Variation

Math 122

Things To Look For in Data

- Shape of data
  - Histogram, Symmetry, Skewness
- Center
  - Mean, Median, Mode, Midrange
- **Variation**

Measures of Variation

- **Range** = maximum – minimum

  Will not use much

- **Standard Deviation** — Sample - s / population - σ

  Think average distance from mean

  Calculate with 1-Var-Stats on TI or summary statistics online

Standard Deviation Formulas

- Sample:

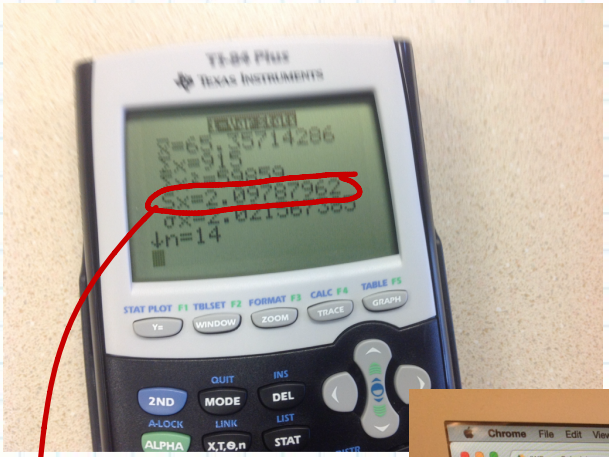
$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

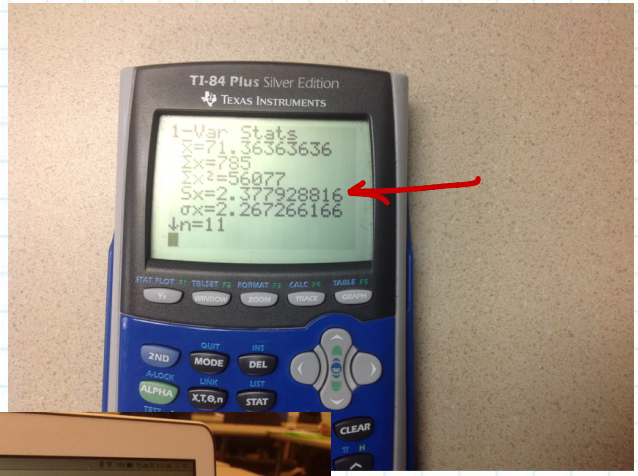
- Population:

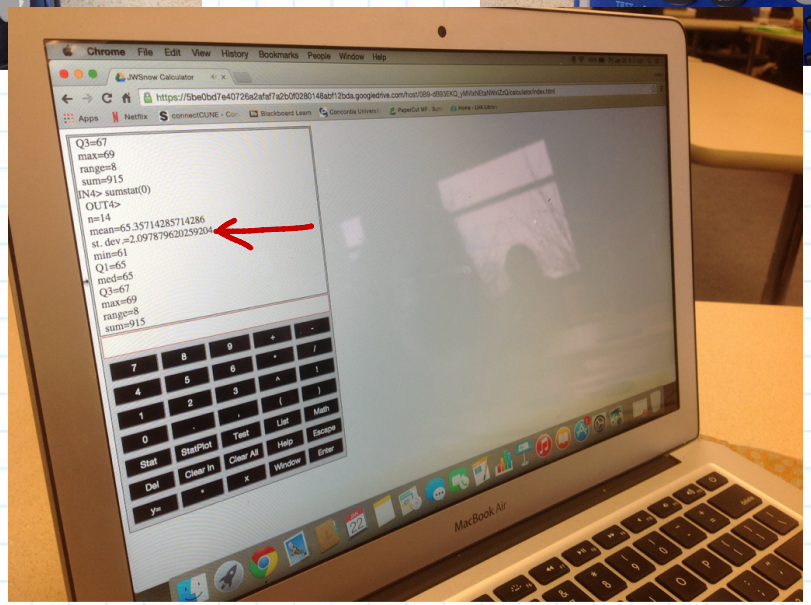
$$\sigma = \sqrt{\frac{\sum (x - \mu)^2}{N}}$$

- Think: Average Distance From the Mean

Find the standard deviation of the heights of males and females in class.







Sx not σx

Variance

- The **variance** is the square of the standard deviation.
- Notation
  - Sample: s^2
  - Population: σ^2
- Find the variance of the heights of females in the class.

$s = 2.09788$

$s^2 = (2.09788)^2 = 4.4$

Some values from an NHANES data set of 1000 college age males and females

	Female		Male	
	Mean	St. Dev.	Mean	St. Dev.
Height	63.8 in	2.7 in	69.6 in	3.2 in
Weight	154.7 lb	43.0 lb	179.7 lb	47.7 lb
Waist	35.2 in	6.7 in	35.2 in	6.3 in
Pulse	79.1	13.2	70.5	10.7

	Mean	Standard Deviation
Annual Snowfall in Lincoln	26.7 in	11.1 in
Annual Mean Temperature in Lincoln	51.5 degrees	1.6 degrees
IQ	100	15
Infant Birth Weight	7.5lb	1.1lb
SAT Area Test	500	100
ACT	18	6

Unusual Data Values

Range Rule of Thumb

- For many types of data (whose histograms are bell-shaped) about 95% of data values are within two standard deviations of the mean.

- **Maximum Usual Value** = $\mu + 2\sigma$
  - Any value larger than this is unusually large.
- **Minimum Usual Value** = $\mu - 2\sigma$
  - Any value smaller than this is unusually small.

$\mu = 26.7$ in

$\sigma = 11.1$ in

Snowfall

- Find the maximum usual annual snowfall for Lincoln.

