PSOT : Mod III

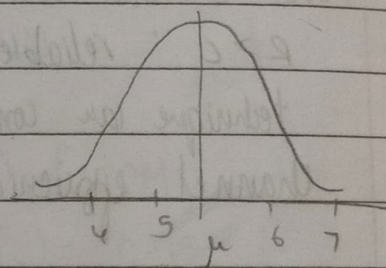
- Population : set / collection / totality of object - animate or inanimate, actual / hypothetical, under study. Thus mainly population consists of sets of numbers, measurements or observations which are of interest
- Size : N. it is the no. of objects / observations in the population. Population is said to be finite / infinite depending on the size N being finite / infinite.

	Population	Sample
Size	N	n
	parameter	statistic
	μ	$\bar{x}$
SD	σ	s
Proportion	P	p

Large
$n \geq 30$
Normal
Z-test
$$Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

Small
$n < 30$
t - distribution
t - test

Sampling

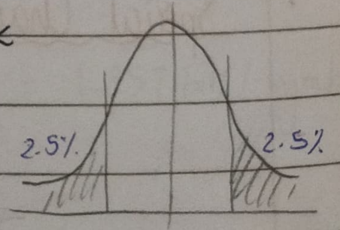


- Estimation i) Point ii) Interval

$$-Z_\alpha < \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} < Z_\alpha \qquad |Z| = \left| \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \right| < Z_\alpha$$

$$\mu - \sigma/\sqrt{n} \; Z_\alpha < \bar{x} < \mu + \sigma/\sqrt{n} \; Z_\alpha$$

- Hypothesis
  H_0 : $\mu = 5.4$ feet
  H_α : $\mu \neq 5.4$



- Level of Significance / Confidence Interval
  ↳ Rejection ↳ Acceptance

95% Confidence
5% Significance

⟶

Page No.
Date

used when % isn't given in ques

Level of significance

	1% (0.01)	5% (0.05)	10% (0.1)
Two tailed test	$\lvert z \rvert = 2.58$	$\lvert z_\alpha \rvert = 1.966$	$\lvert z_\alpha \rvert = 1.645$
Right tailed	$z_\alpha = 2.33$	$z_\alpha = 1.645$	$z_\alpha = 1.28$
Left tailed	$z_\alpha = -2.33$	$z_\alpha = -1.645$	$z_\alpha = -1.28$

	H_0 is accepted	H_0 is rejected
H_0 is true	Correct decision	Type I error
H_0 is false	Type II error	Correct decision

$z = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}}$ where σ : SD

If σ is not known

use $z = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$ where s is the SD of sample

Interval = Confidence limit

$-z_\alpha < z = \dfrac{x - \mu}{\sigma/\sqrt{n}} < z_\alpha$ [same limits from estimation]

- Questions

1 Measurements of the weights of a random sample of 200 small bearings made by a certain machine during one week showed a mean of 0.824 Newton and a SD of 0.042 N. Find 95% confidence limits for the mean weight of all ball bearings

→ Step I : For 95% confidence level, the critical value $z_\alpha = 1.96$

Step II : Since the sample size is large and population standard deviation is not known, the confidence interval is $\bar{x} \pm 1.96\ s/\sqrt{n}$

$0.824 \pm 1.96 \times 0.042/\sqrt{200} = (0.8182, 0.3298)$

$$\left(\mu - \frac{\sigma}{\sqrt{n}} z_\alpha < \bar{x} < \mu + \frac{\sigma}{\sqrt{n}} z_\alpha \right)$$

2 Cardiac patients with implanted pacemakers to control heartbeats. A plastic connector module mounts on top of the pacemaker. A random sample of 75 modules has an average of 0.31 inches. Assuming SD = 0.0015 inches and normal distribution, find 95% confidence interval for the mean size of the connector module.

→ For 95% confidence level the critical value $z_\alpha = 1.96$

Page No.
Date

~~Sample~~ Since the sample is large and population SD is known, the confidence interval is $\bar{x} \pm 1.96\ \sigma/\sqrt{n}$

~~The conf~~ Calculate the limits, $n = 75$, $\sigma = 0.0015$, $\bar{x} = 0.31$

3 A simple random sample of size 65 was drawn in the process of estimating the mean annual income of 950 families of a certain township. The mean and the standard deviation of the sample were found to be ₹4730 and ₹765 respectively. Find a 99% confidence interval for the population mean.

→ Step 1: For 99% confidence level, the critical value $Z_\alpha = 2.58$

Step 2: Since the sample is drawn from a finite population and since $n/N = 65/950 = 0.068$ is greater than 0.05, we use finite population multiplier

$$\sqrt{(N-n)/(N-1)}$$

The ~~conf~~ confidence interval is $\bar{x} \pm 2.58\left(\frac{s}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}}\right)$

4 A machine is set to produce metal plates of thickness 1.5 cm with SD = 0.2 cm. A sample of 100 plates produced by the machine gave an average thickness of 1.52 cm. Is the machine fulfilling the purpose?

