

Cents and the Central Limit Theorem

How the sampling distribution of the mean of a nonnormal distribution looks normal.

Many of the variables that you have studied so far in your statistics class have had a normal distribution. You may have used a table of the normal distribution to answer questions about a randomly selected individual or a random sample taken from a normal distribution. Many distributions, however, are not normal or any other standard shape.

Question

If the shape of a distribution isn't normal, or even symmetrical, what can we infer about its mean from a random sample?

Objectives

In this activity, you will experience the consequences of the Central Limit Theorem as you observe the shape, mean, and standard deviation of the sampling distribution of the mean for random samples taken from a distribution that is decidedly not normal.

Prerequisites

You should know that the mean of a distribution is its "balance point." Also, you should understand the standard deviation as a measure of spread.

Activity

1. You should have a list of the dates from your sample of 25 pennies. Next to each date, write the *age* of the penny by subtracting the date from the current year. What do you think the shape of the distribution of all the ages of the pennies from students in your class will look like?
2. Make a histogram of the ages of all the pennies in the class.
3. Estimate the mean and the standard deviation of the distribution. Confirm these estimates by actual computation.
4. Take a random sample of size 5 from the ages of your pennies and compute the mean age of your sample. Three or four students in your class should place their sample means on a number line.
5. Do you think the mean of the values in this histogram (once it is completed) will be larger than, smaller than, or the same size as the one for the population of all pennies? Regardless of which you choose, try to make an argument to support each choice. Estimate what the standard deviation of this distribution will be.
6. Complete the histogram and determine its mean and standard deviation. Which of the three choices in step 5 appears to be correct?
7. Repeat this experiment for samples of size 10 and 25.
8. Look at the four histograms that your class has constructed. What can you say about the shape of the histogram as n increases? What can you say about the center of the histogram as n increases? What can you say about the spread of the histogram as n increases?

Wrap-Up

1. The three characteristics you examined in “Activity” step 8 (shape, center, and spread of the sampling distribution) make up the *Central Limit Theorem*. Without looking in a textbook, write a statement of what you think the Central Limit Theorem says.
2. The distributions you constructed for samples of size 1, 5, 10, and 25 are called “sampling distributions of the sample mean.” Sketch the sampling distribution of the sample mean for samples of size 36.

Extensions

1. Get a copy of the current *Handbook of United States Coins: Official Blue Book of United States Coins* and graph the distribution of the number of pennies minted in each year. What are the interesting features of this distribution? Compare this distribution with the distribution of the population of ages of the pennies in the class. How are they different? Why? How can you estimate the percentage of coins from year x that are out of circulation?
2. In a geometric distribution, the height of each bar of the histogram is a fixed fraction r of the height of the bar to the left of it. Except for the first bar, is the distribution of the ages of the pennies approximately geometric? Estimate the value of r for this distribution. What does r tell you about how pennies go out of circulation? Why is the height of the first bar shorter than one would expect in a geometric distribution?
3. To cut the standard deviation of the sampling distribution of the sample mean in half, what sample size would you need?

Assessment Questions

1. What is your best guess of the mean age of all nickels in circulation? Explain.
2. Approximately how far off might this guess be? Explain. List the assumptions you are making.