

Lesson 15: Questioning Experimenting

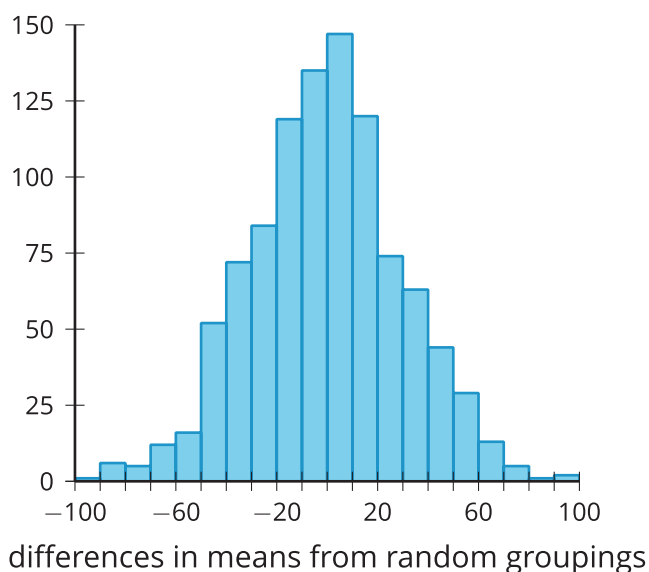
Let's ask the right questions to analyze data from an experiment.

15.1: Is It the Treatment?

A scientist divides 30 strawberry plants into two groups at random. One group of 15 plants will represent the control group and is grown in standard greenhouse conditions. The second group of 15 plants will represent the treatment group and will grow under the same conditions except they are grown in a different type of soil. After 6 weeks, the total weight (in grams) of the strawberries are measured for each plant. The scientist then performs a randomized experiment to compare the groups.

The data are summarized by these statistics and histogram.

- Mean for the control group: 238.67 grams
- Mean for the group with different soil: 347.47 grams
- Mean of differences in means from randomized groupings: -0.540 grams
- Standard deviation of differences in means from randomized groupings: 29.83 grams



Is there evidence that the difference in means from the original groupings is due to the different soil or is it likely that the difference is due to the way the plants were grouped? Explain your reasoning.

15.2: Info Gap: Is There a Difference?

Your teacher will give you either a problem card or a data card. Do not show or read your card to your partner.

If your teacher gives you the data card:

1. Silently read the information on your card.
2. Ask your partner “What specific information do you need?” and wait for your partner to ask for information. Only give information that is on your card. (Do not figure out anything for your partner!)
3. Before telling your partner the information, ask “Why do you need to know (that piece of information)?”
4. Read the problem card, and solve the problem independently.
5. Share the data card, and discuss your reasoning.

If your teacher gives you the problem card:

1. Silently read your card and think about what information you need to answer the question.
2. Ask your partner for the specific information that you need.
3. Explain to your partner how you are using the information to solve the problem.
4. When you have enough information, share the problem card with your partner, and solve the problem independently.
5. Read the data card, and discuss your reasoning.

Pause here so your teacher can review your work. Ask your teacher for a new set of cards and repeat the activity, trading roles with your partner.

15.3: Using Tables for Normal Distribution Areas

A factory produces baseballs. The weights of the baseballs produced are approximately normally distributed with a mean weight of 145 grams and a standard deviation of 2 grams. Official rules require the balls to weigh between 142 and 149 grams.

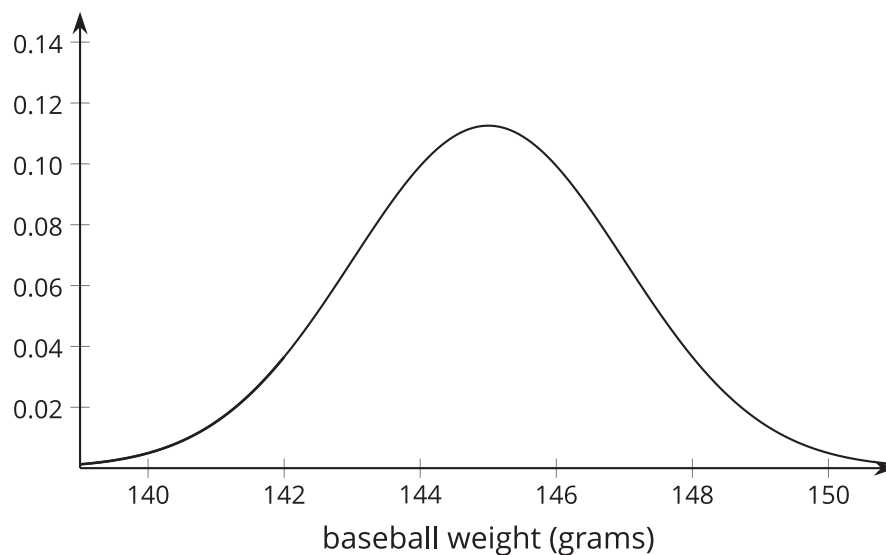
Recall that the proportion of items in an interval of an approximately normally distributed situation is the same as the area under the normal curve. A table can be used to determine the area under a normal curve bounded by an interval.

