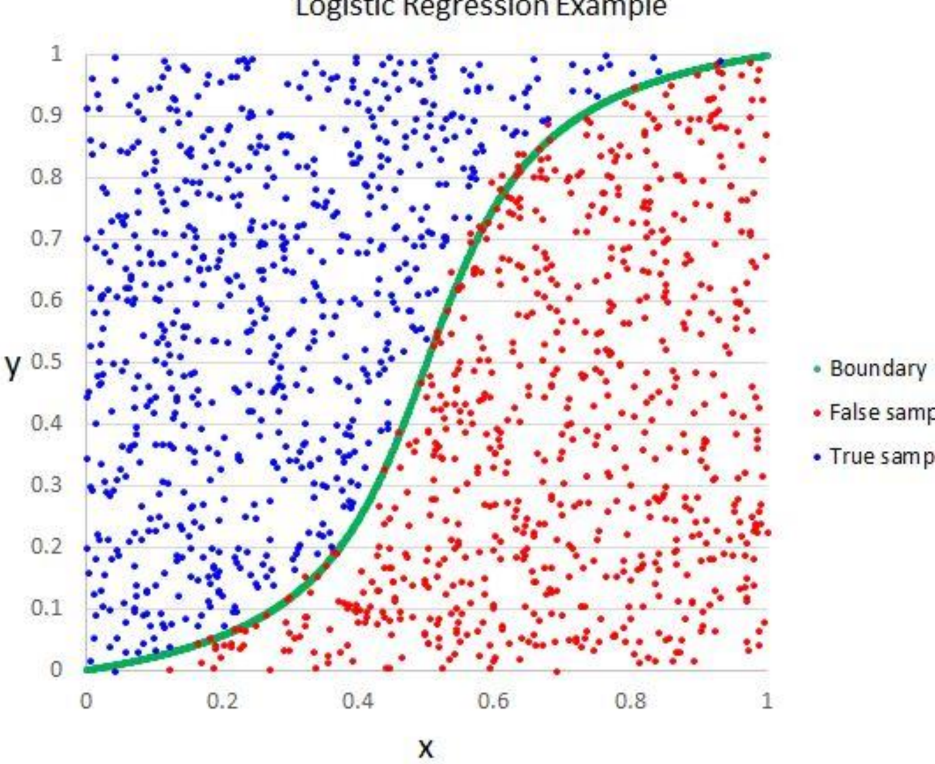
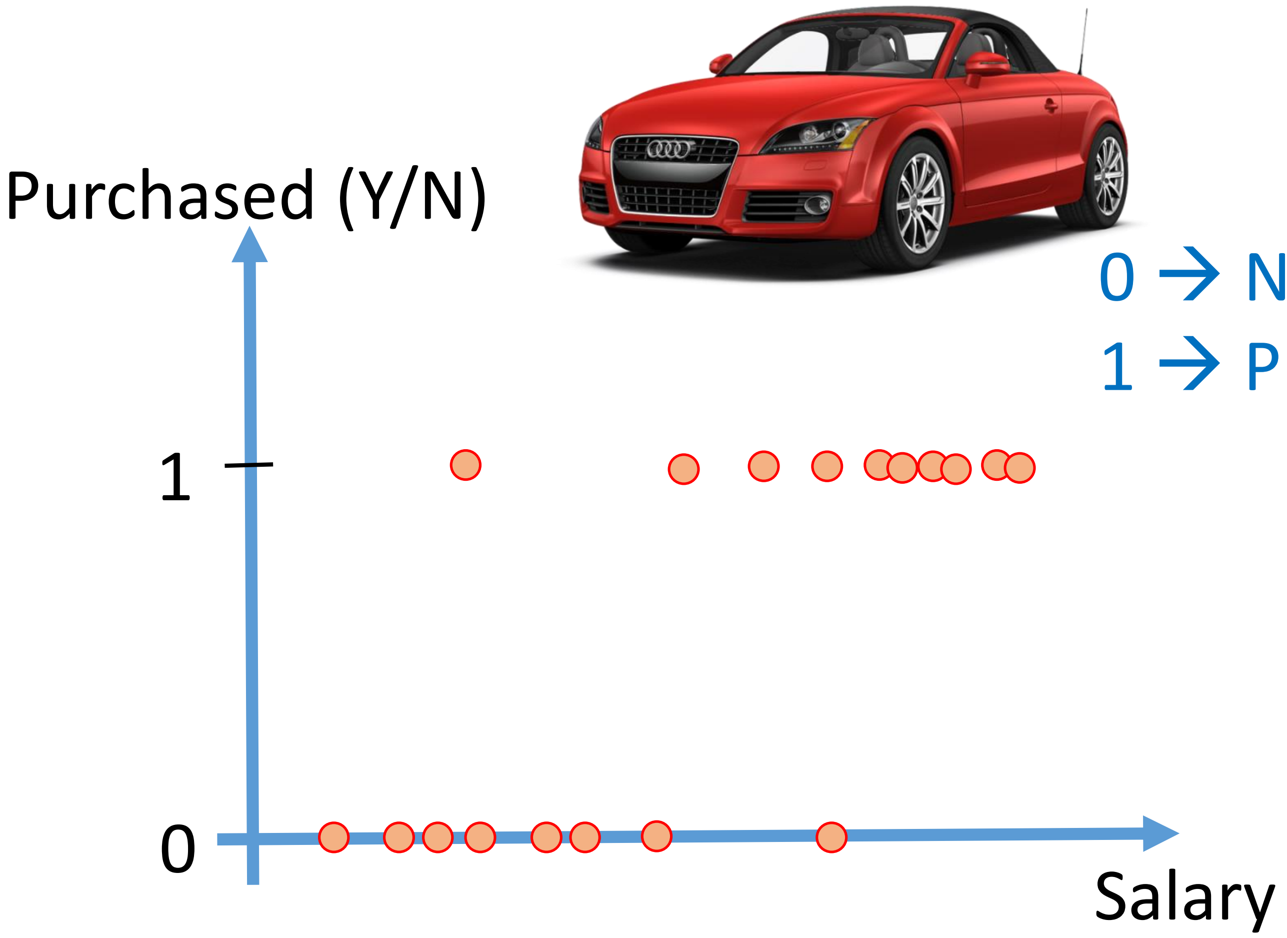


