Unit - 3
Testing of Hypothesis
[Sampling]

POPULATION : collection of individuals

SAMPLE : A finite subset of population.

PARAMETER & STATISTICS :

- Statistical measures calculated on the basis of population are called parameters (mean μ, var σ^2)
- Statistical measures calculated on the basis of sample are called statistics (mean $\bar{x}$, var S^2)
- A sample statistic is denoted by 't'.

SAMPLING DIST^n : The pbty dist^n of a statistic 't'

STANDARD ERROR : The standard deviation of the sampling dist^n of a statistic.

NULL HYPOTHESIS (H_0) : A hypothesis of no difference (ie; no diff b/w pop^n & sample).

ALTERNATE " (H_1) : A hypoth. which is different from H_1

: A procedure to accept / reject null hypoth. is called Testing of hypoth.

TYPE 1 & TYPE 2 ERRORS :

Reject H_0 when its true [Type 1]
Reject H_0 when its false [Type 2]

$\boxed{\text{Pop} > \text{sam}}$ LT

ONE TAIL & TWO TAIL TEST :

Set $H_0 : \theta = \theta_0$

Suppose, $H_1 : \theta \neq \theta_0$ { [$\theta > \theta_0$ or $\theta < \theta_0$] 2 tail

$H_1 : \theta > \theta_0$ { Right Tail, Sample > Pop^n

$H_1 : \theta < \theta_0$ { Left Tail, Sampl < Pop^n.

CRITICAL REGION (CR) : A region where we reject H_0 is called critical region. or Region of Rejection.

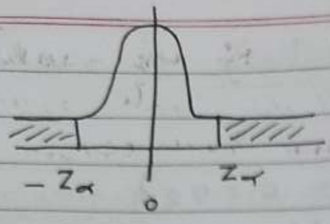
The region complementary to CR is Acceptance Region

CRITICAL VALUE : [Z_α] The value of a statistic Z for which the critical & acceptance region are separated ($\alpha \Rightarrow$ Level of significance)

VALUE OF Z_α

Nature	1%	2%	5%	10%
2 tail	2.58	2.33	1.96	1.645
Right tail	2.33	2.055	1.645	1.28
Left tail	-2.33	-2.055	-1.645	-1.28

([illegible])



LARGE SAMPLE : when size of sample is greater than 30 its called large sample, o/w, small sample.

Procedure for Testing Hypoth :

1) Set H_0
2) Set H_1 (check if its 1 tail or 2)
3) find $|Z|$ (test statistics) & Z_α
4) if $|Z| < |Z_\alpha|$, Accept H_0
   $|Z| > |Z_\alpha|$, Reject H_0

Test 1 :
Test of significance b/w sample proportion & population proportion.

Test statistic $Z = \dfrac{p - P}{\sqrt{\frac{PQ}{n}}}$

$p \rightarrow$ sample proportion
$Q = 1 - P$

$P \rightarrow$ pop^n proportion
$n \rightarrow$ size of sample.

95% confidence limits are :

$$\frac{|p - P|}{\sqrt{\frac{PQ}{n}}} \leq 1.96$$

$$= \left(p - 1.96\sqrt{\frac{PQ}{n}} \, , \; p + 1.96\sqrt{\frac{PQ}{n}} \right)$$

Q. 20% of manufactured product is of top quality. In one day production of 400 articles, only 50 are top quality. Verify hypoth. & find 95% confi. lims.

$P = 20\% = 0.2$ $\quad$ $n = 400$

$p = \frac{50}{400} = 0.125$ $\quad$ $Q = 1 - P = 0.8$

$H_0 : P = p \ (P = 0.2)$

$H_1 : P \neq p$ (2 tail)

Let the level of signifacnce be 5%. (100 − 95%)

$Z_\alpha = 1.96$ [2 tail, 5%]

$$Z = \frac{\{p - P\}}{\sqrt{\frac{PQ}{n}}}$$

$Z = -3.75$

$|Z| > |Z_\alpha|$

∴ Reject H_0

Conclusion $P \neq p$.

∴ There is a significant difference b/w sample & population

1T: claim, comparitive (bigger, grtr)

Large Sample ($n > 30$)

Test-1 : Test of significance of sample & pop. prop.

$$\boxed{z = \frac{p-P}{\sqrt{\frac{PQ}{n}}}}$$

Test 2 :

Q3 Pg 10 (PDF - large sam) typhoid

P = pop^n pro. $= 17.26\% = 0.1726$

$Q = 1 - P = 0.8274$

p = sam. prop. $= \frac{63}{640} = 0.0984$ $\quad n = 640$

$H_0 : p = P$ (Hospital & sufficient)

$H_1 : p < P$ [" ✓ "]

↳ check if p grtr or P.

