

Sec “9”

Confidence Interval (CI)
Of the difference between two
(means, proportions)

Remember: Three values are used to construct all confidence intervals:

Sample
Statistic

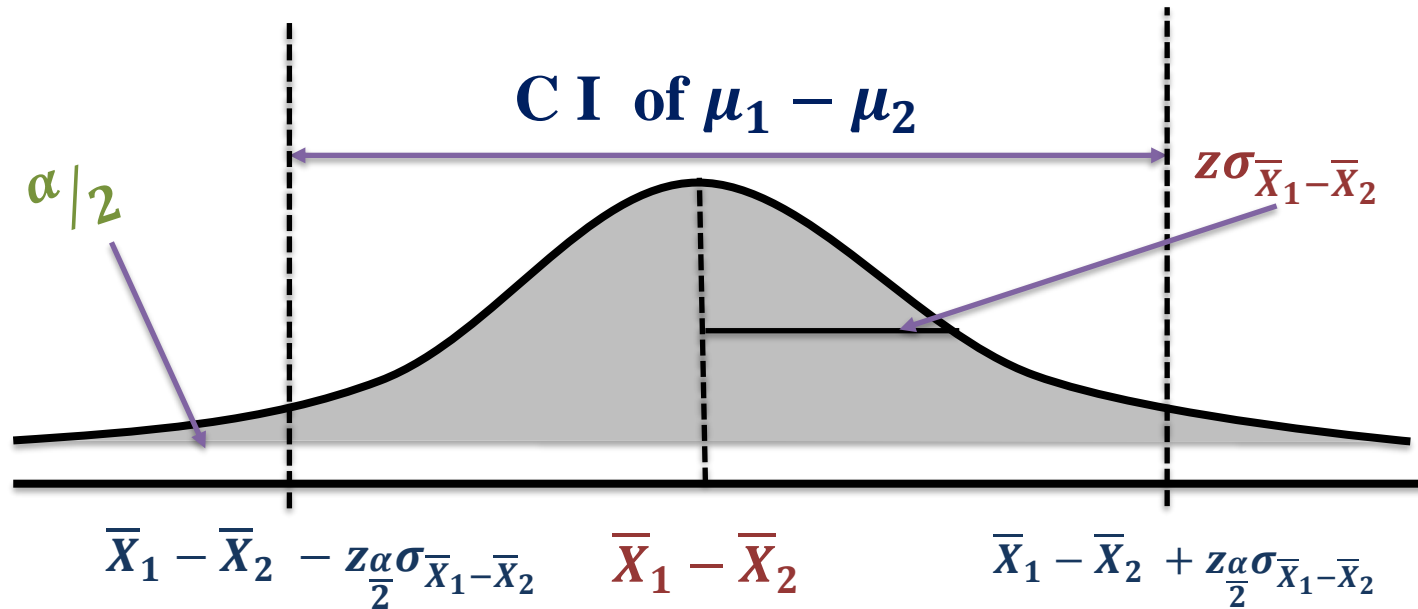
\pm

Z or t-table

\times

Standard error
of sample
statistics

Confidence Interval of the Difference Between two Population Mean " $\mu_1 - \mu_2$ "



- The Lower limit of the confidence interval: $(\bar{X}_1 - \bar{X}_2) - z\frac{\alpha}{2}\sigma_{\bar{X}_1 - \bar{X}_2}$
- The upper limit of the confidence interval: $(\bar{X}_1 - \bar{X}_2) + z\frac{\alpha}{2}\sigma_{\bar{X}_1 - \bar{X}_2}$

\therefore The CI of μ is given by: $(\bar{X}_1 - \bar{X}_2) \mp z\frac{\alpha}{2}\sigma_{\bar{X}_1 - \bar{X}_2}$

Also can be written as follows

$$(\bar{X}_1 - \bar{X}_2) - z\frac{\alpha}{2}\sigma_{\bar{X}_1 - \bar{X}_2} \leq \mu_1 - \mu_2 \leq (\bar{X}_1 - \bar{X}_2) + z\frac{\alpha}{2}\sigma_{\bar{X}_1 - \bar{X}_2}$$

$$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

➤ How to Calculate a Confidence Interval for $\mu_1 - \mu_2$?

