8.1 Confidence Intervals: The Basics

HW: p. 481 (5-13 odd, 17, 19-24)

The Idea of a Confidence Interval

- A _______ _______________ is a statistic that provides an estimate of a population parameter.
- The value of that statistic from a sample is called a __________ __________________.
- Ideally, a point estimate is our "_______ ________" at the value of an unknown parameter.
- Unfortunately, that estimate will vary from sample to sample.
- A _______________ ____________ takes that variation into account to provide an interval of plausible values, based on the statistic, for the true parameter.

Example: Identify the point estimator and the point estimate.
The math department wants to know what proportion of students own a graphing calculator, so they take a random sample of 1001 students and find that 28 own a graphing calculator.

Point Estimator:

Point Estimate:

Components
All confidence intervals have two main components.

1. An _______________ based on the estimate that includes a margin of error

$$estimate \pm margin\ of\ error$$

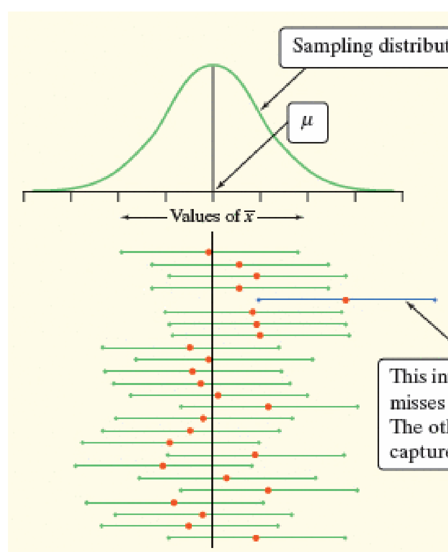
   *The ___________ ___ _____________ tells how close the estimate tends to be to the unknown parameter in repeated random sampling.

2. A __________________ ________ C that reports the success rate of the method used to construct the interval in capturing the parameter in repeated constructions.

   *We usually choose a confidence level of ______ or higher, but the most common choice is _______.

Confidence Intervals: The Basics

- The confidence level describes the ___________________ associated with a sampling method.
- Suppose we used the sampling method to select different samples and to compute a different interval estimate for each sample.
- Some interval estimates would ___________ the true population parameter and some would ______.
- A 95% confidence level means that we would expect 95% of the interval estimates to include the population parameter.



Interpreting

Confidence Level: To say that we are 95% confident is shorthand for
"95% of all possible samples of a given size from this population will result in an interval that captures the unknown parameter."

Confidence Interval: To interpret a $C\%$ confidence interval for an unknown parameter, say:
"We are $C\%$ confident that the interval from _____ to _____ captures the actual value of the ______(population parameter in context)."
***This does NOT mean that we are C% confident that our interval contains the parameter!!!*

Check Your Understanding, *p. 476*
How much does the fat content of Brand X hot dogs vary? To find out researchers measured the fat content (in grams) of a random sample of 10 Brand X hot dogs. A 95% confidence interval for the population standard deviation σ is 2.84 to 7.55.

1. Interpret the confidence interval.
   We are ________ confident that the interval from ______ to _______ captures the ______ standard deviation of the fat content of Brand X hot dogs.
2. Interpret the confidence level.
   In 95% of ____ ________________ ___________ of 10 Brand X hot dogs, the resulting confidence interval would capture the true standard deviation.
3. True or false: The interval from 2.84 to7.55 has a 95% chance of containing the actual population standard deviation σ. Justify your answer.
   _________. *We ___________ know whether our sample is one of the 95% for which the interval catches σ or whether it is one of the unlucky 5%. We are just using a method that gives us correct results _____ of the time.*

Constructing Confidence Intervals

Step 1: Check Conditions

Random **-** The data must come from a well-designed _________ _____________ or randomized experiment.
*For the rest of this chapter, we will limit ourselves to settings that involve random sampling. We will discuss inference for randomized experiments in Chapter 9.

Normal - The ___________ ________________ of the statistic is approximately Normal.
For sample means:
- If the population distribution is Normal, so is the sampling distribution.
- If the population distribution is NOT Normal, then the CLT tells us the sampling distribution will be approximately Normal if $n \geq 30$.

For sample proportions:
- We can use Normal approximations for the sampling distribution as long as $np \geq 10$ and $n(1-p) \geq 10$.

Independent - Check to make sure _________________ observations are independent.
When sampling without replacement, the sample size *n* should be no more than 10% of the population size *N* (10% condition) to use our formula for the standard deviation of a statistic.

Step 2: Construct the Interval

Use the following formula:

$$statistic \pm (critical\ value) \cdot (standard\ deviation\ of\ statistic)$$

- The ______________ is the point estimator for the parameter.
- The ____________ __________ is determined based on the confidence level C.
- The ______________ _______________ is based on the sampling distribution of the statistic.

*This formula is on the AP Exam Formula Sheet.

Step 3: Interpret the Interval

To interpret a confidence interval, use the language learned earlier.
Make sure to ____________________ the wording of the interpretation!

Confidence Interval Goal

Our goal with confidence intervals is to provide as ______________ an ______________ as possible.
That is, we wish to construct a ______________ interval that we are confident captures the ________________ of interest.

The margin of error gets smaller as:

- the confidence level C _________________
- the sample size n ________________

Example

A large company is interested in developing a new bakeware product for consumers. In an effort to determine baking habits of adults, a researcher selects a random sample of 50 addresses in a large, Midwestern, metropolitan area. She calls each selected home in the late morning to collect information on their baking habits. The proportion of adults who bake at least twice a week is calculated and a 90% confidence interval is constructed.

Have each of the conditions for constructing a confidence interval been met? If not, what are the implications on the interpretation of the interval?

Random: ____*! The homes were selected randomly, but they were called ______-____________. Likely those who were home were homemakers who would be more likely to bake at least twice a week. The 90% confidence interval may ___________________ the true proportion of adults who bake at least twice a week.*

Normal: ______. $np \geq 10$ and $n(1-p) \geq 10$.

Independent: ______. *The 10% condition is met (there are more than ________________ adults in the population).*