II. Confidence Interval on Probability (Proportion)

The C.I. is the probability error bound of margin (EBM)

$$C.I. = p' \pm E$$

$\hat{p}$ for regular text and TI-84

where $p' = \frac{x}{n}$ ← sample's prob

$E = Z_{\alpha/2}\sqrt{\frac{p'q'}{n}}$ ← $q' = 1 - p'$

That is, the C.I. for the population's proportion p:

$$\boxed{\begin{array}{c} p' - E < p < p' + E \\ \text{where } p' = \frac{x}{n},\quad E = Z_{\alpha/2}\sqrt{\frac{p'q'}{n}} \end{array}}$$

$Z_{\alpha/2}\sqrt{\frac{p'q'}{n}}$ process skew

C.V.

$p'q' \doteq n$

TI-84: $\boxed{(\hat{p} - E,\ \hat{p} + E)}$

Stat → TESTS → A: 1-PropZInt... :

Eg Earlier, we had that a Pew Research Center poll of 1007 randomly selected adults showed that 85% of respondents know what Twitter is. The sample results are n = 1007.

a. Find the 95% confidence interval estimate of the population proportion p.

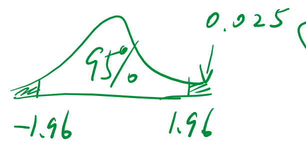
b. Based on the results, can we safely conclude that more than 75% of adults know what Twitter is?

c. Assuming that you are a newspaper reporter, write a brief statement that accurately describes the results and includes all of the relevant information.

S: a. I'd know 95% of the time from the min to max's prob.

$p' = 0.85, \quad q' = 1 - p' = 1 - 0.85 = 0.15$

$\alpha = 0.05$, since C.I = 95%



$\frac{\alpha}{2} = \frac{0.05}{2} = 0.025$

$$E = Z_{\alpha/2}\sqrt{\frac{p'q'}{n}}$$

$$E = 1.96\sqrt{\frac{0.85 \cdot 0.15}{1007}} \approx 0.022$$

Thus, $p' - E < p < p' + E$

$0.85 - 0.022 < p < 0.85 + 0.022$

$\boxed{0.828 < p < 0.872}$ ← 95% of the time

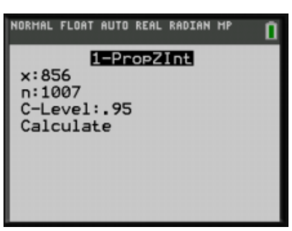
TI-84: Need x : Because $\hat{p}$ $p' = \frac{x}{n}$

$x = np'$

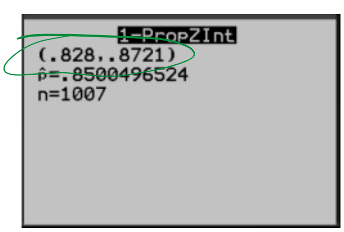
$x = 1007 \cdot 85\%$

$= 855.95$ ← not okay for TI-84

≈ 856 round



→



Eg Earlier, we had that a Pew Research Center poll of 1007 randomly selected adults showed that 85% of respondents know what Twitter is. The sample results are n = 1007.
~~a. Find the 95% confidence interval estimate of the population proportion p.~~

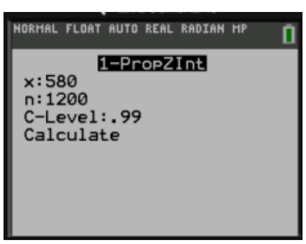
showed that 85% of respondents know what Twitter is. The sample results are n = 1007.
~~a. Find the 95% confidence interval estimate of the population proportion p.~~
b. Based on the results, can we safely conclude that more than 75% of adults know what Twitter is?
c. Assuming that you are a newspaper reporter, write a brief statement that accurately describes the results and includes all of the relevant information.

b. Because the C.I. is $0.828 < p < 0.872$. The entire interval surpass 75%, it surely is yes.

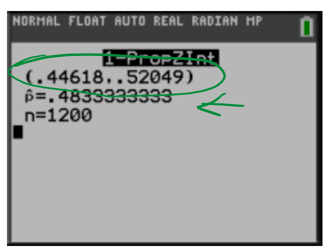
c. Base, 0.02 of margin is low. The C.I. is very reliable due to the narrowness.

Eg Suppose that a sample of 1200 registered voters in California, 580 of them said they plan to vote for Obama back in 2008. Construct the 99% confidence interval for the percentage of registered votes in California who plan to vote for Mr. Obama.

S: $n = 1200$, $x = 580$, ask for 99% C.I.