Step 1:

	Population 1	Population 2
Sample size	n_1	n_2
Sample mean	$\bar{X}_1 = \frac{\sum_{i=1}^{n_1} x_i}{n_1}$	$\bar{X}_2 = \frac{\sum_{i=1}^{n_2} x_i}{n_2}$
Population variance known	σ_1^2	σ_2^2
or Population variance unknown (sample variance known)	$s_1^2 = \frac{\sum_{i=1}^{n_1} (x_i - \bar{X}_1)^2}{n_1 - 1}$	$s_2^2 = \frac{\sum_{i=1}^{n_2} (x_i - \bar{X}_2)^2}{n_2 - 1}$

- **Step 2:** The area in which the population mean $\mu_1 - \mu_2$ is located. $[(1 - \alpha)\%]$ → given in examples → $\alpha \rightarrow \frac{\alpha}{2}$
- **Step 3:** **Populations Variances “ σ_1^2, σ_2^2 ”**

If σ_1^2, σ_2^2 known

Use: Z- table.
CI:

$$\bar{X}_1 - \bar{X}_2 \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$n_1 \geq 30, n_2 \geq 30$

Use: Z- table.
CI:

$$\bar{X}_1 - \bar{X}_2 \pm z_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

If σ_1^2, σ_2^2 unknown → s_1^2, s_2^2 known

$\sigma_1^2 \neq \sigma_2^2$

$n_1 < 30, n_2 < 30$

Use: t- table.
CI:

$$\bar{X}_1 - \bar{X}_2 \pm t_{\frac{\alpha}{2}, v} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$\sigma_1^2 = \sigma_2^2$

Use: t- table.
CI:

$$\bar{X}_1 - \bar{X}_2 \pm t_{\frac{\alpha}{2}, v} \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}$$

Pooled variance

○ Note that:

- $v = v_1 + v_2 = n_1 - 1 + n_2 - 1 = n_1 + n_2 - 2.$

- Pooled variance:

$$s_P^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Exercise

(1) A random sample of size $n_1 = 25$ taken from a normal population with a standard deviation $\sigma_1 = 5$, has a mean $\bar{X}_1 = 80$. A second random sample of size $n_2 = 36$, taken from a different normal population with a standard deviation $\sigma_2 = 3$, has a mean $\bar{X}_2 = 75$. Find a 94% confidence interval for $\mu_1 - \mu_2$.

Solution:

- **Step 1:**

	Population 1	Population 2
Sample size	$n_1 = 25$	$n_2 = 36$
Sample mean	$\bar{X}_1 = 80$	$\bar{X}_2 = 75$
Population variance known	$\sigma_1^2 = 25$	$\sigma_2^2 = 9$

- **Step 2:**

$$(1 - \alpha)\% = 94\% \rightarrow 1 - \alpha = 0.94 \rightarrow \alpha = 1 - 0.94$$

$$\alpha = 0.06$$

- **Step 3:**

$\sigma_1^2, \sigma_2^2 \rightarrow$ known, so we use z-table

$$\frac{z\alpha}{2} = \frac{z_{0.06}}{2} = \frac{z_{0.03}}{2} = 1.88$$

CI of $\mu_1 - \mu_2$, with known σ_1^2, σ_2^2 is given by

$$\bar{X}_1 - \bar{X}_2 \mp \frac{z\alpha}{2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} =$$

$$(80 - 75) \mp (1.88) \left(\sqrt{\frac{25}{25} + \frac{9}{36}} \right)$$

$$= 5 \mp 2.1019 \rightarrow$$

$$5 - 2.1019 \leq \mu_1 - \mu_2 \leq 5 + 2.1019$$

$$\therefore 2.8981 \leq \mu_1 - \mu_2 \leq 7.1019$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0354	0.0349	0.0344	0.0339	0.0334	0.0329	0.0324	0.0319	0.0314
-1.7	0.0446	0.0441	0.0436	0.0431	0.0427	0.0422	0.0418	0.0414	0.0409	0.0405

$$0.0301 - 0.03 = 0.0001$$

$$0.03 - 0.0294 = 0.0006$$

(2) A study was conducted to determine if a certain metal treatment has any effect on the amount of metal removed in a pickling operation. A random sample of 100 pieces was immersed in a bath for 24 hours without the treatment, yielding an average of 12.2 millimeters of metal removed and a sample standard deviation of 1.1 millimeters. A second sample of 200 pieces was exposed to the treatment, followed by the 24-hour immersion in the bath, resulting in an average: removal of 9.1 millimeters of metal with a sample standard deviation of 0.9 millimeter. Compute a 98% confidence interval estimate for the difference between the population means?