Logistic Regression

Retail (Likelihood to Purchase)



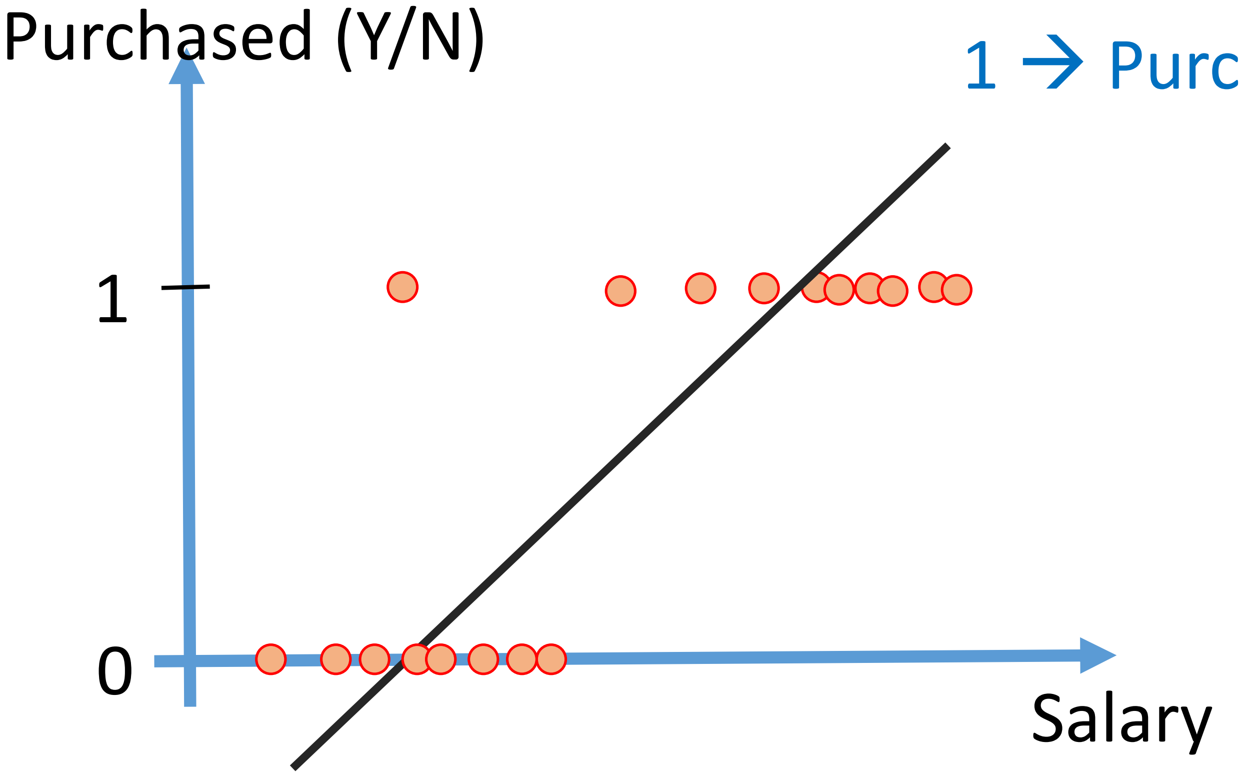
0 → Non-Purchased

