AP STAT DEBRIEF – Question 6

Dr. Debbie M. Kohler, Instructor
Kennesaw State University

2015 QUESTION 6: Tortilla Sampling Investigative Task
Please view the questions here:
http://media.collegeboard.com/digitalServices/pdf/ap/ap15_frq_statistics.pdf

Question 6: INTENT OF THE QUESTION:

This is a question whose main intent was to assess a student's ability to describe sampling data and how it changes based on sampling methods. Students were required to describe a sampling distribution for the mean for two different sampling methods and choose the one that would result in the best estimate for the population mean.

SAMPLE SOLUTION

a) Is the sample obtained using Method 2 representative of the population …?
No, a sample using Method 2 will not be representative of all tortillas made that day. It will be composed of only one production line.

b) What method does the figure represent?
Method 1 was used to select this sample. The bimodal shape in the histogram of sample data indicates that the tortillas were selected from both production lines, which would occur using method 1. (Students may also discuss that method 2 would be unimodal, or approximately normal with a mean near the mean of production line 1 or production line 2).

c) Which is less variable?
Method 2 because it is taken from only one production line OR Method 2 because method 2 has a range of diameters from 5.7-6.1 (0.4) or 5.9 – 6.3 (also 0.4) where Method 1 would contain data values from both production lines and would contain diameters from 5.7-6.3(0.8)

d) What is the sampling distribution……
The sampling distribution is approximately normal with mean = 6 and standard deviation = $\frac{\sigma_x}{\sqrt{n}} = \frac{.11}{\sqrt{200}} = 0.0078$

e) Select one of the two methods every day for 365 days… Which has less variability?
Method 1 because the means are centered around 6. Method 2 would have means centered around 2 different values because some samples come from production line a and some come from production line b, so the means would be centered around 5.9 and 6.1 giving more variability.

f) A government official visits for a day. Which method?
Method 1 will produce a sample mean closer to 6 inches. Since the sample mean is unbiased for both methods, the method with the least variability should be chosen. Method 2 will produce means close to 5. 9 or 6.1 OR
SINGLE DAY APPROACH:

Method 1 would be preferable because the sample mean is close to 6 while Method 2 produces a mean closer to 5.9 or 6. 1 **OR**
Method 1 would be preferable because the sample mean is close to 6 while Method 2 produces a mean smaller or larger than 6

SOLUTION NOTES:

This question contained 6 parts, but was graded as 3 parts. Section 1 consists of parts a, b,c. Section 2 consists of part d. Section 3 consists of parts e & f.

Section 1:

1. *In part (a) the response says " no" AND either argues the sample will only be selected from one production line and not the entire population OR argues that the tortillas from the two production lines are different.*
2. *In part (b) the response choose 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 (and 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 only from one production line (not both) OR justifies by comparing the possible range of diameters for the two methods.*

Section 2:

1. *Shape is approximately normal*
2. *Mean is 6 inches*
3. *Standard deviation is* $\frac{0.11}{\sqrt{200}}$

Section 3:

1. *Part (e)- response chooses Method 1 AND describes the sampling distribution of the sample mean for Method 2 as having <u>some sample means</u> close to the mean of production line A (5.9 inches) and the other sample means close to the mean of production line B*
2. *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***
    - *Indicates that on a single day, it would be preferable to get a sample with a mean around 6 rather than a sample with a mean around 5.9 (less than 6 or a mean around 6.1 (greater than 6)*

COMMON MISTAKES/NOTES:

1. Answer the question! Encourage students to go back and read the question after they have written a solution to ensure they have answered the question.

2. Don't write what's not there! Don't go off on tangents. Students have a tendency to create their own stories about many things that don't pertain to the question. For example---one method may be favored because of the weather on that day.
3. Know correct definitions and apply them at appropriate time. This was a sampling question with a large sample size. Many students talked about **confounding** (variables) and **lurking variables.** This was not appropriate for this question.
4. Review sampling distributions, means of sampling distributions, central limit theorem. Students were all over the place on this part. They definitely could not describe the sampling distribution of a sample of size 200 with standard deviation of 0.11. This should have been quite easy for the students.
5. When describing a sampling distribution, a lot of students missed the mean of the sampling distribution because they said it was ABOUT 6. It was acceptable to say it was centered at 6 since that is the center of the sampling distribution. The student could not say the mean of the sampling distribution was about 6, close to 6, near 6 etc. Those were entirely wrong.
6. Directly address the question, and answer it specifically, and back it up. This may be done by addressing a method, or by describing differences in variability----however if a student thinks that writing the definition of standard deviation is going to get them credit, they are going down the wrong road. There's another method to do the problem.
7. When discussing a sampling distribution, students need to remember that they are discussing a lot of means. Therefore when they talk about the center, they must somehow indicate that the mean for **Method 1 is close to 6, while the mean for Method 2 is close to either 5.9 or 6.1.** It could also be described as Method 2 has **means** that are near the mean for production line A and **means** that are near production line B. [Plural is necessary---and the reader must be able to visualize lots of means that center around the two means for Production Line A and Production Line B].
8. **Work on comparisons with your students! Quite often students will only discuss one of the two methods and this does not compose a comparison.**
9. Watch the use of parallel solutions. If student has justified the answer, they should not have to say "also, because……." If they make a mistake in the second part, the least correct solution is graded.