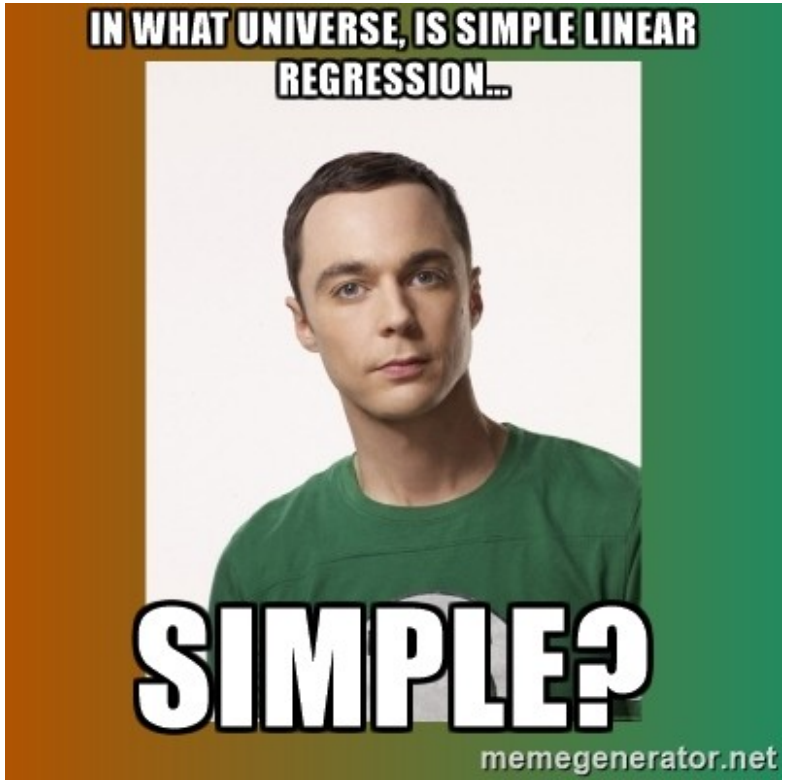
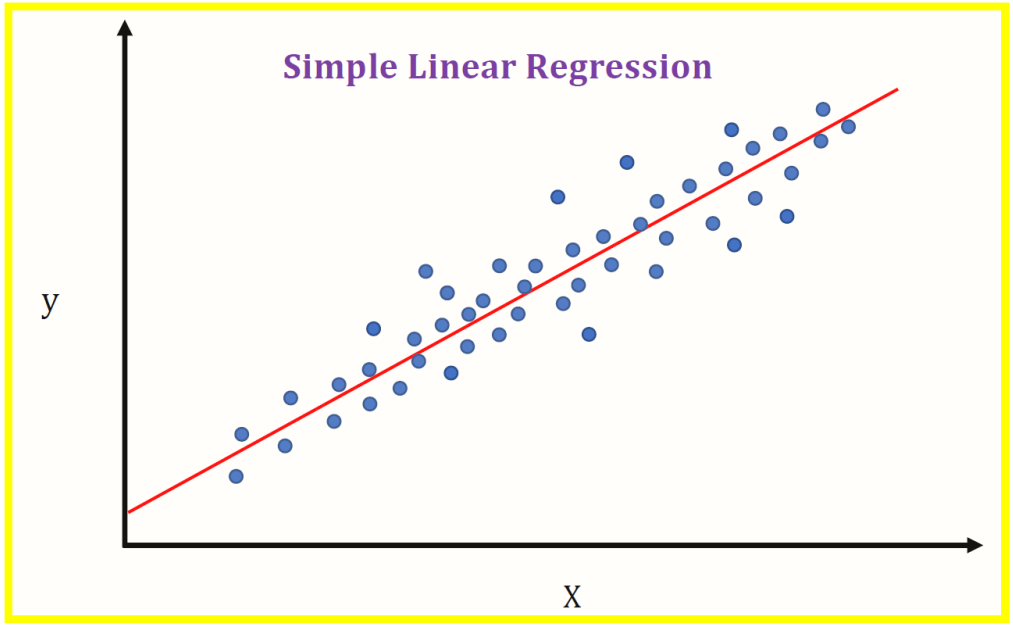
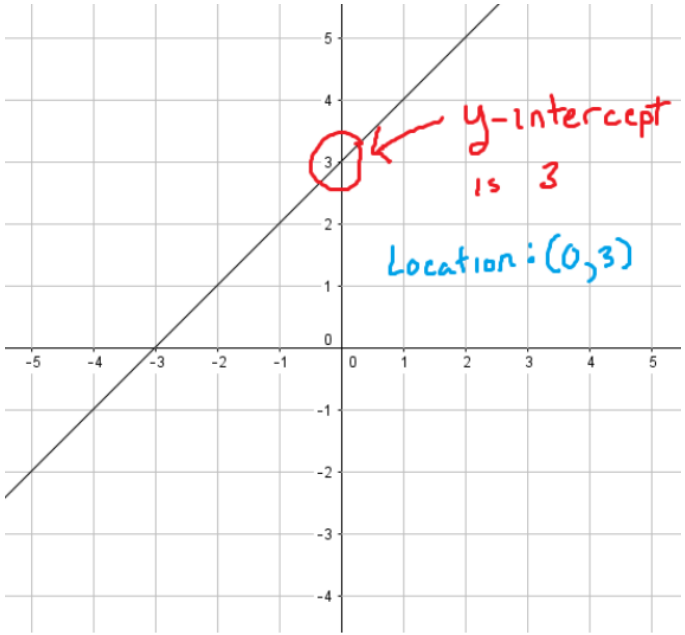
Linear Regression

