7/4/22
9:30pm

– Regression

Predicting Output which is Continous

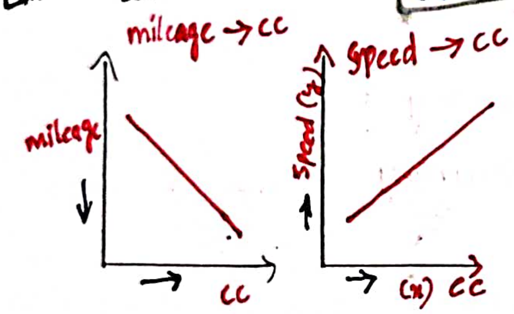
Regression

- Linear Regression
- Non-Linear Regression

("Assuming the relation between x & y is Linear")

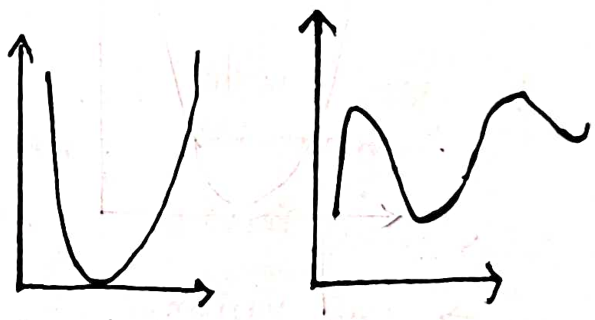
Any line other than straight line is Non-Linear

Ex:- In The Form of $\boxed{(y = mx + c)}$



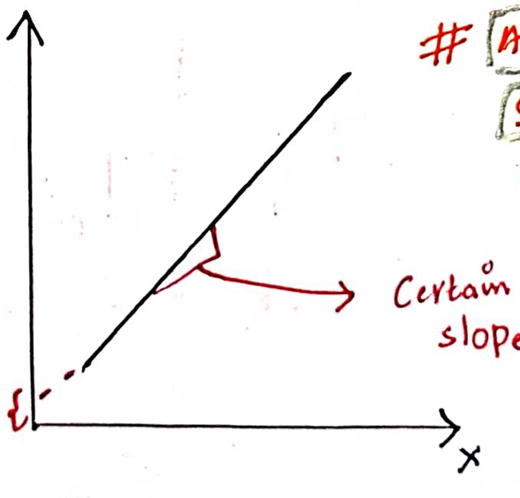
* mileage decreases ↓
cc increases ↑

* cc increases
* speed increases



Quadratic equation

Any relation in terms of Straight Line, it is called as "Linear Regression"



x value ↑, y value ↑
$\boxed{(+ve\ slope)}$

x value ↑, y value ↓
$\boxed{(-ve\ slope)}$

Ex:-

X	Y
1	15
2	25
3	35
4	45
5	55
6	? 35
7	35
8	35.

In Early Ages, They took "Avg" values.

Like

$$\frac{15+25+35+45+55}{5} = 35$$

→ They say, Avg Line is Best Fit Line.

* Best rants

tips (waiter)

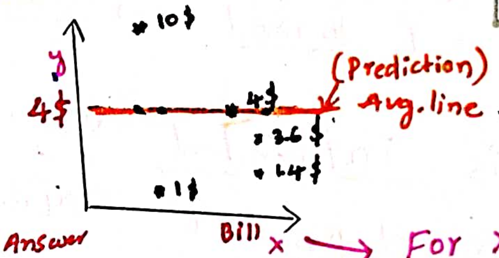
Bill amount (not given)

So, For only One Variable /column - (Avg)

1. 10 $ –
2. 1 $ –
3. 4 $ –
4. 3.6 $ –
5. 1.4 $

6 → $\boxed{\text{Avg}}$ 4$

7 → 4$ Every time Answer is 4$ (because of Avg line)



→ For x value [Predic
Va
c

For $\boxed{\text{Single Variable}}$

Best Fit line is

$\boxed{\text{Average line}}$

H_0 : Average line is best fit line

H_1 : Regression line > Avg. line.

Y	avg.line Y - Prediction	Error	error^2
10	4	6	36
1	4	-3	9
4	4	0	0
3.6	4	-0.4	0.16
1.4	4	-2.6	6.76
		0	51.92

Linear Regression

Simple Linear Regression → 1 input Variable

Multiple Linear Regression → >1 (more than input Variables)

Simple Linear Regression :

→ In Simple Linear regression, We $\boxed{\text{predict the Value}}$ of One Variable $\boxed{\text{"Y"}}$ based on Other Variable $\boxed{\text{"X"}}$

→ "X" is independent Variable (or) input variable (or) Exploratory variable

$\boxed{\text{"y"}}$ is dependent Variable (or) output variable (or) response variable

Ex:-

Weight	CC	Speed	Mileage → O/P
independent	in.var	in.var	(it depends on, weight, cc, Speed)

Que :- Why it is Simple ?

