Lecture 23:-

Distribution function:-

The distribution function of random variable X, denoted by $F(x) = P(X < x)$. The function $F(x)$ gives the Probabilty of the event that X takes a value less than or equal to a specified value.

اس میں ہم 3 سکوں کو ایک ہی وقت میں toss کرتے ہیں اور پھر probabilty نکالیں گے تو ہمیں df ملے گا۔

$S = \{HHH, HHT, HTT, THH, HTH, TTH, HTT, TTT\}$ ٹوٹل 8 ہیں جب Head (0) ہوگا تو probability (1/8) ہوگی۔۔۔ جب Head (1) ہوگا تو Probabilty (3/8) ہوگی جب Head (2) ہوگا تو بھی (3/8) ہوگی اور جب Head (3) ہوں گے تو Probability 1/8 نکلے گی۔ پھر اس کی Commulative Probability نکالیں گے جو نتیجہ نکلے گا وہ یہ ہوگا:

$$P(X < x) = P(X = 0) = \frac{1}{8}$$

- For $1 < x < 2$, we have

$$P(X < x) = P(X = 0) + P(X = 1).$$
$$\frac{1}{8} + \frac{3}{8} = \frac{4}{8}$$

- For $2 < x < 3$ we have,

$$P(X < x) = P(X = 0) + P(X = 1) + P(X = 2)$$
$$\frac{1}{8} + \frac{3}{8} + \frac{3}{8} = \frac{7}{8}$$

- For $x \geq 3$ we have

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

$$\frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1$$

So we concluded that if we toss 3 coins, the ~~probability that one head must is 4/8~~ and must two Head is 7/8.

Mathematical Expectation:-

$E(X)$ is called mathematical expectation or expected value and it is also called mean of (X) and is denoted by μ.

Expression:-

$$E(X) = \sum_{i=1}^{n} x_i f(x_i).$$

Skewness Formula:-

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3}.$$

Kurtosis formula:-

$$\beta_2 = \frac{\mu_4}{\mu_2^2}.$$

Properties of Mathematical Expectation:-

- If c is a constant then $E(C) = C$.
- If X is a discrete random variable and if a and b are constants then $E(aX+b) = aE(X)+b$

جیسے اوپر ہم نے Coins والی Example کی ہے اس کی probabilities کو X سے Multiply کریں گے ہم Expected value نکال سکتے ہیں جو کہ 1.5 بنتی ہے ----

X = no. of Heads, $P(x)$ = Probability.

End.

Lecture # 24.

chebyechev's Inequality.

If x is a random variable having mean μ and variance $\sigma^2 > 0$, and K is any positive integer, then the probability that a value of x falls within K-standard deviation of the mean is at least is:-

$$P(\mu - K\sigma < X < \mu + K\sigma) \geq 1 - \frac{1}{K^2}.$$

اس کا استعمال تب ہوگا جب ہم بہت سی probabilities سے کچھ probabilities کو find کرنا چاہتے ہوں گے...

Example:- Biologist → Petals.

No. of petals (x)	$P(x)$	$x(P(x))$	$x^2(P(x))$
$x_1 = 3$	0.05	0.15	0.45
$x_2 = 4$	0.10	0.4	1.6
$x_3 = 5$	0.20	1	5
$x_4 = 6$	0.30	1.8	10.8
$x_5 = 7$	0.25	1.75	12.25
$x_6 = 8$	0.075	0.6	4.8
$x_7 = 9$	0.025	0.225	2.025
	1	5.925	36.925

$\mu = E(X) = 5.925.$

$E(x)^2 = 36.925.$

ہم نے μ find کر لیا اب دوسری چیز σ find کرنی ہے۔

$$\sigma^2 = E(x^2) - [E(x)]^2$$

$$\sigma^2 = 36.925 - (5.925)^2$$

$$\sqrt{\sigma^2} = \sqrt{36.925 - (5.925)^2}$$

$$\sigma = \sqrt{36.925 - 35.106}$$

$$\sigma = \sqrt{1.819} \Rightarrow 1.3$$

k کی ایک specific value ہوتی ہے $k = 2$

$$1 - \frac{1}{k^2} \Rightarrow 1 - \frac{1}{2^2} \Rightarrow 1 - \frac{1}{4} \Rightarrow \frac{3}{4} \Rightarrow 0.75.$$

اسکا مطلب 75% lie کر رہا ہے تو ہم کہہ سکتے ہیں کہ $\mu+2\sigma$ اور $\mu-2\sigma$ کے درمیان lie کرتا ہے۔

$5.9 - 2(1.3)$ and $5.9 + 2(1.3)$

3.3 and 8.5.

3.3 اور 8.5 x کی value میں دیکھا ہے اور 3.3 سے بڑی اور 8.5 سے چھوٹی x_2, x_3, x_4, x_5, x_6 کو لیں گے۔۔۔

x کے according probability ہے اسکا sum کرنا ہے۔

$$0.10 + 0.20 + 0.30 + 0.25 + 0.075 \Rightarrow 0.925.$$

اب ہمارے پاس 0.925 value نکلی ہے اور اوپر 0.75 نکلی تھی۔۔۔ ہم فارمولا لگاتے ہیں۔

$$P(\mu - k\sigma < x < \mu + k\sigma) \geq 1 - \frac{1}{k^2}.$$

$$0.925 \geq 0.75$$

(because as we calculate $1 - \frac{1}{k^2} = 0.75$ and $P(\mu - k\sigma < x < \mu + k\sigma) = 0.925$).

$0.925 \geq 0.75$ so it is true and Here the Chebyechev's hold.

Example:- Business Venture:-

Find $\mu = E(X)$ and $\sigma = \sqrt{E(X-\mu)^2}$

اب اس میں σ کا فارمولا change ہو گیا ہے۔

X	p(x)	E(X)	X − μ	$(X-\mu)^2 p(x)$
0	0.002	0	−3.5	0.02
1	0.029	0.029	−2.5	0.18
2	0.132	0.264	−1.5	0.30
3	0.309	0.927	−0.5	0.08
4	0.360	1.44	0.5	0.09
5	0.168	0.84	1.5	0.38
		3.5		1.05

a) $E(X) = X\,p(X)$

$\mu = E(X) = 3.5$

b) $\sigma = \sqrt{E(x-\mu)^2}$.

$\sigma = \sqrt{1.05} \Rightarrow 1.02$

یہاں پہ نوٹ کرنے والی بات یہ ہے کہ

$(X-\mu)^2 p_x = E(x-\mu)^2$

c) گراف بناتے وقت X کی values x-axis پہ probability y-axis پہ لیتے ہیں۔

$P(\mu - k\sigma < x < \mu + k\sigma) \Rightarrow (1.46,\ 5.54)$.

$P(1.46 < x < 5.54)$

$p(1.46 < x < 5.54) \Rightarrow 0.132 + 0.309 + 0.360 + 0.168 \Rightarrow 0.969$

$X_2 + X_3 + X_4 + X_5 \Rightarrow 1.32 + 0.3.09 + 0.360 + 0.168$

Successful venture:-

$P(0) + P(1) \Rightarrow 0.002 + 0.029$

$\Rightarrow 0.03$.

Continous Random variable:-

A random variable x is defined to be Continous if can assume every possible value in an interval [a, b] may be $-\infty$ and $+\infty$. The Function $f(x)$ is called Probability density function (p.d.f)

p.d.f has following properties:-

1) $f(x) \geq 0$ for all x
2) $\int_{-\infty}^{\infty} f(x)\,dx = 1$
3) $\int_{c}^{d} f(x)\,dx = p(c < x < d)$.
4) $P(x = k) = \int_{k}^{k} f(x)\,dx = 0$

i) $f(x) > 0$

ii) $F(x) = \int_{-\infty}^{x} f(x)\,dx$

$$\frac{d}{dx}F(x) = f(x).$$

Example:-

$$\text{Given } f(x) = \begin{cases} kx, & 0 < x < 2 \\ 0 & \text{elsewhere.} \end{cases}$$

a) $k = ?$

As we know that $\int_{-\infty}^{\infty} f(x)\,dx = 1$.