→ Null hypothesis $H_0: \mu = 1.5$

Alternative Hypothesis $H_\alpha: \mu \neq 1.5$

Test Statistic: Since the population standard deviation is given

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{1.52 - 1.5}{0.2/\sqrt{100}} = 1$$

Level of significance: $\alpha = 0.05$

Critical value: The value of Z_α at 5% level of significance from the table $= 1.96$

Decision: Since the computed value of $|z| = 1$ is less than critical value $Z_\alpha = 1.96$, the null hypothesis is accepted.

$|z| \leq Z_\alpha$

Page No.
Date

5 A random sample of 50 items gives the mean 6.2 and variance 10.24. Can it be regarded as drawn from a normal population with mean 5.4 at 10% los?

→ $n = 50$ $\quad \bar{x} = 6.2$ $\quad s = \sqrt{\sigma^2} = \sqrt{10.24}$ $\quad \mu = 5.4$

los : $\alpha = 0.1$

$$z = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{6.2 - 5.4}{\sqrt{10.24}/\sqrt{50}} = 1.76$$

los 10% or confidence : $z_\alpha = 1.64$: from table

Decision : $|z| \not< z_\alpha$: not satisfied.

Reject the H_0 (H not)

The sample is not drawn from the population with mean 5.4

6 Can it be concluded that the average life span of an Indian is more than 70 years, if a random sample of 100 indians has an average lifespan of 71.8 years with standard deviation of 8.9 years.

→ ~~n = (small)~~ $\quad \cancel{\bar{x} = 147} \quad \cancel{s = 16}$

~~Step I : Hypothesis~~ :- $\cancel{H_0 : \mu}$ $\quad \cancel{H_\alpha : \mu}$ ~~(one tailed)~~

$n = 100$ $\quad \bar{x} = 71.8$ years $\quad \sigma = 8.9$

Hypothesis : $H_0 : \mu = 70$ $\quad H_\alpha : \mu > 70$ (one tailed right)

$$z = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{71.8 - 70}{8.9/\sqrt{100}} = \frac{1.8}{0.89} = \frac{180}{89} = 2.022$$

los 5% or confidence : $z_\alpha = 1.64$

Decision : Since the computed value of $z = 2.02$ is greater than critical value $z_\alpha = 1.64$. $z > z_\alpha$

Null hypothesis is rejected

Alternate hypothesis on right side is accepted.

∴ It can be concluded that the average life span of an Indian is more than 70 years.

- Distribution of difference b/w means

Standard Deviation : $s = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

$\therefore \; z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$

- Procedure to test the hypothesis

1 Hypothesis : Null $\mu_1 = \mu_2$

2 The Test Statistic : Calculate $\bar{x}_1 - \bar{x}_2$

Calculate error $SE = s = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

Calculate the statistics $|z| = \left| \frac{\bar{x}_1 - \bar{x}_2}{SE} \right|$

3,4 LoS and critical value z_α

5 Decision : if $|z| \leq z_\alpha$; we accept the null hypothesis at given LoS otherwise we reject it.

- Remember

Under the null hypothesis $\mu_1 = \mu_2$ test statistics is given by :-

i) $z = \frac{\bar{x}_1 - \bar{x}_2}{s}$ (if σ_1 and σ_2 are known)

ii) $z = \frac{\bar{x}_1 - \bar{x}_2}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ (samples are from same population) $\sigma_1 = \sigma_2 = \sigma$

iii) $z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ where s_1 and s_2 are sample std. dev^n of the 2 samples and if σ_1 and σ_2 are not known and not equal

(i.e. both the populat^n have different unknown std dev^n)

iv) if σ_1 and σ_2 aren't known and $\sigma_1 = \sigma_2 = \sigma$ then (i.e. both the population have same unknown

Page No.
Date

std dev^n (coming from same population)

$\therefore SE = \sqrt{\frac{s_1^2}{n_2} + \frac{s_2^2}{n_1}}$ and $z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_2} + \frac{s_2^2}{n_1}}}$

- Interval ~~estimation~~ estimation for diff b/w 2 sample means

The confidence limits for $\mu_1 - \mu_2$ at various levels of confidence are $(\bar{x}_1 - \bar{x}_2) \pm 1.96$ SE at 95% level of confidence (5% LOS).

