

Question 3: Focus on Probability and Sampling Distributions

4 points

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	Scoring
A i. $P(\text{Rock Song}) = \frac{100}{1,000} = 0.10$ ii. $P(\text{Both Rock Songs}) = (0.10)(0.10) = 0.01$	Essentially correct (E) if the response satisfies the following four components: - In part A (i) the response calculates the correct probability. - In part A (i) the response provides supporting work for the correct probability. - In part A (ii) the response calculates the correct probability consistent with the answer in part A (i). - In part A (ii) the response provides supporting work consistent with the probability calculated in part A (i). Partially correct (P) if the response satisfies two or three of the four components required for E. Incorrect (I) if the response does not meet the criteria for E or P.

Scoring Notes:

- In part A (ii) a response that does not consider independence when solving (e.g., $\left(\frac{100}{1000}\right)\left(\frac{99}{999}\right) = 0.0099$) does not satisfy component 3 but may satisfy component 4.
- If probabilities are not labeled but are given in the order they are asked, then the response may earn an E. If the probabilities are presented in a different order, there must be a label to satisfy the components. In this case, sufficient labels include probability notation, context, work, or identification of the subparts (i) and (ii).

Model Solution	Scoring
B i. Let the random variable of interest, X , represent the number of the 20 songs played in one hour that are rock songs. It is stated that any song can be replayed at any time, which establishes that each rock song has probability $\frac{100}{1,000} = 0.10$ of being selected each hour and each song is independent from every other song. Therefore, X has a binomial distribution with $n = 20$ independent trials and probability of success $p = 0.10$ for each trial. ii. The expected value for the number of rock songs played in one hour is $np = 20(0.10) = 2$ songs.	Essentially correct (E) if the response satisfies at least three of the following four components: - In part B (i) the response defines the random variable as the number of rock songs played in one hour. - In part B (i) the response describes the distribution as binomial. - In part B (i) or B (ii) the response states that $n = 20$ and $p = 0.10$. - In part B (ii) the response correctly calculates the expected value AND provides supporting work for the calculation of the correct expected value. Partially correct (P) if the response satisfies only two of the four components required for E. Incorrect (I) if the response does not meet the criteria for E or P.

Scoring Notes:

- When defining the random variable, the response must include both “the number of rock songs” and “in one hour” or “out of 20 songs.”
- A response that states $X \sim B(20, 0.1)$ satisfies components 2 and 3.
- If a response states the random variable has a distribution other than binomial (e.g., normal, left skewed, or uniform), part B cannot be scored E.
- Stating that songs are distributed randomly is not a distribution and should be considered extraneous.
- Examples that satisfy components 3 and 4 include:
 - $np = 20(0.10) = 2$
 - $np = 20\left(\frac{100}{1,000}\right) = 2$
 - $n = 20, p = 0.10, np = 2$
- An example that satisfies component 4 only:
 - $20(0.10) = 2$
- An arithmetic or transcription error in a response can be ignored if correct work is shown.

Model Solution	Scoring
C i. The probability that in a particular hour 4 or more rock songs will be played is $P(X \geq 4) = 1 - P(X \leq 3)$ $P(X \geq 4) = 1 - \left[\binom{20}{0}(0.10)^0(0.90)^{20} + \binom{20}{1}(0.10)^1(0.90)^{19} + \binom{20}{2}(0.10)^2(0.90)^{18} + \binom{20}{3}(0.10)^3(0.90)^{17} \right]$ $P(X \geq 4) = 1 - 0.867 = 0.133.$ ii. No, the probability that 4 or more rock songs would be played in an hour is 0.133, which is high enough to be reasonably attributed to chance alone. This probability is not small enough to provide evidence that the selection process was not truly random.	Essentially correct (E) if the response satisfies the following four components: - In part C (i) the response provides a correct probability. - In part C (i) the response shows work that supports the correct probability. - In part C (ii) the response indicates that there is not a strong reason to believe that the selection process was not truly random. - In part C (ii) the response provides an explanation that correctly links the probability to the decision. Partially correct (P) if the response satisfies only two or three of the four components required for E. Incorrect (I) if the response does not meet the criteria for E or P.

Scoring Notes:

- A response may satisfy component 2 by any of the following:
 - Graphical display: Displaying a bar graph of binomial probabilities including axes with scale with appropriate bars shaded.
 - Probability formula: For example,

$$1 - \binom{20}{0}(0.10)^0(0.90)^{20} - \binom{20}{1}(0.10)^1(0.90)^{19} - \binom{20}{2}(0.10)^2(0.90)^{18} - \binom{20}{3}(0.10)^3(0.90)^{17}.$$
 - Calculator function notation: Using calculator function notation with clearly defined arguments. For example:
 - $1 - \text{binomcdf}(n = 20, p = 0.10, \text{upper bound} = 3)$ satisfies component 2 because the boundary value is clearly labeled.
 - $1 - \text{binomcdf}(n = 20, p = 0.10, 3)$ does not satisfy component 2 because the boundary value is not labeled.
 - $\text{Binomcdf}(n = 20, p = 0.10, \text{lower bound} = 4, \text{upper bound} = 20)$ satisfies component 2 because the boundary value is clearly labeled.
 - Random Variable: $P(X \geq 4)$ or $1 - P(X \leq 3)$ with identification of the binomial distribution with correct parameters ($n = 20$ and $p = 0.10$) included in part C satisfies component 2.
- An arithmetic or transcription error in a response can be ignored if correct work is shown.
- A response that indicates that the manager does have a strong argument that the selection process was not truly random (or responds “yes”) that is adequately supported by an explanation based on an incorrectly calculated probability in part C (i) may satisfy components 3 and 4.
- A response that indicates that the manager does have a strong argument that the selection process was not truly random (or responds “yes”) that is supported by a statement claiming the probability is low may satisfy components 3 and 4.

- A response that indicates that the manager does not have a strong argument that the selection process was not random supported by the calculation of the standard deviation of the binomial distribution, 1.342, and an explanation based on 4 being within two standard deviations of the expected value (mean) may earn credit for components 3 and 4.
- If a response gives two arguments, treat them as parallel solutions and score the weaker solution.
- A response that finds the probability of exactly 4 songs playing (e.g., $\text{binompdf}(20, 0.1, 4) = 0.089$) and explains that this is not strong evidence that the selection process was not truly random may still satisfy components 3 and 4.

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

AP[®] STATISTICS
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,

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

AP[®] STATISTICS
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).

AP[®] STATISTICS
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, $0.0077 + 0.10615 = 0.1385$.

AP[®] STATISTICS
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