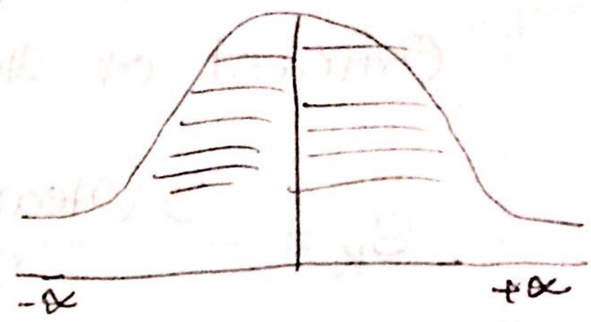
Skewness|

দুই দিকে সমান না হলে

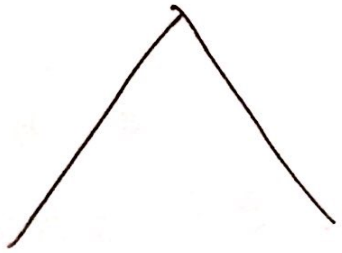
~~Skewne~~ Skewed distribusion



Symmetric distribusion
or,
Equal "

দুই দিকে সমান হলে
symmetric distru

Skewed distribusion

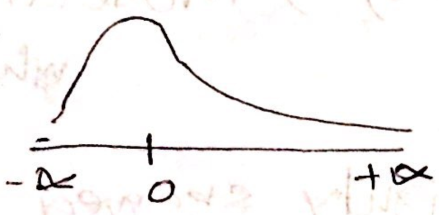


Possitive negative

Possitive: Possitivec দিকে ঝোঁকার tendency দেখা যাবে

negative: negative দি " " " " "

Positive:



negative

c.ofs এর মানগুলো negative এবং Positive দুটোই হতে পারে,

Coficient of skewness :

$$S_K = \frac{3(\text{mean} - \text{median})}{S.D}$$

$\boxed{S_K = (-3 \text{ to } +3)}$ এর মান Am হবে।

if $S_K = 0$; Symmetric dist.

$S_K = (-0.5 \text{ to } 0.5) \rightarrow$ Almost symmetric

$S_K > (0.5 \text{ to } 1) \rightarrow$ Moderately skewed which is positive

$S_K (-1 \text{ to } < -0.5) \rightarrow$ Moderately skewed which is negative

$S_K (< -1) \rightarrow$ Highly skewed which is neg

$K > (1) : \rightarrow$ " " which is positive

Q| ~~D,~~ 10, 2, 11, 20, 1 — Interprete

Calculate Coeff of Skewen & (commt or

or,

Comment on the shape of the distribution

Med = 10

Mean = 8.8

~~S.D~~

Standert Division

$$S.D = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

$$= \sqrt{\frac{(10-8.8)^2 + (2-8.8)^2 + (11-8.8)^2 + (20-8.8)^2 + (1-8.}{5-1}}$$

$=$ ______

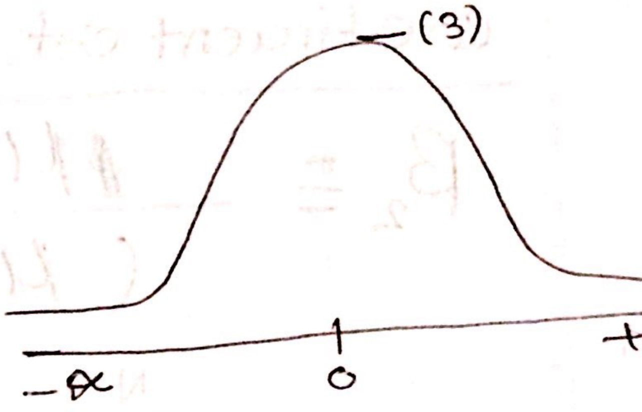
$$S_K = \frac{3(8.8 - 10)}{7.73}$$

$$= -0.47$$

Almost Symtire

Kurtosis

1. Lepto Kurtic
2. Meso Kurtic
3. Platy kurtic



Height 3 হলে Meso Kurtic

Height 3 এর চেয়ে বেশি হলে Lepto Kurtic

ঐ " 3 " " কম " Platy Kurt

μ = moments

$\cancel{\mu}$ μ কে moment বলা হয়

Coefficient of kurtosis:

5, 2, 8

$\bar{x} = 5$

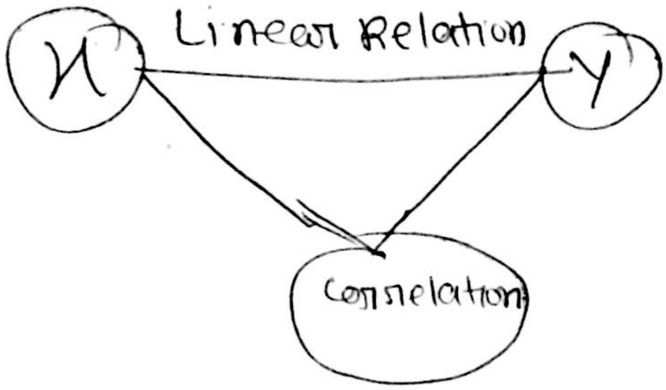
$$\beta_2 = \frac{\mu_4}{(\mu_2)^2}$$

$$\mu_r = \frac{\sum_{i=1}^{N} (x_i - \bar{x})^r}{N}$$

$$\mu_4 = \frac{(5-5)^4 + (2-5)^4 + (8-5)^4}{3}$$

$$=$$

$$\mu_2 = \frac{(5-5)^2 + (2-5)^2 + (8-5)^2}{3}$$



* দুটি variable এর মধ্যে Linear Relation থাকলে তাকে Correlation বলে।

* যখন variable এর মান ~~[illegible]~~ change হলে অন্যটির মান ~~[illegible]~~ change হলে Linear Relation

