Lesson 25: Ruling Out Chance

Student Outcomes

- Given data from a statistical experiment with two treatments, students create a randomization distribution.
- Students use a randomization distribution to determine if there is a significant difference between two treatments.

Lesson Notes

MP.4

In Lessons 25–27, students will be introduced to *randomization testing*. By the end of the three lessons, they will have experienced and executed all steps of this procedure, and they will have engaged in a "start to finish" example with provided data. Once that is accomplished, students will participate in a capstone experience in Lessons 28 and 29 where they will collect their own data, execute their own randomization test, and present their findings in a poster.

Note that randomization testing is somewhat different from more traditional methods for comparing the means of two groups such as the two-sample-t-test and other procedures found in many college-level statistics courses or AP statistics. While a randomization test still requires developing competing hypotheses (null and alternative) and coming to a conclusion based on the probability of obtaining results as extreme as those one obtains in one's sample or experiment, the randomization testing method develops the probability distribution of the statistic of interest via simulation and using the data at hand. As a result, the randomization testing method does not carry with it the usual concerns regarding the characteristics of a population distribution and other assumptions required of traditional methods. The technique is thus also referred to as a *nonparametric* technique.

This lesson starts by reminding students of previous lessons' work regarding randomization distributions. Specifically, students will review that when a single group of observations with any variability is randomly divided into two groups, the means of these two groups will tend to differ just by chance. In some cases, the difference in the means of these two groups may be very small (or "0"), but in other cases, this difference may be quite large. As the difference in the two groups' means is our statistic of interest, students will consider that difference value in context with an experiment regarding the use of a nutrient supplement to encourage tomato growth.

In subsequent lessons, students will estimate a probability distribution for the possible values of this "difference value," and students will assess whether or not the actual difference in means obtained from the experiment (in this case, the difference observed between the treatment group's mean weight and the control group's mean weight) is consistent with usual chance behavior. When the observed difference is not typical of chance behavior, it may indicate a significant difference. Specifically, it may indicate that the nutrient treatment was effective.

Throughout these lessons, experimental design principles are presented, and some questions will prompt students to consider the reasoning behind the principles.

Lesson 27 may require slightly more time than Lessons 25 and 26.

Classwork

Explain that the next three lessons explore some examples of the method called *randomization testing.* Then, after completing these three lessons, students will conduct an experiment and use this randomization testing method to assess if the result observed in the experiment is statistically significant.

Opening Exercise (3 minutes)

Let students answer part (a) in writing and share their responses with a neighbor. Pose the question in part (b) to the class and allow multiple responses.

Opening Exercise

a. **Explain in writing what you learned about randomly divided groups from the last lesson. Share your thoughts with a neighbor.**

In the previous lesson, I saw that when a single group of observations with any variability is randomly divided into two groups, the means of these two groups will tend to differ just by chance. In some cases, the difference in the means of these two groups may be very small (or "0"), but in other cases, this difference may be quite large.

b. **How could simulation be used to understand the "typical" differences between randomized groups?**

Through simulation, you can learn about the values of the difference that would be expected just by chance. You can then use this information to decide if a given set of data from an experiment support a claim of a statistically significant difference between two treatment means.

Explain that you do this by assessing whether or not the actual difference in means obtained from the experiment (such as the difference observed between a control group mean and a treatment group mean or between two treatment groups' means) is consistent with usual chance behavior. When the observed difference is not typical of chance behavior, it may indicate a significant difference in treatment means.

Exercises 1–3 (10 minutes): Random Assignments and Computing the Difference of the Group Means

Read through the example. Let students work independently on the exercises and confirm answers as a class.

Scaffolding:

- Consider using a visual or physical model to explain the experiment clearly.
- Consider asking students to restate, write, or draw a picture of the experiment.
- For advanced learners, consider posing the question: Suppose the nutrient group weighed, on average, 0.25 pounds more than the non-nutrient group. Would that be sufficient evidence that the nutrients caused a difference in weight? Explain why or why not.

Exercises 1–3: Random Assignments and Computing the Difference of the Group Means

Imagine that 10 tomatoes of varying shapes and sizes have been placed in front of you. These 10 tomatoes (all of the same variety) have been part of a nutrient experiment where the application of the nutrient is expected to yield larger tomatoes that weigh more. All 10 tomatoes have been grown under similar conditions regarding soil, water, and sunlight, but 5 of the tomatoes received the additional nutrient supplement. Using the weight data of these 10 tomatoes, you wish to examine the claim that the nutrient yields larger tomatoes on average.

