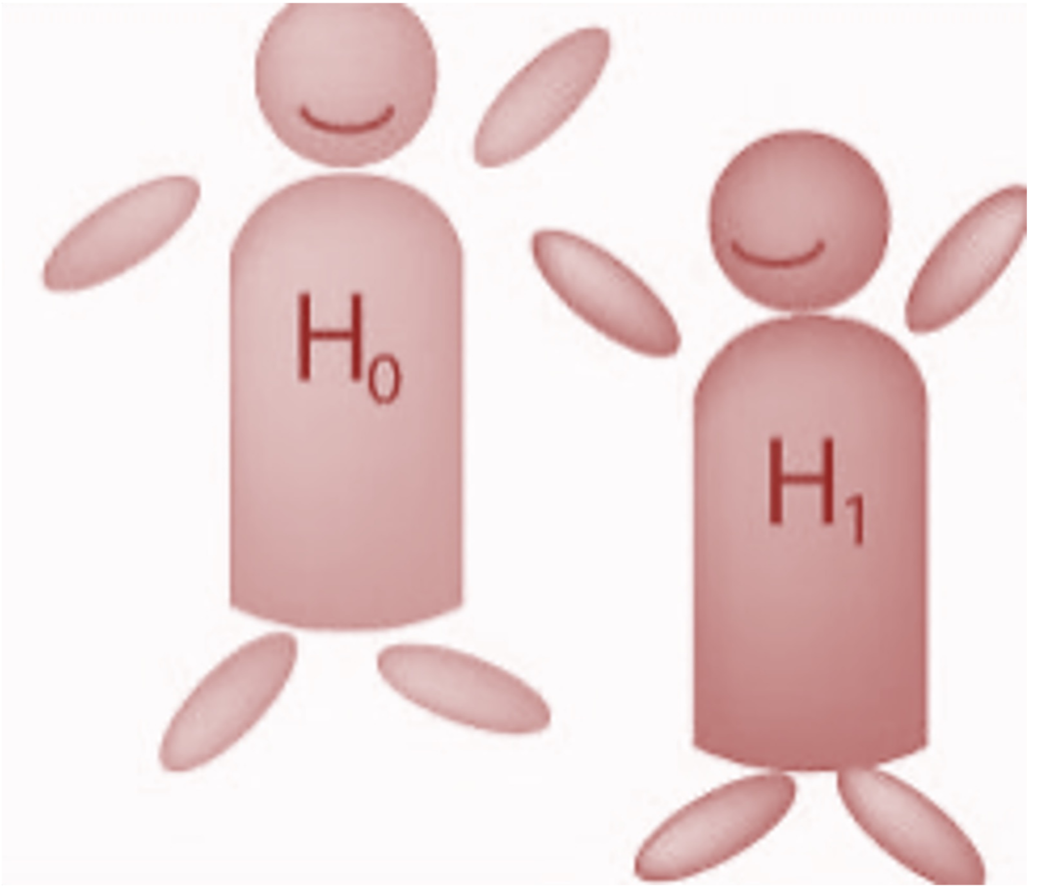
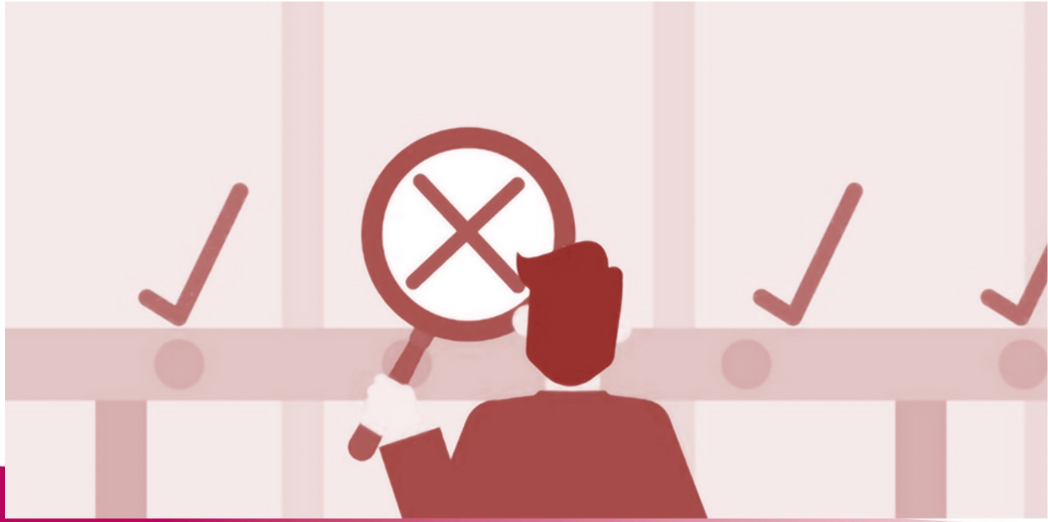
L07:
Hypothesis Testing - One Population





What is hypothesis testing?

A test of hypotheses is a method that uses sample data to decide between two competing claims (hypotheses) about the population characteristic.

We now would have a hypothesized value of the parameter which is assumed to be true and then use the collected data to see if the assumption should be rejected or not be rejected.

Goal and key steps in hypothesis testing

The goal of hypothesis testing is to determine the likelihood that a population parameter, such as the mean, is likely to be true.

Step 1: State the hypotheses.

Step 2: Set the criteria for a decision.

Step 3: Compute the test statistic.

Step 4: Make a decision.

Statistical Hypotheses: $\boldsymbol{H_0}$ and $\boldsymbol{H_1}$

Some key terms:

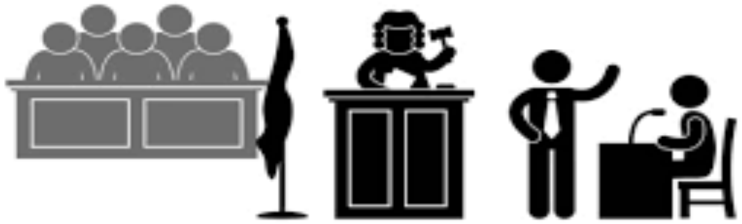
- The null hypothesis $\boldsymbol{H_0}$: It is the hypothesis that is a claim about a population characteristic and is assumed to be true and then tested to be rejected or not to be rejected formally.

The hypothesis statements are ALWAYS about the population – NEVER about a sample!

- The alternative hypothesis $\boldsymbol{H_a}$ or $\boldsymbol{H_1}$: It is the hypothesis that typically represents the underlying research question of the investigator and is the complement of H_0, i.e. it contains the values of parameter we accept if we reject H_0.

A **test of hypotheses** or **test procedure** is a method that uses sample data to decide between two competing claims (hypotheses) about a population characteristic.

Example



Let’s consider a murder trial (guilty or not guilty)

What is the null hypothesis?

H_0: the defendant is innocent

To determine which hypothesis is correct, the jury will listen to the evidence. Only if there is “evidence beyond a reasonable doubt” would the null hypothesis be rejected in favor of the alternative hypothesis.

What is the alternative hypothesis?

H_a: the defendant is guilty

If there is not convincing evidence, then we would “fail to reject” the null hypothesis. Remember that the actually verdict that is returned is “GUILTY” or “NOT GUILTY”. We never end up determining the null hypothesis is true – only that there is not enough evidence to say it’s not true.

So we will make one of two decisions:

- Reject the null hypothesis
- Fail to reject the null hypothesis

Example

An agronomist may want to decide on the basis of experiments whether or not a new fertilizer would produce a higher yield of soybeans than an old one whose mean is known to be 10.

- In this case the agronomist has to test $\mu_X > 10$, where μ_X is the mean of the random variable of the yield of soybean by the new fertilizer, assuming a normal population. Then, we have

$$\boxed{H_0: \mu_X = 10 \text{ against } H_1: \mu_X > 10}$$

Example

A manufacturer of pharmaceutical products may decide on the basis of samples whether or not 90% of all patients given a new medication will recover from a certain disease.

- In this case we might say that the manufacturer has to decide whether or not the parameter *p* of a binomial population equals 0.90.

$$\boxed{H_0: \mu_X = p = 0.9 \text{ against } H_1: \mu_X \neq 0.9}$$

Form of Hypothesis Testing

Null hypothesis

H_0: population characteristic = hypothesized value

Noted that the alternative hypothesis uses the same population characteristic and the same hypothesized value as the null hypothesis.

Alternative hypothesis

One-tailed test since you are interested in one direction

H_1: population characteristic > hypothesized value
H_1: population characteristic < hypothesized value
H_1: population characteristic ≠ hypothesized value

a **two-tailed** test

Example

Compact florescent (cfl) lightbulbs are much more energy efficient than regular incandescent lightbulbs. Ecobulb brand 60-watt cfl lightbulbs state on the package “Average life 8000 hours”. People who purchase this brand would be unhappy if the bulbs lasted less than 8000 hours. A sample of these bulbs will be selected and tested

Example

Compact florescent (cfl) lightbulbs are much more energy efficient than regular incandescent lightbulbs. Ecobulb brand 60-watt cfl lightbulbs state on the package "Average life 8000 hours". People who purchase this brand would be unhappy if the bulbs lasted less than 8000 hours. A sample of these bulbs will be selected and tested

What is the population characteristic of interest?

The true mean (μ) life of the cfl lightbulbs

Example

Compact florescent (cfl) lightbulbs are much more energy efficient than regular incandescent lightbulbs. Ecobulb brand 60-watt cfl lightbulbs state on the package "Average life 8000 hours". People who purchase this brand would be unhappy if the bulbs lasted less than 8000 hours. A sample of these bulbs will be selected and tested

What is the population characteristic of interest?

What is the hypothesized value?

8000 hours

The true mean (μ) life of the cfl lightbulbs

Example

Compact florescent (cfl) lightbulbs are much more energy efficient than regular incandescent lightbulbs. Ecobulb brand 60-watt cfl lightbulbs state on the package "Average life 8000 hours". People who purchase this brand would be unhappy if the bulbs lasted less than 8000 hours. A sample of these bulbs will be selected and tested

What is the population characteristic of interest?

What words indicate the direction of the alternative hypothesis?

What is the hypothesized value?

less than 8000 hours

8000 hours

The true mean (μ) life of the cfl lightbulbs

Example

Compact florescent (cfl) lightbulbs are much more energy efficient than regular incandescent lightbulbs. Ecobulb brand 60-watt cfl lightbulbs state on the package "Average life 8000 hours". People who purchase this brand would be unhappy if the bulbs lasted less than 8000 hours. A sample of these bulbs will be selected and tested

What is the population characteristic of interest?

What words indicate the direction of the alternative hypothesis?

What is the hypothesized value?

State the hypotheses :

less than 8000 hours

8000 hours

The true mean (μ) life of the cfl lightbulbs

$H_0: \mu = 8000$

$H_1: \mu < 8000$

Example

Tennis ball

Suppose the machine was initially calibrated to achieve the specification of μ = 3 inches. However, the manager is now concerned that the diameters no longer conform to this specification. If the mean diameter is not 3 inches, production will have to be halted.

Example

Tennis ball

Suppose the machine was initially calibrated to achieve the specification of μ = 3 inches. However, the manager is now concerned that the diameters no longer conform to this specification. If the mean diameter is not 3 inches, production will have to be halted.

What is the population characteristic of interest?

What words indicate the direction of the alternative hypothesis?

What is the hypothesized value?

Example

Suppose the machine was initially calibrated to achieve the specification of μ = 3 inches. However, the manager is now concerned that the diameters no longer conform to this specification. If the mean diameter is not 3 inches, production will have to be halted.

What is the population characteristic of interest?

State the hypotheses :

What words indicate the direction of the alternative hypothesis?

What is the hypothesized value?

$H_0: \mu = 3$

The true mean μ diameter of tennis balls

$H_1: \mu \neq 3$

Example

For each pair of hypotheses, indicate which are not legitimate and explain why

a) $H_0 : \mu = 15;$ $H_1 : \mu \geq 15$ ✗ must be only greater than!

b) $H_0 : \bar{x} = 4;$ $H_1 : \bar{x} > 4$ ✗ Must use a population characteristic -
Not sample statics

c) $H_0 : p = 0.1;$ $H_1 : p \neq 0.1$ ✓

d) $H_0 : p \neq 0.5;$ $H_1 : p = 0.5$ ✗ H_0 MUST be "=" !

Type of Hypothesis Testing

According to the form of the alternative hypothesis, we can have the following

Four types of tests:

I) SIMPLE TEST

$$\begin{cases} H_0: \mu_X = \mu_0 \\ H_1: \mu_X = \mu_1 \end{cases}$$

II) ONE-SIDED RIGHT TEST

$$\begin{cases} H_0: \mu_X = \mu_0 \\ H_1: \mu_X > \mu_0 \end{cases}$$

III) ONE-SIDED LEFT TEST

$$\begin{cases} H_0: \mu_X = \mu_0 \\ H_1: \mu_X < \mu_0 \end{cases}$$

IV) TWO-SIDED TEST

$$\begin{cases} H_0: \mu_X = \mu_0 \\ H_1: \mu_X \neq \mu_0 \end{cases}$$

Main Concept of Hypothesis Testing

1: Determine H_0 and H_1.

2: Under H_0, define a rare event –the event which happens with a very small probability in one experiment.

3: Collect data and compute the test statistics

4. Make decision: If data contradicts H_0, then reject H_0; otherwise, do NOT reject H_0.

Main Concept of Hypothesis Testing

In hypothesis testing, we only use **the data** to **see if there is enough evidence to reject $\boldsymbol{H_0}$**.

- If we have enough evidence to reject H_0, we can have great confidence that H_0 is false and H_1 is true.

- However, if **we do not have enough evidence to reject $\boldsymbol{H_0}$,** then *it* ***does not mean that*** *we have great confidence in the truth of* H_0. In this case, we should say **"do not reject $\boldsymbol{H_0}$", instead of "accept $\boldsymbol{H_0}$".**

Test Errors and Error Probabilities

- **Note that there is no perfect test statement. Each test statement must lead to the following two kinds of errors.**

	Not reject H_0	Reject H_0
If H_0 is true	No error	**TYPE I ERROR**
If H_0 is false	**TYPE II ERROR**	No error

TYPE I ERROR: the error of rejecting H_0 when it is in fact true.

TYPE II ERROR: the error of not rejecting H_0 when it is in fact false.

Type I and Type II error

The Transportation Statistics reports that for 72% of all domestic passenger flights arrived on time (meaning within 15 minutes of its scheduled arrival time).

Suppose that an airline with a poor on-time record decides to offer its employees a bonus if, in an upcoming month, the airline's proportion of on-time flights exceeds the overall industry rate of .72.

Let p be the actual proportion of the airline's flights that are on time during the month of interest. A random sample of flights might be selected and used as a basis for choosing between

H_0: $p = .72$

H_a: $p > .72$

TypeI: The error of rejecting H_0 when H_0 is true

TypeII: The error of failing to reject H_0 when H_0 is false

Type I error – the airline decides to reward the employees when the proportion of on-time flights doesn't exceeds .72

Type II error – the airline employees do not receive the bonus when they deserve it.

Type I and Type II

In 2004, Vertex Pharmaceuticals, a biotechnology company, issued a press release announcing that it had filed an application with the Food and Drug Administration to begin clinical trials of an experimental drug VX-680 that had been found to reduce the growth rate of pancreatic and colon cancer tumors in animal studies.

Data resulting from the planned clinical trials can be used to test:

Let μ = the true mean growth rate of tumors for patients taking the experimental drug

H_0: μ = mean growth rate of tumors for patients not taking the experimental drug

H_a: μ < mean growth rate of tumors for patients not taking the experimental drug

TypeI: The error of rejecting H_0 when H_0 is true
TypeII: The error of failing to reject H_0 when H_0 is false

Type I error : incorrectly conclude that the experimental drug is effective in slowing the growth rate of tumors

→Consequence: company would continue to devote resources to the development of the drug when it really is not effective.

Type I and Type II

In 2004, Vertex Pharmaceuticals, a biotechnology company, issued a press release announcing that it had filed an application with the Food and Drug Administration to begin clinical trials of an experimental drug VX-680 that had been found to reduce the growth rate of pancreatic and colon cancer tumors in animal studies.

Data resulting from the planned clinical trials can be used to test:

Let μ = the true mean growth rate of tumors for patients taking the experimental drug

H_0: μ = mean growth rate of tumors for patients not taking the experimental drug

H_a: μ < mean growth rate of tumors for patients not taking the experimental drug

TypeI: The error of rejecting H_0 when H_0 is true
TypeII: The error of failing to reject H_0 when H_0 is false

Type II error : concluding that the experimental drug is ineffective when in fact the mean growth rate of tumors is reduced.

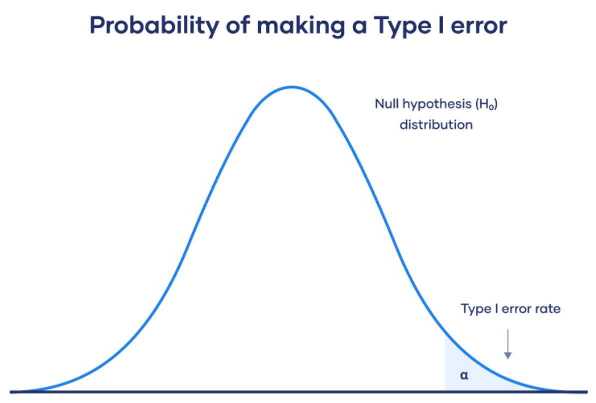
→Consequence : company might abandon development of a drug that was effective.

Test Errors and Error Probabilities

Correspondingly, we have

$\boldsymbol{\alpha}$ also called significance level

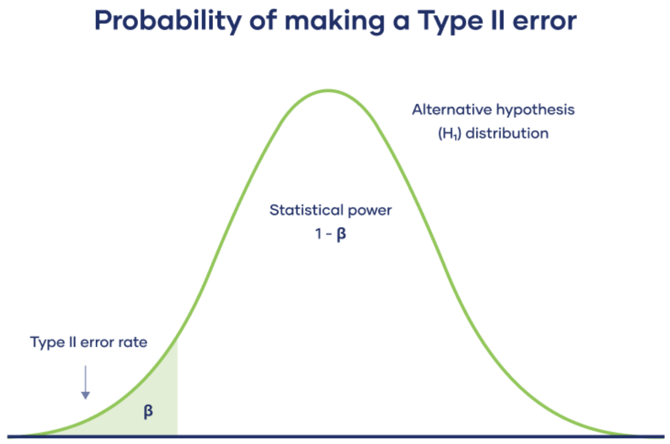
$$\boxed{\alpha = P(Type\ I\ error) = P(reject\ H_0\ if\ H_0\ is\ true)}$$



It is the probability of making a wrong decision to reject H_0.

