Freeport AP Statistics
Chapter 8: Estimating with Confidence
8.1 Confidence Intervals: The Basis

OBJECTIVE(S):
- Students will learn how to interpret a confidence level.
- Students will learn how to interpret a confidence interval in context.
- Students will understand that a confidence interval gives a range of plausible values for the parameter.
- Students will understand why each of the three inference conditions – Random, Normal, and Independent – is important.
- Students will learn to explain how practical issues like nonresponse, under coverage, and response bias can affect the interpretation of a confidence interval.

Point Estimator –

Point Estimate -

1. In each of the following setting, determine the point estimator you would use and calculate the value of the point estimate.

   a. The makers of a new golf ball want to estimate the median distance the new balls will travel when hit by a mechanical driver. They select a random sample of 10 balls and measure the distance each ball travels after being hit by the mechanical driver. Here are the distances (in yards):

285	286	284	285	282	284	287	290	288	285

   b. The golf ball manufacturer would also like to investigate the variability of the distance travelled by the golf balls by estimating the interquartile range.

c. The math department wants to know what proportion of its students owns a graphing calculator, so they take a random sample of 100 students and find that 28 own a graphing calculator.

Confidence Interval

-

    Margin of Error -

- **Confidence Level, *C*, -**

NOTE: Margin of Error in a confidence interval accounts for variability due only to random selection or random assignment; it does not compensate for any bias in the data collection process. If the sample wasn't random or the questions were worded poorly, you shouldn't have much confidence at all!!!!

2. ACTIVITY *The* Confidence interval *applet* (p. 482)
    www.whfreeman.com/tps5e   Click on our book; click on "statistical applets"; click on "confidence intervals"

    a. Click "Sample" to choose an SRS and display the resulting confidence interval. What is the computer doing? Did the interval capture the population mean μ (what the applet calls a "hit")? Do this a total of 10 times. How many of the intervals captured the population mean μ ?

> b. Reset the applet. Click "Sample 50" to choose 50 SRSs and display the confidence intervals based on those samples. How many captured the parameter μ? Keep clicking "Sample 25" and observe the value of "Percent hit." What do you notice?

Confidence Levels v. Confidence Intervals (*Mantras*)

Confidence Level
"If this sampling process were repeated many times, approximately ____% of the resulting confidence intervals would capture the _____ (true unknown parameter)."

Confidence Interval
"We are ____% confident that the interval from _____ to _____ (units) captures the actual value of the _____ (population parameter in context)."

NOTE: Every time you construct a confidence interval, you are required to interpret it using the Mantra. Only interpret Confidence Level when you are asked.

Also, 95% confidence is not the same as a 95% probability or chance the interval contains the parameter. Probability relates to future events and confidence intervals are estimations of past known/unknown parameters.

3. According to www.gallup.com, on August 13, 2010, in a sample of 1500 adults, the 95% confidence interval for the true proportion of Americans who approved of the job Barack Obama was doing as president was 0.44 ± 0.03.
   a. Interpret the confidence interval.

   b. Interpret the confidence level.

4. **ACTIVITY** *The* Confidence interval *applet* (p. 482)
   www.whfreeman.com/tps5e   Click on our book; click on "statistical applets"; click
   on "confidence intervals"
   a. Set the confidence level at 95% and click "Sample 25." Change the
      confidence level to 99%. What happens to the length of the confidence
      intervals? To the "Percent hit"?

   b. Now change the confidence level to 90%. What happens to the length of the
      confidence intervals? To the "Percent hit"?

   c. Finally, change the confidence level to 80%. What happens to the length of
      the confidence intervals? To the "Percent hit"?

   d. Go to page 486 and answer #5-7 below (or on the facing page).

