Statistics:-

- Discriptive
- Inferential
  - Hypothesis Testing
    - Test of Mean
      - When σ is known → Z-test (n large > 30)
        - ...e - Sample
        - Two-Sample
      - When σ is unknown → T-Test (n small < 30.)
        - one sample
        - Two Sample
    - Test variance
      - F-test (two ~~Sample~~ variances)
      - χ^2-test (Single varia...)
  - Estimation

mples.

One Sample Test (mean)

σ known

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

$$\cancel{Z = \frac{\bar{X} - \mu}{S / \sqrt{n}}}$$

σ unknown

$$t = \frac{\bar{X} - \mu}{S / \sqrt{n}}$$

$$\cancel{Z = \frac{\bar{X} - \mu}{S / \sqrt{n}}}$$

Two Sample Test (difference of mean)

σ known

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

where σ_1^2, σ_2^2 are same val b/w Two p

σ unknown.

when: $S_1^2 \neq S_2^2$

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

where $S_1^2 = S_2^2 = S_p^2$ = pooled variance.

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

test for two Samples:-

when $S_1^2 \neq S_2^2$

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

where $\nu = n_1 + n_2 - 2$

when $S_1^2 = S_2^2 = S_p^2 \Rightarrow$ pooled variance

OR Pooled Sample t-test

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

OR

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$S_p^2 = \frac{(n_1 - 1) S_1^2 + (n_2 - 1) S_2^2}{n_1 + n_2 - 2}$$

$\nu = n_1 + n_2 - 2$

F- test :-

$$F = \frac{S_1^2}{S_2^2} \quad \text{when } S_1^2 > S_2^2$$

χ^2 - test :-

$$\chi^2 = \frac{nS^2}{\sigma^2} = \frac{\sum (X_i - \bar{X})^2}{\sigma^2}$$

$\because S^2 = \frac{\sum (X_i - \bar{X}...}{n}$

$nS^2 = \sum (X_i ...$

$\nu = n - 1$

…ting of a single popⁿ mean "σ" is known:-

One Sample Z-test

Example:-

Past record show that average score of student in chemistry is 57 with standard deviation 10.

[redacted]

A new method of teaching is employed in a random sample of 70 students is selected. The Sample average is 60. Can we conclude on the basis of these results at 5% level of significance that the average score has increased.

Sol.

(1) Formulation of Hypothesis:-

∵ claimed always in H_1

$$\left.\begin{aligned} H_0 &: \mu < 57 \\ H_1 &: \mu > 57 \end{aligned}\right\} \text{ one Tailed}$$

(2) Level of Significance:-

$$\alpha = 5\% = 0.05$$

(3) Test statistic:-

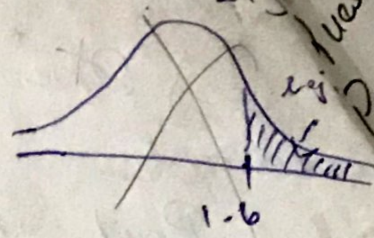
$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

②

5 (4) Critical Region:-

5 $Z_{cal} > Z_{tab}$

$2.51 > 1.6215$



Past Questions

one

$Z_{\alpha/2}$ = For 2 Tail

Z_{α} = For 1 Tail

$Z_{\alpha} = Z_{0.05} = 1.6\cancel{215}5$ = Z-Tabulated.

4 (5) Calculation:-

$$Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} = \frac{60-57}{10/\sqrt{70}} = 2.51 = Z_{cal}$$

(6) Conclusion:-

On the basis of provided evidence we conclude that

$2.51 > 1.6\cancel{2}5$

So, we reject null hypothesis at 5% level of significance & ~~we~~ say that the average score has increased.

$\because Z_{\alpha/2}$

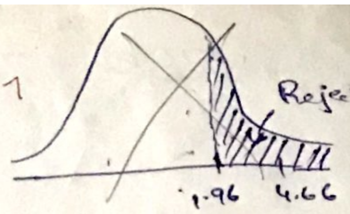
$Z_{0.05/2}$

$Z_{0.025}$

$0.5 - 0.025 = 0.475$

5. Critical Region:-

$4.66 > 1.96$



6. Conclusion:-

From the providing evidence as ~~out~~ $4.66 > 1.96$

So we "Reject H_0" & conclude that $\mu \neq 75$.

