

AP[®] STATISTICS
2015 SCORING GUIDELINES

Question 6

Intent of Question

The primary goals of this question were to assess a student's ability to (1) describe how sample data would differ using two different sampling methods; (2) describe the sampling distribution of the sample mean for two different sampling methods; and (3) choose the sampling method that will result in the best estimate of the population mean.

Solution

Part (a):

No, a sample obtained using Method 2 will not be representative of all tortillas made that day. The sample obtained using Method 2 will only represent the tortillas from one production line, not from the entire population because the distributions of diameters for the two production lines are different.

Part (b):

Method 1 was most likely used to select this sample. The bimodal shape in the histogram of sample data indicates that tortillas were selected from both production lines, which is what would happen using Method 1. Method 2 would be likely to produce a unimodal distribution of diameters centered at either 5.9 inches or 6.1 inches.

Part (c):

Method 2 would result in less variability in the sample of 200 tortillas on a given day because the sample comes from only one production line. Because the distributions of diameters are not the same for the two production lines, selecting tortillas from both lines as in Method 1 would result in more variable sample data.

Part (d):

The sampling distribution of the sample mean diameter for samples obtained using Method 1 would be approximately normal with mean 6 inches and standard deviation $\frac{0.11}{\sqrt{200}} \approx 0.0078$ inch.

Part (e):

Method 1 would result in less variability in the sample means over the 365 days, because with Method 2, roughly half of the sample means will be around 5.9 inches and the other half will be around 6.1 inches. With Method 1, however, the sample means will all be very close to 6 inches, as indicated by the standard deviation in part (d).

Part (f):

Method 1 is more likely to produce a sample mean close to 6 inches. Because the sample mean is an unbiased estimator for both methods, the manager should pick the method that would result in less variability in the distribution of the sample mean. Based on the answer to part (e), Method 1 results in less variability in the distribution of the sample mean.

AP[®] STATISTICS

2015 SCORING GUIDELINES

Question 6 (continued)

Scoring

This question is scored in three sections. Section 1 consists of parts (a), (b), and (c), section 2 consists of part (d), and section 3 consists of parts (e) and (f). 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 includes the following three components:

1. In part (a) the response says no *AND* either argues that the sample will only be selected from one production line and not the entire population (that is, the sample will only represent one production line) *OR* argues that the tortillas from the two productions lines are different.
2. In part (b) the response chooses Method 1 *AND* refers to a relevant characteristic of the histogram (shape, center, or variability) that matches what would be expected when using Method 1 (or that does not match what would be expected when using Method 2).
3. In part (c) the response chooses Method 2 *AND* either justifies by stating that the sample comes from only one production line (does not come from both production lines) *OR* justifies by comparing the possible range of diameters for the two methods.

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.

Note: If a response includes more than one justification in an individual part, score the weaker of the two justifications. For example, a response for part (a) that says “No, because only one line was selected and because the sample size was too small” does not satisfy the first component because the sample size argument is incorrect.

Section 2 is scored as follows:

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

1. The shape is approximately normal.
2. The mean is 6 inches.
3. The standard deviation is $\frac{0.11}{\sqrt{200}}$ inch.

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.

Note: It is not necessary to include units (inches) for the mean or standard deviation.

Section 3 is scored as follows:

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

1. In part (e) the response chooses Method 1 *AND* describes the sampling distribution of the sample mean for Method 2 as having some *sample means* close to the mean of production line A (5.9 inches) and the other *sample means* close to the mean of production line B (6.1 inches).

AP[®] STATISTICS

2015 SCORING GUIDELINES

Question 6 (continued)

2. In part (f) the response chooses Method 1 *AND*
- refers to a correct answer in part (e);
- OR
- describes the sampling distribution of the sample mean for Method 2 as having some *sample means* close to the mean of production line A (5.9, less than 6) and the other *sample means* close to the mean of production line B (6.1, greater than 6);
- OR
- argues that on a single day, it would be preferable to get a sample with a mean around 6 rather than getting a sample that would have a mean around 5.9 (less than 6) or a mean around 6.1 (greater than 6).