5. Oxides of nitrogen (called NOX for short) emitted by cars and trucks are important
   contributors to air pollution. The amount of NOX emitted by a particular model
   varies from vehicle to vehicle. For one light-truck model, NOX emissions vary with
   mean μ that is unknown and standard deviation $\sigma = 0.4$ gram per mile. You test an
   SRS of 50 of these trucks. The sample mean NOX level $\bar{x}$ estimates the unknown μ.
   You will get different values of $\bar{x}$ if you repeat your sampling.
   a. Describe the shape, center, and spread of the sampling distribution of $\bar{x}$.

b. Sketch the sampling distribution $\overline{x}$. Mark its mean and the values one, two, and three standard deviations on either side of the mean.

c. According to the 68-95-99.7 rule, about 95% of all values of $\overline{x}$ lie within a distance *m* of the mean of the sampling distribution. What is *m*? Shade the *region on axis* of your sketch that is within *m* of the mean.

d. Whenever $\overline{x}$ falls in the region you shaded, the unknown population mean μ lies in the confidence interval $\overline{x} \pm m$. For what percent of all possible samples does the interval capture μ?

e. Below the sketch, choose one value of $\overline{x}$ inside the shaded region and draw its corresponding confidence interval. Do the same for one value of $\overline{x}$ outside the shaded region. What is the most important difference between these intervals?

6. A Gallup Poll in November 2008 found that 59% of the people in its sample said "Yes" when asked, "Would you like to lose weight?" Gallup announced: "For results based on the total sample of national adults, one can say with 95% confidence that the margin of (sampling) error (MOE) is ± 3 percentage points."

    a. Explain what the margin of error means in this setting.

    b. State and interpret the 95% confidence interval.

    c. Interpret the confidence level.

    d. As Gallup indicates, the 3 percentage point margin of error for this poll includes only sampling variability (what they call "sampling error"). What other potential sources of error (Gallup calls these "nonsampling errors") could affect the accuracy of the 59% estimate?

Calculating a Confidence Interval

7. What effects does increasing the sample size have on the:

    a. Standard Deviation of the Statistic?

    b. MOE?

    c. Confidence Interval?

<u>**DAY 1**</u>

8. Many teens have posted profiles on sites such as Facebook. A sample survey asked random samples of teens with online profiles if they included false information in their profiles. Of 170 younger teens (ages 12 to 14) polled, 117 said "Yes." Of 317 older teens (ages 15 to 17) polled, 152 said "Yes." A 95% confidence interval for the difference in the population proportions (younger teens – older teens) is 0.120 to 0.297.
   a. Interpret the confidence interval.

   b. Interpret the confidence level.

9. The admissions director from Big City University found that (107.8, 116.2) is a 95% confidence interval for the mean IQ score of all freshmen. Comment on whether or not each of the following explanations is correct.
   a. There is a 95% probability that the interval from 107.8 to 116.2 contains μ.

   b. There is a 95% chance that the interval (107.8, 116.2) contains $\bar{x}$.

    c.   This interval was constructed using a method that produces intervals that capture the true mean in 95% of all possible samples.

    d.   95% of all possible samples will contain the interval (107.8, 116.2).

    e.   The probability that the interval (107.8, 116.2) captures μ is either 0 *or* 1, but we don't know which.

10. An online poll posed the following question:

> *It is now possible for school students to log on to Internet sites and download homework. Everything from book reports to doctoral dissertations can be downloaded free for a fee. Do you believe that giving a student who is caught plagiarizing an F for their assignment is the right punishment?*

Of the 20,125 people who responded, 14,793 clicked "Yes." That's 73.5% of the sample. Based on this sample, a 95% confidence interval for the percent of the population who would say "Yes" is $73.5\% \pm 0.61\%$. Which of the three inference conditions is violated? Why is this confidence interval worthless?

<div style="border: 1px solid black;">

Freeport AP Statistics
Chapter 8: Estimating with Confidence
8.2 Estimating a Population Proportion

</div>

OBJECTIVE(S):
- Students will learn how to construct and interpret a confidence interval for a population proportion.
- Students will learn how to determine critical values for calculating a confidence interval using a table or calculator.
- Students will learn how to carry out the steps in contstructing a confidence interval for a population proportion: PANIC.
- Students will learn how to determine the sample size required to obtain a level *C* confidence interval for a population proportion with a specified MOE.
- Students will learn to understand how the MOE of a confidence interval changes with the sample size and the level of confidence *C*.
- Students will learn to understand why each of the three inference conditions – Random, Normal, and Independent – is important.

