TEST OF HYPOTHESIS

Hypothesis: A hypothesis is an educated guess about something in the world around you. It should be testable, either by experiment or observation. In other words, a statement about the parameters describing a population (not a sample). For example:

- A new medicine you think might work.
- A way of teaching you think might be better.
- A possible location of new species.
- A fairer way to administer standardized tests.

Statistic: A value calculated from a sample without any unknown parameters, often to summarize the sample for comparison purposes.

Simple hypothesis: Any hypothesis which specifies the population distribution completely.

Composite hypothesis: Any hypothesis which does not specify the population distribution completely.

Null hypothesis (H_0): A hypothesis is speculation or theory based on insufficient evidence that lends itself to further testing and experimentation. With further testing, a hypothesis can usually be proven true or false. A null hypothesis is a type of hypothesis used in statistics that proposes that there is no difference between certain characteristics of a population.

Alternative hypothesis (H_1): A hypothesis (often composite) associated with a theory one would like to prove. An alternative hypothesis simply is the inverse, or opposite, of the null hypothesis. The alternative hypothesis is a position that states something is happening, a new theory is true instead of an old one.

It is standard practice for statisticians to conduct tests in order to determine whether or not a "tentative hypothesis" concerning the observed phenomena of the world can be supported.

Positive data: Data that enable the investigator to reject a null hypothesis.

TEST OF HYPOTHESIS

Statistical test: A procedure whose inputs are samples and whose result is a hypothesis.

Hypothesis testing: It is used to test the validity of a ***null hypothesis (claim)*** that is made about a population using sample data. The ***alternative hypothesis (complement)*** is the one you would believe if the null hypothesis is concluded to be untrue. In other words, we'll make a ***null hypothesis (claim)*** *and* use sample data to check if the claim is valid. If the claim isn't valid, then we'll choose our ***alternative hypothesis (complement)*** instead.

Statistical analysts test a hypothesis by measuring and examining a random sample of the population being analyzed. All analysts use a random population sample to test two different hypotheses: the null hypothesis and the alternative hypothesis.

Region of rejection: The set of values of the test statistic for which the null hypothesis is rejected. The **rejection region** (also called a critical region) is a part of the testing process. Specifically, it is an area of probability that tells you if your theory (claim) is probably true.

Region of acceptance: The set of values of the test statistic for which we fail to reject the null hypothesis. The **acceptance region** is basically the **complement** of the rejection region. If your result does not fall into the rejection region, it must fall into the acceptance region.

Critical value: The threshold value delimiting the regions of acceptance and rejection for the test statistic.

Type I error: The first kind of error is the rejection of a true null hypothesis as the result of a test procedure. This sort of error is called a type I error and is sometimes called an error of the first kind.

Type II error: The second kind of error is the failure to reject a false null hypothesis as the result of a test procedure. This sort of error is called a type II error and is sometimes called an error of the second kind.

TEST OF HYPOTHESIS

	H_0 is true	H_1 is true
Accept the null hypothesis	Right decision	Wrong decision (Type II Error)
Reject the null hypothesis	Wrong decision (Type I Error)	Right decision

False-positive error: A false positive error is a type I error where the test is checking a single condition, and wrongly gives an affirmative decision.

False-negative error: A false negative error is a type II error occurring in a test where a single condition is checked for and the result of the test is erroneous that the condition is absent.

The type I error rate or significance level is the probability of rejecting the null hypothesis given that it is true. It is denoted by the Greek letter α (alpha) and is also called the alpha level. The rate of the type II error is denoted by the Greek letter β (beta) and related to the power of a test, which equals $1-\beta$. These two types of error rates are traded off against each other: for any given sample set, the effort to reduce one type of error generally results in increasing the other type of error.

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 = $1-\beta$)

Power: The power of a binary hypothesis test is the probability that the test rejects the null hypothesis ($\boldsymbol{H_0}$) when a specific alternative hypothesis ($\boldsymbol{H_1}$) is true.

Power of a test $(\boldsymbol{1-\beta})$: The test's probability of correctly rejecting the null hypothesis. The complement of the false-negative rate, β.