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

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

Notes:

- In part (e), the response must be clear that there is *more than one* mean being described (365 sample means).
 - Correct: The sample means will be around 5.9 or 6.1.
 - Not correct: The sample mean will be around 5.9 or 6.1.
- Parts (e) and (f) are not satisfied if the response does not imply the sample means *vary* from the population means in both methods (or does not imply that the sample mean varies from the population mean when making the single day argument in part (f)). However, if the mistake is made in part (e), do not penalize the response in part (f) for the same mistake.
 - Correct: The sample means will be around 5.9 or 6.1 instead of close to 6.
 - Not correct: The sample means will be 5.9 or 6.1 instead of close to 6.
 - Not correct: The sample means will be 6 instead of around 5.9 or 6.1.

4 Complete Response

All three sections essentially correct

3 Substantial Response

Two sections essentially correct and one section partially correct

2 Developing Response

Two sections essentially correct and one section incorrect

OR

One section essentially correct and either one or two sections partially correct

OR

All three sections partially correct

1 Minimal Response

One section essentially correct and two sections incorrect

OR

Two sections partially correct and one section incorrect

6A1

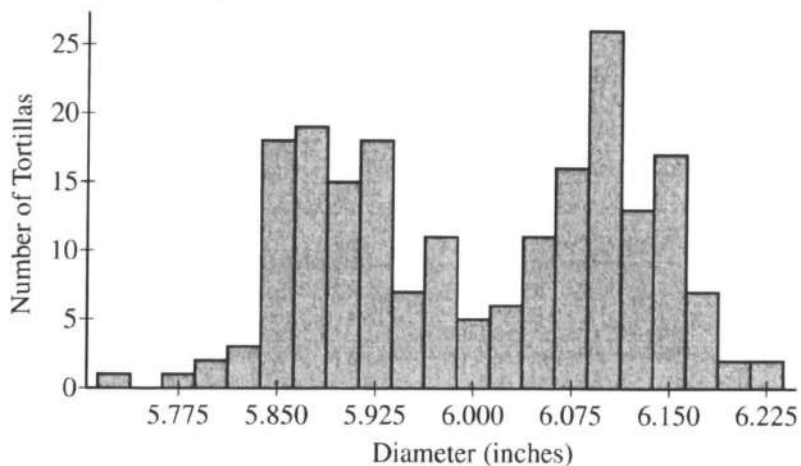
6A1 of 3

6A1

- (a) Will a sample obtained using Method 2 be representative of the population of all tortillas made that day, with respect to the diameters of the tortillas? Explain why or why not.

No. Method 2 suffers from selection bias because the sample will not be obtained from the entire population of tortillas, and the tortillas that are not sampled from will tend to have a different diameter.

- (b) The figure below is a histogram of 200 diameters obtained by using one of the two sampling methods described. Considering the shape of the histogram, explain which method, Method 1 or Method 2, was most likely used to obtain a such a sample.



Method 1

The histogram is bimodal, with a peak at ≈ 5.9 and a peak at ≈ 6.1 inches. This means the sample was likely obtained from both production lines, or Method 1.

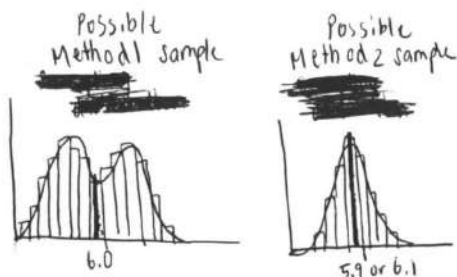
6A2

6A2

6A2 of 3

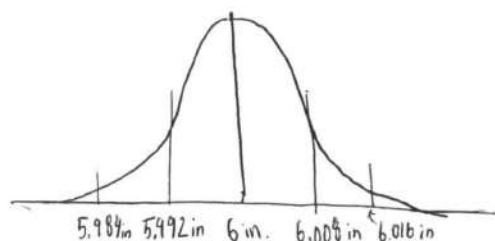
- (c) Which of the two sampling methods, Method 1 or Method 2, will result in less variability in the diameters of the 200 tortillas in the sample on a given day? Explain.

