AP® STATISTICS
2007 SCORING GUIDELINES

Question 3

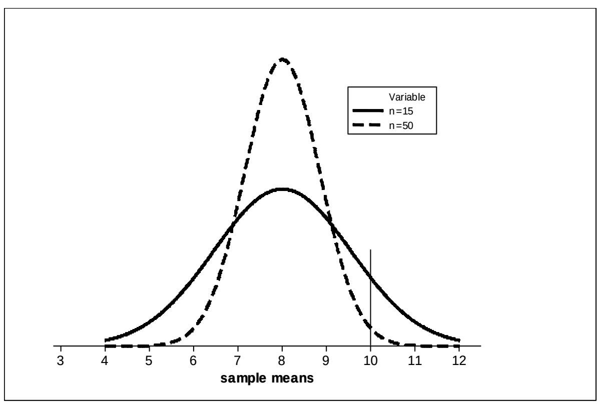
Intent of Question

This question was developed to assess a student's understanding of the sampling distribution of the sample mean: in particular, a student's ability to: (1) compare probabilities concerning sample means from different sample sizes; (2) compute an appropriate probability; and (3) recognize that an application of the Central Limit Theorem is being evaluated.

Solution

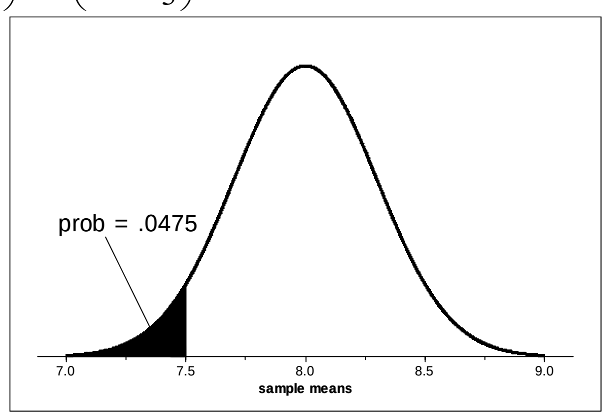
Part (a):

The random sample of $n = 15$ fish is more likely to have a sample mean length greater than 10 inches. The sampling distribution of the sample mean $\bar{x}$ is normal with mean $\mu = 8$ and standard deviation $\sigma/\sqrt{n}$. Thus, both sampling distributions will be centered at 8 inches, but the sampling distribution of the sample mean when $n = 15$ will have more variability than the sampling distribution of the sample mean when $n = 50$. The tail area ($\bar{x} > 10$) will be larger for the distribution that is less concentrated about the mean of 8 inches when the sample size is $n = 15$, as shown in the following graph.



Part (b):

$$P(\bar{x} < 7.5) = P\left(z < \frac{7.5 - 8}{0.3}\right) = P\left(z < -\frac{5}{3}\right) = P(z < -1.67) = 0.0475$$



Question 3 (continued)

Part (c):

Yes. The Central Limit Theorem says that the sampling distribution of the sample mean will become approximately normal as the sample size *n* increases. Since the sample size is reasonably large ($n = 50$), the calculation in part (b) will provide a good approximation to the probability of interest even though the population is nonnormal.

Scoring

Parts (a), (b), and (c) are scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is scored as essentially correct (E) if the student says that the sample of 15 fish is more likely to have a mean length that is greater than 10, AND the justification is based on *variability* in the *sampling* distributions.

Part (a) is scored as partially correct (P) if:
the student makes correct statements about the sampling distribution of the sample mean or the probabilities but does not specifically refer to the variability in these two sampling distributions;
OR
the student remarks that the sample mean approaches the population mean as the sample size increases (an argument based on the Law of Large Numbers).

Some examples of partially correct statements are:

- With the smaller sample size we will be more likely to get an extreme value for the sample mean.
- Variability in the smaller sample is larger.
- Variability in the larger sample is smaller.
- The sample mean approaches the population mean as the sample size increases.

Part (a) is scored as incorrect (I) if an answer is provided with no justification or incorrect justification.

Note: If a student chooses a particular value for a standard deviation and goes through the correct calculations, or shows the result algebraically based on a generic standard deviation, then the response should be scored essentially correct.

Part (b) is scored essentially correct (E) if the probability is calculated correctly and a reasonable sketch or evidence of calculation is shown.

Part (b) is scored partially correct (P) if:
an incorrect but plausible calculation is shown. Examples include using an incorrect standard deviation (such as $0.3/\sqrt{50}$) to obtain the probability;
OR
the student switches the sample mean and the population mean to obtain a standardized *z* value of 1.67.

Question 3 (continued)

Part (b) is scored incorrect (I) if an answer is provided with no justification or incorrect justification.

