

2018

AP[®]

CollegeBoard

AP Statistics

Scoring Guidelines

2018 SCORING GUIDELINES**Question 1****Intent of Question**

The primary goals of this question were to assess a student's ability to (1) identify various values in regression computer output; (2) interpret the intercept of a regression line in context; (3) interpret the coefficient of determination (r^2) in context; and (4) identify an outlier from a scatterplot.

Solution**Part (a):**

The estimate of the intercept is 72.95. It is estimated that the average time to finish checkout if there are no other customers in line is 72.95 seconds.

Part (b):

The coefficient of determination is $r^2 = 73.33\%$. This value indicates that 73.33% of the variability in the times it takes customers to finish checkout, including time waiting in line, can be explained by knowing how many customers are in line in front of the selected customer.

Part (c):

The outlier is the point with $x = 3$ and y close to 0. This point is considered an outlier because the combination of x and y values differs from the pattern of the rest of the data. Specifically, the value of y (time to finish checkout) is much lower than would be expected when there are $x = 3$ customers in line in front of the selected customer, given the remaining data.

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 three components:

1. Correctly identifies 72.95 as the intercept.
2. Communicates the concept of a y -intercept in a context that includes both time and zero customers.
3. Indicates that the value of the intercept is a prediction by using language such as “predicted,” “estimated,” or “average” value of y .

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

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

2018 SCORING GUIDELINES

Question 1 (continued)

Notes:

- Regression equations (such as $\hat{y} = 72.95 + 174.40x$) cannot be used to satisfy identification of the intercept in component 1, unless the intercept is explicitly labeled as such.
- A regression equation cannot be used to satisfy component 3.
- Incorrect regression equations are treated as extraneous and do not affect the scoring of any component.
- A response that interprets 72.95 as a slope does not satisfy components 1 or 2.

Part (b) is scored as follows:

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

1. Correctly identifies 73.33% as the coefficient of determination.
2. Provides a correct (possibly generic) interpretation of r^2 .
3. Interpretation includes context.

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

OR

if the response satisfies the three components, but reverses the roles of number of customers in line and time to finish checkout in the interpretation.

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

Notes:

- In component 2 the correct interpretation of the coefficient of determination can take any of several equivalent forms, such as:
 - The percent variability in y that is attributed to the linear relationship between y and x or between x and y .
 - The proportion of the total variability in the dependent variable y that is explained by the independent variable x .
 - The proportion of variation in y that is accounted for by the linear model.
 - The proportionate reduction of total variation of the y values that is associated with the use of the independent variable x .
 - The proportionate reduction in the sum of the squares of vertical deviations obtained by using the least-squares line instead of the naïve prediction of $\bar{y}$.
- In component 2 common *incorrect* interpretations of the coefficient of determination include:
 - The percent variability in the *predicted* y values that is explained by the linear relationship between y and x .
 - The percent variability in the *data* that is explained by the linear relationship between y and x .
 - The percent variability that is explained by the linear relationship between y and x .
 - The percent variability in y that is *on average* explained by the linear relationship between y and x .
- For component 3 context must include mention of time or customers.

2018 SCORING GUIDELINES**Question 1 (continued)**

Part (c) is scored as follows:

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

1. Correctly identifies the outlier.
2. Describes an unusual feature of the identified scatter plot point, relative to the remaining data points, that is sufficient to identify it as the outlier. Examples include:
 - The combination of x and y values is unusual compared to the other points.
 - The value of y is much lower than would be expected (or predicted), given the remaining data.
 - The residual for the point is unusually large relative to the other residuals.

Partially correct (P) if the response satisfies component 1 but does not satisfy component 2.

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

Notes:

- In the absence of any point being circled on the graph, component 1 can still be satisfied by explicitly referring to the coordinates of the outlier. Valid coordinates for outlier identification must specify an x value of 3 and a y value that is strictly between 0 and 250.
- A response that does not make a comparison to the remaining data points, such as stating the outlier has a large residual or is nowhere near the regression line, does not satisfy component 2.
- A response that makes a comparison to the remaining data points based upon an unusual feature that is *insufficient* for outlier identification, such as stating the point is the only point with that particular y value, does not satisfy component 2.
- In the absence of explicit numerical calculation, a response that appeals to the influence that the outlier has on the regression coefficient estimates or on the sample correlation coefficient does not satisfy component 2.

2018 SCORING GUIDELINES**Question 1 (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 two parts partially correct

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

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 “ ≈ 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.

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.

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 “ ≈ 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.

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

2018 SCORING GUIDELINES

Question 3

Intent of Question

The primary goals of this question were to assess a student's ability to (1) compute a probability based on a weighted mixture of two populations; (2) compute a conditional probability; and (3) recognize a binomial random variable and compute the probability associated with it.

Solution**Part (a):**

Let L denote left-handed, M denote multiple birth, and S denote single birth.
The probability that a randomly selected child born in the region is left-handed is:

$$P(L) = P(M)P(L | M) + P(S)P(L | S) = (0.035)(0.22) + (0.965)(0.11) = 0.0077 + 0.10615 = 0.11385.$$

Part (b):

From part (a), $P(L) = 0.11385$. Therefore,

$$P(M | L) = \frac{P(L \text{ and } M)}{P(L)} = \frac{(0.035)(0.22)}{0.11385} = \frac{0.0077}{0.11385} \approx 0.0676.$$

Part (c):

Let X represent the number of children who are left-handed in a random sample of 20 children from the region. X has a binomial distribution with $n = 20$ and $p = 0.11385$ (found in part (a)). Using the binomial distribution,

$$\begin{aligned} P(X \geq 3) &= 1 - P(X \leq 2) \\ &= 1 - \binom{20}{0}(0.11385)^0(0.88615)^{20} - \binom{20}{1}(0.11385)^1(0.88615)^{19} - \binom{20}{2}(0.11385)^2(0.88615)^{18} \\ &\approx 1 - 0.598 \approx 0.402. \end{aligned}$$

2018 SCORING GUIDELINES**Question 3 (continued)****Scoring**

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

Part (a) is scored as follows:

Essentially correct (E) if the probability is computed correctly, *AND* work is shown that includes correct numerical values using a formula, end results from a tree diagram, or some other appropriate strategy.

Partially correct (P) if the response provides a reasonable strategy for finding the probability, such as a formula or tree diagram, but uses one or more inappropriate values;

OR

if the response gives the correct probability but not enough work is shown to determine how that probability was found.

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

Note: A reasonable strategy needs to include summing the results of two multiplications.

Part (b) is scored as follows:

Essentially correct (E) if the probability is computed correctly, with work shown that includes appropriate numerical values for both the numerator and denominator.

Partially correct (P) if the response includes a numerator and denominator in calculating the conditional probability, with one appropriate term (numerator or denominator) and the other inappropriate.

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

Note: Appropriate values include incorrectly calculated values from part (a).

2018 SCORING GUIDELINES

Question 3 (continued)

Part (c) is scored as follows:

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

1. Uses a calculation based on the binomial distribution to find the probability of the number of children in the sample who are left-handed.
2. Specifies appropriate values for n and p .
3. Uses correct endpoint value for the probability.
4. Uses correct direction to calculate the probability of at least three left-handed children.
5. Correctly calculates a binomial probability consistent with the previous work.