Ans :- Because, it Examines $\boxed{\text{Relation ship}}$ between $\boxed{\text{two}}$ $\boxed{\text{Variables Only.}}$

X	Y

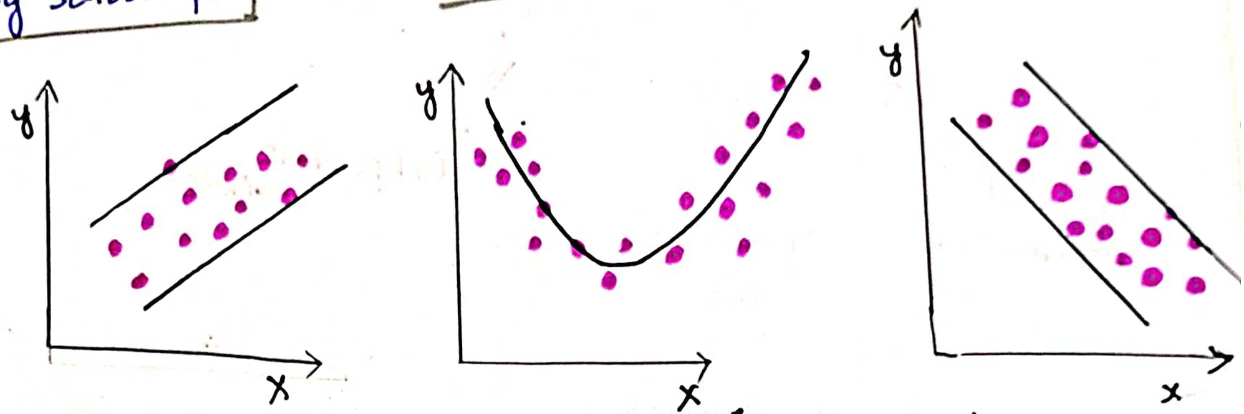
$y = x$

Que :- Why linear ?

Ans :- When the $\boxed{\text{independent Variable increases}}$ (or Decreases) the $\boxed{\text{depedent Variable}}$ Increases (or Decreases) in a linear Fashion.

Que: How to check Linear Regression ? (or) not

Ans:- $\boxed{\text{By scatter plot}}$ — Linearity, direction



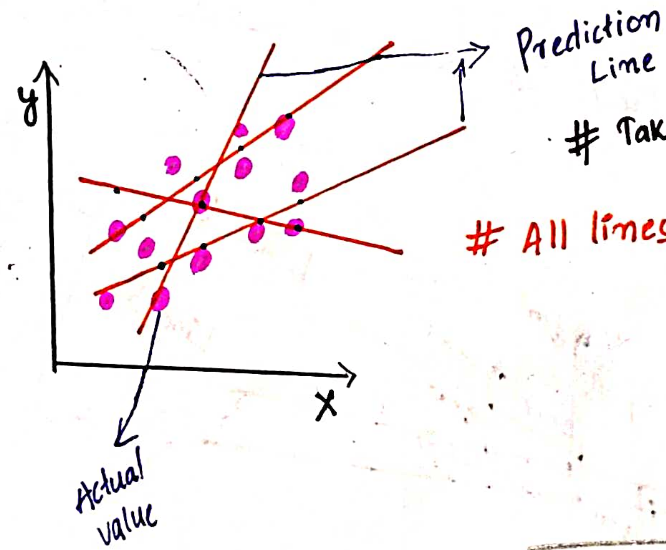
Linear — As "x" value increasing "y" value also increasing $\boxed{\text{" Positive Direction"}}$

Quadratic line

Linear But down ward direction — As "x" value increasing "y" value also decreases $\boxed{\text{"Negative direction"}}$

Que:- Which Line to be Considered ?

Ans:-



Taking Relation → linear

All lines are straight line

Que:- Different notations ? But $\boxed{\text{Same Concept}}$?

Ans:-

$y = mx + c$ — m: slope, c: "y" intercept

$y = ax + b$ — a: slope, b: "y" intercept

$y = b_0 + b_1x$ — b_0: y Intercept, b_1: slope

$y = \beta_0 + \beta_1 x$ — β_0: "y" Intercept, β_1: slope

Que:- What is Standard notation ?

Ans :-

$$y = \underbrace{\beta_0}_{\text{"y intercept" Value}}^{\text{Constant}} + \overbrace{\beta_1}^{\text{coefficient}} * \underbrace{x_1}_{\text{Independent Variable}}$$

y → (dependent Variable)