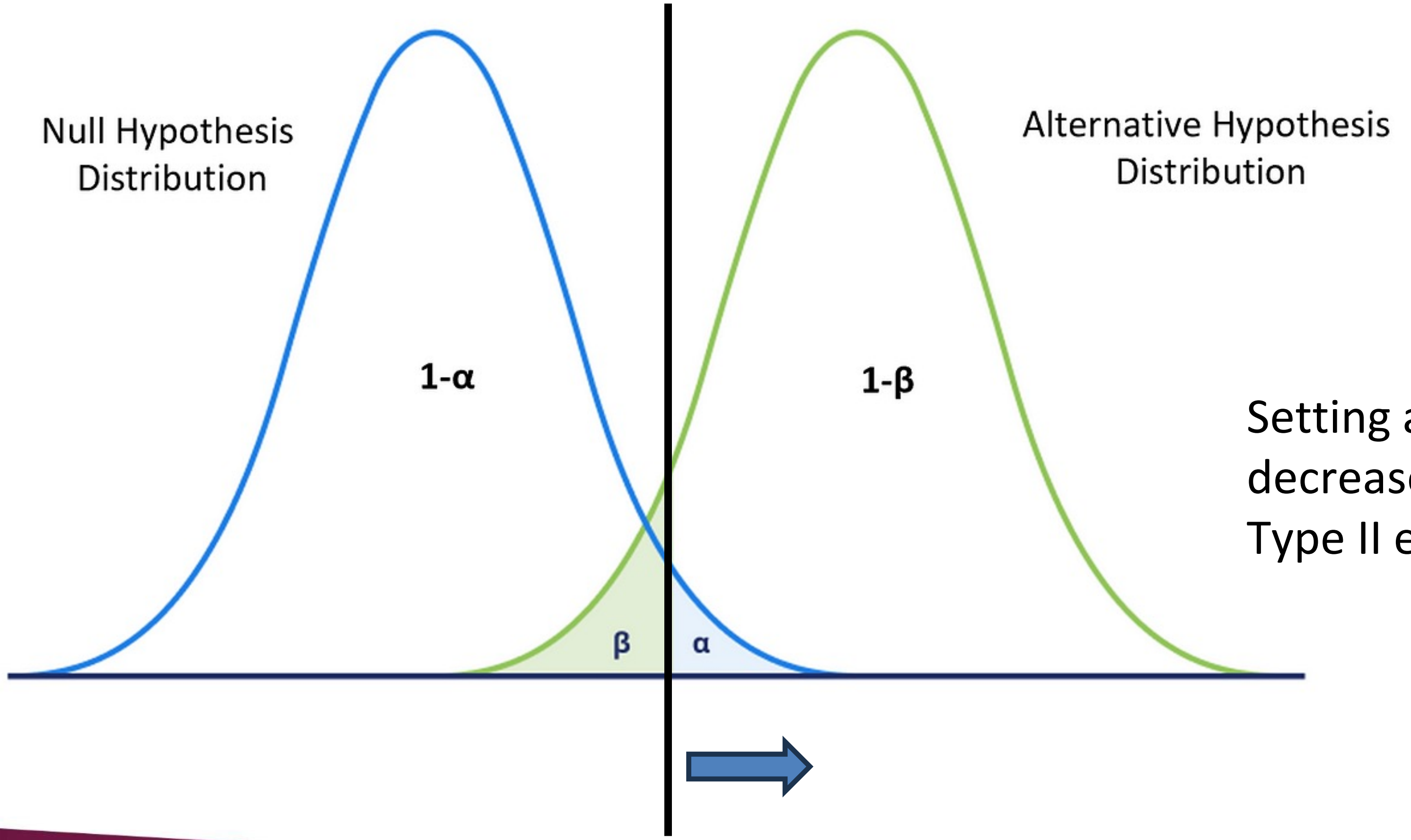
$$\boxed{\beta = P(Type\ II\ error) = P(Not\ reject\ H_0\ if\ H_0\ is\ false)}$$

It is the probability of making a wrong decision not to reject H_0.



Trade-off between Type I and Type II errors

The Type I and Type II error rates influence each other



The alpha level α (the significance level) represents the maximum probability of making a Type I error that the researcher is willing to accept.

Setting a lower significance level (α) decreases a Type I error risk; but increases a Type II error (β) risk. And vice versa

Determination of a Critical Value

So, in designing a test statement, we normally guarantee $\boldsymbol{\alpha}$ in a desired low value (often choose 0.01, 0.05 or 0.1), and then find a test statement with $\boldsymbol{\beta}$ as small as possible.

How to design a test statement with this restriction of α?

A Probability-Value Approach

Hypothesis test:

1. A well-organized, step-by-step procedure used to make a decision.

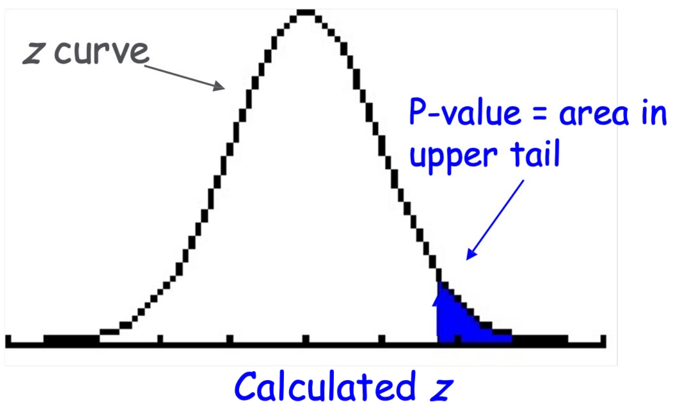
2. Probability-value approach (*p*-value approach)

What is P-value?

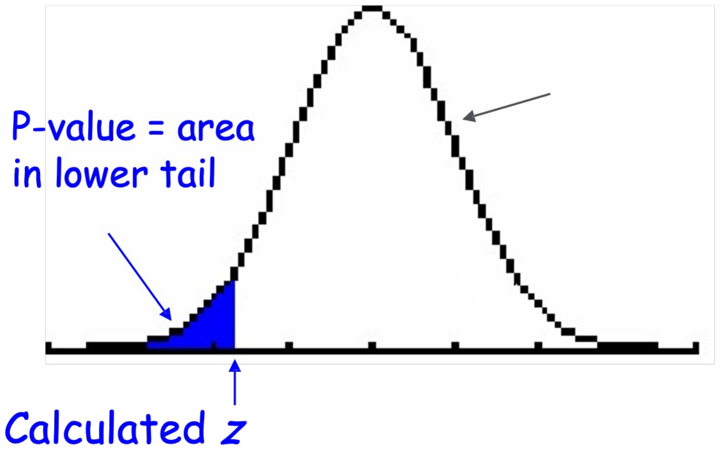
The ***P*-value** (also sometimes called the **observed significance level**) measures of the **strength of the evidence against the null hypothesis (H_0).** It's calculated from the observed data and represents the probability of obtaining results **at least as extreme as the observed results**, assuming that the null hypothesis is true.

The calculation of the P-value depends on the form of the inequality in the alternative hypothesis.

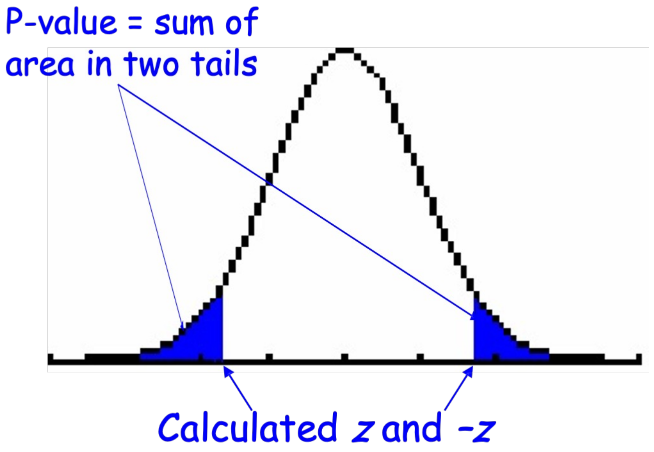
- H_1: *p > hypothesize value*



- H_1: *p < hypothesize value*



- H_a: *p ≠ hypothesize value*



The smaller the p-value, the stronger the evidence against H_0 provided by the data.

Decision-making after computing the P-value

A decision about whether to reject or to fail to reject H_0 results from comparing the P-value to the chosen α:

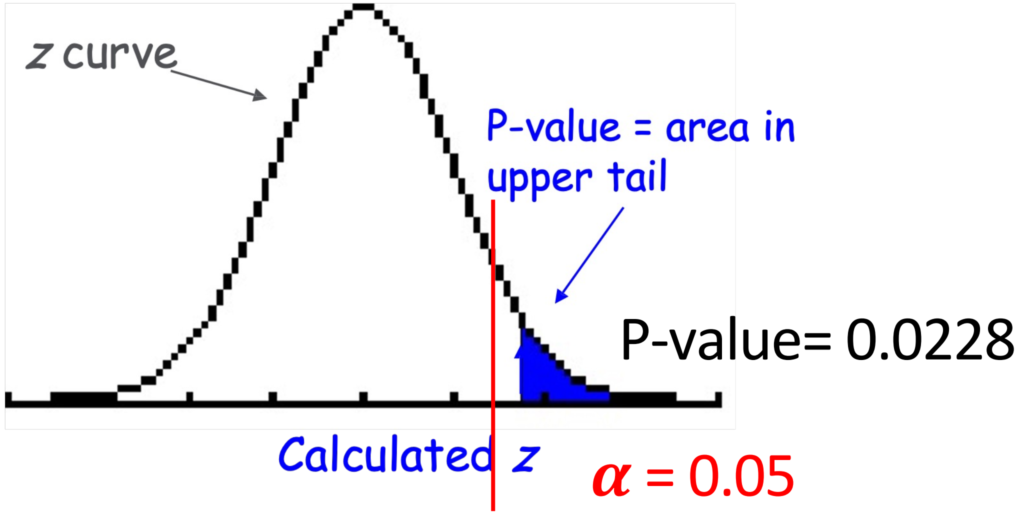
H_0 should be rejected if P-value $\leq \alpha$.
H_0 should not be rejected if P-value $> \alpha$.

Decision-making after computing the P-value

A decision about whether to reject or to fail to reject H_0 results from comparing the *P*-value to the chosen α:

H_0 should be rejected if *P*-value $\leq \alpha$. (data is more extreme than the threshold)
H_0 should not be rejected if *P*-value $> \alpha$.

The *P*-value measures of the strength of the evidence against the null hypothesis (H_0).



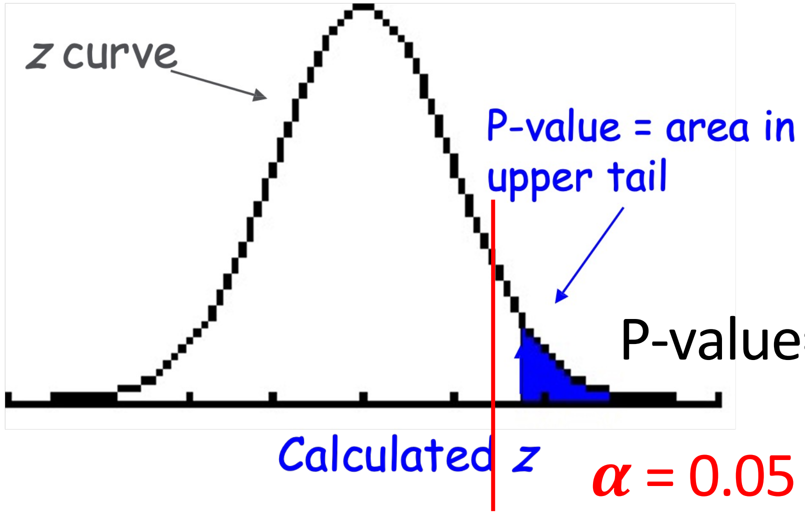
The alpha α (significance level) represents the maximum probability of making a Type I error that the researcher is willing to accept.

Decision-making after computing the P-value

A decision about whether to reject or to fail to reject H_0 results from comparing the *P*-value to the chosen α:

H_0 should be rejected if *P*-value $\leq \alpha$.
H_0 should not be rejected if *P*-value $> \alpha$.

The *P*-value measures of the strength of the evidence against the null hypothesis (H_0).



P-value= 0.0228

α = 0.05

α = 0.05
P-value= 0.0228

P-value < α

Reject the null hypothesis (H_0)

The alpha α (significance level) represents the maximum probability of making a Type I error that the researcher is willing to accept.

Hypothesis Test of Mean μ (σ Known):

A Probability-Value Approach

The Probability-Value Hypothesis Test: A Five-Step Procedure:

1. The Set-Up:

   a. Describe the population parameter of concern.

   b. State the null hypothesis (H_0) and the alternative hypothesis (H_1).

Hypothesis Test of Mean μ (σ Known):
A Probability-Value Approach

The Probability-Value Hypothesis Test: A Five-Step Procedure:

1. The Set-Up:

   a. Describe the population parameter of concern.

   b. State the null hypothesis (H_0) and the alternative hypothesis (H_1).

2. The Hypothesis Test Criteria:

   a. Check the assumptions.

   b. Identify the probability distribution and the test statistic formula to be used.

   c. Determine the level of significance, α.

Hypothesis Test of Mean μ (σ Known):

A Probability-Value Approach

The Probability-Value Hypothesis Test: A Five-Step Procedure:

1. The Set-Up:
   a. Describe the population parameter of concern.
   b. State the null hypothesis (H_0) and the alternative hypothesis (H_1).
2. The Hypothesis Test Criteria:
   a. Check the assumptions.
   b. Identify the probability distribution and the test statistic formula to be used.
   c. Determine the level of significance, α.
3. The Sample Evidence:
   a. Collect the sample information.
   b. Calculate the value of the test statistic.

Hypothesis Test of Mean μ (σ Known):

A Probability-Value Approach

The Probability-Value Hypothesis Test: A Five-Step Procedure:

1. The Set-Up:

   a. Describe the population parameter of concern.

   b. State the null hypothesis (H_0) and the alternative hypothesis (H_1).

2. The Hypothesis Test Criteria:

   a. Check the assumptions.

   b. Identify the probability distribution and the test statistic formula to be used.

   c. Determine the level of significance, α.

3. The Sample Evidence:

   a. Collect the sample information.

   b. Calculate the value of the test statistic.

4. The Probability Distribution:

   a. Calculate the p-value for the test statistic.

   b. Determine whether or not the p-value is smaller than α.

CityU

Hypothesis Test of Mean μ (σ Known):

A Probability-Value Approach

The Probability-Value Hypothesis Test: A Five-Step Procedure:

1. The Set-Up:
   a. Describe the population parameter of concern.
   b. State the null hypothesis (H_0) and the alternative hypothesis (H_1).
2. The Hypothesis Test Criteria:
   a. Check the assumptions.
   b. Identify the probability distribution and the test statistic formula to be used.
   c. Determine the level of significance, α.
3. The Sample Evidence:
   a. Collect the sample information.
   b. Calculate the value of the test statistic.
4. The Probability Distribution:
   a. Calculate the *p*-value for the test statistic.
   b. Determine whether or not the *p*-value is smaller than α.
5. The Results:
   a. State the decision about H_0.
   b. State a conclusion about H_1.

Example

A company advertises the net weight of its cereal is 24 ounces. A consumer group would like to check this claim. They cannot check every box of cereal, so a sample of cereal boxes will be examined. A decision will be made about the true mean weight based on the sample mean. State the consumer group's null and alternative hypotheses. A random sample of 40 boxes showed a sample mean 23.95 ounces. Is there any evidence to suggest that the net weight is less than 24 ounces at $\alpha = 0.05$? Assume $\sigma = 0.2$.

A company advertises the net weight of its cereal is 24 ounces. A consumer group would like to check this claim. They cannot check every box of cereal, so a sample of cereal boxes will be examined. A decision will be made about the true mean weight based on the sample mean. State the consumer group's null and alternative hypotheses. A random sample of 40 boxes showed a sample mean 23.95 ounces. Is there any evidence to suggest that the net weight is less than 24 ounces at $\alpha = 0.05$? Assume $\sigma = 0.2$.

Solution:

1. The Set-Up:

a) Describe the population parameter of concern.

The population parameter of interest is the mean μ, the mean weight of the cereal boxes.

b) State the null hypothesis (H_0) and the alternative hypothesis (H_1).

Formulate two opposing statements concerning μ.

H_0: $\mu = 24$ ($\geq$) (the mean is at least 24)

H_1: $\mu < 24$ (the mean is less than 24)

A company advertises the net weight of its cereal is 24 ounces. A consumer group would like to check this claim. They cannot check every box of cereal, so a sample of cereal boxes will be examined. A decision will be made about the true mean weight based on the sample mean. State the consumer group's null and alternative hypotheses. A random sample of 40 boxes showed a sample mean 23.95 ounces. Is there any evidence to suggest that the net weight is less than 24 ounces at $\alpha = 0.05$? Assume $\sigma = 0.2$.

Solution:

1. The Set-Up:

a) Describe the population parameter of concern.

The population parameter of interest is the mean μ, the mean weight of the cereal boxes.

b) State the null hypothesis (H_0) and the alternative hypothesis (H_1).

Formulate two opposing statements concerning μ.

H_0: $\mu = 24$ ($\geq$) (the mean is at least 24)

H_1: $\mu < 24$ (the mean is less than 24)

2. The Hypothesis Test Criteria:

a) Check the assumptions.

A sample size of 40 should be sufficient for the CLT to apply. The sampling distribution of the sample mean can be expected to be normal.

A company advertises the net weight of its cereal is 24 ounces. A consumer group would like to check this claim. They cannot check every box of cereal, so a sample of cereal boxes will be examined. A decision will be made about the true mean weight based on the sample mean. State the consumer group's null and alternative hypotheses. A random sample of 40 boxes showed a sample mean 23.95 ounces. Is there any evidence to suggest that the net weight is less than 24 ounces at $\alpha = 0.05$? Assume $\sigma = 0.2$.

Solution:

1. The Set-Up:

a) Describe the population parameter of concern.

The population parameter of interest is the mean μ, the mean weight of the cereal boxes.

b) State the null hypothesis (H_0) and the alternative hypothesis (H_1).

Formulate two opposing statements concerning μ.

H_0: $\mu = 24$ ($\geq$) (the mean is at least 24)

H_1: $\mu < 24$ (the mean is less than 24)

2. The Hypothesis Test Criteria:

a) Check the assumptions.

