② Poisson distribution :-

The discrete random variable $X$ is said to follow the poisson distribution if it has the probability

$$P(X=x) = \frac{e^{-\lambda}\lambda^{x}}{x!} \quad ; \lambda > 0$$

NOTE :-

① here $n$ is infinetly large

② Mean :- $E(X) = \sum x_i \, P(X=x_i)$

$$= \sum x \cdot \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$\cancel{= \sum e^{-\lambda}}$$

$$= e^{-\lambda} \sum_{x=1} \frac{\lambda^{x}}{(x-1)!}$$

$$= e^{-\lambda} \sum \lambda \cdot \frac{\lambda^{x-1}}{(x-1)!}$$

$$= \lambda e^{-\lambda} \sum \frac{\lambda^{x-1}}{(x-1)!}$$

$$= \lambda e^{-\lambda} \left[1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \dots\right]$$

$$= \lambda e^{-\lambda} \cdot [e^{\lambda}]$$

$$\boxed{E(X) = \lambda}$$

③ Variance $= \lambda$

④ mean $=$ variance

⑤ MGF $= m_x(t) = e^{\lambda(e^t-1)}$

⑥ $\boxed{\lambda = np}$

Ex. ① If $X$ and $Y$ are independent poisson variates such that $P(X=1) = P(X=2)$, $P(Y=1) = P(Y=3)$. Find variance of $(X-2Y)$

$\Rightarrow$ $P(X=x) = \dfrac{e^{-\lambda}\lambda^{x}}{x!}$

① $P(X=1) = \dfrac{e^{-\lambda}\lambda}{1!}$

② $P(X=2) = \dfrac{e^{-\lambda}\lambda^2}{2!}$

$P(X=1) = P(X=2)$

$\therefore \dfrac{\cancel{e^{-\lambda}}\lambda}{1!} = \dfrac{\cancel{e^{-\lambda}}\lambda^2}{2!}$

$\boxed{\lambda = 2}$

③ $P(Y=3) = \dfrac{e^{-\lambda}\lambda^3}{3!}$

④ $P(Y=1) = \dfrac{e^{-\lambda}\lambda}{1!}$

$P(Y=3) = P(Y=1)$

$\dfrac{\cancel{e^{-\lambda}}\lambda^3}{6} = \dfrac{\cancel{e^{-\lambda}}\lambda}{1}$

$\boxed{\lambda'^2 = 6}$

$\boxed{\lambda' = \sqrt{6}}$

$\boxed{\lambda \neq -ve}$

$Var(aX \pm bY) = a^2 Var(X) + b^2 Var(Y) \pm 2ab\, Cov(X,Y)$

$Var(X-2Y) = Var(X) + 4\,Var(Y) - \cancel{2(2)\,Cov(X,Y)}$

$= 2 + 4(\sqrt{6})$

$Cov(X,Y) = E(XY) - E(X) \cdot E(Y)$

$= 0 - \cancel{\lambda^2}$

Ex② In a component manufacturing Industry there is a small chance of $\frac{1}{500}$ for any component to be defective.
The components are supplied in packets of 10.
Use poisson distribution to calculate the app. no. of packets containing
i) no defective
ii) One defective
iii) Two defective components respectively in a consignment of 10,000 packets

$p \Rightarrow$ defective

$\lambda = np$

$= \cancel{10 \times \left(1 - \frac{1}{500}\right)}$ $\quad 10 \times \frac{1}{500}$

$= \cancel{\frac{499}{50}}$ $\quad \frac{1}{50}$

$=$

$P(X=0) = \frac{e^{-\lambda}\lambda^0}{0!} = e^{-1/50} = \frac{1}{e^{1/50}}$

$P(X=1) = \frac{e^{-\lambda}\lambda^1}{1!} = \frac{1}{50} e^{-1/50}$

$P(X=2) = \frac{e^{-\lambda}\lambda^2}{2!} = \frac{1}{(50)^2} \cdot \frac{e^{-1/50}}{2}$

① packet containing zero defective

$= 10{,}000 \times e^{-1/50}$

$= 10000 \times 0.980$

$\approx 9800$ packet

② Packet containing 1 defective

$= \cancel{10000}^{200} \times \frac{1}{\cancel{50}} \times e^{-1/50}$

$= 200 \times 0.98$

$= 196$ packet

③ Packet containing 2 defective

$= \cancel{10000}^{\cancel{20}\ \cancel{4}\ 2} \times \frac{1}{\cancel{50} \times \cancel{50} \times \cancel{2}} \quad e^{-1/50}$

$= 2 \times 0.98$

$\approx 2$ packet

Ex. 3 An insurance company ensure 4000 people against loss of both eyes in a car accident. based on previous data. The rates where computed or assumption that on average of 10 persons in 1 lakh will have car accident each year that result in this type of injury. what is prob. That more than 3 of injured will collect on their policy in a given year.

$\Rightarrow$ $p = \dfrac{10}{100000} = \dfrac{1}{10000} = 0.0001$

$n = 4000$

$\lambda = 4000 \times 0.0001$

$= 0.4000$

$\lambda = 0.4$

$P(x \geq 3) = 1 - P(x \leq 3)$

$= 1 - [P(x=0) + P(x=1) + P(x=2) + P(x=3)]$

$$= 1 - \left[\frac{e^{-\lambda}\lambda^0}{0!} + \frac{e^{-\lambda}\lambda^1}{1!} + \frac{e^{-\lambda}\lambda^2}{2!} + \frac{e^{-\lambda}\lambda^3}{3!}\right]$$

$$= 1 - e^{-\lambda}\left[1 + \lambda + \frac{\lambda^2}{2} + \frac{\lambda^3}{3}\right]$$

$$= 1 - e^{-0.4}\left[1 + 0.4 + \cancel{0.16}\ 0.08 + \cancel{0.213}\ 0.0106\right]$$

$= 1 - e^{-0.4}[1.693]$

$= 1 - \cancel{1.13}\ 0.999$

$= 0.0008$

Ex. 4 Fit a Poisson distribution to the following data.

| No. of deaths (X) | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| Frequency (f) | 123 | 59 | 14 | 3 | 1 |

Sol$^n$:

~~Mean~~

Note that Poisson distribution has the probability

$$P(x=x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

• Mean $= \lambda = \dfrac{\Sigma f_i x_i}{\Sigma f_i} = \dfrac{59 + 28 + 9 + 4}{123 + 59 + 14 + 3 + 1}$