Method 2. Because Method 2 obtains the sample from just one production line, the sample will only have one peak, while Method 1 will have 2 peaks. Thus, Method 2 produces samples with smaller spreads/standard deviations.



Each day, the distribution of the 200,000 tortillas made that day has mean diameter 6 inches with standard deviation 0.11 inch.

- (d) For samples of size 200 taken from one day's production, describe the sampling distribution of the sample mean diameter for samples that are obtained using Method 1.



- Because $n=200 \geq 30$, the central limit theorem says that the sampling distribution will be approximately normal.
- The mean of the sampling distribution is equal to the population mean.
- the standard deviation is:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.11}{\sqrt{200}} = 0.00778 \text{ in}$$

GA3

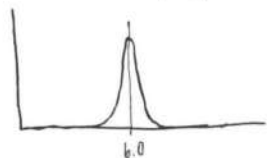
GA3
GA3 of 3

- (e) Suppose that one of the two sampling methods will be selected and used every day for one year (365 days). The sample mean of the 200 diameters will be recorded each day. Which of the two methods will result in less variability in the distribution of the 365 sample means? Explain.

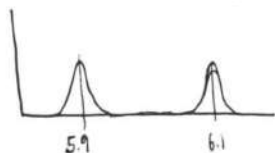
Method 1. The sampling distribution for Method 2 is bimodal, with a peak

at 5.9 in and a peak at 6.1 in, because the samples have means close to either of these two values. This results in much greater variability in sample means than the unimodal Method 1 sampling distribution.

Method 1 sampling distribution



Method 2 sampling distribution



- (f) A government inspector will visit the facility on June 22 to observe the sampling and to determine if the factory is in compliance with the advertised mean diameter of 6 inches. The manager knows that, with both sampling methods, the sample mean is an unbiased estimator of the population mean. However, the manager is unsure which method is more likely to produce a sample mean that is close to 6 inches on the day of sampling. Based on your previous answers, which of the two sampling methods, Method 1 or Method 2, is more likely to produce a sample mean close to 6 inches? Explain.

Method 1.

Method 2 is likely to produce a sample mean close to either 5.9 inches or 6.1 inches, depending on the chosen production line, but not close to 6.0 inches. Method 1 ~~tends~~ selects tortillas from both production lines, so the tortilla diameter tends to average out very close to 6.0 inches.

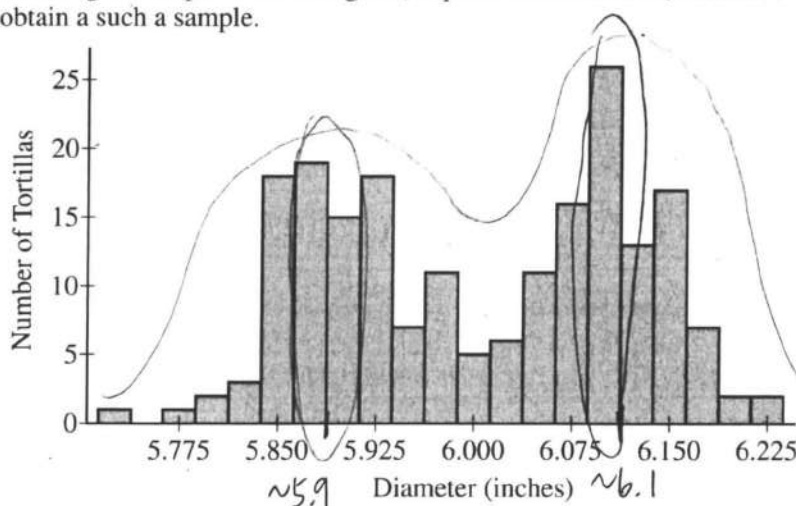
6B1

6B1.f.3 6B1

- (a) Will a sample obtained using Method 2 be representative of the population of all tortillas made that day, with respect to the diameters of the tortillas? Explain why or why not.

Method 2 will not produce a sample that is representative of the population of all tortillas made that day because before the sample is taken, only 1 of the 2 production lines is selected. So the data from Method 2 would only provide information on the tortilla diameters of that selected production line.

