Bernoulli Distribution:

A random variable X which takes two values 0 and 1, with probability q and p, resp. i.e $P(X=1) = p$, $P(X=0) = q = 1-p$ is called a Bernoulli variate and is said to have a Bernoulli distribution.

Mean $= p$. Variance $= p - p^2 = p(1-p)$
$= pq$.

Binomial distribution:

Consider a set of n independent Bernoulli trials (n - finite), in which the probability 'p' of success in any trial is constant. Then $q = 1-p$, is the probability of failure.

∴ The probability of x successes in n trials is given by: ${}^nC_x \cdot p^x q^{n-x}$.

The probabilities of 0, 1, 2 -- n successes are the successive terms of binomial expansion $(q+p)^n$.

Def: A random variable X is said to follow Binomial distribution if it assumes only non-negative values and its prob. mass function is given by:

$$P(X=x) = p(x) = \begin{cases} {}^nC_x\, p^x q^{n-x} & : x = 0, 1, \ldots n. \\ & \quad q = 1-p. \\ 0 & \text{otherwise.} \end{cases}$$

Suppose that 'n' trials constitute an experiment. Then if experiment is repeated N times, the frequency (function) of the binomial distribution is:

$$\boxed{f(x) = N \cdot p(x)}$$

The Binomial distribution is applied when:

1) Each trial results in two mutually disjoint outcomes, termed as success and failure.
2) The number of trials 'n' is finite.
3) The trials are independent of each other.
4) The prob. of success 'p' is constant for each trial.

Q: Ten coins are thrown simultaneously. Find the probability of getting at least seven heads.

Sol: $p = 1/2 = q$ $\quad$ $p \rightarrow$ getting head

$q \rightarrow$ not getting head.

$P(X \geq 7) = p(7) + P(8) + P(9) + P(10)$

$= {}^{10}C_7 (1/2)^7 (1/2)^3 + {}^{10}C_8 (1/2)^{10} + {}^{10}C_9 (1/2)^{10} + {}^{10}C_{10} (1/2)^{10}$

$= (\frac{1}{2})^{10} \cdot [120 + 45 + 10 + 1] = \frac{176}{1024}$

Eg: A and B play a game in which their chances of winning are in the ratio 3:2. Find A's chance of winning at least three games out of the five games played.

Sol: Let p be the prob. that A wins.

$\therefore p = 3/5$ $\quad$ $q = 2/5$.

Calculate $P(X \geq 3)$: $\quad$ Ans = 0.68.

Eg: An irregular six faced (die) is thrown and the expectation that in 10 throws it will give five even numbers is twice the expectation that it will give four even numbers. How many times in 10,000 sets of 10 throws each, would you expect it to give no even number.

Hint: $P(X=5) = 2\,P(X=4)$ $X \to$ x even no's. $p \to$ prob. of getting even no..

ie $\frac{p}{5} = \frac{2}{3}$.

ie $3p = 5(1-p) \Rightarrow p = 5/8,\ q = 3/8$.

In 10,000 sets, no even no.

$= 10000 \times f(0) = 10000 \times (3/8)^{10} = 1$ (approx)

Moments:

$$\mu_1 = E(X) = \sum_{x=0}^{n} x\ {}^nC_x\ p^x q^{n-x}.$$

$$= \sum_{1}^{n} \cancel{x}\ {}^{n-1}C_{x-1} \frac{(n)}{\cancel{x}} p^{x-1}.p.q^{n-x}$$

$$= np\,[p+q]^{n-1}$$

$$= np\,(1)$$

$$\boxed{{}^nC_x = \frac{\lfloor n}{\lfloor x \lfloor n-x} = \frac{n.\lfloor n-1}{x\lfloor x-1.\lfloor n-x} = \frac{n}{x}.{}^{n-1}C_{x-1}}$$

$$\mu_2' = E(X^2) = \sum x^2 \binom{n}{x} p^x q^{n-x}$$

$$= \sum (x(x-1)+x)\binom{n}{x} p^x q^{n-x}$$

$$= n(n-1)p^2 + np.$$

$v(x) = E[(x-E(x))^2]$. ③

$V(X) = E(x^2) - (E(x))^2$
$= npq.$

Ex: The mean and variance of binomial distribution are 4 and 4/3. Find $P(x \geq 1)$

Ans: $p = 2/3$: $n = 6$: $P(x \geq 1) = 0.99863$.

