Name: ______________________________________ Date: ___________________

THE DISTRIBUTION OF SAMPLE MEANS
COMMON CORE ALGEBRA II

In order to be able to make **inferences** about population parameters based on sample statistics we first must understand how sample statistics, like the sample mean and sample proportion, are distributed. For example, if we take many samples from a population, how will the means of those samples distribute? In this lesson we will investigate this both with simulation and formally.

***Exercise* #1:** Using the Normal Distribution simulator, run a simulation for a population with a mean of $\mu = 50$ and a standard deviation of $\sigma = 15$ for a sample size of 30. Run 100 simulations.

(a) Does the distribution of sample means appear normal (i.e. like a normal distribution)? Explain.

(b) What is the mean of the sample means, symbolized by $\mu_{\bar{x}}$?

(c) What is the standard deviation of the sample means, symbolized by $\sigma_{\bar{x}}$, rounded to the nearest tenth? How does it compare with the standard deviation of the population?

(d) Based on this simulation alone, how likely would it be that a sample of this size taken from this population would have a mean greater than 2 standard deviations, $\sigma_{\bar{x}}$, above the mean?

THE CENTRAL LIMIT THEOREM

When a sample size is fairly large, say 30 or more, then the **distribution of all sample means** of a given size *n* will be **normally distributed** with:

1. **A mean**: $\mu_{\bar{x}} = \mu$ (the mean of the sample means will just be the mean of the population)

2. **A standard deviation**: $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$ (the variation of sample means is smaller than that of the population)

***Exercise* #2:** Do the results of your simulation agree with the Central Limit Theorem. Explain.

***Exercise* #3:** The mean height of adult American males is 177 cm with a standard deviation of 7.3 cm. What is the standard deviation of the distribution of samples means from this population with a sample size of 50?

(1) 0.15 (3) 3.54

(2) 1.03 (4) 4.72

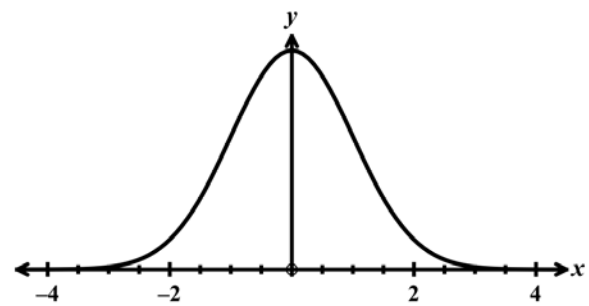
____

Because the distribution of sample means follows a normal distribution, we can determine how likely a sample mean would be given population parameters.

***Exercise* #4:** Jumbo eggs have a mean weight of 71 grams and a standard deviation of 3 grams. A sample of three dozen jumbo eggs was taken at a local egg processing plant and found to have a mean weight of 70 grams. Should this be of concern? Let's explore this question.

(a) What is the standard deviation of sample means of this size from the described population?

(b) What is the z-score for this particular sample mean? Illustrate this on the standard normal curve shown below.



(c) Using either tables or your calculator, determine the probability that a sample of this size would have a mean of 70 grams **or lower**. Round to the nearest tenth of a percent. Shade this area on the normal graph shown in (c).

(d) Why does it make sense in part (c) to determine the probability of having a sample with 70 grams or lower? What does the probability from part (c) tell you?

***Exercise* #5:** Given a population with a mean of 58 and a standard deviation of 12, which of the following represents the probability of getting a sample mean of 61 or greater with a sample size of 50? Show the analysis that leads to your choice.

(1) 7.2% (3) 18.0%

(2) 3.9% (4) 24.2%

____

***Exercise* #6:** In a normal distribution, approximately 95% of all data lie within two standard deviations of the mean. This includes normal distributions of sample means. If a population has a mean of 130, a standard deviation of 8 and samples of size 30 are taken, find the sample mean two standard deviations below the mean and two standard deviations above. Round both means to the nearest hundredth.

