Tips from the past Chief Reader of AP Statistics

Christine Franklin

University of Georgia

chris@stat.uga.edu

Outline

- Introductory comments
- Six questions on 2011 operational exam
  - Intent
  - Question
  - Solution
  - Common errors
  - Teaching tips
- Overall student performance
- Q & A

2011 AP Statistics Exam

- Taken by about 143,500 students
- Free response scored by 620 readers in Daytona Beach last June

I'll only discuss operational questions – using these questions to offer helpful tips

#1: Intent

- Assess a student's ability to:
  - Relate summary statistics to the shape of a distribution;
  - Calculate and interpret a *z*-score;
  - Make and justify a decision that involves comparing variables that are recorded on different scales.

#1(a): Question

A professional sports team evaluates potential players for a certain position based on two main characteristics, speed and strength. Speed is measured by the time required to run a distance of 40 yards, with smaller times indicating more desirable (faster) speeds. From previous speed data for all players in this position, the times to run 40 yards have a mean of 4.60 seconds and a standard deviation of 0.15 seconds, with a minimum time of 4.40 seconds, as shown in the table below.

#1(a): Question (cont.)

	Mean	Std Dev	Minimum
Time to run 40 yards	4.60 sec	0.15 sec	4.40 sec

Based on the relationship between the mean, standard deviation, and minimum time, is it reasonable to believe that the distribution of 40-yard running times is approximately normal? Explain.

#1(a): Solution

No, it is not reasonable to believe that the distribution of running times is approximately normal, because the minimum time is only 1.33 standard deviations below the mean:

$z = (4.4-4.6)/.15 \approx -1.33$

In a normal distribution, approximately 10% of the *z*-values are below -1.33. However, there are no running times below 4.4 seconds, which has a z-score of -1.33. Therefore, the distribution of running times is not approximately normal.

7

#1(b): Question

Strength is measured by the amount of weight lifted, with more weight indicating more desirable (greater) strength. From previous strength data for all players in this position, the amount of weight lifted has a mean of 310 pounds and a SD of 25 pounds. Calculate and interpret the *z*-score for a player in this position who can lift a weight of 370 pounds.

8

#1(b): Solution

The *z*-score for a player who can lift 370 pounds is:

$z = (370 – 310) / 25 = 2.4$

This says that the player's weight lifting amount is 2.4 standard deviations above the mean for all previous players at this position.

9

#1(c): Question

The characteristics of speed and strength are considered to be of equal importance to the team in selecting a player for the position. Based on the information about the means and SDs of the speed and strength data for all players and the measurements listed in the table below for Players A and B, which player should the team select if the team can only select one of the two players? Justify your answer.

10

#1(c): Question (cont)

	Player A	Player B
Time to run 40 yards	4.42 seconds	4.57 seconds
Amount of weight lifted	370 pounds	375 pounds

11

#1(c): Solution

Because these two variables, running time and weight lifting amounts, are recorded on different scales, it is important not only to compare the players' values but also to take into account the standard deviations of the distributions of these variables. One reasonable way to do this is with *z*-scores:

12

#1(c): Solution (cont.)

The *z*-scores for the running times are:

Player A: (4.42 – 4.60) / 0.15 = -1.2

Player B: (4.57 – 4.60) / 0.15 = -0.2

The *z*-scores for the weights lifted are:

Player A: (370 – 310) / 25 = 2.4

Player B: (375 – 310) / 25 = 2.6

13

#1(c): Solution (cont.)

These *z*-scores indicate that both players are well above average in weight lifting, and both are faster than average in running time. Player A is better in running time, and player B is better in weight lifting.

But the *z*-scores indicate that the difference in their weight lifting (0.2 standard deviations) is quite small, and the difference in their running times (one standard deviation) being much larger.

Therefore, player A is a better choice, because Player A is much faster than player B and only slightly less strong.

14

#1(a): Common errors

- Explanations were generally quite weak
- Difficulty in clearly communicating why there was a problem with the minimum running time being 1.33 standard deviations below the mean

15

#1(b): Common errors

- Neglecting to provide a reasonable interpretation for the *z*-score
- Mentioning only *distance* from mean without specifying *direction* from mean

16

#1(c): Common errors

- Not adequately explaining the connection between calculations and choice of player
- Performing normal probability calculations for speed
  - Despite having argued in (a) that distribution of speed is not approximately normal
- Comparing the players on only one variable
- Not adjusting for different scales

17

#1: Teaching tips

- Emphasize, and provide practice and feedback, with communicating statistical arguments and justifications
- Make clear the distinction between *data* (such as running times) and *model* (such as normal distribution)
- Use *z*-scores for identifying relative location in a distribution
  - Not only for normal probability calculations

18

#2: Intent

- Assess a student's ability to:
  - Determine a conditional probability from a table of data;
  - Use a table of data to determine whether or not two events are independent;
  - Demonstrate an understanding of the concept of independence by constructing a graph that displays independence between two variables.

