Chapter-3

Random variables and probability

Random variable:- If the numerical values assumed by a variable are the result of some chance factors so that a particular value cannot be exactly predicted in advance. The variable is then called random variable and is also called stochastic variable.

a) **Continuous Random variable:-**

It is variable which can assume any value within an interval.

b) **Discrete Random variable:-**

It is variable which can assume only isolated values.

ex: No. of heads in 4 tosses of a coin.

Random experience:-

The occurance which can be repeated a number of times essentially under the same conditions and whose results cannot be predicted before hand.

Discrete probability distribution:-

Let a random variable x assumes values $x_1, x_2, \ldots x_n$ with probabilities $p_1, p_2, \ldots p_n$ respectively.

$$p(x = x_i) = p_i \geq 0 \quad \text{and}$$

$$p_1 + p_2 + \cdots + p_n = \sum_{i=1}^{n} p_i = 1$$

then

X	x_1	x_2	x_3	--	x_n
$p(x)$	p_1	p_2	p_3	--	p_n

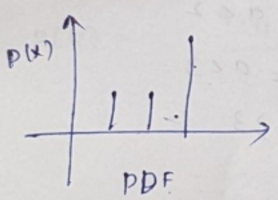
is called discrete probability distribution for x.

The function p(x) is called probability density function (PDF).

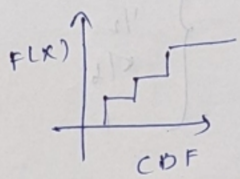
The value of p(x) for $x = t$ i.e $p(x = t)$

probability of x taking the value $\le t$ denoted by $p(x \le t)$ is defined by $p(x \le t) = p_1 + p_2 + \dots + p(t)$.

The function $p(x \le t)$ is called cumulative distribution function (C.D.F)



PDF



CDF

$$\boxed{\text{Mean} = \frac{\Sigma p_i x_i}{\Sigma p_i} = \Sigma p_i x_i}$$

$[\Sigma p_i = 1]$

$$\boxed{\text{variance} = \Sigma p_i x_i^2 - (\text{Mean})^2}$$

Ex: Consider a random variable x that is equal to 1, 2, 3 and given $p(1) = \frac{1}{2}$, $p(2) = \frac{1}{3}$ then find mean, variance of distribution and also plot PDF and CDF.

x	1	2	3
p(x)	1/2	1/3	1/6

$\text{mean} = \frac{1}{2} + \frac{2}{3} + \frac{1}{2}$

$= 1.66.$

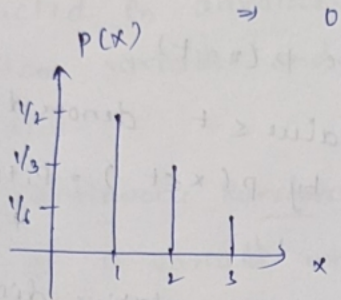
Variance $= \Sigma p_i x_i^2 - (\text{mean})^2$

$= \frac{1}{2}(1) + \frac{1}{3}(2)^2 + \frac{1}{6}(3)^2 - (1.66)^2$

$= \frac{1}{2} + \frac{4}{3} + \frac{9}{6} = 2.56$

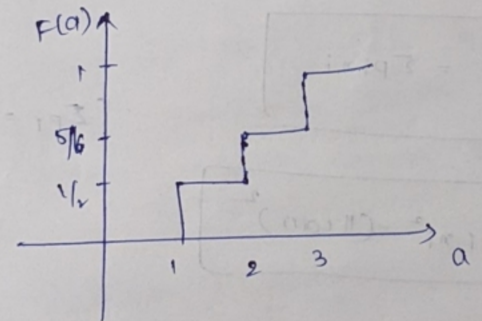
$\Rightarrow 0.5 + 1.33 + 1.5 = 2.56$

$\Rightarrow 0.77$



PDF

$$F(a) = \begin{cases} 0 & a < 1 \\ 1/2 & 1 \le a < 2 \\ 5/6 & 2 \le a < 3 \\ 1 & a \ge 3 \end{cases}$$



Ex: 2

x = No. of defective up-chips

No. of defective upchips = 5

No. of good upchips = 20

Total no. of upchips = 25

Four up Chips are drawn,

x	0	1	2	3	4
p(x)	0.38	0.45	0.15	0.015	4×10^{-4}

$p(0) = p(\text{No defective}) = \dfrac{{}^{20}C_4 \times {}^{5}C_0}{{}^{25}C_4} = \dfrac{4845}{12650} = 0.38$

$p(1) = p(\text{one defective}) = \dfrac{{}^{20}C_3 \times {}^{5}C_1}{{}^{25}C_4} = \dfrac{1140\times5}{12650} = 0.45$

