Population: Pop Mean $\mu$ (mu) $\mu = \frac{\sum x}{N}$

Sample: Mean $\bar{x}$ $\bar{x} = \frac{\sum x}{n}$

Sample Standard Deviation $s_x$ $s_x = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$
Look at nice algebraic properties of $\sigma_x$ calc

Also

$\sigma_x \propto \sqrt{\overline{x^2} - \overline{x}^2}$

If you apply this to sample data you can get $\sigma_x$ out of it

Just

$\sigma_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2}$

As

$\sigma_x = \sqrt{\frac{m}{m-1} \frac{1}{n} \overline{x^2} - \overline{x}^2}$

Likewise

$\sigma_y = \sqrt{M_{y^2} - \overline{y}^2}$

$\sigma_y = \sqrt{\frac{m}{m-1} \overline{y^2} - \overline{y}^2}$

So (example) given data $m = 400$

$\overline{x} = 18.2 \quad \overline{x^2} = 500 \quad \sigma_x = \sqrt{\frac{400}{399} \sqrt{500 - 18.2^2}}$
\[ y = 2.7 \text{ AVG CLASS LEVEL} \quad y^2 = 11 \text{(Day)} \]

\[ y = \sqrt{\frac{400}{399}} \sqrt{11 - 2.7^2} \]

Note: Close to 1. Beware precision issues.

If also \( xy = 65 \) \( \Rightarrow n = \frac{xy - \bar{x} \bar{y}}{\bar{x} \bar{y}} = \frac{65 - (18.2)(2.7)}{\sqrt{500 - 18.2^2} \sqrt{11 - 2.7^2}} \)

All good for \( xy, x^2, y^2, xy \).

\[ M_{ax+b} = a \quad M_{x+b} \]

\[ \bar{x} \bar{y} = \frac{1}{|a|} \sigma_x \]

\[ ax + b, cy + d \quad \text{AVGS.} \]

\[ \bar{x}, \bar{y} \quad \text{if} \quad ac > 0 \]
LOOKEE! SAMPLE PLOT M = 10 POINTS GIVES RED SAMPLE REGRESSION

CLOSELY COINCIDES WITH THE POPULATION REGRESSION!!

IMPORTANT! M = 10 PLOT CAN KEEP YOU ON TRACK.

RED LINE IS SAMPLE REGRESSION LINE

POP LINE NAIVE
Now to ch 23 readings (syllabus)

Matter of estimating population mean $\mu$. From random eq. probability with replacement.

"One possible estimator of $\mu$ is $\bar{x}$.

Other possibilities include

Trimmed mean — e.g. toss out $X_{(1)} + X_{(m)}$

(earliest, censored)

And avg $n-2$ remaining.

Non-parametric weights $w_1 \ldots w_n$ on $X_{(1)} \leq X_{(m)}$

Method of assigning

(esp) Margin of error $1.96 \frac{\sigma}{\sqrt{n}}$

Margin of error for $\bar{x}$

95% confidence interval for $\mu$:

$\bar{x} \pm 1.96 \frac{\sigma}{\sqrt{n}}$

Claim: Probability $\bar{x} \pm 1.96 \frac{\sigma}{\sqrt{n}}$ covers $\mu$ is $\approx 0.95$.
Why so??

Imagine a pop of scores 16 17 18 19 20

Pop

\[ \text{Suppose} \]

\[ \mu_x = 17.7 \quad \text{(say)} \]

\[ \sigma_x = 1.8 \quad \text{!! (say)} \]

Take \( m = 100 \)

\[ \bar{x} \]

\[ \pm 1.96 \times \sigma/\sqrt{m} \]

One sample of \( m \)

Random

With REPL

Sample

\[ \bar{x}, \sigma_x, m \]

One sample

\[ \bar{x} \pm 1.96 \times \sigma/\sqrt{m} \]

\[ \bar{x} \]

An est of

Central limit theorem

\[ \frac{\bar{x}}{\mu_x} \pm \frac{\sigma}{\sqrt{m}} \]

\[ \bar{x} \]

Possible
CENTRAL LIMIT THEOREM

Dist of $\frac{\bar{X} - M_x}{\sigma_{X/Nn}} \rightarrow Z \text{ Distn } N(0,1)$

Also (more useful form):

$(\bar{X} - M_x) / (\sigma_X / \sqrt{n}) \rightarrow Z \text{ Distn}$

Applied:

$P\left( \left| \frac{\bar{X} - M_x}{\sigma_X / \sqrt{n}} \right| < 1.00 \right) \approx P(1.21 < 1.00) = .68$

This is equal to $M_x$ covered by $\bar{X} \pm 1.00 \frac{\sigma_X}{\sqrt{n}}$.
App: proby $M_x$ in
**NOTE:** If your sample size is without replacement, modify as follows:

\[ n' = \frac{n}{1 - \frac{N-n}{N}} \]

**Entry covers \( \mu \sim (\text{Table E}) 98\% \)**

So, reject \( H_0: \mu = 3.23 \) if \( z > 2.326 \).
Students: How to make do with small $n$.

Looked at $\frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \sim Z$ as $n \to \infty$ regardless of form of $\mu_0n$!!

If $x_1, \ldots, x_n$ are from a normal population

If population is (approximately) normal

Think $x_i = \xi_i + \mu$ persists thus thru

$\frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{\bar{z}}{\sigma / \sqrt{n}}$ tabulate this dist $N$
For \( m = 2, 3, 4 \ldots \text{ad \textit{ad}} \) \( m = \infty \) \textit{back}.

Say \( B_0 \) is normal eq to \( \mu = 100 \), \( \sigma = 15 \)

Sample \( n = 4 \) persons.

\[
\bar{x} = 106 \quad \sigma_x = 13.8 \quad n = 4
\]

Use 95% interval \( \bar{x} \pm t \frac{\sigma_x}{\sqrt{n}} \)

\[
106 \pm \frac{2.776}{2.776} \frac{13.8}{\sqrt{4}}
\]

\( \text{Deg} = n-1 = 3 \)

So \( t-\text{interval} \):

\[
\bar{x} \pm (t \text{ score}) \frac{\sigma_x}{\sqrt{n}} \text{ having proby of exactly proby 95% of cover}
\]