#2(a): Question

The table below shows political party registration by gender for all 500 registered voters in Franklin Township:

	Party W	Party X	Party Y	Total
Female	60	120	120	300
Male	28	124	48	200
Total	88	244	168	500

a) Given that a randomly selected voter is a male, what is the probability that he is registered for Party Y?

#2(a): Solution

Of the 200 male registered voters in Franklin Township, 48 are registered for Party Y. Therefore the conditional probability that a randomly selected voter is registered for Party Y, given that the voter is male, is

48 / 200 = 0.24

#2(b): Question

Among the registered voters of Franklin Township, are the events "is a male" and "is registered for Party Y" independent? Justify your answer based on probabilities calculated from the table above.

#2(b): Solution

No, the events "is a male" and "is registered for Party Y" are not independent.

One justification is to note that

P(Party Y | Male) = 0.24 and

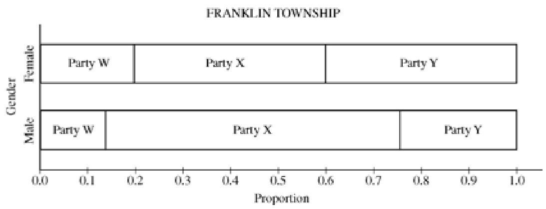
P(Party Y) = 168 / 500 = 0.336.

Because 0.336 ≠ 0.24, the two events are not independent.

#2(c): Question

One way to display the data in the table is to use a segmented bar graph. The following segmented bar graph, constructed from the data in the party registration–Franklin Township table, shows party-registration distributions for males and females in Franklin Township.

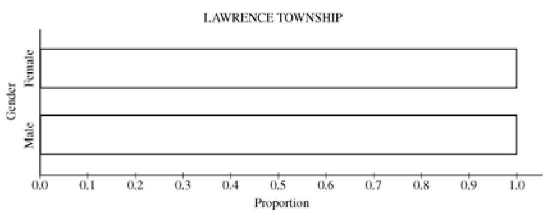
#2(c): Question (cont.)



#2(c): Question (cont.)

In Lawrence Township, the proportions of all registered voters for Parties W, X, and Y are the same as for Franklin Township, and party registration is independent of gender. Complete the graph below to show the distributions of party registration by gender in Lawrence Township.

#2(c): Question (cont.)



#2(c): Solution

The marginal proportions registered for the three political parties (without regard to gender) are:

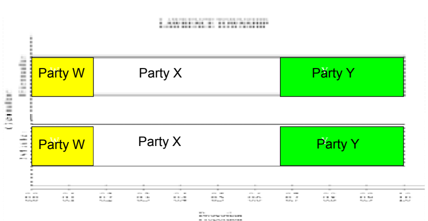
Party W: 88 / 500 = 0.176

Party X: 244 / 500 = 0.488

Party Y: 168 / 500 = 0.336

Because party registration is independent of gender in Lawrence Township, the proportions of males and females registered for each party must be identical to each other and also identical to the marginal proportion of voters registered for that party. Thus, the graph for Lawrence Township must be:

#2(c): Solution (cont.)



#2(a): Common errors

- There were few common errors on part (a). Nearly every student who attempted part (a) got it essentially correct.
- Most common incorrect answer: computing the reverse conditional probability.
  - *i.e.*, P(male | Party Y) instead of P(Party Y | male).

#2(b): Common errors

- Happy surprises:
  - Students showed little confusion between independent events and disjoint events.
  - Few chi-square tests were performed.

#2(b): Common errors (cont.)

- Most students knew an appropriate way to determine whether two events are independent using a table of data.
- Of those who didn't, the most common mistakes were these:
  - Checking for P(M|Y) = P(Y|M)
  - Checking for P(Y) = P(M)

#2(b): Common errors (cont.)

- We frequently saw very weak communication, such as:
  - Symbols introduced without identifying what they meant in this context.
  - Equations posited without any explanation of what they were for—and a general aversion to sentences with a subject and a predicate.

