AP[°] **Statistics** Sample Student Responses and Scoring Commentary

Inside:

Free-Response Question 3

- \square Scoring Guidelines
- ☑ Student Samples
- **☑** Scoring Commentary

© 2022 College Board. College Board, Advanced Placement, AP, AP Central, and the acorn logo are registered trademarks of College Board. Visit College Board on the web: collegeboard.org. AP Central is the official online home for the AP Program: apcentral.collegeboard.org.

Question 3: Focus on Probability and Sampling Distributions

General Scoring Notes

- Each part of the question (indicated by a letter) is initially scored by determining if it meets the criteria for essentially correct (E), partially correct (P), or incorrect (I). The response is then categorized based on the scores assigned to each letter part and awarded an integer score between 0 and 4 (see the table at the end of the question).
- The model solution represents an ideal response to each part of the question, and the scoring criteria • identify the specific components of the model solution that are used to determine the score.

Model Solution

Random variable A, which represents the amount **(a)** of shampoo in a randomly selected bottle, follows a normal distribution with mean 0.6 liter and standard deviation 0.04 liter. Then, the probability that a randomly selected bottle is underfilled is

$$P(A < 0.5) = P\left(Z < \frac{0.5 - 0.6}{0.04} = -2.5\right) \approx 0.0062.$$

Scoring

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

- Indicates the use of a normal (or 1. approximately normal) distribution and identifies the correct parameter values (mean 0.6 and standard deviation 0.04)
- 2. Specifies the correct event (boundary value and direction), or an event consistent with values reported in component 1
- 3. Provides the correct probability of 0.0062 or probability consistent with components 1 and 2

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

if the response fails to satisfy component 1 and 2, but shows the correct z-score formula, z-score value, and correct probability (e.g.,

 $\frac{0.5 - 0.6}{0.04} = -2.5$, resulting in a probability of 0.04 0.0062).

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

Additional Notes: Component 1

- A response may satisfy component 1 by any of the following or a combination of the following:
 - Graphical: Displaying a graph of a normal density function with the appropriate scale on the horizontal axis showing the mean and standard deviation for the distribution of shampoo amount.

- <u>Calculator function syntax</u>: Labeling correct values of the mean and standard deviation in a "normalcdf" statement, such as normalcdf (lower = -∞, upper = 0.5, mean = 0.6, standard deviation = 0.04).
 Correct specification of the upper and lower bounds is not required to satisfy component 1.
- Words: Using a statement such as "normal distribution with mean 0.6 and standard deviation 0.04."
- <u>Standard Notation</u>: Using standard notation such as N(0.6, 0.04) or $N(0.6, (0.04)^2)$.
- <u>Z-score</u>: Displaying the correct mean and standard deviation in a z-score calculation that includes "z," such as $z = \frac{0.5 0.6}{0.04}$.

Component 2

- A response may satisfy component 2 by any of the following or a combination of the following:
 - <u>Graphical</u>: Displaying a graph of a normal density function with the region of interest (A < 0.5 or Z < -2.5) clearly identified. The shaded area does not need to be proportional, but the boundary should be on the proper side of the mean, and the shading should be in the proper direction.
 - <u>Calculator function syntax</u>: Identifying the lower and upper bounds of the region of interest in a "normalcdf" statement, such as:
 - normalcdf (lower = $-\infty$, upper = 0.5, mean = 0.6, standard deviation = 0.04)
 - normalcdf (lower = $-\infty$, upper = $-2.\beta$, = 0, σ = 1)

Correct specification of the mean and standard deviation is not required to satisfy component 2.

• <u>Words</u>: Specifying the correct event in words with correct numerical values for the boundary value and correct direction, such as "the probability that the amount of shampoo is less than 0.5 liter" or P(amount of shampoo < 0.5).

• <u>Standard Notation</u>: Using standard notation such as: P(A < 0.5) or $P\left(z < \frac{0.5 - 0.6}{0.04}\right)$ or

P(Z < -2.5).

General

- It is not necessary to define the random variable *A* because it is defined in the stem. It is not necessary to define the random variable *Z* because it is standard notation. Any other random variable must be defined correctly.
- An error in statistical notation, such as using s instead of σ for the population standard deviation or using \overline{x} instead of μ for the population mean, does not satisfy component 1.
- If the only error in the response to part (a) is the reversal of the numerator for the z-score (0.6 0.5), the response is scored P.
- An arithmetic or transcription error in a response can be ignored if correct work is shown.

