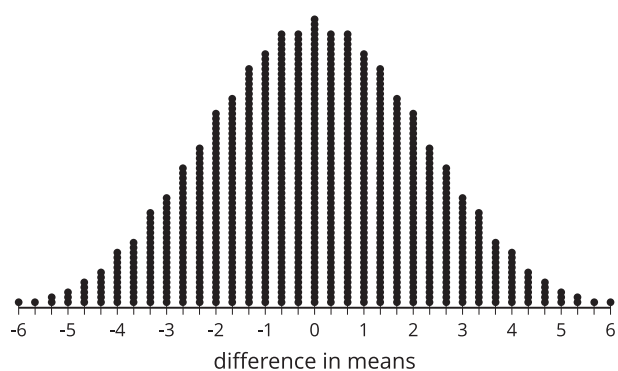
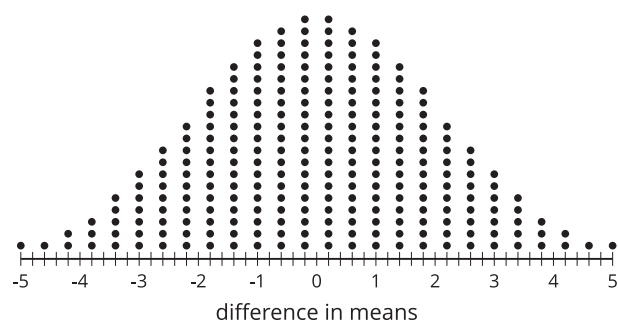
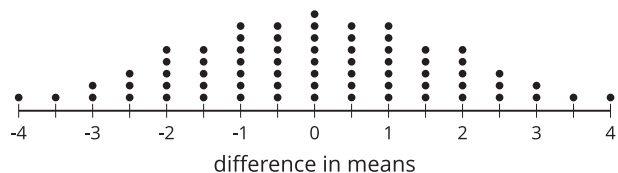
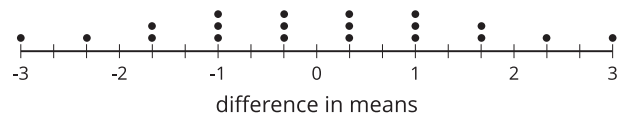


Unit 7 Lesson 14: Using Normal Distributions for Experiment Analysis

1 Notice and Wonder: Some Distributions (Warm up)

Student Task Statement

What do you notice? What do you wonder?



2 A Theoretical Experiment

Student Task Statement

To see what might be happening when we regroup data, consider an experiment that takes 12 subjects and divides them into 2 groups at random. The control group contains 6 subjects and the treatment group contains 6 subjects. To explore what's possible, assume the control group results in the data: 1, 3, 4, 6, 8, and 10. The treatment group results in the data: 2, 5, 7, 9, 11, and 12.

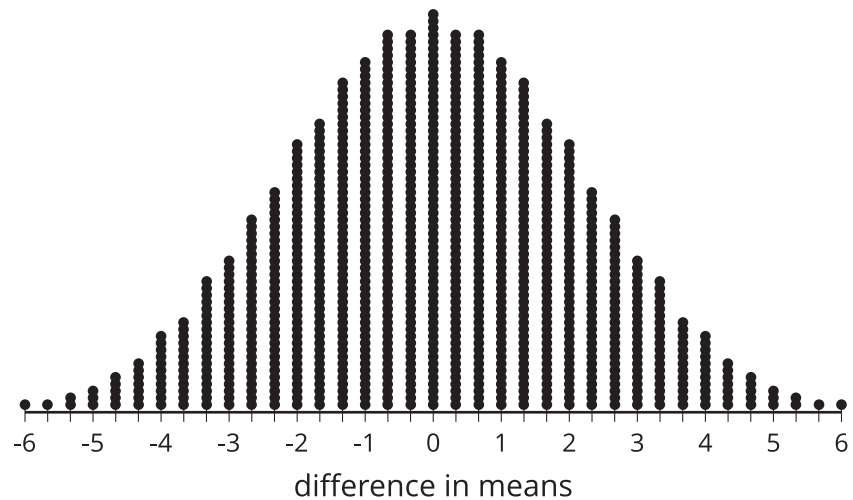
1. Find the difference in means for the original groups by subtracting the control group mean from the treatment group mean.
2. With a smaller data set like this, we can actually consider all of the different arrangements of the data. There are 924 distinct ways to separate the 12 values into 2 groups of 6. The frequency table shows all the possible differences in means and how often they occur. Notice that a difference in means of 4.33 occurs 7 times and a difference of -4.33 also occurs 7 times. The dot plot shows the same information.