$p(2) = \dfrac{{}^{5}C_2 \times {}^{20}C_2}{{}^{25}C_4} = \dfrac{10\times190}{12650} = 0.15$

$p(3) = \dfrac{{}^{5}C_3 \times {}^{20}C_1}{{}^{25}C_4} = \dfrac{200}{12650} = 0.015$

$p(4) = \dfrac{{}^{5}C_4 \times {}^{20}C_0}{{}^{25}C_4} = \dfrac{5}{12650} = 4\times10^{-4}$

$$F_a = \begin{cases} 0.38 & a \le 0 \\ 0.83 & 0 < a \le 1 \\ 0.98 & 1 < a \le 2 \\ 0.995 & 2 < a \le 3 \\ 0.9954 & 3 < a \le 4 \end{cases}$$

Binomial distribution:-

$p(X = r) = N\, {}^{n}C_r\, p^r q^{n-r}$

$p(X = r) = \dfrac{N e^{-\lambda} \lambda^r}{r!}$

mean = np,

variance = npq.

r is number of success.

$\lambda = np = \text{mean}$

LP:-

(6) p = probability that bomb strike the target $= 0.5$

i-e $p = q \Rightarrow q = 0.5$ $(\because p+q=1)$

when p - large, use binomial

p - small $\Rightarrow$ n - large, use poisson.

$$p(r \geq 2) \geq 0.99 \quad \text{(given)}$$

$$1 - p(r < 2) \geq 0.99$$

$$1 - \{p(r=0) + p(r=1)\} \geq 0.99$$

$$1 - \{ {}^nc_0 (0.5)^0 (0.5)^{n-0} + {}^nc_1 (0.5)^1 (0.5)^{n-1} \} \geq 0.99$$

$$\Rightarrow 1 - \{(0.5)^n + n(0.5)(0.5)^{n-1}\} \geq 0.99$$

$$1 - 0.99 \geq \left(\tfrac{1}{2}\right)^n + n\left(\tfrac{1}{2}\right)^n$$

$$0.01 \geq \frac{1}{2^n}(1+n)$$

$$2^n \geq \frac{1+n}{0.01}$$

$$2^n \geq 100 + 100n$$

By inspection method,

$$2^n \geq 100 + 100n$$

when $n = 11$, equality holds good.

$$2048 \geq 1200$$

minimum number of bombs to destroy completely is 11.

LP

(9):-

$p = 0.001$

As p is small, use poisson distribution

$n = ?$

$p(r \geq 1) = 0.9$ (given)

$1 - p(r < 1) = 0.9$

$1 - p(r = 0) = 0.9$

$1 - \left[\frac{e^{-\lambda} \lambda^0}{0!}\right] = 0.9$

$1 - e^{-\lambda} = 0.9$

$e^{-\lambda} = 1 - 0.9$

$e^{-\lambda} = 0.1$

$e^{-np} = 0.1$

$e^{-0.001n} = 0.1$

$-0.001n = \ln(0.1)$

$n = \frac{-2.3025}{-0.001}$

$\therefore \boxed{n = 2302.5}$

(8) 0.1 is not a small probability.

$p = 0.1$ $\downarrow$ big, $n = 20$ $\downarrow$ small, use B.D

a) $p(r = 0) = {}^{n}C_r \, p^r q^{n-r}$

$= {}^{20}C_0 (0.1)^0 (0.9)^{20} = 0.1215$

b) p(at least one silent error) $= p(r \geq 1)$

$= 1 - p(r = 0)$

$= 1 - 0.122$

$= 0.878$

c) p(more than 4 strictly errors) $= p(r > 4)$

$= 1 - p(r \leq 4)$

$= 1 - \{p(r=0) + p(r=1) + p(r=2) + p(r=3) + p(r=4)\}$

$\Rightarrow 1 - \{0.122 + 0.270 + 0.285 + 0.190 + 0.089\}$

$p(r=2) = {}^nC_r\, p^r q^{n-r} = {}^{20}C_2\,(0.1)^2\,(0.9)^{18}$

$= 190 \times 0.01 \times 0.15$

$p(r=3) = {}^{20}C_3\,(0.1)^3\,(0.9)^{17} = 1140 \times 0.001 \times 0.1667$

$p(r=4) = {}^{20}C_4\,(0.1)^4\,(0.9)^{16} = 4845 \times 0.0001 \times 0.1853$

$\therefore\ p = 1 - 0.956$

$\Rightarrow p = 0.044 = 4\%$

As probablity is small, occurance is rare.

(7)

$n = 8,\quad N = 256,\quad p = 0.5,\quad \therefore q = 0.5$

No. of heads :	0	1	2	3	4	5	6	7	8
freq :	2	6	24	63	64	50	36	10	1

mean = np.