R.15 Question:- when σ is unknown.

$n = 9$; $\alpha = 0.05$; $\bar{X} = \frac{\Sigma X}{n} = 49.11$; $S = 2.491$, $\mu = 47.5$

1- Formulation of Hypothesis:-

$H_0 : \mu = 47.5$

$H_1 : \mu \neq 47.5$ (Two Tailed)

2- Level of Significance:-

$\alpha = 0.05 \Rightarrow \alpha/2 = 0.025 = t_{0.025(8)} = 2.306$

3- Test Statistic:-

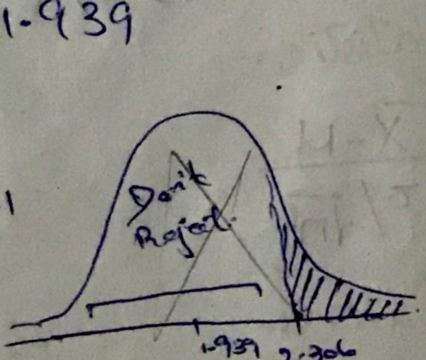
$$t = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

4:- Calculation:-

$$t = \frac{49.11 - 47.5}{2.491/\sqrt{9}} = \cancel{1.244}\ 1.939$$

5.- Critical Region:- $t(\alpha, \nu) \Rightarrow \nu = n-1$

$\cancel{1.244}$ 1.939 > 2.306



6. Conclusion:-

Don't Reject H_0

×

18

∴ Significant = Reject H_0

In Sig = Don't Reject H_0

Two-Sample Z-test.

Two astronomers recorded observation of certain star. The mean of 30 obs. obtained by 1st astronomist is 8.85 & mean of 40 obs. made by 2nd astronomist is 8.20. Experiment show that each astronomist obtained recording with variance of 1.25. Using $\alpha = 0.01$. Can we say that the diff b/w two results is significant.

1:- Formulation of Hypothesis:-

$H_0 : \mu_1 = \mu_2$

$H_0 : \mu_1 - \mu_2 = 0$

$H_1 : \mu_1 - \mu_2 \neq 0$

2 Level of Sig:-

$\alpha = 0.01$

3. Test-Statistic:-

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

when $\sigma_1^2 = \sigma_2^2$. Same variances b/w two popⁿ

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

4:- Calculation:-

$$Z = \frac{(8.85 - 8.20) - (0)}{1.095\sqrt{\frac{1}{30} + \frac{1}{40}}}$$

$$Z_{cal} = 2.472$$

5:- Critical Region:-

$$Z_{\alpha/2} = Z_{0.01/2} = Z_{0.005} = 2.575$$

$$Z_{cal} > Z_{tab}$$

$$2.472 \ngtr 2.575$$

6:- Conclusion:-

Difference b/w two results are insignificant.

Question Num 16.19 (b) :-

2 - Sample t-test

10

Question:-

Two Samples are randomly selected from 2 classes of students who have been taught by different methods. An examination is given & the results are shown as follow.

	Class I	Class II:
Sample size	$n_1 = 8$	$n_2 = 10$
Mean $\bar{X}$	$\bar{X}_1 = 95$	$\bar{X}_2 = 97$
Variance	$S_1^2 = 47$	$S_2^2 = 30$

On the assumption that the test course of the two classes of students have IDENTICAL variances. Determine wether the 2 diff methods of teaching are equally affective at $\alpha = 0.01$

$$S_1^2 = S_2^2 = S_p^2$$

1- Formulation of Hypothesis:-

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 - \mu_2 \neq 0$$

2- Level of Sig:-

$$\alpha = 0.01$$

3 Test- Statistic:-

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Calculation:-

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} = 37.437$$

$$t = \frac{(95-97)-(0)}{\sqrt{37.437\left(\frac{1}{8}+\frac{1}{10}\right)}} = -0.689$$

Critical Region:-

$t_{cal} < -t_{tab}$

$t_{cal} < -t_{(\alpha/2,\, \vartheta)}$

$-0.689 < -2.921$

$\vartheta = n_1+n_2-2$

$= 8+10-2$

$= 16$

Conclusion:-

Don't Reject H_0.