Moment generating function:

$$M_x(t) = E[e^{tx}] = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x q^{n-x}$$

$$= \sum (pe^t)^x q^{n-x} = (q + e^t p)^n.$$

Coeff of $t = \mu_1 = E(x)$

Coef of $\frac{t^2}{\lfloor 2} = \mu_2' = E(x^2)$

$$\frac{d}{dt}[M_x(t)]_{t=0} = n(q+e^t p)^{n-1} (e^t \cdot p) \Big|_{t=0}$$

$$= n(q+p)^{n-1}(p)$$

$$= np.$$

$$\frac{d^2}{dt^2}[M_x(t)] = n(n-1)(q+e^t p)^{n-2} (pe^t) pe^t$$
$$+ pe^t \cdot n(q+e^t p)^{n-1} \text{ at } t=0$$
$$= n(n-1)p^2 + np = np((n-1)p+1)$$
$$= np(np+q)$$

$\Sigma^{0}/$

Poisson Distribution

[Limiting Case of Binomial Distribution]

1) n, the no. of trials is indefinite i.e. $n \to \infty$
2) p, the constant prob. of success for each trial is indefinitely small i.e. $p \to 0$
3) $np = \lambda$ is finite ; $p = \lambda/n$; $q = 1 - \lambda/n$.

The probability of x success in a series of n independent trials is:

$$^{n}C_{x}\, p^{x} q^{n-x} = \frac{\lfloor n}{\lfloor x \lfloor n-x}\, p^{x} q^{n-x}$$

$$= \frac{\lfloor n}{\lfloor x \lfloor n-x} \left(\lambda/n\right)^{x} \left(1 - \lambda/n\right)^{n-x}$$

$\to 1$

$$= \frac{n(n-1) - (n-(x-1))}{\lfloor x} \cdot \frac{\lambda^{x}}{n^{x}} \left(1 - \lambda/n\right)^{n} \left(1 - \lambda/n\right)$$

as $n \to \infty$.

$$= \frac{\lambda^{x}}{\lfloor x} e^{-\lambda}. \quad x = 0, 1, 2 - \infty.$$

i.e. $X \sim P(\lambda) \Rightarrow P(X = x) = \dfrac{e^{-\lambda} \cdot \lambda^{x}}{\lfloor x}$: $x = 0, 1, \ldots$

0 otherwise

λ is known parameter.

Moments:

$$\mu_1 = E(X) = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

$$= \lambda e^{-\lambda} \sum^{\infty} \frac{\lambda^{x-1}}{\lfloor x-1}$$

$$= \lambda e^{-\lambda} e^{\lambda} = \boxed{\lambda = E(X)}$$

$$\mu_2' = \lambda^2 + \lambda.$$

$$\therefore \boxed{V(X) = \lambda.}$$

MGF: $= e^{\lambda(e^t - 1)}$. (Check it)

Eg: A Car hire firm has two Cars which it hires out day by day. The no. of demand for a car on each day is distributed Poisson variate with mean 1.5. Calculate the proportion of days on which

(i) Neither car is used

(2) Some demand is refused.

Sol:

$$P(X=x) = \frac{e^{-\lambda}\lambda^x}{\lfloor x}.$$

$\lambda = 1.5$

$x \to$ dem

$$P(X=0) = e^{-1.5} = 0.2231$$

$$P(X>2) = 1 - P(X \le 2)$$

$$= 1 - e^{1.5}\left[1 + 1.5 + \frac{(1.5)^2}{\lfloor 2}\right]$$

$$= 0.19126.$$

Q: In a book of 520 pages, 390 typo-graphical errors occurs. Assuming Poisson law for the number of errors per page, find the probability that a random sample of 5 pages contains no error.

Sol: $\lambda = \frac{390}{520} = 0.75$ $\qquad P(x) = \frac{e^{-\lambda}\lambda^x}{\lfloor x}$.

Req. prob: $(P(X=0))^5 = (e^{-0.75})^5 = e^{-3.75}$.

Continuous Distributions

Uniform distribution: A random variable X is said (Rectangular) to have a continuous uniform distribution over an interval (a,b) if its pdf is constant 'k' over the entire range of X ie

$$f(x) = \begin{cases} K : & a<x<b \\ 0 : & \text{otherwise} \end{cases}$$

$$\int_a^b f(x)\,dx = 1 \Rightarrow K = \frac{1}{b-a}.$$

Cdf: $$F(x) = \begin{cases} 0 & \text{if } -\infty < x < a \\ \frac{x-a}{b-a} & \text{if } a \le x \le b. \\ 1 & \text{if } b < x < \infty \end{cases}$$

Find its moments.

Def: A random variable X is said to have a normal distribution with parameters: μ - mean, $\sigma^2 \rightarrow$ variance; if its pdf is:

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

$$-\infty < x < \infty$$
$$-\infty < \mu < \infty$$
$$\sigma > 0.$$

we write $X \sim N(\mu, \sigma^2)$.

and $Z = \dfrac{X-\mu}{\sigma}$ is a standard normal variate

with $E(Z) = 0$ $\quad var(Z) = 1$,

ie $Z \sim N(0,1)$.