$$\int_{-\infty}^{+\infty} k(x)\,dx = 1.$$

$$\int_{-\infty}^{0} kx\,dx + \int_{0}^{2} kx\,dx + \int_{2}^{\infty} kx\,dx = 1.$$

$$0 + k\int_{0}^{2} x\,dx + 0 = 1$$

$$k\left|\frac{x^{1+1}}{1+1}\right|_0^2 = 1.$$

$$k\left|\frac{x^2}{2}\right|_0^2 = 1 \Rightarrow \frac{k}{2}\left((2)^2-(0)^2\right) = 1.$$

$$\frac{k}{2}(4) = 1 \Rightarrow 2k = 1 \Rightarrow k = \frac{1}{2}.$$

b) $P(X > 1) = \int_1^2 f(x)\,dx$.

$$\int_1^2 \frac{x}{2}\,dx = \left|\frac{x^2}{4}\right|_1^2 \Rightarrow \frac{(2)^2-(1)^2}{4} \Rightarrow \frac{4-1}{4} \Rightarrow \frac{3}{4}.$$

c) For any X such that $-\infty < X \leq 0$

$$F(x) = \int_{-\infty}^{x} 0\,dx = 0.$$

d) If $0 < x \leq 2$ we have.

$$F(x) = \int_{-\infty}^{0} 0\,dx + \int_0^x \left(\frac{x}{2}\right)dx.$$

$$0 + \int_0^x \frac{x^2}{4} \Rightarrow \frac{x^2}{4}.$$

For $x > 2$

we have $\int_0^2 \frac{x}{2}\,dx = 1.$

Hence $F(x) = 0$ for $x < 0$.

$= \frac{x^2}{4}$ for $0 \leq x \leq 2$.

$= 1$ for $x > 2$ —

Lecture # 25:-

لیکچر 24 کا e پارٹ کرنا ہے۔

e) $P(X < 1/2 \mid 1/3 < x < 2/3)$.

$$= \frac{P(\frac{1}{3} \le x \le \frac{1}{2})}{P(\frac{1}{3} \le x \le \frac{2}{3})} \Rightarrow \frac{\int_{1/3}^{1/2} \frac{x}{2}\,dx}{\int_{1/3}^{2/3} \frac{x}{2}\,dx}$$

$$\Rightarrow \frac{\left[\frac{x^2}{4}\right]_{1/3}^{1/2}}{\left[\frac{x^2}{4}\right]_{1/3}^{2/3}} \Rightarrow \frac{5}{36} \times \frac{3}{1} = \frac{5}{12}.$$

Example:- اِس example میں $\hat{F}(x)$ کی value دی گئی ہیں۔ ہم جانتے ہیں کہ جب $F(x)$ ہو تو وہ $f(x)$ کا derivative ہوتا ہے۔ So ہم Given function پہ Derivative لگا کر Example کو solve کریں گے۔ جیسا کہ ہے:-

$$F(X) = \frac{2x^2}{5}, \text{ for } x < X \le 1.$$

$$F(X) = \frac{d}{dx}\frac{2x^2}{5}.$$

$$\Rightarrow \frac{2}{5}(2x) \Rightarrow \frac{4x}{5}.$$

Example:-

$$f(x) = 2(1-x),\ 0 < x < 1$$
$$= 0 \text{ elsewhere}.$$

Now, $E(x) = \int_{-\infty}^{+\infty} x f(x)\, dx$.

$E(x) = \int_{0}^{1} x\, 2(x-1)\, dx$.

$Ex = 2\int_{0}^{1} x(x-1)\, dx \Rightarrow 2\int_{0}^{1} x - x^2\, dx$

$2\Big|_0^1 \frac{x^2}{2} - \frac{x^3}{3} \Rightarrow 2\left[\frac{1}{2} - \frac{1}{3}\right] \Rightarrow \frac{1}{3} \Rightarrow 0.333$

$E(ax+b) = aE(x)+b$.

Suppose $a=3$, $b=5$.

R.H.S $= aE(x)+b \Rightarrow 3\left(\frac{1}{3}\right)+5 \Rightarrow 6$.

L.H.S $\Rightarrow E(3x+5)$.

$\Rightarrow \int_{0}^{1} (3x+5)\, 2(1-x)\, dx$.

$\Rightarrow 2\int_{0}^{1} (3x+5)(1-x)\, dx$

$\Rightarrow 2\int_{0}^{1} (5-2x-3x^2)\, dx$.

$\Rightarrow 2\Big|_0^1 5x - x^2 - x^3$.

$\Rightarrow 2(5-1-1) \Rightarrow 2(3) \Rightarrow 6$.

If $b=0$

$E(ax) = a\,E(x)$.

Example:-

$$f(x) = \frac{3}{4}x(2-x), \quad 0 \le x \le 2.$$

Find moment and momment ratios

Fist of all we compute first four momments about zero.

$$\mu'_1 = E(X) = \int_{-\infty}^{+\infty} x f(x)\, dx$$

$$\Rightarrow \frac{3}{4}\int_0^2 x(2x - x^2)\, dx \Rightarrow \frac{3}{4}\int_0^2 2x^2 - x^3.$$

$$\Rightarrow \frac{3}{4}\left|\frac{2x^3}{3} - \frac{x^4}{4}\right|_0^2 \Rightarrow \frac{3}{4}\left(\frac{16}{3} - \frac{16}{4}\right) \Rightarrow \frac{3}{4}\left(\frac{16}{12}\right)$$

$$\Rightarrow \frac{4}{4} \Rightarrow 1$$

$\mu'_2 = E(X^2).$ اسی طرح compute کر لیں

$= \frac{6}{5}$ value

$$\mu'_3 = E(X^3) = \frac{8}{5}$$

$$\mu'_4 = E(X^4). = \frac{16}{7}$$

Momment about the mean:-

1) $\mu_1 = 0.$

2) $\mu_2 = \mu'_2 - (\mu'_1)^2 \Rightarrow \frac{6}{5} - (1)^2 \Rightarrow \frac{1}{5}.$

3) $\mu_3 = \mu'_3 - 3\mu'_1\mu'_2 + 2(\mu'_1)^3.$

$\mu_3 = \frac{8}{5} - 3(1)\left(\frac{6}{5}\right) + 2(1)^3.$

$\Rightarrow \frac{8}{5} - \frac{18}{5} + 2 \Rightarrow$

$$\mu_4 = \mu'_4 - 4\mu'_1\mu'_3 + 6(\mu'_1)^2\mu'_2 - 3(\mu'_1)^4.$$

$$\mu_4 = \frac{16}{7} - 4(1)\left(\frac{8}{5}\right) + 6(1)^2\left(\frac{6}{5}\right) - 3(1)^4.$$

$$\Rightarrow \frac{16}{7} - \frac{32}{5} + \frac{36}{5} - 3 = \frac{3}{35}.$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(0)^2}{\left(\frac{1}{5}\right)^3} \Rightarrow 0$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} \Rightarrow \frac{3/35}{(1/5)^2} \Rightarrow 2.14.$$

Joint probability distribution:-

It is also called Bivariate probability function-

$$f(x_i, y_j) = P(X = x_i \text{ and } Y = y_j).$$

X-values کو ہم 1st کالم میں رکھیں گے

اور Y-values کو ہم 1st رو میں رکھیں گے

اور ان سب کا sum 1 ہوگا۔

X-values کے sum کو g(x) کہیں گے اور

Y-values کے sum کو h(y) کہیں گے۔

Properties:-

1) $\sum_i \sum_j f(x_i, y_j) = 1$

i) $f(x_i, y_j) \geq 0$ for all
(x_i, y_j) for $i = 1, 2, 3 \ldots m$ and $j = 1, 2 \ldots m$

Marginal Probability:-

Let $f(x, y)$ be joint probability function of two discrete r.v.'s X and Y. Then the marginal Probability function of X is defined as

$$g(x_i) = \sum_{j=1}^{n} f(x_i, y_j)$$

In other words $= f(x_i, y_1) + f(x_i + y_2) \ldots f(x_i, y_n)$

پھر x_i کی جگہ 1,2,3 put کر کے solve کریں گے

Example: Balls:-

$f(0,0)$ $P(X=0$ and $y=0)$.

$\binom{3}{0}\binom{2}{0}\binom{3}{2} = 3.$

The total number of ways in which 2 balls drawn out of total 8 $\binom{8}{2} = 28$.

So $f(0,0) = 3/28$.

Lecture 26.:-

26 میں کچھ اہم نہیں ہے بارش کے دن ہیں ...

lecture 26:-

2) compute $P(X+Y \leq 1)$ for the cells $(0,0)$, $(0,1)$ and $(1,0)$.

