AP® STATISTICS
2015 SCORING GUIDELINES

Question 2

Intent of Question

The primary goals of this question were to assess a student's ability to (1) use confidence intervals to test a question about a proportion and (2) understand the relationship between sample size and margin of error in a confidence interval for a proportion.

Solution

Part (a):

(i) No. The confidence interval is $(0.09, 0.21)$, which includes the value of 0.20. Therefore, it is plausible that the computer program is generating discounts with a probability of 0.20, and the confidence interval does not provide convincing statistical evidence that the program is not working as intended.

(ii) No. The confidence interval includes values from 0.09 to 0.21, so any value in that interval is a plausible value for the probability that the computer is using to generate discounts.

Part (b):

The formula for computing the margin of error for a proportion includes the square root of the sample size in the denominator. For a random sample that is four times the size of the original sample, the margin of error can be determined by dividing the margin of error of the original sample by two. Therefore, the new margin of error is 0.03.

Part (c):

Using the margin of error of 0.03 obtained from the second sample, the confidence interval for p is 0.15 ± 0.03 or $(0.12, 0.18)$. The interval does not include 0.20, and therefore, there is convincing evidence that the computer program is not working as intended and is not generating discounts with a probability of 0.20.

Scoring

This question is scored in four sections. Section 1 consists of part (a-i); section 2 consists of part (a-ii); section 3 consists of part (b); and section 4 consists of part (c). Sections 1, 2, 3, and 4 are scored as essentially correct (E), partially correct (P), or incorrect (I).

Section 1 is scored as follows:

Essentially correct (E) if the response states that because the interval contains 0.20, it does not provide convincing statistical evidence that the computer program is not working as intended.

Partially correct (P) if the response indicates that it is necessary to check whether the value of 0.20 is in the computed interval, but there are errors in implementation. Examples of errors include:

- The response notes that 0.20 is within the interval but does not draw a conclusion.
- The response has an arithmetic error in the computation of the endpoints of the interval but provides a correct conclusion with justification that is consistent with the computed interval.

Question 2 (continued)

Incorrect (I) if the response does not recognize how to use the confidence interval to check whether the computer is working correctly;
OR
if the response states that the interval shows that the proportion is equal to 0.20;
OR
if the response notes that 0.20 is within the interval and concludes that the program is not working as intended;
OR
if the response otherwise does not meet the criteria for E or P.

Section 2 is scored as follows:

Essentially correct (E) if the response concludes that there is not convincing statistical evidence that the computer program generates the discount with a probability of 0.20 *AND* justifies the conclusion by noting that there are values other than 0.20 in the interval.

Partially correct (P) if the response correctly concludes that there is not convincing evidence that the computer program generates the discount with a probability of 0.20, but provides incomplete reasoning to justify the conclusion.

Examples of incomplete reasoning include:

- stating that 0.20 is a plausible value for the proportion of discounts without giving further explanation;
- indicating that having 0.20 in the interval does not prove that 0.20 is the true proportion; and
- providing a generic statistical argument, such as reference to the fact that the null hypothesis should never be accepted, or stating that not rejecting the null hypothesis is not proof that 0.20 is the true proportion of bills discounted.

Note: If an incorrect interval is computed in part (a-i) that does not contain 0.20, and in part (a-ii) the response concludes that the program is not generating discounts with a probability of 0.20 because 0.20 is not in the interval, section 2 is scored as P.

Incorrect (I) if the response states that there is evidence that the computer program generates the discount with a probability of 0.20;
OR
if the response correctly concludes that there is not convincing evidence that the computer program generates the discount with a probability of 0.20 *AND* provides incorrect or no justification;
OR
if the response otherwise does not meet the criteria for E or P.

Section 3 is scored as follows:

Essentially correct (E) if the response gives the correct value of 0.03 as the new margin of error by using the correct formula or by providing a correct explanation that recognizes that quadrupling the sample size divides the margin of error from the original sample by two.

Note: If the response relies on the margin of error formula and calculates a value that would round to 0.030, the response is scored as E.

Question 2 (continued)

Partially correct (P) if the response uses the correct margin of error formula or provides a correct explanation that recognizes that quadrupling the sample size divides the margin of error from the original sample by two, but calculates an incorrect margin of error that is less than 0.06;
OR
if the response does not calculate a margin of error but provides a correct explanation that recognizes that quadrupling the sample size divides the margin of error from the original sample by two;
OR
if the response gives the correct margin of error without correct justification.