TEST OF HYPOTHESIS

The significance level of a test (α): A study's defined **significance level**, denoted by α, is the probability of the study rejecting the null hypothesis, given that the null hypothesis was assumed to be true. It is the maximum exposure to erroneously rejecting $\boldsymbol{H_0}$. Testing $\boldsymbol{H_0}$ at the significance level, α means the testing $\boldsymbol{H_0}$ with a test whose size does not exceed α.

The Probability value (*p*-value): In statistical hypothesis testing, the ***p*-value** or **probability value** is, for a given statistical model, the probability that, when the null hypothesis is true, the statistical summary would be greater than or equal to the actual observed results.

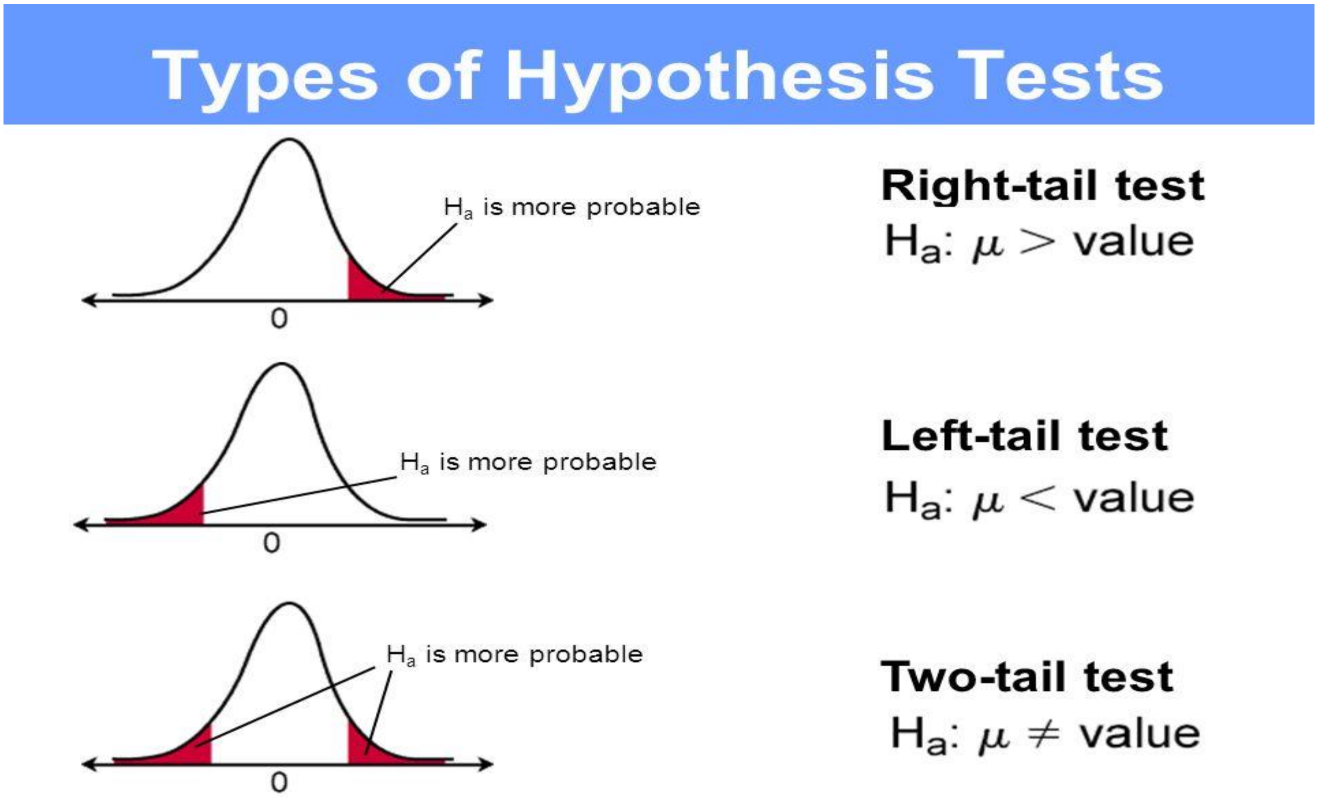
The *p*-value is used as an alternative to rejection points to provide the smallest level of significance at which the null hypothesis would be rejected. A smaller *p*-value means that there is stronger evidence in favor of the alternative hypothesis. The *p*-value approach to hypothesis testing uses the calculated probability to determine whether there is evidence to reject the null hypothesis. The level of statistical significance is often expressed as a *p*-value between 0 and 1.