$= \dfrac{100}{200}$

$\boxed{\lambda = 0.5}$

$\boxed{N = \Sigma f_i = 200}$

$P(X=x) = \dfrac{e^{-0.5} \cdot (0.5)^x}{x!}$, $x = 0, 1, 2, 3, 4$

Note: Expected frequency $= N \times p$

(i) $x = 0$, $P(X=0) = \dfrac{e^{-0.5}(0.5)^0}{0!} = e^{-0.5}$

Expected freq. $= \underset{p}{e^{-0.5}} \times \underset{N}{200} = 121$

(ii) $x = 1$, $P(X=1) = \dfrac{e^{-0.5}(0.5)^1}{1!} = 0.5 \times 0.6065 = 0.3032$

Expected freq $= 0.3032 \times 200 = 61$

(iii) $X=2$, $P(X=2) = \frac{e^{-0.5}(0.5)^2}{2!} = 0.6065 \times 0.25 = 0.0758$

Exp. freq. $= 200 \times 0.0758 = 15$

(iv) $X=3$, $P(X=3) = \frac{e^{-0.5} \times (0.5)^3}{3!} = \frac{0.6065 \times 0.125}{6} = 0.0126$

Exp. freq $= 200 \times 0.0126 = 3$

(v) $X=4$, $P(X=4) = \frac{e^{-0.5} \times (0.5)^4}{4!} = \frac{0.6065 \times 0.0625}{24}$

$= 0.0015$

Exp. freq $= 200 \times 0.0015 = 0$

The required poison distribution is

| No. of death | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| expeted frequencies | 121 | 61 | 15 | 3 | 0 |

## * Continous Distribution

### (1) Normal Distribution

The random variable $X$ is said to follow the Normal Distribution if $X$ has the probability density function.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2}$$

Note :-

(1) mean of $X = \bar{X} = m$
(2) Variance $= Var(X) = \sigma^2$
(3) mean = median = mode
(4) MGF $= M_X(t) = e^{\left(mt + \frac{t^2\sigma^2}{2}\right)}$

* Standard Normal variable :-

A R.V. X with the parameter m and $\sigma$,

$\boxed{Z = \frac{X-m}{\sigma}}$ is said to be standard normal variable.

* Area Property :-

If $X = m \Rightarrow Z = 0$

If $X = x_1 \Rightarrow Z = \frac{x_1 - m}{\sigma} = z_1$ (say)

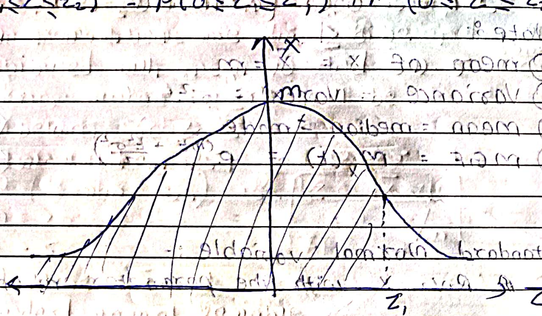
~~The area will between~~
The area under the normal variates X between $x=m$ and $x=x_1$ is equals to the area under the standard normal curve Z

between $Z=0$ to $Z=z_1$

i.e., $P(m \le X \le x_1) = P(0 \le Z \le z_1) = \frac{1}{\sqrt{2\pi}} \int_0^{z_1} e^{-\frac{1}{2}x^2} dx$

Note :-

$P(-z_1 \le Z \le z_2) = P(0 \le Z \le z_1) + P(0 \le Z \le z_2)$




Ex. ① If $X$ is normal variates with mean 10 and S.D. 4 find

① $P(5 \le X \le 18)$   ii) $P(X \le 12)$   iii) $P(|X-14| < 1)$

$\Rightarrow$ $m = 10$
$\sigma = 4$

Note that the standard normal variable is

$Z = \frac{X - m}{\sigma}$

① $x = 5$

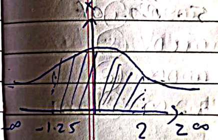
$z_1 = \frac{5-10}{4} = \frac{-5}{4} = -1.25$

$x = 18 \Rightarrow$

$z_2 = \frac{18-10}{4} = 2$

$\therefore P(5 \le X \le 18) = P(-1.25 \le Z \le 2)$

$= P(-\infty \le Z \le 2) - P(-\infty \le Z \le -1.25)$




$= 0.9772 - 0.1056$

$= 0.8716$

② $P(X \le 12) = P(-\infty \le X \le 12)$

$x = 12 \Rightarrow \frac{12-10}{4} = \frac{2}{4} = 0.5$

$\therefore P(-\infty \le Z \le 0.5) = 0.6915$

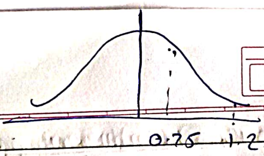
③ $P(|X-14| < 1) = P(-1 \le X-14 \le 1)$

~~At $x-14$, $z_1 = \frac{x-14-10}{}$~~

$= P(13 \le X \le 15)$

At $x=13$, $z_1 = \frac{13-10}{4} = \frac{3}{4} = 0.75$

At $x = 15$, $z_1 = \frac{15-10}{4} = \frac{5}{4} = 1.25$




$\therefore P(0.75 \le Z \le 1.25)$

$= P(-\infty \le Z \le \cancel{0.75}\ 1.25) - P(-\infty \le Z \le \cancel{1.25}\ 0.75)$

$= \cancel{0.7734} - 0.8944 - 0.7734$

$= 0.121$

Q) The marks obtained by 1,000 students in an examination are found to be normally distributed with mean 70 and S.D. 5. estimate No. of students whose marks will be
i) between 60 & 75
ii) more than 75

$\Rightarrow$ $m = 70$
$\sigma = 5$
$N = 1000$

Standard normal variable

$$z = \frac{X - m}{\sigma}$$

1) $X = 60$

$$z_1 = \frac{60 - 70}{5} = -2$$

$X = 75$

$$z_2 = \frac{75 - 70}{5} = 1$$

$P(-2 \le z_1 \le 1) = P(-\infty \le z \le 1) - P(-\infty \le z \le -2)$

$= 0.8413 - \cancel{0.4207}\ 0.0228$

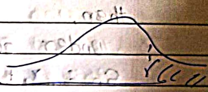
$= \cancel{0.4206}\ 0.8185$

$\therefore$ No. of student bet$^n$ 60 & 75 $= \cancel{0.4206}\ 0.8185 \times 1000$

$= \cancel{421 \text{ students}}$
819 student

2) $X = 75$

$$z = \frac{75 - 70}{5} = 1$$




$\therefore P(z > 1) = 1 - P(z < 1)$

$= 1 - P(-\infty < z < 1)$

$= 1 - 0.8413$

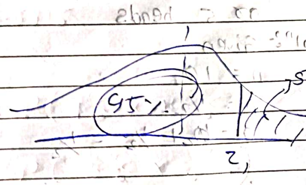
$= 0.1587$

No. of students $= 0.1587 \times 1000 = 159$ students

Q) Monthly salary X in an big organisation is normally distributed with the mean Rs 3000 and S.D. of Rs. 250, what should be the min. salary of worker in this organisation so that the probability that he belongs to top 5% workers.

