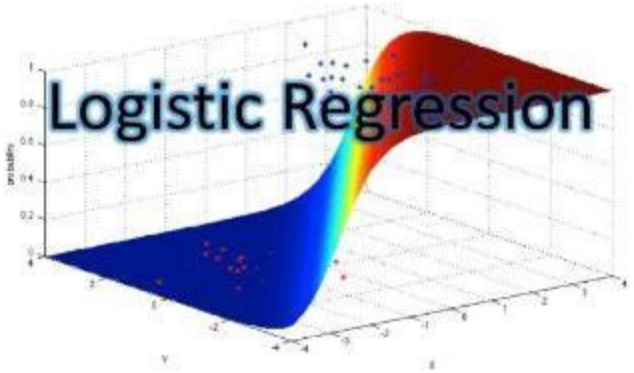
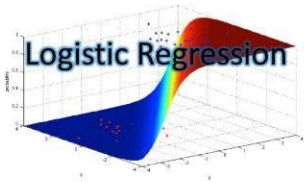
INTERNSHIPSTUDIO

Introduction to Logistic Regression



Agenda



INTERNSHIPSTUDIO

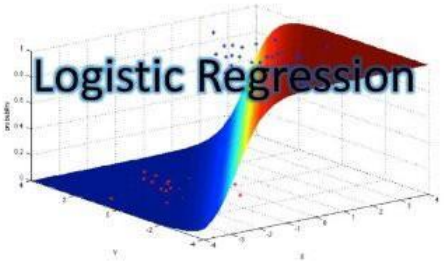
- Intro to Logistic Regression
- Cross Validation
- ROC Curve
- Confusion Matrix
- Implementing Logistic Regression Model
- Analyzing result parameters

What is Logistic Regression?

INTERNSHIPSTUDIO

- Logistic regression is a supervised learning classification algorithm used to predict the probability of a target variable.

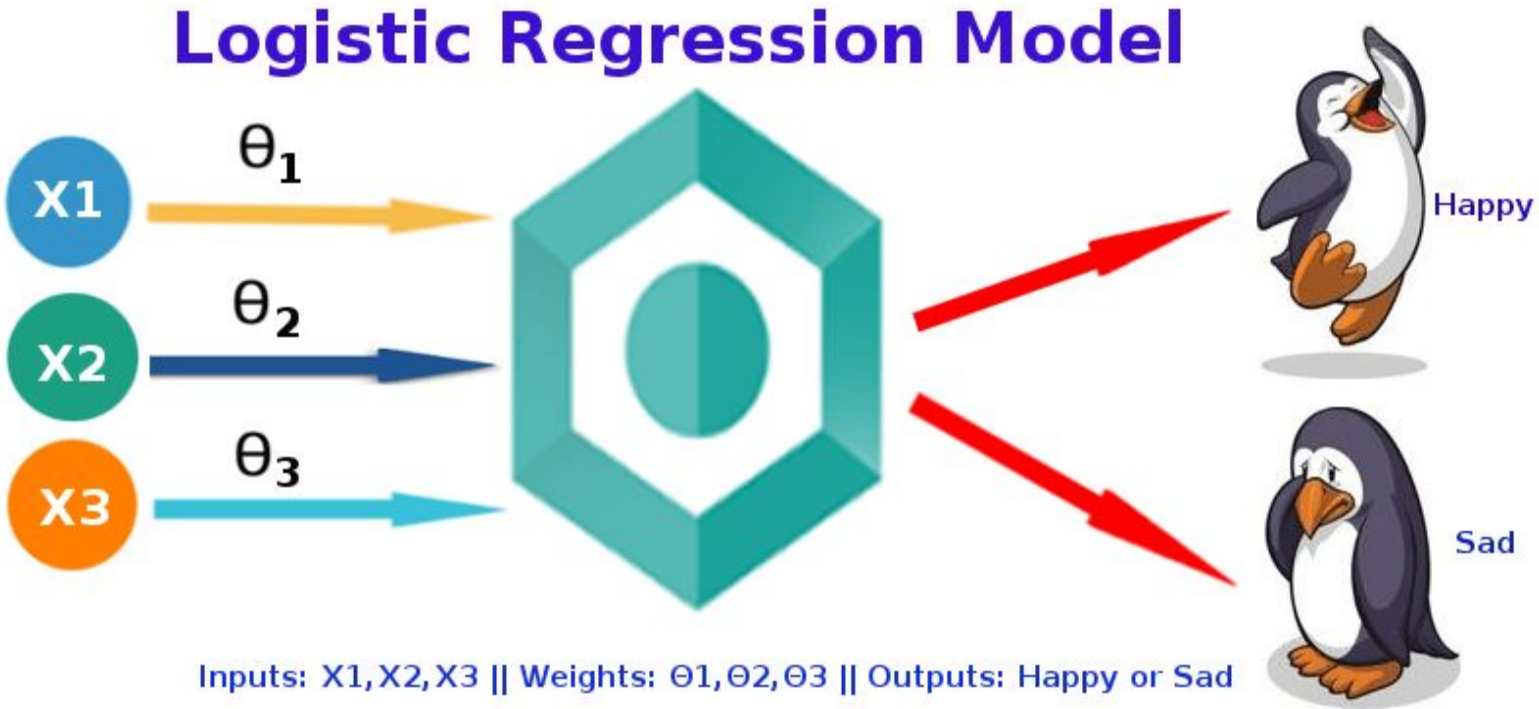
- The nature of target or dependent variable is binary, and there are only two possible classes.