Conditions for Constructing a Confidence Interval

-

-

-

Standard Error –

One-Sample *z* Interval for a Population Proportion

11. Find the critical value z^* for:
μ

    a.   a 96% confidence interval. Assume that the normal condition is met.

    b. a 92% confidence interval. Assume that the normal condition is met

When constructing confidence intervals, you must **PANIC**!!!!!

    P

    A

    N

    I

    C

12. Ms. Smith's class took an SRS of 102 pennies and discovered that 57 of the pennies were more than 10 years old.

     a. Calculate and interpret a 99% confidence interval for p = the true proportion of pennies from the collection that are more than 10 years old.

*P*arameter of interest

*A*re the conditions met?
- **Random:**

- **Normal:**

- **Independent:**

*N*ame of the interval

*I*nterval

*C*onclusion

b. Is it plausible that exactly 60% of all the pennies in the collection are more than 10 years old? Explain.

13. In her first-grade social studies class, Jordan learned that 70% of the Earth's surface was covered in water. She wondered if this was really true and asked her dad for help. To investigate, he tossed an inflatable globe to her 50 times, being careful to spin the globe each time. When she caught it, he recorded where her right index finger was pointing. In 50 tosses, her finger was pointing to water 33 times. Construct and interpret a 95% confidence interval for the proportion of the Earth's surface that is covered in water.

*P*arameter of interest

*A*re the conditions met?
- **Random:**

- **Normal:**

- **Independent:**

*N*ame of the interval

*I*nterval

*C*onclusion

Sample Size for Desired Margin of Error

14. Suppose that you wanted to estimate p = the true proportion of students at your school who have a tattoo with 95% confidence and a margin of error of no more than 0.10. Determine how many students should be surveyed to estimate p within 0.10 with 95% confidence.

15. A college student organization wants to start a nightclub for students under the age of 21. To assess support for this proposal, they will select an SRS of students and ask each respondent if he or she would patronize this type of establishment. They expect that about 70% of the student body would respond favorably. What sample size is required to obtain a 90% confidence interval with an approximate margin of error of 0.04? Show your work.

FROM THIS DAY FORWARD, WHENEVER YOU ARE ASKED TO CALCULATE AN INTERVAL ON A HOMWORK ASSIGNMENT, TEST, AP EXAM, AT LUNCH, ETC., YOU MUST PANIC!!!!!!!!

Freeport AP Statistics
Chapter 8: Estimating with Confidence
8.3 Estimating a Population Mean

OBJECTIVE(S):

- Students will learn how to construct and interpret a confidence interval for a population mean.
- Students will learn how to determine the sample size required to obtain a level *C* confidence interval for a population mean with a specified margin of error.
- Students will learn how to carry out the steps in constructing a confidence interval for a population mean: PANIC.
- Students will learn how to determine sample statistics from a confidence interval.
- Students will understand why each of the three inference conditions – Random, Normal, and Independent- is important.

One-Sample _z_ Interval for a Population Mean (_Never Used_)

Choosing Sample Size for a Desired Margin of Error When Estimating μ

-

-

-

16. Administrators at your school want to estimate how much time students spend on homework, on average, during a typical week. They want to estimate μ at the 90% confidence level with a margin of error of at most 30 minutes. A pilot study indicated that the standard deviation of time spent on homework per week is about 154 minutes. How many students need to be surveyed to estimate the mean number of minutes spent on homework per week with 90% confidence and a margin of error of at most 30 minutes?

17. High school students who take the SAT Math exam a second time generally score higher than on their first try. Past data suggest that the score increase has a standard deviation of about 50 points. How large a sample of high school students would be needed to estimate the mean change in SAT score to within 2 points with 95% confidence? Show your work.

The *t* Distributions; Degrees of Freedom

Three facts about the t distributions:

-

-

-

18. Suppose you wanted to construct a 90% confidence interval for the mean μ of a Normal population based on an SRS of size 10. What critical value t^* should you use?

19. What critical value t^* from Table B should be used for a confidence interval for the population mean in each of the following situations?
    a. A 90% confidence interval based on $n = 12$ observations.

    b. A 95% confidence interval from an SRS of 30 observations.