$(\bar{x}_1 - \bar{x}_2) \pm 2.58$ SE at 99% LoC and (1% los)

$(\bar{x}_1 - \bar{x}_2) \pm 3$ SE at 99.73% LoC and (0.27% los)

$(\bar{x}_1 - \bar{x}_2) - z_\alpha(SE) < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) - z_\alpha(SE)$

- Questions

1 2 samples drawn from 2 different populations gave the following results. Find 95% confidence limits for the diff b/w populat^n means

	Size	Mean	SD
Sample 1 :	400	124	14
Sample 2 :	250	120	12

Case 3

→ Given: $n_1 = 400$ $\bar{x}_1 = 124$ $s_1 = 14$

$n_2 = 250$ $\bar{x}_2 = 120$ $s_2 = 12$

Calculate: 95% confidence limit $z_\alpha = 1.96$

$\bar{x}_1 - \bar{x}_2 = 124 - 120 = 4$

$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{14^2}{400} + \frac{12^2}{250}} = \sqrt{\frac{196}{400} + \frac{144}{250}} = 1.03$

Interval estimation : $(\bar{x}_1 - \bar{x}_2) \pm 1.96$ SE $\rightarrow (1.981, 6.018)$

2 The mean of 2 samples of sizes 1000 and 2000 respectively are 67.50 and 68.0 inches. Can the samples be regarded as drawn from the $\boxed{\text{same population}}$ of std. deviation 2.5 inches at 0.0027 LOS? → case 2

Page No.
Date

→ Null hypothesis : H_0 : $\mu_1 = \mu_2$
Alternative hypothesis : H_a : $\mu_1 \neq \mu_2$

$n_1 = 1000$ $\bar{x}_1 = 67.5$ $\mu_1 = \mu_2$

$n_2 = 2000$ $\bar{x}_2 = 68$

$\sigma_1 = \sigma_2 = \sigma = 2.5$

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{SE}$$ $(\bar{x}_1 - \bar{x}_2) = -0.5$

$$SE = \sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 2.5\sqrt{\frac{3}{2000}} = 0.0968$$

$$z = \frac{-0.5}{0.0968} = -5.165$$

los = 0.0027 = 0.27%

$Z_\alpha = 3$

Decision : $|z| = 5.165 > 3$ Z_α

H_0 is rejected.

3 Test the significance of difference b/w the means of 2 normal population with $\boxed{\text{same SD}}$ from the following :-

case 3 ← unknown

	Size	Mean	SD
Sample 1 :	100	64	6
Sample 2 :	200	67	8

→ Null hypothesis : H_0 : $\mu_1 = \mu_2$
Alt hypothesis : H_a : $\mu_1 \neq \mu_2$
Calculation of statistics :-

$\bar{x}_1 - \bar{x}_2 = 64 - 67 = -3$

Since the SD of 2 populations are equal but unknown [case 4]

$$SE = S = \sqrt{\frac{s_1^2}{n_2} + \frac{s_2^2}{n_1}} = 0.9055$$

$$z = \frac{-3}{0.9055} = -3.313$$

LoS = 0.05 $Z_\alpha = 1.96$

$|z| > Z_\alpha$
Null hypothesis is rejected
Difference b/w the means of 2 normal population is ~~significant~~ significant.

Page No.
Date

4 The avg. of marks scored by 32 boys is 72 with SD 8 while that of 36 girls is 70 with std devⁿ 6. Test at 1% LOS whether the boys perform better than the girls. SD known & not equal

Null Hypothesis : $H_0 : \mu_1 = \mu_2$

Alt. Hypothesis : $H_\alpha : \mu_1 > \mu_2$

Calculate the statistic : $\bar{x}_1 - \bar{x}_2 = 72 - 70 = 2$

$$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 1.732$$

(We assume that the SD σ_1 and σ_2 of the 2 populations aren't =)

$$z = \frac{\bar{x}_1 - \bar{x}_2}{SE} = \frac{2}{\sqrt{3}} = 1.15$$ los $\alpha = 1\%$

Critical value : value of z_α at 1% los from the table is 2.33

Decision : since $z = 1.15 < z_\alpha = 2.33$

The null hypo is accepted. Boys perform better than girls.

5 2 samples drawn from 2 different populations gave the following results

	Size	Mean	SD
Sample 1 :	125	340	25
Sample 2 :	150	380	30

Test the hypothesis at 5% LoS that the difference of the means of the two population is 45.

Null Hypothesis : $H_0 : \mu_1 - \mu_2 = 45$

Alternate hypothesis : $H_\alpha : \mu_1 - \mu_2 \neq 45$

Calculation of statistic : $\bar{x}_1 - \bar{x}_2 = -40$ $\mu_1 - \mu_2 = 45$

$$SE = s = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{25^2}{125} + \frac{30^2}{150}} = 3.32$$

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{SE} = -25.60$$

los 5% , $z_\alpha = 1.96$

$|z| < z_\alpha$

Null hypothesis is accepted.

Difference b/w means of 2 normal population is insignificant.

- Curve is given by $y = C\left(1 + \frac{t^2}{v}\right)^{-(v+1)/2}$

Student's t distribution

where $t = \frac{\bar{x} - \mu}{S/\sqrt{n}}$ C: constant required to make area under curve unity

$v = n-1$, the no. of degrees of freedom.