Logistic Regression

INTERNSHIPSTUDIO

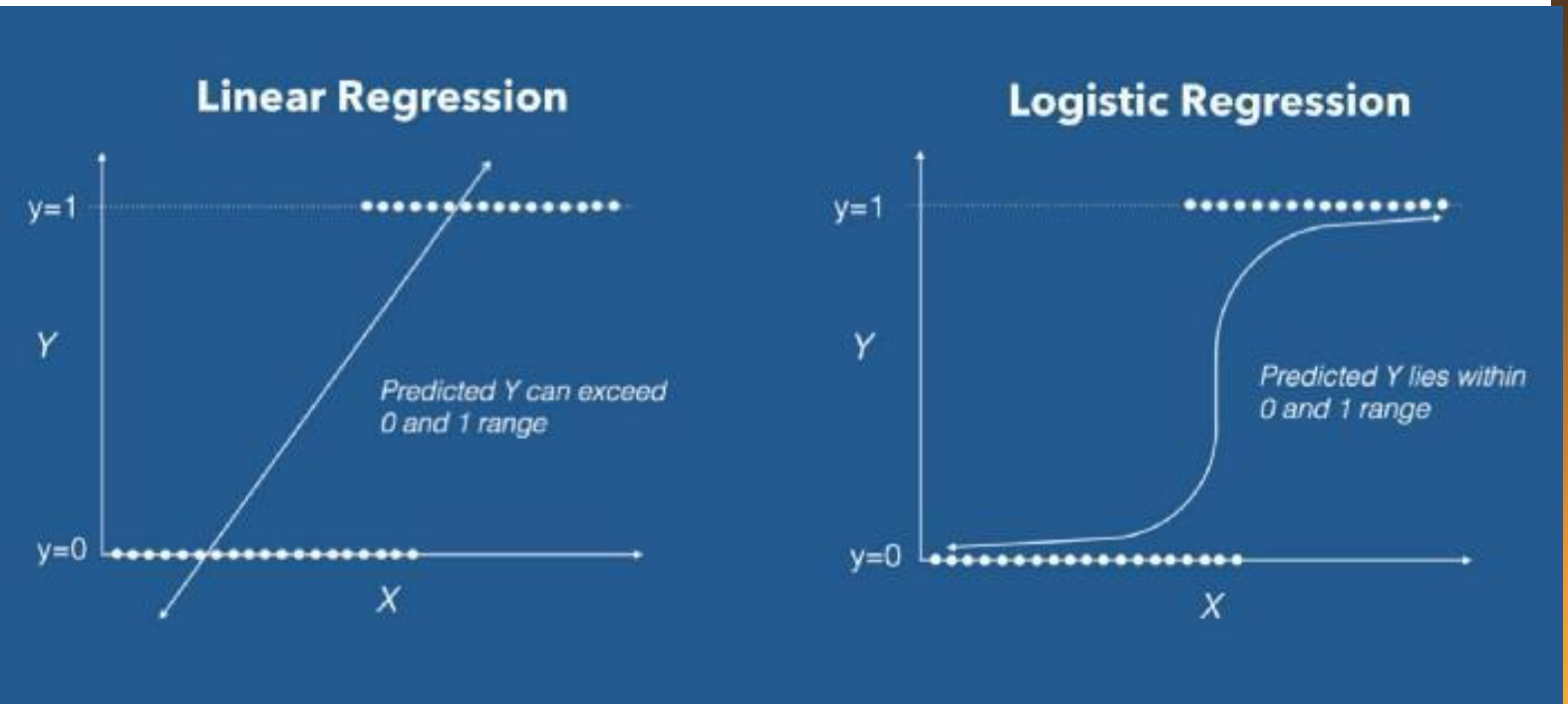
- Logistic regression is a classification algorithm used to assign observations to a discrete set of classes.
- Some of the examples of classification problems are **Email spam or not spam, Online transactions Fraud or not Fraud, Cats or Dogs**



Logistic Vs. Linear Regression

INTERNSHIPSTUDIO

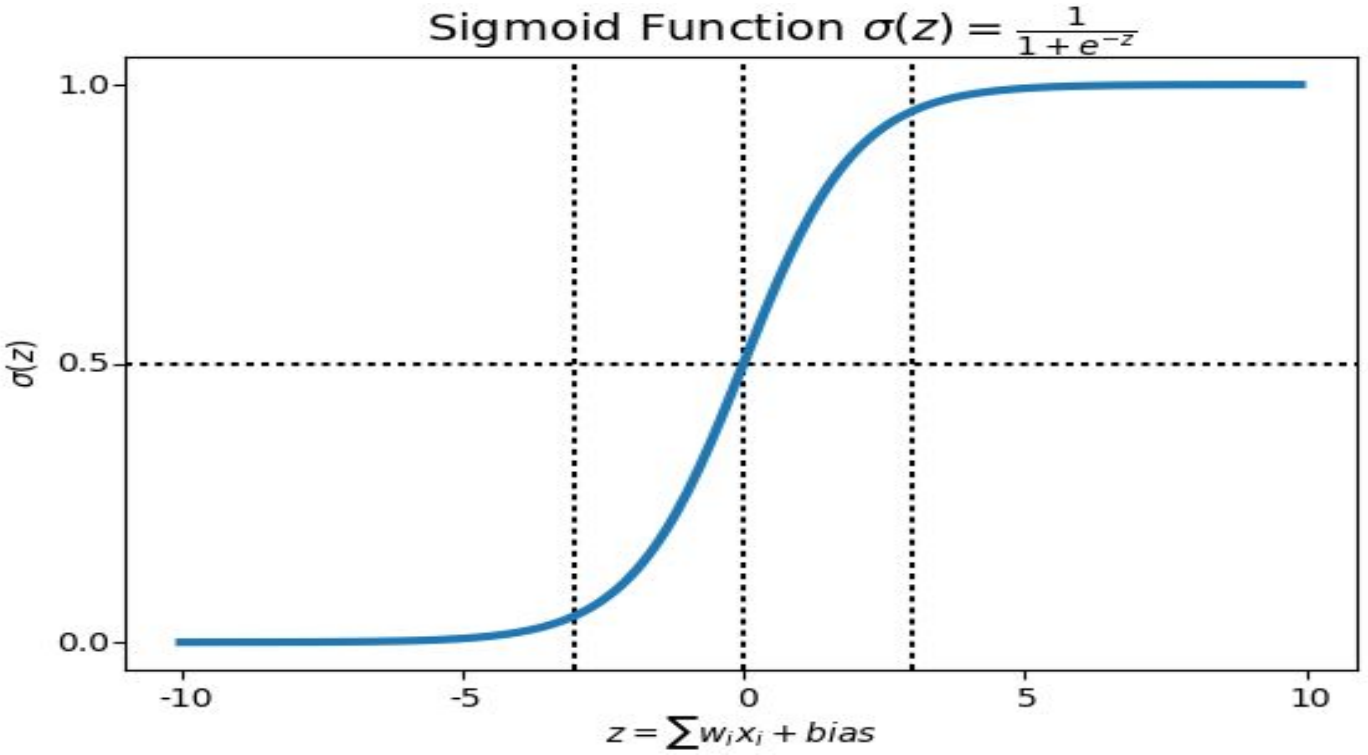
- In **linear regression**, the outcome (dependent variable) is continuous. It can have any one of an infinite number of possible values.
- In **logistic regression**, the outcome (dependent variable) has only a limited number of possible values. **Logistic regression** is used when the response variable is categorical in nature.