→



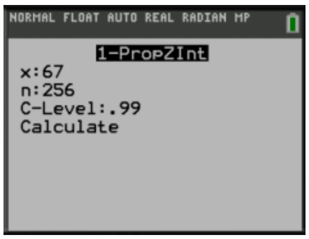
Eg A long-term study followed 256 women throughout their lifetime. Among the various data collected, one interesting result was that 26.2% of these women had suffered from a migraine at some point in their life.
Use the results of this study and a 99% confidence interval to estimate the percentage of women in the population that have suffered from a migraine at some point in their life.

S: $n = 256$, $p' = 26.2\%$

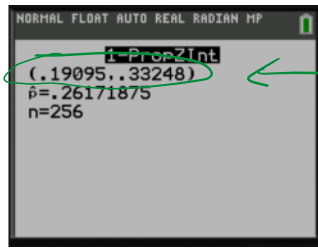
$p' = \frac{x}{n}$

$x = np'$

$x = 256 \cdot 0.262 \approx 67$



$\rightarrow$



$\leftarrow$ Wow

III. Confidence Interval on Mean

$\leftarrow$ $87 \text{ pt} < \mu < 92 \text{ pt}$

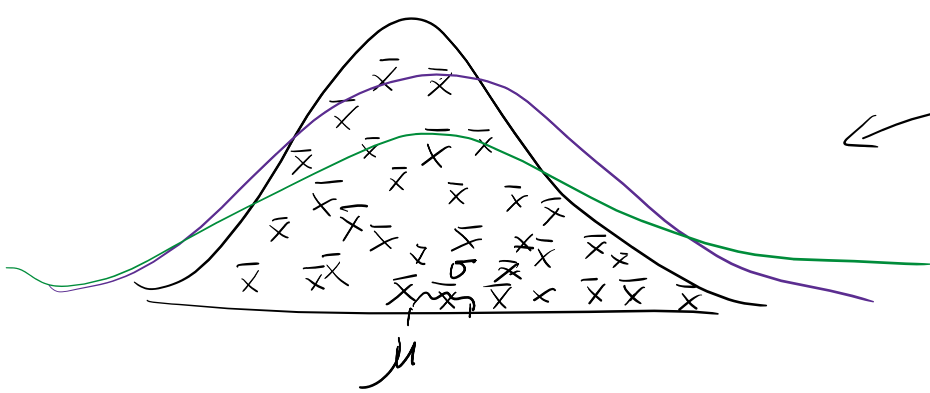
$\leftarrow$ $55 \text{ mph} < \mu < 60 \text{ mph}$

The C.I. on population mean is the range of value that consist of the value of the sample mean.

There are 2 parts:

A. σ is given $\leftarrow$ automatically N.D. for any size n

B. σ is <u>not</u> given. Find s $\leftarrow$ our focus! B includes A.



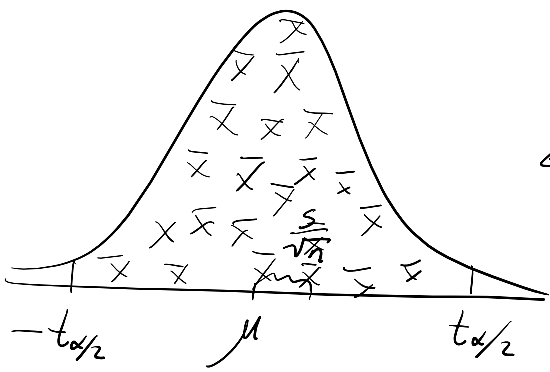
$\leftarrow$ Size n is small

Now, we have a pre-made "bell-curve" that is for any small size n.

The new distribution is called t-distribution (Student distribution) small size n.

The new distribution is called t-distribution (Student distribution)

$\Rightarrow$



$\leftarrow$ t-distribution

This would work for any size n, with

$$\boxed{t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}}$$

where s is sample's std. dev.

The Confidence Interval for the population mean is

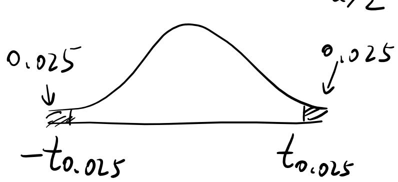
$$\boxed{\begin{array}{l} \bar{x} - E < \mu < \bar{x} + E \\ \text{where } E = t_{\alpha/2}\frac{s}{\sqrt{n}},\ d.f. = n-1 \end{array}}$$

d.f.: degree of freedom

association:
1 2 3 4 5

Eg A sample of size 12 is a simple random sample selected from a normally distributed population. Find the critical value $t_{\alpha/2}$ corresponding to a 95% confidence level.