Partially correct (P) if the response satisfies component 1 and only two or three of the other four components;
OR

if components 2, 3, 4, and 5 are met, and the response does not explicitly indicate the binomial distribution is used by name or formula.

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

Notes:

- “Appropriate” values include incorrectly calculated values from part (a) or a recalculated probability from part (b).
- An unlabeled numerical value in a calculator statement cannot be used to satisfy a component.
- A response which calculates $P(X \leq 3)$ satisfies component 3 but does not satisfy component 4.
- A normal approximation to the binomial is not appropriate because $np = 20 \times 0.11385 = 2.277 < 5$.
A response using the normal approximation can score at most P. To earn a score of P, the response must include all of the following:
 - a correct mean and standard deviation based on the binomial parameters
 - clear indication of boundary and direction with a z-score or diagram
 - the probability computed correctly

Notes for all parts:

- If the resulting probability or part of the calculation of the probability uses a value that is not between 0 and 1, inclusive, the score for that part is lowered by one level (that is, from E to P, or from P to I).
- An arithmetic or transcription error in a response can be ignored if correct work is shown. For example, $0.0077 + 0.10615 = 0.1385$.

2018 SCORING GUIDELINES**Question 3 (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 two parts partially correct

2018 SCORING GUIDELINES**Question 4****Intent of Question**

The primary goals of this question were to assess a student's ability to (1) determine whether a cause-and-effect conclusion can be made based on how a study was conducted and (2) set up, perform, and interpret the results of a hypothesis test, in the context of the problem.

Solution**Part (a):**

Yes, it would be reasonable to conclude that the new procedure causes a reduction in recovery time, for patients similar to those in the study. The patients in the study were randomly assigned to the two procedures, which reduces the chance that confounding variables will affect the results. Therefore the statistically significant reduction in mean recovery time can be attributed to the new procedure being superior to the standard procedure.

Part (b):

Step 1: State a correct pair of hypotheses.

Let μ_S represent the mean recovery time among all patients similar to those in the study if they were to receive the standard treatment.

Let μ_N represent the mean recovery time among all patients similar to those in the study if they were to receive the new treatment.

The hypotheses to be tested are $H_0 : \mu_S = \mu_N$ versus $H_a : \mu_S > \mu_N$.

Step 2: Identify a correct test procedure (by name or by formula) and check appropriate conditions.

The appropriate procedure is a two-sample t -test for a difference between means.

Because this is an experiment, the first condition is that subjects were randomly assigned to one treatment group or the other. In this case the condition is satisfied because we were told that the subjects were randomly assigned to either the standard or new procedure.

The second condition is that the recovery times of the two populations are normally distributed or the sample sizes are sufficiently large to presume that the distribution of the difference in the sample means is approximately normal. In this case the condition is met because the sample sizes of 110 and 100 are both sufficiently large.

2018 SCORING GUIDELINES

Question 4 (continued)

Step 3: Correct mechanics, including the value of the test statistic, degrees of freedom, and p -value (or rejection region).

$$\text{The test statistic is } t = \frac{\bar{x}_S - \bar{x}_N}{\sqrt{\frac{s_S^2}{n_S} + \frac{s_N^2}{n_N}}} = \frac{217 - 186}{\sqrt{\frac{34^2}{110} + \frac{29^2}{100}}} \approx 7.13.$$

The p -value is the area greater than 7.13 for a t -distribution with $df = 207.18$, which is essentially 0 (8.36×10^{-12}).

Step 4: State a correct conclusion in the context of the problem, using the result of the statistical test.

Because the p -value is very small, we have sufficient evidence to conclude that for patients similar to the ones in the study, those receiving the new procedure would have less recovery time, on average, than those receiving the standard procedure.

Scoring

This question is scored in three sections. Section 1 consists of part (a); section 2 consists of step 1, step 2, and the test statistic in step 3 in part (b); and section 3 consists of the p -value in step 3 and step 4 in part (b). Sections 1, 2, and 3 are each scored essentially correct (E), partially correct (P), or incorrect (I).

Section 1 is scored as follows:

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

1. Correctly states that it is reasonable to make a causal conclusion.
2. Justifies the causal conclusion based on random assignment of patients to procedures (or procedures to patients);
OR
justifies the causal conclusion by stating that a randomized experiment was conducted.
3. Includes the context of the situation.

Partially correct (P) if the response satisfies component 1 *AND* provides WEAK justification of the causal conclusion by stating that there was random assignment or a randomized experiment was conducted, but with no context;

OR

by stating that an experiment was conducted or there was assignment (without the word “randomized”) *AND* the response includes context of the situation;

OR

by stating that the study design reduces the chance of confounding variables or balances the effects of uncontrolled variables across both groups in context without explicitly referring to the random assignment.

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

2018 SCORING GUIDELINES

Question 4 (continued)

Notes:

- If the response states that it is *not* reasonable to make a causal conclusion because the result could have been due to random chance *AND* explains that there is evidence for a causal conclusion based on random assignment of patients to procedures or by stating that a randomized experiment was conducted, then the response is scored E.
- If the response discusses aspects of an experiment other than random assignment (such as, control, replication, or large samples), then those aspects are considered extraneous and the response can be scored E unless those aspects are incorrect for this study (such as, blocking is a requirement, or the study used blocking, or the study used a placebo) in which case the score should be lowered one level (that is, from E to P, or from P to I).
- If the response correctly states in context that it is reasonable to make a causal conclusion but includes incorrect or contradictory justification (such as, random selection of patients), then the response is scored I.

Section 2 is scored as follows:

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

1. Parameters are defined correctly.
2. Hypotheses imply equality in the null and correct direction in the alternative.
3. Correct test is identified by name or formula.
4. Correct test statistic for a difference in means is calculated.

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

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

Notes:

- If standard symbols are used for the parameters with appropriate group labels (such as, μ_S, μ_N), component 1 is satisfied.
- If the correct test is identified, but the response states an incorrect formula or uses incorrect notation in the formula, component 3 is not satisfied.
- A pooled two-sample t -test is acceptable for component 3, but the student must also state and comment on the plausibility of the equal population variances assumption.
- If the response identifies a z -test for equal means as the correct test identification, component 3 is not satisfied but component 4 could be satisfied.

Confidence Interval approach:

- If a single two-sample t -interval for the difference in means is used, components 3 and 4 can be satisfied. Component 3 is satisfied if the t -interval is correctly identified by name or formula. Component 4 is satisfied if the correct interval is calculated. If an alpha level is stated, then an appropriate adjustment to the confidence level must be made because the appropriate test is one-sided.
- If two one-sample t -intervals are used, while not a recommended approach, component 3 is not satisfied but component 4 could be satisfied. Component 4 is satisfied if both intervals are calculated correctly.

2018 SCORING GUIDELINES

Question 4 (continued)

Section 3 is scored as follows:

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

1. Makes reference to an approximately correct p -value that is consistent with the test statistic and alternative hypothesis for a difference in means.
2. Correctly justifies the conclusion based on the size of the p -value or the test statistic.
3. Correctly states the conclusion in context.

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

Incorrect (I) if the response does not meet the criteria for E or P or includes a justification not based on the inferential results.