- (b) The figure below is a histogram of 200 diameters obtained by using one of the two sampling methods described. Considering the shape of the histogram, explain which method, Method 1 or Method 2, was most likely used to obtain a such a sample.



Method 1 was most likely used because the histogram displays 2 apparent peaks—one centered about 5.9 inches and another centered about 6.1 inches. Both of these diameters are the population means for production lines A and B, respectively. This shows that both production lines had to be included, which means Method 2 was not used.

6B2

6B2
6B2.0f3

- (c) Which of the two sampling methods, Method 1 or Method 2, will result in less variability in the diameters of the 200 tortillas in the sample on a given day? Explain.

Method 2 would result in less variability in the diameters of the tortillas because Method 2 involves sampling tortillas from only one of the production lines. Tortillas from one production line will vary less than tortillas from both production lines b/c tortillas from one production line will be more similar in diameter to each other.

Each day, the distribution of the 200,000 tortillas made that day has mean diameter 6 inches with standard deviation 0.11 inch.

total

- (d) For samples of size 200 taken from one day's production, describe the sampling distribution of the sample mean diameter for samples that are obtained using Method 1.

$$\frac{0.11}{\sqrt{200}}$$

The sampling distribution of the sample mean diameter for samples obtained using Method 1 will have a mean of 6 inches and a standard deviation of 0.007778 inches ($200,000 \geq 10(200)$ ✓). The sampling distribution will be normally distributed because by the Central Limit Theorem, the sample size of 200 is large enough. ($200 \geq 30$ ✓)

6B3

365 trials
of $n=200$
(sample size)
6B3.f3

- (e) Suppose that one of the two sampling methods will be selected and used every day for one year (365 days). The sample mean of the 200 diameters will be recorded each day. Which of the two methods will result in less variability in the distribution of the 365 sample means? Explain.

Method 2 will result in less variability in the distribution of the 365 sample means because Method 2 involves sampling from only one of the production lines. Tortillas made from one production tend to be more similar in diameter to each other. Therefore, Method 2 would produce less variable sample means and less variability overall.

- (f) A government inspector will visit the facility on June 22 to observe the sampling and to determine if the factory is in compliance with the advertised mean diameter of 6 inches. The manager knows that, with both sampling methods, the sample mean is an unbiased estimator of the population mean. However, the manager is unsure which method is more likely to produce a sample mean that is close to 6 inches on the day of sampling. Based on your previous answers, which of the two sampling methods, Method 1 or Method 2, is more likely to produce a sample mean close to 6 inches? Explain.

Method 1 is more likely to produce a sample mean close to 6 inches because Method 1 involves sampling from both production lines and as shown in the histogram of part b, there will be 2 apparent peaks in the histogram—one about 5.9 in. and another about 6.1 in. With these 2 apparent peaks, the sample mean will be roughly the average of 5.9 and 6.1 at about 6 in, which is the desired sample mean. Method 2 would most likely provide a sample mean closer to 5.9 in. or to 6.1 in. b/c it only takes measurements from 1 production line.

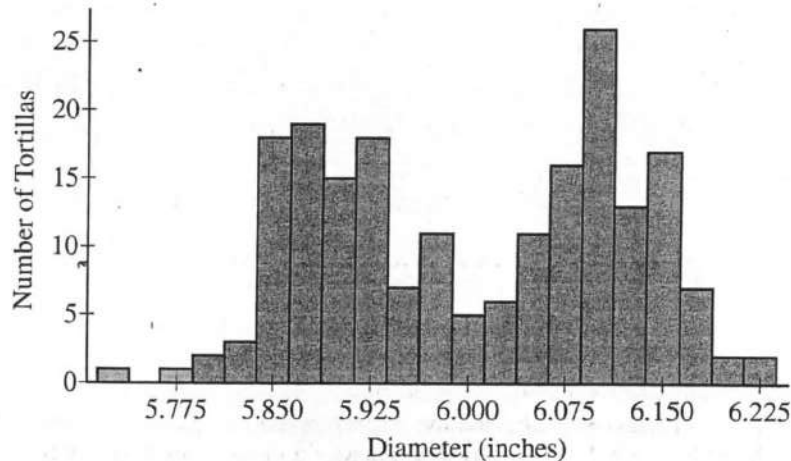
6C1

6C1 of 3

- (a) Will a sample obtained using Method 2 be representative of the population of all tortillas made that day, with respect to the diameters of the tortillas? Explain why or why not.

