Hypothesis testing

Example:

Jeffrey, as an eight-year-old, established a mean time of 16.43 seconds for swimming the 25-yard freestyle, with a standard deviation of 0.8 seconds. His dad, Frank, thought that Jeffrey could swim the 25-yard freestyle faster using goggles. Frank bought Jeffrey a new pair of expensive goggles and timed Jeffrey for 15, 25-yard freestyle swims. For the 15 swims, Jeffrey's mean time was 16 seconds. **Frank thought that the goggles helped Jeffrey to swim faster than the 16.43 seconds**.

Hypothesis testing is a statistical procedure or form of statistical inference that uses data from a sample to draw conclusions about a population parameter. **OR**

Hypothesis testing is really a systematic way to test claims or ideas about a group or population.

Steps of Hypothesis Testing

1. Setting null and alternative hypothesis
2. Setting the level of significance
3. Calculating the Test Statistics
4. Calculating probability value (*p*-value), or finding the rejection region
5. Making a decision about the null hypothesis
6. Stating an overall conclusion

Step 1 Set the null and alternative hypothesis:

Each hypothesis test includes two hypotheses about the population. One is the null hypothesis and is denoted by $\boldsymbol{H_0}$, which is a statement of a particular parameter value. This hypothesis is assumed to be true until there is evidence to suggest otherwise.

The second hypothesis is called the alternative hypothesis. An **alternative hypothesis** (denoted $\boldsymbol{H_1}$**)**, is the opposite of what is stated in the null hypothesis. The alternative hypothesis is a statement of a range of alternative values in which the parameter may fall.

Example:

We want to test whether the mean GPA of students in UIU is different from 2.5 (out of 4.0). The null and alternative hypotheses are:

$H_0 : \mu = 2.5$

$H_1 : \mu \neq 2.5$ $H_1 : \mu < 2.5$ $H_1 : \mu > 2.5$ based on wording the question

In our Case (Starting Example)

$H_0 : \mu = 16.43$ $H_1 : \mu < 16.43$

Step 2 Setting the level of significance

The level of significance is defined as the fixed probability of wrong elimination of null hypothesis when in fact, it is true. The level of significance is preset by the researcher with the outcomes of error. The significance level is denoted by the Greek letter **alpha – α**. This value is used as a probability cutoff for making decisions about the null hypothesis. This alpha value represents the probability we are willing to place on our test for making an incorrect decision regarding rejecting the null hypothesis.

Level of significance measure of the strength of the evidence that must be present in your sample before you will reject the null hypothesis and conclude that the effect is statistically significant.

A 5% level of significance means the chance that you will accept your alternative hypothesis when your null hypothesis is true. The smaller the significance level, the greater the burden of proof needed to reject the null hypothesis, or in other words, to support the alternative hypothesis.

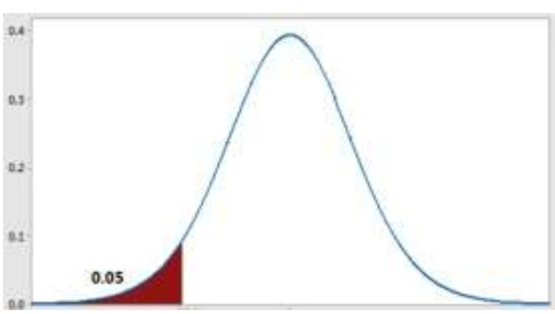
Rejection region: The rejection region is the values of test statistic for which the null hypothesis is rejected.

Acceptance region: The set of all possible values for which the null hypothesis is not rejected is called the acceptance region.

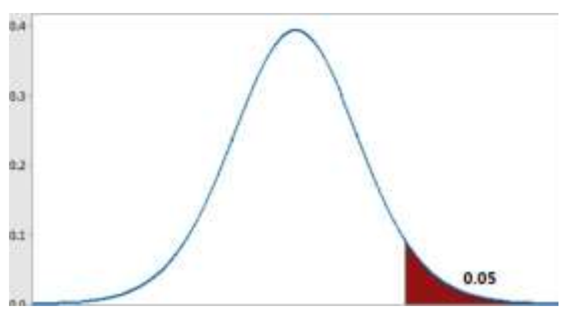
One-tailed hypothesis tests

One-tailed hypothesis tests are also known as directional and one-sided tests because you can test for effects in only one direction

The rejection region for one-tailed test is shown below:



$H_0 : \mu = 16.43$
$H_1 : \mu < 16.43$

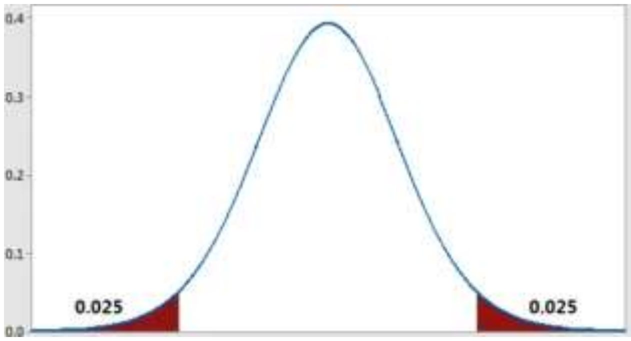


$H_0 : \mu = 16.43$
$H_1 : \mu > 16.43$

Two-tailed hypothesis tests

Two-tailed hypothesis tests are also known as nondirectional and two-sided tests because you can test for effects in both directions.

The rejection region for **two-tailed** test is shown below:



$$H_0 : \mu = 16.43$$
$$H_1 : \mu \neq 16.43$$

Step 3 **Calculating the Test-statistic**

$$p(x < 16) = p\left(z < \frac{16 - 16.43}{0.8}\right) = p(z < -0.5375)$$

Step 4 **Drawing Conclusion**

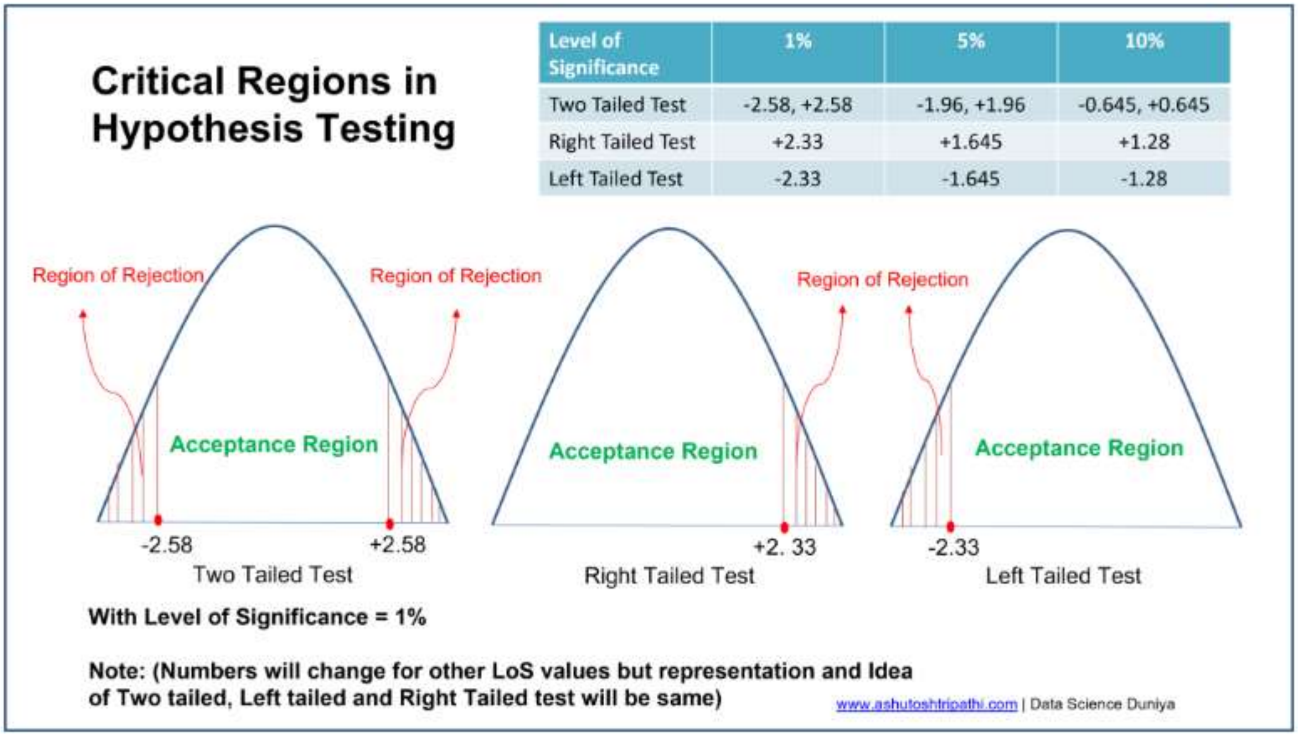
Type I error and type II error

A type I error (false-positive) occurs if an investigator rejects a null hypothesis that is actually true in the population; a type II error (false-negative) occurs if the investigator fails to reject a null hypothesis that is actually false in the population.