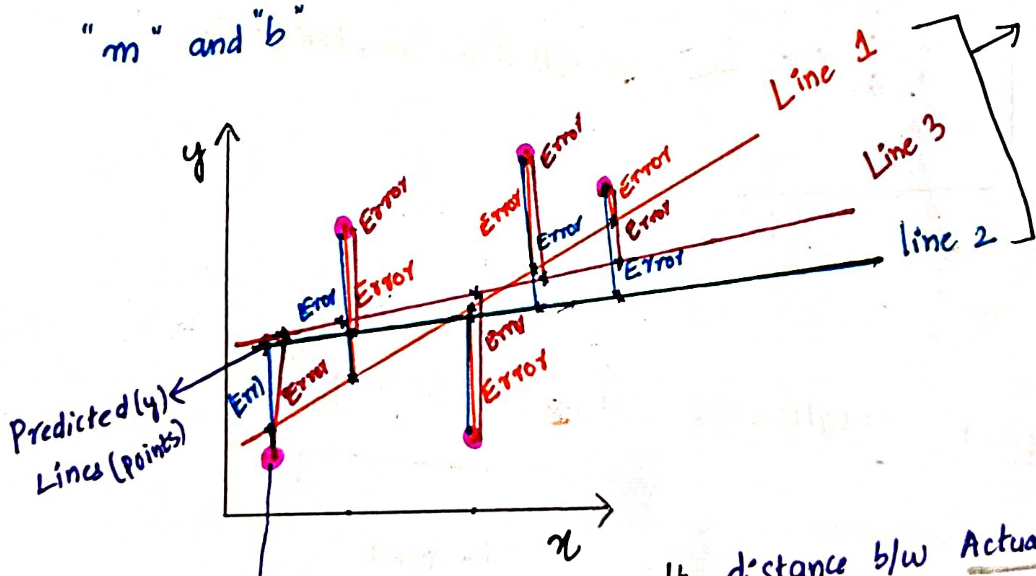
$\beta_1 * x_1$ → slope Value

Ordinary Least squares :

→ Least squares fitting is a way to find The best fit Curve (or) Line for a set of points

→ Sum of Squares of Residuals / Errors are used to estimate the best fit curve (or) Line

→ Least squares method is used to obtain The coefficient of "m" and "b"



best fit line ?

→ Where The sum of squares of Error is Minimum

distance b/w Actual data point - Predicted data point [Error]

Square The errors [each error do square]

after Squaring Error, Than Sum Error

Que :- What is SSE ?

Ans :- SSE $\Rightarrow$ Some of squares of Error

Denoted by $\boxed{\Sigma(y-\hat{y})^2}$ $\rightarrow$ squaring them = SSE_{min}

Sum

Actual Value

Predicted Value

$\Sigma(y-\hat{y})^2_{min} \rightarrow$ This value should be Minimum

* you Have identify the line such a way $\boxed{\Sigma(y-\hat{y})^2_{min}}$

Que :- How to identify The Exact line ?

Ans :- Explanation :- of Calcus

Minimum Point : $y = x^2$ (what is Least "y" value) $= 0$

$\boxed{y = 0}$

if 0.00001 also it get squared

so, "0" is Least y value

how it is identified

$\boxed{\frac{d(y)}{dx} \text{ at } x=0}$ $\Rightarrow \frac{dy}{dx}$ at $x=0$ $\Rightarrow \frac{d(x^2)}{dx}$ at $x=0$ $\Rightarrow 2x \Rightarrow 2(0) = 0$

Ex :- $y = 4x^2+5 \longrightarrow$ $4\times2=8$

$\boxed{\frac{dy}{dx} \text{ at } x=0}$ $\Rightarrow 8x+0 \Rightarrow 8(0)+0 \Rightarrow 0$

* $\frac{d}{dx}(4x^2) = 2\times4 = [8x]$

* $\frac{d}{dx}(5) = 0$

Mimimum Value $= \frac{dy}{dx}$ at $x=0$

$SSE_{min} = \Sigma[y-\hat{y}]^2$

So, in order to get minimum value For S.S.E

Calculate $\boxed{\frac{d[\Sigma[y-\hat{y}]^2]}{dx} \text{ at } x=0}$ $\boxed{\hat{y} = \beta_0 + \beta_1 x}$

$\therefore \frac{d}{dx} \Sigma(y - \beta_0 + \beta_1 x)$

Expand This Equation

$$\boxed{\beta_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}$$

$$\boxed{\beta_0 = \bar{y} - \beta_1 \bar{x}}$$

"y" Mean

"X" Mean

Example :

H_0 : Avg. line Errors ≤ Regression Line
H_1 : Regression line < Avg. line

X	Y
4	11
6	4
7	6
8	5
12	8

identifying best fit line for These Values

By using

$y = \beta_0 + \beta_1 x$ [Linear line] ✓

not in $y = \log x, e^x, ax^2, \beta x^{-2} x$

For Average Fit line

Avg line

X	Y	y-Pred avg	Error	$Error^2$
4	11	6.8	4.2	17.64
6	4	6.8	-2.8	7.84
7	6	6.8	-0.8	0.64
8	5	6.8	-1.8	3.24
12	8	6.8	1.2	1.44
				30.8

SSE

For Random Value 2+5x (Equ)

For 2+5x Equation

X	Y	y-Pred	Error	$Error^2$
4	11	22	-11	121
6	4	32	-28	784
7	6	37	-31	961
8	5	42	-37	1369
12	8	62	-54	2916
				6151