1. Why would it be important in this experiment for the 10 tomatoes to all be of the same variety and grown under the same conditions (except for the treatment applied to 5 of the tomatoes)?

   Our variable is the weight of the tomatoes, and our objective is to see if the nutrient treatment yields heavier tomatoes. We don't want any other factors in our experiment that might cause some tomatoes to grow better than others; their presence might distort our findings. So, we prefer that the tomatoes all have similar growing conditions such as the same soil, the same amount of water, and the same amount of sunlight.

   Note: When we examine data collected from an experiment, we assume that the experiment was conducted in a reasonable way with respect to important experimental design principles such as proper control measures. In most cases, if these principles aren't incorporated, we may come to inaccurate conclusions.

Here are the 10 tomatoes with their weights shown. They have been ordered from largest to smallest based on weight.

9.1	8.4	8.0	7.7	7.3	6.4	5.9	5.2	4.4	3.8
#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
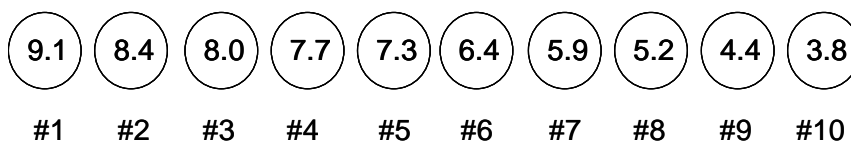
For now, do not be concerned about which tomatoes received the additional nutrients. The object here is to randomly assign the tomatoes to two groups.

Imagine that someone assisting you uses a random-number generator or some other impartial selection device and randomly selects Tomatoes #1, 4, 5, 7, and 10 to be in Group A. By default, Tomatoes #2, 3, 6, 8, and 9 will be in Group B. The result is illustrated below.

Group A	Group B
9.1	8.4
7.7	8.0
7.3	6.4
5.9	5.2
3.8	4.4
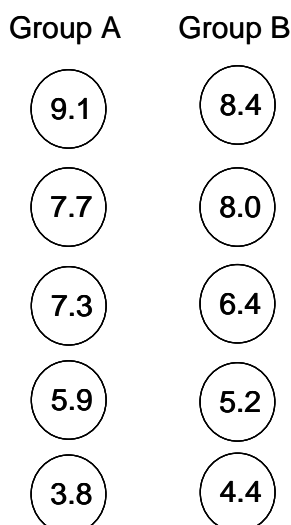
2. Confirm that the mean for Group A is 6.76 ounces, and calculate the mean for Group B.

    The mean for Group B is 6.48 ounces.

3. Calculate the difference between the mean of Group A and the mean of Group B (that is, calculate $\bar{x}_A - \bar{x}_B$).

    $6.76 - 6.48 = 0.28$ *ounces.*

Exercises 4–6 (5-7 minutes): Interpreting the Value of a Difference

In these exercises, students interpret the difference between the group means. Let students work independently on the exercises and confirm answers as a class.

Exercises 4–6: Interpreting the Value of a Difference

The statistic of interest that you care about is the difference between the mean of the 5 tomatoes in Group A and the mean of the 5 tomatoes in Group B. For now, call that difference "Diff." "Diff" = $\bar{x}_A - \bar{x}_B$.

4. Explain what a "Diff" value of "1.64 ounces" would mean in terms of which group has the larger mean weight and the number of ounces by which that group's mean weight exceeds the other group's mean weight.

    Group A has a mean weight that is 1.64 ounces higher than Group B's mean weight.

5. Explain what a "Diff" value of "-0.4 ounces" would mean in terms of which group has the larger mean weight and the number of ounces by which that group's mean weight exceeds the other group's mean weight.

    Group A has a mean weight that is 0.4 ounces lower than Group B's mean weight.

6. Explain what a "Diff" value of "0 ounces" would mean regarding the difference between the mean weight of the 5 tomatoes in Group A and the mean weight of the 5 tomatoes in Group B.

    Group A's mean weight is the same as Group B's mean weight.

Exercises 7–8 (10 minutes): Additional Random Assignments

Let students work independently on the exercises and confirm answers as a class.

Exercises 7–8: Additional Random Assignments

7. Below is a second random assignment of the 10 tomatoes to two groups. Calculate the mean of each group, and then calculate the value of "Diff" for this second case. Also, interpret the "Diff" value in context using your responses to the previous questions as a guide.