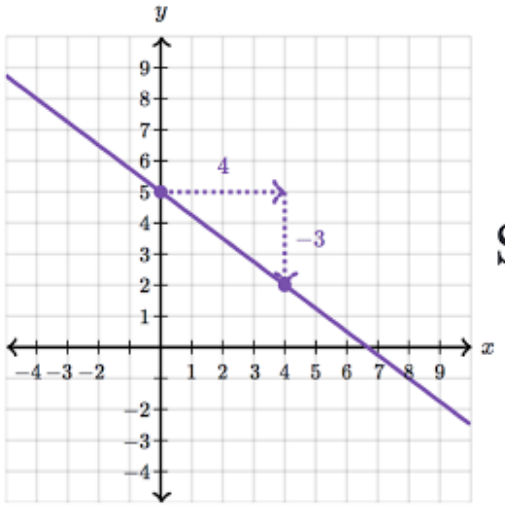


Why?



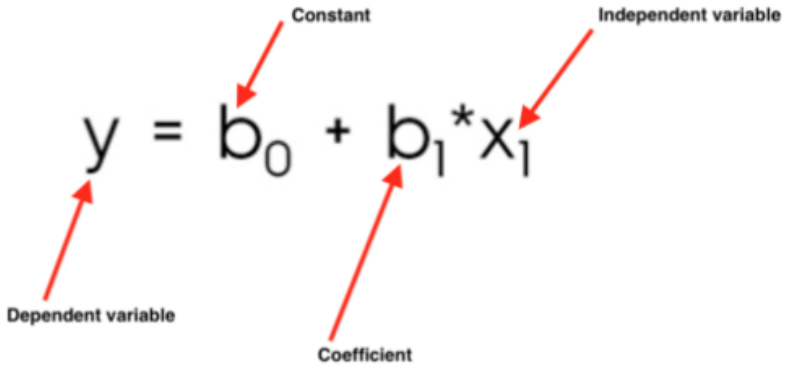
Slope and Intercept

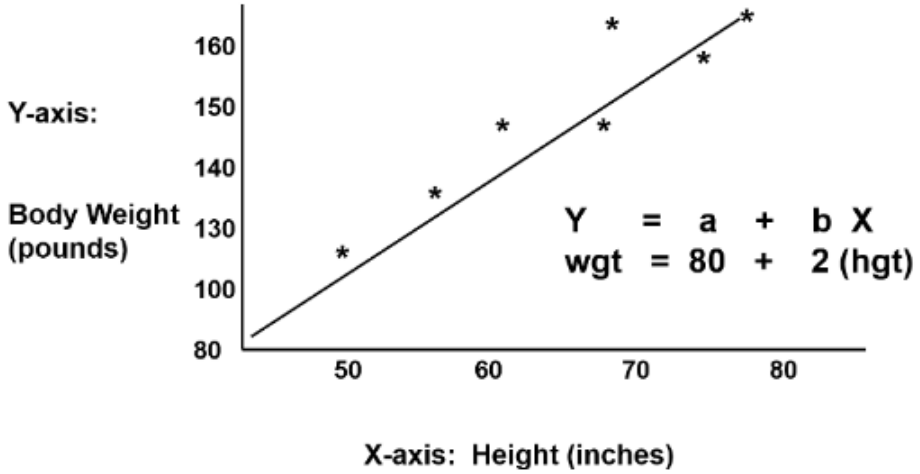




$$\text{Slope} = \frac{\Delta y}{\Delta x}$$

Simple


Constant
Independent variable
y = b0 + b1*x1
Dependent variable
Coefficient


Y-axis:
Body Weight
(pounds)
160
150
140
130
100
80
50
60
70
80
Y = a + b X
wgt = 80 + 2 (hgt)
X-axis: Height (inches)

