26/8/17

MODULE - 4

TESTING OF HYPOTHESIS

Statistical hypothesis : To make decisions about population on the basis of sample information, we make certain assumptions (or statements) about the population. Such assumptions which may / may not be true are called statistical hypothesis.

→ A hypothesis in statistics is simply a ~~qualita~~ quantitative stmt about a population.

Example :- The average marks of a student in a class test is 8. i.e $H_s : \mu = 8$

→ In general, a statistical hypothesis, may be simple or composite.

Simple Statistical Hypothesis : If the statistical (definite) hypothesis specifies the population completely, then it is called as simple statistical Hypothesis.

Example :-

(i) $H_s : \mu = \mu_0$ (Average)

(ii) $H_s : \sigma = \sigma_0$ (S.D)

(iii) $H_s : \mu = \mu_0 , \sigma = \sigma_0$

Composite Statistical Hypothesis :- A statistical hypothesis is said to be C.S.H, if its a compliment of ~~sy~~ simple statistical hypothesis.

Example :- (i) $H_c : \mu \neq \mu_0$

(ii) $H_c : \sigma \neq \sigma_0$

(iii) $H_c : \mu \neq \mu_0, \sigma \neq \sigma_0$

NOTE :- If the statistical hypothesis, specifies the population completely, then it is simple statistical hypothesis otherwise, its called C.S.H.

* 28/8/17

NULL Hypothesis :- A definite hypothesis or statement about unknown population parameter is known as statistical Null hypothesis.

→ A statistical Null hypothesis is always simple. and denoted by H_0

Example :- "95% students of second year C.S.E B-14 are intelligent" is called NULL Hypothesis

Alternative Hypothesis :- Any hypothesis which is complimentary to the NULL Hypothesis is called an alternative hypothesis. and is usually denoted by symbol H_a (or) H_1.

Example :- If H_0 : $\mu = 165$ cm

then H_1 : $\mu \neq 165$ cm. (i.e H_1 : $\mu < 165$ & $\mu > 165$)

NOTE :- The testing of hypothesis is a technique or tool to test the statistical hypothesis or any statement about the population parameter.

Types of Errors :- In hypothesis testing, in order to take a decision about statistical Null Hypothesis either accepting / rejecting involves two types of errors known as Type 1 and Type 2 error.

4/10/17

TYPE : 1 ERROR :- If we reject the null hypothesis H_0, while it should have been accepted, we say that a type - I error has been made

→ The probability of committing a type - I error is denoted by 'α'. i.e

$$P(\text{Reject } H_0 \text{ when it is true}) = \alpha$$

TYPE : 2 ERROR :- If we accept the null hypothesis while it should have been rejected, we say that a type - 2 error has been made.

→ The probability of committing a type - II error is denoted by 'β' i.e

$$P(\text{Accept } H_0, \text{ when it is wrong}) = \beta$$

True

	H_0	H_1
Reject H_0	Type-I error	Correct decision
Accept H_0	Correct decision	Type-II error

Critical and Acceptial Region :-

→ Let S be a sample space (or) whole sample region from an experiment and it can be divided into two complement regions i.e w and w^c.

$W \cup W^c = S$

$W \cap W^c = \phi$

→ If the sample points fall in w region and null hypothesis (H_0) is rejected, is known as "CRITICAL | RJECTION REGION".

→ If the sample points fall in w^c region and null hypothesis (H_0) is acceptable, is known as "ACCEPTIAL REGION".

Size of the test or (Level of Significance) :- (LOS):-

→ The probability of Type-1 error is known as size of the test | LOS and is denoted by α.

i.e $\alpha = P(\text{Type - I error})$

$= P(\text{reject } H_0, \text{ where its true})$.

$= P(\{x \in \omega \mid H_0 \text{ is true}\})$

$= \int_{\omega} L_0 \, dx$ where L_0 = likely hood function (Pdf) under H_0

→ In general, the level of significance α gives us an idea to reject the solution of a hypothesis testing problem.

Power of the test :- The probability of taking correct decision is called power of the test and is denoted by $1-\beta$, where β is the probability of type - 2 error. i.e

$\beta = P(\text{Type- II error})$

$= P(\text{accept } H_0, \text{ when wrong})$

$= P(\{x \in \omega^c \mid H_0 \text{ is false}\})$

$= \int_{\omega^c} L_1 \, dx$ where L_1 = likely hood function under H_1.

5|9|17

PROCEDURE FOR HYPOTHESIS TESTING PROBLEM :

→ A hypothesis testing problem can be solved by using the following five steps :

STEP: 1 Set up the statistical ~~a~~ null and alternate hypothesis, from the given problem.

STEP: 2 List out the complete sample data which is available from the testing problem.

STEP: 3 Use an appropriate test statistic to test the statistical hypothesis.

$$Z = \frac{t - E(t)}{S.E(t)} \sim N(0,1) \qquad t : \text{statistic}$$

STEP: 4 Using sample data compute the test statistic and compare with its table (or) critical value of the test statistic. at the given level of significance (α)

STEP: 5 Give the conclusion of the problem.

If the $\boxed{\text{test statistic value} \leq \text{table value}}$ then "accept H_0". Otherwise "reject H_0".

1. A coin is tossed 5 times and it follows binomial probability law. The Null hypothesis is rejected if more than 3 heads appear under

$H_0 : P = 1/2$ against $H_1 : P = 3/4$

Find:

(i) P (Type I error)

(ii) P (Type II error)

(iii) Power of test.

Sol: Let $X \sim B(n, p)$

X: no of heads

$n = 5$

$x = 0, 1, 2, 3, 4, 5$

$\therefore P(X = x) = ({}^nC_x)(p)^x(q)^{5-x}$; $x = 0$ to 5.

Given $H_0 : P = 1/2$

vs

$H_1 : P = 3/4$

H_0 is rejected if more than 3 Head appear.

$\therefore$ Critical region $= \{x \in w \,/\, H_0$ is rejected$\}$

$= \{x \in w \,/\, x \geq 3\}$

Now

(i) P (Type - I error) $= \alpha$

$\alpha = P$ (reject H_0, when H_0 is true)

P (reject H_0, $H_0 : P = 1/2$)

$P(X = x$ for $x > 3$, $P = 1/2$, $q = 1/2)$

$= P(X > 3 \,/\, P = 1/2,\ q = 1/2)$

$= \sum_{x=4}^{5} P(X = x)$

$= P(X = 4) + P(X = 5)$

${}^5C_4 \left(\frac{1}{2}\right)^4 \left(\frac{1}{2}\right)^1 + {}^5C_5 \left(\frac{1}{2}\right)^5$

$\Rightarrow \frac{5!}{1! \times 4!} \left(\frac{1}{32}\right) + \frac{1}{32}$

$= \frac{5}{32} + \frac{1}{32}$

$= \frac{\cancel{6}\ 3}{\cancel{32}\ 16}$

$= \frac{3}{16}$

$$\boxed{\alpha = 0.188}$$

$16 \overline{)3}$ 0.1

0

30

16

P(Type - I error) = 0.188.

