AP® STATISTICS
2011 SCORING GUIDELINES

Question 6

Intent of Question

The primary goals of this question were to assess students' ability to (1) construct and interpret a confidence interval for a population proportion; (2) create a probability tree to represent a particular random process; (3) use a probability tree to calculate a probability; and (4) integrate provided information to create a confidence interval for an atypical parameter.

Solution

Part (a):

The appropriate inference procedure is a one-sample z-interval for a population proportion p, where p is the proportion of all United States twelfth-grade students who would answer the question correctly.

The conditions for this inference procedure are satisfied because:

1. The question states that the students are a random sample from the population, and
2. $n \times \hat{p} = 9{,}600 \times 0.28 = 2{,}688$ and $n \times (1 - \hat{p}) = 9{,}600 \times 0.72 = 6{,}912$ are both much larger than 10.

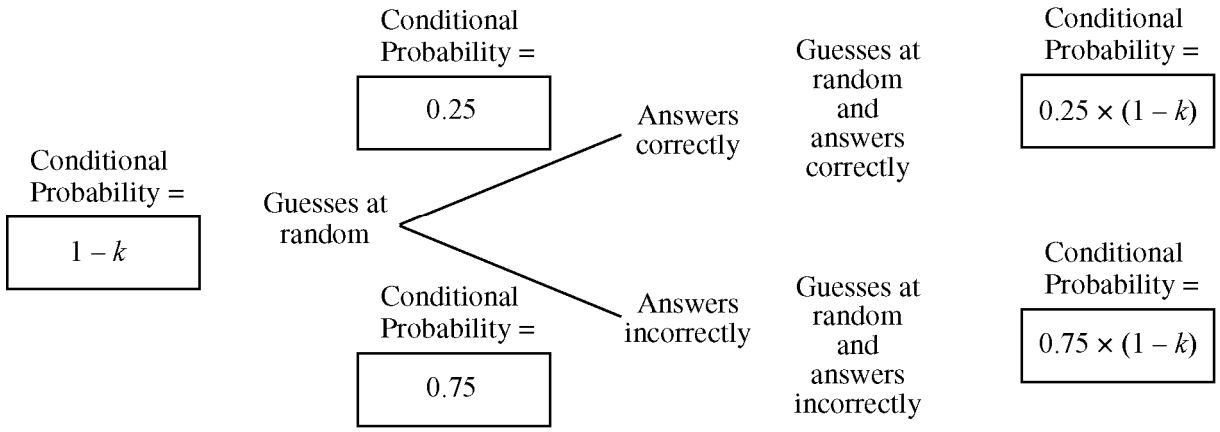
A 99 percent confidence interval for the population proportion p is constructed as follows:

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.28 \pm 2.576 \sqrt{\frac{0.28(0.72)}{9{,}600}}$$
$$= 0.28 \pm 0.012$$
$$\rightarrow (0.268, 0.292)$$

We are 99 percent confident that the interval from 0.268 to 0.292 contains the population proportion of all United States twelfth-grade students who would answer this question correctly.

Part (b):

The five probabilities to be filled in the boxes are shown below.



AP® STATISTICS
2011 SCORING GUIDELINES

Question 6 (continued)

Part (c):

P(answers correctly) =
P(knows correct answer and answers correctly) + P(guesses at random and answers correctly) =

$k + 0.25 \times (1 - k)$, which simplifies to $0.25 + 0.75k$, or $\frac{3k+1}{4}$.

Part (d):

We want to estimate k, the proportion of all United States twelfth-grade students who actually know the answer to the history question.

From part (c) the probability that a randomly selected student correctly answers the question is $0.25 + 0.75k$. From part (a) we are 99 percent confident that this probability is between 0.268 and 0.292. Thus the endpoints for a confidence interval for k can be found by equating the expression $0.25 + 0.75k$ from part (c) to the endpoints of the interval from part (a) as follows:

$$0.25 + 0.75k = 0.268 \qquad 0.25 + 0.75k = 0.292$$
$$k = 0.024 \qquad k = 0.056$$

We are 99 percent confident that the interval from 0.024 to 0.056 contains the proportion of all United States twelfth-grade students who actually know the answer to the history question.

Scoring

This question is scored in four sections. Sections 1 and 2 are based on part (a), section 3 consists of parts (b) and (c) and section 4 consists of part (d). 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 response correctly includes the following three components:
1. Identifies the correct inference procedure.
2. Checks the randomness condition.
3. Checks the large sample size condition.

Partially correct (P) if the response correctly includes exactly two of the three components listed above.

