Freeport AP Statistics

Chapter 7: Sampling Distribution

7.1 What is a Sampling Distribution?

OBJECTIVE(S):

- Students will learn how to distinguish between a parameter and a statistic.
- Students will learn to understand the definition of a sampling distribution.
- Students will learn to distinguish between population distribution, sampling distribution, and the distribution of sample data.
- Students will learn to determine whether a statistic is an unbiased estimator of a population parameter.
- Students will learn to understand the relationship between sample size and the variability of an estimator.

Parameter –

Statistic -

	Parameter	Statistic
Mean		
Variance		
Standard Deviation		
Proportion		

1. Identify the population, the parameter, the sample, and the statistic in each of the following settings.
   a. A pediatrician wants to know the 75^{th} percentile for the distribution of heights of 10-year-old boys, so she takes a sample of 50 patients and calculates Q_3 = 56 inches.

   b. A Pew Research Center Poll asked 1102 12- to 17-year-olds in the United States if they have a cell phone. Of the respondents, 71% said Yes.

   c. Each month, the Current Population Survey interviews a random sample of individuals in about 55,000 U.S. households. One of their goals is to estimate the national unemployment rate. In December 2009, 10.0% of those interviewed were unemployed.

   d. How much do gasoline prices vary in a large city? To find out, a reporter records the price per gallon of regular unleaded gasoline at a random sample of 10 gas stations in the city on the same day. The range (maximum – minimum) of the prices in the sample is 25 cents.

2. For each **boldface** number (1) state whether it is a parameter or a statistic and (2) use appropriate notation to describe each number; for example $p = 0.65$.
    a. Florida has played a key role in recent presidential elections. Voter registration records show that **41%** of Florida voters are registered as Democrats. To test a random digit dialing device, you use it to call 250 randomly chosen residential telephones in Florida. Of the registered voters contacted, **33%** are registered Democrats.

    b. A random sample of female college students has a mean height of **64.5 inches**, which is greater than the **63-inch** mean height of all adult American women.

Sampling Variability –

Sampling Distribution –

Population Distribution -

3. **See Figure 7.3 on p. 428.** What is the difference between a Distribution of Sample Data and a Sampling Distribution?

Unbiased Estimator -

Biased Estimator –

Variability of a Statistic -

4. Will increasing the sample size eliminate bias?

5. **p. 437 #10**.
   a. There is one dot on the graph at 62.4. Explain what this value represents.

   b. Describe the distribution. Are there any obvious outliers?

   c. Suppose that the average height of the 20 girls in the class's actual sample is $\overline{x} = 64.7$. What would you conclude about the population mean height μ for the 16-year-old females at the school? Explain.

   d. Make a graph of the population distribution.

6. **p. 438 #14**.
   a. Describe the approximate sampling distribution.

   b. Suppose that the minimum of an actual sample is 40 degrees Fahrenheit. What would you conclude about the thermostat manufacturer's claim? Explain.

7. A statistics teacher fills a large container with 1000 white and 3000 red beads and then mixes the beads thoroughly. She then has her students take repeated SRSs of 50 beads from the container. After many SRSs, the values of the sample proportion $\hat{p}$ of red beads are approximated well by a |Normal distribution with mean 0.75 and standard deviation 0.06.
   a. What is the population? Describe the population distribution.

   b. Describe the sampling distribution of $\hat{p}$. How is it different from the population distribution?

8. Just before a presidential election, a national opinion poll increases the size of its weekly random sample from the usual 1500 people to 4000 people.
   a. Does the larger random sample reduce the bias of the poll result? Explain.

   b. Does it reduce the variability of the result? Explain.

9. A study of the health of teenagers plans to measure the blood cholesterol levels of an SRS of 13- to 16-year-olds. The researchers will report the mean $\bar{x}$ from their sample as an estimate of the mean cholesterol level μ in this population.
   a. Explain to someone who knows no statistics what it means to say that $\bar{x}$ is an unbiased estimator of μ.

   b. The sample result $\bar{x}$ is an unbiased estimator the population mean μ no matter what size SRS the study chooses. Explain to someone who knows no statistics why a large random sample gives more trustworthy results than a small random sample.

Freeport AP Statistics

Chapter : Sampling Distribution

7.2 Sample Proportions

OBJECTIVE(S):

- Students will learn how to find the mean and standard deviation of the sampling distribution of a sample proportion $\hat{p}$ for an SRS of size n from a population having proportion p of successes.
- Students will learn how to check whether the 10% and Normal conditions are met in a given setting.
- Students will learn how to use Normal approximation to calculate probabilities involving $\hat{p}$.
- Students will learn how to use the sampling distribution of $\hat{p}$ to evaluate a claim about a population proportion.