$f(0,0) + f(0,1) + f(1,0)$

$3/28 + 6/28 + 9/28$

$18/28 \Rightarrow 9/14$

3) marginal probability:-

$g(x)$ کو اور $h(y)$ کو marginal probability کہتے ہیں۔

So

X	0	1	2
$g(x)$	10/28	15/28	3/28

Y	0	1	2
$h(y)$	15/28	12/28	1/28

4) Conditional probability:-

$f(x|1) = P(X = x \mid Y = 1)$

$\Rightarrow$ X represents Random variable.

x represents any particular value of variable.

1) $f(0,1) = \frac{7}{3} f(0,1) = \frac{7}{3}\left(\frac{6}{28}\right) = \frac{1}{2}$

2) $f(1,1) = \frac{7}{3} f(1,1) = \frac{7}{3}\left(\frac{6}{28}\right) = \frac{1}{2}$

3) $f(2,1) \Rightarrow \frac{7}{3} f(2,1) = \frac{7}{3}(0) = 0$

X and Y are ^not^ independent.
because $\frac{6}{28} \neq \frac{12}{28} \times \frac{10}{28}$

Bivariate Continous distribution:-

i) $f(x,y) \geq 0$ for all (x,y)

ii) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$. (Joint density function).

iii) $P(a \leq x \leq b,\ c \leq y \leq d)$
$= \int_{a}^{b}\int_{c}^{d} f(x,y)\,dy\,dx$.

Remember:- Marginal Probability of x and Y.

M.d.f X = $g(x) = \int_{-\infty}^{\infty} f(x,y)\,\underline{dy}$.

m.d.f Y = $h(y) = \int_{-\infty}^{\infty} f(x,y)\,\underline{dx}$.

Remember:- Conditional Probability:-

$$f(X \mid Y) = \frac{f(x,y)}{h(y)}$$

$$f(Y \mid X) = \frac{f(x,y)}{g(x)}$$

First one is conditional Probability of X when $Y = y$.
Second one is Conditional Probability of Y when $X = x$.

Example 1 main sary formulas jo btaye hn wo put kr k simplify krna hai -

Property No1:-

1) $E(X+Y) = E(X) + E(Y)$

2) $E(X-Y) = E(X) - E(Y)$

3) $E(XY) = E(X)\,E(Y)$

Table دیکھنا ہے

Example:-

Formula used:-

1) $E(X) = \sum x_j\, g(x_j)$

2) $E(Y) = \sum y_i\, h(y_i)$

$E(X) = \sum x_j\, g(x_j)$

$\Rightarrow 2 \times 0.40 + 4 \times 0.60 \Rightarrow 3.2$

$E(Y) = \sum y_i\, h(y_i)$

$\Rightarrow 1 \times 0.25 + 3 \times 0.50 + 5 \times 0.25 \Rightarrow 3.0$

$E(X) + E(Y) \Rightarrow 3.2 + 3.0 \Rightarrow 6.2$

Lecture 27:-

x\y	1	3	5	g(x)
2	0.10	0.20	0.10	0.40
4	0.15	0.30	0.15	0.60
h(y)	0.25	0.50	0.25	1.00

Find $E(X)$, $E(Y)$, $E(X+Y)$ and $E(XY)$.

Now,

$E(X) = \sum x_i g(x_i)$.

$2 \times 0.40 + 4 \times 0.60 \Rightarrow 0.80 + 2.40 = 3.2$.

$E(Y) = \sum y_j h(y_j)$.

$= 1 \times 0.25 + 3 \times 0.50 + 5 \times 0.25 \Rightarrow 3.0$.

$E(X+Y) = E(X) + E(Y)$.

So, $3.2 + 3.0 \Rightarrow 6.2$

$E(XY) \Rightarrow (3.2)(3.0) \Rightarrow 9.6$.

other Method.

$E(X+Y) = \sum_i \sum_j (x_i + y_j) f(x_i, y_j)$.

$\Rightarrow (2+1)(0.10) + (2+3)(0.20) + (2+5)(0.10) +$
$(4+1)(0.15) + (4+3)(0.30) + (4+5)(0.15) \Rightarrow 6.20$.

So, it is proved.

other method of product $E(XY)$.

$E(XY) = \sum_i \sum_j (x_i, y_j) f(x_i, y_j)$.

$\Rightarrow (2\times1)(0.10) + (2\times3)(0.20) + (2\times5)(0.10) + (4\times1)(0.15)$
$+ (4\times3)(0.30) + (4\times5)(0.15) \Rightarrow 9.6$.

So, it is proved That
Same answer.

Example 2:-

Let x and y are independent r.v.'s with joint p.d.f.

$$f(x,y) = \frac{x(1+3y^2)}{4},$$

$0<x<2, 0<y<1$ 0 (elsewhere).

Find $E(X)$, $E(Y)$, $E(X+Y)$, $E(XY)$.

$$g(x) = \int_{-\infty}^{\infty} f(x,y)\,dy.$$

$$\Rightarrow \int_0^1 \frac{x(1+3y^2)}{4}\,dy$$

$\frac{d}{dy}x \Rightarrow xy$

$3xy^2 \Rightarrow \frac{3xy^3}{3}$

xy^3.

$$\Rightarrow \int_0^1 \frac{x+3xy^2}{4}\,dy$$

$$\Rightarrow \frac{1}{4}\left[xy + xy^3\right]_0^1 \Rightarrow \frac{1}{4}\left[x(1-0) + x(1-0)^3\right].$$

$$\Rightarrow \frac{1}{4}\left[x+x\right] \Rightarrow \frac{2x}{4} \Rightarrow \frac{x}{2}.$$

$$h(y) = \int_{-\infty}^{\infty} f(x,y)\,dx.$$

$$\int_0^2 \frac{x(1+3y^2)}{4}\,dx -$$

$$\Rightarrow \int_0^2 \frac{x+3xy^2}{4}\,dx \Rightarrow \frac{1}{4}\left[\frac{x^2}{2} + 3xy^2\right]_0^2$$

$$\Rightarrow \frac{1}{4}\left[\frac{(2-0)^2}{2} + 3(2-0)y^2\right].$$

$$\Rightarrow \frac{1}{4}\left[2+6y^2\right] \Rightarrow \frac{1}{4\cdot 2}(1+3y^2)$$

$$E(X) = \int_{-\infty}^{\infty} x\, g(x)\, dx.$$

$$\int_0^2 x\left(\frac{x}{2}\right) dx = \frac{1}{2}\left[\frac{x^3}{3}\right]_0^2 \Rightarrow \frac{1}{2}\left[\frac{8}{3}\right] \Rightarrow \frac{4}{3}$$

$$E(Y) = \int_{-\infty}^{\infty} y\, h(y)\, dy.$$

$$\Rightarrow \frac{1}{2}\int_0^1 y(1+3y^2)\, dy$$

$$\frac{1}{2}\left[\frac{y^2}{2} + \frac{3y^4}{4}\right]_0^1 \Rightarrow \frac{1}{2}\left[\frac{1}{2} + \frac{3}{4}\right] \Rightarrow \frac{5}{8}.$$

$$E(X+Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y) f(x,y)\, dx\, dy.$$

$$\int_0^2\int_0^1 (x+y)\, x\frac{(1+3y^2)}{4}\, dy\, dx.$$

$$\Rightarrow \int_0^2\int_0^1 \frac{x^2+3x^2y^2}{4} + \int_0^2\int_0^1 \frac{xy+3xy^3}{4}\, dy\, dx$$

$$\Rightarrow \int_0^2 \frac{1}{4}\left[x^2y + x^2y^3\right]_0^1 dx + \int_0^2 \frac{1}{4}\left[\frac{xy^2}{2} + \frac{3xy^4}{4}\right]_0^1 dx$$

$$\Rightarrow \int_0^2 \frac{1}{4}(2x^2)\, dx + \int_0^2 \frac{1}{4}\left(\frac{x}{2} + \frac{3x}{4}\right) dx$$

$$\Rightarrow \frac{1}{2}\left[\frac{x^3}{3}\right]_0^2 = \frac{1}{4}\left[\frac{x^2}{4} + \frac{3x^2}{8}\right]_0^2 \Rightarrow \frac{4}{3}.$$

$$\frac{4}{3} + \frac{5}{8} \Rightarrow \frac{47}{24}$$

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f(x,y)\, dx\, dy.$$

$$\Rightarrow \int_0^2\int_0^1 (x\,y)\,x\frac{(1+3y^2)}{4}\, dy\, dx.$$

$$\Rightarrow \int_0^2\int_0^1 \frac{x^2y+3x^2y^3}{4}\, dy\, dx.$$

$$\Rightarrow \int_0^2 \frac{1}{4}\left[\frac{x^2y^2}{2}+\frac{3x^2y^4}{4}\right]_0^1 dx.$$

