
AP[®] Statistics

Sample Student Responses and Scoring Commentary

Inside:

Free-Response Question 4

- Scoring Guidelines**
- Student Samples**
- Scoring Commentary**

Question 4: Focus on Inference**4 points****General Scoring Notes**

- This two-part question is scored in four sections. Each section is initially scored by determining if it meets the criteria for essentially correct (E), partially correct (P), or incorrect (I). Part (a) includes three sections that may appear in any order in the response. The first section includes identification of the appropriate confidence interval in part (a). The second section includes verifying the conditions for inference in part (a) and calculating the values of the endpoints of the confidence interval. The third section includes the interpretation of the confidence interval in part (a). The fourth section includes the response to part (b). The response is then categorized based on the scores assigned to each section and awarded an integer score between 0 and 4 (see the table at the end of the question).
- The model solution represents an ideal response to each section of the question, and the scoring criteria identify the specific components of the model solution that are used to determine the score.

	Model Solution	Scoring
Section 1	(a) The appropriate procedure is a one-sample z -interval for the proportion of all teenagers in the United States who would respond that they use a video streaming service every day.	<p>Essentially correct (E) if the response satisfies the following two components:</p> <ol style="list-style-type: none"> 1. Identifies the appropriate procedure as a one-sample z-interval by name or formula or by the calculations of the correct confidence interval endpoint values 2. States that the parameter of interest is the population proportion <p>Partially correct (P) if the response satisfies only one of the two components.</p> <p>Incorrect (I) if the response does not meet the criteria for E or P.</p>

Additional Notes:

- The response to component 2 concerning the statement of “population proportion” can be found in any of the three sections of part (a).
- Any notation used to represent sample proportion or population proportion should remain consistent throughout part (a).

	Model Solution	Scoring
Section 2 (a)	<p>This survey selected a random sample of 920 teenagers in the United States, which enables the interval to be generalized to the population of interest. This sample of 920 teenagers is less than 10% of the total number of teenagers in the United States, which is required as sampling was conducted without replacement from a finite population. In addition, there were more than 10 successes and 10 failures as $(920)(0.59) = 542.8$ (or 543) responded that they use a streaming service daily and $(920)(0.41) = 377.2$ (or 377) responded that they did not. Thus, the sample size is large enough to support the assumption that the sampling distribution of \hat{p} is approximately normal.</p> <p>Therefore, a 95% confidence interval for the population proportion is given by</p> $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = 0.59 \pm 1.96 \sqrt{\frac{(0.59)(0.41)}{920}},$ <p>which is 0.59 ± 0.032, and the interval is $(0.558, 0.622)$.</p>	<p>Essentially correct (E) if the response satisfies the following four components:</p> <ol style="list-style-type: none"> States that a random sample was selected Indicates 920 is less than ten percent of all teenagers in the United States Verifies that there are at least 10 successes and failures by calculating the following values $n\hat{p} = (920)(0.59) \approx 542.8$ and $n(1 - \hat{p}) = (920)(1 - 0.59) \approx 377.2$ Reports the values for a correct interval consistent with the procedure stated in Section 1 <p>Partially correct (P) if the response satisfies either component 3 or component 4 and at least one of the other three components.</p> <p>Incorrect (I) if the response does not meet the criteria for E or P.</p>

Additional Notes:

- Stating the large sample condition without verification is not sufficient for component 3.
- If the response includes an inappropriate check of conditions, such as $n > 30$, then component 3 is not satisfied.
- Supporting work, showing formulas or calculations, is not required for component 4.
- If the interval values are correct, the use of a one-sample z procedure for proportion can be used to satisfy component 1 of Section 1.
- A response that uses the value of $x = 543$ will result in an interval of $(0.5584, 0.6219)$, and a response that uses the value of $x = 542$ will result in an interval of $(0.5573, 0.6209)$. These interval endpoint values may be used to satisfy component 4.
- If the response includes supporting work for calculating the confidence interval that displays a correct formula with correct values inserted for \hat{p} , n , and z , then component 4 is satisfied even if values for the endpoints of the confidence interval are not displayed or calculated incorrectly.
- A response that computes an interval in percentages rather than proportions may satisfy component 4 if the response correctly indicates the use of percentages, $(55.8\%, 62.2\%)$.
- Minor errors or omissions when checking assumptions may be considered if holistic scoring is required.