Solution:

• **Step 1:**

	Without treatment	With treatment
Sample size	$n_1 = 100$	$n_2 = 200$
Sample mean	$\bar{X}_1 = 12.2$	$\bar{X}_2 = 9.1$
Sample variance known	$s_1^2 = (1.1)^2 = 1.21$	$s_2^2 = (0.9)^2 = 0.81$

- **Step 2:** $(1 - \alpha)\% = 98\% \rightarrow 1 - \alpha = 0.98 \rightarrow \alpha = 1 - 0.98$

$$\alpha = 0.02$$

- **Step 3:** $\sigma_1^2, \sigma_2^2 \rightarrow$ unknown,
so we use z-table

$$z_{\frac{\alpha}{2}} = z_{\frac{0.02}{2}} = z_{0.01} = 2.33$$

$$0.0102 - 0.01 = 0.0002$$

$$0.01 - 0.0099 = 0.0001$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143

CI of $\mu_1 - \mu_2$, with unknown σ_1^2, σ_2^2 and $n_1, n_2 > 30$ is given by

$$\bar{X}_1 - \bar{X}_2 \mp z_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$= (12.2 - 9.1) \mp (2.33) \left(\sqrt{\frac{1.21}{100} + \frac{0.81}{200}} \right)$$

$$= 3.1 \mp 0.296 \rightarrow$$

$$3.1 - 0.296 \leq \mu_1 - \mu_2 \leq 3.1 + 0.296$$

$$\therefore 2.804 \leq \mu_1 - \mu_2 \leq 3.396$$

(3) Students may choose between a 3-semester-hourcourse in physics without labs and a 4-semester-hourcourse with labs. The final written examination is the same for each section. If 12 students in the section with labs made an average examination grade of 84 with a standard deviation of 4, and 18 students in the section without labs made an average: grade of 77 with a standard deviation of 6, find a 99% confidence interval for the difference between the average grades for the two courses. Assume the populations to be approximately normally distributed with equal variances.

Solution:

• **Step 1:**

	3-semester without labs	4-semester with labs
Sample size	$n_1 = 12$	$n_2 = 18$
Sample mean	$\bar{X}_1 = 84$	$\bar{X}_2 = 77$
Sample variance known	$s_1^2 = (4)^2 = 16$	$s_2^2 = (6)^2 = 36$

- **Step 2:** $(1 - \alpha)\% = 99\% \rightarrow 1 - \alpha = 0.99 \rightarrow \alpha = 1 - 0.99$

$$\alpha = 0.01$$

- **Step 3:** $\sigma_1^2, \sigma_2^2 \rightarrow$ unknown,
so we use **t-table**

$$t_{\frac{\alpha}{2}, v} = t_{\frac{\alpha}{2}, n_1 + n_2 - 2} = t_{\frac{0.01}{2}, 12 + 18 - 2}$$

$$= t_{0.005, 28} = 2.763$$

v	α						
	0.02	0.015	0.01	0.0075	0.005	0.0025	0.0005
26	2.162	2.296	2.479	2.605	2.779	3.067	3.707
27	2.158	2.291	2.473	2.598	2.771	3.057	3.689
28	2.154	2.286	2.467	2.592	2.763	3.047	3.674
29	2.150	2.282	2.462	2.586	2.756	3.038	3.660
30	2.147	2.278	2.457	2.581	2.750	3.030	3.646