A sample size of 40 should be sufficient for the CLT to apply. The sampling distribution of the sample mean can be expected to be normal.

b) Identify the probability distribution and the test statistic to be used. To test the null hypothesis, ask how many standard deviations away from μ is the sample mean.

$$\text{test statistic: } z^* = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

c) Determine the level of significance.

Consider the four possible outcomes and their consequences. Let $\alpha = 0.05$.

A company advertises the net weight of its cereal is 24 ounces. A consumer group would like to check this claim. They cannot check every box of cereal, so a sample of cereal boxes will be examined. A decision will be made about the true mean weight based on the sample mean. State the consumer group's null and alternative hypotheses. A random sample of 40 boxes showed a sample mean 23.95 ounces. Is there any evidence to suggest that the net weight is less than 24 ounces at $\alpha = 0.05$? Assume $\sigma = 0.2$.

Solution:

1. The Set-Up:

a) Describe the population parameter of concern.

The population parameter of interest is the mean μ, the mean weight of the cereal boxes.

b) State the null hypothesis (H_0) and the alternative hypothesis (H_1).

Formulate two opposing statements concerning μ.

H_0: $\mu = 24$ ($\geq$) (the mean is at least 24)

H_1: $\mu < 24$ (the mean is less than 24)

2. The Hypothesis Test Criteria:

a) Check the assumptions.

A sample size of 40 should be sufficient for the CLT to apply. The sampling distribution of the sample mean can be expected to be normal.

b) Identify the probability distribution and the test statistic to be used. To test the null hypothesis, ask how many standard deviations away from μ is the sample mean.

$$\text{test statistic}: \quad z^* = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

c) Determine the level of significance.

Consider the four possible outcomes and their consequences. Let $\alpha = 0.05$.

3. The Sample Evidence:

a) Collect the sample information.

A random sample of 40 cereal boxes is examined.

$$\bar{x} = 23.95 \quad \text{and} \quad n = 40$$

b) Calculate the value of the test statistic. ($\sigma = 0.2$)

$$z^* = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{23.95 - 24}{.2 / \sqrt{40}} = -1.5811$$

A company advertises the net weight of its cereal is 24 ounces. A consumer group would like to check this claim. They cannot check every box of cereal, so a sample of cereal boxes will be examined. A decision will be made about the true mean weight based on the sample mean. State the consumer group's null and alternative hypotheses. A random sample of 40 boxes showed a sample mean 23.95 ounces. Is there any evidence to suggest that the net weight is less than 24 ounces at α = 0.05? Assume σ = 0.2.

Solution:

1. The Set-Up:

a) Describe the population parameter of concern.

The population parameter of interest is the mean μ, the mean weight of the cereal boxes.

b) State the null hypothesis (H_0) and the alternative hypothesis (H_1).

Formulate two opposing statements concerning μ.

H_0: $\mu = 24$ ($\geq$) (the mean is at least 24)

H_1: $\mu < 24$ (the mean is less than 24)

2. The Hypothesis Test Criteria:

a) Check the assumptions.

A sample size of 40 should be sufficient for the CLT to apply. The sampling distribution of the sample mean can be expected to be normal.

b) Identify the probability distribution and the test statistic to be used. To test the null hypothesis, ask how many standard deviations away from μ is the sample mean.

$$\text{test statistic}: \ z^* = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$$

c) Determine the level of significance.

Consider the four possible outcomes and their consequences. Let α = 0.05.

3. The Sample Evidence:

a) Collect the sample information.

A random sample of 40 cereal boxes is examined.

$$\bar{x} = 23.95 \quad \text{and} \quad n = 40$$

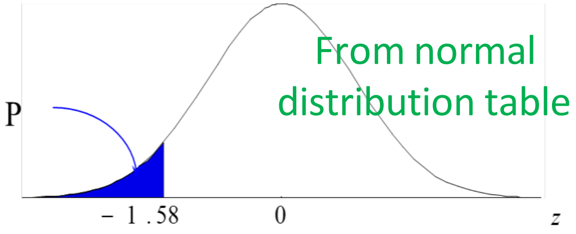
b) Calculate the value of the test statistic. (σ = 0.2)

$$z^* = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = \frac{23.95-24}{.2/\sqrt{40}} = -1.5811$$

4. The Probability Distribution:

a) Calculate the p-value for the test statistic.

b) Determine whether or not the p-value is smaller than α. The p-value (0.0571) is greater than α (0.05).



$$P = P(z < z^*) = P(z < -1.58) = P(z > 1.58)$$
$$= 0.0571$$

A company advertises the net weight of its cereal is 24 ounces. A consumer group would like to check this claim. They cannot check every box of cereal, so a sample of cereal boxes will be examined. A decision will be made about the true mean weight based on the sample mean. State the consumer group's null and alternative hypotheses. A random sample of 40 boxes showed a sample mean 23.95 ounces. Is there any evidence to suggest that the net weight is less than 24 ounces at $\alpha = 0.05$? Assume $\sigma = 0.2$.

Solution:

1. The Set-Up:

a) Describe the population parameter of concern.

The population parameter of interest is the mean μ, the mean weight of the cereal boxes.

b) State the null hypothesis (H_0) and the alternative hypothesis (H_1).

Formulate two opposing statements concerning μ.

H_0: $\mu = 24$ ($\geq$) (the mean is at least 24)

H_1: $\mu < 24$ (the mean is less than 24)

2. The Hypothesis Test Criteria:

a) Check the assumptions.

A sample size of 40 should be sufficient for the CLT to apply. The sampling distribution of the sample mean can be expected to be normal.

b) Identify the probability distribution and the test statistic to be used. To test the null hypothesis, ask how many standard deviations away from μ is the sample mean.

$$\text{test statistic}: \quad z^* = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

c) Determine the level of significance.

Consider the four possible outcomes and their consequences. Let $\alpha = 0.05$.

3. The Sample Evidence:

a) Collect the sample information.

A random sample of 40 cereal boxes is examined.

$$\bar{x} = 23.95 \quad \text{and} \quad n = 40$$

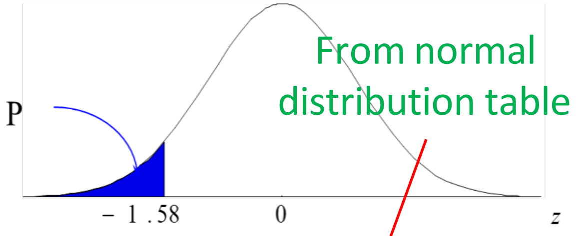
b) Calculate the value of the test statistic. ($\sigma = 0.2$)

$$z^* = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{23.95 - 24}{.2 / \sqrt{40}} = -1.5811$$

4. The Probability Distribution:

a) Calculate the p-value for the test statistic.

b) Determine whether or not the p-value is smaller than α. The p-value (0.0571) is greater than α (0.05).



$$P = P(z < z^*) = P(z < -1.58) = P(z > 1.58)$$
$$= 0.0571$$

5. **Results → Decision Rule**:

a) If the p-value is **less than or equal** to the level of significance α, then the decision must be to **reject H_0**.

b) If the p-value is **greater than** the level of significance α, then the decision must be to **fail to reject H_0**.

6. Decision:

a) State the decision about H_0.

Decision about H_0 : Fail to reject H_0.

b) Write a conclusion about H_1.

There is not sufficient evidence at the 0.05 level of significance to show that the mean weight of cereal boxes is less than 24 ounces.

Note:

1. If we fail to reject H_0, there is no evidence to suggest the null hypothesis is false. This does not mean H_0 is true.
2. The *p*-value is the area, under the curve of the probability distribution for the test statistic, that is more extreme than the calculated value of the test statistic.

6. Decision:

a) State the decision about H_0.

Decision about H_0 : Fail to reject H_0.

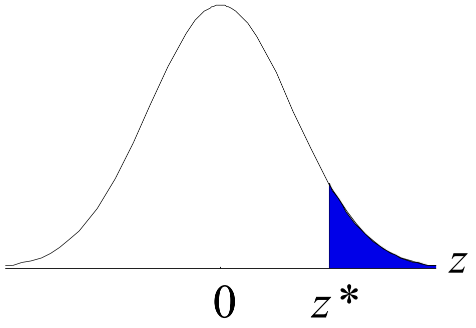
b) Write a conclusion about H_1.

There is not sufficient evidence at the 0.05 level of significance to show that the mean weight of cereal boxes is less than 24 ounces.

Finding *p*-values:

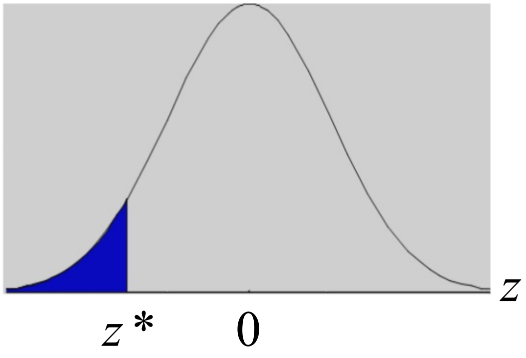
1. H_1 contains $>$ (Right tail)

p-value = $P(z > z^*)$



2. H_1 contains $<$ (Left tail)

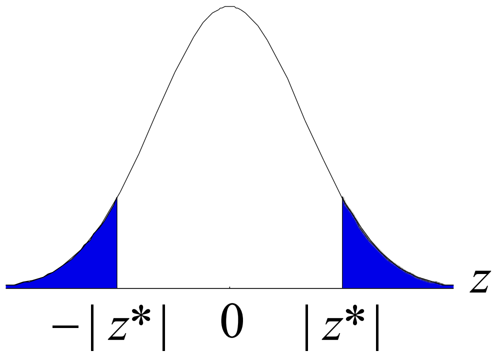
p-value = $P(z < z^*)$



3. H_1 contains (Two-tailed)

p-value = $P(z < -|z^*|) + P(z > |z^*|)$

$= 2 \times P(z > |z^*|)$



Note:

1. If we fail to reject H_0, there is no evidence to suggest the null hypothesis is false. This does not mean H_0 is true.
2. The p-value is the area, under the curve of the probability distribution for the test statistic, that is more extreme than the calculated value of the test statistic.
3. There are 3 separate cases for p-values. The direction (or sign) of the alternative hypothesis (H_1) is the key.

Another Approach for the Hypothesis Test of Mean μ (σ Known)

Hypothesis Test of Mean μ (σ Known):

A Classical Approach

The sampling distribution of $\bar{x}$ has a normal distribution.

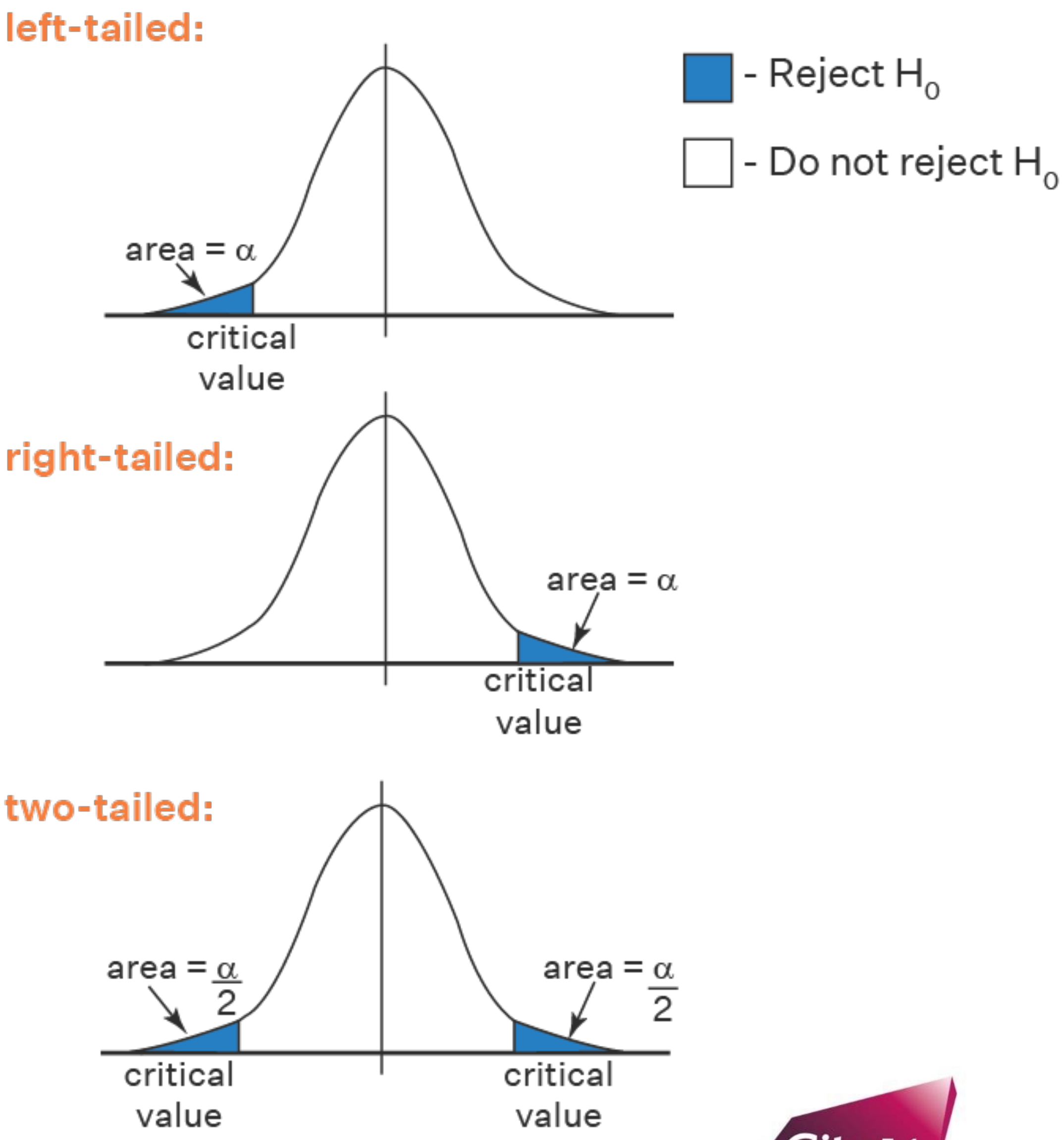
Hypothesis test:

1. A well-organized, step-by-step procedure used to make a decision.
2. The **classical approach** is the hypothesis test process that has enjoyed popularity for many years.
   - Determine the critical region(s) and critical value(s).
   - Determine the critical region(s) and critical value(s) calculated test statistic is in the critical region.

Z critical value

The z critical value can be calculated as follows

- Find the alpha level.
- Subtract the alpha level from 1 for a two-tailed test. For a one-tailed test subtract the alpha level from 0.5 (or 1- $\alpha/2$).
- Look up the **area from the z distribution** table to obtain the z critical value.
- For a left-tailed test, a negative sign needs to be added to the critical value at the end of the calculation.



Test statistic $$z^* = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

- The sampling distribution of $\bar{x}$ has a normal distribution.
- σ Known

Decision Criteria

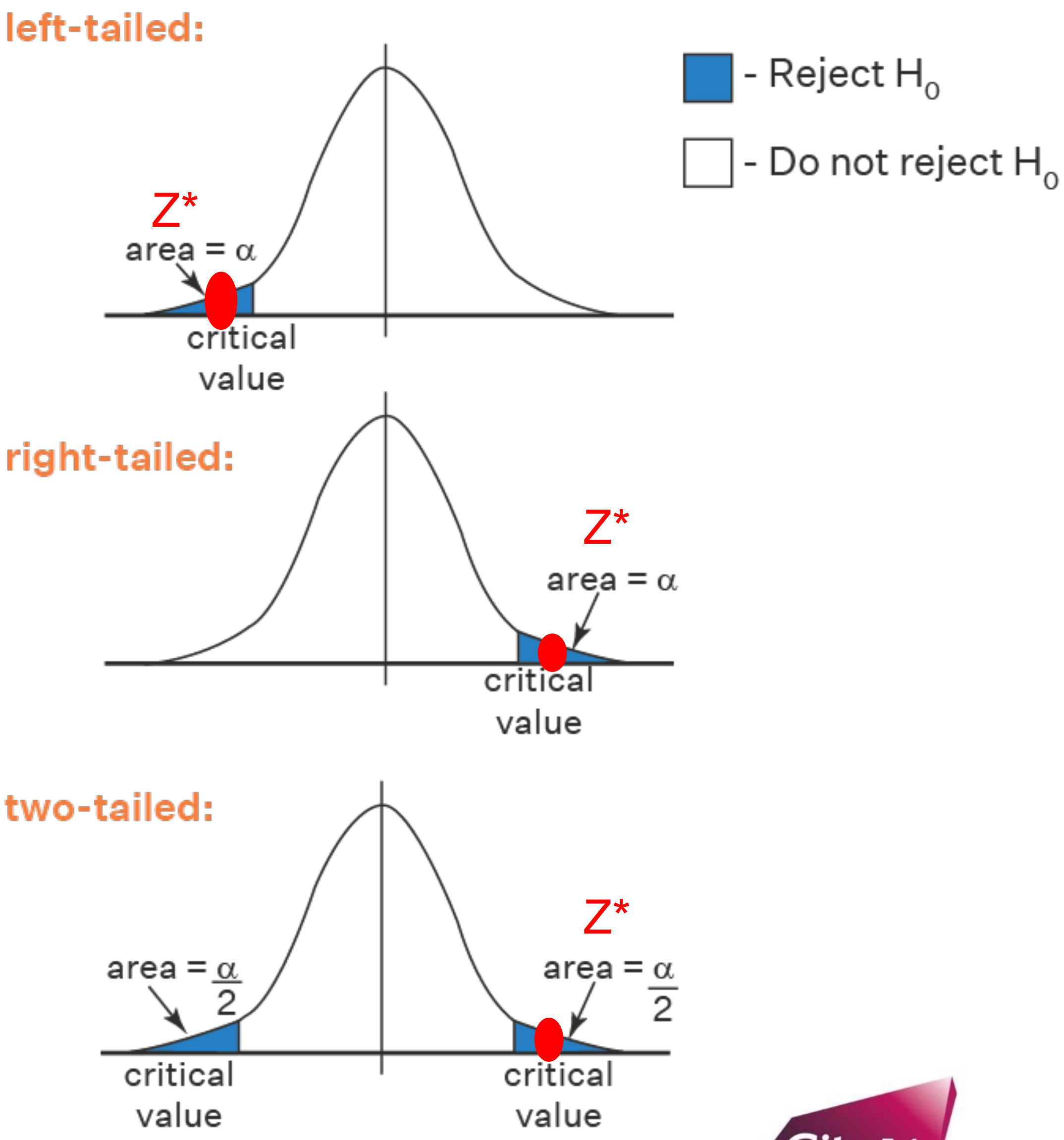
Test statistic $$z^* = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

•**Reject** the null hypothesis if test statistic < Z critical value (left-tailed hypothesis test) (data outside the acceptable region or in the rejection region).