Incorrect (I) if the response fails to meet the criteria for E or P.

Notes
- The identification of the procedure must include "z," "proportion," and "interval."
- Stating the correct formula for a confidence interval for a proportion is sufficient for the first component.
- "Random sample given" is sufficient for the second component.

Question 6 (continued)

- To satisfy the third component, the response:
  - Must check both the number of successes and the number of failures.
  - Must use a reasonable criterion (for example, ≥ 5 or ≥ 10).
  - Must provide numerical evidence (for example, $2{,}688 \geq 10$ and $6{,}912 \geq 10$, or $9{,}600 \times 0.28 \geq 10$ and $9{,}600 \times 0.72 \geq 10$).
- Any statement of hypotheses, definitions of parameters, statements of populations, etc. should be considered extraneous. However, if such statements are included and incorrect, this should be considered poor communication in terms of holistic scoring.
- Any checks of reasonable conditions, such as independence of observations, sample size less than 10 percent of population size, $9{,}600 > 30$, etc. should be considered extraneous. However, if a response includes an incorrect condition, such as population normality, reduce the score in section 1 from E to P or from P to I.
- Any reference to the central limit theorem should be treated as extraneous and not sufficient for the large sample size condition.

Section 2 is scored as follows:

Essentially correct (E) if the response correctly includes the following two components:
1. Calculates the interval.
2. Interprets the interval, including a confidence statement and correct parameter, in context.

Notes
- The critical value for the confidence interval must be for 99 percent confidence.
- If the response includes an incorrect formula or has incorrect values substituted into the formula, then the response does not earn credit for the calculation component, even if the final interval is correct.
- A response that makes minor arithmetic mistakes in the calculation of the interval is considered correct, as long as the resulting interval is reasonable.
- A correct interval that is stated only in the interpretation is considered sufficient for the first component.
- To identify the parameter, the response must refer to the proportion "who would answer the question correctly" or include a modifier for the proportion such as "population" or "true." An interpretation about the sample proportion (for example, "the proportion of students who answered correctly") is not sufficient for the second component.
- If the response provides only an interpretation of the confidence level instead of the confidence interval, the second component is considered incorrect. If an interpretation of the confidence level is given along with an interpretation of the confidence interval, both must be correct to be considered sufficient.
- A correct interpretation with an incorrect interval is sufficient for the second component.

Partially correct (P) if the response correctly includes exactly one of the two components.

Incorrect (I) if the response fails to meet the criteria for E or P.

Question 6 (continued)

Section 3 is scored as follows:

Essentially correct (E) if the response correctly includes the following two components:
1. In part (b) completes the tree diagram in terms of k.
2. In part (c) adds the correct results from the tree diagram.

Notes
- If a response states "not k" or "k^C" in the first box, the first component is considered incorrect.
- If the response to part (b) is incorrect, then part (c) is considered correct if the response is consistent with the response from part (b) or if the response to part (c) is correct.
- The response to part (c) does not need to show a simplified expression.
- The response to part (c) can be expressed as a fraction with the sum of the four branches in the denominator.
- A response to part (c) that adds the appropriate probabilities from the tree but has an error in the simplification of the sum is still considered correct.
- If the response to part (c) is expressed as $P(0.25 + 0.75k)$ or equivalent, the second component is considered incorrect.
- If the tree diagram includes numbers only, adding the appropriate values is sufficient for the second component, provided that the sum is between 0 and 1.

Partially correct (P) if the response correctly includes exactly one of the two components.

Incorrect (I) if the response fails to meet the criteria for E or P.

Section 4 is scored as follows:

Essentially correct (E) if the response correctly includes the following three components:
1. Equates the expression from part (c) to a numerical estimate from part (a).
2. Uses the endpoints from part (a) to calculate a reasonable interval.
3. Interprets the resulting interval, including a confidence statement and correct parameter, in context.

Partially correct (P) if the response correctly includes exactly two of the three components.

Incorrect (I) if the response fails to meet the criteria for E or P.

Notes
- Using the point estimate $\hat{p} = 0.28$ from part (a) or the endpoints of the interval $(0.268, 0.292)$ from part (a) is sufficient for the first component.
- A response that makes minor arithmetic mistakes in the calculation of the interval is considered correct, as long as the resulting interval is reasonable.
- For the third component, the parameter must be the proportion of students who actually know the answer to the history question.
- A response that creates a correct interval using linear transformations (of the point estimate and standard error/margin of error) is equivalent to transforming the endpoints and therefore is sufficient for the first two components.

