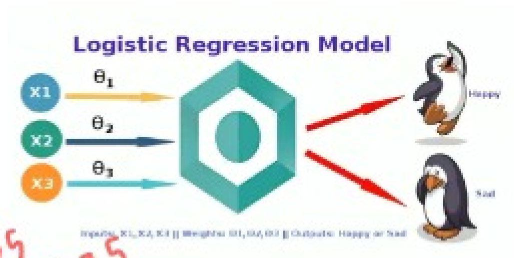
Assumptions of linear regression:

1. Linearity
2. No autocorrelation(optional)
3. No or little multicollinearity
4. homoscedasticity(The variance of errors should be constant for all X levels)

Logistic Regression

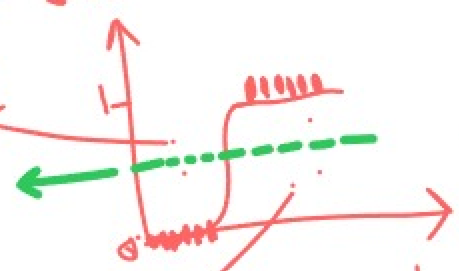


$P(1) = 0.75$

$P(0) = 1 - P(1) = 0.25$

0.35

0.65



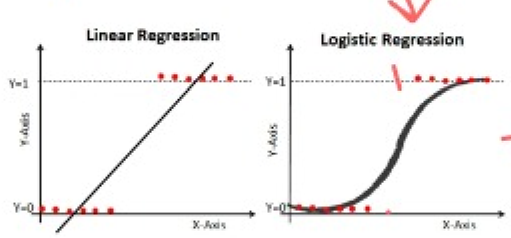
Sigmoid graph
bernoulli's PD

Threshold value=0.5

$$P(y) = \frac{1}{1+e^{-y}} \quad y = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n$$

e Euler's Constant

0.23



Cost function (Log Loss) | Binary Cross Entropy

$$= \text{Log loss} = \frac{1}{N}\sum_{i=1}^{N} -(y_i * \log(\underbrace{\hat{Y}_i}_{P_1}) + (1-y_i) * \log(\underbrace{1-\hat{Y}_i}_{P_0}))$$

MLE (Maximum Likelihood Estimator) is used in Logistic regression

Study Hours (x)	Pass (y)
1	0
2	0
3	1
4	1

P(y) = 1/1+e^-z

z= B1x1 + B0

x=2 hours , y=0
B1,B0=0

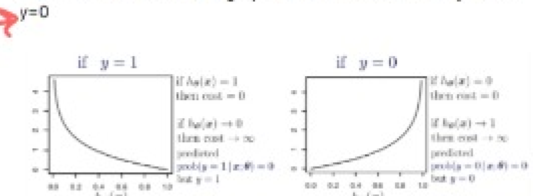
Step 1:
P(y)= 0.5

Step 2:
Loss (After applying maximum likelihood estimation my log loss formula will become like this

$$\text{Loss} = -[y \cdot \log(\hat{y}) + (1-y) \cdot \log(1-\hat{y})]$$

- p(y) is the probability of 1.
- 1-p(y) is the probability of 0.

Let's see what will be the graph of cost function when y=1 and y=0



Loss=-[0log(0.5) + (1-0)log()1-0.5]
Loss=0.69

Step 3 Calculate gradients
Gradient for feature X1=(0.5-0).2=1.0
Gradient for intercept = 0.5-0= 0.5

Step 4 update parameters
learning rate =0.1

B1'=0-0.1(1.0)=-0.1
B0'=0-0.1(0.5)=-0.05