(ii) Acceptible region = $\{ x \in \omega^c / H_0$ is accepted $\}$

$= \{ x \in \omega^c / x \leq 3 \}$

P(Type II error) = $\beta = P$(acceptible H_0, where H_0 is false)

= P(accept H_0, H_1 is true)

$P(X \leq 3 \,/\, H_1 : P = 3/4$ is true)

$P(X \leq 3 \,/\, P = 3/4,\ q = 1/4)$

$= \sum_{x=3}^{5} P(X = x)$

$$= {}^5C_3 \left(\frac{3}{4}\right)^3 \left(\frac{1}{4}\right)^2 + {}^5C_2 \left(\frac{3}{4}\right)^2 \left(\frac{1}{4}\right)^3 + {}^5C_1 \left(\frac{3}{4}\right)^1 \left(\frac{1}{4}\right)^4$$

$$+ {}^5C_0 \left(\frac{1}{4}\right)^5$$

$$= \frac{5 \times 4 \times \cancel{3!}}{(2 \times 1)\ \cancel{3!}} \left(\frac{3^3}{4^5}\right) + \frac{5 \times 4 \times \cancel{3!}}{\cancel{3!} \times 2 \times 1} \left(\frac{3^2}{4^5}\right) + \frac{5 \times 4!}{4!} \left(\frac{3}{4^5}\right) +$$

$$\frac{1}{4^5}$$

$$= \frac{20}{2} \left(\frac{3^3}{4^5}\right) + \frac{20}{2} \left(\frac{3^2}{4^5}\right) + 5 \left(3/4^5\right) + \frac{1}{4^5}$$

$$= \frac{270 + 90 + 15 + 1}{4^5} = \underline{376}$$

$$= \frac{47}{128}$$

2. 8 coins are tossed at a time and follows Binomial population. If the statistical Null Hypothesis is rejected atmost 3 times a head appears under $H_0 : P = 1/3$ vs $H_1 : P = 3/4$. Hence find size of the test (ii) power of test

Sol: Let $X \sim B(n, P)$

$n = 8$

$P(X = x) = {}^nC_x\, p^x\, q^{n-x}$

Critical region $= \{x \in W / H_0 \text{ is rejected}\}$

$= \{x \in W / x \leq 3\}$

(i) Size of the test $= \alpha = P(\text{Type - I error})$

$= P(\text{Reject } H_0,\ H_0 \text{ is true})$

$= P(\text{Reject } H_0,\ P = 1/3\ ;\ q = 2/3)$

$P(X \leq 3,\ p = 1/3\ \ q = 2/3)$

$$= \sum_{x=0}^{3} P(X = n)$$

$$= P(X=0) + P(X=1) + P(X=2) + P(X=3)$$

$$= {}^8C_0 \left(\tfrac{1}{3}\right)^0 \left(\tfrac{2}{3}\right)^8 + {}^8C_1 \left(\tfrac{1}{3}\right)^1 \left(\tfrac{2}{3}\right)^7 +$$
$${}^8C_2 \left(\tfrac{1}{3}\right)^2 \left(\tfrac{2}{3}\right)^6 + {}^8C_3 \left(\tfrac{1}{3}\right)^3 \left(\tfrac{2}{3}\right)^5.$$

$\Rightarrow 0.7413$

(ii) Power ~~of thy~~ the test $\Rightarrow 1 - \beta$

$\beta = P(\text{Type - II error})$

Accentible region = $\{x \in w^c /\ H_0$ is accepted$)$

$= \{x \in w^c /\ x \geq 3\}$

$\beta = P(H_0$ is acceptable ; H_1 is true$)$

$= P(x \in w^c\ ;\ p = 3/4\ ;\ q = 1/4)$

$= P(x > 3,\ p = 3/4,\ q = 1/4)$

$P(X > 3) \Rightarrow 1 - P(X \leq 3)$

$$\Rightarrow 1 - \left[P(X=0) + P(X=1) + P(X=2) + P(X=3)\right]$$

$$\Rightarrow 1 - \left[{}^8C_0 \left(\tfrac{3}{4}\right)^0 \left(\tfrac{1}{4}\right)^8 + {}^8C_1 \left(\tfrac{3}{4}\right)^1 \left(\tfrac{1}{4}\right)^7 + {}^8C_2 \left(\tfrac{3}{4}\right)^2 \left(\tfrac{1}{4}\right)^6\right.$$
$$\left. + {}^8C_3 \left(3/4\right)^3 \left(\tfrac{1}{4}\right)^5\right] \cancel{{}^8C_4 (3/4)^2 \left(\tfrac{1}{4}\right)^6 +}$$

7/9/17 (Text book)

Central Limit Theorem :- Let us consider a population, and sample size n. If independent random samples drawn from the population with mean μ and variance σ^2, then as n increases each population tends to a normal distribution.

Large Sample Test :-

Sampling :- It is a technique of ~~to~~ drawing a sample, from the population to study (to test) the characteristic or hypothesis, about the population.

Sampling test :- A test statistic which is used to test the statistical hypothesis (Null / Alternate) is known as Sampling test.

→ Depending on the nature of sample size and test statistic, sampling test can be split into

(i) Large Sample test (ii) Small Sample Test

LARGE SAMPLING TEST :- A test statistic which is used to test the statistical hypothesis by selecting a large sample ($n > 30$) is known as Large Sampling test.

