# LOGISTIC REGRESSION

# LINEAR VS LOGISTIC

While Linear regression is used to make predictions on continuous data, logistic is used to determine the probability of an event happening if the event only has two outcomes.

$y = mx + b$

is the equation of linear regression with one independent variable.

If we have more than one independent variable then we will have more intercepts.

But we can't use this to find probability.

# WHEN DO WE USE LOGISTIC REGRESSION?

Whenever we have two outcomes.

Examples:

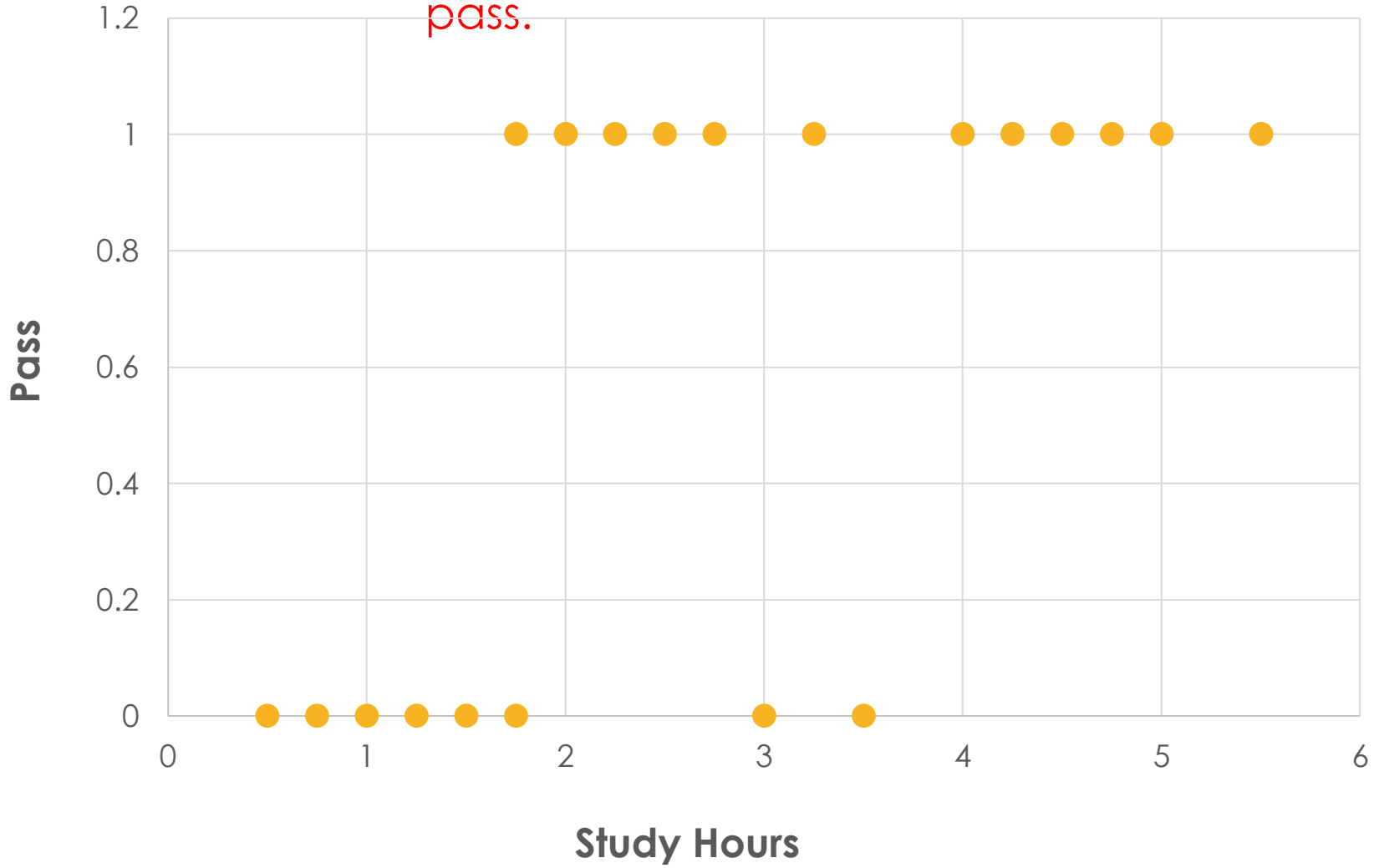
If students studied for a certain time, would they pass an exam or not?