[LT : pop^n > Sam]

$$z = \frac{p-P}{\sqrt{PQ/n}} = -4.98$$

Let LOS be 5%. [1% or 5% level of signi. if x Gi, any]

$Z_\alpha = 1.645$ (table)

$|Z| > |Z_\alpha|$

Rej. H_0.

H_1 ✓ Hosp. ✓ suff & typh. patient.

Q. P14 (

Q4 Pg 12 (salesman)

$P = 60\% = 0.6$

$Q = 0.4$

$p = \frac{35}{50} = 0.7$ $\quad (n = 50)$

$H_0 : p = P$ (pop = sam)

$H_1 : p > P$ (RT)

$$Z = \frac{p-P}{\sqrt{PQ/n}} = 1.443$$

pop grtr LT

Let LOS be 5%.

$Z_\alpha = 1.645$

$|Z| < |Z_\alpha|$

Accept H_0

Test 2 : Test of signif. of diff b/w 2 sample prop

$$Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$P \to$ pop^n prop

$$\boxed{P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}}$$

Q.P14 (cities)

2 cities, 2% es $\therefore$ 2Tail

$n_1 = 900$	$n_2 = 1600$
$p_1 = 20\% = 0.2$	$p_2 = 18.5\% = 0.185$

$H_0 : p_1 = p_2$ [same prop)

$H_1 : p_1 \neq p_2$

$\hookrightarrow$ 2T coz no C word.

$P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = 0.1904$ $\quad q = 0.8096$

$$Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = 0.92$$

Let LOS be 5%.

$Z_\alpha = 1.96$

$|Z| < |Z_\alpha|$

accept H_0 (no signif. difference)

P2 grtr LT

classmate
Date
Page

Q2 Pg15 (smokers)

$n_1 = 1600$ $\quad\quad$ $n_2 = 900$

$p_1 = 15.5\% = 0.155$ $\quad\quad$ $p_2 = 0.2$

$H_0 : p_1 = p_2$

$H_1 : p_1 < p_2$ $\longrightarrow$ (LT) $\quad$ c.word, ∴ 1T.

$$Z = \frac{p_1 - p_2}{\sqrt{PQ(1/n_1 + 1/n_2)}} = -2.87$$

Let LOS = 5%.

$Z_\alpha = -1.645$

$|Z| > |Z_\alpha|$

Rej. H_0. $\quad\quad$ (UG < PG) smokers.

Q3 P16 (tea)

$n_1 = 1000$ $\quad\quad$ $n_2 = 1200$

$p_1 = \frac{800}{1000} = \frac{4}{5} = 0.8$ $\quad\quad$ $p_2 = \frac{800}{1200} = \frac{2}{3} = 0.67$

$H_0 : p_1 = p_2$

$H_1 : p_1 > p_2$ [RT]

$$P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = 0.7236.$$

LOS = 1%.

$Z = 6.82$

$|Z| > |Z_\alpha|$

Rej H_0.

Test 5 Test of signif. b/w (sample & pop^n) mean

$$\boxed{Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{\bar{x} - \mu}{s/\sqrt{n}}}$$ (*)

$\bar{x} \rightarrow$ sample mean σ = S.D. of pop^n
$\mu \rightarrow$ pop^n mean $s \rightarrow$ SD of samp

$$(95\% \text{ confid. interval}) = \left(\bar{x} - 1.96\frac{\sigma}{\sqrt{n}},\ \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}\right)$$

$$(99\% \text{ '' '' }) = \left(\bar{x} - 2.58\frac{\sigma}{\sqrt{n}},\ \bar{x} + 2.58\frac{\sigma}{\sqrt{n}}\right)$$

Q A sample of 100 students is taken. The mean ht. of this sample is 160 cm. Can it reasonably regardd that it in pop^n the mean ht is 165 & SD is 10 cm @ 1% LOS.