Notes:

Component 1:

- Is satisfied if the response makes reference to a large test statistic without referring to a p -value.

Component 2:

- No alpha level is needed to provide justification of the conclusion based on the size of the p -value.
- Is satisfied if the response states the p -value without reference to size, but it is contiguous to the conclusion and clearly indicates a continuous train of thought.
- A correct interpretation of the p -value with a complete explanation that obtaining a test statistic at least this extreme is unlikely due to chance alone is considered justification based on the size of the p -value.
- If an incorrect interpretation of the p -value is given, the score is lowered one level (that is, from E to P, or from P to I).
- A decision about the null hypothesis (reject H_0 or fail to reject H_0) is not required, but if an incorrect decision is stated based on the given p -value then component 2 is not satisfied.
- If a rejection region approach is used, a reasonable critical value replaces the p -value.

Component 3:

- A correct conclusion must be related to the alternative hypothesis in order to satisfy component 3.
- The following responses do not satisfy component 3:
 - States or implies that the null hypothesis is *accepted*
 - States or implies that the alternative hypothesis has been *proven*
 - States the conclusion in past tense (unless the response did not satisfy a component of section 2 for the use of past tense)

2018 SCORING GUIDELINES

Question 4 (continued)

Confidence Interval approach:

- If a single two-sample t -interval for the difference in means is used:
 - Component 1 is satisfied if the response indicates that zero is either included or not included in the calculated interval.
 - Component 2 is satisfied if the response indicates that the bounds are either both above or both below zero (consistent with alternative hypothesis) and uses that as justification for the conclusion.
 - Component 3 is satisfied if the conclusion is stated in context.
- If two one-sample t -intervals are used (which is not recommended) the response is scored at most P if all three components are satisfied, otherwise scored I:
 - Component 1 is satisfied if the response states that the intervals do not overlap.
 - Component 2 is satisfied if the conclusion indicates that the confidence interval for the new procedure lies below the confidence interval for the standard procedure.
 - Component 3 is satisfied if the conclusion is stated in context.

Note: If the three sections of the response are scored as E, to earn a score of 4 as a complete response, both conditions in step 2 must be correctly stated and justified. Additional condition(s) inappropriate for a two-sample t -test must not be stated. Otherwise, the response earns a score of 3 a substantial response.

2018 SCORING GUIDELINES

Question 4 (continued)

4 Complete Response

Three sections essentially correct with conditions for inference

3 Substantial Response

Three sections essentially correct without conditions for inference

OR

Two sections essentially correct and one section partially correct

2 Developing Response

Two sections essentially correct and no sections partially correct

OR

One section essentially correct and one or two sections partially correct

OR

Three sections partially correct

1 Minimal Response

One section essentially correct

OR

No sections essentially correct and one or two sections partially correct

2018 SCORING GUIDELINES

Question 5

Intent of Question

The primary goals of this question were to assess a student's ability to (1) determine which of two histograms represents data with a larger median; (2) calculate the mean of a combined data set when the separate means and sample sizes are known; and (3) calculate the probability that an individual randomly chosen from a finite population will have a value within one standard deviation of the mean, when provided with values for the mean, standard deviation, and all members of the population.

Solution**Part (a):**

The median teaching year for High School A is any value with 100 data values at or below it and 100 data values at or above it. The median teaching year for High School B is the 111th value in the ordered list of values. For High School A the median is in the interval that starts at 7 and ends just before 10, because there are only 94 data values below 7 and 106 data values of at least 7. Therefore the median cannot be less than 7. For High School B the median is in the interval that starts at 4 and ends just before 7 because there are more than half (113) of the data values less than 7. Therefore the median must be less than 7. So High School A must be the one with a median of 7, and High School B must be the one with a median of 6.

Another way to determine which school has the median of 7 is to notice that the distribution for High School B is highly skewed to the right, whereas the distribution for High School A is bimodal with a few possible outliers on the right. A distribution that is highly right-skewed is likely to have a substantially larger mean than median. The mean of both distributions is given as 8.2 years, so it makes sense that the highly right-skewed distribution (High School B) is the one with the bigger gap between the mean and median and, therefore, the one with the lower median of 6.

Part (b):

The mean for the original 200 teachers was given as 8.2 years, and the mean for the additional 18 teachers is 2.5 years. Therefore the mean for the combined data set is:

$$\frac{(200)(8.2) + (18)(2.5)}{200 + 18} = \frac{1,640 + 45}{218} \approx 7.73 \text{ years.}$$

Part (c):

The interval mean plus or minus 1 standard deviation on either side of the mean is 8.2 ± 7.2 , or from 1.0 year to 15.4 years. Because teaching year is recorded as an integer, the interval includes teaching years 1 to 15. The number of teachers in that interval can be found by adding the heights of the five bars in the histogram for the intervals from 1 to 16, which includes $79 + 34 + 28 + 29 + 19 = 189$. Therefore the probability is $\frac{189}{221} \approx 0.8552$.

2018 SCORING GUIDELINES**Question 5 (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 three components:

1. States that the median is 6 for High School B and the median is 7 for High School A.
2. Provides a reasonable explanation of how the decision was made.
3. Provides the definition of the median or explicitly applies the definition of a median as a criterion in reaching their decision.

OR

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

1. States that the median is 6 for High School B and the median is 7 for High School A.
2. States that High School B shows a skewed distribution (or High School A shows a less skewed distribution).
3. Provides a reasonable explanation of how the more skewed distribution (High School B) would be the one with a larger separation between the mean and median.

Partially correct (P) if the response satisfies the first component and only one of the other two components required for E.

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

Note: An incorrect statistical statement in the response will result in E being lowered to P, but not P being lowered to I. For example,

- If either distribution is described as left skewed, normal, or approximately normal;
- If the discussion would indicate a median different than 7 for High School A or a median different than 6 for High School B.

Part (b) is scored as follows:

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

1. The correct answer that the mean is 7.73.
2. Enough work to show that the answer was obtained as a weighted average of the two individual means.

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

Incorrect (I) if the response does not satisfy the requirements for E or P.

2018 SCORING GUIDELINES**Question 5 (continued)**

Part (c) is scored as follows:

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

1. Calculates that the appropriate interval is 1 to 15.4 or 1 to 15 teaching years.
2. Correctly sums the counts of data values in the numerator based on the intervals provided.
3. Computes the probability using 221 as the denominator.

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

OR

if the response reports the correct probability (0.8552) without supporting work.

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

Notes:

- If the response attempts to use the Empirical Rule or normal distribution to provide the desired probability, the response is scored I.
- If an incorrect count is shown in component 2, for instance by including the interval from 16 to 19, then component 3 is satisfied if that incorrect count is divided by 221 to find the reported probability.
- It is acceptable if the count is slightly off because of difficulty reading the exact heights of the bars in the histogram.
- If only one of component 2 or component 3 is missing, but the correct probability (0.8552) is reported, the response can be scored E.
- If the response recognizes that all values in the histogram bins up to 16 fall within one standard deviation of the mean and reports the interval as 1 to 16, component 1 is satisfied.

