Point and Interval Estimation for population mean (large sample $n > 30$)

Why to estimate? Due to our limited resources (Time, Money, manpower, destruction of tested subjects, widely scattered data, hardly accessible subjects).

What is the confidence interval?

In statistics, a confidence interval is a range of values that is determined using observed data, calculated at a desired confidence level that may contain the true value of the parameter being studied. The **confidence level**, for example, a 95% confidence level, relates to how reliable the estimation procedure is, not the degree of certainty that the computed confidence interval contains the true value of the parameter being studied. Confidence intervals for **population mean** typically written as **(sample mean) ± (error)** such as $\mu = \overline{x} \pm E$. The range can be written as an actual value or a percentage. It can also be written as simply the range of values. For example, the following are all equivalent confidence intervals:

Example: Using Formula $\mu = \overline{x} \pm E$ $\quad \overline{x}$ **= Point estimate (Sample Mean) ,** **E = Margin of error**

1. Estimate the **average** life of Diehard batteries in months? $\mu = 50 \pm 10$ or $40 < \mu < 60$ or $(40, 60)$
2. Estimate the **average** waiting time at a store register in minutes? $\mu = 4 \pm 2$ or $2 < \mu < 6$ or $(2, 6)$
3. Estimate the **average** time to finish the final exam in minutes? $\mu = 75 \pm 15$ or $60 < \mu < 90$ or $(60, 90)$

Learning Objectives

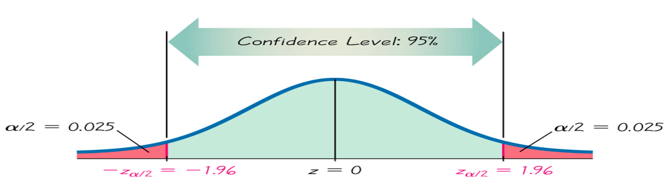
What do we estimate? **Population Mean** ($\mu = ?$ **)**
Know all the new **terminologies** and related **notations** (Point estimate $\overline{x}$, Margin of error)
Know in estimating **population mean** ($\mu = ?$) when to use **normal distribution** versus **t- distribution**.
Know how to use TI (**option 7**) or (**formula** $\mu = \overline{x} \pm E$) to estimate **population mean** ($\mu = ?$).

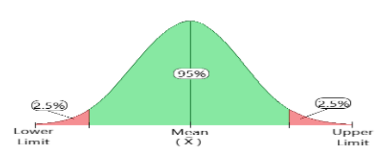
Definitions:

Point estimate: Sample statistics such as $(\overline{x})$

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 $(\alpha = 5\%)$.





Margin of error: (also called error, error bound or maximum error) is the maximum likely difference observed between point estimates $(\overline{x})$ and population parameter (μ).

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

Estimating One Population **Mean** $\mu = \bar{x} \pm E$		
$\bar{X}$ = **Point** estimate (Sample Mean)	E = Margin of error(error bound)	
Margin of Error	If $n > 30$ then use $z_{\alpha/2}$, from table on page 2 $\quad E = z_{\alpha/2}\frac{s}{\sqrt{n}}$	
Interval Estimate	$\mu = \bar{x} \pm E$	$\mu = \bar{x} \pm E$
TI-83/84	*stat* → *tests* → *Option 7(Z-interval)*	*stat* → *tests* → *Option 8(t-interval)*
Important relationships	**Width** (difference between upper and lower bounds) $= 2E = UB - LB$ **Margin of Error** $= E = (UB - LB)/2$ **Point Estimate** $= \bar{x} = (UB + LB)/2$	

Confidence Level	Out Side Area On left or right Cut-off Point	Z - Value (±) Critical Value $= Z_{\alpha/2}$
99%	.005	± 2.5758
98%	.01	±2.3263
97%	.015	±2.1701
96%	.02	±2.0537
95%	.025	**±1.9600**
94%	.03	±1.8808
92%	.04	±1.7507
90%	.05	±1.6450

Practice Problems with solutions

a) **Find the margin of error** $E = Z_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$ **for the following problems by using the z-table on this page.**

1) Sample size $n = 36$, $s = 4$ and 95% confidence level? So $E = 1.96\frac{4}{\sqrt{36}} = 1.31$

2) Sample size $n = 49$, $s = 14$ and 99% confidence level? Answer: **5.1516**
3) Sample size $n = 81$, $s = 36$ and 90% confidence level? Answer: **6.58**
4) Sample size $n = 100$, $s = 15$ and 95% confidence level? Answer: **2.94**
5) Sample size $n = 64$, $s = 24$ and the 90% confidence level? Answer: **4.935**