Can a person with a certain credit score, get a home loan or not?

In case of student study time, there are only two outcomes. Either the student passes the exam or does not.

# AN EXAMPLE

Graph of study hours and pass, 0 indicates fail and 1 indicate pass.



# GOALS OF LOGISTIC REGRESSION

**Model and use the model for prediction.**

Create a model using the independent variables to determine the probability of an event occurring

Use the model to predict probability that an observation would fall into one category or other.

# SIGMOID FUNCTION

We consider sigmoid function that would help us find the probability while satisfying the most important properties of probability that is

$$0 \leq p \leq 1$$

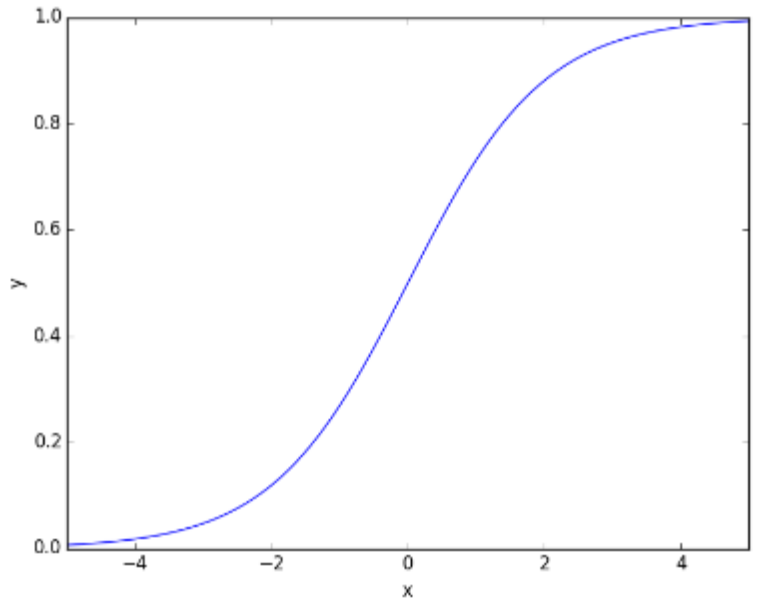
$$\log \text{it}(p) = \log\left(\frac{\text{p}}{1-\text{p}}\right) = y = mx + b$$

Solving this equation for p will give us

$$p = \frac{e^{mx+b}}{1+e^{mx+b}} = \frac{1}{1+e^{-(mx+b\ )}}$$

# SIGMOID FUNCTION GRAPH

Sigmoid function looks like an ‘S’ with values between 0 and 1 making it ideal for logit function.



# CLASSIFICATION

After determining the probability, we can choose a cut-off value to categorize observations

If $p < value$, outcome is 0

If $p \geq value$, outcome is 1

Considering this data from Wikipedia

https://en.wikipedia.org/wiki/Logistic_regre

| Study Hours | Pass |
|---|---|
| 0.5 | 0 |
| 0.75 | 0 |
| 1.00 | 0 |
| 1.25 | 0 |
| 1.50 | 0 |
| 1.75 | 0 |
| 1.75 | 1 |
| 2 | 1 |
| 2.25 | 1 |
| 2.50 | 1 |
| 2.75 | 1 |
| 3.00 | 0 |
| 3.25 | 1 |
| 3.50 | 0 |
| 4.00 | 1 |
| 4.25 | 1 |
| 4.50 | 1 |
| 4.75 | 1 |
| 5.00 | 1 |
| 5.50 | 1 |