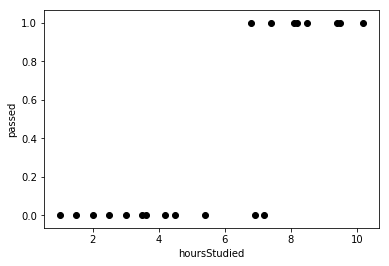
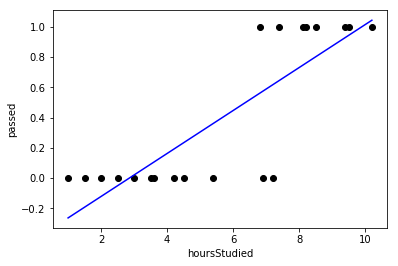
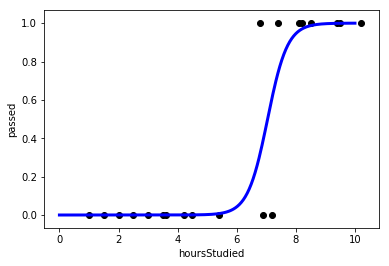
**Logistic Regression**

**What is Logistic Regression :** Logistic regression is used to predict the outcome variable which is categorical.

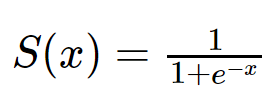
**What is a Categorical variable :** A categorical variable is a variable that can take only specific and limited values.

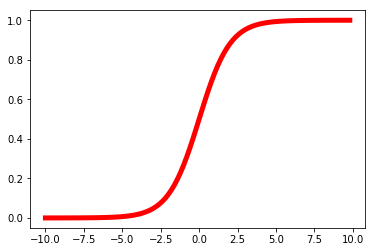
**Example:** Gender : Male/Female | yes/no , 0/1 etc,.

*We know that output should be either 0 or 1.* *We can see that this regression is producing all sort of values between 0 and 1. That's not the actual problem.* *It is also producing impossible values: negative values and values greater than 1 which has no meaning.* *So we need a better regression line than this. Logistic Regression is something we should use here.* *The Logistic regression will fit our data points something like this:*



**The Logistic Function:** Most often, we would want to predict our outcomes as YES/NO (1/0).

 **Where**  
 L - Curve's maximum value  
 k - Steepness of the curve  
 - x value of Sigmoid's midpoint

A standard logistic function is called **sigmoid function** (k=1,x0=0,L=1)

**The sigmoid curve**

*The sigmoid function gives an 'S' shaped curve.* *This curve has a finite limit of:*

*'0' as x approaches*−∞

*'1' as x approaches*+∞

*The output of sigmoid function when x=0 is 0.5*. *Thus, if the output is more tan 0.5 , we can classify the outcome as 1 (or YES) and if it is less than 0.5 , we can classify it as 0(or NO) .*

*For example: If the output is 0.65, we can say in terms of probability as:* *"There is a 65 percent chance that your favorite football team is going to win today”.*

*Thus the output of the sigmoid function cannot be just used to classify YES/NO, it can also be used to determine the probability of YES/NO.*