Pdf of Z is: $\dfrac{1}{\sqrt{2\pi}} e^{-1/2 z^2}$: $-\infty < z < \infty$

Properties:

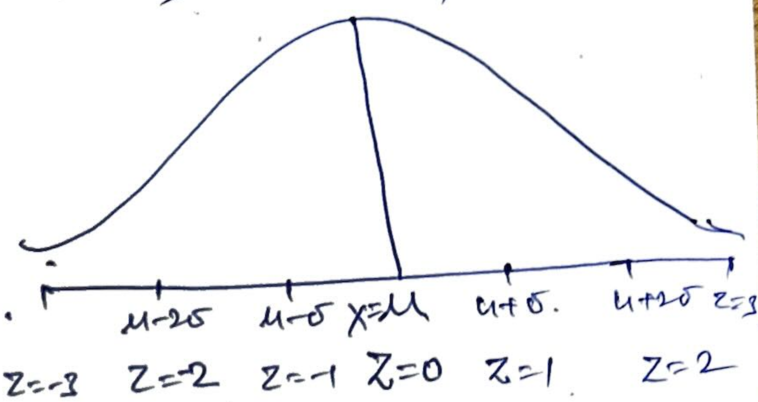
1) Symmetric, bell shaped.
2) $-\infty < X < \infty$
3) Parameters: μ and σ.
4) 2/3 of all cases fall within one standard deviation of the mean, ie

$$P(\mu-\sigma \le X \le \mu+\sigma) = 0.6826$$

5) 95% of cases lies within 2 standard deviations of the mean

$$P(\mu-2\sigma \le X \le \mu+2\sigma) = .9544$$

$$P(\mu-3\sigma < x < \mu+3\sigma)$$
$$= 0.9973$$



6) Mean, Median, Mode Coincide.

) $P(Z \le a)$

$= F(a)$

$= 1 - F(-a)$

$Z \rightarrow$ Standardised normal variate

ie $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

ie $X = \mu + \sigma Z$.

$$P(a \le Z \le b) = F(b) - F(a)$$

Total Area under normal prob. Curve is unity.

$P(|Z| < 3) = 0.9973$

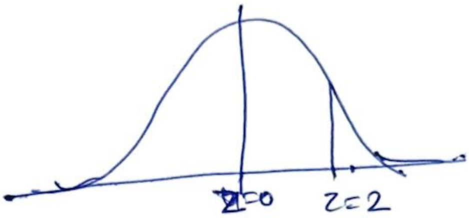
$P(|Z| > 3) = 0.0027$

Eg: X is normally distributed and the mean of X is 12, S.D is 4.

a) Find the prob. of

1) $X \geq 20$; 2) $X \leq 20$ 3) $0 \leq X \leq 12$.

4) Find x' s.t $P(X > x') = 0.24$

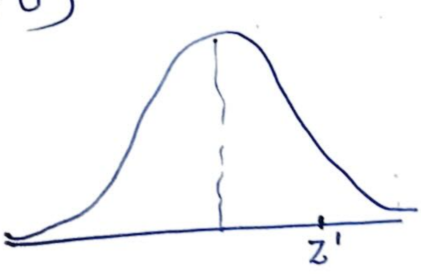


Sol: $Z = \frac{X-\mu}{\sigma} = \frac{X-12}{4}$

i.e 1) $P(X \geq 20) \doteq P(Z \geq 2) = 0.5 - P(0 \leq z \leq 2)$
$= 0.5 - 0.4772$
$= 0.0228$

2) $P(X \leq 20) = P(z \leq 2) = 0.5 + 0.4772$
$= 0.9772$

3) $P(0 \leq X \leq 12) = P(-3 \leq z \leq 0)$
$= 0.49865.$



4) $P(X > x') = 0.24$

Say $P(Z > z') = 0.24 \Rightarrow$ P

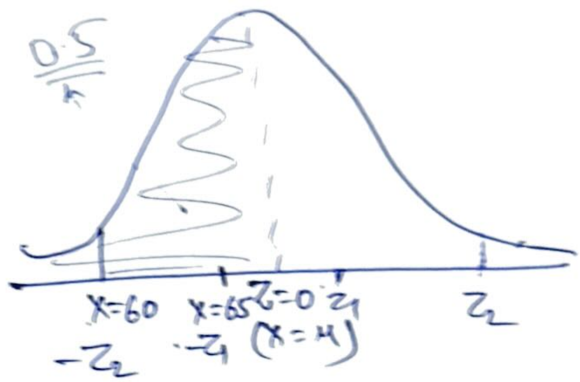
$0.5 - P(0 < z < z') = 0.24$

$P(0 < z < z') = 0.5 - 0.24 = 0.26$

$\Rightarrow z' = 0.71.$ $\Rightarrow x' = \mu + \sigma z'$
$= 12 + 4(0.71) = 14.84$

Q: In a large group of men, 5% are under 60 inches in height and 40% are b/w 60 and 65 inches. Assuming a normal distribution, find the mean height and S.D.

Sol: $X \sim N(\mu, \sigma^2)$.