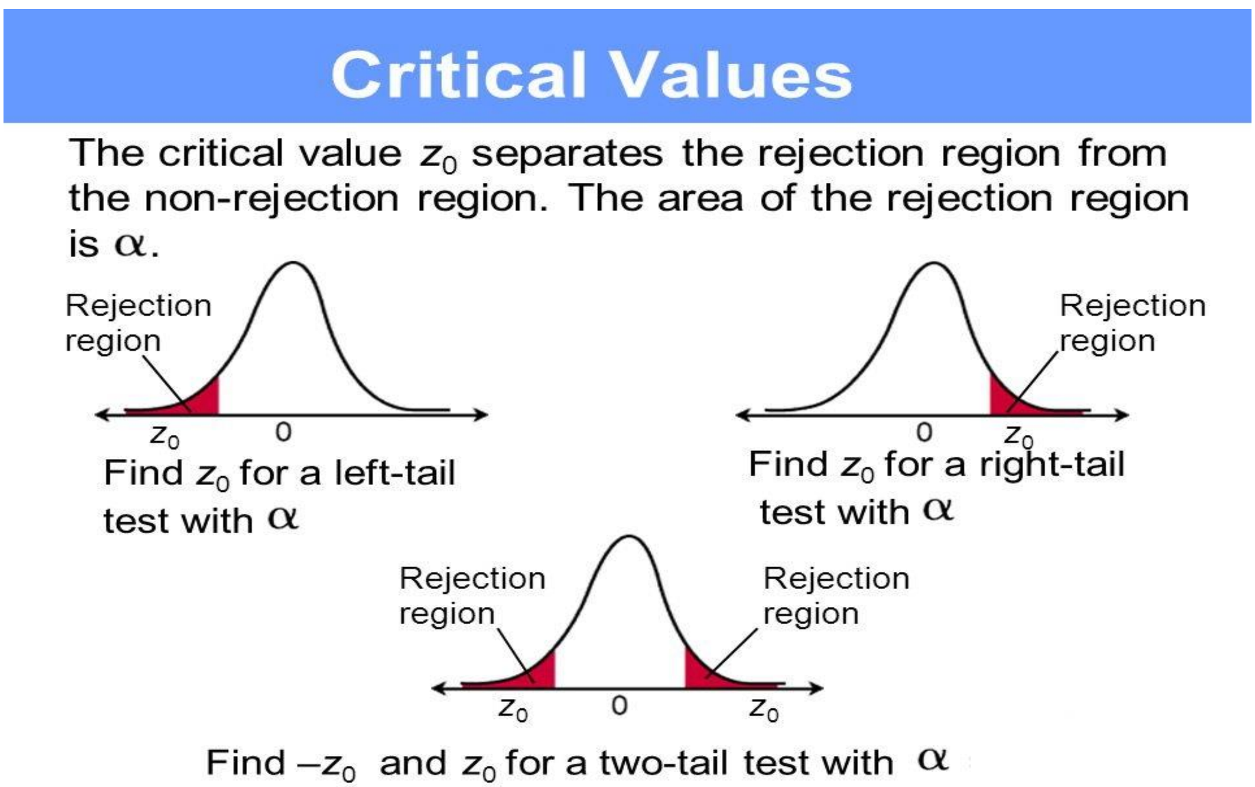
One-Tailed Test: A one-tailed test is a statistical test in which the critical area of a distribution is one-sided so that it is either greater than or less than a certain value, but not both. If the sample being tested falls into the one-sided critical area, the alternative hypothesis will be accepted instead of the null hypothesis.

