Confidence Intervals

An interval of **4 plus or minus 2**

A Confidence Interval is a **range of values** we are fairly sure our **true value** lies in.



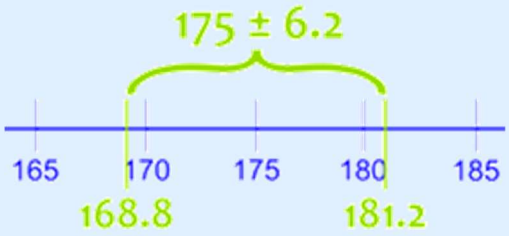
Example: Average Height

We measure the heights of **40** randomly chosen men, and get a mean height of **175cm**,

We also know the standard deviation of men's heights is **20cm**.

The **95% Confidence Interval** (we show how to calculate it later) is:

175cm ± 6.2cm



This says the **true mean** of ALL men (if we could measure all their heights) is likely to be between 168.8cm and 181.2cm.

But it might not be!

The "95%" says that 95% of experiments like we just did will include the true mean, but **5% won't**.

So there is a 1-in-20 chance (5%) that our Confidence Interval does NOT include the true mean.

Calculating the Confidence Interval

Step 1: find the number of observations **n**, calculate their mean $\bar{X}$, and standard deviation **s**

Using our example:

- Number of observations: **n = 40**
- Mean: $\bar{X}$ **= 175**
- Standard Deviation: **s = 20**

Note: we should use the standard deviation of the entire **population**, but in many cases we won't know it.

We can use the standard deviation for the **sample** if we have enough observations (at least n=30, hopefully more).

Step 2: decide what Confidence Interval we want: 95% or 99% are common choices. Then find the "Z" value for that Confidence Interval here:

Confidence Interval	Z
80%	1.282
85%	1.440
90%	1.645
95%	1.960
99%	2.576
99.5%	2.807
99.9%	3.291

For 95% the Z value is **1.960**

Step 3: use that Z in this formula for the Confidence Interval

$$\bar{X} \pm Zs/\sqrt{n}$$

Where:

- **X** is the mean
- **Z** is the chosen Z-value from the table above
- **s** is the standard deviation
- **n** is the number of observations

And we have:

$$175 \pm 1.960 \times 20/\sqrt{40}$$

Which is:

175cm ± 6.20cm

In other words: from 168.8cm to 181.2cm

The value after the ± is called the **margin of error**
The margin of error in our example is 6.20cm

Formula for Population standard deviation:

The "**Population** Standard Deviation": $\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$

Here are the two formulas, explained at Standard Deviation Formulas if you want to know more:

The "**Population** Standard Deviation": $\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$

The "**Sample** Standard Deviation**":** $s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2}$

Looks complicated, but the important change is to
divide by **N-1** (instead of **N**) when calculating a Sample Variance.