In-Class Quiz (the best 4 out of 6)

Week 9 (Mar 19)	• In-class Quiz 4: L07-L08
Week 10 (Mar 26)	• In-class Quiz 5: L09-L10
Week 11 (Apr 2)	• In-class Quiz 6: L11

Taking the six quizzes is not mandatory, you can skip two if you are satisfied with your scores

Continuation on L07:
Hypothesis Testing - One Population

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 <u>the complement of H_0</u>, 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.

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

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

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 $\boldsymbol{H_0}$ when it is in fact true.

TYPE II ERROR: the error of not rejecting $\boldsymbol{H_0}$ when it is in fact false.

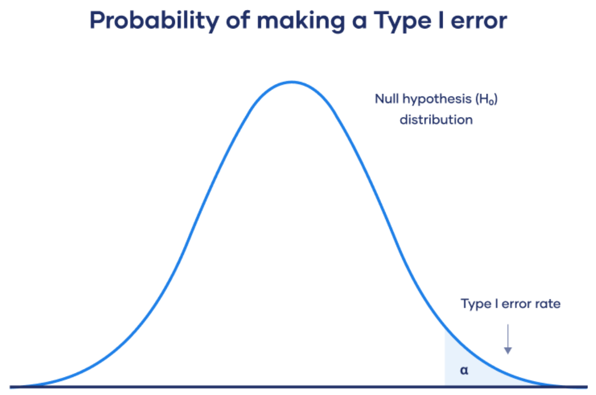
Test Errors and Error Probabilities

Correspondingly, we have

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

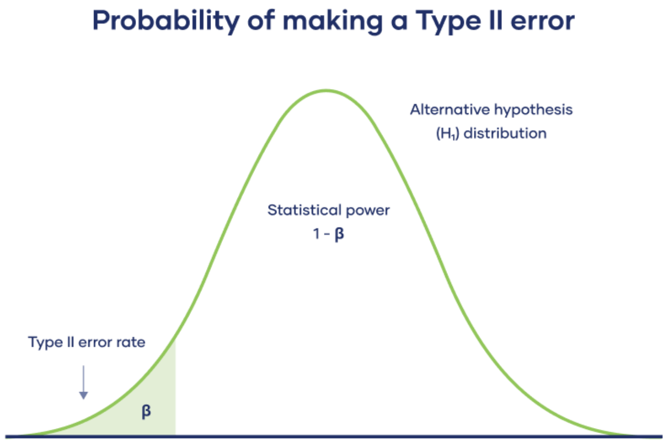
$$\alpha = P\big(Type\ I\ error\big) = P(reject\ H_0\ if\ H_0\ is\ true)$$

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



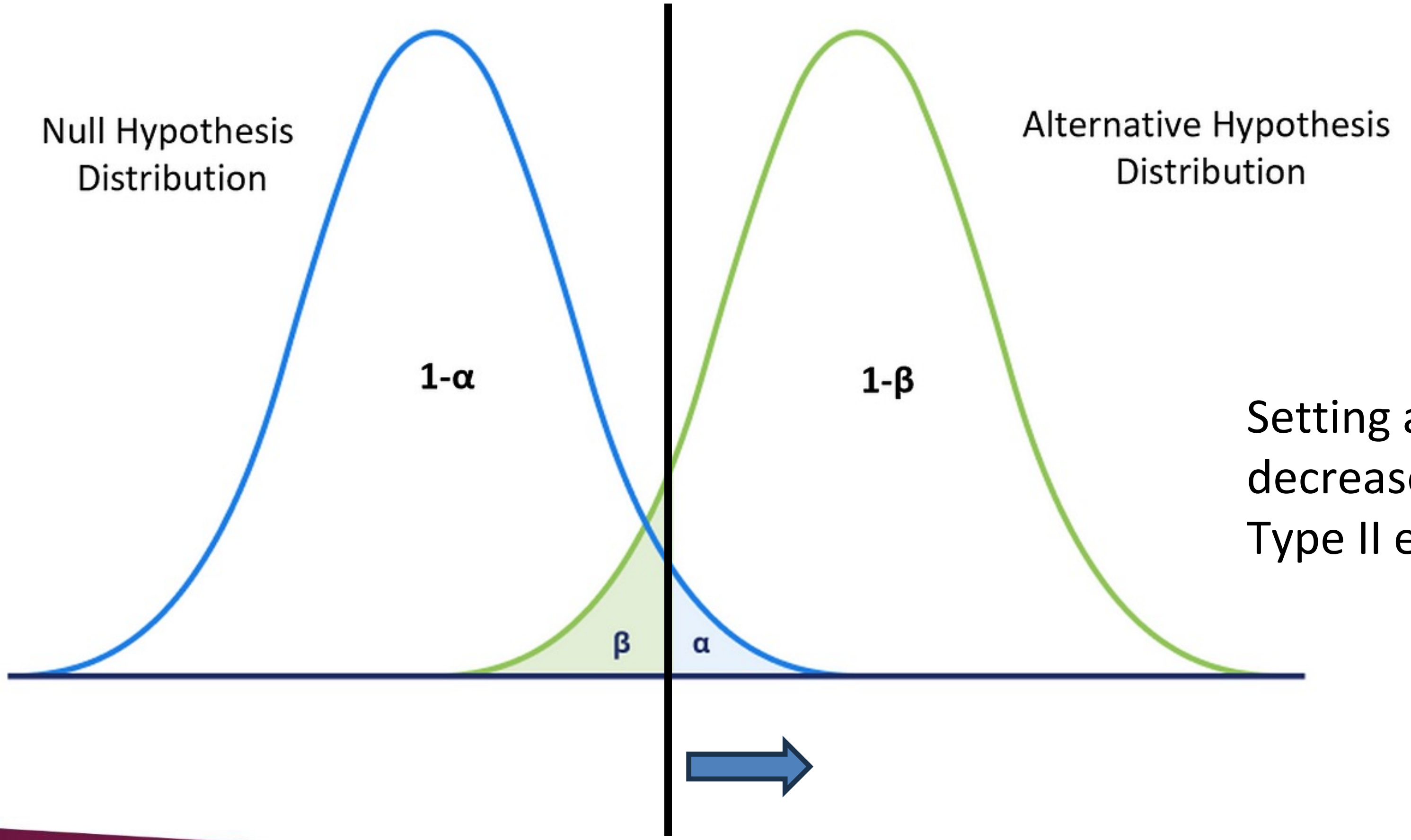
$$\beta = P\big(Type\ II\ error\big) = P(Not\ reject\ H_0\ if\ H_0\ is\ false)$$

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.

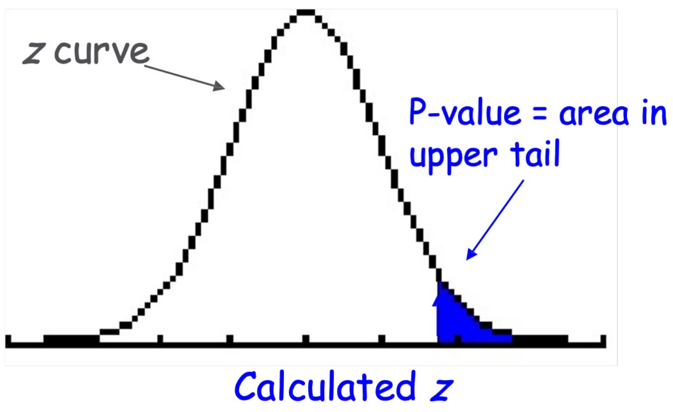
2. Probability-value approach (*p*-value approach)

What is P-value?

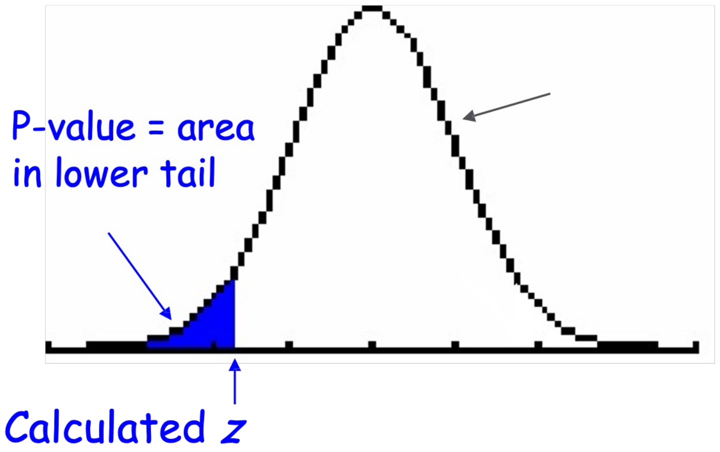
The ***P*-value** (also sometimes called the **observed significance level**) measures of the **strength of the evidence against the null hypothesis (H_0).** It's calculated from the observed data and represents the probability of obtaining results **at least as extreme as the observed results**, assuming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

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

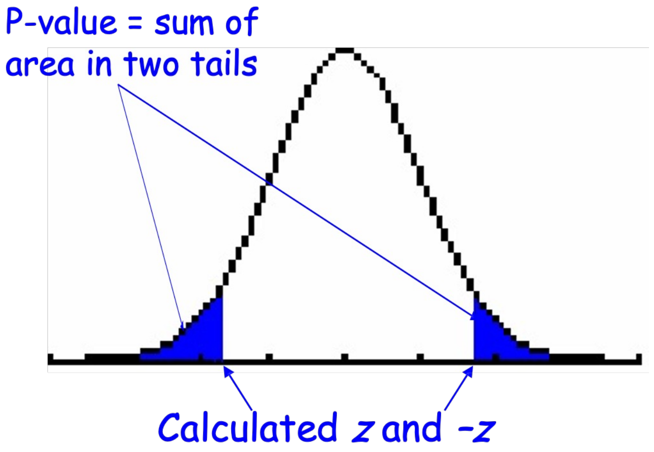
- H_1: *p > hypothesize value*



- H_1: *p < hypothesize value*



- H_a: *p ≠ hypothesize value*



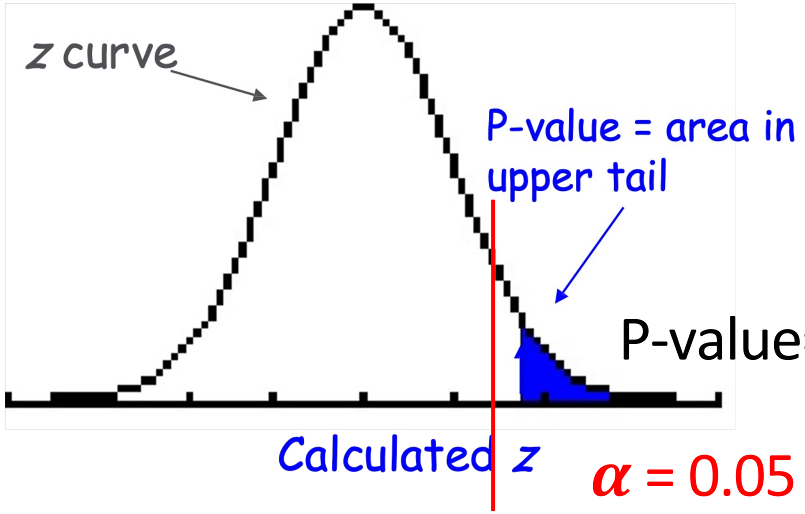
The smaller the p-value, the stronger the evidence against H_0 provided by the data.

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 $> \alpha$.

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



P-value= 0.0228

$\alpha = 0.05$

$\alpha = 0.05$
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.

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

A Classical Approach

- 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.

Decision Criteria

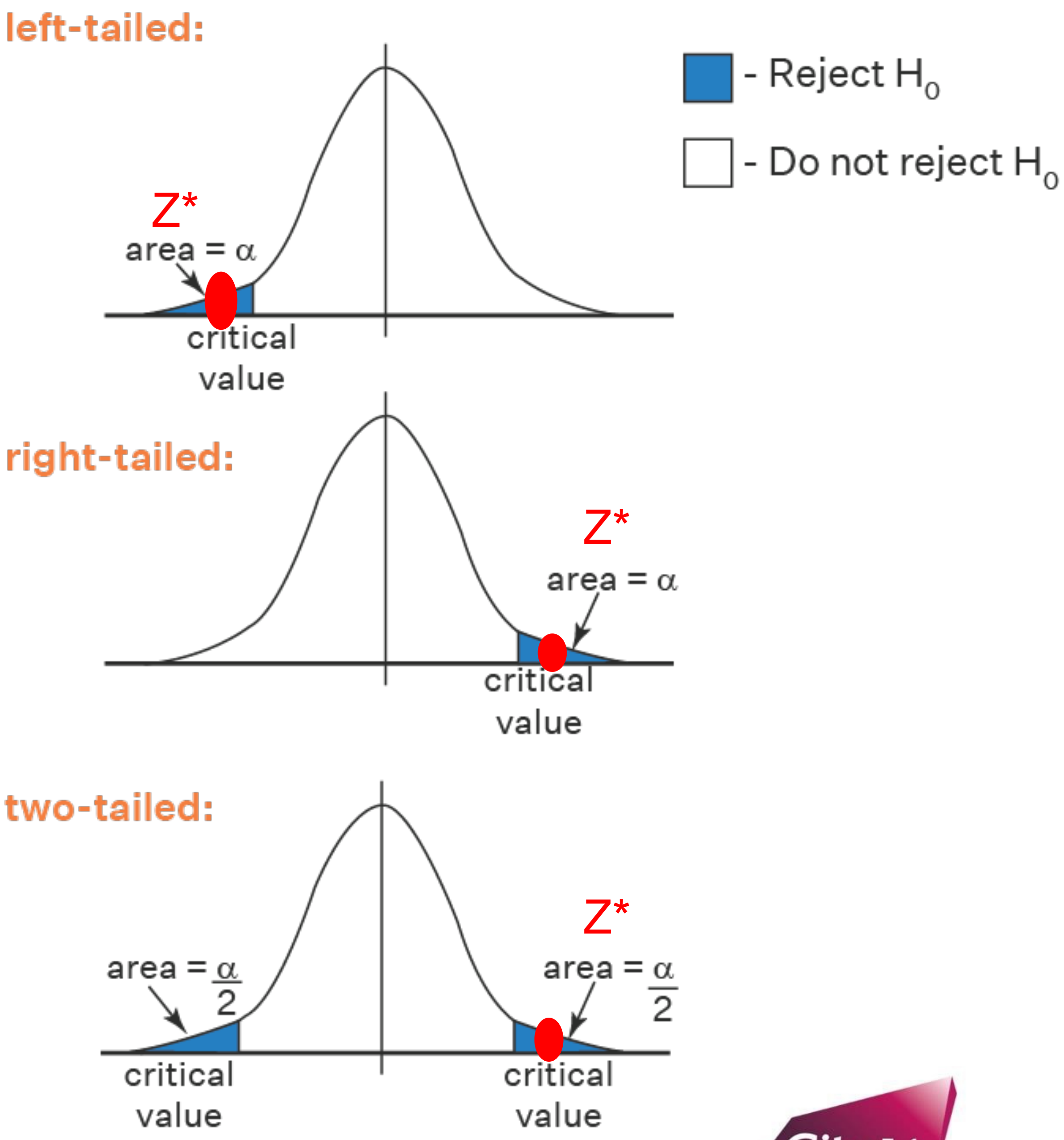
Test statistic $$z^* = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

•**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).



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.