$n = 100$ $\sigma = 10$
$\bar{x} = 160$
$\mu = 165$

$H_0 : \bar{x} = \mu$
$H_1 : \bar{x} \neq \mu$ (no compare. ∴ 2T)

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = -5$$

$Z_\alpha = 2.58$
$|Z| > |Z_\alpha|$
Rej H_0. ∴ There maybe a significant difference

Q The mean life time of sample of 45 bulbs is 1550 hrs & SD of 120 hrs. The company manufacturing the bulbs claims that the avg. life of bulbs is 1600 hrs. Is the claim acceptable @ 5% LOS.

Q1 P24

$n = 45$ $\quad s = 120$

$\bar{x} = 1550$ $\quad \mu = 1600$

$H_0 : \bar{x} = \mu$

$H_1 :$ 1T (w.r.t) $\mu > \bar{x}$ (LT)

$$z = \frac{\bar{x} - \mu}{s/\sqrt{n}} = -2.79$$

$z_\alpha = -1.645$

$|z| > |z_\alpha|$

reject H_0 avg. life of bulb is

Q P2

Test-4 Test of signif of diff b/w 2 means

$$\boxed{z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}}$$ (*)

If $\sigma_1 = \sigma_2 = \sigma$ } $\rightarrow$ $z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$

$$\boxed{\sigma^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2}}$$ (*)

Q P

$\bar{x}_1 \rightarrow$ Ist sample mean. $\quad \sigma_1 =$ SD of Ist popⁿ

$\bar{x}_2 \rightarrow$ IInd sample mean. $\quad \sigma_2 =$ SD of IInd popⁿ

$s_1 \rightarrow$ SD of Ist sample

$s_2 \rightarrow$ SD of IInd sample.

$H_0 : \bar{x}_1 = \bar{x}_2$ (2 sample mean equal)

hrs 2
e bulbs
s the
f bulb is ↑
pop^n
pop^t

x_2 gets LT

Q P24 (inches)

$n_1 = 1000$ $\quad$ $n_2 = 2000$

$\overline{x_1} = 67.5$ $\quad$ $\overline{x_2} = 68$

$\sigma = 2.5$

$H_0 : \overline{x_1} = \overline{x_2}$

$H_1 : \overline{x_1} \neq \overline{x_2}$ $\quad$ (2T)

$$Z = \frac{\overline{x_1} - \overline{x_2}}{\sigma \left[\frac{1}{n_1} + \frac{1}{n_2}\right]} = 5.96$$

Let LOS be 5%.

$Z_\alpha = 1.96$

$|Z| > |Z_\alpha|$, Rej H_0.

Q P29 (men hts)

1st sample (eng men) $\quad$ 2nd sample [Am. men]

$n_1 = 6400$ $\quad$ $n_2 = 1600$

$\overline{x_1} = 170$ $\quad$ $\overline{x_2} = 17.2$

$s_1 = 6.4$ $\quad$ $s_2 = 6.3$

$H_0 : \overline{x_1} = \overline{x_2}$

$H_1 : \overline{x_2} > \overline{x_1}$ $\quad$ [LT]

$$Z = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = 11.32$$

Let LOS 5%.

$Z_\alpha = 2.33$

$|Z| > |Z_\alpha|$, Rej H_0.

Q P9 30

$n_1 = 32$ $\quad$ $n_2 = 36$

$\overline{x_1} = 72$ $\quad$ $\overline{x_2} = 70$

$s_1 = 8$ $\quad$ $s_2 = 6$

$H_0 : \overline{x_1} = \overline{x_2}$

$H_1 : \overline{x_1} > \overline{x_2}$ $\quad$ [RT]

Small Sample ($n < 30$)

Test 1 : Test of signif. of sample & pop^n mean

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n-1}}$$

$\bar{x} \to$ samp. mean
$\mu \to$ pop mean
$s \to$ sam. SD.

- degree of freedom $\gamma = n-1$

Q P2 (machine)

$\mu = 0.025$ $\quad\bar{x} = 0.024$

$n = 10$ $\quad s = 0.02$

$H_0 : \quad \bar{x} = \mu$

$H_1 : \quad \bar{x} \neq \mu$ (2T)

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n-1}} = -1.5$$

2T	0.05 value
1T	0.10 "

dof $\nu = n-1 = 9$

tab t = 2.26

$|t| < |\text{tab } t|$ Accept H_0.