Incorrect (I) if the response does not recognize in any way that the new margin of error depends on the square root of the sample size;
OR
if the response calculates a margin of error that is greater than or equal to 0.06;
OR
if the response otherwise does not meet the criteria for E or P.

Section 4 is scored as follows:

Essentially correct (E) if the conclusion states that there is convincing evidence that the computer program is not working as intended because, based on the new margin of error, the interval does not contain 0.20;
OR
if the conclusion states the intended value of 0.20 is greater than the upper boundary of 0.18;
OR
if the conclusion states the margin of error is smaller than the difference between the sample proportion and the intended long-run proportion of 0.20.

Notes:
- If the margin of error was computed incorrectly in part (b), but a correct answer to part (c) is consistent with this incorrect margin of error, section 4 is scored as E.
- If no specific margin of error or confidence interval was given in parts (b) or (c), but the response provides a complete and correct conclusion with reference to a smaller margin of error (or narrower confidence interval), section 4 is scored as E.

Partially correct (P) if an interval is incorrectly constructed using the margin of error from part (b), but a correct conclusion is given and justified for the computed interval;
OR
if an interval is computed correctly using the margin of error from part (b) and an argument is made based on whether or not 0.20 is in the interval, but the conclusion is incorrect;
OR
if no specific margin of error or confidence interval has been given in parts (b) or (c), but the response concludes that because the margin of error has decreased (or the confidence interval is narrower), 0.20 is not in the interval.

Incorrect (I) if an interval is computed correctly using the margin of error from part (b) and a correct conclusion is given, but no argument is made based on whether or not 0.20 is in the interval;
OR
if the response otherwise does not meet the criteria for E or P.

Question 2 (continued)

Notes:

- If a response includes a confidence level, the level can be ignored because no confidence level was provided.
- If the response bases a conclusion on the relative location of 0.20 within the interval (for example, 0.20 is near the edge of the interval), the response is scored as I.
- A response that provides additional incorrect explanation lowers the score in section 4 by one level (that is, from E to P, or from P to I).

Each essentially correct (E) section counts as 1 point. Each partially correct (P) section counts as ½ point.

4 Complete Response

3 Substantial Response

2 Developing Response

1 Minimal Response

If a response is between two scores (for example, 2½ points), use a holistic approach to decide whether to score up or down, depending on the overall strength of the response and communication.

2. To increase business, the owner of a restaurant is running a promotion in which a customer's bill can be randomly selected to receive a discount. When a customer's bill is printed, a program in the cash register randomly determines whether the customer will receive a discount on the bill. The program was written to generate a discount with a probability of 0.2, that is, giving 20 percent of the bills a discount in the long run. However, the owner is concerned that the program has a mistake that results in the program not generating the intended long-run proportion of 0.2.

   The owner selected a random sample of bills and found that only 15 percent of them received discounts. A confidence interval for p, the proportion of bills that will receive a discount in the long run, is 0.15 ± 0.06. All conditions for inference were met.

   (a) Consider the confidence interval 0.15 ± 0.06.

      (i) Does the confidence interval provide convincing statistical evidence that the program is not working as intended? Justify your answer.

No. The interval, [.09, .21] contains the intended probability, p = .2. Therefore, the interval doesn't provide statistically significant evidence ~~that there is the program~~ that the program is not working as intended

      (ii) Does the confidence interval provide convincing statistical evidence that the program generates the discount with a probability of 0.2 ? Justify your answer.

No. It does not yield any projection of the actual probability, only a range of acceptable probabilities. Because .2 is in the interval it is an acceptable value, but this does not provide evidence that it is the actual value. ~~[illegible]~~ A confidence interval does not define the actual proportion of a population, only a range of plausible proportions.

GO ON TO THE NEXT PAGE.

A second random sample of bills was taken that was four times the size of the original sample. In the second sample 15 percent of the bills received the discount.

(b) Determine the value of the margin of error based on the second sample of bills that would be used to compute an interval for p with the same confidence level as that of the original interval.