1 → Purchased

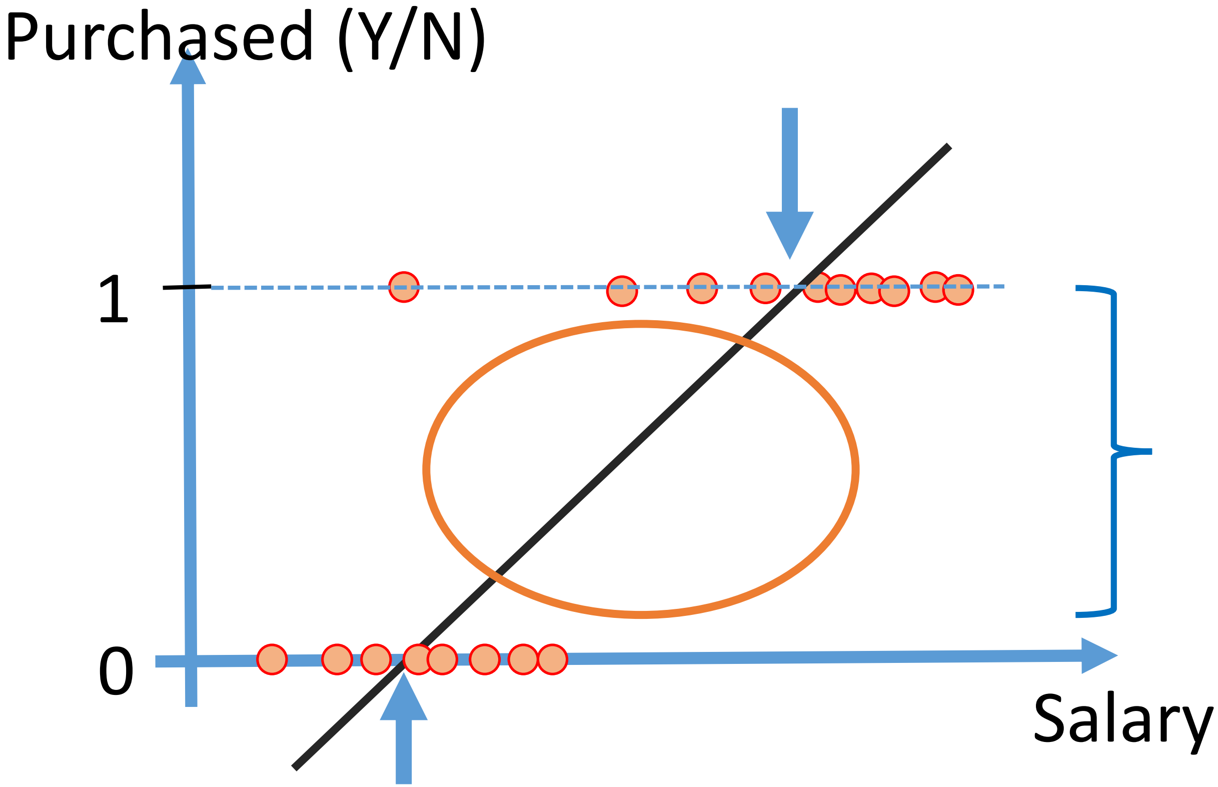
Linear Regression Prediction in Retail (Like hood to purchase)