What proportion of possible groupings have a difference at least as great as the difference in means for the original groups? Explain or show your reasoning.

difference in means	± 6	± 5.67	± 5.33	± 5	± 4.67	± 4.33	± 4
frequency	1	1	2	3	5	7	11

difference in means	± 3.67	± 3.33	± 3	± 2.67	± 2.33	± 2
frequency	13	18	22	28	32	39

difference in means	± 1.67	± 0.33	± 1	± 0.67	± 0.33	0
frequency	42	48	51	55	55	58



3. The proportion you calculate represents the probability that the original difference in means could be due to the groupings themselves. Based on the proportion you calculated for this situation, which description is most accurate? Explain your reasoning.
- Because the proportion is so low, it is unlikely that the difference in means is due to the randomized groupings. This means that the difference in means is most likely caused by the treatment.
 - Because the proportion is not that low, it is still rather possible that the original difference in means is due to the random groupings. This means that there is not enough evidence to determine that the difference in means is likely caused by the treatment.

3 Simulating to Decide

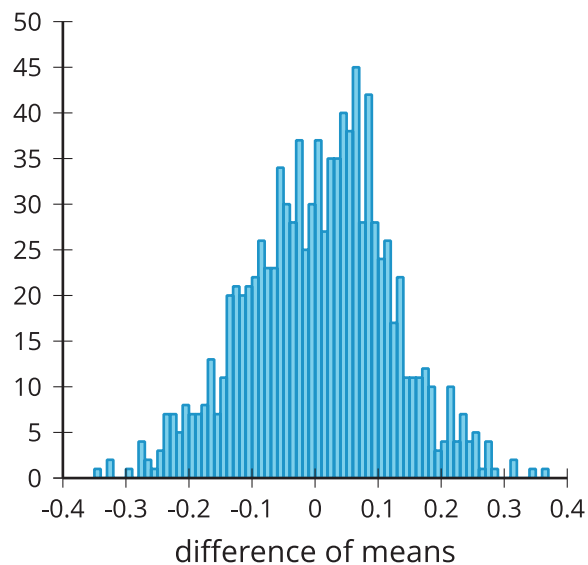
Student Task Statement

Researchers want to know the effect of captively raising birds on the weight of the birds. The researchers begin with 100 birds divided into 2 groups of 50 each. One group of 50 will be raised in captivity and the other 50 are tagged and released into the wild. After 5 years, all 100 birds are collected and weighed.



There are more than 10^{29} different ways to regroup the 100 birds into groups of 50 again, so looking at all the combinations would be too time consuming to reproduce. In this case, we can run simulations to determine how the original difference in means compares to those from regrouping the data.

The original groups have a difference of means of 0.27 grams. Researchers run 1,000 simulations regrouping the data into 2 groups at random and record the differences in means for the groups in each simulation. The histogram shows the differences in means from the simulations.



They determine that the mean of the differences of means from the simulations is 0.0021 grams and the standard deviation for the differences of means from the simulations is 0.112 grams.

1. What features of the distribution in the histogram let you know that modeling with a normal distribution is reasonable?
2. Model the simulations using a normal distribution with a mean of 0.0021 and a standard deviation of 0.112. What is the area under this normal curve that is more extreme than 0.27?
3. How can this area be used to compare the difference of means from the simulations to the difference of means from the original groups?

4. Based on the area under the normal curve, is there evidence that the original difference in means is likely due to where the birds spent the 5 years? Explain your reasoning.