#2(b): Common errors (cont.)

- We frequently saw very weak communication, such as:
  - Sloppy use of mathematical notation, which sometimes made it difficult or impossible to tell what the student meant.
  - Interpretations of probabilities that were so muddled that it was difficult or impossible to tell what the student meant.

#2(c): Common errors

- Most common two responses:
  - A correct graph.
  - A copy of the graph given in the problem for the other township.

#2: Teaching tips

- Give students more practice articulating statistical arguments and interpretations of probabilities.
- Ensure that we are teaching students concepts and not just skills.

#3: Intent

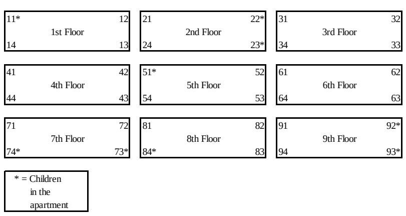
- Assess a student's ability to:
  - Understand and describe a process for implementing cluster sampling;
  - Describe a statistical advantage of stratified sampling over cluster sampling in a particular situation.

#3: Question

An apartment building has nine floors and each floor has four apartments. The building owner wants to install new carpeting in eight apartments to see how well it wears before she decides whether to replace the carpet in the entire building.

The figure below shows the floors of apartments in the building with their apartment numbers.

#3: Question (cont.)



Only the nine apartments indicated with an asterisk (*) have children in the apartment.

#3(a): Question

For convenience, the apartment building owner wants to use a cluster sampling method, in which the floors are clusters, to select the eight apartments. Describe a process for randomly selecting eight different apartments using this method.

#3(a): Solution

Generate a random integer between 1 and 9, inclusive, using a calculator or computer software or a table of random digits.

Select all four apartments on the floor corresponding to this random integer.

Then select another random integer between 1 and 9, inclusive; if the same integer appears again, continue until a different random integer appears.

#3(a): Solution (cont.)

Again select all four apartments on the floor corresponding to this second random integer.

The cluster sample consists of the eight apartments on these two randomly selected floors.

#3(b): Question

An alternative sampling method would be to select a stratified random sample of eight apartments, where the strata are apartments with children and apartments with no children. A stratified random sample of size eight might include two randomly selected apartments with children and six randomly selected apartments with no children.

#3(b): Question (cont.)

In the context of this situation, give one statistical advantage of selecting such a stratified sample as opposed to a cluster sample of eight apartments using the floors as clusters.

#3(b): Solution

Because the amount of wear on the carpets in apartments with children could be different from the wear on the carpets in apartments without children, it would be advantageous to have apartments with children represented in the sample.

The cluster sampling procedure in part (a) could produce a sample with no children in the selected apartments; for example, a cluster sample of the apartments on the 3^{rd} and 6^{th} floors would consist entirely of apartments with no children.

#3(b): Solution (cont.)

Stratified random sampling, where the two strata are apartments with children and apartments without children, *guarantees* a sample that includes apartments with and without children and would thus yield sample data that are representative of both types of apartment.

#3(a): Common errors

- Many students did not know what a cluster sample is. Their responses were often an attempt to do something other than a simple random sample.
- Many students selected two floors at random but failed to state that all four apartments from each floor form the sample.

#3(a): Common errors (cont.)

- Students' descriptions of the random selection process were often so sparsely (or poorly) written that it was not clear how the sampling process was to be implemented.
- More generally, many students showed very weak written communication skills.

#3(b): Common errors

- Many students do not know what a "statistical advantage" is.
- Students often failed to link children with carpet wear.
- Students often focused on what is wrong with the cluster—potential lack of apartments with children—without saying anything about the stratified sample and why it would be better.

#3(b): Common errors (cont.)

- Students often parroted the stem—*i.e.* simply stated that there would be 2 apartments with children and 6 apartments without children and doing little or nothing else.
- More generally, many students showed very weak written communication skills.

#3: Teaching tips

- Show students examples of cluster sampling methods side-by-side with stratified sampling methods and simple random sampling methods. Contrast them, and be sure students understand (and can articulate!) advantages and drawbacks of each.

#3: Teaching tips (continued)

- A large number of our students need much more practice articulating statistical ideas. On a regular basis, have students explain statistical principles, such as advantages and disadvantages of cluster sampling compared to stratified sampling. Give them quick feedback.

