CLASSIFICATION
LOGISTIC REGRESSION

Mr.Gangadhar Immadi

immadi.gangadhar@gmail.com

9986789040

Logistic Regression - Introduction

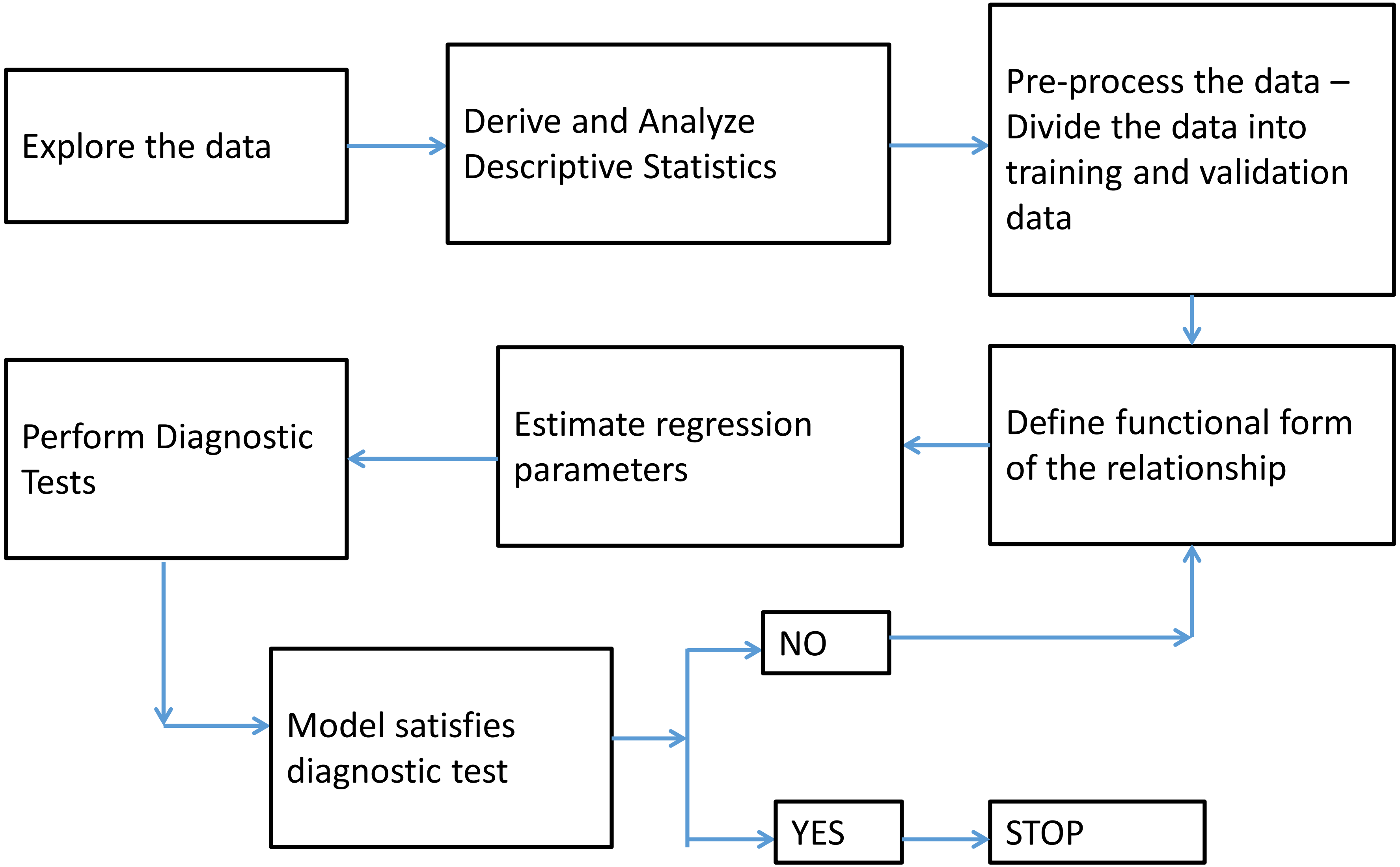
- Logistic regression models estimate how probability of an event may be affected by one or more explanatory variables.

- Logistic regression is a technique used for predicting "**class probability**", that is the probability that the case belongs to a particular class.

- Binomial (or binary) logistic regression is a model in which the dependent variable produces 2 outputs.