0 → Non-Purchased

1 → Purchased



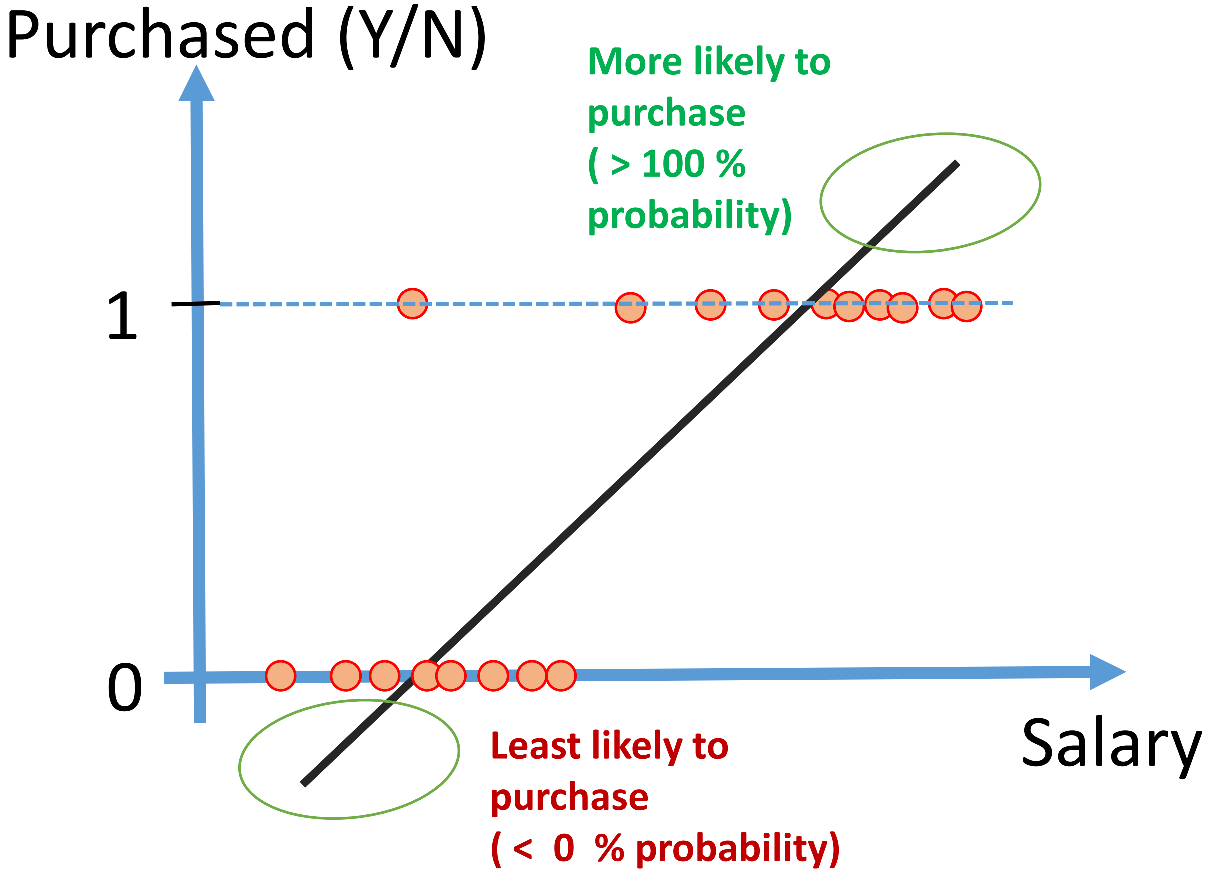
Probability of Likely hood to purchase



Gives probability scores,

- As “Salary” increases there is