Geometric Distribution

One of discrete rand var dist.

It also uses Ber. trails (Yes/No) 2 possible outcomes.

$$P(X=x) = pq^{x-1} \; ; x = 1, 2, \ldots$$

$$p(X=x) = p(1-p)^{x-1} = pq^{x-1} \; ; x = 1, 2 \ldots$$

No derivation

Mean & Var of G.D

$\mu = E(X) = 1/p$

$\sigma^2 = Var(X) = \frac{1-p}{p^2}$.

Prob

1) Prob that target is destroy on any shot is 0.5, what is prob that it would be destroyed on 6th attempt?

$\Rightarrow$

$p(x) = 0.5$

$= (0.5)(0.5)^{6-1}$

$p(x) = \cancel{0.5} \; 0.015625$

2) Suppose that a traine soldier shoots a target in independent fashion.
If prob to hit a ~~targ~~ shot is 0.8

a) prob to hit a target on 6^{th} attempt?

b) prob that it takes him less than 5 shots.

c) prob of taking even no. of shots.

$\Rightarrow$ $P = 0.8 \quad q = 0.2$

a) $p(x=x) = 0.8\,(0.2)^{6-1} = 0.00032$

$= 0.000256$

b) $p(x<5) \overset{1,2,3,4}{=} 0.9984$

c) ~~$p(x = 2,4,6 \quad = 0.166$~~.

$$\sum_{x=1}^{\infty} (0.8)\,(0.2)^{2x-1}$$

$$= (0.8)\,(0.2) \sum_{n=1}^{\infty} (0.2)^{2n-2})$$

$$= 0.16 \times \frac{1}{1-(0.2)^2} = \frac{0.16}{0.96}$$

$= 0.1666$

Hyper - Geometry Dist.

Negative Binomial dist.

HGD: Hypergeometric Prob dist:

The x success (defective) can be choosen in $\binom{m}{x}$ ways:

The n-x failures (non-def.) can be choosen in $\binom{N-m}{n-x}$ ways.

The n obj. can be chosen from a set of N-obj in $\binom{N}{n}$ ways.

If we consider as equally likely, it follows that for sampling without replacement the prob of getting x successes in n trails is

$$h(x; n, m, N) = \frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}$$

$x = 0, 1, \ldots n; \quad x \leq m; \quad n-x \leq N-m,$

Exponential Distributn :

distance b/w

pathhole problem

no. of

Poisson Dist :

pathholes → no. of pathh → poisson

→ distance betn pathhol. → expo. dist

7.7.1) pdf :

pdf of expo dist.

$$f_X(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

λ is called rate parameter. $\lambda \propto \frac{1}{\mu}$

$\mu \rightarrow$ expected duration

Eg: $\mu = 5$ mins $\lambda = 0.2$ $\lambda = \frac{1}{\mu}$.

cdf :

$$F_X(x;\lambda) = \begin{cases} 1-e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$p(X \leq x) = 1-e^{-\lambda x}$$

Eg : $\lambda = 0.2$ $x = 10$

$$p(x \leq 10) = 1-e^{-2}.$$

Mean & Variance of expo dist.

$$E[X] = \frac{1}{\lambda} = \mu.$$

$$Var[X] = \frac{1}{\lambda^2} = \mu^2.$$

7.7.8) Assume length of phone call in mins, $\lambda = 1/10$. If a customer arrives at a phone booth just before you arrive, find the prob that you have to wait

a) less than 5 min b) greater than 10 mins

c) btw 5 & 10 mins

also compute the Expectation & variance

→ $\lambda = 1/10$ $\mu = 10$. $\lambda = 0.1$.

cdf $1 - e^{-\lambda x}$

a) $p(X \leq 5) = 1-e^{-\lambda x}$ $p(x=1) + p(x=2) +$

$p(x=3) + p(x=4)$.