(data is more extreme than the threshold)

•**Reject** the null hypothesis if test statistic > Z critical value (right-tailed hypothesis test) (data outside the acceptable region or in the rejection region).

•**Reject** the null hypothesis if the test statistic does not lie in the acceptance region/ in the rejection region (two-tailed hypothesis test).



Hypothesis Test of Mean μ (σ Known):

The Classical Hypothesis Test: A Five-Step Procedure:

1. The Set-Up:
   a. Describe the population parameter of concern.
   b. State the null hypothesis (H_0) and the alternative hypothesis (H_1).
2. The Hypothesis Test Criteria:
   a. Check the assumptions.
   b. Identify the probability distribution and the test statistic to be used.
   c. Determine the level of significance, α.
3. The Sample Evidence:
   a. Collect the sample information.
   b. Calculate the value of the test statistic.
4. The Probability Distribution:
   a. Determine the critical region(s) and critical value(s).
   b. Determine whether or not the calculated test statistic is in the critical region.
5. The Results:
   a. State the decision about H_0.
   b. State the conclusion about H_1.

Example

An elementary school principal claims students receive no more than 30 minutes of homework each night. A random sample of 36 students showed a sample mean of 36.8 minutes spent doing homework (assume $\sigma = 7.5$). Is there any evidence to suggest the mean time spent on homework is greater than 30 minutes? Use $\alpha = 0.01$.

Example

Right tailed example

An elementary school principal claims students receive no more than 30 minutes of homework each night. A random sample of 36 students showed a sample mean of 36.8 minutes spent doing homework (assume $\sigma = 7.5$). Is there any evidence to suggest the mean time spent on homework is greater than 30 minutes? Use $\alpha = 0.01$.

Solution:

1. The parameter of concern: μ, the mean time spent doing homework each night.

   H_0: $\mu = 30$

   H_1: $\mu > 30$

2. The Hypothesis Test Criteria:

a) The sample size is $n = 36$, the CLT applies.

b) The test statistic is z^*.

c) The level of significance is given: $\alpha = 0.01$.

3. The Sample Evidence: $\bar{x} = 36.8, \quad n = 36$

4. The Probability Distribution:

$$z^* = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{36.8 - 30}{7.5 / \sqrt{36}} = 5.44$$

Example

Right tailed example

An elementary school principal claims students receive no more than 30 minutes of homework each night. A random sample of 36 students showed a sample mean of 36.8 minutes spent doing homework (assume $\sigma = 7.5$). Is there any evidence to suggest the mean time spent on homework is greater than 30 minutes? Use $\alpha = 0.01$.

Solution:

1. The parameter of concern: μ, the mean time spent doing homework each night

   H_0: $\mu = 30$

   H_1: $\mu > 30$

STANDARD NORMAL DISTRIBUTION: Table Values Represent AREA to the LEFT of the Z score.

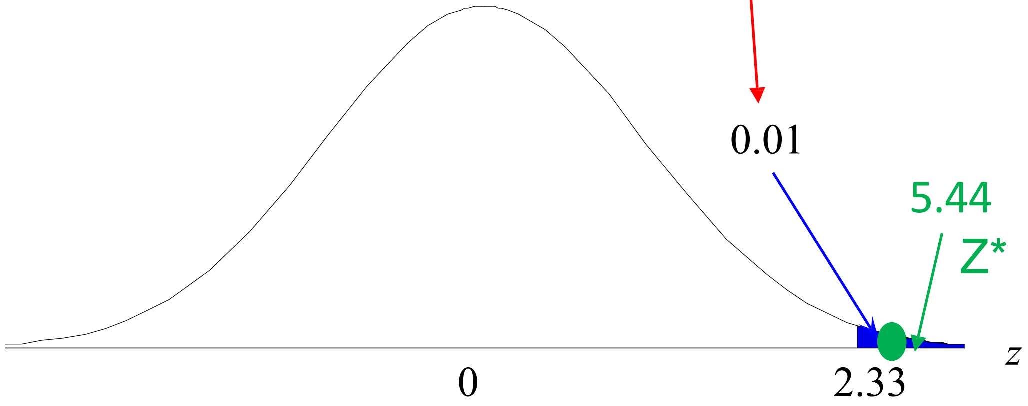
Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.50000	.50399	.50798	.51197	.51595	.51994	.52392	.52790	.53188	.53586
0.1	.53983	.54380	.54776	.55172	.55567	.55962	.56356	.56749	.57142	.57535
0.2	.57926	.58317	.58706	.59095	.59483	.59871	.60257	.60642	.61026	.61409
0.3	.61791	.62172	.62552	.62930	.63307	.63683	.64058	.64431	.64803	.65173
0.4	.65542	.65910	.66276	.66640	.67003	.67364	.67724	.68082	.68439	.68793
0.5	.69146	.69497	.69847	.70194	.70540	.70884	.71226	.71566	.71904	.72240
0.6	.72575	.72907	.73237	.73565	.73891	.74215	.74537	.74857	.75175	.75490
0.7	.75804	.76115	.76424	.76730	.77035	.77337	.77637	.77935	.78230	.78524
0.8	.78814	.79103	.79389	.79673	.79955	.80234	.80511	.80785	.81057	.81327
0.9	.81594	.81859	.82121	.82381	.82639	.82894	.83147	.83398	.83646	.83891
1.0	.84134	.84375	.84614	.84849	.85083	.85314	.85543	.85769	.85993	.86214
1.1	.86433	.86650	.86864	.87076	.87286	.87493	.87698	.87900	.88100	.88298
1.2	.88493	.88686	.88877	.89065	.89251	.89435	.89617	.89796	.89973	.90147
1.3	.90320	.90490	.90658	.90824	.90988	.91149	.91309	.91466	.91621	.91774
1.4	.91924	.92073	.92220	.92364	.92507	.92647	.92785	.92922	.93056	.93189
1.5	.93319	.93448	.93574	.93699	.93822	.93943	.94062	.94179	.94295	.94408
1.6	.94520	.94630	.94738	.94845	.94950	.95053	.95154	.95254	.95352	.95449
1.7	.95543	.95637	.95728	.95818	.95907	.95994	.96080	.96164	.96246	.96327
1.8	.96407	.96485	.96562	.96638	.96712	.96784	.96856	.96926	.96995	.97062
1.9	.97128	.97193	.97257	.97320	.97381	.97441	.97500	.97558	.97615	.97670
2.0	.97725	.97778	.97831	.97882	.97932	.97982	.98030	.98077	.98124	.98169
2.1	.98214	.98257	.98300	.98341	.98382	.98422	.98461	.98500	.98537	.98574
2.2	.98610	.98645	.98679	.98713	.98745	.98778	.98809	.98840	.98870	.98899
2.3	.98928	.98956	.98983	.99010	.99036	.99061	.99086	.99111	.99134	.99158
2.4	.99180	.99202	.99224	.99245	.99266	.99286	.99305	.99324	.99343	.99361

$\alpha = 0.01$

Z = 2.33

2. The Hypothesis Test Criteria:
   a) The sample size is $n = 36$, the CLT applies.
   b) The test statistic is z^*.
   c) The level of significance is given: $\alpha = 0.01$.
3. The Sample Evidence: $\bar{x} = 36.8, \quad n = 36$
4. The Probability Distribution:

$$z^* = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{36.8 - 30}{7.5 / \sqrt{36}} = 5.44$$



The calculated value of z* = 5.44 >Z, is in the *rejection region (outside the acceptable region)*

Example

Right tailed example

An elementary school principal claims students receive no more than 30 minutes of homework each night. A random sample of 36 students showed a sample mean of 36.8 minutes spent doing homework (assume $\sigma = 7.5$). Is there any evidence to suggest the mean time spent on homework is greater than 30 minutes? Use $\alpha = 0.01$.

Solution:

1. The parameter of concern: μ, the mean time spent doing homework each night.

   H_0: $\mu = 30$

   H_1: $\mu > 30$

2. The Hypothesis Test Criteria:

a) The sample size is $n = 36$, the CLT applies.

b) The test statistic is z^*.

c) The level of significance is given: $\alpha = 0.01$.

3. The Sample Evidence: $\bar{x} = 36.8, \quad n = 36$

4. The Probability Distribution:

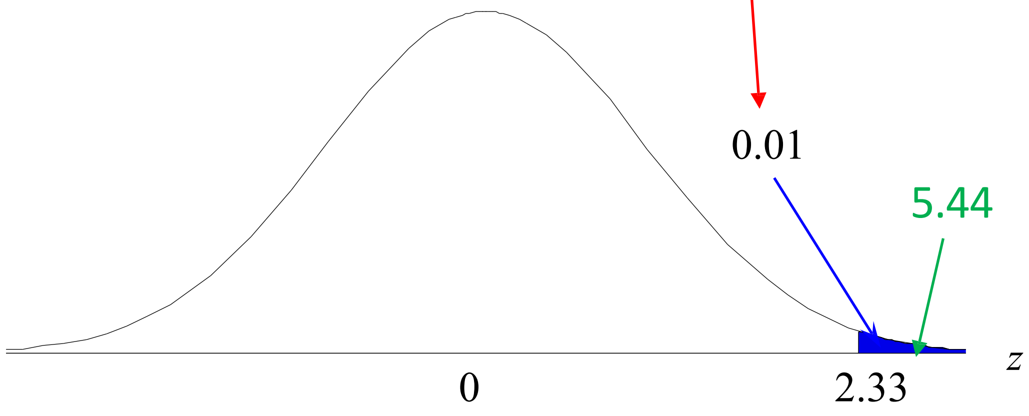
$$z^* = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{36.8 - 30}{7.5 / \sqrt{36}} = 5.44$$

The calculated value of z* = 5.44 >Z, is in the *rejection region (outside the acceptable region)*

5. The Results:

Decision: *Reject H_0*.

Conclusion: There is sufficient evidence at the 0.01 level of significance to conclude the mean time spent on homework by the elementary students is more than 30 minutes.



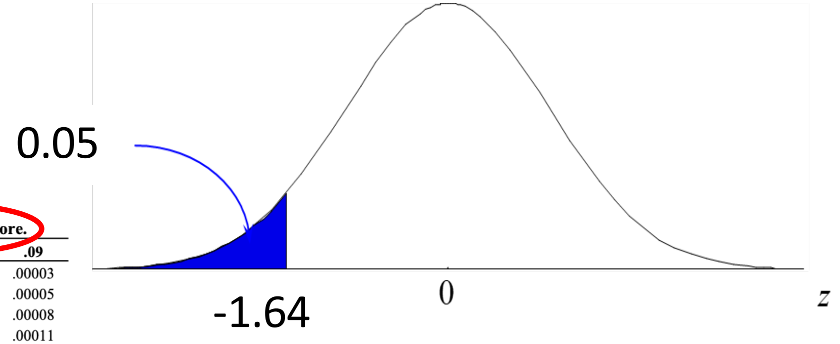
Example

Left tailed example

All cigarettes on the market have an average nicotine content is at least 1.6 mg ($\geq$1.6). A firm claims that they discovered a new formula, that will result the average nicotine content of a cigarette being less than 1.6 (< 1.6). Suppose the standard deviation of nicotine content is known to be 0.8. What conclusions can be drawn, at the 5% significance level, if the average nicotine of a sample of 20 cigarettes from this firm is 1.54?

We are testing

$$H_0: \mu = 1.6; \quad H_1: \mu < 1.6$$



STANDARD NORMAL DISTRIBUTION: Table Values Represent AREA to the LEFT of the Z score.

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.9	.00005	.00005	.00004	.00004	.00004	.00004	.00004	.00004	.00003	.00003
-3.8	.00007	.00007	.00007	.00006	.00006	.00006	.00006	.00005	.00005	.00005
-3.7	.00011	.00010	.00010	.00010	.00009	.00009	.00008	.00008	.00008	.00008
-3.6	.00016	.00015	.00015	.00014	.00014	.00013	.00013	.00012	.00012	.00011
-3.5	.00023	.00022	.00022	.00021	.00020	.00019	.00019	.00018	.00017	.00017
-3.4	.00034	.00032	.00031	.00030	.00029	.00028	.00027	.00026	.00025	.00024
-3.3	.00048	.00047	.00045	.00043	.00042	.00040	.00039	.00038	.00036	.00035
-3.2	.00069	.00066	.00064	.00062	.00060	.00058	.00056	.00054	.00052	.00050
-3.1	.00097	.00094	.00090	.00087	.00084	.00082	.00079	.00076	.00074	.00071
-3.0	.00135	.00131	.00126	.00122	.00118	.00114	.00111	.00107	.00104	.00100
-2.9	.00187	.00181	.00175	.00169	.00164	.00159	.00154	.00149	.00144	.00139
-2.8	.00256	.00248	.00240	.00233	.00226	.00219	.00212	.00205	.00199	.00193
-2.7	.00347	.00336	.00326	.00317	.00307	.00298	.00289	.00280	.00272	.00264
-2.6	.00466	.00453	.00440	.00427	.00415	.00402	.00391	.00379	.00368	.00357
-2.5	.00621	.00604	.00587	.00570	.00554	.00539	.00523	.00508	.00494	.00480
-2.4	.00820	.00798	.00776	.00755	.00734	.00714	.00695	.00676	.00657	.00639
-2.3	.01072	.01044	.01017	.00990	.00964	.00939	.00914	.00889	.00866	.00842
-2.2	.01390	.01355	.01321	.01287	.01255	.01222	.01191	.01160	.01130	.01101
-2.1	.01786	.01743	.01700	.01659	.01618	.01578	.01539	.01500	.01463	.01426
-2.0	.02275	.02222	.02169	.02118	.02068	.02018	.01970	.01923	.01876	.01831
-1.9	.02872	.02807	.02743	.02680	.02619	.02559	.02500	.02442	.02385	.02330
-1.8	.03593	.03515	.03438	.03362	.03288	.03216	.03144	.03074	.03005	.02938
-1.7	.04457	.04363	.04272	.04182	.04093	.04006	.03920	.03836	.03754	.03673
-1.6	.05480	.05370	.05262	.05155	.05050	.04947	.04846	.04746	.04648	.04551
-1.5	.06681	.06552	.06426	.06301	.06178	.06057	.05938	.05821	.05705	.05592

Example

Left tailed example

All cigarettes on the market have an average nicotine content is at least 1.6 mg (≥1.6). A firm claims that they discovered a new formula, that will result the average nicotine content of a cigarette being less than 1.6 (< 1.6). Suppose the standard deviation of nicotine content is known to be 0.8. What conclusions can be drawn, at the 5% significance level, if the average nicotine of a sample of 20 cigarettes from this firm is 1.54?

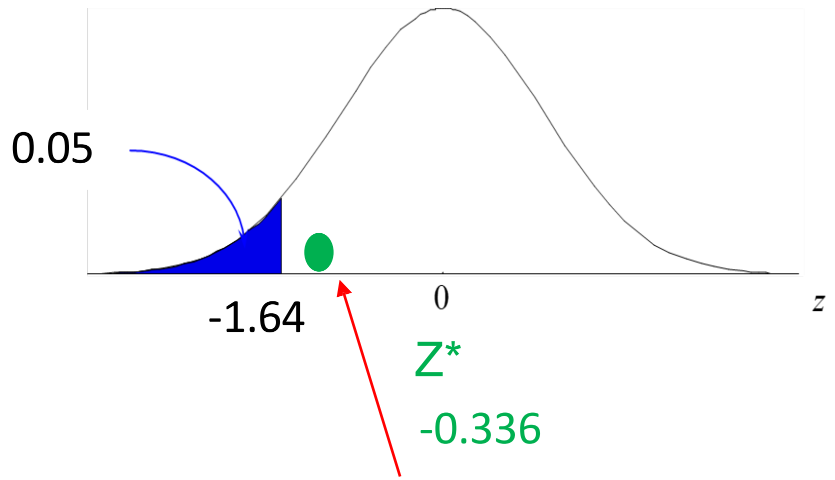
Solution: We are testing

$$H_0: \mu = 1.6; \qquad H_1: \mu < 1.6$$

From the sample data

$$z^* = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

$$z^* = \frac{1.54 - 1.6}{\frac{0.8}{\sqrt{20}}} = -0.336$$



But (Z) $-1.64 < -0.336$ (Z*).

Not in the rejection region, but in the acceptable region, So, fail to reject H_0.

Example

Two tailed example

Air crew escape systems are powered by a solid propellant. The burning rate of this propellant is an important product characteristic. Specifications require that the mean burning rate must be 50 centimeters per second. We know that the standard deviation of burning rate is σ_X = 2 centimeters per second. The experimenter decides to specify a type I error probability or significance level of α = 0.05 (95%) and selects a random sample of n = 25 and obtains a sample average burning rate of $\bar{x}$ of 51.3 centimeters per second. What conclusions should be drawn?

Example Two tailed example

Air crew escape systems are powered by a solid propellant. The burning rate of this propellant is an important product characteristic. Specifications require that the mean burning rate must be 50 centimeters per second. We know that the standard deviation of burning rate is $\sigma_X = 2$ centimeters per second. The experimenter decides to specify a type I error probability or significance level of $\alpha = 0.05$ (95%) and selects a random sample of $n = 25$ and obtains a sample average burning rate of $\bar{x}$ of 51.3 centimeters per second. What conclusions should be drawn?

Solution:

1. Parameter of interest: The parameter of interest is μ, the mean burning rate.
2. Null hypothesis: H_0 : $\mu_X = 50$ centimeters per second
3. Alternative hypothesis: H_1 : $\mu_X \neq 50$ centimeters per second

Example

Two tailed example

Air crew escape systems are powered by a solid propellant. The burning rate of this propellant is an important product characteristic. Specifications require that the mean burning rate must be 50 centimeters per second. We know that the standard deviation of burning rate is $\sigma_X = 2$ centimeters per second. The experimenter decides to specify a type I error probability or significance level of α = 0.05 (95%) and selects a random sample of $n = 25$ and obtains a sample average burning rate of $\bar{x}$ of 51.3 centimeters per second. What conclusions should be drawn?

