

---

# AP Statistics

## Sample Student Responses and Scoring Commentary

### **Inside:**

#### **Free Response Question 2**

- Scoring Guideline**
- Student Samples**
- Scoring Commentary**

**AP<sup>®</sup> STATISTICS**  
**2018 SCORING GUIDELINES**

**Question 2**

**Intent of Question**

The primary goals of this question were to assess a student's ability to (1) calculate the sample size when given the endpoints of a confidence interval for a proportion; (2) explain how bias could be present in a particular survey method; and (3) estimate a proportion from sample data collected using a method designed to decrease bias.

**Solution**

**Part (a):**

Using the standard formula for a confidence interval for one proportion, the interval (0.584 to 0.816) is found as follows.  $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$  where  $\hat{p} = \frac{0.584 + 0.816}{2} = 0.7$ , the margin of error is  $0.816 - 0.7 = 0.116$ , and  $z^* = 1.96$ .

Solving  $1.96\sqrt{\frac{0.7(1-0.7)}{n}} = 0.116$  yields  $n = \frac{(1.96)^2(0.7)(1-0.7)}{(0.116)^2} \approx 59.95$ . The sample size was 60.

**Part (b):**

Bias might have been introduced because students responded directly to the environmental science teacher. Because the students would know that an environmental science teacher cares about the environment, they might say yes when they actually don't recycle. This would result in a point estimate that is greater than the proportion of all students who would respond yes to the question.

**Part (c):**

(i) The expected number is  $(300)\left(\frac{1}{2}\right) = 150$ .

(ii) The point estimate is based on expecting 150 students to be required to say no and 150 students to truthfully answer the question. Of the 213 answers of no, we expect that  $213 - 150 = 63$  were from students who truthfully answered the question. That means we expect that the remaining  $150 - 63 = 87$  students truthfully answered the question and responded yes. So the point estimate for the proportion of all students at the high school who would respond yes to the question is  $\frac{87}{150} = 0.58$ .

**AP<sup>®</sup> STATISTICS**  
**2018 SCORING GUIDELINES**

**Question 2 (continued)**

**Scoring**

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

**Part (a)** is scored as follows:

Essentially correct (E) if the response satisfies the following five components:

1. Uses a standard error in the form  $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$  where  $\hat{p}$  is between 0 and 1.
2. Shows evidence that  $\hat{p} = 0.7$  was correctly used in the standard error.
3. Shows evidence that 0.116 was correctly used as the margin of error in the calculation.
4. Shows evidence that  $z^* = 1.96$  was correctly used as the critical value in the calculation.
5. Includes a single, positive whole-number answer.

Partially correct (P) if the response satisfies only three or four of the five components.

Incorrect (I) if the response satisfies at most two of the five components.

*Notes:*

- Using an equation in the form  $n = \frac{z^2 \hat{p}(1-\hat{p})}{\text{MOE}^2}$  satisfies component 1.
- A value of 0.21 in the numerator of the standard error implies that  $\hat{p} = 0.7$  was correctly used in the standard error and satisfies component 2.
- An equation such as  $0.816 = 0.7 + \text{MOE}$  implies that 0.116 was correctly used for the margin of error and satisfies component 3.
- Statements that suggest a whole-number answer is approximate (such as, “about 60” or “ $\approx 60$ ”) satisfy component 5.
- Algebraic work between the set-up and final answer does not need to be shown to satisfy component 5.
- When calculating the values 0.7, 0.116, or 1.96, ignore minor arithmetic errors or transcription errors if they can be identified by the work shown.

# AP<sup>®</sup> STATISTICS

## 2018 SCORING GUIDELINES

### Question 2 (continued)

**Part (b)** is scored as follows:

Essentially correct (E) if the response satisfies the following three components:

1. Explains why the responses to the survey might differ from the truth about student recycling in this context (for example, the survey was not anonymous, the question was asked by an authority figure).
2. Explains how the responses to the survey might differ from the truth about student recycling (for example, “students might say yes when they actually don’t recycle,” “students lie and say yes,” “students don’t recycle but lie to the teacher”).
3. Describes the effect of the bias on the point estimate (or the proportion, percentage, number of yes responses in the sample) and doesn’t contradict the bias described.