CI of $\mu_1 - \mu_2$, with unknown σ_1^2, σ_2^2 and $\sigma_1^2 = \sigma_2^2$ is given by

$$\bar{X}_1 - \bar{X}_2 \mp t_{\frac{\alpha}{2}, v} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$= (84 - 77) \mp (2.763)(5.305) \left(\sqrt{\frac{1}{12} + \frac{1}{18}} \right)$$

$$= 7 \mp 5.463 \rightarrow$$

$$7 - 5.463 \leq \mu_1 - \mu_2 \leq 7 + 5.463$$

$$\therefore 1.537 \leq \mu_1 - \mu_2 \leq 12.463$$

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

$$= \sqrt{\frac{(12-1)(16) + (18-1)(36)}{12+18-2}}$$

$$= 5.305$$

(4) In a study conducted at (lie Virginia Polytechnic Institute and State University on the development of ectomycorrhizal, a symbiotic relationship between the roots of trees and a fungus in which minerals are transferred from the fungus to the trees and sugars from the trees to the fungus, 20 northern red oak seedlings with the fungus *Pisolithus tinctorius* were grown in a greenhouse. All seedlings were planted in the same type of soil and received the same amount of sunshine and water. Half received no nitrogen at planting time: to serve as a control and the other half received 368 ppm of nitrogen in the form NaNO_3 . The stem weights recorded in grams, at the end of 140 days were recorded as follows:

No Nitrogen	Nitrogen
0.32	0.26
0.53	0.43
0.28	0.47
0.37	0.49
0.47	0.52
0.43	0.75
0.30	0.79
0.42	0.86
0.38	0.62
0.43	0.46

Construct a 95% confidence interval for the difference in the mean stem weights between seedlings that receive no nitrogen and those that receive 308 ppm of nitrogen. Assume the populations to be normally distributed with equal variances.

Solution:

• Step 1:

$$\square n_1 = n_2 = 10$$

$$\square \bar{X}_1 = \frac{\sum_{i=1}^{n_1} x_i}{n_1} = \frac{0.32+0.52+\dots+0.43}{10} = 0.393$$

$$\square \bar{X}_2 = \frac{\sum_{i=1}^{n_2} x_i}{n_2} = \frac{0.26+0.43+\dots+0.46}{10} = 0.565$$

$$\square s_1^2 = \frac{\sum_{i=1}^{n_1} (x_i - \bar{X}_1)^2}{n_1 - 1} = \frac{(0.32 - 0.393)^2 + (0.52 - 0.393)^2 + \dots + (0.43 - 0.393)^2}{9} \\ = 0.00618$$

$$\square s_2^2 = \frac{\sum_{i=1}^{n_2} (x_i - \bar{X}_2)^2}{n_2 - 1} = \frac{(0.26 - 0.565)^2 + (0.43 - 0.565)^2 + \dots + (0.46 - 0.565)^2}{9} \\ = 0.03487$$

- **Step 2:** $(1 - \alpha)\% = 95\% \rightarrow 1 - \alpha = 0.95 \rightarrow \alpha = 1 - 0.95$

$$\alpha = 0.05$$

- **Step 3:** $\sigma_1^2, \sigma_2^2 \rightarrow$ unknown, $s_1^2, s_2^2 \rightarrow$ known, $\sigma_1^2 = \sigma_2^2$, so we use t-table

$$t_{\frac{\alpha}{2}, v} = t_{\frac{\alpha}{2}, n_1 + n_2 - 2} = t_{\frac{0.05}{2}, 10 + 10 - 2}$$

$$= t_{0.025, 18} = 2.101$$

v	α						
	0.40	0.30	0.20	0.15	0.10	0.05	0.025
16	0.258	0.535	0.865	1.071	1.337	1.746	2.120
17	0.257	0.534	0.863	1.069	1.333	1.740	2.110
18	0.257	0.534	0.862	1.067	1.330	1.734	2.101
19	0.257	0.533	0.861	1.066	1.328	1.729	2.093
20	0.257	0.533	0.860	1.064	1.325	1.725	2.086

CI of $\mu_1 - \mu_2$, with unknown σ_1^2, σ_2^2 and $\sigma_1^2 = \sigma_2^2$ is given by

$$\bar{X}_1 - \bar{X}_2 \mp t_{\frac{\alpha}{2}, v} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$= (0.565 - 0.393) \mp (2.101)(0.1433) \left(\sqrt{\frac{1}{10} + \frac{1}{10}} \right)$$