$$\Rightarrow \int_0^2 \frac{1}{4}\left(\frac{5x^2}{4}\right) dx \Rightarrow \frac{1}{4}\left[\frac{5x^3}{12}\right]_0^2$$

$$\Rightarrow \frac{1}{4}\left[\frac{5(2-0)^3}{12}\right] \Rightarrow \frac{1}{4}\left(\frac{5(8)}{12}\right) \Rightarrow \frac{1}{4}\left(\frac{\cancel{40}^{10}}{12}\right)$$

$$\Rightarrow \frac{10}{12} \Rightarrow \frac{5}{6}.$$

It should be <u>Noted</u> that:-

i) $E(X) + E(Y) \Rightarrow \frac{4}{3} + \frac{5}{8}$

$\Rightarrow 47/24.$

ii) $E(X)E(Y) \Rightarrow \left(\frac{4}{3}\right)\left(\frac{5}{8}\right)$

$\Rightarrow \frac{20}{24} \Rightarrow \frac{5}{6}$ — — —

Covariance of Random variable:-

Main formula:- $Cov(X,Y) = E(XY) - E(X)E(Y)$ -

when the Cov is 0 then variables are independent otherwise not -

Correlation coefficient of Two r.v.'s:-

Main Formula:-

$$\rho_{xy} = \frac{Cov(x,y)}{\sqrt{Var(X)\,Var(Y)}}$$

Example:-

Main Formulas which are used in this example are:-

$E(X) = \sum x_i g(x_i)$

$E(Y) = \sum y_j h(y_j)$.

$E(X^2) = \sum x_i^2 g(x_i)$.

$E(Y^2) = \sum y_j^2 h(y_j)$.

$Var(X) = E(X^2) - [E(X)]^2$ -

$Var(Y) = E(Y^2) - [E(Y)]^2$ -

The formula of Cov and ρ is mentioned above

Example:-

اس میں $g(x)$ کو find out کرتے ہوئے y limits
دی گئی ہیں اور dy کے respect سے اس کو integrate کرنا
پھر $h(y)$ میں dx کے Respect اور x limits والی دی گئی
جب یہ دونوں find out ہو جاتی ہیں تو پھر $E(X)$
اور $E(Y)$ نکال لیں گے۔ پھر variance and correlation کو
Find کریں گے۔

Lecture 28:-

Binomial distribution-

Example:-

Suppose we toss a coin 5 times, and we are interested in determining the probability of x, where X represents numbers of Heads-

* Probability of Head is 1/2 and remains constant.

$$P(X=x) = \binom{n}{x} p^x q^{n-x}.$$

n = total no. of trials.

p = Probability of success in each trial

q = Probability of failure in each trial-

In this example,

$n=5$ and $p = 1/2$ and $q = 1-p \Rightarrow q = 1/2$.

Putting values:-

$$P(X=x) = \binom{5}{x}\left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{5-x}$$

$$X = 0, 1, 2, 3 \ldots\ldots$$

Subtituting all values of x step by step.

put $x=0$.

$$P(X=0) = \binom{5}{0}\left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^{5-0}$$

$$\Rightarrow \frac{5!}{0!\,5!}(1)\left(\frac{1}{2}\right)^5 \Rightarrow \frac{1}{32}.$$

for $x=1$ $\quad p(x=1) = \binom{5}{1}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^{5-1}$

$$\frac{5!}{1!\,5!}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^4 \Rightarrow 5\left(\frac{1}{2}\right)^5 = 5\left(\frac{1}{32}\right) = \frac{5}{32}$$

اسی طرح 5 تک ہم اس کی Probability نکالیں گے اور پھر جب اس کا Sum کریں گے تو 1 آئے گا

For a binomial distribution:-

$E(x) = np$

and variance $(x) = npq$.

$S.D(x) = \sqrt{npq}$ -

Mean = $E(x) = np$ $\Rightarrow 5\left(\frac{1}{2}\right) = 2.5$.

$S.D(x) = \sqrt{npq}$ $\Rightarrow \sqrt{5\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)}$

$\Rightarrow \sqrt{\frac{5}{4}} \Rightarrow 1.12$ -

Coefficient of variance -

$C.V = \frac{\sigma}{\mu} \times 100$ $\Rightarrow \frac{1.12}{2.5} \times 100 \Rightarrow 44.8\%$

Remember:-

1) If $p = q$ then we obtained absolutely symmetric distribution -

If $p < q$ (p is less than 1/2 and q is greater than 1/2) then our distribution is positively skewed -

3) If $\underline{p > q}$ (p is greater than 1/2 and q is less than 1/2) then our distribution is **negatively skewed**.

Hyper geometric Random variable:-

when the hyper geometric random variable x assumes a value x, the hyper geometric probability is given

$$P(X=x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$$

N = number of units in Population.

n = number of units in Sample.

K = number of Success.

(K is Success and N-K is Failure)

Lecture 29:-

Example:- 6 men and 5 women:-

Since $N=10$, $k=5$ and $n=4$, So,

$$p(X=x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}.$$

$$\Rightarrow \frac{\binom{5}{x}\binom{10-5}{4-x}}{\binom{10}{4}}$$

For $p(X=2) \Rightarrow \frac{\binom{5}{2}\binom{10-5}{4-2}}{\binom{10}{4}} \Rightarrow \frac{\binom{5}{2}\binom{5}{2}}{\binom{10}{4}}$

$$\Rightarrow \frac{10\times 10}{210} \Rightarrow \frac{10}{21}.$$

Properties of Hypergeometric distri[bution]

- The mean and variance of hyper geometric distribution are:-

  $\mu = n\frac{K}{N}$, $\sigma^2 = n\frac{k}{N}\frac{N-k}{N}\frac{N-n}{N-1}$,

- If N becomes indefinetly large, the hyper geometric p.d tends to Binomial distribution.
- There are Two ways of drawing a Sample From Population, Sampling with replacement and Sampling Without Replacement.

- Sample can be drawn from either finite Population or an infinite Population.
- when n is greater than 5 percent of N, the hyper geometric formula should be used.

(Here we use without replacement from a finite population.)

- With replacement either the Population is finite or infinite we use Binomial formula for calculation. They remains independent.
- Without replacement from a infinite Population we use Binomial formula because various trial becomes independent.

Poisson Distribution:

Two Situations:

a) It is a limiting approximation to Binomial distribution, when p is very small but n is so large that the $np = \mu$ is of a moderate size.

b) A Distribution in its own right by considering a POISSON PROCESS where events occur randomly over a specified interval of time, space or length.

When we use Poisson:-

1) p is 0.05 or less.
2) n is 20 or more.

مثال کے طور پر ایک ٹریفک روڈ پہ 24 گھنٹوں میں کتنے accident ہوتے ہیں ان کی Probability Find کرنے کے لیے Poisson distribution استعمال کرتے ہیں کیونکہ ہمیں پتا نہیں ہے اور نہ ہی اندازہ کر سکتے ہیں کہ کتنے ہوں گے کیونکہ یہ ایسا event ہے کہ کبھی بھی occur کر سکتا ہے اس کو ہم Random کہتے ہیں۔

Formula:-

$$\lim_{p \to 0} b(x; n, p) = \frac{e^{-\mu} \mu^{x}}{x!}, \quad x = 0, 1, 2 \ldots$$

where $e = 2.71828$.

Example:- Passanger of Airplane:-

Here $n = 200$, and $p = 0.01$.

In example, three will not show up

So, we use $P(X=3)$.

First of all we ~~use~~ find μ.

$$\mu = np \Rightarrow (200)(0.01) \Rightarrow 2.$$

$$P(X=3) = \frac{e^{-2}(2)^3}{3!}$$

$$\Rightarrow \frac{(0.1353)(8)}{3 \times 2 \times 1} \Rightarrow 0.1804.$$

Poisson Process Physically:-

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

λ = average no. of occurance of outcomes per unit of time

t = number of time units under consideration.

x = number of occurance of outcomes in t units.

Steps:-

1) Identify the unit of time
2) Identify λ.
3) Identify t.
4) Compute λt.
5) Apply the Poisson formula.