$\Rightarrow$ mean $= 3000$
S.D. $= 250$

$$z = \frac{x - m}{z}$$




1) $\cancel{X = 0.5}$

$P(-\infty < z < z_1) = 0.95$

$\boxed{1.605 = z_1}$

$$1.65 = \frac{x - 3000}{250}$$

$\boxed{x = 3413 \text{ Rs}}$

* **Central limit theorem :-**

Statement :-

Let $X_1, X_2, \ldots, X_n$ be the $n$ independent identically distributed random variables with $E(X_i) = \mu_i$ and $V(X_i) = \sigma_i^2$

then,

Under certain condition random variable $S_n = X_1 + X_2 + \ldots + X_n$ is asymptotically normal with mean $\mu = \sum_{i=1}^{n} \mu_i$ and

variance $\sigma^2 = \sum_{i=1}^{n} \sigma_i^2$

Note :- The normal variate for the random variable $S_n$ is $\boxed{Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}}$

where $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$

Ex. (1) A coin is tossed 10 times using central limit theorem find probability of getting 3 or 4 or 5 heads

Sol$^n$:- Given

$n = 10$

$p = 1/2$

$q = 1 - 1/2 = 1/2$

$\therefore$ mean $= \mu = np = 10 \cdot \frac{1}{2} = 5$

variance $= \sigma^2 = npq = 10 \cdot \frac{1}{2} \cdot \frac{1}{2} = 2.5$

$\therefore \sigma = \sqrt{2.5} = 1.581$

To find :- $P(3 \le X \le 5)$

~~$P(3 \le X \le 5) =$~~

here $X$ is discrete random variable so we need to find $P(2.5 \le X \le 5.5)$

if $x = 2.5$ then $Z = \frac{x - \mu}{\sigma} = \frac{2.5 - 5}{1.58} = -1.58$

if $x = 5.5$ then $Z = \frac{5.5 - 5}{1.58} = \cancel{3.16}\ 0.316$

$P(2.5 \le X \le 5.50) = P(-1.58 \le Z \le 0.32)$

$= P(-\infty \le Z \le 0.32) - P(-\infty \le Z \le -1.58)$

$= 0.6255 - \cancel{0.9429}\ 0.0571$

$= 0.5684$

DATE

PAGE NO.

Ex. 2 A Random sample of size 100 is taken from a population whose mean is 60 and Variance is 400 Using C.L.T. find the probability such that the mean of sample will not differ from $\mu = 60$ by more than 4

$\Rightarrow$ Given :-

$n = 100$

$\mu = 60$

Variance $= 400 = \sigma^2$

$\sigma = 20$

To find :- probability such that $\bar{x}$ will not differ from $\mu = 60$ by more than 4

$\cancel{P(|\bar{x}-60| \leq X \leq 4)} \; P(|\bar{x}-60| \leq 4)$

$= P\left(\frac{|\bar{x}-60|}{\sigma/\sqrt{n}} \leq \frac{4}{\sigma/\sqrt{n}}\right)$

$= P\left(|z| \leq \frac{4\sqrt{n}}{\sigma}\right)$

$= P\left(|z| \leq \frac{4 \cdot 10}{20}\right)$

$= P(|z| \leq 2)$

$= P(-2 \leq z \leq 2)$

$= P(-\infty \leq z < 2) - P(-\infty < z < -2)$

$= 0.9772 - 0.0228$

$= 0.9544$

DATE

PAGE NO.

Ex. 3 The lifetime of certain ~~grant~~ brand of electric bulb may be consider as random variable with the mean 1200 hrs. and S.D. 250 hr. Find the probability using CLT that the average lifetime of 60 bulbs exceeds 1250 hrs

$\Rightarrow$

$n = 60$

$\mu = 1200$

$\sigma^2 = 250$

$\cancel{\sigma = \sqrt{250}} =$

$\cancel{P(}\; P(\bar{x} \geq 1250) = 1 - P(\bar{x} < 1250)$

$z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = \left[\frac{1250 - 1200}{(250)}\right] \times \sqrt{60}$

$z = \cancel{0.00195} \; 1.549) = 1.55$

$\therefore P(\bar{x} \geq 1250) = 1 - P(z < \cancel{0.00195}\; 1.549)$

$= 1 - P(-\infty < z < \cancel{0.00195}\; 1.549)$

$= 1 - (\cancel{0.5359} \; \cancel{0.5000} \; 0.9394)$

$= \cancel{0.0618} \; 0.0606$

Ex. 4 A distribution with unknown mean $\mu$ has variance 1.5 Using CLT find how large a sample should be taken from the distribution in order that the probability will be at least 0.95 such that this prob. sample mean will be within 0.5 of population mean.

$\Rightarrow$ $\sigma^2 = 1.5$

$\cancel{P(-a \le z \le a)} =$

$P(x-0.5 < \bar{x} < x+0.5) \ge 0.95$

$P(-\infty < \bar{x} < \mu+0.5) - P(-\infty < \bar{x} < \mu-0.5) \ge 0.95$

$z_1 = \dfrac{\mu+0.5-\mu}{\sigma/\sqrt{n}} = \dfrac{0.5\sqrt{n}}{1.2} \quad 0.4082\sqrt{n}$

$z_2 = \dfrac{\mu_1 - 0.5 - \mu}{\sigma/\sqrt{n}} = -0.4082\sqrt{n}$

$\therefore P(-\infty < z < 0.4082\sqrt{n}) - P(-\infty < z < -0.4082\sqrt{n})$

$\ge 0.95$

$\therefore P(-\infty < z < 0.4082\sqrt{n}) - P(-\infty < z < -0.4082\sqrt{n}) \ge P(-\infty < z < 1.65)$

$\therefore P(-0.4082\sqrt{n} < z < 0.4082\sqrt{n}) \ge P($

$P(z \le 0.4082\sqrt{n}) \ge 0.95$

i.e. $4082\sqrt{n} = 1.65 \quad 16.33 \quad \boxed{n = 17}$

Ex. 5 The burning time of certain type of lamp is an experimental r.v. with mean 30 hr what is probability that 144 of this lamps will be provide a total of more than 4500 hrs of burning time