→ By C.L.T, the good example for the large sample test is ~~termed~~ Normal distribution and is termed as "Z test" i.e $\boxed{Z = \dfrac{t - E(t)}{S.E(t)}}$ ★ $\sim N(0,1)$

The following tests are good examples for the large sample test :-

1. TEST FOR SINGLE MEAN (μ)

2. TEST FOR DIFFERENCE OF MEAN ($\mu_1 - \mu_2$)

3. TEST FOR STANDARD DEVIATION (σ)

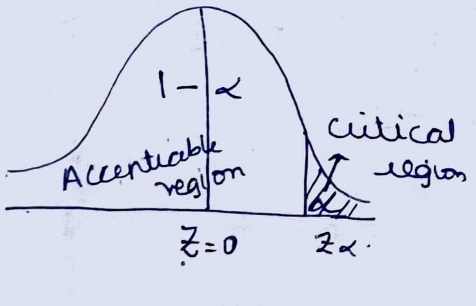
4. TEST FOR DIFFERENCE OF ST DEVIATION ($\sigma_1 - \sigma_2$)

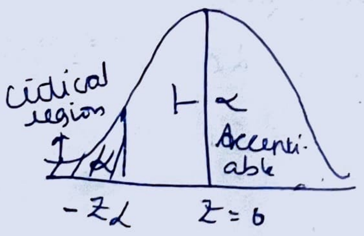
5. TEST FOR SINGLE PROPORTION (P)

6. TEST FOR DIFFERENCE OF PROPORTION ($P_1 - P_2$)

ONE TAIL & TWO TAILED TESTS :- To test the statistical hypothesis, to a testing of hypothesis problem, the critical region is more stretched towards the left and the right tailed is known as Two tailed test.

→ If the critical region, is more stretched towards the left or right is known as One Tailed test.





$H_0 : \mu = \mu_0$

$H_1 : \mu > \mu_0$

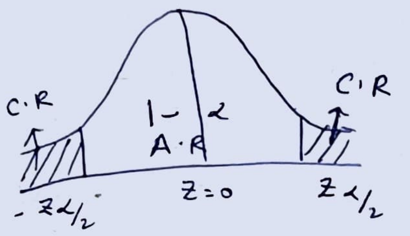
(Right)

$H_0 : \mu = \mu_0$

$H_1 : \mu < \mu_0$

(Left)

Two tailed :-



$H_0 : \mu = \mu_0$

$H_1 : \mu \neq \mu_0 \; (\mu > \mu_0 ; \mu < \mu_0)$

20/9/17

Example :- If we give assignment to the students then 90% i.e $(1-\alpha)$ students says that they will submit assignment before cut-off date but 10% (α) are not sure.

CRITICAL VALUES OF Z AT GIVEN LOS(α) :-

Two Tailed $P(Z \le Z_{\alpha/2}) = 1 - \frac{\alpha}{2} = 1 - 0.005 = 0.995$

$P(0 \le Z \le Z_{\alpha/2}) = 0.995 - 0.5 = 0.495$

from table $Z_{\alpha/2} = 2.58$.

Test	$\alpha = 1\%$	2%	5%	10%
Two tailed	$Z_{\frac{\alpha}{2}} = 2.58$	2.33	1.96	1.65
One tailed	$Z_\alpha = \pm 2.33$	± 1.96	± 1.65	± 1.28

NOTE :- If α is not given in the problem, then we will consider $\alpha = 5\%$ LOS.

TEST FOR SINGLE MEAN (μ) :-

Aim :- This large sample test can be used to test the ~~significant di~~ Hypothetical population mean μ (or) random sample is drawn from same population or not.

PROCEDURE :-

→ Set up H_0 : Population mean μ is known.

i.e $H_0 : \mu = \mu_0$

(vs)

$H_1 : \mu \neq \mu_0$

STEP:2 Sample data is given by $\bar{x} = \frac{\Sigma x}{n}$,

σ^2, $\alpha = 1\%$ (or) 2% (or) 5% (or) 10%.

STEP:3 Test statistic $z = \frac{\bar{x} - \mu}{(\sigma/\sqrt{n})}$

STEP:4 using the sample data, find calculated value of z and compare with table value of z at given level of significance α

STEP:5 Conclusion;

If $z_{cal} \leq z_{tab}$; then H_0 is acceptable

otherwise H_0 is rejected. (i.e H_1 : accept)

PROBLEMS :-

1. From a class of a college, a sample of 40 students, selected at random with sample mean marks 26 and standard deviation of marks 2.5. Test the hypothesis that, the average marks of the students in this college is 24 at 5% LOS.

Sol: H_0 : The average marks of students in clg is 24

i.e $H_0 : \mu = 24$

vs

$H_1 : \mu \neq 24$ (Two tailed test)

Sample Data :- $n = 40$ ($> 30 \Rightarrow$ Large sample test)

$\bar{x} = 26$

$\sigma = 2.5$

LOS $\alpha = 5\% = Z_{\alpha/2} = 1.96$

Test statistic $\Rightarrow Z = \dfrac{\bar{x} - \mu}{(\sigma/\sqrt{n})} \Rightarrow \dfrac{(26-24)\sqrt{40}}{2.5}$

$\Rightarrow \dfrac{2 \times \cancel{10}^{2}}{\cancel{25}_{5}} \times 2\sqrt{10} \Rightarrow \dfrac{\cancel{2 \times 10}^{4}}{\cancel{25}_{5}} (6.324)$

$= \dfrac{4}{5} \, 8\sqrt{10} \Rightarrow \dfrac{25.298}{5}$

$\Rightarrow 5.0596$

$Z_{calculated} = 5.0596$; $Z_{tab} = 1.96$

$Z_{cal} > Z_{tab} \Rightarrow$ reject H_0

Conclusion : The average marks of students in college is not 24.

NOTE :- Confidence limits for 'μ'

→ If population mean μ is unknown, then it is estimated by interval estimators i.e

$$\mu \in \bar{x} \pm Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

i.e $\mu \in \left(\bar{x} - Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} , \bar{x} + Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \right)$

For example, from above problem $\mu \in \left(26 - \frac{(1.96)\,2.5}{\sqrt{90}} , 26 + \frac{(1.96)\,2.5}{\sqrt{90}} \right)$

$\mu \in (25.22, 26.77)$

11/9/17

2. If $n = 140$; $\bar{x} = 26.5$; $\sigma = 3.9$; $\alpha = 1\%$ LOS, then test $H_0: \mu = 24.9$, $H_1: \mu > 24.9$

sol:- given $H_0: \mu = 24.9$

$H_1: \mu > 24.9$

sample data :- $n = 140$ (> 30) ⇒ large sample Test

$\bar{x} = 26.5$; $\sigma = 3.9$

LOS $\alpha = 1\%$ ⇒ $Z_\alpha = 2.33$

Test statistic $\Rightarrow z = \dfrac{(\bar{x} - \mu)\sqrt{n}}{\sigma}$

$$= \frac{[(26.5) - 24.9]\sqrt{140}}{3.9}$$

$$= \frac{(1.6)\ 11.832}{3.9}$$

$$= \frac{18.9312}{}$$

~~Zcal~~ $Z = 4.8541$

$Z_{cal} > Z_{tab} \Rightarrow H_0$ Rejected

3. If $n = 1000$; $\bar{x} = 410$, $\sigma = 40$; $\alpha = 5\%$,

$H_0: \mu = 420$; $H_1: \mu \neq 420$.

