Aim #98: How do we describe the shape, center and spread of data using the standard deviation?

Kickoff: For the histograms shown below, state if the data is symmetrical, skewed left, or skewed right. Notice the relationship between the mean and median for each set.

| Data set A: Mean = 2.64, Median = 2.00 |
| Data set B: Mean = 4.36, Median = 5.00 |
| Data set C: Mean = 3.00, Median = 3.00 |

Recall!

- The standard deviation of a data set is the square root of the average squared deviation from the mean. It measures on average how spread out data is from the mean. It is a measure of variability.

According to the U.S. Census Bureau, the average female in her twenties is 5 feet 4 inches tall. If we took a large sample of twenty-year-old women and made a histogram of their heights, it would look like this:

What do you notice about the mean and the median?

about the same

If we connect the midpoints of each interval, we have the normal curve.

Data that models this curve is said to have a normal distribution.
A normal curve is a smooth curve that is symmetric and bell shaped. Data distributions that are mound shaped are often modeled using a normal curve, and we say that such a distribution is approximately normal.

For a normal curve:
- Approximately 68.2% of data falls within 1 standard deviation of the mean.
- Approximately 95.4% of data falls within 2 standard deviations of the mean.
- Approximately 99.8% of data falls within 3 standard deviations of the mean.

When a distribution is skewed, it is not appropriate to model the data distribution with a normal curve.
Mod Prob: A set of data is normally distributed with a mean of 40 and a standard deviation of 2.

a) What percent of the data values are less than 40?

b) What percent of the data values are between 38 and 42?

c) What percent of the data values are between 36 and 44?

d) What percent of the data values are less than or equal to 36?

1. On a standardized test, the test scores are normally distributed with a mean of 60 and a standard deviation of 6.

a) Of the data, 84% of the scores are at or below what score?

b) Of the data, 16% of the scores are at or below what score?

2. The mean score of normally distributed data is 80 and the standard deviation is 5.

a) What is the probability that a student had a test score above 90?

b) What is the probability that a student had a test score between 75 and 85?
Do You Speak Math?

z-score - the number of standard deviations a piece of data is from the mean

\[ z = \frac{x - \bar{x}}{s_x} \]

A positive z-score corresponds to a value above the mean.
A negative z-score corresponds to a value below the mean.

In a normal distribution with a mean of 50 and a standard deviation of 6, what is the z-score of:

a) 56  
b) 47  
c) 41  
d) 65

A population has a mean of 102.8 and a standard deviation of 15.4. If a data point has a z-value of 1.87 then what is the value of the data point?

In a normal distribution with a mean of 56 and a standard deviation of 8, what percent of the scores lie between:

a. 48-64  
b. 52-60  
c. 56-68
How can we do this on the calculator?

2nd VARS (DIST)

To calculate normal probabilities without z-scores:
normalcdf (left bound, right bound, \( \bar{x}, s_x \))

To calculate normal probabilities given z-scores:
normalcdf (low Z score, high Z score)

**Mod Prob:** The playing life of an iPod is normally distributed with a mean of 30,000 hours and a standard deviation of 500 hours. Savannah's iPod lasted for 31,250 hours. Her player lasted longer than what percent of other iPods?

This is called the percentile!

**Try:** The Brite Lites R Us Company manufactures lightbulbs. They advertise that the "average lightbulb" can burn for 1,000 hours. The times that the lights can burn are normally distributed with a standard deviation of 200 hours. What percent of the bulbs can be expected to last 700 hours or less?
The mean age of the entering freshman class of 1600 at a certain university is 18.5, with a standard deviation of 0.75 years. If the data produces a normal distribution, find

a. The percent of students who are between 19.25 and 17.75 years of age
b. The percentile of a student who is 20 years old
c. The number of students who are expected to be younger than 17 years old

Practice

1. The mean score of normally distributed data is 80 and the standard deviation is 5.
   a) What is the probability that a student had a test score above 90?
   b) What is the probability that a student had a test score between 75 and 85?

2. In the accompanying diagram, the shaded area represents approximately 95% of the scores on a standardized test with normally distributed results. If these scores ranged from 78 to 92, which could be the standard deviation?
   1) 3.5  2) 7.0  3) 14.0  4) 20.0
3. Each year, the College Board publishes the mean SAT score and the standard deviation for students taking the test. SAT scores are normally distributed. Assume for a group of students that the mean SAT score is 500 with a standard deviation of approximately 100 points.

a) Find a score that is one standard deviation above the mean

b) Find a score that is two standard deviations below the mean

c) What percentile would a student be in if he or she scored
   i) 500   ii) 450   iii) 750

d) In 2007, approximately 1,500,000 students took the SAT. Approximately how many of them would have been expected to above 750?