Power of a Test

Power of a Test

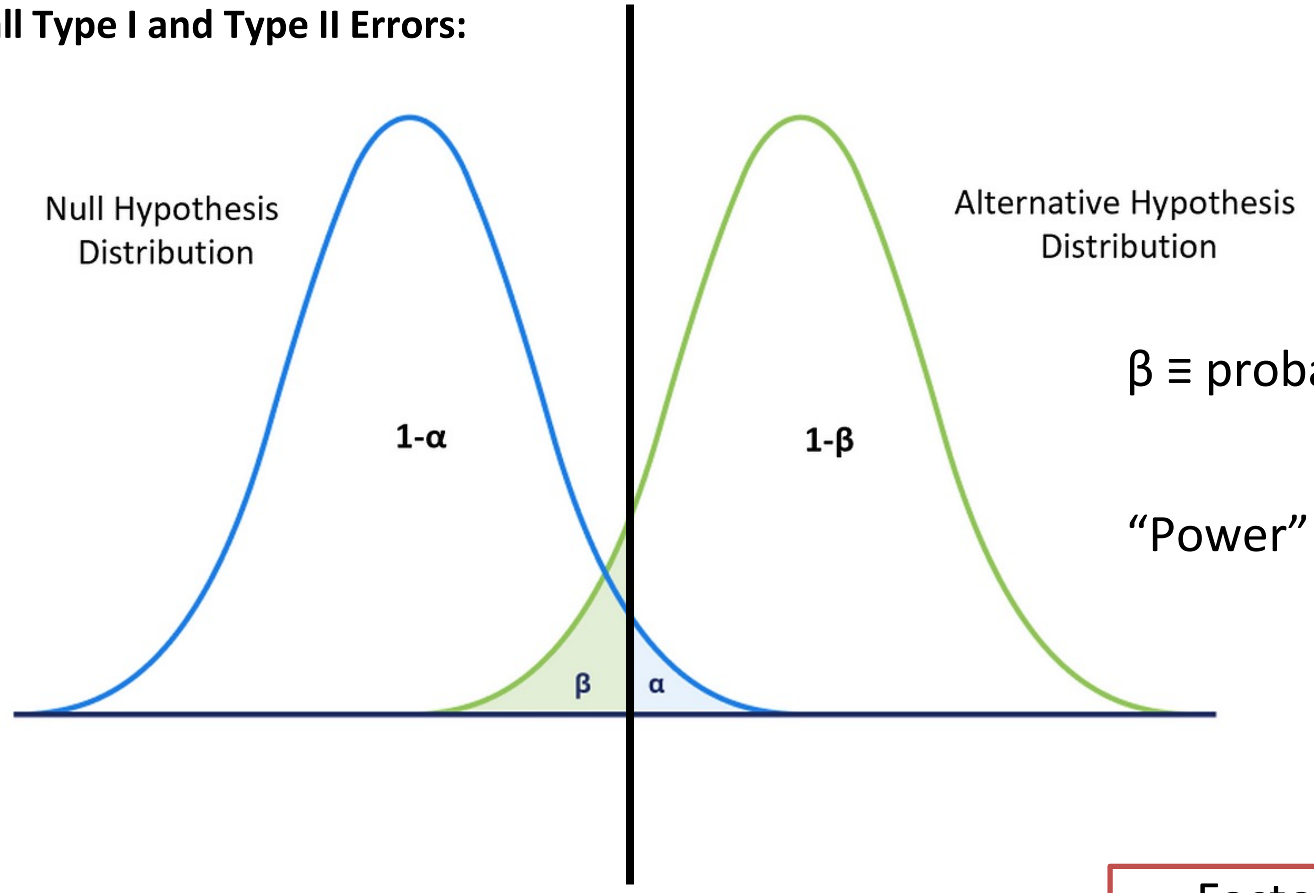
The test procedures allow us to directly control the probability of rejecting a true H_0 by our choice of the significance level $\boldsymbol{\alpha}$.

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:



$\beta \equiv$ probability of a Type II error

"Power" $\equiv$ probability of avoiding a Type II error= $1-\beta$

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

Factors affect the power of a test?

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

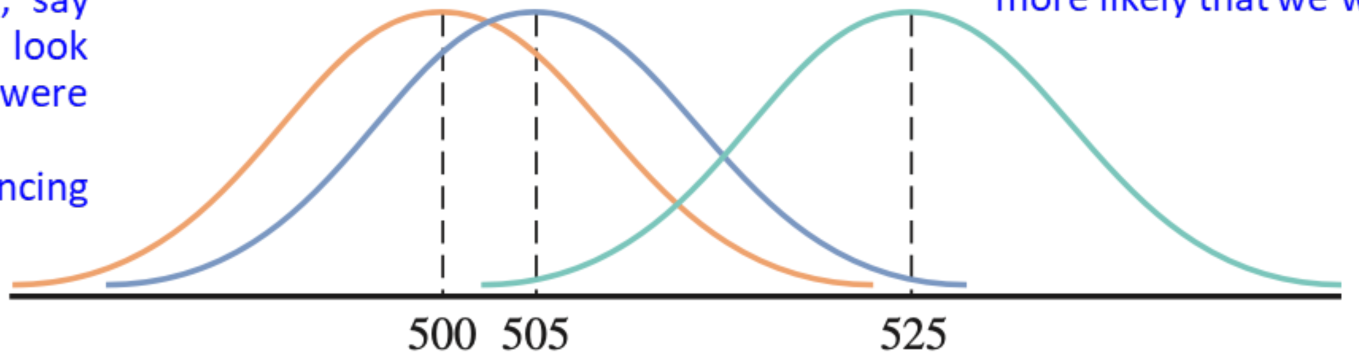
The director of financial aid services believes that average amount spent on textbooks is \$500 each semester, and uses this to determine the amount of financial aid for which a student is eligible. The student body president plans to ask each student in a random sample how much he or she spent on books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

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

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

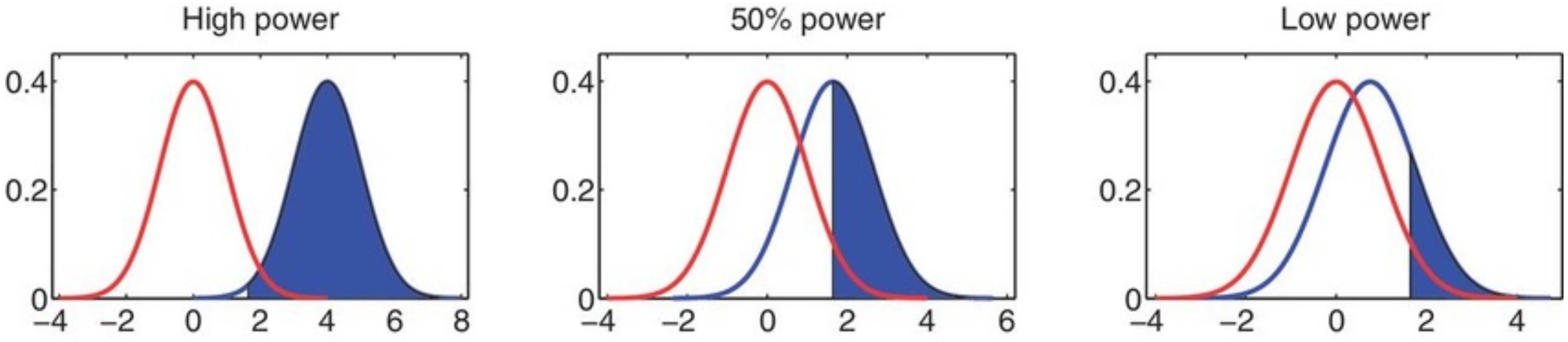
If the true mean is greater than \$500, then we should reject H_0. BUT, if the true mean is ONLY a little greater, say \$505, then the sample mean might look like we expect if the true mean were \$500.
Thus we wouldn't have convincing evidence to reject H_0.

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.

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



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.)

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:

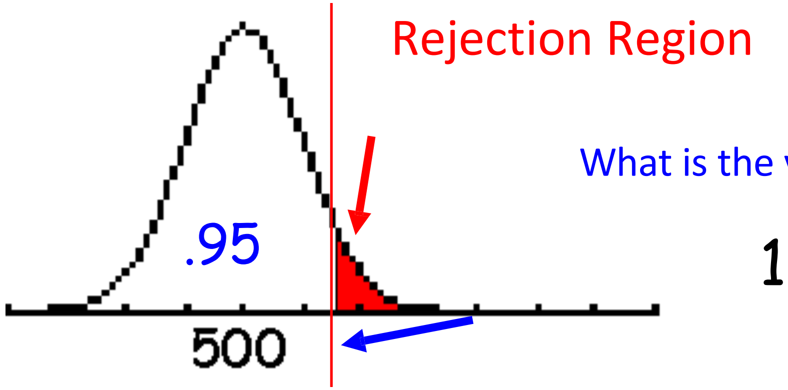
H_0: $\mu = 500$ versus H_a: $\mu > 500$

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

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 α =.05) the following hypotheses:

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

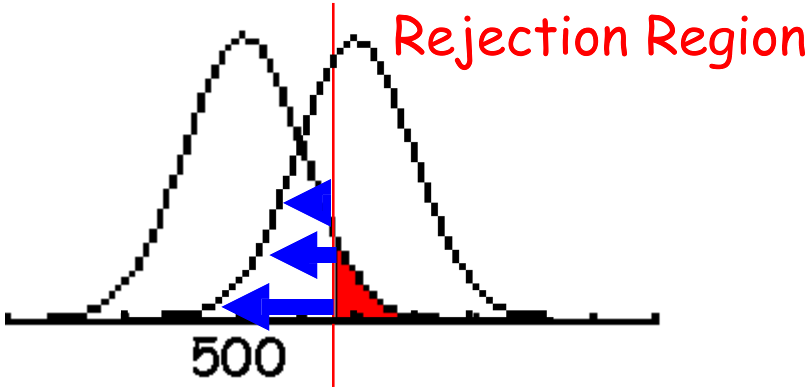
H_0: $\mu = 500$ versus H_a: $\mu > 500$

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 is the probability of a Type II error (β)?

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



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: $\mu > 500$

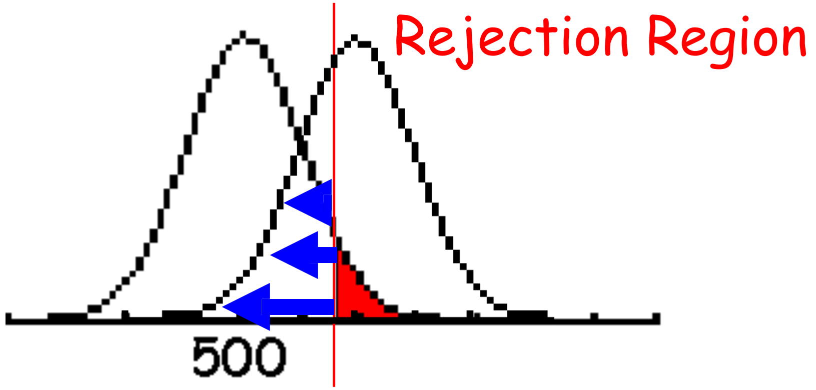
Suppose that s = \$85 and n = 100.

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

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

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

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 from 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 $\bar{x} \geq 513.98$.

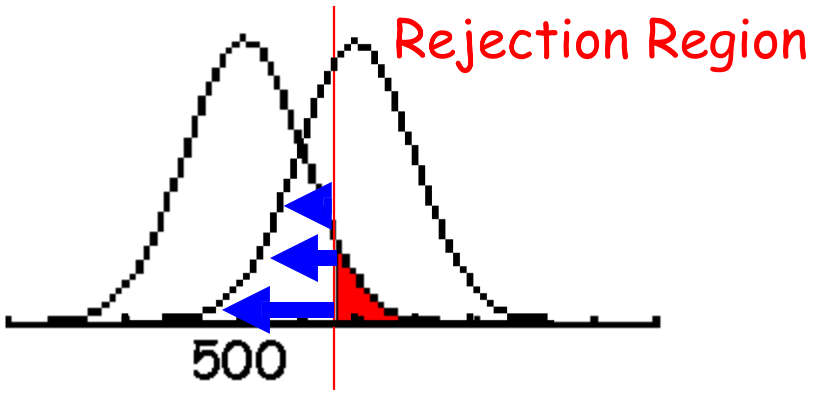
If the null hypothesis is false, then $\mu > 500$.
What if μ=520?

What is the power of the test?

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 from z table

β =0.239

Power = 1 - .239 = .761

CityU

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.

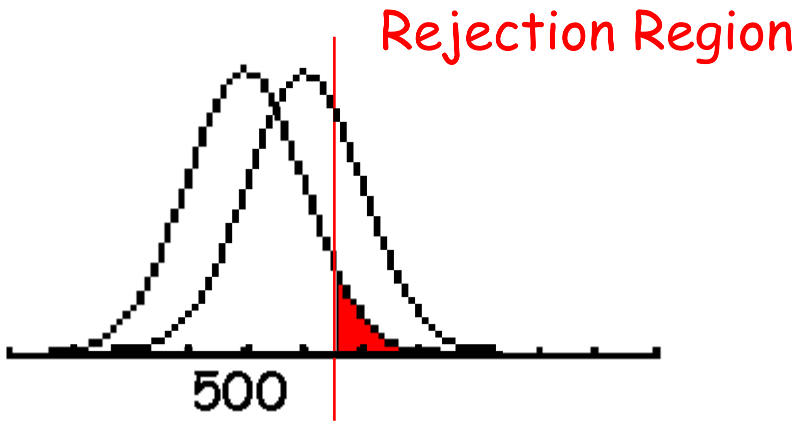
We would reject H_0 for $x \geq 513.98$.

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

What if μ = 510?

Find β and power.

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



β = .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 $x \geq 513.98$.

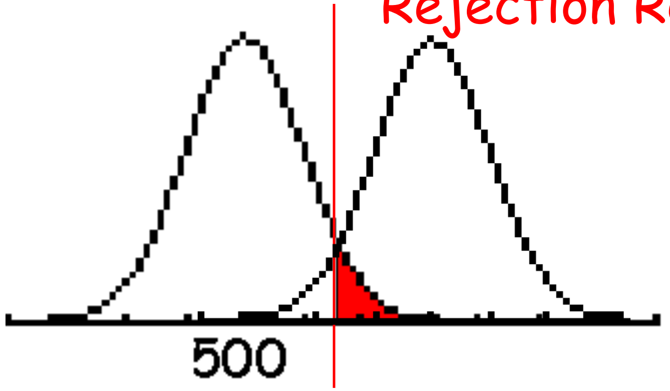
If the null hypothesis is false, then $\mu > 500$.

What if $\mu = 520$?

What if $\mu = 530$?

Find β and power.

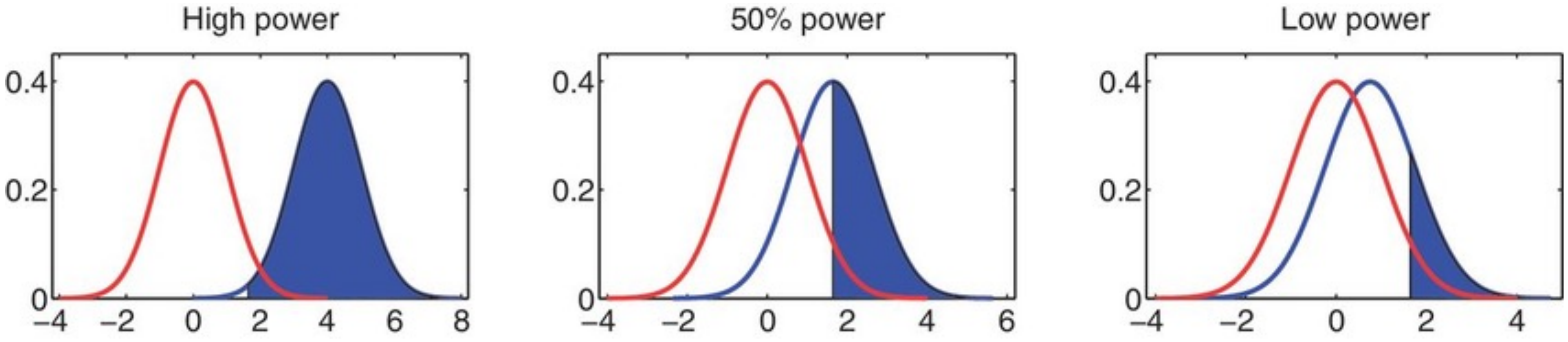
$$z = \frac{513.98 - 530}{85/\sqrt{100}} = -1.884$$



$\beta = .03$ 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 larger the size of the discrepancy between the hypothesized value and the actual value of the population characteristic, the higher the power.