- An interval estimate of the populat^n mean is given by $\bar{x} \pm t\hat{\sigma}_{\bar{x}}$

$$\bar{x} - t_\alpha\left(\frac{s}{\sqrt{n}}\right) < \mu < \bar{x} + t_\alpha\left(\frac{s}{\sqrt{n}}\right)$$

$$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{n-1} \quad n-1 \rightarrow \text{dof}$$

$\rightarrow$ unbiased estimates of sample SD is given

- For Small Samples

$$t = \frac{\bar{x} - \mu}{S/\sqrt{n-1}} \quad \text{where} \quad S^2 = \frac{\Sigma(x_i - \bar{x})^2}{n} \text{ : sample SD}$$

- Questions :

1. Given : $n = 10$: sample size, has mean 40 and SD 10. Construct the 99% confidence interval for the population mean.

$\rightarrow$ Given: $n = 10$ (small) $\bar{x} = 40$ $s = 10$

Step I : confidence interval 99% los = 1% = 0.01

dof $= n-1 = 9$

table value $= t_\alpha = 3.250$

$$\mu \in \left[\bar{x} \pm t_\alpha\left(\frac{s}{\sqrt{n-1}}\right)\right]$$

$$\in [40 \pm 2.25 \times 10/\sqrt{9})$$

$$\mu \in (29.166, 50.833)$$

2. A sample of size 12 from a normal population gave $\bar{x} = 15.8$ and $s_x^2 = 10.3$. Find 99% interval of population mean.

$\rightarrow$ $n = 12$ $\bar{x} = 15.8$ $S_x^2 = 10.3$ $S_x = 3.209$

Confidence level : 99% los = 1% = 0.01

dof $= n-1 = 11$

table value $= t_\alpha = 3.106$

Page No.
Date

$\mu \in \left[\bar{x} \pm t_\alpha \left(\frac{s}{\sqrt{n-1}}\right)\right] \in \left[15.8 \pm 3.106 \times \frac{3.209}{\sqrt{11}} \right]$

$\mu \in (12.655, 18.805)$

3 A random sample of 16 values from a normal population showed a mean of 41.6 inches and the sum of squares of deviations from this mean equal to 135. Obtain 95% and 99% fiducial limits for the mean.

→ $\mu = 16$ $\quad \bar{x} = 41.6$ $\quad \Sigma(x_i - \bar{x})^2 = 135$

~~Hypothesis : $H_0 : \mu = 1$~~

Confidence interval 99% los = 1% = 0.01

95% los = 5% = 0.05

dof = 16 - 1 = 15 = n - 1

$t_{\alpha_1} = 2.947$ $\quad t_{\alpha_5} = 2.131$

$\mu \in \left[\bar{x} \pm t_\alpha \left(\frac{s}{\sqrt{n-1}}\right)\right]$ → μ_1, μ_5

$s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n}} = \sqrt{\frac{135}{16}} = 2.904$

4 A random sample of size 16 from a normal population showed a mean = 103.75 cm and sum of squares of deviations from mean 843.75 cm^2. Can we say that the population has a mean of 108.75 cm?

→ Given: n = 16 (small)

$\bar{x} = 103.75$ $\quad \Sigma(x_i - \bar{x})^2 = 843.75$

Step I : Hypothesis : $H_0 : \mu = 108.75$

$H_\alpha : \mu \neq 108.75$ (Two Tailed)

Step II : Test Statistics : $s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n}} = \sqrt{\frac{843.75}{16}} = 7.261$

$t = \frac{\bar{x} - \mu}{s/\sqrt{n-1}} = \frac{103.75 - 108.75}{7.261/\sqrt{15}} = -2.667$

Step III : los = 0.05 $\quad$ dof = 16 - 1 = 15

Step IV : Critical value : $z_\alpha = 2.131$

Page No.

Date

Step V : Decision : $|t| \nless t_\alpha$ H_0 is rejected

5 Large no. of retail shops. Before a heavy advertisement campaign, the mean sales per week per shop was 140 dozens. After the campaign, a sample of 26 shops was taken and the mean sale = 147 dozens with SD of 16. Can you consider the advertisement effective?

→ $n = 26$ (small) $\bar{x} = 147$ $S = 16$

Step I : Hypothesis

Null hypothesis: $H_0 : \mu = 140$

Alternate hypothesis: $H_\alpha : \mu > 140$ (one tailed)

Step II : Calculation of test statistics :-

Sample is small, SD is not known,

$$t = \frac{\bar{x} - \mu}{S/\sqrt{n-1}} = \frac{147 - 140}{16/\sqrt{26-1}} = 2.19$$

Step III : los = 0.05 = α $t_\alpha = 1.708$ dof = 26-1 = 25

Step IV : Decision :-

Since $|t| = 2.19$ is greater than critical value $t_\alpha = 1.708$

$|t| \geq t_\alpha$. Null hypothesis is rejected

Alt hypothesis is accepted (right side)

The advertisement have increased average sales.