No, Method 2 will only be representative of the population of tortillas at the production line that is picked. Because Method 2 uses only one production line that differs from the other production line ^{in mean diameter}, the results in Method 2 would not produce results that represent ^{both} Production line A and B.

- (b) The figure below is a histogram of 200 diameters obtained by using one of the two sampling methods described. Considering the shape of the histogram, explain which method, Method 1 or Method 2, was most likely used to obtain a such a sample.



Method one was used to obtain this sample because the histogram is bimodal around 5.9 and 6.1, which accounts for production line A and production line B. If Method B was used, the histogram would be unimodal at either 5.9 or 6.1, because only one production line would be tested.

6C2

6C2 of 3

- (c) Which of the two sampling methods, Method 1 or Method 2, will result in less variability in the diameters of the 200 tortillas in the sample on a given day? Explain.

Method 2 will result in less variability in the diameters of the 200 tortillas in the sample on a given day. This is because Method 2 tests one production line with a standard deviation of the production line that's chosen. Method 1, however, tests both production lines whose variances would combine.

Method 2: $SD = \text{standard deviation of production line B}$

Method 1: $\text{Var}(A+B) = \text{Var } A + \text{Var } B$

$$SD^2(A+B) = SD^2 A + SD^2 B$$

$$SD(A+B) = \sqrt{SD^2 A + SD^2 B}$$

$$\frac{5}{\sqrt{n}}$$

Each day, the distribution of the 200,000 tortillas made that day has mean diameter 6 inches with standard deviation 0.11 inch.

- (d) For samples of size 200 taken from one day's production, describe the sampling distribution of the sample mean diameter for samples that are obtained using Method 1.

$$\text{Mean} = 6 \quad SD = 0.11$$

For sample size 200, the mean diameter is the same as the population diameter at $\text{mean} = 6$. The standard error is equal to the population standard deviation divided by \sqrt{n} $\left(\frac{5}{\sqrt{n}}\right)$, which is equal to $\frac{0.11}{\sqrt{200}}$,

$$\text{or } 0.007778.$$

6C3

6C3.6C3

- (e) Suppose that one of the two sampling methods will be selected and used every day for one year (365 days). The sample mean of the 200 diameters will be recorded each day. Which of the two methods will result in less variability in the distribution of the 365 sample means? Explain.

Method 2 because each production line has a population standard deviation of less than .1, while the population standard deviation of all 200,000 tortillas is .11.

- (f) A government inspector will visit the facility on June 22 to observe the sampling and to determine if the factory is in compliance with the advertised mean diameter of 6 inches. The manager knows that, with both sampling methods, the sample mean is an unbiased estimator of the population mean. However, the manager is unsure which method is more likely to produce a sample mean that is close to 6 inches on the day of sampling. Based on your previous answers, which of the two sampling methods, Method 1 or Method 2, is more likely to produce a sample mean close to 6 inches? Explain.

Method 1 is more likely to produce a sample mean close to 6 inches because the mean of method one should be 6, while method 2 should produce a mean of either 5.9 or 6.1 inches.

AP[®] STATISTICS

2015 SCORING COMMENTARY

Question 6

Overview

The primary goals of this question were to assess a student's ability to (1) describe how sample data would differ using two different sampling methods; (2) describe the sampling distribution of the sample mean for two different sampling methods; and (3) choose the sampling method that will result in the best estimate of the population mean.

