Spearman Rank correlation

formula = $r_s = \dfrac{Cov(R_X R_Y)}{\sigma R_X \, \sigma R_Y}$

Eg.

Ecogrowth X	SO index Y
2.1	8
2.5	12
3.6	10
4.0	14

In spearman rank correlation we put rank of given data.

	Rank x		Rank y
2.1	4	8	4
2.5	3	12	2
3.6	2	10	3
4.0	1	14	1

$$\text{mean } x = \frac{4+3+2+1}{4}$$

$$= 2.5$$

$$\text{mean } y = \frac{4+3+2+1}{4}$$

$$= 2.5$$

$(x-\bar{x})$		$(x-\bar{x})$	$y-\bar{y}$	$(y-\bar{y})$
$4-2.5$	=	1.5	$4-2.5$	1.5
$3-2.5$	=	0.5	$2-2.5$	-0.5
$2-2.5$	=	-0.5	$3-2.5$	$\cdot 0.5$
$1-2.5$	=	-1.5	$1-2.5$	-1.5

$$\rho_{(x,y)} = \sum_{i=1}^{n} \frac{(x-\bar{x})(y-\bar{y})}{n-1}$$

$$\Rightarrow \frac{(1.5)\times(1.5)+(0.5)\times(-0.5)+(-0.5)\times(0.5)+(-1.5)\times(-1.5)}{4-1}$$

$$\Rightarrow \frac{2.25 - 0.25 - 0.25 + 2.25}{3}$$

$$\Rightarrow \frac{4}{3} \Rightarrow 1.33$$

$$S.D \quad \sigma_X = \sqrt{\frac{(1.5)^2 + (0.5)^2 + (-0.5)^2 + (-1.5)^2}{4-1}}$$

$$= \sqrt{1.66} \Rightarrow 1.288$$

$$SD. \quad \sigma_Y = \sqrt{\frac{(1.5)^2 + (-0.5)^2 + (0.5)^2 + (-1.5)^2}{4-1}}$$

$$= \sqrt{1.66} = 1.288$$

$$r_s = \frac{R_X \cdot R_Y}{\sigma R_X \, \sigma R_Y}$$

$$\Rightarrow \frac{1.33}{(1.288)(1.288)} \Rightarrow 0.81 = \boxed{81\%}$$

* Probability Distribution.

1. Discrete pr. Dist.
2. Continuous pr. Dist.

1. Discrete pro. Dist.

1. Bernoulli Dist
2. Binomial Dist
3. Poisson Dist

1. Bernoulli Dist

no. of experiment = 1
no of outcome = fix

pass = p
fail = $1-p$

(ii) Binomial Dist.

no of experiment = fix time 10, 20,

output = fix

pass = p

fail = $1 - p$

(iii) Poisson Dist.

for Example :- A small busines recieve an avg 12 customer per day

① what is the probability that the business will recieve exact 8 custome in a day?