$= 1-e^{-0.1(1)} + 1-e^{-0.1(2)} + 1-e^{-0.1(3)} +$

$1-e^{-0.1(4)}$

$= 0.864$.

b) $p(x > 10) = 0.3678$

c) 0.966

7) Calls to a customer service are to a poisson dist with an avg of 6 calls per min. From any initial point of observation, let x denote the time until the 1st call.

a) Give the dist of X, its pdf and CDF?

b) What is prob that 1st call answer

i) within 15 sec

ii) b/w 6 to 12 sec from init observation

iii) atleast 10 sec after init observation

c) avg & std deviation of time needed for 1st call?

d) what amount of time t is such that 1st call answer within $t \leq 90\%$ of time?

$\Rightarrow \quad \lambda = \frac{1}{6}$

$$pdf = \lambda e^{-\lambda x}$$

$$= \frac{1}{6} e^{-x/6}$$

b)

i) 0.777

ii)

iii)

Uniform Distribution :-
(Rectangular Dist)

Relaiability and failure rate:

Consider a fixed no. of identical components.

Instantaneous failure Rate

(active Hazard)

$f(t)$ & $h(t)$
failure density   failure Rate

$R(t)$

1) $h(t) = \lambda_0 t \quad \lambda_0 > 0$

$h(t) = \lambda_0 t^{1/2}$

2) $h(t) = \alpha u t^{\alpha - 1} + \beta \gamma^{\beta - 1}$

Hypoexponential Distribution :-

$$F(t) = 1 - \frac{\lambda_2}{\lambda_2-\lambda_1} e^{-\lambda_1 t} + \frac{\lambda_1}{\lambda_2-\lambda_1} e^{-\lambda_2 t} \; ; \; t \geq 0. \quad \text{— (2)}$$

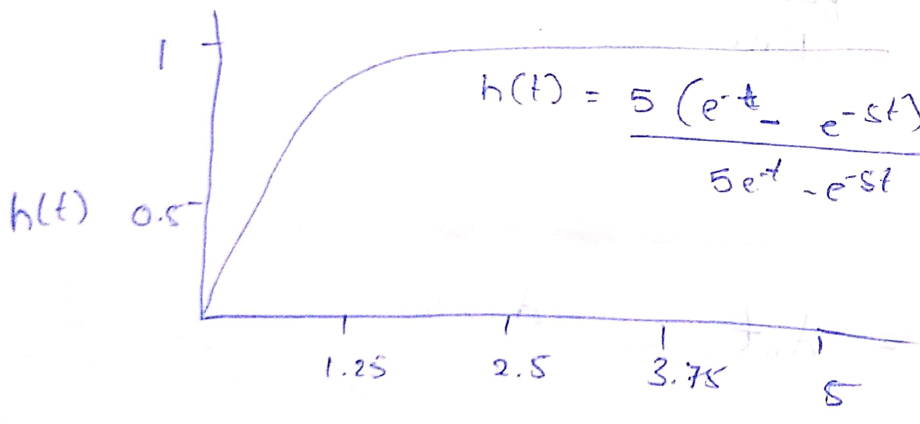
cdf func for HD.

$$f(t) = \frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1} \left(e^{-\lambda_1 t} - e^{-\lambda_2 t} \right), \quad t > 0 \quad \text{— (1)}$$

* Rate of 1 phase is not equal to rate of other phases.

Hazard Rate :-

$$h(t) = \frac{\lambda_1 \lambda_2 (e^{-\lambda_1 t} - e^{-\lambda_2 t})}{\lambda_2 e^{-\lambda_1 t} - \lambda_1 e^{-\lambda_2 t}} \quad \text{— (5)}$$



$$h(t) = \frac{5(e^{-t} - e^{-5t})}{5e^{-t} - e^{-5t}}$$

Erlang & Gamma Distribution :-

↳ $$f(t) = \frac{\lambda^r t^{r-1}}{(r-1)!} e^{-\lambda t}, \quad t > 0, \; \lambda > 0$$