Properties of Poisson Distribution:-

- The mean and variance of Poisson distribution are same as: $E(X) = \mu$, $Var(X) = \mu$.
- The shape of Poisson Distribution is positevly skewed. It tends to symmetrical as μ becomes larger.

* Binomial distribution is symmetrical, positive skewed or negative skewed but Poisson distribution can never be negative skewed.

Uniform Distribution:—

Remember:—

Uniform distribution is also called Rectangular Distribution.

Definition:—

A Random variable x is Said to be uniformly distributed if its density function is defined as:—

$$f(x) = \frac{1}{b-a}, \quad a \le x \le b.$$

1) $\int_{-\infty}^{\infty} f(x)\,dx = 1$ -

2) Area of Recatangle = Base x Hight.

Parameters of Uniform Distribution are Two.

1) **Mean:—** $\mu = \frac{a+b}{2}$.

2) **Variance** = $\sigma^2 = \frac{(b-a)^2}{12}$ -

Lecture 30
Normal Distribution:-
A continous random variable is said to be normally distributed with mean (μ) and standard deviation (σ) if its probability destiny function is given by:

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

where $\pi = 3.1416$.
and $e = 2.71828$.
It has Two parameters:-
μ and σ.

1) The interval $\mu \pm \sigma$ will contains 68.26% of total area
2) $\mu + 2\sigma$ contains 95.44% Total area.
3) $\mu + 3\sigma$ contains 99.73% of Total area.

Point of inflection:-
$$\left(\mu-\sigma, \frac{1}{\sigma\sqrt{2\pi e}}\right) \text{ and } \left(\mu+\sigma, \frac{1}{\sigma\sqrt{2\pi e}}\right).$$

Standard Normal Distribution:-
A Normal distribution whose mean is zero and standard deviation is 1 is called standard Normal Distribution.
Process of standardization:-
$$Z = \frac{X-\mu}{\sigma}.$$

Example:- Dishwasher:-

Solution:-

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

X represents the warranty period is 1.0.

$\mu = 3.5$ while $\sigma = 1$, putting values.

$$Z = \frac{1.0 - 3.5}{1.0} \Rightarrow \frac{-2.5}{1} = -2.5.$$

Area from 0 to 2.5 = 0.4938.

Area from $-\infty$ to -2.5 is 0.0062 it means

Probability is 0.62% even less than 1%.

(we obtained 0.0062 by subtracting 0.4938 - 0.5)

Properties of Normal Distribution:-

1) For Normal Distribution $N(\mu, \sigma^2)$ where μ represents mean and σ^2 represent the standard deviation.

2) The normal curve is asymptotic to x-axis as $x \to \pm\infty$

(Asymptotic means curve approaches to line but actually not touches the line.

3) Normal distribution is absolutly symmetrical around the mean, Therefore 50% of the area is to the right of verticle line and 50% is to the left.

4) The density function attains its maximum value at $x = \mu$ and falls off symmetrically on each side of μ. This is why the mean, median and mode of the normal distribution are equal to μ.

5) Since it is symmetrical, hence μ_3 the third moment about the mean is zero.

6) For the normal distribution, forth moment about mean $= \mu_4 = 3\sigma^4$.

7) moment ratios of Normal distribution are 0 and 3 respectively.

$\beta_1 = 0$ and $\beta_2 = 3$.

As $\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(0)^2}{(\sigma^2)^3} \Rightarrow 0$.

$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3\sigma^4}{(\sigma^2)^2} = \frac{3\sigma^4}{\sigma^4} = 3$ -

Example:- Hights of applicants:-

ہمیں اس مثال میں mean (3.8) دیا گیا ہے اور
$\sigma = 170$ ہے جو 170 بالکل
mid میں ہوگا اور 50% کو
left پہ ہوگا اور 50% Right پہ ہوگا۔ لیکن ہمیں
left پہ 30% چاہیے اور
Right پہ 70% ۔ تو ہم فارمولا لگائیں گے

$$Z = \frac{X - \mu}{\sigma}.$$

can be written as $X = \mu + \sigma Z$ -

$X = 170 + 3.8 Z$ -

The area between $Z = 0$ and $Z = 0.52$ is 0.1985 -

The area between $Z = 0$ and $Z = 0.53$ is 0.2019.

Since 0.1985 is closer to 0.2000 than 0.2019, so 5.2 is taken as approximate Z-value -

$$X = 170 + 3.8 Z.$$

$$X = 170 + 3.8(-0.52)$$

$$X = 168.024 \cong 168\text{cm}$$

Hence, minimum acceptable Hight is 168cm -

Example:- Malaria -

Regarding "death from malaria" as Success.

$n = 500, \quad p = 0.20$

$P(70 \leq x \leq 80)$ -

$np = 500(0.2) \Rightarrow 100 > 5$

$nq = 500(0.8) \Rightarrow 400 > 5$

یہ دونوں 5 سے بڑے ہیں تو ہم Normal approximation
لگا سکتے ہیں ---

$\mu = np = 500(0.2) = 100.$

$\sigma^2 = npq = 500(0.2)(0.8) \Rightarrow 80.$

$\sqrt{\sigma^2} = \sqrt{80} \Rightarrow \sigma = 8.94$ -

Applying the Continuity correction.

$P(70 \leq x \leq 80)$

is replaced by $P(69.5 \leq x \leq 80.5)$.

So, $Z = \dfrac{x - \mu}{\sigma}$

$$Z = \frac{x - 100}{8.94}$$

Lecture 31:- Sampling Distribution

Example:- Cars and Faults.

No. of faulty items	Probablity $f(x)$	$x f(x)$	$x^2 f(x)$
0	1/5	0	0
1	1/5	1/5	1/5
2	1/5	2/5	4/5
3	1/5	3/5	9/5
4	1/5	4/5	16/5
Total	1	10/5 = 2	30/5 = 6

$$\mu = E(X) = \sum x f(x) = \underline{2}$$

$$\sigma^2 = var(X) = E(X)^2 - \left[E(X)\right]^2$$

$$\Rightarrow 6 - (2)^2 \Rightarrow 6 - 4 = 2$$

$$\sqrt{\sigma^2} = \sqrt{2} \Rightarrow \sigma = \sqrt{2} \Rightarrow \underline{\sigma = 1.414}$$

Factor

$$\sqrt{\frac{N-n}{N-1}}$$ is known as the finite population correction. (fpc).

In case of without replacement from a finite population.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

Remember:-

$$\mu_{\bar{x}} = \mu$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Central Limit Theorem:-

The theorem states that; If a Variable x from a Population has mean μ and finite variance σ^2, Then the Sampling distribution of Sample mean $\bar{x}$ approaches a Normal distribution with mean μ and variance σ^2/n as Sample Size "n" approaches infinity.

Note:- If the Population Sample is normally distributed, then sampling distribution of $\bar{x}$ will also be normal regardless of Sample Size.

In other words, $\bar{x}$ will be normally distributed with mean μ and variance σ^2/n.

Lecture 32:-

- Sampling distribution of $\hat{p}$.
- Sampling distribution of $\bar{x}_1 - \bar{x}_2$.

Example:- Given $n = 100$, $\sigma = 5000$

$N = 310$, $\mu = 24{,}000$.

we have to find standard deviation

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} \Rightarrow \frac{5000}{\sqrt{100}}\sqrt{\frac{310-100}{310-1}}$$

$\Rightarrow 412.20$ Rupees.

When fcp is not required $\Rightarrow Z = \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ -

When fcp is required, $Z = \dfrac{X-\mu}{\sigma/\sqrt{n}\sqrt{\frac{N-n}{N-1}}}$

Normal distribution $N(24000, 412.20)$ -

therefore,

$$Z = \frac{\bar{X}-\mu_{\bar{x}}}{\sigma_{\bar{x}}}$$

$$Z = \frac{\bar{X}-24000}{412.20}$$

as $p(\bar{x} > 24,500)$ So

$$Z = \frac{24500-24000}{412.20}$$

$Z = 1.21$

Hence, $p(\bar{x} > 24,500)$

$p(Z > 1.21) \Rightarrow 0.5 - p(0 < Z < 1.21)$

$\Rightarrow 0.5 - 0.3869 \Rightarrow 0.1131$ -

Chances are that 89% that, average salary not exeed 24,500. because the value comes out is 0.1131 which is 11%.

Sampling Distribution of Sample proportion:-

The proportion of success in population is $P = \dfrac{X}{N}$.