2018 SCORING GUIDELINES**Question 5 (continued)****4 Complete Response**

Three parts essentially correct

3 Substantial Response

Two parts essentially correct and one part partially correct

OR

Part (a) essentially correct and two parts partially correct

2 Developing Response

Two parts essentially correct and no parts partially correct

OR

Part (b) or part (c) 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

2018 SCORING GUIDELINES

Question 6

Intent of Question

The primary goals of this question were to assess a student's ability to (1) describe what constitutes a Type II error for a specific hypothesis test; (2) specify a rejection region in terms of values of the sample mean; (3) compute the power of a test for a specific value in the alternative hypothesis; (4) recognize the definition of power; and (5) understand the impact of increasing the sample size on the power of a test.

Solution**Part (a):**

A Type II error occurs when the alternative hypothesis is true, but the null hypothesis is not rejected. In this situation a Type II error would happen if the mean systolic blood pressure of the population of employees is greater than 122 mmHg, but the null hypothesis that it is 122 mmHg is not rejected. In other words a Type II error would happen if the mean blood pressure for the population of employees is higher than the national average, but the test does not conclude that it is higher.

Part (b):

The test is one-sided and the standard deviation is known, so the null hypothesis will be rejected if the test statistic $z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} > 1.645$. With $\mu = 122$, $\sigma = 15$, and $n = 100$, we get $\frac{\bar{x} - 122}{1.5} > 1.645$. Therefore, $\bar{x} > 124.4675$.

Part (c):

If the actual population mean is 125, with $\sigma = 15$ and $n = 100$, then the sampling distribution of $\bar{x}$ is approximately normal with mean of 125 and standard deviation $\frac{15}{\sqrt{100}} = 1.5$. Therefore,

$$P(\bar{x} > 124.4675) = P\left(\frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} > \frac{124.4675 - 125}{1.5}\right) = P(z > -0.355) = 0.64.$$

Part (d):

The probability found in part (c) is called the power of the test.

Part (e):

If the sample size is increased from 100 to something larger, the probability of rejecting the null hypothesis when the population mean is 125 will be higher than it is for a sample of size 100. Intuitively, more data provide a higher probability of a correct conclusion. The technical explanation is that the rejection region will still be $z > 1.645$, but the sampling distributions of the sample mean will have a smaller standard deviation; therefore, the minimum value of $\bar{x}$ for which we would reject the null hypothesis would be lower and, in return, the probability the null hypothesis is rejected will increase.

2018 SCORING GUIDELINES

Question 6 (continued)

Scoring

This question is scored in three sections. Section 1 consists of part (a) and part (b), section 2 consists of part (c), and section 3 consists of part (d) and part (e). Sections 1, 2, and 3 are scored as essentially correct (E), partially correct (P), or incorrect (I).

Section 1 is scored as follows:

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

1. Part (a) includes in the description of a Type II error the fact that the alternative hypothesis is true, either generically or in context of the situation.
2. Part (a) includes in the description of a Type II error the fact that the null hypothesis is not rejected, either generically or in context of the situation.
3. Part (b) includes a correct z-score for the upper 5 percent tail and indicates the correct direction for the rejection region.
4. Part (b) includes $\mu_{\bar{x}} = 122$, $\sigma_{\bar{x}} = 1.5$, and the resulting $\bar{x}$ value.

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

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

Notes:

- If the response in part (a) does not include context, the number of components satisfied is reduced by one (that is, from four to three, or from three to two, and so on). Context includes a reference to units, to blood pressure, to employees, etc.
- If a response in part (a) is clearly referring to an individual's blood pressure as opposed to the mean blood pressure of all employees, neither components 1 nor 2 are satisfied.

Section 2 is scored as follows:

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

1. Recognizes that the null hypothesis will be rejected when $\bar{x} \geq 124.4675$, as found in part (b).
2. Provides the correct sampling distribution for the sample mean when the true mean is 125, including correct values for the mean and standard deviation, either explicitly or by plugging them into the test statistic formula.
3. Provides evidence of using the normal curve and finds the correct probability value.

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

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

Notes:

- Components 1 and 3 can still be satisfied if errors made in finding the rejection region in part (b) are carried into part (c).
- A calculator statement that does not include labels for input values does not satisfy component 2 but may still satisfy components 1 and 3.

2018 SCORING GUIDELINES

Question 6 (continued)

Section 3 is scored as follows:

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

1. Part (d) specifies power as the name of the probability.
2. Part (e) correctly states that the probability would be greater.
3. Part (e) correctly implies that the standard deviation of the sampling distribution decreases, either explicitly or by substituting values into a formula.
4. Part (e) indicates the minimum value of $\bar{x}$ for which the null hypothesis is rejected decreases, either explicitly or by substituting values into a formula.

Note: Component 4 can still be satisfied if a response indicates that a maximum value of $\bar{x}$ for which the null hypothesis is rejected increases if this direction is consistent with answers in parts (b) and (c).

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

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

2018 SCORING GUIDELINES

Question 6 (continued)

4 Complete Response

Three sections essentially correct

3 Substantial Response

Two sections essentially correct and one section partially correct

2 Developing Response

Two sections essentially correct and no sections partially correct

OR

One section essentially correct and one or two sections partially correct

OR

Three sections partially correct

1 Minimal Response

One section essentially correct

OR

No sections essentially correct and two sections partially correct

OR

Sections 1 and 2 incorrect, and section 3 partially correct with exactly three of the four components satisfied

OR

Section 1 partially correct with exactly three of the four components satisfied, and sections 2 and 3 incorrect

2017**AP[®]** **CollegeBoard**

AP Statistics

Scoring Guidelines

2017 SCORING GUIDELINES

Question 1

Intent of Question

The primary goals of this question were to assess a student's ability to (1) explain statistical terms used when describing the relationship between two variables; (2) interpret the slope of a linear regression equation; and (3) calculate a value of y when given a regression equation, a value of x , and a residual.

Solution**Part (a):**

In the context of a scatterplot in which y represents weight and x represents length, the following are defined.

A positive relationship means that wolves with higher values of length also tend to have higher weights.

A linear relationship means that as length increases by one meter, weight tends to change by a constant amount, on average.

A strong relationship means that the data points fall close to a line (or curve).

Part (b):

The slope of 35.02 indicates that two wolves that differ by one meter in length are predicted to differ by 35.02 kilograms in weight, with the longer wolf having the greater weight.

Part (c):

In general, a residual is equal to actual weight minus predicted weight, or equivalently,

$$\text{actual weight} = \text{predicted weight} + \text{residual}.$$

For the wolf with length 1.4 meters and residual of -9.67 , the predicted weight is

$$-16.46 + 35.02(1.4) = 32.568 \text{ kilograms}.$$

Therefore, the actual weight of the wolf is $32.568 + (-9.67) = 22.898$ kilograms.

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 includes the following four components:

1. A reasonable definition of positive
2. A reasonable definition of linear
3. A reasonable definition of strong
4. At least one definition in context

2017 SCORING GUIDELINES

Question 1 (continued)

Partially correct (P) if the response includes only three of the four components.