$P(X < 60) = 0.05$

$P(60 < X < 65) = 0.40$

ie $P(X < 65) = 0.45$

From diag $X = 60$, $X = 65$ both located left of $X = \mu$.

when $X = 65$ ie $Z = \frac{X-\mu}{\sigma} = \frac{65-\mu}{\sigma} = -z_1$ (say)

$X = 60$ ie $Z = \frac{60-\mu}{\sigma} = -z_2$ (say)

$P(0 < Z < z_2) = 0.45 \quad (0.5 - 0.05)$

and $P(0 < Z < z_1) = 0.5 - 0.45 = 0.05$

ie $z_2 = 1.645$, $z_1 = 0.13$.

ie $\frac{60-\mu}{\sigma} = -1.645$ and $\frac{65-\mu}{\sigma} = -0.13$.

$$\Rightarrow \boxed{\begin{aligned} \mu &= 65.42 \\ \sigma &= 3.29. \end{aligned}}$$

Eg: Students passes if secures min 30%.
first if more than 60%.
second if more 45%, less than 60%.
third if b/w 30% to 45%.
distinction if more than 80%.
10% fails, 5% obtain distinction.
Calculate the % of students placed in the second division.

Sol: Let X denotes the marks (say out of 100).

$P(X \le 30) = 0.10$

$P(X \ge 80) = 0.05$

$z = \frac{30 - \mu}{\sigma} = -z_1$ (say) when $X = 30$.

$z = \frac{80 - \mu}{\sigma} = z_2$ (say) when $X = 80$.

$$P(0 < z < z_1) = 0.5 - P(0 < z < z_1)$$
$$= 0.5 - 0.10$$
$$= 0.40$$
$$\Rightarrow z_1 = 1.29$$

$$P(0 < z < z_2) = \cancel{0.05} \; \cancel{0.50} \; 0.45$$
$$\Rightarrow z_2 = 1.65$$

$$\Rightarrow \frac{30 - \mu}{\sigma} = -1.29 ; \quad \frac{80 - \mu}{\sigma} = 1.65$$

i.e $\left.\begin{aligned} \frac{\mu - 30}{\sigma} &= 1.29 \\ \frac{80 - \mu}{\sigma} &= 1.65 \end{aligned}\right\} \Rightarrow \frac{50}{\sigma} = 2.94$

$$\sigma = \frac{50}{2.94} = 17 \text{ (approx)}$$

$$\Rightarrow \mu = 30 + 1.29(17) \approx 30 + 22 = \underline{\underline{52}}$$

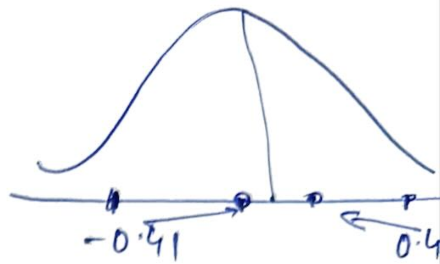
$$45 < X < 60) = P\left(\cancel{\otimes}\,\frac{45-52}{17} < \frac{X-\mu}{\sigma} < \frac{60-52}{17}\right)$$

↓

second division

$$= P(-0.41 < Z < 0.47)$$

$$= P(0 < Z < 0.41) + P(0 < Z < 0.47)$$



$$= 0.1591 + 0.1808$$

$$= 0.3399 \approx 0.34.$$

ie 34% students got second division.

x ——————————— x

Joint probability function:

Two random variables X and Y are said to be jointly distributed if they are defined on the same prob. space. Their joint pdf: is Defined and denoted as:

$X(S) = \{x_1, \dots x_n\}$

$Y(S) = \{y_1, y_2 \dots y_m\}$

$$P_{XY}(x, y) = P_{XY}(X = x, Y = y)$$

X \ Y	y_1	y_2	—	—	y_m	$p_{1.}$
x_1	p_{11}	p_{12}	—		p_{1m}	$p_{2.}$
x_2						
$\vdots$	—	—	—			
	—	—	—			
x_n	—	—	—			$p_{n.}$
	$p_{.1}$	—	—	—	$p_{.m}$	1

$$\sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} = 1.$$

Marginal $P(Y = y_j) = \sum_{i=1}^{n} p_{ij}$ and $P(X = x_i) = \sum_{j=1}^{m} p_{ij}$

Conditional $P(X = x_i \mid Y = y_j) = \dfrac{P(X = x_i \cap Y = y_j)}{P(Y = y_j)} = \dfrac{P(x_i, y_j)}{P(y_j)}$

$$= \frac{p_{ij}}{p_{.j}}$$

Joint Distribution function: is defined as:

$$F_{XY}(x, y) = P(X \le x, Y \le y)$$

$$\left.\begin{aligned} F_X(x) &= \sum_{y} P(X \le x, Y = y) \\ F_Y(y) &= \sum_{x} P(X = x, Y \le y) \end{aligned}\right\} \text{Marginal}$$

In Continuous Case: (Marginal pdf's)

$$F_X(x) = \int_{-\infty}^{x}\left(\int_{-\infty}^{\infty} f_{XY}(x,y)\,dy\right)dx$$

$$F_Y(y) = \int_{-\infty}^{y}\left(\int_{-\infty}^{\infty} f_{XY}(x,y)\,dx\right)dy.$$