(p not given) $\text{mean} = \dfrac{\Sigma p_i x_i}{\Sigma p_i} = \dfrac{\Sigma f_i x_i}{\Sigma f_i}$

p(given), $\text{mean} = np = (8)(0.5) = 4$

$p(r=0) = N \times {}^nC_0\, p^0 q^{n-0}$

$= 256 \times 1 \times 1 \times q^8$

$= 256 \times (0.5)^8 = 1$

$p(r=1) = 256 \times {}^8C_1 \times (0.5)^1 (0.5)^7$

$= 8$

$p(x=2) = 256 \times {}^8C_2 \times (0.5)^2 (0.5)^6$

$= 28$

$p(x=3) = 256 \times {}^8C_3 (0.5)^3 (0.5)^5$

$= 56$ //

Only mean is available

mean $= \lambda$

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

(i) p(two will arrive) $= p(r=2)$

$$= \frac{e^{-\lambda}\lambda^2}{2!} \quad \text{where } \lambda = 1.5$$

$$= \frac{e^{-1.5}(1.5)^2}{2!} = \boxed{0.251}$$

(ii) p(at least three will arrive during an interval of two minutes) $= p(r \geq 3)$ where $\lambda = 1.5 \times 2$

$\boxed{\lambda = 3.0}$

$$= 1 - p(r < 3)$$

$$= 1 - \{p(r=0) + p(r=1) + p(r=2)\}$$

$$= \cancel{1-\{\frac{e^{-0.3}(0.3)^0}{0!} + \frac{e^{-0.3}(0.3)^1}{1!} + e^{-0.3}}$$

$$\Rightarrow 1 - \left\{\frac{e^{-3}(3)^0}{0!} + \frac{e^{-3}(3)^1}{1!} + \frac{e^{-3}(3)^2}{2!}\right\}$$

$$= 1 - e^{-3}\left(1 + 3 + \frac{9}{2}\right)$$

$$= 0.5768$$

(iii) p(at most 8 will arrive during an interval of six minutes) $\Rightarrow p(r \leq 8)$

$\lambda = 1.5 \times 6$

$\boxed{\lambda = 9}$

$$\Rightarrow p(r=0) + p(r=1) + p(r=2) + p(r=3) + p(r=4) + p(r=5) + p(r=6) + p(r=7) + p(r=8)$$

Ex:Q: Fit a poisson distribution,

x:	0	1	2	3	4
f:	122	60	15	2	1

Soln: Mean $= \frac{\Sigma f_i x_i}{\Sigma f_i}$

$= \frac{0(122)+1(60)+2(15)+3(2)+4(1)}{122+60+15+2+1}$

$\boxed{N = 200}$

$= \frac{100}{200}$

$\boxed{\text{mean}, \lambda = 0.5}$

$p(r=0) \Rightarrow N\left(\frac{e^{-\lambda}\lambda^r}{r!}\right) = \left(\frac{e^{-0.5}(0.5)^0}{0!}\right)N = \left(\frac{e^{-0.5}(1)}{1}\right)200 \Rightarrow 0.606 \times 200$

$= 120.$

$p(r=1) \Rightarrow \frac{e^{-\lambda}\lambda^r}{r!} \Rightarrow \left(\frac{e^{-0.5}(0.5)^1}{1!}\right) \times 200.$

$= 60.65 \Rightarrow 61$

$p(r=2) = \left(\frac{e^{-\lambda}\lambda^r}{r!}\right)N = \frac{e^{-0.5}(0.5)^2}{2} \times 200$

$= 15.16.$

$p(r=3) = \left(\frac{e^{-\lambda}(\lambda)^r}{r!}\right)N = \frac{e^{-0.5}(0.5)^3}{3!} \times 200$

$= 2.527.$

$p(r=4) = \frac{e^{-0.5}(0.5)^4}{4!} \times 200.$

$= 0.315$

$p(r=0) + p(r=1) + p(r=2) + p(r=3) + p(r=4) = 200.$

* Continuous probability distribution:-

Probability function $f(x)$ Satisfying the conditions

(1) $f(x) \geq 0$

(2) $\int_{-\infty}^{\infty} f(x)\,dx = 1$ (geometrically means that the area bounded by the curve and x-axis is equal to unity).

$f(x)$ is called a continuous probability function (or) probability density function.

If 'X' is a continuous random variable with probability density function $f(x)$ then the function $f(x)$ is defined by,

$$\boxed{F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)\,dx.}$$

is called cummulative distribution function. (cdf).