AP® STATISTICS
2011 SCORING GUIDELINES

Question 6 (continued)

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, particularly in parts (a) and (d). However, a response that earns a P or an I in section 4 cannot receive a score of 4.

6A1 6A1

STATISTICS
SECTION II
Part B
Question 6
Spend about 25 minutes on this part of the exam.
Percent of Section II score—25

Directions: Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

6. Every year, each student in a nationally representative sample is given tests in various subjects. Recently, a random sample of 9,600 twelfth-grade students from the United States were administered a multiple-choice United States history exam. One of the multiple-choice questions is below. (The correct answer is C.)

> In 1935 and 1936 the Supreme Court declared that important parts of the New Deal were unconstitutional. President Roosevelt responded by threatening to
>
> (A) impeach several Supreme Court justices
> (B) eliminate the Supreme Court
> (C) appoint additional Supreme Court justices who shared his views
> (D) override the Supreme Court's decisions by gaining three-fourths majorities in both houses of Congress

Of the 9,600 students, 28 percent answered the multiple-choice question correctly.

(a) Let p be the proportion of all United States twelfth-grade students who would answer the question correctly. Construct and interpret a 99 percent confidence interval for p.

$\hat{p}=.28$
$n=9600$

assumptions: - SRS: random sample stated
- pop at least 10x sample size: reasonable to assume 9600(10) or 96000 12th graders in USA.
- $np \geq 10$ $9600(.28)$ $2688 \geq 10$ ✓
- $n(1-p) \geq 10$ $9600(1-.28)$ $6912 \geq 10$ ✓

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = .28 \pm 2.576 \sqrt{\frac{.28(1-.28)}{9600}} = (.2682, .2918)$$

I am 99% confident that the true proportion p, of 12th grade USA students who would answer the question correctly, is between .2682 and .2918.

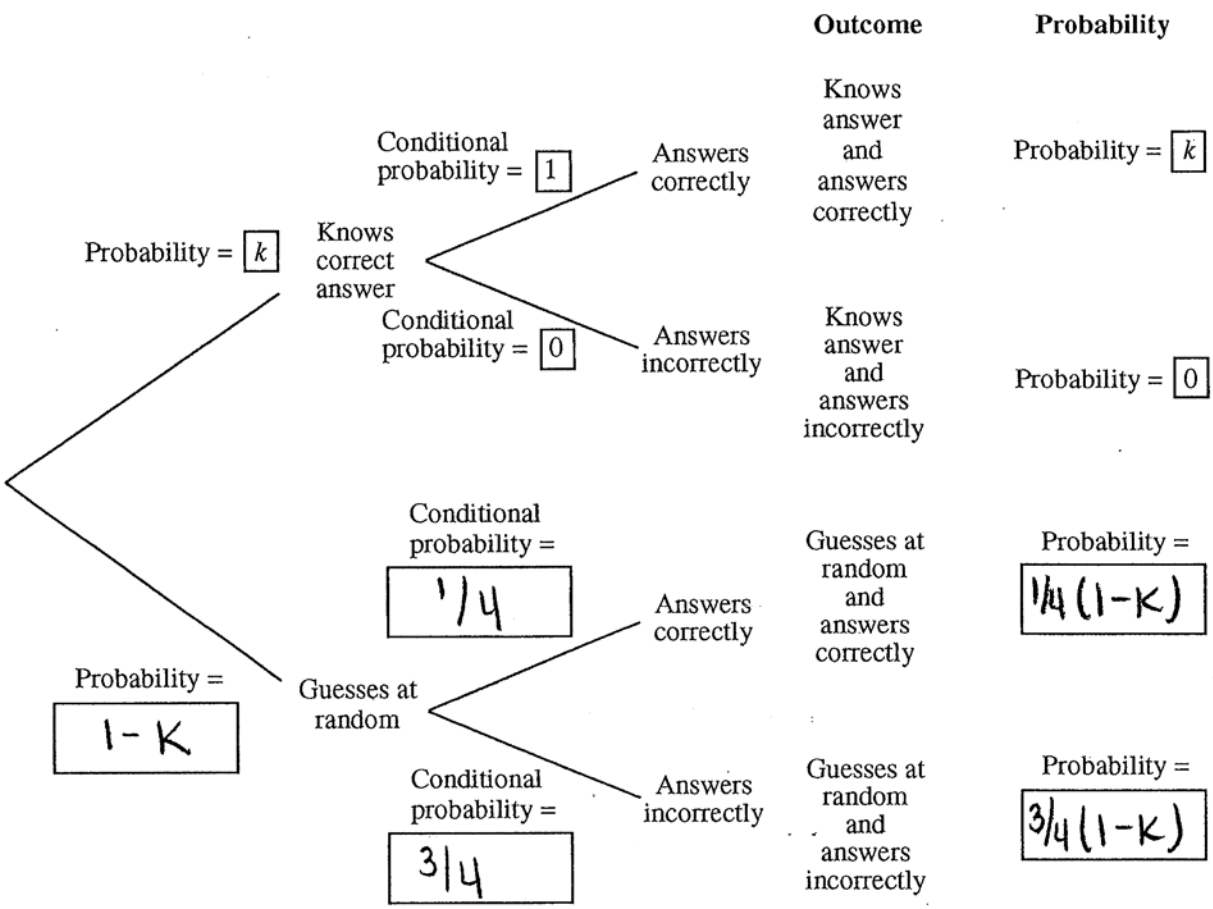
GO ON TO THE NEXT PAGE.