$\Rightarrow$ $n = 144$

$\cancel{\sigma^2}$

$\mu = \cancel{30} \; \cancel{4320} \; 30$

$np = \cancel{30} \; 4320$

$\cancel{p = \dfrac{30}{144} = 0.208}$

$q = 0.7918$

$\therefore \sigma^2 = npq = 30 \cdot \cancel{0.79} \; 0.8 = 24$

$\sigma = 4.89$

$\cancel{\dfrac{4500}{30} = 150}$

$\dfrac{4500}{144} = 31.25$

$P(\bar{x} > 4500) =$

$\cancel{z = \left(\dfrac{4500-30}{4.89}\right)\sqrt{144} = 10969}$

$P(\bar{x} > 150) =$

$\cancel{z = \left(\dfrac{150-30}{4.89}\right)\sqrt{144}} =$

$P(\bar{x} > 31.25)$

$z = \left(\dfrac{31.25-30}{4.89}\right)\sqrt{144} = 3.06$

DATE

PAGE NO.

$\mu = 144 \times 30 = 4320$

$P(\bar{x} \geq 4500)$

$\therefore z = \dfrac{4500-4320}{\phantom{0}}$

DATE

PAGE NO.

## Module 4 :- Sampling Distribution

* **Population :-**
The group of individuals under study is called population or universe.

* **Sampling :-**
A part selected from the population is called sample. And the process of selection of sample is called sampling.

* **Notations :-**
1] $\mu$ = mean of population
2] $\bar{x}$ = mean of sample
3] $N$ = size of population
4] $n$ = size of sample
5] $\sigma$ = S.D. of population
6] $s$ = S.D. of sample

* **The Testing of Hypothesis :-**
On the basis of sample information we make certain decision about population. In taking such a decision we make Assumption. This assumption is called statastical Hypothesis.

1] NULL Hypothesis ($H_0$)
(i) In this hypothesis we shall persume that their is no significant difference betn observed value and expected value.

2] Alternate Hypothesis ($H_a$)
Its specify the range of value rather than one value.

* level of significance :-

It is usually employed in testing of hypothesis are 5% OR 1%

· It is always fixed in advance before collecting the sample information

* Critical region :-

The level marked by probabilities 0.05 or 0.01 which decide the significance of an event and corresponding region are called critical region

⇒ A region in sample space 'S' which amongst to rejection of $H_0$ is called critical region of rejection

* Error in Sampling :-

1] Type I Error :- Reject $H_0$ when it is true

2] Type II Error :- Accept $H_0$ when it is wrong

* One tailed and two tailed test

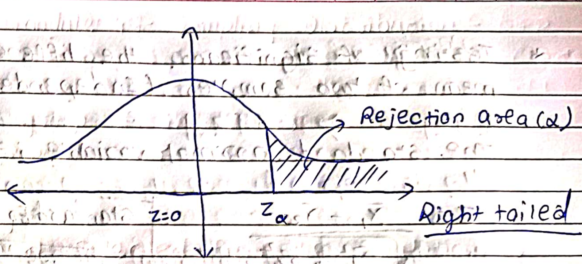
In Alternate Hypothesis $H_1$:

① $\mu > \mu_0$ ⇒ one tailed (Right tailed)

② $\mu < \mu_0$ ⇒ one tailed (Left tailed)

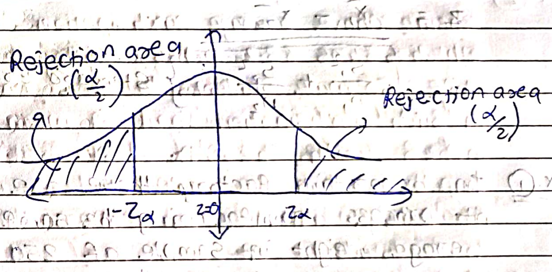
③ $\mu \neq \mu_0$ (i.e. $\mu > \mu_0$ or $\mu < \mu_0$) ⇒ two tailed

①

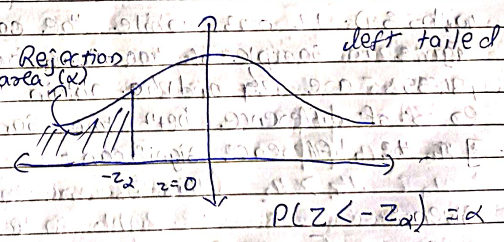



$P(Z > Z_\alpha) = \alpha$

②




③




$P(Z < -Z_\alpha) = \alpha$

* Testing of significance the difference betⁿ mean of two samples (Independent sample)

The standard normal variable is

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

Standard Error:-

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

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\underbrace{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}_{\text{Standard Error}}}$$

Ex. ① In a certain factory there are two independent process manufacturing the same item. The average weight in sample of 250 items produced in one process is found to be 120 ozs with S.D. 12 ozs while the corresponding figure in sample of 400 items from other process are 124 and 14. obtain the stand error of difference betⁿ two sample mean.
] Is the difference significant

⇒ Given:-

$n_1 = 250$, $\bar{x}_1 = 120$, $S_1 = 12$

$n_2 = 400$, $\bar{x}_2 = 124$, $S_2 = 14$

* $H_0 : \mu_1 = \mu_2$

* $H_a : \mu_1 \neq \mu_2$ (Two Tail)

↳ ($\mu_1 > \mu_2$ or $\mu_1 < \mu_2$)

* Calculation:-

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

] Standard Error:- $\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$

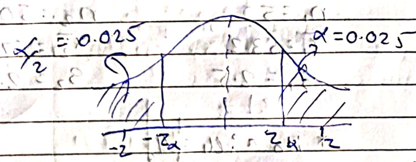
$$= \sqrt{\frac{(12)^2}{250} + \frac{(14)^2}{400}}$$

$$= \sqrt{0.576 + 0.490}$$

$$= \cancel{0.78} \quad 1.032$$

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{S.E.} = \frac{120 - 124}{1.034} = -3.86$$

$|Z| = 3.86$




* Level of significance $\alpha = 5\% = 0.05$ (considered)

Critical value the corresponding entry in z-table of $\alpha = 5\%$ (two-tailed)

$Z_\alpha = |-1.96| = 1.96$

Desesion :-

$|Z| > |Z_\alpha|$

$H_0$ is rejected $\therefore$ $H_a$ is Accepted

that is there is significant difference between samples mean

Ex. (2) The mean height of 50 male students to show above average participation in college athletics was 68.2 inches with the S.D 2.5 inches while 50 male students to show no interest in such participation had a mean height 67.5 inches with the S.D 2.8 inches. Test the hypothesis that the male student who participate in college athletics is taller than other male students at 5% los (level of significance)

