Relative Standing

Math 122

Notation

- Σ (Greek Sigma) denotes a SUM
- x - variable to denote individual data values
- Σx - sum of all values of x
- n - number of data values in a SAMPLE
- N - number of data values in a POPULATION
- $\bar{x}$ - the mean or average of a SAMPLE
- μ - the mean or average of a POPULATION

Standard Deviation

- Sample: s
- Population: σ
- Think: Average Distance From the Mean

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

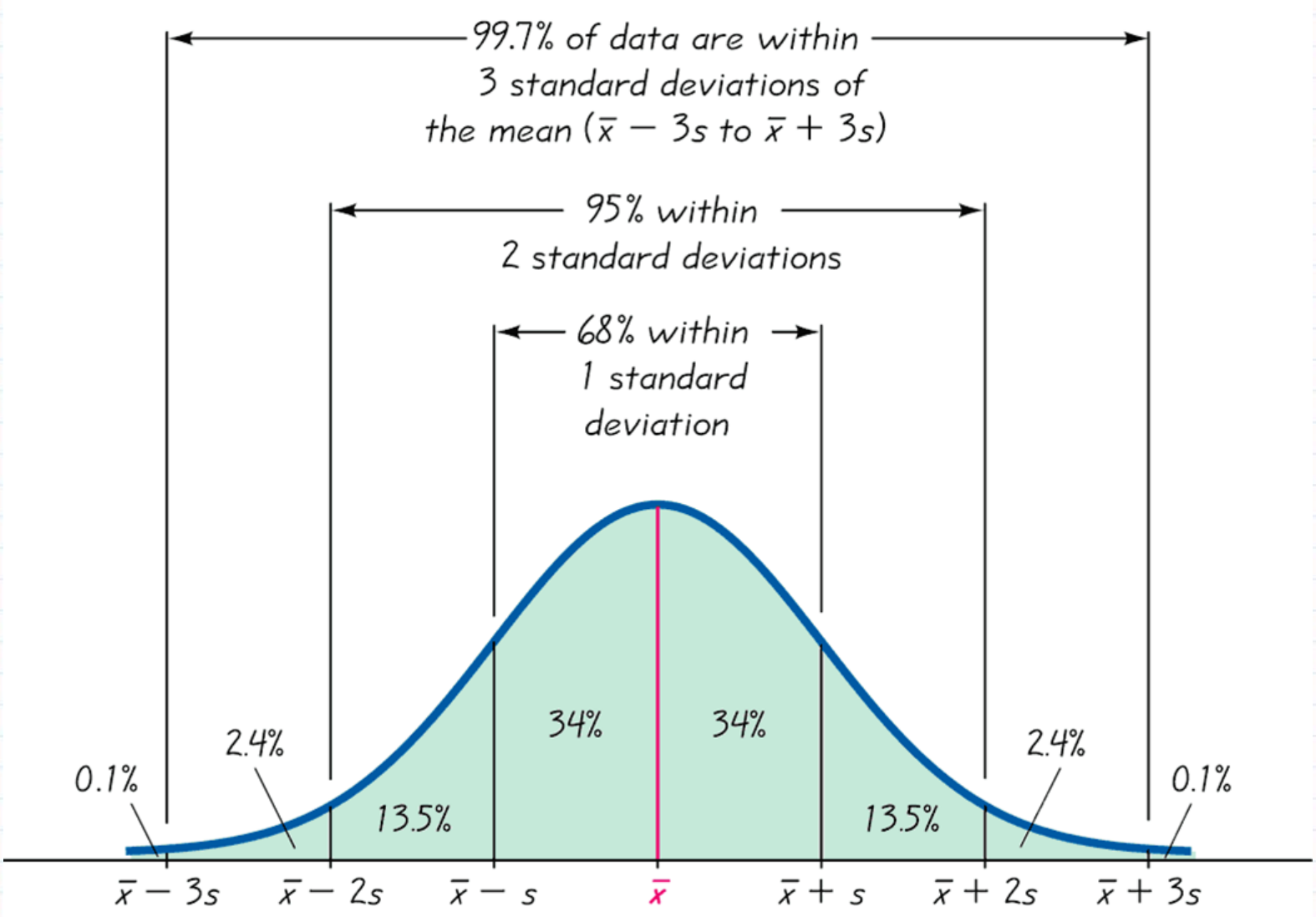
	Female		Male	
	Mean	St. Dev.	Mean	St. Dev.
Height	63.8 in	2.7 in	69.6 in	3.2 in
Weight	154.7 lb	43.0 lb	179.7 lb	47.7 lb
Waist	35.2 in	6.7 in	35.2 in	6.3 in
Pulse	79.1	13.2	70.5	10.7