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

Notes:

- The description of a positive relationship should clearly indicate that relatively low values of one variable tend to appear with relatively low values of the other variable, and relatively high values of the first variable tend to appear with relatively high values of the other variable.
 - Examples of acceptable responses:
 - As length increases, so does weight.
 - Longer wolves weigh more.
 - The points on the graph go up as you move from left to right.
 - Examples of unacceptable responses:
 - As length goes up, weight changes.
 - Both length and weight get bigger.
 - The correlation is greater than 0.
- The description of a linear relationship can take one of two approaches: the data pattern (data points exhibit the pattern of a line in the graph) or the constant rate of change (as the explanatory variable changes, the response variable exhibits a constant rate of change).
 - Examples of acceptable responses:
 - The points generally follow a straight line.
 - The relationship between x and y is straight.
 - Length and width have a constant slope.
 - Examples of unacceptable responses:
 - The points all line up.
 - You can draw a straight line through the points.
 - There is a positive correlation.
 - Every increase in x yields a 35.02 increase in y .
- The description of strong should indicate how close points are to a line.
 - Examples of acceptable responses:
 - Observed values are close to predicted values.
 - Deviations from the least-squares regression line are small.
 - The correlation coefficient is close to 1.
 - Examples of unacceptable responses:
 - All the points are close together.
 - The scatterplots are clustered together.
 - There is a high positive correlation.
- Context can be shown by referring to length and weight or by using meters and kilograms.
- Sketches and graphs can be used to help clarify definitions, but a sketch alone cannot satisfy a definition component.

Part (b) is scored as follows:

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

1. The correct value of 35.02 for the slope.
2. An interpretation that includes an increase of a specified amount of weight for each unit increase in length.
3. An indication that the relationship is not exact by using words such as “on average” or “predicted weight.”

2017 SCORING GUIDELINES

Question 1 (continued)

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

Note: If the response identifies the slope as -16.46 (the intercept value), the second component is satisfied only if the response states that for each one-meter increase in length there is a *decrease* in predicted or average weight of 16.46 kilograms.

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

Part (c) is scored as follows:

Essentially correct (E) if the response includes the following two components:

1. A correct computation for the predicted value 32.568 kilograms.
2. A correct computation for the actual weight 22.9 kilograms using the given residual and the predicted value.

Partially correct (P) if the response provides a correct computation for the predicted value but is not able to complete the correct calculation of the actual weight, including if the residual is defined in the wrong direction as $(\text{predicted weight}) - (\text{actual weight})$ to give an answer of 42.24 kilograms;

OR

if the response provides an incorrect value for the predicted weight, but then uses that value correctly to determine the actual weight as $(\text{predicted weight}) + \text{residual} = (\text{predicted weight}) + (-9.67)$;

OR

if the response provides a correct answer for the actual weight but does not give sufficient information to determine how it was calculated.

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

Notes:

- The expression $-16.46 + 35.02(1.4)$ is enough to satisfy the first component.
- The equation $-16.46 + 35.02(1.4) - 9.67 = 22.9$ satisfies both components.
- Arithmetic mistakes are overlooked if they do not lead to an unreasonable answer (such as a negative value). For example, $32.568 + (-9.67) = 21.9$ satisfies the second component.

2017 SCORING GUIDELINES**Question 1 (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 two parts partially correct

2017 SCORING GUIDELINES

Question 2

Intent of Question

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

Solution

Part (a):

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

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

Conditions:

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

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

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

Step 2: Correct mechanics

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

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

Step 3: Interpretation

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

Part (b):

Using the confidence interval in part (a), a 95 percent interval estimate for the number of customers in June who asked for a water cup but then filled it with a soft drink is $3,000 \times 0.1883$ to $3,000 \times 0.3867$, or 565 to 1,160. At a cost of \$0.25 per customer, a 95 percent interval estimate for the cost to the restaurant in June is \$141.25 to \$290.00.

2017 SCORING GUIDELINES

Question 2 (continued)

Scoring

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

Section 1 is scored as follows:

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

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

OR

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

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

Notes:

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

Section 2 is scored as follows:

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

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

OR

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

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

Notes:

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

2017 SCORING GUIDELINES

Question 2 (continued)

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

Section 3 is scored as follows:

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

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

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

OR

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

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

Notes:

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

Section 4 is scored as follows:

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

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

OR

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

OR

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

2017 SCORING GUIDELINES**Question 2 (continued)**

OR

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

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

Notes:

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

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

- | | |
|----------|-----------------------------|
| 4 | Complete Response |
| 3 | Substantial Response |
| 2 | Developing Response |
| 1 | Minimal Response |

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

2017 SCORING GUIDELINES

Question 3

Intent of Question

The primary goals of this question were to assess a student's ability to (1) calculate a probability from a normal distribution; (2) calculate a weighted probability from two individual probabilities; and (3) calculate a conditional probability for dependent events when individual and joint probabilities are provided.

Solution**Part (a):**

Let X denote the diameter of a randomly selected melon from Distributor J. X has an approximately normal distribution with mean 133 mm and standard deviation 5 mm.

The z-score for a diameter of 137 mm is $z = \frac{137 - 133}{5} = \frac{4}{5} = 0.8$.

Therefore, $P(X > 137) = P(Z > 0.8) = 1 - 0.7881 = 0.2119$.

Part (b):

Define events:

J : melon is from Distributor J

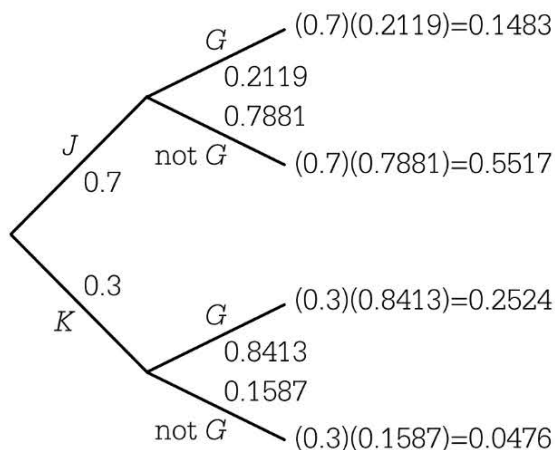
K : melon is from Distributor K

G : melon diameter is greater than 137 mm

$$\begin{aligned} P(G) &= P(G | J) \times P(J) + P(G | K) \times P(K) \\ &= (0.2119)(0.7) + (0.8413)(0.3) \\ &= 0.1483 + 0.2524 \\ &= 0.4007 \end{aligned}$$

For a randomly selected melon from the grocery store,

OR



From the tree diagram, $P(G) = P(G \text{ and } J) + P(G \text{ and } K) = 0.1483 + 0.2524 = 0.4007$.

2017 SCORING GUIDELINES

Question 3 (continued)

Part (c):

Using the events defined in part (b), the requested probability is

$$P(J | G) = \frac{P(J \text{ and } G)}{P(G)} = \frac{P(G | J)P(J)}{P(G)} = \frac{(0.2119)(0.7)}{0.4007} = \frac{0.1483}{0.4007} = 0.3701.$$

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 provides the following three components:

1. Normality and parameters: Indicates use of a normal (or approximately normal) distribution and clearly identifies the correct parameter values. (Showing correct components in a z-score, labeling the mean and standard deviation in a normalcdf calculator statement, or drawing a normal curve with sufficient values marked to indicate the mean and standard deviation are all sufficient methods for identifying the correct parameter values.)
2. Boundary and direction: Uses the correct boundary value of $x = 137$ or $z = 0.8$ and the correct direction. (Showing correct boundary and direction in a probability statement using “greater than,” labeling the lower limit and upper limit in a normalcdf calculator statement, drawing a normal curve with 137 labeled and an indication that the area of interest is to the right of 137, or a conclusion in words using “greater than.”)
3. Probability: Reports the correct normal probability consistent with the response’s setup described in components 1 and 2.