~~Mean = Explanation~~

Mean = Expectation of x $E(x) = \int_{-\infty}^{\infty} x\, f(x)\,dx.$

$$\text{Variance} = E(x^2) - [E(x)]^2$$

where, $E(x^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx.$

Q. A probability density function is given by,

$$f(x) = \begin{cases} 0 & \text{for } x < 1 \\ b/x^2 & \text{for } 1 < x < 5 \\ 0 & \text{for } x > 5 \end{cases}$$

a) what is the value of b?

b) obtain the probability that x is in between 2 & 4.

c) what is the probability that x is exactly '2'.

d) Find CDF of x?

a) By using the definition of probability density function

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

$$\int_{-\infty}^{1} \cancel{f(x)\,dx} + \int_{1}^{5} f(x)\,dx + \int_{5}^{\infty} \cancel{f(x)\,d(x)} = 1$$

$$\int_{1}^{5} b/x^2\,dx = 1$$

$$\Rightarrow b \int_{1}^{5} \frac{1}{x^2}\,dx = 1$$

$$\Rightarrow b \int_{1}^{5} x^{-2} \cdot dx = 1 \quad \Rightarrow b \left[\frac{x^{-2+1}}{-2+1} \right]_{1}^{5} = 1$$

$$\Rightarrow b \left[\frac{x^{-1}}{-1} \right]_{1}^{5} = 1$$

$$\Rightarrow b\left(-\frac{1}{x}\right)_{1}^{5} = 1$$

$$\Rightarrow b\left(\frac{-1}{5} + \frac{1}{1}\right] = 1$$

$$\Rightarrow b\left[1 - \frac{1}{5}\right] = 1$$

$$b\left(\frac{4}{5}\right) = 1$$

$$\boxed{b = \frac{5}{4}} = 1.25.$$

b) $P(2 < x < 4)$

$$\Rightarrow \int_{2}^{4} f(x) \cdot dx = \int_{2}^{4} b/x^2 \cdot dx.$$

$$= \frac{5}{4}\left[\frac{-1}{x}\right]_{2}^{4}$$

$$= -\frac{5}{4}\left[\frac{1}{4} - \frac{1}{2}\right]$$

$$= -\frac{5}{4}\left[\frac{1-2}{4}\right] = \frac{-5}{4} \times \frac{-1}{4} \Rightarrow \frac{5}{16}$$

$$\Rightarrow 0.3125$$

c) $p(x=2) \Rightarrow \int_2^2 \frac{b}{x^2}\,dx = 0.$

Probability at exactly one point ~~is~~ 0.

Note :-

$\Rightarrow$ probability that any continuous random variable is exactly equal to single quantity is zero.

d) CDF :-

for $x_1 < 1$ $\quad f(x_1) = \int_{-\infty}^{x_1} 0\,dx = 0$

for $1 < x_1 < 5$ $\quad f(x_1) = 0 + \int_1^{x_1} \frac{(1.25)}{x^2}\,dx.$

$$= 1.25\left[1 - \frac{1}{x_1}\right].$$

for $x > 5$
$5 < x < \infty.$ $\quad = \int_{-\infty}^{x_1}$

$$= 0 + \int_1^5 \frac{1.25}{x^2}\,dx + \int_5^x 0\,dx.$$

$$= 1.25\left[1 - \frac{1}{5}\right]$$

$$= \frac{4}{5}(1.25)$$

$$= \frac{4}{5}\left(\frac{5}{4}\right) = 1$$

$\Rightarrow$ Cummulative distribution function (CDF) :-

CDF for,

$$f(x) = \begin{cases} 0 & \text{for } x_1 < 1 \\ 1.25\left(1 - \frac{1}{x_1}\right) & \text{for } 0 < x_1 < 5 \\ 1 & \text{for } x_1 > 5 \end{cases}$$

LP:-
(11) average = mean = $E(x) = \int_{-\infty}^{\infty} x\, f(x)\cdot dx$.

Soln:- $f(x) = ke^{-x}$; $x > 0$.

Since,

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

$\int_{0}^{\infty} ke^{-x}\cdot dx = 1$

$k\left[\frac{e^{-x}}{-1}\right]_0^{\infty} = 1$

$\Rightarrow -k\left[e^{-\infty} - e^{-0}\right] = 1 \Rightarrow \boxed{k=1}$

$= \int_{-\infty}^{0} 0 + \int_{0}^{\infty} x(ke^{-x})dx$

mean $= \int_{0}^{\infty} x \times 1 e^{-x}\cdot dx$

$\Rightarrow \left[\frac{x\cdot e^{-x}}{-1} - 1\left(\frac{e^{-x}}{-1(-1)}\right)\right]_0^{\infty}$