	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

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.

Empirical Rule



Range Rule of Thumb

A data value x is not unusual if

$$\mu - 2\sigma \leq x \leq \mu + 2\sigma$$

Z-scores

The z score of a data value x is

$$z = \frac{x - \mu}{\sigma}$$

Interpreting Z-scores

- If x is above average $Z>0$
- If x is below average $Z<0$

- If x is close to average, Z is close to 0

Comparing Z scores

- In the last 5 years, **male** NAIA pole vaulters averaged 3.92m with a standard deviation of 0.63m
- In the last 5 years, **female** NAIA pole vaulters averaged 3.02m with a standard deviation of 0.42m
- Bob can vault 4.75m

- Sue can vault 3.65m

- Who is better?

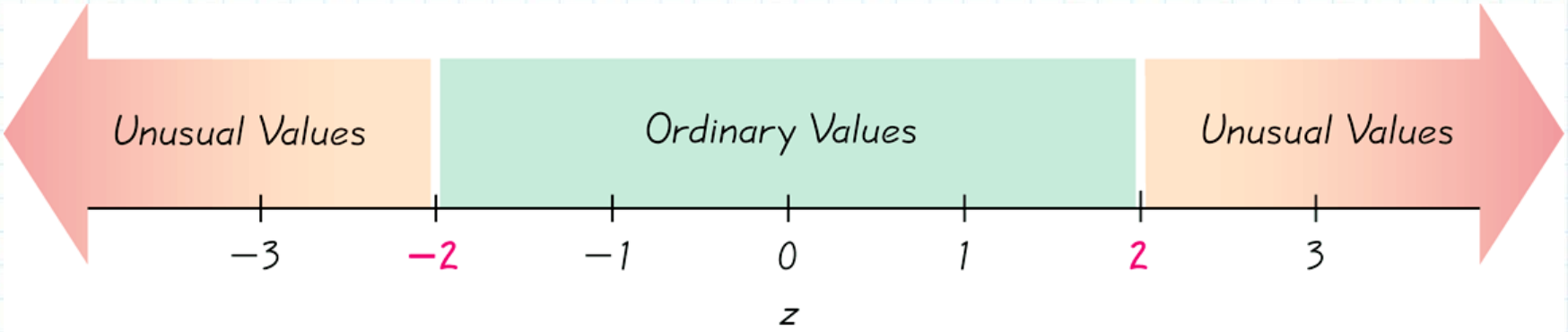
Comparing Z scores

- **Male** NAIA pole Vaulters average 3.92m with a standard deviation of 0.63m
- **Female** NAIA sprinters averaged 13.31s in the 100m with a standard deviation of 0.875s.
- Frank can vault 4.85m.

- Jane ran the 100m in 11.73s.

- Who is better?

Range Rule of Thumb



Range Rule of Thumb

- Is 69in an exceptionally tall height for a female?

Other measures of relative standing

Percentiles

- Percentiles are numbers P_1, P_2, P_3,…,P_{99} which divided data into 100 sets which each contain about 1% of the data.
- About 1% of the data is less than P_1.
- About 2% of the data is less than P_2.
- About k% of the data is less than P_k.