$$= 0.172 \mp 0.1346 \rightarrow$$

$$0.172 - 0.1346 \leq \mu_1 - \mu_2 \leq 0.172 + 0.1346$$

$$\therefore 0.0374 \leq \mu_1 - \mu_2 \leq 0.3066$$

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

$$= \sqrt{\frac{(10-1)(0.00618) + (10-1)(0.03487)}{10+10-2}}$$

$$= 0.1433$$

(5) A study was conducted by the Department of Zoology at the Virginia Polytechnic Institute and State University to estimate the difference in the amount of chemical orthophosphoric measured at two different stations on the James River. Orthophosphoric is measured in milligrams per liter. Fifteen samples were collected from station 1 and 12 samples were obtained from station 2. The 15 samples from station 1 had an average orthophosphoric content of 3.84 milligrams per liter and a standard deviation of 3.07 milligrams per liter, while the 12 samples from station 2 had an average orthophosphoric content of 1.49 milligrams per liter and a standard deviation of 0.8 milligram per liter. Find a 95% confidence interval for the difference in the true average orthophosphoric content at these two stations, assuming that the observations came from normal populations with different variances ?

Solution:

- **Step 1:**

	Station 1	Station 2
Sample size	$n_1 = 15$	$n_2 = 12$
Sample mean	$\bar{X}_1 = 3.84$	$\bar{X}_2 = 1.49$
Sample variance known	$s_1^2 = (3.07)^2 = 9.425$	$s_2^2 = (0.8)^2 = 0.64$

- **Step 2:** $(1 - \alpha)\% = 95\% \rightarrow 1 - \alpha = 0.95 \rightarrow \alpha = 1 - 0.95$

$$\alpha = 0.05$$

- **Step 3:** $\sigma_1^2, \sigma_2^2 \rightarrow$ unknown, $s_1^2, s_2^2 \rightarrow$ known, $n_1, n_2 < 30$, so we use t-table

$$t_{\frac{\alpha}{2}, v} = t_{\frac{\alpha}{2}, n_1 + n_2 - 2} = t_{\frac{0.05}{2}, 15 + 12 - 2} = t_{0.025, 25} = 2.060$$

v	α						
	0.40	0.30	0.20	0.15	0.10	0.05	0.025
21	0.257	0.532	0.859	1.063	1.323	1.721	2.080
22	0.256	0.532	0.858	1.061	1.321	1.717	2.074
23	0.256	0.532	0.858	1.060	1.319	1.714	2.069
24	0.256	0.531	0.857	1.059	1.318	1.711	2.064
25	0.256	0.531	0.856	1.058	1.316	1.708	2.060

CI of $\mu_1 - \mu_2$, with unknown σ_1^2, σ_2^2 and $n_1, n_2 < 30$ is given by

$$\bar{X}_1 - \bar{X}_2 \mp t_{\frac{\alpha}{2}, v} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$= (3.84 - 1.49) \mp (2.060) \left(\sqrt{\frac{9.425}{15} + \frac{0.64}{12}} \right)$$

$$= 2.35 \mp 1.7008 \rightarrow$$

$$2.35 - 1.7008 \leq \mu_1 - \mu_2 \leq 2.35 + 1.7008$$

$$\therefore 0.6492 \leq \mu_1 - \mu_2 \leq 4.0508$$

More
Examples

1. Fast food companies want their products to be as similar as possible across locations. Miguel is tasked with making sure that two locations of a fast food chain have burger patties of the same size. He takes a random sample of 30 patties from each location. Here are summary statistics for his samples:

	Mass of burger patties	
	Location A	Location B
Sample mean	$\bar{x}_A = 115 \text{ g}$	$\bar{x}_B = 117 \text{ g}$
Sample standard deviation	$s_A = 2.8 \text{ g}$	$s_B = 2.1 \text{ g}$
Sample size	$n_A = 30$	$n_B = 30$

Construct a 90% confidence interval for the difference in mean weight (in grams) of the patties from the two restaurants $\mu_B - \mu_A$?