Eg: Let X, Y be continuous variables with joint pdf:

$$f(x,y) = x+y : \quad 0 \le x, y \le 1.$$

Then marginal pdf's of X and Y are:

$$f_X(x) = \int_0^1 f(x,y)\,dy$$

$$= \int_0^1 (x+y)\,dy = \left(xy + \frac{y^2}{2}\right)\Big|_0^1$$

$$= x + 1/2 \quad ; \quad 0 \le x \le 1$$

$$f_Y(y) = y + 1/2. \quad 0 \le y \le 1.$$

Also for

$$g(x,y) = (x+1/2)(y+1/2)$$

$$0 \le x, y \le 1$$

$$g_1(x) = x+1/2 \quad 0 \le x \le 1$$

$$g_2(y) = y+1/2 \quad 0 \le y \le 1$$

Two different joint pdf have the same marginal pdf's.

Two variables X, Y with joint pdf $f(x,y)$ and marginal pdf's $f_X(x)$, $g_Y(y)$ resp. are said to be independent iff

$$f_{XY}(x,y) = f_X(x)\, g_Y(y).$$

Eg: For given distribution table:

Y \ X	−1	0	1	
0	1/15	2/15	1/15	4/15
1	3/15	2/15	1/15	6/15
2	2/15	1/15	2/15	5/15
	6/15	5/15	4/15	1

Find (i) marginal distribution of X, Y

(ii) the Conditional distribution of X given $Y=2$

Sol: i) $P(X=-1) = \sum_{y=0}^{2} P(X=-1, Y=y) = 6/15$

$P(X=0) = 5/15$

$P(X=1) = 4/15$

$P(Y=0) = 4/15$, $P(Y=1) = 6/15$ $P(Y=2) = 5/15$.

ii) $P(X=x \mid Y=2) = \dfrac{P(X=x \cap Y=2)}{P(Y=2)}$

ie $P(X=-1 \mid Y=2) = \dfrac{P(X=-1, Y=2)}{P(Y=2)} = \dfrac{2/15}{1/3} = 2/5$

etc

Eg: If $f(x,y) = \frac{1}{27}(x+y)$ $x = 0,1,2$, $y = 0,1,2$

Find $P(Y=y \mid X=x)$.

Hint: $P(Y=y_0 \mid X=x) = \dfrac{P(Y=y \cap X=x)}{P(X=x)}$

$X \downarrow$ / $Y \rightarrow$	0	1	2
0	0	1/3	
1			
2			

etc

$P(Y=0 \mid X=0) = \dfrac{P(Y=0, X=0)}{P(X=0)}$

$= 0.$

$P(Y=0 \mid X=1) = \dfrac{P(Y=0, X=1)}{P(X=1)}$

$= \dfrac{1/27}{1/27 + \frac{2}{27}} = \frac{1}{3}$

Ex: $f(x,y) = \dfrac{9(1+x+y)}{2(1+x)^4(1+y)^4}$ $0 \le y < \infty$, $0 \le x < \infty$

Find marginal distribution of X and Y.

Ex: $f(x,y) = 2$ $0 < x < 1$, $0 < y < x$
$\quad\quad\quad\;\; 0$ elsewhere

(0) Check if $f(x,y)$ is pdf

(i) Find the marginal density function of X, Y

(ii) Find the conditional density func of Y given $X = x$.

$$f(x,y) = e^{-(x+y)} \quad : \begin{matrix} 0 < x < \infty \\ 0 < y < \infty \end{matrix}$$

(i) $P(X>1)$

(ii) $P(X<Y \mid X<2Y)$

(iii) $P(1 < X+Y < 2)$.

Sol:

$$f(x,y) = e^{-x} \cdot e^{-y} = f_X(x)\, f_Y(y)$$

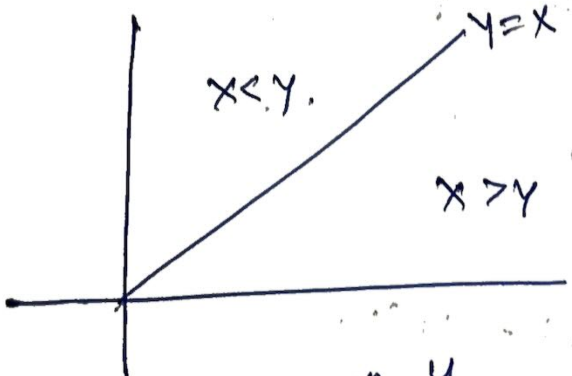
$\Rightarrow$ X, Y are independent.

i.e. $f_X(x) = e^{-x} \;:\; x \geq 0$

$f_Y(y) = e^{-y} \;:\; y \geq 0$

1) $P(X>1) = \int_1^\infty e^{-x}\,dx = 1/e$

2) $P(X<Y \mid X<2Y) = \dfrac{P(X<Y \cap X<2Y)}{P(X<2Y)} = \dfrac{P(X<Y)}{P(X<2Y)}$