Name: ____________________________________ Date: __________________

THE DISTRIBUTION OF SAMPLE MEANS
COMMON CORE ALGEBRA II HOMEWORK

FLUENCY

1. For each of the population standard deviations given, calculate the standard deviation of the sample means, i.e. $\sigma_{\bar{x}}$, with a sample size of the specified value of n. Show your calculation. Round all answers to the nearest tenth.

(a) $\sigma = 12.5$ and $n = 40$

$\sigma_{\bar{x}} =$

(b) $\sigma = 22$ and $n = 65$

$\sigma_{\bar{x}} =$

(c) $\sigma = 2.7$ and $n = 127$

$\sigma_{\bar{x}} =$

(d) $\sigma = 35$ and $n = 237$

$\sigma_{\bar{x}} =$

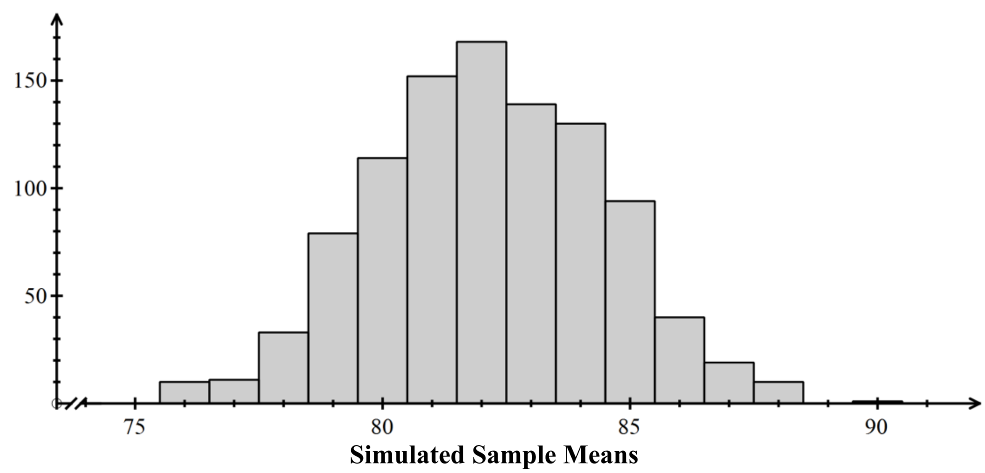
(e) $\sigma = 23.8$ and $n = 30$

$\sigma_{\bar{x}} =$

(f) $\sigma = 18.5$ and $n = 35$

$\sigma_{\bar{x}} =$

2. A simulation is done using a population with a mean of 82, a standard deviation of 15 and samples of size 40. When 1000 samples are simulated, the distribution of simulated sample means is created and shown. The mean of these sample means is 82.09 and the standard deviation is 2.33. Does this simulation support the conclusions of the Central Limit Theorem? Explain.



3. On a standardized test, the standard deviation of the scores was 14.8 points. If samples of size 60 were taken from this test's results, which of the following would be closest to the standard deviation of the means of these samples?

(1) 0.25
(2) 0.48
(3) 1.91
(4) 2.29

____

APPLICATIONS

4. In 2014, new cars had an average fuel efficiency of 27.9 miles per gallons with a standard deviation of 6.8 miles per gallon. A sample of 30 new cars is taken.

   (a) What is the probability the sample has a mean gas mileage between 27 and 29 miles per gallon?

   (b) What is the probability that the sample has a mean gas mileage greater than 30 miles per gallon?

   (c) If a sample of 30 trucks had a sample mean gas mileage of 24.3 miles per gallon, why is it reasonable to assume that all trucks have a lower overall gas mileage than cars? Explain

5. The average length of a cell phone call in 2012 was 1.80 minutes with a standard deviation of 0.32 minutes. A sample of 50 cell phone calls made by users less than 20 years old was taken and had a mean call length of 1.89 minutes.

   (a) What is the probability that a sample of 50 from a population with a mean of 1.80 and a standard deviation would have a mean call length of 1.89 minutes or longer.

   (b) What conclusion can you make about the average phone call length of users younger than 20 compared to the general population? Explain.

REASONING

6. A population has a standard deviation of 37. If a researcher is designing a study so that the distribution of sample means has a standard deviation of less than 5, what is the smallest sample size that can be used?