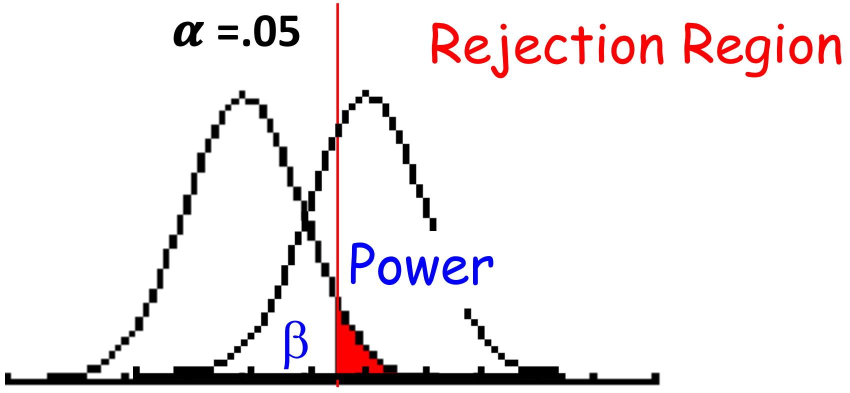
The director of financial aid services believes that average amount spent on textbooks is \$500 each semester, and uses this to determine the amount of financial aid for which a student is eligible. The student body president plans to ask each student in a random sample how much he or she spent on books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

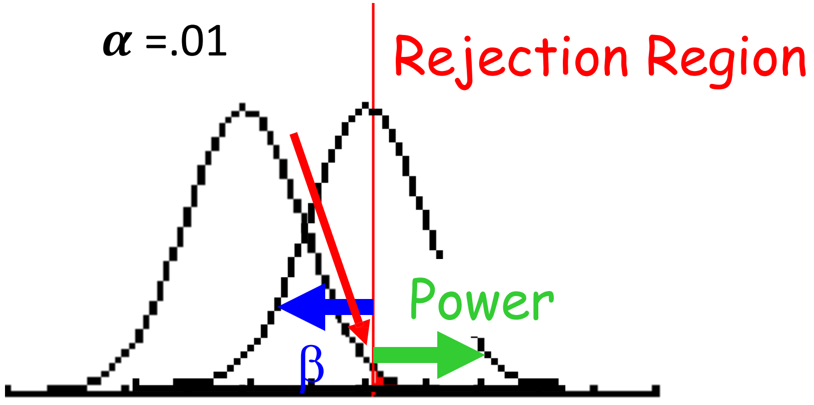
H_0: $\mu = 500$ versus H_a: $\mu > 500$

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

What happens if we use α = .01?

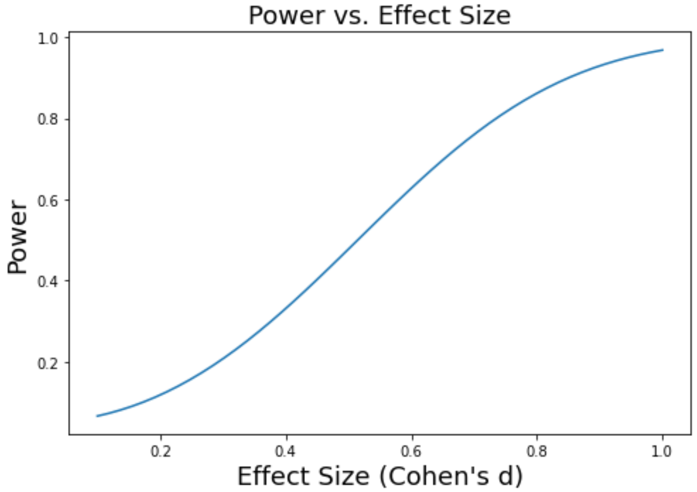
β will increase and power will decrease.



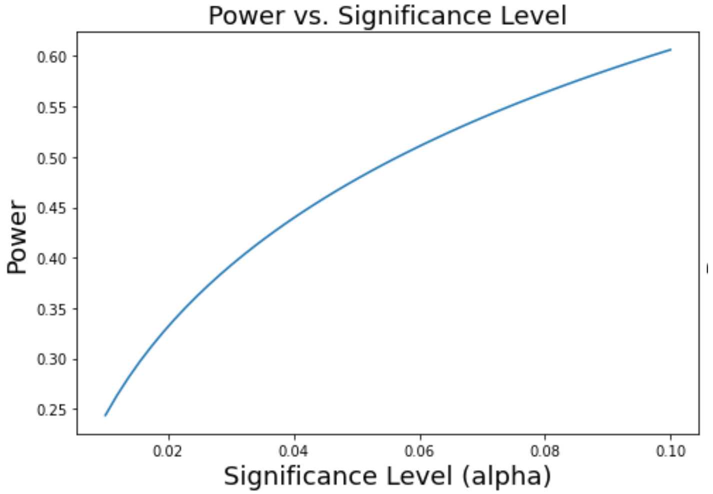


Effect of Various Factors on the Power of a Test

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



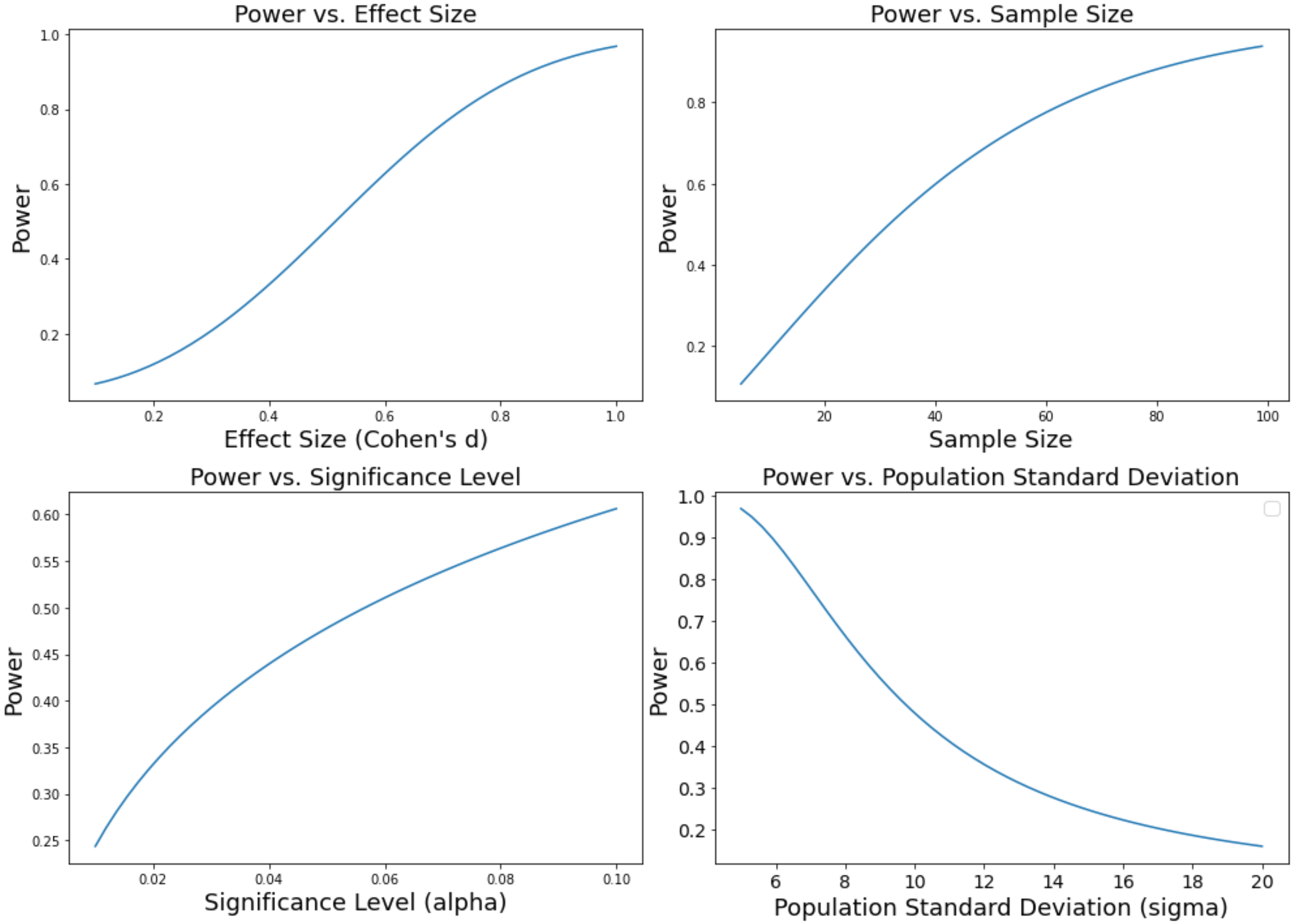
- The larger the significance level, α , the higher the power of the test.



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.

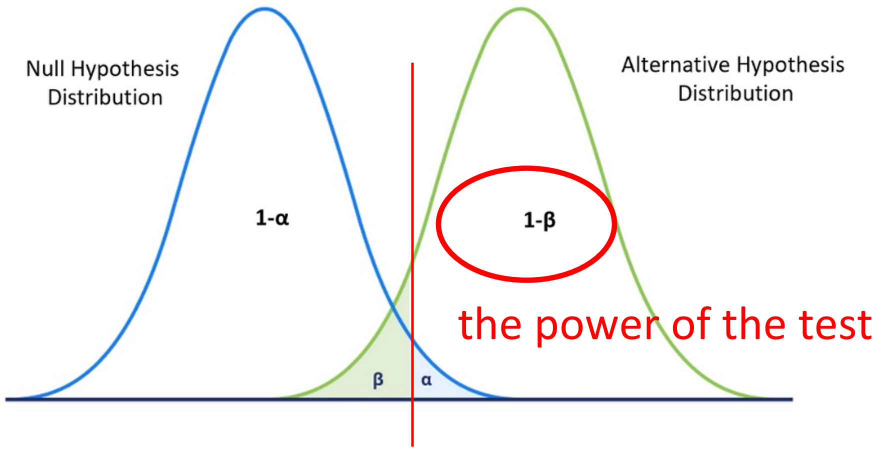


- The larger the sample size, the higher the power of the test.

As the sample size gets larger, the z value increases therefore we will be more likely to reject the null hypothesis.

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

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

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 μ = 10.5 or μ = 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.

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 μ = 10.5 or μ = 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, $H_1 : \mu < 12$.

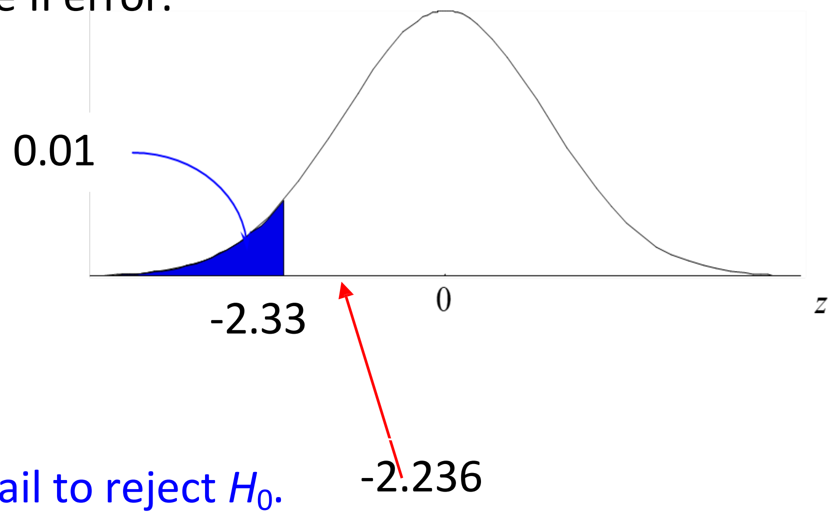
b) Find β, the probability of the type II error.

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 do not reject region, we fail to reject H_0.



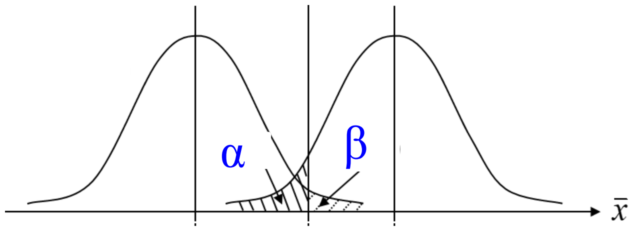
Example

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 μ = 10.5 or μ = 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, $H_1 : \mu < 12$.

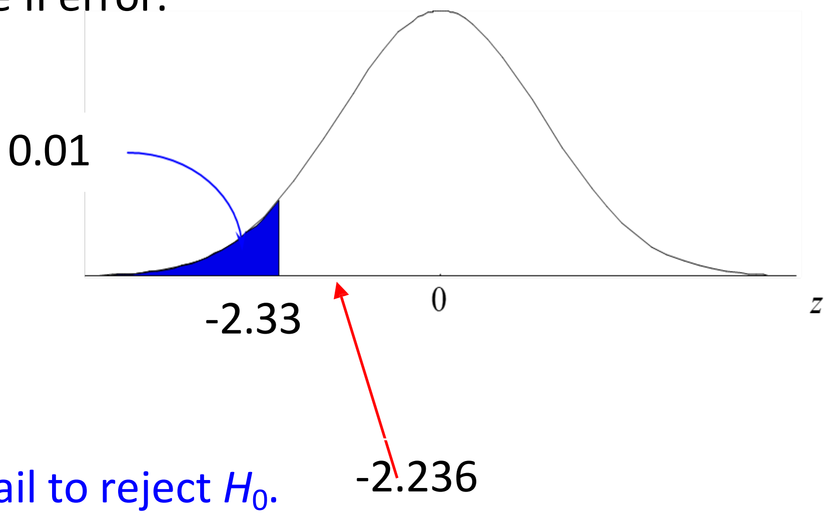
b) Find β, the probability of the type II error.



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 do not reject region, we fail to reject H_0.

b) If H_0 is true, the critical value $\bar{x}$ of is

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

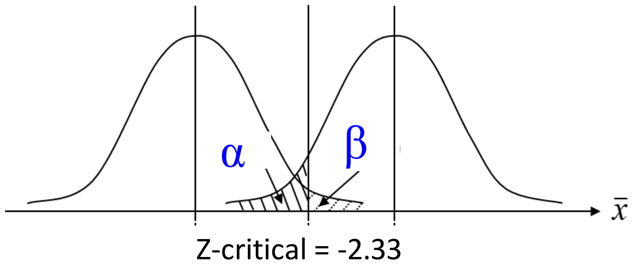
Example

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,

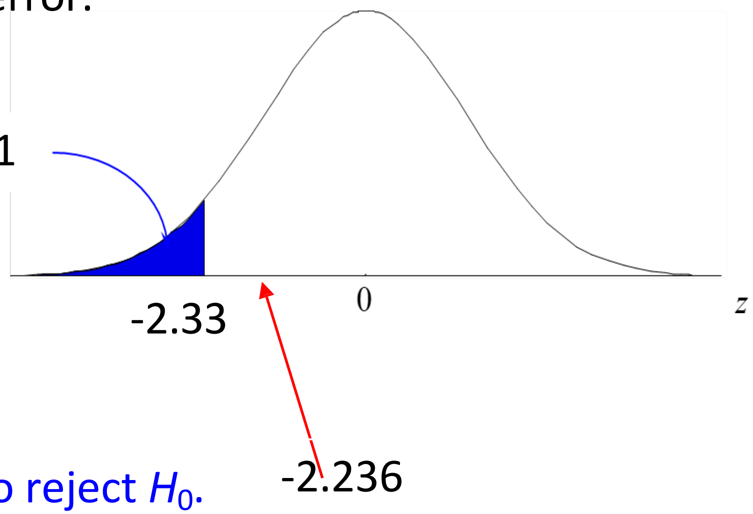
$H_0 : \mu = 12$ against the alternative hypothesis, $H_1 : \mu < 12$.

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 do not reject region, we fail to reject H_0.

b) If H_0 is true, the critical value $\bar{x}$ of is

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

If H_1 is true,

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

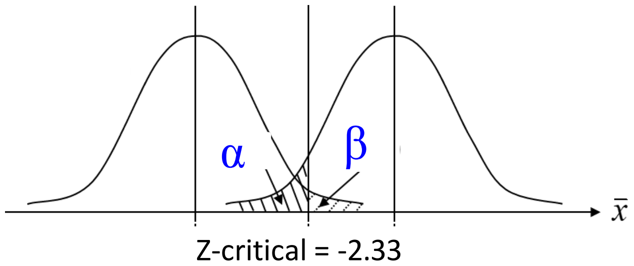
Example

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 μ = 10.5 or μ = 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, $H_1 : \mu < 12$.