Assume that students who actually know the correct answer have a 100 percent chance of answering the question correctly, and students who do not know the correct answer to the question guess completely at random from among the four options.

Let k represent the proportion of all United States twelfth-grade students who actually know the correct answer to the question.

(b) A tree diagram of the possible outcomes for a randomly selected twelfth-grade student is provided below. Write the correct probability in each of the five empty boxes. Some of the probabilities may be expressions in terms of k.

TREE DIAGRAM OF OUTCOMES FOR A RANDOMLY SELECTED TWELFTH-GRADE STUDENT



(c) Based on the completed tree diagram, express the probability, in terms of k, that a randomly selected twelfth-grade student would correctly answer the history question.

$K + 1/4(1-K)$

GO ON TO THE NEXT PAGE.

6A3

(d) Using your interval from part (a) and your answer to part (c), calculate and interpret a 99 percent confidence interval for k, the proportion of all United States twelfth-grade students who actually know the answer to the history question. You may assume that the conditions for inference for the confidence interval have been checked and verified.

$.2682 = K + 1/4(1-K)$ $\qquad$ $.2918 = K + 1/4(1-K)$

$K = .0243$ $\qquad$ $K = .0557$

(.0243, .0557)

I am 99% confident that the true value for K, the proportion of all U.S. 12th graders who actually know the answer to the question, is between .0243 and .0557.

STOP

END OF EXAM

THE FOLLOWING INSTRUCTIONS APPLY TO THE COVERS OF THE SECTION II BOOKLET.

- **MAKE SURE YOU HAVE COMPLETED THE IDENTIFICATION INFORMATION AS REQUESTED ON THE FRONT AND BACK COVERS OF THE SECTION II BOOKLET.**
- **CHECK TO SEE THAT YOUR AP NUMBER LABEL APPEARS IN THE BOX(ES) ON THE COVER(S).**
- **MAKE SURE YOU HAVE USED THE SAME SET OF AP NUMBER LABELS ON ALL AP EXAMS YOU HAVE TAKEN THIS YEAR.**

6B1

6B1

STATISTICS
SECTION II
Part B
Question 6
Spend about 25 minutes on this part of the exam.
Percent of Section II score—25

Directions: Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

6. Every year, each student in a nationally representative sample is given tests in various subjects. Recently, a random sample of 9,600 twelfth-grade students from the United States were administered a multiple-choice United States history exam. One of the multiple-choice questions is below. (The correct answer is C.)

In 1935 and 1936 the Supreme Court declared that important parts of the New Deal were unconstitutional. President Roosevelt responded by threatening to (A) impeach several Supreme Court justices (B) eliminate the Supreme Court (C) appoint additional Supreme Court justices who shared his views (D) override the Supreme Court's decisions by gaining three-fourths majorities in both houses of Congress

Of the 9,600 students, 28 percent answered the multiple-choice question correctly.

(a) Let p be the proportion of all United States twelfth-grade students who would answer the question correctly. Construct and interpret a 99 percent confidence interval for p.

I. population: all US 12th grade students
parameter: proportion (p) who could answer the question correctly

II SRS? yes, given Normality? yes, sample is 9600, which is > 30

III. 1-sample-z proportion 99% Confidence Interval

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

$.28 \pm 2.576 \left(\sqrt{\frac{(.28)(.72)}{9600}}\right)$

$(.2682, .2918)$

IV. We are 99% confident that the true proportion of all 12th grade US students who would answer the question correctly is between .2682 and .2918.

GO ON TO THE NEXT PAGE.