#4: Intent

- Assess a student's ability to conduct a test of significance:
  - State hypotheses;
  - Select an appropriate test and check conditions for its use;
  - Compute a test statistic and *p*-value;
  - Use the *p*-value to justify a conclusion written in context.

#4: Question

High cholesterol levels in people can be reduced by exercise, diet, and medication. Twenty middle-aged males with cholesterol readings between 220 and 240 milligrams per deciliter (mg/dL) of blood were randomly selected from the population of such male patients at a large local hospital.

#4: Question (cont.)

Ten of the 20 males were randomly assigned to group A, advised on appropriate exercise and diet, and also received a placebo. The other 10 males were assigned to group B, received the same advice on appropriate exercise and diet, but received a drug intended to reduce cholesterol instead of a placebo.

#4: Question (cont.)

After three months, posttreatment cholesterol readings were taken for all 20 males and compared to pretreatment cholesterol readings. The tables below give the reduction in cholesterol level (pretreatment reading minus posttreatment reading) for each male in the study.

#4: Question (cont.)

Group A (placebo)

2	19	8	4	12	8	17	7	24	1

Mean Reduction: 10.20 SD of Reductions: 7.66

Group B (cholesterol drug)

30	19	18	17	20	-4	23	10	9	22

Mean Reduction: 16.40 SD of Reductions: 9.40

#4: Question (cont.)

Do the data provide convincing evidence, at the $\alpha = 0.01$ level, that the cholesterol drug is effective in producing a reduction in mean cholesterol level beyond that produced by exercise and diet?

#4: Solution – hypotheses

H_0: $\mu_A = \mu_B$ H_A: $\mu_A < \mu_B$

where μ_A represents the mean cholesterol reduction if all such male patients at this hospital get exercise/diet advice and placebo,

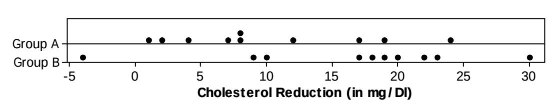
and μ_B represents the mean cholesterol reduction if all such male patients at this hospital get exercise/diet advice and drug.

#4: Solution – choice, conditions

Two-sample *t*-test

Randomness condition: satisfied because subjects were randomly *assigned* to treatment

Normality: satisfied because of no severe skewness or outliers in graphs of sample cholesterol reductions



#4 Solution – calculations

The test statistic is:

$$t = \frac{\overline{x}_A - \overline{x}_B}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}} = \frac{10.20 - 16.40}{\sqrt{\frac{7.66^2}{10} + \frac{9.40^2}{10}}} \approx -1.62$$

The *p*-value is 0.062

#4 Solution – conclusion

Because the p-value of .062 is not less than the significance level (.01), we do not have sufficient evidence to conclude that the drug significantly increases cholesterol reduction as compared to placebo.

#4 Common errors

- Not including all four components of a significance test
- Expressing hypotheses in terms of sample statistics
- Commenting on random *sampling* rather than random *assignment*
- Not providing both graphs and commentary to check normality condition

#4 Common errors (cont.)

- Applying a *z*-test rather than a *t*-test
- Using a paired *t*-test
- Providing correct calculations from calculator but also providing incorrect calculations by hand
- Not justifying conclusion explicitly based on magnitude of *p*-value
- Expressing conclusion as equivalent to "accepting null hypothesis"

#4 Teaching tips

- Emphasize distinction between random *sampling* and random *assignment*
  - Different types of randomness, purposes, scopes of conclusion, inference conditions to check
  - Random sampling enables generalizing sample results to larger population
  - Random assignment allows for drawing cause/effect conclusions

#4 Teaching tips (cont.)

- Provide frequent reminders of basic ideas of significance tests
  - Hypotheses are about population parameters
  - Inference conditions must be checked
  - Conclusion must be justified from p-value
  - Conclusion must be expressed in context
  - Data never support null hypothesis

#5: Intent

- Assess a student's ability to:
  - Determine the equation of a least squares line from computer output;
  - Use the slope of a least squares line to compare expected values of the response for different values of the explanatory variable;
  - Recognize how to determine the proportion of variability in the response explained by a least squares line;
  - Use output to determine whether the relationship between two quantitative variables is statistically significant.

#5: Question

Windmills generate electricity by transferring energy from wind to a turbine. A study was conducted to examine the relationship between wind velocity in miles per hour (mph) and electricity production in amperes for one particular windmill. For the windmill, measurements were taken on twenty-five randomly selected days, and the computer output for the regression analysis for predicting electricity production based on wind velocity is given below.