SSE

For 5+4x Equation

For 5+4x Equation

X	Y	y-P	Error	$Error^2$
4	11	21	-10	100
6	4	29	-25	625
7	6	33	-27	729
8	5	37	-32	1024
12	8	53	-45	2025
				4503

SSE

* Avg $= \frac{11+4+6+5+8}{5} = 6.8$

* Error ⇒ (Actual - Predicted Answer) y y_{pred}

$11 - 6.8 = 4.2$
$4 - 6.8 = -2.8$
$6 - 6.8 = -0.8$
$5 - 6.8 = -1.8$
$8 - 6.8 = 1.2$

* $Error^2 = (4.2)^2 = 17.64$
$(2.8)^2 = 7.84$
$(0.8)^2 = 0.64$
$(1.8)^2 = 3.24$
$(1.2)^2 = 1.44$

* Some of square of Error $= 17 + 7.84 + 0.64 + 3.24 + 1.44 = 30.8$

For, 2+5x

SSE

⇒ y Pred ⇒
$2+5(4) = 22$
$2+5(6) = 32$
$2+5(7) = 37$
$2+5(8) = 42$
$2+12(12) = 62$

⇒ Error $(y-\hat{y})$ =
$11 - 22 = -11$
$4 - 32 = -28$
$6 - 37 = -31$
$5 - 42 = -37$
$8 - 62 = -54$

⇒ $Error^2$ =
$(-11)^2 = 121$
$(-28)^2 = 784$
$(-31)^2 = 961$
$(-37)^2 = 1369$
$(-54)^2 = 2916$

⇒ SSE $= 121+784+961+1369+2916$
$= 6151$

For 5+4x

SSE

• y pred ⇒ $5+4(4) = 2$
$5+4(6) = 29$
$5+4(7) = 33$
$5+4(8) = 37$
$5+4(12) = 53$

• Error $y - y_{pred}$ = 11-21
4-29
6-33
5-31
8-53

* $Error^2$ =
$(-10)^2 = 10$
$(25)^2 = 62$
$(27)^2 = 7$
$(32)^2 = 10$
$(45)^2 = 20$

SSE = 100+625+729+
2025 =

For Regression line

$$y = \beta_0 + \beta_1 x$$

$$\therefore \beta_0 = \bar{y} - \beta_1(\bar{x})$$

$$\beta_1 = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\Sigma(x-\bar{x})^2}$$

X	Y	$x-\bar{x}$ mean	$y-\bar{y}$	$\bar{x} * \bar{y}$	$(x-\bar{x})^2$
4	11	-3.4	4.2	-14.28	11.56
6	4	-1.4	-2.8	3.92	1.96
7	6	-0.4	-0.8	0.32	0.16
8	5	0.6	-1.8	-1.08	0.36
12	8	4.6	1.2	5.52	21.16
				Σ -5.6	Σ 35.2

X Avg: 7.4

y Avg. 6.8

* Avg. of X $= \frac{4+6+7+8+12}{5} = 7.4$

* Avg. of Y $= \frac{11+4+6+5+8}{5} = 6.8$

* $X - \bar{X}$(mean) $=$
$4 - 7.4 = -3.4$
$6 - 7.4 = -1.4$
$7 - 7.4 = -0.4$
$8 - 7.4 = 0.6$
$12 - 7.4 = 4.6$

* $Y - \bar{Y}$(mean) $=$
$11 - 6.8 = 4.2$
$4 - 6.8 = -2.8$
$6 - 6.8 = -0.8$
$5 - 6.8 = -1.8$
$8 - 6.8 = 1.2$

* $\bar{x}_{mean} * \bar{y}_{mean} =$
$-3.4 \times 4.2 = -14.28$
$-1.4 \times -2.8 = 3.92$
$-0.4 \times -0.8 = 0.32$
$+0.6 \times -1.8 = -1.08$
$4.6 \times 1.2 = 5.52$

* $\Sigma \bar{x}_{mean} * \bar{y}_{mean}$
$= -14.28 + 3.92 + 0.32 - 1.08 + 5.52 = -5.6$

* $(x - \bar{x})^2$
$= (3.4)^2 = 11.56$
$(-1.4)^2 = 1.96$
$(-0.4)^2 = 0.16$
$(0.6)^2 = 0.36$
$(4.6)^2 = 21.16$

* $\Sigma(x-\bar{x})^2$
$= 11.56 + 1.96 + 0.16 + 0.36 + 21.16 = 35.2$

$$\boxed{\beta_1 = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\Sigma(x-\bar{x})^2}} \Rightarrow \frac{-5.6}{35.2} = \boxed{0.15909}$$

Slope

$$\boxed{\beta_0 = \bar{y} - \beta_1(\bar{x})} = 6.8 - (-0.159)(7.4)$$

$6.8 - (-0.159)$

$= \boxed{7.977}$

y intercept

Now, The equation is

$$y = \beta_0 + \beta_1 x$$

$$y = 7.977 - 0.15909\,[x]$$

Here $\boxed{y = 7.977273 - 0.159[x]}$ Equation.

