7.10: Chapter Homework

7.2 The Central Limit Theorem for Sample Means

49 Previously, De Anza statistics students estimated that the amount of change daytime statistics students carry is exponentially distributed with a mean of $0.88. Suppose that we randomly pick 25 daytime statistics students.

a. In words, X = ____________
b. X $\sim$ _____(_____,_____)
c. In words, $\overline{X}$ = ____________
d. $\overline{X}$ $\sim$ ______ (______, ______)
e. Find the probability that an individual had between $0.80 and $1.00. Graph the situation, and shade in the area to be determined.
f. Find the probability that the average of the 25 students was between $0.80 and $1.00. Graph the situation, and shade in the area to be determined.
g. Explain why there is a difference in part e and part f.

Answer

a. X = amount of change students carry
b. X $\sim E(0.88, 0.88)$
c. $\overline{X}$ = average amount of change carried by a sample of 25 students.
d. $\overline{X} \sim N(0.88, 0.176)$
e. 0.0819
f. 0.1882
g. The distributions are different. Part 1 is exponential and part 2 is normal.

50. Suppose that the distance of fly balls hit to the outfield (in baseball) is normally distributed with a mean of 250 feet and a standard deviation of 50 feet. We randomly sample 49 fly balls.

a. If $\overline{X}$ = average distance in feet for 49 fly balls, then $\overline{X} \sim$ _______(_______,_______)
b. What is the probability that the 49 balls traveled an average of less than 240 feet? Sketch the graph. Scale the horizontal axis for $\overline{X}$. Shade the region corresponding to the probability. Find the probability.
c. Find the 80th percentile of the distribution of the average of 49 fly balls.

51. According to the Internal Revenue Service, the average length of time for an individual to complete (keep records for, learn, prepare, copy, assemble, and send) IRS Form 1040 is 10.53 hours (without any attached schedules). The distribution is unknown. Let us assume that the standard deviation is two hours. Suppose we randomly sample 36 taxpayers.

a. In words, X $=$ _____________
b. In words, $\overline{X}$ = _____________
c. $\overline{X}$ $\sim$ _____(_____,_____)
d. Would you be surprised if the 36 taxpayers finished their Form 1040s in an average of more than 12 hours? Explain why or why not in complete sentences.
e. Would you be surprised if one taxpayer finished his or her Form 1040 in more than 12 hours? In a complete sentence, explain why.

52. Suppose that a category of world-class runners are known to run a marathon (26 miles) in an average of 145 minutes with a standard deviation of 14 minutes. Consider 49 of the races. Let $\overline{X}$ the average of the 49 races.

a. $\overline{X}$ $\sim$ _____(_____,_____)
b. Find the probability that the runner will average between 142 and 146 minutes in these 49 marathons.
c. Find the 80^{th} percentile for the average of these 49 marathons.
d. Find the median of the average running times.

53. The length of songs in a collector's iTunes album collection is uniformly distributed from two to 3.5 minutes. Suppose we randomly pick five albums from the collection. There are a total of 43 songs on the five albums.

a. In words, X = _________
b. X ∼ _____________
c. In words, $\overline{X}$ = _____________
d. $\overline{X}$ ∼ _____(_____,_____)
e. Find the first quartile for the average song length.
f. The IQR (interquartile range) for the average song length is from _______–_______.

54. In 1940 the average size of a U.S. farm was 174 acres. Let's say that the standard deviation was 55 acres. Suppose we randomly survey 38 farmers from 1940.

a. In words, X = _____________
b. In words, $\overline{X}$ = _____________
c. $\overline{X}$ ∼ _____(_____,_____)
d. The IQR for $\overline{X}$ is from _______ acres to _______ acres.

55. Determine which of the following are true and which are false. Then, in complete sentences, justify your answers.

a. When the sample size is large, the mean of $\overline{X}$ is approximately equal to the mean of X.
b. When the sample size is large, $\overline{X}$ is approximately normally distributed.
c. When the sample size is large, the standard deviation of $\overline{X}$ is approximately the same as the standard deviation of X.

56. The percent of fat calories that a person in America consumes each day is normally distributed with a mean of about 36 and a standard deviation of about ten. Suppose that 16 individuals are randomly chosen. Let $\overline{X}$ = average percent of fat calories.

a. $\overline{X}$ ∼ ______(______, ______)
b. For the group of 16, find the probability that the average percent of fat calories consumed is more than five. Graph the situation and shade in the area to be determined.
c. Find the first quartile for the average percent of fat calories.

57. The distribution of income in some Third World countries is considered wedge shaped (many very poor people, very few middle income people, and even fewer wealthy people). Suppose we pick a country with a wedge shaped distribution. Let the average salary be $2,000 per year with a standard deviation of $8,000. We randomly survey 1,000 residents of that country.

a. In words, X = _____________
b. In words, $\overline{X}$ = _____________
c. $\overline{X}$ ∼ _____(_____,_____)
d. How is it possible for the standard deviation to be greater than the average?
e. Why is it more likely that the average of the 1,000 residents will be from $2,000 to $2,100 than from $2,100 to $2,200?

