

Activity 2: Confidence Intervals for Proportions

Confidence Interval - Is an interval of values computed from sample data that is likely to include the true population value.

If we consider all possible randomly selected samples of the same size from a population, the confidence level is the fraction or percent of those samples for which the confidence interval includes the population parameter. Common confidence levels are 80%, 90%, 95% and 99%.

On March 4, 1999, *New England Journal of Medicine* published a research article by Dr. Douglas Jorenby and colleagues comparing different methods used to stop smoking.

Treatment	Subjects participating	Subjects not smoking after Six Months
Placebo	160	30
Nicotine patch	244	52
Zyban	244	85
Zyban and nicotine patch	245	95

1. Let's look at how the confidence intervals change width for the same sample data as the confidence level changes. Use the nicotine patch only data.

a) Calculate the 80% confidence interval.

b) Calculate the 90% confidence interval.

c) Calculate the 95% confidence interval.

d) Calculate the 99% confidence interval.

e) What is happening the widths of the intervals as the confidence level increases? Why do you think this is occurring?

2. Calculate the 95% confidence intervals for each of the four treatments:

Treatment	Approximate 95% CI
Placebo only	
Nicotine patch	
Zyban	
Zyban and nicotine patch	

Sketch the graphs of the four CI's on the same axis below letting the horizontal axis be the proportions not smoking after 6 months:

According to Utts and Heckard (2002) in *Minds on Statistics* the two principles for using confidence intervals to Guide Decision Making are:
Principle 1: A value not in the confidence interval can be rejected as a possible value of the population proportion. A value in the confidence interval is an "acceptable" possibility for the value of a population proportion.
Principle 2: When the confidence intervals for proportions in two different populations do not overlap, it is reasonable to conclude that the two populations are different.

Using the CI's from question 2,

a) Is Zyban effective? Why or why not?

b) Is there a significant difference between the Zyban only and Zyban with nicotine patch groups? Why or why not?

c) Is there a significant difference between the placebo and the nicotine patch only groups? Why or why not?

Keystrokes for the *fx-9750G Plus*

From the Main Menu, press 2 for STAT.

If there are data in List 1, follow these directions:

- *Press F6 (make sure that the highlighted cell is List 1 by pressing the right/left arrow)*
- *Press F4 (delete all) then press F1 (yes)*

Find Confidence Intervals for Proportions: In the STAT mode

- Press F4 for Interval
- Press F1 for Z, then press F3 for 1-proportion
- Enter the level of confidence as a decimal value
- Enter the number of participants not smoking after 6 months for x value since these are considered the "successes" in the proportion calculation.
- Enter the number of subjects for the n value since these are the considered the sample size in the proportion calculation.
- Press EXECUTE to calculate the confidence interval

Activity 3: Hypothesis Test for a Proportion

Using the same data from Activity 2, we will now conduct hypothesis testing and see if our conclusions are the same as in Activity 2 using confidence intervals.

We can use a z -test for a sample if it is a random sample from a population and if n (sample size) times p_0 (calculated proportion) as well as n times $(1 - p_0)$ are both at least 10.

1. Make sure these assumptions are true for all four treatments. Assume the study was a random sample for this data.

2. Now conduct a hypothesis test to see if the placebo and the nicotine patch only produce a statistically significant result.

a) Determine null and alternative hypotheses:

b) Use the calculator to perform the two sample z proportion hypothesis test.

z -value =

p -value =

c) What conclusions can you make and why?

3. Now conduct a hypothesis test to see if the Zyban only and the Zyban and nicotine patch together produce a statistically significant result.

a) Determine null and alternative hypotheses:

b) Use the calculator to perform the two sample z proportion hypothesis test.

z -value =

p -value =

c) What conclusions can you make and why?

4. Now conduct a hypothesis test to see if the nicotine patch only and Zyban only produce a statistically significant result.

a) Determine null and alternative hypotheses:

b) Use the calculator to perform the two sample z proportion hypothesis test.

z -value =

p -value =

c) What conclusions can you make and why?

5. Compare you hypothesis testing results from this activity with the results from activity 2.

Keystrokes for the fx-9750G Plus

From the Main Menu, press 2 for STAT.

If there are data in List 1, follow these directions:

- *Press F6 (make sure that the highlighted cell is List 1 by pressing the right/left arrow)*
- *Press F4 (delete all) then press F1 (yes)*

Calculate Hypothesis Tests for two sample Proportions: In the STAT mode

- Press F3 for Tests
- Press F1 for Z, then press F4 for 2-proportion
- Enter the alternative hypothesis
- Enter the number of participants not smoking after 6 months for x value since these are considered the "successes" in the proportion calculation.
- Enter the number of subjects for the n value since these are the considered the sample size in the proportion calculation.
- Press EXECUTE to calculate the confidence interval

Answers for 3rd quarter activities:

ACTIVITY 1. Answers will vary depending on data. The student social security number simulation should produce a uniform distribution. The means should produce a normal distribution and the mean and standard deviation of the means should be very close to the CLT values.

ACTIVITY 2:

1. a) .179 to .247

b). .170 to .256

c). .162 to .264

d). .146 to .281

5. As the confidence level gets higher the width increases. As you try to be more accurate with the confidence level the width increases because of sampling error.

2. Placebo only: .127 to .248

Nicotine patch: .162 to .264

Zyban: .289 to .408

Zyban and nicotine patch: .327 to .449

a) Zyban is effective. The two Zyban groups confidence intervals did not overlap the two groups not using Zyban.

b) The nicotine patch does not appear to be particularly effective. The placebo and the nicotine patch groups overlap so you cannot conclude the patch is better than the placebo.

c) There is also an overlap between the Zyban only and Zyban plus nicotine patch so you cannot conclude the patch helps the Zyban.

ACTIVITY 3:

1. Placebo: $.1875 \times 893 = 167.4$

Nicotine patch only: $.213 \times 893 = 190.3$

Zyban only : $.348 \times 893 = 311$

Zyban and nicotine patch: $.388 \times 893 = 346$

2. a) $H_0: p_1 = p_2$

$H_a: p_1$ does not equal p_2

b) z-value = -.626

p-value = .53126

c) You can conclude that there is not a statistically significant difference between the placebo group and the nicotine patch group. The p-value is greater than .05.

3 a) $H_0: p_1 = p_2$

$H_a: p_1$ does not equal p_2

b) z-value = -.90313

p-value = .36645

c) You can conclude that there is not a statistically significant difference between the Zyban only and the Zyban with nicotine patch group. The p-value is greater than .05.

4. a) $H_0: p_1 = p_2$

$H_a: p_1$ does not equal p_2

b) z-value = -3.2948

p-value = .00098476

c) You can conclude that there is a statistically significant difference between the nicotine patch group and the Zyban group. The p-value is less than .01.

5. The hypothesis testing results and confidence interval results agree. They show that the Zyban appears to be effective in this study and that the nicotine patch does not appear to be effective.