Confusing Notations:

- $\mu_{\hat{p}}$ -
- $\sigma_{\hat{p}}$ -

Sampling Distribution of a Sample Proportion

-
-
-

10. Suppose a large candy machine has 15% orange candies (see **p. 435**).
    a. Would you be surprised if a sample of 25 candies from the machine contained 8 orange candies (that's 32% orange)? How about 5 orange candies (20% orange)? Explain.

    b. Which is more surprising: getting a sample of 25 candies in which 32% are orange or getting a sample of 50 candies in which 32% are orange? Explain.

11. Suppose a large candy machine has 15% orange candies. Imagine taking an SRS of 25 candies from the machine and observing the sample proportion $\hat{p}$ of orange candies.
    a. What is the mean of the sampling distribution of $\hat{p}$? Why?

    b. Find the standard deviation of the sampling distribution of $\hat{p}$. Check to see if the 10% condition is met.

c. Is the sampling distribution of $\hat{p}$ approximately Normal? Check to see if the Normal condition is met.

d. If the sample size were 75 rather than 25, how would this change the sampling distribution of $\hat{p}$?

12. In the game of Scrabble, each player begins by drawing 7 tiles from a bag containing 100 tiles. There are 42 vowels, 56 consonants, and 2 blank tiles in the bag. Cait chooses an SRS of 7 tiles. Let $\hat{p}$ be the proportion of vowels in her sample.

a. Is the 10% condition met in this case? Justify your answer.

b. Is the Normal condition met in this case? Justify your answer.

13. The Gallup Poll asked a random sample of 1785 U.S. adults whether they attended church or synagogue during the past week. Of the respondents, 44% said they did attend. Suppose that 40% of the U.S. adult population actually went to church or synagogue last week. Let $\hat{p}$ be the proportion of people in the sample who attended church or synagogue.
    a. What is the mean of the sampling distribution of $\hat{p}$? Why?

    b. Find the standard deviation of the sampling distribution of $\hat{p}$.

    c. Is the sampling distirubtion of $\hat{p}$ approximately Normal?

    d. Find the probability of obtaining a sample of 1785 adults in which 44% or more say they attended church or synagogue last week. Do you have any doubts about the result of this poll?

e. What sample size would be required to reduce the standard deviation of the sampling distribution to one-third the value you found above? Justify your answer.

14. Harley-Davidson motorcycles makeup 14% of all the motorcycles registered in the United States. You plan to interview an SRS of 500 motorcycle owners. How likely is your sample to contain 20% or more who own Harleys?

Freeport AP Statistics

Chapter 7: Sampling Distribution

7.3 Sample Means

OBJECTIVE(S):

- Students will learn how to find the mean and standard deviation of the sampling distribution of a sample mean $\bar{x}$ from an SRS of size *n*.
- Students will learn how to calculate probabilities involving a sample mean $\bar{x}$ when the population distribution is Normal.
- Students will learn how to explain how the shape of the sampling distribution of $\bar{x}$ is related to the shape of the population distribution.
- Students will learn how to use the central limit theorem to help find probabilities involving a sample mean $\bar{x}$.

15. Why do nurses/doctors typically measure a patient's blood pressure more than once when performing a physical?

Mean and Standard Deviation of the Sampling Distribution $\bar{x}$

-

-

<u>Notation/Symbols Matter (Define these symbols using words, not formulas)</u>

- μ -
- σ -
- p -
- $\hat{p}$ -
- $\bar{x}$ -
- $\mu_{\hat{p}}$ -
- $\sigma_{\hat{p}}$ -
- $\mu_{\bar{x}}$ -
- $\sigma_{\bar{x}}$ -

16. Suppose that the number of movies viewed in the last year by high school students has an average of 19.3 with a standard deviation of 15.8. Suppose we take an SRS of 100 high school students and calculate the mean number of movies viewed by the members of the sample.

    a. What is the mean of the sampling distribution of $\bar{x}$?

    b. What is the standard deviation of the sampling distribution of $\bar{x}$? Check whether the 10% condition is satisfied.