Standard Error of the Sample Mean $\bar{x}$ (SEM) -

The One-Sample t Interval for a Population Mean

20. A study of commuting times reports the travel times to work of a random sample of 20 employed adults in New York State. The mean is $\bar{x} = 31.25$ minutes, and the standard deviation is $s_x = 21.88$ minutes. What is the standard error of the mean? Interpret this value in context.

Conditions for Inference about a Population Mean

- **Random:**

- **Normal:**

- **Independent:**

21. Several years ago, the U.S. Agency for International Development provided 238,300 metric tons of corn-soy blend (CSB) for emergency relief in countries throughout the world. CSB is a highly nutritious, low-cost fortified food. As part of a study to evaluate appropriate vitamin C levels in this food, measurements were taken on samples of CSB produced in a factory. The following data are the amounts of vitamin C, measured in milligrams per 100 grams (mg/100 g) of blend, for a random sample of size 8 from one production run:

26	31	23	22	11	22	14	31

Construct and interpret a 95% confidence interval for the mean amount of vitamin C μ in the CSB from this production run.

*P*arameter of interest

*A*re the conditions met?
- **Random:**

- **Normal:**

- **Independent:**

*N*ame of the interval

*I*nterval

*C*onclusion

22. As part of their final project in AP Statistics, Christina and Rachel randomly selected
    18 rolls of a generic brand of toilet paper to measure how well this brand could
    absorb water. To do this, they poured ¼ cup of water onto a hard surface and counted
    how many squares it took to completely absorb the water. Here are the results from
    their 18 rolls:

29	20	25	29	21	24	27	25	24
29	24	27	28	21	25	26	22	23

Construct and interpret a 99% confidence interval for μ = the mean number of
squares of generic toilet paper needed to absorb ¼ cup of water.

*P*arameter of interest

*A*re the conditions met?
- **Random:**

- **Normal:**

- **Independent:**

*N*ame of the interval

*I*nterval

*C*onclusion

Table B does not include every possible sample size. If you are insistent on using Table B rather than using your calculator and want to calculate the interval by hand, how do you determine the *df* to use when *n* is not listed?

23. The principal at a large high school claims that students spend at least 10 hours per week doing homework, on average. To investigate this claim, an AP Statistics class selected a random sample of 250 students from their school and asked them how long they spent doing homework during the last week. The sample mean was 10.2 hours and the sample standard deviation was 4.2 hours.
    a. Construct and interpret a 95% confidence interval for the mean time that students at this school spent doing homework in the last week.

*P*arameter of interest

*A*re the conditions met?
- **Random:**

- **Normal:**

- **Independent:**

*N*ame of the interval

*I*nterval

*C*onclusion

  b. Based on your interval in part a., what can you conclude about the principal's claim?

Robust-

NOTE: z procedures for proportions are not robust when the Normal condition is violated. That is, when the Normal condition is violated, the actual capture rate of a "95% confidence interval" for a population proportion may be very different from 95%.

Why do larger sample improve the accuracy of critical values from *the t* distributions when the population is not normal?

   •

   •

24. Trace metals found in wells affect the taste of drinking water, and high concentrations can pose a health risk. Researchers measured the concentration of zinc (in milligrams/liter) near the top and the bottom of 10 randomly selected wells in a large region. The data are provided in the table below.

Well:	1	2	3	4	5	6	7	8	9	10
Bottom:	0.430	0.266	0.567	0.531	0.707	0.716	0.651	0.589	0.469	0.723
Top:	0.415	0.238	0.390	0.410	0.605	0.609	0.632	0.523	0.411	0.612

      a. Construct and interpret a 95% confidence interval for the mean difference μ in the zinc concentrations from these two locations in the wells.

*P*arameter of interest

*A*re the conditions met?
- **Random:**

- **Normal:**

- **Independent:**

*N*ame of the interval

*I*nterval

*C*onclusion

b. Does your interval in part a. give convincing evidence of a difference in zinc concentrations at the top and bottom of wells in the region? Justify your answer.

CHAPTER 8 SUMMARY
Statistic $\pm$ MOE
Statistic $\pm$ (**Critical Value**)(**Standard Deviation**)

*P*arameter of interest

*A*re the conditions met?
- **Random:**

- **Normal:**

- **Independent:**

*N*ame of the interval

*I*nterval

*C*onclusion

Sample Size