58. Which of the following is NOT TRUE about the distribution for averages?

a. The mean, median, and mode are equal.
b. The area under the curve is one.
c. The curve never touches the x-axis.
d. The curve is skewed to the right.

59. The cost of unleaded gasoline in the Bay Area once followed an unknown distribution with a mean of $4.59 and a standard deviation of $0.10. Sixteen gas stations from the Bay Area are randomly chosen. We are interested in the average cost of gasoline for the 16 gas stations. The distribution to use for the average cost of gasoline for the 16 gas stations is:

a. $\overline{X} \sim N(4.59, 0.10)$

b. $\overline{X} \sim N\left(4.59, \frac{0.10}{\sqrt{16}}\right)$

c. $\overline{X} \sim N\left(4.59, \frac{16}{0.10}\right)$

d. $\overline{X} \sim N\left(4.59, \frac{\sqrt{16}}{0.10}\right)$

7.3 Using the Central Limit Theorem

60. A large population of 5,000 students take a practice test to prepare for a standardized test. The population mean is 140 questions correct, and the standard deviation is 80. What size samples should a researcher take to get a distribution of means of the samples with a standard deviation of 10?

61. A large population has skewed data with a mean of 70 and a standard deviation of 6. Samples of size 100 are taken, and the distribution of the means of these samples is analyzed.

a. Will the distribution of the means be closer to a normal distribution than the distribution of the population?
b. Will the mean of the means of the samples remain close to 70?
c. Will the distribution of the means have a smaller standard deviation?
d. What is that standard deviation?

62. A researcher is looking at data from a large population with a standard deviation that is much too large. In order to concentrate the information, the researcher decides to repeatedly sample the data and use the distribution of the means of the samples? The first effort used sample sized of 100. But the standard deviation was about double the value the researcher wanted. What is the smallest size samples the researcher can use to remedy the problem?

63. A researcher looks at a large set of data, and concludes the population has a standard deviation of 40. Using sample sizes of 64, the researcher is able to focus the mean of the means of the sample to a narrower distribution where the standard deviation is 5. Then, the researcher realizes there was an error in the original calculations, and the initial standard deviation is really 20. Since the standard deviation of the means of the samples was obtained using the original standard deviation, this value is also impacted by the discovery of the error. What is the correct value of the standard deviation of the means of the samples?

64. A population has a standard deviation of 50. It is sampled with samples of size 100. What is the variance of the means of the samples?

7.4 The Central Limit Theorem for Proportions

65. A farmer picks pumpkins from a large field. The farmer makes samples of 260 pumpkins and inspects them. If one in fifty pumpkins are not fit to market and will be saved for seeds, what is the standard deviation of the mean of the sampling distribution of sample proportions?

66. A store surveys customers to see if they are satisfied with the service they received. Samples of 25 surveys are taken. One in five people are unsatisfied. What is the variance of the mean of the sampling distribution of sample proportions for the number of unsatisfied customers? What is the variance for satisfied customers?

67. A company gives an anonymous survey to its employees to see what percent of its employees are happy. The company is too large to check each response, so samples of 50 are taken, and the tendency is that three-fourths of the employees are happy. For the mean of the sampling distribution of sample proportions, answer the following questions, if the sample size is doubled.

a. How does this affect the mean?
b. How does this affect the standard deviation?
c. How does this affect the variance?

68. A pollster asks a single question with only yes and no as answer possibilities. The poll is conducted nationwide, so samples of 100 responses are taken. There are four yes answers for each no answer overall. For the mean of the sampling distribution of sample proportions, find the following for yes answers.

a. The expected value.
b. The standard deviation.
c. The variance.

69. The mean of the sampling distribution of sample proportions has a value of p of 0.3, and sample size of 40.

a. Is there a difference in the expected value if p and q reverse roles?
b. Is there a difference in the calculation of the standard deviation with the same reversal?

Finite Population Correction Factor

70. A company has 1,000 employees. The average number of workdays between absence for illness is 80 with a standard deviation of 11 days. Samples of 80 employees are examined. What is the probability a sample has a mean of workdays with no absence for illness of at least 78 days and at most 84 days?

71. Trucks pass an automatic scale that monitors 2,000 trucks. This population of trucks has an average weight of 20 tons with a standard deviation of 2 tons. If a sample of 50 trucks is taken, what is the probability the sample will have an average weight within one-half ton of the population mean?

72. A town keeps weather records. From these records it has been determined that it rains on an average of 12% of the days each year. If 30 days are selected at random from one year, what is the probability that at most 3 days had rain?

73. A maker of greeting cards has an ink problem that causes the ink to smear on 7% of the cards. The daily production run is 500 cards. What is the probability that if a sample of 35 cards is checked, there will be ink smeared on at most 5 cards?

74. A school has 500 students. Usually, there are an average of 20 students who are absent. If a sample of 30 students is taken on a certain day, what is the probability that at least 2 students in the sample will be absent?