more likely hood that he/she can purchase a CAR

Probability of Likely hood to purchase

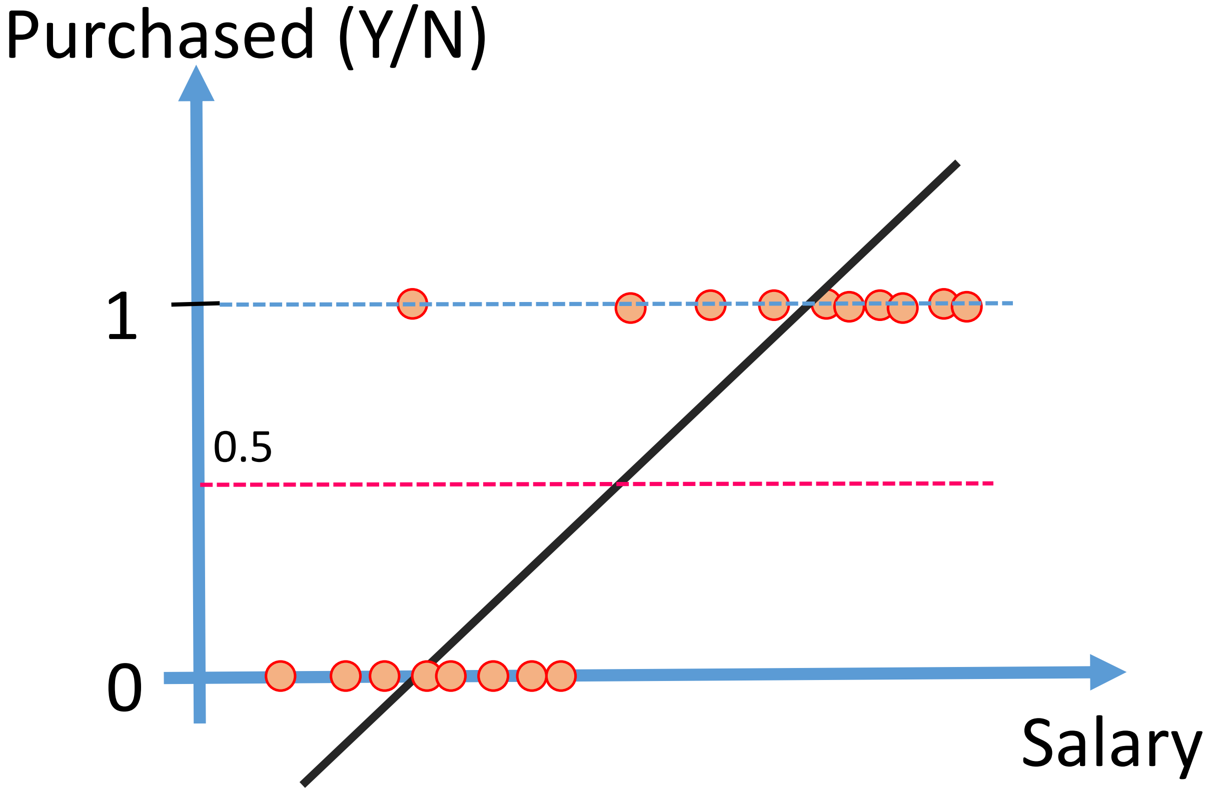


More likely to purchase
(> 100 % probability)