Ans:- Given :-

$n_1 = 50$ $n_2 = 50$

$\bar{x}_1 = 68.2$ $\bar{x}_2 = 67.5$

$S_1 = 2.5$ $S_2 = 2.8$

* $H_0 : \mu_1 = \mu_2$
* $H_a : \mu_1 > \mu_2$ (This is one tailed test) i.e. Right Tailed

Calculation :-

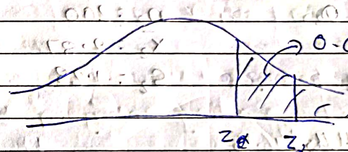
1] ~~Standar~~ $Z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}} = \dfrac{68.2 - 67.5}{\sqrt{\dfrac{(2.5)^2}{50} + \dfrac{(2.8)^2}{50}}} = 1.318$

$Z = 1.31$

* los $= \alpha = 5\% = 0.05$

Critical value corresponding entry in Z-table of $\alpha = 5\%$ (~~Two tailed~~) is

$Z_\alpha = |-1.64| = 1.64$




Decision :- $Z < Z_\alpha$

$H_0$ is Accepted

Hence male student who participate in college athletics is not taller than other male students.

~~Ex 3~~

Ex. 3 The avg Hourly ~~basis~~ wages of a sample of 150 workers in a plant A was Rs 2.56 with a S.D. Rs 1.08. The avg hourly wage of a sample of 200 workers in plant B is Rs 2.87 with S.D. 1.28 Rs.

i] Can an applicant safely assume that hourly wages paid by plant B are higher than those paid by plant A

Sol$^{n}$: Given :-

$n_1 = 150$, $n_2 = 200$

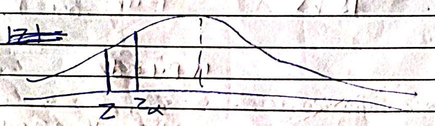
$\bar{x}_1 = 2.56$, $\bar{x}_2 = 2.87$

$s_1 = 1.08$, $s_2 = 1.28$

* $H_0 \Rightarrow \mu_1 = \mu_2$
* $H_a \Rightarrow \mu_1 < \mu_2$ (left-Tailed Test)

* Calculation :-

1] $Z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} = \dfrac{2.56 - 2.87}{\sqrt{\dfrac{(1.08)^2}{150} + \dfrac{(1.28)^2}{200}}} = -2.45$




loc = 5% = 0.05

corresponding $Z = Z_\alpha = -1.64$ (= Critical value

Decision :- $Z_\alpha > Z$ i.e. $Z < Z_\alpha$

$H_0$ is accepted

Hence applicant can not safely assume that hourly wages paid by plant B are higher than those paid by plant A.

* Student - t - distribution (t-Test)

This test is useful when size of sample is 30 or less (small samples)

i) Test the significance for difference between sample mean and population mean.

Formulae :-

* $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$ if s is not directly provided

* $s^2 = \dfrac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$

* $\bar{x} = \dfrac{1}{n} \sum_{i=1}^{n} x_i$

* $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n-1}}$ If s is directly provided in data

* degree of freedom ($\nu$)

No. of Independent parameters required to specify the location of particle in the space is called d.o.f

Note :- For single ~~mean~~ sample $\boxed{dof = n-1}$

Ex① Weekly sale mean of soap bars in departmental store was 146.3 bars per store. After an adv. campaign the mean weekly sales in 22 stores for a typical week increased to 153.7 ~~bo~~ and shown a S.D. of 17.2. Was the ad campaign successful?

Sol$^n$:- Given :-

$n = 22$

$\bar{x} = 153.7$

$S = 17.2$

$\mu = 146.3$

* $H_0$ : $\mu = 146.3$
* $H_a$ : $\mu > 146.3$ (Right Tailed test)

*Calculation :-

$$t = \frac{\bar{x} - \mu}{S/\sqrt{n-1}} = \frac{153.7 - 146.3}{17.2/\sqrt{21}} = 1.971$$

$t = 1.971$

* LOS :- $\alpha = 5\% = 0.05$
* DOF :- $n - 1 = 22 - 1 = 21$
* Critical value : The corresponding entry in t-table for $\alpha = 0.05$ & dof $= 21$ is $t_{\alpha,\nu} = 1.721$

* Decision :- $t > t_{\alpha,\nu}$

$H_0$ is rejected

$H_a$ is accepted

~~Adv~~ Ad Campaign was Successful

Ex.② A random sample of 10 boys of following IQ's : 70, 120, 110, 101, 88, 83, 95, 98, 107, 100

Do this data supports the assumption of population mean IQ of 100?

Sol$^n$:- Given :-

$n = 10$, $\mu = 100$,

$$\bar{x} = \frac{\sum x_i}{n} = \frac{972}{10} = 97.2$$

S

| IQ | 70 | 120 | 110 | 101 | 88 | 83 | 95 | 98 | 107 | 100 |
|---|---|---|---|---|---|---|---|---|---|---|
| $x_i - \bar{x}$ | -27.2 | 22.8 | 12.8 | 3.8 | -9.2 | -14.2 | -2.2 | 0.8 | 9.8 | 2.8 |
| $(x_i - \bar{x})^2$ | 739.84 | 519.84 | 163.84 | 14.44 | 84.64 | 201.64 | 4.84 | 0.64 | 96.04 | 7.84 |

$\sum (x_i - \bar{x})^2 = 1833.6$

$$S^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2 = \frac{1833.6}{10-1} = 203.73$$

$\therefore$ $\boxed{\cancel{S = 4.757}}$ $\boxed{S = 14.27}$

* $H_0$ : $\mu = 100$
* $H_a$ : $\mu \neq 100$ (Two tailed test)

~~LOS~~

* Calculation :-

$$t = \frac{\bar{x} - \mu}{S/\sqrt{n}} = \frac{97.2 - 100}{14.27/\sqrt{10}} = -0.62$$

$|t| = \cancel{0.082}\ 0.62$

* los :- ~~5%~~ $\alpha = 5\% = 0.05$
* DOF : $\nu = n-1 = 9$

* Critical value :
The corresponding entry in t table for $\alpha = 0.05$ & $\nu = 9$ is $|t_{\alpha,\nu}| = 2.262$

* Decision :-
$|t| < |t_{\alpha,\nu}|$

$H_0$ is accepted

Therefore data supports Assumption that $\mu = 100$.