$\Rightarrow -\left[e^{-\infty} - e^{-0}\right]$

$\boxed{\text{mean} = 1}$

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

$= \int_{0}^{\infty} x^2\cdot e^{-x} dx - (1)^2$

$\Rightarrow \left[x^2\cdot\frac{e^{-x}}{-1} - \frac{2x\, e^{-x}}{(-1)(-1)} + 2\,\frac{e^{-x}}{-1}\right]_0^{\infty} - 1$

$\Rightarrow -2\left[e^{-\infty} - e^{-0}\right] - 1$

$= +2 - 1$

$= 1$

LP:- (12) prob(tube will last for first 150 hrs) = $P(X \le 150)$

$\Rightarrow P(0 < x < 100) + P(100 \le x \le 150)$

$\Rightarrow \int_{0}^{100} f(x)\cdot dx + \int_{100}^{150} f(x)\cdot dx$.

$$\Rightarrow \int_{100}^{150} \frac{100}{x^2}\, dx$$

$$-100\left[\frac{1}{x}\right]_{100}^{150}$$

$$\Rightarrow -100\left[\frac{1}{150}-\frac{1}{100}\right]$$

$$\Rightarrow \frac{1}{3}$$

b)

Hence,

P(None of those '3' tubes will have to be replaced during the first 150 hrs) $\Rightarrow \frac{1}{3}\times\frac{1}{3}\times\frac{1}{3}$

$$= \frac{1}{27}$$

a) P (all of '3' tubes will have to be replaced during the first 150 hrs) $\Rightarrow (1-\frac{1}{3})\times(1-\frac{1}{3})\times(1-\frac{1}{3})$

$$\Rightarrow \frac{2}{3}\times\frac{2}{3}\times\frac{2}{3}$$

$$\Rightarrow \frac{8}{27}$$

* Normal Distribution:-

Normal distribution is a probability distribution of continous Random variable x

Let μ and σ be two ordinary real constants such that $-\infty < \mu < \infty$ and $\sigma > 0$,

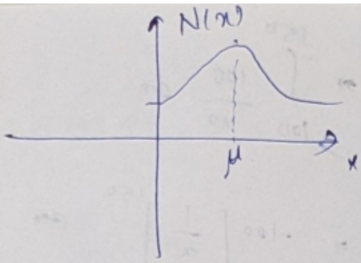
The continuous probability distribution,

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

μ – mean

σ – S.D

Properties:-



→ Symmetric about y-axis.

→ It is bell-shaped.

→ The mean, median and mode coincide and therefore normal curve is unimodal.

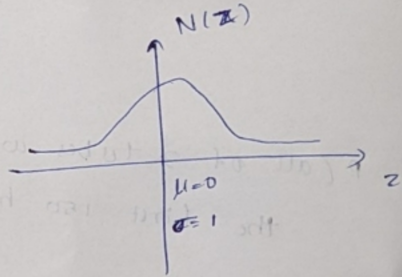
i-e it has only one maximum point

→ Area under the normal curve is unity.

* Standard form of N.D:-

variant of the standard Normal curve is denoted by 'z'.

→ The Random variable, $\boxed{z = \frac{x-\mu}{\sigma}}$



has the normal distribution

with mean = 0 and

standard deviation = 1

Probability Density function for the Normal distribution is

$$\boxed{f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}}$$

ex:- x is a normal variate with mean 30 and S.D = 5, find (1) $P(26 \le x \le 40)$

(2) $P(x \ge 45)$

Soln:-

$$z = \frac{x-\mu}{\sigma} \qquad \mu = 30, \ \sigma = 5$$

(i) $26 \le x \le 40$

when $x = 26$,

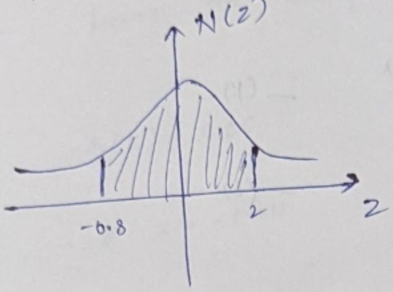
* $z = \frac{26-30}{5} = -0.8$

when $x = 40$,

$$z = \frac{40-30}{5} = 2$$

$$\boxed{z = 2}.$$

i) $P(26 \le x \le 40) = P(-0.8 \le z \le 2)$



$\Rightarrow A(2) - A(-0.8)$

$\Rightarrow 0.97725 - 0.21186$

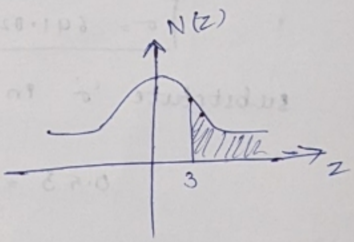
$\Rightarrow 0.76540$ // (76%).