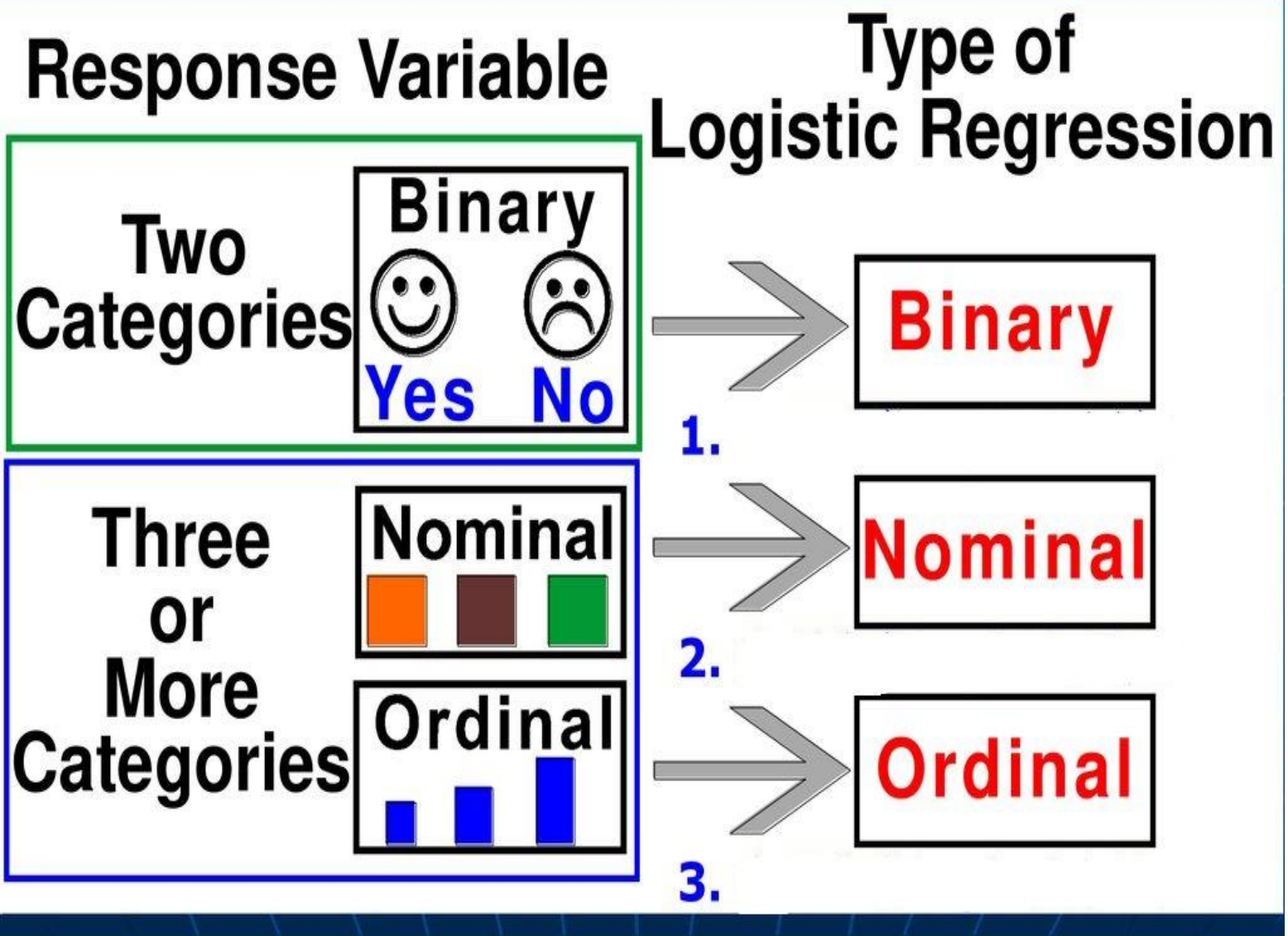
Sigmoid or Logistic Function

INTERNSHIPSTUDIO

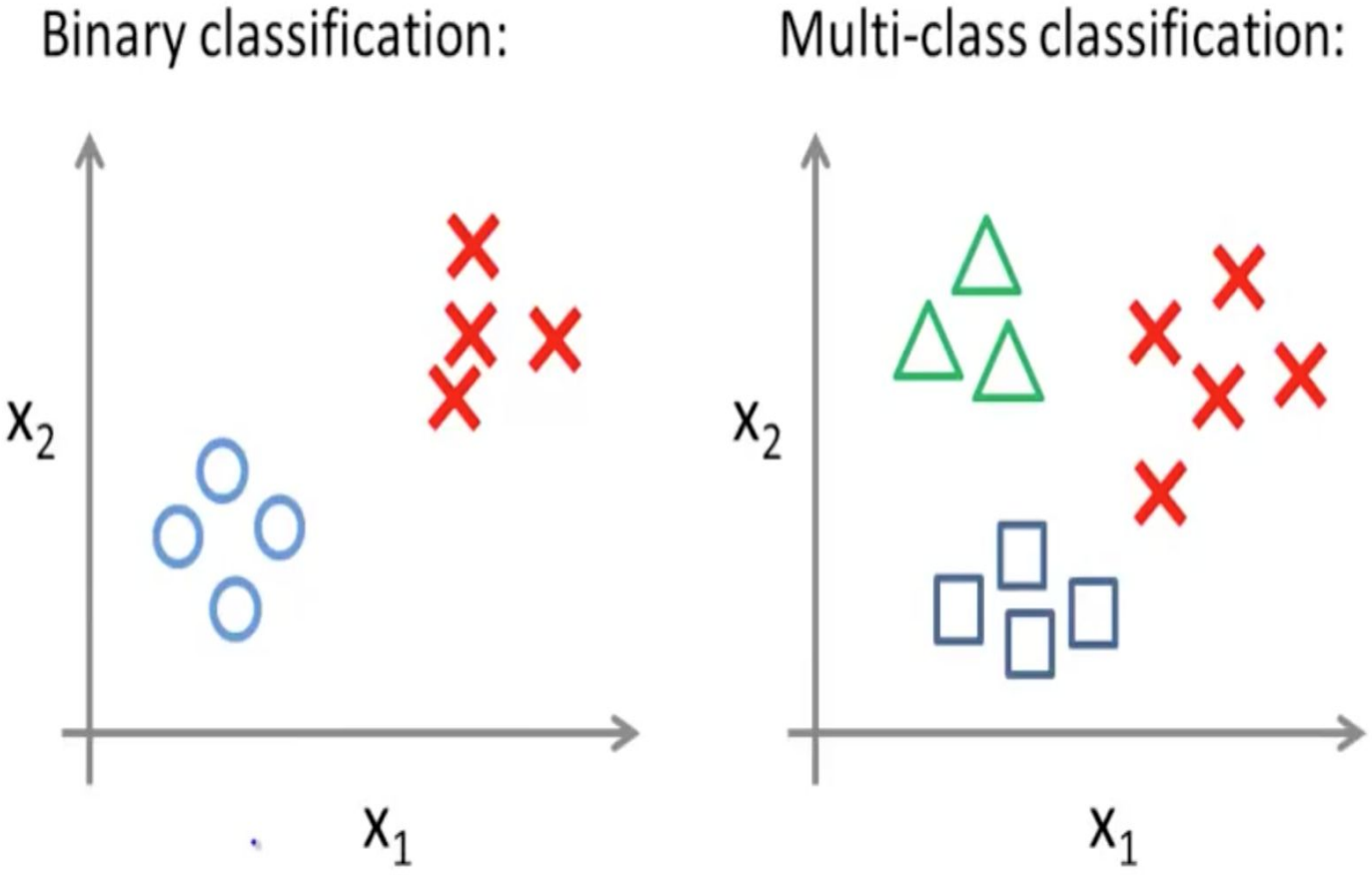
- In order to map predicted values to probabilities, we use the **Sigmoid function**. The function maps any real value into another value between 0 & 1.
- In machine learning, we use sigmoid to map predictions to probabilities.