Table of error types		Null hypothesis ($\boldsymbol{H_0}$) is	
		True	False
Decision about null hypothesis ($\boldsymbol{H_0}$)	Accept	Correct inference (True negative) (Probability = $1-\alpha$)	Type II error (False negative) (probability = $\boldsymbol{\beta}$)
	Reject	Type I error (False positive) (Probability = α)	Correct inference (True positive) (Probability = $1-\beta$)

Critical Regions

In hypothesis testing, critical region is represented by set of values, where null hypothesis is rejected. So it is also known as region of rejection. It takes different boundary values for different level of significance. Below info graphics shows the region of rejection that is critical region and region of acceptance with respect to the level of significance 1%.



Example 1:

A Telecom service provider claims that individual customers pay on an average 400 rs. per month with standard deviation of 25 rs. A random sample of 50 customers' bills during a given month is taken with a mean of 250 and standard deviation of 15. What to say with respect to the claim made by the service provider?

Solution:

First thing first, Note down what is given in the question:

H0 (Null Hypothesis) : $\mu = 400$

H1 (Alternate Hypothesis): $\mu \neq 400$ (Not equal means either $\mu > 400$ or $\mu < 400$ Hence it will be validated with two tailed test)

$\sigma = 25$ (Population Standard Deviation)

LoS (α) = 5% (Take 5% if not given in question)

n = 50 (Sample size)

xbar $\bar{x}$ = 250 (Sample mean)

s = 15 (sample Standard deviation)

n > = 30 hence will go with z-test

Step 1:

Calculate z using z-test formula as below:

$z = (\bar{x} - \mu)/ (\sigma/\sqrt{n})$

$z = (250 - 400) / (25/\sqrt{50})$

z = -42.42

Step 2:

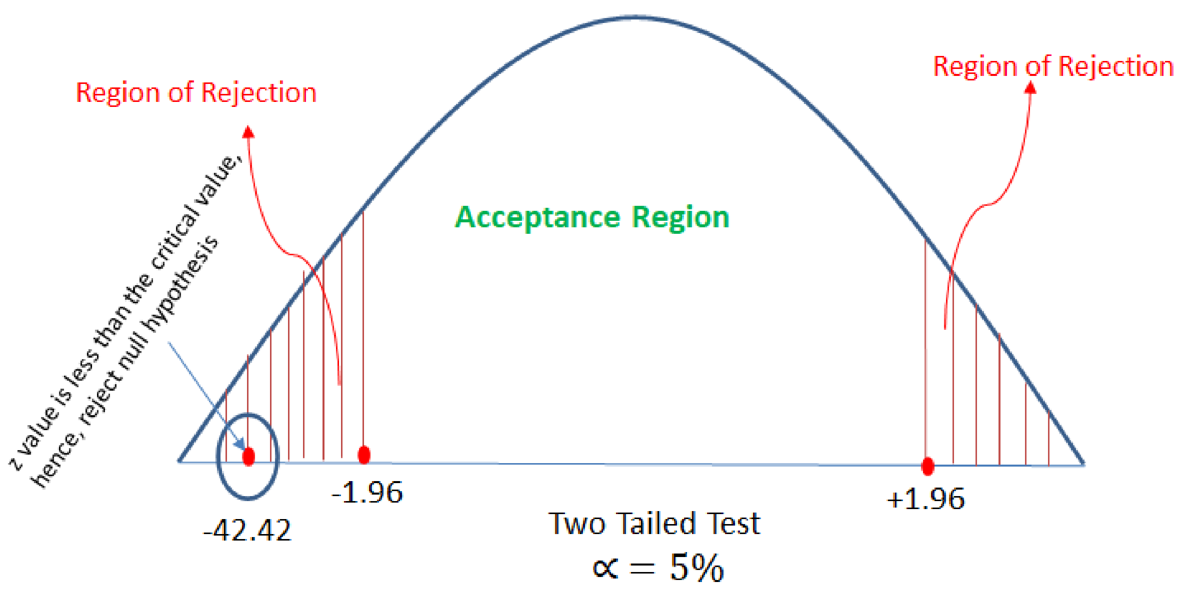
get z critical value from z table for α = 5%

z critical values = (-1.96, +1.96)

to accept the claim (significantly), calculated z should be in between

-1.96 < z < +1.96

but calculated z (-42.42) < -1.96 which mean reject the null hypothesis



z-test example 1

Example 2:

It is claimed that the mean of the population is 67 at 5% level of significance. Mean obtained from a random sample of size 100 is 64 with SD 3. Validate the claim.

