MATH 3307
Lesson 27

Confidence Interval from a Proportion

Before any inferences can be made about a proportion, certain conditions must be satisfied:

1. The sample must be an SRS from the population of interest.
2. The population must be at least 10 times the size of the sample.
3. The number of successes and the number of failures must each be at least 10 (both $n\hat{p} \geq 10$ and $n(1-\hat{p}) \geq 10$).

Sample Statistics

The sample statistic for a population proportion is $\hat{p}$, so based on the formula for a CI, we have

$$\hat{p} \pm margin\ of\ error$$

How do we find the margin of error if it is not given to us?
The margin of error is equal to the **critical value** (a number based on our level of confidence) and the **standard deviation** (or **standard error** when needed) of the statistic.

Critical Value: When the distribution is assumed to be normal, our critical value is found from the z table (or using invNorm on calculator or qnorm in R). If it is not normal, we will use the t distribution (discussed later).

Standard Deviation/Error: When working with proportions, the standard deviation of the statistic $\hat{p}$ is $\sqrt{p(1-p)/n}$. Since p is unknown, we will use the standard error. To calculate the standard error of $\hat{p}$, use the formula $\sqrt{\hat{p}(1-\hat{p})/n}$.

The formula for confidence intervals of proportions:

This is taken from the Formula Sheet for Test 3, already posted on the casa calendar!

Confidence Intervals

General formula: *statistic ± margin of error*

One-sample z-test: $\bar{x} \pm z^* \dfrac{\sigma}{\sqrt{n}}$

Two-proportion z-test: $(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\dfrac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \dfrac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

One-sample t-test: $\bar{x} \pm t^* \dfrac{s}{\sqrt{n}}$

One-proportion z-test: $\hat{p} \pm z^* \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$

Two-sample z-test: $(\bar{x}_1 - \bar{x}_2) \pm z^* \sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$

Two-sample t-test: $(\bar{x}_1 - \bar{x}_2) \pm t^* \sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}$

R Studio / Calculator Assistance!

For a confidence level of 95%, you are interested in the middle 95% of a standard deviation:

0.5 + 0.95/2

½ + 0.95/2

1.95/2

Essentially, you will always use:

qnorm(1.##/2)

or

InvNorm(1.##/2)

Example:

In the first eight games of this year's basketball season, Lenny made 25 free throws in 40 attempts.
  a. What is $\hat{p}$, Lenny's sample proportion of successes?
  b. Find and interpret the 90% confidence interval for Lenny's proportion of free-throw success.

Sometimes we are asked to find the minimum sample size needed to produce a particular margin of error given a certain confidence level. When working with a one-sample proportion, we can use the formula:

$$ME = z * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Note: If a preliminary proportion (p) is not given in the problem, use p = 0.50

Example:
It is believed that 35% of all voters favor a particular candidate. How large of a simple random sample is required so that the margin of error of the estimate of the percentage of all voters in favor is no more than 3% at the 95% confidence level?

Changes in the Confidence Interval:

The following changes will result in an *increase* in the width of the confidence interval:

Increase of Margin of Error

Increase of Confidence Level

Change in sample proportion nearer to $\hat{p} = 0.50$

Decrease of Sample Size

A simple random sample, of size 150, is selected from a large population. It is determined that, within this sample, 85% of people own their own vehicles.

Are the number of successes and failures appropriate for this method?

Determine the standard deviation or standard error for the sample selected.

Determine the margin of error with a confidence level of 90%.

Determine the confidence interval for this statistic.

If the sample size were increased, how would the width of the confidence interval change?

Interpret the confidence interval.