Proportion Notation

p Population proportion. This is what we will be estimating.

$\hat{p}$ Sample proportion of successes. This is also known as a point estimate for the population parameter. Read this as "p-hat." This is calculated by $\frac{x}{n}$, where x is number of successes and n is the sample size. Remember, a success is defined as finding what you are looking for, even if what you are looking for is a negative action.

$\hat{q}$ Sample proportion of failures. Read this as "q-hat." This is calculated by $1-\hat{p}$. Remember, a failure is defined as not finding what you are looking for.

To make a prediction about p, simply calculate $\hat{p}N$, where N is the population size.

Confidence Intervals

We can make predictions about the population proportion using a number of different samples, and each time we are likely to get a different prediction.

We want to be confident about our prediction. So, we construct a range of proportions such that we are confident that the actual population proportion is within that range.

How confident do we want to be? Common confidence levels are 90%, 95%, and 99%.

To say that we are 90% confident that the actual population proportion falls within the range we calculated means the same as 90% of our point estimates will fall within that range.

To calculate the confidence interval, we will be using the Standard Normal Curve, Table A-2, and the following formulas:

$\hat{p}-E<p<\hat{p}+E$ where $E=z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$

E is called the margin of error.

Confidence Intervals (cont.)

In order to be able to use the Standard Normal Curve, we need to ensure that the distribution is normal. We do that by requiring:

$n\hat{p} > 5$ and $n\hat{q} > 5$

We can get $z_{\alpha/2}$ using the Standard Normal Curve and Table A-2:

1. Calculate α using $1 - CL$.
2. For these confidence intervals, the α needs to be split evenly between both tails of the curve, so calculate $\alpha/2$ and write that in each tail of your curve.
3. Look up $\alpha/2$ in the body of the Negative z-Scores portion of Table A-2.
4. Ignoring the sign, add together corresponding the z-Row and z-Column values to get the z-score.

Determining Sample Size

Suppose we want to know what size our sample should be in order to be sure our point estimate will be within a specific confidence interval?

Simply solve the error formula for n:

$$n = \frac{[z_{\alpha/2}]^2 \hat{p}\hat{q}}{E^2}$$

If $\hat{p}$ is unavailable, we can replace $\hat{p}\hat{q}$ with 0.25