Model Solution

- (b) (i) The random variable of interest, X, is the number of underfilled bottles in a box of 10 bottles. The distribution of X is binomial with parameters n = 10 and p = 0.0062.
 - (ii) The crate will be rejected by the warehouse if two or more underfilled bottles are found in the box. The probability of that is

$$P(X \ge 2) = 1 - P(X \le 1)$$

= $1 - {\binom{10}{1}} (0.0062)^1 (0.9938)^9$
 $- {\binom{10}{0}} (0.0062)^0 (0.9938)^{10}$
 $\approx 0.0017.$

Scoring

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

- 1. Defines a random variable as the number of underfilled bottles in a box of 10 bottles in the response to part (b-i)
- 2. Indicates that the random variable has a binomial distribution with parameters n = 10 and p = 0.0062 (or the probability from part (a)). The parameters may be located in the response to either part (b-i) or part (b-ii)
- 3. Provides supporting work for the calculation of the probability in part (b-ii) that identifies the event of interest
- Calculates the correct probability of approximately 0.0017, or a probability consistent with the response to part (a) or part (b-i)

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

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

Additional Notes:

Component 1

• A response may satisfy component 1 if the response indicates that the random variable is the number of underfilled bottles and n = 10 is used in the description of its distribution.

Component 2

- A response may satisfy component 2 by any of the following:
 - <u>Binomial formula</u>: Using the binomial formula with correct *n* and *p* values. For example:

$$1 - {\binom{10}{1}} (0.0062)^1 (0.9938)^9 - {\binom{10}{0}} (0.0062)^0 (0.9938)^{10}$$

- Words or standard notation: Using a statement such as "binomial distribution with n = 10 and p = 0.0062," or using standard notation such as X = B(10, 0.0062).
- <u>Calculator function syntax</u>: Labeling correct parameter values in a "binomcdf" or "binompdf" statement such as:
 - 1 binomcdf(n = 10, p = 0.0062, upper bound = 1)
 - 1 binompdf(n = 10, p = 0.0062, x = 0) binompdf(n = 10, p = 0.0062, x = 1)
- Referring to a "box" does not satisfy the requirement for parameter n = 10.

Component 3

- A response may satisfy component 3 by any of the following:
 - o <u>Graphical display</u>: Displaying a bar graph of binomial probabilities with appropriate bars shaded.

- <u>Words or standard notation</u>: Specifying the correct event in words with identification of the correct numerical boundary and correct direction, such as "probability that X is at least two" or "probability that at least two bottles are underfilled" or P(at least two bottles are underfilled). Identification of the distribution and parameters may be obtained from the response to part (b-i).
- <u>Random variable</u>: $P(X \ge 2)$ or $1 P(X \le 1)$. Identification of the distribution and parameters may be obtained from the response to part (b-i).
- $\circ \quad \underline{\text{Probability formula}}: \text{ e.g., } 1 \binom{10}{1} (0.0062)^1 (0.9938)^9 \binom{10}{0} (0.0062)^0 (0.9938)^{10}.$
- <u>Calculator function notation</u>: Using calculator function notation with clearly defined arguments. For example:
 - "1 binomcdf (n = 10, p = 0.0062, upper bound = 1)" satisfies component 3 because the binomial parameters and the boundary value are clearly labeled.
 - "1 binomcdf (n = 10, p = 0.0062, 1)" does not satisfy component 3 because the boundary value is not labeled.
 - "1 binomcdf (10, 0.0062, upper bound = 1)" does not satisfy component 3 because the binomial parameters are not labeled.
- Because np = (10)(0.0062) = 0.062 is less than 5, the normal approximation to the binomial distribution is not an appropriate method to calculate the probability, and a response that uses this method does not satisfy component 3. However, a response that uses the normal approximation to the binomial distribution may satisfy component 4 if it displays the correct mean and standard deviation of the binomial distribution AND provides a clear indication of the appropriate collection of possible outcomes included

in the event using a diagram or a z-score, e.g., $P\left(Z \ge \frac{2 - (10)(0.0062)}{\sqrt{(10)(0.0062)(0.9938)}}\right)$ or

$$1 - P\left(Z \le \frac{1 - (10)(0.0062)}{\sqrt{(10)(0.0062)(0.9938)}}\right). \text{ (Note that } \sqrt{(10)(0.0062)(0.9938)} \approx 0.248. \text{)}$$

• An arithmetic or transcription error in a response can be ignored if correct work is shown.