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

Incorrect (I) if the response gives the answer with no work shown or otherwise does not satisfy the criteria for E or P.

Notes:

- An inconsistency in calculations lowers the score for part (a) by one level (from E to P, or from P to I). For instance, if the response states the requested probability as $P(X > 137)$, but actually computes $P(X < 137)$.
- An error in statistical notation, such as using s instead of σ for the standard deviation or $\bar{x}$ instead of μ for the mean, does not satisfy component 1.
- A correct direction can be obtained by using a left direction and an upper bound of 137 and then subtracting the value of the cumulative probability from 1. However, simply showing the calculation $1 - 0.7881 = 0.2119$ does NOT give a correct direction.
- Consider any steps of a hypothesis test as extraneous work.
- Standard notations such as $N(133,5)$ or $N(133,25)$ satisfy the first component.
- A sketch of the normal curve with 133 and 138 in the appropriate positions satisfies the first component.
- Use of a value other than 137 does not satisfy the second component.
- If the only error in part (a) is the reversal of the numerator for the z-score $(133 - 137)$, the response is scored as P.

2017 SCORING GUIDELINES

Question 3 (continued)

- For calculator notation:
 - ncdf is sufficient to indicate use of the normal distribution.
 - Stating the lower and upper bounds is sufficient for identifying boundary and direction.
 - Minimum notation, such as “u” for upper, is sufficient for identifying boundary.
 - The minimum value for an upper bound must be at least 4 standard deviations above the mean.

Part (b) is scored as follows:

Essentially correct (E) if the probability is computed correctly, *AND* work is shown that includes correct numerical values using a formula, end results from a tree diagram, or some other appropriate method.

Partially correct (P) if the response provides a correct strategy for finding the probability, such as a formula or tree diagram, but uses one or more incorrect values or does not use the end result from the tree;

OR

if the response gives the correct probability but not enough work is shown to determine how it was found.

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

Notes:

- A response that uses 0.3 for J and 0.7 for K earns a P if no other mistakes are made.
- Tree diagrams:
 - If used, the multiplication is implied, but the addition must be indicated.
 - If not used, both the multiplication and the addition must be indicated.
- The shown work might be found in part (c).
- Probabilities can be shown in a table. Addition is implied if “Total” is labeled, as shown.

Intersection	G	Not G	Total
J	0.1483	0.5517	0.7
K	0.2524	0.0576	0.3
Total	0.4007	0.5993	

Part (c) is scored as follows:

Essentially correct (E) if the probability is computed correctly *AND* work is shown that illustrates how the probability was found.

Partially correct (P) if the response provides a strategy for finding the probability for dependent variables, such as the appropriate formula or end results from the correct tree diagram, but uses one or more incorrect values or does not use the end result from the tree;

OR

if the response gives the correct probability but not enough work is shown to determine how it was found.

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

2017 SCORING GUIDELINES**Question 3 (continued)**

Notes:

- Part (c) earns an E if a correct answer follows from incorrect values used from part (a) and/or part (b), provided all values are between 0 and 1, inclusive.
- A transcription error in a response can be ignored if correct work is shown.
- For any part, if the resulting probability or part of the calculation of the probability uses a value that is not between 0 and 1, inclusive, the score is lowered by one level (from E to P, or from P to I).
- In any part, unsupported answers are scored as I.

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 two parts partially correct

2017 SCORING GUIDELINES

Question 4

Intent of Question

The primary goals of this question were to assess a student's ability to use boxplots to (1) compare multiple sets of data; (2) identify which set of data is most likely to have produced a particular summary value; and (3) determine which variable is most useful for classifying a new observation.

Solution**Part (a):**

The median value for the percent of chemical Z in the pottery pieces is similar for all three sites, at about 7 percent. The ranges for the percent of chemical Z are much different for the three sites, with the smallest range of about 2 percent (from 6 percent to 8 percent) at Site II, a range of about 6 percent (from about 4 percent to 10 percent) at Site I, and the largest range of about 8 percent (from about 3 percent to 11 percent) at Site III.

Part (b):

- (i) The piece most likely originated at Site III. Although values outside of the range of data observed in the samples would be possible, using the available data results in approximate minimum and maximum sums of the percents for the three chemicals as shown in the table below. Site III is the only site in which 20.5 falls between the sums of the minimum and maximum values.

	Site I		Site II		Site III	
Chemical	Min	Max	Min	Max	Min	Max
X	6	8	5	7	5	7.5
Y	11	15	1.9	4	6	8
Z	4	10	6	8	3	11
Sum	21	33	12.9	19	14	26.5

- (ii) Chemical Y would be most useful, because the distribution of the percentages of total weights at the three sites do not overlap. The distributions of chemicals X and Z have substantial overlap.

Scoring

This question is scored in three sections. Section 1 consists of part (a), section 2 consists of part (b-i), and section 3 consists of part (b-ii). 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 includes the following three components:

1. Recognition that the medians or centers are almost the same for the three sites
2. Recognition that the variability (ranges, IQRs, spread) is different across the three sites
3. Context is included

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

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

2017 SCORING GUIDELINES

Question 4 (continued)

Notes:

- In all sections, comments about shape should be ignored because complete shape information is not obtainable from boxplots.
- Responses are not required to give numerical values. If responses provide numerical values, any reasonable approximation from the boxplots is acceptable.
- Because the boxplots are all symmetric, it is acceptable if the response discusses means instead of medians.
- Any discussion of chemical X and chemical Y is considered extraneous.
- Context is satisfied by any of the following references: site, chemical, weight, total weight, X, Y, or Z.

Section 2 is scored as follows:

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

1. Site III is chosen.
2. Sums of the minimum and maximum are computed for the three chemicals at each site.
3. A reasonable numerical justification is given involving sums of a statistical measure across the three chemicals to choose Site III.

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

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

Notes:

- If the response computes only the sum of the minimums for Site I and the sum of the maximums for Site II and recognizes that this is sufficient, the response is scored E.
- If an alternative measure is used that involves sums of the three chemicals, such as the sum of the medians or the sums of the first quartiles and sums of the third quartiles, instead of the minimum and maximum sums, the second component is not satisfied, but the third component might be satisfied.
 - If the response explicitly or implicitly compares the alternate sum to the other two sites (for example, by indicating that the sum is the closest to 20.5 percent or by listing the sums for all three sites) the response is scored P.
 - If the response does not have an implicit or explicit comparison, the response is scored I.
- If either Site I or Site II is identified as the correct choice, no matter how that choice is justified, the response is scored I.
- The approximate sums of the medians are 27.5 for Site I, 16 for Site II, and 20 for Site III.

Section 3 is scored as follows:

Essentially correct (E) if the response chooses chemical Y *AND* gives a reasonable justification based on the fact that the distributions of chemical Y are distinctive across sites.