* Test the significance for difference between two samples means (Dependent Samples)

Formulae :-

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

where $S^2 = \frac{1}{n_1+n_2-2}\left[\sum_{i=1}^{n_1}(x_i - \bar{x})^2 + \sum_{i=1}^{n_2}(y_i - \bar{y})^2\right]$

* $S^2 = \frac{1}{n_1+n_2-2}\left[n_1 s_1^2 + n_2 s_2^2\right]$ ... for Unbiased condition

* $S^2 = \frac{1}{n_1+n_2-2}\left[(n_1-1)s_1^2 + (n_2-1)s_2^2\right]$ ... for biased condⁿ

* Degree of freedom $(\nu) = n_1 + n_2 - 2$

Ex ① Sample of 2 types of electric bulbs were tested for length of life and following data were obtained

| Sample no. | Type I | Type II |
|---|---|---|
| Sample size | $n_1 = 8$ | $n_2 = 7$ |
| Sample mean | $\bar{x}_1 = 1234$ hrs | $\bar{x}_2 = 1036$ hrs |
| Sample S.D. | $s_1 = 36$ hr | $s_2 = 40$ hrs |

Is the difference in mean sufficient to warrant that Type I is superior than Type 2 regarding length of life?

Sol$^n$:

* $H_0 : \mu_1 = \mu_2$
* $H_a : \mu_1 > \mu_2$ (~~Left~~ Right Tailed Test)

* Calculation:

$$S^2 = \frac{1}{n_1+n_2-2}\left[n_1 s_1^2 + n_2 s_2^2\right]$$

$$= \frac{1}{8+7-2}\left[8(36)^2 + 7(40)^2\right]$$

$$= \frac{1}{13}\left[10368 + 11200\right]$$

$$S^2 = 1659$$

$$S = 40.73 \text{ hrs}$$

$$* \; t = \frac{\bar{x} - \bar{y}}{\sqrt{S^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{1234 - 1036}{\sqrt{1659\left[\frac{1}{8}+\frac{1}{7}\right]}} = \frac{196}{\sqrt{444.375}}$$

$$= \frac{196}{21.08}$$

$$= 9.29$$

* LOS :- $5\% = 0.05 = \alpha$
* DOF :- $\nu = n_1 + n_2 - 2 = 7 + 8 - 2 = 13$

* Critical value corresponding to $t_{\alpha,\nu} = t_{0.05,13}$

$$t_{\alpha,\nu} = 1.771$$

Descision :-

$t > t_{\alpha,\nu}$

$\therefore H_0$ is rejected

$\therefore H_a$ is accepted

Conclusion :

The difference in mean is sufficient to warrant that Type I is superior than Type 2 regarding length of life.

Ex. 2 : The height of 6 randomly choosen sellers are in Inches 63, 65, 68, 69, 71 and 72. Those of 10 randomly chosen soldiers are 61, 62, 65, 66, 69, 69, 70, 71, 72, 73. Discuss the ~~life~~ life that this data show on suggestion that sellers are on the average ~~not~~ taller than Soldiers. (los = 1%)

⇒ Given :-

$n_1 = 6$, $n_2 = 10$

* $H_0 : \mu_1 = \mu_2$
* $H_a : \mu_1 > \mu_2$ (~~Left~~ Right Tailed Test)

* Calculation :

$$\bar{x}_1 = \frac{63+65+68+69+71+72}{6} = 68$$

$$\bar{x}_2 = \frac{61+62+65+66+69+69+70+71+72+73}{10}$$

$$= 67.8$$

| Sailor height ($x_i$) | $(x_i - \bar{x})^2$ | Soldier height ($y_i$) | $(y_i - \bar{y})^2$ |
|---|---|---|---|
| 63 | 25 | 61 | 46.24 |
| 65 | 9 | 62 | 33.64 |
| 68 | 0 | 65 | 7.84 |
| 69 | 1 | 66 | 3.24 |
| 71 | 9 | 69 | 1.44 |
| 72 | 16 | 69 | 1.44 |
| | | 70 | 4.84 |
| | | 71 | 10.24 |
| | | 72 | 17.64 |
| | | 73 | 27.04 |
| Total | 60 | Total | 153.6 |

$$S^2 = \frac{1}{n_1+n_2-2}\left[\sum_{i=1}^{6}(x_i-\bar{x})^2 + \sum_{i=1}^{10}(y_i-\bar{y})^2\right]$$

$$= \frac{1}{6+10-2}\left[60 + 153.6\right]$$

$$= \frac{1}{14}$$

$$= 15.25$$

$$t = \frac{\bar{x}-\bar{y}}{\sqrt{S^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{68 - 67.8}{\sqrt{15.25\left(\frac{1}{6}+\frac{1}{10}\right)}} = 0.099$$

los :- 1% $\Rightarrow \alpha = 0.01$

DOF :- $\nu = n_1+n_2-2 = 14$

Critical value $t_{\alpha,\nu} = 2.624$

Decision :- $t < t_{\alpha,\nu}$

$H_0$ is Accepted

$\therefore$ The data are Inconsistent with the suggestion that sailors are on an average taller than Soldier.

* Test the significance for difference between the two samples mean (Independent samples)

formulae :-

$$t = \frac{(\bar{x}-\bar{y}) - (\mu_x - \mu_y)}{\sqrt{S^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

Ex. (1) To test the claim that the resistor of collecting wire can be reduced by at least 0.05 ohm by Alloying, 25 values obtained for each alloyed wire and standard wire produced following results

| | Mean wire | S.D. |
|---|---|---|
| Alloyed | 0.083 ohm | 0.003 ohm |
| Standard | 0.136 ohm | 0.002 ohm |

Test that 5% los whether or not the claim is substantiated.