6 Nine items of a sample had the values 45, 47, 50, 52, 48, 47, 49, 53, 51. Does the mean of 9 items differ significantly from the assumed population mean 47.5?

→ $\bar{x} = \frac{\sum x_i}{n}$ $S^2 = \frac{\sum (x_i - \bar{x})^2}{n}$

$\bar{x} = 49.111$

x_i	45	47	50	52	48	47	49	53	51
$x_i - \bar{x}$	-4.111	-2.111	0.889	2.889	-1.111	-2.111	-0.111	3.889	1.889
$(x_i - \bar{x})^2$	16.9	4.456	0.790	8.346	1.232	4.456	0.012	15.124	3.562

$\sum (x_i - \bar{x})^2 = 54.878$

$S^2 = \frac{54.878}{9} = 6.097$ $S = 2.4692$

Page No.
Date

The null hypothesis $H_0 : \mu = 47.5$
Alt hypothesis $H_\alpha : \mu \neq 47.5$
Calculation of test statistic : $|t| = \left| \frac{\bar{x} - \mu}{s/\sqrt{n-1}} \right| = 1.84$

los = 0.05 dof = 9 − 1 = 8 $t_\alpha = 2.306$
Decision: since $|t| < t_\alpha$, null hypothesis is accepted. The mean of nine items doesn't differ significantly from assumed population mean 47.5

- Testing the Difference b/w means

If sample size $(n_1 + n_2 - 2)$ is small, the test statistic is computed as, $t = \frac{\bar{x}_1 - \bar{x}_2}{SE}$.

The statistic t so computed follows Student's t - distribution
The standard error of difference b/w means is

$$SE = s_p \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

If we are given unbiased Estimator for SD of the two samples then, $S_p = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$

where $S_1 = \sqrt{\frac{\Sigma(x_{i1} - \bar{x}_1)^2}{n_1 - 1}}$ and $S_2 = \sqrt{\frac{\Sigma(x_{i2} - \bar{x}_2)^2}{n_2 - 1}}$

When SD σ_1 and σ_2 are known, we can assume $\bar{X}_1 - \bar{X}_2$ to be normal with mean zero and SE = $\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ and hence use Z-distribution.

- Questions

1. If two independent random samples of sizes 15 and 8 have respectively the following means and population SD, $\bar{x}_1 = 980$, $\bar{x}_2 = 1012$, $\sigma_1 = 75$, $\sigma_2 = 80$. Test the hypothesis that $\mu_1 - \mu_2$ at 5% los

→ Given: $n_1 = 15$ $\quad n_2 = 8$ $\quad$ Null Hypothesis: $H_0 : \mu_1 = \mu_2$

$\bar{x}_1 = 980$ $\quad \bar{x}_2 = 1012$ $\quad$ Alt Hypothesis: $H_\alpha : \mu_1 \neq \mu_2$

$\sigma_1 = 75$ $\quad \sigma_2 = 80$ $\quad$ los: $\alpha = 0.05$

SD are known, we can assume $\bar{x}_1 - \bar{x}_2$ to be normal with mean zero and $SE = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ and hence use z-distribution

$SE = 34.28$ $\quad Z_\alpha = 1.96$

Decision: Since $|Z| = \left| \frac{\bar{x}_1 - \bar{x}_2}{SE} \right| = 0.93 < Z_\alpha = 1.96$

The hypothesis is accepted.

The population means are equal $\mu_1 = \mu_2$

2 The means of two random samples of size 9 and 7 are 196.42 and 198.82 respectively. The sum of squares of the deviation from the means are 26.94 and 18.73 respectively. Can the samples be considered to have been drawn from the same population.

→ Given: $n_1 = 9$ $\quad \bar{x}_1 = 196.42$ $\quad \Sigma(x_{i1} - \bar{x}_1)^2 = 26.94$

$n_2 = 7$ $\quad \bar{x}_2 = 198.82$ $\quad \Sigma(x_{i2} - \bar{x}_2)^2 = 18.73$

Hypothesis: $H_0 : \mu_1 = \mu_2$ $\quad H_\alpha : \mu_1 \neq \mu_2$

$$S_p = \sqrt{\frac{\Sigma(x_i - \bar{x})^2 + \Sigma(y_i - \bar{y})^2}{n_1 + n_2 - 2}} = \sqrt{\frac{26.94 + 18.73}{9 + 7 - 2}} = \sqrt{3.262}$$

$= 1.80$[illegible]9

$$SE = S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 1.80\text{[illegible]}9 \sqrt{\frac{1}{9} + \frac{1}{7}} = 1.80\text{[illegible]} \times \sqrt{0.253}$$