$r = 1, 2, \ldots$

— (4)

The CDF is,

$$F(t) = 1 - \sum_{k=0}^{r-1} \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad t>0, \lambda>0, r=1,2,\ldots \quad (5)$$

also,

$$h(t) = \frac{\lambda^r t^{r-1}}{(r-1)! \sum_{k=0}^{r-1} \frac{(\lambda t)^k}{k!}}$$

Hyperexponential Distribution :-

$$f(t) = \sum_{i=1}^{k} \alpha_i \lambda_i e^{-\lambda_i t}, \quad t>0, \lambda_i>0, \alpha_i>0$$

$$\sum_{i=1}^{k_i} \alpha_i = 1 \quad (10)$$

and the CDF is,

$$F(t) = \sum_i \alpha_i (1 - e^{-\lambda_i t}), \quad t \geq 0 \quad (11)$$

The failure rate is,

$$h(t) = \frac{\sum \alpha_i \lambda_i e^{-\lambda t}}{\sum \alpha_i e^{-\lambda_i t}}, \quad t \geq 0 \quad (12).$$

has a DFR from $\sum \alpha_i \lambda_i$ down to min $\{\alpha_1, \alpha_2, \ldots\}$

Weibull Distribution :-

The density fn is:

$$f(t) = \lambda \alpha t^{\alpha - 1} e^{-\lambda t^\alpha} \quad (13)$$

The CDF is:

$$F(t) = 1 - e^{-\lambda t^{\alpha}} \quad \text{—— (14)}$$

failure rate is,

$$h(t) = \lambda \alpha t^{\alpha - 1} \quad \text{—— (15)}.$$

2D Random Variables :-

Conditional pmf :-

$$P_{Y|X}(y|x) = P(Y=y \mid X=x) \quad (5.1)$$

$$= \frac{P(Y=y,\ X=x)}{P(X=x)} = \frac{P(x,y)}{P_x(x)} \quad (5.2)$$

X & Y are discrete random var, having a joint pmf (x,y) The conditional pmf of Y given X is.

$$p(x,y) = P_x(x)\, P_{y|x}(y|x) = P_y(y)\, P_{x|y}(x/y) \quad \text{—— (5.3)}$$

↳ If x & y are not independent.

If they are independent,

$$P_{y/x}(y|x) = P_y(y). \quad \text{—— (5.4)}$$

marginal probability is,
for y,

$$P_y(y) = \sum_{\text{all } x} p(x,y) = \sum_{\text{all } x} P_{y|x}(y|x) P_x(x) \quad \text{—— (5.5)}$$

for x,

$$P_x(x) = \sum_{\text{all } y} p(x,y) = \sum_{\text{all } y} P_{x/y}(x/y)\, P_y(y).$$

Conditional distributio for $F_{y/x}(y/x)$.

$$F_{y/x}(y/x) = P(Y \leq y \mid X = x) = \frac{P(Y \leq y \text{ and } X = x)}{P(X = x)} \quad \text{— } (5.6)$$

CDF for $F_{x/y}(x/y)$.

$$F_{x/y}(x/y) = P(X \leq x \mid Y = y) = \frac{P(X \leq x \text{ and } Y = y)}{P(Y = y)}.$$

Conditional pdf:-

$$f_{y/x}(y/x) = \frac{f(x,y)}{f(y)}, \text{ if } 0 < f_x(x) < \infty$$

It follows the def of conditio density that

$$f(x,y) = f_x(x)\, f_{y/x}(y/x) = f_y(y)\, f_{x/y}(x,y)$$

$\rightarrow$ X & Y are not independent.

$$f(x,y) = f_x(x)\, f_y(y) \quad \text{— } (5.10)$$

$$f_{y/x}(y/x) = f_y(y)$$

$\rightarrow$ X & Y are independent rand var

$$\cancel{f_{x/y}}(x/y) = \frac{f(x,y)}{\cancel{f_y(y)}} \quad \text{— } (5.12)$$