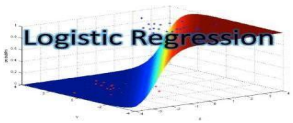


Types of Logistic Regression



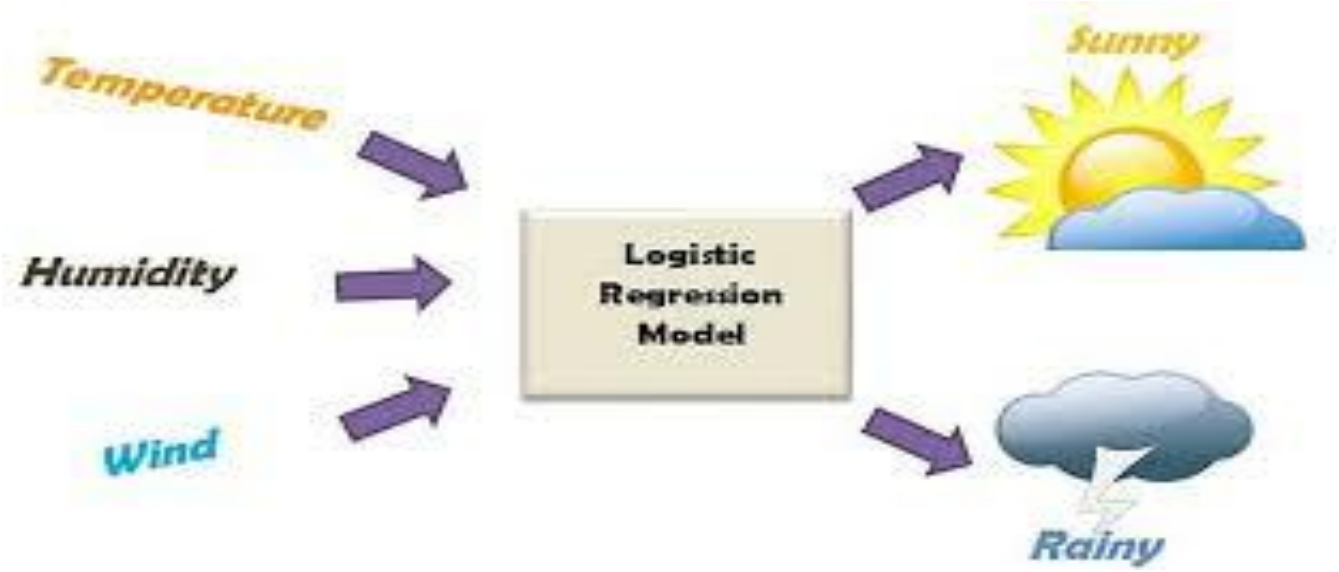
Types Of Logistic Regression





Multi Linear Logistic Regression

- Dependent variable column will contain more than two values but in the repeated format.
- Here the Outlook can be sunny, cloudy or overcast



