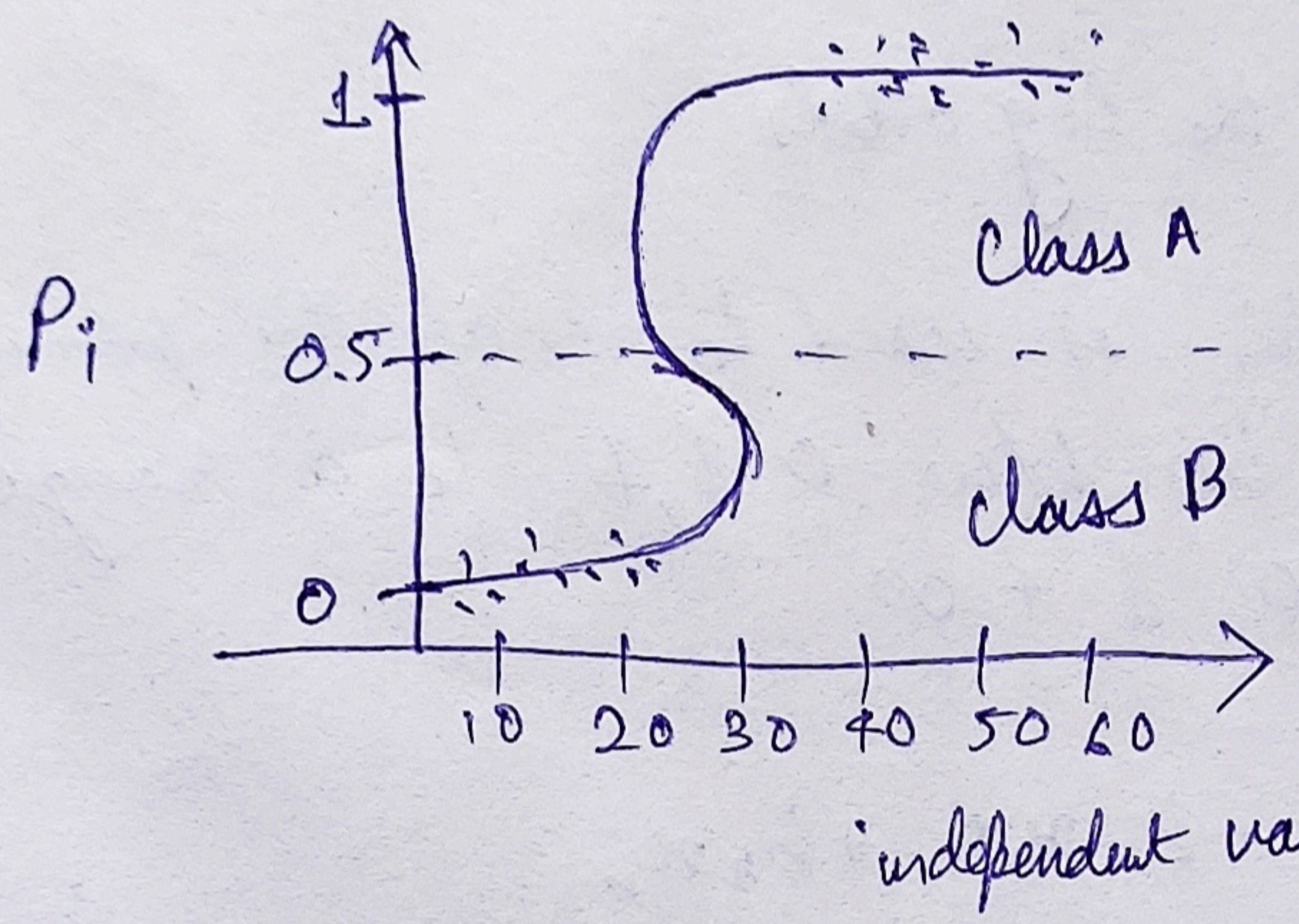
Logistic Regression

\# used in categorical dataset. (yes/No, +ve/-ve, true/False)
\# fraud detection, disease diagnosis, Emergency detection, spam/no spam

Sigmoid fn = $y = \frac{1}{1+e^{-x}}$ → independent variable

e → 2.718

\# Sigmoid fn is trying to convert the independent var. into a expression of probabely that ranges b/w 0 & 1 wrt. dependent var



$$\frac{1}{1+e^{-x}} = \frac{1}{1+e^{-mx+c}}$$

m = give stretching
c = gives the moment of the sigmoid up or down

\# To calculate the error we use the log loss fn.

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

→ Goes to gradient descent for calculation

↓

goal is to minimize this to a large -ve number

Logistic Model

Features $x_1, x_2, \dots x_n$
Y = binary output
$p = P(y=1)$ (Probability of $y=1$)

$$\boxed{\ln\left(\frac{p}{1-p}\right) = b_0 + b_1 x_1 + b_2 x_2 \dots b_n x_n}$$

$b_0, b_1, b_2 \dots$ parameter of weight that will be estimated by training

$\frac{p}{1-p}$ = the odds

$\ln(p/1-p)$ = the log odds

↳ This is used to map the ~~prob~~ probability that lies b/w 0 & 1 to a range b/w $-\infty$ & $+\infty$

$$\frac{p}{1-p} = e^{b_0 + b_1 x_1 + b_2 x_2 + \dots + b_n x_n}$$

$$\Rightarrow p = \frac{e^{b_0 + b_1 x_1 + b_2 x_2 \dots b_n x_n}}{1 + e^{b_0 + b_1 x_1 + b_2 x \dots b_n x_n}}$$

$$\Rightarrow p = \frac{1}{1 + e^{-(b_0 x_0 + b_1 x_1 \dots b_n x_n)}} \quad // \quad S(x) = \frac{1}{1+e^{-x}}$$

Loss fn :-

used to calculate error in the predicted value
The L2 loss fn :-

$$L = \sum_{i=1}^{n} (y_i - \bar{y}_i)^2$$

Our goal is to minimize the entire loss using ~~gra~~ gradient descent algorithm.

Gradient Descent Algorithm for logistic Reg :-

$$\bar{y}_i = p = \frac{1}{1+e^{-(b_0+b_1x_1)}} = \frac{1}{1+e^{-b_0-b_1x_1}}$$

① Initially $b0 = 0$, $b1 = 0$, $L = 0.001$

② Calculate partial derivatives :-

$$D_{b_0} = -2\sum_{i=1}^{n}(y_i - \bar{y}_i) \times \bar{y}_i \times (1-\bar{y}_i)$$

$$D_{b_1} = -2\sum_{i=1}^{n}(y_i - \bar{y}_i) \times \bar{y}_i \times (1-\bar{y}_i) \times x_i$$

③ update b_0 & b_1

$$b_0 = b_0 - L \times D_{b_0}$$

$$b_1 = b_1 - L \times D_{b_1}$$

④ Repeat until loss is 0