UNIT No:5 Estimation Theory

Syllabus: Unbiased & efficient estimates, Point estimates & interval estimates. Confidence interval for means, Confidence interval for proportions, Confidence interval for differences & sums of mean & proportions.

* "Estimate" → We mean judgement or opinion of approximate size or amount.

* "Estimation" → In statistics, estimation refers to the process by which one makes inference about a population, based on information obtained from a sample.

* "Unbiased Estimator" → A statistic is called an unbiased estimator of a population parameter, if the mean or expectation of statistics is equal to the mean of parameter of popul^n

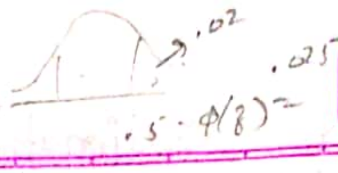
* Types of Estimation *

An estimate of populat^n parameter may be express in 2-ways.

Point Estimate | Interval Estimate

(1) **Point Estimate** :- A pt estimate of populat^n parameter is a single value of a statistics

Ex: Distance of b/w College & house → (18 km).
Cello griper ~~ga~~ pen (10/5 Rs)



1) Point Es

2) <u>Interval Estimate</u> :- An interval estimate is defined by two no's, b/w which a population parameter is said to lie.

Ex Predictⁿ of % will get in 10th std → 90 ↔ 95%.
or " " " " " in 12th std → 80% to 85%.

* Confidence Interval Estimates of Population Parameters *.

Table Learn:

Confidence Level:	99.73%	99%	98%	96%	95.45%
Z_c : (critical values)	3.00	2.58	2.33	2.05	2.00

Confidence Level:	95%	90%	80%	68.27%	50%
Z_c	1.96	1.645	1.28	1.00	0.6745

* Confidence Interval for Means :-

For Large Sample ($n \geq 30$):

Let Mean $= \bar{X}$, n = no. of samples, σ = std. deviatⁿ, then

Confidence interval for Mean is gⁿ by

$$C.I = \bar{X} \pm Z_c \frac{\sigma}{\sqrt{n}} = [\bar{X} - E, \bar{X} + E],$$

where $E = Z_c \left[\frac{\sigma}{\sqrt{n}}\right]$

Confidence Int → (guarentee | type) → Ex 2nd hand moble phone

(In this sampling is from an ∞-populatn or sampling is with replacement from finite populatn)

and :

$$C.I = \bar{x} \pm z_c \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} , = [\bar{x} - E, \bar{x} + E];$$

where $E = z_c \frac{\sigma}{\sqrt{n}} \left(\frac{N-n}{N-1}\right)^{1/2}$

(Sampling is without replacement from a populatn of finite size N).

For t-distributn Wened table + No. of degree of freedom ie $\nu = n-1$

For sample is small ($n < 30$) In this case, we use t-distribution to obtain confidence levels

$$C.I = \bar{x} \pm t_c \frac{\hat{s}}{\sqrt{n-1}} \quad \because \hat{s} \text{ is std. deviatn of sample.}$$

$$C.I = \left[\bar{x} - t_c \frac{\hat{s}}{\sqrt{n-1}}, \bar{x} + t_c \frac{\hat{s}}{\sqrt{n-1}}\right]$$

Q Measurements of a diameter of a random sample of 200 ball bearing made by a certain machine during one week showed a mean of 0.824 inch & standard deviatn of 0.042 inch. Find (a) 95% (b) 99% Confidence limits for mean diameter of all ball bearings.

Soln

Since ; $n = 200$ is large

Mean ; $\bar{x} = 0.824$ & $\sigma = 0.042$

(a) 95% Confidence limit / Interval are

Scanned with CamScanner

$P(z<0) = .025$

$$C.I = \bar{X} \pm z_c \frac{\sigma}{\sqrt{n}} \qquad 95\% \to z_c = 1.96$$

$$= 0.824 \pm 1.96\left(\frac{0.042}{\sqrt{200}}\right)$$

$$= 0.824 \pm 0.0058 \text{ inches}$$

$$CI = [0.824 - 0.0058,\ 0.824 + 0.0058]$$

(b) 99% Confidence interval are

$$CI = \bar{X} \pm z_c \cdot \frac{\sigma}{\sqrt{n}} \qquad 99\% \to z_c = 2.58$$

$$= 0.824 \pm 2.58\left(\frac{0.042}{\sqrt{200}}\right)$$

$$= 0.824 \pm 0.0077 \text{ inch}$$

$$CI = [0.824 - 0.0077,\ 0.824 + 0.0077]$$

(c) Same que solve for 99.73%.