$= \dfrac{1/2}{2/3}$

$= 3/4$

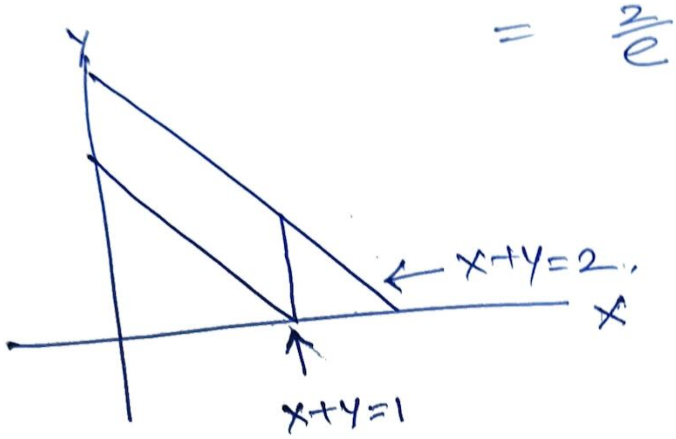
$$P(X<Y) = \int_0^\infty \left(\int_0^y f(x,y)\,dx \right) dy = 1/2$$

$$P(X<2Y) = \int_0^\infty \left(\int_0^{2y} f(x,y)\,dx \right) dy = 2/3$$

(iii) $P(1 < X+Y < 2)$

$$= \int_0^1 \left(\int_{1-x}^{2-x} f\, dy \right) dx + \int_1^2 \left(\int_0^{2-x} f\, dy \right) dx$$

$$= \frac{2}{e} - 3/e^2.$$



Ex $f(x,y) = 1/8\,(6-x-y) \quad : \quad 0<x<2,\ 2<y<4$

$0 \quad$ elsewhere

Find (i) $P(X<1 \cap Y<3)$ (ii) $P(X+Y<3)$

(iii) $P(X<1 \mid Y<3)$

Ans: (i) $3/8$ (ii) $5/24$ (iii) $3/5$.

Transformation: If $u = u(x,y)$
$v = v(x,y)$

$$J = \frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{vmatrix}$$

then $g_{UV}(u,v) = f_{XY}(x,y)\,|J|.$

Q: Given the joint density function of X, Y as:

$$f(x,y) = \begin{cases} \frac{1}{2} x e^{-y} & : 0 < x < 2 \quad y > 0 \\ 0 & \text{elsewhere} \end{cases}$$

Find the distribution of $X+Y$.

Sol: Let $u = x+y$ $\quad v = y$

$\Rightarrow y = v$ and $x = u - v$.

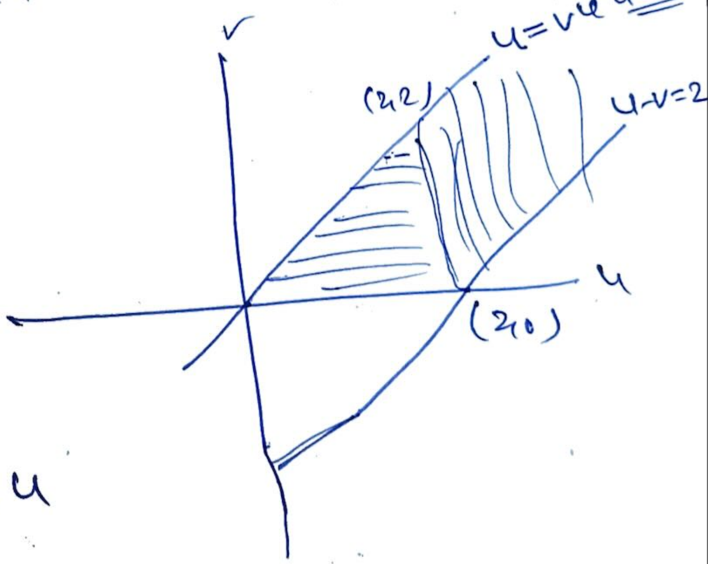
$$J = \frac{\partial(x,y)}{\partial(u,v)} = 1.$$