Solution:

1. Parameter of interest: The parameter of interest is μ, the mean burning rate.
2. Null hypothesis: $H_0 : \mu_X = 50$ centimeters per second
3. Alternative hypothesis: $H_1 : \mu_X \neq 50$ centimeters per second
4. Test statistic: The test statistic is

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma_X / \sqrt{n}}$$

Example

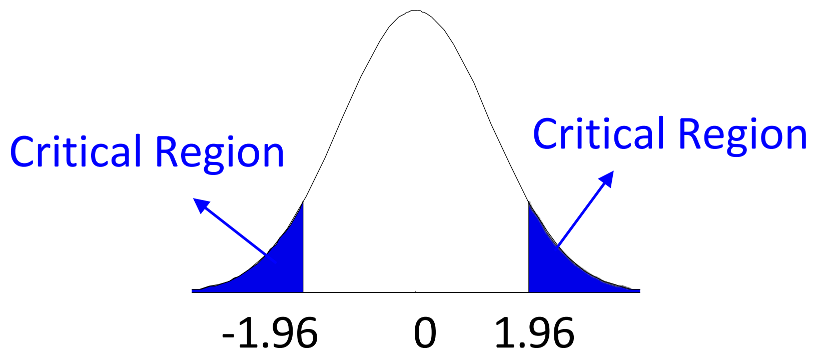
Two tailed example

Air crew escape systems are powered by a solid propellant. The burning rate of this propellant is an important product characteristic. Specifications require that the mean burning rate must be 50 centimeters per second. We know that the standard deviation of burning rate is σ_X = 2 centimeters per second. The experimenter decides to specify a type I error probability or significance level of α = 0.05 (95%) and selects a random sample of n = 25 and obtains a sample average burning rate of $\bar{x}$ of 51.3 centimeters per second. What conclusions should be drawn?

Solution:

1. Parameter of interest: The parameter of interest is μ, the mean burning rate.
2. Null hypothesis: H_0 : μ_X = 50 centimeters per second
3. Alternative hypothesis: H_1 : $\mu_X \neq$ 50 centimeters per second
4. Test statistic: The test statistic is

$$z^* = \frac{\bar{x} - \mu_0}{\sigma_X / \sqrt{n}}$$



Two tailed: (1- $\alpha/2) = 97.5\% \rightarrow$ $z_{0.025}$ = 1.96

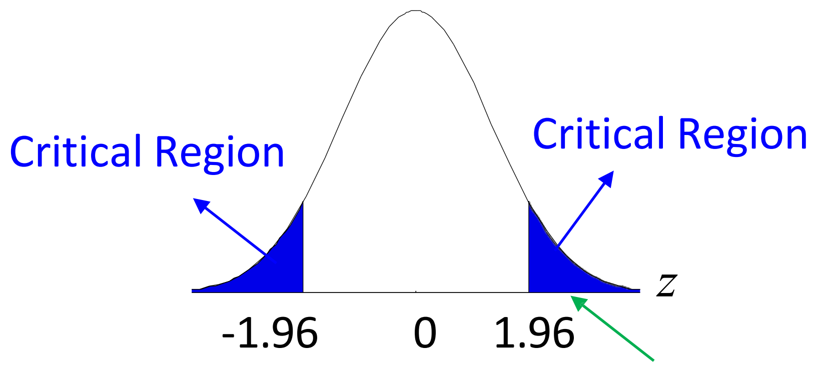
5. **To use a fixed significance level test, the boundaries of the rejection region would be $z_{0.025}$ = 1.96 and $z_{-0.025}$ = –1.96.**

Example

Air crew escape systems are powered by a solid propellant. The burning rate of this propellant is an important product characteristic. Specifications require that the mean burning rate must be 50 centimeters per second. We know that the standard deviation of burning rate is σ_X = 2 centimeters per second. The experimenter decides to specify a type I error probability or significance level of α = 0.05 (95%) and selects a random sample of n = 25 and obtains a sample average burning rate of $\bar{x}$ of 51.3 centimeters per second. What conclusions should be drawn?

Solution:

1. Parameter of interest: The parameter of interest is μ, the mean burning rate.
2. Null hypothesis: H_0 : μ_X = 50 centimeters per second
3. Alternative hypothesis: H_1 : $\mu_X \neq$ 50 centimeters per second
4. Test statistic: The test statistic is



$$z^* = \frac{\bar{x} - \mu_0}{\sigma_X / \sqrt{n}}$$

Two tailed: (1- $\alpha/2) = 97.5\%$ → $z_{0.025}$ = **1.96**

5. **To use a fixed significance level test, the boundaries of the critical region would be $z_{0.025}$ = 1.96 and $z_{-0.025}$ = –1.96.**
6. Computations: Because $\bar{x}$ =51.3 and σ_X = 2,

$$z^* = \frac{51.3 - 50}{2/\sqrt{25}} = 3.25$$

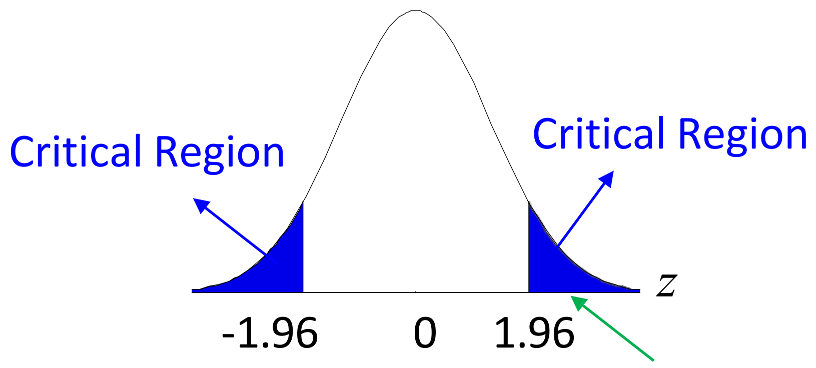
Example

Air crew escape systems are powered by a solid propellant. The burning rate of this propellant is an important product characteristic. Specifications require that the mean burning rate must be 50 centimeters per second. We know that the standard deviation of burning rate is $\sigma_X = 2$ centimeters per second. The experimenter decides to specify a type I error probability or significance level of $\alpha = 0.05$ (95%) and selects a random sample of $n = 25$ and obtains a sample average burning rate of $\bar{x}$ of 51.3 centimeters per second. What conclusions should be drawn?

Solution:

1. Parameter of interest: The parameter of interest is μ, the mean burning rate.
2. Null hypothesis: $H_0 : \mu_X = 50$ centimeters per second
3. Alternative hypothesis: $H_1 : \mu_X \neq 50$ centimeters per second
4. Test statistic: The test statistic is

$$z^* = \frac{\bar{x} - \mu_0}{\sigma_X / \sqrt{n}}$$



Two tailed: (1- $\alpha/2) = 97.5\% \rightarrow$ **$z_{0.025}$ = 1.96**

5. **To use a fixed significance level test, the boundaries of the critical region would be $z_{0.025} = 1.96$ and $z_{-0.025} = -1.96$.**
6. Computations: Because $\bar{x} = 51.3$ and $\sigma_X = 2$,

$$z^* = \frac{51.3 - 50}{2/\sqrt{25}} = 3.25$$

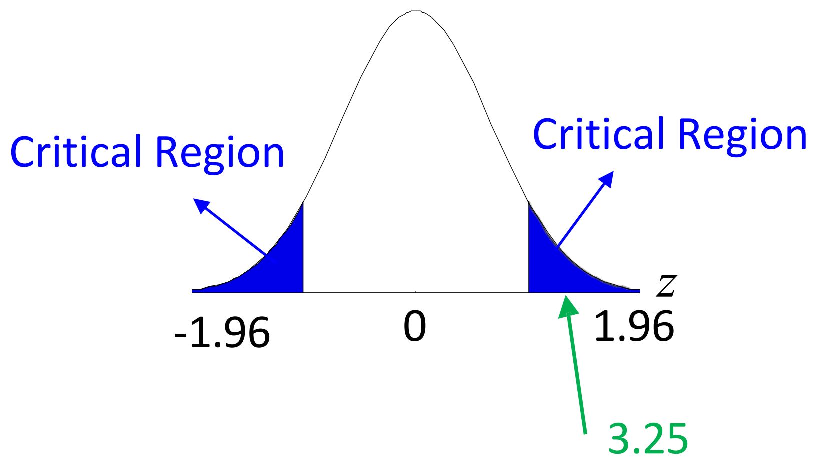
Example

Air crew escape systems are powered by a solid propellant. The burning rate of this propellant is an important product characteristic. Specifications require that the mean burning rate must be 50 centimeters per second. We know that the standard deviation of burning rate is $\sigma_X = 2$ centimeters per second. The experimenter decides to specify a type I error probability or significance level of $\alpha = 0.05$ (95%) and selects a random sample of $n = 25$ and obtains a sample average burning rate of $\bar{x}$ of 51.3 centimeters per second. What conclusions should be drawn?

Solution:

1. Parameter of interest: The parameter of interest is μ, the mean burning rate.
2. Null hypothesis: $H_0 : \mu_X = 50$ centimeters per second
3. Alternative hypothesis: $H_1 : \mu_X \neq 50$ centimeters per second
4. Test statistic: The test statistic is

$$z^* = \frac{\bar{x} - \mu_0}{\sigma_X / \sqrt{n}}$$



5. **To use a fixed significance level test, the boundaries of the critical region would be $z_{0.025} = 1.96$ and $z_{-0.025} = -1.96$.**
6. Computations: Because $\bar{x} = 51.3$ and $\sigma_X = 2$,

$$z^* = \frac{51.3 - 50}{2/\sqrt{25}} = 3.25$$

7. *z^* is in the rejection region ($Z^* < Z$) , so we reject $H_0 : \mu_X = 50$ at the 0.05 level of significance.*

Practical Interpretation: We conclude that the mean burning rate differs from 50 centimeters per second, based on a sample of 25 measurements. In fact, there is strong evidence that the mean burning rate exceeds 50 centimeters per second.

Example

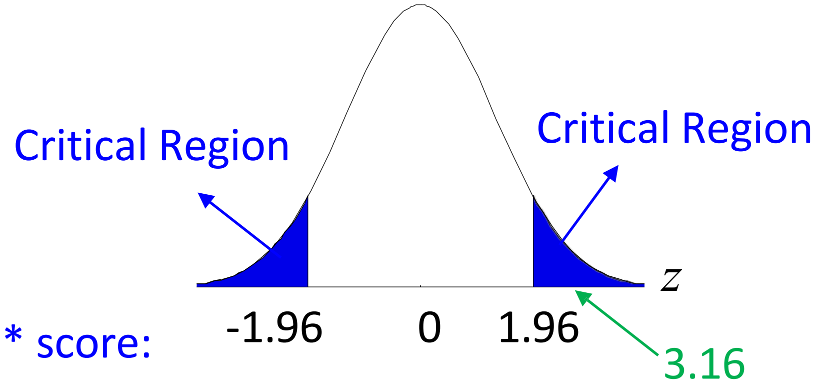
Try it out!

Suppose that we want to test the hypothesis with a significance level of 0.05 that the climate has changed since industrialization. Suppose that the mean temperature throughout history is 50 degrees. During the last 40 years, the mean temperature has been 51 degrees and suppose the population standard deviation is 2 degrees. What can we conclude?

Example

Suppose that we want to test the hypothesis with a significance level of 0.05 that the climate has changed since industrialization. Suppose that the mean temperature throughout history is 50 degrees. During the last 40 years, the mean temperature has been 51 degrees and suppose the population standard deviation is 2 degrees. What can we conclude?

We have
H_0: $\mu = 50$
H_1: $\mu \neq 50$



We Compute the z* score:

$$z^* = \frac{51 - 50}{\frac{2}{\sqrt{40}}} = 3.16$$

95% $\rightarrow$ $z_{0.025} = 1.96$ and $z_{-0.025} = -1.96$.

z^* is in the rejection region, so we reject H_0: $\mu_X = 50$ at the 0.05 level of significance

We can conclude that there has been a change in temperature

Example

Try it out!

The mean water pressure in the main water pipe from a town well should be kept at 56 psi. Anything less and several homes will have an insufficient supply, and anything greater could burst the pipe. Suppose the water pressure is checked at 47 random times. The sample mean is 57.1. (Assume σ = 7) Is there any evidence to suggest the mean water pressure is different from 56? Use $\alpha = 0.01$.

Example

The mean water pressure in the main water pipe from a town well should be kept at 56 psi. Anything less and several homes will have an insufficient supply, and anything greater could burst the pipe. Suppose the water pressure is checked at 47 random times. The sample mean is 57.1. (Assume σ = 7) Is there any evidence to suggest the mean water pressure is different from 56? Use $\alpha = 0.01$.

Solution:

1. The Set-Up:

a) Describe the parameter of concern: The mean water pressure in the main pipe.

b) State the null and alternative hypotheses.

H_0: $\mu = 56$

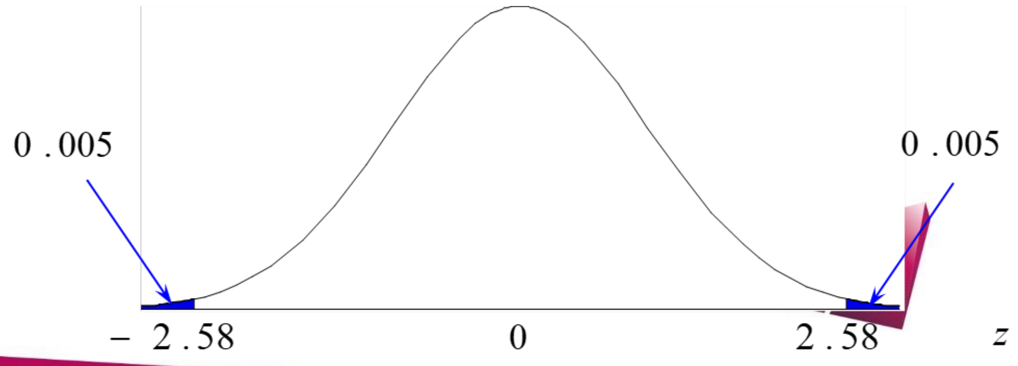
H_1: $\mu \neq 56$

2. The Hypothesis Test Criteria:

a) Check the assumptions: A sample of n = 47 is large enough for the CLT to apply.

b) Identify the test statistic. The test statistic is z^*.

c) Determine the level of significance: α = 0.01



Example

The mean water pressure in the main water pipe from a town well should be kept at 56 psi. Anything less and several homes will have an insufficient supply, and anything greater could burst the pipe. Suppose the water pressure is checked at 47 random times. The sample mean is 57.1. (Assume σ = 7) Is there any evidence to suggest the mean water pressure is different from 56? Use $\alpha = 0.01$.

Solution:

1. The Set-Up:

a) Describe the parameter of concern: The mean water pressure in the main pipe.

b) State the null and alternative hypotheses.

H_0: $\mu = 56$

H_1: $\mu \neq 56$

2. The Hypothesis Test Criteria:

a) Check the assumptions: A sample of n = 47 is large enough for the CLT to apply.

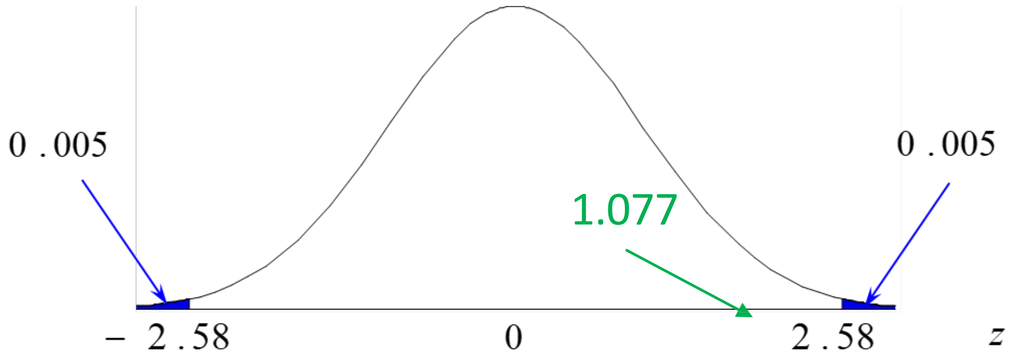
b) Identify the test statistic. The test statistic is z^*.

c) Determine the level of significance: α = 0.01

3. The Sample Evidence:

a) The sample information: $\bar{x} = 57.1, \quad n = 47$

b) Calculate the value of the test statistic: $z^* = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \dfrac{57.1 - 56}{7/\sqrt{47}} = 1.077$



Example

The mean water pressure in the main water pipe from a town well should be kept at 56 psi. Anything less and several homes will have an insufficient supply, and anything greater could burst the pipe. Suppose the water pressure is checked at 47 random times. The sample mean is 57.1. (Assume σ = 7) Is there any evidence to suggest the mean water pressure is different from 56? Use $\alpha = 0.01$.

Solution:

1. The Set-Up:

a) Describe the parameter of concern: The mean water pressure in the main pipe.

b) State the null and alternative hypotheses.

H_0: $\mu = 56$

H_1: $\mu \neq 56$

2. The Hypothesis Test Criteria:

a) Check the assumptions: A sample of n = 47 is large enough for the CLT to apply.

b) Identify the test statistic. The test statistic is z^*.

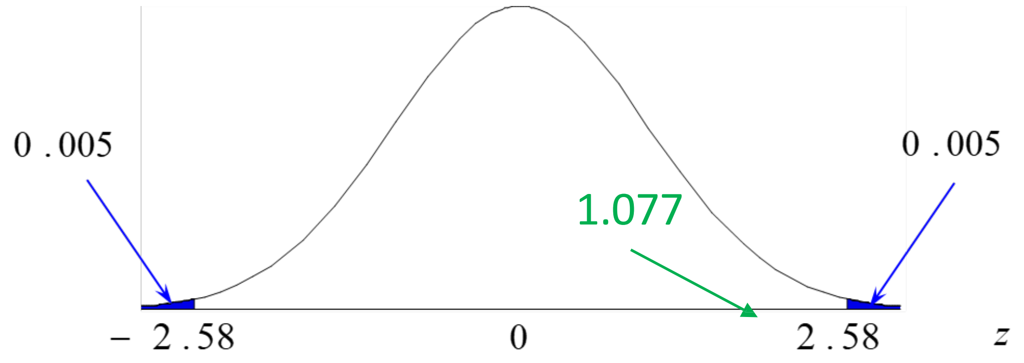
c) Determine the level of significance: $\alpha = 0.01$

3. The Sample Evidence:

a) The sample information: $\bar{x} = 57.1, \quad n = 47$

b) Calculate the value of the test statistic: $z^* = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{57.1 - 56}{7/\sqrt{47}} = 1.077$

4. The calculated value of z^* = 1.077 is in the acceptable region.



Example

The mean water pressure in the main water pipe from a town well should be kept at 56 psi. Anything less and several homes will have an insufficient supply, and anything greater could burst the pipe. Suppose the water pressure is checked at 47 random times. The sample mean is 57.1. (Assume σ = 7) Is there any evidence to suggest the mean water pressure is different from 56? Use $\alpha = 0.01$.

Solution:

1. The Set-Up:

a) Describe the parameter of concern: The mean water pressure in the main pipe.

b) State the null and alternative hypotheses.