Solution:

First thing first, Note down what is given in the question:

H0 (Null Hypothesis) : $\mu = 67$

H1 (Alternate Hypothesis): $\mu \neq 67$ (Not equal to means either $\mu > 67$ or $\mu < 67$ Hence it will be validated with two tailed test)

LoS (α) = 5%

n = 100 (Sample size)

xbar $\bar{x} = 64$ (Sample mean)

s = 3 (sample Standard deviation)

n > = 30 hence will go with z-test

Step 1: Calculate z using z-test formula as below:

$z = (\bar{x} - \mu)/ (\sigma/\sqrt{n})$

$z = (64 - 67) / (3/\sqrt{100})$ (in question population standard deviation is not given, in that case take sample standard deviation)

$z = -10$

step 2:

calculate z critical value for α = 5% from z-table.

so from z-table Z critical value = -1.96, +1.96 (will get two values due two tailed test)

step 3:

check if calculated z value is in between z critical value then accept the null hypothesis if z calculated is outside z critical then reject the null hypothesis.

Here, z calculated value = -10 which is much lesser than the left side z critical value -1.96, hence will reject the null hypothesis.

Conclusion:

with given data it is significantly proven that population mean is not equal to 67.

Example 3:

In the past, the mean height of plants of a particular species has been 2.3 m. A random sample of 60 plants of this species was treated with fertiliser and the mean height of these 60 plants was found to be 2.4 m. Assume that the standard deviation of the heights of plants treated with fertiliser is 0.4 m.

Carry out a test at the 2.5% significance level of whether the mean height of plants treated with fertiliser is greater than 2.3 m. [5]

Solution:

H_0: Pop mean height = 2.3 H_1: Pop mean height > 2.3
$\dfrac{2.4-2.3}{\frac{0.4}{\sqrt{60}}}$
1.936 or 1.937 or 1.94
'1.936' < 1.96
[Do not reject H_0] No evidence that (mean) height (with fertiliser) is more than without

Example 4:

Arvind uses an ordinary fair 6-sided die to play a game. He believes he has a system to predict the score when the die is thrown. Before each throw of the die, he writes down what he thinks the score will be. He claims that he can write the correct score more often than he would if he were just guessing. His friend Laxmi tests his claim by asking him to write down the score before each of 15 throws of the die. Arvind writes the correct score on exactly 5 out of 15 throws.

Test Arvind's claim at the 10% significance level. [5]

Solution:

H_0: P(correct) = $\frac{1}{6}$ H_1: P(correct) > $\frac{1}{6}$
$1-({}^{15}C_4\times(\frac{5}{6})^{11}\times(\frac{1}{6})^4+{}^{15}C_3\times(\frac{5}{6})^{12}\times(\frac{1}{6})^3+{}^{15}C_2\times(\frac{5}{6})^{13}\times(\frac{1}{6})^2+15\times(\frac{5}{6})^{14}\times\frac{1}{6}+(\frac{5}{6})^{15})$
0.0898 or 0.0897 (3 sf)
0.0898 < 0.1
[Reject H_0] There is evidence (at the 10% level) that Arvind can predict scores

Example 5:

The percentage of people having a different medical condition is thought to be 30%. A researcher suspects that the true percentage is less than 30%. In a medical trial a random sample of 28 people was selected and 4 people were found to have this condition.

Use a binomial distribution to test the researcher's suspicion at the 2% significance level. [5]

Solution:

H_0: $p = 0.3$
H_1: $p < 0.3$

$0.7^{28} + 28 \times 0.7^{27} \times 0.3 + {}^{28}C_2 \times 0.7^{26} \times 0.3^2 + {}^{28}C_3 \times 0.7^{25} \times 0.3^3 + {}^{28}C_4 \times 0.7^{24} \times 0.3^4$

0.0474

$0.0474 > 0.02$ [Not reject H_0]

No evidence that suspicion is true.

Example 6:

A machine is supposed to produce random digits. Bob thinks that the machine is not fair and that the probability of it producing the digit 0 is less than $\frac{1}{10}$. In order to test his suspicion he notes the number of times the digit 0 occurs in 30 digits produced by the machine. He carries out a test at the 10% significance level.

(a) State suitable null and alternative hypotheses. [1]

(b) Find the rejection region for the test.

Solution:

(a)	H_0: $P(0) = \frac{1}{10}$ H_1: $P(0) < \frac{1}{10}$
(b)	For B(30,0.1)
	$P(X = 0) = 0.9^{30}$ [= 0.0424] [<0.1]
	$P(X = 0 \text{ or } 1) = 0.9^{30} + 30\times0.9^{29}\times0.1=0.184$ [>0.1]
	Rejection region is 0 zeros

Example 7:

Accidents on a stretch of road occur at the rate of seven each month. New traffic measures have been put in places to reduce the number of accidents. In the following month, there were only two accidents.

(a) Test at the 5% level of significance if there is evidence that the new traffic measures have significantly reduced the number of accidents.

(b) Over a period of 6 months, there are 32 accidents. It is claimed that new traffic measures are now longer reducing the number of accidents. Test this at 5% level of significance