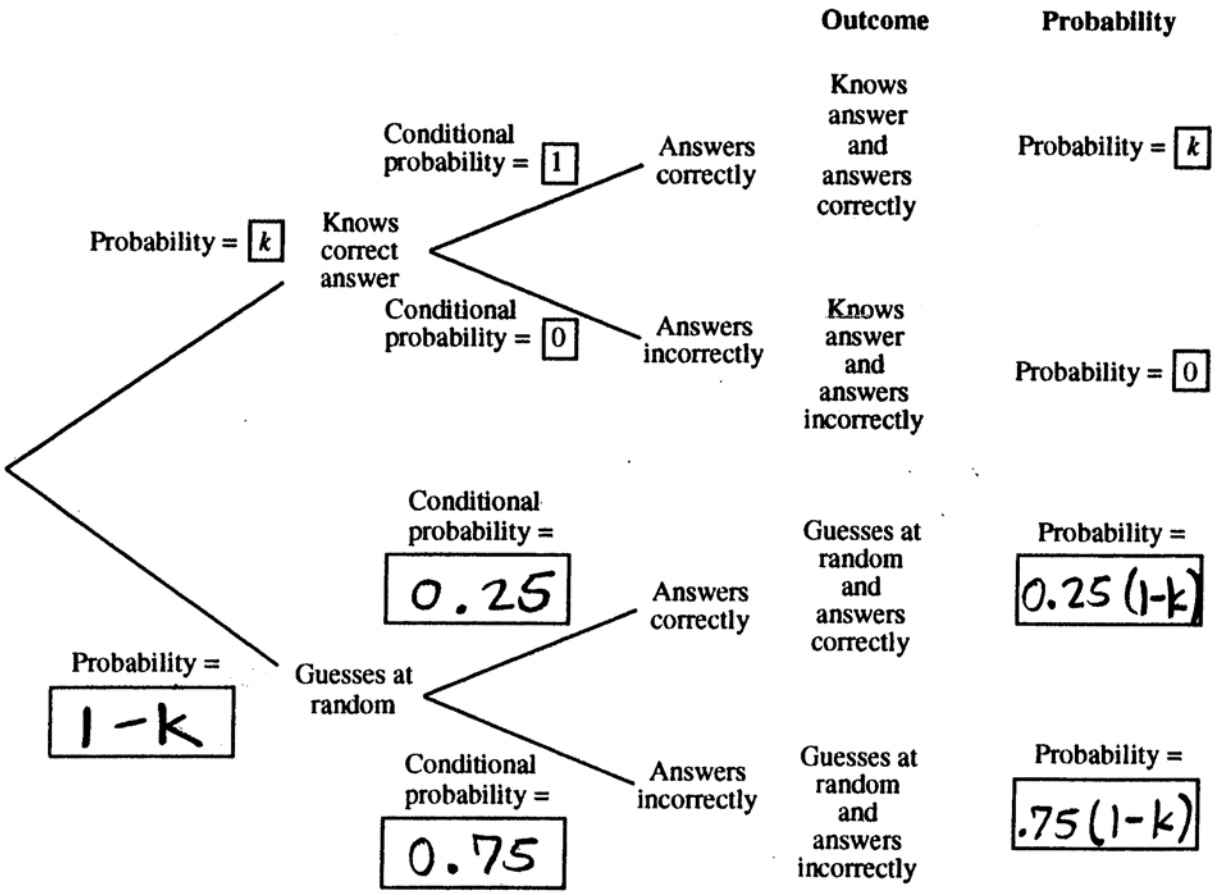
GB2

Assume that students who actually know the correct answer have a 100 percent chance of answering the question correctly, and students who do not know the correct answer to the question guess completely at random from among the four options.

Let k represent the proportion of all United States twelfth-grade students who actually know the correct answer to the question.

(b) A tree diagram of the possible outcomes for a randomly selected twelfth-grade student is provided below. Write the correct probability in each of the five empty boxes. Some of the probabilities may be expressions in terms of k.

TREE DIAGRAM OF OUTCOMES FOR A RANDOMLY SELECTED TWELFTH-GRADE STUDENT



(c) Based on the completed tree diagram, express the probability, in terms of k, that a randomly selected twelfth-grade student would correctly answer the history question.

p(random 12th grade student answers correctly) =
p(knows answer and answers correctly) +
p(guesses and answers correctly)

$p = k + 0.25(1 - k)$

GO ON TO THE NEXT PAGE.