Partially correct (P) if the response satisfies only two of the three components.

Incorrect (I) if the response satisfies at most one of the three components.

*Notes:*

- To satisfy component 1 the response must provide a reason that is based on a bias created by the teacher asking students in person. For example, a response that addresses the wording of the question, voluntary response, or sampling variability does not satisfy component 1.
- To satisfy component 2 the response needs to explicitly contrast what the students say with what they do.
- Evidence used to address component 3 cannot also be used to address component 2. For example, a response that says “Students might lie, producing an estimate that is too high” addresses the effect of the bias on the point estimate but should not be combined with the statement about students lying to infer that students do not actually recycle. However, a response that says “Students may lie and say yes, producing an estimate that is too high” satisfies both components 2 and 3.
- If the response is clearly about the population proportion and not about the point estimate, component 3 cannot be satisfied.
- Statements such as “the interval will be too high” do not satisfy component 3 because they don’t specifically address the point estimate.

**AP<sup>®</sup> STATISTICS**  
**2018 SCORING GUIDELINES**

**Question 2 (continued)**

**Part (c)** is scored as follows:

Essentially correct (E) if the response gives an answer of 150 in (c-i) and gives an answer of 0.58 (or equivalent) in (c-ii).

Partially correct (P) if the response gives an answer of 150 in (c-i) and gives an answer of 0.42 (or equivalent) in (c-ii);

*OR*

if the response does not give an answer of 150 in (c-i) but gives an answer of 0.58 (or equivalent) with supporting work in (c-ii).

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

*Notes:*

- In part (c-i) the answer must be a single number. Responses such as “at least 150” or “147–153” are incorrect. However, responses such as “about 150” or “ $\approx 150$ ” are acceptable.
- In part (c-ii) the proportion can be described verbally (e.g., “87 out of 150”).
- In part (c-ii) if the response clearly indicates that 0.58 (or 0.42) is the population proportion, lower the overall score in part (c) by one level (that is, from E to P, or from P to I). Using probability notation such as  $P(\text{yes})$  does not clearly indicate a population proportion.
- In part (c-ii) if the response includes a point estimate of 0.58 or 0.42 but uses a confidence interval as the final answer, lower the overall score in part (c) by one level (that is, from E to P, or from P to I).
- If the answer is incorrect in part (c-i) and the answer in part (c-ii) uses numerator = 87 and denominator = 300 – answer to (c-i), the response should be scored P.

**AP<sup>®</sup> STATISTICS  
2018 SCORING GUIDELINES**

**Question 2 (continued)**

**4 Complete Response**

Three parts essentially correct

**3 Substantial Response**

Two parts essentially correct and one part partially correct

**2 Developing Response**

Two parts essentially correct and no parts partially correct

*OR*

One part essentially correct and one or two parts partially correct

*OR*

Three parts partially correct

**1 Minimal Response**

One part essentially correct

*OR*

No parts essentially correct and one or two parts partially correct

2A1

2A1

2. An environmental science teacher at a high school with a large population of students wanted to estimate the proportion of students at the school who regularly recycle plastic bottles. The teacher selected a random sample of students at the school to survey. Each selected student went into the teacher's office, one at a time, and was asked to respond yes or no to the following question.

Do you regularly recycle plastic bottles?

Based on the responses, a 95 percent confidence interval for the proportion of all students at the school who would respond yes to the question was calculated as (0.584, 0.816).