$= 0.9$[illegible]102

$$t = \frac{\bar{x}_1 - \bar{x}_2}{SE} = \frac{-2.4}{0.9\text{[illegible]}102} = -2.6\text{[illegible]}367$$

los $= 0.05$ $\quad Z_\alpha = 1.96$

$$Z = \frac{-32}{34.28} = -0.9334$$ idk where tf this is from

$|Z| < Z_\alpha$, hypothesis accepted

3 Six guinea pigs injected ~~[illegible]~~ with 0.5mg of a medication took on an average 15.4 secs to fall asleep with an unbiased SD 2.2 secs, while 6 other guinea pigs injected with 1.5mg

of medication took over an average 11.2 secs to fall asleep with an unbiased SD ~~~~ 2.6 ms. Use 5% los to test the null hypothesis that the difference in dosage has no effect.

$n_1 = 6 \qquad \bar{x}_1 = 15.4 \qquad s_1 = 2.2$

$n_2 = 6 \qquad \bar{x}_2 = 11.2 \qquad s_2 = 2.6$

Null hypothesis : $H_0 : \mu_1 = \mu_2$

Alt hypothesis : $H_\alpha : \mu_1 \neq \mu_2$

$\alpha = 0.05$; los $\qquad$ dof $= n_1 + n_2 - 2 = 12 - 2 = 10$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \approx 3.022$$

$t_\alpha = 2.228$

$|t| > t_\alpha$, we reject the null hypothesis

There is significant difference in time it takes for guinea pigs to fall asleep b/w the 2 dosages of medication.

- Find the difference of corresponding values of the 2 sets of data then find the mean. Difference $\bar{x}$ and SD of differences s. We then define

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

$\mu = 0$ is null hypothesis

s = SD of sample

S = unbiased estimator of σ

Note : Taking the null hypothesis $\mu = 0$ for differences amounts to the null hypothesis of equality of means $\mu_1 = \mu_2$ of the 2 populations.

- Questions :

1. A certain injection administered to 12 patients resulted in the following changes of blood pressure : 5, 2, 8, -1, 3, 0, 6,

-2, 1, 5, 0, 4 can it be concluded that the injection will be in general accompanied by an increase will blood pressure?

x_i	5	2	8	-1	3	0	6	-2	1	5	0	4
$x_i - \bar{x}$												
$(x_i - \bar{x})^2$												

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

$$s = \sqrt{\frac{\Sigma (x_i - \bar{x})^2}{n}} = \text{[illegible]}$$

los : $\alpha = 0.05$

Critical value : The value of t_α for 5% los for

dof = $n-1 = 11$ for one tailed test

For 2 tailed test, $t_\alpha = 1.796$ for 10% los

$$t = \frac{\bar{x}}{s/\sqrt{n-1}} = \text{[illegible]}\ 2.89$$

The null hypothesis is rejected.

There is a rise in B.P.

2 10 school boys were given a test in statistics and their scores were recorded. They were given a month special coaching and a second test was given to them in the same subject at the end of the coaching period. Test if the marks given below give evidence to the fact that the students are benefited by coaching.

Marks in Test I	70	68	56	75	80	90	68	75	56	58
Marks in Test II	68	70	52	73	75	78	80	92	54	55
$x = b - a$	-2	2	-4	-2	-5	-12	12	17	-2	-3
$x_i - \bar{x}$	-2.1	1.9	-4.1	-2.1	-5.1	-12.1	11.9	16.9	-2.1	-3.1

$$\bar{x} = \frac{\Sigma x_i}{n} = 0.1 \qquad \Sigma (x_i - \bar{x})^2 = 678.999$$

$$s^2 = \frac{\Sigma (x_i - \bar{x})^2}{n} = 67.90$$

Null Hypothesis $H_0 : \mu = 0$

Alternate Hypothesis $H_\alpha : \mu > 0$

As the sample is small,

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

$|t| = 0.036$

Page No.
Date

los : $\alpha = 0.05$

Critical value : The value of t_α for 5% los for $v = 12-1 = 9$ degrees of freedom is 1.83 for one-tailed test

Decision : Since the calculated value of $|t| = 0.036$ is less than the critical value $t_\alpha = 1.83$, the hypothesis is accepted. The students are not benefitted by coaching.

3 In a certain experiment to compare 2 types of pig-foods A and B, the following results of increasing weights were obtained. Assuming that the 2 sample of pigs are independent, can we conclude that food B is better than food A. Examine the case if the same set of pigs were used in both the cases.

Pig No.	1	2	3	4	5	6	7	8
↑ in weight X kg by A	49	53	51	52	47	50	52	53
↑ in weight Y kg by B	52	55	52	53	50	54	54	53

→ Calculate $\bar{x}_1, \bar{x}_2, \Sigma(x_1-\bar{x}_1)^2, \Sigma(x_2-\bar{x}_2)^2$ by making the table.

