MATH 3850   Fall 2025

Homework 1
Due Thursday, September 11

1. Given below are the average wind speeds (in miles per hour) for 45 selected U.S. cities:

```
 8.9  12.4   8.6  11.3   9.2  8.8  35.1   6.2  7.0
 7.1  11.8  10.7   7.6   9.1  9.2   8.2   9.0  8.7
 9.1  10.9  10.3   9.6   7.8 11.5   9.3   7.9  8.8
 8.8  12.7   8.4   7.8   5.7 10.5  10.5   9.6  8.9
10.2  10.3   7.7  10.6   8.3  8.8   9.5   8.8  9.4
```

(The value 35.1 was recorded at Mt. Washington, New Hampshire.)

(a) The average wind speed for Chicago is 10.3 miles per hour. What percentage of the cities have average wind speeds in excess of Chicago's? Do you think that Chicago is unusually windy and deserves "The Windy City" nickname?

(b) Calculate $\overline{Y}$ and S for the data given.

(c) Calculate the interval $\overline{Y} \pm k \cdot S$ for k= 1, 2, and 3. Count the number of measurements that fall within each interval and compare this result with the number that you would expect according to the empirical rule.

2. The top 40 stocks on the over-the-counter (OTC) market, ranked by percentage of outstanding shares traded on one day last year are as follows:

```
11.88 6.27 5.49 4.81 4.40 3.78 3.44 3.11 2.88 2.68
 7.99 6.07 5.26 4.79 4.05 3.69 3.36 3.03 2.74 2.63
 7.15 5.98 5.07 4.55 3.94 3.62 3.26 2.99 2.74 2.62
 7.13 5.91 4.94 4.43 3.93 3.48 3.20 2.89 2.69 2.61
```

(a) What proportion of these top 40 stocks traded more than 4% of the outstanding shares?

(b) Calculate $\overline{Y}$ and S for the data given.

(c) Calculate the interval $\overline{Y} \pm k \cdot S$ for k= 1, 2, and 3. Count the number of measurements that fall within each interval and compare this result with the number that you would expect according to the empirical rule.

(d) Note that there is one extremely large value (11.88). Eliminate this value and repeat parts (b) and (c) for the remaining 39 observations. Compare the answers here to those found in parts (b) and (c), noting the effect of a single large observation on $\overline{Y}$ and S.

3. The (sample) range of a set of measurements is the difference between the largest and the smallest observed values (see also Lecture 1, page 20). The empirical rule suggests that the standard deviation of a set of measurements may be roughly approximated by one-fourth of the range (that is, range/4). Calculate this approximation to S for the data sets in (a) Problem 1, (b) Problem 2, full data, and (c) Problem 2, with data value 11.88 eliminated. Compare the result in each case to the actual, calculated value of S.

4. Show that
$$\sum_{i=1}^{n}(Y_i - \overline{Y}) = 0.$$

5. Show that
$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})Y_i.$$