Using Confidence Intervals to Estimate the Difference $(p_1 - p_2)$ in 2 Population Proportions p_1 and p_2 using Two Independent Samples

If $p_1 - p_1 = 0$ then there is no difference in the 2 Population Proportions at a significance level of α

Requirements

1. An independent random sample of each population is taken. The sample proportion for the sample of one population is $\hat{p}_1$ and the sample proportion for the sample of a second population is $\hat{p}_2$. The equations are easier to use if you select the largest $\hat{p}$ as $\hat{p}_1$ and the smallest $\hat{p}$ as $\hat{p}_2$.

2. The number of success and failures for each sample is greater than or equal to 5.

Notation for the Samples of Two Population Proportions

Population 1

p_1 = the population one proportion

n_1 = sample size

x_1 = number of successes in the sample

$\hat{p}_1 = \dfrac{x_1}{n_1}$ (the sample one proportion)

Population 2

p_2 = the population two proportion

n_2 = sample size

x_2 = number of successes in the sample

$\hat{p}_2 = \dfrac{x_2}{n_2}$ (the sample two proportion)

Creating a Confidence Interval to Estimate the value of $(p_1 - p_2)$

$$(\hat{p}_1 - \hat{p}_2) - E \;<\; (p_1 - p_2) \;<\; (\hat{p}_1 - \hat{p}_2) + E \quad \textbf{if} \quad \hat{p}_1 > \hat{p}_2$$

or

$$(\hat{p}_2 - \hat{p}_1) - E \;<\; (p_2 - p_1) \;<\; (\hat{p}_2 - \hat{p}_1) + E \quad \textbf{if} \quad \hat{p}_2 > \hat{p}_1$$

Where $E = z_{\alpha/2} \bullet \sqrt{\left(\dfrac{\hat{p}_1 \bullet \hat{q}_1}{n_1}\right) + \left(\dfrac{\hat{p}_2 \bullet \hat{q}_2}{n_2}\right)}$

Note: It does not matter which population is **p1 and p2. It is important that the difference between them is positive.** The **first formula** is based **on** $\hat{p}_1$ **being larger than** $\hat{p}_2$**.** The **second formula** is based on $\hat{p}_2$ **being larger than** $\hat{p}_1$**.**

You are free to chose which population is selected to be population 1 or population 2 . If you always **chose the population with the largest** $\hat{p}$ **as population 1 then you can always use the first formula.**

Example 1

Is there a difference in the proportion of FLC students and ARC students who bring a laptop to school?

A random sample of 200 FLC students shows that 88 of them use a personal laptop at school. A random sample of 300 ARC students shows that 120 of them use a personal laptop at school. Construct a 96% confidence interval for the difference between the two population proportions. Does it appear there is a difference in the two proportions? **How can you tell?**

Sample from Population 1 FLC

$n_1 = 200 \qquad x_1 = 88$

$$\hat{p}_1 = \frac{x_1}{n_1} = \frac{88}{200} = .44$$

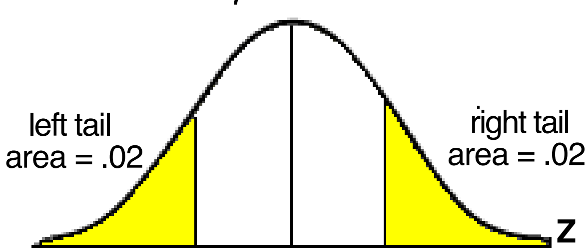
Sample from Population 2 ARC

$n_2 = 300 \qquad x_2 = 120$

$$\hat{p}_2 = \frac{x_2}{n_2} = \frac{120}{300} = .40$$

Find the critical value $+z_{\alpha/2}$

$\alpha = .04 \quad \text{so} \quad \alpha/2 = .02$



$-z_{\alpha/2}$ **= – 2.05** $\qquad +z_{\alpha/2}$ **= 2.05**

Find the Maximum Error

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

$$E = \ 2.05 \bullet \sqrt{\left(\frac{.44 \bullet .56}{200}\right) + \left(\frac{.40 \bullet .60}{300}\right)}$$

$$E = \ .09$$

Find the Confidence Interval to Estimate the value of $p_1 - p_2$

$$(\hat{p}_1 - \hat{p}_2) - E \ < (p_1 - p_2) \ < (\hat{p}_1 - \hat{p}_2) + E \quad \textbf{if} \quad \hat{p}_1 > \hat{p}_2$$

$$(.44 - .40) - .09 \ < (p_1 - p_2) \ < (.44 - .40) + .09$$

$$-.05 < (p_1 - p_2) \ < .13$$

Conclusion based on the problem: The confidence interval does contain zero.