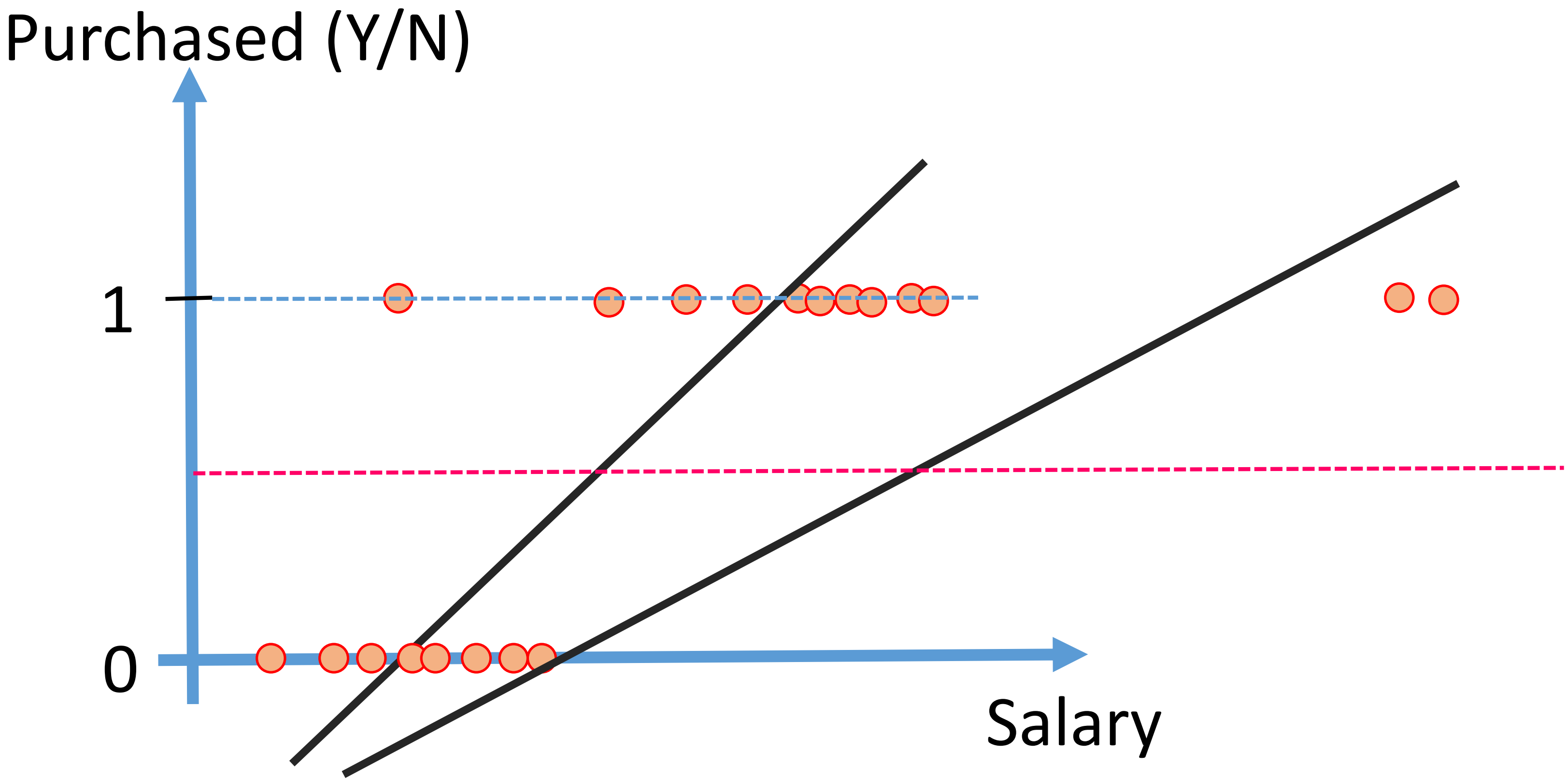
Least likely to purchase
(< 0 % probability)

The Line above and below the probability limit is telling that there is more likely to purchased and not-purchased

Linear Regression



Linear Regression could fail



Attention:

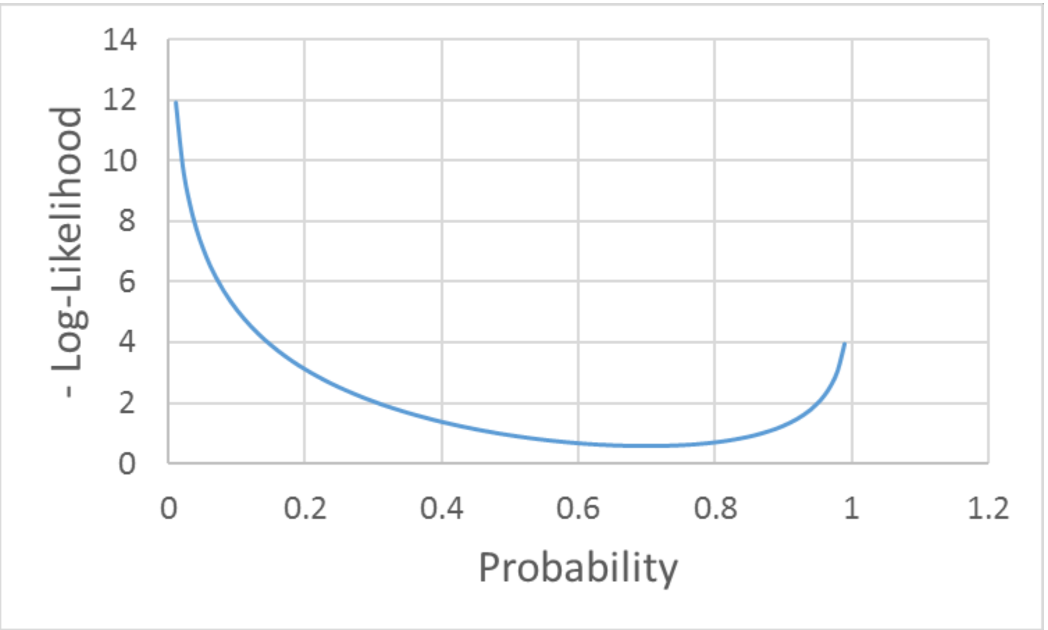
Linear regression slopes can be much larger than 1 or much smaller than zero and hence thresholding becomes difficult.

Attention

- Error terms do not follow normal distribution.
- Error terms are not independent.
- Error variances are heteroscedastic.
- Least Squares is inappropriate. Maximum Likelihood Estimation (MLE) is used instead.

MLE

- Goal is to maximize likelihood.
- In most Data Science optimizations, the goal is to find minima using calculus (minimize sum of squared errors in linear regression, and so on) or numerical techniques like Gradient Descent (minimize deviance in logistic regression, and so on).
- Maximum Likelihood => Minimum of Negative Log-Likelihood.

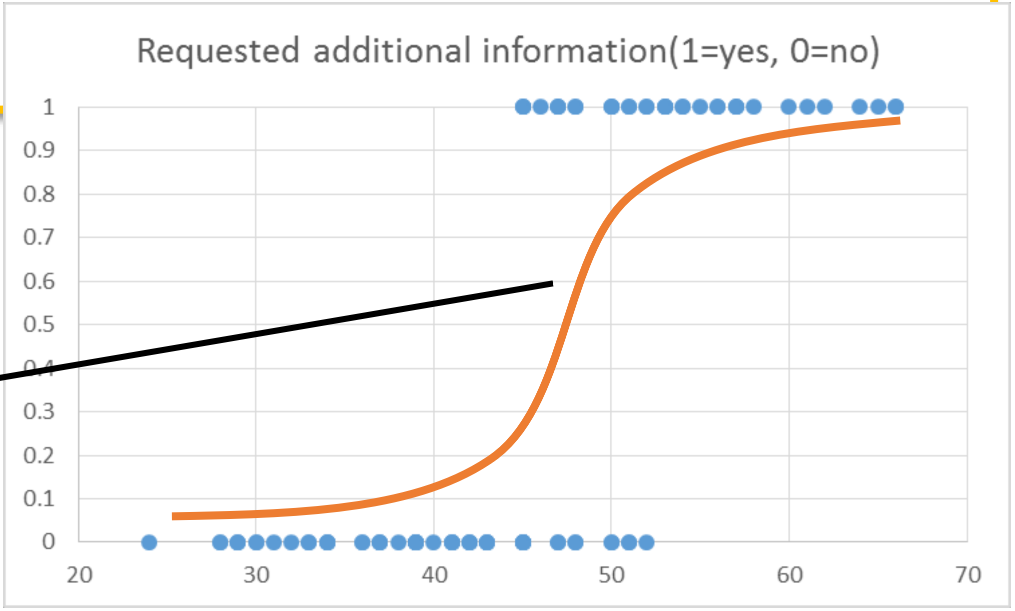


Example

An auto club mails a flier to its members offering to send more information regarding a supplemental health insurance plan if the member returns a brief enclosed form.

