## **Measures of Variation**

## **Measures of Variation**

- We are in section 2.7 of the textbook
- Variation describes how widely a data set is spread out.
- The *range* is the difference between the maximum and minimum values, range = max min.
- The *standard deviation* uses all the data points to quantify spread. Calculation:

$$s = \sqrt{\frac{\sum (x - \overline{x})^2}{N - 1}}$$

- Compute the mean  $\bar{x}$
- Subtract the mean from each data point,  $x \bar{x}$
- Square each of the differences,  $(x \overline{x})^2$
- Sum all of the squares,  $\sum (x \overline{x})^2$
- Divide by N-1,  $\frac{\sum (x-\overline{x})^2}{N-1}$
- Take the square root
- The population standard deviation is computed by dividing by N, instead of N-1. Notation: sigma  $\sigma$
- Compare the standard deviations from two data sets with different scales using the *coefficient of variation*, the standard deviation divided by the mean.

$$\frac{s}{\overline{x}}$$
 or  $\frac{\sigma}{\mu}$ 

- Caution! If two data sets have different mean and standard deviation values, then comparing the data values directly can be misleading.
  - Compare data values by computing the number of standard deviations the value is from the mean.
- Chebyshev's Theorem: for any data set (yes, really any!)
  - At least 75% of the data is within two standard deviations of the mean.
  - At least 89% of the data is within three standard deviations of the mean.
  - At least 93.8% of the data is within four standard deviations of the mean.

1. Compute the mean, range, and sample standard deviation of the following data set. 8 12 15 16 20 Range: Mean: \_\_\_\_\_ Sample St. Dev: \_\_\_\_\_ 2. Compute the mean, range, and population standard deviation of the following data set. 40 70 74 80 90 100

Range: \_\_\_\_\_

Mean: \_\_\_\_\_

Pop. St. Dev:

3. Compute the coefficient of variation for the data set in problem 1.

4. Suppose we have a population with a mean of 250 and a standard deviation of 30. Use Chebyshev's theorem to find an interval around the mean that includes at least 75% of the population.

5. Suppose we have a population with a mean of 60 and a standard deviation of 8. Use Chebyshev's theorem to find an interval around the mean that includes at least 89% of the population.

6. Mutual Funds are investments that take money from a large number of investors, purchase stocks, and then distribute the gains or losses back to the investors. The following table shows the earnings, in percent, of two mutual funds. The fund 'JCBUX' is a JP Morgan fund focusing on bonds, while the fund 'RERGX' is an American Funds mutual fund focusing on European stocks.<sup>1</sup>

Fund	2014	2015	2016	2017	2018
JCBUX	5.41%	1.01%	2.51%	3.87%	0.35%
RERGX	-2.29%	-0.48%	1.01%	31.17%	-14.91%

(a) On separate paper, or using SPSS, compute the mean return and the standard deviation for each of the two funds. Fill in the following table. Round your answers to two decimal places.

Fund	Mean return	Std. Dev.
JCBUX		
RERGX		

(b) When you compare the mean returns, what do you see? Explain in one or two sentences.

(c) The coefficient of variation is one of many factors used by investors to evaluate mutual funds. Compute the coefficient of variation for each fund. When you compare these coefficients, what do you see? Explain in one or two sentences.

<sup>&</sup>lt;sup>1</sup>This data was collected from Yahoo Finance in September 2019. You can the ticker symbols, such as 'JCBUX' and 'RERGX', to search for almost any mutual fund.