Section 3.2

Standard Deviation

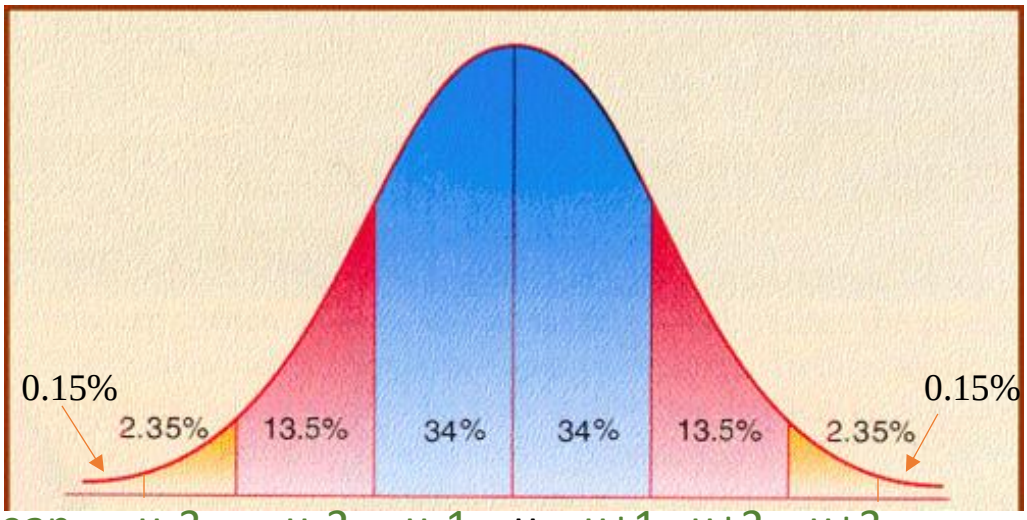
The sum of the deviations about the mean is always equal to zero.
It describes how far each observation is from the typical value
(measures the center of the distribution)

The standard deviation measures the spread of the distribution.
A larger standard deviation means that the observations are more distant from the typical value and more dispersed.

Chebyshev's inequality is less precise than the empirical rule, but will work for any distribution, while the empirical rule only works for bell-shaped

σ means standard deviation of the population Unadj. std. dev.
σ^2 means variance of a population Unadj. variance
s means standard deviation of a sample so use Std. dev.
s^2 means variance of a sample so use variance

Empirical Rule



Value away
From the mean μ-3 μ-2 μ-1 μ μ+1 μ+2 μ+3

Standard ------3--------2-------1----0----1------2-------3--------
Deviations |-----68%-----|
|-------------95%-------------|
|-----------------99.7%-----------------|

- Approximately 68% of the data will lie within 1 standard deviation of the mean.
- Approximately 95% of the data will lie within 2 standard deviations of the mean.
- Approximately 99.7% of the data will lie within 3 standard deviations of the mean.

Chebyshev"s inequality – for a data set, regardless of the shape of distribution, at least $\left(1-\frac{1}{k^2}\right)100\%$ of the observations will lie within k standard deviations of the mean.

Sample Variance

$$s^2 = \frac{\sum(x-\bar{x})^2}{n-1}$$

Sample Standard Deviation

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