	Model Solution	Scoring
(a) Section 3	We can be 95% confident that the proportion of all teenagers in the United States who would respond that they use a streaming service every day is between 0.558 and 0.622.	<p>Essentially correct (E) if the response satisfies the following two components:</p> <ol style="list-style-type: none"> 1. Indicates 95% confidence and interprets the interval using words such as “we are 95% confident” or “with 95% confidence” and provides interval endpoint values consistent with calculations in Section 2 2. Conveys inference about a population proportion in the proper context, i.e., “the true proportion of U.S. teenagers,” “the population proportion of U.S. teenagers,” or “the proportion of all U.S. teenagers” <p>Partially correct (P) if the response satisfies only one of the two components.</p> <p>Incorrect (I) if the response does not meet the criteria for E or P.</p>

Additional Notes:

- For component 1, responses that use the same incorrect proportion interval endpoint values calculated in Section 2 satisfies this portion of component 1.
- Section 3 is scored I if the response includes unreasonable values for a proportion or percentage.
- Responses must be consistent with the use of the terms proportion and percentage in component 2. A response that refers to the population percentage must use percentage values to satisfy component 2, and a response that refers to the population proportion must use proportion values.
- Any interpretation of the level of the interval, whether correct or incorrect, is considered extraneous and does not satisfy component 1.
- Clear indication of an inference to the sample of 920 teenagers, rather than the population of teenagers, does not satisfy component 2.
- Section 3 is scored I if it implies that the population proportion is a random variable. For example, it indicates that there is a 95 percent chance that the population proportion is between 0.558 and 0.622.
- A response that implies that we have found the actual percentage of U.S. teenagers who respond that they use a streaming service (deterministic) does not satisfy component 1.

	Model Solution	Scoring
Section 4	<p>(b) The 95% confidence interval of (0.558, 0.622) indicates that any value between 0.558 and 0.622 is a plausible value for the proportion of all teenagers in the United States who use video streaming every day. Because the value 0.5 is not contained in the interval, the sample data provide convincing statistical evidence that the proportion of all teenagers in the United States who would use a streaming service every day is not 0.5.</p>	<p>Essentially correct (E) if the response satisfies the following two components:</p> <ol style="list-style-type: none"> 1. Provides a correct conclusion (there is convincing evidence or there is not convincing evidence that 0.50 is a plausible value), consistent with the interval calculated in part (a) 2. Provides correct justification based on whether the value of 0.5 is contained in the interval calculated in part (a) <p>Partially correct (P) if the response satisfies only one of the two components.</p> <p>Incorrect (I) if the response does not meet the criteria for E or P.</p>

Additional Notes:

- Any use of a hypothesis test to answer Section 4 cannot be used to satisfy component 2. However, the statement of a conclusion that 0.50 is not a plausible value based on the results of the hypothesis test can be used to satisfy component 1.
- A response that correctly uses incorrectly calculated proportion interval endpoint values can satisfy components 1 and 2 if the response is consistent with interval values.
- A response that includes a correct interpretation of the confidence interval in Section 4 can receive credit for the interpretation in Section 3 if no interpretation was provided in Section 3.
- A response that presents a correct conclusion in the context of the population percentage may satisfy both component 1 and component 2.
- A response that indicates that there is evidence that the proportion of teenagers in the United States who would use a streaming service every day is greater than 0.5 based on the fact that the entire confidence interval is above 0.5 may satisfy both component 1 and component 2.
- A response that does not provide the correct conclusion and justification but does recognize that the value is not in the interval may be considered a positive if holistic scoring is required.

Scoring for Question 4

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

Score**Complete Response****4****Substantial Response****3****Developing Response****2****Minimal Response****1**

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 quality of the communication.

Question 4

Begin your response to **QUESTION 4** on this page.

4. A survey conducted by a national research center asked a random sample of 920 teenagers in the United States how often they use a video streaming service. From the sample, 59% answered that they use a video streaming service every day.

(a) Construct and interpret a 95% confidence interval for the proportion of all teenagers in the United States who would respond that they use a video streaming service every day.

$$1 \text{ Prop Z Int } 920 \cdot .59 = 542.8$$

Parameter: True proportion of teenagers in the United States who use a video streaming service every day.

- ① The sample of 920 teenagers in the United States is random
- ② $920 < 10\%$ of all teenagers in the United States
- ③ $920 \cdot .59 \geq 10$ $542.8 \geq 10$
 $920 \cdot .41 \geq 10$ $377.2 \geq 10$

$$(.55844, .622) = \text{confidence interval}$$

$$\hat{p} = .59$$

We are 95% confident that the true proportion of teenagers in the United States who use a video streaming service every day is between the interval .55844 to .622

Question 4

Continue your response to **QUESTION 4** on this page.