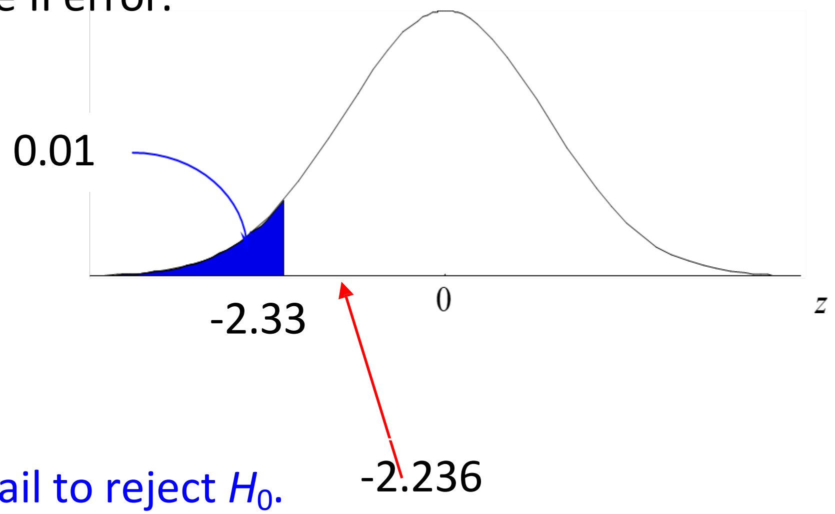
b) Find β, the probability of the type II error.



a) $H_0 : \mu = 12$,

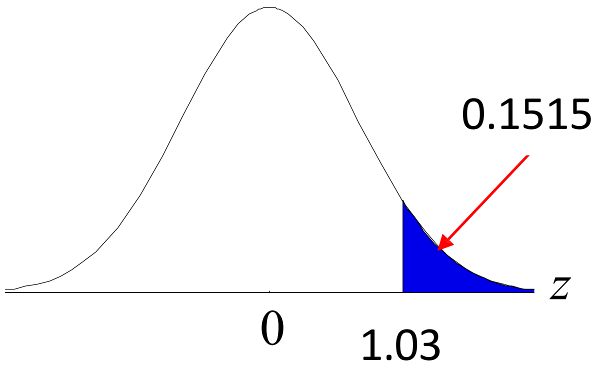
$H_1 : \mu < 12$.

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



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

As z^* is in the do not reject region, we fail to reject H_0.



b) If H_0 is true, the critical value $\bar{x}$ of is

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

If H_1 is true,

Check z table P(Z<1.03)= 0.84849

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

Then, $\beta = P(z > 1.03) = 1 - 0.84849 = 0.1515$.

Summary of Key Concepts and Formulas

TERM OR FORMULA	COMMENT
Hypothesis	A claim about the value of a population characteristic.
Null hypothesis, H_0	The hypothesis initially assumed to be true. It has the form H_0: population characteristic = hypothesized value.
Alternative hypothesis, H_a	A hypothesis that specifies a claim that is contradictory to H_0 and is judged the more plausible claim when H_0 is rejected.
Type I error	Rejecting H_0 when H_0 is true; the probability of a Type I error is denoted by α and is referred to as the significance level for the test.
Type II error	Not rejecting H_0 when H_0 is false; the probability of a Type II error is denoted by β.
Test statistic	A value computed from sample data that is then used as the basis for making a decision between H_0 and H_a.

TERM OR FORMULA	COMMENT
P-value	The probability, computed assuming H_0 to be true, of obtaining a value of the test statistic at least as contradictory to H_0 as what actually resulted. H_0 is rejected if P-value $\leq \alpha$ and not rejected if P-value $> \alpha$, where α is the chosen significance level.
$z = \dfrac{\bar{x} - \text{hypothesized value}}{\dfrac{\sigma}{\sqrt{n}}}$	A test statistic for testing H_0: μ = hypothesized value when σ is known and either the population distribution is normal or the sample size is large. The P-value is determined as an area under the z curve.
$t = \dfrac{\bar{x} - \text{hypothesized value}}{\dfrac{s}{\sqrt{n}}}$	A test statistic for testing H_0: μ = hypothesized value when σ is unknown and either the population distribution is normal or the sample size is large. The P-value is determined from the t curve with df $= n - 1$.
Power	The power of a test is the probability of rejecting the null hypothesis. Power is affected by the size of the difference between the hypothesized value and the actual value, the sample size, and the significance level.

L08: Comparing Two Populations or Treatments

Many investigations are carried out to compare two populations or treatments.

- The efficiency of solar panels differs between two geographical locations.
- Energy consumption differs significantly between green buildings and conventional buildings.
- Air quality between two different locations.
- Implementing a specific air quality policy reduces pollution levels in an urban area.

Comparing Two Populations or Treatments

- When comparing the means of two populations, consider the difference between their means: $\mu_1 - \mu_2$.

- Inferences based on $\overline{x_1} - \overline{x_2}$

    - Point Estimate and Confidence Interval

    - Hypothesis Testing

Independent and Dependent Samples

Before developing inferential procedures concerning $\mu_1 - \mu_2$, we must consider how the two samples, one from each population, are selected.

Independent and Dependent Samples

Source:

Can be a person, an object, or anything that yields a piece of data.

Dependent Sampling:

The same set of sources or related sets are used to obtain the data representing both populations.

Independent Sampling:

Two unrelated sets or sources are used, one set from each population.

Example

The number of items produced by two different assembly lines is to be compared. 25 days are randomly selected for assembly line 1 and the number of items produced on each day is recorded. 40 randomly selected days for assembly line 2 are selected and the number of items produced on each day is recorded.

This sampling plan illustrates independent sampling. The sources (assembly lines) used for each sample (line 1 and 2) were selected separately.

Dependent Samples

- When dependent samples are involved, the data is **paired data**.

Two samples are dependent (or paired) if the members of one sample are related to the members of the other sample in some way.

This typically occurs in "**before and after**" scenarios, matched pairs, or repeated measures on the same subjects.

Is this an example of paired samples?

An engineering association wants to see if there is a difference in the mean annual salary for electrical engineers and chemical engineers. A random sample of electrical engineers is surveyed about their annual income. Another random sample of chemical engineers is surveyed about their annual income.

No, there is no pairing of individuals, you have two **independent** samples

Is this an example of paired samples?

A pharmaceutical company wants to test its new weight-loss drug. Before giving the drug to volunteers, company researchers weigh each person. After a month of using the drug, each person's weight is measured again.

Yes, you have two observations on each individual, resulting in paired data.

Inferences concerning the Mean Difference using Two Dependent Samples

Paired difference: $d = x_1 - x_2$

1. Using the differences *removes* the dependence in the data.
2. Also removes the effect of otherwise uncontrolled factors.
3. The difference between the two population means, when dependent samples are used, is equivalent to the mean of the paired differences.
4. An inference about the mean of the paired differences is an inference about the difference of two means.
5. The mean of the sample paired differences is used as a point estimate for these inferences.

Inferences concerning the Mean Difference using Two Dependent Samples

Sampling Distribution of $\overline{d}$

1. When paired observations are randomly selected from normal populations, the paired difference, $d = x_1 - x_2$, will be approximately normally distributed about a mean μ_d with a standard deviation σ_d.

Inferences concerning the Mean Difference using Two Dependent Samples

Sampling Distribution of $\bar{d}$

1. When paired observations are randomly selected from normal populations, the paired difference, $d = x_1 - x_2$, will be approximately normally distributed about a mean μ_d with a standard deviation σ_d.
2. Use a *t*-test for one mean: make an inference about an unknown mean (μ_d) where d has an approximately normal distribution with an unknown standard deviation (σ_d).
3. Inferences are based on a sample of n dependent pairs of data and the t distribution with $n - 1$ degrees of freedom.

Inferences concerning the Mean Difference using Two Dependent Samples

Sampling Distribution of $\bar{d}$

1. When paired observations are randomly selected from normal populations, the paired difference, $d = x_1 - x_2$, will be approximately normally distributed about a mean μ_d with a standard deviation σ_d.
2. Use a *t*-test for one mean: make an inference about an unknown mean (μ_d) where d has an approximately normal distribution with an unknown standard deviation (σ_d).
3. Inferences are based on a sample of n dependent pairs of data and the t distribution with $n - 1$ degrees of freedom.

The assumption for inferences about the mean of paired differences (μ_d): The paired data are randomly selected from normally distributed populations.

The One-Sample Confidence Interval for μ

Recall

The general formula for a confidence interval for a population mean $\boldsymbol{\mu}$ is

$$\bar{x} \pm (\text{z}\, critical\ value)\frac{\sigma}{\sqrt{n}}$$

When σ is known

$$\bar{x} \pm (t\ critical\ value)\frac{s}{\sqrt{n}}$$

When σ is unknown, we use the sample standard deviation ***s*** to estimate σ.

Inferences concerning the Mean Difference using Two Dependent Samples

Paired difference: $d = x_1 - x_2$

Confidence Interval:

The confidence interval for estimating the mean difference μ_d is found using the formula:

$$\bar{d} \pm (t\ critical\ value)\frac{s_d}{\sqrt{n}}$$

Where $\bar{d}$ is the mean of the sample differences: $\bar{d} = \frac{\sum d}{n}$

and s_d is the standard deviation of the sample differences:

$$s_d = \sqrt{\frac{\sum d^2 - \left[\frac{(\sum d)^2}{n}\right]}{n-1}}$$

Example

Salt-free diets are often prescribed for people with high blood pressure. The following data was obtained from an experiment designed to estimate the reduction in diastolic blood pressure as a result of following a salt-free diet for two weeks. Assume diastolic readings to be normally distributed.

Before	93	106	87	92	102	95	88	110
After	92	102	89	92	101	96	88	105
Difference	1	4	-2	0	1	-1	0	5

Find a 99% confidence interval for the mean reduction.

CityU

Example

Salt-free diets are often prescribed for people with high blood pressure. The following data was obtained from an experiment designed to estimate the reduction in diastolic blood pressure as a result of following a salt-free diet for two weeks. Assume diastolic readings to be normally distributed.

Before	93	106	87	92	102	95	88	110
After	92	102	89	92	101	96	88	105
Difference	1	4	-2	0	1	-1	0	5

Find a 99% confidence interval for the mean reduction.

Solution:

1. Population Parameter of Interest: The mean reduction (difference) in diastolic blood pressure.
2. The Confidence Interval Criteria:
   a) Assumptions: Both sample populations are assumed normal.
   b) Test statistic: t with df = 8 - 1 = 7.
   c) Confidence level: $1 - \alpha = 0.99$

Example

Salt-free diets are often prescribed for people with high blood pressure. The following data was obtained from an experiment designed to estimate the reduction in diastolic blood pressure as a result of following a salt-free diet for two weeks. Assume diastolic readings to be normally distributed.

Before	93	106	87	92	102	95	88	110
After	92	102	89	92	101	96	88	105
Difference	1	4	-2	0	1	-1	0	5

Find a 99% confidence interval for the mean reduction.

Solution:

1. Population Parameter of Interest: The mean reduction (difference) in diastolic blood pressure.
2. The Confidence Interval Criteria:
   a) Assumptions: Both sample populations are assumed normal.
   b) Test statistic: *t* with df = 8 - 1 = 7.
   c) Confidence level: 1 - α = 0.99
3. Sample evidence:

Sample information: $n = 8,\quad \overline{d} = 1.0,\quad \text{and}\quad s_d = 2.39$

Example

Salt-free diets are often prescribed for people with high blood pressure. The following data was obtained from an experiment designed to estimate the reduction in diastolic blood pressure as a result of following a salt-free diet for two weeks. Assume diastolic readings to be normally distributed.

Before	93	106	87	92	102	95	88	110
After	92	102	89	92	101	96	88	105
Difference	1	4	-2	0	1	-1	0	5

Find a 99% confidence interval for the mean reduction.

Solution:

1. Population Parameter of Interest: The mean reduction (difference) in diastolic blood pressure.
2. The Confidence Interval Criteria:
   a) Assumptions: Both sample populations are assumed normal.
   b) Test statistic: t with df = 8 - 1 = 7.
   c) Confidence level: 1 - α = 0.99
3. Sample evidence:

Sample information: $n = 8, \quad \bar{d} = 1.0, \quad \text{and} \quad s_d = 2.39$

4. The Confidence Interval:

a) Confidence coefficients:

Two-tailed situation, $\alpha/2 = 0.005$

$t(\text{df}, \alpha/2) = t(7, 0.005) = 3.499$

b) Maximum error:

$$E = t\left(\text{df}, \frac{\alpha}{2}\right)\frac{s_d}{\sqrt{n}} = 3.499\left(\frac{2.39}{\sqrt{8}}\right) = (3.499)(0.845) = 2.957$$

***t* Table**

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01
df									
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355

Example

Salt-free diets are often prescribed for people with high blood pressure. The following data was obtained from an experiment designed to estimate the reduction in diastolic blood pressure as a result of following a salt-free diet for two weeks. Assume diastolic readings to be normally distributed.

Before	93	106	87	92	102	95	88	110
After	92	102	89	92	101	96	88	105
Difference	1	4	-2	0	1	-1	0	5

Find a 99% confidence interval for the mean reduction.

Solution:

1. Population Parameter of Interest: The mean reduction (difference) in diastolic blood pressure.
2. The Confidence Interval Criteria:
   a) Assumptions: Both sample populations are assumed normal.
   b) Test statistic: t with df = 8 - 1 = 7.
   c) Confidence level: 1 - α = 0.99
3. Sample evidence:

Sample information: $n = 8, \quad \bar{d} = 1.0, \quad \text{and} \quad s_d = 2.39$

4. The Confidence Interval:

a) Confidence coefficients:

Two-tailed situation, α/2 = 0.005

t(df, α /2) = t(7, 0.005) = 3.499

b) Maximum error:

$$E = t\left(\text{df}, \frac{\alpha}{2}\right)\frac{s_d}{\sqrt{n}} = 3.499\left(\frac{2.39}{\sqrt{8}}\right) = (3.499)(0.845) = 2.957$$

c) Confidence limits:

$$\bar{d} - E \quad \text{to} \quad \bar{d} + E$$

$$1.0 - 2.957 \quad \text{to} \quad 1.0 + 2.957$$

$$-1.957 \quad \text{to} \quad 3.957$$

5. The Results: -1.957 to 3.957 is the 99% confidence interval for μ_d.

***t* Table**

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01
df									
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355

Example

In environmental study, an important procedure is the analysis of standard specimens that contain known amounts of a substance. These specimens are usually introduced into the laboratory routine in a way that keeps the analysts blind to the identity of the sample. Often the analyst is blind to the fact that quality assurance samples are included in the assigned work. In this example, the analysts were asked to measure the dissolved oxygen (DO) concentration of the same specimen using two different methods.

Example

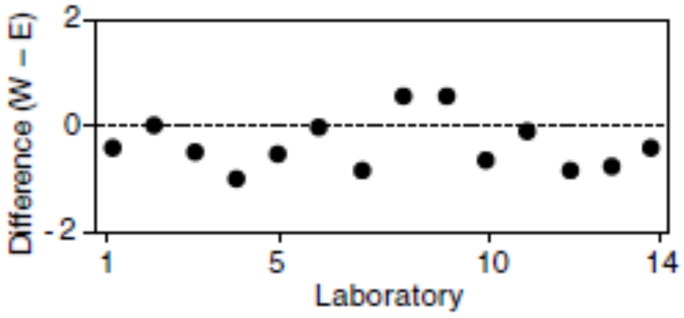
In environmental study, an important procedure is the analysis of standard specimens that contain known amounts of a substance. These specimens are usually introduced into the laboratory routine in a way that keeps the analysts blind to the identity of the sample. Often the analyst is blind to the fact that quality assurance samples are included in the assigned work. In this example, the analysts were asked to measure the dissolved oxygen (DO) concentration of the same specimen using two different methods.

Fourteen laboratories were sent a test solution that was prepared to have a low dissolved oxygen concentration (1.2 mg/L). Each laboratory made the measurements using the Winkler method (a titration) and the electrode method. The question is whether the two methods predict different DO concentrations.

Table below shows the data (Wilcock et al., 1981). The observations for each method may be assumed random and independent as a result of the way the test was designed. The differences plotted in Figure below suggest that the Winkler method may give DO measurements that are slightly lower than the electrode method.

Laboratory	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Winkler	1.2	1.4	1.4	1.3	1.2	1.3	1.4	2	1.9	1.1	1.8	1	1.1	1.4
Electrode	1.6	1.4	1.9	2.3	1.7	1.3	2.2	1.4	1.3	1.7	1.9	1.8	1.8	1.8
Diff. (W – E)	-0.4	0	-0.5	-1	-0.5	0	-0.8	0.6	0.6	-0.6	-0.1	-0.8	-0.7	-0.4

Determine the 95% confidence interval of the true difference is.