- In multinomial logistic regression model, the dependent variable can take more than two values.
- Linear Regression v/s Logistic Regression

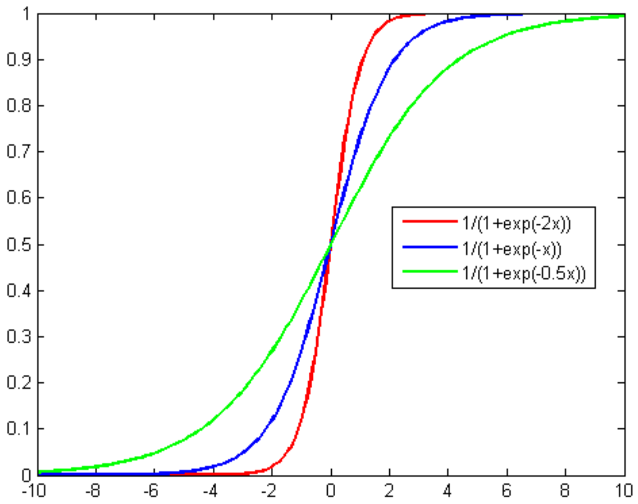
Logistic Regression Model Development



Logistic Regression - Introduction

- The name logistic regression emerges from logistic distribution function.

$$\frac{e^{Z}}{1+e^{Z}}$$



- Mathematically, logistic regression attempts to estimate conditional probability of an event (or class probability).

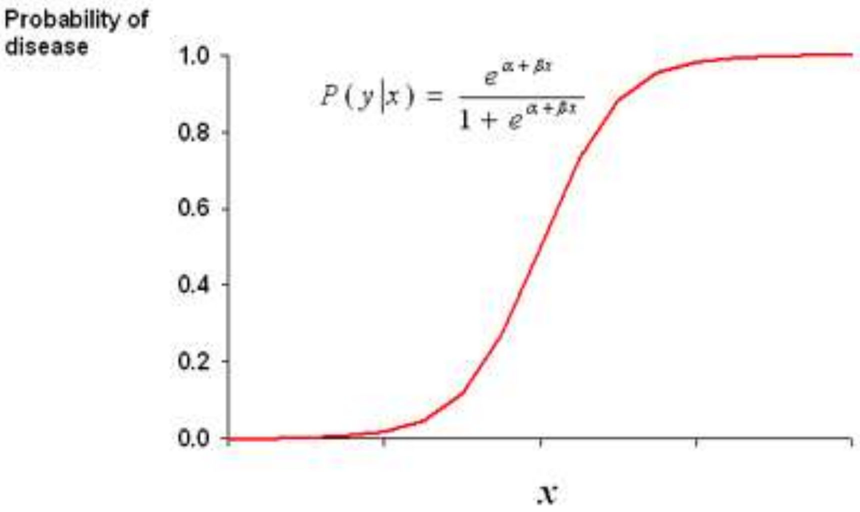
Logistic Function (Sigmoidal function)

$$P(Y=1)=\pi(z)=\frac{1}{1+e^{-z}}=\frac{e^{z}}{1+e^{z}}$$

$$z=\beta_0+\beta_1 x_1+\beta_2 x_2+\ldots+\beta_n x_n$$

Logistic Regression with one Explanatory Variable

$$P(Y=1\mid X=x)=\pi(x)=\frac{\exp(\alpha+\beta x)}{1+\exp(\alpha+\beta x)}$$



b = 0 implies that P(Y|x) is same for each value of x
b > 0 implies that P(Y|x) is increases as the value of x increases
b < 0 implies that P(Y|x) is decreases as the value of x increases

Logistic Transformation

- The logistic regression model is given by:

$$\pi_i = \frac{e^{(\beta_0 + \beta_1 X_i)}}{1 + e^{(\beta_0 + \beta_1 X_i)}}$$

$$\frac{\pi_i}{1 - \pi_i} = e^{(\beta_0 + \beta_1 X_i)}$$

$$P(\text{Heart Disease} = 1) = \frac{e^{\beta_0 + \beta_1 Age}}{1 + e^{\beta_0 + \beta_1 Age}}$$

$$\ln\left(\frac{\pi_i}{1 - \pi_i}\right) = \beta_0 + \beta_1 X_i$$

Function with linear properties (Link Function)

Logistic Regression of challenger data

Let:

- $Y_i = 0$ denote no damage
- $Y_i = 1$ denote damage to the O-ring
- $P(Y_i = 1) = \Pi_i$ and $P(Y_i = 0) = 1 - \Pi_I$.
- We have to estimate $P(Y_i = 1 | X_i)$.