Model SolutionScoringThe company should use the original programming for
the filling machine. For the original machine of
the filling machine. For the original machine of
the filling machine.Essentially correct (E) if the response satisfi
the filling machine.

(c) The company should use the original programming for the filling machine. For the original programming of the filling machine, the probability of an underfilled bottle is

$$P(A < 0.5) = P\left(Z < \frac{0.5 - 0.60}{0.04}\right)$$
$$= P(Z < -2.5) \approx 0.0062.$$

For the adjusted programming of the filling machine, the probability of an underfilled bottle is

$$P(A < 0.5) = P\left(Z < \frac{0.5 - 0.56}{0.03}\right)$$
$$= P(Z < -2.0) \approx 0.02275$$

Because the probability of an underfilled bottle is greater for the adjusted programming, this would result in more rejected shipments. The company should continue with the original machine programming. **Essentially correct (E)** if the response satisfies the following two components by comparing either probabilities or *z*-scores:

Comparing probabilities:

- Correctly calculates the probability of underfilling a bottle as 0.023 for the adjusted programming of the filling machine
- 2. Provides a correct conclusion about which programming (adjusted or original) should be recommended based on a comparison of the probabilities calculated for the original and adjusted programming

OR

Comparing *z*-scores:

- 1. Correctly calculates the *z*-score for the adjusted programming
- 2. Provides a correct conclusion about which programming (adjusted or original) should be recommended based on a comparison of the *z*-scores (e.g., a higher *z*-score results in more bottles being underfilled) calculated for the original and adjusted programming

Partially correct (P) if the response satisfies only one of the two components required for an E.

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

Additional Notes:

• A response that correctly uses the binomial distribution to find the probability that a crate will be rejected with correct values and justification should be scored E. For the original programming, this probability is 0.0017, and for the adjusted programming, this probability is 0.0206.

Adjusted programming:

Let *Y* represent the number of underfilled shampoo bottles in a box of 10 using the adjusted programming. $P(Y \ge 2) = 1 - P(Y \le 1)$

$$= 1 - {\binom{10}{1}} (0.02275)^{1} (0.97725)^{9} - {\binom{10}{0}} (0.02275)^{0} (0.97725)^{10} \approx 0.0206$$

• A response that incorrectly computes the probability that a crate will be rejected, with or without justification, should be scored P if it provides a correct conclusion based on comparing that probability to the probability computed in part (b-ii).

- Component 2 is not satisfied if no recommendation is made for choice of programming. A response stating "yes" or "no" is not sufficient for indicating a choice of programming.
- An arithmetic or transcription error in a response can be ignored if correct work is shown.

Scoring for Question 3	Score
Complete Response	4
Three parts essentially correct	
Substantial Response	3
Two parts essentially correct and one part partially correct	
Developing Response	2
Two parts essentially correct and no part partially correct	
OR	
One part essentially correct and one or two parts partially correct	
OR	
Three parts partially correct	
Minimal Response	1
One part essentially correct and no part partially correct	
OR	
No part essentially correct and two parts partially correct	

Sample 3A

Question 3

Begin your response to QUESTION 3 on this page.

3. A machine at a manufacturing company is programmed to fill shampoo bottles such that the amount of shampoo in each bottle is normally distributed with mean 0.60 liter and standard deviation 0.04 liter. Let the random variable A represent the amount of shampoo, in liters, that is inserted into a bottle by the filling machine.

(a) A bottle is considered to be underfilled if it has less than 0.50 liter of shampoo. Determine the probability that N(0,6,0.04) a randomly selected bottle of shampoo will be underfilled. Show your work.

 $z = \frac{0.5 - 0.6}{0.04} = -2.5$ -P(A<0.5) = P(z < -2.5) = norreal cdf(-99, -2.5) = 0.0062 L of sharpoo There is about a 0.62% chance a random bottle of shampoo will be underfilled.

After the bottles are filled, they are placed in boxes of 10 bottles per box. After the bottles are placed in the boxes, several boxes are placed in a crate for shipping to a beauty supply warehouse. The manufacturing company's contract with the beauty supply warehouse states that one box will be randomly selected from a crate. If 2 or more bottles in the selected box are underfilled, the entire crate will be rejected and sent back to the manufacturing company.