More
Examples

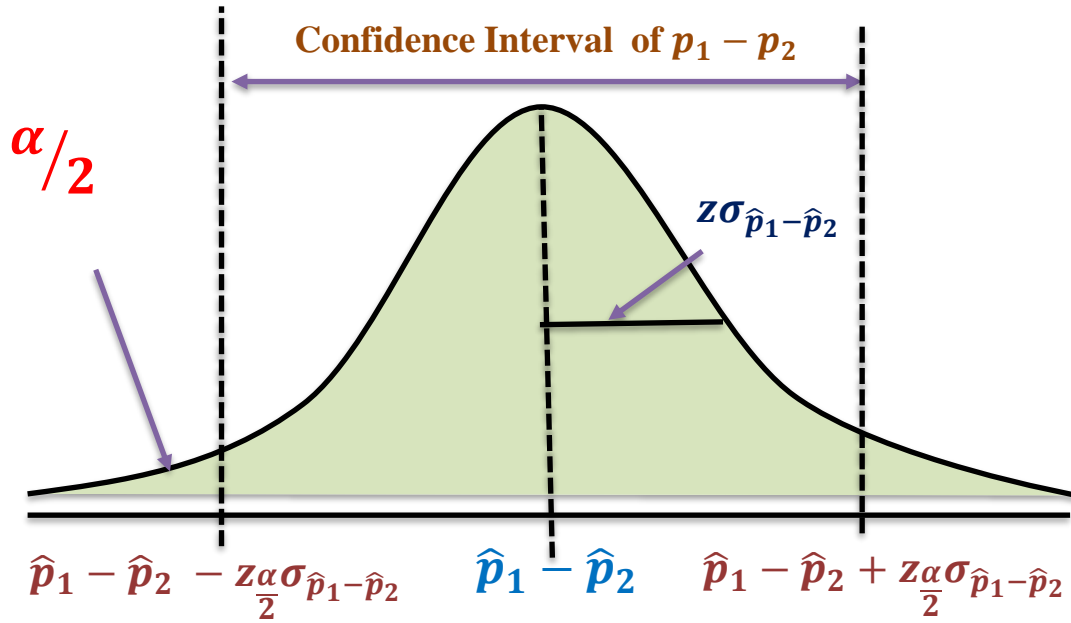
2. A sample of 40 pine trees grown on the north side of a hill has a mean of 25.4 meters and a standard deviation of 2.1 meters. A second sample of 40 trees from the south side has a mean of 23.2 meters and a standard deviation of 1.7 meters.

Find the 95% confidence interval for the difference in the mean heights of the two populations of trees?

More
Examples

3. Weekly expenditure of students of two Universities are studied. Variances of the two populations are 90.1 and 97.7, respectively. From population 1, a sample of 15 students is selected with sample mean being \$204.20. From population 2, a sample of 10 students is selected and the mean is \$184.60. What is the 95% Confidence Interval for the difference between the two population means?

Confidence Interval of the Population Proportion " $p_1 - p_2$ "



\therefore The CI of $p_1 - p_2$ is given by

$$(\hat{p}_1 - \hat{p}_2) \mp z_{\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

Also can be written as follows

$$(\hat{p}_1 - \hat{p}_2) - z_{\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}} \leq p_1 - p_2 \leq (\hat{p}_1 - \hat{p}_2) + z_{\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

- The Lower limit of the confidence interval:

$$(\hat{p}_1 - \hat{p}_2) - z_{\alpha/2} \sigma_{\hat{p}_1 - \hat{p}_2}$$

$$= (\hat{p}_1 - \hat{p}_2) - z_{\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

- The upper limit of the confidence interval:

$$(\hat{p}_1 - \hat{p}_2) + z_{\alpha/2} \sigma_{\hat{p}_1 - \hat{p}_2}$$

$$= (\hat{p}_1 - \hat{p}_2) + z_{\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

□ How to Calculate a Confidence Interval for a Population Proportion?

Step 1:

	Population 1	Population 2
Sample size	n_1	n_2
Sample success	X_1	X_2
Sample proportion “probability of success”	$\hat{p}_1 = \frac{X_1}{n_1}$	$\hat{p}_2 = \frac{X_2}{n_2}$
Probability of failure for sample	$\hat{q}_1 = 1 - \hat{p}_1$	$\hat{q}_2 = 1 - \hat{p}_2$

Step 2: $[(1 - \alpha)\%]$ → given in examples, $\alpha \rightarrow \frac{\alpha}{2}$. Compute $z_{\frac{\alpha}{2}}$

Step 3: The confidence interval for a population proportion $p_1 - p_2$ is given by:

$$(\hat{p}_1 - \hat{p}_2) - z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}} \leq p_1 - p_2 \leq (\hat{p}_1 - \hat{p}_2) + z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

Exercise

(1) A certain geneticist is interested in the proportion of males and females in the population that have a certain minor blood disorder. In a random sample of 1000 males, 250 are found to be afflicted, whereas 275 of 1000 females tested appear to have the disorder. Compute a 95% confidence interval for the difference between the proportion of males and females that have the blood disorder ?