6B3

(d) Using your interval from part (a) and your answer to part (c), calculate and interpret a 99 percent confidence interval for k, the proportion of all United States twelfth-grade students who actually know the answer to the history question. You may assume that the conditions for inference for the confidence interval have been checked and verified.

I. population: all US 12th grade students
parameter: proportion who actually know answer, (k)

II. SRS? yes given Normality? yes, given

III. 1-sample - z proportion 99% CI

$\hat{p} = \hat{k} + 0.25(1-\hat{k})$
$.28 = \hat{k} + 0.25 - 0.25\hat{k}$
$.28 = 0.25 + .75\hat{k}$
$.03 = .75\hat{k}$
$\hat{k} = 0.04$

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

$0.04 \pm 2.576\left(\sqrt{\frac{.04(.96)}{9600}}\right)$

$(0.03485, 0.04515)$

IV. We are 99% confident that the true proportion of US 12th grade students who actually know the answer to the history question is between 0.03485 and 0.04515.

STOP

END OF EXAM

THE FOLLOWING INSTRUCTIONS APPLY TO THE COVERS OF THE SECTION II BOOKLET.

- **MAKE SURE YOU HAVE COMPLETED THE IDENTIFICATION INFORMATION AS REQUESTED ON THE FRONT AND BACK COVERS OF THE SECTION II BOOKLET.**
- **CHECK TO SEE THAT YOUR AP NUMBER LABEL APPEARS IN THE BOX(ES) ON THE COVER(S).**
- **MAKE SURE YOU HAVE USED THE SAME SET OF AP NUMBER LABELS ON ALL AP EXAMS YOU HAVE TAKEN THIS YEAR.**

6C 1

6C 1

STATISTICS
SECTION II
Part B
Question 6
Spend about 25 minutes on this part of the exam.
Percent of Section II score—25

Directions: Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

6. Every year, each student in a nationally representative sample is given tests in various subjects. Recently, a random sample of 9,600 twelfth-grade students from the United States were administered a multiple-choice United States history exam. One of the multiple-choice questions is below. (The correct answer is C.)

In 1935 and 1936 the Supreme Court declared that important parts of the New Deal were unconstitutional. President Roosevelt responded by threatening to (A) impeach several Supreme Court justices (B) eliminate the Supreme Court (C) appoint additional Supreme Court justices who shared his views (D) override the Supreme Court's decisions by gaining three-fourths majorities in both houses of Congress

Of the 9,600 students, 28 percent answered the multiple-choice question correctly.

(a) Let p be the proportion of all United States twelfth-grade students who would answer the question correctly. Construct and interpret a 99 percent confidence interval for p.

I. pop – US 12th graders
para – ~~proportion~~ who answered correctly

II. SRS – given
normality normally
$n \cdot p > 10$ $.28 \cdot 9600 > 10$ $2688 > 10$ ✓
$n(1-\hat{p}) > 10$ $9600(1-.28) > 10$ $6912 > 10$ ✓
pop is 10x sample size
more than 96000 12th graders in the US.

III 1 proportion z int

$.28 \pm 2.576 \cdot \frac{\sqrt{}}{\sqrt{9600}}$ $.28 \pm 2^{**} \cdot \frac{\sqrt{}}{\sqrt{N}}$

(.2682, .2918)

IV. We are 99% ~~CONF~~ confident that the true ~~proton~~ proportion of U.S. 12th graders who answered correctly is between 26.82% & 29.18%.

GO ON TO THE NEXT PAGE.