(ii) $P(x \ge 45)$

$$z = \frac{45-30}{5} = \frac{15}{5} = 3$$

$$\boxed{z = 3}.$$

$P(x \ge 45) \Rightarrow P(z \ge 3)$



$P(z \ge 3) = 1 - A(3)$

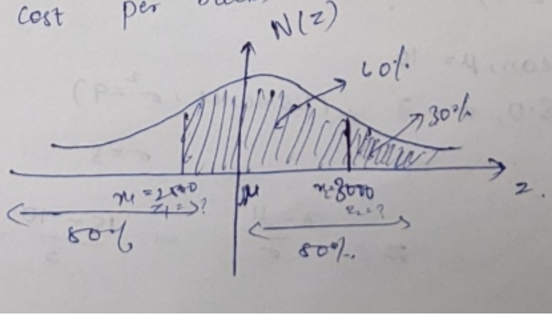
$= 1 - 0.9986$

$= 0.0014$ (Area).

L.P

(4) For a certain type of floorscent light in a large building, the cost per bulb,

Soln:



$A(z_1) = 1 - 0.6 = 0.4$

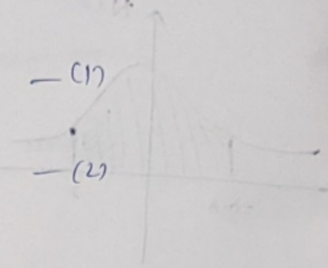
$z_1 = -0.25$ (from the table) (-ve table)

$A(z_2) = 1 - 0.3 = 0.7$

$z_2 = 0.53$ (from the table)

$z_1 = \frac{x_1 - \mu}{\sigma} \Rightarrow -0.25 = \frac{2500 - \mu}{\sigma}$ — (1)

$z_2 = \frac{x_2 - \mu}{\sigma} \Rightarrow 0.53 = \frac{3000 - \mu}{\sigma}$ — (2)



$$\begin{array}{rcl} -0.25\sigma &=& 2500 - \cancel{\mu} \\ 0.53\sigma &=& 3000 - \cancel{\mu} \\ (-) && (-) \quad (+) \\ \hline -0.78\sigma &=& -500 \end{array}$$

$\sigma = \frac{500}{0.78}$

$\boxed{\sigma = 641.025}$

substitute 'σ' in (2)

$0.53 = \frac{3000 - \mu}{641.025}$

$\mu = 3000 - 0.53(641.025)$

$= 3000 - 339.743$

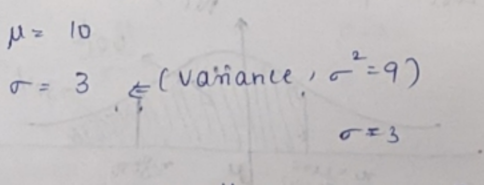
$\boxed{\mu = 2660.257}$

L.P (13) :-

i) $P(x \le 15) = ?$

mean, $\mu = 10$

S.D, $\sigma = 3$ ⇐ (variance, $\sigma^2 = 9$)



$\sigma = 3$

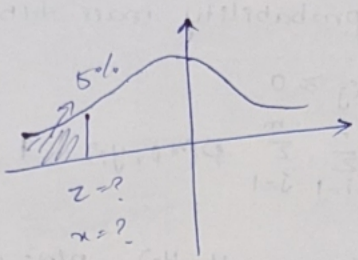
$z = \frac{x - \mu}{\sigma} \Rightarrow \frac{15 - 10}{3} = \frac{5}{3} = 1.66$

$\boxed{z = 1.66}$

$\Rightarrow P(z \le 1.66)$.

$A(z) = 0.9515$ (from table)

(ii) fastest 5% of repairs take at most how many hours to complete).



$A(z) = 0.05$

$z = -1.64$

$z = \frac{x-\mu}{\sigma}$

$-1.64 = \frac{x-10}{3}$

$x = 3(-1.64)+10$

$= -4.92+10$

$\boxed{x = 5.08 \text{ hrs}}$

L.P:-
15 :-

$\mu = 4{,}300$

$\sigma = 750$

$P(2500 < x < 4{,}200) = ?$

* Joint probability Distribution:-

This distance is associated with two random variable both of which are defined on the same sample space.