Sol$^n$ ⇒ Given :-

$n_1 = 25$ $n_2 = 25$

$\bar{x}_1 = 0.083\ \Omega$ $\bar{x}_2 = 0.136$ ohm

$S_1 = 0.003\ \Omega$ $S_2 = 0.002\ \Omega$

$\mu_x - \mu_y \geq 0.05$

* $H_0 : \mu_x - \mu_y \geq 0.05$
* $H_a : \mu_x - \mu_y < 0.05$ (left - Tailed Test)

* Calculations :-

$$S^2 = \frac{1}{n_1+n_2-2}\left[n_1 S_1^2 + n_2 S_2^2\right]$$

$$= \frac{1}{25+25-2}\left[25(0.003)^2 + 25(0.002)^2\right]$$

$$= 6.77 \times 10^{-6}$$

* los :- $\alpha = 0.05$
* DOF :- $n_1+n_2-2 = 50-2 = 48$

* Critical value corresponding $t_{\alpha,\nu} = -1.684$

* $$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_x - \mu_y)}{\sqrt{S^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

$$= \frac{(0.083 - 0.136) - (0.05)}{\sqrt{6.77 \times 10^{-6}\left(\frac{1}{25}+\frac{1}{25}\right)}}$$

$$= -140$$

* Decision :- $t < t_{\alpha,\nu}$

$H_0$ is Accepted

Hence, the claim is substantiated

Q) In an random experiment to compare two types of animal foods A and B following results of increased in weights are observed in animals.

| Animal no. | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| Increase in wt | Food A | 49 | 53 | 51 | 52 | 47 | 50 | 52 | 53 |
| | Food B | 52 | 55 | 52 | 53 | 50 | 54 | 54 | 53 |

Assuming that the two samples of animals are independent, can we conclude that food B is better than food A?

$H_0$ : ~~$\mu_1 \leq \mu_2$~~ $\mu_1 = \mu_2$

$\Rightarrow$ $H_a$ : ~~$\mu_1 \neq \mu_2$~~ A → B (~~Right~~ left hand tail)

$\mu_1 < \mu_2$ ~~mean~~ → B > A

$\bar{x}_1 = \frac{49+53+51+52+47+50+52+53}{8}$

$\bar{x}_1 = 50.875$

$\bar{x}_2 = 52.875$

$(x_i - \bar{x})^2 = 3.515 \mid 4.515 \mid 0.015 \mid 1.26 \mid 15.01 \mid 0.076 \mid 1.26$

$(y_i - \bar{y}) = 0.765 \mid 4.515 \mid 0.765 \mid 0.015 \mid 8.26 \mid 1.265 \mid 1.265$

$\Sigma(x_i - \bar{x})^2 = 30.85$

$\Sigma(y_i - \bar{y})^2 = 16.865$

$s^2 = \frac{1}{8+8-2}[30.85 + 16.865]$

$s^2 = 3.408$

$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_x - \mu_y)}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

$t = -2.166$

~~los~~ los = 5%

$\alpha = 0.05$

DOF ($\nu$) = $n_1 + n_2 - 2 = 14$

$t_{\alpha,\nu} = t_{0.05,14} = \mp 1.761$

$\therefore$ ~~$t > t_{\alpha,\nu}$~~

$\therefore$ $H_0$ is ~~accepted~~ Rejected

$H_a$ is accepted

$\therefore$ we can conclude that food B is ~~not~~ better than food A

* $\chi^2$ - Distribution (chi-square distribution)

① Goodness to fit

* $$\chi^2 = \sum_{i=1}^{n} \left[ \frac{(f_i - e_i)^2}{e_i} \right]$$

Note: ① $\sum_{i=1}^{n} f_i = \sum_{i=1}^{n} e_i$

② Degree of freedom ($\nu$) = $n-1$

Ex. ① The demand of particular spare part in a factory was found to vary from day-to-day. In a sample study the following information was obtained

| days | Mon | Tue | wed | Thu | Fri | Sat |
|---|---|---|---|---|---|---|
| no. of parts demanded | 1124 | 1125 | 1110 | 1120 | 1126 | 1115 |

Test the hypothesis that the no. of parts demanded does not depends on days of the week. (los = 5%)

Sol$^n$: $H_0$: No. of spare part demanded does not depends on days of week.

$H_a$: The no. of parts demanded depends on days of week.

* Calculation:

• The average ~~spare~~ no. of spare part demanded on a day is

Expected frequency $(e_i) = \frac{\sum f_i}{n} = 1120$

| Days | frequencies $f_i$ | $e_i$ | $(f_i - e_i)^2$ | $\frac{(f_i - e_i)^2}{e_i}$ |
|---|---|---|---|---|
| Mon | 1124 | 1120 | 16 | ~~0.007~~ 0.014 |
| Tue | 1125 | 1120 | 25 | 0.022 |
| Wed | 1110 | 1120 | 100 | 0.089 |
| Thu | 1120 | 1120 | 0 | 0 |
| Fri | 1126 | 1120 | 36 | 0.032 |
| Sat | 1115 | 1120 | 25 | 0.022 |

$$\chi^2 = \sum_{i=1} \frac{(f_i - e_i)^2}{e_i} = 0.179 \quad \cancel{0.18}$$

$\nu = n - 1 = 6 - 1 = 5$

$\chi^2_{\alpha,\nu} = \chi^2_{0.05,5} = 11.07$

* Decision:

$\chi^2 > \chi^2_{\alpha,\nu}$

$H_0$ is accepted

Ex.② A Sample Analysis of Examination results of 200 MBA students was. It was found that 46 students was failed, 68 secured 3rd div, 62 secured 2nd div, rest placed in 1st div. Are this figures commensurate with the general examination results which is in ratio 4:3:2:1 for various categories respec?

=> ~~Ho: Given data commensurate with~~

* Ho: Observed data do not differ significanly from hypothetical frequencies which are in ration 4:3:2:1

* Ha: ——— differ significantly

* Calculation:
Table for $\chi^2$

| category | frequencies $f_i$ | $e_i$ | $(f_i-e_i)^2$ | $\frac{(f_i-e_i)^2}{e_i}$ |
|---|---|---|---|---|
| ~~faied~~ | | | | |
| failed | 46 | 80 | 1156 | 14.45 |
| III division | 68 | 60 | 64 | 1.06 |
| II division | 62 | 40 | 484 | 12.1 |
| I division | 24 | 20 | 16 | 0.8 |

$$\boxed{\chi^2 = \sum \frac{(f_i-e_i)^2}{e_i} = 28.41}$$

* los: $\alpha = 5\% = 0.05$
* DOF: $\nu = n-1 = 4-1 = 3$
* Critical value: $\chi^2_{\alpha,\nu} = 7.815$

* Decision: $\chi^2 > \chi^2_{\alpha,\nu}$

Ho is rejected
Ha is Accepted

observed data differ significanty from hypothetical frequencies which are in ratio 4:3:2:1

② Independent of Attributes:-

for 2x2 contigency table $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$

$$\chi^2 = \frac{N(ad-bc)^2}{(a+c)(b+d)(a+b)(c+d)}$$

* DOF: $\nu = (m-1)(n-1)$ $\boxed{m \Rightarrow \text{rows},\ n \Rightarrow \text{colums}}$

Ex.① Two sample polls of votes for Two candidate of A & B for a public office are taken, one from among the resident of ruler areas. The results are given in adjoining table. examine whether the nature of area is related to voting preference in this election

| Area | Votes for | |
|---|---|---|
| | A | B |
| Rural | 620 | 380 |
| Urban | 550 | 450 |