$$C.I = \bar{X} \pm z_c \frac{\sigma}{\sqrt{n}} = 0.824 \pm 3 \cdot \left(\frac{0.042}{\sqrt{200}}\right)$$

$$C.I = 0.824 \pm 0.0089 \text{ inch}$$

$$CI = [0.824 - 0.0089,\ 0.824 + 0.0089]$$

Note: Confidence Int. for mean:

$$95\% \longrightarrow C.I = \bar{x} \pm t_{0.975}\left(\frac{\hat{s}}{\sqrt{n-1}}\right)$$

$$99\% \longrightarrow CI = \bar{x} \pm t_{0.995}\left(\frac{\hat{s}}{\sqrt{n-1}}\right).$$

Note: For 95% C.I for T is g^n by $C.I = \bar{x} \pm t_{0.975}\left(\frac{\bar{s}}{\sqrt{n-1}}\right)$

H.W. Q. A sample of 10 measurements of a diameter of a sphere gave a mean $\bar{x} = 4.38$ inches and a standard deviation $(s) = 0.06$ inches. Find (a) 95% (b) 99% (interval) C.I for actual diameter.

Here
$n = 10$
$\bar{x} = 4.38$
$\bar{s} = 0.06$

$\nu = n-1$
$= 10-1$
$\nu = 9$

Sol^n(a) The 95% of Confidence limit are g^n by

$$C.I = \bar{x} \pm t_{0.975}\left(\frac{\bar{s}}{\sqrt{n-1}}\right)$$

$\Rightarrow$ ~~C.I = 46.4~~

$$\Rightarrow C.I = 4.38 \pm (2.26)\left(\frac{0.06}{\sqrt{9}}\right)$$

$$= 4.38 \pm 0.0452 \text{ inch.}$$

$\therefore$ We can be confident that true mean lies b/w $4.38 - 0.0452 = 4.335$ inches &
$4.38 + 0.0452 = 4.425$ inches

i.e. $C.I = [4.335, 4.425]$

(b) $\rightarrow$ 99% Confi. limit are $\bar{x} \pm t_{0.995}\left(\frac{\bar{s}}{\sqrt{n-1}}\right)$

$= 4.38 \pm 0.0650$ inches

$C.I = (4.315, 4.445)$.

Q. Determine 99% Confidence interval for the mean of contents of soft drinks bottles, if contents of 7 such soft drinks bottles are 10.2, 10.4, 9.8, 10.0, 9.8, 10.2, 9.6 ml. and std. deviat^n is 0.283.

Sol^n. Let Mean; $\bar{x} = 10$ ml, $n = 7$ & $\bar{s} = 0.283$

Degree of freedom; $\nu = n-1$
$= 7-1 = 6$ $(n < 30)$

Confidence interval for mean is given by

$$C.I = \bar{x} \pm t_c\left[\frac{\bar{s}}{\sqrt{n-1}}\right] = \bar{x} \pm t_{0.995}\left[\frac{\bar{s}}{\sqrt{n-1}}\right]$$

9.572, 10.428

$$C.I = 10 \pm (3.707)\left(\frac{0.283}{\sqrt{6}}\right)$$

~~CI = 10 ± 2.65 → [07.35, 12.65]~~

Q Five measurements of the reaction time of an individual to certain stimuli were recorded as 0.28, 0.30, 0.27, 0.33 & 0.31 second. Find (a) 95% (b) 99% Confidence limits for actual mean reaction time.

Solⁿ. Here; $n = 5 \rightarrow$ (use t-distributⁿ) $(n<30)$

Mean; $\bar{x} = \frac{0.28+0.30+0.27+0.33+0.31}{5} = 0.298$

$$S^2 = \frac{\Sigma(x_i-\bar{x})^2}{5} = 0.000456$$

$$\therefore S = 0.02135$$

For 95% Confidence interval for mean;

$$C.I = \bar{x} \pm t_{0.975}\frac{S}{\sqrt{n-1}} = 0.298 \pm \frac{2.78 \times 0.02135}{\sqrt{5-1}}$$

$$CI = 0.298 \pm 0.02968$$

$$0.2684 \leq \mu \leq 0.3276$$

For 99%; $C.I = \bar{x} \pm t_{0.995}\frac{S}{\sqrt{n-1}}$

$$C.I = 0.298 \pm \frac{4.60 \times 0.02135}{\sqrt{4}}$$

* Confidence Interval for proportions *

Suppose statistics is the proportion of successes in sample of size $n \geq 30$ drawn from binomial population in which p is prob. of success, then Confidence interval for population proportions are given by

$$C.I = P \pm z_c \sqrt{\frac{PQ}{n}}$$ → sampling with replacement from finite populatⁿ

or

$$C.I = P \pm z_c \sqrt{\frac{PQ}{n}\left(\frac{N-n}{N-1}\right)}$$ → sampling is without replacement from finite popu. of size 'N'