Sampling Distribution of a Sample Mean from a Normal Population

17. At the P. Nutty Peanut Company, dry-roasted, shelled peanuts are placed in jars by a machine. The distribution of weights in the jars is approximately Normal, with a mean of 16.1 ounces and a standard devaition of 0.15 ounces.
    a. Without doing any calculations, explain which outcome is more likely: randomly selecting a single jar and finding that the contents weigh less than 16 ounces or randomly selecting 10 jars and finding that the average contents weigh less than 16 ounces.

b. Find the probability of each event described above.

18. A grinding machine in an auto parts plant prepares axles with a target diameter μ = 40.125 millimeters (mm). The machine has some variability, so the standard deviation of the diameters is σ = 0.002 mm. The machine operator inspects a random sample of 4 axles each hour for quality control purposes and records the sample mean diameter $\overline{x}$.
    a. Assuming that the process is working properly, what are the mean and standard deviation of the sampling distribution of $\overline{x}$? Explain.

    b. How many axles would you need to sample if you wanted the standard deviation of the sampling distribution of $\overline{x}$ to be 0.0005 mm? Justify your answer.

19. A car company has found that the lifetime of its disc brake pads varies from car to car according to a Normal distribution with mean μ = 55,000 miles and standard deviation σ = 4500 miles. The company installs a new brand of brake pads on an SRS of 8 cars.
    a. If the new brand has the same lifetime distribution as the previous type of brake pad, what is the sampling distribution of the mean lifetime $\overline{x}$?

    b. The average life of the pads on these 8 cars turns out to be $\overline{x}$ = 51,800 miles. Find the probability that the sample mean lifetime is 51,800 or less if the lifetime distribution is unchanged. What conclusion would you draw?

20. The composite scores of individual students on the ACT college entrance examination in 2009 followed a Normal distribution with mean 21.1 and standard deviation 5.1.
    a. What is the probability that a single student randomly chosen from all those taking the test scores 23 or higher? Show your work.

    b. Now take an SRS of 50 students who took the test. What is the probability that the mean score $\bar{x}$ of these students is 23 or higher? Show your work.

Central Limit Theorem (CLT) – When it comes to statistics, this theorem is the Grand Daddy of Them All

When we are using the Normal curve to approximate probability, we must somehow conclude that our sampling distribution is ***APPROXIMATELY*** normal. The two ways to conclude normality regarding our sampling distribution are:

-
-

21. Suppose that the number of texts sent during a typical day by a randomly selected high school student follows a right-skewed distribution with a mean of 15 and a standard deviation of 35. Assuming that students at your school are typical texters, how likely is it that a random sample of 50 students will have sent more than a total of 1000 texts in the last 24 hours?

22. The number of lightning strikes on a square kilometer of open ground in a year has mean 6 and standard deviation 2.4. (These values are typical of much of the |United States.) The National Lightning Detection Network (NLDN) uses automatic sensors to watch for lightning in a random sample of 10 one-square-kilometer plots of land.
    a. What are the mean and standard deviation of $\overline{x}$, the sample mean number of strikes per square kilometer?

    b. Explain why you cannot safely calculate the probability that $\overline{x} < 5$ based on a sample size 10.

    c. Suppose the NLDN takes a random sample of n = 50 square kilometers instead. Explain how the central limit theorem allows us to find the probability that the mean number of lightning strikes per square kilometer is less than 5. Then calculate this probability. Show your work.

23. A study of rush-hour traffic in San Francisco counts the number of people in each car entering a freeway at a suburban interchange. Suppose that this count has mean 1.5 and standard deviation 0.75 in the population of all cars that enter at this interchange during rush hours.
    a. Could the exact distribution of the count be Normal? Why or why not?

    b. Traffic engineers estimate that the capacity of the interchange is 700 cars per hour. Find the probability that 700 cars will carry more than 1075 people. Show your work. (Hint: Restate this event in terms of the mean number of people $\bar{x}$ per car.)

24. The number of flaws per square yard in a type of carpet material varies with mean 1.6 flaws per square yard and standard deviation 1.2 flaws per square yard. The population distribution cannot be Normal, because a count takes only whole-number values. An inspector studies 200 square yards of the material, records the number of flaws found in each square yard, and calculates $\bar{x}$, the mean number of flaws per square yard inspected. Find the probability that the mean number of flaws exceeds 2 per square yard.

SAMPLING DISTRIBUTIONS

NB: You must describe the sampling distribution completely before calculating your probability.

	$\hat{p}$	$\overline{x}$
Chapter		
Shape:		
Center:		
Spread:		