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

$$= \int_{-\infty}^{\infty} f_X(x)\, f_{Y|X}(y|x)\,dx \quad \text{— (5.11)}$$

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$$

$$f_{X|Y}(x|y) = \frac{f_X(x)\, f_{Y|X}(y|x)}{\int_{-\infty}^{\infty} f_X(x)\, f_{Y|X}(y|x)\,dx} \quad \text{— (5.1}$$

$$P(a \le Y \le b \mid X = x) = \int_a^b f_{Y|X}(y|x)\,dy, \quad a \le b \quad \text{— (5.}$$

$$F_{Y|X}(y|x) = P(Y \le y \mid X = x) = \frac{\int_{-\infty}^{y} f(x,t)\,dt}{f_X(x)}$$

$$= \int_{-\infty}^{y} f_{Y|X}(t|x)\,dt \quad \text{— (5.14)}$$

Until 5.14 is important for exams

Exercises

1) Joint pmf of X & Y is given

a) $P(X \le 1)$.

$$P(X \le 1) = P(X=0) + P(X=1)$$

$$= 8/32 + 20/32 = 28/32 = 7/8$$

b) $P(Y \le 3) = P(Y=1) + P(Y=2) + P(Y=3)$

$$= 3/32 + 3/32 + 4/64 = 23/64$$

c) $P(X \le 1, Y \le 3) = P(0,1) + P(0,2) + P(0,3) + P(1,1) + P(1,2) + P(1,3)$

$$= 0 + 0 + 1/32 + 1/16 + 1/16 + 1/8 = 9/32$$

d) $P(X \le 1 \mid Y \le 3) = \dfrac{P(X \le 1,\ Y \le 3)}{P(Y \le 3)}$

$= \dfrac{9/32}{23/64} = 9/32 \times 64/32 = 18/23$

	1	2	3
1	1/12	1/6	0
2	0	1/9	1/5
3	1/18	1/4	2/15

	1	2	3	p(x)
1				1/4
2				14/45
3				79/180
p(x)	5/36	19/36	1/3	1

e) $P(Y \le 3 \mid X \le 1) = \dfrac{P(X \le 1,\ Y \le 3)}{P(X \le 1)}$

$= \dfrac{9/32}{7/8} = 9/32 \times 8/7 = \dfrac{9}{28}$

f) $P(X+Y \le 4) = P(0,1) + P(0,2) + P(0,3) + P(0,4)$
$+ P(1,1) + P(1,2) + P(1,3) + P(2,1) + P(2,2)$

$= 13/32$

2)

	1	2	3
1	1/12	1/6	0
2	0	1/9	1/5
3	1/18	1/4	2/15

→ Ans

	1	2	3	p(x)
1				1/4
2				14/45
3				79/180
p(x)	5/36	19/36	1/3	1

Joint pdf for discrete rand var (X, Y).

a) Evaluate the marginal dist of X & Y

b) find the cond dist^n of X given Y = 2

c) " " " " " Y given X = 3.

Marginal Distribution of X and Y is:

Ex-3) The joint pdf of 2 rand var X & Y

$$f(x,y) = \begin{cases} \frac{1}{8}x(x-y) & , 0 < x \leq 2 ; -x < y < x \\ 0 & \text{otherwise} \end{cases}$$

find $f(y/x)$.

$\Longrightarrow$

from cdf of Y

$$f_{XY}(y/x) = \frac{f_{XY}(x,y)}{f_X(x)}$$, $f_X(x)$ is marginal density fn of X.