* Confidence Interval for Sum & differences *

(1) **Sum & difference for mean:-**

(1) Sum

$$C.I = (\bar{x}_1 + \bar{x}_2) \pm z_c \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$\boxed{C.I = \bar{x} \pm z_c\left(\frac{\sigma}{\sqrt{n}}\right)}$$

Large sample Space/Size

(2) Difference

$$C.I = (\bar{x}_1 - \bar{x}_2) \pm z_c \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

(2) **Sum & difference for proportion:-**

(1) Sum

$$C.I = \cancel{(\bar{x}_1 + \bar{x}_2)}\ (P_1 + P_2) \pm z_c \sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}$$

(2) Difference

$$C.I = (P_1 - P_2) \pm z_c \sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}$$

Q. In 40 tosses of a coin, 24 heads were obtained. Find (a) 95% & (b) 99.73% Confidence limits for proportion of heads that would be obtained in an unlimited no. of tosses of the coin.

Sol^n. (a)

At 95% level; $z_c = 1.96$ &

prob. of 24 heads; $p = \frac{24}{40} = 0.6$

& $n = 40$

Confidence limit for proportion of heads;

$C.I = P \pm z_c \sqrt{\frac{p.q}{n}}$; where $q = 1 - p = 1 - 0.6 = 0.4$

$C.I = 0.6 \pm 1.96 \sqrt{\frac{(0.6)(0.4)}{40}}$

$= 0.60 \pm 0.15 = [0.60 - 0.15, 0.60 + 0.15]$

$C.I = [0.45, 0.75]$.

(b) At, 99.73% $z_c = 3.0$ & $p = 0.60$, $q = 0.40$

$C.I = P \pm z_c \sqrt{\frac{P.q}{n}} = 0.60 \pm (3) \sqrt{\frac{0.6 \times 0.4}{40}}$

$= 0.60 \pm 0.23$

$= [0.60 - 0.23, 0.60 + 0.23]$

$C.I = [0.37, 0.79]$

$\bar{x} \pm z_c \frac{\sigma}{\sqrt{n}}$ $\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ $n > 30 \to z$

t

$\bar{x} \pm t_c \frac{s}{\sqrt{n-1}}$

H.W.

Q. A sample poll of 100 voters chosen at random from all voters in the given districts indicates that 55% of them were favor of particular candidate. Find (a) 95% (b) 99% (c) 99.73% confidence limits for proportion of all voters in favor of this candidates.

→

(a) The prob. $p = 55\% = 0.55$; & $q = 1 - p$

$q = 1 - 0.55 = \underline{0.45}$ & here $n = 100$

95% confidence limits for population p

$$C.I = p \pm z_c \sqrt{\frac{p.q}{n}}$$

$95\% \to z_c \to 1.96$

$$C.I = 0.55 \pm (1.96) \sqrt{\frac{(0.45)(0.55)}{100}}$$

$$C.I = 0.55 \pm 0.10$$

$$C.I = [0.55 - 0.10, 0.55 + 0.10] = [\quad , \quad]$$

(b) 99% ; $z_c = 2.58$

$$C.I = p \pm z_c \sqrt{\frac{pq}{n}} = 0.55 + 2.58 \sqrt{\frac{(0.45)(0.55)}{100}}$$

$$C.I = 0.55 \pm 0.13 = [0.55 - 0.13, 0.55 + 0.13]$$

$$C.I = [\quad , \quad]$$

(c) 99.73% ; $z_c = 3$, confidence limits for p

$$C.I = p \pm z_c \sqrt{\frac{pq}{n}}$$

$$= 0.55 \pm 3 \sqrt{\frac{0.45 \times 0.55}{100}} = 0.55 \pm 0.15$$

$$C.I = [0.55 - 0.15, 0.55 + 0.15]$$

Imp
Ex A sample of 200 bolts from one machines showed that 15 were defective, while a sample of 100 bolts from another machine showed that 12 were defective.

Find: (a) 95% (b) 99% (c) 99.73% Confidence limit for differences in proportions of defective bolts from two machines.