Example

Solution:

The differences were calculated by subtracting the electrode measurements from the Winkler measurements. The average of the paired differences is:

$$\bar{d} = \frac{(-0.4) + (0) + (-0.5) + \cdots + (-0.4)}{14} = -0.329\ \mathrm{mg/L}$$

and the s_d of the paired differences is:

$$s_d = \sqrt{\frac{[-0.4-(-0.329)]^2 + [0-(-0.329)]^2 + [-0.5-(-0.329)]^2 + \cdots + [-0.4-(-0.329)]^2}{14-1}} = 0.494$$

Example

Solution:

The differences were calculated by subtracting the electrode measurements from the Winkler measurements. The average of the paired differences is:

$$\bar{d} = \frac{(-0.4) + (0) + (-0.5) + \cdots + (-0.4)}{14} = -0.329 \text{ mg/L}$$

and the s_d of the paired differences is:

$$s_d = \sqrt{\frac{[-0.4-(-0.329)]^2 + [0-(-0.329)]^2 + [-0.5-(-0.329)]^2 + \cdots + [-0.4-(-0.329)]^2}{14-1}} = 0.494$$

The (1 –α)100% confidence interval is computed using the t distribution with 13 degrees of freedom at the α /2 probability point. For (1 − α) = 0.95, $t_{13,0.025}$= 2.160, and the 95% confidence interval of the mean difference μ_d is:

$$\bar{d} - t_{13,0.025}\frac{s_d}{\sqrt{n}} < \mu_d < \bar{d} + t_{13,0.025}\frac{s_d}{\sqrt{n}}$$

***t* Table**

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	$\boxed{0.05}$	0.02	0.01
df									
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055
$\boxed{13}$	0.000	0.694	0.870	1.079	1.350	1.771	$\boxed{2.160}$	2.650	3.012
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977

Example

Solution:

The differences were calculated by subtracting the electrode measurements from the Winkler measurements. The average of the paired differences is:

$$\bar{d} = \frac{(-0.4) + (0) + (-0.5) + \cdots + (-0.4)}{14} = -0.329 \text{ mg/L}$$

and the s_d of the paired differences is:

$$s_d = \sqrt{\frac{[-0.4-(-0.329)]^2 + [0-(-0.329)]^2 + [-0.5-(-0.329)]^2 + \cdots + [-0.4-(-0.329)]^2}{14-1}} = 0.494$$

The $(1-\alpha)100\%$ confidence interval is computed using the t distribution with 13 degrees of freedom at the $\alpha/2$ probability point. For $(1-\alpha) = 0.95$, $t_{13,0.025} = 2.160$, and the 95% confidence interval of the mean difference μ_d is:

$$\bar{d} - t_{13,0.025}\frac{s_d}{\sqrt{n}} < \mu_d < \bar{d} + t_{13,0.025}\frac{s_d}{\sqrt{n}}$$

For the particular values of this example:

$$-0.329 - 2.160\frac{0.494}{\sqrt{14}} < \mu_d < -0.329 + 2.160\frac{0.494}{\sqrt{14}}$$

$$-0.614 < \mu_d < -0.044$$

We are highly confident that the difference between the two methods is not zero because the confidence interval does not include the difference of zero. The methods give different results and, furthermore, the electrode method has given higher readings than the Winkler method.

***t* Table**

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01
df									
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977

Hypothesis test for Comparing Two Population

Null Hypothesis: H_0: μ_d = hypothesized value

The hypothesized value is usually 0 – meaning that there is no difference.

Where μ_d is the mean of the differences in the paired observations

Alternative Hypothesis:

H_a: $\mu_d >$ hypothesized value

H_a: $\mu_d <$ hypothesized value

H_a: $\mu_d \neq$ hypothesized value

Hypothesis test for Comparing Two Population

Null Hypothesis: H_0: μ_d = hypothesized value

The hypothesized value is usually 0 – meaning that there is no difference.

Where μ_d is the mean of the differences in the paired observations

Test Statistic:

$$t^* = \frac{\bar{x}_d - \text{hypothesized value}}{s_d / \sqrt{n}}$$

Where n is the number of sample differences and $\bar{x}_d$ and s_d are the mean and standard deviation of the sample differences.

This test is based on df = $n - 1$.

Alternative Hypothesis:

H_a: $\mu_d >$ hypothesized value

H_a: $\mu_d <$ hypothesized value

H_a: $\mu_d \neq$ hypothesized value

Hypothesis test for Comparing Two Population

Null Hypothesis: H_0: μ_d = hypothesized value

Where μ_d is the mean of the differences in the paired observations

The hypothesized value is usually 0 - meaning that there is no difference.

Test Statistic:

$$t^* = \frac{\bar{x}_d - \text{hypothesized value}}{s_d / \sqrt{n}}$$

Where n is the number of sample differences and $\bar{x}_d$ and s_d are the mean and standard deviation of the sample differences.
This test is based on df = $n - 1$.

Alternative Hypothesis:

H_a: $\mu_d >$ hypothesized value

H_a: $\mu_d <$ hypothesized value

H_a: $\mu_d \neq$ hypothesized value

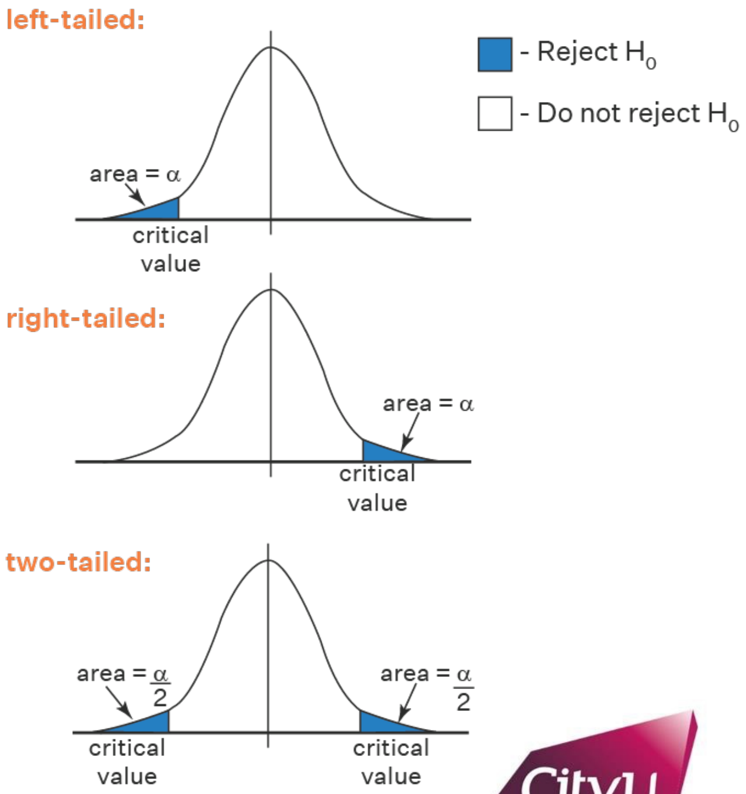
P-value:

Area to the right of calculated t

Area to the left of calculated t

2(area to the right of t) if $+t$ or

2(area to the left of t) if $-t$



Example

The corrosive effects of various chemicals on normal and specially treated pipes were tested by using a dependent sampling plan. The data collected is summarized by

$$n = 17, \quad \bar{d} = 5.7, \quad s_d = 4.8$$

where *d* is the amount of corrosion on the treated pipe subtracted from the amount of corrosion on the normal pipe.

Does this sample provide sufficient evidence to conclude the specially treated pipes are more resistant to corrosion? Use $\alpha = 0.05$.

The corrosive effects of various chemicals on normal and specially treated pipes were tested by using a dependent sampling plan. The data collected is summarized by

$$n = 17, \quad \bar{d} = 5.7, \quad s_d = 4.8$$

where *d* is the amount of corrosion on the treated pipe subtracted from the amount of corrosion on the normal pipe.

Does this sample provide sufficient evidence to conclude the specially treated pipes are more resistant to corrosion? Use α = 0.05.

Solution:

1. The Set-up:

a) Population parameter of concern: The mean difference in corrosion, normal pipe - treated pipe.

b) The null and alternative hypothesis:

H_0: $\mu_d = 0$ (≤) (did not lower corrosion)

H_1: $\mu_d > 0$ (did lower corrosion)

2. The Hypothesis Test Criteria:

a) Assumptions: Assume corrosion measures are approximately normal.

b) Test statistic:

$$t^* = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}}, \qquad \text{where df} = n - 1 = 16$$

c) Level of significance: α = 0.05.

The corrosive effects of various chemicals on normal and specially treated pipes were tested by using a dependent sampling plan. The data collected is summarized by

$$n = 17, \quad \bar{d} = 5.7, \quad s_d = 4.8$$

where *d* is the amount of corrosion on the treated pipe subtracted from the amount of corrosion on the normal pipe.

Does this sample provide sufficient evidence to conclude the specially treated pipes are more resistant to corrosion? Use α = 0.05.

Solution:

1. The Set-up:

a) Population parameter of concern: The mean difference in corrosion, normal pipe - treated pipe.

b) The null and alternative hypothesis:

H_0: $\mu_d = 0$ (≤) (did not lower corrosion)

H_1: $\mu_d > 0$ (did lower corrosion)

2. The Hypothesis Test Criteria:

a) Assumptions: Assume corrosion measures are approximately normal.

b) Test statistic:

$$t^* = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}}, \qquad \text{where df} = n - 1 = 16$$

c) Level of significance: α = 0.05.

3. The Sample Evidence:

a) Sample information:

$$n = 17, \quad \bar{d} = 5.7, \quad s_d = 4.8$$

b) Calculate the value of the test statistic:

$$t^* = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{5.7 - 0.0}{4.8 / \sqrt{17}} = \frac{5.7}{1.164} = 4.896$$

The corrosive effects of various chemicals on normal and specially treated pipes were tested by using a dependent sampling plan. The data collected is summarized by

$$n = 17, \qquad \bar{d} = 5.7, \qquad s_d = 4.8$$

where d is the amount of corrosion on the treated pipe subtracted from the amount of corrosion on the normal pipe.

Does this sample provide sufficient evidence to conclude the specially treated pipes are more resistant to corrosion? Use α = 0.05.

Solution:

1. The Set-up:

a) Population parameter of concern: The mean difference in corrosion, normal pipe - treated pipe.

b) The null and alternative hypothesis:

H_0: $\mu_d = 0$ (≤) (did not lower corrosion)

H_1: $\mu_d > 0$ (did lower corrosion)

2. The Hypothesis Test Criteria:

a) Assumptions: Assume corrosion measures are approximately normal.

b) Test statistic:

$$t^* = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}}, \qquad \text{where df} = n - 1 = 16$$

c) Level of significance: α = 0.05.

3. The Sample Evidence:

a) Sample information:

$$n = 17, \qquad \bar{d} = 5.7, \qquad s_d = 4.8$$

b) Calculate the value of the test statistic:

$$t^* = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{5.7 - 0.0}{4.8 / \sqrt{17}} = \frac{5.7}{1.164} = 4.896$$

4. The Probability Distribution (Classical Approach):

a) Critical value: $t(16, 0.05) = 1.746$

b) t^* is in the critical region.

One-sides problem

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05
two-tails	1.00	0.50	0.40	0.30	0.20	0.10
df						
1	0.000	1.000	1.376	1.963	3.078	6.314
2	0.000	0.816	1.061	1.386	1.886	2.920
3	0.000	0.765	0.978	1.250	1.638	2.353
4	0.000	0.741	0.941	1.190	1.533	2.132
5	0.000	0.727	0.920	1.156	1.476	2.015
6	0.000	0.718	0.906	1.134	1.440	1.943
7	0.000	0.711	0.896	1.119	1.415	1.895
8	0.000	0.706	0.889	1.108	1.397	1.860
9	0.000	0.703	0.883	1.100	1.383	1.833
10	0.000	0.700	0.879	1.093	1.372	1.812
11	0.000	0.697	0.876	1.088	1.363	1.796
12	0.000	0.695	0.873	1.083	1.356	1.782
13	0.000	0.694	0.870	1.079	1.350	1.771
14	0.000	0.692	0.868	1.076	1.345	1.761
15	0.000	0.691	0.866	1.074	1.341	1.753
$\boxed{16}$	0.000	0.690	0.865	1.071	1.337	$\boxed{1.746}$
17	0.000	0.689	0.863	1.069	1.333	1.740

The corrosive effects of various chemicals on normal and specially treated pipes were tested by using a dependent sampling plan. The data collected is summarized by

$$n = 17, \quad \bar{d} = 5.7, \quad s_d = 4.8$$

where d is the amount of corrosion on the treated pipe subtracted from the amount of corrosion on the normal pipe.

Does this sample provide sufficient evidence to conclude the specially treated pipes are more resistant to corrosion? Use $\alpha = 0.05$.

Solution:

1. The Set-up:

a) Population parameter of concern: The mean difference in corrosion, normal pipe - treated pipe.

b) The null and alternative hypothesis:

H_0: $\mu_d = 0$ (≤) (did not lower corrosion)

H_1: $\mu_d > 0$ (did lower corrosion)

2. The Hypothesis Test Criteria:

a) Assumptions: Assume corrosion measures are approximately normal.

b) Test statistic:

$$t^* = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}}, \qquad \text{where df} = n - 1 = 16$$

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

3. The Sample Evidence:

a) Sample information:

$$n = 17, \quad \bar{d} = 5.7, \quad s_d = 4.8$$

b) Calculate the value of the test statistic:

$$t^* = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{5.7 - 0.0}{4.8 / \sqrt{17}} = \frac{5.7}{1.164} = 4.896$$

One-sides problem

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05
two-tails	1.00	0.50	0.40	0.30	0.20	0.10
df						
1	0.000	1.000	1.376	1.963	3.078	6.314
2	0.000	0.816	1.061	1.386	1.886	2.920
3	0.000	0.765	0.978	1.250	1.638	2.353
4	0.000	0.741	0.941	1.190	1.533	2.132
5	0.000	0.727	0.920	1.156	1.476	2.015
6	0.000	0.718	0.906	1.134	1.440	1.943
7	0.000	0.711	0.896	1.119	1.415	1.895
8	0.000	0.706	0.889	1.108	1.397	1.860
9	0.000	0.703	0.883	1.100	1.383	1.833
10	0.000	0.700	0.879	1.093	1.372	1.812
11	0.000	0.697	0.876	1.088	1.363	1.796
12	0.000	0.695	0.873	1.083	1.356	1.782
13	0.000	0.694	0.870	1.079	1.350	1.771
14	0.000	0.692	0.868	1.076	1.345	1.761
15	0.000	0.691	0.866	1.074	1.341	1.753
16	0.000	0.690	0.865	1.071	1.337	1.746
17	0.000	0.689	0.863	1.069	1.333	1.740

4. The Probability Distribution (Classical Approach):

a) Critical value: $t(16, 0.05) = 1.746$

b) t^* is in the rejection region.

5. The Results:

a) Decision: Reject H_0.

b) Conclusion: At the 0.05 level of significance, there is evidence to suggest the treated pipes do not corrode as much as the normal pipes when subjected to chemicals.

Example : Native Shrubs at Reclaimed and Reference Sites

Try it out!

A mining company has paid a bond to a government agency to guarantee the successful reclamation of a strip mining site. Having carried out the necessary work, the company wants the bond released. However, the agency requires the company to provide evidence that the mined site is equivalent to an untouched control site with respect to the density of native shrubs.

A consultant has designed and carried out a study that involved randomly selecting eight plots from the treated site and matching them up on the basis of slope, aspect, and soil type with eight plots from the control site. The densities of native shrubs that were obtained are shown in below Table (95% confidence interval).

Comparison between the Vegetation Density on Eight Paired Plots from an Undamaged Control Site and a Site Where Mining Has Occurred

Plot pair	1	2	3	4	5	6	7	8
Control site	0.94	1.02	0.80	0.89	0.88	0.76	0.71	0.75
Mined site	0.75	0.94	1.01	0.67	0.75	0.88	0.53	0.89
Difference	0.19	0.08	−0.21	0.22	0.13	−0.10	0.18	−0.14

Example : Native Shrubs at Reclaimed and Reference Sites

A mining company has paid a bond to a government agency to guarantee the successful reclamation of a strip mining site. Having carried out the necessary work, the company wants the bond released. However, the agency requires the company to provide evidence that the mined site is equivalent to an untouched control site with respect to the density of native shrubs.

A consultant has designed and carried out a study that involved randomly selecting eight plots from the treated site and matching them up on the basis of slope, aspect, and soil type with eight plots from the control site. The densities of native shrubs that were obtained are shown in below Table (95% confidence interval).

Solution:

Population parameter of concern: control site, mined site

The null and alternative hypothesis:

H_0: $\mu_d = 0$ (≤) (mean density of native shrubs is the same on paired plots at the two sites)

H_a: $\mu_d > 0$ (density is higher on the control site)

Test Statistic: $t^* = \dfrac{\bar{d} - \mu_d}{s_d / \sqrt{n}}$

Level of Significance: α = 0.05

Comparison between the Vegetation Density on Eight Paired Plots from an Undamaged Control Site and a Site Where Mining Has Occurred

Plot pair	1	2	3	4	5	6	7	8
Control site	0.94	1.02	0.80	0.89	0.88	0.76	0.71	0.75
Mined site	0.75	0.94	1.01	0.67	0.75	0.88	0.53	0.89
Difference	0.19	0.08	−0.21	0.22	0.13	−0.10	0.18	−0.14

Example : Native Shrubs at Reclaimed and Reference Sites

A mining company has paid a bond to a government agency to guarantee the successful reclamation of a strip mining site. Having carried out the necessary work, the company wants the bond released. However, the agency requires the company to provide evidence that the mined site is equivalent to an untouched control site with respect to the density of native shrubs.