The proportion of success in sample is $\hat{p} = \frac{X}{n}$ -

Here X is a binomial random variable -

Formula given in example:-

$\mu_{\hat{p}} = p.$

$Var(\hat{p}) = \frac{pq}{n} \cdot \frac{N-n}{N-1}$

See the table and Now,

$\mu_{\hat{p}} = \Sigma \hat{p} f(\hat{p}) = \frac{10}{20} = 0.5.$

$\sigma^2_{\hat{p}} = \Sigma \hat{p}^2 f(\hat{p}) - [\Sigma \hat{p} f(\hat{p})]^2$ -

$\Rightarrow \frac{2}{60} - \left(\frac{10}{20}\right)^2 \Rightarrow \frac{1}{20} \Rightarrow 0.05$ -

we know that $p = \frac{X}{N} \Rightarrow p = \frac{3}{6} \Rightarrow 0.5$

Hence we say that $\mu_{\hat{p}} = p = 0.5$ -

Then we find:- $\frac{pq}{n} \cdot \frac{N-n}{N-1}$

$\Rightarrow \frac{0.25}{3} \cdot \frac{6-3}{6-1} \Rightarrow \frac{0.25}{5} \Rightarrow 0.05 = Var(\hat{p})$ -

Hence, both properties of sampling distribution are verefied -

Properties of Sampling distribution

1) The mean of sampling distribution denoted by $\mu_{\hat{p}}$ is equal to Population proportion P as $\mu_{\hat{p}} = P$.

2) The standard deviation of Sampling distribution called the standard error and denoted by $\sigma_{\hat{p}}$.

- $$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}},$$
- When sampling is done with replacement.
- For without replacement $\sqrt{\frac{N-n}{N-1}}$ is used.

3) The sampling distribution of $\hat{p}$ is binomial distribution.

This will be approximatly normal when ever both np and nq are equal or greater than 5.

Sampling Distribution of differences Between mean.

If μ_1 and μ_2 and σ_1^2 and σ_2^2 and n_1 and n_2 are sample size and difference is $\bar{X}_1 - \bar{X}_2$ is computed, Such as distribution is called distribution of differences of Sample mean.

Note:-
we say that $\bar{x}_1 - \bar{x}_2$ is d-
Formula:-

$$\mu_{\bar{x}_1 - \bar{x}_2} = \sum (\bar{x}_1 - \bar{x}_2) f(\bar{x}_1 - \bar{x}_2)$$

$$\text{or} \sum d f(d) -$$

$$\sigma^2_{\bar{x}_1 - \bar{x}_2} = \sum d^2 f(d) - \left(\sum d f(d)\right)^2 -$$

properties of Sampling Distribution of $\bar{X}_1 - \bar{X}_2$ -.

1) $\mu_{\bar{x}_1 - \bar{x}_2} = \mu_1 - \mu_2$.
It means the it is equal to difference b/w Population Mean-

2) In case of Sampling with or without replacment from two infinite Population the S.D is:-

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

3) If the POPULATIONS are normaly distributed then regardless of Sample Size, it will be normal with mean and variance-

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

Lecture 33:- Sampling Distribution.

Example:- car batteries-

Given:- Population A

$\mu_1 = 4.3$ years, $\sigma_1 = 0.6$ years, $n_1 = 49$ -

Population B = $\mu_2 = 4.0$ years, $\sigma_2 = 0.4$ years, $n_2 = 36$ -

Mean:-

$\mu_{\bar{X}_1 - \bar{X}_2} = \mu_1 - \mu_2 = 4.3 - 4.0 \Rightarrow 0.3$ years -

Standard deviation:-

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \Rightarrow \sqrt{\frac{0.36}{49} + \frac{0.16}{36}}$$

$\Rightarrow 0.1086$ years -

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

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (0.3)}{0.1086}$$

we are required to find:-

$P(\bar{X}_1 - \bar{X}_2 \geq 0.5)$ -

$\bar{X}_1 - \bar{X}_2 \Rightarrow 0.5$

$$Z = \frac{0.5 - 0.3}{0.1086} = 1.84$$

$P(\bar{X}_1 - \bar{X}_2 \geq 0.5) = P(Z \geq 1.84)$

$P(\bar{x}_1 - \bar{x}_2 \geq 0.5) = P(Z \geq 1.84)$

$0.5 - P(0 < Z < 1.84)$

$0.5 - 0.4671 \Rightarrow 0.0329$ -

Sampling Distribution Between proportions:-

There are Two binomial Populations with proportions of Successes P_1 and P_2 respectively, Let independent random sample of Sizes n_1 and n_2 be drawn from the respective Populations, and the differences $\hat{P}_1 - \hat{P}_2$ between the proportions of all possible pairs of sample be computed.

In order to convert:, $\hat{P}_A - \hat{P}_B$ to Z, we need the value of $\mu_{\hat{P}_A - \hat{P}_B}$ as well as $\sigma_{\hat{P}_A - \hat{P}_B}$ -

Properties of $\hat{P}_1 - \hat{P}_2$ -

1) $\mu_{\hat{P}_1 - \hat{P}_2} = P_1 - P_2$ -

2) $\sigma_{\hat{P}_1 - \hat{P}_2} = \sqrt{\dfrac{P_1 q_1}{n_1} + \dfrac{P_2 q_2}{n_2}}$,

$(q = 1 - P)$ -

3) when we are fully confident to Say that our Population is homogenous then we use random Sampling -

* The Standard deviation of $\bar{x}$ gives us "standard" value of the error, this is Called Standard **error** -

Point Estimation:-

Point estimation of a Population Parameter Provides as an estimate a Single value calculated from the Sample that is likely to be close in magnitude to unknow Parameter-

Sometimes it is denoted by θ and Sometimes by T-

An estimator is a random variable and has probabilty distribution, which is known as, Sampling distribution-

Qualities of Good Point estimator:-

1) Unbiasedness-
2) Consistency-
3) Efficiency-

Unbiasdness:-

If the expected value is equal to the true value of Population parameter being estimated-

$E(\hat{\theta}) = \theta$ (Unbiased estimator)

$E\,\hat{\theta} \neq \theta$ (Biased estimator)-

- $E(\bar{X}) = \mu$ implies that the distribution of $\bar{X}$ is centered at μ –
- $E(\hat{p}) = E\left(\frac{X}{n}\right) = \frac{1}{n} E(X)$.

  $\Rightarrow \frac{nP}{n} = P$.
- $E(s^2) \neq \sigma^2$ –

Hence we mean that mean and sample proportion P is unbiased estimator, while the sample variance σ^2 is a biased estimator – But later on the

Modified formula

$$s^2 = \frac{\sum (x - \bar{x})^2}{n - 1}$$

Proves that s^2 is an unbiased estimator –

Consistency:-

An estimator $\hat{\theta}$ is said to consistent of parameter θ if, for any arbitrarily small positive quantity e –

$$\lim_{n \to \infty} P\left[|\hat{\theta} - \theta| \leq e \right] = 1 -$$

Sample mean $= \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ (is consistent).

Sample variance $= s^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$ –

(is consistent) –

An estimator is said to be Consistent if the Probability that estimator is very close to the parameter if this probability increases and tends to 1, as the sample size increases.

Lecture 34:-

Efficiency:-

An unbiased estimator is defined to be **efficient** if the variance of its sampling distribution is smaller than that of the sampling distribution of any other unbiased estimator of same parameter.

- when the variance is less, the sample is more efficient.

Methods of Estimation:-

MLE:- Maximum Likelihood estimation is a Method that determines the values for parameters of the model. It is the statistical Method of estimating the parameters of the probability distribution by maximizing the likelihood function.

Example:-

① For the geometric distribution given by $P(X=x) = pq^{x-1}$, $x = 1, 2, 3 \ldots$ the MLE of p is $\frac{1}{\bar{X}}$

Hence,

the **MLE** of p is equal to the reciprocal of **sample mean**.

② For the Bernoulli distribution given by:-

$P(X=x) = p^{x} q^{1-x}$, $x = 0, 1$.

The **MLE** of p is $\bar{X}$ equal to the mean of sample.

③ For the exponential distribution given by:-

$f(x) = \theta e^{-\theta x}$, $x > 0$, $\theta > 0$,

the MLE of θ is $\frac{1}{\bar{X}}$.