#5: Question (cont.)

The regression model assumptions were checked and determined to be reasonable over the interval of wind speeds represented in the data, which were from 10 miles per hour to 40 miles per hour.

Predictor	Coef	SE Coef	T	P
Constant	0.137	0.126	1.09	0.289
Wind Velocity	0.240	0.019	12.63	0.000

S = 0.237 R-Sq = 0.873 R-Sq (adj) = 0.868

#5(a): Question

Use the computer output above to determine the equation of the least squares regression line. Identify all variables used in the equation.

#5(a): Solution

The equation of the least squares line is:

predicted electricity production =

0.137 + 0.240 x wind velocity

or

$$\widehat{\text{electrical production}} = 0.137 + 0.240(\text{wind velocity})$$

#5(b): Question

How much more electricity would the windmill be expected to produce on a day when the wind velocity is 25 mph than on a day when the wind velocity is 15 mph? Show how you arrived at your answer.

#5(b): Solution

The slope coefficient of 0.240 indicates that for each additional mile per hour of wind speed, the expected electricity production increases by 0.240 amperes. Thus, the expected electricity production is $10 \times 0.240 = 2.40$ amperes higher on a day with 25 mph wind velocity as compared to a day with 15 mph wind velocity.

#5(b): Solution (alternative)

When the wind velocity is 25 mph, the windmill is expected to produce 0.24(25) + 0.137 = 6.137 amps of electricity.

When the wind velocity is 15 mph, the windmill is expected to produce 0.24(15) + 0.137 = 3.737 amps of electricity.

The difference between these two predictions is 6.137 – 3.737 = 2.4 amps.

#5(c): Question

What proportion of the variation in electricity production is explained by its linear relationship with wind velocity?

#5(c): Solution

The proportion of variation in electricity production that is explained by the linear relationship with wind speed is r^2, which the regression output reports to be 0.873, or 87.3%.

#5(d): Question

Is there statistically convincing evidence that electricity production by the windmill is related to wind velocity? Explain.

#5(d): Solution

Yes, there is very strong statistical evidence that the population slope differs from zero, so electricity production is related to wind speed. For testing the hypotheses H_0: $\beta = 0$ vs. H_a: $\beta \neq 0$, the output reveals that the test statistic is $t = 12.63$ and the p-value (to three decimal places) is 0.000. Because this p-value is so small (much less than .05 and even .01, for example), the sample data provide very strong statistical evidence that electricity production is related to wind speed.

#5(a): Common errors

- Missing "predicted" or "expected" in the definition of $\hat{y}$.
- Putting "^" on both variables.
- Defining coefficients rather than variables.
- Interchanging variables.
- Unable to read the computer output.

CollegeBoard

79

#5(b): Common errors

- Missing units, or giving incorrect units.

CollegeBoard

80

#5(c): Common errors

- Some students were unable to read the computer output.
- Some students reported r^2 adjusted or computed r.

CollegeBoard

81

#5(d): Common errors

- Some students didn't explain that the small p-value indicates statistically significant evidence of a relationship between the variables.
- Some students based their answer on r or on r^2 alone.

CollegeBoard

82

#5: Teaching tips

- If possible, have students perform linear regression using a standard statistical package (some are inexpensive or free), and teach them to read the output.
- Use examples to illustrate why r^2 may be close to 1 even while a graph illustrates little relationship between the variables. (For example, this can occur when there is a single extreme outlier in both variables.)

CollegeBoard

83

#6: Intent

- Assess a student's ability to:
  - Construct and interpret a confidence interval for a population proportion;
  - Create a probability tree to represent a particular random process;
  - Use a probability tree to calculate a probability;
  - Integrate provided information to create a confidence interval for an atypical parameter.

CollegeBoard

84

#6: Question

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.)

#6: Question (cont.)

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.

#6(a): Question

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.

#6(a): Solution

The appropriate procedure is a one-sample z interval for a population proportion p, where p is the proportion of all U.S. 12^{th} grade students who *would* answer the question correctly.

The conditions for this inference procedure are satisfied because:

1) This question states that the students are a random sample from the population, and

2) $n \times \hat{p} = 9600 \times 0.28 = 2688$ and
$n \times (1-\hat{p}) = 9600 \times 0.72 = 6912$

are much larger than 10.

#6(a): Solution (cont.)