- (a) How many students were in the sample selected by the environmental science teacher?

$$\begin{aligned}
 &0.816 - 0.584 = 0.232 \quad \leftarrow \text{margin of error} \\
 &0.232 \div 2 = 0.116 \\
 &0.584 + 0.116 = 0.7 = p \\
 &1 - 0.7 = 0.3 = q \\
 &\frac{1 - 0.95}{2} = 0.025 \\
 &\text{inv Norm} \\
 &\left. \begin{array}{l} \text{area: } 0.025 \\ \mu: 0 \\ \sigma: 1 \end{array} \right\} 1.96 = z^* \\
 &0.7 \pm z^* \sqrt{\frac{pq}{n}} \\
 &0.7 \pm 1.96 \sqrt{\frac{(0.7)(0.3)}{n}} \\
 &0.116 = 1.96 \sqrt{\frac{(0.7)(0.3)}{n}} \\
 &n \approx 60 \text{ students}
 \end{aligned}$$

60 students were in the sample selected by the environmental science teacher.

- (b) Given the method used by the environmental science teacher to collect the responses, explain how bias might have been introduced and describe how the bias might affect the point estimate of the proportion of all students at the school who would respond yes to the question.

Bias might have been introduced in terms of response bias. The environmental science teacher is likely to support regularly recycling plastic bottles, so students might have been more likely to respond yes to the question, even if that was not their true answer. This bias might affect the point estimate of the proportion of all students at the school who would respond yes to the question by causing the point estimate to be too high. Due to the pressure to answer a certain way (yes) that the students in the sample likely felt, it is likely that the point estimate was too high.

2A2

2A2

(c) The statistics teacher at the high school was concerned about the potential bias in the survey. To obtain a potentially less biased estimate of the proportion, the statistics teacher used an alternate method for collecting student responses. A random sample of 300 students was selected, and each student was given the following instructions on how to respond to the question.

- In private, flip a fair coin.
  - If heads, you must respond no, regardless of whether you regularly recycle.
  - If tails, please truthfully respond yes or no.
- (i) What is the expected number of students from the sample of 300 who would be required to respond no because the coin flip resulted in heads?

$$300 \left( \frac{1}{2} \right) = 150$$

The expected number of students from the sample of 300 who would be required to respond no because the coin flip resulted in heads, is 150 students. Since there are only 2 potential outcomes to the coin flip, about half of the sample of 300 would be expected to get a coin flip of heads and be required to respond no.

- (ii) The results of the sample showed that 213 of the 300 selected students responded no. Based on the results of the sample, give a point estimate for the proportion of all students at the high school who would respond yes to the question.

$$213 - 150 = 63$$

About 63 students got a coin flip of tails, meaning they truthfully responded no.

$$150 - 63 = 87$$

$$\frac{87}{150} = 0.58$$

Out of the 150 students who got a coin flip of tails, about 63 of them responded no. Thus, the other 87 students responded yes truthfully (a proportion of 0.58).

A point estimate for the proportion of all students at the high school who would respond yes to the question is 0.58



2B1

2B1

2. An environmental science teacher at a high school with a large population of students wanted to estimate the proportion of students at the school who regularly recycle plastic bottles. The teacher selected a random sample of students at the school to survey. Each selected student went into the teacher's office, one at a time, and was asked to respond yes or no to the following question.

Do you regularly recycle plastic bottles?

Based on the responses, a 95 percent confidence interval for the proportion of all students at the school who would respond yes to the question was calculated as (0.584, 0.816).

- (a) How many students were in the sample selected by the environmental science teacher?

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

$$\hat{p} = \frac{.584 + .816}{2} = 0.7$$

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

$$MOE = .816 - .7 = .116$$

$$.116 = 1.96 \sqrt{\frac{.7(.3)}{n}}$$

$$\frac{.116}{1.96} = \sqrt{\frac{.7(.3)}{n}}$$

$$\frac{.116}{1.96} = \frac{.7(.3)}{n} \Rightarrow n = \frac{(.7(.3))}{\left(\frac{.116}{1.96}\right)^2} = 59.95$$

$\approx 60$  students

- (b) Given the method used by the environmental science teacher to collect the responses, explain how bias might have been introduced and describe how the bias might affect the point estimate of the proportion of all students at the school who would respond yes to the question.

The students might feel pressure to answer yes since they were being surveyed by an environmental science teacher who would likely promote recycling. Thus, the point-estimate might overestimate the <sup>true</sup> proportion of all students who recycle plastic bottles.