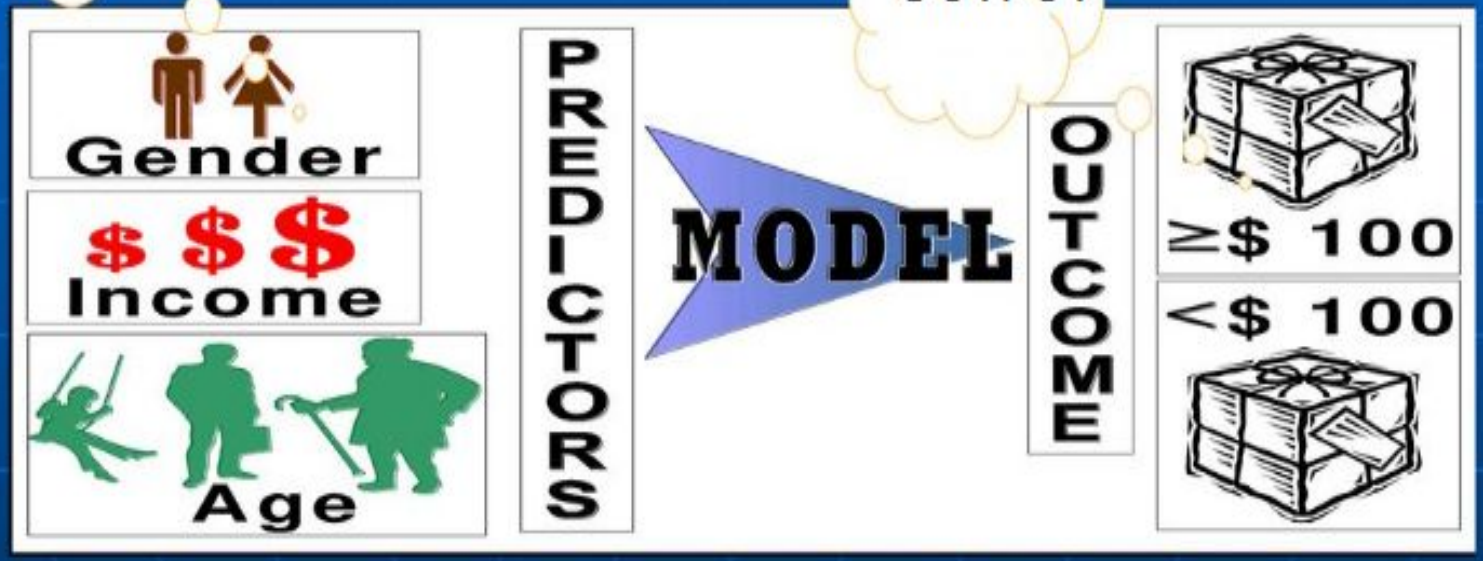
Temp-in ˚C	Humidity	Wind speed	Outlook
40	Hot	6.0	Sunny
43	Hot	11.1	Sunny
36	Hot	9.3	Overcast
39	Humidity	8.1	Cloudy

What Does Logistic Regression Do?

to predict the probability of specific outcomes.

INPUT

OUTPUT



Predictor variables	Predicted variable
Explanatory variables	Response variable
Covariables	Outcome variable
Independent variables	Dependent variable

INTERNSHIPSTUDIO

Q.1 What is Logistic Regression?

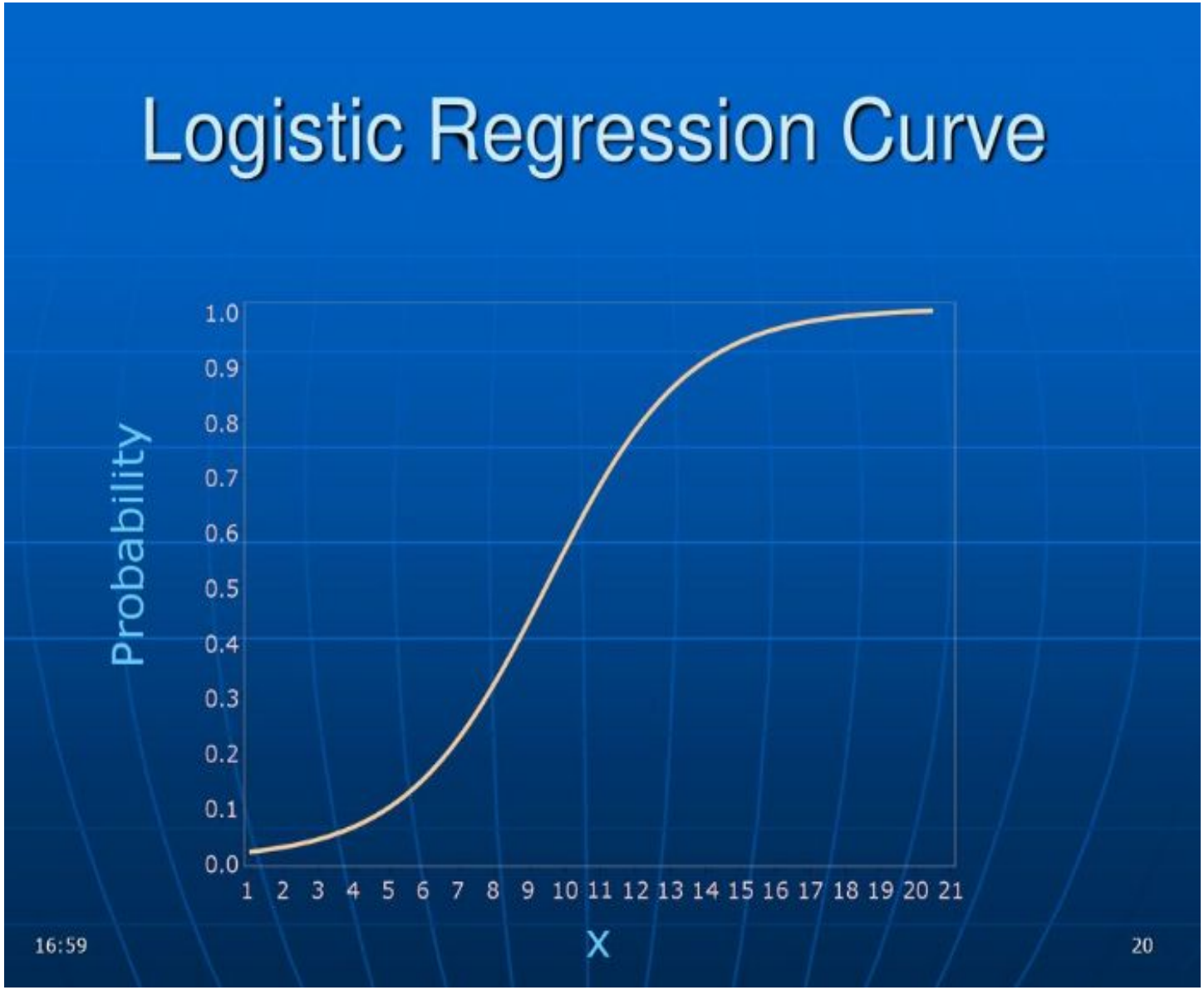
Q.2 What is the difference between logistic regression and linear regression?