A 99% confidence interval for the population proportion p is given by:

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \rightarrow 0.28 \pm 2.576\sqrt{\frac{0.28(0.72)}{9600}}$$

$$\rightarrow 0.28 \pm 0.012 \rightarrow (0.268, 0.292)$$

We are 99% confident that the interval from 0.268 to 0.292 contains the population proportion of all U.S. 12^{th} graders who *would* answer this question correctly.

#6: Question (cont.)

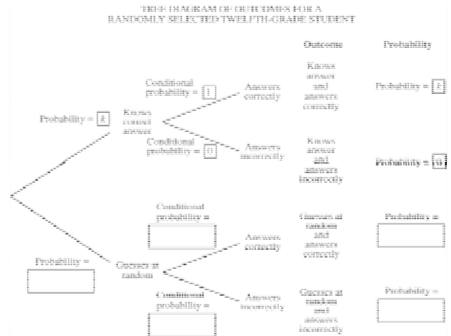
Assume that students who actually know the correct answer have a 100% chance of answering the question correctly and students who do not know the answer to the question guess completely at random from among the four options.

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

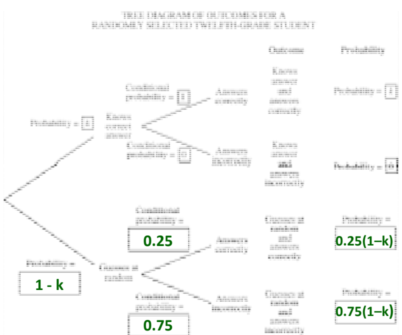
#6(b): Question

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*.

#6(b): Question (cont.)



#6(b): Solution



#6(c): Question

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.

#6(c): Solution

$$
\begin{aligned}
&P(\text{answers correctly}) \\
&= P(\text{knows correct answer and answers correctly}) \\
&+ P(\text{guesses at random and answers correctly}) \\
&= k + (1-k) \times 0.25
\end{aligned}
$$

which simplifies to:

#6(d): Question

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.

#6(d): Solution

From part (c), the probability that a randomly selected student correctly answers the question is 0.25 + 0.75 *k*.

From part (a), we are 99% 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 answer from part (c) to the endpoints of the interval from part (a):

#6(d): Solution (cont.)

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

We are 99% confident that the interval from 0.024 to 0.056 contains the proportion of all U.S. 12th grade students who actually *know* the correct answer.

#6(a): Common errors

- Omitting name of procedure or giving ambiguous name such as "*z*-interval"
- Not checking conditions for inference
- Incorrectly checking sample size condition
  - E.g. $n \geq 30$ or "CLT satisfied"
- Interpreting parameter as proportion who answer*ed* correctly (statistic)
- Interpreting confidence *level*, not interval
- Not recognizing impractical values in CI

#6(b),(c): Common errors

- Confusing sets with probabilities
  - E.g., "not *k*" or *k*^*c*" rather than "1 – *k*"
- Plugging in a particular value of *k*
  - Such as *k* = 0.28
- Multiplying, rather than adding, relevant probabilities
- Using poor notation
  - E.g., P(0.25+0.75*k*)

#6(d): Common errors

- Not using interval from part (a)
  - E.g., First determine point estimate of *k*, then use conventional CI procedure
- Confusing what *p* and *k* mean in interpreting confidence interval
- Making same errors of interpretation as in (a)
  - E.g., confusing parameter and statistic

#6: Teaching tips

- Recognize the difficulty of identifying parameters, provide practice and feedback
- Provide guidance and practice for identifying which inference procedure to use when
- Emphasize that confidence interval procedures have conditions to be checked
  - Just as test of significance procedures do
- Develop habit of checking reasonable-ness of calculations

#6: Teaching tips (cont.)

- Use good notation with probability calculations
  - Recognize difference between set operations, probability calculations
- Emphasize that confidence intervals must be accompanied by interpretations

Overall student performance

	#1	#2	#3	#4	#5	#6
Mean	1.33	1.47	1.41	1.53	1.59	1.31
SD	1.13	1.21	1.16	1.26	1.25	1.18

Score	1	2	3	4	5
%	23.8%	17.7%	25.1%	21.3%	12.1%

Final Tips

- Students need more and more practice with interpreting simulation results with regard to the issue of assessing statistical significance
- Students continue to struggle with sampling distribution questions
- Confidence intervals – students forget to check assumptions; students mix the interpretation of the interval with the interpretation of the confidence level
- No bald answers – even if the answer is obvious, student should show some work for how the answer was obtained. We especially see this with probabilities found from a contingency table