

# AP<sup>®</sup> STATISTICS

## 2017 SCORING GUIDELINES

### Question 2

#### Intent of Question

The primary goals of this question were to assess a student's ability to (1) construct and interpret a confidence interval for a population proportion and (2) use a confidence interval for a proportion to find a confidence interval for a dollar amount that can be calculated using that proportion.

#### Solution

##### Part (a):

Step 1: Identify the appropriate confidence interval by name or formula and check appropriate conditions.

The appropriate procedure is a one-sample z-interval for a population proportion  $p$ . In this case, the population is all customers of the restaurant who ask for a water cup, and  $p$  is the proportion of that population who will fill the cup with a soft drink. The appropriate formula is  $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ .

Conditions:

1. Random sample
2. Large sample (number of successes  $n\hat{p} \geq 10$  and number of failures  $n(1-\hat{p}) \geq 10$ )

For condition 1 the stem of the problem states that a random sample of customers who asked for a water cup was used.

For condition 2 the number of successes (filled cup with soft drink) is 23 and the number of failures is 57, both of which are greater than 10.

Step 2: Correct mechanics

The sample proportion is  $\hat{p} = \frac{23}{80} = 0.2875$ . The confidence interval is

$$\begin{aligned} 0.2875 \pm 1.96 \sqrt{\frac{0.2875(1-0.2875)}{80}} \\ = 0.2875 \pm 1.96(0.0506) \quad \text{or } 0.1883 \text{ to } 0.3867. \\ = 0.2875 \pm 0.0992 \end{aligned}$$

Step 3: Interpretation

We can be 95 percent confident that in the population of all customers of the restaurant who ask for a water cup, the proportion who will fill it with a soft drink is between 0.1883 and 0.3867.

##### Part (b):

Using the confidence interval in part (a), a 95 percent interval estimate for the number of customers in June who asked for a water cup but then filled it with a soft drink is  $3,000 \times 0.1883$  to  $3,000 \times 0.3857$ , or 565 to 1,160. At a cost of \$0.25 per customer, a 95 percent interval estimate for the cost to the restaurant in June is \$141.25 to \$290.00.

# 2017 SCORING GUIDELINES

### Question 2 (continued)

#### Scoring

This question is scored in four sections. Section 1 consists of step 1 in part (a), section 2 consists of step 2 in part (a), section 3 consists of step 3 in part (a), and section 4 consists of part (b). Each section is scored as essentially correct (E), partially correct (P), or incorrect (I).

**Section 1** is scored as follows:

Essentially correct (E) if the one-sample z-interval for a proportion is identified (either by name or formula) *AND* both conditions (random sampling and large sample) are adequately addressed.

Partially correct (P) if the response identifies the correct procedure *BUT* adequately addresses only one of the two required conditions;

*OR*

if the response does not identify the correct procedure *BUT* adequately addresses both required conditions.

Incorrect (I) if the response does not meet the criteria for E or P.

*Notes:*

- Stating the large sample condition without verifying it is not sufficient. The response must use specific numerical values to adequately address the condition.
- If the response includes additional inappropriate conditions, such as  $n \geq 30$  or requiring a normal population, then the response earns at most P for section 1.
- Stating and checking a condition about the size of the sample relative to the size of the population is appropriate but not required.
- Any statement of hypotheses, description of the population, or definition of the parameter should be considered extraneous. However, if such statements are included and incorrect, they are considered as poor communication in terms of holistic scoring.

**Section 2** is scored as follows:

Essentially correct (E) if the response gives the correct 95 percent confidence interval. Supporting work is not required, but if included, it must be correct.

Partially correct (P) if the response gives a correct confidence interval with incorrect (but appropriate) supporting work shown;

*OR*

if the response gives an incorrect but reasonable confidence interval with appropriate supporting work shown — for instance, if a value other than 1.96 is used for the critical value or a value other than  $\frac{23}{80}$  is used for  $\hat{p}$ .

Incorrect (I) if the response does not meet the criteria for E or P.

*Notes:*

- Appropriate supporting working must have the form:  
proportion  $\pm$  (critical value)(SE/SD of proportion).

## 2017 SCORING GUIDELINES

## Question 2 (continued)

- A confidence interval that has both endpoints outside the interval from 0 to 1 is considered unreasonable.

Section 3 is scored as follows:

Essentially correct (E) if the response provides an appropriate interpretation of the interval that includes the following three components:

1. Conveys inference about a population proportion
2. Demonstrates a clear understanding that the parameter is the proportion of the water cup population that fills the cup with a soft drink
3. Mentions 95 percent confidence and interprets it correctly using words such as “We can be 95 percent confident” or “With 95 percent confidence”

Partially correct (P) if the response provides an appropriate interpretation of the interval that includes the first component *AND* only one of the other two components;

*OR*

if the response provides a correct interpretation of the confidence *level* in context without interpreting the specific interval.

Incorrect (I) if the response does not meet the criteria for E or P.

Notes:

- Clear indication of an inference to the random sample of 80 customers rather than to the population does not satisfy the first component and is scored I.
- Stating values that are unrealistic as proportions or percentages (including blanks) in the interpretation lowers the score one level (from E to P, or from P to I).
- When both the interpretation and the level of the interval are given, only the interpretation is scored. If the interpretation of the confidence level is incorrect, it is considered as poor communication in terms of holistic scoring.
- Any interpretation that implies the interval has a 95 percent chance (or possibility or probability) of capturing the population proportion is scored I.

Section 4 is scored as follows:

Essentially correct (E) if the response gives a correct interval estimate for the cost to the restaurant *AND* shows enough work to indicate how the interval was found.

Partially correct (P) if the response gives the correct interval estimate for the *number* of customers (565 to 1,160) who would fill the water cup with soda, including showing work, but does not multiply by \$0.25 to find the interval for the cost;

*OR*

if the response gives the correct interval estimate for the *expected cost* to the restaurant for an individual who asked for a water cup (\$0.05 to \$0.10), including showing work, but does not multiply by 3,000 to find the interval for the total cost;

*OR*

if the response makes a reasonable attempt to find the endpoints of the interval as  $(0.1883)(3,000)(0.25)$  and  $(0.3867)(3,000)(0.25)$ , but makes an error such as using the sample size of 80 instead of the population size of 3,000;

## 2017 SCORING GUIDELINES

## Question 2 (continued)

*OR*

if the response gives the correct interval without showing how it was found.

Incorrect (I) if the response does not meet the criteria for E or P.

Notes:

- The response in section 4 earns an E if it follows correctly from an incorrect interval found in part (a), even if the interval is unreasonable.
- Units (\$) are not required to earn an E.

Each essentially correct (E) section counts as 1 point, and a partially correct (P) section counts as  $\frac{1}{2}$  point.

**4 Complete Response**

**3 Substantial Response**

**2 Developing Response**

**1 Minimal Response**

If a response is between two scores (for example,  $2\frac{1}{2}$  points), use a holistic approach to decide whether to score up or down, depending on the strength of the response and communication.

**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  $0.15 \pm 0.03$  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.

**2015 SCORING GUIDELINES****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.

## 2015 SCORING GUIDELINES

## 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.

## 2015 SCORING GUIDELINES

## 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.