Question 18.19 (b).

4.- $Y = a + bX$

$\Sigma Y = na + b\Sigma X \Rightarrow$

$\Sigma Y = na + b(0) \Rightarrow \Sigma Y = na$

5. $\Sigma XY = a\Sigma X + b\Sigma X^2$

$\Sigma XY = a(0) + b\Sigma X^2$

6. $\Sigma XY = b\Sigma X^2$

10 + 10 - 2 = 18

0.025, 18

(18.20) 2.101

.0017 2.101

Don't

uestions:-

12

Given 2 random sample of size $n_1 = 12$ & $n_2 = 10$ for two independent normal popⁿ with $S_1 = 2.3$ & $S_2 = 1.5$. Test at 0.05 level of sig. The hyp $H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1$

$H_1: \frac{\sigma_1^2}{\sigma_2^2} > 1$

1. Formulation of Hyp:-

$$H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1$$

$$H_1: \frac{\sigma_1^2}{\sigma_2^2} > 1$$

2.- Level of Sig:-

$$\alpha = 0.05$$

3.- Test - Statistic:-

$$F = \frac{S_1^2}{S_2^2}$$

4.- Calculation:-

$$F = \frac{(2.3)^2}{(1.5)^2} = 2.351$$

5.- Critical Region:-

$$2.351 \not> 3.14$$

6.- Conclusion:- Don't Reject H_0.

Testing of Two - Pop^n variances:-

F-test:- (It is a ratio of 2 chi-square)

$H_0: \sigma_1^2 = \sigma_2^2$ $\qquad$ $H_0: \sigma_1^2 \leq \sigma_2^2$

$H_1: \sigma_1^2 \neq \sigma_2^2$ $\qquad$ $H_1: \sigma_1^2 > \sigma_2^2$

Test - Statistic:-

Since it the ratio of 2 chi-square

$$\frac{nS_1^2}{\sigma^2} \div \frac{nS_2^2}{\sigma^2} \Rightarrow \frac{\cancel{n}S_1^2}{\cancel{\sigma^2}} \times \frac{\cancel{\sigma^2}}{\cancel{n}S_2^2}$$

$$F = \frac{S_1^2}{S_2^2}$$

Critical Regions:-

$$F > F_{\alpha/2}(v_1, v_2)$$

$v_1 = n_1 - 1$; $v_2 = n_2 - 1$

Hence the ... limits on the population variance are (8.3, 80.5).

Example 17.4. Given that X_i are normally distributed and given the sample values $\bar{x} = 42$, $S = 5$ and $n = 20$. Test the hypothesis that $\sigma = 8$.

(i) Our null hypothesis is $H_0 : \sigma = 8$. Let the alternative hypothesis be $H_1 : \sigma \neq 8$.

(Two-tailed Test)

(ii) We choose the significance level at $\alpha = 0.05$.

(iii) The test-statistic is $\frac{nS^2}{\sigma_0^2} = \chi^2_{(n-1)}$, since X_i's are normally distributed, and this statistic under null hypothesis has a chi-square distribution with $(n-1)$ degrees of freedom.

The critical region is $\chi^2 > \chi^2_{0.025,(19)} = 32.85$ and $\chi^2 < \chi^2_{0.975,(19)} = 8.91$.

Now we compute the value of χ^2 from the given data as

$$\chi^2 = \frac{nS^2}{\sigma_0^2} = \frac{20\,(5)^2}{(8)^2} = \frac{20 \times 25}{64} = 7.81$$

(vi). **Conclusion.** Since the computed value of χ^2 falls in the critical region, we therefore reject the hypothesis, and conclude that there is no evidence to accept the hypothesis that $\sigma = 8$.