Q.3 What is Sigmoid Function?

Q.4 What are the types of Logistic Regression?

Q.5 What are Nominal Variables?

Regression Curve



Logit Transformation

Logistic regression models transform probabilities called ***logits***.

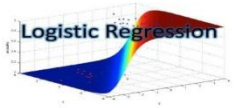
$$\text{logit}(\hat{p}_i) = \ln\left(\frac{\hat{p}_i}{1-\hat{p}_i}\right)$$

where

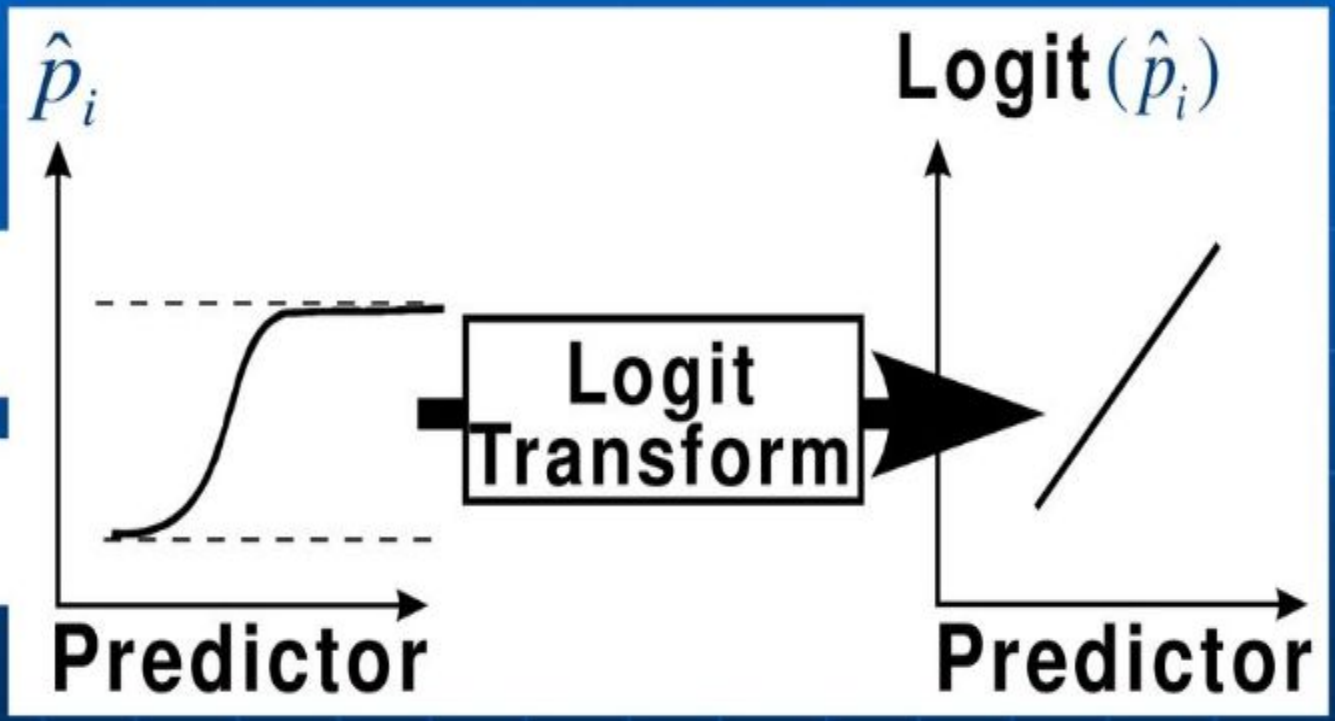
i indexes all cases (observations).

$\hat{p}_i$ is the probability the event (a sale, for example) occurs in the *i*th case.