A consultant has designed and carried out a study that involved randomly selecting eight plots from the treated site and matching them up on the basis of slope, aspect, and soil type with eight plots from the control site. The densities of native shrubs that were obtained are shown in below Table (95% confidence interval).

Solution:

Population parameter of concern: control site, mined site

The null and alternative hypothesis:

H_0: $\mu_d = 0$ (≤) (mean density of native shrubs is the same on paired plots at the two sites)

H_a: $\mu_d > 0$ (density is higher on the control site)

Test Statistic: $t^* = \dfrac{\bar{d} - \mu_d}{s_d / \sqrt{n}}$

Level of Significance: α = 0.05

Sample Information:

$n = 8, \bar{d} = 0.041, s_d = 0.171$; df = 8-1 = 7

Calculate the value of the test statistic:

$$t^* = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{0.041 - 0}{0.171/\sqrt{8}} = 0.68$$

Comparison between the Vegetation Density on Eight Paired Plots from an Undamaged Control Site and a Site Where Mining Has Occurred

Plot pair	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Control site	0.94	1.02	0.80	0.89	0.88	0.76	0.71	0.75
Mined site	0.75	0.94	1.01	0.67	0.75	0.88	0.53	0.89
Difference	0.19	0.08	–0.21	0.22	0.13	–0.10	0.18	–0.14

CityU

Example : Native Shrubs at Reclaimed and Reference Sites

A mining company has paid a bond to a government agency to guarantee the successful reclamation of a strip mining site. Having carried out the necessary work, the company wants the bond released. However, the agency requires the company to provide evidence that the mined site is equivalent to an untouched control site with respect to the density of native shrubs.

A consultant has designed and carried out a study that involved randomly selecting eight plots from the treated site and matching them up on the basis of slope, aspect, and soil type with eight plots from the control site. The densities of native shrubs that were obtained are shown in below Table (95% confidence interval).

One-sides problem

Comparison between the Vegetation Density on Eight Paired Plots from an Undamaged Control Site and a Site Where Mining Has Occurred

Plot pair	1	2	3	4	5	6	7	8
Control site	0.94	1.02	0.80	0.89	0.88	0.76	0.71	0.75
Mined site	0.75	0.94	1.01	0.67	0.75	0.88	0.53	0.89
Difference	0.19	0.08	−0.21	0.22	0.13	−0.10	0.18	−0.14

Solution:

Population parameter of concern: control site, mined site

The null and alternative hypothesis:

H_0: $\mu_d = 0$ (≤) (mean density of native shrubs is the same on paired plots at the two sites)

H_a: $\mu_d > 0$ (density is higher on the control site)

Test Statistic: $t^* = \dfrac{\bar{d} - \mu_d}{s_d / \sqrt{n}}$

Level of Significance: $\alpha = 0.05$

Sample Information:

$n = 8, \bar{d} = 0.041, s_d = 0.171$; $df = 8 - 1 = 7$

Calculate the value of the test statistic:

$$t^* = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{0.041 - 0}{0.171/\sqrt{8}} = 0.68$$

The Probability Distribution (Classical Approach):

a) Critical value: $t(7, 0.05) = 1.895$

b) t^* is not in the rejection region.

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05
two-tails	1.00	0.50	0.40	0.30	0.20	0.10
df						
1	0.000	1.000	1.376	1.963	3.078	6.314
2	0.000	0.816	1.061	1.386	1.886	2.920
3	0.000	0.765	0.978	1.250	1.638	2.353
4	0.000	0.741	0.941	1.190	1.533	2.132
5	0.000	0.727	0.920	1.156	1.476	2.015
6	0.000	0.718	0.906	1.134	1.440	1.943
7	0.000	0.711	0.896	1.119	1.415	1.895
8	0.000	0.706	0.889	1.108	1.397	1.860

The Results:

a) Decision: Do not reject H_0.

b) Conclusion: At the 0.05 level of significance, there is no evidence to suggest the shrub density is higher on the control site as compared to the mined site.

Independent and Dependent Samples

Source:

Can be a person, an object, or anything that yields a piece of data.

Dependent Sampling:

The same set of sources or related sets are used to obtain the data representing both populations.

Independent Sampling:

Two unrelated sets or sources are used, one set from each population.

Inferences concerning the Difference between Means using Two Independent Samples

- When comparing the means of two populations, consider the difference between their means: $\mu_1 - \mu_2$.

- Inferences based on $\overline{x_1} - \overline{x_2}$

Varies from sample to sample

→ sampling distribution

Properties of the Sample Distribution of $\bar{x}_1$-$\bar{x}_2$

If the random samples on which $\bar{x}_1$ and $\bar{x}_2$ are based are selected independently of one another, then

1. $\mu_{\bar{x}_1 - \bar{x}_2} = \begin{pmatrix} \text{mean value} \\ \text{of } \bar{x}_1 - \bar{x}_2 \end{pmatrix} = \mu_{\bar{x}_1} - \mu_{\bar{x}_2} = \mu_1 - \mu_2$

   The sampling distribution of $\bar{x}_1 - \bar{x}_2$ is always centered at the value of $\mu_1 - \mu_2$, so $\bar{x}_1 - \bar{x}_2$ is an unbiased statistic for estimating $\mu_1 - \mu_2$.

Properties of the Sample Distribution of $\bar{x}_1$-$\bar{x}_2$

If the random samples on which $\bar{x}_1$ and $\bar{x}_2$ are based are selected independently of one another, then

1. $\mu_{\bar{x}_1 - \bar{x}_2} = \begin{pmatrix} \text{mean value} \\ \text{of } \bar{x}_1 - \bar{x}_2 \end{pmatrix} = \mu_{\bar{x}_1} - \mu_{\bar{x}_2} = \mu_1 - \mu_2$

   The sampling distribution of $\bar{x}_1 - \bar{x}_2$ is always centered at the value of $\mu_1 - \mu_2$, so $\bar{x}_1 - \bar{x}_2$ is an unbiased statistic for estimating $\mu_1 - \mu_2$.

2. $\sigma^2_{\bar{x}_1 - \bar{x}_2} = \begin{pmatrix} \text{variance of} \\ \bar{x}_1 - \bar{x}_2 \end{pmatrix} = \sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2} = \dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}$

   and

   $\sigma_{\bar{x}_1 - \bar{x}_2} = \begin{pmatrix} \text{standard deviation} \\ \text{of } \bar{x}_1 - \bar{x}_2 \end{pmatrix} = \sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$

Properties of the Sample Distribution of $\bar{x}_1$-$\bar{x}_2$

3. If n_1 and n_2 are both large or the population distributions are (at least approximately) normal, $\bar{x}_1$ and $\bar{x}_2$ each have (at least approximately) a normal distribution.

This implies that the sampling distribution of $\bar{x}_1$ - $\bar{x}_2$ is also normal or approximately normal.

→can be standardized to obtain a variable with a sampling distribution that is approximately to the standard normal distribution.

$$z = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

When $\sigma 1$ *and* $\sigma 2$ are known

$$t = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

When $\sigma 1$ *and* $\sigma 2$ *are un*known, we must estimate them using the corresponding sample variances, $s_1{}^2$ and $s_2{}^2$

Degrees of freedom (df)

When conducting an Independent Samples t-test, an important step is to determine the degrees of freedom (df) for the test.

Welch's t-Test: Adjusts for unequal variances (Welch-Satterthwaite Equation)

$$df = \frac{(V_1 + V_2)^2}{\dfrac{V_1^2}{n_1 - 1} + \dfrac{V_2^2}{n_2 - 1}} \text{ where } V_1 = \frac{s_1^2}{n_1} \text{ and } V_2 = \frac{s_2^2}{n_2}$$

The computed value of df should be truncated (rounded down) to obtain an integer value of df.

s_1^2 and s_2^2 and are the sample variances of the two groups

where df is the **smaller** of df_1 or df_2 when computing t^* without the aid of a computer or calculator.

Two-Sample t Test for Comparing Two Populations

Assumptions:

1) The two samples are independently selected random samples from the populations of interest
2) The sample sizes are large (generally 30 or larger) or the population distributions are (at least approximately) normal.

When comparing two treatment groups, use the following assumptions:

1) Individuals or objects are randomly assigned to treatments (or vice versa)
2) The sample sizes are large (generally 30 or larger) or the treatment response distributions are approximately normal.

The Two-Sample t Confidence Interval

The general formula for a confidence interval for : $\mu_1 - \mu_2$ when

1) The two samples are independently selected random samples from the populations of interest
2) The sample sizes are large (generally 30 or larger) or the population distributions are (at least approximately) normal.

is $(\bar{x}_1 - \bar{x}_2) \pm (t \text{ criticalvalue}) \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

$$\text{df} = \frac{(V_1 + V_2)^2}{\frac{V_1^2}{n_1 - 1} + \frac{V_2^2}{n_2 - 1}} \quad \text{where } V_1 = \frac{s_1^2}{n_1} \text{ and } V_2 = \frac{s_2^2}{n_2}$$

where df is the **smaller** of df_1 or df_2 when computing t^* without the aid of a computer or calculator.

Example

A recent study reported the longest average workweeks for non-supervisory employees in private industry to be chef and construction.

Industry	n	Average Hours/Week	Standard Deviation
Chef	18	48.2	6.7
Construction	12	44.1	2.3

Find a 95% confidence interval for the difference in mean length of workweek between chef and construction. Assume normality for the sampled populations and that the samples were selected randomly.

A recent study reported the longest average workweeks for non-supervisory employees in private industry to be chef and construction.

Industry	n	Average Hours/Week	Standard Deviation
Chef	18	48.2	6.7
Construction	12	44.1	2.3

Find a 95% confidence interval for the difference in mean length of workweek between chef and construction. Assume normality for the sampled populations and that the samples were selected randomly.

Solution:

1. Parameter of interest: The difference between the mean hours/week for chefs and the mean hours/week for construction workers, $\mu_1 - \mu_2$

A recent study reported the longest average workweeks for non-supervisory employees in private industry to be chef and construction.

Industry	n	Average Hours/Week	Standard Deviation
Chef	18	48.2	6.7
Construction	12	44.1	2.3

Find a 95% confidence interval for the difference in mean length of workweek between chef and construction. Assume normality for the sampled populations and that the samples were selected randomly.

Solution:

1. Parameter of interest: The difference between the mean hours/week for chefs and the mean hours/week for construction workers, $\mu_1 - \mu_2$
2. The Confidence Interval Criteria:

a) Assumptions: Both populations are assumed normal and the samples were random and independently selected.

b) Test statistic: t with df = 11; (choose a smaller one)

the smaller of $n_1 - 1 = 18 - 1 = 17$ or $n_2 - 1 = 12 - 1 = 11$.

c) Confidence level: $1 - \alpha = 0.95$.

A recent study reported the longest average workweeks for non-supervisory employees in private industry to be chef and construction.

Industry	n	Average Hours/Week	Standard Deviation
Chef	18	48.2	6.7
Construction	12	44.1	2.3

Find a 95% confidence interval for the difference in mean length of workweek between chef and construction. Assume normality for the sampled populations and that the samples were selected randomly.

Solution:

1. Parameter of interest: The difference between the mean hours/week for chefs and the mean hours/week for construction workers, $\mu_1 - \mu_2$
2. The Confidence Interval Criteria:

a) Assumptions: Both populations are assumed normal and the samples were random and independently selected.

b) Test statistic: t with df = 11; (choose a smaller one)

the smaller of $n_1 - 1 = 18 - 1 = 17$ or $n_2 - 1 = 12 - 1 = 11$.

c) Confidence level: $1 - \alpha = 0.95$.

3. The Sample Evidence:

Sample information given in the table.

Point estimate for $\mu_1 - \mu_2$: $\bar{x}_1 - \bar{x}_2 = 48.2 - 44.1 = 4.1$

A recent study reported the longest average workweeks for non-supervisory employees in private industry to be chef and construction.

Industry	n	Average Hours/Week	Standard Deviation
Chef	18	48.2	6.7
Construction	12	44.1	2.3

Find a 95% confidence interval for the difference in mean length of workweek between chef and construction. Assume normality for the sampled populations and that the samples were selected randomly.

Solution:

1. Parameter of interest: The difference between the mean hours/week for chefs and the mean hours/week for construction workers, $\mu_1 - \mu_2$
2. The Confidence Interval Criteria:

a) Assumptions: Both populations are assumed normal and the samples were random and independently selected.

b) Test statistic: t with df = 11; (choose a smaller one)

the smaller of $n_1 - 1 = 18 - 1 = 17$ or $n_2 - 1 = 12 - 1 = 11$.

c) Confidence level: $1 - \alpha = 0.95$.

3. The Sample Evidence:

Sample information given in the table.

Point estimate for $\mu_1 - \mu_2$: $\bar{x}_1 - \bar{x}_2 = 48.2 - 44.1 = 4.1$

4. The Confidence Interval:

a) Confidence coefficients:

$t(\text{df}, \alpha/2) = t(11, 0.025) = 2.201$

Two-sides problem

b) Maximum error:

$$E = t(\text{df}, \alpha/2)\sqrt{\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)}$$

$$= (2.201)\sqrt{\left(\frac{6.7^2}{18}\right) + \left(\frac{2.3^2}{12}\right)} = (2.201)(1.7131) = 3.77$$

***t* Table**

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	$\boxed{0.05}$
df							
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228
$\boxed{11}$	0.000	0.697	0.876	1.088	1.363	1.796	$\boxed{2.201}$
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179

A recent study reported the longest average workweeks for non-supervisory employees in private industry to be chef and construction.

Industry	n	Average Hours/Week	Standard Deviation
Chef	18	48.2	6.7
Construction	12	44.1	2.3

Find a 95% confidence interval for the difference in mean length of workweek between chef and construction. Assume normality for the sampled populations and that the samples were selected randomly.

Solution:

1. Parameter of interest: The difference between the mean hours/week for chefs and the mean hours/week for construction workers, μ_1 - μ_2
2. The Confidence Interval Criteria:

a) Assumptions: Both populations are assumed normal and the samples were random and independently selected.

b) Test statistic: t with df = 11; (choose a smaller one)

the smaller of $n_1 - 1 = 18 - 1 = 17$ or $n_2 - 1 = 12 - 1 = 11$.

c) Confidence level: $1 - \alpha = 0.95$.

3. The Sample Evidence:

Sample information given in the table.

Point estimate for μ_1 - μ_2: $\bar{x}_1 - \bar{x}_2 = 48.2 - 44.1 = 4.1$

4. The Confidence Interval:

Two-sides problem

a) Confidence coefficients:

$t(\text{df}, \alpha/2) = t(11, 0.025) = 2.201$

b) Maximum error:

$$E = t(\text{df}, \alpha/2)\sqrt{\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)}$$

$$= (2.201)\sqrt{\left(\frac{6.7^2}{18}\right) + \left(\frac{2.3^2}{12}\right)} = (2.201)(1.7131) = 3.77$$

c) Confidence limits:

$$(\bar{x}_1 - \bar{x}_2) \pm E = 4.1 \pm 3.77$$

$$4.1 - 3.77 \quad \text{to} \quad 4.1 + 3.77$$

$$0.33 \quad \text{to} \quad 7.87$$

5. The Results: 0.33 to 7.87 is a 95% confidence interval for the difference in mean hours/week for chefs and construction workers.

***t* Table**

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05
df							
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179

Example

Try it out!

Water specimens collected from a residential area that is served by the city water supply are compared with the specimens taken from a residential area that is served by private wells. The mean and variances of the samples are

	n	Mean Mercury Concentration (μg/L)	Variances
City	13	0.157	0.0071
Private	10	0.151	0.0076

Find a 95% confidence interval for the difference in mean mercury concentration between city water supply and private wells. Assume normality for the sampled populations and that the samples were selected randomly.

Example

Water specimens collected from a residential area that is served by the city water supply are compared with the specimens taken from a residential area that is served by private wells. The mean and variances of the samples are

	n	Mean Mercury Concentration (μg/L)	Variances
City	13	0.157	0.0071
Private	10	0.151	0.0076

Find a 95% confidence interval for the difference in mean mercury concentration between city water supply and private wells. Assume normality for the sampled populations and that the samples were selected randomly.

