⊛ Hypothesis Testing for Difference between Two populations means:

Let $\bar{x}_1, \bar{x}_2$ be the means of two independent samples of sizes n_1 and n_2 (both n_1 and n_2 are large) from two different populations with standard deviations σ_1 and σ_2 respectively,

$$\therefore \bar{x}_1 \sim N\left(\mu_1, \frac{\sigma_1^2}{n_1}\right), \quad \bar{x}_2 \sim N\left(\mu_2, \frac{\sigma_2^2}{n_2}\right)$$

The difference $(\bar{x}_1 - \bar{x}_2)$ is also a normal variate.

The test statistic is given by

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

Under the null hypothesis H_0: There is no difference between the population means.
i.e $H_0 : \mu_1 = \mu_2$

<u>Note</u>: 1) If $\sigma_1 = \sigma_2 = \sigma$ then $z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$

2) If σ_1 & σ_2 are unknown & $\sigma_1 \neq \sigma_2$ the test statistic in this case is $z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$

3) If σ is not known σ_1 & σ_2, we use $\sigma^2 = \dfrac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2}$ to calculate σ.

Ex-1 A college conducted both day and night clesses intended to be identical. A sample of 100 days stududents yields examination results as under.

$\bar{x}_1 = 72.4$ and $s_1 = 14.8$

A sample of 200 night students yields examination results as under.

$\bar{x}_2 = 73.9$ and $s_2 = 17.9$

Are the two means statistically equal to 1% level?

Solⁿ

Here $n_1 = 100$, $n_2 = 200$

$\bar{x}_1 = 72.4$, $\bar{x}_2 = 73.9$

$\sigma_1 = s_1 = 14.8$, $\sigma_2 = s_2 = 17.9$

1) Null hypothesis H_0 : The two means are statistically equal

i.e $H_0 : \mu_1 = \mu_2$

Alternative hypo. $H_1 : \mu_1 \neq \mu_2$ (two tailed test)

2) calculation of statistics.

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

(since σ_1 & σ_2 are unknown)

$$= \frac{72.4 - 73.9}{\sqrt{\frac{(14.8)^2}{100} + \frac{(17.9)^2}{200}}}$$

$$= \frac{72.4 - 73.9}{1.95}$$

$$= -0.7704$$

Scanned with CamScanner

3) Critical region : Given that the significance level $\alpha = 10\% = 0.10$

$\therefore$ $z < -1.64$ & $z > 1.64$

4) Conclusion : As the calculated value of $z = -0.7704$ is bet^n -1.64 & 1.64

i.e $-1.64 < z = -0.7704 < 1.64$

$\therefore$ Null hypothesis H_0 would not be rejected.

i.e The two means are statistically equal.

Ex-2. In a random sample of 100 light bulbs manufactured by Company. A, the mean lifetime of light bulb is 1190 hours with S.D. of 90 hours. Also, in a random sample of 75 light bulbs manufactured by company B, the mean lifetime of light bulb is 1230 hours with S.D of 120 hours. Is there a difference between the mean lifetimes of the two brands of lightbulbs at a significance level of (a) 0.05 & (b) 0.01 ?

Sol^n: (a) Given $n_A = 100$, $n_B = 75$

$\bar{X}_A = 1190$, $\bar{X}_B = 1230$

$\sigma_A = S_A = 90$, $\sigma_B = S_B = 120$

1) Null hypothesis H_0 : Two means are statistically equal.

i.e $H_0 : \mu_A = \mu_B$

Alternative hypothesis $H_1 : \mu_A \neq \mu_B$. (two tailed test)

2) Calculation of statistic:

$$Z = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}} = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{S_A^2}{n_A} + \frac{S_B^2}{n_B}}}$$

(since σ_A & σ_B are unknown.)

$$= \frac{1190 - 1230}{\sqrt{\frac{(90)^2}{100} + \frac{(120)^2}{75}}}$$

$$= \frac{-40}{\sqrt{\frac{8100}{100} + \frac{14400}{75}}}$$

$$= \frac{-40}{16.5227} = -2.421$$

3) Critical region: given significance level $\alpha = 0.05$.

$\therefore \quad Z < -1.96 \quad \& \quad Z > 1.96$.

4) Decision: Reject the null hypothisis Ho since.