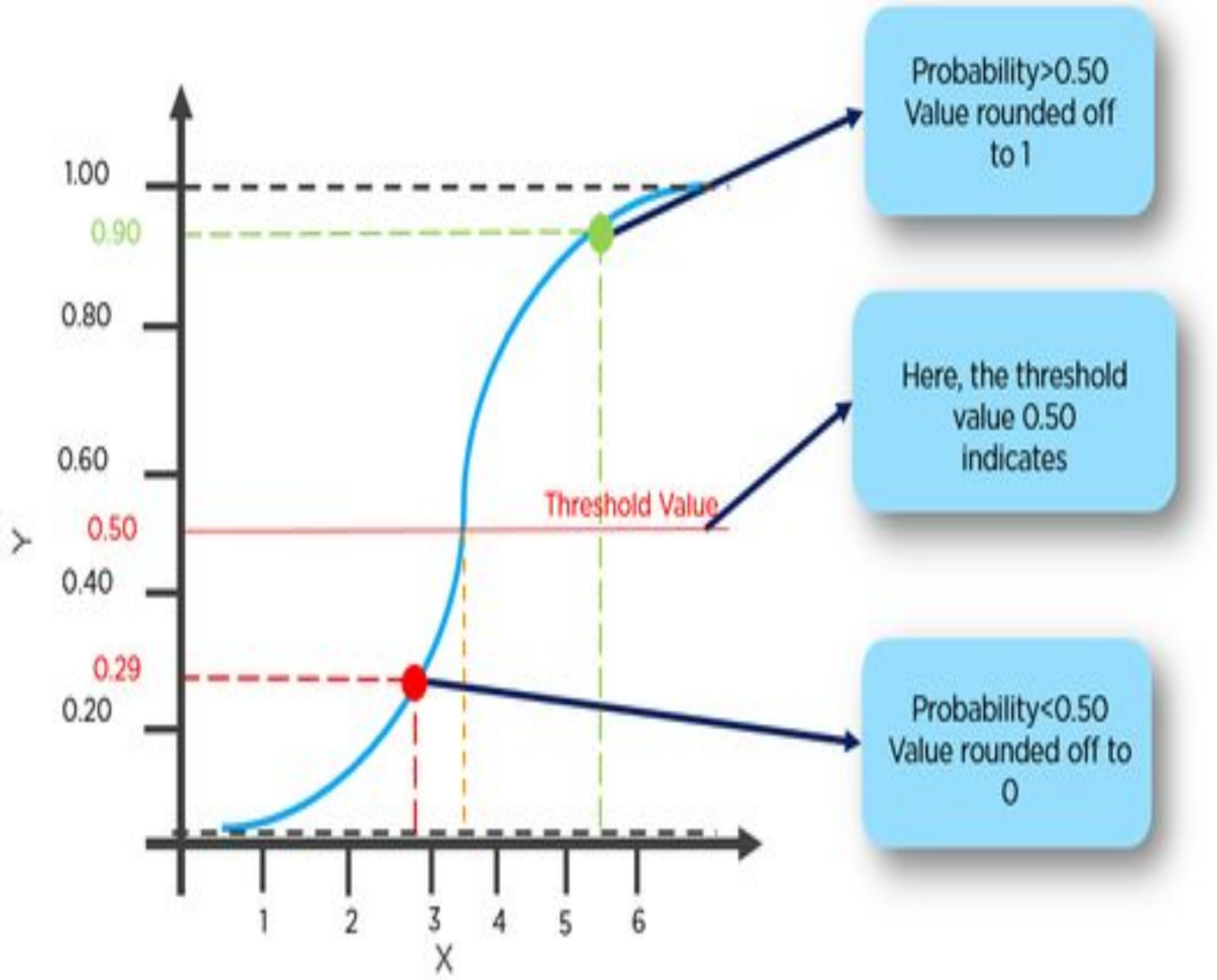
ln is the natural log (to the base e).