Region

$0 < x < 2 \quad\quad y > 0$

$\downarrow \quad\quad \downarrow$

$0 < u - v < 2 \quad\quad v > 0$



$$g(u,v) = |J| \frac{1}{2}(u-v) e^{-v}$$

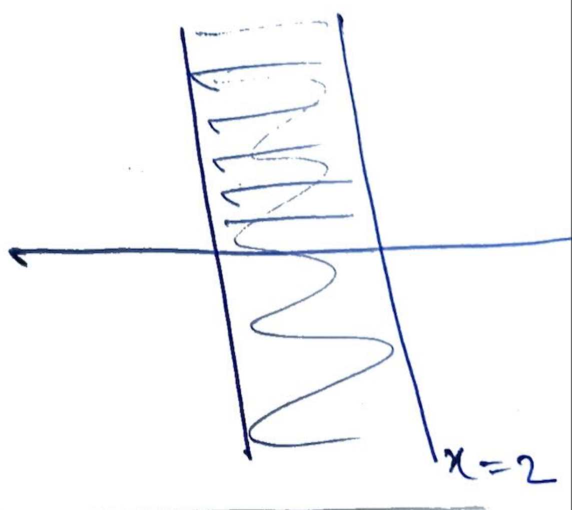
$: 0 < v < u, \quad u > 0, \ u < 2$

$u-2 < v < u \quad u > 2$

Region $= 0 < u \leq 2$ i.e. $0 < v < u$

$$h(u) = \int_0^u g(u,v)\, dv$$

$$= \frac{1}{2}(e^{-u} + u - 1)$$



$R = 2 < u < \infty$

$$h(u) = \frac{1}{2}\int_{u-2}^{2} (u-v)\, e^{-v}\, dv$$

$$= \frac{1}{2} e^{-u}(1+e^{2})$$

ie

$$f(u) = \begin{cases} 1/2\,(e^{-u}+u-1) & : 0 \le u \le 2 \\ 1/2\,(e^{-u})(1+e^{2}) & 2 < u < \infty \\ 0 & \text{else} \end{cases}$$

X ———— X

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

2) If X, Y are independent, then

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

3) $E(X+a) = E(X) + a$

4) $E(aX) = a\,E(X)$

Var 1) $V(aX+b) = a^2 V(X)$ | $V(b) = 0$.

2) $V(X+b) = V(X)$

Co-variance

$$\operatorname{Cov}(X,Y) = E\big[[X - E(X)][Y - E(Y)]\big].$$

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

If X, Y independent $\Rightarrow$ $\operatorname{Cov}(X,Y) = 0$.

\# $\text{Cov}(X,Y) = E[(X-E(X))(Y-E(Y))]$

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

If X, Y are independent, then $E(XY) = E(X)E(Y)$

ie $\text{Cov}(X,Y) = 0$

\# $\text{Cov}(X+a, Y+b) = \text{Cov}(X,Y)$

\# $\text{Cov}(X+Y, Z) = \text{cov}(X,Z) + \text{cov}(Y,Z)$.

<u>Correlation Coefficient:</u> The correlation coefficient ρ_{XY}, b/w the variables X and Y is defined as:

$$\rho_{XY} = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}$$

\# $V(X_1+X_2) = V(X_1) + V(X_2) + 2\,\text{Cov}(X_1, X_2)$

$V(X_1-X_2) = V(X_1) + V(X_2) - 2\,\text{Cov}(X_1, X_2)$

OR $V(X_1+X_2+\dots X_n) = \sum_{i=1}^{n} V(X_i) + 2\sum_{i}\sum_{j,\, i<j} \text{Cov}(X_i, X_j)$

Eg! A deck of n numbered Cards is thoroughly Shuffled and the Cards are inserted into 'n' numbered Cells one by one. If the Card number 'i' falls in the Cell 'i', we count it as a match, otherwise not. Find mean and variance of total number of such matches.

Sol: Let X_i be the r.v with i^{th} draw defined

$$X_i = \begin{cases} 1 & i^{th} \text{ Card on } i^{th} \text{ cell} \\ 0 & \text{otherwise} \end{cases}$$

$$S = X_1 + X_2 + \cdots X_n = \sum_{i=1}^{n} X_i.$$

$$E(S) = \sum_{i=1}^{n} E(X_i) = n.\left[1 \cdot P(X_i=1) + 0 \cdot P(X_i=0)\right]$$

$$= n.\left(1/n\right) = 1.$$

$$\left[\because P(X_i=1) = 1/n\right]$$

$$V(S) = V(X_1 + X_2 + \cdots X_n)$$

$$= \sum_{i=1}^{n} V(X_i) + 2 \sum_i \sum_j Cov(X_i, X_j)$$

where

$$V(X_i) = E[X_i^2] - [E(X_i)]^2$$

$$= 1^2 \cdot P(X_i=1) + 0 - 1/n^2$$

$$= 1/n - 1/n^2$$

$$Cov(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j)$$

$$= 1 \cdot P(X_i X_j = 1) + 0 \cdot P(X_i X_j = 0) - 1/n^2$$

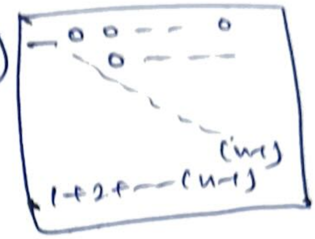
$$= \left(\frac{\lfloor n-2}{\lfloor n}\right) \frac{1}{n(n-1)} - 1/n^2$$

[$X_i X_j = 1$ iff both Cards i^{th}, j^{th} are on their resp. matching places
and there are $\lfloor n-2$ arrangements of remaining Cards

$$V(S) = \sum_{i=1}^{n}\left(\frac{n-1}{n^2}\right) + 2\sum\sum\frac{1}{n^2(n-1)}.$$

$$= \left(\frac{n-1}{n}\right) + \frac{2}{n^2(n-1)}\sum_{i}\sum_{\substack{j \\ i<j}} 1.$$

$$= \left(\frac{n-1}{n}\right) + \frac{2}{n^2(n-1)}\left({}^{n}C_2 = \frac{n(n-1)}{2}\right)$$



$$= \frac{(n-1)}{n} + \frac{1}{n} = 1.$$

<u>Exer</u>: A man with n keys wants to open his door and tries the keys independently and at random. Find the mean and variance of the number of trials required to open the door.

