

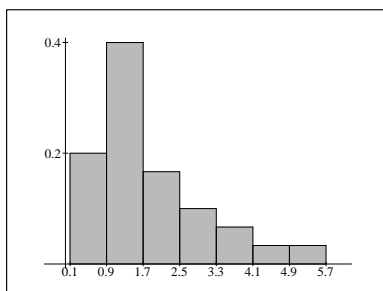
Descriptive Statistics Example

Consider the data set

3.6 1.1 1.4 0.6 1.1 1.6 3.7 2.7 1.3 2.5 2.6 5.2 1.6 1.7 1.9
1.3 2.2 0.3 1.4 1.1 4.5 1.3 1.7 0.5 0.2 1.2 2.3 1.2 0.7 0.3

- (a) Construct a relative frequency histogram for the data set using classes starting at 0.1 with a width of 0.8 (i.e., 0.1 to less than 0.9, etc.).

Class Boundaries	Frequency	Relative Frequency
0.1 to < 0.9	6	$6/30 = 1/5$
0.9 to < 1.7	12	$12/30 = 2/5$
1.7 to < 2.5	5	$5/30 = 1/6$
2.5 to < 3.3	3	$3/30 = 1/10$
3.3 to < 4.1	2	$2/30 = 1/15$
4.1 to < 4.9	1	$1/30$
4.9 to < 5.7	1	$1/30$
TOTAL	30	1



- (b) Find the sample mean \bar{x} , median m , and mode.

$$\bar{x} = \frac{1}{30} \sum_{k=1}^{30} x_k = 1.76$$

Sorted data:

0.2, 0.3, 0.3, 0.5, 0.6, 0.7, 1.1, 1.1, 1.1, 1.2, 1.2, 1.3, 1.3, 1.3, 1.4
1.4, 1.6, 1.6, 1.7, 1.7, 1.9, 2.2, 2.3, 2.5, 2.6, 2.7, 3.6, 3.7, 4.5, 5.2

Median: $m = 1.4$. The modes are 1.1 and 1.3, (each has the highest frequency 3).

- (c) Discuss the skewness of the distribution.
Since $\bar{x} > m$, the distribution is skewed to the right.
- (d) Find the range. The range is $R = 5.2 - 0.2 = 5$.
- (e) Find the sample variance s^2 and sample standard deviation s .

$$s^2 = \frac{1}{29} \sum_{k=1}^{30} (x_k - \bar{x})^2 \approx 1.4749, \quad s = \sqrt{s^2} \approx 1.2144$$

- (f) Find the percentage of data in the interval $\bar{x} \pm 2s$ and compare it with what Tchebysheff's Theorem predicts.
 $(\bar{x} - 2s, \bar{x} + 2s) \approx (-0.669, 4.189)$. There are 28 out of 30, i.e. about 93.3%, data points in the interval. Tchebysheff's Theorem predicts that at least 75% of the data should be in the interval.