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

$$= \int_{-x}^{x} \frac{1}{8} x(x-y)\,dy.$$

$$= \frac{1}{8}\left(x^2y - \frac{xy^2}{2}\right)_{-x}^{x} = \frac{1}{8}\left(x^3 - \frac{x^3}{2} + x^3 - \frac{x^3}{2}\right)$$

$$= x^3/8 \quad , 0 < x < 2.$$

$$f_{XY}(y/x) = \frac{\frac{1}{8}x(x-y)}{x^3/8} \quad , 0 < x < 2 ; -x < y < x$$

$$f(y/x) = \begin{cases} \frac{x-y}{x^2} & , -x < y < x \\ 0 & ; \text{otherwise} \end{cases}$$

4) The joint pdf of rand var (X, Y) is given by

$$f(x,y) = kxy\,e^{-(x^2+y^2)} \quad ; x>0, y>0$$

Find the value of k & P.T X, Y are independent.

$\Longrightarrow$ The range space is the entire 1st quadrant of xy-plane

By the prop of joint pdf,

$$\int_{x>0}\int_{y>0} kxy\, e^{-(x^2+y^2)}\, dx\, dy = 1.$$

$$k\int_0^\infty\int_0^\infty y e^{-y^2}\, x e^{-x^2}\, dx\, dy = 1.$$

put $x^2 = t$ so $2x\,dx = dt$.

then $k/2 \int_0^\infty y e^{-y^2} x e^{-x^2}\, dt\, dy = 1.$

$$k/2 \int_0^\infty y e^{-y^2} \left(e^{-t}/_{-1}\right)_0^\infty dy = 1.$$

$$k/2 \int_0^\infty y e^{-y^2} (1)\, dy = 1.$$

put $y^2 = v$; so $2y\,dy = dv$

then $k/2 \cdot 1/2 \int_0^\infty e^{-v}\, dv = 1.$

$$k/2\; 1/2 \left(\frac{e^{-v}}{-1}\right)_0^\infty = 1$$

$$k/2 \cdot 1/2\, (1) = 1.$$

$$k/4 = 1 \Rightarrow \boxed{k=4}.$$

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

$$= 4x \int_0^\infty y e^{-(x^2+y^2)}\, dy.$$

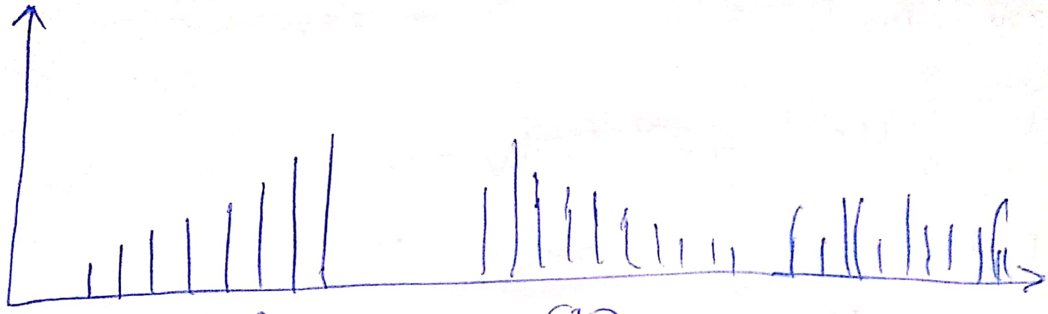
$$= 4x e^{-x^2} \int_0^\infty y e^{-y^2} dy$$

$$= 4x e^{-x^2} \int_0^\infty e^{-t}\, 1/2\, dt.$$

$$= 4x e^{-x^2} \left[\frac{e^{-t}}{-1}\right]_0^\infty$$

$$= 2x e^{-x^2}, \quad x>0$$

Normal Distribution :- [Gaussian or Z]

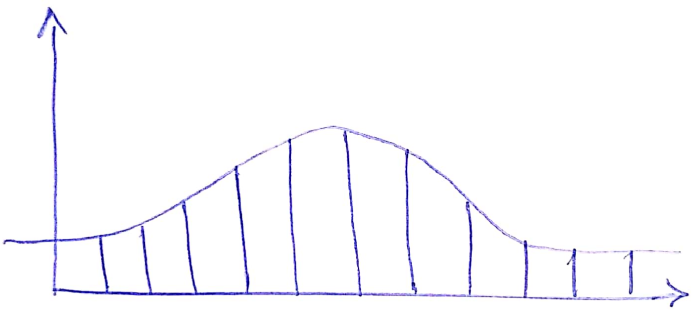


(a) (b) (c)

a) Spread out more on right.

b) " " " " left.

c) Jumbled - up.