I am 96% confident that there is no difference in the **proportion** of FLC and ARC students that bring a personal laptop to school.

Negative Z Scores									
Standard Normal (Z) Distribution: Cumulative Area to the LEFT of Z									
Z	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
–2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183

Example 2

Is there a difference in the proportion of FLC students and FLC Faculty who own a I Phone?

A random sample of 300 FLC students shows that 210 of them own an I phone. A random sample of 100 FLC faculty shows that 80 of them own an I phone. Construct a 90% confidence interval for the difference between the two proportions. Does it appear there is a difference in the two population proportions? **How can you tell?**

Sample from Population 1 Faculty students

n_1 = 100 x_1 = 80

$$\hat{p}_1 = \frac{x_1}{n_1} = \frac{80}{100} = .80$$

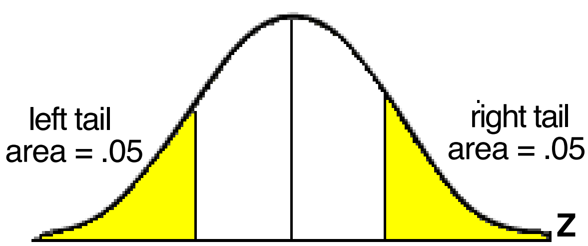
Sample from Population 2 FLC

n_2 = 300 x_2 = 210

$$\hat{p}_2 = \frac{x_2}{n_2} = \frac{210}{300} = .70$$

Find the critical value $+z_{\alpha/2}$

$\alpha = .10 \text{ so } \alpha/2 = .05$



$-z_{\alpha/2}$ **= − 1.645** $+z_{\alpha/2}$ **= 1.645**

Find the Maximum Error

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

$$E = 1.645 \bullet \sqrt{\left(\frac{.80 \bullet .20}{100}\right) + \left(\frac{.70 \bullet .30}{300}\right)}$$

$$E = .08$$

Find the Confidence Interval to Estimate the value of $(p_1 - p_2)$

$$(\hat{p}_1 - \hat{p}_2) - E < (p_1 - p_2) < (\hat{p}_1 - \hat{p}_2) + E \quad \textbf{if} \quad \hat{p}_1 > \hat{p}_2$$

$$(.80 - .70) - .08 < (p_1 - p_2) < (.80 - .70) + .08$$

$$+.02 < (p_1 - p_2) < .18$$

Conclusion based on the problem: The confidence interval does not contain zero

I am 90% confident at the .05 significance level that there **is a difference** in the **proportion** of FLC students and FLC faculty that own an I phone. From 2% to 18 % more FLC faculty own an I phone compared to FLC students

Negative Z Scores										
Standard Normal (Z) Distribution: Cumulative Area to the LEFT of Z										
				AREA	**Z Score**		**AREA**	**Z Score**	**0.08**	**0.09**
				0.0500	−1.645		0.0050	−2.575	0.0146	0.0143

Example 3

Are men below more likely to have a tattoo than women?

A random sample of 400 men showed that 176 of them had a tattoo. A random sample of 500 women showed that 185 of them had a tattoo. Construct a 92% confidence interval for the difference between the two population proportions. Does it appear there is a difference in the two proportions? **How can you tell?**

Sample from Population 1
below age 24

$n_1 = 400 \qquad x_1 = 176$

$$\hat{p}_1 = \frac{x_1}{n_1} = \frac{176}{400} = .44$$

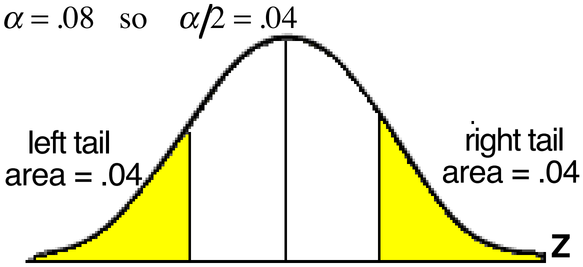
Sample from Population 2
between the ages of 24 and 30

