2017

AP CollegeBoard

AP Statistics

Sample Student Responses and Scoring Commentary

Inside:

- ☑ Free Response Question 3
- ☑ Scoring Guideline
- ☑ Student Samples
- ☑ Scoring Commentary

AP Central is the official online home for the AP Program: apcentral.collegeboard.org

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

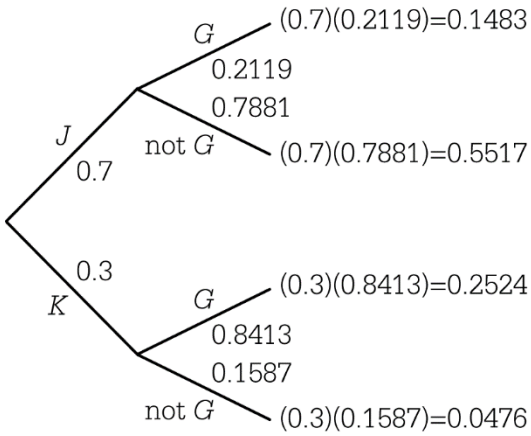
Part (b):

Define events:
J: melon is from Distributor J
K: melon is from Distributor K
G: melon diameter is greater than 137 mm

For a randomly selected melon from the grocery store,

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

OR



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

Question 3 (continued)

Part (c):

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

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

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.

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.

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

3A

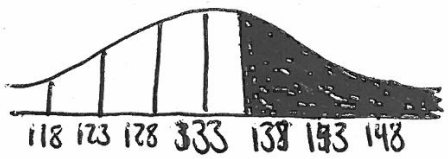
3. A grocery store purchases melons from two distributors, J and K. Distributor J provides melons from organic farms. The distribution of the diameters of the melons from Distributor J is approximately normal with mean 133 millimeters (mm) and standard deviation 5 mm.

(a) For a melon selected at random from Distributor J, what is the probability that the melon will have a diameter greater than 137 mm?

$$P(d>137\text{ mm}) = P\left(z > \frac{137-133}{5}\right) = 1 - P\left(z < \frac{137-133}{5}\right)$$

$$= P(z > .8000) = 1 - (.7881)$$

$$= \boxed{.2119}$$



Distributor K provides melons from nonorganic farms. The probability is 0.8413 that a melon selected at random from Distributor K will have a diameter greater than 137 mm. For all the melons at the grocery store, 70 percent of the melons are provided by Distributor J and 30 percent are provided by Distributor K.

(b) For a melon selected at random from the grocery store, what is the probability that the melon will have a diameter greater than 137 mm?

$$P(\text{Dist K} \cap d>137) + P(\text{Dist. J} \cap d>137) = P(d>137)$$

$$(.30 \times .8413) + (.70 \times .2119) = \boxed{.40072}$$

$$P(\text{Dist K})P(d>137 \mid \text{Dist K}) + P(\text{Dist J})P(d>137 \mid \text{Dist. J})$$

(c) Given that a melon selected at random from the grocery store has a diameter greater than 137 mm, what is the probability that the melon will be from Distributor J?

$$P(\text{Dist J} \mid d>137) = \frac{P(\text{Dist. J} \cap d>137)}{P(d>137)}$$

$$= \frac{(.70 \times .2119)}{(.40072)}$$

$$= \boxed{.3702}$$

GO ON TO THE NEXT PAGE.

3B

3. A grocery store purchases melons from two distributors, J and K. Distributor J provides melons from organic farms. The distribution of the diameters of the melons from Distributor J is approximately normal with mean 133 millimeters (mm) and standard deviation 5 mm.

(a) For a melon selected at random from Distributor J, what is the probability that the melon will have a diameter greater than 137 mm?

Distributor J
$\mu = 133$ mm
$SD = 5$ mm

$z = \frac{x-\mu}{SD}$

$z = \frac{137-133}{5} = .8$

Normal CDF (.8, 999) =

P(diameter > 137 mm) = .2119

Distributor K provides melons from nonorganic farms. The probability is 0.8413 that a melon selected at random from Distributor K will have a diameter greater than 137 mm. For all the melons at the grocery store, 70 percent of the melons are provided by Distributor J and 30 percent are provided by Distributor K.

(b) For a melon selected at random from the grocery store, what is the probability that the melon will have a diameter greater than 137 mm?

$(.212)(.70) + (.8413)(.30) = .4008$

(c) Given that a melon selected at random from the grocery store has a diameter greater than 137 mm, what is the probability that the melon will be from Distributor J?

Let J = Distributor J
G = Greater than 137

$P(J/G137) = \frac{P(J \cap G)}{P(G)} = \frac{(.30)(.8413)}{.4008} = \frac{.25239}{.4008}$

$= .6297$

3C

3. A grocery store purchases melons from two distributors, J and K. Distributor J provides melons from organic farms. The distribution of the diameters of the melons from Distributor J is approximately normal with mean 133 millimeters (mm) and standard deviation 5 mm.

(a) For a melon selected at random from Distributor J, what is the probability that the melon will have a diameter greater than 137 mm?

normalcdf (137, ∞, 133, 5) = 0.3446

The probability that the melon selected at random from Distributor J will have a diameter greater than 137 mm is 0.3446.

Distributor K provides melons from nonorganic farms. The probability is 0.8413 that a melon selected at random from Distributor K will have a diameter greater than 137 mm. For all the melons at the grocery store, 70 percent of the melons are provided by Distributor J and 30 percent are provided by Distributor K.