4. A small company has salaries for their 50 employees as given in the table below.

   (a) Find the mean and standard deviation of the salary range.

   (b) What is the median of this data set? Why is the median considerably lower than the mean in this data set?

   (c) Does more or less than 50% of the data set fall within one standard deviation of the mean? Show the analysis that leads to your answer.
5. Biologists are studying the weights of Red King Crabs in the Alaskan waters. They sample 16 crabs and compiled their weights, in pounds, as shown below.

9.8, 10.1, 11.1, 12.4, 11.8, 13.2, 12.8, 12.5, 13.7, 11.6, 13.4, 12.3, 12.6, 14.8, 14.2, 15.1

a) Determine the mean and sample standard deviation for this sample of crabs. Round both statistical measures to the nearest tenth of a pound.

b) Compare the median with the mean. What does this comparison indicate about how the data is distributed?

c) Assuming your mean and standard deviation from part (a) apply to a normally distributed population of crabs caught in Alaska, what percent will fall between 9.6 pounds and 15.6 pounds?

6. On a standardized test, the test scores are normally distributed with a mean of 60 and a standard deviation of 6. What is the z-score of a test grade of 48?

7. For a normal distribution of weights, the mean weight is 160 pounds and a weight of 186 pounds has a z-score of 2.
   a) What is the standard deviation of the set of data?
   b) What percent of the weights are between 155 and 165?

8. In a normally distributed set of data, the mean is 48 and the standard deviation is 6.
   a) What percent of the scores will lie between 42 and 54?
   b) What percent of the scores will lie between 40 and 49?
   c) What percentile is a score of 57?
9. A swimmer named Amy specializes in the 50-meter backstroke. In competition, her mean time for the event is 39.7 seconds, and the standard deviation of her times is 2.3 seconds. Assume that Amy’s times are approximately normally distributed.

a. Estimate the probability that Amy’s time is between 37 and 44 seconds.

b. Using a graphing calculator, find the probability that Amy’s time in her race is between 37 and 44 seconds.

c. Estimate the probability that Amy’s time is more than 45 seconds.

d. Using a graphing calculator, find the probability that Amy’s time in her next race is more than 45 seconds.

e. Using a graphing calculator, find the probability that Amy’s time in her next race is less than 36 seconds.

10. The prices of the printers in a store have a mean of $240 and a standard deviation of $50. The printer that you eventually choose costs $340.

   i. What is the z-score for the price of your printer?

   ii. How many standard deviations above the mean was the price of your printer?

11. The U.S. Department of Agriculture (USDA), in its Official Food Plans states that the average cost of food for a 14- to 18-year-old male (on the Moderate-Cost Plan) is $261.50 per month. Assume that the monthly food cost for a 14- to 18-year-old male is approximately normally distributed with a mean of $261.50 and a standard deviation of $16.25.

   Find the probability that the monthly food cost for a randomly selected 14- to 18-year-old male is

   i. less than $280

   ii. more than $270

   iii. more than $250

   iv. between $240 and $275

   v. Explain the meaning of the probability that you found in part (iv).

12. The average weight of full grown beef cows is 1470 pounds with a standard deviation of 230 pounds. If the weights are normally distributed, what is the percentile rank of a cow that weighs 1750 pounds?

13. A population has a mean of 102.8 and a standard deviation of 15.4. If a data point has a z-value of 1.87 then what is the value of the data point?

14. The data below shows the quiz grades for an AP History class at MHS. Using the data find the:

   a. mean
   b. median
   c. standard deviation
   d. What percent of the scores fall within one standard deviation of the mean?
   e. Is this data normally distributed?

<table>
<thead>
<tr>
<th>Time</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>55</td>
<td>1</td>
</tr>
<tr>
<td>60</td>
<td>2</td>
</tr>
<tr>
<td>65</td>
<td>8</td>
</tr>
<tr>
<td>70</td>
<td>9</td>
</tr>
<tr>
<td>75</td>
<td>12</td>
</tr>
<tr>
<td>90</td>
<td>10</td>
</tr>
</tbody>
</table>

   Apr 10-15 AM
13. The data below shows the quiz grades for an AP History class at MHS. Using the data find the:

a. mean  
b. median  
c. standard deviation  

d. What percent of the scores fall within one standard deviation of the mean?

e. Is this data normally distributed?

<table>
<thead>
<tr>
<th>Time</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>55</td>
<td>1</td>
</tr>
<tr>
<td>60</td>
<td>3</td>
</tr>
<tr>
<td>65</td>
<td>8</td>
</tr>
<tr>
<td>70</td>
<td>9</td>
</tr>
<tr>
<td>75</td>
<td>12</td>
</tr>
<tr>
<td>80</td>
<td>10</td>
</tr>
</tbody>
</table>