The joint probability function $P(x,y)$ is said to be joint probability mass function if all

i) $P_{ij} \geq 0$

(ii) $\sum_{i=1}^{n} \sum_{j=1}^{m} P(x_i, y_j) = 1$

$P_{ij} = P(X = x_i,\ Y = y_j) = P(x_i, y_j)$

⇒ It is usually represented in the table form as follows:-

x \ Y	y_1	y_2	---	y_m	
x_1	P_{11}	P_{12}	---	P_{1m}	P_1
x_2	P_{21}	P_{22}	---	P_{2m}	P_2
x_3					⋮
⋮					⋮
x_n	x_{n1}	x_{n2}	---	x_{nm}	P_n
	Q_1	Q_2	---	Q_m	

* Marginal distribution of X:-

X	x_1	x_2	---	x_n
$P(X=x_i)$	P_1	P_2	---	P_n

* Marginal distribution of Y:-

Y	y_1	y_2	---	y_m
$P(Y=y_j)$	Q_1	Q_2	---	Q_m

* Mean for discrete Random variable:-

$E(x) = \sum_{i=1}^{n} x\, p(x_i)$

$E(y) = \sum_{j=1}^{m} y\, p(y_j)$

$E(xy) = \sum_{i=1}^{n} \sum_{j=1}^{m} x_i y_j P_{ij}$

* Covariance:-

$\text{cov}(x,y) = E(xy) - E(x)\,E(y)$

Correlation Coefficient, $\boxed{\rho = \frac{\text{cov}(x\,y)}{\sigma_x \sigma_y}}$

where $(\sigma_x)^2 = E(x^2) - (E(x))^2$

$= \sum x^2 p(x_i) - (\text{mean})^2$

$(\sigma_y)^2 = E(y^2) - (E(y))^2$

The Random variable ~~x & y~~ x and y stochastically independent if $\boxed{P_{ij} = P_i Q_j}$

OR,

$\text{cov}(xy) = 0$

$E(xy) = E(x)\,E(y).$

ex: The distribution of two stochastically independent variables x and y defined on the same sample space are given in the following table

x	1	2
P(x)	0.2	0.8

y	1	2	3
P(y)	0.1	0.4	0.5

Q: find joint distribution of x and y, and compute E(x), E(y), and E(xy).

Soln: Since x and y are stochastically independent.

$p_{11} = P_1 Q_1 = 0.2 \times 0.1 = 0.02$

$p_{12} = P_1 Q_2 = 0.2 \times 0.4 = 0.08$

x \ y	1	2	3
1	0.02	0.03	0.1
2	0.08	0.32	0.4

$E(X) = \Sigma x \cdot p(x_i)$

$= 1 \times 0.2 + 2 \times 0.8$

$= 0.2 + 1.6$

$= 1.8$

$E(Y) = \Sigma y p(y_j)$

$= 1 \times 0.1 + 2 \times 0.4 + 3 \times 0.5$

$= 0.1 + 0.8 + 1.5$

$= 2.4$

$E(XY) = \sum_{i=1}^{m} \sum_{j=1}^{m} x_i y_i p_{ij}$

$= 1 \times 1 \times 0.02 + 1 \times 2 \times 0.08 + 1 \times 3 \times 0.1 + 2 \times 1 \times 0.08 + 2 \times 2 \times 0.32 + 2 \times 3 \times 0.4 =$

$= 4.3$

Covariance $= E(XY) - E(X)\,E(Y)$

$= 4.3 - (2.4)(1.8)$

$= -0.02$

Q: A coin is tossed 3 times, Let x denote 0 or 1 according as a tail or head occurs on the first toss. Let y denote the total number of tails which occur. Determine

(1) the marginal distribution of x and y

(2) the joint distribution of x and y, and find expected values of $E(XY)$.

Soln: $S = \{HHH, HHT, HTH, THH, HTT, TTH, TTT, THT\}$

$x = 0 \Rightarrow$ first occurance is a tail,

$\{THH, THT, TTH, TTT\}$

$x = 1 \Rightarrow$ First occurance is a head.

$\{HHH, HHT, HTH, HTT.\}$

x	0	1
$P(x)$	$4/8$	$4/8$

$y = 0$, No tail.

$\{H, H, H\}$

$y = 1$ One tail.

$\{THH, HTH, HHT\}$

$y = 2$ two tail $\Rightarrow \{TTH, THT, HTT\}$.

$y = 3$ 3 tails $\Rightarrow \{TTT\}$

y	0	1	2	3
$P(y)$	$1/8$	$3/8$	$3/8$	$1/8$

Now, joint distribution:

$P_{11} \Rightarrow P(x = 0, y = 0)$

(There is not outcome for tail should occur first and no tail occurs).

$\boxed{P_{11} = 0}$.

$P_{12} \Rightarrow P(x = 0, y = 1) \Rightarrow \{THH\} \Rightarrow P_{12} = \frac{1}{8}$

$P_{13} = P(X=0, Y=2) \Rightarrow \{THT, TTH\} \Rightarrow P_{13} = 2/8$

$P_{14} = P(X=0, Y=3) \Rightarrow \{TTT\} \Rightarrow P_{14} = \frac{1}{8}$.