(b) For a melon selected at random from the grocery store, what is the probability that the melon will have a diameter greater than 137 mm?

$0.7(0.3446) + 0.3(0.8413) = 0.49361$

For a melon selected at random from the grocery store, the probability that the melon will have a diameter greater than 137 mm is 0.49361.

(c) Given that a melon selected at random from the grocery store has a diameter greater than 137 mm, what is the probability that the melon will be from Distributor J?

$P(J \mid D) = \frac{P(J \cap D)}{P(D)} = \frac{0.3446}{0.49361} = 0.698$

Given that a melon selected at random from the grocery store has a diameter greater than 137 mm, the probability that the melon will be from Distributor J is 0.648.

AP® STATISTICS
2017 SCORING COMMENTARY

Question 3

Overview

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.

Sample: 3A
Score: 4

In part (a) the response indicates the use of the normal distribution and identifies the mean and standard deviation using the standard statistical notation inside the probability statement $P\left(z > \frac{137 - 133}{5}\right)$, satisfying component 1. The graph of the normal distribution, with the mean of 133 in the appropriate position on the horizontal axis, and the standard deviation identified by the sequence of values 118, 123, 128, 133, 138, 143, 148 on the axis, also satisfies component 1. The response satisfies component 2 several ways. Both $P(d > 137 \text{ mm})$ and $P\left(z > \frac{137 - 133}{5}\right)$ give the boundary with direction. The graph alone seems to indicate direction, but because 137 is not labeled on the axis, the boundary is not clear. Finally, the response gives the final answer of .2119 showing supporting work and so satisfies component 3. Because each of the three components is satisfied, part (a) was scored as essentially correct. In part (b) the response includes the work $(.30 \times .8413) + (.70 \times .2119) = .40072$. The equation has the required work of the appropriate multiplications and addition and produces the correct answer, so part (b) was scored as essentially correct. The response is enhanced by the supporting symbolic probabilities, but they are not necessary for the score of essentially correct. In part(c) the response produces the correct answer and includes the supporting work $\frac{.70(.2119)}{.40072} = .3702$, so part (c) was scored as essentially correct. Again, the response is enhanced by the supporting symbolic probabilities, but they are not necessary for the score of essentially correct. Because three parts were scored as essentially correct, the response earned a score of 4.

Sample: 3B
Score: 3

In part (a) the response indicates the use of the normal distribution and identifies the mean and standard deviation using the standard statistical notation $z = \frac{137 - 133}{5}$, satisfying component 1. The graph of the normal distribution, with the mean of 133 in the appropriate position on the horizontal axis and the standard deviation identified by the sequence of values 128, 133, 138, 143, 148 on the axis, also satisfies component 1. Finally, the combination of "Normal CDF" with $\mu = 133$ mm and $\text{SD} = 5$ mm satisfies component 1. The response places the value of 137 in the proper position of the z-score $z = \frac{137 - 133}{5}$, giving the boundary. The graph shades in an area to the right of 137, giving both the boundary and direction. The response also includes "P(diameter > 137mm)", which gives both the boundary and the direction. Because of the multiple ways the response gives the boundary and direction, the response satisfies component 2. Finally, the response produces a correct probability of .2119 and so satisfies component 3. Because each of the three components is satisfied, part (a) was scored as essentially correct. In part (b) the response includes the work $(.212)(.70) + (.8413)(.30) = .4008$. The equation has the required work of the appropriate multiplications and addition and produces the correct answer, so part (b) was scored as essentially correct.

Question 3 (continued)

In part (c) the response calculates $\frac{.30(.8413)}{.4008} = \frac{.25239}{.4008} = .6297$. The numerator is not the correct product for calculating $P(J \text{ and } G)$. The numerator in the response is actually the product that calculates $P(K)P(G \mid K) = P(K \text{ and } G)$. Because the response provides a reasonable strategy for finding the probability for dependent variables, but uses incorrect values, part(c) was scored as partially correct. Because two parts were scored as essentially correct, and one part was scored as partially correct, the response earned a score of 3.

Sample: 3C
Score: 2

In part (a) the response uses a "normalcdf" command, which satisfies the normality component. But the values of 133 and 5 are not labeled as the mean and standard deviation, so the parameters are not identified, and component 1 is not satisfied. The response earns the boundary and direction component from the final sentence, which includes the phrase "diameter greater than 137 mm." The given probability of 0.3446 is incorrect, so component 3 is not satisfied. Because only one of the three components is satisfied, part (a) was scored as incorrect. In part (b) the response includes the equation $0.7(0.3446) + 0.3(0.8413) = 0.49361$. Although the probability is not correct, the approach is correct. The error in the probability is caused by the use of 0.3446 calculated in part (a). If the numerical value had been correct, the probability of 0.49361 would be correct, so part (b) was scored as essentially correct. In part (c) the response uses the ratio $\frac{0.3446}{0.49361}$ to obtain the probability of 0.698. Using the incorrect values from parts (a) and (b), a correct response should be in the form $\frac{(0.7)(\text{answer from part (a)})}{(\text{answer from part (b)})}$. The response is missing only the 0.7, so part (c) was scored as partially correct. Because one part was scored as essentially correct, one part was scored as partially correct, and one part was scored as incorrect, the response earned a score of 2.