Note: Normalcdf (…) with no additional work is at best partially correct. If an appropriate sketch accompanies the calculator command, *OR* if the components of the calculator command are clearly identified/labeled, then the solution should be scored essentially correct.

Part (c) is scored as essentially correct (E) if the student says that the probability is a reasonable approximation because of the Central Limit Theorem and also refers to the large sample size in this case.

Part (c) is scored partially correct (P) if the student indicates that the response in part (b) would not change but provides a weak justification. Examples of a weak justification include mentioning CLT without reference to sample size, and mentioning sample size without reference to CLT.

Part (c) is scored incorrect (I) if an answer is provided with no justification or incorrect justification.

Note: An E counts for 2 points in part (a), and an E counts for 1 point in each of parts (b) and (c). Similarly, a P counts for 1 point in part (a), and a P counts for ½ point in parts (b) and (c). When the total number of points earned is not an integer, the final score earned will be rounded down to the integer value.

4 Complete Response

4 points earned

3 Substantial Response

3 or 3½ points earned

2 Developing Response

2 or 2½ points earned

1 Minimal Response

1 or 1½ points earned

3. Big Town Fisheries recently stocked a new lake in a city park with 2,000 fish of various sizes. The distribution of the lengths of these fish is approximately normal.

(a) Big Town Fisheries claims that the mean length of the fish is 8 inches. If the claim is true, which of the following would be more likely?

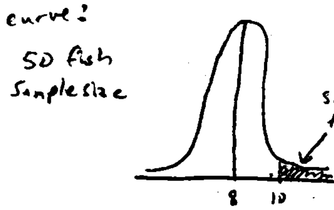
- A random sample of 15 fish having a mean length that is greater than 10 inches

or

- A random sample of 50 fish having a mean length that is greater than 10 inches

Justify your answer.

The random sample of 15 fish with a mean that is greater than 10 is more likely. This is true because as the sample grows larger, it reflects the population better. The stadard deviation of a distribution of sample means is $\frac{\sigma}{\sqrt{n}}$. Thus, as n increases, the stadard deviation decreases, and a smaller standard deviation means less variance. Smaller standard deviations also thin out the normal curve:



Shaded area for sample size of 50 is much smaller than that of 15

(b) Suppose the standard deviation of the sampling distribution of the sample mean for random samples of size 50 is 0.3 inch. If the mean length of the fish is 8 inches, use the normal distribution to compute the probability that a random sample of 50 fish will have a mean length less than 7.5 inches.

$$z = \frac{7.5 - 8}{.3} = -1.6\bar{6}$$

$$\boxed{p = .0485}$$

(c) Suppose the distribution of fish lengths in this lake was nonnormal but had the same mean and standard deviation. Would it still be appropriate to use the normal distribution to compute the probability in part (b) ? Justify your answer.

Yes. The central limit theorem states that the distribution of sample means (if the sample size is large enough) will be normally distributed. Therefore, since we are using a sample mean and the sample size is large enough (50), a normal distribution can be used.

GO ON TO THE NEXT PAGE.

3. Big Town Fisheries recently stocked a new lake in a city park with 2,000 fish of various sizes. The distribution of the lengths of these fish is approximately normal.

(a) Big Town Fisheries claims that the mean length of the fish is 8 inches. If the claim is true, which of the following would be more likely?

- A random sample of 15 fish having a mean length that is greater than 10 inches

or

- A random sample of 50 fish having a mean length that is greater than 10 inches

Justify your answer.

If the claim is true, the random sample of 15 fish with a mean length of 10 inches would be more likely. The smaller the sample size, the greater chance of significant variability between samples. In addition, the Central Limit Theorem states that the sample size should be greater than 30. In this case, with n=15, there is a chance the entire population is not adequately represented.

(b) Suppose the standard deviation of the sampling distribution of the sample mean for random samples of size 50 is 0.3 inch. If the mean length of the fish is 8 inches, use the normal distribution to compute the probability that a random sample of 50 fish will have a mean length less than 7.5 inches.

$$z = \frac{7.5-8}{.3} = {}^{-}1.67 = .0475$$

P(random sample of 50 fish will have a mean length less than 7.5 in.) = .0475

(c) Suppose the distribution of fish lengths in this lake was nonnormal but had the same mean and standard deviation. Would it still be appropriate to use the normal distribution to compute the probability in part (b)? Justify your answer.

Yes, however the sample would have to be large enough (n>30) and the sample would need to be an SRS. According to the CLT, if you have these two conditions it doesn't matter if the population is nonnormal.

GO ON TO THE NEXT PAGE.

3. Big Town Fisheries recently stocked a new lake in a city park with 2,000 fish of various sizes. The distribution of the lengths of these fish is approximately normal.

