Mod 20 Distribution of Sample Means

Upon successful completion of this activity and associated homework, students will be able to:

- Recognize when to use a hypothesis test or a confidence interval to draw a conclusion about a population mean
- Describe the sampling distribution of sample means
- Draw conclusions about a population mean from a simulation.
- Estimate the probability of an event using a normal model of the sampling distribution.

Mod 20 Warm-up problems: Mix and Match

Read the following examples and circle whether they are categorical or quantitative and then whether they are sample proportions or sample means.

1) What proportion of U.S. adults have watched *Survivor*?
   Categorical Quantitative

   Sample proportions Sample means

2) What is New York City's average household income?
   Categorical Quantitative

   Sample proportions Sample means

3) What is the lifetime of car brake pads?
   Categorical Quantitative

   Sample proportions Sample means

4) What percent of the adult population attended a sports event last week?
   Categorical Quantitative

   Sample proportions Sample means

5) What is the average blood pressure for women over 40?
   Categorical Quantitative

   Sample proportions Sample means

For this activity, students should be organized into groups of three.

Activity: *Penny for your thoughts*

1) Your instructor will display a dotplot of the population of pennies from a jar of various ages. Please describe the:
   Shape:

   Center:

   Spread:

2) While you are answering the above question the instructor will come around so that each student writes down the dates of 5 random pennies taken from the jar. Be sure to replace the pennies before the next student samples.

3) Calculate the mean age $\bar{x}$ of the 5 pennies in your sample.

4) A class dotplot of these means will be made and displayed please describe the
   Shape:

   Center:

   Spread:

5) What differences do you notice from the original dotplot?

6) Steps 2-5 will be repeated with a sample size of 10 pennies. Find the mean age $\bar{x}$ of the 10 pennies ages in your sample. A class dotplot of these means will be made and displayed. Please describe the:
   Shape:

   Center:

   Spread:

7) What differences do you notice from the original dotplot? From the dotplot with samples of size 5?

8) Steps 2-5 will be repeated with a sample size of 25 pennies. Find the mean age $\bar{x}$ of the 25 pennies ages in your sample. A class dotplot of these means will be made and displayed. Please describe the:
Shape:

Center:

Spread:

9) What differences do you notice form the original dotplot? From the dotplot with samples of size 5? With samples of size 10?

10) Let's explore what the penny activity suggests about choosing many simple random samples (SRSs) of size *n* drawn from a large population with mean μ and standard deviation σ.

a) The mean of the sampling distribution of $\bar{x}$ is

b) The standard deviation of the sampling distribution of $\bar{x}$ is

This standard deviation should only be used when the population is at least 10 times as large as the sample (the 10% condition).

11) **Example: *This Wine Stinks***

Sulfur compounds such as dimethyl sulfide (DMS) are sometimes present in wine. DMS causes "off-odors" in wine, so winemakers want to know the odor threshold, the lowest concentration of DMS that the human nose can detect. Extensive studies have found that the DMS odor threshold of adults follows roughly a Normal distribution with mean μ = 25 micrograms per liter and standard deviation σ = 7 micrograms per liter. Suppose we take an SRS of 10 adults and determine the mean odor threshold $\bar{x}$ for the individuals in the sample.

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

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

Central Limit theorem (CLT) (Write the Central Limit Theorem in the space below.)

12) **Example: *Lightning Strikes***

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 $\bar{x}$, the sample mean number of strikes per square kilometer?

b) Explain why you cannot safely calculate the probability that $\bar{x} < 5$ based on a sample size of size 10?

c) Suppose the NLDN takes a random sample of size $n = 50$ 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.