm()` function to find the Z-value for an (alpha) = 0.005 in the left tail.

qnorm(0.005)

Using either method we can find that the critical Z-value in the left tail is (approx underline{-2.5758})

Since a normal distribution i symmetric, we know that the critical Z-value in the right tail will be the same number, only positive: (underline{2.5758})

For a **two-tailed** test we need to check if the test statistic (TS) is **smaller** than the negative critical value (-CV), **or bigger** than the positive critical value (CV).

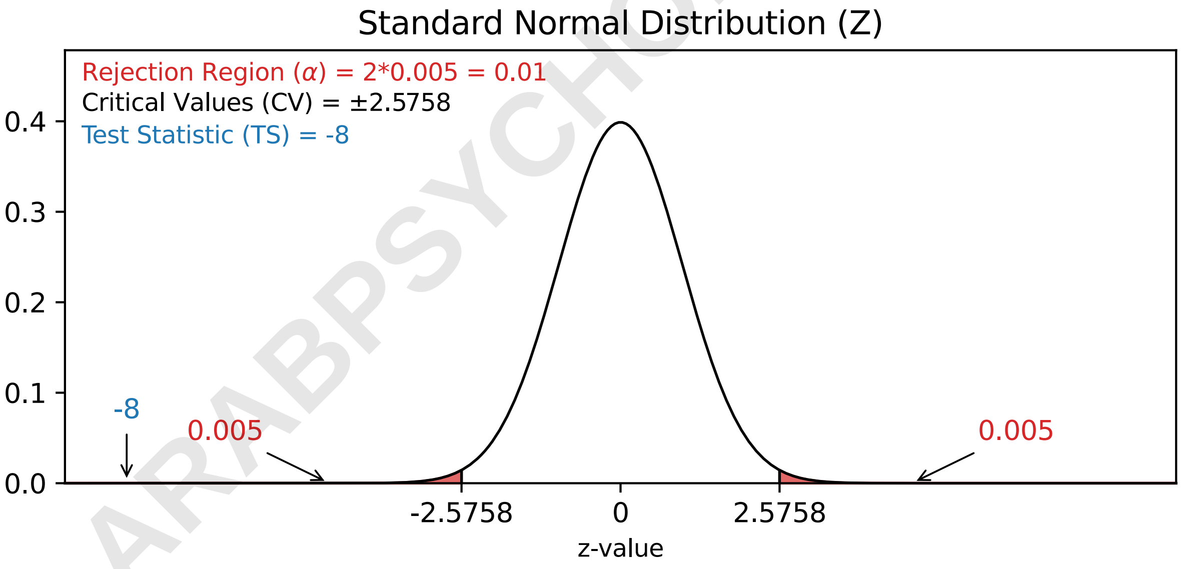
If the test statistic is smaller than the **negative** critical value, the test statistic is in the **rejection region**.

If the test statistic is bigger than the **positive** critical value, the test statistic is in the **rejection region**.

When the test statistic is in the rejection region, we **reject** the null hypothesis ((H_{0})).

Here, the test statistic (TS) was (approx underline{-8}) and the critical value was (approx underline{-2.5758})

Here is an illustration of this test in a graph:



Since the test statistic was **smaller** than the negative critical value we **reject** the null hypothesis.

This means that the sample data supports the alternative hypothesis.

And we can summarize the conclusion stating:

The sample data **supports** the claim that "The share of Nobel Prize winners that are women is **not** 50%" at a **1% significance level**.

The P-Value Approach

For the P-value approach we need to find the **P-value** of the test statistic (TS).

If the P-value is **smaller** than the significance level ((alpha)), we **reject** the null hypothesis ((H_{0})).

The test statistic was found to be (approx underline{-8})

For a population proportion test, the test statistic is a Z-Value from a standard normal distribution.

Because this is a **two-tailed** test, we need to find the P-value of a Z-value **smaller** than -8 and **multiply it by 2**.

We can find the P-value using a Z-table, or with a programming language function:

Example

With Python use the Scipy Stats library `norm.cdf()` function find the P-value of a Z-value smaller than -8 for a two tailed test:

```
import scipy.stats as stats

print(2*stats.norm.cdf(-8))
```

