9.2 Measures of Variation

Definition 1. Given a set of data $x_1, x_2, \ldots, x_n$, the **range** is the length of the smallest interval that contains all of the data. That is, the largest value minus the smallest value.

$$\text{range} = \max\{x_1, x_2, \ldots, x_n\} - \min\{x_1, x_2, \ldots, x_n\}$$

1. Find the range of the following data set.

$$20, \quad 21, \quad 21, \quad 26, \quad 21, \quad 19, \quad 20, \quad 19, \quad 24, \quad 20, \quad 22, \quad 22, \quad 22, \quad 18, \quad 27$$

2. Find the range of the following data set.

$$61.6332, \quad 79.4085, \quad 68.3606, \quad 55.4502, \quad 31.8612, \quad 40.5185, \quad 64.1433, \quad 88.1797$$

3. Show that the following two data sets have the same range.

$$\text{Data set } A = 1, 5, 5, 5, 5, 5, 5, 9$$
$$\text{Data set } B = 1, 1, 1, 1, 9, 9, 9, 9$$

Definition 2. Given a data set $x_1, x_2, \ldots, x_n$ with mean $\overline{x}$, the corresponding set of **deviations from the mean** is

$$x_1 - \overline{x}, \quad x_2 - \overline{x}, \quad \ldots, \quad x_n - \overline{x}$$

and the corresponding set of **deviations from the mean squared** is

$$(x_1 - \overline{x})^2, \quad (x_2 - \overline{x})^2, \quad \ldots, \quad (x_n - \overline{x})^2.$$

Definition 3. For a sample of n measurements $x_1, x_2, \ldots, x_n$ with sample mean $\overline{x}$, the **variance** of the sample is

$$s^2 = \frac{\sum (x - \overline{x})^2}{n - 1} \qquad \text{(classic)}$$
$$= \frac{\sum x^2 - n\overline{x}^2}{n - 1}, \qquad \text{(alternative)}$$

and the **standard deviation** of the sample is

$$s = \sqrt{s^2} = \sqrt{\frac{\sum (x - \overline{x})^2}{n - 1}} \qquad \text{(classic)}$$
$$= \sqrt{\frac{\sum x^2 - n\overline{x}^2}{n - 1}}. \qquad \text{(alternative)}$$

4. Given the following sample data,

$$20, \quad 22, \quad 17, \quad 21, \quad 25$$

calculate the variance and the standard deviation of the sample using *both* formulas given above.

5. Find the standard deviation for the sample data summarized in the table below.

Interval	Frequency
0-24	4
25–49	8
50–74	5
75 –99	10
100 –124	4
125 –149	5

Theorem 1 (Chebyshev's Theorem)**.** *Given any set of measurements $x_1, x_2, \ldots, x_n$, and any number $k > 0$, the proportion of measurements that lie within k standard deviations of the mean is at least*

$$1 - \frac{1}{k^2}.$$

6. Given a sample of 300 measurements with mean $\overline{x} = 72$ and standard deviation $s = 7$, use Chebyshev's Theorem to answer the following questions.

   (a) At least what proportion of measurements lie between 58 and 86?

   (b) At least how many measurements lie between 51 and 93?