2B2

2B2

(c) The statistics teacher at the high school was concerned about the potential bias in the survey. To obtain a potentially less biased estimate of the proportion, the statistics teacher used an alternate method for collecting student responses. A random sample of 300 students was selected, and each student was given the following instructions on how to respond to the question.

- In private, flip a fair coin.
  - If heads, you must respond no, regardless of whether you regularly recycle.
  - If tails, please truthfully respond yes or no.
- (i) What is the expected number of students from the sample of 300 who would be required to respond no because the coin flip resulted in heads?

$$p(\text{heads}) (300) = 150 \text{ students}$$

(ii) The results of the sample showed that 213 of the 300 selected students responded no. Based on the results of the sample, give a point estimate for the proportion of all students at the high school who would respond yes to the question.

$$213 - 150 = 63$$

→ expected number of "no" responses from flipping heads

We expect 63 students to have responded no after flipping tails, and these would be the truthful responses

$$\hat{p}_{\text{no}} = \frac{63}{150} = .42$$

$$\hat{p}_{\text{yes}} = 1 - .42 = \boxed{.58}$$

201

201

2. An environmental science teacher at a high school with a large population of students wanted to estimate the proportion of students at the school who regularly recycle plastic bottles. The teacher selected a random sample of students at the school to survey. Each selected student went into the teacher's office, one at a time, and was asked to respond yes or no to the following question.

Do you regularly recycle plastic bottles?

Based on the responses, a 95 percent confidence interval for the proportion of all students at the school who would respond yes to the question was calculated as (0.584, 0.816).

- (a) How many students were in the sample selected by the environmental science teacher?

$$\text{point estimate} = (0.816 + 0.584) / 2 = 0.7$$

$$(0.584, 0.816) = 0.7 \pm 1.96 \sqrt{\frac{(0.7)(0.3)}{n}}$$

$$0.584 = 0.7 - 1.96 \sqrt{\frac{(0.7)(0.3)}{n}}$$

$$\boxed{n \approx 60 \text{ students}}$$

- (b) Given the method used by the environmental science teacher to collect the responses, explain how bias might have been introduced and describe how the bias might affect the point estimate of the proportion of all students at the school who would respond yes to the question.

Since the teacher asking the question is an environmental science teacher, the students may feel that they need to show that they recycle regularly when in reality they might not. Due to this bias the point estimate of the proportion may be higher than it actually is. This is an example of a response bias.

202

202

- (c) The statistics teacher at the high school was concerned about the potential bias in the survey. To obtain a potentially less biased estimate of the proportion, the statistics teacher used an alternate method for collecting student responses. A random sample of 300 students was selected, and each student was given the following instructions on how to respond to the question.
- In private, flip a fair coin.
  - If heads, you must respond no, regardless of whether you regularly recycle.
  - If tails, please truthfully respond yes or no.
- (i) What is the expected number of students from the sample of 300 who would be required to respond no because the coin flip resulted in heads?

$$300 (0.5) = \boxed{150 \text{ students}}$$

- (ii) The results of the sample showed that 213 of the 300 selected students responded no. Based on the results of the sample, give a point estimate for the proportion of all students at the high school who would respond yes to the question.

$$\frac{(300 - 213)}{300} = \boxed{0.29}$$

# AP<sup>®</sup> STATISTICS

## 2018 SCORING COMMENTARY

### Question 2

#### Overview

The primary goals of this question were to assess a student's ability to (1) calculate the sample size when given the endpoints of a confidence interval for a proportion; (2) explain how bias could be present in a particular survey method; and (3) estimate a proportion from sample data collected using a method designed to decrease bias.