Percentile of x

$$\frac{\text{number of values less than } x}{\text{number of data values}} \cdot 100$$

Find the percentile of 10 in this data

0	1	2	2	2	3	3	4	4	4
5	5	5	6	8	8	10	10	11	12
12	12	14	15	15	15	16	16	17	17
19	19	21	22	22	23	24	25	25	26
27	28	28	28	29	30	30	30	32	32

Finding P_k

- Let $m = \frac{k}{100} \cdot n$
- We want a number which separates the bottom m data values from the other data values.
- If m is a whole number, P_k is the

  average of the m^{th} and $(m + 1)^{st}$ data values.
- If m is not a whole number, let M be the next whole number greater than m. Then P_k is the M^{th} data value.

Find P_{26}

0	1	2	2	2	3	3	4	4	4
5	5	5	6	8	8	10	10	11	12
12	12	14	15	15	15	16	16	17	17
19	19	21	22	22	23	24	25	25	26
27	28	28	28	29	30	30	30	32	32

Find P_{75}

0	1	2	2	2	3	3	4	4	4
5	5	5	6	8	8	10	10	11	12
12	12	14	15	15	15	16	16	17	17
19	19	21	22	22	23	24	25	25	26
27	28	28	28	29	30	30	30	32	32

Quartiles

- Quartiles are numbers Q_1, Q_2, and Q_3 which divide data into 4 sets which each contain about 25% of the data.
- About 25% of the data is below Q_1.
- About 50% of the data is below Q_2.
- About 75% of the data is below Q_3.

Quartiles

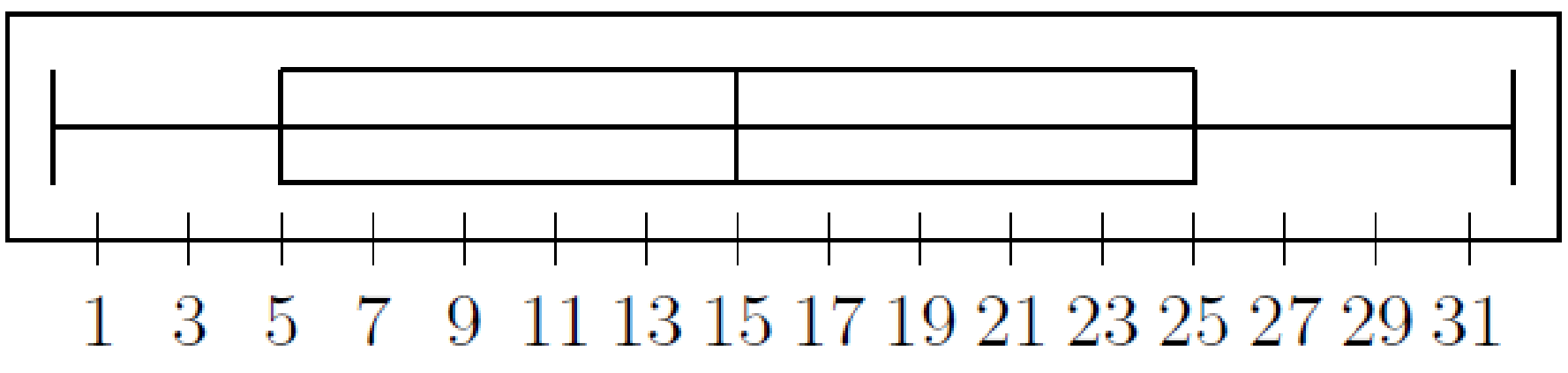
- $Q_1 = P_{25}$
- $Q_2 = P_{50}$
- $Q_3 = P_{75}$

Five Number Summary

- Minimum
- Q_1
- Q_2
- Q_3
- Maximum

We will get the five number summary from a calculator.

Box Plot



Female and Male Heights