Sol:- Its a two tailed test.

$\therefore$ $H_0: \mu = 420$

$H_1: \mu \neq 420$ ($\mu > 420$ / $\mu < 420$)

Sample data :- $n = 1000$;

$\bar{x} = 410$; $\sigma = 40$

LOS $\alpha = 5\%$

$\Rightarrow Z_{\alpha/2} = 1.96$

$\boxed{Z_{cal} < Z_{tab}}$

Test static $\Rightarrow$ $z = \dfrac{(410 - 420)\sqrt{1000}}{40}$

$$= \frac{(-10)(31.622)}{40} = \frac{-316.22}{40} = -7.905$$

$\therefore$ Accepted H_0

Note :- $\mu \in \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

i.e $\mu \in \left(\bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} , \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \right)$

$\mu \in \left[410 - 1.96 \left(\frac{40}{31.622} \right), 410 + 1.96 \left(\frac{40}{31.622} \right) \right]$

$\mu \in [$

H.W

4. If $n = 2000$, $\bar{x} = 560$, $\sigma = 25$, then test

$H_0 : \mu = 550$, $H_1 : \mu < 550$

sol :- given $H_0 : \mu = 550$

$H_1 : \mu < 550$

Sample data :- $n = 2000$; $\bar{x} = 560$; $\sigma = 25$

LOS $\alpha = 5\% \Rightarrow z_{\alpha} = -1.65$

Total statistic $\Rightarrow$ $z = \frac{(\bar{x} - \mu)\sqrt{n}}{\sigma}$

$= \frac{(560 - 550)\sqrt{2000}}{25}$

$= \frac{(10)(44.72)}{25}$

$= \frac{447.2}{25}$

$z_{cal} = 17.888$

$\therefore z_{cal} > z_{Tab}$

$\Rightarrow$ Reject H_0

5. A light bulb company manufactures standardized light bulbs. From a lot, a sample of 2500 light bulbs are selected at random. On testing process, it was found that the average lifetime of light bulbs is 365 days. and the S.D of ~~of~~ ~~2~~ of whole light bulbs is 26 days. Test the hypothesis that the average life time of all light bulbs produced by the company is 345 days at 10% LOS (α).

Sol:- The average light bulbs.

$H_0 : \mu = 345$

(vs)

$H_1 : \mu \neq 345.$

Sample data :- $n = 2500$; $\bar{x} = 365$;

$\sigma = 26$; $\frac{\alpha}{2} = 10\% = Z_{\alpha/2} = 1.65$

$$Z = \frac{(\bar{x} - \mu)\sqrt{n}}{\sigma} \Rightarrow \frac{(365 - 345)\sqrt{2500}}{26}$$

$$= \frac{20 \times 50}{26}$$

$$= \frac{1000}{26}$$

$$Z_{cal} = 38.461$$

$\therefore$ $Z_{cal} > Z_{\alpha/2} \Rightarrow$ Reject H_0

Now

$\mu \in \left(\bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} , \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \right)$

$\mu \in \left(365 - 1.65 \left(\frac{26}{50} \right) , 365 + 1.65 \left(\frac{26}{50} \right) \right)$

$\mu \in (365 - 1.65(0.52), 365 + 1.65(0.52))$

$\mu \in [(365 - 0.858), (365 + 0.858)]$

$\mu \in [364.12, 365.858]$

6. A sample of 900 members is found to have mean 3.4 cm. Can it be reasonably regarded as a truely random sample, from a large population with mean 3.25 cm, S.D = 1.61 cm,

Sol:- $H_0 : \mu = 3.25$

H_0: ~~$\mu > 3.25$~~ sample can be regarded as random sample.
(us)

~~Sample~~:- $H_1 : \mu \neq 3.25$ (Two tailed)

Sample data :- $n = 900$; $\bar{x} = 3.4$;

$\sigma = 1.61$; $z_{\alpha/2} = 1.96$

$\alpha = 5\%$

$$\begin{array}{r} 3.40 \\ 3.25 \\ \hline 0.15 \end{array}$$

$z = \frac{(3.4 - 3.25)\sqrt{900}}{1.61} = \frac{(0.15)(30)}{1.61}$

$= \frac{4.5}{1.61} = 2.79$

$\therefore Z_{cal} > Z_{tab} \Rightarrow$ Reject H_0 i.e there is significant difference between sample mean and population mean. Hence sample cannot be regarded as a random sample.

2/9/17

TEST FOR DIFFERENCE OF MEANS :-

Aim :- These large sample test can be used to test the significant difference between 2 hypothetical population means of 2 distinct and independent populations.

→ It used ~~weather~~ whether the samples are drawn from the same population or not.

STEP: 1 $H_0 : \mu_1 - \mu_2 = 0$ (or) $\mu_1 = \mu_2$.

(vs) (Two Tailed)

$H_1 : \mu_1 - \mu_2 \neq 0$ (or) $\mu_1 \neq \mu_2$

Step : 2

	I	II
Sample Data		
Sample size	n_1	n_2
Variance	σ_1^2	σ_2^2
Sample means	$\bar{x}$	$\bar{y}$

Step : 3 Test statistic

$$\boxed{Z = \frac{t - E(t)}{S.E(t)} = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}}$$

where

$$S.E(t) = S.E(\bar{x} - \bar{y})$$

$$= \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$t = \bar{x} - \bar{y}$$

$$E(t) = E(\bar{x}) - E(\bar{y})$$

Step : 4 Using the sample data compute z test statistic and then compare with its table value of given α.

Step : 5 If $z_{cal} \leq z_{tab}$ then accept Ho otherwise reject Ho.

NOTE :-

1. If σ_1^2, σ_2^2 are unknown, then $\sigma_1^2 \approx s_1^2$, $\sigma_2^2 \approx s_2^2$.

$$\therefore \boxed{z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}}$$

2. If $\sigma_1 = \sigma_2$ ($= \sigma$ say) then

$$\boxed{z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} = \frac{\bar{x} - \bar{y}}{\sqrt{\sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{\bar{x} - \bar{y}}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}}$$

Confidence limits (or) Interval Estimates :

If the difference between means $\mu_1 - \mu_2$ is unknown then it is estimated by

$$\left[\mu_1 - \mu_2\right] \in \left[\bar{x} - \bar{y} - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}},\ \bar{x} - \bar{y} + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right]$$

1. From 2 colleges A and B, a sample of 40 and 50 students are selected at random with their average marks 12 and 10, standard deviation of marks 2 and 3 respectively. Test their hypothesis that, if their any significant difference between the average performance between the marks of students from these two colleges A and B at 1% LOS.