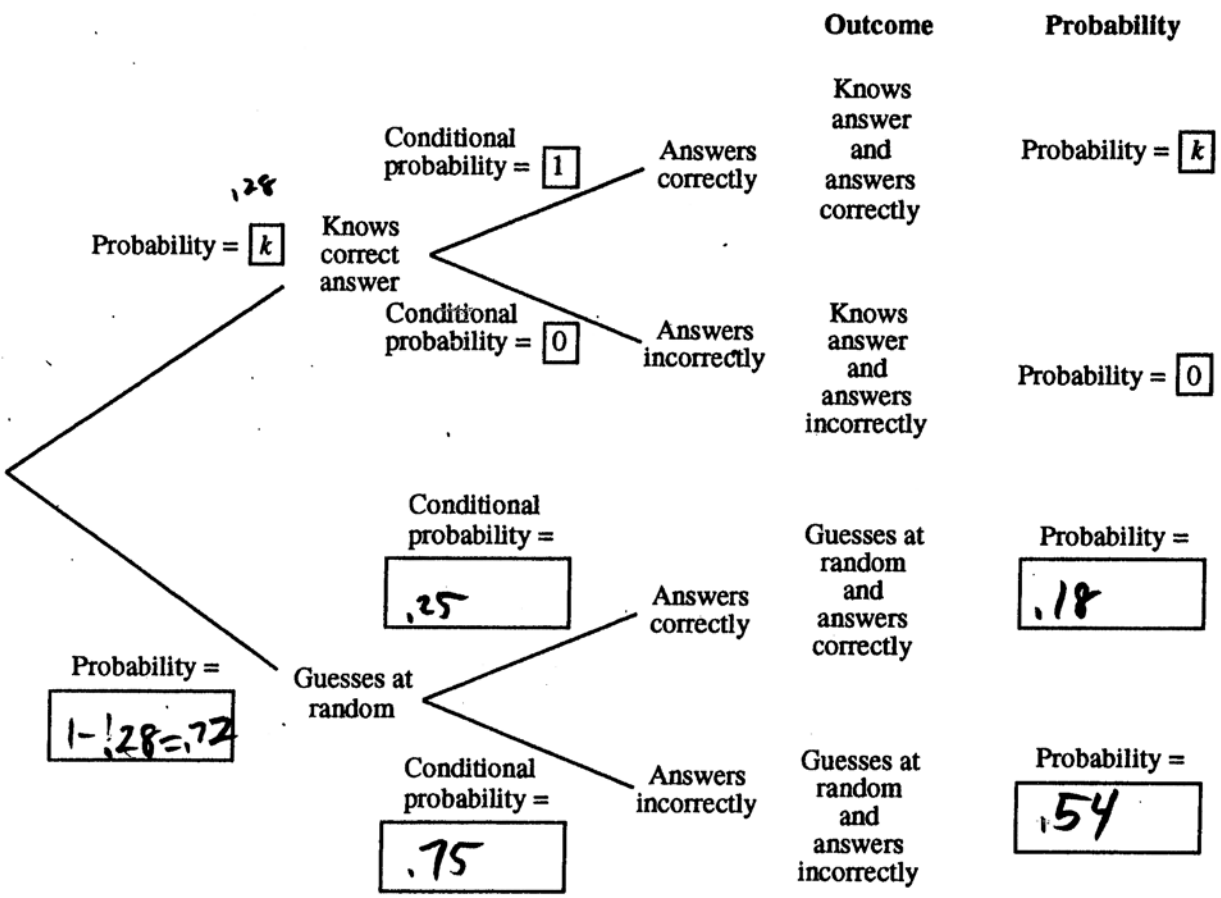
6C2

Assume that students who actually know the correct answer have a 100 percent chance of answering the question correctly, and students who do not know the correct answer to the question guess completely at random from among the four options.

Let k represent the proportion of all United States twelfth-grade students who actually know the correct answer to the question.

(b) A tree diagram of the possible outcomes for a randomly selected twelfth-grade student is provided below. Write the correct probability in each of the five empty boxes. Some of the probabilities may be expressions in terms of k.

TREE DIAGRAM OF OUTCOMES FOR A RANDOMLY SELECTED TWELFTH-GRADE STUDENT



(c) Based on the completed tree diagram, express the probability, in terms of k, that a randomly selected twelfth-grade student would correctly answer the history question.

.18 + .28 = .46

GO ON TO THE NEXT PAGE.

6C3

(d) Using your interval from part (a) and your answer to part (c), calculate and interpret a 99 percent confidence interval for k, the proportion of all United States twelfth-grade students who actually know the answer to the history question. You may assume that the conditions for inference for the confidence interval have been checked and verified.

III. $\bar{x} \pm 2^{*} \cdot \frac{\sigma}{\sqrt{N}}$

(.2682 .2418)
−.18 /−.18
(.0882, .1118)

IV. We are 99% confident that the true proportion of US 12th graders is between 8.82% & .1118

STOP

END OF EXAM

THE FOLLOWING INSTRUCTIONS APPLY TO THE COVERS OF THE SECTION II BOOKLET.

- **MAKE SURE YOU HAVE COMPLETED THE IDENTIFICATION INFORMATION AS REQUESTED ON THE FRONT AND BACK COVERS OF THE SECTION II BOOKLET.**
- **CHECK TO SEE THAT YOUR AP NUMBER LABEL APPEARS IN THE BOX(ES) ON THE COVER(S).**
- **MAKE SURE YOU HAVE USED THE SAME SET OF AP NUMBER LABELS ON ALL AP EXAMS YOU HAVE TAKEN THIS YEAR.**

AP® STATISTICS
2011 SCORING COMMENTARY

Question 6

Overview

The primary goals of this question were to assess students' ability to (1) construct and interpret a confidence interval for a population proportion; (2) create a probability tree to represent a particular random process; (3) use a probability tree to calculate a probability; and (4) integrate provided information to create a confidence interval for an atypical parameter.

