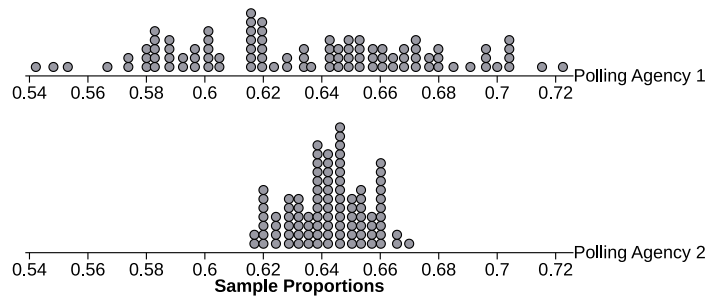
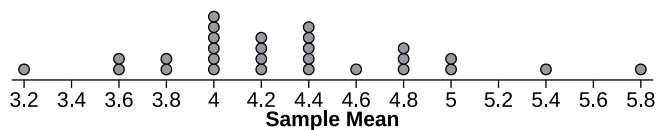


1. There are two polling agencies that are trying to estimate the proportion of voters who will be voting for a certain candidate in an upcoming election. Here are simulated sampling distributions for polling agency 1 and polling agency 2. The true proportion of voters who will be voting for the candidate is 0.64.



- a. What is similar about the two sampling distributions?
- b. What is different about the two sampling distributions?
- c. Which polling agency is likely to be using a larger sample size? Explain.

2. A class of AP Stats students tried to estimate the average word length for Beyonce's song "Crazy in Love". On a printout of the lyrics, each student circled a sample of 5 words and calculated the mean of the sample. The sample means are displayed in this partial sampling distribution.



- The true mean word length for Crazy in Love is 3.53 letters. Based on the dotplot, is the sampling method used by the students producing unbiased estimates? Explain.
- Suppose that the students circled 20 words and found the mean, instead of just circling 5. How would this change the sampling distribution?
- How would the increase in sample size affect the bias of the estimates? Explain.
- What change could the students make to lower the bias of the estimates?

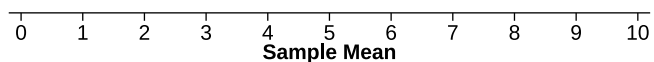
3. Here is the list of ages (in years) for the 20 employees at a small startup company. The average age for all 20 employees is 33.5 years.

21, 22, 23, 24, 25, 25, 29, 30, 30, 30, 33, 34, 35, 35, 35, 39, 40, 42, 43, 75

- a. Two employees are to be randomly selected. List one possible sample that would have an average age greater than 45 years.
- b. Five employees are to be randomly selected. List one possible sample that would have an average age greater than 45 years.
- c. Is it more or less likely for a sample of five employees to have an average age greater than 45 years than a sample of two employees? Explain.

4. The mean of a certain population distribution is 7. Suppose that random samples are taken from this population and the mean is calculated for each sample. Create a dotplot with 10 dots to represent a simulated sampling distribution with the following descriptions.

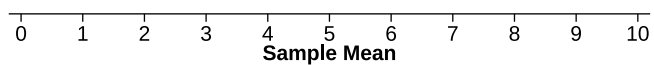
a. Low bias and low variability



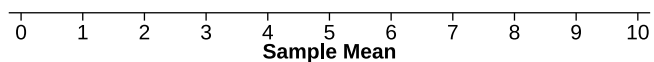
b. Low bias and high variability



c. High bias and low variability



d. High bias and high variability



5. The “Fab Five” was the 1991 University of Michigan basketball team. The starting lineup was all freshmen and is considered to be the best college basketball recruiting class of all time. Here is the number of points per game scored by each of the Fab Five.

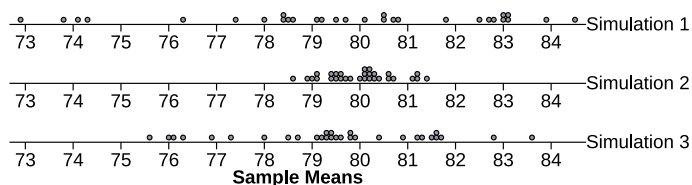
Player	Jalen Rose	Chris Webber	Juwan Howard	Jimmy King	Ray Jackson
Points per game	17.6	15.5	11.1	9.9	4.6

- a. What is the mean points per game for the population of these five players?
- b. Complete the table by listing all 10 possible samples of size $n = 2$ from this population and calculate the sample mean points per game for each sample. The first sample mean is calculated for you.

Sample	17.6, 15.5									
Mean	16.55									

- c. Create a dotplot of the sampling distribution of the sample mean points per game for samples of size 2.
- d. What is the mean of all the values in the sampling distribution?
- e. Is the sample mean an unbiased estimator of the population mean? Justify your answer.
- f. Suppose we had taken samples of size 3 instead of size 2. How would this change the sampling distribution? Justify your answer.

6. A statistician wants to reduce the variability of the sampling distribution of the sample proportion. Which of the following is the most appropriate approach?
- A) Verify the shape of the population distribution is approximately normal.
 - B) Decrease the sample size.
 - C) Verify the population size is at least 10 times larger than the sample size.
 - D) Increase the sample size.
7. A certain population is approximately normal with an unknown mean and standard deviation. Three simulations are performed with technology, each with a different sample size n and each with the same number of trials. For each simulation, the sample mean was calculated for each sample. The dotplots show the approximate sampling distribution of the sample mean for each simulation.



Which of the following is the correct ordering of the dotplots from the smallest value of n to the largest value of n ?

- A) 1, 2, 3
- B) 1, 3, 2
- C) 2, 1, 3
- D) 2, 3, 1