Soln :- $\mu = 12$ $x = 8$

Formula =

$$P_r(X = x) = \frac{\mu^x e^{-\mu}}{x!}$$

$$P_r(X = 8) = \frac{12^8 e^{-12}}{8!}$$

$e = 2.718$

$8! = 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$

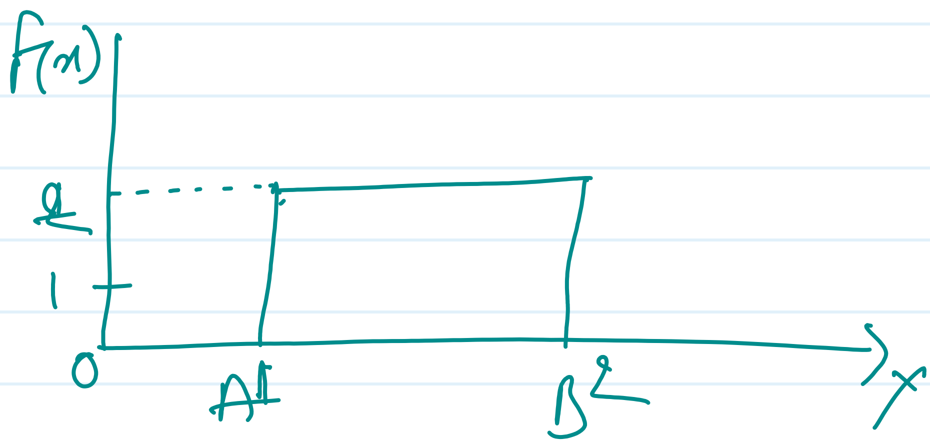
$\Rightarrow 0.0655$

$\Rightarrow 6.55\%$

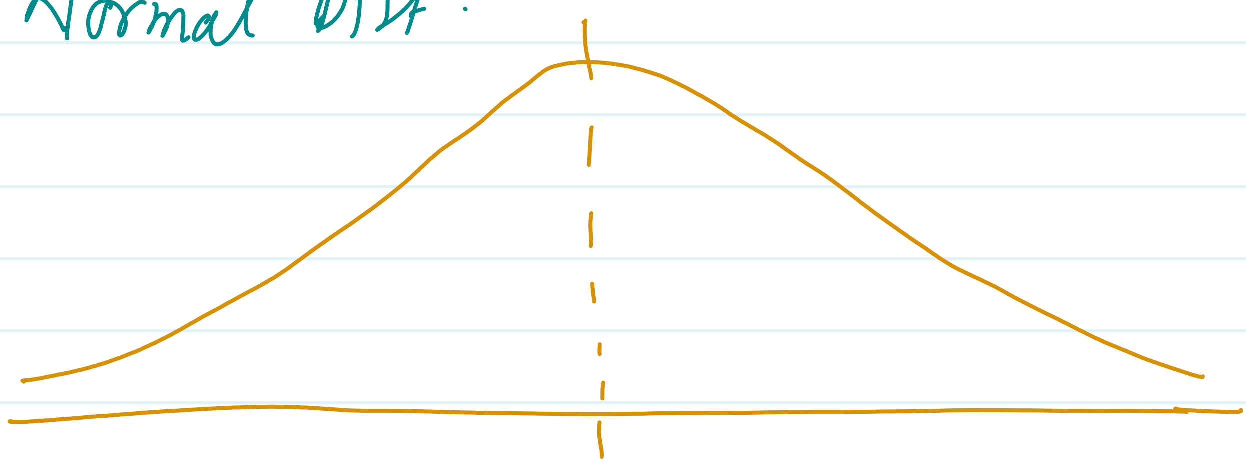
(11) Continous probability Dist.