Check next page for solution

2) Sample size $n = 49$, $s = 14$, and 99% confidence level? ***Z-value*** when *n > 30* $\quad \mathrm{E} = 2.5758 \frac{14}{\sqrt{49}} = 5.1516$

3) Sample size $n = 81$, $\sigma = 36$, and 90% confidence level? ***Z-value*** when *n > 30* $\quad \mathrm{E} = 1.645 \frac{36}{\sqrt{81}} = 6.58$

4) Sample size $n = 100$, $s = 15$, and 95% confidence level? ***Z-value*** when *n > 30* $\quad \mathrm{E} = 1.96 \frac{15}{\sqrt{100}} = 2.94$

5) Sample size $n = 64$, $s = 24$, and 90% confidence level? Z-***value*** when *n > 30* $\quad \mathrm{E} = 1.645 \frac{24}{\sqrt{64}} = 4.935$

Important: Notice that **higher confidence levels** correspond to **larger z-values**, which leads to **wider confidence intervals**. This means that, for example, a 95% confidence interval will be wider than a 90% confidence interval for the same set of data.

Practice Problems with solutions

1) A random sample of 36 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. $n = 36 \qquad \overline{x} = 340 \qquad s = 24$

Because sample size n > 30 , we use normal distribution

$$E = z\left(s / \sqrt{n}\right) = 1.96 \frac{24}{\sqrt{36}} = 7.84 \qquad \mu = 340 \pm 7.84 \qquad \$332.16 < \mu < \$347.84$$

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 64 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 48 milligrams.

$n = 64 \qquad \overline{x} = 100 \qquad s = 48$ ***Because sample size n > 30 , we use normal distribution***

$$E = z\left(s / \sqrt{n}\right) = 1.645 \frac{48}{\sqrt{64}} = 9.87 \qquad \mu = 100 \pm 9.87 \qquad 90.13 < \mu < 109.87$$

3) In estimating the population mean, we have these answers (10, 40), Find the **width**, **margin of error** and **the point estimate**

Use the formula from page 2

.
Solution:

Width = $40 - 10 = 30$ **Margin of error** = $E = \frac{40-10}{2} = 15$ **Point estimate** $= \overline{x} = \frac{40+10}{2} = 25$

Rewrite them back $\mu = 25 \pm 15 = (10, 40)$

4) In estimating the population mean, we have these answers (2, 14), Find **width**, **margin of error** and **the point estimate. Use the formula from page 2**

.

Solution:

Width = $14-2=12$ **Margin of error** = $\frac{14-2}{2}=6$ Point estimate $\bar{x}=\frac{14+2}{2}=8$

Rewrite them back $\mu = 8 \pm 6 = (2,14)$

Using TI 83/84 calculator

Estimating the $\mu=$average life of Diehard batteries by using **95%** confidence Level. A sample of **64** batteries has $\mathbf{\bar{x}=50}$ **months** .From prior study, we know $\sigma=$**10 months**

Solution by Formula

n > 30 , $\rightarrow E=z_{\alpha/2}\frac{\sigma}{\sqrt{n}}=1.96\frac{10}{\sqrt{64}}$ $E=2.45$ $\mu=50\pm2.45$ $47.55<\mu<52.45$

By 95% confidence, the average life of Diehard batteries is between 47.55 to 52.45 months

Solution by TI 83/84 Calculator

σ **known** $\rightarrow$ **TI-83/84** ***stat→tests→Option 7***

```
ZInterval
 Inpt:Data
 σ:10
 x̄:50
 n:64
 C-Level:.95
 Calculate
```

```
ZInterval
 (47.55,52.45)
 x̄=50
 n=64
```

$E = (UB - LB)/2 = (52.45 - 47.55)/2 = 2.45$

Questions: Answers on the last page.

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

b) Why do we need to estimate? Cite some reasons?

c) What is the point estimate?

d) What is the confidence level?

e) What is the margin of error formulas for it?

f) Where can you find the z table and under what condition you will be using this table?

g) What is the width of a confidence interval?

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

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

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

Answers:

a) Population mean (μ) or sample mean ($\bar{x}$) : Population mean (μ)

b) Why do we need to estimate? Because population is big, to save money and time

c) What is the point estimate? $\bar{x}$

d) What is the confidence level? how reliable the estimation procedure

e) What is the margin of error formulas for it? $E = z_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$

f) Where can you find the z table and under what condition you will be using this table? P 3. When n>30

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

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

$$(UB + LB) / 2$$

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

$$E = (UB - LB) / 2$$

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

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

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