First, the relevant values need to be converted to a *z-score*. A value's *z-score* represents the number of standard deviations it is above the mean. In the baseball example, the value 147 grams has a *z-score* of 1 since it is 1 standard deviation above the mean. The value 140 grams has a *z-score* of -2.5 since it is 2.5 standard deviations below the mean.

In general, a z-score can be found using

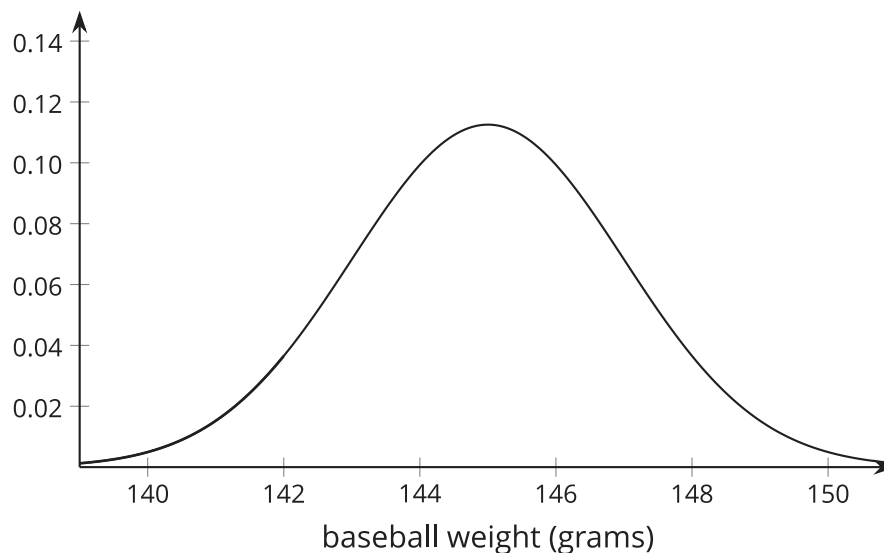
$$z = \frac{\text{value} - \text{mean}}{\text{standard deviation}}$$

1. Find the z-score for 142 grams.
2. Find the z-score for 149 grams.
3. What value would have a z-score of 1.45?
4. The table gives the area under the normal curve that is less than the given value. Shade the region that is given by the table for the area related to 142 grams.



5. Use the z-score for 142 grams and the table to find the area under the normal curve that is less than 142 grams.

6. Shade the region that is given by the table for the area related to 149 grams.



7. Use the z-score for 149 grams and the table to find the area under the normal curve that is less than 149 grams.

8. Use the two areas to find the area under the normal curve between 142 and 149 grams. Explain or show your reasoning.

9. What proportion of the baseballs that the factory makes are within the official rules?

Are you ready for more?

There are 2 different distributions. Distribution A has a mean of 55 and a standard deviation of 8. Distribution B has a mean of 6 and a standard deviation of 1.6.

From distribution A, a person is interested in a value of 70 and from distribution B, the person is interested in a value of 10. How can z-scores be used to determine which is more relatively extreme?

Lesson 15 Summary

After collecting data from an experiment, it is important to analyze the data to determine whether there is evidence that the difference in means for the groups is due to the treatment or whether the difference might be explained by the random groupings. There are several things that an experimenter needs to know to determine the possible cause of the differences.

First, the difference in the means for the two groups is important to know. Then the difference can be compared to the differences in means collected from regrouping the data into groups at random. The proportion of differences in means that are more extreme than the original difference can help determine how likely it is that the original difference was due to the random grouping.

The proportion can be determined either from counting the actual number of differences that are more extreme or modeling the differences with a normal distribution.