Group A   Group B

Group A	Group B
9.1	7.7
8.4	5.9
8.0	5.2
7.3	4.4
6.4	3.8

Group A has a mean weight of 7.84 ounces, and that is 2.44 ounces higher than Group B's mean weight (5.4 ounces).

8. Here is a third random assignment of the 10 tomatoes. Calculate the mean of each group, and then calculate the value of "Diff" for this case. Interpret the "Diff" value in context using your responses to the previous questions as a guide.

Group A   Group B

Group A	Group B
9.1	8.4
7.7	8.0
7.3	6.4
5.2	5.9
3.8	4.4

There is no difference between Group A's mean weight and Group B's mean weight; each mean is 6.62 ounces.

T

Closing (10 minutes)

MP.3

- Using the 10 tomatoes, what arrangement of tomatoes would yield the largest value of "Diff"? How would you explain your reasoning without performing calculations to verify your answer? How big is the largest value of "Diff"?

    □ *Sample Response: The largest "Diff" value would occur when Group A had the 5 largest tomatoes. That would create a situation where the mean of Group A has the largest possible mean for any grouping of the 5 tomatoes, and, consequently, Group B would have the smallest possible mean for any grouping of the 5 tomatoes. The "Diff" value would be $8.10 - 5.14 = 2.96$ ounces.*

- Using your logic above, what arrangement of tomatoes would yield the most negative value of "Diff"? How would you explain your reasoning without performing calculations to verify your answer? What is this "Diff" value?

    □ *Sample Response: The most negative "Diff" value would occur in the exact opposite case from above – that is when Group A had the 5 smallest tomatoes. That would create a situation where the mean of Group A has the smallest possible mean for any grouping of the 5 tomatoes, and, consequently, Group B would have the largest possible mean for any grouping of the 5 tomatoes. The "Diff" $value = 5.14 - 8.10 = -2.96$ ounces.*

- What would it mean for the mean of all "Diff" values to be close to zero?

    □ *Sample Response: A mean close to zero would indicate that the tomatoes in each group were mixed, or neither group had only the largest or smallest values. The resulting differences were both above and below the mean of the differences. Use Exercise 8 as a way to highlight the answer to this question.*

- Ask students to summarize the main ideas of the lesson in writing or with a neighbor. Use this as an opportunity to informally assess comprehension of the lesson. The Lesson Summary below offers some important ideas that should be included.

> **Lesson Summary**
>
> In this lesson, when the single group of observations was randomly divided into two groups, the means of these two groups differed by chance. These differences have a context based on the purpose of the experiment and the units of the original observations.
>
> The differences varied. In some cases, the difference in the means of these two groups was very small (or "0"), but in other cases, this difference was larger. However, in order to determine which differences were typical and ordinary vs. unusual and rare, a sense of the center, spread, and shape of the distribution of possible differences is needed. In the following lessons, you will develop this distribution by executing repeated random assignments similar to the ones you saw in this lesson.

Exit Ticket (7 minutes)

Name _____ Date _____

Lesson 25: Ruling Out Chance

Exit Ticket

Six ping-pong balls are labeled as follows: $0, 3, 6, 9, 12, 18$. Three ping-pong balls will be randomly assigned to Group A; the rest will be assigned to Group B. "Diff" = $\bar{x}_A - \bar{x}_B$.

1. Calculate "Diff," the difference between the mean of the numbers on the balls assigned to Group A and the mean of the numbers on the balls assigned to Group B (i.e., $\bar{x}_A - \bar{x}_B$), when the 3 ping-pong balls selected for Group A are 3, 6, and 12.

2. Calculate "Diff," the difference between the mean of the numbers on the balls assigned to Group A and the mean of the numbers on the balls assigned to Group B (i.e., $\bar{x}_A - \bar{x}_B$), when the 3 ping-pong balls selected for Group A are 3, 12, and 18.

3. What is the greatest possible value of "Diff", and what selection of ping-pong balls for Group A corresponds to that value?

4. What is the smallest (most negative) possible value of "Diff", and what selection of ping-pong balls for Group A corresponds to that value?

5. If these 6 observations represent the burn times of 6 candles (in minutes), explain what a "Diff" value of "6" means in terms of (a) which group (A or B) has the longer average burn time and (b) the amount of time by which that group's mean exceeds the other group's mean.