* যখনটির মান change হলে অন্যটির opposite change হলে negative correlation

Correlation

Karl pearson correlation $r = \dfrac{\sum_{i=1}^{n}\left[(x_i-\bar{x})(y_1-\bar{y})\right]}{\sqrt{\sum_{i=1}^{n}(x_1-\bar{x})^2 \cdot \sum_{i=1}^{n}(y_i-\bar{y})^2}}$

r এর value $(-1 \text{ to } +1)$ এর মধ্যে থাকবে

$r = 0$

Correlation

Correlation: A correlation is a linear relationship between two variables.

Coefficient of correlation

Karl Pearson's coefficient of correlation

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$$

Properties

* -1 indiating a perfect negative relationship

* 0 indicating no relationship

* The size of the correlation indicate the strength of the relationship;
  Ex. -0.89 indicates a stronger relationship than a cofficient of +0.60

* The closer to 1, the stronger positive linear relations

* The closer to 0, The weake any positie " "

Cofficient Range	Strength of Relationship
0.00 - 0.20	Very low
0.21 - 0.40	Low
0.41 - ~~6~~0.60	Moderate
0.61 - 0.80	High Moderate
0.81 - 1.00	Very High

$$r = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(Y_i-\bar{Y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2 \sum_{i=1}^{n}(Y_i-\bar{Y})^2}}$$

$\bar{X} = 4.3$

$\bar{Y} = 3.7$

X Supply	Y Demand	$x_i-\bar{x}$	$Y_i-\bar{Y}$	$(x_i-\bar{x})(Y_i-\bar{Y})$	$(x_i-\bar{x})^2$	$(Y_i-\bar{Y})^2$
5	2	$5-4.3=0.7$	$2-3.7=-1.7$	-1.19	0.49	2.89
6	4	$6-4.3=1.7$	$4-3.7=0.3$	0.51	2.89	0.90
2	5	$2-4.3=-2.3$	$5-3.7=1.3$	-2.99	5.29	1.69
				-3.67	8.67	5.48

$$r = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(Y_i-\bar{X})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2 \sum_{i=1}^{n}(Y_i-\bar{Y})^2}}$$

$$= \frac{-3.67}{\sqrt{8.67 * 5.48}}$$

weaker any positive linear relationship

$$= \frac{-3.67}{6.9} = -0.53$$

Scanned by CamScanner

Regression Analysis

① Simple R.A (two variables) → I Indep vari (Edu) → Dep. vari (salary)

② Multiple R.A (More than two) → More then → Indep [Edu + Ex]

1. Dependent var [salary]

Simple R.A

• Estimated simple linear Regression education Model/line

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

ind vari

Dep vari

Regression Parameter

$$\hat{\beta}_1 = \text{Slope} = \frac{\sum_{i=1}^{n} [(x_i - \bar{x})(y_i - \bar{y})]}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

$$\beta_0 = \text{Intercept} = \bar{y} - \hat{\beta}_1 \bar{x}$$

$\bar{x} = 13.5$
$\bar{y} = 12.5$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}[(x_i-\bar{x})(y_i-\bar{y})]}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\hat{y} = \hat{\beta}_0 + \hat{\beta} x_i$$

Depen (y) inde (x)

Income	Expenditure
20	18
7	6
15	15
12	11

$\bar{x} = 13.5$ 12.5

Q: Determine the Regression line/on the base of expenditure.

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}[(x_i-\bar{x})(y_i-\bar{y})]}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

$$= (20-13.5)(18-12.5) + (7-13.5)(6-12.5) + (15-13.5)(15-12.5) + (12-13.5)(11-12.5)$$

$$= 84$$

Scanned by CamScanner

Skewness |

$$S_K = \frac{3(\text{Mean} - \text{Median})}{\text{Standerd Division}}$$

$$\left[\; S.D = \sqrt{\frac{\sum_{i=1}^{n} (x_1 - \bar{x})^2}{n-1}} \;\right]$$

~~Coe~~ Kurtosis |

$$\beta_2 = \frac{\mu_4}{(\mu_2)^2}$$

$$\left[\; \mu_{\cancel{a}p} = \frac{\sum_{i=1}^{N} (x_1 - \bar{x})^p}{N} \right.$$

Correrelation |

$$\rho = \frac{\sum_{i=1}^{n} \left[(x_1 - \bar{x})(Y_i - \cancel{a}\bar{Y}) \right]}{\sqrt{\sum_{i=1}^{n} (x_1 - \bar{x})^2 \sum_{i=1}^{n} (Y_1 - \cancel{a}\bar{Y})^2}}$$

Regression analysis |

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

$$\hat{\beta}_1 = \frac{\sum \left[(x_1 - \bar{x})(Y_1 - \bar{x}) \right]}{\sum (x_1 - \bar{x})^2}$$

$$\hat{\beta}_0 = \beta \bar{Y} * \hat{\beta}_1 \bar{x}$$

$$\cancel{S_{xy} = \sqrt{\frac{\sum (x_1 - \bar{x})(Y_1 - \bar{x})}{n-2}}}$$

$$S_{xy} = \sqrt{\frac{\sum (Y_1 - \cancel{a}\bar{Y})^2}{n-2}}$$