Chapter 19: Estimating with Confidence

Suppose I want to know what proportion of teenagers typically goes to the movies on a Friday night.

Suppose I take an SRS of 25 teenagers and calculate the sample proportion to be \( \hat{p} = 0.40 \).

The sample proportion _____ is an unbiased estimator of the unknown population proportion _____, so I would estimate the population proportion to be approximately ____. However, using a different sample would have given a different sample proportion, so I must consider the amount of variation in the sampling model for ____.  

Based on one sample, it would _____ be correct to conclude that 40% of all teenagers typically go to the movies on a Friday night.

But don’t despair!... based on my one sample, I can come up with an ____________ that *may* contain the true proportion of teenagers who typically go to the movies on a Friday night.

Not only will I tell you what that interval is, but I will also tell you how ____________ I am that the true proportion falls somewhere in that interval.

Remember...

- The sampling model for \( \hat{p} \) is ______________ assuming __________ and __________.
- The mean of the sampling model is _____.
- The standard deviation of the sampling model is __________ assuming the population size is at least _____ times larger than the sample size \( (N \geq 10n) \).

Since we don’t know \( p \), we cannot calculate the standard deviation of the sampling model. We can, however, use _____ to estimate the value of \( p \) and calculate the standard error \( \sqrt{\frac{p\hat{q}}{n}} \) instead.

So the standard error for the sampling model for the proportion of teenagers who typically go to the movies on a Friday night is:

According to the 68-95-99.7 Rule, _____ of all possible samples of size 25 will produce a statistic \( \hat{p} \) that is within _____ standard errors of the mean of our sampling model.

This means that, in our example, 95% of the \( \hat{p} \)’s will be between __________ and __________.

So the distance between the actual _____ value and the statistic _____ will usually (95% of the time) be less than or equal to ________.

Thus, in 95% of our samples, the interval between \( \hat{p} - 0.196 \) and \( \hat{p} + 0.196 \) will contain the parameter \( p \).
We say that the ________________ is 0.196

For our sample of 25 teenagers, \( \hat{p} = 0.40 \). Because the margin of error is 0.196, then we are 95% confident that the true population proportion lies somewhere in the interval ________________, or ________________.

The interval [0.21, 0.59] is called a ____ ________________ ________________ because we are 95% ________________ that the true proportion of teenagers who typically go to the movies on a Friday night is between about ____ and ____.

CAUTION!!

This does NOT mean the probability that \( p \) is between 0.21 and 0.59 is 95%. ________________ does not mean the same thing as ________________.

We ________________ calculate the probability that \( p \) is within a given interval without using a Normal model, and we ________________ draw a Normal model because we don’t know the center, \( p \).

If you assume that \( p = \hat{p} \), then \( P(x_1 \leq p \leq x_2) \) is either ______ or ______.

So... maybe you’re not happy with the interval we constructed. Too wide? Would you prefer a more precise conclusion? One way of changing the length of the interval is to change the ________________ ________________.

So how do we construct 90% confidence intervals? 99% confidence intervals? C% confidence intervals?

Since the sampling model of the sample proportion \( \hat{p} \) is ________________ ________________, we can use normal calculations to construct confidence intervals.

- For a 95% confidence interval, we want the interval corresponding to the ________________ 95% of the normal curve.

- For a 90% confidence interval, we want the interval corresponding to the ________________ 90% of the normal curve.

- And so on...

If we are using the standard normal curve, we want to find the interval using ________________.

Suppose we want to find a 90% confidence interval for a standard normal curve. If the middle 90% lies within our interval, then the remaining ______ lies ________________ our interval. Because the curve is symmetric, there is _____ below the interval and _____ above the interval. Find the ________________ with area 5% below and 5% above.

These \textit{z-values} are denoted __________. Because they come from the standard normal curve, they are centered at mean ____.
is called the _________________, with probability p lying to its ___________ under the standard normal curve.

To find the upper critical p value, we find the complement of C and divide it in half, or find:

For a 95% confidence interval, we want the z-values with upper p critical value __________.

For a 99% confidence interval, we want the z-values with upper p critical value __________.

Remember that z-values tell us how many _______________ _______________ we are above or below the mean.

To construct a 95% confidence interval, we want to find the values __________ standard deviations below the mean and 1.96 standard deviations above the mean, or:

Using our sample data, this is \( \hat{p} \pm 1.96 \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \), assuming the population is at least _______ times as large as the sample size, ________.

In general, to construct a level C confidence interval using our sample data, we want to find:

The margin of error is \( z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \). Note that the margin of error is a positive _________. It is not an interval.

We would like __________ confidence and a __________ margin of error.

A higher confidence level means a higher percentage of all samples produce a statistic close to the true value of the parameter. Therefore we want a __________ level of confidence.

A smaller margin of error allows us to get closer to the true value of the parameter (length of the interval is small), so we want a __________ margin of error.

So how do we reduce the margin of error?

- __________ the confidence level (by decreasing the value of \( z^* \))
- __________ the standard deviation
- __________ the sample size. To cut the margin of error in half, increase the sample size by __________ times the previous size.

You can have __________ confidence and a __________ margin of error if you choose the right sample size.

To determine the sample size \( n \) that will yield a confidence interval for a population proportion with a specified margin of error \( m \), set the expression for the margin of error to be equal to \( m \) and solve for \( n \). Always round \( n \) up to the next greatest integer.
CAUTION!!

These methods only apply to certain situations. In order to construct a level C confidence interval using the formula $\hat{p} \pm z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$, for example, the data must come from a ________________ sample. Also, we want to eliminate (if possible) any ________________.

The margin of error only covers random sampling errors. Things like ________________, ________________, and ________________ can cause additional errors.

*Remember, if you are asked to construct a Confidence Interval, you must PANIC!!

P: _________________________________

A: _________________________________

N: _________________________________

I: _________________________________

C: _________________________________