Sol :-

	A	B
sample	$n_1 = 40$	$n_2 = 50$
$\bar{x}$ =	12	$10 = \bar{y}$
σ_1 =	2	$\sigma_2 = 3$

$\alpha = 1\%$

Let H_0: There is no significant difference between average performance between the marks of both colleges.

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

$H_0 : \mu_1 - \mu_2 = 0$

(Als) (Two tailed)

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

Sample data | Clg A | Clg B

	Clg A	Clg B
	$n_1 = 40$	$n_2 = 50$
	$\bar{x} = 12$	$\bar{y} = 10$
	$\sigma_1 = 2$	$\sigma_2 = 3$

$\alpha = 1\%$ LOS

$$\boxed{Z_{\alpha/2} = 2.56}$$

Test statistic :-

$$Z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$= \frac{12 - 10}{\sqrt{\frac{4}{40} + \frac{9}{50}}}$$

$$= \frac{2}{\sqrt{\frac{200 + 360}{2000}}}$$

$$= \frac{2}{\sqrt{\frac{56\cancel{0}}{200\cancel{0}}}}$$

$$= \frac{2}{\sqrt{\frac{56}{200}}} = \frac{2}{\sqrt{0.28}} = \frac{2}{0.5291}$$

$\boxed{Z_{cal} = 3.7800}$; $\boxed{Z_{tab} = 2.56}$

$Z_{cal} > Z_{tab} \Rightarrow$ Reject H_0 //

Conclusion :- There's no evidence to accept H_0, i.e there's a significant difference between the average performance of the students from 2 clg's A and B at 1% LOS.

3/9/17

NOTE: Confidence limits:

$$\mu_1 - \mu_2 = \left[(\bar{x}-\bar{y}) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}},\ (\bar{x}-\bar{y}) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$$

$$= \left[(12-10) - (2.56)\sqrt{0.28}\ ;\ (12-10) + 2.56\sqrt{0.28}\right)$$

$$= \left[2 - 2.56(0.529),\ 2 + 2.56(0.529)\right)$$

$$= (2 - 1.354,\ 2 + 1.354)$$

$$= (0.646,\ 3.354).$$

2. Test $H_0 : \mu_1 = \mu_2$

vs

$H_1 : \mu_1 < \mu_2$ (One Tailed test)

$\bar{x} = 26.6$; $\bar{y} = 24.8$

$s_1^2 = 4$; $s_2^2 = 9$

$n_1 = 100$; $n_2 = 200$

$\alpha = 1\%$ LOS.

Sol:

$H_0 : \mu_1 = \mu_2$

vs

$H_1 : \mu_1 < \mu_2$

$\bar{x} = 26.6$; $\bar{y} = 24.8$

$s_1^2 = 4$; $s_2^2 = 9$

$n_1 = 100$; $n_2 = 200$

$\alpha = 1\%$ LOS

Test statistic $z = \dfrac{\bar{x} - \bar{y}}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$

$$= \frac{26.6 - 24.8}{\sqrt{\frac{4}{100} + \frac{9}{200}}}$$

$$= \frac{1.8}{\sqrt{\frac{800+900}{20000}}}$$

$$= \frac{1.8}{\sqrt{\frac{17\cancel{00}}{200\cancel{00}}}}$$

$$= \frac{1.8}{\sqrt{0.085}}$$

$$= \frac{1.8}{0.291}$$

$Z_{cal} = 6.185$ $\qquad Z_{\alpha\,tab} = -2.33$

$\therefore$ $Z_{cal} > Z_{tab} \Rightarrow$ ~~Reject~~ Accept H_0

3. $\bar{x} = 105.5$; $\bar{y} = 110.6$; $\sigma_1 = 2.5$; $\sigma_2 = 2.5$

$n_1 = 90$; $n_2 = 65$; $\alpha = 5\%$.

Sol :-

$H_0 : \mu_1 - \mu_2 = 0$

$H_1 : \mu_1 \neq \mu_2$ (Two tailed test)

$\bar{x} = 105.5$; $\bar{y} = 110.6$

$\sigma_1 = 2.5 = \sigma_2$

Test statistic $z = \dfrac{\bar{x}-\bar{y}}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$

$$= \frac{105.5 - 110.6}{2.5\sqrt{\frac{1}{90}+\frac{1}{65}}}$$

$$= \frac{-5.1}{2.5\sqrt{\frac{155}{5850}}}$$

$$= \frac{-5.1}{2.5(\sqrt{0.026})} = \frac{-5.1}{(2.5)(0.161)}$$

$$= \frac{-5.1}{0.4025}$$

$Z_{cal} = -12.670$; $Z_{\alpha/2} = 1.96$

$Z_{cal} < Z_{\alpha/2}$

$\therefore$ Ho is accepted.

4. $H_0: \mu_1 - \mu_2 = 0$
(vs)

$H_1: \mu_1 - \mu_2 > 0$; $\bar{x} = 365$; $\bar{y} = 345$

$\sigma_1^2 = 16$; $\sigma_2^2 = 25$ (Right tailed)

$n_1 = 1000$; $n_2 = 2000$

$\alpha = 2\%$ LOS.

test statistic $(z) = \dfrac{\bar{x}-\bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$

$$= \frac{365 - 345}{\sqrt{\frac{16}{1000}+\frac{25}{2000}}} = \frac{20}{\sqrt{\frac{32+25}{2000}}}$$

$$= \frac{20}{\sqrt{\frac{57}{2000}}}$$

$$= \frac{20}{\sqrt{0.028}}$$

$$= \frac{20}{0.167}$$

$$Z_{cal} = 119.760$$

Z ~~cal~~ $tab = +1.96$

$\therefore$ $Z_{cal} > Z_{tab}$

Reject Ho

5. A sample of ~~1500~~ 1000 and 1500 individuals selected at random from the two populations male and female with their average heights 5'6" and 5'5". Under standard deviation of heights 0.5 and 1.5 inches respectively. Test the claim that the male average height is more than the ~~fe~~ average height of female at 5% LOS.