Residuals

$$S(\hat{\beta}_0, \hat{\beta}_1) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$$S(\hat{\beta}_0, \hat{\beta}_1) = \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$

Perceptron

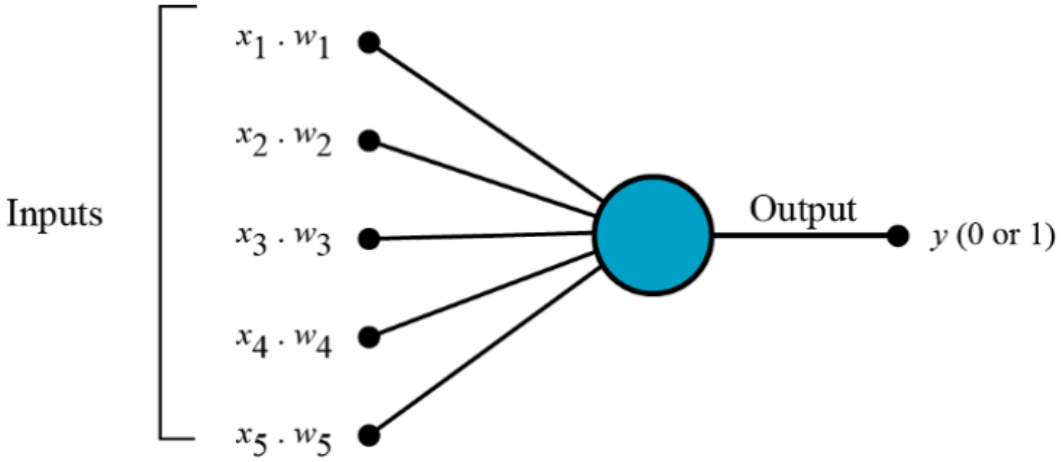


Fig: Multiplying inputs with weights for 5 inputs

Not Simple

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_p x_{ip} + \epsilon$$

where, for $i = n$ observations:

y_i = dependent variable

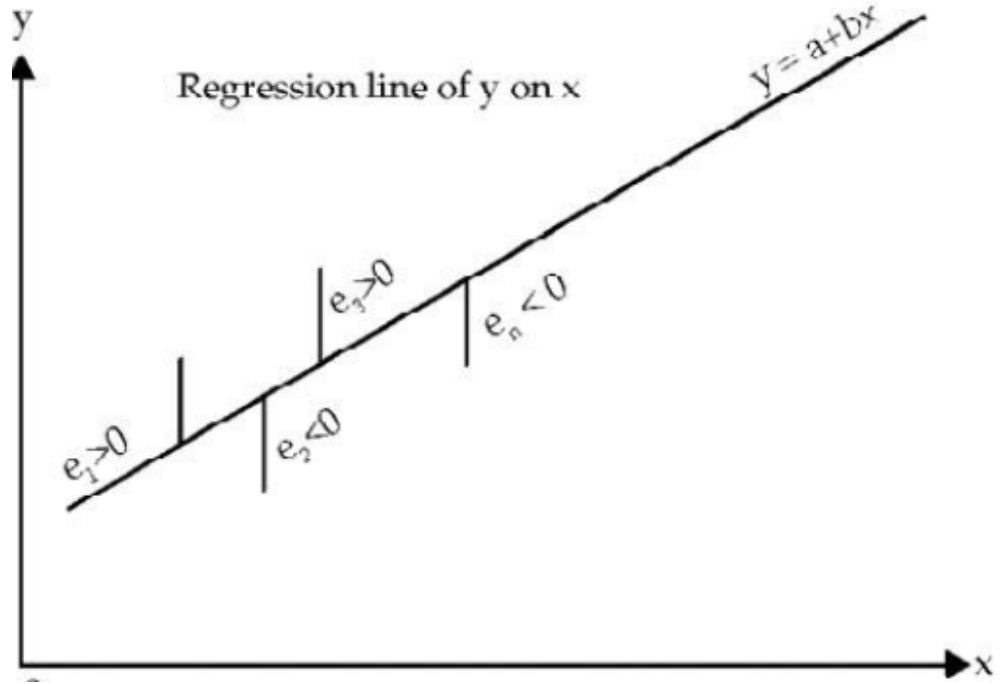
x_i = expanatory variables

β_0 = y-intercept (constant term)

β_p = slope coefficients for each explanatory variable

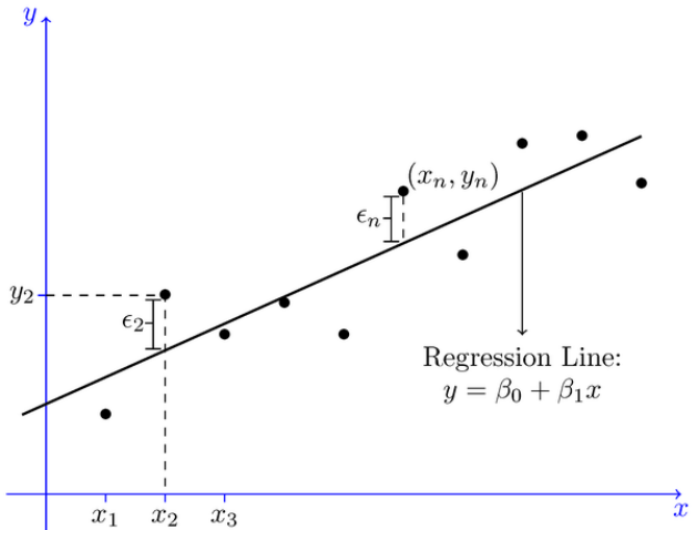
ϵ = the model's error term (also known as the residuals)

Regression line



Loss

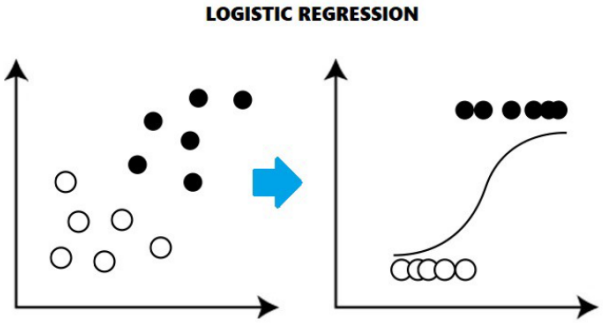
$$J = \frac{1}{n}\sum_{i=1}^{n}(pred_i - y_i)^2$$

