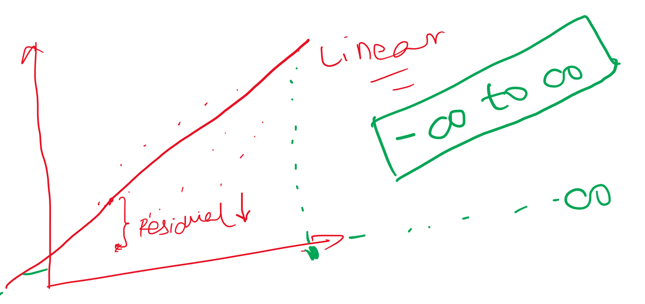
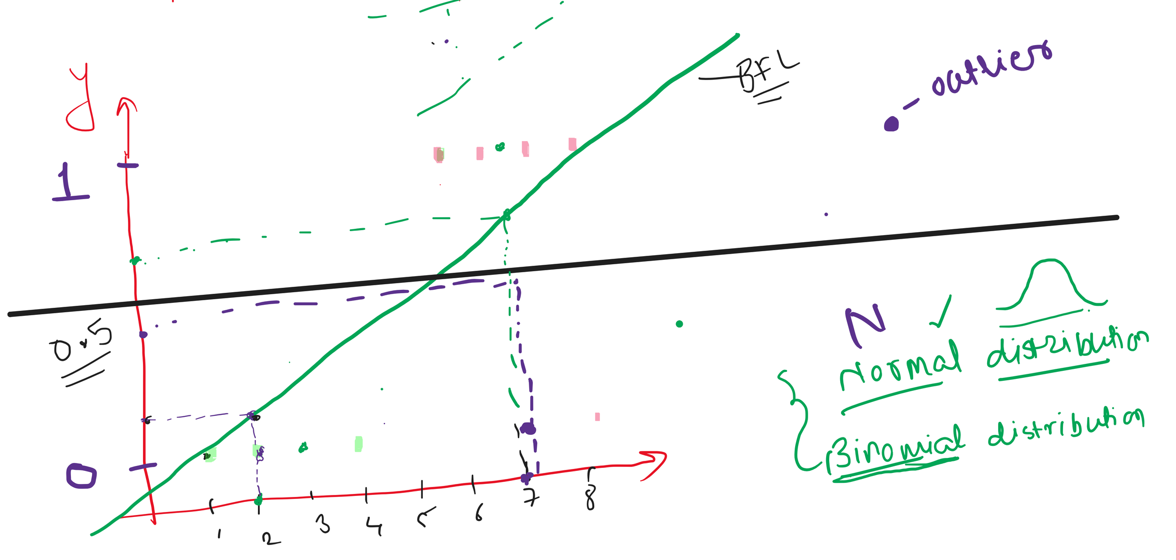
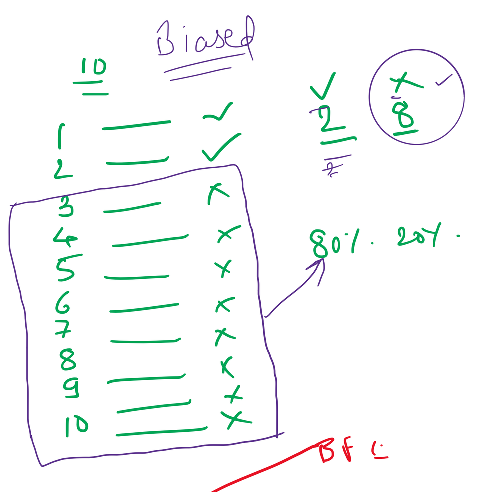
	x_1	x_2	x_3	y
1				
2	cat.	conti	conti	0
3				1
4				0
5				0
6				