Exit Ticket Sample Solutions

> Six ping-pong balls are labeled as follows: $0, 3, 6, 9, 12, 18$. Three ping-pong balls will be randomly assigned to Group A; the rest will be assigned to Group B. "Diff" = $\bar{x}_A - \bar{x}_B$.
>
> 1. Calculate "Diff," the difference between the mean of the numbers on the balls assigned to Group A and the mean of the numbers on the balls assigned to Group B (i.e., $\bar{x}_A - \bar{x}_B$), when the 3 ping-pong balls selected for Group A are 3, 6, and 12.
>
> $7 - 9 = -2$
>
> 2. Calculate "Diff," the difference between the mean of the numbers on the balls assigned to Group A and the mean of the numbers on the balls assigned to Group B (i.e., $\bar{x}_A - \bar{x}_B$), when the 3 ping-pong balls selected for Group A are 3, 12, and 18.
>
> $11 - 5 = 6$
>
> 3. What is the greatest possible value of "Diff", and what selection of ping-pong balls for Group A corresponds to that value?
>
>   *If 9, 12, and 18 (the three highest numbers) are selected for Group A, the "Diff" would be $13 - 3 = 10$.*
>
> 4. What is the smallest (most negative) possible value of "Diff", and what selection of ping-pong balls for Group A corresponds to that value?
>
>   *If 0, 3, and 6 (the three lowest numbers) are selected for Group A, the "Diff" would be $3 - 13 = -10$.*
>
> 5. If these 6 observations represent the burn times of 6 candles (in minutes), explain what a "Diff" value of "6" means in terms of (a) which group (A or B) has the longer average burn time and (b) the amount of time by which that group's mean exceeds the other group's mean.
>
>   *Group A has a mean burn time that is 6 minutes higher (longer) than Group B's mean burn time.*

Problem Set Sample Solutions

Six ping-pong balls are labeled as follows: $0, 3, 6, 9, 12, 18$. Three ping-pong balls will be randomly assigned to Group A; the rest will be assigned to Group B. "Diff" = $\bar{x}_A - \bar{x}_B$.

In the Exit Ticket problem, 4 of the 20 possible randomizations have been addressed.

1. Develop the remaining 16 possible random assignments to two groups, and calculate the "Diff" value for each.

    (Note: Avoid redundant cases; selecting "0, 3, and 6" for Group A is NOT a distinct random assignment from selecting "6, 0, and 3" – so do not record both!)

    Note: Reference notes on the far right of the table below refer to the corresponding question in the "Exit Ticket."

Group A Selection			"A" Mean	"B" Mean	"Diff"	
0	3	6	3	13	-10	Question #4
0	3	9	4	12	-8	
0	3	12	5	11	-6	
0	3	18	7	9	-2	
0	6	9	5	11	-6	
0	6	12	6	10	-4	
0	6	18	8	8	0	
0	9	12	7	9	-2	
0	9	18	9	7	2	
0	12	18	10	6	4	
3	6	9	6	10	-4	
3	6	12	7	9	-2	Question #1
3	6	18	9	7	2	
3	9	12	8	8	0	
3	9	18	10	6	4	
3	12	18	11	5	6	Question #2
6	9	12	9	7	2	
6	9	18	11	5	6	
6	12	18	12	4	8	
9	12	18	13	3	10	Question #3

2. Create a dot plot that shows the 20 "Diff" values obtained from the 20 possible randomizations. By visual inspection, what is the mean and median value of the distribution?

    The mean and median appear to be "0" due to the center and symmetry.

3. Based on your dot plot, what is the probability of obtaining a "Diff" value of "8 or higher"?

    Two of the 20 values are "8 or higher." $\frac{2}{20} = 0.10 = 10\%$.

4. Would a "Diff" value of "8 or higher" be considered a difference that is likely to happen or one that is unlikely to happen? Explain.

   A "Diff" value of "8 or higher" would be considered unlikely to happen since a difference this large would happen only 10% of the time.

5. Based on your dot plot, what is the probability of obtaining a "Diff" value of "−2 or smaller"?

   Nine of the 20 values are "−2 or smaller." $\dfrac{9}{20} = 0.45 = 45\%.$

6. Would a "Diff" value of "−2 or smaller" be considered a difference that is likely to happen or one that is unlikely to happen? Explain.

   A "Diff" value of "−2 or smaller" would be considered likely to happen since a difference this small would happen 45% of the time.