Point and Interval Estimation for population mean (small samples $n \le 30$)

Learning Objectives

What do we estimate? **Population Mean** ($\mu = ?$ **)**
Know all the new **terminologies** and related **notations** (Point estimate $\bar{x}$)
Know all the new **formulas** on **formula sheet** and their related **TI commands**.
Know in estimating **population mean** ($\mu = ?$) when to use **normal distribution** versus **t- distribution**.
Know how to use TI (**option 8**) or (**formula** $\mu = \bar{x} \pm E$) to estimate **population mean** ($\mu = ?$).

Definitions:

Point estimate: Sample statistics such as $(\bar{X})$

Confidence Interval: A confidence interval (or interval estimate) is a range (or an interval) of values used to estimate the true value of a population parameter. A confidence interval is sometimes abbreviated as CI.

***A confidence level**: a* confidences level is the probability $(1-\alpha)$ (often expressed as the equivalent percentage value) usually 90%, 95%, or 99%.that is the proportion of times that the confidence interval actually does contain the population parameter, assuming that the estimation process is repeated a large number of times. Percentage outside confidence level is called **critical area** (α).
So for example with $95\% = (1-\alpha)$ confidence level then the critical **area** will be $(\alpha = 5\%)$.

Estimating Population **Mean** $\mu = \bar{x} \pm E$

Important: *If* ***confidence level is not given*** *use* ***95% as a default***.

$\bar{x}$ = **Point** estimate (Sample Mean)		E = Margin of error(error bound)
$n \le 30$		
Margin of Error	If $n \le 30$ use t, from table on page 3 $E = t\frac{s}{\sqrt{n}}$	
Interval Estimate	$\mu = \bar{x} \pm E$	$\mu = \bar{x} \pm E$
TI-83/84	*stat* $\rightarrow$ *tests* $\rightarrow$ ***Option 7(Z-interval)***	*stat* $\rightarrow$ *tests* $\rightarrow$ ***Option 8(t-interval)***
Important relationship.	**Width (**difference between upper and lower bounds**)** $= 2E = UB - LB$ **Margin of Error** $= E = (UB - LB)/2$ **Point Estimate** $= \bar{x} = (UB + LB)/2$	

T-Distribution vs. the Normal Distribution for Confidence Interval for Means

Main Point to Remember:
You must use the t-distribution table when working problems when the population standard deviation (σ) is not known and the sample size is small $n \leq 30$.

General Correct Rule:
If σ is not known, then using t-distribution is correct. If σ is known, then using the normal distribution is correct.

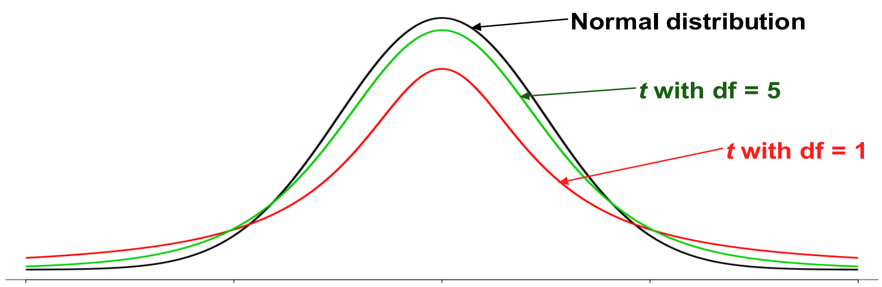
What is Most Common Practice:
Since people often prefer to use the normal, and since the t-distribution becomes equivalent to the normal when the number of cases becomes large, common practice often is:

- If σ known, then use normal.
- If σ not known:
  - If n is large, then use normal.
  - If n is small, then use t-distribution.