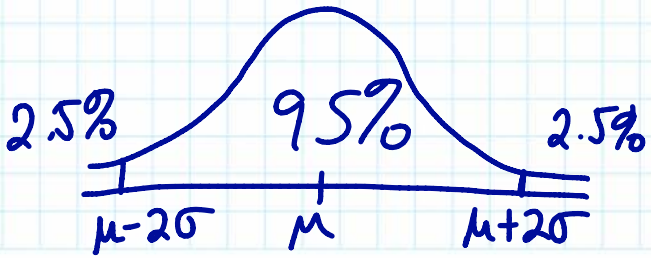
$\mu + 2\sigma = 48.9$ in

- Find the minimum usual annual snowfall for Lincoln.

$\mu - 2\sigma = 4.5$ in

- Would 40 inches in one year be unusual?

No

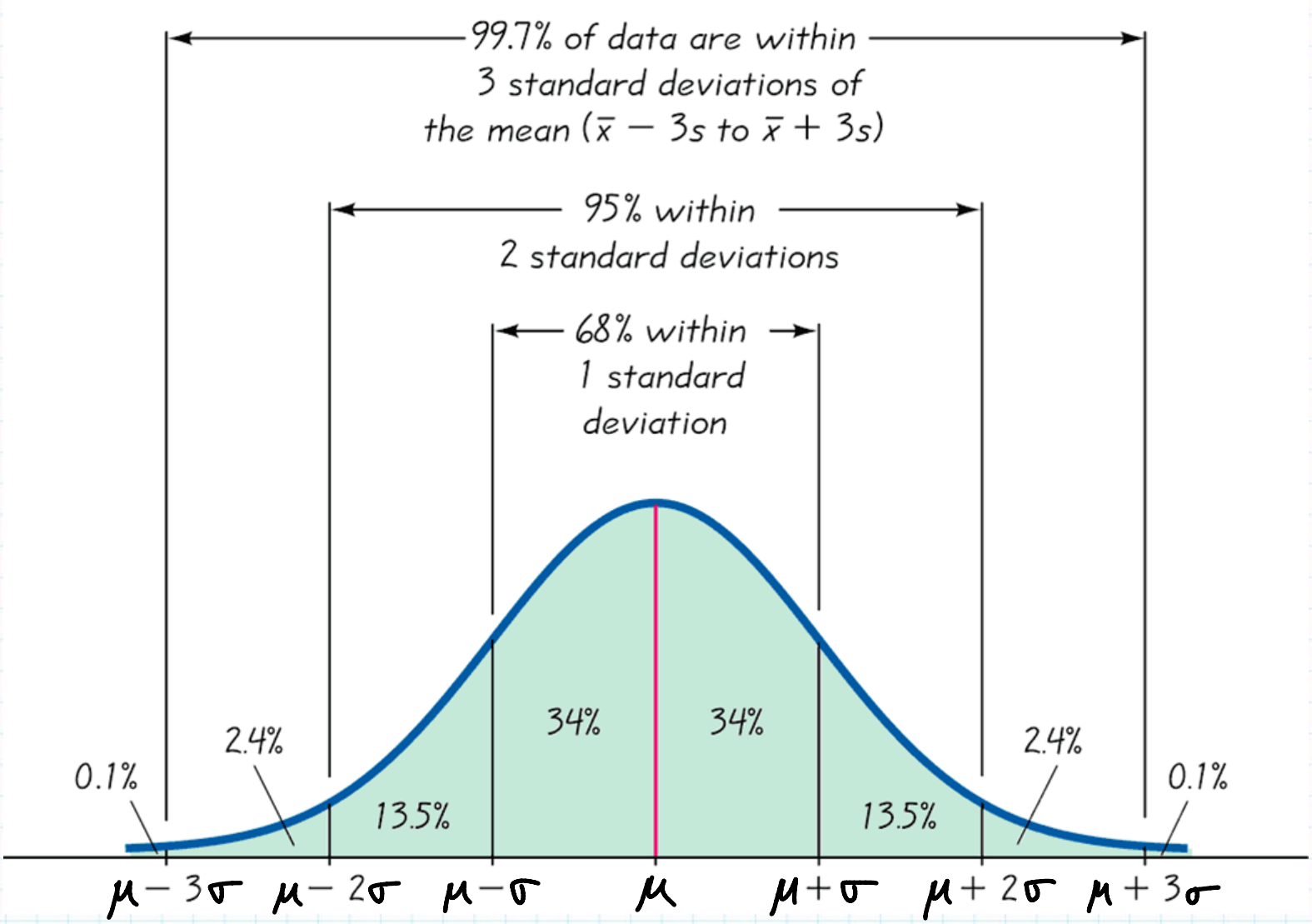


Empirical Rule

- For many types of data (whose histograms are bell-shaped)
  - About 68% of the data values are within one standard deviation of the mean.
  - About 95% of the data values are within two standard deviations of the mean.
  - About 99.7% of the data values are within three standard deviations of the mean.

RBT

Empirical Rule



$\mu = 7.5$ lbs

$\sigma = 1.1$ lbs

Infant Birth Weights

- Use the Empirical Rule to find an interval which contains 68% of infant birth weights.

$\mu - \sigma =$ $\mu + \sigma =$

- Find an interval which contains 95% of infant birth weights.

$\mu - 2\sigma = 5.3$ lbs $\mu + 2\sigma = 9.7$ lbs

- Find an interval which contains 99.7% of infant birth weights.

$\mu - 3\sigma =$ $\mu + 3\sigma$

Summary

- Range = max – min
- Standard Deviation
  - Think average distance from the mean
- Range Rule of Thumb
  - Find max and min usual values
  - Identify unusual values
- Empirical Rule
  - Find intervals for 68%, 95%, 99.7%