(Reciprocal of sample mean)

Method of Least square:-

The method of least square is a standard approach in regression analysis to approximate the solution of overdetermined

Systems by minimizing the Sum of the squares of residual made in result of every singal equation —

Method of mommments:-

It starts by expressing the Population moments (The expected value of Powers of random variable under Consideration) as function of parameters of interest. Those expressions are equal to sample moments —

Confidence interval:-

A confidence interval is an interval Computed from the Sample observations $X_1, X_2 \cdots X_n$ with a statment of How confident we are that the interval does contain the Population parameter —

For example:-

we take confidence interval of 95% then confidence interval for μ is,

$$\left(\overline{X} - 1.96\frac{\sigma}{\sqrt{n}}, \overline{X} + 1.96\frac{\sigma}{\sqrt{n}}\right) -$$

(when σ is known) —

Also written as: $\left(\overline{X} \pm 1.96\frac{\sigma}{\sqrt{n}}\right)$ —

In other case when σ is unknown then we use this formula for the 95% confidence interval for μ is:-

$$\left(\bar{X} - 1.96\frac{S}{\sqrt{n}} < \mu < \bar{X} + 1.96\frac{S}{\sqrt{n}}\right)$$

Lecture 35:-

Confidence interval for μ, The mean of infinite population-

Formula = $\bar{x} \pm Z_{\alpha/2} \frac{s}{\sqrt{n}}$ -

To compute $\bar{x} = \frac{\sum x}{n}$

and $s = \sqrt{\frac{\sum (x-\bar{x})^2}{n-1}}$ -

Example:- The punjab highway department is studying the traffic pattern on G.T Road. The random sample of 64 days gives $x = 5410$ and $s = 680$. Find 90% C.I estimate for μ, the average number of vhicles per day-

Given:- $\bar{x} = 5410$, $s = 880$ and $Z = 1.645$ -

put in formula:-

$$\bar{x} \pm Z_{\alpha/2} \frac{s}{\sqrt{n}}$$

$$5410 \pm (1.645)\left(\frac{680}{\sqrt{64}}\right) \Rightarrow 5410 \underset{(-)}{\pm} (1.645)(85)$$

$$5410 \pm 139.8 \Rightarrow 52\overset{(+)}{7}0 \text{ and } 5549.8.$$

(we have 5270 and 5550)

Somewhere between 5270 and 5550 vhichles that pass bridge and this is based on 90% confident-

Confidence interval for difference b/w means of Two Population -

$$(\bar{X}_1 - \bar{X}_2) \pm Z_{\alpha/2} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

Lecture # 36:-

Confidence interval for a population Proportion p

$$\hat{p} \pm Z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

or $$\hat{p} \pm Z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

Confidence Interval for $P_1 - P_2$

$$\hat{P}_1 - \hat{P}_2 \pm Z_{\alpha/2} \sqrt{\frac{\hat{P}_1 \hat{q}_1}{n_1} + \frac{\hat{P}_2 \hat{q}_2}{n_2}}$$

Hypothesis Testing.

Hypothesis Testing is an act in statistics whereby an analyst tests an assumption regarding a population Parameter. Hypothesis testing is used to assess the plausiblity of hypothesis by using sample Data.

اس میں ہم دو طریقوں سے گزریں گے پہلے ہم
interval نکالتے ہیں اس کے لیے فارمولا ہوتا ہے
پھر ہم Test کرتے ہیں اس کے لیے دوسرا
فارمولا لگاتے ہیں ----

Steps in finding Interval:-

1) Finding the number of Samples (n).
2) Calculate the mean ($\bar{x}$) of the Samples.
3) Calculate the standard deviation (s) of Samples.
4) Decide the confidence interval that will be used.
   95 and 99 percent Confidence interval are the most common choices.
5) Find the z-value for the Selected Confidence interval.

As:-

Confidence Interval	Z
80%	1.282
85%	1.440

Confidence interval	Z-value
90%	1.645
95%	1.960
99%	2.576
99.5%	2.807
99.9%	3.291

6) calculate the formula

$$\bar{X} \text{ (any parameter)} \pm Z_{\alpha/2} \frac{s}{\sqrt{n}}$$

To (use this formula regarding Parameter) —

7) Draw a Conclusion —

NoTe :—

ہمیں یہ کیسے پتا چلے گا کہ ہم Z-Value دیکھیں یا T-value ؟ ....

یاد رکھیں کہ جب σ یعنی standard devia ہمیں معلوم ہو گا مطلب (σ is Known) تو ہم Z-Table لگائیں گے ....

لیکن جب (σ is unknown) اور S دیا گیا ہو تو ہم T-Table لگائیں گے ....

How to find critical value :—

$$\text{critical value} = 1 - \frac{\alpha}{2}.$$

$$\Rightarrow 1 - \frac{\text{Confidence interval}}{100} = \alpha$$

(Suppose we have 99% confidence interval)

$$1 - \frac{90}{100} = \alpha \Rightarrow 0.1 = \alpha$$

Steps in Hypothesis Testing:-

1) State Null and alternate Hypothesis.
2) Select a level of significance.
3) Identify the test-statistic.
4) Formulate a decision rule.
5) Do not reject H_0 or reject H_0 and accept H_1 (make decision)

In other words:-

1) Set up H_0 and H_1.
where H_0 is Null Hypothesis and H_1 is **alternate** Hypothesis.
If we take H_0 $(=)$ then H_1 $(\neq)$
If we take H_0 $(\geq)$ then H_1 $(<)$
If we take H_0 $(\leq)$ then H_1 $(>)$
2) Determine the critical numbers.
جب n 30 سے بڑا یا 30 کے برابر ہو گا اور
σ known ہو یا σ unknown ہو ہم Z-Table
لگائیں گے ... لیکن اگر n 30
سے چھوٹا ہے اور (σ unknown)
ہو تو ہم T-Table لگائیں گے اور اگر
(σ known) ہو تو ہم وہی Z-Table ہی
استعمال کریں گے ۔
3) Draw a curve and plot number.
(لیکن ہم ٹوٹ نہیں کریں گے ...)

4) Determine the Test-statistic -

5) Reject of do not reject the H_0.

- Reject H_0 if $\alpha \geq$ P-value.
- Donot Reject H_0 if $\alpha \leq$ P-value -

6) Write the Final Conclusion —

Tail type -

H_0	H_1	Type
$\geq$	$<$	left 1-tail
$\leq$	$>$	right 1-tail
$=$	$\neq$	both 2-tail

Type 1 error and Type 2 error -

If we accept H_1 we say it is Type 1 error &

when we accept H_0 we say it is Type 2 error -

Type I error:-

Type I error صحیح ہوتا ہے لیکن پھر بھی Reject کر دیا جاتا ہے اس لیے اس کو False Positive بھی کہتے ہیں ... اسکی مثال یہ ہے کہ ایک معصوم بندے کو مجرم بنا کر جیل ڈال دینا ....

Type II error :-

یہ غلط ہوتا ہے لیکن پھر بھی Accept کر لیا جاتا ہے اس کو False Negative کہا جاتا ہے۔ اسکی مثال یہ ہے کہ ایک بندہ مجرم ہوتے ہوئے بھی جیل میں نہیں ڈالا گیا

Hypothesis Test Statistics and Confidence Intervals

$1-\alpha$ Confidence Interval **Point Estimate $\pm$ Maximum Error E**	**Hypothesis Test Value (Statistic)** **NULL Hypothesis:** Use the statement containing the condition of equality either directly or implied, as the Null Hypothesis H_o.