Two-Tailed Test: In statistics, a two-tailed test is a method in which the critical area of a distribution is two-sided and tests whether a sample is greater than or less than a certain range of values. It is used in null-hypothesis testing and testing for statistical significance. If the sample being tested falls into either of the critical areas, the alternative hypothesis is accepted instead of the null hypothesis.

Two-Tailed vs One-Tailed Rejection Regions: Which type of test is determined by your null hypothesis statement? For example, if your statement asks "Is the average growth rate greater than 10cm a day?" that's a one-tailed test because you are only interested in one direction (greater than 10cm a day). You could also have a single rejection region for "less than". For example, "Is the growth rate less than 10cm a day?" A two-tailed test, with two rejection regions, would be used when you want to know if there's a difference in both directions (greater than and less than).

TEST OF HYPOTHESIS





TEST OF HYPOTHESIS

1. Design a decision rule to test the hypothesis that a die is fair if we take a sample of 150 trials of the die to get even/odd faces and use 0.05 as the significance level. Predict the acceptance and critical region.

Solution: Let us assume the following hypotheses

Null hypothesis H_0: $p = 0.5$
Alternative hypothesis H_1: $p \neq 0.5$

Here, the number of samples $n = 150$ and the claimed probability of success $p = 0.5$ and hence $q = 0.5$. This is a two-tailed test and the given significance level $\alpha = 0.05$. The required $z\alpha/_2$ can be estimated from the Normal table as

$$P\left(-z\alpha/_2 \leq Z \leq z\alpha/_2\right) = 1 - \alpha$$

$$or, \qquad \varphi\left(z\alpha/_2\right) = \frac{2-\alpha}{2}$$

$$or, \qquad \varphi\left(z\alpha/_2\right) = \frac{2-0.05}{2}$$

$$or, \qquad \varphi\left(z\alpha/_2\right) = 0.975$$

$$or, \qquad z\alpha/_2 = 1.96$$

Now, $\mu = np = 75$ and $\sigma = \sqrt{npq} = 6.124$. Then we will accept the claim (H_0) of that the die is fair if

$$-z\alpha/_2 \leq Z \leq z\alpha/_2$$

$$or, \qquad -z\alpha/_2 \leq \frac{X-\mu}{\sigma} \leq z\alpha/_2$$

$$or, \qquad \mu - z\alpha/_2\sigma \leq X \leq \mu + z\alpha/_2\sigma$$

$$or, \qquad 75 - 1.96 * 6.124 \leq X \leq 75 + 1.96 * 6.124$$

$$or, 62.997 \leq X \leq 87.003$$

So, the claim of the die is fair will be accepted if the number of even/odd faces will appear within 63 and 87 (faces of the die cannot be measured by a fraction) and rejected otherwise.

TEST OF HYPOTHESIS

2. A pharmaceutical company produces a new medicine and they claimed that it will reduce the migraine pain very fast with 95% accuracy. Now design decision rule for the process with significance 0.01 by applying the medicine to 120 people. Make a decision if 115 people get relief from the migraine pain by using the medicine.

Solution: Let us assume the following hypotheses

Null hypothesis H_0: $p = 0.95$
Alternative hypothesis H_1: $p < 0.95$

Here, the number of samples $n = 120$ and the claimed probability of success $p = 0.95$ and hence $q = 0.05$. This is a one-tailed test and the given significance level $\alpha = 0.01$. Since the medicine is newly developed the privilege of the significance will go in favor of the pharmaceutical company. So, the required z_α can be estimated from the Normal table as

$$P(Z > -z_\alpha) = 1 - \alpha$$
$$or, \quad P(Z > -z_\alpha) = 1 - 0.01$$
$$or, \quad P(Z < z_\alpha) = 0.99$$
$$or, \quad \varphi(z_\alpha) = 0.99$$
$$or, \quad z_\alpha = 2.33$$

Now, $\mu = np = 114$ and $\sigma = \sqrt{npq} = 2.39$. Then we will accept the claim (H_0) of the pharmaceutical company if

$$Z > -z_\alpha$$
$$or, \quad \frac{X - \mu}{\sigma} > -z_\alpha$$
$$or, \quad X > \mu - z_\alpha \sigma$$
$$or, \quad X > 114 - 2.33 * 2.39$$
$$or, \quad X > 108.43$$

So, the claim of the pharmaceutical company will be accepted if more than 109 people (people cannot be measured by a fraction) get relief from the migraine pain and rejected otherwise. Since 115 people relieved from the migraine pain, the claim of the pharmaceutical company should be accepted.