Sol:- $n_1 = 1000$; $n_2 = 1500$

$\sigma_1 = 0.5$; $\sigma_2 = 1.5$

$\bar{x} = 5.6$; $\bar{y} = 5.5$

$\alpha = 5\%$

$H_0: \mu_1 - \mu_2 = 0$

$H_1: \mu_1 > \mu_2$ (Right tailed problem)

Test statistic $z = \dfrac{\bar{x} - \bar{y}}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$

$$= \frac{5.6 - 5.5}{\sqrt{\dfrac{0.25}{1000} + \dfrac{2.25}{\cancel{2000}\ 1500}}}$$

$$= \frac{0.1}{\sqrt{\dfrac{375 + 2250}{1500000}}}$$

$$= \frac{0.1}{\sqrt{\dfrac{2625}{1500000}}}$$

$$= \frac{0.1}{\sqrt{0.00175}} = \frac{0.1}{0.0418}$$

$$z_{cal} = 2.392$$

$$z_{\alpha} = 1.65$$

$$z_{cal} > z_{\alpha}$$

$\therefore$ Ho is Rejected

Conclusion :- That implies that male average height is greater than female average height which indicates H_1 is true.

TEST FOR SINGLE PROPORTION (P)

Aim :- Its used to test the proportion of a population.

Procedure :- To test the hypothetical population, P, let us define Ho : population proportion is known i.e P

i.e $H_0 : P = P_0$ (say) $(\Rightarrow Q = 1 - P)$

v/s

$H_1 : P \neq P_0$ $(\Rightarrow$ Two tailed)

Sample data :- n = sample size (> 30)

x = having attribute (success)

$p = x/n$ = sample proportion

i.e proportion of success

$\Rightarrow q = 1 - p$

~~q = 1 - p~~

α = 1% (or) 2% (or) 5% (or) 10% LOS

test statistic :-

$$\boxed{Z_{cal} = \frac{p - P}{\sqrt{\frac{PQ}{n}}} \sim N(0,1)}$$

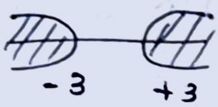
P : population proportion ; $Q = 1 - P$

→ Compute and Compare Z values

→ If $Z_{cal} \leq Z_{tab} \Rightarrow$ accept Ho, otherwise reject Ho

14/9/22

NOTE: without any reference to the level of significance we may reject the null hypothesis H_0 when $|Z| > 3$ (i.e $Z < -3$, $Z > 3$)



From a class of a clg a sample of 100 students selected at random of them 60 students are intelligent. Test hypothesis that 90% of students are intelligent in that clg at 95% confidence level (i.e $\alpha = 5\%$ LOS)

Sol: Let $H_0: P = 0.90 = P_0$ $(Q = 1 - P = 0.10)$

vs

$H_1: P \neq 0.90$ (-Two tailed)

Sample data :- $n = 100$ (> 30)

$x = 60 \Rightarrow$ having attributes

$$p = \frac{x}{n} = \frac{60}{100} = 0.6$$

Test statistic : $$Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}} = \frac{0.6 - 0.90}{\sqrt{\frac{0.90(0.10)}{100}}}$$

$$= -10$$

i.e $|Z| = 10 > 3$

$(-3, 3)$

Since $|Z| > 3$, then reject H_0

$\Rightarrow$ 90% of students are not intelligent in the college.

Confidence limits for proportion (P):

If P is not known then taking the p(sample proportion) as an estimate of P, the probable limits for the proportion in population are

$$P \in p \pm Z_{\alpha/2} \sqrt{\frac{pq}{n}} \; ; \quad p = x/n$$
$$q = 1 - x/n$$

i.e

$$\boxed{P \in \left(p - Z_{\alpha/2}\sqrt{\frac{pq}{n}},\; p + Z_{\alpha/2}\sqrt{\frac{pq}{n}}\right)}$$

NOTE:- If α is not known then the limits for proportion in the formulation are :-

$$\boxed{p \pm 3\sqrt{\frac{pq}{n}}}$$

Example:- A dice are thrown 9000 times of these 3220 yielded a 3 or 4. ST the dice can't be regarded as an unbiased one and find limits.

Sol:- $H_0 : P = 0.33$ $(Q = 0.67)$ (P = probability of getting 3/4

$1/6 + 1/6 = 2/6 = 1/3 = 0.33)$

(vs)

$H_1 : P \neq 0.33$ (Two-tail).

sample data :- $n = 9000$

$x = 3220$

$$p = \frac{3220}{9000} = 0.3578$$

$$q = 1 - p = 0.6422$$

Test statistic : $$Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}} = \frac{0.3578 - 0.6422}{\sqrt{\frac{(0.33)(0.67)}{9000}}}$$

$$Z = 4.94$$

$|Z| = 4.94 > 3$ $\therefore$ Reject Ho i.e $P \neq 0.33$

Confidence limits are :-

$$P \in \left(p - 3\sqrt{\frac{pq}{n}}, \; p + 3\sqrt{\frac{pq}{n}}\right)$$

i.e $$P \in \left(0.3578 - 3\sqrt{\frac{(0.3578)(0.6422)}{9000}}, \; 0.3578 + 3\sqrt{\frac{(0.357)(0.6422)}{9000}}\right)$$

i.e $P \in (0.345, 0.375)$

*Q. In a sample of 1000 people in Maharashtra 540 are rice eaters and the rest are wheat eaters. Can we assume that both rice and wheat eaters are equally popular in this state at 1% level of significance.

Sol :-

Ho : $P = P_0$ (NS)

H_1 : $P \neq P_0$ (Two tailed test)

Sample data :- $n = 1000$

$x = 540$

$p = \cancel{1000}\ \frac{54\cancel{0}}{100\cancel{0}} = 0.54$

$\alpha = 1\%$

$Z_{\alpha/2} = 2.58$

sol: Let

Ho: Both rice and wheat eaters are equally popular in the state.

i.e Ho: $P = 1/2$ ($\cancel{q = 1/2}$) (probability of equality

$H_1 : P \neq \frac{1}{2}$ (Two tailed)

$Q = 1 - P = 0.5$

Sample data :- $n = 1000$

$x = 540$

(Sample proportion) $p = x/n = \frac{540}{1000} = 0.54$

$q = 0.46$

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

$= \frac{0.54 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{1000}}}$

0.54
0.50

$= \frac{0.04}{\sqrt{\frac{25}{100000}}} = \frac{0.04 \times 10 \times 10}{\frac{5}{\sqrt{10}}}$

$Z_{cal} = 2.532$

For $\alpha = 1\%$; $Z_{\alpha/2} = 2.58$

$Z_{cal} = 2.532$; $Z_{tab} = 2.58$

$Z_{cal} < Z_{tab} \Rightarrow$ accept H_0

$\therefore$ we conclude that rice and wheat eaters are equally popular in Maharashtra state.