X	Y	ypred	Err	Error²
4	11	7.34	3.66	13.39
6	4	7.02	-3.02	9.12
7	6	6.86	-0.86	0.73
8	5	6.70	-1.7	2.89
12	8	6.06	1.94	3.76
				29.89

29.89 → SSE

Equation: $7.977 - 0.159(x)$

$7.977 - 0.159(4) = 7.341$
$7.977 - 0.159(6) = 7.023$
$7.977 - 0.159(7) = 6.864$
$7.977 - 0.159(8) = 6.705$
$7.977 - 0.159(12) = 6.069$

Error

$11 - 7.34 = 3.66$
$4 - 7.02 = -3.02$
$6 - 6.86 = -0.86$
$5 - 6.70 = -1.7$
$8 - 6.06 = 1.94$

Error^2

$= (3.66)^2 = 13.39$
$(-3.02)^2 = 9.12$
$(-0.86)^2 = 0.73$
$(-1.7)^2 = 2.89$
$(1.94) = 3.78$

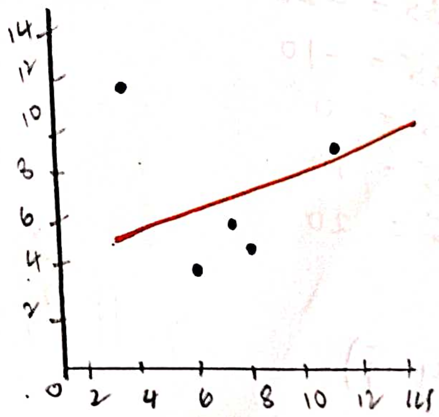
SSE $= 13.39 + 9.12 + 0.73 + 2.89 + 3.76$
$= 29.89$

Comparing with Avg. line and Other Equations

$y = 7.977 - 0.159[x]$

Best fit line.

$[SSE]_{min}$



Ex [For Interv

For, Average →

X	Y	$\bar{y}$	Error ($y-\bar{y}$)	$error^2$
1	15	35	+20	400
2	25	35	+10	100
3	35	35	0	0
4	45	35	-10	100
5	55	35	-20	400

SST = 1000

$$\therefore R^2 = 1 - \frac{0_{Reg}}{1000_{Av}}$$

$\boxed{R^2 = 1}$ ← Maximum value

For, 10x+5

X	Y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x} \times y-\bar{y})$	$(x-\bar{x})^2$	$\hat{y}$	Error	$Error^2$
1	15	-2	-20	40	-4	15	0	0
2	25	-1	-10	10	-1	25	0	0
3	35	0	0	0	0	35	0	0
4	45	1	10	10	1	45	0	0
5	55	2	20	40	4	55	0	0
				100	0			SSE=0

$\bar{x} = 3$ $\bar{y} = 35$

$$X\ Avg = \frac{1+2+3+4+5}{5} = 3$$

$$y\ Avg = \frac{15+25+35+45+55}{5} = 35$$

* $x - \bar{x} = 1-3 = -2$

$2-3 = -1$

$3-3 = 0$

$4-3 = 1$

$5-3 = 2$

* $y-\bar{y} = 15-35 = -20$

$25-35 = -10$

$35-35 = 0$

$45-35 = 10$

$55-35 = 20$

* $(x-\bar{x}) \times (y-\bar{y})$

$= -2 \times (-20) = 40$

$-1 \times (-10) = 10$

$1 \times 10 = 10$

$2 \times 20 = 40$

* $\Sigma(x-\bar{x} \times y-\bar{y})$

$= 40+10+0+10+40 = 100$

* $(x-\bar{x})^2$ # For squaring don't consider $(+,-)$

$= (-2)^2 = +4$

$(-1)^2 = +1$

$(1)^2 = 1$

$(2)^2 = 4$

* $\Sigma[x-\bar{x}]^2 = -4-1+1+4 = 10$

$$\beta_0 = \bar{y} - \beta_1 \bar{x}$$

$$\beta_1 = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\Sigma(x-\bar{x})^2}$$

$\beta_1 = \frac{\cancel{100}^{10}}{\cancel{10}_1} = 10$ → slope

$\beta_0 = 35 - 100(3) = 5$ → y intercept

Equation is

$y = \beta_0 + \beta_1 x$

$y = 5 + 10x$ ✓

8/4/22
3:30am

8/4/22
1:00 pm

* BIVARIATE STATISTICS *

X	Y

⇒ Simple linear regression :-

it is **2 Sample "T" test**

H_0 : Avg. line ≥ Regression line

H_1 : Regression line > Avg. line.

Q :- How can This can be proved by statistically

A :- Where $(SSE)_{\text{Regression line}}$ < $(SSE)_{\text{Average line}}$

30 < Ex:- 120

⇒ Multiple linear regression :-

x_1	x_2	x_3	x_4	Y

Where we have multiple "x" Values and only single "y". We use **"ANOVA"** Statistical Test.

applied when Output variable is "Continous".