→

Here; for 1st machine, Let A; $n_A = 200$

prob; $P_1 = \frac{15}{200} = 0.75$ & $q_1 = 1 - P_1 = 0.25$

& 2nd machine be; B; $n_B = 100$ &
prob; $P_2 = \frac{12}{100} = 0.12$ & $q_2 = 1 - P_2 = 0.88$.

Now; Confidence limit for difference in proportion;

$$C.I = P_1 - P_2 \pm Z_c \sqrt{\frac{P_1 \cdot q_1}{n_A} + \frac{P_2 \cdot q_2}{n_B}}$$

$$C.I = 0.75 - 0.12 \pm Z_c \sqrt{\frac{(0.75)(0.25)}{200} + \frac{(0.75)(0.25)}{100}}$$

$$= 0.045 \pm Z_c (0.03736)$$

(a) 95% Confi. int. for diff. of proporth; $Z_c = \cancel{2.58}$
$Z_c = 1.96$

$$C.I = 0.045 \pm (1.96)(0.03736)$$
$$= 0.045 \pm 0.0723$$

$$\Rightarrow -0.0273 < P_1 - P_2 < 0.1173$$

(b) 99%; $Z_c = 2.58$

$$\therefore C.I = 0.045 \pm (2.58)(0.03736)$$

$\Rightarrow \quad -0.0514 < P_1 - P_2 < 0.1414$

(c) 99.73% ; $z_c = 3$

$\therefore C.I = 0.045 \pm (3)(0.03736) = 0.045 \pm 0.11208$

$\Rightarrow \quad -0.06708 < P_1 - P_2 < 0.15708$

Imp Ex A sample of 150 brand A light bulbs showed a mean lifetime of 1400 hours & std. deviatn of 120 hours. A sample of 200 brand B light bulbs showed a mean lifetime of 1200 hours & std. deviatn of 80 hours.

Find, (a) 95% (b) 99% confidence limits for difference of mean lifetimes of populatn of brands A & B.

Solⁿ.

For light bulbs 'A' brand

$n_A = 150$; $\bar{X}_A = 1400$ hrs ; $\sigma_A = 120$ hours

For light bulbs 'B' brand,

$n_B = 200$; $\bar{X}_B = 1200$ hrs ; $\sigma_B = 80$ hrs.

Confidence limits for difference of mean lifetime of populatn for brands A & B is given by

$$C.I = \bar{X}_A - \bar{X}_B \pm z_c \sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}$$

Scanned with CamScanner

DELUXE
PAGE NO. :
DATE:

& $z_c = 1.96$

(a) The 95% confidence limits are

$$C.I = 1400 - 1200 \pm (1.96)\sqrt{\frac{(120)^2}{150} + \frac{(80)^2}{100}}$$

$$C.I = 200 \pm 24.8$$

∴ $\cancel{167} < \cancel{\mu} \cancel{<}$

$$175 \; \cancel{167} < \bar{x}_A - \bar{x}_B < \cancel{233} \; 225$$

ie We can be 95% confident that diff. of populatn means lies b/w 175 & 225 hrs.

(b) The 99% & $z_c = 2.58$, confidence limits are

$$C.I = 1400 \pm 1200 \pm (2.58)\sqrt{\frac{(120)^2}{150} + \frac{(80)^2}{100}}$$

$$C.I = 200 \pm 32.6$$

ie $167 < \bar{x}_A - \bar{x}_B < 233$

∴ We can be 99% confident that diff of populatn means lies b/w 167 & 233 hours.

~~Q~~. In a random Sample of 400 adults & 600 teenagers who watched a certain TV program 100 adults and 300 teenagers indicated that they liked it. Construct (a) 95% (b) 99% Confidence limits for difference in proportions of all adults & all teenagers who watched program & liked it.

Soln

(a)

(b

Solⁿ Confidence limits for difference in proportion of two groups are given by

$$C.I = p_1 - p_2 \pm z_c \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

For Adults; prob. of watching television;

$$p_2 = \frac{100}{400} = 0.25 \text{ ; where } n_2 = 400$$

$$q_2 = 1 - p_2 = 1 - 0.25 = 0.75$$

For teenagers ; $n_1 = 600$ &

prob. of watching TV ; $p_1 = \frac{300}{600} = 0.50$

$$q_1 = 1 - p_1 = 0.50$$

(a) For 95% ; $z_c = 1.96$, confidents limits are

$$C.I = p_1 - p_2 \pm z_c \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

$$C.I = 0.50 - 0.25 \pm (1.96) \sqrt{\frac{(0.5)(0.5)}{600} + \frac{(0.25)(0.75)}{400}}$$

$$C.I = 0.25 \pm 0.06$$

ie $0.19 < p_1 - p_2 < 0.31$

∴ We can be 95% confident that true diff. in proportion lies b/w 0.19 & 0.31.