Q P5

$\mu = 146.3$ $\quad n = 22$ $\quad \bar{x} = 153.7$ $\quad s = 17.2$

$H_0 : \quad \bar{x} = \mu$

$H_1 : \quad \bar{x} > \mu$ (RT)

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n-1}} = 1.97$$

dof $\nu = n-1 = 21$

tab t $= 1.72$

$|t| >$ tab t

Rej H_0

After ad, sales ↑, ∴ successful

Q A ce
foll
Can

Q. A certain injection given to 12 patients resulted in the following increase of BP 5, 2, 8, -1, 3, 0, 6, -2, 1, 5, 0, 4. Can it be concluded that the injection ↑es the BP.

x	x^2
5	25
2	4
8	64
-1	1
3	9
0	0
6	36
-2	4
1	1
5	25
0	0
4	16
Σ 31	185

$\bar{x} = \frac{\Sigma x}{n} = \frac{31}{12} = 2.58$ $\quad n = 12$

SD = ? we need var(x)

$s^2 = \frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2 = \frac{185}{12} - \left(\frac{31}{12}\right)^2$

$\downarrow$

$var(x) = E(X^2) - [E(X)]^2$

$= 8.74$

$s^2 = 8.74$

$s = \sqrt{8.74} = 2.95$

$H_0: \bar{x} = \mu$

$H_1: \bar{x} > \mu$ (RT)

$t = \frac{\bar{x} - \mu}{s/\sqrt{n-1}} \Rightarrow \mu = ?$

$\mu = 0$ (bec. they're telling how much its ↑ing)

$t = \frac{2.58 - 0}{2.95/\sqrt{11}} = 2.85$

dof $\nu = n - 1 = 11$

tab t = 1.8

$|t| >$ tab t

Rej H_0.

Difference

Q P8

x_i	$d_i = x_i - A$	d_i^2
70	-30	900
120	20	400
110	10	100
101	1	1
88	-12	144
83	-17	289
95	-5	25
98	-2	4
107	7	49
100	0	0
	-28	1912

A : 70 - 110 choose mid value ie, 95 ≃ 100

$x_i = d_i + A$

$\bar{x} = \bar{d} + A$

$\bar{d} = \frac{\Sigma d_i}{n} = \frac{-28}{10} = -2.8$

$\bar{x} = -2.8 + 100$

$\bar{x} = 97.2$

$$s^2 = \frac{\Sigma d_i^2}{n} - \left(\frac{\Sigma d_i}{n}\right)^2$$

$$= \frac{1912}{10} - \left(\frac{-28}{10}\right)^2 = 191.2 - (-2.8)^2$$

(var) $s^2 = 183.36$

(SD) $s = \sqrt{183.36} = 13.36$

$H_0 : \bar{x} = \mu$

$H_1 : \bar{x} \neq \mu$ 2T

$\mu = 100$ (G)

classmate

Date

Page

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n-1}} = -0.62$$

dof $\nu = 9 \Rightarrow$ tab t =

Test-2 : Test of significance of the difference b/w 2 sample mean.

large

$$\boxed{t = \frac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}}$$

where

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

3 100

dof $= n_1 + n_2 - 2$

If $n_1 = n_2 = n$ then,

small.

$$\boxed{t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2 + s_2^2}{n-1}}}}$$

dof $2n-2$

Q 2 Independant samples from nomal population is given by

sample	size	mean	SD
1	16	23.4	2.5
2	12	24.9	2.8

$8)^2$

Is the difference between the mean significant.

Sol

$n_1 = 16 \quad \bar{x}_1 = 23.4 \quad s_1 = 2.5$

$n_2 = 12 \quad \bar{x}_2 = 24.9 \quad s_2 = 2.8$

$H_0 : \bar{x}_1 = \bar{x}_2$

$H_1 : \bar{x}_1 \neq \bar{x}_2$ (2T)

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n-1}} = -0.62$$

dof $\nu = 9 \Rightarrow$ tab t =

Test-2 : Test of significance of the difference b/w 2 sample mean.

large

$$\boxed{t = \frac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}}$$

where

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

dof $= n_1 + n_2 - 2$

small.

If $n_1 = n_2 = n$ then, $\boxed{t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2 + s_2^2}{n-1}}}}$ dof $2n-2$

Q 2 Independant samples from nomal population is given by

sample	size	mean	SD
1	16	23.4	2.5
2	12	24.9	2.8

Is the difference between the mean significant.