Variables in the Equation

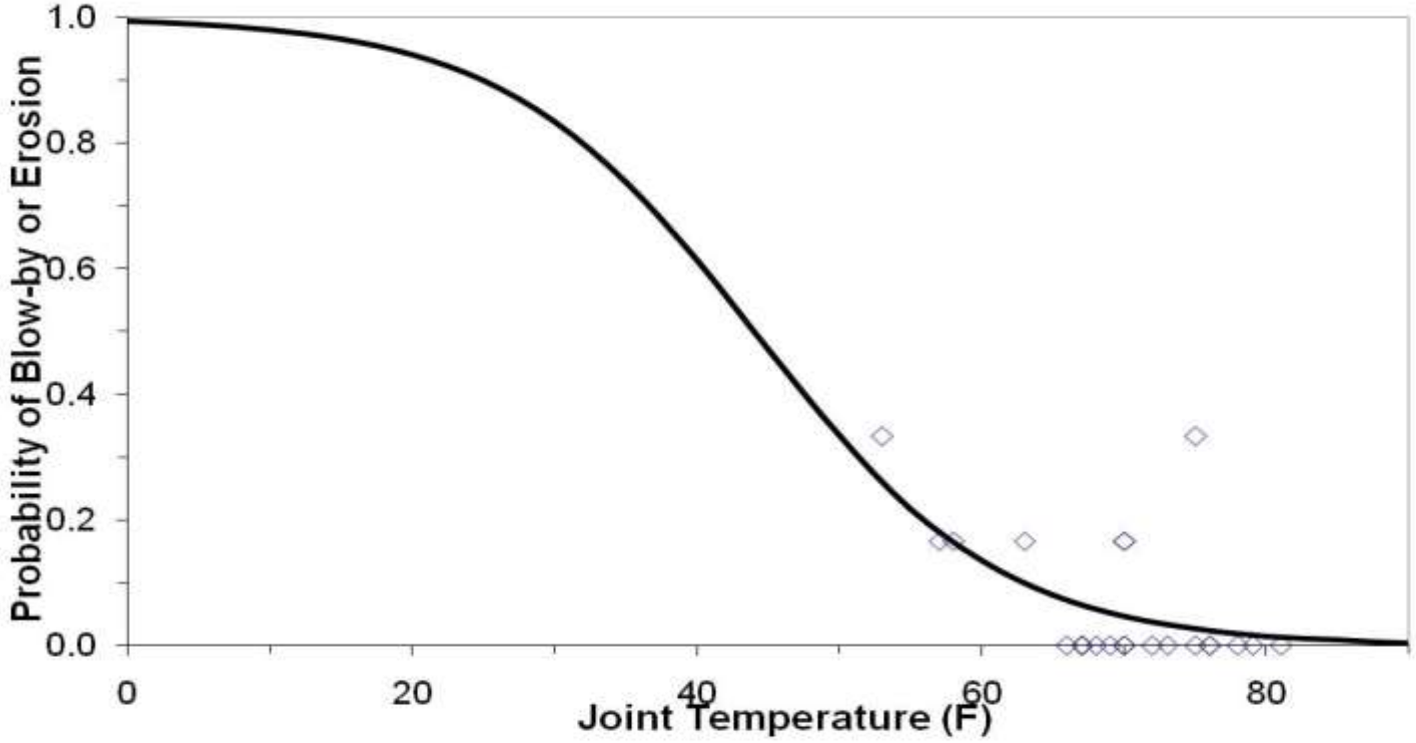
		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	LaunchTemperature	-.236	.107	4.832	1	.028	.790
	Constant	15.297	7.329	4.357	1	.037	4398676

a. Variable(s) entered on step 1: LaunchTemperature.

$$\ln\left(\frac{\pi_i}{1-\pi_i}\right) = 15.297 - 0.236 X_i$$

Probability of failure estimate

$$\pi_i = \frac{e^{15.297 - 0.236 X_i}}{1 + e^{15.297 - 0.236 X_i}}$$



Using Logistic Regression Predict the class label for the following dataset as Positive or Negative at Cut-off probability 0.5, & parameters B0= - 42.54,B1=2.95 and B2 = 10.4,

Observation#	**1**	**2**	**3**	**4**	**5**
X1	1	2	3	4	5
X2	3	3	4	4	3
Y (class label)	?	?	?	?	?

1. 0.00001237 = -ve class label
2. 3.53×10^{-24} = -ve class label
3. 0.99 = +ve class label
4. 0.91 = +ve class label
5. 0.95 = +ve class lebel