$ME = z^* (SE)$

$SE_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{.2(.8)}{4n_1}} = \frac{1}{2} SE_{\hat{p}_{n_1}}$

z^* constant

$SE_{\hat{p}_{n_2}} = \frac{1}{2} SE_{\hat{p}_{n_1}}$

$ME_2 = \frac{1}{2} ME_1$

the margin of error is $\pm .03$

(c) Based on the margin of error in part (b) that was obtained from the second sample, what do you conclude about whether the program is working as intended? Justify your answer.

Confidence interval: $[.12, .18]$ .2 is not in the confidence interval

We are confident, to the degree of the confidence interval in part (a), that the program is not working as intended because .2 does not fall between .12 and .18. It is implausible that a random sample of this size from a population with $p = .2$ would yield this interval.

2. To increase business, the owner of a restaurant is running a promotion in which a customer's bill can be randomly selected to receive a discount. When a customer's bill is printed, a program in the cash register randomly determines whether the customer will receive a discount on the bill. The program was written to generate a discount with a probability of 0.2, that is, giving 20 percent of the bills a discount in the long run. However, the owner is concerned that the program has a mistake that results in the program not generating the intended long-run proportion of 0.2.

   The owner selected a random sample of bills and found that only 15 percent of them received discounts. A confidence interval for p, the proportion of bills that will receive a discount in the long run, is 0.15 ± 0.06. All conditions for inference were met.

   (a) Consider the confidence interval 0.15 ± 0.06.

      (i) Does the confidence interval provide convincing statistical evidence that the program is not working as intended? Justify your answer.

Because .2 is included in the confidence interval, there is not enough convincing evidence that the program is not working as intended.

      (ii) Does the confidence interval provide convincing statistical evidence that the program generates the discount with a probability of 0.2 ? Justify your answer.

Because .2 is on the upper edge of the confidence interval, there isn't convincing statistical evidence that the program generates the discount with probability of .2

A second random sample of bills was taken that was four times the size of the original sample. In the second sample 15 percent of the bills received the discount.

(b) Determine the value of the margin of error based on the second sample of bills that would be used to compute an interval for p with the same confidence level as that of the original interval.

4x the original, so divide the original by 2
original margin of error is .06, so the new margin of error is .03

(c) Based on the margin of error in part (b) that was obtained from the second sample, what do you conclude about whether the program is working as intended? Justify your answer.

The new interval would be .15±.03, and this time .2 isn't included in the interval, so we'd conclude that the owner's concern is correct, the program doesn't appear to be running as it should be running.

2C1 of 2

2. To increase business, the owner of a restaurant is running a promotion in which a customer's bill can be randomly selected to receive a discount. When a customer's bill is printed, a program in the cash register randomly determines whether the customer will receive a discount on the bill. The program was written to generate a discount with a probability of 0.2, that is, giving 20 percent of the bills a discount in the long run. However, the owner is concerned that the program has a mistake that results in the program not generating the intended long-run proportion of 0.2.

The owner selected a random sample of bills and found that only 15 percent of them received discounts. A confidence interval for p, the proportion of bills that will receive a discount in the long run, is 0.15 ± 0.06. All conditions for inference were met.

(a) Consider the confidence interval 0.15 ± 0.06.

(i) Does the confidence interval provide convincing statistical evidence that the program is not working as intended? Justify your answer.

No, the confidence interval is (0.09, 0.21). 0.20 is contained within the interval

(ii) Does the confidence interval provide convincing statistical evidence that the program generates the discount with a probability of 0.2 ? Justify your answer.

No, 0.15 is the center of the interval which means 0.15 is the most probable outcomes. The confidence interval provides convincing statistical evidence that 15% receive discounts, not 20%.

A second random sample of bills was taken that was four times the size of the original sample. In the second sample 15 percent of the bills received the discount.

(b) Determine the value of the margin of error based on the second sample of bills that would be used to compute an interval for p with the same confidence level as that of the original interval.

$0.06(100)\left(\frac{1}{400}\right) = 0.015$ *Assuming 1st sample size = 100

$\therefore$ margin of error $\boxed{\pm 0.015}$

Margin of error decreases with an increase in n value.

(c) Based on the margin of error in part (b) that was obtained from the second sample, what do you conclude about whether the program is working as intended? Justify your answer.

Would conclude that the program is not working as intended. Interval would be $.15 \pm 0.015$ $\therefore$ $(0.135, 0.165)$. 0.20 is not included in the interval & $\therefore$ is highly unlikely that the 20% are getting discounts.