Q :- What is The hypothesis Test apply For Regression ?

A :- **Anova Test** (>2 samples) Apply For More Than 2 Samples

if we apply Only For 2 samples it is "2 sample T Te

* **"Correlation"** is high gives "good" results → Linea Regression

"Relation between two variables"

Modulus of 'r' ← |r| = ± any value

→ $|r| \geq \boxed{0.8}$ (strong) correlation

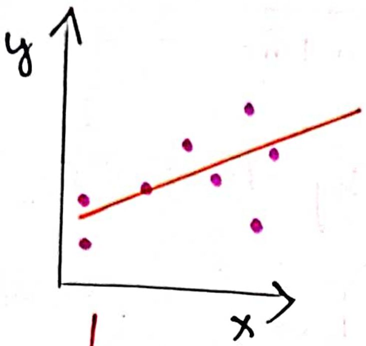
$\boxed{0.5} < |r| < \boxed{0.8}$ (moderate)

$|r| < \boxed{0.5}$ (weak)

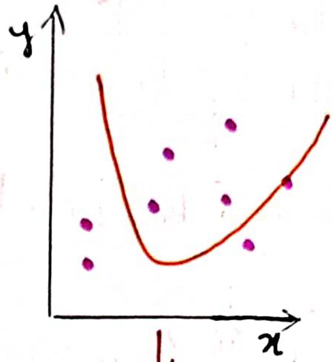
The Value of "One Variable", is a function of "Other Variable"

The Value of $\boxed{y}$, is a Function of $\boxed{x}$;

$$\boxed{y = f(x)}$$



$\therefore y = mx + c$



$y = ax^2 + bx + c$

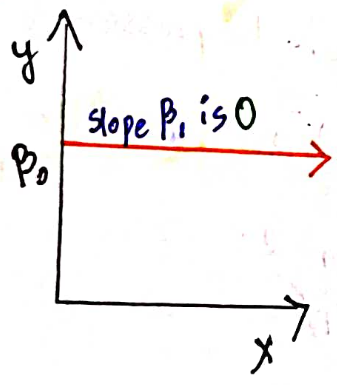
The Value of dependent Variable, is function of independent Variable

$$\Rightarrow \quad E(y) = \beta_0 + \beta_1 x$$

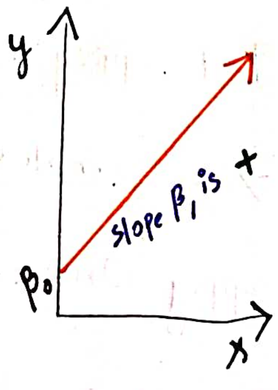
$\beta_1 \rightarrow$ "Slope"

$\beta_0 \rightarrow$ "y" intercept

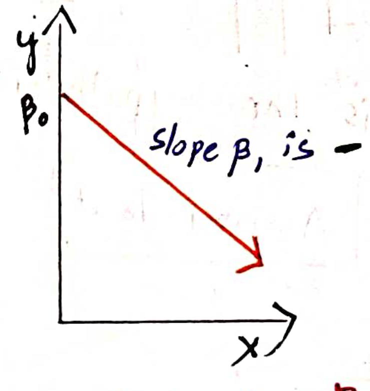
$E(y) \rightarrow$ Estimated Value



$E(y) = \beta_0 + 0(x)$



$E(y) = \beta_0 + \beta_1 x$



$E(y) = \beta_0 - \beta_1 x$

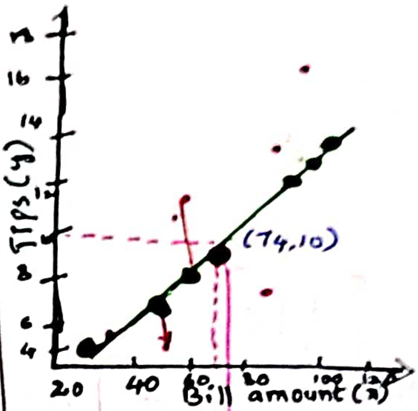
* Calculations

Deviations

$y = 0.1462[x] - 0.8188$

Equation

X	Y	$X-\bar{X}$	$Y-\bar{Y}$	$(x-\bar{x})^*[y-\bar{y}]$	$[x-x]^2$	$y = 0.1462x - 0.8188$	$\hat{y}$
		34-74	5-10	-40×-5			
34	5	-40	-5	200	1600	0.1462[34]-0.8188	4.1505
108	17	34	7	238	1156	0.1462[108]-0.8188	14.9693
64	11	-10	1	-10	100	0.1462[64]-0.8108	8.5365
88	8	14	-2	-28	196	0.1462[88]-0.81088	12.0453
99	14	25	4	100	625	0.1462[99]-0.81088	13.6535
51	5	-23	-5	115	529	0.1462[51]-0.8108	6.6359
$\bar{X} = 74$	$\bar{Y} = 10$			$\Sigma = 615$	$\Sigma = 4206$		



Centroid

The line should pass through centroid.