TEST OF HYPOTHESIS

3. Design a decision rule to test the hypothesis that a coin is fair if we take a sample of 120 trials of the die to get head/tail and use 0.1 as the significance level. Predict the acceptance and critical region.

Solution: Try yourself.

4. A pharmaceutical company produces a new medicine and they claimed that it will reduce the migraine pain very fast with 85% accuracy. Now design decision rule for the process with significance 0.05 by applying the medicine to 200 people.

Solution: Try yourself.

5. A company produces electric bulbs whose average lifetime is 180 days and an average variation of 15 days. It is claimed that by using a newly developed process the mean lifetime of the bulbs can be increased.
   (i) Design a decision rule for the process at the 0.01 significance level to test 100 bulbs. Also, what about the decision if the average lifetime of a bulb 182 days and 197 days.
   (ii) If the new process has increased the mean lifetime to 192 days and assuming the estimated sample mean 187 days, find α and β for 30 samples.
   (iii) Suppose for 50 samples average lifetime is observed as 185 days, estimate the p-value of the claim of the manufacturer.

Solution: Let us assume the following hypotheses

Null hypothesis H_0: $\mu = 180$ days.
Alternative hypothesis H_1: $\mu > 180$ days.

(i) Here, the number of samples $n = 100$ and the standard deviation is $\sigma = 15$.

This is a one-tailed test and the given significance level $\alpha = 0.01$. The required z_α can be estimated from the Normal table as

$$P(Z > z_\alpha) = \alpha$$

$$or, \quad P(Z > z_\alpha) = 0.01$$

$$or, \quad 1 - \varphi(z_\alpha) = 0.01$$

$$or, \quad \varphi(z_\alpha) = 0.99$$

$$or, \quad z_\alpha = 2.33$$

Now, we will accept the claim (H_1) of the company if

$$Z > z_\alpha$$

$$or, \quad \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > z_\alpha$$

$$or, \quad \bar{X} > \mu + z_\alpha \, \sigma/\sqrt{n}$$

$$or, \quad \bar{X} > 180 + 2.33 * 15/\sqrt{100}$$

$$or, \quad \bar{X} > 183.495$$

TEST OF HYPOTHESIS

So, the claim of the company will be accepted if the lifetime increased more than 183.495 days and rejected otherwise. Now, the claim will be rejected and accepted for the average sample lifetime 182 days and 197 days, respectively.

(ii) In this case alternative hypothesis H_1: $\mu = 192$ years, sample size $n = 30$, and the estimated sample mean $\bar{x} = 187$. Then we have

$$\begin{aligned}\alpha &= P(E_1)\\ &= P(\bar{X} \geq 187;\ \mu = 180)\\ &= P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \geq \frac{187-180}{15/\sqrt{30}}\right)\\ &= P(Z \geq 2.56)\\ &= 1-\varphi(2.56)\\ &= 1-0.9948\\ &= 0.0052\end{aligned}$$

$$\begin{aligned}\beta &= P(E_2)\\ &= P(\bar{X} < 187;\ \mu = 192)\\ &= P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < \frac{187-192}{15/\sqrt{30}}\right)\\ &= P(Z < -1.83)\\ &= P(Z > 1.83)\\ &= 1-\varphi(1.83)\\ &= 1-0.9664\\ &= 0.0336\end{aligned}$$

(iii) The p-value can be estimated as

$$\begin{aligned}p-value &= P(\bar{X} > 185;\ \mu = 180)\\ &= P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} > \frac{185-180}{15/\sqrt{50}}\right)\\ &= P(Z > 2.36)\\ &= 1-\varphi(2.36)\\ &= 1-0.9909\\ &= 0.0091\end{aligned}$$