$P_{21} = P(X=1, Y=0) \Rightarrow \{HHH\} \Rightarrow P_{21} = 1/8$

$P_{22} = P(X=1, Y=1) \Rightarrow \{HT\cancel{H}, HHT\} \Rightarrow P_{22} = 2/8$

$P_{23} = P(X=1, Y=2) \Rightarrow \{HTT\} \Rightarrow P_{23} = 1/8$

$P_{24} = P(X=1, Y=3) \Rightarrow \{\phi\} \Rightarrow P_{24} = 0$.

x \ y	0	1	2	3
0	0	1/8	2/8	1/8
1	1/8	2/8	1/8	0

$\Rightarrow E(XY) = ?$

$E(XY) = \Sigma \Sigma x_i y_j P_{ij}$

$= (1 \times 1 \times \frac{2}{8}) + (1 \times 2 \times \frac{1}{8}) \cancel{+ \times 3}$

$= \frac{2}{8} + \frac{2}{8} \Rightarrow \frac{4}{8} = \frac{1}{2}$

$\boxed{E(XY) = 0.5}$

LP:-
16

X = No. of Red balls

Y = No. of white balls

X + Y = 3 (3 balls are drawn)

(0,3) (1,2), (2,1) (3,0)

P(red ball) = 3/8 = p

P(white ball) = $1 - \frac{3}{8} = \frac{5}{8} = q$

using Binomial distribution,

$B(p,q,r) = {}^{n}C_{r}\, p^{r} q^{n-r}$

$p(0,3) = {}^{3}C_{0}\, (3/8)^{0}\, (5/8)^{3}$ → (No red balls).

$p(1,2) = {}^{3}C_{1}\, (3/8)^{1}\, (5/8)^{2}$

$p(2,1) = {}^{3}C_{2}\, (3/8)^{2}\, (5/8)^{1}$

$p(3,0) = {}^{3}C_{3}\, (3/8)^{3}\, (5/8)^{0}$

$$p(x,y) = \begin{cases} 0.24 & x=0,\ y=3 \\ 0.439 & x=1,\ y=2 \\ 0.26 & x=2,\ y=1 \\ 0.05 & x=3,\ y=0 \end{cases}$$

x \ y	0	1	2	3
0	0	0	0	0.24
1	0	0	0.439	0
2	0	0.26	0	0
3	0.05	0	0	0

LP:

(18) $f(x,y) = k(x+y)$.

$0 \le x \le 2, \quad 0 \le y \le 3$.

x \ y	0	1	2	3	P_i
0	0	k	2k	3k	6k
1	k	2k	3k	4k	10k
2	2k	3k	4k	5k	14k
Q_j	3k	6k	9k	12k	30k

$30k = 1$

$\boxed{k = \frac{1}{30}}$

Marginal distribution of x :

x	0	1	2
P(x)	6/30	10/30	$\frac{14}{30}$

Marginal distribution of y

y	0	1	2	3
P(y)	3/30	6/30	9/30	12/30

Corelation coefficient, $\rho = \frac{cov(xy)}{\sigma_x \sigma_y}$

$$Cov(xy) = E(xy) - E(x)E(y)$$

$$E(xy) = ((1 \times 1 \times 2k) + (1 \times 2 \times 3k) + (1 \times 3 \times 4k) + (2 \times 2 \times 4k) + (2 \times 3 \times 5k)$$

$$+ (2 \times 1 \times 3k) + (2 \times 2 \times 4k) + (2 \times 3 \times 5k)$$

$$(\cancel{3 \times 1 \times 4k}) + \cancel{(3 \times 2 \times 5k) + (3 \times 3 \times 6k}$$

$$E(xy) =$$

$$E(x) = 1 \times \frac{10}{30} + 2 \times \frac{14}{30} = 1.26$$

$$E(y) = 1 \times \frac{6}{30} + \frac{9}{30} \times 2 + 3 \times \frac{12}{30} = 2$$

$$(\sigma_x)^2 = \Sigma(x^2) - (E(x))^2$$

$$E(x^2) = 0^2 \times \frac{6}{30} + 1^2 \times \frac{10}{30} + 2^2\left(\frac{14}{30}\right)$$

$$= \frac{10}{30} + 4\left(\frac{14}{30}\right)$$

$$= \frac{66}{30} = 2.2$$

$$(E(x))^2 = (1.26)^2 = 1.5876$$

$$= 2.2 - 1.58$$

$$= 0.624$$

$$\boxed{\sigma_x = 0.783}$$