Solution:

1. Parameter of interest: The difference between the mean mercury concentration for city water supply and private wells, $\mu_1 - \mu_2$

Example

Water specimens collected from a residential area that is served by the city water supply are compared with the specimens taken from a residential area that is served by private wells. The mean and variances of the samples are

	n	Mean Mercury Concentration (μg/L)	Variances
City	13	0.157	0.0071
Private	10	0.151	0.0076

Find a 95% confidence interval for the difference in mean mercury concentration between city water supply and private wells. Assume normality for the sampled populations and that the samples were selected randomly.

Solution:

1. Parameter of interest: The difference between the mean mercury concentration for city water supply and private wells, $\mu_1 - \mu_2$
2. The Confidence Interval Criteria:

a) Assumptions: Both populations are assumed normal and the samples were random and independently selected.

b) Test statistic: t with df = 9;

the smaller of $n_1 - 1 = 13 - 1 = 12$ or $n_2 - 1 = 10 - 1 = 9$.

c) Confidence level: $1 - \alpha = 0.95$.

Example

Water specimens collected from a residential area that is served by the city water supply are compared with the specimens taken from a residential area that is served by private wells. The mean and variances of the samples are

	n	Mean Mercury Concentration (µg/L)	Variances
City	13	0.157	0.0071
Private	10	0.151	0.0076

Find a 95% confidence interval for the difference in mean mercury concentration between city water supply and private wells. Assume normality for the sampled populations and that the samples were selected randomly.

Solution:

1. Parameter of interest: The difference between the mean mercury concentration for city water supply and private wells, $\mu_1 - \mu_2$
2. The Confidence Interval Criteria:

a) Assumptions: Both populations are assumed normal and the samples were random and independently selected.

b) Test statistic: t with df = 9;

the smaller of $n_1 - 1 = 13 - 1 = 12$ or $n_2 - 1 = 10 - 1 = 9$.

c) Confidence level: $1 - \alpha = 0.95$.

3. The Sample Evidence:

Sample information given in the table.

Point estimate for $\mu_1 - \mu_2$:

$$\overline{x_1} - \overline{x_2} = 0.157 - 0.151 = 0.006 \text{ µg/L}$$

Example

Water specimens collected from a residential area that is served by the city water supply are compared with the specimens taken from a residential area that is served by private wells. The mean and variances of the samples are

	n	Mean Mercury Concentration (µg/L)	Variances
City	13	0.157	0.0071
Private	10	0.151	0.0076

Find a 95% confidence interval for the difference in mean mercury concentration between city water supply and private wells. Assume normality for the sampled populations and that the samples were selected randomly.

Solution:

1. Parameter of interest: The difference between the mean mercury concentration for city water supply and private wells, $\mu_1 - \mu_2$
2. The Confidence Interval Criteria:

a) Assumptions: Both populations are assumed normal and the samples were random and independently selected.

b) Test statistic: t with df = 9;

the smaller of $n_1 - 1 = 13 - 1 = 12$ or $n_2 - 1 = 10 - 1 = 9$.

c) Confidence level: $1 - \alpha = 0.95$.

3. The Sample Evidence:

Sample information given in the table.

Point estimate for $\mu_1 - \mu_2$:

$$\overline{x_1} - \overline{x_2} = 0.157 - 0.151 = 0.006 \text{ µg/L}$$

***t* Table**

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05
df							
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228

Two-sides problem

4. The Confidence Interval:

a) Confidence coefficients: $t_{9,0.025} = 2.262$

b) Maximum error:

$$E = t_{df,\alpha/2}\sqrt{\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)}$$

$$E = t_{9,0.025}\sqrt{\left(\frac{0.0071}{13}\right) + \left(\frac{0.0076}{10}\right)}$$

$$E = 2.262(0.0361) = 0.0818$$

c) Confidence limits:

$$(\overline{x_1} - \overline{x_2}) \pm E = 0.006 \pm 0.0818$$

$$-0.0758 \; to \; 0.0878$$

Example

Water specimens collected from a residential area that is served by the city water supply are compared with the specimens taken from a residential area that is served by private wells. The mean and variances of the samples are

	n	Mean Mercury Concentration (µg/L)	Variances
City	13	0.157	0.0071
Private	10	0.151	0.0076

Find a 95% confidence interval for the difference in mean mercury concentration between city water supply and private wells. Assume normality for the sampled populations and that the samples were selected randomly.

Solution:

1. Parameter of interest: The difference between the mean mercury concentration for city water supply and private wells, $\mu_1 - \mu_2$
2. The Confidence Interval Criteria:

a) Assumptions: Both populations are assumed normal and the samples were random and independently selected.

b) Test statistic: t with df = 9;

the smaller of $n_1 - 1 = 13 - 1 = 12$ or $n_2 - 1 = 10 - 1 = 9$.

c) Confidence level: $1 - \alpha = 0.95$.

3. The Sample Evidence:

Sample information given in the table.

Point estimate for $\mu_1 - \mu_2$:

$$\overline{x_1} - \overline{x_2} = 0.157 - 0.151 = 0.006 \text{ µg/L}$$

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05
df							
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228

Two-sides problem

4. The Confidence Interval:

a) Confidence coefficients: $t_{9,0.025} = 2.262$

b) Maximum error:

$$E = t_{df,\alpha/2}\sqrt{\left(\frac{s_1{}^2}{n_1}\right) + \left(\frac{s_2{}^2}{n_2}\right)}$$

$$E = t_{9,0.025}\sqrt{\left(\frac{0.0071}{13}\right) + \left(\frac{0.0076}{10}\right)}$$

$$E = 2.262(0.0361) = 0.0818$$

c) Confidence limits:

$$(\overline{x_1} - \overline{x_2}) \pm E = 0.006 \pm 0.0818$$

$$-0.0758 \text{ to } 0.0878$$

5. The Results: -0.0758 to 0.0878 is a 95% confidence interval for the difference in mean mercury concentration between city water supply and private wells. **<u>This confidence interval includes zero so there is no persuasive evidence in these data that the mercury contents are different in the two residential areas</u>**. Future sampling can be done in either area without worrying that the water supply will affect the outcome.

Two-Sample t Test for Comparing Two Populations

Null Hypothesis: H_0: $\mu_1 - \mu_2$ = hypothesized value

Alternative Hypothesis:

The hypothesized value is often 0, but there are times when we are interested in testing for a difference that is not 0.

H_a: $\mu_1 - \mu_2 >$ hypothesized value

H_a : $\mu_1 - \mu_2 <$ hypothesized value

H_a: $\mu_1 - \mu_2 \neq$ hypothesized value

Another Way to Write Hypothesis Statements:

$H_0: \mu_1 = \mu_2$

$H_0: \mu_1 - \mu_2 = 0$

$Ha: \mu_1 < \mu_2$

$H_a: \mu_1 - \mu_2 < 0$

$Ha: \mu_1 > \mu_2$

$H_a: \mu_1 - \mu_2 > 0$

$H_a: \mu_1 \neq \mu_2$

$Ha: \mu_1 - \mu_2 \neq 0$

CityU

Two-Sample t Test for Comparing Two Populations

Test statistic: $t = \dfrac{\bar{x}_1 - \bar{x}_2 - \text{hypothesized value}}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$

The appropriate df for the two-sample t test is

$$\text{df} = \frac{(V_1 + V_2)^2}{\dfrac{V_1^2}{n_1 - 1} + \dfrac{V_2^2}{n_2 - 1}} \quad \text{where } V_1 = \frac{s_1^2}{n_1} \text{ and } V_2 = \frac{s_2^2}{n_2}$$

The computed number of degrees of freedom should be truncated (rounded down) to an integer.

The hypothesized value is often 0, but there are times when we are interested in testing for a difference that is not 0.

where df is the **smaller** of df_1 or df_2 when computing t^* without the aid of a computer or calculator.

Example

A recent study compared a new drug to ease post-operative pain with the leading brand. Independent random samples were obtained and the number of hours of pain relief for each patient were recorded. The summary statistics are given in the table below.

Pain Reliever	n	Mean	St.Dev.
New Drug	10	4.350	0.542
Leading Brand	17	3.929	0.169

Is there any evidence to suggest the new drug provides longer relief from post-operative pain? Use $\alpha = 0.05$.

Example

A recent study compared a new drug to ease post-operative pain with the leading brand. Independent random samples were obtained and the number of hours of pain relief for each patient were recorded. The summary statistics are given in the table below.

Pain Reliever	n	Mean	St.Dev.
New Drug	10	4.350	0.542
Leading Brand	17	3.929	0.169

Is there any evidence to suggest the new drug provides longer relief from post-operative pain? Use $\alpha = 0.05$.

Solution:

1. The Set-up:

a) Parameter of concern: The difference between the mean time of pain relief for the new drug and that for the leading brand.

b) The null and alternative hypotheses:

H_0: $\mu_1 - \mu_2 = 0$ (new drug relieves pain no longer)

H_a: $\mu_1 - \mu_2 > 0$ (new drug works longer to relieve pain)

Example

A recent study compared a new drug to ease post-operative pain with the leading brand. Independent random samples were obtained and the number of hours of pain relief for each patient were recorded. The summary statistics are given in the table below.

Pain Reliever	n	Mean	St.Dev.
New Drug	10	4.350	0.542
Leading Brand	17	3.929	0.169

Is there any evidence to suggest the new drug provides longer relief from post-operative pain? Use $\alpha = 0.05$.

Solution:

1. The Set-up:

a) Parameter of concern: The difference between the mean time of pain relief for the new drug and that for the leading brand.

b) The null and alternative hypotheses:

H_0: $\mu_1 - \mu_2 = 0$ (new drug relieves pain no longer)

H_a: $\mu_1 - \mu_2 > 0$ (new drug works longer to relieve pain)

2. The Hypothesis Test Criteria:

a) Assumptions: Both populations are assumed to be approximately normal. The samples were random and independently selected.

b) Test statistic: t^*, df = 9

df = smaller of $n_1 - 1 = 10 - 1 = 9$ or $n_2 - 1 = 17 - 1 = 16$

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

Example

A recent study compared a new drug to ease post-operative pain with the leading brand. Independent random samples were obtained and the number of hours of pain relief for each patient were recorded. The summary statistics are given in the table below.

Pain Reliever	n	Mean	St.Dev.
New Drug	10	4.350	0.542
Leading Brand	17	3.929	0.169

Is there any evidence to suggest the new drug provides longer relief from post-operative pain? Use $\alpha = 0.05$.

Solution:

1. The Set-up:

a) Parameter of concern: The difference between the mean time of pain relief for the new drug and that for the leading brand.

b) The null and alternative hypotheses:

H_0: $\mu_1 - \mu_2 = 0$ (new drug relieves pain no longer)

H_a: $\mu_1 - \mu_2 > 0$ (new drug works longer to relieve pain)

2. The Hypothesis Test Criteria:

a) Assumptions: Both populations are assumed to be approximately normal. The samples were random and independently selected.

b) Test statistic: t^*, df = 9

df = smaller of $n_1 - 1 = 10 - 1 = 9$ or $n_2 - 1 = 17 - 1 = 16$

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

3. The Sample Evidence:

a) Sample information: Given in the table.

b) Test statistic:

$$t^* = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)}} = \frac{(4.350 - 3.929) - (0.00)}{\sqrt{\left(\frac{0.542^2}{10}\right) + \left(\frac{0.169^2}{17}\right)}}$$

$$= \frac{0.421}{\sqrt{0.0294 + 0.0017}} = \frac{0.421}{0.1763} = 2.39$$

CityU

Example

A recent study compared a new drug to ease post-operative pain with the leading brand. Independent random samples were obtained and the number of hours of pain relief for each patient were recorded. The summary statistics are given in the table below.

Pain Reliever	n	Mean	St.Dev.
New Drug	10	4.350	0.542
Leading Brand	17	3.929	0.169

Is there any evidence to suggest the new drug provides longer relief from post-operative pain? Use $\alpha = 0.05$.

Solution:

1. The Set-up:

a) Parameter of concern: The difference between the mean time of pain relief for the new drug and that for the leading brand.

b) The null and alternative hypotheses:

H_0: $\mu_1 - \mu_2 = 0$ (new drug relieves pain no longer)

H_a: $\mu_1 - \mu_2 > 0$ (new drug works longer to relieve pain)

2. The Hypothesis Test Criteria:

a) Assumptions: Both populations are assumed to be approximately normal. The samples were random and independently selected.

b) Test statistic: t^*, df = 9

df = smaller of $n_1 - 1 = 10 - 1 = 9$ or $n_2 - 1 = 17 - 1 = 16$

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

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$
one-tail	0.50	0.25	0.20	0.15	0.10	$\boxed{0.05}$	0.025
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05
df							
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306
$\boxed{9}$	0.000	0.703	0.883	1.100	1.383	$\boxed{1.833}$	2.262
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228

One-sides problem

3. The Sample Evidence:

a) Sample information: Given in the table.

b) Test statistic:

$$t^* = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)}} = \frac{(4.350 - 3.929) - (0.00)}{\sqrt{\left(\frac{0.542^2}{10}\right) + \left(\frac{0.169^2}{17}\right)}}$$

$$= \frac{0.421}{\sqrt{0.0294 + 0.0017}} = \frac{0.421}{0.1763} = 2.39$$

4. The Probability Distribution (Classical Approach):

a) Critical value: t(df, 0.05) = t(9, 0.05) = 1.833

b) t* is in the critical region.

5. The Results:

a) Decision: Reject H_0.

b) Conclusion: There is evidence to suggest that the new drug provides longer relief from post-operative pain.

Example

For an example of a comparison between a reference site and a potentially contaminated site, some data were extracted from a much larger set described by Gore and Patil (1994). Their study involved two phases of sampling of polychlorinated biphenyl (PCB) at the site of the Armagh compressor station in Indiana County, Pennsylvania. The phase 1 sampling was in areas close to sources of PCB, while the phase 2 sampling was away from these areas.

Example

For an example of a comparison between a reference site and a potentially contaminated site, some data were extracted from a much larger set described by Gore and Patil (1994). Their study involved two phases of sampling of polychlorinated biphenyl (PCB) at the site of the Armagh compressor station in Indiana County, Pennsylvania. The phase 1 sampling was in areas close to sources of PCB, while the phase 2 sampling was away from these areas.

For the present purpose, a random sample of 30 observations was extracted from the phase 2 sampling results to represent a sample from a reference area, and a random sample of 20 observations was extracted from the phase 1 sample results to represent a sample from a possibly contaminated area. The values for the PCB concentrations in parts per million (ppm) are shown in the right-hand side of Table. However, for data of this type, it is common to find that distributions are approximately lognormal, suggesting that the comparison between samples is best made on the logarithms of the original results, which should be approximately normally distributed, with the variation being more similar in different samples. This turns out to be the case here, as shown by the right-hand sides of Table. Suppose that it is decided that the two areas are equivalent in practical terms, provided that the ratio of the mean PCB concentration in the possibly contaminated area to the mean in the reference area is above
0.5. Then this corresponds to a difference between the logarithms of means for log(0.5) = −0.301, using logarithms to base 10. These tests will be carried out here using the 5% level of significance.

PCB Concentrations in a Reference Area and a Possibly Contaminated Area around the Armagh Compressor Station, and Results Transformed to Logarithms to Base 10

	Original PCB Concentration (ppm)		After Log Transformation	
	Reference	Contaminated	Reference	Contaminated
	3.5	2.6	0.54	0.41
	5.0	18.0	0.70	1.26
	36.0	110.0	1.56	2.04
	68.0	1300.0	1.83	3.11
	170.0	6.9	2.23	0.84
	4.3	1.0	0.63	0.00
	7.4	13.0	0.87	1.11
	7.1	1070.0	0.85	3.03
	1.6	661.0	0.20	2.82
	3.8	8.9	0.58	0.95
	35.0	34.0	1.54	1.53
	1.1	24.0	0.04	1.38
	27.0	22.0	1.43	1.34
	19.0	74.0	1.28	1.87
	64.0	80.0	1.81	1.90
	40.0	1900.0	1.60	3.28
	320.0	2.4	2.51	0.38
	1.7	1.5	0.23	0.18
	7.8	1.6	0.89	0.20
	1.6	140.0	0.20	2.15
	0.1		−1.30	
	0.1		−1.30	
	2.2		0.34	
	210.0		2.32	
	300.0		2.48	
	1.1		0.04	
	4.0		0.60	
	31.0		1.49	
	7.5		0.88	
	0.1		−1.30	
Mean	46.0	273.5	0.859	1.489
SD	86.5	534.7	1.030	1.025

Example

For an example of a comparison between a reference site and a potentially contaminated site, some data were extracted from a much larger set described by Gore and Patil (1994). Their study involved two phases of sampling of polychlorinated biphenyl (PCB) at the site of the Armagh compressor station in Indiana County, Pennsylvania. The phase 1 sampling was in areas close to sources of PCB, while the phase 2 sampling was away from these areas. For the present purpose, a random sample of 30 observations was extracted from the phase 2 sampling results to represent a sample from a reference area, and a random sample of 20 observations was extracted from the phase 1 sample results to represent a sample from a possibly contaminated area. The values for the PCB concentrations in parts per million (ppm) are shown in the right-hand side of Table. However, for data of this type, it is common to find that distributions are approximately lognormal, suggesting that the comparison between samples is best made on the logarithms of the original results, which should be approximately normally distributed, with the variation being more similar in different samples. This turns out to be the case here, as shown by the right-hand sides of Table. Suppose that it is decided that the two areas are equivalent in practical terms, provided that the ratio of the mean PCB concentration in the possibly contaminated area to the mean in the reference area is above 0.5. Then this corresponds to a difference between the logarithms of means for log(0.5) = −0.301, using logarithms to base 10. These tests will be carried out here using the 5% level of significance.