Single Population

(TI-84) (TI-84)

One Sample for mean μ (σ is known)

(ZInterval)		(Z-Test)
$\bar{x} \pm z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$	Use the *Normal* z-Table for the critical value z	$z = \dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}}$

One Sample for mean μ (σ is unknown)

(TInterval)		(T-Test)
$\bar{x} \pm t_{\alpha/2}\dfrac{s}{\sqrt{n}}$	$df = n - 1$ Use the t-distribution Table for the critical value t	$t = \dfrac{\bar{x}-\mu}{s/\sqrt{n}}$

One Sample for Proportion p

(1-PropZInt)		(1-PropZTest)
$\hat{p} \pm z_{\alpha/2}\sqrt{\dfrac{\hat{p}\hat{q}}{n}}$	Use the *Normal* z-Table for the critical value z	$z = \dfrac{\hat{p}-p}{\sqrt{pq/n}}$

Dual Population

(TI-84) (TI-84)

Dependent Paired for μ_d

(TInterval)		(T-Test)
$\bar{d} \pm t_{\alpha/2}\dfrac{s_d}{\sqrt{n}}$	$df = n - 1$ Use the t-distribution Table for the critical value t	$t = \dfrac{\bar{d}-\mu_d}{s_d/\sqrt{n}}$ Use $H_o: \mu_d = 0$

Two Independent Samples for $\mu_1 - \mu_2$ (σ_1, σ_2 are known)

(2-SampZInt)		(2-SampZTest)
$(\bar{x}_1-\bar{x}_2) \pm z_{\alpha/2}\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}$	Use the *Normal* z-Table for the critical value z	$z = \dfrac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$ Use $H_o: \mu_1-\mu_2=0$

Two Independent Samples for $\mu_1 - \mu_2$ (σ_1, σ_2 are unknown)

(2-SampTInt)		(2-SampTTest)
$(\bar{x}_1-\bar{x}_2) \pm t_{\alpha/2}\sqrt{\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}}$	*df = smaller of* n_1-1 *or* n_2-1 Use the t-distribution Table for the critical value t **Use "NOT POOLED" on the calculator.**	$t = \dfrac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}}}$ Use $H_o: \mu_1-\mu_2=0$

Two Independent Samples for Proportions $p_1 - p_2$

(2-PropZInt)		(2-PropZTest)
$(\hat{p}_1-\hat{p}_2) \pm z_{\alpha/2}\sqrt{\dfrac{\hat{p}_1\hat{q}_1}{n_1}+\dfrac{\hat{p}_2\hat{q}_2}{n_2}}$	Use the *Normal* z-Table for the critical value z	$z = \dfrac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\dfrac{\bar{p}\bar{q}}{n_1}+\dfrac{\bar{p}\bar{q}}{n_2}}}$ Use $H_o: p_1-p_2=0$

where, $\bar{p} = \dfrac{x_1+x_2}{n_1+n_2}$ or $\bar{p} = \dfrac{n_1\hat{p}_1+n_2\hat{p}_2}{n_1+n_2}$ $\bar{q} = 1-\bar{p}$

1-Prop: $p = \dfrac{X}{N}$ $\hat{p} = \dfrac{x}{n}$ $q = 1-p$ $\hat{q} = 1-\hat{p}$

Dual Prop: $p_1 = \dfrac{X_1}{N_1}$ $p_2 = \dfrac{X_2}{N_2}$ $\hat{p}_1 = \dfrac{x_1}{n_1}$ $\hat{p}_2 = \dfrac{x_2}{n_2}$ $\bar{p} = \dfrac{x_1+x_2}{n_1+n_2}$ or $\bar{p} = \dfrac{n_1\hat{p}_1+n_2\hat{p}_2}{n_1+n_2}$

$q_1 = 1-p_1$ $q_2 = 1-p_2$ $\hat{q}_1 = 1-\hat{p}_1$ $\hat{q}_2 = 1-\hat{p}_2$ $\bar{q} = 1-\bar{p}$

Sample Size Determination

for Mean μ		for Proportion p
$n = \dfrac{z_{\alpha/2}^2\sigma^2}{E^2} = \left(\dfrac{z_{\alpha/2}\sigma}{E}\right)^2$	(round up)	$n = \dfrac{z_{\alpha/2}^2 pq}{E^2}$ or use $n = \dfrac{z_{\alpha/2}^2(.25)}{E^2}$ (if p, q unknown)

Uses of the Chi-Square Test

- One of the most useful properties of the chi-square test is that it tests the null hypothesis "the row and column variables are not related to each other" whenever this hypothesis makes sense for a two-way variable.

Uses of the Chi-Square Test

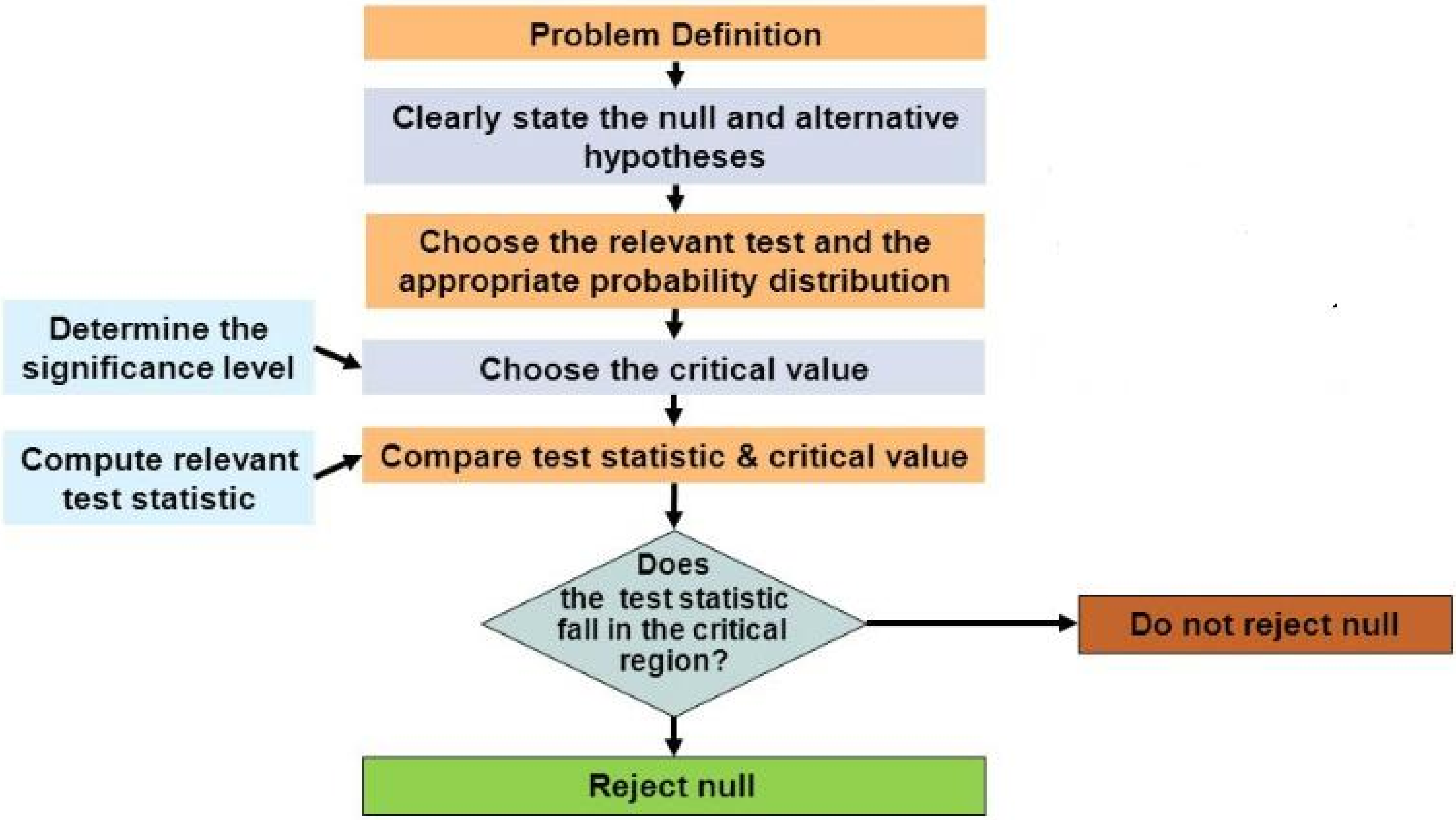
Use the chi-square test to test the null hypothesis

H_0: there is no relationship between two categorical variables

when there is a two-way table from one of these situations:

- Independent random samples from two or more populations, with each individual classified according to one categorical variable.
- A single random sample, with each individual classified according to both of two categorical variables.

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="2">Decision</td></tr>
<tr><td>Accept H_0</td><td>Reject H_0 (or accept H_1)</td></tr>
<tr><td rowspan="2">True Situation</td><td>H_0 is true</td><td>Correct decision (No error)</td><td>Wrong decision (Type-I error)</td></tr>
<tr><td>H_0 is false</td><td>Wrong decision (Type-II error)</td><td>Correct decision (No error)</td></tr>
</table>


Problem Definition
Clearly state the null and alternative hypotheses
Choose the relevant test and the appropriate probability distribution
Determine the significance level
Choose the critical value
Compute relevant test statistic
Compare test statistic & critical value
Does the test statistic fall in the critical region?
Do not reject null
Reject null

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Testing کے Basic آپ کو بتا دیتے ہیں

یہ Practice سے آئے گی آپ کو so آپ

یہ خود کریں ...... Best of luck.

She believed, she could,
So she did ....

Written by:— Mysterious —