#### Sample: 2A

##### Score: 4

In part (a) the response includes the formula  $\sqrt{\frac{pq}{n}}$  for the standard error, satisfying component 1. Because the difference between a point estimate and population proportion is assessed in parts (b) and (c), using  $p$  and  $q$  instead of  $\hat{p}$  and  $\hat{q}$  in the formula was considered a minor error. In the equation  $0.116 = 1.96\sqrt{\frac{(0.7)(0.3)}{n}}$ , the response correctly uses  $\hat{p} = 0.7$  in the standard error, which satisfies component 2, correctly uses 0.116 as the margin of error, which satisfies component 3, and correctly uses 1.96 as the critical value, which satisfies component 4. Finally, the response includes a single, positive whole-number answer (60), satisfying component 5. Because the response includes all five components, part (a) was scored as essentially correct. In part (b) the response provides a reason why the survey responses might be different from the truth (“Due to the pressure to answer a certain way”), satisfying component 1. The response explains how the survey responses differ from the truth (“more likely to respond yes to the question, even if that was not their true answer”), satisfying component 2. Finally, the response describes the effect of the bias on the point estimate (“causing the point estimate to be too high”), satisfying component 3. Because the response includes all three components, part (b) was scored as essentially correct. In part (c-i) the response gives an answer of 150. In part (c-ii) the response provides an answer of 0.58. Because the response includes both correct answers, part (c) was scored as essentially correct. Because three parts were scored as essentially correct, the response earned a score of 4.

#### Sample: 2B

##### Score: 3

In part (a) the response includes the equation  $0.116 = 1.96\sqrt{\frac{.7(.3)}{n}}$ , which uses a standard error in the correct form, satisfying component 1. In the same equation, the response also correctly uses  $\hat{p} = 0.7$  in the standard error, which satisfies component 2, correctly uses 0.116 as the margin of error, which satisfies component 3, and correctly uses 1.96 as the critical value, satisfying component 4. Finally, the response includes a single, positive whole-number answer (60), satisfying component 5. Because the response includes all five components, part (a) was scored as essentially correct. In part (b) the response provides a reason why the survey responses might be different than the truth (“the students might feel pressure to answer yes”), satisfying component 1. While the response indicates that the students are likely to say yes (“pressure to answer yes”), the response does not indicate that the students do not actually recycle. Because the response does not describe how the survey responses would differ from the truth, component 2 is not satisfied. Finally, the response describes the effect of the bias on the point estimate (“the point-estimate might overestimate the true proportion of all students who recycle plastic bottles”), satisfying component 3. Because the response includes two of the three components, part (b) was scored as partially correct. In part (c-i) the response gives an answer of 150. In part (c-ii) the response provides an answer of 0.58. Because the response includes both correct answers, part (c) was scored as essentially correct. Because two parts were scored as essentially correct, and one part was scored as partially correct, the response earned a score of 3.

**AP<sup>®</sup> STATISTICS**  
**2018 SCORING COMMENTARY**

**Question 2 (continued)**

**Sample: 2C**

**Score: 2**

In part (a), the response includes the formula  $0.584 = 0.7 - 1.96\sqrt{\frac{(0.7)(0.3)}{n}}$ , which uses a standard error in the correct form, satisfying component 1. In the same equation, the response also correctly uses  $\hat{p} = 0.7$  in the standard error, which satisfies component 2, and correctly uses 1.96 as the critical value, which satisfies component 4. Furthermore, the equation implies that 0.116 was correctly used for the margin of error (described in the third scoring note of the scoring guidelines), satisfying component 3. Finally, the response includes a single, positive whole-number answer (60), satisfying component 5. Because the response includes all five components, part (a) was scored as essentially correct. In part (b) the response provides a reason why the survey responses might be different from the truth (“Since the teacher asking the question is an environmental science teacher”), satisfying component 1. The response explains how the survey responses differ from the truth (“show that they recycle regularly when in reality they might not”), satisfying component 2. Finally, the response describes the effect of the bias on the point estimate (“the point estimate of the proportion may be higher than it actually is”), satisfying component 3. Because the response includes all three components, part (b) was scored as essentially correct. In part (c-i) the response gives an answer of 150. In part (c-ii) the response provides an answer of 0.29, which is the overall proportion of “yes” responses in the sample. Because the response in part (c-ii) does not account for the 150 students who would be expected to say “no” because of the coin flip, the response was scored as incorrect. Because two parts were scored as essentially correct, and one part was scored as incorrect, the response earned a score of 2.