PCB Concentrations in a Reference Area and a Possibly Contaminated Area around the Armagh Compressor Station, and Results Transformed to Logarithms to Base 10

	Original PCB Concentration (ppm)		After Log Transformation	
	Reference	Contaminated	Reference	Contaminated
	3.5	2.6	0.54	0.41
	5.0	18.0	0.70	1.26
	36.0	110.0	1.56	2.04
	68.0	1300.0	1.83	3.11
	170.0	6.9	2.23	0.84
	4.3	1.0	0.63	0.00
	7.4	13.0	0.87	1.11
	7.1	1070.0	0.85	3.03
	1.6	661.0	0.20	2.82
	3.8	8.9	0.58	0.95
	35.0	34.0	1.54	1.53
	1.1	24.0	0.04	1.38
	27.0	22.0	1.43	1.34
	19.0	74.0	1.28	1.87
	64.0	80.0	1.81	1.90
	40.0	1900.0	1.60	3.28
	320.0	2.4	2.51	0.38
	1.7	1.5	0.23	0.18
	7.8	1.6	0.89	0.20
	1.6	140.0	0.20	2.15
	0.1		−1.30	
	0.1		−1.30	
	2.2		0.34	
	210.0		2.32	
	300.0		2.48	
	1.1		0.04	
	4.0		0.60	
	31.0		1.49	
	7.5		0.88	
	0.1		−1.30	
Mean	46.0	273.5	0.859	1.489
SD	86.5	534.7	1.030	1.025

Solution:

1. The Set-up: Parameter of concern: two areas are equivalent in practical terms.

The null and alternative hypotheses:

H_0: $\mu_1 - \mu_2 = 0$ (two areas are non-equivalent; difference between the logarithms of means is below -0.301)

H_a: $\mu_1 - \mu_2 > 0$ (two areas are equivalent; difference between the logarithms of means is above -0.301)

Example

For an example of a comparison between a reference site and a potentially contaminated site, some data were extracted from a much larger set described by Gore and Patil (1994). Their study involved two phases of sampling of polychlorinated biphenyl (PCB) at the site of the Armagh compressor station in Indiana County, Pennsylvania. The phase 1 sampling was in areas close to sources of PCB, while the phase 2 sampling was away from these areas. For the present purpose, a random sample of 30 observations was extracted from the phase 2 sampling results to represent a sample from a reference area, and a random sample of 20 observations was extracted from the phase 1 sample results to represent a sample from a possibly contaminated area. The values for the PCB concentrations in parts per million (ppm) are shown in the right-hand side of Table. However, for data of this type, it is common to find that distributions are approximately lognormal, suggesting that the comparison between samples is best made on the logarithms of the original results, which should be approximately normally distributed, with the variation being more similar in different samples. This turns out to be the case here, as shown by the right-hand sides of Table. Suppose that it is decided that the two areas are equivalent in practical terms, provided that the ratio of the mean PCB concentration in the possibly contaminated area to the mean in the reference area is above 0.5. Then this corresponds to a difference between the logarithms of means for log(0.5) = −0.301, using logarithms to base 10. These tests will be carried out here using the 5% level of significance.

Solution:

1. The Set-up: Parameter of concern: two areas are equivalent in practical terms.

The null and alternative hypotheses:

H_0: $\mu_1 - \mu_2 = 0$ (two areas are non-equivalent; difference between the logarithms of means is below -0.301)

H_a: $\mu_1 - \mu_2 > 0$ (two areas are equivalent; difference between the logarithms of means is above -0.301)

2. The Hypothesis Test Criteria:

a) Assumptions: Both populations are assumed to be approximately normal. The samples were random and independently selected.

b) Test statistic: t^*, df = 19

df = smaller of $n_1 - 1 = 20 - 1 = 19$ or $n_2 - 1 = 30 - 1 = 29$

c) Level of significance: α = 0.05.

PCB Concentrations in a Reference Area and a Possibly Contaminated Area around the Armagh Compressor Station, and Results Transformed to Logarithms to Base 10

	Original PCB Concentration (ppm)		After Log Transformation	
	Reference	Contaminated	Reference	Contaminated
	3.5	2.6	0.54	0.41
	5.0	18.0	0.70	1.26
	36.0	110.0	1.56	2.04
	68.0	1300.0	1.83	3.11
	170.0	6.9	2.23	0.84
	4.3	1.0	0.63	0.00
	7.4	13.0	0.87	1.11
	7.1	1070.0	0.85	3.03
	1.6	661.0	0.20	2.82
	3.8	8.9	0.58	0.95
	35.0	34.0	1.54	1.53
	1.1	24.0	0.04	1.38
	27.0	22.0	1.43	1.34
	19.0	74.0	1.28	1.87
	64.0	80.0	1.81	1.90
	40.0	1900.0	1.60	3.28
	320.0	2.4	2.51	0.38
	1.7	1.5	0.23	0.18
	7.8	1.6	0.89	0.20
	1.6	140.0	0.20	2.15
	0.1		−1.30	
	0.1		−1.30	
	2.2		0.34	
	210.0		2.32	
	300.0		2.48	
	1.1		0.04	
	4.0		0.60	
	31.0		1.49	
	7.5		0.88	
	0.1		−1.30	
Mean	46.0	273.5	0.859	1.489
SD	86.5	534.7	1.030	1.025

Example

For an example of a comparison between a reference site and a potentially contaminated site, some data were extracted from a much larger set described by Gore and Patil (1994). Their study involved two phases of sampling of polychlorinated biphenyl (PCB) at the site of the Armagh compressor station in Indiana County, Pennsylvania. The phase 1 sampling was in areas close to sources of PCB, while the phase 2 sampling was away from these areas. For the present purpose, a random sample of 30 observations was extracted from the phase 2 sampling results to represent a sample from a reference area, and a random sample of 20 observations was extracted from the phase 1 sample results to represent a sample from a possibly contaminated area. The values for the PCB concentrations in parts per million (ppm) are shown in the right-hand side of Table. However, for data of this type, it is common to find that distributions are approximately lognormal, suggesting that the comparison between samples is best made on the logarithms of the original results, which should be approximately normally distributed, with the variation being more similar in different samples. This turns out to be the case here, as shown by the right-hand sides of Table. Suppose that it is decided that the two areas are equivalent in practical terms, provided that the ratio of the mean PCB concentration in the possibly contaminated area to the mean in the reference area is above 0.5. Then this corresponds to a difference between the logarithms of means for log(0.5) = −0.301, using logarithms to base 10. These tests will be carried out here using the 5% level of significance.

PCB Concentrations in a Reference Area and a Possibly Contaminated Area around the Armagh Compressor Station, and Results Transformed to Logarithms to Base 10

	Original PCB Concentration (ppm)		After Log Transformation	
	Reference	Contaminated	Reference	Contaminated
	3.5	2.6	0.54	0.41
	5.0	18.0	0.70	1.26
	36.0	110.0	1.56	2.04
	68.0	1300.0	1.83	3.11
	170.0	6.9	2.23	0.84
	4.3	1.0	0.63	0.00
	7.4	13.0	0.87	1.11
	7.1	1070.0	0.85	3.03
	1.6	661.0	0.20	2.82
	3.8	8.9	0.58	0.95
	35.0	34.0	1.54	1.53
	1.1	24.0	0.04	1.38
	27.0	22.0	1.43	1.34
	19.0	74.0	1.28	1.87
	64.0	80.0	1.81	1.90
	40.0	1900.0	1.60	3.28
	320.0	2.4	2.51	0.38
	1.7	1.5	0.23	0.18
	7.8	1.6	0.89	0.20
	1.6	140.0	0.20	2.15
	0.1		−1.30	
	0.1		−1.30	
	2.2		0.34	
	210.0		2.32	
	300.0		2.48	
	1.1		0.04	
	4.0		0.60	
	31.0		1.49	
	7.5		0.88	
	0.1		−1.30	
Mean	46.0	273.5	0.859	1.489
SD	86.5	534.7	1.030	1.025

Solution:

1. The Set-up: Parameter of concern: two areas are equivalent in practical terms.

The null and alternative hypotheses:

H_0: $\mu_1 - \mu_2 = 0$ (two areas are non-equivalent; difference between the logarithms of means is below -0.301)

H_a: $\mu_1 - \mu_2 > 0$ (two areas are equivalent; difference between the logarithms of means is above -0.301)

2. The Hypothesis Test Criteria:

a) Assumptions: Both populations are assumed to be approximately normal. The samples were random and independently selected.

b) Test statistic: t^*, df = 19

df = smaller of n_1 - 1 = 20 - 1 = 19 or n_2 - 1 = 30 - 1 = 29

c) Level of significance: α = 0.05.

3. The Sample Evidence:

a) Sample information: Given in the table.

b) Test statistic:

$$t^* = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)}} = \left(\frac{(1.489 - 0.859) - (-0.301)}{\sqrt{\left(\frac{1.025^2}{20}\right) + \left(\frac{1.03^2}{30}\right)}}\right) = 3.137$$

Example

For an example of a comparison between a reference site and a potentially contaminated site, some data were extracted from a much larger set described by Gore and Patil (1994). Their study involved two phases of sampling of polychlorinated biphenyl (PCB) at the site of the Armagh compressor station in Indiana County, Pennsylvania. The phase 1 sampling was in areas close to sources of PCB, while the phase 2 sampling was away from these areas. For the present purpose, a random sample of 30 observations was extracted from the phase 2 sampling results to represent a sample from a reference area, and a random sample of 20 observations was extracted from the phase 1 sample results to represent a sample from a possibly contaminated area. The values for the PCB concentrations in parts per million (ppm) are shown in the right-hand side of Table. However, for data of this type, it is common to find that distributions are approximately lognormal, suggesting that the comparison between samples is best made on the logarithms of the original results, which should be approximately normally distributed, with the variation being more similar in different samples. This turns out to be the case here, as shown by the right-hand sides of Table. Suppose that it is decided that the two areas are equivalent in practical terms, provided that the ratio of the mean PCB concentration in the possibly contaminated area to the mean in the reference area is above 0.5. Then this corresponds to a difference between the logarithms of means for log(0.5) = −0.301, using logarithms to base 10. These tests will be carried out here using the 5% level of significance.

PCB Concentrations in a Reference Area and a Possibly Contaminated Area around the Armagh Compressor Station, and Results Transformed to Logarithms to Base 10

	Original PCB Concentration (ppm)		After Log Transformation	
	Reference	Contaminated	Reference	Contaminated
	3.5	2.6	0.54	0.41
	5.0	18.0	0.70	1.26
	36.0	110.0	1.56	2.04
	68.0	1300.0	1.83	3.11
	170.0	6.9	2.23	0.84
	4.3	1.0	0.63	0.00
	7.4	13.0	0.87	1.11
	7.1	1070.0	0.85	3.03
	1.6	661.0	0.20	2.82
	3.8	8.9	0.58	0.95
	35.0	34.0	1.54	1.53
	1.1	24.0	0.04	1.38
	27.0	22.0	1.43	1.34
	19.0	74.0	1.28	1.87
	64.0	80.0	1.81	1.90
	40.0	1900.0	1.60	3.28
	320.0	2.4	2.51	0.38
	1.7	1.5	0.23	0.18
	7.8	1.6	0.89	0.20
	1.6	140.0	0.20	2.15
	0.1		−1.30	
	0.1		−1.30	
	2.2		0.34	
	210.0		2.32	
	300.0		2.48	
	1.1		0.04	
	4.0		0.60	
	31.0		1.49	
	7.5		0.88	
	0.1		−1.30	
Mean	46.0	273.5	0.859	1.489
SD	86.5	534.7	1.030	1.025

Solution:

1. The Set-up: Parameter of concern: two areas are equivalent in practical terms.

The null and alternative hypotheses:

H_0: $\mu_1 - \mu_2 = 0$ (two areas are non-equivalent; difference between the logarithms of means is below -0.301)

H_a: $\mu_1 - \mu_2 > 0$ (two areas are equivalent; difference between the logarithms of means is above -0.301)

2. The Hypothesis Test Criteria:

a) Assumptions: Both populations are assumed to be approximately normal. The samples were random and independently selected.

b) Test statistic: t^*, df = 19

df = smaller of n_1 - 1 = 20 - 1 = 19 or n_2 - 1 = 30 - 1 = 29

c) Level of significance: α = 0.05.

3. The Sample Evidence:

a) Sample information: Given in the table.

b) Test statistic:

$$t^* = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)}} = \left(\frac{(1.489 - 0.859) - (-0.301)}{\sqrt{\left(\frac{1.025^2}{20}\right) + \left(\frac{1.03^2}{30}\right)}}\right) = 3.137$$

4. The Probability Distribution (Classical Approach):

a) Critical value: t(df, 0.05) = t(19, 0.05) = 1.729

b) t^* is in the critical region.

5. The Results:

a) Decision: Reject H_0.

b) Conclusion: There is evidence to suggest that the two areas are equivalent.

***t* Table**

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05
two-tails	1.00	0.50	0.40	0.30	0.20	0.10
df						
1	0.000	1.000	1.376	1.963	3.078	6.314
2	0.000	0.816	1.061	1.386	1.886	2.920
3	0.000	0.765	0.978	1.250	1.638	2.353
4	0.000	0.741	0.941	1.190	1.533	2.132
5	0.000	0.727	0.920	1.156	1.476	2.015
6	0.000	0.718	0.906	1.134	1.440	1.943
7	0.000	0.711	0.896	1.119	1.415	1.895
8	0.000	0.706	0.889	1.108	1.397	1.860
9	0.000	0.703	0.883	1.100	1.383	1.833
10	0.000	0.700	0.879	1.093	1.372	1.812
11	0.000	0.697	0.876	1.088	1.363	1.796
12	0.000	0.695	0.873	1.083	1.356	1.782
13	0.000	0.694	0.870	1.079	1.350	1.771
14	0.000	0.692	0.868	1.076	1.345	1.761
15	0.000	0.691	0.866	1.074	1.341	1.753
16	0.000	0.690	0.865	1.071	1.337	1.746
17	0.000	0.689	0.863	1.069	1.333	1.740
18	0.000	0.688	0.862	1.067	1.330	1.734
19	0.000	0.688	0.861	1.066	1.328	1.729
20	0.000	0.687	0.860	1.064	1.325	1.725

Summary of Key Concepts and Formulas

TERM OR FORMULA	COMMENT
Independent samples	Two samples where the individuals or objects in the first sample are selected independently from those in the second sample.
Paired samples	Two samples for which each observation in one sample is paired in a meaningful way with a particular observation in a second sample.
$t = \dfrac{(\bar{x}_1 - \bar{x}_2) - \text{hypothesized value}}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$	The test statistic for testing H_0: $\mu_1 - \mu_2 =$ hypothesized value when the samples are independently selected and the sample sizes are large or it is reasonable to assume that both population distributions are normal.
$(\bar{x}_1 - \bar{x}_2) \pm (t \text{ critical value})\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}$	A formula for constructing a confidence interval for $\mu_1 - \mu_2$ when the samples are independently selected and the sample sizes are large or it is reasonable to assume that the population distributions are normal.
$\text{df} = \dfrac{(V_1 + V_2)^2}{\dfrac{V_1^2}{n_1 - 1} + \dfrac{V_2^2}{n_2 - 1}}$ where $V_1 = \dfrac{s_1^2}{n_1}$ and $V_2 = \dfrac{s_2^2}{n_2}$	The formula for determining df for the two-sample t test and confidence interval.