(b) For 99% , $z_c = 2.58$; confident limits are

$$C.I = p_1 - p_2 \pm (2.58) \sqrt{\frac{p_1 \cdot q_1}{n_1} + \frac{p_2 \cdot q_2}{n_2}}$$

$$= 0.25 \pm 0.08 .$$

∴ We can be 99% confident that true diff in proportions lies b/w 0.17 & 0.33 .

* For Small sample *

In difference of Means

→(1) 95% Confidence Interval for diff. of means

$$C.I = x_1 - x_2 \pm t_{0.975} \, \sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$; where

$$\sigma = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}}$$

Since; s^2 is Variance

& $v = n_1 + n_2 - 2$ is free variable (no. of degree of freedom)

→(2) 99% Confidence Interval for diff of means,

$$C.I = x_1 - x_2 \pm t_{0.995} \, \sigma \sqrt{1/n_1 + 1/n_2}$$ where

$$\sigma = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}}$$ &

$v = n_1 + n_2 - 2$ is free variable

s^2 = Variance

Imp

Q. A random sample of size '6' selected from a normal population has mean 25 & variance 4. A second random sample of size 8 selected from normal population has mean 16 & variance 2.5. Construct 95% confidence interval for difference of ~~proportion~~ population means.

Soln

Let; Size of samples; $n_1 = 6$ & $n_2 = 8$
Means; $x_1 = 25$ & $x_2 = 16$
& Variance; $s_1^2 = 4$ & $s_2^2 = 2.5$

For 95% confidence interval for diff. of means given by (Small sample size)

$$C.I = x_1 - x_2 \pm t_{0.975}\, \sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Where;

$$\sigma = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}}$$

$$\Rightarrow \sigma = \sqrt{\frac{6(4) + 8(2.5)}{6+8-2}} = \sqrt{\frac{44}{12}}$$

$$\Rightarrow \sigma = \sqrt{3.667} = 1.914$$

$$\& \quad v = n_1 + n_2 - 2 = 6 + 8 - 2 = 12$$

Now to see $v = 12$ in $t_{0.975}$ (ie 2.18)

$$\therefore C.I = 25 - 16 \pm (2.18)(1.914)\sqrt{\frac{1}{6} + \frac{1}{8}}$$

$$= 9 \pm 4.17252\,(\sqrt{0.2916})$$

$$= 9 \pm 4.17252\,(0.54) = 9 \pm 2.2531$$

$$\therefore 6.7469 < x_1 - x_2 < 11.2531$$

Confidence Interval for

$\boxed{\text{Large Sample}}$ use Z_c

1) **Means** →

$$C.I = \bar{x} \pm Z_c \frac{\sigma}{\sqrt{n}} \rightarrow \text{(With Replacement)}$$

$$C.I = \bar{x} \pm Z_c \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} \rightarrow \text{(Without replacement)}.$$

2) **Proportion** →

$$C.I = P \pm Z_c \sqrt{\frac{pq}{n}} \rightarrow \text{(With R)}$$

$$C.I = P \pm Z_c \sqrt{\frac{pq}{n}} \sqrt{\left(\frac{N-n}{N-1}\right)} \rightarrow \text{(Without R)}$$

(3) **Differences & sum of means** ⇒

$$\cancel{C.I = \bar{x}_1 - \bar{x}_2 \pm Z_c \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$C.I = \bar{x}_1 - \bar{x}_2 \pm Z_c \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$\boxed{\text{Small Sample}}$ use t-distribution

1) **Means** →

$$C.I = \bar{x} \pm t_c \frac{s}{\sqrt{n-1}} \quad ; \quad s = \text{std. deviation}$$

$\nu = n-1$ (No. of freevariables)

$95\% \rightarrow t_{0.975}$ &

$99\% \rightarrow t_{0.995}$

(2) **Differences of means**

$$C.I = \bar{x}_1 - \bar{x}_2 \pm t_c \cdot \sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} ,$$

where $\sigma = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}}$

& $\nu = n_1 + n_2 - 2$

(4) **Differences & sum of proportions** ⇒

$$C.I = (P_1 - P_2) \pm Z_c \sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}} \rightarrow \text{(difference)}$$

$$C.I = (P_1 + P_2) \pm Z_c \sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}} \rightarrow \text{(sum)}.$$

Where $q = 1 - p$.

Confi level →	99.73%	99%	98%	96%	95.45%	95%	90%
Z_c →	3	2.58	2.33	2.05	2	1.96	1.645