- (b) Based on the confidence interval in part (a), do the sample data provide convincing statistical evidence that the proportion of all teenagers in the United States who would respond that they use a video streaming service every day is not 0.5? Justify your answer.

Yes, the data provides statistical evidence that the proportion of all teenagers in the United States who would respond that they use a streaming service every day is not 0.5 because the confidence interval is greater than 0.5 at .55844 to .622

Use a pencil or a pen with black or dark blue ink. Do NOT write your name. Do NOT write outside the box.

0052968



Question 4

Begin your response to **QUESTION 4** on this page.

4. A survey conducted by a national research center asked a random sample of 920 teenagers in the United States how often they use a video streaming service. From the sample, 59% answered that they use a video streaming service every day.

(a) Construct and interpret a 95% confidence interval for the proportion of all teenagers in the United States who would respond that they use a video streaming service every day.

State: Let p equal the true proportion of United States teenagers who use video streaming services everyday. We will construct a 95% confidence interval for this situation.

Plan: Random \rightarrow "Random Sample" stated \checkmark

$$\text{Normal} \rightarrow (p)(n) = (.59)(920) = 542.8 \rightarrow 543 \geq 30 \checkmark$$

$$(n)(1-p) = (.41)(920) = 377.2 \rightarrow 377 \geq 30 \checkmark$$

Independent $\rightarrow n \leq 10\%$ of the population of teenagers in the United States \checkmark

Do: 1 prop z int (542, 920, .95) = $\hat{p} = 0.58913$
(0.55734, 0.62092)

Conclude: We are 95% confident that the true mean of the proportion of all United States teenagers who would respond that they use a video streaming service everyday is between (0.55734 and 0.62092)

Question 4

Continue your response to QUESTION 4 on this page.

- (b) Based on the confidence interval in part (a), do the sample data provide convincing statistical evidence that the proportion of all teenagers in the United States who would respond that they use a video streaming service every day is not 0.5? Justify your answer.

Yes, the 1.15 ^{confidence} evidence that the proportion of all teenagers in the US who would respond that they use a video streaming service everyday is not 0.5 as it does not lay within our 95% confidence interval.

Question 4

Begin your response to **QUESTION 4** on this page.

4. A survey conducted by a national research center asked a random sample of 920 teenagers in the United States how often they use a video streaming service. From the sample, 59% answered that they use a video streaming service every day.

- (a) Construct and interpret a 95% confidence interval for the proportion of all teenagers in the United States who would respond that they use a video streaming service every day.

$$920 \cdot 0.59 = 542.8$$

$$df = 920 - 1 = 919$$

$$z^* \approx 1.96$$

$$C = 0.59 \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$= 0.59 \pm 1.96 \sqrt{\frac{0.59(1-0.59)}{920}}$$

$$C = (0.562, 0.618)$$

Question 4

Continue your response to **QUESTION 4** on this page.

- (b) Based on the confidence interval in part (a), do the sample data provide convincing statistical evidence that the proportion of all teenagers in the United States who would respond that they use a video streaming service every day is not 0.5? Justify your answer.

Based on the confidence interval of $(0.562, 0.618)$, there is a probable chance that the proportion of all teens in the US who would respond that they use a video streaming service is not 0.5. With the confidence interval we can claim that the true proportion of teens in the US who would respond that they use a video streaming service is within the interval $(0.562, 0.618)$ with 95% confidence.

Use a pencil or a pen with black or dark blue ink. Do NOT write your name. Do NOT write outside the box.

0168325

Question 4

Note: Student samples are quoted verbatim and may contain spelling and grammatical errors.

Overview

The primary goals of the question were to assess a student's ability to (1) identify an appropriate procedure for constructing a confidence interval for a population proportion; (2) check conditions required for accurate application of the identified procedure; (3) calculate the confidence interval; (4) interpret the confidence interval in the context of the proportion of the relevant population that would give a specific response to a survey, and (5) use the confidence interval to determine if a statement about the population is justified.

This question primarily assesses skills in skill category 1: Selecting Statistical Methods, skill category 3: Using Probability and Simulation, and skill category 4: Statistical Argumentation. Skills required for responding to this question include (1.D) Identify an appropriate inference method for confidence intervals, (3.D) Construct a confidence interval, provided conditions for inference are met, (4.B) Interpret statistical calculations and findings to assign meaning or assess a claim, (4.C) Verify that inference procedures apply in a given situation, and (4.D) Justify a claim based on a confidence interval.