H.W

Q. 20 people were ~~obtained~~ attacked by a disease and only 18 survived. Will you reject the hypothesis that, the survival rate if attacked by the disease is 85% in favour of hypothesis that it is more. at 5% level

Sol: Let

$H_0 : P = 0.85$; $Q = 1 - 0.85 = 0.15$

$H_1 : P > 0.85$ (~~Two tailed~~) Right tailed

$n = 20$; $\bar{x} = 18$

$$p = \frac{\bar{x}}{n} = \frac{18}{20} = \frac{9}{10} = 0.9$$

Total statistic (z) $= \dfrac{p - P}{\sqrt{\dfrac{PQ}{n}}}$

$$= \frac{0.9 - 0.85}{\sqrt{\dfrac{(0.85)(0.15)}{20}}}$$

i.e $= \dfrac{0.05}{\sqrt{\dfrac{(0.85)(0.15)}{20}}}$

$\alpha = 5\%$

$= \dfrac{0.05}{\sqrt{\dfrac{0.1275}{20}}}$

$Z_{\alpha} \Rightarrow 1.65.$

$= \dfrac{0.05}{\sqrt{0.0063}}$

$= \dfrac{0.05}{0.0793}$

$Z_{cal} = 0.630$

$\therefore$ accept H_0

$Z_{cal} < Z_{tab}$

H.W

Q If in a random sample of 600 cars, taking a right turn at a certain ~~Profit~~ Traffic junction, 157 above into the wrong way. Test whether actually 30% of all drivers make this mistake or not at this given junction. Use

a) 0.05 (LOS) $\Rightarrow$ 5%

b) 0.01 (LOS) $\Rightarrow$ 1%

Sol:- Let

$P = 0.30 \Rightarrow Q = 1 - P \Rightarrow 0.7$

Q. In a big city, 325 men out of 600 men were found to be smokers. Does this information support the conclusion that the majority of men in this city are smokers.

Sol:- Let

H_0: Number of smokers & non-smokers are equal in the city

i.e $H_0: P = 1/2$

v/s

$H_1: P > 1/2$ (Right tailed test)

$Q = 1 - P = 0.5$

Sample data

$n = 600$; $x = 325$

$p = \frac{\cancel{325}^{65}\ 13}{\cancel{600}_{\cancel{120}\ 24}} = 0.5417$

Test statistic :- $Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}} = \frac{0.5417 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{600}}}$

$= \frac{0.0417}{\frac{25}{60000}}$

$Z = 2.04$

$Z_{cal} = 2.04$; $Z_\alpha = 1.65$

$Z_{cal} > Z_{\alpha\text{, tab}}$ ∴ Reject H_0

∴ Accept H_1 i.e smokers more than

Conclusion :- The majority men in the city are smokers

Q. A manufacturer claimed that atleast 95% of the equipment which is supplied to a factory confirmed to specifications. An examination of sample of 200 pieces of equipment revealed that 182 were faulty. Test this claim at 5% LOS.

Sol:- Let

$H_0 : P = 0.95$

V/S

$H_1 : P > 0.95$ (Right - tailed)

$Q = 1 - P$
$= 0.05$

sample :-

$n = 200$

$x = 182$

$p = \frac{x}{n} = \frac{182}{200}$

$p = 0.91$

test statistic $z = \frac{p - P}{\sqrt{\frac{PQ}{n}}} = \frac{0.91 - 0.95}{\sqrt{\frac{(0.95)(0.05)}{200}}}$

$= \frac{-0.04}{\sqrt{\frac{0.0475}{200}}}$

$= \frac{-0.04}{\sqrt{2.375 \times 10^{-4}}}$

$z_{cal} = \frac{-0.04}{0.0154} = -2.59$

$\alpha = 5\%$

$Z_{tab} = 1.65$

$Z_{cal} = -2.59$

$\therefore$

$Z_{cal} < Z_{tab} \Rightarrow$ accept H_0

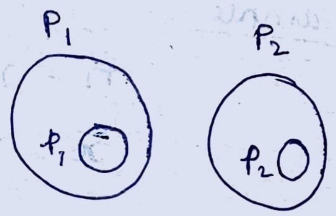
18/9/17

TEST FOR DIFFERENCE OF PROPORTIONS $(P_1 - P_2)$:

Aim :- It is used to test the population proportion.

Let H_0: $P_1 = P_2 = P_{(say)}$

i.e $P_1 - P_2 = 0$



V/S

H_1: $P_1 - P_2 \neq 0$ (Two tailed test)

Sample data

	I	II
sample size	$n_1\ (>30)$	$n_2\ (>30)$
sample having attribute	x_1	x_2
sample proportion	$p_1 = \frac{x_1}{n_1}$	$p_2 = \frac{x_2}{n_2}$

Test Statistic :-

$$Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim N(0,1)$$

→ If $Z_{cal} \leq Z_{tab}$, then accept H_0

otherwise reject H_0.

1. Random samples of 400 men and 600 women were asked whether, they would like to have a flyover, near their residence. 200 men and 325 women were in favour of the proposal. Test the hypothesis that, the proportions of men and women is in favour of the proposal are same against that, they are not at 5% LOS.

Sol: Let H_0: There's no significant difference b/w opinion of men & women as far as flyover is concerned.

~~$n_1 = 400$; $n_2 = 600$~~

~~$x_1 = 200$; $x_2 = 325$~~

i.e

$$H_0 : P_1 = P_2 = P \text{(say)}$$

v/s

$$H_1 : P_1 \neq P_2$$

Sample data :

$n_1 = 400$; $n_2 = 600$

$x_1 = 200$; $x_2 = 325$

$$p_1 = \frac{200}{400} ; \quad p_2 = \frac{325}{600}$$

$$p_1 = 0.5 ; \quad p_2 = 0.541$$

$$Z = \frac{0.5 - 0.541}{\sqrt{\left(\frac{1}{400} + \frac{1}{600}\right)}} = \frac{-0.041}{\sqrt{PQ\left(\frac{1000}{240000}\right)}}$$

$$= \frac{-0.041}{\sqrt{PQ(0.0\overset{0}{}41)}}$$

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

$$= \frac{(0.5)(400) + 0.541(600)}{1000}$$

$$= \frac{200 + 324.6}{1000}$$

$$= \frac{524.6}{1000}$$

$$P = 0.5245$$

$$Q = 0.4755 \quad (Q = 1 - P)$$

$$Z_{cal} = \frac{0.041}{\sqrt{(0.525)(0.475)(0.0\overset{0}{}41)}}$$

$$= \frac{0.041}{\sqrt{0.0102}} \quad \frac{0.041}{\sqrt{0.0009}} = \frac{0.041}{0.03}$$

$$= \frac{0.041}{0.100} = \cancel{0.41} \Rightarrow$$

$Z = -1.269$; $Z_{\alpha/2} = 1.96$

$\therefore$ we conclude that men & women do not differ significantly as regarding proposal of flyover is concerned.