Avg. of x :- $\dfrac{34+108+64+88+99+51}{6} = \bar{x} = 74$

Avg of y :- $\dfrac{5+17+11+8+14+5}{6} = \bar{y} = 10$

• Descriptive Statistics / Centroid

$\bar{x}, \bar{y} = 74, 10$

$b_0 [\bar{y} - b_1 \bar{x}]$

* line $[\hat{y} = b_0 + b_1 x]$ min

$$\boxed{b_1 = \frac{\Sigma[x_i - x][y_i - \bar{y}]}{\Sigma[x_i - \bar{x}]^2}}$$

$\boxed{\text{Slope}} = \dfrac{\Sigma[x-\bar{x}][y-\bar{y}]}{\Sigma[x-\bar{x}]^2}$

How This Equation is Formed ?

**

$\boxed{\text{Variance }(x)} = \dfrac{\Sigma(x-\bar{x})^2}{(n-1)} \to s$

Varience(x) can be written as.

$$\frac{\Sigma(x-\bar{x})(x-\bar{x})}{(n-1)}$$

***

$\boxed{\text{Covarience }(x,y)}$ written as

$$\frac{\Sigma(x-\bar{x})(y-\bar{y})}{(n-1)}$$

Substuting & Equating

$$\frac{\text{cov}(x,y)}{V(x)} = \frac{\dfrac{\Sigma[x-\bar{x}][y-\bar{y}]}{\cancel{n-1}}}{\dfrac{\Sigma[x-\bar{x}]^2}{\cancel{n-1}}}$$

(S.S.E)

* Sum of Squares of Error $= \Sigma[y-\hat{y}]^2$

$$= \Sigma(y-[\beta_0+\beta_1 x)]^2$$

$$= \Sigma\underbrace{[y-\beta_0-\beta_1 x]^2}_{\text{should be minimum}}$$

$\boxed{\text{Slope}}$

$$b_1 = \frac{\Sigma[x-\bar{x}][Y-\bar{Y}]}{\Sigma[x-\bar{x}]^2}$$

$$b_1 = \frac{615}{4206}$$

$$b_1 = 0.1462$$

$\boxed{\text{y intercept}}$ $\Rightarrow b_0 = \bar{y} - b_1\bar{x}$

$$b_0 = 10 - 0.1462[74]$$

$$b_0 = -0.8188$$

$\boxed{\text{Final Equation}}$ $\Rightarrow$ $\boxed{y = b_0 + b_1 x}$

$$y = -0.8188 + 0.1462\,x$$

$$y = \beta_0 + \beta_1 x$$

(or)

$$y = \underbrace{0.1462\,x}_{mx} \underbrace{-\,0.8188}_{+\,c}$$

"*** "Accuracy" Called as

Coeffient of Determination"

$$R^2 = 1 - \frac{SSE_{(reg)}}{SST_{(Avg)}}$$

"Determing The variation of "y" with respect To variation of "x""

(or)

→ As "x" value changes, how "y" value changes.

(or)

"y" variable $\boxed{\text{Explained by}}$ "x" variable.

Que:- What is Max. value of R^2

Ans: $R^2 = 1$ [Max. value]

$$(SSE)_{Reg} = 0$$

$R^2 = -$Infinite $\boxed{\therefore R^2 < 0 \text{ (Worst model)}}$

****

$\boxed{\text{Range of } R^2}$

$$[\underbrace{\alpha}_{\text{Infinite}}, 1]$$

"Regression Squared Error"

$\hat{y}$	Error $y-\hat{y}$	Square Error $(y-\hat{y})^2$
4.1505	(5-4.1505) 0.8495	$(0.8495)^2$ 0.7217
14.9693	2.0307	4.1237
8.5365	2.4635	6.0688
12.0453	-4.0453	16.3645
13.6535	0.3465	0.1201
6.6359	-1.6359	2.6762

SSE $\Rightarrow \sum = 30.075$

* Avg. line = 120
* Regression line = 30.075

→ difference between these are $\frac{30}{120} \Rightarrow 75\%$.

Co-efficient of Determination.

***

(or)

$$\boxed{R^2 = 1 - \frac{SSE_{(Reg)}}{SST_{(Avg)}}}$$

→ $\dfrac{SSE \text{ [Sum of squares of Error [Regression line]}}{SST \text{ [Sum of squares of Error [Average line]}}$