H_0: $\mu = 56$

H_1: $\mu \neq 56$

2. The Hypothesis Test Criteria:

a) Check the assumptions: A sample of n = 47 is large enough for the CLT to apply.

b) Identify the test statistic. The test statistic is z^*.

c) Determine the level of significance: α = 0.01

3. The Sample Evidence:

a) The sample information: $\bar{x} = 57.1, \quad n = 47$

b) Calculate the value of the test statistic: $z^* = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \dfrac{57.1 - 56}{7/\sqrt{47}} = 1.077$

5. The Results:

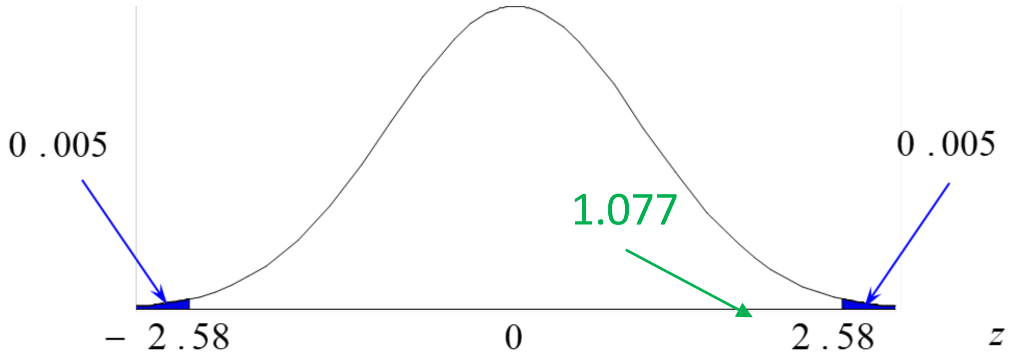
a) State the decision about H_0.

Fail to reject H_0.

b) State the conclusion about H_1.

There is no evidence to suggest the water pressure is different from 56.

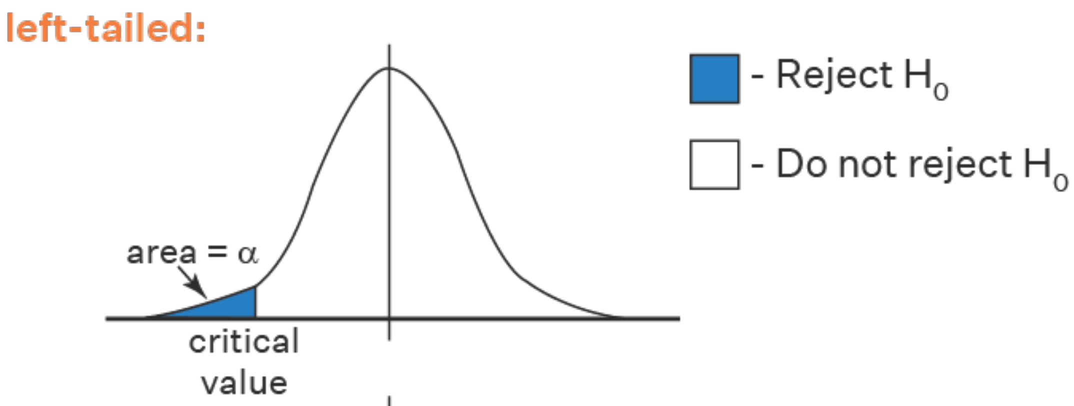
4. The calculated value of z^* = 1.077 is in the acceptable region.

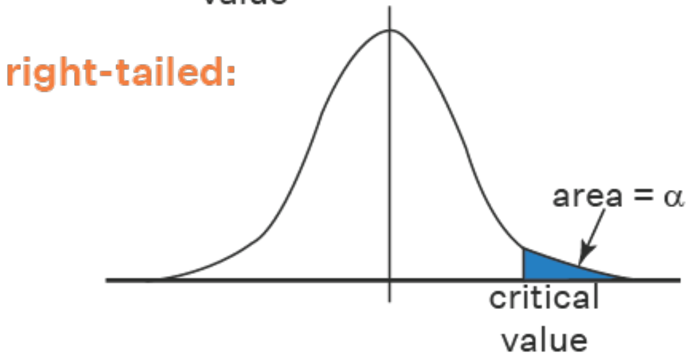


Z critical value

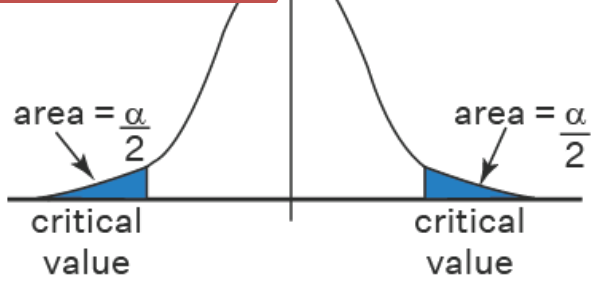
The z critical value can be calculated as follows

- Find the alpha level.
- Subtract the alpha level from 1 for a two-tailed test. For a one-tailed test subtract the alpha level from 0.5 (or 1- $\alpha/2$).
- Look up the **area from the z distribution** table to obtain the z critical value.
- For a left-tailed test, a negative sign needs to be added to the critical value at the end of the calculation.





What about σ Unknown???



Test statistic $$z^* = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

- The sampling distribution of $\bar{x}$ has a normal distribution.
- σ Known

Formulation of a Test Statement about $\boldsymbol{\mu_X}$ (Normal Case)

In last week, when population σ^2 is Unknown and we used sample standard deviation to replace σ^2 had a t distribution to derive the formulas of the random and confidence interval for μ_X. Similarly, in the following we would also use the t distribution to get the test statement of the hypothesis about μ_X when σ_X^2 is Unknown.

Interference about Mean μ (σ unknown)

Hypothesis-Testing Procedure:

1. The t-statistic is used to complete a hypothesis test about a population mean μ.

2. The test statistic:

$$t^* = \frac{\bar{x} - \mu}{s/\sqrt{n}} \; with \; df = n - 1$$

3. The calculated t is the number of estimated standard errors of $\bar{x}$ from the hypothesized mean μ.

Example

[σ^2 is Unknown]

A random sample of 25 students registering for classes showed the mean waiting time in the registration line was 22.6 minutes and the standard deviation was 8.0 minutes. Is there any evidence to support the student newspaper's claim that registration time takes longer than 20minutes? Use α = 0.05 and assume waiting time is approximately normal.

Example

A random sample of 25 students registering for classes showed the mean waiting time in the registration line was 22.6 minutes and the standard deviation was 8.0 minutes. Is there any evidence to support the student newspaper's claim that registration time takes longer than 20minutes? Use $\alpha = 0.05$ and assume waiting time is approximately normal.

Solution:

[σ_X^2 is Unknown]

1. The Set-up:

a) Population parameter of concern: the mean waiting time spent in the registration line.

b) State the null and alternative hypotheses:

H_0: $\mu = 20$ ($\leq$) (no longer than)

H_1: $\mu > 20$ (longer than)

2. The Hypothesis Test Criteria:

a) Check the assumptions: The sampled population is approximately normal.

b) Test statistic: t^* with df = $n - 1 = 24$

c) Level of significance: $\alpha = 0.05$

Example

A random sample of 25 students registering for classes showed the mean waiting time in the registration line was 22.6 minutes and the standard deviation was 8.0 minutes. Is there any evidence to support the student newspaper's claim that registration time takes longer than 20minutes? Use $\alpha = 0.05$ and assume waiting time is approximately normal.

Solution:

$[\sigma_X^2$ is Unknown$]$

1. The Set-up:

a) Population parameter of concern: the mean waiting time spent in the registration line.

b) State the null and alternative hypotheses:

H_0: $\mu = 20$ ($\leq$) (no longer than)

H_1: $\mu > 20$ (longer than)

2. The Hypothesis Test Criteria:

a) Check the assumptions: The sampled population is approximately normal.

b) Test statistic: t^* with df = $n - 1 = 24$

c) Level of significance: $\alpha = 0.05$

***t* Table**

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
$\boxed{24}$	0.000	0.685	0.857	1.059	1.318	$\boxed{1.711}$	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725

Example

A random sample of 25 students registering for classes showed the mean waiting time in the registration line was 22.6 minutes and the standard deviation was 8.0 minutes. Is there any evidence to support the student newspaper's claim that registration time takes longer than 20minutes? Use $\alpha = 0.05$ and assume waiting time is approximately normal.

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725

Solution:

[σ_X^2 is Unknown]

1. The Set-up:

a) Population parameter of concern: the mean waiting time spent in the registration line.

b) State the null and alternative hypotheses:

H_0: $\mu = 20$ ($\leq$) (no longer than)

H_1: $\mu > 20$ (longer than)

2. The Hypothesis Test Criteria:

a) Check the assumptions: The sampled population is approximately normal.

b) Test statistic: t^* with df = $n - 1 = 24$

c) Level of significance: $\alpha = 0.05$

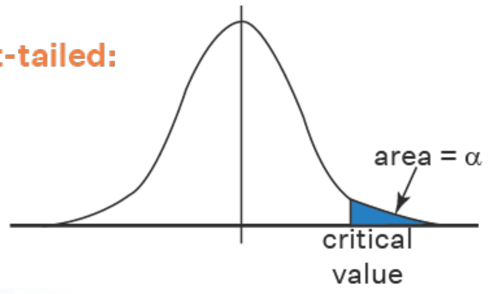
3. The Sample Evidence:

a) Sample information: $n = 25$, $\bar{x} = 22.6$, and $s = 8$

b) Calculate the value of the test statistic:

$$t^* = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{22.6 - 20}{8/\sqrt{25}} = \frac{2.6}{1.6} = 1.625$$

right-tailed:



Example

A random sample of 25 students registering for classes showed the mean waiting time in the registration line was 22.6 minutes and the standard deviation was 8.0 minutes. Is there any evidence to support the student newspaper's claim that registration time takes longer than 20minutes? Use $\alpha = 0.05$ and assume waiting time is approximately normal.

Solution: [σ_X^2 is Unknown]

1. The Set-up:

a) Population parameter of concern: the mean waiting time spent in the registration line.

b) State the null and alternative hypotheses:

H_0: $\mu = 20$ ($\leq$) (no longer than)

H_1: $\mu > 20$ (longer than)

2. The Hypothesis Test Criteria:

a) Check the assumptions: The sampled population is approximately normal.

b) Test statistic: t^* with df = $n - 1 = 24$

c) Level of significance: $\alpha = 0.05$

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725

3. The Sample Evidence:

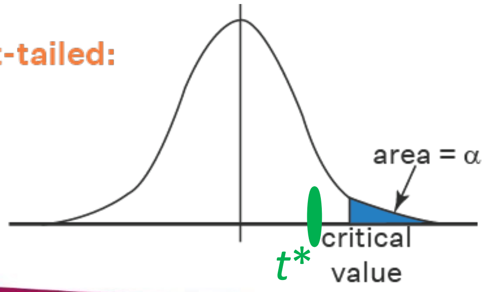
a) Sample information: $n = 25$, $\bar{x} = 22.6$, and $s = 8$

b) Calculate the value of the test statistic:

$$t^* = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{22.6 - 20}{8/\sqrt{25}} = \frac{2.6}{1.6} = 1.625$$

The critical value: $t(24, 0.05) = 1.71$

right-tailed:



t^* is *not in the rejection* region

→Fail to reject H_0.

Example

[σ^2 is Unknown]

Try it out!

Average nicotine content is at least 1.6 mg (1.6). A firm claims that the average nicotine content of their cigarettes is less than 1.6 (< 1.6). Test at 5% significance level. The average nicotine of a sample of 20 cigarettes from this firm is 1.54, with sample standard deviation 1.2.

Example

Average nicotine content is at least 1.6 mg (1.6). A firm claims that the average nicotine content of their cigarettes is less than 1.6 (< 1.6). Test at 5% significance level. The average nicotine of a sample of 20 cigarettes from this firm is 1.54, with sample standard deviation 1.2.

Solution: We are testing

$$H_0: \mu_x = 1.6 \quad v.s. \quad H_1: \mu_x < 1.6$$

From the sample data

$$\sqrt{n}\frac{\bar{X} - \mu_0}{S} = \frac{\sqrt{20}(1.54 - 1.6)}{1.2} = -0.224$$

But $-t_{\alpha,19} = -1.73 < -0.224$.

t^* is *not* in the rejection region → fail to reject H_0

***t* Table**

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
$\boxed{\text{one-tail}}$	0.50	0.25	0.20	0.15	0.10	$\boxed{0.05}$	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
$\boxed{19}$	0.000	0.688	0.861	1.066	1.328	$\boxed{1.729}$	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850

Example

Average nicotine content is at least 1.6 mg (1.6). A firm claims that the average nicotine content of their cigarettes is less than 1.6 (< 1.6). Test at 5% significance level. The average nicotine of a sample of 20 cigarettes from this firm is 1.54, with sample standard deviation 1.2.

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850

Solution: We are testing

$$H_0: \mu_x = 1.6 \quad v.s. \quad H_1: \mu_x < 1.6$$

From the sample data

$$\sqrt{n}\,\frac{\bar{X} - \mu_0}{S} = \frac{\sqrt{20}(1.54 - 1.6)}{1.2} = -0.224$$

But $-t_{\alpha,19} = -1.73 < -0.224.$

t^* is *not* in the rejection region → fail to reject H_0

What If the same sample mean and sd are from a sample of 1100??

Example

Average nicotine content is at least 1.6 mg (1.6). A firm claims that the average nicotine content of their cigarettes is less than 1.6 (< 1.6). Test at 5% significance level. The average nicotine of a sample of 20 cigarettes from this firm is 1.54, with sample standard deviation 1.2.

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850

Solution: We are testing

$$H_0: \mu_x = 1.6 \quad v.s. \quad H_1: \mu_x < 1.6$$

From the sample data

$$\sqrt{n}\,\frac{\bar{X} - \mu_0}{S} = \frac{\sqrt{20}(1.54 - 1.6)}{1.2} = -0.224$$

But $-t_{\alpha,19} = -1.73 < -0.224.$

*t** is *not* in the rejection region → fail to reject H_0

If the same sample mean and sd are from a sample of 1100,

$$TS = \frac{\sqrt{1100}\,(1.54 - 1.6\,)}{1.2} = -1.658 < -t_{\alpha,1099} = -1.646$$

(from t-distribution table)

t^* is in the rejection region → Reject H_0

Power of a Test

Power of a Test

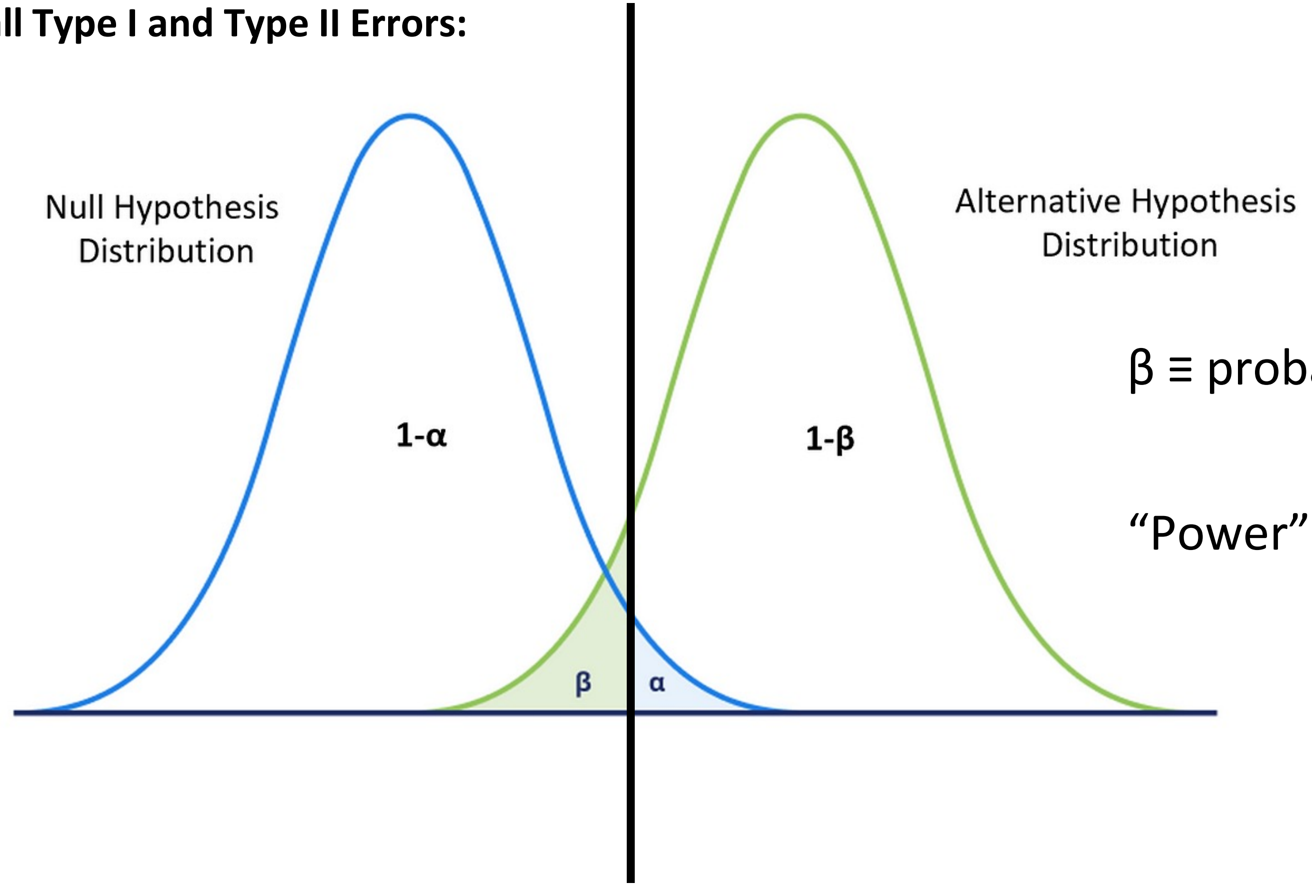
The test procedures allow us to directly control the probability of rejecting a true H_0 by our choice of the significance level $\boldsymbol{\alpha}$.

But what about the probability of rejecting H_0 when it is false? As we will see, several factors influence this probability.

Power of a Test

When we consider the probability of rejecting the null hypothesis, we are looking at what statisticians refer to as the **power** of the test.

Recall Type I and Type II Errors:



β ≡ probability of a Type II error

"Power" ≡ probability of avoiding a Type II error= 1-β

That is, the power of a hypothesis test is the probability that it will reject when it's supposed to.