This question covers content from Unit 6: Inference for Categorical Data: Proportions of the course framework in the AP Statistics Course and Exam Description. Refer to topics 6.2 and 6.3 and learning objectives UNC-4.A, UNC-4.B, UNC-4.D, UNC-4.F, and UNC-4.G.

Sample: 4A

Score: 4

The response earned the following: Section 1 – E; Section 2 – E; Section 3 – E; Section 4 – E.

In part (a) section 1 the response satisfies component 1 as the response identifies the appropriate procedure as a “1 prop z int.” The response satisfies component 2 as the response states the parameter of interest is the “true proportion” when defining the parameter. Section 1 was scored essentially correct (E).

In part (a) section 2 the response satisfies component 1 as the response states that the sample is random. The response satisfies component 2 as the response states that “920 < 10% of all teenagers in the United States.” The response satisfies component 3 as the response verifies, by showing both the calculations and values, that there are at least 10 successes and 10 failures. The response satisfies component 4 as the response shows the correct confidence interval endpoint values. Section 2 was scored essentially correct (E).

In part (a) section 3 the response satisfies component 1 as the response states the confidence level by stating “we are 95% confident” and includes the interval endpoint values calculated in section 2. The response satisfies component 2 as the response includes the population parameter with context. Section 3 was scored essentially correct (E).

In part (b) section 4 the response satisfies component 1 as the response states, “Yes.” The response satisfies component 2 as the response notes that the interval values are all above 0.50. Section 4 was scored essentially correct (E).

Question 4 (continued)**Sample: 4B****Score: 3**

The response earned the following: Section 1 – P; Section 2 – P; Section 3 – E; Section 4 – E.

In part (a) section 1 the response satisfies component 1 as the response identifies the appropriate procedure as a “1 prop z int.” The response does not satisfy component 2 as the response states that “we will construct a 95% confidence interval for this observation” even though they also refer to the population. As only one component is satisfied, section 1 was scored partially correct (P).

In part (a) section 2 the response satisfies component 1 as the response states that the sample is random. The response satisfies component 2 as the response states that “ $n \leq 10\%$ of the population.” The response does not satisfy component 3 as the response uses the population proportion symbol in the calculation and proceeds to compare the value to 30. The response provides interval endpoint values consistent with using $x = 542$, which satisfies component 4. As component 4 and at least one other component are satisfied, section 2 was scored partially correct (P).

In part (a) section 3 the response correctly references the 95 percent confidence and the interval of proportions, satisfying component 1 while including an appropriate reference to the population through the phrase “all United States teenagers,” satisfying component 2. Section 3 was scored essentially correct (E).

In part (b) section 4 the response indicates “yes” there is statistically significant evidence, satisfying component 1. The response satisfies component 2 from the statement “[.5] does not lay within our 95% confidence interval.” Section 4 was scored essentially correct (E).

Sample: 4C**Score: 2**

The response earned the following: Section 1 – E; Section 2 – I; Section 3 – E; Section 4 – I.

In part (a) section 1 the response satisfies component 1 as the response identifies the appropriate procedure by writing the formula for a z -interval. The response satisfies component 2 by stating that the parameter of interest is the “true proportion” in the interpretation of the confidence interval, which is shown in part (b) of the response along with the confidence interval interpretation. Section 1 was scored as essentially correct (E).

In part (a) section 2 the response does not check any of the three inference conditions; therefore, it does not satisfy components 1, 2, or 3. The response satisfies component 4 by showing the values of the sample proportion, sample size, and z^* inserted into the formula, although the lower endpoint value is incorrectly calculated. Only one of the three components are satisfied, so section 2 was scored incorrect (I).

In part (a) section 3 the response to this section is provided within the part (b) section 4 response. Note that a response in section 4 that includes a correct interpretation of the confidence interval can receive credit for section 3 if no interpretation was provided in section 3. The response satisfies component 1 as the response states the confidence level by stating “with 95% confidence” and includes the interval endpoint values calculated in section 2. The response satisfies component 2 as the response includes the population parameter with context. Section 3 was scored as essentially correct (E).

Question 4 (continued)

In part (b) section 4 the response does not satisfy component 1 as the response never states if there is convincing statistical evidence. The response does not satisfy component 2 as the response does not use the interval to justify a decision about the value of 0.50. Section 4 was scored incorrect (I).