LOGISTIC REGRESSION

→ It is a machine Learning method used in any classification problem when we want to distinguish one class from another. The most simplistic case is binary classification. We can answer either question with 'yes' or 'No'.

→ Neural Networks are built on Top of it.

→ It is statistical method used for predicting the PROBABILITY of a binary outcome based on 1 or more predictor variables. It's type of regression analysis where the dependent variable is categorical (usually binary), and the independent var can be <u>continuous</u> or <u>categorical</u>.

→ It uses the logistic function (Sigmoid function) to model the relationship b/w the dependent var & independent var. It is defined as.

$$\boxed{P(Y=1|X) = \frac{1}{1+e^{-Z}}}$$

where, → $P(Y=1|X)$ is the prob of dependent var (Y) being 1 given the independent var (X).

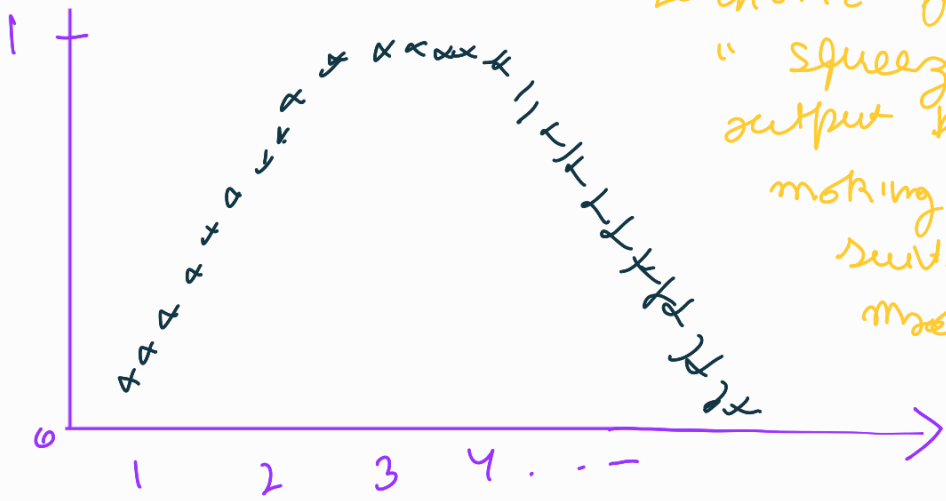
→ e is the base of natural logarithm.

→ Z is the linear combinations of the

independent vars & their coefficients

$$Z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$$

The coefficient β are estimated using MAXIMUM LIKELIHOOD ESTIMATOR or other optimization techniques.



LOGISTIC function "squeezes" the output b/w 0 & 1 making it suitable for modelling probabilities.

Logistic regression has a LINEAR APPROACH, but predictions are transformed using the Logistic Function. Bcz of this we can no longer understand the prediction as a Linear comb. of input variables as we can with linear regression.

Model can be stated as:

$$P(Y=1|x) = \frac{e^{(\beta_0 + \beta_1 x)}}{e^{(\beta_0 + \beta_1 x)} + 1}$$

we can change the above equation as follows:

$$\ln(P(Y=1|x)) = \underbrace{b_0 + b_1 * X}_{\text{like LINEAR regression}}$$

Q: Why do we use Logistic or Sigmoid function in Logistic Regression.

Answer: To map any real-valued input to

bounded output 0 and 1, which is suitable for representing probability. Reasons for using logistic

1) PROBABILITY INTERPRETATION

→ helps in interpreting prob. b/w 0 & 1.

2) S-shaped curve

Logistic function has S-shaped curve, which is desirable for modeling probabilities. As the input variable changes, the output smoothly transitions from near 0 to near 1.

3) Montonicity

→ Logistic func is monotonic, means it always increases or always decreases. This property ensures that as the input variable changes, the predicted prob of the event occuring moves monotonically from 0 to 1.

4) Continuity & Differentiability:

continuous & differentiable for all real numbers. This property is essential for optimization algo to converge efficiently during parameter estimation.

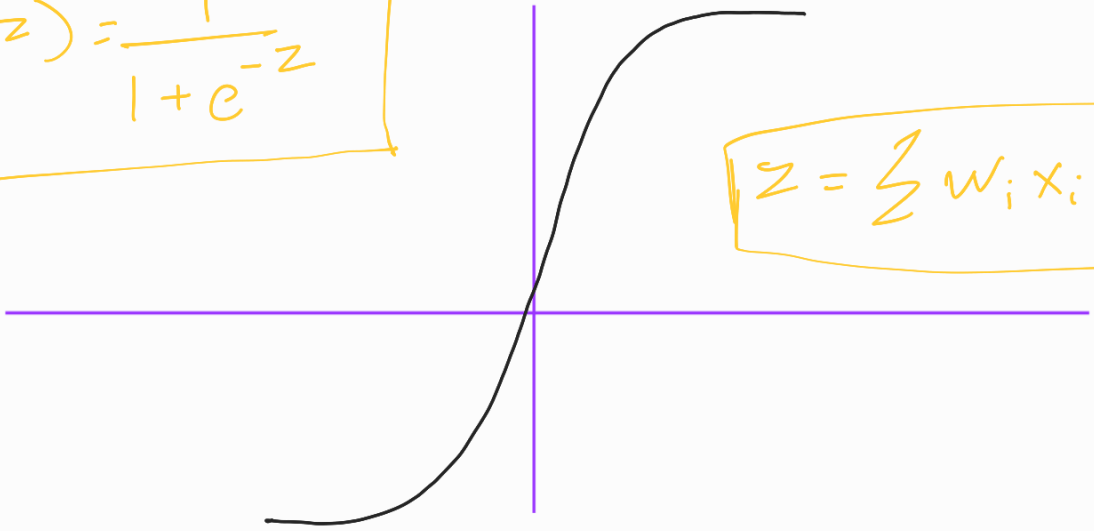
5) Easy to compute Inverse

inverse → logit function → convert prob back to odds, useful for interpreting coefficients

---

Function Plot - Sigmoid function

$$\boxed{\sigma(z) = \frac{1}{1+e^{-z}}}$$



$$\boxed{z = \sum w_i x_i + \text{bias}}$$