S:



$\alpha/2$

$\alpha = 0.05$

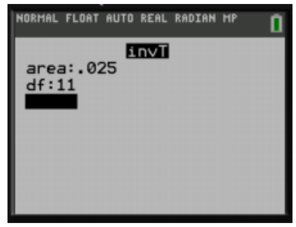
$$\frac{1-0.95}{2} = \frac{0.05}{2} = 0.025$$

$n = 12$, $d.f. = n-1$

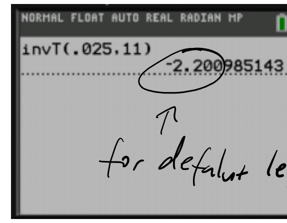
$n = 12, \quad d.f. = n-1$
$= 12-1$
$= 11$

Table: $t_{0.025} = \boxed{2.201}$

TI-84: 2nd $\rightarrow$ distr $\rightarrow$ 4: invT(



$\rightarrow$



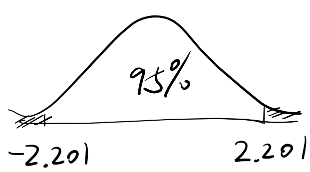
$\uparrow$ for default left tail

Eg Listed below are speeds (mi/h) measured from southbound traffic on I-280 near Cupertino, California (based on data from SigAlert). This simple random sample was obtained at 3:30 P.M. on a weekday. The speed limit for this road is 65 mi/h. Use the sample data to construct a 95% confidence interval for the mean speed. What does the confidence interval suggest about the speed limit?

62 61 61 57 61 54 59 58 59 69 60 67 $\leftarrow n = 12$

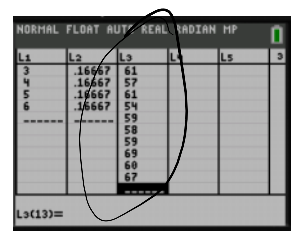
S: Want 95% C.I. $\bar{x} - E < \mu < \bar{x} + E$

$E = t_{\alpha/2} \frac{s}{\sqrt{n}}, \quad d.f. = n-1 = 12-1 = 11$

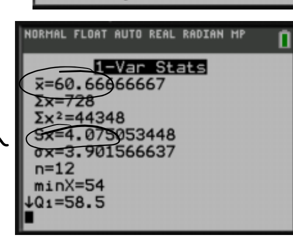


$\frac{1-0.95}{2} = \frac{0.05}{2} = 0.025, \quad d.f. = 11$

Put the list and calculate $\bar{x}$, s: $\bar{x} = 60.7, \quad s \approx 4.1$



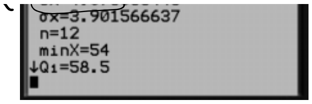
then, $E = 2.201 \cdot \frac{4.1}{\sqrt{12}} \approx 2.6$



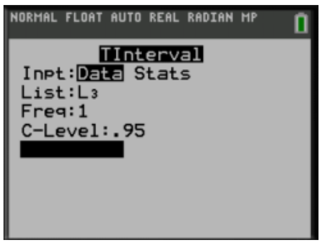
Thus, $60.7 - 2.6 < \mu < 60.7 + 2.6$

$\boxed{58.1 < \mu < 63.3}$

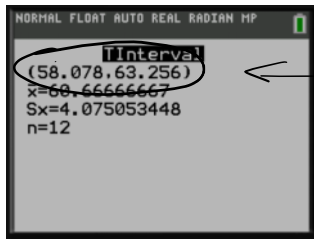
$\boxed{58.1 < \mu < 63.3}$



TI-84: Stat → TESTS → 8: TInterval...



→



← min to max

Very Cool !

need the list of data value

The C.I. shows the entire interval under the speed limit of 65 mph.

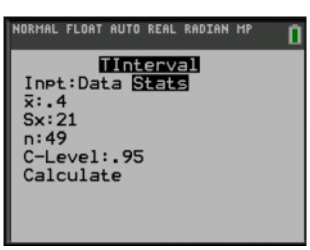
n

Eg A ~~common claim~~ is that garlic lowers cholesterol levels. In a test of the effectiveness of garlic, 49 subjects were treated with doses of raw garlic, and their cholesterol levels were measured before and after the treatment. The changes in their levels of LDL cholesterol (in mg/dL) have a mean of 0.4 and a standard deviation of 21.0. Use this statistics to construct a 95% confidence interval estimate of the mean net change. What does the confidence interval suggest about the effectiveness of garlic in reducing LDL cholesterol?

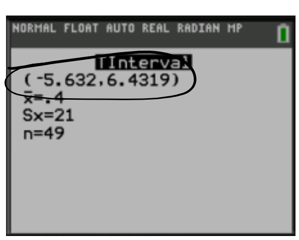
S: $n = 49$, $\bar{x} = 0.4$, $s = 21.0$, ask 95% C.I.

← $\bar{x} = 0.4$, $s = 21.0$

red flag!



→



$-5.6 < \mu < 6.4$ contains 0, where -5.6 to 0 is 0 to 6.4 is close. It has 0 effect.

(0 in the C.I. means 0 effect!)