Soln: Ho : The nature of areas is independent of voting preference in this selection

Ha : —"— is not independent of —"—

* Calculation :

$n$ = Total frequency = Total no. of voters
$= \Sigma f_i$
$= 620 + 380 + 550 + 450$
$= 2000$

$$\chi^2 = \frac{n(ad-bc)^2}{(a+c)(b+d)(a+b)(c+d)}$$

$$= \frac{2000\,[(620 \times 450) - (380 \times 550)]^2}{(620+550)(380+450)(620+380)(550+450)}$$

$= 10.09$

* los : $\alpha = 5\% = 0.05$
* DOF $= \nu = (m-1)(n-1) = (2-1)(3-1) = 1$

* Critical value $= \chi^2_{\alpha,\nu} = 3.84$

* Decision $\chi^2 > \chi^2_{\alpha,\nu}$

Ho is rejected
Ha is Accepted

Ex.3 Two researcher adapted different sampling ~~taking~~ while the same group of student to investigating find no. of students following different level

| Researches | No. of students in each level | | | | Total |
|---|---|---|---|---|---|
| | Below Avg | Avg | above avg | Genius | |
| X | 86 | 60 | 40 | 10 | 200 |
| Y | 40 | 33 | 25 | 2 | 100 |
| | 126 | 93 | 69 | 12 | 300 |

Would you say that the sampling technique adapted

Ho : The observed data is independent of the sampling technique adapted

Ha : —"— is not independent of —"—

* Calculation :- This 2x4 cont. table we find expected freq.

$$E(86) = \frac{126 \times 200}{300} = 84$$

$$E(60) = \frac{93 \times 200}{300} = 62$$

$$E(44) = \frac{69 \times 200}{300} = 46$$

Expected freq. table

| Researchers | No. of student | | | | Total |
|---|---|---|---|---|---|
| X | ~~80~~ 84 | 62 | 46 | 8 | 200 |
| Y | 42 | 31 | 23 | 4 | 100 |
| Total | 126 | 93 | 69 | 12 | 300 |

## * Properties of eigenvalues and eigenvectors
[A] For matrix A

1] Sum of all the eigenvalues = Trace [A]

2] Product of all eigenvalues = $|A|$

3] The matrices A and $A^T$ have same eigenvalues

4] The eigenvalues of diagonal or triangular matrices are its diagonal element.

5] Eigenvalues of symmetric matrix are real.

6] Eigenvalues of skew-symmetric matrix are either ~~'0' or~~ zero or purely imaginary.

7] let, $f(x)$ be a algebra in polynomial with real coefficient and A be the square matrix of order n, with eigen values $\lambda_1, \lambda_2 \cdots \lambda_n$ then the eigen values of a matrix $f(A)$ are $f(\lambda_1), f(\lambda_2), f(\lambda_3) \ldots f(\lambda_n)$

i] The eigen values of a matrix $kA$ are $k\lambda_1, k\lambda_2, \ldots k\lambda_n$ where k is scalar.

ii] The eigenvalues of a matrix $A^m$ are $\lambda_1^m, \lambda_2^m, \lambda_3^m, \ldots, \lambda_m^m$ if m is +ve integer

iii] Eigenvalues of a matrix $A^{-1}$ are $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \ldots \frac{1}{\lambda_n}$

iv] Eigen vectors of a matrix A and $f(A)$ are same

8] The eigenvalues of a matrix adj(A) are $\frac{|A|}{\lambda_1}, \frac{|A|}{\lambda_2}, \ldots, \frac{|A|}{\lambda_n}$

Imp

* Result :-

① The characterstics polynomial of 2x2 matrix A is of the form $\lambda^2 - Tr(A).\lambda + |A| = 0$

② The characterstic eqn of 3x3 matrix A is of the form $\lambda^3 - tr(A).\lambda^2 + (A_{11}+A_{22}+A_{33})\lambda - |A| = 0$

where $A_{11}, A_{22}, A_{33} \Rightarrow$ minors of diagonal elements

Ex. ③ Find eigenvalues and eigenvectors of a matrices $A^3 + 2A^1 - A^{-1}$ and adj (adj A)) where

$$A = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix}$$

Ans:- for eigen value of A, consider

$$|A - \lambda I| = 0$$

$$\begin{vmatrix} 2-\lambda & 2 & 1 \\ 1 & 3-\lambda & 1 \\ 1 & 2 & 2-\lambda \end{vmatrix} = 0$$

$$\Rightarrow \lambda^3 - tr(A).\lambda^2 + (A_{11}+A_{22}+A_{33})\lambda - |A| = 0$$

$$= \lambda^3 - 7\lambda^2 + (\cancel{5}\,4 + 3 + 4)\lambda - 5 = 0$$

$$= \cancel{-6\lambda^2}\ \lambda^3 - 7\lambda^2 + 11\lambda - 5 = 0$$

$\lambda = 5, 1, 1$ are eigen values

① For $\lambda = \lambda_1 = 1$

Consider $[A - \lambda I] = 0$

$[A - \lambda I] = 0$

$$\begin{bmatrix} 1 & 2 & 1 \\ 1 & 2 & 1 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\rightarrow x_1 + 2x_2 + x_3 = 0$

$\therefore$ here $x_2, x_3$ are non-leading

$\therefore$ Put $x_2 = s$, $x_3 = t$

$\therefore x_1 = -2s - t$

$$\therefore X = \begin{bmatrix} -2s-t \\ s \\ t \end{bmatrix} = s\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

$$\therefore X_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \quad X_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

② For $\lambda = \lambda_2 = 5$

Consider $[A - \lambda I]X = 0$

$[A - 5I]X = 0$

$$\begin{bmatrix} -3 & 2 & 1 \\ 1 & -2 & 1 \\ 1 & 2 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_3 \rightarrow R_3 - R_2$, ~~$R_2 \rightarrow R_2 + 2$~~ $R_2 \rightarrow R_2 + R_1$

$$\begin{bmatrix} -3 & 2 & 1 \\ -2 & 0 & 2 \\ 0 & 4 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$4x_2 = 4x_3$

$-2x_1 = -2x_3 \quad x_2 = x_3 = t$

$x_1 = x_3 = t$

$-3x_1 + 2x_2 + x_3 = 0$, $-2x_1 + 2x_3 = 0$, $4x_2 - 4x_3 = 0$

$$x = \begin{bmatrix}1\\1\\1\end{bmatrix}$$