$\bar{x}_1 = 50.875$ $\quad\quad$ $\bar{x}_2 = 52.875$

$\Sigma(x_1-\bar{x}_1)^2 = 30.875$ $\quad\quad$ $\Sigma(x_2-\bar{x}_2)^2 = 16.875$

Null hypothesis H_0 : $\mu_1 = \mu_2$

Alternate hypothesis H_α : $\mu_1 < \mu_2$

Test statistic : $S_p = \sqrt{\dfrac{\Sigma(x_1-\bar{x}_1)^2 + \Sigma(x_2-\bar{x}_2)^2}{n_1+n_2-2}} = \sqrt{3.41}$

$SE = S_p \times \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}} = 0.92$

$t = \dfrac{\bar{x}_1 - \bar{x}_2}{SE} = -2.17$

For first part, los : $\alpha = 0.05$

Critical value : for $\alpha = 0.05$ & $v = 8+8-2 = 14$, dof $= 1.76$

Decision : Since the computed value of $t = -2.17$ is less than the table value $t_\alpha = -1.76$

The hypothesis is rejected at 5% los (one tailed test)

Page No.
Date

If the same set of pigs were used in the two tests:
we first calculate the differences b/w the weights in the 2 tests and we calculate $\bar{x}$ and s^2.

$\bar{x} = -2$ $\quad \sum (x_i - \bar{x})^2 = 12$

$s^2 = \sum (x_i - \bar{x})^2 / n = 1.5$

$H_0 : \mu = 0$ $\quad H_\alpha : \mu < 0$

$t = \dfrac{\bar{x} - \mu}{s/\sqrt{n-1}} = -4.32$

los: $\alpha = 0.05$ $\quad v = 8-1 = 7$ $\quad$ dof $= -1.89$ (one tailed test)

Decision: Since the calculated value $t = -4.32 < t_\alpha = -1.89$
The hypothesis is rejected.
Food B is superior to Food A.

4. The heights of 6 randomly chosen sailors are in inches: 63, 65, 68, 69, 71 and 72. The heights of 10 randomly chosen soldiers are: 61, 62, 65, 66, 69, 69, 70, 71, 72, and 73. Discuss in the light that these data throw on the suggestion that the soldiers on an average are taller than sailors.

→ We first calculate the mean and standard deviation of heights of both sailors and soldiers. $\bar{x}_1, \bar{x}_2, \sum (x_1 - \bar{x}_1)^2, \sum (x_2 - \bar{x}_2)^2$

$\bar{x}_1 = \dfrac{\sum x_1}{N} = \dfrac{408}{6} = 68$ $\quad \bar{x}_2 = \dfrac{\sum x_2}{N} = \dfrac{678}{10} = 67.8$

$\sum (x_1 - \bar{x}_1)^2 = 60$ $\quad \sum (x_2 - \bar{x}_2)^2 = 153.60$

$s_p = \sqrt{\dfrac{\sum (x_1 - \bar{x}_1)^2 + \sum (x_2 - \bar{x}_2)^2}{n_1 + n_2 - 2}} = 3.9$

Null Hypothesis $H_0 : \mu_1 = \mu_2$
Alternative Hypothesis $H_\alpha : \mu_1 < \mu_2$
Test statistic:- $t = \dfrac{\bar{x}_1 - \bar{x}_2}{SE}$

$SE = s_p \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}} = 2.014$

$t = 0.099$

$\alpha = 0.05$ for $v = 6 + 10 - 2 = 14$ dof : $t_\alpha = -1.761$ (one tailed)

Decision : $t = 0.099$ is greater than the table value $t_\alpha = -1.761$

The null hypothesis is accepted.

The means are equal i.e. the suggestion that the soldiers on the average are taller than sailors cannot be accepted

- Chi-Square (x^2)

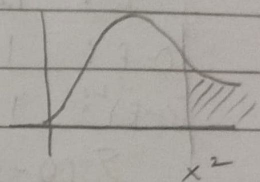
x^2 is calculated on the assumption of status quo (no change)

$$x^2 = \Sigma \left(\frac{(O-E)^2}{E} \right)$$

O : observed freq

E : Expected freq

Examples :-



i) coin or die is unbiased

ii) there is no association b/w the attributes

iii) the accident occurs evenly on all days.

iv) digits occur evenly on all pages

v) the events occur in the given ratio

vi) the events occur according to the given distribution (Binomial, Poisson, Normal)

$x^2_{calculated} < x^2_{table}$ ↳ accept H_0

This is called testing goodness of fit.

Conditions :-

i) N>50

ii) Cell frequency >10 : Combine to make it 10

iii) No. of classes : $4 \leq n \leq 16$

To test the goodness of fit :-

x^2 - test enables us to ascertain how well the theoretical distributions such as Binomial, Poisson, or Normal fit the observed frequencies.