Partially correct (P) if the response chooses chemical Y *AND* provides justification based on the boxplots, but does not clearly explain that the distributions of chemical Y are distinctive across sites;

OR

if the response correctly discusses that the distributions of chemical Y are distinctive across sites, but never explicitly chooses chemical Y as the best choice, for instance, by stating only that there is substantial overlap across sites for chemicals X and Z but no overlap for chemical Y.

2017 SCORING GUIDELINES

Question 4 (continued)

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

Notes:

- To justify that the distributions of chemical Y are distinctive across sites, the justification must address both location and variability of the boxplots; for example, by stating that the boxplots do not overlap for chemical Y.
- If the response chooses chemical X or Z *OR* chooses chemical Y with no reasonable justification, the response is scored I.
- The justification that the distributions of chemical Y are distinctive across sites:
 - The following are acceptable because both location and variability are addressed. Such responses are scored E.
 - The boxplots for chemical Y do not overlap, or the boxplots for chemicals X and Z overlap.
 - All values of Site I are high, all values of Site II are low, and all values of Site III are in the middle.
 - The ranges never intersect.
 - The boxplots share no data.
 - Has completely different percentages at each site.
 - The following are incomplete justifications and are scored P.
 - The boxplots vary.
 - Chemical Y varies the most.
 - Chemical Y has the greatest variation.
 - The variation between/among sites is the largest.
 - The boxplots are different.
 - The medians/means differ.
 - The medians/means are most variable.
 - There is a difference in the percentages of chemical Y for each site.
 - The distribution of percents differs the most among the sites.

2017 SCORING GUIDELINES**Question 4 (continued)****4 Complete Response**

Three sections essentially correct

3 Substantial Response

Two sections essentially correct and one section partially correct

2 Developing Response

Two sections essentially correct and no sections partially correct

OR

One section essentially correct and one or two sections partially correct

OR

Three sections partially correct

1 Minimal Response

One section essentially correct

OR

No sections essentially correct and two sections partially correct

2017 SCORING GUIDELINES

Question 5

Intent of Question

The primary goal of this question was to assess a student's ability to identify, set up, perform, and interpret the results of an appropriate hypothesis test to address a particular question. More specific goals were to assess a student's ability to (1) state appropriate hypotheses; (2) identify the appropriate statistical test procedure and check appropriate conditions for inference; (3) calculate the appropriate test statistic and p -value; and (4) draw an appropriate conclusion, with justification, in the context of the study.

Solution

Step 1: State a correct pair of hypotheses.

The null hypothesis is that age group at diagnosis and gender are independent (that is, they are not associated) for the population of people currently being treated for schizophrenia.

The alternative hypothesis is that age group at diagnosis and gender are not independent for the population of people currently being treated for schizophrenia.

Step 2: Identify a correct test procedure (by name or formula) and check appropriate conditions.

The appropriate test is a chi-square test of independence.

The conditions for this test are satisfied because:

1. The question states that the sample was randomly selected.
2. The expected counts for the eight cells of the table are at least 5, as seen in the following table, with expected counts shown below observed counts.

		Age at Diagnosis				
		20 to 29	30 to 39	40 to 49	50 to 59	Total
Women	46	40	21	12	119	
	56.91	36.22	17.25	8.62		
Men	53	23	9	3	88	
	42.09	26.78	12.75	6.38		

Step 3: Find the value of the test statistic and the p -value.

The test statistic is calculated as $\chi^2 = \sum \frac{(O - E)^2}{E}$, or

$$\begin{aligned}\chi^2 &= 2.093 + 0.395 + 0.817 + 1.322 \\ &\quad + 2.830 + 0.534 + 1.105 + 1.788 \\ &= 10.884.\end{aligned}$$

The p -value is $P(\chi^2 \geq 10.884) = 0.012$, based on $(4 - 1) \times (2 - 1) = 3$ degrees of freedom.

2017 SCORING GUIDELINES

Question 5 (continued)

Step 4: State the conclusion in context, with linkage to the p -value.

Because the p -value is very small (for instance much smaller than $\alpha = 0.05$), we reject the null hypothesis and conclude that the sample data provide strong evidence that there is an association between age group at diagnosis and gender for the population currently being treated for schizophrenia.

Scoring

This question is scored in three sections. Section 1 consists of steps 1 and 2 (stating the correct hypotheses, identifying the appropriate test procedure, and checking the technical conditions); section 2 consists of step 3 (performing the correct mechanics); and section 3 consists of step 4 (stating a correct conclusion with justification). Sections 1, 2, and 3 are 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. States BOTH hypotheses correctly with context included in at least one of them
2. Identifies a chi-square test of independence by name or formula
3. Verifies appropriate conditions that minimally include the condition for the expected counts and do not include any incorrect conditions (such as normality)

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

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

Notes:

- Stating the expected count condition is not sufficient; the condition must be checked by reporting the expected counts, or minimally by showing that the smallest expected count is at least 5.
- The random sample condition was stated in the stem so need not be explicitly checked.
- If the null and alternative hypothesis are correctly stated in terms of population proportions, component 1 is satisfied. For example:

$H_0 : p_1 = p_2 = p_3 = p_4$, where p_i is the population proportion of women at each indicated age group, 1, 2, 3, or 4, who are currently being treated for schizophrenia.

H_a : At least one of the population proportions, p_1, p_2, p_3, p_4 , differs from the other three.

OR

H_a : The population proportions for the four age groups are not all the same.

Section 2 is scored as follows.

Essentially correct (E) if the response correctly calculates the following two values:

1. The value of the chi-square test statistic
2. The p -value, critical value, or p -value range from chi-square table

Partially correct (P) if the response correctly calculates only one of the two values.

2017 SCORING GUIDELINES

Question 5 (continued)

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

Notes:

- If the response makes an error in one calculation, subsequent calculations are considered correct if they follow correctly from the initial miscalculation.
- With 3 degrees of freedom, the correct critical value is 7.81 for a significance level of 0.05 and 11.34 for a significance level of 0.01.
- Work does not have to be shown for calculations of test statistic or p -value. However, if incorrect work (other than minor arithmetic/transcription errors) is shown it is considered to be an incorrect calculation of the respective component, even if the correct value is given.
- If a response provides a test statistic that is not a chi-square test statistic, section 2 is scored I.

Section 3 is scored as follows.

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

1. A correct conclusion about the alternative hypothesis.
2. Justification of the conclusion based on linkage between the p -value and a reasonable alpha (or linkage between test statistic and critical value).
3. The conclusion is stated in context.

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

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

Notes:

- If the response provides a correct decision, in context, with linkage to the p -value, but the decision is stated in terms of the null hypothesis with no conclusion about the alternative hypothesis, component 1 is not satisfied.
- Incorrect statistical statements are considered incorrect conclusions for the hypothesis test and do not satisfy component 1.
- If the conclusion is consistent with the p -value from section 2, and also in context with justification based on linkage to the p -value, section 3 is scored E.
- If no alpha level is given, the solution must be explicit about the linkage by giving a correct interpretation of the p -value or explaining how the conclusion follows from the p -value. For example, stating that because the p -value is small, we reject the null hypothesis or stating that because the p -value is large, we do not reject the null hypothesis.
- A decision about the null hypothesis (reject H_0 or fail to reject H_0) is not required, but if such a statement is given the scoring of the decision is considered in component 2.