(i) if unsuccessful keys are not eliminated for ~~from~~ further selection.

(ii) if they are.

<u>Hint:</u> Suppose the man gets the first success at the x^{th} trial. ie he is unable to open the door in the first $(x-1)$ trials.

<u>(i)</u> X can take the values $1, 2, 3, \ldots$ infinity.

Prob. of success at first trial $= 1/n$.

failure ——— $= 1 - 1/n$.

Hence

$$p(x) = (1-1/n)^{x-1} \cdot 1/n.$$ (unsuccessful keys are not eliminated)

(ii) If unsuccessful keys are eliminated.

$\therefore$ X can take values: 1 to n.

Prob at first trial = $1/n$.
(of success)

2nd trial $= \frac{1}{n-1}$.

3rd trial $= \frac{1}{n-2}$. and so on...

Hence first success at 2nd trial $= (1-1/n)\frac{1}{n-1} = \frac{1}{n}$.

first success at 3rd trial $= (1-\frac{1}{n})(1-\frac{1}{n-1})\frac{1}{n-2}$

$= 1/n$.

$\therefore$ $p(x) =$ prob of first success at x^{th} trial

$= 1/n$.

$\therefore$ $E(X) = \frac{n+1}{2}$

$V(X) = \frac{n^2-1}{12}$ } Calculate it.

Exer: Find E(X), V(X) for (i) part.

Ans: $E(X) = n$

$V(X) = n(n-1)$.

Additive property: (of Binomial variates)

Let $X \sim B(n_1, p_1)$ and $Y \sim B(n_2, p_2)$ be independent r.v's. Then $M_X(t) = (q_1 + p_1 e^t)^{n_1}$

$M_Y(t) = (q_2 + p_2 e^t)^{n_2}$

X, Y independent $\Rightarrow$ $M_{X+Y}(t) = M_X(t) M_Y(t)$

$= (q_1 + p_1 e^t)^{n_1} (q_2 + p_2 e^t)^{n_2}$

not in the form $(q + pe^t)^n$.

$\therefore$ from Uniqueness theorem of mgf's;

$X+Y$ is not a binomial variate

But if $p_1 = p_2 = p$. $\Rightarrow$ $M_{X+Y}(t) = (q + pe^t)^{n_1 + n_2}$

ie $X+Y \sim B(n_1 + n_2, p)$.

Eg: If the independent binomial random variables, X and Y, distributed with $p_1 = p_2 = 1/3$ and $n = 3$ and $n = 5$, resp.

Find prob. that $X+Y \geq 1$.

Hint: $X \sim B(3, 1/3)$: $Y \sim B(5, 1/3)$.

$X+Y \sim B(8, 1/3)$.

$\therefore$ $P(X+Y = r) = {}^8C_r (1/3)^r (2/3)^{8-r}$

$\therefore P(X+Y \geq 1) = 1 - P(X+Y=0) = 1 - \left(\frac{2}{3}\right)^8$.

Additive property of Poisson variates.

Let $X_1, X_2, - X_n$ are Poisson independent variates with parameter λ_i. Then

$\sum x_i$ is also a Poisson variate with parameter $\lambda = \sum_{i=1}^{n} \lambda_i$

$X_i \sim P(\lambda_i) : i = 1, 2, - n.$

$$M_{X_i}(t) = e^{\lambda_i (e^t - 1)}$$

and $M_{X_1 + X_2 + - X_n}(t) = M_{X_1}(t) - M_{X_n}(t)$

$$= e^{(\lambda_1 + \lambda_2 + - \lambda_n)(e^t - 1)}$$

Hence the result.

Check: $X_1 - X_2$ is not a Poisson variate (Exercise)

$\because \left[M_{cX}(t) = M_X(ct) \right]$

Eg: Suppose that the no. of telephone Calls Coming into a telephone exchange b/w 10 to 11 AM ;

say $X_1 \sim P(\lambda = 2)$; | X_1, X_2 are independent

From 11 to 12 : $X_2 \sim P(\lambda = 6)$.

Find the prob. that more than 5 Calls Come in 10 to 12.

Hint: $X = X_1 + X_2 \Rightarrow X \sim P(2 + 6 = 8)$.

$$P(X = x) = \frac{e^{-8} \lambda^8}{\lfloor x}, \quad x = 0, 1, 2 --$$

$$P(X > 5) = 1 - P(X \le 5) = 1 - \sum_{x=0}^{5} \frac{e^{-8} 8^x}{\lfloor x} = 1 - 0.1912$$

$$= 0.8088.$$