- (b) The beauty supply warehouse manager is interested in the probability that a crate shipped to the warehouse will be rejected. Assume that the amounts of shampoo in the bottles are independent of each other.
- X = * of bottles in 10 (one box) which will be wider filled (i) Define the random variable of interest for the warehouse manager and state how the random variable is distributed.

(ii) Determine the probability that a crate will be rejected by the warehouse manager. Show your work.

Binon (n=10, p=,0062)

 $P(X \ge 2) = |-P(X \le 1) = |-P(X = 0 \le 1) = |-b|_{NONCO} f(n=10, p=0.0062, x=1) =$ =1-0.9983 =0.0017

There is about a 0.10%, chance a crate will be rejected by the Manager.

Unauthorized copying or reuse of this page is illegal.

Page 6

GO ON TO THE NEXT PAGE.

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

合适 化合理

Question 3

Continue your response to QUESTION 3 on this page.

To reduce the number of crates rejected by the beauty supply warehouse manager, the manufacturing company is considering adjusting the programming of the filling machine so that the amount of shampoo in each bottle is normally distributed with mean 0.56 liter and standard deviation 0.03 liter.

Unauthorized copying or reuse of this page is illegal.

Q5149/7

0179836

Page 7

GO ON TO THE NEXT PAGE.

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



=

Question 3

Continue your response to QUESTION 3 on this page.

To reduce the number of crates rejected by the beauty supply warehouse manager, the manufacturing company is considering adjusting the programming of the filling machine so that the amount of shampoo in each bottle is normally distributed with mean 0.56 liter and standard deviation 0.03 liter.

(c) Would you recommend that the manufacturing company use the original programming of the filling machine or the adjusted programming of the filling machine? Provide a statistical justification for your choice.

The manufacturing company should use th se beca filling programming of t original P crat Deine CX. OW - 0 ex 0.1 emelu 6 was a Ole TRY DOT e1 SIL A ine or 0062096799 00 ander wa alr toli antikel(1) DI remei 10w or ld entu mmc no led, as the c elinood und CT a bottle with the adjusted nolei ore at 0, 822 be mi 10 al 0 obabilituof crate estn d(90 Grause 71 a use compand show line. nning

Unauthorized copying or reuse of this page is illegal.

Q5149/7

0021714

Page 7

GO ON TO THE NEXT PAGE

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

Sample 3C

Question 3 Begin your response to QUESTION 3 on this page. A machine at a manufacturing company is programmed to fill shampoo bottles such that the amount of shampoo in each bottle is normally distributed with mean 0.60 liter and standard deviation 0.04 liter. Let the random variable A represent the amount of shampoo, in liters, that is inserted into a bottle by the filling machine. (a) A bottle is considered to be underfilled if it has less than 0.50 liter of shampoo. Determine the probability that a randomly-selected bottle of shampoo will be underfilled. Show your work. X=. 60 liters O= . OU liturs one sample z test P(Shappo will be underfilled): .50-.60 = -2.5 L> .0062 There is a . 62-1- chance that a bottle of shapped will be under filled .62-1- chance problem states it is a normal distribution After the bottles are filled, they are placed in boxes of 10 bottles per box. After the bottles are placed in the boxes, several boxes are placed in a crate for shipping to a beauty supply warehouse. The manufacturing company's contract with the beauty supply warehouse states that one box will be randomly selected from a crate. If 2 or more bottles in the selected box are underfilled, the entire crate will be rejected and sent back to the manufacturing company. (b) The beauty supply warehouse manager is interested in the probability that a crate shipped to the warehouse will be rejected. Assume that the amounts of shampoo in the bottles are independent of each other. (i) Define the random variable of interest for the warehouse manager and state how the random variable is distributed. The vandom voriable of interest is the amount of under filled bottles within a book. This variable is randomly distributed into boxes which ap into crotes, (ii) Determine the probability that a crate will be rejected by the warehouse manager. Show your work. paperally empty possibility: 0062 -(12 Possible mess ups .0062 .0000 (1.35 - 26 There is almost a zero percent chance of a crate getting rijected by the manager. Unauthorized copying or reuse of this page is illegal. Page 6 GO ON TO THE NEXT PAGE. Use a pencil or a pen with black or dark blue ink. Do NOT write your name. Do NOT write outside the box. Q5149/6

Sample 3C



Question 3

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

Overview

The primary goals of the question were to assess a student's ability to (1) calculate the probability that a bottle filling machine would underfill a bottle of shampoo using a specified normal distribution; (2) define a random variable as the number of underfilled bottles in a box of ten shampoo bottles; (3) describe the distribution of that random variable; (4) use the identified distribution to compute a probability, showing work; and (5) identify and compare relevant quantities, e.g., probabilities or *z*-scores, to justify a recommendation about whether a specific adjustment to the bottle filling machine should be made.