What is Another Common Way Textbooks Teach This:
Textbooks often simplify this to “large-sample” vs. “small-sample” methods; use normal distribution with large samples and t-distribution with small samples. This is right almost all the time, because in real sampling problems we seldom have a basis for knowing σ. However, there can be some situations when we do have a basis for assuming a value for σ, such as using a σ based on past data, and in those situations even if sample size is small the correct procedure would be to use the normal distribution, so the simplified “large-sample” vs. “small sample” approach would lead to an error.

t distribution

- *t* distribution looks like a normal distribution, but has “thicker” tails. The tail thickness is controlled by the **degrees of freedom**



- The smaller the degrees of freedom, the thicker the tails of the *t* distribution
- If the degrees of freedom is large (if we have a large sample size), then the t distribution is pretty much identical to the normal distribution

t distribution for small sample $n < 30$ Complete table on last page

df = n-1	<-------------------- alpha α -------------------->							
2-Tailed	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.005
1-Tailed	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.0025
Conf. Levl.	60%	70%	80%	90%	95%	98%	99%	99.5%
1	1.376	1.963	3.078	6.314	12.706	31.821	63.656	127.321
2	1.061	1.386	1.886	2.920	4.303	6.965	9.925	14.089
3	0.978	1.250	1.638	2.353	3.182	4.541	5.841	7.453
4	0.941	1.190	1.533	2.132	2.776	3.747	4.604	5.598
5	***0.920***	***1.156***	***1.476***	***2.015***	***2.571***	***3.365***	***4.032***	***4.773***
6	0.906	1.134	1.440	1.943	2.447	3.143	3.707	4.317
7	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.029
8	0.889	1.108	1.397	1.860	2.306	2.896	3.355	3.833
9	0.883	1.100	1.383	1.833	2.262	2.821	3.250	3.690
10	***0.879***	***1.093***	***1.372***	***1.812***	***2.228***	***2.764***	***3.169***	***3.581***
11	0.876	1.088	1.363	1.796	2.201	2.718	3.106	3.497
12	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.428
13	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.372
14	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.326
15	***0.866***	***1.074***	***1.341***	***1.753***	***2.131***	***2.602***	***2.947***	***3.286***
16	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.252
17	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.222
18	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.197
n>30 $\Rightarrow$ *Z*	*0.842*	*1.036*	*1.282*	*1.645*	*1.96*	*2.326*	*2.576*	*2.807*

Practice Problems 1

Find the **margin of error** for the following problems by using the t-table, be sure you **subtract 1 from n** and then use (n-1) row to find the t-value

$$E = t\left(s/\sqrt{n}\right) \quad n \le 30$$

1) Sample size $n = 9$, $s = 6$ and 95% confidence level? **Answer: 4.612**

Degree of freedom $= df = ? = 8$, $t = 2.306$ $\qquad E = 2.306\frac{6}{\sqrt{9}} = 4.612$

2) Sample size $n = 16$, $s = 8$ and 99% confidence level? **Answer: 5.894**
3) Sample size $n = 8$, $s = 20$ and 90% confidence level? **Answer: 13.4**
4) Sample size $n = 10$, $s = 4$ and 97% confidence level? **Answer: 3.568**
5) Sample size $n = 14$, $s = 10$ and the 95% confidence level? **Answer: 5.773**

2)Sample size $n = 16$, $s = 8$, and 99% confidence level? $df = ? = 15$, $t = 2.947$ $\quad E = 2.947\frac{8}{\sqrt{16}} = 5.894$

3)Sample size $n = 8$, $s = 20$, and 90% confidence level? $df = ? = 7$, $t = 1.895$ $\quad E = 1.895\frac{20}{\sqrt{8}} = 13.40$

4)Sample size $n = 10$, $s = 4$, and 98% confidence level? $df = ? = 9$, $t = 2.821$ $\quad E = 2.821\frac{4}{\sqrt{10}} = 3.568$

5)Sample size $n = 14$, $s = 10$, and 95% confidence level? $df = ? = 13$, $t = 2.16$ $\quad E = 2.16\frac{10}{\sqrt{14}} = 5.77$

Practice Problems 2

A) For the following problems based on sample size decide to use z or t value or neither?

1) Sample size $n = 12$, $s = 4$ and the population is normally distributed? z or t value : **t value**
2) Sample size $n = 39$, $s = 4$ and the population is normally distributed? z or t value: **z value**
3) Sample size $n = 18$, $s = 4$ and the population is normally distributed? z or t value : **t value**
4) Sample size $n = 40$, $\sigma = 4$ and the population is normally distributed? z or t value: **z value**
5) Sample size $n = 100$, $s = 4$ and the population is normally distributed? z or t value: **z value**

Practice Problems 3 with solution

When $n \le 30$ $\qquad \mu = \overline{x} \pm E \qquad E = t\left(\frac{s}{\sqrt{n}}\right)$