2. A company has a head office at Kolkata and a branch office at Mumbai. The personal director wanted to know if worker at 2 places would like the introduction of new plan of work. A survey was conducted for this purpose, out of a sample of 500 workers at Kolkata, 62% ~~favour~~ favoured the new plan at Mumbai out of a sample of 400 workers 41% were against the new plan. ($p_2 = 59\%$). Is there any significant difference b/w two group in their attitude towards the new plan at 5% LOS.

Sol: ~~$p_1 = 0.62$; $p_2 = 0.59$~~

H_0: $P_1 = P_2$

v/s

H_1: $P_1 \neq P_2$

$n_1 = 500$; $n_2 = 400$

$p_1 = 62\%$; $p_2 = 59\%$

$p_1 = 0.62$; $p_2 = 0.59$

$$Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{0.62 - 0.59}{\sqrt{PQ\left(\frac{1}{500}+\frac{1}{400}\right)}}$$

$$\begin{array}{r} 0.62 \\ 0.59 \\ \hline 0.03 \end{array}$$

$$= \frac{0.03}{\sqrt{PQ\left(\frac{900}{200000}\right)}}$$

$$= \ominus$$

$$P = \frac{p_1 n_1 + p_2 n_2}{n_1 + n_2} \Rightarrow \frac{310 + 236}{900} \Rightarrow \frac{546}{900}$$

$$P \Rightarrow 0.606$$

$Q = 0.394$.

$$\therefore Z = \frac{0.03}{\sqrt{0.606 \times 0.394 \times 0.0045}}$$

$$= \frac{0.03}{\sqrt{0.00107}}$$

$\alpha = 5\% = 1.96$

$$= \frac{0.03}{0.032}$$

$$Z_{cal} = 0.9375$$

$\therefore$ $Z_{cal} < Z_{tab}$

$\Rightarrow$ Accept H_0

$\therefore$ We conclude that there is no significant difference between the two groups in their attitude towards the new plan.

20/9/17

CONFIDENCE LIMITS :- (Internal Estimators / Fiducial ~~Fiducial~~ ~~Judicial~~ Limits)

→ If the difference of population proportion is unknown, then it is estimated by

$$\boxed{|P_1 - P_2| \in p_1 - p_2 \pm Z_{\alpha/2} \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}}$$

3. On the basis of their total scores, 200 candidates of a civil service examination are divided into two groups, the upper 30% and the remaining 70%. Consider the 1st question of the examination. Among the first group 40 had the correct answer, in the second group 80 had the correct answer. On the

basis of these results, can one conclude that, the first question is not good at discriminating ability of type being examined here.

Sol:- given:- $n_1 = 200 = n_2$

Let H_0: The first question is not good at discriminating ability of type being examined.

H_0: ~~P~~ $P_1 = P_2 = P_0$ (say)

v/s

H_1: $P_1 \neq P_2$

sample data :- $n_1 = n_2 = 200$

$\bar{x} = 140$; $\bar{y} = 80$

$p_1 = 30\%$; $p_2 = 70\%$

$\alpha = 5\%$

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

$$= \frac{0.3 - 0.7}{\sqrt{PQ\left(\frac{2}{200}\right)}}$$

$$= \frac{-0.4}{\sqrt{PQ\left(\frac{1}{100}\right)}}$$

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

$$= \frac{(0.3)(200) + (0.7)(200)}{400}$$

$$P = \frac{60 + 140}{400} = \frac{200}{400} = 0.5$$

$\frac{3}{14} \times 200$

$\frac{7}{10} \times 200$

$Q = 0.5$

$$\Rightarrow \frac{-0.4}{\sqrt{0.25 \times \frac{1}{100}}}$$

(WRONG)

$$\Rightarrow \frac{-0.4}{\sqrt{\frac{25}{100 \times 100}}} \quad \frac{\sqrt{25}}{\sqrt{100}\sqrt{100}}$$

$$= \frac{-0.45 \times 100}{\frac{5}{100}} = \frac{45}{5} = 9$$

sol :- ~~given~~ sample data :-

$$n_1 = 30 \times \frac{200}{100} = 60$$

$$n_2 = 70 \times \frac{200}{100} = 140$$

$n_1 = 60$; $n_2 = 140$.

$\bar{x} = 40$; $\bar{y} = 80$

$p_1 = \frac{40}{60}$; $p_2 = \frac{80}{140}$ $\quad \frac{40}{70}$

$p_1 = 0.666$; $p_2 = 0.57$

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

$$P = \frac{p_1 n_1 + p_2 n_2}{n_1 + n_2} = \frac{(0.666)(60) + (0.57)(140)}{200}$$

$$P = \frac{39.96 + 79.94}{200} = 0.5995$$

$Q = 1 - 0.5995 \Rightarrow 0.4005$

$$Z \Rightarrow \frac{0.666 - 0.57}{\sqrt{(0.599)(0.400) \times \left(\frac{1}{60} + \frac{1}{140}\right)}}$$

$$\Rightarrow \frac{0.096}{\sqrt{0.2396 \times \frac{200}{8400}}}$$

$$= \frac{0.096}{\sqrt{0.2396 \times 0.023}}$$

$$= \frac{0.096}{\sqrt{5.5108 \times 10^{-3}}}$$

$$= \frac{0.096}{0.074}$$

$Z \Rightarrow 1.297$ $\quad \therefore$ $Z_{cal} < Z_{tab}$

Conclusion :- Accept H_0 i.e the first question is not good at discriminating ability of

---

Degrees of freedom :- Degrees of freedom (in the sense of Physics) means number of dimension an object can move, e.g. train can move in one dimension viz. 1. forward-backward, a mosquito can move in three dimensions viz. 1. forward-backward, 2. left-right and 3. up-down, in the above case of mosquito, there are three independent directions and no restriction so mosquito's degrees of freedom is 3.
In the case of train, there are 3 independent directions but train can not move in up-down and left-right dimension as it will go forward-backward on the track hence 2 restrictions are there so train's degrees of freedom is 3-2.