This question primarily assesses skills in skill category 3: Using Probability and Simulation, and skill category 4: Statistical Argumentation. Skills required for responding to this question include (3.A) Determine relative frequencies, proportion, or probabilities using simulation or calculations, and (4.B) Interpret statistical calculations and findings to assign meaning or assess a claim.

This question covers content from Unit 4: Probability, Random Variables, and Probability Distributions, and Unit 5: Sampling Distributions of the course framework in the AP Statistics Course and Exam Description. Refer to topics 4.3, 4.10, and 5.2, and learning objectives VAR-6.A, UNC-3.B, UNC-3.A, and VAR-4.B.

Sample: 3A Score: 4

The response earned the following: Part (a) - E; Part (b) - E; Part (c) - E.

In part (a) the response satisfies component 1 with the notation N(0.6, 0.04). Component 1 is also satisfied by the use of the z formula clearly identified with "z" and the correct mean and standard deviation. The response satisfies component 2 by using the notation P(A < 0.5) and again in the sketch of the normal distribution by labeling the correct boundary value and shading to the left. The correct probability satisfies component 3. Part (a) was scored essentially correct (E).

In part (b) the response correctly identifies the random variable as "# of bottles in 10 (one box) which will be underfilled," which satisfies component 1. The statement "binom(n = 10, p = .0062)" satisfies component 2. Component 3 is satisfied using the calculator function notation with clearly defined arguments, and component 4 is satisfied with the correct probability. Part (b) was scored essentially correct (E).

In part (c) the correct probability for the adjusted programming satisfies component 1, and the correct conclusion based on a comparison of probabilities satisfies component 2. Part (c) was scored essentially correct (E).

Sample: 3B Score: 3

The response earned the following: Part (a) - E; Part (b) - P; Part (c) - E.

In part (a) the response satisfies component 1 by using calculator function syntax with labels for the correct values of the mean and standard deviation. Component 1 is also satisfied by using a graph of the normal distribution with the appropriate scale on the *x*-axis showing the mean and standard deviation. The response satisfies component 2 by labeling the upper and lower bounds in the calculator function syntax. Component 2 is not satisfied with the sketch of the normal distribution because the boundary value is not clearly identified. The correct probability satisfies component 3. Part (a) was scored essentially correct (E).

© 2022 College Board. Visit College Board on the web: collegeboard.org.

Question 3 (continued)

In part (b) the response does not correctly describe the random variable, so component 1 is not satisfied. The response to part (b-i) identifies the distribution as normal, which is incorrect, so component 2 is not satisfied. In part (b-ii) the binomial distribution is correctly identified by using a correctly labeled calculator syntax, but this is considered a parallel solution, so component 2 is not satisfied. Component 3 is satisfied by using the calculator function notation with the binomial parameters and the boundary value clearly labeled. Component 3 is also satisfied using the notations $P(X \ge 2)$ and $1 - P(X \le 1)$. Component 4 is satisfied with the correct probability. Part (b) was scored partially correct (P).

In part (c) the correct probability for the adjusted programming is eventually stated, so component 1 is satisfied. There is no work shown on how to get this correct probability, but it is unnecessary to get component 1. The response gives the correct conclusion based on a comparison of probabilities, so component 2 is also satisfied. Part (c) was scored essentially correct (E).

Sample: 3C Score: 2

The response earned the following: Part (a) - P; Part (b) - I; Part (c) - E.

In part (a) component 1 is not satisfied because the z formula is not identified with "z," and the mean is identified with the incorrect symbol \bar{x} . The response satisfies component 2 in the sketch of the normal distribution by labeling the correct boundary value and shading to the left. The correct probability satisfies component 3. Part (a) was scored partially correct (P).

In part (b) the response does not satisfy component 1 because the number of trials, 10, is not included. The binomial distribution is not mentioned, so component 2 is not satisfied. There is no supporting work, so component 3 is not satisfied. The probability is incorrect, so component 4 is not satisfied. Part (b) was scored incorrect (I).

In part (c) the correct probability satisfies component 1, and the correct conclusion based on a comparison of probabilities satisfies component 2. Part (c) was scored essentially correct (E).