$n_2 = 500 \qquad x_2 = 185$

$$\hat{p}_2 = \frac{x_2}{n_2} = \frac{185}{500} = .37$$

Find the critical value $+z_{\alpha/2}$

$\alpha = .08 \quad \text{so} \quad \alpha/2 = .04$



$-z_{\alpha/2}$ **= – 1.75** $\qquad +z_{\alpha/2}$ **= 1.75**

positive critical value

Find the Maximum Error

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

$$E = 1.75 \bullet \sqrt{\left(\frac{.44 \bullet .56}{400}\right) + \left(\frac{.37 \bullet .63}{500}\right)}$$

$$E = .06$$

Find the Confidence Interval to Estimate the value of $p_1 - p_2$

$$(\hat{p}_1 - \hat{p}_2) - E < (p_1 - p_2) + E < (\hat{p}_1 - \hat{p}_2) \quad \textbf{if} \quad \hat{p}_1 > \hat{p}_2$$

$$(.44 - .37) - .06 < (p_1 - p_2) < (.44 - .37) + .06$$

$$-.01 < (p_1 - p_2) < .13$$

Conclusion based on the problem: The confidence does contain zero.
I am 92% confident that there is **no difference** in the proportion of men who have a tattoo and the proportion of women who have a tattoo.

Negative Z Scores										
Standard Normal (Z) Distribution: Cumulative Area to the LEFT of Z										
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	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367

Example 4

Are the proportion students that pass the bar the same for McGeorge's Law School and Berkley Law School?

In 2009 a review of 300 randomly selected students from McGeorge Law school found that 102 of the 120 students who took the bar exam passed on the first attempt. A similar 2009 review of 300 randomly selected students from Berkley Law school found that 225 of the 300 students who took the bar exam passed on the first attempt. Construct a 90% confidence interval for the difference between the two proportions. Does it appear there is a difference in the two population proportions? **How can you tell?**

Population 1 McGeorge Law

x_1 = 102

n_1 = 120

$\hat{p}_1 = \frac{102}{120} = .85$ (the sample proportion)

Population 2 Berkley Law

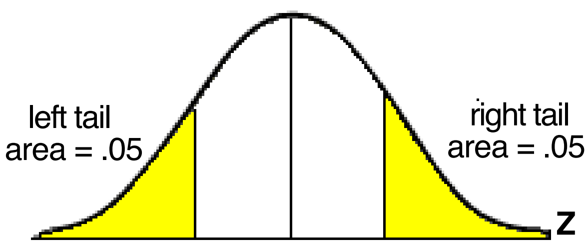
x_2 = 225

n_2 = 300

$\hat{p}_2 = \frac{225}{300} = .75$ (the sample proportion)

Find the critical value $+z_{\alpha/2}$

$\alpha = .10$ so $\alpha/2 = .05$



$-z_{\alpha/2}$ **= – 1.645** $+z_{\alpha/2}$ **= 1.645**

Find the Maximum Error

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

$$E = 1.645 \bullet \sqrt{\left(\frac{.85 \bullet .15}{120}\right) + \left(\frac{.75 \bullet .25}{300}\right)}$$

$$E = .07$$

Find the Confidence Interval to Estimate the value of $p_1 - p_2$

$$(\hat{p}_1 - \hat{p}_2) - E < (p_1 - p_2) < (\hat{p}_1 - \hat{p}_2) + E \quad \textbf{if} \quad \hat{p}_1 > \hat{p}_2$$

$$(.85 - .75) - .06 < (p_1 - p_2) < (.85 - .75) + .06$$

$$+.04 < (p_1 - p_2) < .16$$

Conclusion based on the problem:

The confidence does not contain zero. There is sufficient evidence at the $\alpha = .10$ **to reject** the claim that the proportion of **McGeorge students that passed the Bar Exam is equal** to the proportion of **McGeorge students that passed the Bar Exam**.

Negative Z Scores										
Standard Normal (Z) Distribution: Cumulative Area to the LEFT of Z										
Z scores of –3.5 or less use .0001					AREA	Z Score		AREA	Z Score	
					0.0500	–1.645		0.0050	–2.575	