Assumption

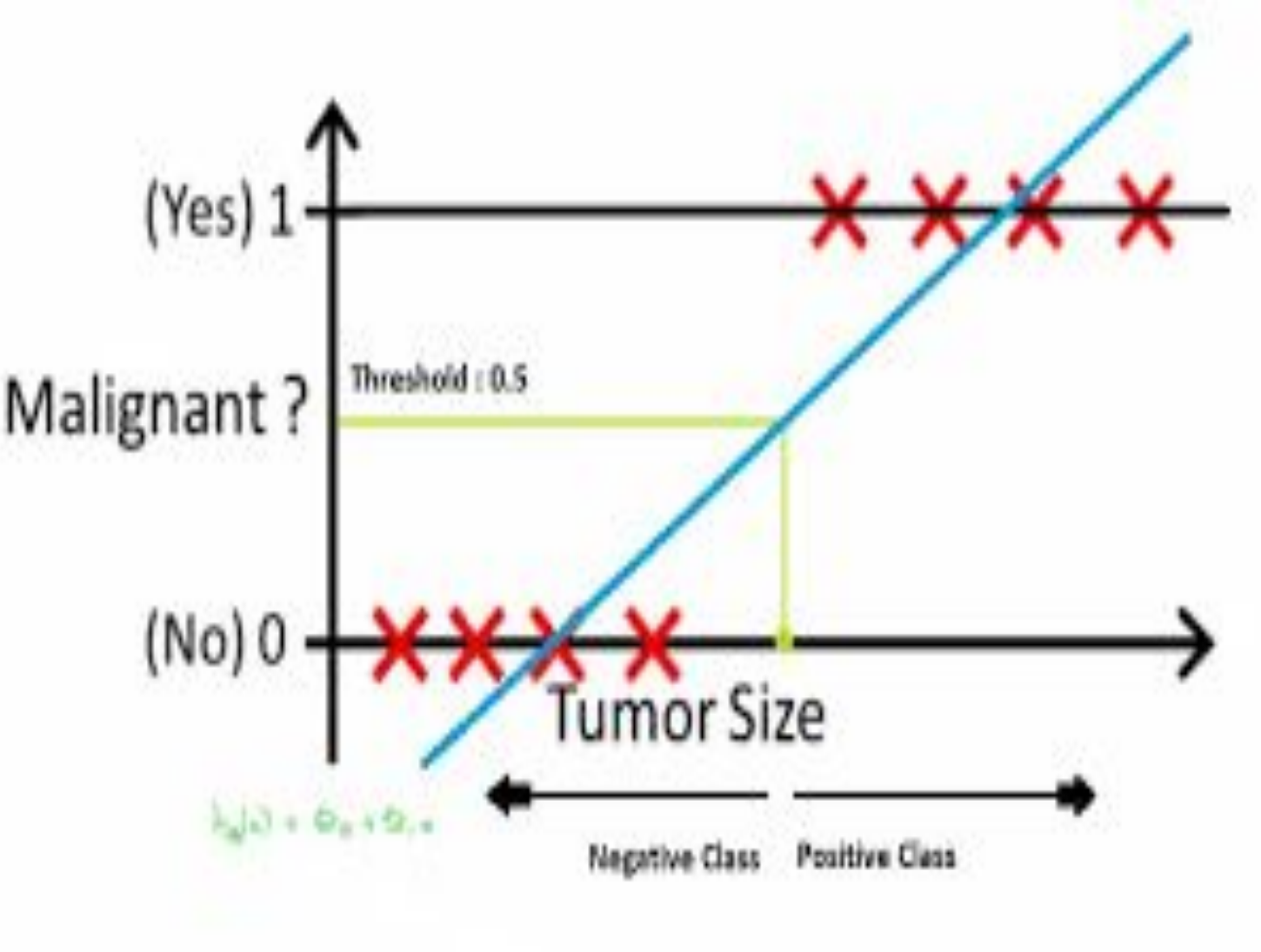


Understanding LR

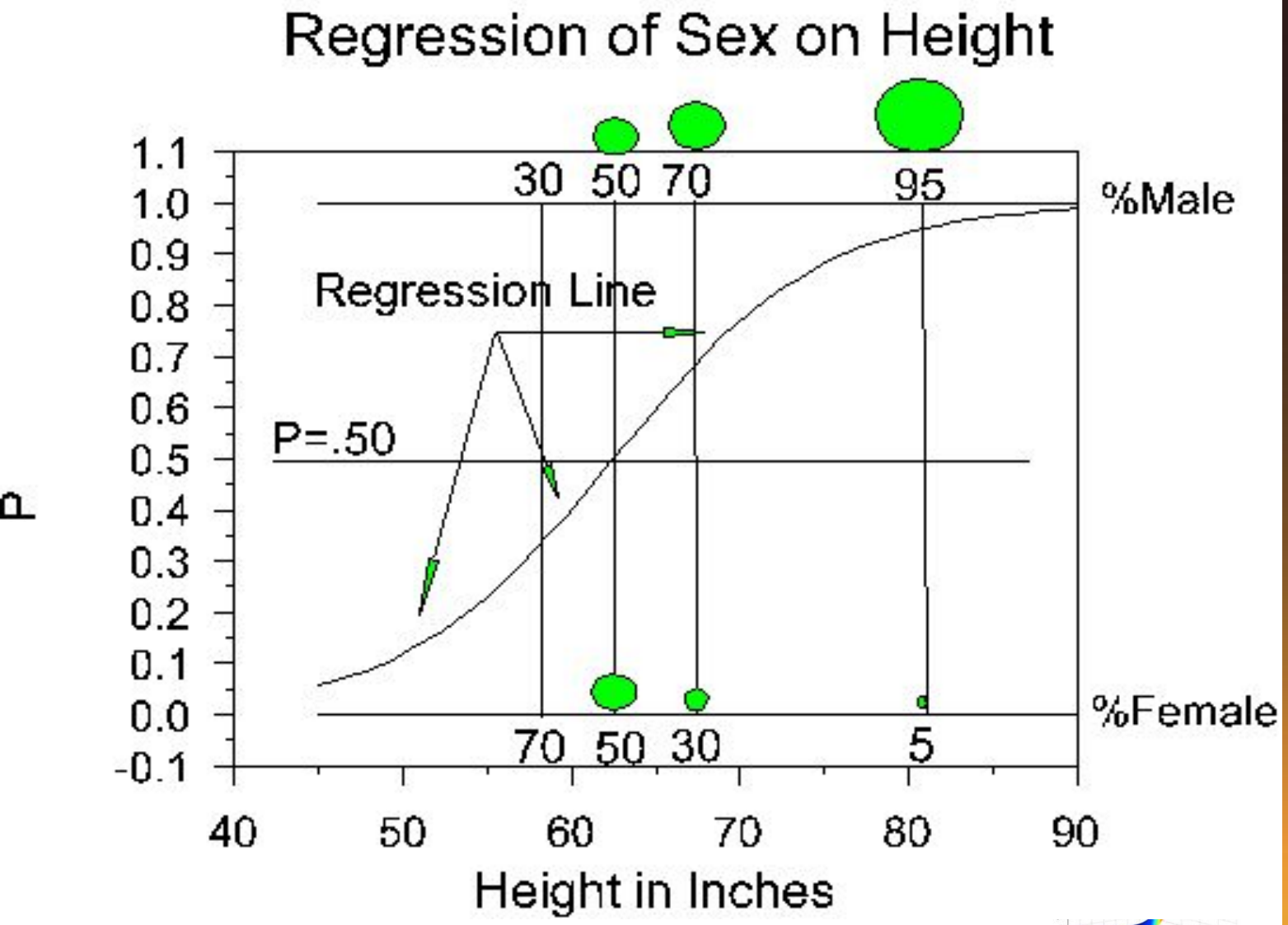


INTERNSHIPSTUDIO

Example- Tumor or not?



Example- Male/Female?



INTERNSHIPSTUDIO

Q.1 How do Logistic Regression works?

Q.2 What is Logit Function?

Q.3 What is Odds ratio?

Q.4 What is Logistic Regression curve?

Q.5 What are Ordinal Variables?