Sol

$n_1 = 16$ $\quad \bar{x}_1 = 23.4$ $\quad s_1 = 2.5$

$n_2 = 12$ $\quad \bar{x}_2 = 24.9$ $\quad s_2 = 2.8$

$H_0 : \bar{x}_1 = \bar{x}_2$

$H_1 : \bar{x}_1 \neq \bar{x}_2$ (2T)

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

39, 40

$$\sigma = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1+n_2-2}} = 2.73$$

now substitute to find t.

$t = -1.44$ $\quad dof = n_1+n_2-2 = 2.6$

$\hookrightarrow |t| < |tab\,t| \quad tab\,t = 2.06$

Accept H_0

Q 2 horses A & B were tested according to time to run a particular race

Horse A	28	30	32	33	33	29	34
Hors B	29	30	30	24	27	29	

Test whether horse A is running faster than horse B

× dir. so calc. $\quad A = 33 \quad B = 30$

⇒

x_1	$d_1 = x_1 - A$	d_1^2	x_2	$d_2 = x_2 - B$	d_2^2
28	-5	25	29	-1	1
30	-3	9	30	0	0
32	-1	1	30	0	0
33	0	0	24	-6	36
33	0	0	27	-3	9
29	-4	16	29	-1	1
34	1	1		-11	47
	-12	52			

$\bar{d_1} = \frac{\Sigma d_1}{n_1} = \frac{-12}{7} = -1.7.$

$Mean_1 = \bar{x_1} = A + \bar{d_1} = 33 - \frac{12}{7} = 31.28$

$$s_1^2 = \frac{\Sigma d_1^2}{n_1} - \left(\frac{\Sigma d_1}{n_1}\right)^2 = \frac{52}{7} - \ldots = 4.48$$

for small sample LT&RT table
value are same.

For IInd sample,

$\overline{d_2} = \frac{\Sigma d_2}{n_2} = \frac{-11}{6} = -1.83$

$\overline{x_2}$ = mean 2 = $B + \overline{d_2} = 30 - 1.83 = 28.16$

s_2^2 = var = $\frac{\Sigma d_2^2}{n_2} - \left(\frac{\Sigma d_2}{n}\right)^2 = \frac{47}{6} - \left(\frac{-11}{6}\right)^2$

$= 4.47$

$s_2 = 2.11$

$H_0: \overline{x_1} = \overline{x_2}$

$H_1: \overline{x_1} > \overline{x_2}$ (RT)

$$\sigma = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}} = 2.29$$

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = 2.24$$

dof = $n_1 + n_2 - 2 = 7 + 6 - 2 = 11$

tabt = 1.80

$|t| > |tabt|$

Rej H_0 (A > B) Horse.

→ ek hi sample ke 2 chijein

Test-3 [Paired t test] If $n_1 = n_2 = n$ & the pair values of x_1 & x_2 are associated in some way then

$$t = \frac{\bar{d}}{s/\sqrt{n-1}} \qquad d = \bar{x} - \bar{y}$$

$$d_i = x_i - y_i$$

dof : $n-1$

$$\boxed{s^2 = \frac{\Sigma d^2}{n} - \left(\frac{\Sigma d}{n}\right)^2}$$

Q. IQ tests were administrated to 5 persons before & after training the results are given below

Candidate	1	2	3	4	5
IQ b. training	110	120	123	132	125
IQ a. training	120	118	125	136	121

$n = 5$

A.

x_i	y_i	$d_i = y_i - x_i$	d_i^2
110	120	10	100
120	118	-2	4
123	125	2	4
132	136	4	16
125	121	-4	16
		10	140

$$t = \frac{\bar{d}}{s/\sqrt{n-1}}$$

$\bar{d} = ?$ $\quad s = ?$

$$\bar{d} = \frac{\Sigma d_i}{n} = \frac{10}{5} = 2$$

$$s^2 = \frac{\Sigma d_i^2}{n} - \left(\frac{\Sigma d_i}{n}\right)^2 = \frac{140}{5} - 2^2 = 24$$

$s = \sqrt{24} = 4.89$

$$t = \frac{2}{\frac{4.89}{2}} \quad \frac{2 \times 2}{4.89} = 0.817$$

~~test~~ $H_0 : \bar{x} = \bar{y}$ $[\bar{d} = 0]$ → states that no diff b/w before & after train

$H_1 : \bar{x} < \bar{y}$ $(\bar{d} > 0)$ → states that there maybe a change.

~~test~~ dof $= n-1 = 4$

tab $t = 2.13$ (1T)

$|t| < |tab\ t|$

Accept H_0.