Lesson 21: Why Worry About Sampling Variability?

Classwork

There are three bags, Bag A, Bag B, and Bag C, with 100 numbers in each bag. You and your classmates will investigate the population mean (the mean of all 100 numbers) in each bag. Each set of numbers has the same range. However, the population means of each set may or may not be the same. We will see who can uncover the mystery of the bags!

Exercises

1. To begin your investigation, start by selecting a random sample of ten numbers from Bag A. Remember to mix the numbers in the bag first. Then, select one number from the bag. Do not put it back into the bag. Write the number in the chart below. Continue selecting one number at a time until you have selected ten numbers. Mix up the numbers in the bag between each selection.

Selection	1	2	3	4	5	6	7	8	9	10
Bag A										

   a. Create a dot plot of your sample of ten numbers. Use a dot to represent each number in the sample.

   b. Do you think the mean of all the numbers in Bag A might be 10? Why or why not?

   c. Based on the dot plot, what would you estimate the mean of the numbers in Bag A to be? How did you make your estimate?

   d. Do you think your sample mean will be close to the population mean? Why or why not?

e. Is your sample mean the same as your neighbors' sample means? Why or why not?

2. Repeat the process by selecting a random sample of ten numbers from Bag B.

Selection	1	2	3	4	5	6	7	8	9	10
Bag B										

a. Create a dot plot of your sample of ten numbers. Use a dot to represent each of the numbers in the sample.

b. Based on your dot plot, do you think the mean of the numbers in Bag B is the same or different from the mean of the numbers in Bag A? Explain your thinking.

3. Repeat the process once more by selecting a random sample of ten numbers from Bag C.

Selection	1	2	3	4	5	6	7	8	9	10
Bag C										

a. Create a dot plot of your sample of ten numbers. Use a dot to represent each of the numbers in the sample.

b. Based on your dot plot, do you think the mean of the numbers in Bag C is the same or different from the mean of the numbers in Bag A? Explain your thinking.

4. Are your dot plots of the three bags the same as the dot plots of other students in your class? Why or why not?

5. Calculate the mean of the numbers for each of the samples from Bag A, Bag B, and Bag C.

	Mean of the sample of numbers
Bag A	
Bag B	
Bag C	

   a. Are the sample means you calculated the same as the sample means of other members of your class? Why or why not?

   b. How do your sample means for Bag A and for Bag B compare?

   c. Calculate the difference of sample mean for Bag A minus sample mean for Bag B ($\text{Mean}_{\text{A}} - \text{Mean}_{\text{B}}$). Based on this difference, can you be sure which bag has the larger population mean? Why or why not?

6. Based on the class dot plots of the sample means, do you think the mean of the numbers in Bag A and the mean of the numbers in Bag B are different? Do you think the mean of the numbers in Bag A and the mean of the numbers in Bag C are different? Explain your answers.

7. Based on the difference between the sample mean of Bag A and the sample mean of Bag B ($\text{Mean}_\text{A} - \text{Mean}_\text{B}$) that you calculated in Exercise 5, do you think that the two populations (Bags A and B) have different means, or do you think that the two population means might be the same?

8. Based on this difference, can you be sure which bag has the larger population mean? Why or why not?

9. Is your difference in sample means the same as your neighbors' differences? Why or why not?

10. Plot your difference of the means ($\text{Mean}_\text{A} - \text{Mean}_\text{B}$) on a class dot plot. Describe the distribution of differences plotted on the graph. Remember to discuss center and spread.

11. Why are the differences in the sample means of Bag A and Bag B not always 0?

12. Does the class dot plot contain differences that were relatively far away from 0? If yes, why do you think this happened?

13. Suppose you will take a sample from a new bag. How big would the difference in the sample mean for Bag A and the sample mean for the new bag ($\text{Mean}_{\text{A}} - \text{Mean}_{\text{new}}$) have to be before you would be convinced that the population mean for the new bag is different from the population mean of Bag A? Use the class dot plot of the differences in sample means for Bags A and B (which have equal population means) to help you answer this question.

The differences in the class dot plot occur because of sampling variability—the chance variability from one sample to another. In Exercise 13, you were asked about how great the difference in sample means would need to be before you have convincing evidence that one population mean is larger than another population mean. A "*meaningful*" difference between two sample means is one that is unlikely to have occurred by chance if the population means are equal. In other words, the difference is one that is greater than would have been expected just due to sampling variability.

14. Calculate the sample mean of Bag A minus the sample mean of Bag C ($\text{Mean}_{\text{A}} - \text{Mean}_{\text{C}}$).

15. Plot your difference ($\text{Mean}_{\text{A}} - \text{Mean}_{\text{C}}$) on a class dot plot.

16. How do the centers of the class dot plots for $\text{Mean}_{\text{A}} - \text{Mean}_{\text{B}}$ and $\text{Mean}_{\text{A}} - \text{Mean}_{\text{C}}$ compare?

17. Each bag has a population mean that is either 10.5 or 14.5. State what you think the population mean is for each bag. Explain your choice for each bag.

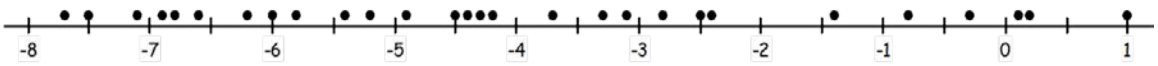
Lesson Summary

- Remember to think about sampling variability—the chance variability from sample to sample.
- Beware of making decisions based just on the fact that two sample means are not equal.
- Consider the distribution of the difference in sample means when making a decision.

Problem Set

Below are three dot plots. Each dot plot represents the differences in sample means for random samples selected from two populations (Bag A and Bag B). For each distribution, the differences were found by subtracting the sample means of Bag B from the sample means of Bag A (sample mean A − sample mean B).

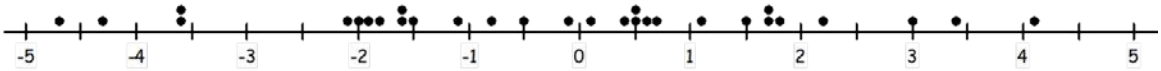
1. Does the graph below indicate that the population mean of Bag A is larger than the population mean of Bag B? Why or why not?



Sample Mean A − Sample Mean B

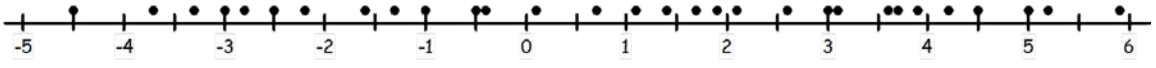
2. Use the graph above to estimate the difference in the population means (Mean A − Mean B).

3. Does the graph below indicate that the population mean of Bag A is larger than the population mean of Bag B? Why or why not?



Sample Mean A − Sample Mean B

4. Does the graph below indicate that the population mean of Bag A is larger than the population mean of Bag B? Why or why not?



Sample Mean A − Sample Mean B

5. In the above graph, how many differences are greater than 0? How many differences are less than 0? What might this tell you?

6. In Problem 4, the population mean for Bag A is really larger than the population mean for Bag B. Why is it possible to still get so many negative differences in the graph?