Example 17.5. The manager of a bottling plant is anxious to reduce the variability in net weight of fruit bottled. Over a long period, the standard deviation has been 15.2 gm. A new machine is introdcued and the net weights (in grams) in randomly selected bottles (all of the same nominal weight) are 987, 966, 955, 977, 981, 967, 975, 980, 953, 972. Would you report to the manager that the new machine has a better performance? (M.Sc., P.U., 1989, I.U., 1993, 96)

(i) We have to decide between the hypotheses

$H_0 : \sigma = 15.2$, *i.e.* the standard deviation is 15.2 gm

$H_1 : \sigma < 15.2$ *i.e.* the standard deviation has been reduced.

(ii) We choose the significance level at $\alpha = 0.05$.

(iii) The test-statistic is

$S^2 = \sqrt{\frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2}$

$$\chi^2 = \frac{nS^2}{\sigma_0^2} = \frac{\Sigma (X_i - \bar{X})^2}{\sigma_0^2}.$$

$S = \frac{1}{n}\left(\Sigma x^2 - \frac{(\Sigma x)^2}{n}\right.$

which under H_0, has a χ^2-distribution with $(n-1)$ degrees of freedom, assuming that the weights are normally distributed.

(iv) The critical region is $\chi^2 \leq \chi^2_{0.95,(9)} = 3.32$ (the lower 5% point)

(v) Computations. $n = 10$, $\Sigma X_i = 9713$, $\Sigma X_i^2 = 9435347$

Now $nS^2 = \Sigma(X_i - \bar{X})^2 = \Sigma X_i^2 - (\Sigma X_i)^2/n$

$$= 9435347 - (9713)^2/10 = 1110.1$$

$$\therefore \quad \chi^2 = \frac{1110.1}{(15.2)^2} = \frac{1110.1}{231.04} = 4.81$$

(vi) **Conclusion.** Since the calculated value of $\chi^2 = 4.81$ does not fall in the critical region, we therefore cannot reject the null hypothesis that the standard deviation is 15.2 gm and hence we would not report to the manager that the new machine has a better performance.

... use the left hand tail (two-sided test) also, then the critical region will be $F \geq F_{\alpha/2}(\upsilon_1, \upsilon_2)$ and $F \leq 1/F_{\alpha/2}(\upsilon_2, \upsilon_1)$.

(vi) Decide as below:

Reject H_0 if the computed value of F falls in the critical region, accept H_0 otherwise.

Example 19.2. Given two random samples of size $n_1=12$ and $n_2=10$ from two independent normal populations, with $s_1=2.3$ and $s_2=1.5$, test at 0.05 level of significance, the hypothesis $H_0:\sigma_1^2/\sigma_2^2=1$ against the alternative $H_1:\sigma_1^2/\sigma_2^2>1$.

(i) We state our null and alternative hypotheses as

$$H_0 : \sigma_1^2 / \sigma_2^2 = 1 \quad (\text{that is, } H_0 : \sigma_1^2 = \sigma_2^2), \text{ and}$$

$$H_1 : \sigma_1^2 / \sigma_2^2 > 1 \quad (\text{that is, } H_1 : \sigma_1^2 > \sigma_2^2).$$

(ii) The level of significance is set at $\alpha=0.05$.

(iii) The test-statistic to use is

$$F = \frac{s_1^2}{s_2^2}, \text{ where } s_1^2 \text{ is larger than } s_2^2,$$

which, if H_0 is true, has an F-distribution with $\upsilon_1=11$ and $\upsilon_2=9$ degrees of freedom.

(iv) Computations. Substituting the values, we get

$$F = \frac{(2.3)^2}{(1.5)^2} = \frac{5.29}{2.25} = 2.35.$$

(v) The critical region is $F > F_{0.05}(11, 9) = 3.10$.

(vi) **Conclusion.** Since the computed value of F does not fall in the critical region, we therefore do not reject H_0 at $\alpha=0.05$ and may conclude that there is sufficient evidence to indicate that the two variances are equal.

Example 19.3. Two random samples drawn from two normal populations are:

Sample I: 20, 16, 26, 27, 23, 22, 18, 24, 25 and 19.

Sample II: 27, 33, 42, 35, 32, 34, 38, 28, 41, 43, 30, and 37.