Solution:

- **Step 1:**

	Male	Female
Sample size	$n_1 = 1000$	$n_2 = 1000$
Sample success	$X_1 = 250$	$X_2 = 275$
Sample proportion	$\hat{p}_1 = \frac{250}{1000} = 0.25$	$\hat{p}_2 = \frac{275}{1000} = 0.275$
Probability of failure of sample	$\hat{q}_1 = 1 - 0.25 = 0.75$	$\hat{q}_2 = 1 - 0.275 = 0.725$

- Step 2: $(1 - \alpha)\% = 95\% \rightarrow 1 - \alpha = 0.95 \rightarrow \alpha = 1 - 0.95$

$$\alpha = 0.05$$

- Step 3:

$$Z_{\frac{\alpha}{2}} = Z_{\frac{0.05}{2}} = Z_{0.025} = 1.96$$

CI of $p_1 - p_2$, is given by

$$(\hat{p}_1 - \hat{p}_2) \mp Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

$$= (0.25 - 0.275) \mp (1.96) \sqrt{\frac{(0.25)(0.75)}{1000} + \frac{(0.275)(0.725)}{1000}}$$

$$= -0.025 \mp 0.038552$$

$$\rightarrow -0.025 - 0.038552 \leq p_1 - p_2 \leq -0.025 + 0.038552$$

$$\therefore -0.0636 \leq p_1 - p_2 \leq 0.0136$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0269	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367

(2) Ten engineering schools in the United States were surveyed. The sample contained 250 electrical engineers, 80 being women; 175 chemical engineers, 40 being women. Compute a 90% confidence interval for the difference between the proportions of women in these two fields of engineering.

Solution:

• **Step 1:**

Sample size	$n_1 = 250$	$n_2 = 175$
Sample success	$X_1 = 80$	$X_2 = 40$
Sample proportion	$\hat{p}_1 = \frac{80}{250} = 0.32$	$\hat{p}_2 = \frac{40}{175} = 0.229$
Probability of failure of sample	$\hat{q}_1 = 1 - 0.32 = 0.68$	$\hat{q}_2 = 1 - 0.229 = 0.771$

- **Step 2:** $(1 - \alpha)\% = 90\% \rightarrow 1 - \alpha = 0.90 \rightarrow \alpha = 1 - 0.90$

$$\alpha = 0.1$$

- **Step 3:**

$$\begin{aligned} \frac{z\alpha}{2} &= \frac{z_{0.1}}{2} = z_{0.05} \\ &= \frac{1.64 + 1.65}{2} = 1.645 \end{aligned}$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-1.7	0.0446	0.0438	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0540	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681

CI of $p_1 - p_2$, is given by

$$(\hat{p}_1 - \hat{p}_2) \mp z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

$$0.0505 - 0.05 = 0.0005$$

$$0.05 - 0.0495 = 0.0005$$

$$= (0.32 - 0.229) \mp (1.645) \sqrt{\frac{(0.32)(0.68)}{250} + \frac{(0.229)(0.771)}{175}}$$

$$= 0.0914 \mp 0.07131$$

$$\rightarrow 0.0914 - 0.07131 \leq p_1 - p_2 \leq 0.0914 + 0.07131$$

$$\therefore 0.02009 \leq p_1 - p_2 \leq 0.16271$$



More
Examples

1. A certain change in a process for manufacture of component parts is being considered. Samples are taken using both the existing and the new procedure so as to determine if the new process results in an improvement. If 75 of 1500 items from the existing procedure were found to be defective and 80 of 2000 items from the new procedure were found to be defective.

Find a 90% confidence interval for the true difference in the fraction of defective between the existing and the new process?

More
Examples

2. Suppose the Cartoon Network conducts a nation-wide survey to assess viewer attitudes toward Superman. Using a simple random sample, they select 400 boys and 300 girls to participate in the study. Forty percent of the boys say that Superman is their favorite character, compared to thirty percent of the girls.

What is the 90% confidence interval for the true difference in attitudes toward Superman?

More
Examples

3.

	African Women	Women from Americas
Sample size	2100	1900
Number with anemia	840	323

Find a 95% confidence interval for the difference in proportions of all African women with anemia and all women from the Americas with anemia ?