Example

The director of financial aid services believes that average amount spent on textbooks is $500 each semester, and uses this to determine the amount of financial aid for which a student is eligible. The student body president plans to ask each student in a random sample how much he or she spent on books this semester and use the data to test (using α =.05) the following hypotheses:

H_0: $\mu = 500$ versus H_a: $\mu > 500$

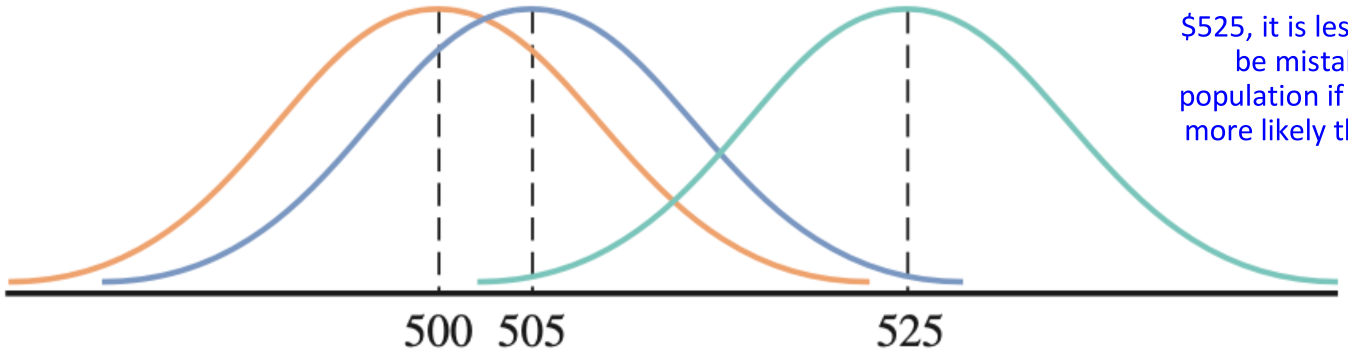
Example

The director of financial aid services believes that average amount spent on textbooks is $500 each semester, and uses this to determine the amount of financial aid for which a student is eligible. The student body president plans to ask each student in a random sample how much he or she spent on books this semester and use the data to test (using $\boldsymbol{\alpha}$ =.05) the following hypotheses:

H_0: $\mu = 500$ versus H_a: $\mu > 500$

The power of a test depends on the value of the mean!

If the true mean is greater than $500, then we should reject H_0. BUT, if the true mean is ONLY a little greater, say $505, then the sample mean might look like we expect if the true mean were $500.
Thus we wouldn't have convincing evidence to reject H_0.



However, if the true mean was $525, it is less likely that the sample would be mistaken for a sample from the population if the mean were $500. So, it is more likely that we will correctly reject H_0.

Sampling distribution of x when μ = 500, 505, 525.

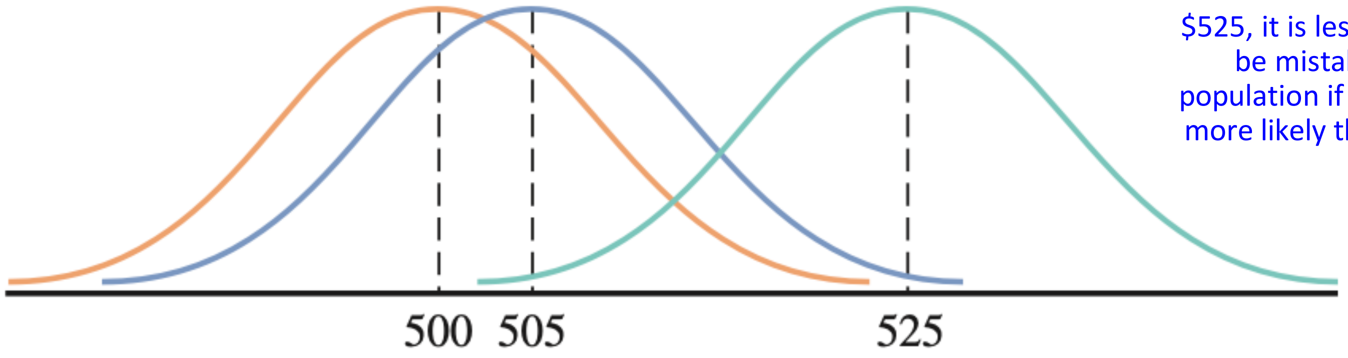
Example

The director of financial aid services believes that average amount spent on textbooks is \$500 each semester, and uses this to determine the amount of financial aid for which a student is eligible. The student body president plans to ask each student in a random sample how much he or she spent on books this semester and use the data to test (using α =.05) the following hypotheses:

H_0: $\mu = 500$ versus H_a: $\mu > 500$

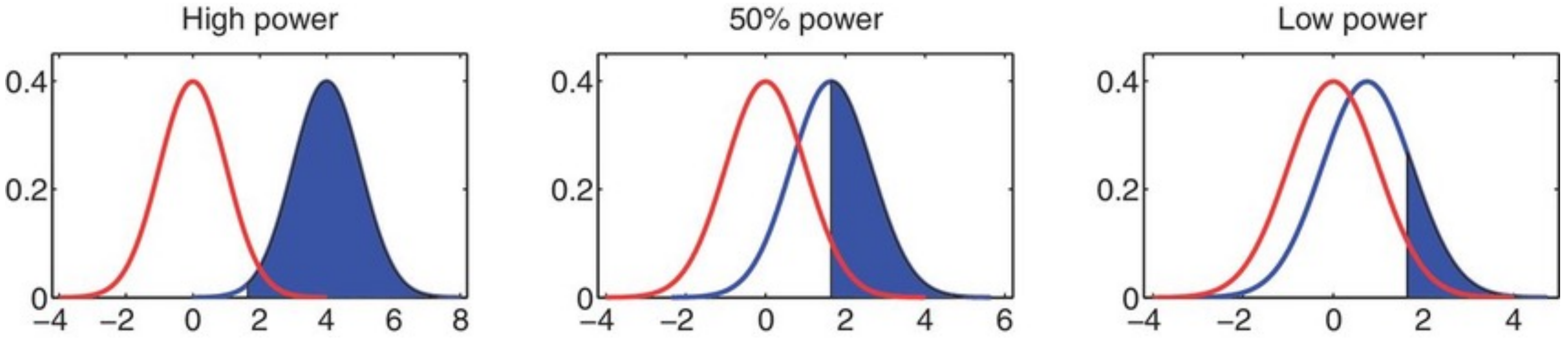
The power of a test depends on the value of the mean!

If the true mean is greater than \$500, then we should reject H_0. BUT, if the true mean is ONLY a little greater, say \$505, then the sample mean might look like we expect if the true mean were \$500.
Thus we wouldn't have convincing evidence to reject H_0.



However, if the true mean was \$525, it is less likely that the sample would be mistaken for a sample from the population if the mean were \$500. So, it is more likely that we will correctly reject H_0.

Sampling distribution of x when μ = 500, 505, 525.


High power
50% power
Low power
0.4
0.2
0
−4
−2
0
2
4
6
8

The director of financial aid services believes that average amount spent on textbooks is \$500 each semester, and uses this to determine the amount of financial aid for which a student is eligible. The student body president plans to ask each student in a random sample how much he or she spent on books this semester and use the data to test (using α =.05) the following hypotheses:

H_0: $\mu = 500$ versus H_a: $\mu > 500$

Suppose that s = \$85 and n = 100. (Since n is large, the sampling distribution of $\bar{x}$ is approximately normal.)

What is the probability of committing a Type I error?

If $\mu = 500$ is true, for what values of the sample mean would you reject the null hypothesis?

The director of financial aid services believes that average amount spent on textbooks is \$500 each semester, and uses this to determine the amount of financial aid for which a student is eligible. The student body president plans to ask each student in a random sample how much he or she spent on books this semester and use the data to test (using α =.05) the following hypotheses:

H_0: $\mu = 500$ versus H_a: $\mu > 500$

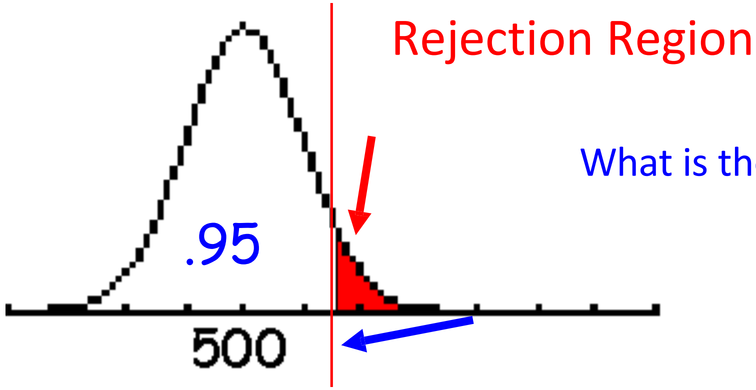
Suppose that s = \$85 and n = 100. (Since n is large, the sampling distribution of $\bar{x}$ is approximately normal.)

What is the probability of committing a Type I error?

If $\mu = 500$ is true, for what values of the sample mean would you reject the null hypothesis?

α =.05

This is the z value with .95 area to its left. (z value=1.645)



What is the value of this $\bar{x}$?

$$1.645 = \frac{\bar{x} - 500}{85/\sqrt{100}}$$

$\bar{x} = 513.98$

We would reject H_0 for $\bar{x} \geq 513.98$.

The director of financial aid services believes that average amount spent on textbooks is \$500 each semester, and uses this to determine the amount of financial aid for which a student is eligible. The student body president plans to ask each student in a random sample how much he or she spent on books this semester and use the data to test (using $\boldsymbol{\alpha}$ =.05) the following hypotheses:

H_0: $\mu = 500$ versus H_a: $\mu > 500$

Suppose that s = \$85 and n = 100.

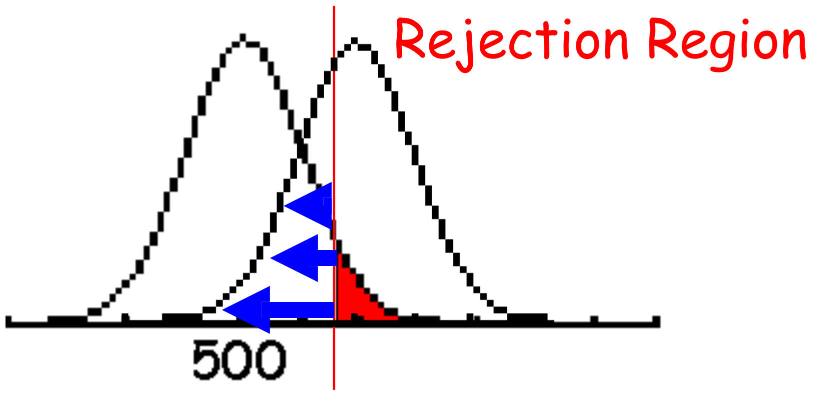
We would reject H_0 for $\bar{x} \geq 513.98$.

If the null hypothesis is false, then $\mu > 500$.

What if μ=520?

What is the probability of a Type II error (B)?

This area (to the left of $\bar{x}$ = 513.98) is β .



$$Z^* = \frac{513.98 - 520}{85/\sqrt{100}} = -0.708$$

Check the probability form z table

β =0.239

The director of financial aid services believes that average amount spent on textbooks is \$500 each semester, and uses this to determine the amount of financial aid for which a student is eligible. The student body president plans to ask each student in a random sample how much he or she spent on books this semester and use the data to test (using α =.05) the following hypotheses:

H_0: $\mu = 500$ versus H_a: $\mu > 500$

Suppose that $s = \$85$ and $n = 100$.

We would reject H_0 for $\overline{x} \geq 513.98$.

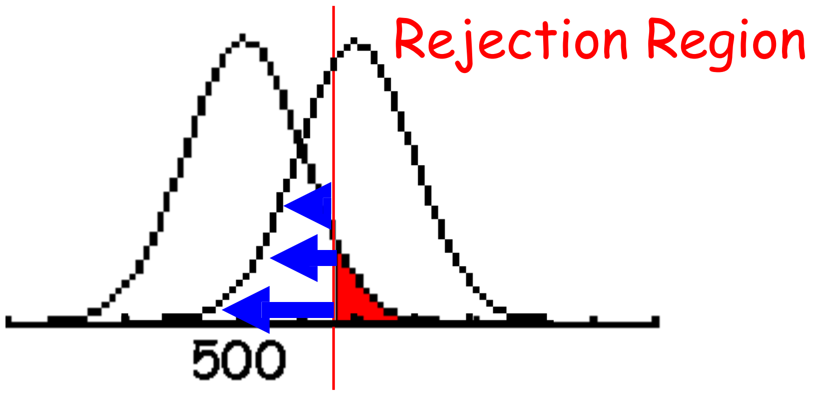
If the null hypothesis is false, then $\mu > 500$.
What if $\mu = 520$?

What is the power of the test if $\mu = 520$?

Power is the probability of correctly rejecting H_0.

Notice that power is in the SAME curve as β

Power = 1 – β



$$Z^* = \frac{513.98 - 520}{85/\sqrt{100}} = -0.708$$

Check the probability form z table

β =0.239

Power = 1 - .239 = .761

The director of financial aid services believes that average amount spent on textbooks is \$500 each semester, and uses this to determine the amount of financial aid for which a student is eligible. The student body president plans to ask each student in a random sample how much he or she spent on books this semester and use the data to test (using α =.05) the following hypotheses:

H_0: $\mu = 500$ versus H_a: $\mu > 500$

Suppose that s = \$85 and n = 100.

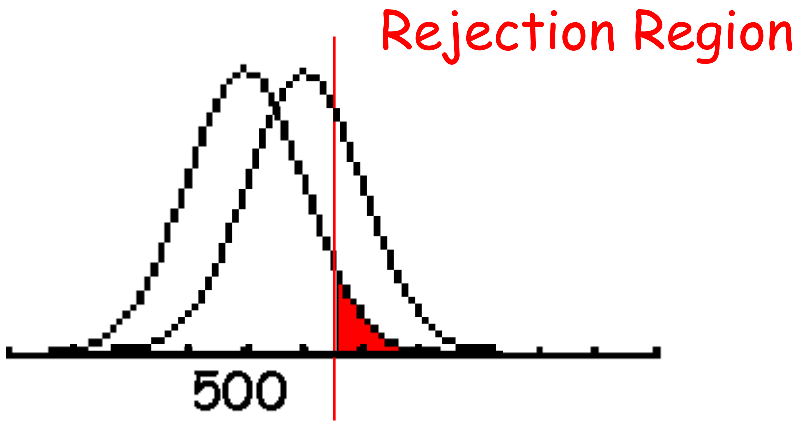
We would reject H_0 for $x \geq 513.98$.

If the null hypothesis is false, then $\mu > 500$.

What if $\mu = 520$?

What if $\mu = 510$?

Find β and power.



$$z = \frac{513.98 - 510}{85/\sqrt{100}} = 0.4817$$

$\beta = .685$

power = .315

Notice that, as the distance between the null hypothesized value for μ and our alternative value for μ decreases, β increases AND power decreases.

The director of financial aid services believes that average amount spent on textbooks is \$500 each semester, and uses this to determine the amount of financial aid for which a student is eligible. The student body president plans to ask each student in a random sample how much he or she spent on books this semester and use the data to test (using α =.05) the following hypotheses:

H_0: $\mu = 500$ versus H_a: $\mu > 500$

Suppose that s = \$85 and n = 100.

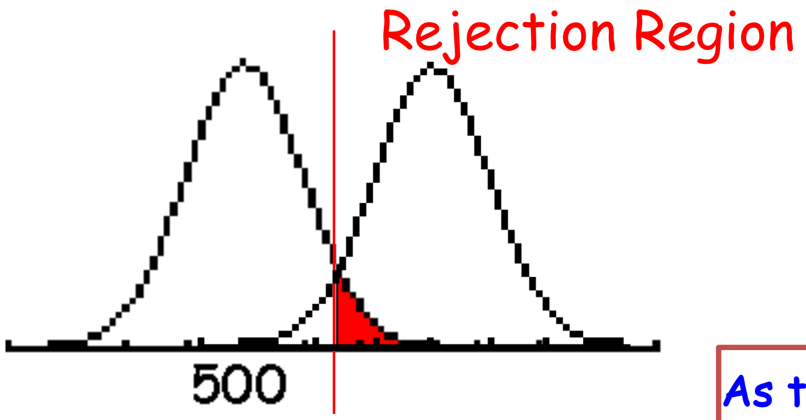
We would reject H_0 for $\bar{x} \geq 513.98$.

If the null hypothesis is false, then $\mu > 500$.
What if $\mu = 520$?

What if μ = 530?

Find β and power.

$$z = \frac{513.98 - 530}{85/\sqrt{100}} = -1.884$$



β = .03

power = .97

As the distance between the null hypothesized value for μ and our alternative value for μ increases, β decreases AND power increases.

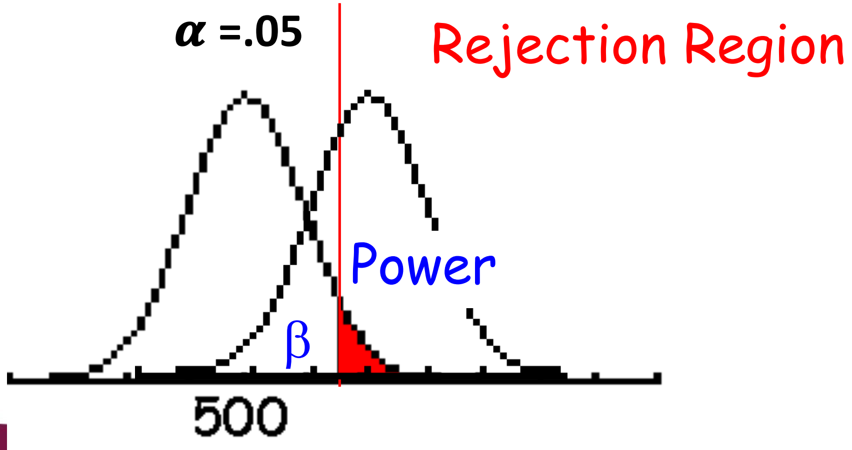
The director of financial aid services believes that average amount spent on textbooks is \$500 each semester, and uses this to determine the amount of financial aid for which a student is eligible. The student body president plans to ask each student in a random sample how much he or she spent on books this semester and use the data to test (using α =.05) the following hypotheses:

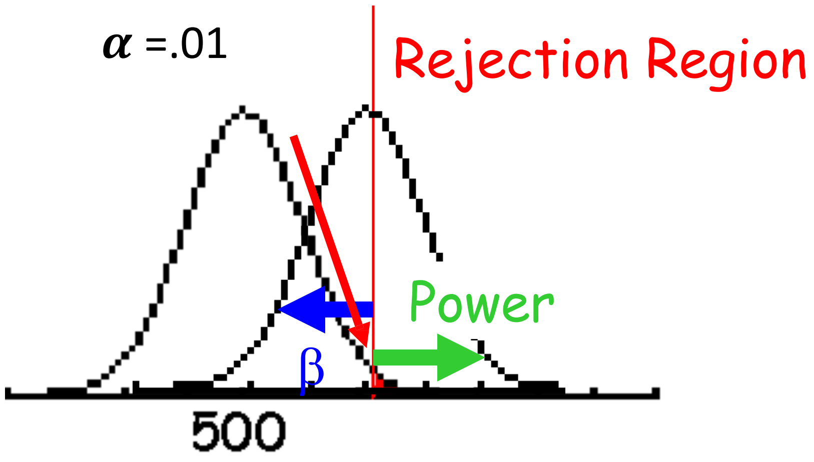
H_0: $\mu = 500$ versus H_a: $\mu > 500$

Suppose that s = \$85 and n = 100.