2017 SCORING GUIDELINES**Question 5 (continued)****4 Complete Response**

Three sections essentially correct

3 Substantial Response

Two sections essentially correct and one section partially correct

2 Developing Response

Two sections essentially correct and no sections partially correct

OR

One section essentially correct and one or two sections partially correct

OR

Three sections partially correct

1 Minimal Response

One section essentially correct

OR

No sections essentially correct and two sections partially correct

OR

Section 1 partially correct and the other two sections incorrect

2017 SCORING GUIDELINES

Question 6

Intent of Question

The primary goals of this question were to assess a student's ability to (1) calculate probabilities associated with treatment and control group memberships for two different methods of random assignment and (2) justify which method of random assignment is more appropriate in a given situation.

Solution**Part (a):**

- (i) Let T (tail) represent being assigned to the treatment group and H (head) represent being assigned to the control group for each coin flip. The process stops when either the treatment group or the control group has two members. The outcomes and their probabilities are as follows.

Arrangement	A	B	C	D	E	F
Chip outcomes	TT	THT	THH	HH	HTH	HTT
Calculation	$\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)$	$\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)$	$\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)$	$\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)$	$\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)$	$\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)$
Probability	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$

- (ii) Man 1 and Man 2 are assigned to the same group for arrangements A and D, so the probability is

$$P(A) + P(D) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

Part (b):

- (i) Let T represent being assigned to the treatment group and C represent being assigned to the control group for each chip drawn. The process stops when either the treatment group or the control group has two members. The probabilities differ from the coin flip method because chips are drawn *without* replacement. The outcomes and their probabilities are as follows.

Arrangement	A	B	C	D	E	F
Chip outcomes	TT	TCT	TCC	CC	CTC	CTT
Calculation	$\left(\frac{2}{4}\right)\left(\frac{1}{3}\right)$	$\left(\frac{2}{4}\right)\left(\frac{2}{3}\right)\left(\frac{1}{2}\right)$	$\left(\frac{2}{4}\right)\left(\frac{2}{3}\right)\left(\frac{1}{2}\right)$	$\left(\frac{2}{4}\right)\left(\frac{1}{3}\right)$	$\left(\frac{2}{4}\right)\left(\frac{2}{3}\right)\left(\frac{1}{2}\right)$	$\left(\frac{2}{4}\right)\left(\frac{2}{3}\right)\left(\frac{1}{2}\right)$
Probability	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

- (ii) Man 1 and Man 2 are assigned to the same group for arrangements A and D, so the probability is

$$P(A) + P(D) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}.$$

2017 SCORING GUIDELINES

Question 6 (continued)

Part (c):

Use the chip method. The chip method gives equal probability to all possible arrangements, but the coin method does not, as shown in the tables from parts (a-i) and (b-i). Furthermore, the coin method is more likely to result in imbalanced treatment groups with regard to students and teachers, based on the probabilities in parts (a-ii) and (b-ii). If food preferences for teachers are different than for students, the imbalance is a problem. For example, if one treatment group consists entirely of students, it would be impossible to know if a difference in the response variable is due to the treatment (type of meal) or the role of the person at the school (teacher or student).

Scoring

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

Part (a) is scored as follows:

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

1. The correct probability for each arrangement in the table of (i).
2. Appropriate justification is shown in (i).
3. The correct probability is given in (ii).

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

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

Notes:

- For component 2, examples of appropriate justification include the following:
 - Shows calculation of a correct probability for at least one of arrangements A and D *AND* shows calculation of a correct probability for at least one of arrangements B, C, E, and F. Calculations must be linked to a specific arrangement with a label, by proximity to an arrangement in the table, or by following the structure of the table or display of arrangements.
 - Lists correct coin outcomes for at least one of arrangements A and D *AND* lists correct coin outcomes for at least one of arrangements B, C, E, and F (for example, A: TT and B: THT). Lists of outcomes must be linked to a specific arrangement with a label, by proximity to an arrangement in the table, or by following the structure of the table or display of arrangements.
 - Shows a tree diagram that ends with six branches corresponding to the six arrangements. Probabilities or coin outcomes must be listed on the branches, but arrangements do not need to be identified.
 - Shows a tree diagram that ends with eight (or 16) branches where the branches leading to at least one of arrangements A and D are identified, and the branches leading to at least one of arrangements B, C, E, and F are identified. Probabilities or coin outcomes must be listed on the branches.
- If an incorrect probability is given in part (ii), component 3 is satisfied if the probability is between 0 and 1 and:
 - is the result of a minor arithmetic error; *OR*
 - is the result of adding the probabilities of arrangements A and D from the table in (i).

2017 SCORING GUIDELINES

Question 6 (continued)

Part (b) is scored as follows:

Essentially correct (E) if the response includes the following two components:

1. The correct probability for each arrangement in the table of (i).
2. The correct probability is given in (ii).

Partially correct (P) if the response includes only one of the three components.

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

Note: If an incorrect probability is given in (ii), component 2 is satisfied if the probability is between 0 and 1 and

- is the result of a minor arithmetic error; *OR*
- is the result of adding the probabilities of arrangements A and D from the table in (i); *OR*
- work is shown that demonstrates the same incorrect approach is used in both parts (a-ii) and (b-ii). For example, component 2 is satisfied if the response forgets to include $P(D)$ and says

$P(A) = \frac{1}{4}$ in (a-ii) and $P(A) = \frac{1}{6}$ in part (b-ii). However, unlabeled answers of $\frac{1}{4}$ in (a-ii) and $\frac{1}{6}$ in (b-ii) would not satisfy component 2 because it is unclear if the same approach is being used.

Part (c) is scored as follows:

Essentially correct (E) if the response chooses the chip method AND includes the following three components:

1. Provides a statistical benefit to the chip method. For example, stating that all arrangements are equally likely.
2. Provides a statistical drawback to the coin method. For example, stating that the coin method is more likely to result in imbalanced treatment groups.
3. States that the responses (opinions, food preferences) of teachers and students might be different.

Partially correct if the response chooses the chip method and includes only one or two of the three components.

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

Notes:

- A correct comparative statement can be used to satisfy both components 1 and 2. For example, stating that the coin method is *more likely* to result in a disproportionate number of students in one treatment group, or that the chip method has a *greater chance* of providing roughly equivalent treatment groups.
- Benefits and drawbacks must be about the probabilities of the arrangements (from the tables in (i)), or the probabilities of imbalanced (balanced) groups (from the questions in (ii)), or both.
- If the response chooses the coin method, look back to see if the coin method should be the preferred method based on incorrect work from parts (a) and (b). If so, score part (c) using the corresponding three components.

2017 SCORING GUIDELINES**Question 6 (continued)**

- If the response says that it does not matter which method to use, look back to see if the tables are the same in parts (a-i) and (b-i) or if the probabilities are the same in parts (a-ii) and (b-ii). If so, part (c) is partially correct if the response justifies the decision by stating that the two tables or the two probabilities are the same. To be essentially correct, the response also needs to satisfy component 3.
- If the response does not choose a method, deduct one component from the number of correct components.

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 two parts partially correct