1. Uniform Dist.
2. Normal Dist
3. Standard normal Dist
4. Log normal Dist.
5. Exponential Dist.

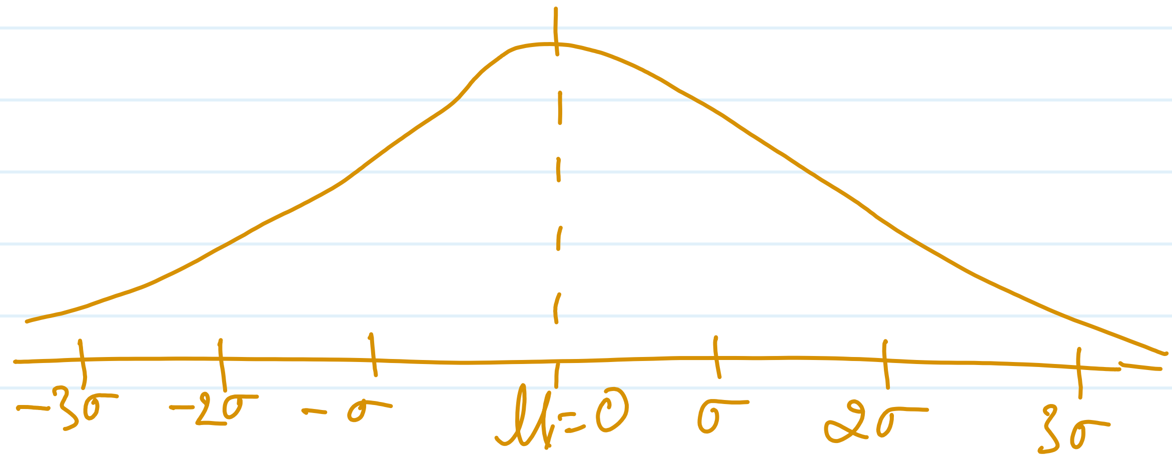
(1) Uniform Dist



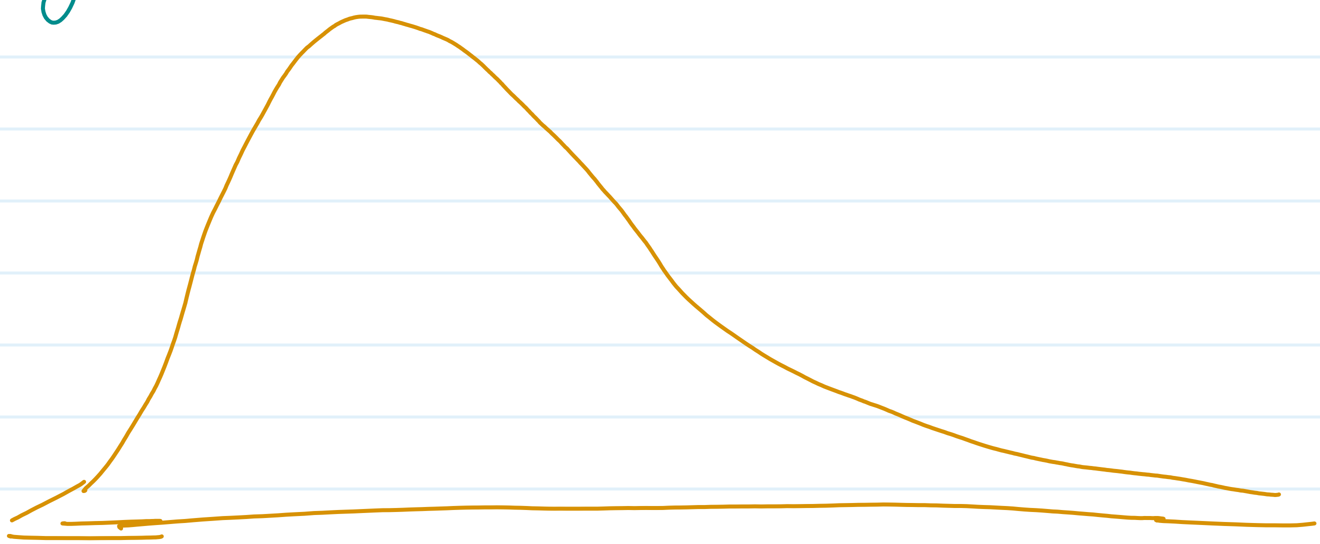
(2) Normal Dist.



③ standard normal Dist.

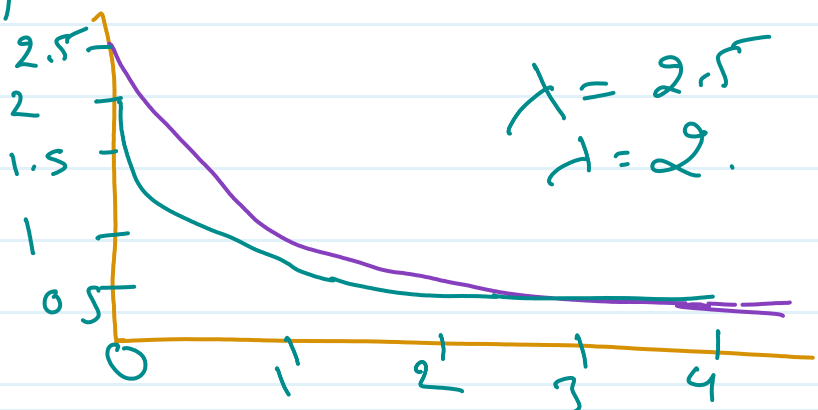


④ Log normal Dist.



mean > median > mode

⑤ Exponential Dist



* Chebyshev's Inequality

$X \approx$ Gaussian Dist

$$Pr(\mu - \sigma \le X \le \mu + \sigma) = 68\%$$

$$Pr(\mu - 2\sigma \le X \le \mu + 2\sigma) = 95\%$$

$$Pr(\mu - 3\sigma \le X \le \mu + 3\sigma) = 99.7\%$$

$Y \not\approx$ Gaussian Dist

$$Pr(\mu - K\sigma \le Y \le \mu + K\sigma) \ge 1 - \frac{1}{K^2}$$

$\because$ K is SD. value of K will be 2 or more than 2

If $K = 2$

$$Pr(\mu - 2\sigma \le Y \le \mu + 2\sigma) \ge 1 - \frac{1}{2^2}$$

$$\ge 1 - \frac{1}{4}$$

$$\ge \frac{4-1}{4} = \frac{3}{4} = 75\%$$

If $K = 3$

$$P_r(\mu - 3\sigma \leq Y \leq \mu + 3\sigma) \geq 1 - \frac{1}{3^2}$$

$$\geq 1 - \frac{1}{9}$$

$$\geq \frac{9-1}{9} \Rightarrow \frac{8}{9} = 88.9\%$$

If $K = 4$

$$P_r(\mu - 4\sigma \leq Y \leq \mu - 4\sigma) \geq 1 - \frac{1}{4^2}$$

$$\geq \frac{15}{16} \Rightarrow 93\%$$