1) A random sample of 9 life insurance policy holders showed that the average premiums paid on their life insurance policies was \$340 per year with a standard deviation of \$24. Construct a 95% confidence interval for the population mean. $n = 9$ $\quad \overline{x} = 340$ $\quad s = 24$

Because sample size is less than 30, we use t distribution

$E = t\left(s/\sqrt{n}\right) = 2.306\frac{24}{\sqrt{9}} = 18.45 \qquad \mu = \overline{x} \pm E \qquad \mu = 340 \pm 18.45 \qquad \$321.55 < \mu < \$358.45$

---

2) A company that produces white bread is concerned about the distribution of the amount of sodium in its bread. The company takes a simple random sample of 25 slices of bread and computes the sample mean to be 100 milligrams of sodium per slice. Construct a 90% confidence interval for the unknown mean sodium level assuming that the sample standard deviation is 10 milligrams.

$n = 25$ $\quad \overline{x} = 100$ $\quad s = 10$ ***Because sample size is less than 30, we use t distribution***

$E = t\left(s/\sqrt{n}\right) = 1.711\frac{10}{\sqrt{25}} = 3.42 \qquad \mu = \overline{x} \pm E \qquad \mu = 100 \pm 3.42 \qquad 96.58 < \mu < 103.42$

3) The football coach randomly selected eight players and timed how long it took to perform a certain drill. The times in minutes were: 12, 9, 13, 7, 8, 13, 16, 10. Assuming that the times follow a normal distribution, find a 90% confidence interval for the population mean. $n = 8$ $\quad \overline{x} = 11$ $\quad s = 3.02$

Because sample size is less than 30, we use t distribution

$E = t\left(s/\sqrt{n}\right) = 1.895\frac{3.02}{\sqrt{8}} = 2.02 \qquad \mu = \overline{x} \pm E \qquad \mu = 11 \pm 2.02 \qquad 8.98 < \mu < 13.02$

4) The actual time it takes to cook a ten-pound turkey is a normally distributed. Suppose that a random sample of 9 ten-pound turkeys is taken. Given that an average of 2.9 hours and a standard deviation of .24 hours was found for a sample of 9 turkeys, calculate a 95% confidence interval for the average cooking time of a ten-pound turkey. $n = 9 \quad \overline{x} = 2.9 \quad s = 2.4$

Because sample size is less than 30, we use t distribution

$E = t\left(s/\sqrt{n}\right) = 2.306\dfrac{0.24}{\sqrt{9}} = 0.18 \qquad \mu = 2.9 \pm 0.18 \qquad 2.72 < \mu < 3.08$

Estimating based on raw data

Estimating the $\mu =$ average life of Diehard batteries by using **95%** confidence Level when a sample of **6** batteries provides these data 48,54,57,45, 56,52

Solution by Formula

Hint: to use the formula, you need to calculate $\overline{\mathbf{x}} = ?\,.\, s = ?$ $\overline{\mathbf{x}} = \mathbf{52}$ **months** . $s = \mathbf{4.69}$ **months**

($n \le 30$) **(for t- value use table page 4)** $df = 6-1 \quad t_{\alpha/2} = 2.571 \quad E = 2.571\dfrac{4.69}{\sqrt{6}} \quad E = 4.92$

$\mu = 52 \pm 4.92 \qquad 47.08 < \mu < 56.92$

Solution by TI 83/84 Calculator

input data in L1 then, ($n \le 30$) $\rightarrow$ **TI-83/84** ***stat*** $\longrightarrow$ ***tests*** $\longrightarrow$ ***Option 8***

$$E = (UB - LB)/2 = (56.92 - 47.08)/2 = 4.92$$

Estimating one population Mean $\mu = \overline{x} \pm E$

a) What do we estimate? Population mean (μ) or sample mean ($\overline{x}$) or both?

b) What is the point estimate?

c) What is the confidence level?

d) What is the criteria of t-distribution?

e) Under what condition we use t-distribution?

f) What is the formula for degree of freedom $df = ?$

g) What is the formula for margin of error?

h) Where can you find the t table and under what condition you will be using this table? By usinf

i) What is the width of a confidence interval?

j) How can we use the upper and lower boundaries of a confidence interval to find point estimate?

k) How can we use the width of a confidence interval to find margin of error?

l) How to use **TI calculator** to find the boundaries of a confidence interval when we use **normal distribution**?

m) How to use **TI calculator** to find the boundaries of a confidence interval when we use **t-distribution**?