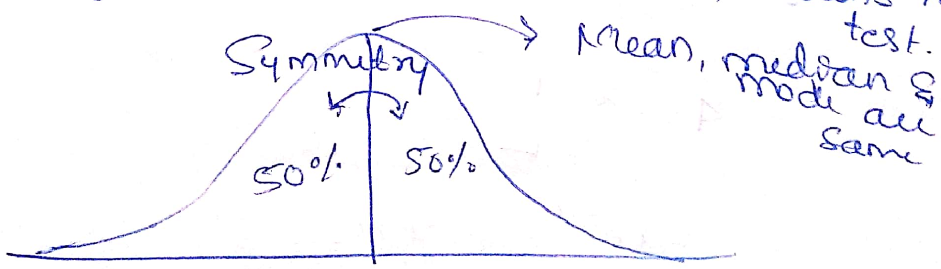


(d)

Data tends to be around central value with no bias left or right, and it gets close to a 'normal distribution'.

Normal distⁿ eg:

Heights of people, blood press., marks in a test.

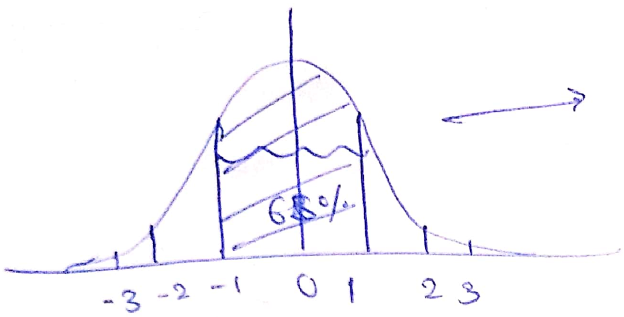


50% values of are < mean

50% values are > mean

Standard deviations :-

It is a measure of how spread out members are



34 on left, 34 on right
↑
68% of values are within 1SD of the mean
or
95% - 2 SD
99.7% - 3 SD

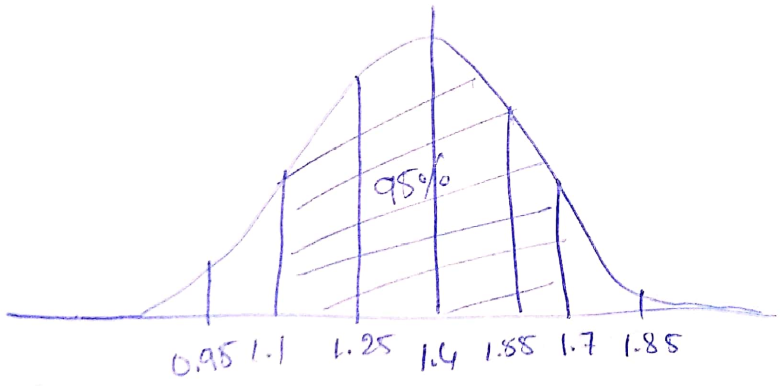
Prob

1) 95% of stud are b/w 1.1m & 1.7m tall data is norm dist. What is average & S.D?

⇒ mean = $(1.1+1.7)/2 = 1.4m$

95% of 2 SD either side of the mean (a total of 2 SD)

$1 SD = (1.7 - 1.1) \; 4 = 0.15m$



* likely to be within 1SD (68 out of 100)
* Very likely to be within 2SD (95/100)
* Certainly to be within 3SD (99.7/100)

The no. of SD from the mean is also called the "standard score" or "Sigma" or 'Z-score'