# Measures of Dispersion

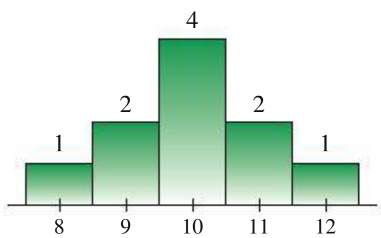
**Definition 1.** The range for a set of data is the difference between the largest and the smallest values in the data set. The range for a frequency distribution is the difference between the upper boundary of the highest class and the lower boundary of the lowest class.

**Example 1.** Find the range of the following data sets.

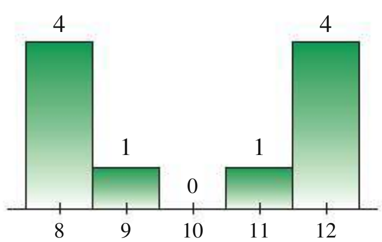
1. 54 70 50 16 92 82 24 58 76 23

2.

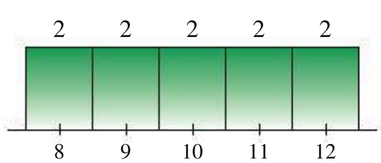
| Class Interval | Frequency |
|---|---|
| 14.5-19.5 | 2 |
| 19.5-24.5 | 4 |
| 24.5-29.5 | 5 |
| 29.5-34.5 | 6 |
| 34.5-39.5 | 3 |
| 39.5-44.5 | 1 |




(A) Mean = 10
Range = 12.5 − 7.5 = 5




(B) Mean = 10
Range = 5




(C) Mean = 10
Range = 5

**DEFINITION Variance: Ungrouped Data***

The **sample variance $s^2$** of a set of $n$ sample measurements $x_1, x_2, \ldots, x_n$ with mean $\bar{x}$ is given by

$$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n - 1} \tag{3}$$

If $x_1, x_2, \ldots, x_n$ is the whole population with mean $\mu$, then the **population variance $\sigma^2$** is given by

$$\sigma^2 = \frac{\sum_{i=1}^{n} (x_i - \mu)^2}{n}$$

*In this section, we restrict our interest to the sample variance.

**DEFINITION Standard Deviation: Ungrouped Data†**

The **sample standard deviation $s$** of a set of $n$ sample measurements $x_1, x_2, \ldots, x_n$ with mean $\bar{x}$ is given by

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n - 1}} \tag{4}$$

If $x_1, x_2, \ldots, x_n$ is the whole population with mean $\mu$, then the **population standard deviation $\sigma$** is given by

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \mu)^2}{n}}$$

†In this section, we restrict our interest to the sample standard deviation.

**Example 2.** Calculate the variance and standard deviation for the following set of data.

5 7 11 15 13 10 8 19

**Example 3.** Using the data set from the previous example, what percent of the data lies within 1 standard deviation of the mean? 2 standard deviations?

# Empirical Rule

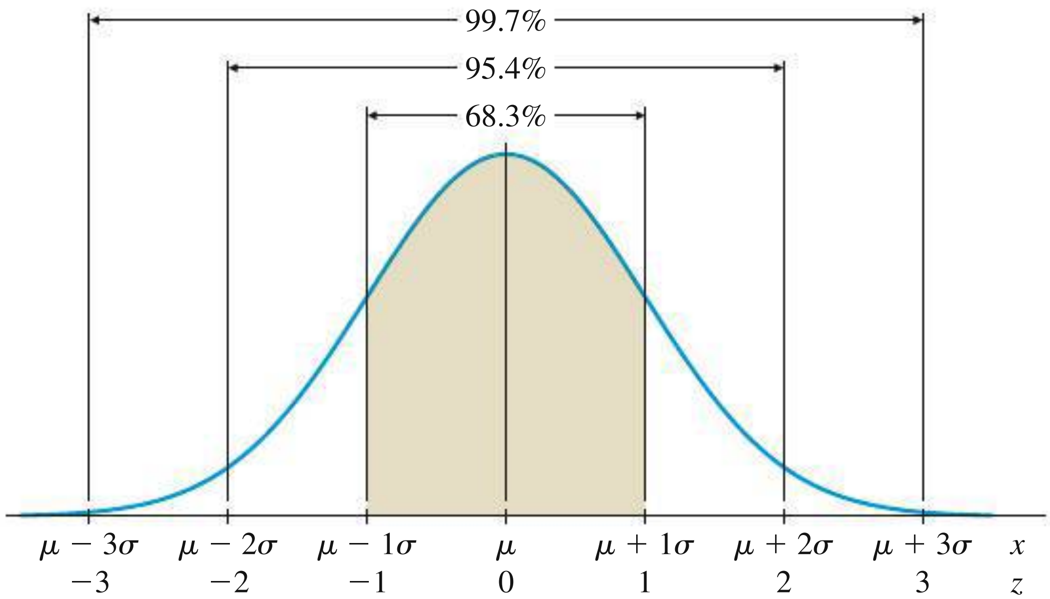



Figure 1: Symmetric distributions such that values near the mean are common and values far from the mean are rare are bell-shaped. This type of distriution appears naturally, and when it does we can approximate percentages of data that lie within intervals centered around the mean.

**Example 4.** In order to learn how much sleep American high school students are getting, researchers asked 15,000 randomly selected American high school students how much sleep they got last night. The researchers calculated the sample mean and sample standard deviation to be 7.6 hours and 0.5 hours, respectively. Assume the collected data has a bell-shaped distribution. Approximately what proportion (how many) of the students sampled

1. slept between 7.1 and 8.1 hours last night?
2. slept between 6.6 and 8.6 hours last night?
3. slept less than 7.1 hours?
4. slept between 7.6 and 8.1 hours?
5. slept between 6.1 and 6.6 hours?