$Z = -2.421 < -1.96$.

Thus, we conclude that there is difference between the mean lifetimes of the light bulbs manufactured by company A and company B.

Similarly; solve (b) $\alpha = 0.01$.

Ex-3 A random sample of 1000 workers from South India show that their mean wages are Rs. 47 per week with a standard deviation of Rs 28. A random sample of 1500 workers from North India gives a mean wage of Rs. 49 per week with a standard deviation of Rs. 40. Is there any significant difference between their mean level of wages?

Solⁿ: Here $n_1 = 1000$, $n_2 = 1500$

$\bar{X}_1 = 47$, $\bar{X}_2 = 49$

$S_1 = 28$, $S_2 = 40$.

1) Null hypothesis H_0: There is no significant differences two mean level of wages.

$H_0 : \mu_1 = \mu_2$

Alternative hypothesis $H_1 : \mu_1 \neq \mu_2$ (two tailed test.)

2) Calculation of statistic:

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$ (since σ_1 & σ_2 are unknown)

$$= \frac{47 - 49}{\sqrt{\frac{(28)^2}{1000} + \frac{(40)^2}{1500}}}$$

$$= \frac{-2}{1.36}$$

$$= -1.47$$

Scanned with CamScanner

3) Critical region : take significance level
$\alpha = 5\% = 0.05$

$z < -1.96$ & $z > 1.96$.

4) Decision : As the calculated value of
$z = -1.47$ between -1.96 & 1.96.

i.e $-1.96 < z = -1.47 < 1.96$.

$\therefore$ Null hypothesis H_0 would not be rejected.
i.e there is not significant differences betn two mean level of wages.

Ex-4 The research investigator was interested in studying whether there is a significant difference in the salaries of MBA grades in two metropolitan cities. A random sample size 100 from Mumbai yields an average income of Rs. 20,150. Another random sample of 60 from Chennai results in an average income of Rs. 20,250 if the Variances of both the populations are given as $\sigma_1^2 = 40000$ Rs. and $\sigma_2^2 =$ Rs 32,400 respectively.

Soln :

From the given data, 1's related to MBA grades in Mumbai and 2's related to MBA grades in Chennai

$n_1 = 100$, $\bar{x}_1 = 20,150$, $\sigma_1^2 = 40,000$

$n_2 = 60$, $\bar{x}_2 = 20,250$, $\sigma_2^2 = 32,400$

1) Null hypothesis $H_0 : \mu_1 = \mu_2$
Alternative hypothesis $H_1 : \mu_1 \neq \mu_2$ (two tailed test).

Scanned with CamScanner

2) Calculation of statistics:-

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{20150 - 20250}{\sqrt{\frac{40,000}{100} + \frac{32400}{60}}} = -3.26$$

3) Critical region: take significance level $\alpha = 5\%$
$= 0.05$

$Z < -1.96$ & $Z > 1.96$.

4) Decision: As the calculated value of $Z = -3.26$
< -1.96

$\therefore$ H_0 is rejected.

$\therefore$ there is a significant differences between mean values.

Ex-5 IQ test on two groups of boys and girls gave the following results:

Mean of Girls = 78, S.D = 10, n = 30.
Mean of Boys = 78, S.D = 13, n = 70.

Is there any significance in the mean score of girls and boys at 5% level of significance?

Solⁿ: From the given data, 1's related to girls and 2's related to boys.

$n_1 = 30$, $\bar{x}_1 = 78$, $S_1 = 10$.
$n_2 = 70$, $\bar{x}_2 = 78$, $S_2 = 13$.

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

Alternative hypothesis $H_1 : \mu_1 \neq \mu_2$.

2) Calculation of statistics:

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

(since σ_1 & σ_2 are unknown)

$$= \frac{78 - 78}{\sqrt{\frac{100}{30} + \frac{169}{70}}}$$

$$= 0.$$

3) Critical region: Given level of significance $\alpha = 5\% = 0.05$

$\therefore$ $z < -1.96$ & $z > 1.96$

4) Conclusion: The calculated value is less than the table value of z at 0.05 level of significance.

i.e $-1.96 < z = 0 < 1.96$

$\therefore$ we need not reject the Null hypothesis. Hence, we conclude that there is no significant difference between the two groups girls and boys.

Scanned with CamScanner