TEST OF HYPOTHESIS

6. A company produces mechanical tools whose average lifetime is 18 years and variance 16 years. It is claimed that in a new process the mean lifetime can be increased.
   (i) Design a decision rule for the process at the 0.1 significance level to test 25 tools.
   (ii) If the estimated average lifetime for 20 samples is 20.5 years, find the p-value of the claim of the producer.
   (iii) If the new process has increased the mean lifetime to 21.25 years. Find α and β for the estimated sample mean 19.75 years for 30 samples.

Solution: Let us assume the following hypotheses

Null hypothesis H_0: $\mu = 18$ years.
Alternative hypothesis H_1: $\mu > 18$ years.

(i) Here, the number of samples $n = 25$ and the standard deviation is $\sigma = 4$.

This is a one-tailed test and the given significance level $\alpha = 0.1$. The required z_α can be estimated from the Normal table as

$$P(Z > z_\alpha) = \alpha$$
$$or, \quad P(Z > z_\alpha) = 0.1$$
$$or, \quad 1 - \varphi(z_\alpha) = 0.1$$
$$or, \quad \varphi(z_\alpha) = 0.9$$
$$or, \quad z_\alpha = 1.28$$

Now, we will accept the claim (H_1) of the company if

$$Z > z_\alpha$$
$$or, \quad \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > z_\alpha$$
$$or, \quad \bar{X} > \mu + z_\alpha \, \sigma/\sqrt{n}$$
$$or, \quad \bar{X} > 18 + 1.28 * 4/\sqrt{25}$$
$$or, \quad \bar{X} > 19.024$$

So, the claim of the company will be accepted if the lifetime increased more than 19.024 years and rejected otherwise.

(ii) The p-value can be estimated as

$$p - value = P(\bar{X} > 20.5;\ \ \mu = 18)$$

$$= P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > \frac{20.5 - 18}{4/\sqrt{20}}\right)$$

$$= P(Z > 2.80)$$

$$= 1 - \varphi(2.80)$$

$$= 1 - 0.9974$$

$$= 0.0026$$

(iii) In this case alternative hypothesis H_1: $\mu = 21.25$ years, sample size $n = 30$, and the estimated sample mean $\bar{x} = 19.75$. Then we have

$$\alpha = P(E_1)$$

$$= P(\bar{X} \geq 19.75;\ \ \mu = 18)$$

$$= P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \geq \frac{19.75 - 18}{4/\sqrt{30}}\right)$$

$$= P(Z \geq 2.40)$$

$$= 1 - \varphi(2.40)$$

$$= 1 - 0.9918$$

$$= 0.0082$$

$$\beta = P(E_2)$$

$$= P(\bar{X} < 19.75;\ \ \mu = 21.25)$$

$$= P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < \frac{19.75 - 21.25}{4/\sqrt{30}}\right)$$

$$= P(Z < -2.05)$$

$$= P(Z > 2.05)$$

$$= 1 - \varphi(2.05)$$

$$= 1 - 0.9798$$

$$= 0.0202$$

<u>**TEST OF HYPOTHESIS**</u>

7. The breaking strengths of cables produced by a manufacturer have a mean of 2000 pounds and a standard deviation of 300 pounds. By a new technique in the manufacturing process, it is claimed that the breaking strength can be increased.
   (i) Can we support the claim at the 0.1 significance level? Construct a hypothesis test for the 200 samples.
   (ii) To test this claim, a sample of 100 cables is tested and it is found that the mean breaking strength is 2050 pounds. Find the p-value of the test.

Solution: Try yourself.

8. A company produces an electric tool whose average lifetime is 200 days and average variation 30 days. It is claimed that in a newly developed process the mean lifetime can be increased. Design a decision rule for 160 samples with 0.05 significance. If the new process has increased the mean lifetime to 208 days, assuming a sample of 80 bulbs with estimated lifetime 205 days, find α and β.

Solution: Try yourself.