$\sum[y-\hat{y}]^2$

$\sum[y-\bar{y}]^2$

***

$$\boxed{R^2 = 1 - \frac{\sum[y-\hat{y}]^2_{reg}}{\sum[y-\bar{y}]^2_{Avg}}}$$

$$R^2 = 1 - \frac{\cancel{30}\,1}{\cancel{120}\,4}$$

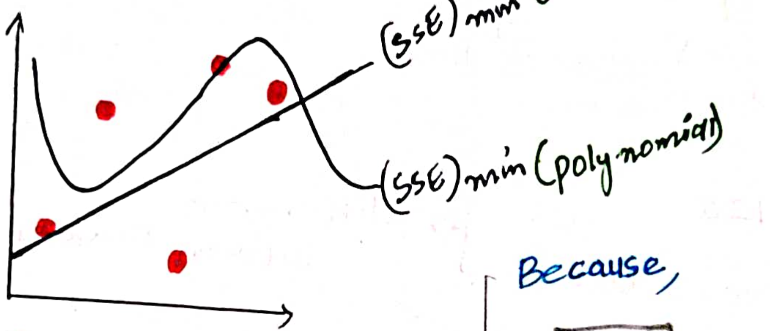
$$R^2 = \frac{4-1}{4} = \frac{3}{4} \Rightarrow 0.75 \Rightarrow \boxed{75\% \text{ accuracy}}$$

8/4/22
4:30 PM

Dt:- 9/4/22
3:00 pm

Evaluation Metrics in Regression

After Model fitting, we would like to assess the performance of model by $\boxed{\text{Comparing model Predictions}}$ to actual $\boxed{\text{[True] data.}}$



1. $\boxed{\text{MAE}}$ [Mean Absolute Error] → $\boxed{\dfrac{\sum |y-\hat{y}|}{n}}$

Bad → differentiation

squaring of units

2. $\boxed{\text{MSE}}$ [Mean Squared Error] → $\boxed{\dfrac{\sum [y-\hat{y}]^2}{n}}$

Because,

$\boxed{\text{MAE}} \rightarrow \dfrac{d[|x|]}{dx}$

$\dfrac{d[|x|]}{dx}$ at $x=0$ $\rightarrow \dfrac{x}{|x|} = \dfrac{0}{|0|} \Rightarrow \infty$ → Infinite

$\boxed{\text{MSE}} \rightarrow \dfrac{dx^2}{dx}$

$\dfrac{dx^2}{dx}$ at $x=0$ $= 2x$

**** Only way To make Error Minimum * * is $(\text{error})_{min}$

$\boxed{\dfrac{d[\ ]}{dx} \text{ at } x=0}$

x is only notation it, z, anyword $\dfrac{d(\ )}{dy}$

Example

milage of a car

Y	$\hat{Y}$	$y-\hat{y}$
18.2	18.9	0.7

Here 18.2 Kmph

18.9 Kmph

Error $= (0.7)^2 = 0.49 \dfrac{\text{Km}^2}{\text{h}^2}$

If we square it

Problem is here units also Squaring

Que:- What are Evaluation Metrics used for Regression?

Ans:- We have 1. Mean Absolute Error (MAE] and 2. Mean Square error [MSE].

Que ⇒ Mean absolute Error (MAE) Where, we are going to use?

Ans:- MAE is "Absolute Values of Mean" is called mean absolute Error.

But, Problem with $\frac{d}{dx}$ of (Mod) $|x|$ is $\frac{x}{|x|}$ and at $x=0$, The answer is infinite

So, we are not going to consider "MAE"

Que ⇒ Mean Square Error (MSE)

Ans :- Problem with MSE is with calculating of $\frac{d}{dx}$ of x^2 we will get some value of $2x$. But, by squaring up squaring the Error. Units are also squaring up. So, we not going to consider it gives wrong prediction. So, we not going to consider "MSE".

And Solution For This is :-

$$RMSE = \sqrt{\frac{\Sigma(y-\hat{y})^2}{N}}$$

Same as in Variance

we Take Standard deviation

↓

"Root Mean Square Error"

Que :- What is "Root Mean Square Error" RMSE ?

A :- "RMSE" represents Standard deviation of residuals [Errors]

difference between model predictions and True

Values (Training data)

* RMSE can be easily intrepreted compared To MSE

because RMSE units match units of Output.

$$\boxed{RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}[y-\hat{y}]^2}}$$

(or)

$$\boxed{\sqrt{\frac{\sum[y-\hat{y}]^2}{n}}}$$

Example

squaring of Errors

Converting negetaive Values To positive Values by adding | |

X	Ŷ	Error	MAE	MSE
1	1.4	0.4	0.4	0.16
2	2.2	0.2	0.2	0.04
3	3.8	0.8	0.8	0.64
4	4.1	0.1	0.1	0.01
5	5.6	0.6	0.6	0.36

$\frac{\Sigma}{n} = 0.42$ $\frac{\Sigma}{n} = 0.242$

Root Mean Squar Error $= \sqrt{0.242}$

⇒ **Mean percentage Error [MPE]**

$$MPE = \frac{100\%}{n} \sum_{i=1}^{n} (y-\hat{y}) / y_i$$

In this we calculate percentage of Error. But we don't consider "positive" or "Negative" value, without absolute operation.

⇒ **Mean absolute Percentage Error [MAPE]**

In This "MAPE" we convert Every value To positive by applying "Absolute" to the Equation.

$$MAPE = \frac{100\%}{n} \sum_{i=1}^{n} |y-\hat{y}| / y_i |$$

9/4/22
5:10pm