02-02-2022

02 February 2022 13:38

$$\beta = (X^T X)^{-1} X^T y$$

$$\hat{y} = X\beta = X(X^T X)^{-1} X^T y$$

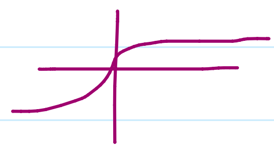
$$Hy$$

Linear
multiple
Polynomial
Logistic

Logistic Regression → Classification

→ o/p → categorical data

→ Sigmoid function



S-shaped
0 – 1

i) Binomial → 2 classes → Target var → 2 possible vals

ii) Multinomial → More than 2 possible vals - Target var. (Multi-class) (Not ordered)

iii) Ordinal → two or more categories for target var. (ordered)

Eg:

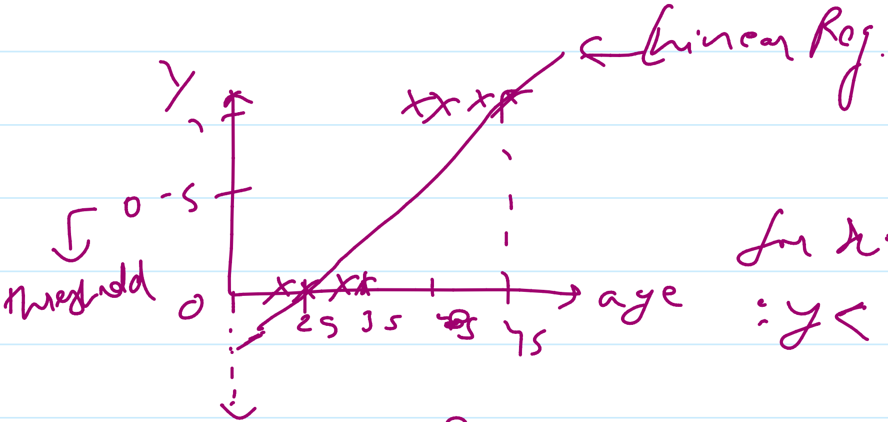
Dengue – High or Low → Binomial

Why use Logistic Regression & not Linear Regression?

Why use Logistic Regression & not Linear Regression?

x — age

y — subscribe



for x = 17

: y < 0

∴ $P_{rob} < 0$ for age < 25

Prob > 1 for age > 45

but P should be b/w 0 & 1

for age 38, y = 0.7

$y \geq 0.5 \quad \therefore P = 1$ //

$y < 0.5 \quad \therefore P = 0$ //

Problem with Linear Regression

$$P = \frac{\text{outcome of interest}}{\text{all possible outcomes}}$$

(Prob.)

$$\text{Odds } (\theta) = \frac{P \text{ of event}}{P \text{ of } \overline{\text{event}}} = \frac{P}{1-P}$$

If $P = 0.7$ $\quad \theta = \frac{0.7}{0.3} = 2.33$

If $P = 0.7$ $\theta =$ 0.3 2.33

$$P \in [0,1] \quad [\bar{P} = 1 - P]$$

$$\theta \in [0, \infty) \quad [\text{No such relation here}]$$

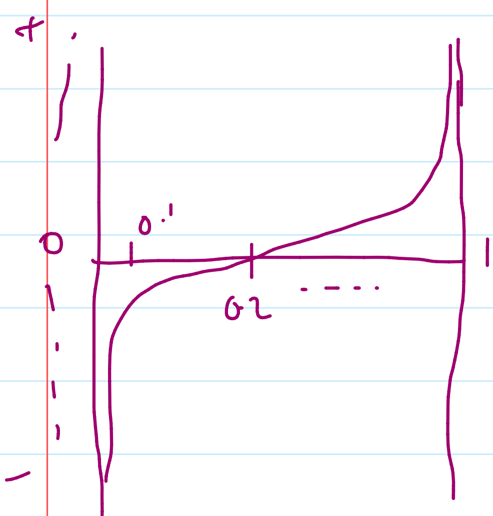
for odds we calculate

$$Logit(P) = \ln(Odds(y_i))$$

$$Logit(P) = \ln(\theta) = \ln\left(\frac{P}{1-P}\right)$$

Input : $[0, 1]$

Output : $(-\infty, \infty)$



$$\underline{\beta_0 + \beta_1 x_1} = logit(P)$$

Linear combo. of input var.

$$= \ln\left(\frac{P}{1-P}\right)$$

$$e^{\ln(P/1-P)} = e^{\beta_0 + \beta_1 x_1}$$

$$\frac{P}{1-P} = e^{\beta_0 + \beta_1 x_1}$$

$\beta + \beta_1 x_1$

$$1-p$$

$$p = (1-p)\, e^{\beta_0 + \beta_1 x_1}$$

$$p = e^{\beta_0 + \beta_1 x_1} - p\, e^{\beta_0 + \beta_1 x_1}$$

$$p + p\, e^{\beta_0 + \beta_1 x_1} = e^{\beta_0 + \beta_1 x_1}$$

$$p\left(1 + e^{\beta_0 + \beta_1 x_1}\right) = e^{\beta_0 + \beta_1 x_1}$$

$$\therefore \left[p = \frac{e^{\beta_0 + \beta_1 x_1}}{1 + e^{\beta_0 + \beta_1 x_1}} \right]$$

or

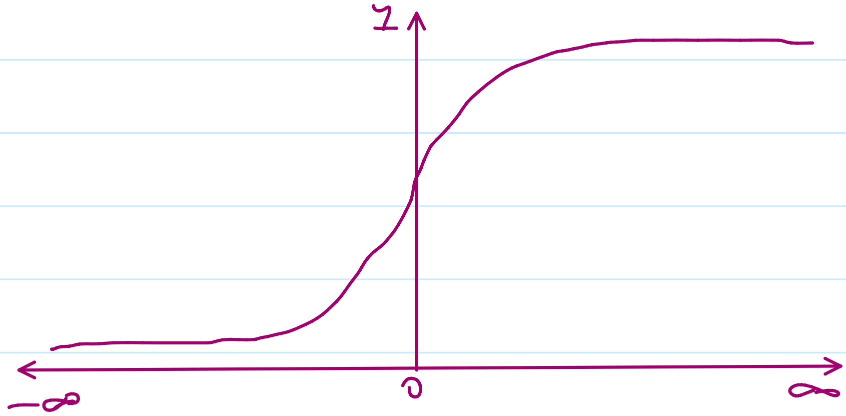
$$\left[p = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1)}} \right]$$

$$\text{Logit}(p) = \beta_0 + \beta_1 x_1$$ ← In linear eq. this is p

→ Inverse Logit func / sigmoid func.

$x \in (-\infty, \infty)$

$P \in [0, 1]$



i) For P to be in $[0, 1]$

func we choose should be +ve.
(to calc. P)

$abs(x)$, e^x, x^2 etc.

ii) func should yield < 1