Example

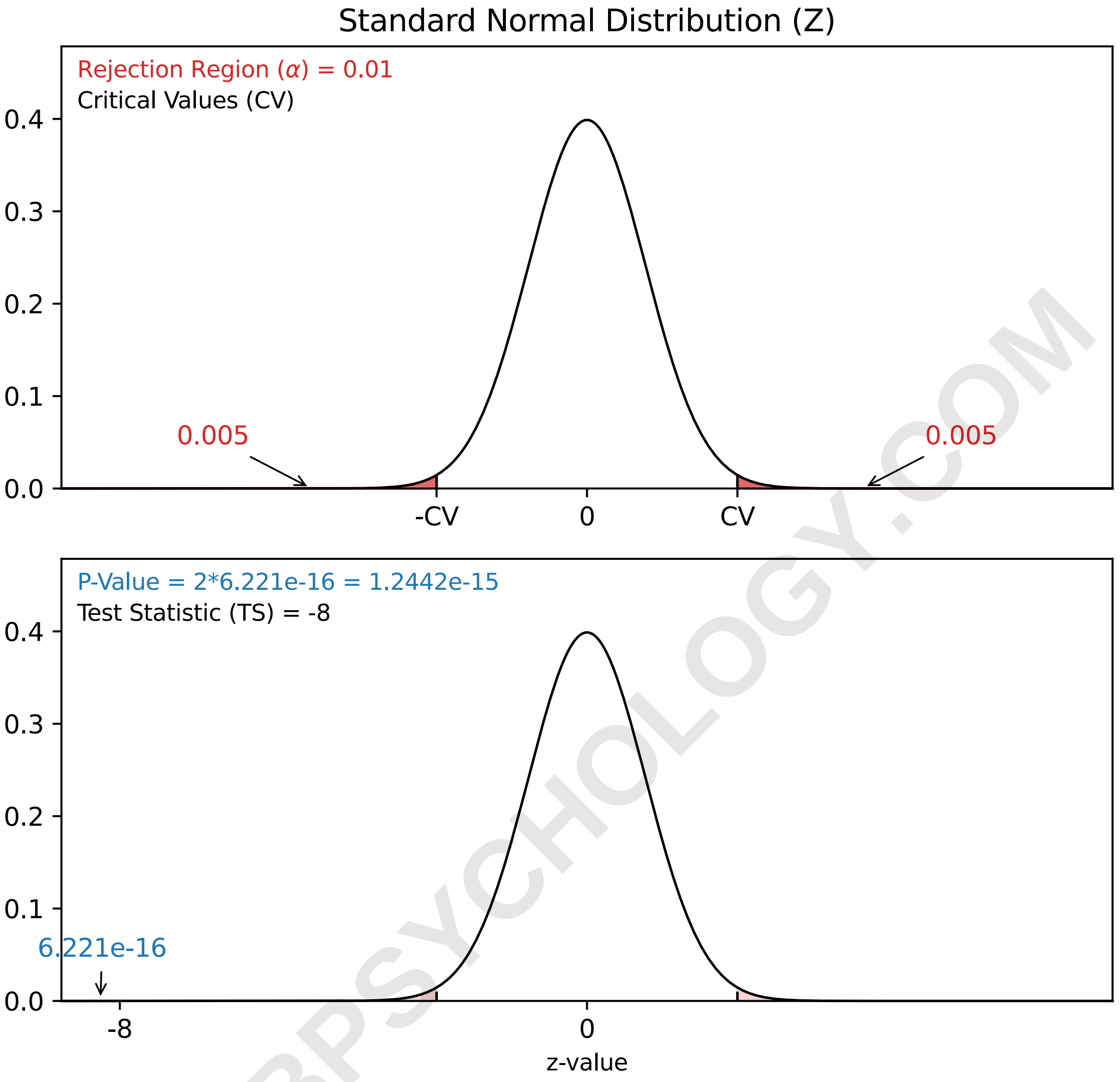
With R use the built-in `pnorm()` function find the P-value of a Z-value smaller than -8 for a two tailed test:

```
2*pnorm(-8)
```

Using either method we can find that the P-value is (approx underline{1.25 cdot 10^{-15}}) or (0.00000000000000125)

This tells us that the significance level ((alpha)) would need to be bigger than 0.000000000000125%, to **reject** the null hypothesis.

Here is an illustration of this test in a graph:



This P-value is **smaller** than any of the common significance levels (10%, 5%, 1%).

So the null hypothesis is **rejected** at all of these significance levels.

And we can summarize the conclusion stating:

The sample data **supports** the claim that "The share of Nobel Prize winners that are women is not 50%" at a **10%, 5%, and 1% significance level**.

Calculating a P-Value for a Hypothesis Test with Programming

Many programming languages can calculate the P-value to decide outcome of a hypothesis test.

Using software and programming to calculate statistics is more common for bigger sets of data, as calculating manually becomes difficult.

The P-value calculated here will tell us the **lowest possible significance level** where the null-hypothesis can be rejected.

Example

With Python use the scipy and math libraries to calculate the P-value for a two-tailed tailed hypothesis test for a proportion.

Here, the sample size is 100, the occurrences are 10, and the test is for a proportion different from than 0.50.

```
import scipy.stats as stats

import math

# Specify the number of occurrences (x), the sample size (n), and the proportion claimed in the null-hypothesis (p)

x = 10

n = 100

p = 0.5

# Calculate the sample proportion

p_hat = x/n

# Calculate the test statistic

test_stat = (p_hat-p)/(math.sqrt((p*(1-p))/(n)))

# Output the p-value of the test statistic (two-tailed test)

print(2*stats.norm.cdf(test_stat))
```

Example

With R use the built-in `prop.test()` function find the P-value for a left tailed hypothesis test for a proportion.

Here, the sample size is 100, the occurrences are 10, and the test is for a proportion different from 0.50.

```
# Specify the sample occurrences (x), the sample size (n), and the null-hypothesis claim (p)

x <- 10

n <- 100

p <- 0.5

# P-value from left-tail proportion test at 0.01 significance level

prop.test(x, n, p, alternative = c("two.sided"), conf.level = 0.99, correct = FALSE)$p.value
```

Note: The `conf.level` in the R code is the reverse of the significance level.

Here, the significance level is 0.01, or 1%, so the conf.level is 1-0.01 = 0.99, or 99%.

Left-Tailed and Two-Tailed Tests

This was an example of a **two** tailed test, where the alternative hypothesis claimed that parameter is **different** from the null hypothesis claim.

You can check out an equivalent step-by-step guide for other types here:

Right-Tailed TestLeft-Tailed Test
★+1 W3schools PathfinderTrack your progress - it's free!

Log in

Sign Up