What happens if we use α = .01?

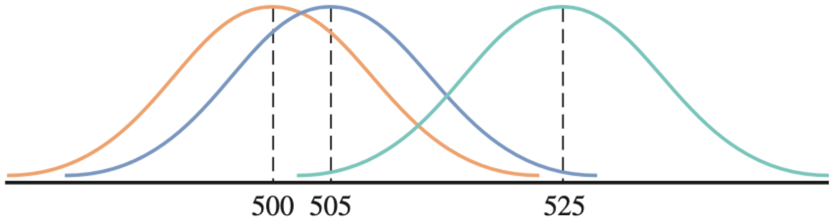
β will increase and power will decrease.





Example

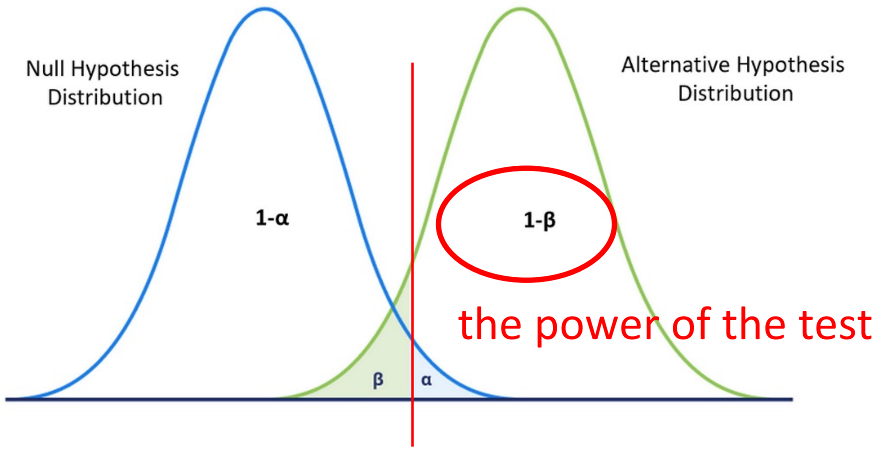
H_0: μ = 500 versus H_a: μ > 500



For any value of μ exceeding 500, the power of the test is greater if

- If the actual mean is 525 than if the actual mean is 505
- Using a larger significance level $\boldsymbol{\alpha}$
- Using a larger sample size

Effect of Various Factors on the Power of a Test



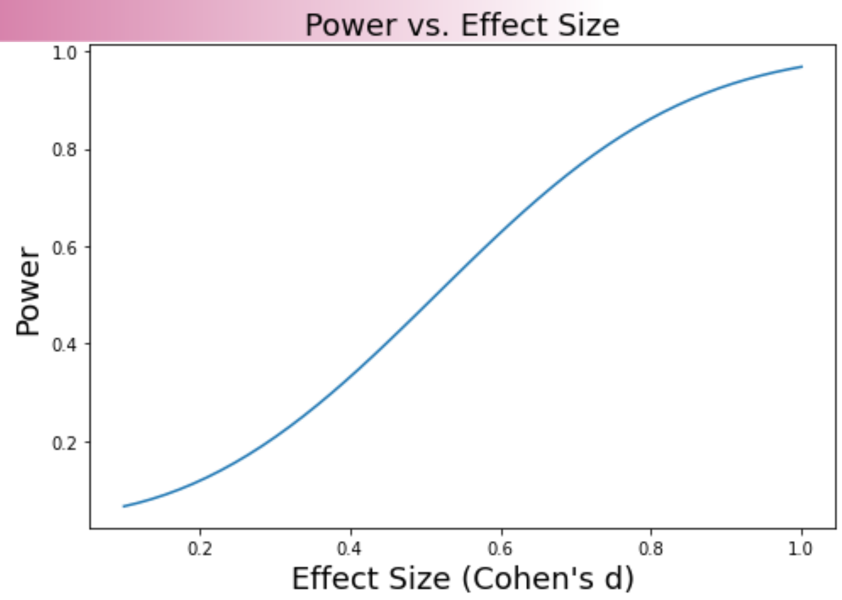
- The larger the size of the discrepancy between the hypothesized value and the actual value of the population characteristic, the higher the power.

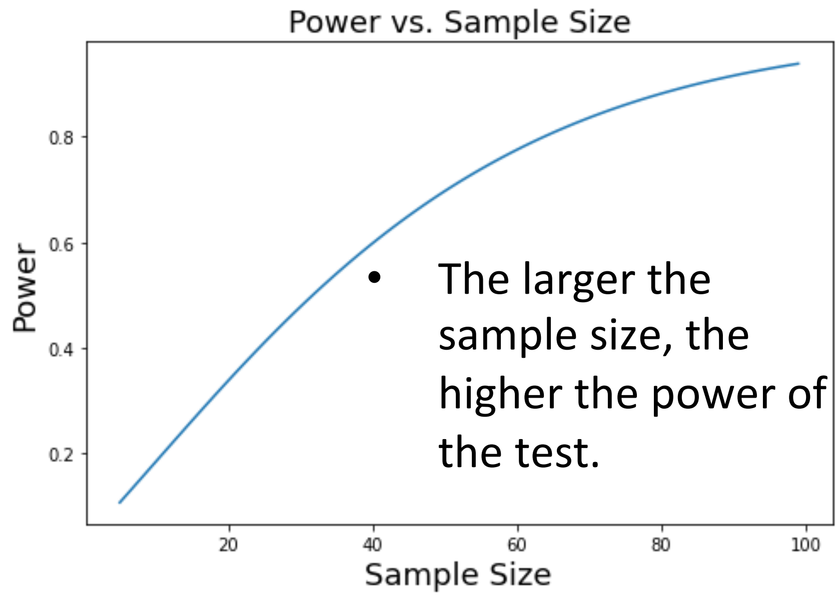
- The larger the significance level, α , the higher the power of the test.

- The larger the sample size, the higher the power of the test.

- The population standard deviation increases, resulting in an incorrect conclusion, the power will decrease

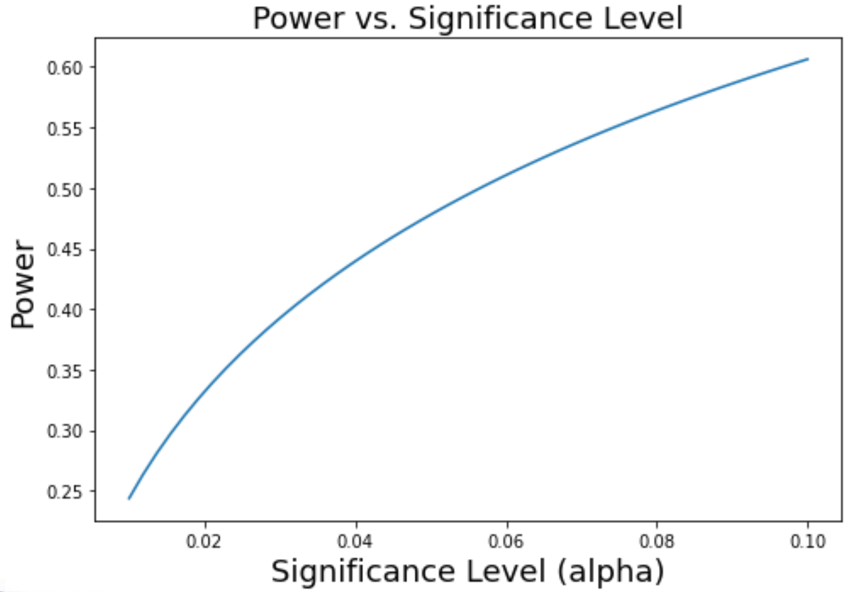
- The larger the size of the discrepancy between the hypothesized value and the actual value of the population characteristic, the higher the power.

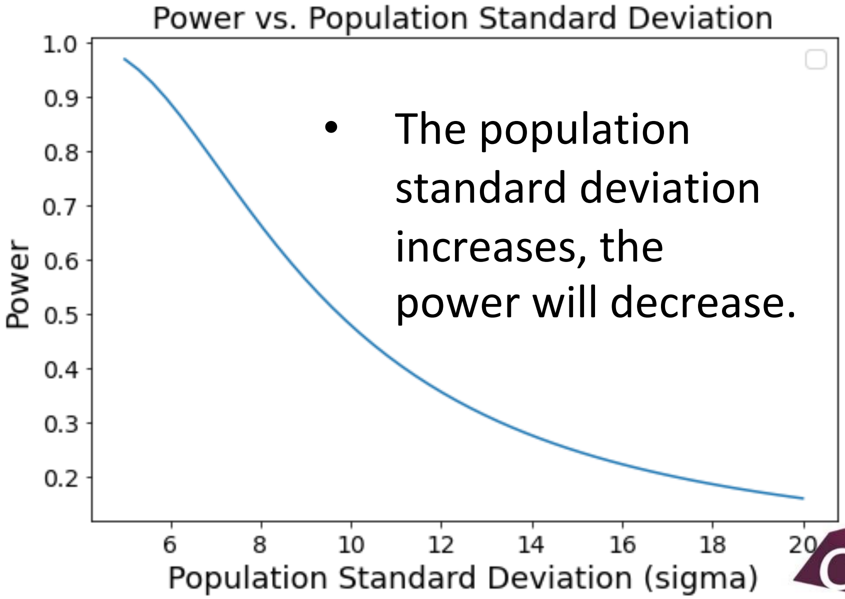




- The larger the sample size, the higher the power of the test.

- The larger the significance level, α , the higher the power of the test.





- The population standard deviation increases, the power will decrease.

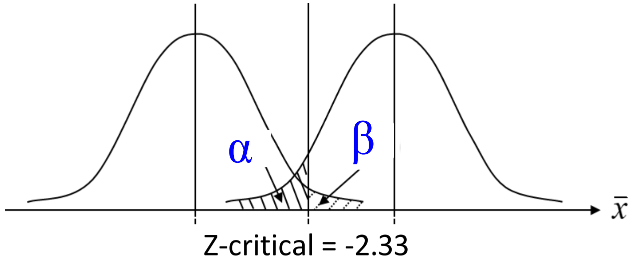
Example

A normally distributed population is known to have a standard deviation of $\sigma = 2$, but its mean is in question. It has been argued to be either $\mu = 10.5$ or $\mu = 12$. A random sample of size 20 is drawn from the population and the sample mean is 11.

a) At $\alpha = 0.01$ level of significance, test the null hypothesis,

$H_0 : \mu = 12$ against the alternative hypothesis, $H_1 : \mu < 12$.

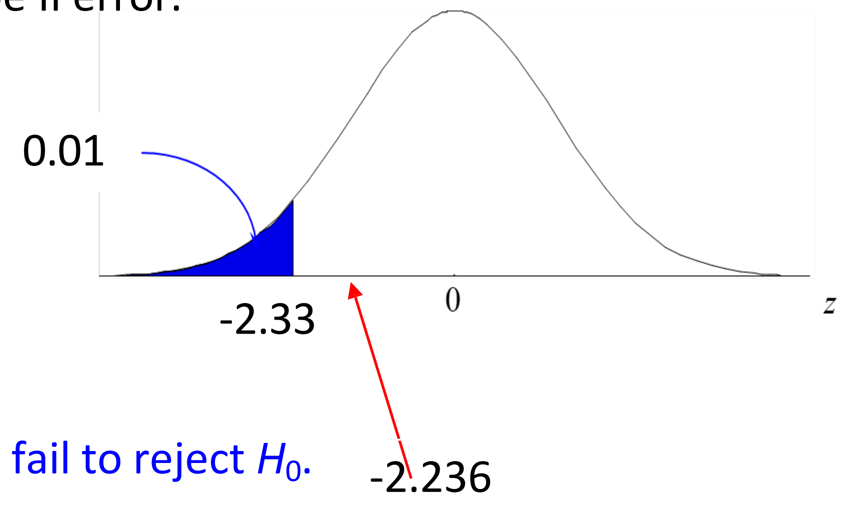
b) Find β, the probability of the type II error.



a) $H_0 : \mu = 12$,

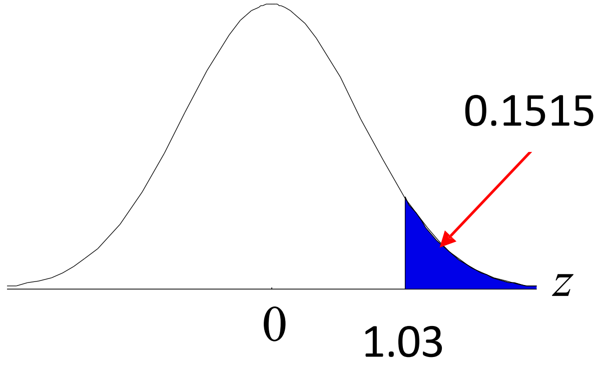
$H_1 : \mu < 12$.

$$z^* = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{11 - 12}{2/\sqrt{20}} = -2.236.$$



z-critical = $-z(0.01) = -2.33$.

As z^* is in the acceptable region, we fail to reject H_0.



b) If H_0 is true, the critical value of is

$$\frac{2}{\sqrt{20}} \times (-2.33) + 12 = 10.96.$$

If H_1 is true,

$$z^* = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{10.96 - 10.5}{2/\sqrt{20}} = 1.03.$$

Then, $\beta = P(z > 1.03) = 0.1515$.

CityU

Python for hypothesis test

Python for hypothesis test

To perform a one-sample hypothesis test in Python, you can use the scipy.stats library, which provides a function called ttest_1samp.

```
# Perform one-sample t-test
t_statistic, p_value = ttest_1samp(data, popmean)
```

When you execute a one-sample t-test using ttest_1samp, you provide it with the sample data and the hypothesized population mean.

It then returns two values:

- t_statistic: A value that represents the number of standard errors that the sample mean deviates from the hypothesized population mean. A larger absolute value of the t-statistic indicates a greater deviation between the sample mean and the population mean.

- p_value: The probability of observing data as extreme as those observed if the null hypothesis is true. If the p-value is less than the chosen significance level (commonly 0.05), you reject the null hypothesis, suggesting there is a statistically significant difference between the sample mean and the population mean.

Python for hypothesis test

Example

```
import numpy as np
from scipy.stats import ttest_1samp

# Sample data - replace this with your actual data
data = np.array([2.5, 3.1, 2.8, 3.6, 2.9, 3.2, 3.0, 2.7, 2.8, 3.3])

# Null hypothesis mean value
popmean = 3.0

# Perform one-sample t-test
t_statistic, p_value = ttest_1samp(data, popmean)

# Output the t-statistic and the p-value
print(f"t-statistic: {t_statistic:.3f}")
print(f"p-value: {p_value:.3f}")

# Typically, we compare the p-value with a significance level (alpha), e.g., 0.05
alpha = 0.05
if p_value < alpha:
    print(f"Since the p-value ({p_value:.3f}) is less than alpha ({alpha}), we reject the null hypothesis.")
else:
    print(f"Since the p-value ({p_value:.3f}) is greater than alpha ({alpha}), we fail to reject the null hypothesis
```

```
t-statistic: -0.098
p-value: 0.924
Since the p-value (0.924) is greater than alpha (0.05), we fail to reject the null hypothesis.
```

To calculate the power of a one-sample t-test, you would typically need to know:

1. The effect size (the standardized difference you expect to detect),
2. The sample size (number of observations),
3. The significance level (alpha, usually 0.05), and
4. The population standard deviation (to calculate the effect size if not directly given).

The statistical power is the probability that the test will correctly reject the null hypothesis when it is false (i.e., detect an effect if there is one).

Python's statsmodels

Python's statsmodels library has functions for conducting power analysis.

```
statsmodels.stats.power.tt_ind_solve_power(
    effect_size=None,
    nobs1=None,
    alpha=None,
    power=None,
    ratio=1.0,
    alternative='two-sided'
)
```

https://www.statsmodels.org/stable/generated/statsmodels.stats.power.tt_ind_solve_power.html

Parameters

effect_size : `float`

standardized effect size, difference between the two means divided by the standard deviation. *effect_size* has to be positive.

nobs1 : `int` or `float`

number of observations of sample 1. The number of observations of sample two is ratio times the size of sample 1, i.e. `nobs2 = nobs1 * ratio`

alpha : `float` `in` `interval` (0,1)

significance level, e.g. 0.05, is the probability of a type I error, that is wrong rejections if the Null Hypothesis is true.

power : `float` `in` `interval` (0,1)

power of the test, e.g. 0.8, is one minus the probability of a type II error. Power is the probability that the test correctly rejects the Null Hypothesis if the Alternative Hypothesis is true.

ratio : `float`

ratio of the number of observations in sample 2 relative to sample 1. see description of nobs1 The default for ratio is 1; to solve for ratio given the other arguments it has to be explicitly set to None.

alternative : `str`, 'two-sided' (`default`), 'larger', 'smaller'

extra argument to choose whether the power is calculated for a two-sided (default) or one sided test. The one-sided test can be either 'larger', 'smaller'.

Python's statsmodels

Example:

```python
from statsmodels.stats.power import tt_ind_solve_power

# Parameters for the power calculation
effect_size = 0.5  # Cohen's d, small=0.2, medium=0.5, large=0.8
sample_size = 30  # Number of observations in the sample
alpha = 0.05  # Significance level
std_dev = 10  # Population standard deviation
mean_difference = effect_size * std_dev

# Calculate the effect size from the mean difference and standard deviation
cohen_d = mean_difference / std_dev

# Calculate the power
power = tt_ind_solve_power(effect_size=cohen_d, nobs1=sample_size, alpha=alpha, ratio=1, alternative='two-sided')

print(f"The calculated power is: {power:.4f}")
```

```
The calculated power is: 0.4779
```