(a) Big Town Fisheries claims that the mean length of the fish is 8 inches. If the claim is true, which of the following would be more likely?

- A random sample of 15 fish having a mean length that is greater than 10 inches

or

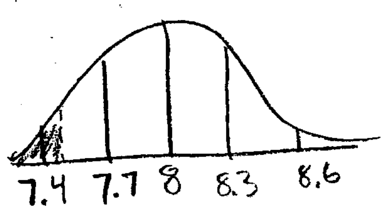
- A random sample of 50 fish having a mean length that is greater than 10 inches

Justify your answer.

A random sample of 15 fish having a mean of 10 inches, because the law of large numbers states: as a sample size increases the sample mean $\bar{x}$ grows closer to the true mean μ.

(b) Suppose the standard deviation of the sampling distribution ~~of the sample mean~~ for random ~~sample~~s of size 50 is 0.3 inch. If the mean length of the fish is 8 inches, use the normal distribution to compute the probability that a random sample of 50 fish will have a mean length less than 7.5 inches.

$\sigma_{\bar{x}} = .3$



$\frac{7.5-8}{.3} = -1.6667$

$P(Z < -1.6667) = .0475$

(c) Suppose the distribution of fish lengths in this lake was nonnormal but had the same mean and standard deviation. Would it still be appropriate to use the normal distribution to compute the probability in part (b) ? Justify your answer.

No, because standardized values, and standard normal probabilities are only applicable if a sample is normally distributed

GO ON TO THE NEXT PAGE.

Question 3

Overview

This question was developed to assess a student's understanding of the sampling distribution of the sample mean; in particular, a student's ability to: (1) compare probabilities concerning sample means from different sample sizes; (2) compute an appropriate probability based on the normal distribution; and (3) recognize that for large size samples the Central Limit Theorem assures that the sampling distribution of the sample mean still follows an approximately normal distribution even when the population distribution is nonnormal.

Sample: 3A
Score: 4

The response to part (a) is correct and very well justified. The verbal description, along with the $\sigma/\sqrt{n}$ expression for the standard deviation of $\bar{x}$, is very clear. The accompanying sketches are particularly illuminating, and the shaded area is clearly labeled as the probability that the sample mean is greater than 10 inches, $\bar{x} > 10$. The probability calculation in part (b) is correct, supported by a *z*-score calculation. This is sufficient for an essentially correct response, but a stronger response would have included more information such as a sketch or P ($\bar{x} < 7.5$) notation. As with part (a) the response to part (c) is correct and very well justified. The response not only cites the Central Limit Theorem (CLT) as the justification for answering "yes" but goes on to describe what the CLT says about the distribution of sample means. A minor quibble with part (c) is that an ideal response would note that the sampling distribution of $\bar{x}$ is *approximately* normal when the sample size is large enough.

Sample: 3B
Score: 3

The correct sample size (n = 15) is selected in part (a). The justification for this choice is based on "greater chance of significant variability between samples" with a smaller sample size. This response has the right idea about decreased variability with larger samples, and the phrase "between samples" reveals a partial understanding of sampling variability, but this does not clearly indicate that variability between sample *means* is the important consideration. The last two sentences of the part (a) response are not relevant, because with a normally distributed population of fish lengths, the distribution of sample means is normal even with small sample sizes. The part (b) response provides a correct probability calculation, with support given by the intermediate calculation of a *z*-score. The notation is poor in equating the *z*-score of -1.67 with the probability of 0.0475, but the probability statement about a "random sample of 50 fish will have a mean length less than 7.5 inches" is very clear and precise. The response provides a correct answer to part (c) and justifies it by appealing to the Central Limit Theorem and noting the need for a large sample of more than 30 fish. The part (c) response would be strengthened by more definitively declaring that the sample size is indeed greater than 30 in this problem, but it has all of the components to be considered essentially correct.

Question 3 (continued)

Sample: 3C
Score: 2

The response chooses the correct sample size in part (a). The justification is based on the Law of Large Numbers, however, which asserts that the sample mean $\overline{x}$ tends to approach the population mean μ as the sample size increases. This justification is partially correct, but it does not refer to variability among sample means under repeated random sampling with a fixed sample size. The response to part (b) arrives at the correct probability value, including a *z*-score calculation and shaded sketch to support the answer. The "$\sigma_x = 0.3$" notation on the left side of the response is unfortunate, because 0.3 is the standard deviation of the sampling distribution of sample means, not of the distribution of fish lengths; a better notation would be "$\sigma_{\overline{x}} = 0.3$". The response gives an incorrect answer for part (c), implying that the probability statement is about the distribution of fish lengths and not about the distribution of sample means.