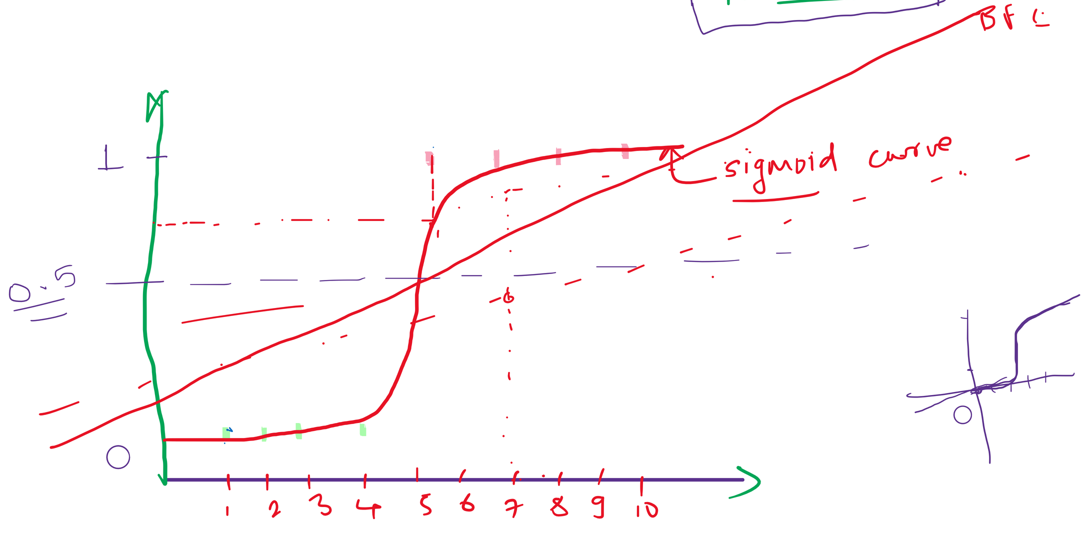




Imbalance data :-

{ 80% & 20%
70% & 30% }





$y = mx + c$

sigmoid fun = $\frac{1}{1+e^{-y}}$

logistic fun = $\frac{1}{1+e^{-(mx+c)}}$

Assumptions.
1) No multicolinearity
2) Linearity → independent variable & log of odds
3) Req. large sample size.
4) Independent →

Cost function :- Logistic reg.

i] log-loss

$$\text{log-loss} = -\frac{1}{N}\sum_{i=1}^{n}\left[(y_i \cdot \log(P_i)) + (1-y_i)\cdot\log(1-P_i)\right]$$

$y_i = y_a = y_{actual}$ (actual class)

P_i = Prob. of 1

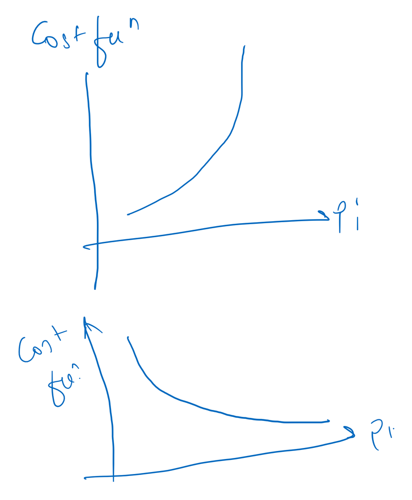
$1-P_i$ = Prob. of 0

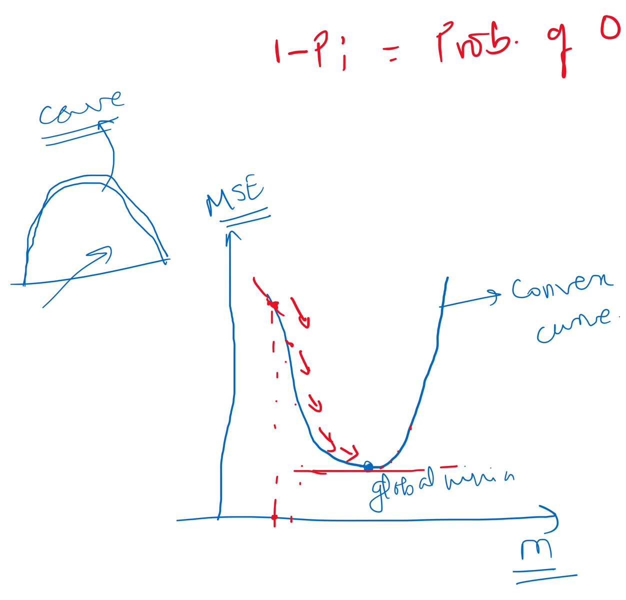
when $y_i = 0$

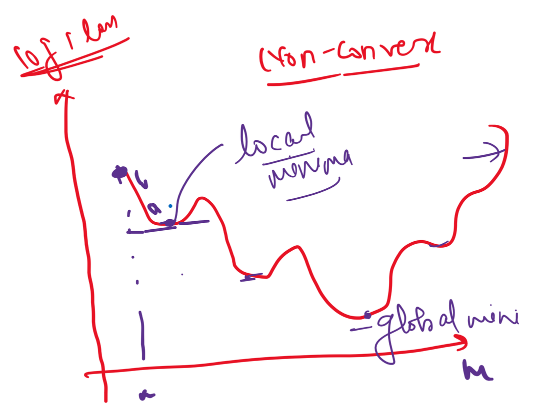
cost fun $= -\frac{1}{N}\sum(\log(1-P_i))$

when $y_i = 1$

cost fun $= -\frac{1}{n}\sum \log(P_i)$







Case-I ; $y_a = 1$ & $y_p = 0$ (wrong. Pred.)

$MSE = (y_a - y_p)^2 = (1-0)^2 = 1$

log-loss $= -\log(P_i) = -\log(0) = -\infty$

log-loss >>> MSE

Case-II $y_i = 1$, $y_p = 1$

$MSE = (y_i - y_p)^2 = 0$ ✓

log loss $= -\log(1) = 0$ ✓

Revision:-

i) Logistic reg. — classification
- Binary
- Multiclass — ordinal - / multinomial

2] What happen if we use Linear reg for classification
- outlier - sensitive
- Assumptions - violates

3] diff. betn Linear reg & logist

{ $-\infty$ to ∞ / Continuous } 0 to 1 / catego. }

4) Cost fun → log. loss

$$= -\frac{1}{n}\sum_{i=1}^{n}(y_i \log(P_i) + (1-y_i)\log(1-P_i))$$

MSE as cost fun → logistic.

i) Non convex → stuck at local min

ii) log-loss - strongly penalize miscalli