Sample: 6A
Score: 4

In section 1 the student correctly checks the randomness condition ("random sample stated"), checks the large sample size condition with numerical evidence, and identifies the correct inference procedure by stating the formula for a one-sample *z*-interval for a proportion. Although not required by the scoring guidelines, the student also correctly checks that the population size is at least 10 times the sample size. Because the response includes all three required components, section 1 was scored as essentially correct. In section 2 the student correctly calculates and interprets the interval, using the correct parameter (the proportion of students who would answer the question correctly). Because the response includes both required components, section 2 was scored as essentially correct. In section 3 the student correctly completes the tree diagram in part (b) in terms of k, and in part (c) adds the appropriate values from the tree diagram. Because the response includes both required components, section 3 was scored as essentially correct. In section 4 the student equates the expression $k+\frac{1}{4}(1-k)$ from part (c) to the endpoints of the interval from part (a) and produces a reasonable interval. The student also correctly interprets the resulting interval, using the correct parameter (the proportion of students who actually know the answer to the question). Because the response includes all three required components, section 4 was scored as essentially correct. Because all four sections were scored as essentially correct, the response earned a score of 4 points.

Sample: 6B
Score: 3

In section 1 the student identifies the correct inference procedure ("1-sample -*z* proportion 99% Confidence Interval") and correctly checks the randomness condition ("SRS? yes, given"). However, the student does not adequately check the large sample size condition, stating only that 9,600 is greater than 30. Because the response includes two of the three required components, section 1 was scored as partially correct. In section 2 the student correctly calculates and interprets the interval, using the correct parameter (the proportion of students who would answer the question correctly). Because the response includes both required components, section 2 was scored as essentially correct. In section 3 the student correctly completes the tree diagram in part (b) in terms of k, and in part (c) adds the appropriate values from the tree diagram. Because the response includes both required components, section 3 was scored as essentially correct. In section 4 the student equates the expression $\hat{k}+0.25(1-\hat{k})$ from part (c) to the point estimate $\hat{p}=0.28$ from part (a), solves the resulting equation for $\hat{k}$, and constructs a new confidence interval, using the point estimate $\hat{k}=0.04$. The student also correctly interprets the resulting interval, using the correct parameter (the proportion of students who actually know the answer to the question). However, because the *endpoints* from the interval in part (a) are not used to create the new interval, the student does not complete the second required component. Because the response includes two of the three required components, section 4 was scored as partially correct. Because two sections were scored as essentially correct and two sections were scored as partially correct, the response earned a score of 3 points.

Question 6 (continued)

Sample: 6C
Score: 2

In section 1 the student correctly checks the randomness condition ("SRS–given"), checks the large sample size condition with numerical evidence, and identifies the correct inference procedure ("1 proportion z int"). Although not required by the scoring guidelines, the student also correctly checks that the population size is at least 10 times the sample size. Because the response includes all three required components, section 1 was scored as essentially correct. In section 2 the student provides the correct interval. However, because the standard error in the formula is incorrect, the response did not earn credit for the calculation component. In the interpretation, the reference to the proportion of students who answered correctly might be referring to the sample proportion, but the inclusion of the modifier "true" indicates that the student is trying to make an inference about the population proportion. Because the response includes one of the two required components, section 2 was scored as partially correct. In section 3 the student in part (b) incorrectly assumes that $k = 0.28$ and fills in the rest of the tree accordingly, so no credit was earned for the tree component. However, in part (c) the student adds the appropriate values from the tree. Because the response includes only one of the two required components, section 3 was scored as partially correct. In section 4, because the student has a numerical answer to part (c), the response cannot earn credit for equating the expression in part (c) to an estimate from part (a). However, the student does use the endpoints of the interval from part (a) to create a reasonable interval by subtracting 0.18 (the proportion of students who guess and answer correctly according to the student's tree diagram) from each endpoint, showing an understanding of the relationship between p (the proportion who answer correctly) and k (the proportion who actually know the answer). The interpretation provided is not correct, however, because the parameter is not completely identified. Because the response includes only one of the three required components, section 4 was scored as incorrect. Because one section was scored as essentially correct, two sections were scored as partially correct, and one section was scored as incorrect, the response earned a score of 2 points.