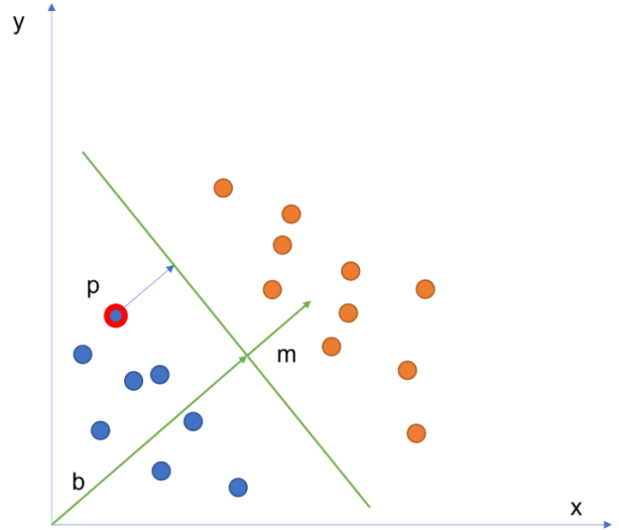


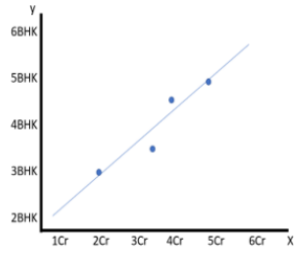
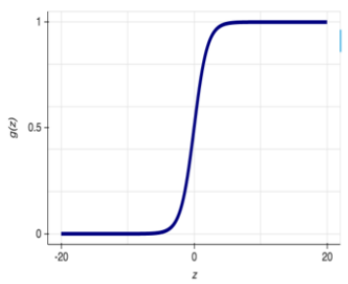
Workflow

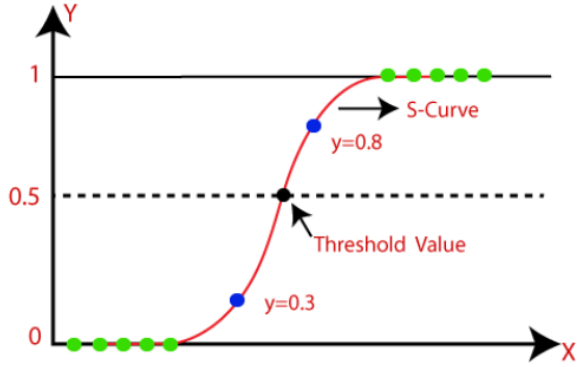

Stop using linear regression
You know what? I'm gonna start using linear regression even harder

Logistic Regression

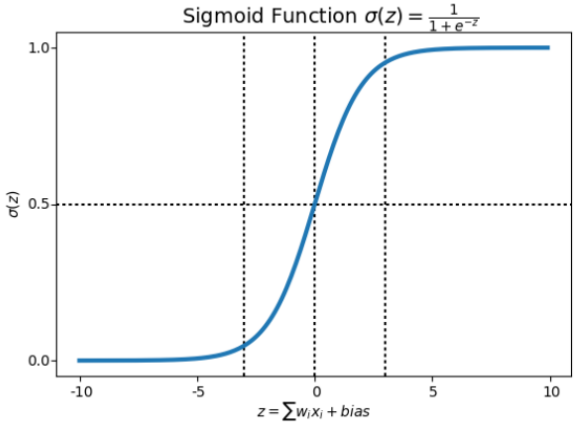




Linear Regression	Logistic Regression
Target is an interval variable	Target is discrete (binary or ordinal) variable
Predicted values are the mean of the target variable at the given values of the input variable	Predicted values are the probability of the particular levels of the given values of the input variable
Solve regression problems	Solve classification problems
Example : What is the Temperature?	Example : Will it rain or not?
Graph is straight line	Graph is S-curve
	



Sigmoid



Sigmoid Function Graph

$$f(x) = \frac{1}{1+e^{-(x)}}$$