In such cases we proceed on the null hypothesis that the theory supports the observations, i.e. the fit is good.

It can be used to test whether proportions p_1, p_2, p_3, p_4 in different populations are equal

Page No.
Date

i.e. χ^2 test can also be used to test the null hypothesis that $p_1 = p_2 = p_3 = p_4$

• Questions

1 The following table gives the no. of accidents in a city during a week. Find ~~whether~~ whether the accidents are uniformly distributed over a week.

Day :	Sun	Mon	Tue	Wed	Thur	Fri	Sat	Total
No. of accidents :	13	15	9	11	12	10	14	84

O :	13	15	9	11	12	10	14
E :	12	12	12	12	12	12	12
O-E :	1	3	-3	-1	0	-2	2
$(O-E)^2$:	1	9	9	1	0	4	4

$\Sigma (O-E)^2 = 28$

$\chi^2 = \Sigma (O-E)^2 / E = 28/12 = 2.33$

los : $\alpha = 0.05$ dof $= n-1 = 7-1 = 6$

Critical value : los table value of χ^2 is 12.59

Decision: $\chi^2_{cal} < \chi^2_{tab}$: Hypothesis accepted.

The accidents occur equally on all working days.

2 A die was thrown 132 times and the following frequencies were observed. Test the hypothesis that the die is unbiased.

No. obtained :	1	2	3	4	5	6	Total
Frequency :	15	20	25	15	29	28	132

$E = \frac{\Sigma O}{n} = \frac{132}{6} = 22$ O : Frequency

Calculate $(O-E)$ and its square.

$\chi^2 = \frac{\Sigma (O-E)^2}{E} = \frac{196}{22} = 8.909$

H_0 : die is unbiased los : $\alpha = 0.05$

H_1 : die is biased. dof : $n-1 = 5$

Critical value : table value of χ^2 is 11.07

Page No.
Date

Decision: Calculated value is less than table value of x^2
The hypothesis is accepted. The die is unbiased.

3 300 digits were chosen at random from a table of random numbers. The frequency of digits was as follows. Using x^2-test examine the hypothesis that the digits were distributed in equal numbers in the table.

Digit:	0	1	2	3	4	5	6	7	8	9	Total
Freq:	28	29	33	31	26	35	32	30	31	25	300

$E = \Sigma O / n = 300/10 = 30$ O: Freq

Calculate $(O-E)$ and $(O-E)^2$

$-2, -1, 3, 1, -4, 5, 2, 0, 1, -5$

$\Sigma (O-E)^2 = 86$

$x^2 = \Sigma (O-E)^2 / E = 86/30 = 2.86$

los: $\alpha = 0.05$ dof: $n-1 = 10-1 = 9$

Critical value: los table value of x^2 is 16.91

Decision: $x^2_{cal} < x^2_{tab}$: Hypothesis accepted.

The digits were equally distributed.

4 Theory predicts that the proportion of beans in the four groups A, B, C, D should be 9:3:3:1 in an experiment among 1600 beans the numbers in the 4 groups were 882, 313, 287 and 118. Do the experimental results support the theory?

Null hypothesis H_0: beans in A, B, C, D are in 9:3:3:1

Alt hypothesis H_α: Proportion is not as given in H_0

Calculation of test statistics: sum is $9+3+3+1 = 16$

The number of beans in 4 groups will be :-

$1600/16 = 100$

O:	882	313	287	118
E:	900	300	300	100
O-E:	-18	13	-13	18
$(O-E)^2$:	324	169	169	324

$x^2 = \Sigma \frac{(O-E)^2}{E}$

$= 4.72$

los: $\alpha = 0.05$

dof : $n-1 = 3$

Critical value : table value of $\chi^2 = 7.81$

Decision : $\chi^2_{cal} < \chi^2_{tab}$ $\therefore$ Null hypothesis accepted.

The proportion 9:3:3:1 is correct.

5 Investigate the association b/w the darkness of eye colour in father and son from the following data

		Color of father's eyes :-		
		Dark	Not dark	Total
color	Dark	48	90	138
of son's	Not dark	80	782	862
eyes	Total	128	872	1000

O : 48 80 90 782

E : 18 110 120 752

$E_{11} = \frac{R_1 \times C_1}{Total} = \frac{138 \times 128}{1000}$

Calculate $(O-E)$ and $(O-E)^2/E$

do this seperately for each O-E as the E is differend.

$$\chi^2 = \Sigma\left[\frac{(O-E)^2}{E}\right] = 66.88$$

los : $\alpha = 0.05$

dof : $n-1 = 2-1 = 1$

Critical value : $\chi^2_{tab} = 3.841$

Decision : $\chi^2_{cal} > \chi^2_{tab}$ $\therefore$ hypothesis rejected

There is no association b/w darkness of eye color of fathers and sons.