C) Important properties about the **relationship** of sample size and **confidence level** and increase, ***decrease***, of

$$E = z\frac{\sigma}{\sqrt{n}}.$$

a) As the ***sample size (n) decreases***, the ***margin of error (E) increases***

b) ***As the confidence level (C) decreases***, the ***margin of error (E) decreases***

Estimating one population **Mean** $\mu = \overline{x} \pm E$

a) What do we estimate? Population mean (μ) or sample mean ($\bar{x}$) or both? Population mean (μ)

b) What is the point estimate? $\overline{x}$

c) What is the confidence level? **the percentage of probability, or certainty**, that the confidence interval would contain the true population parameter when you draw a random sample many times.

d) What is the criteria of t-distribution? Like the normal distribution, the t-distribution has **a smooth shape**. Like the normal distribution, the t-distribution is symmetric. If you think about folding it in half at the mean, each side will be the same. Like a standard normal distribution (or z-distribution), the t-distribution has a mean of zero.

e) Under what condition we use t-distribution? When $n \le 30$

f) What is the formula for degree of freedom $df = ?$ $df = n-1$

g) What is the formula for margin of error? $E = t\frac{s}{\sqrt{n}}$

h) Where can you find the t table and under what condition you will be using this table? $n \le 30$

i) What is the width of a confidence interval? $= UB - LB$

j) How can we use the upper and lower boundaries of a confidence interval to find point estimate? $\dfrac{UB + LB}{2}$

k) How can we use the width of a confidence interval to find margin of error? $\dfrac{UB - LB}{2}$

l) How to use **TI calculator** to find the boundaries of a confidence interval when we use **normal distribution**?

stat $\longrightarrow$ *tests* $\longrightarrow$ ***Option 7(Z-interval)***

m) How to use **TI calculator** to find the boundaries of a confidence interval when we use **t-distribution**?

stat $\longrightarrow$ *tests* $\longrightarrow$ ***Option 8(t-interval***

t distribution for small sample $n \leq 30$

df = n-1	<-------------------- alpha α -------------------->						
2-Tailed	**0.40**	**0.30**	**0.20**	**0.10**	**0.05**	**0.02**	**0.01**	**0.005**
1-Tailed	**0.20**	**0.15**	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**	**0.0025**
Conf. Levl.	**60%**	**70%**	**80%**	**90%**	**95%**	**98%**	**99%**	**99.5%**
1	1.376	1.963	3.078	6.314	12.706	31.821	63.656	127.321
2	1.061	1.386	1.886	2.920	4.303	6.965	9.925	14.089
3	0.978	1.250	1.638	2.353	3.182	4.541	5.841	7.453
4	0.941	1.190	1.533	2.132	2.776	3.747	4.604	5.598
5	***0.920***	***1.156***	***1.476***	***2.015***	***2.571***	***3.365***	***4.032***	***4.773***
6	0.906	1.134	1.440	1.943	2.447	3.143	3.707	4.317
7	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.029
8	0.889	1.108	1.397	1.860	2.306	2.896	3.355	3.833
9	0.883	1.100	1.383	1.833	2.262	2.821	3.250	3.690
10	***0.879***	***1.093***	***1.372***	***1.812***	***2.228***	***2.764***	***3.169***	***3.581***
11	0.876	1.088	1.363	1.796	2.201	2.718	3.106	3.497
12	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.428
13	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.372
14	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.326
15	***0.866***	***1.074***	***1.341***	***1.753***	***2.131***	***2.602***	***2.947***	***3.286***
16	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.252
17	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.222
18	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.197
19	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.174
20	***0.860***	***1.064***	***1.325***	***1.725***	***2.086***	***2.528***	***2.845***	***3.153***
21	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.135
22	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.119
23	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.104
24	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.091
25	***0.856***	***1.058***	***1.316***	***1.708***	***2.060***	***2.485***	***2.787***	***3.078***
26	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.067
27	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.057
28	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.047
29	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.038
30	***0.854***	***1.055***	***1.310***	***1.697***	***2.042***	***2.457***	***2.750***	***3.030***
n>30 $\Rightarrow$ *Z*	***0.842***	***1.036***	***1.282***	***1.645***	***1.96***	***2.326***	***2.576***	***2.807***