AP® STATISTICS
2015 SCORING COMMENTARY

Question 2

Overview

The primary goals of this question were to assess a student's ability to (1) use confidence intervals to test a question about a proportion and (2) understand the relationship between sample size and margin of error in a confidence interval for a proportion.

Sample: 2A
Score: 4

In part (a-i) the student correctly concludes that the confidence interval does not provide statistically significant evidence that the program is not working as intended because the confidence interval contains 0.20. Section 1 was scored as essentially correct. In part (a-ii) the student provides a thorough explanation of why having 0.20 in the interval does not provide evidence that the program generates the discount with a probability of 0.20. The student explains that although 0.20 is a plausible value, there are other plausible values and having 0.20 in the interval is not evidence to prove that the true proportion is likely to be 0.20. The student says that "a confidence interval does not define the actual proportion of a population, only a range of plausible proportions." This is a thorough and correct explanation, and section 2 was scored as essentially correct. The student calculates a correct new margin of error in part (b) and uses the formula for margin of error to show how quadrupling the sample size results in the margin of error being divided by 2. The student substitutes 0.2 for p and 0.8 for $1 - p$, which is incorrect, as it should be 0.15 and 0.85, but this mistake does not result in the lowering of the score. Section 3 was scored as essentially correct. In part (c) the student calculates a new confidence interval based on the new margin of error calculated in part (b) and correctly concludes that because 0.20 is not contained in the confidence interval, there is evidence that the program is not working as intended. The student continues to explain the rationale for this decision by stating that, "It is implausible that a random sample of this size from a population with $p = .2$ would yield this interval." This is a complete statistical argument of the reasoning for the decision and section 4 was scored as essentially correct. Because all four sections were scored as essentially correct, the response earned a score of 4.

Sample: 2B
Score: 3

In part (a-i) the student correctly explains that because 0.20 lies in the interval, there is no convincing evidence that the program is not working as intended. Section 1 was scored as essentially correct. However, in part (a-ii) although the student correctly concludes that there is no convincing evidence that the program correctly generates a discount with a probability of 0.20, the justification is incorrect. The student is basing a conclusion on the location of 0.20 within the confidence interval and indicates that because 0.20 is near the upper boundary of the interval that it leads to no evidence. A particular value within a confidence interval is not more or less likely to be the correct estimate of a parameter based on its location within the interval. The response also neglects to note that the confidence interval gives other plausible values for the proportion of discounts on bills that the computer program could generate. The reasoning in the response is incorrect; therefore, section 2 was scored as incorrect. In part (b) the response clearly demonstrates an understanding that quadrupling the sample size divides the margin of error by 2 and provides the correct new margin of error. Section 3 was scored as essentially correct. The student then uses the new margin of error in part (c) to construct a new confidence interval that does not contain 0.20. The student correctly concludes that because 0.20 lies outside of the confidence interval, there is convincing evidence that the program is not working as intended. Section 4 was scored as essentially correct. Because three sections were scored as essentially correct, and one section was scored as incorrect, the response earned a score of 3.

AP® STATISTICS
2015 SCORING COMMENTARY

Question 2 (continued)

Sample: 2C
Score: 2

In part (a-i) the response correctly concludes that because 0.20 lies within the given confidence interval, there is no convincing evidence that the program is not working as intended. Although this response has minimal communication, it is considered complete, and section 1 was scored as essentially correct. In part (a-ii) the student incorrectly states that the confidence interval provides evidence that the program is generating discounts with a probability of 0.15, not 0.20. The response also incorrectly states that because 0.15 is in the center of the interval, it is a more likely estimate of the parameter. A particular value within a confidence interval is not more or less likely to be the correct estimate of a parameter based on its location within the interval. Section 2 was scored as incorrect. An incorrect margin of error was calculated in part (b). The student did not demonstrate an understanding of the inverse square root relationship between sample size and margin of error. The student divided the margin of error by 4 instead of $\sqrt{4}$ to obtain an incorrect value of 0.015. Section 3 was scored as incorrect. The student uses the margin of error in part (c) to construct a new confidence interval. This confidence interval is correctly constructed based on the incorrect margin of error, and the student notes that 0.20 falls outside of this newly constructed confidence interval. The student correctly concludes, based on the interval, that there is convincing evidence that the program is not working as intended. Section 4 was scored as essentially correct. Because two sections were scored as essentially correct, and two sections were scored as incorrect, the response earned a score of 2.