Sample: 6A

Score: 4

In part (a) the response says no and justifies the choice by stating that “the sample will not be obtained from the entire population of tortillas, and the tortillas that are not sampled from will tend to have a different diameter,” satisfying the first component of section 1. In part (b) the response says Method 1 because “the histogram is bimodal,” satisfying the second component of section 1. In part (c) the response says Method 2 because the sample is “from just one production line,” satisfying the third component of section 1. The picture in part (c), although not required, strengthens the explanation. Because the response satisfies all three components, section 1 was scored as essentially correct. In part (d) the response states that the sampling distribution will be “approximately normal,” satisfying the first component of section 2. The justification using the Central Limit Theorem is appropriate but not required. The response also correctly identifies the mean and standard deviation of the sampling distribution of the sample mean, satisfying the second and third components of section 2. The picture in part (d), although not required, is a nice way to summarize the response. Because the response satisfies all three components, section 2 was scored as essentially correct. In part (e) the response says Method 1 and correctly describes the sampling distribution of the sample mean for Method 2 as having some means close to 5.9 and other means close to 6.1, satisfying the first component of section 3. The picture in part (e), although not required, strengthens the already correct explanation. In part (f) the response says Method 1 because on a single day, “Method 2 is likely to produce a sample mean close to either 5.9 inches or 6.1 inches” while the average for Method 1 will be “very close to 6.0 inches.” Because the response satisfies both components, section 3 was scored as essentially correct. Because all three sections were scored as essentially correct, the response earned a score of 4.

Sample: 6B

Score: 3

In part (a) the response says “Method 2 will not produce a sample that is representative of the population” because “only 1 of the 2 production lines is selected,” satisfying the first component of section 1. In part (b) the response says Method 1 because the histogram “displays 2 apparent peaks,” satisfying the second component of section 1. In part (c) the response says Method 2 because the sample comes from “only one of the production lines,” satisfying the third component of section 1. Because the response satisfies all three components, section 1 was scored as essentially correct. In part (d) the response correctly identifies the mean and standard deviation of the sampling distribution of the sample mean, satisfying the second and third components of section 2. The response also states that the sampling distribution “will be normally distributed,” satisfying the first component of section 2. It would have been better if the response said the distribution should be *approximately* normally distributed, but in the scope of a six-part investigative task, this was considered a minor error. Because the response satisfies all three components, section 2 was scored as essentially correct. In part (e) the response says Method 2, which is incorrect. It appears that the response is describing the distribution of sample data from Method 2, as in part (c), rather than the sampling distribution of the sample mean. In part (f) the response says Method 1 and argues that on a single day, Method 1 would produce a mean “at about 6 in.” while Method 2 would likely “provide a sample mean closer to 5.9 in. or to 6.1 in.,” satisfying the second component of section 3. Because the response satisfies only one

AP[®] STATISTICS
2015 SCORING COMMENTARY

Question 6 (continued)

of the two components, section 3 was scored as partially correct. Because two sections were scored as essentially correct, and one section was scored as partially correct, the response earned a score of 3.

Sample: 6C

Score: 2

In part (a) the response says no and justifies the choice by stating that the tortillas “will only be representative of the population of tortillas at the production line that is picked,” satisfying the first component of section 1. Although not necessary, the response continues the justification by describing how the two lines are different. In part (b) the response says Method 1 because “the histogram is bimodal” and provides a second justification by describing how the histogram would look if Method 2 was used. Either of these justifications is sufficient, and the second component of section 1 is satisfied. In part (c) the response says Method 2 because it “tests one production line.” Although the response also includes an incorrect application of the formula $\sqrt{\sigma_A^2 + \sigma_B^2}$, this was not an essential part of the justification, and the third component is satisfied. Because the response satisfies all three components, section 1 was scored as essentially correct. In part (d) the response correctly identifies the values of the mean and standard deviation of the sampling distribution of the sample mean, satisfying the second and third components of section 2. However, the response does not state the shape of the sampling distribution, so the first component is not satisfied. Because the response satisfies only two of the three components, section 2 was scored as partially correct. In part (e) the response says Method 2, which is incorrect. In part (f) the response argues that on a single day, the mean using Method 1 will result in a “sample mean close to 6 inches” while “method 2 should produce a mean of either 5.9 or 6.1 inches.” Because there is no indication that the sample means will vary from the population means in Method 2, the response does not satisfy the second component. Because the response satisfies neither of the two components, section 3 was scored as incorrect. Because one section was scored as essentially correct, one section was scored as partially correct, and one section was scored as incorrect, the response earned a score of 2.