Can a model be built to predict if a member will return the form or not?

Example



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

Where $\mu = \beta_0 + \beta_1 x_1$(also know as the systematic or the structural component or linear predictor.

- This is a logistic model. The function is also known as the inverse link function, which links the response with the systematic component.
- *p* is the probability that a club member fits into group 1 (returns the form; success; P(Y=1|X)).

Logistic Model

$$f(x) = p = \frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n}}{1 + e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n}}$$

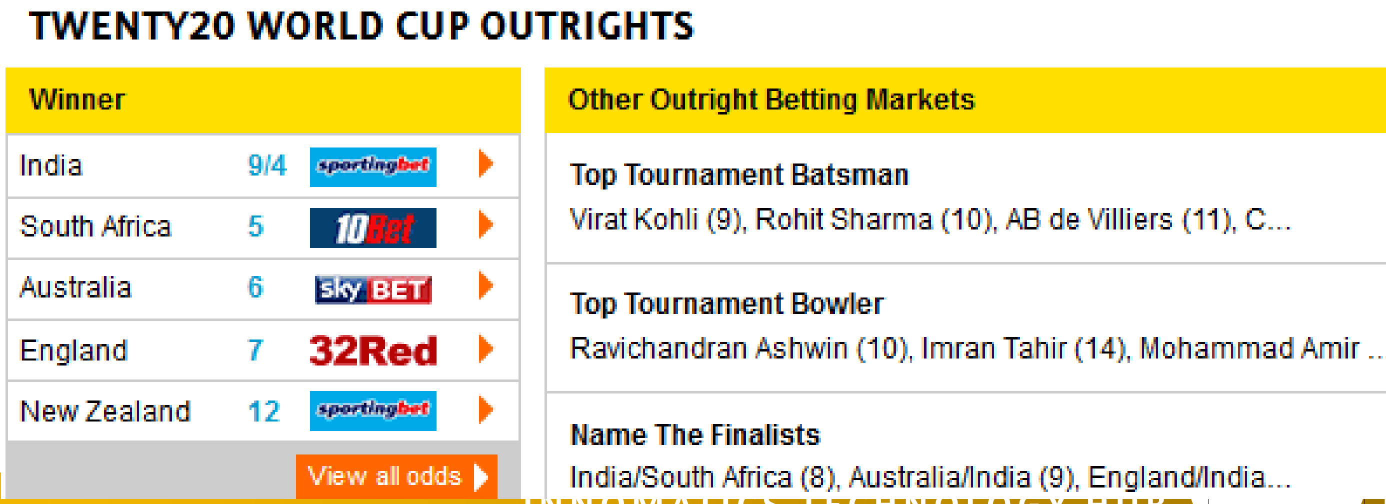
Odds Ratio is obtained by the probability of an event occurring divided by the probability that it will not occur.

Logistic model can be transformed into an odds ratio:

$$S = Odds\ ratio = \frac{p}{1-p}$$

Attention Check – Probability and Odd

If the probability of winning is 6/12, what are the odds of winning ?	1:1 (Note, the probability of losing also is 6/12)
If the odds of winning are 13:2, what is the probability of winning?	13/15
If the odds of winning are 3:8, what is the probability of losing?	8/11
If the probability of losing is 6/8, what are the odds of winning?	2:6 or 1:3



TWENTY20 WORLD CUP OUTRIGHTS

Winner		
India	9/4	sportingbet
South Africa	5	10Bet
Australia	6	sky BET
England	7	32Red
New Zealand	12	sportingbet

Other Outright Betting Markets

Top Tournament Batsman
Virat Kohli (9), Rohit Sharma (10), AB de Villiers (11), C...

Top Tournament Bowler
Ravichandran Ashwin (10), Imran Tahir (14), Mohammad Amir ..

Name The Finalists
India/South Africa (8), Australia/India (9), England/India...

Logistic Model

S = odds ratio = $\frac{p}{1-p}$

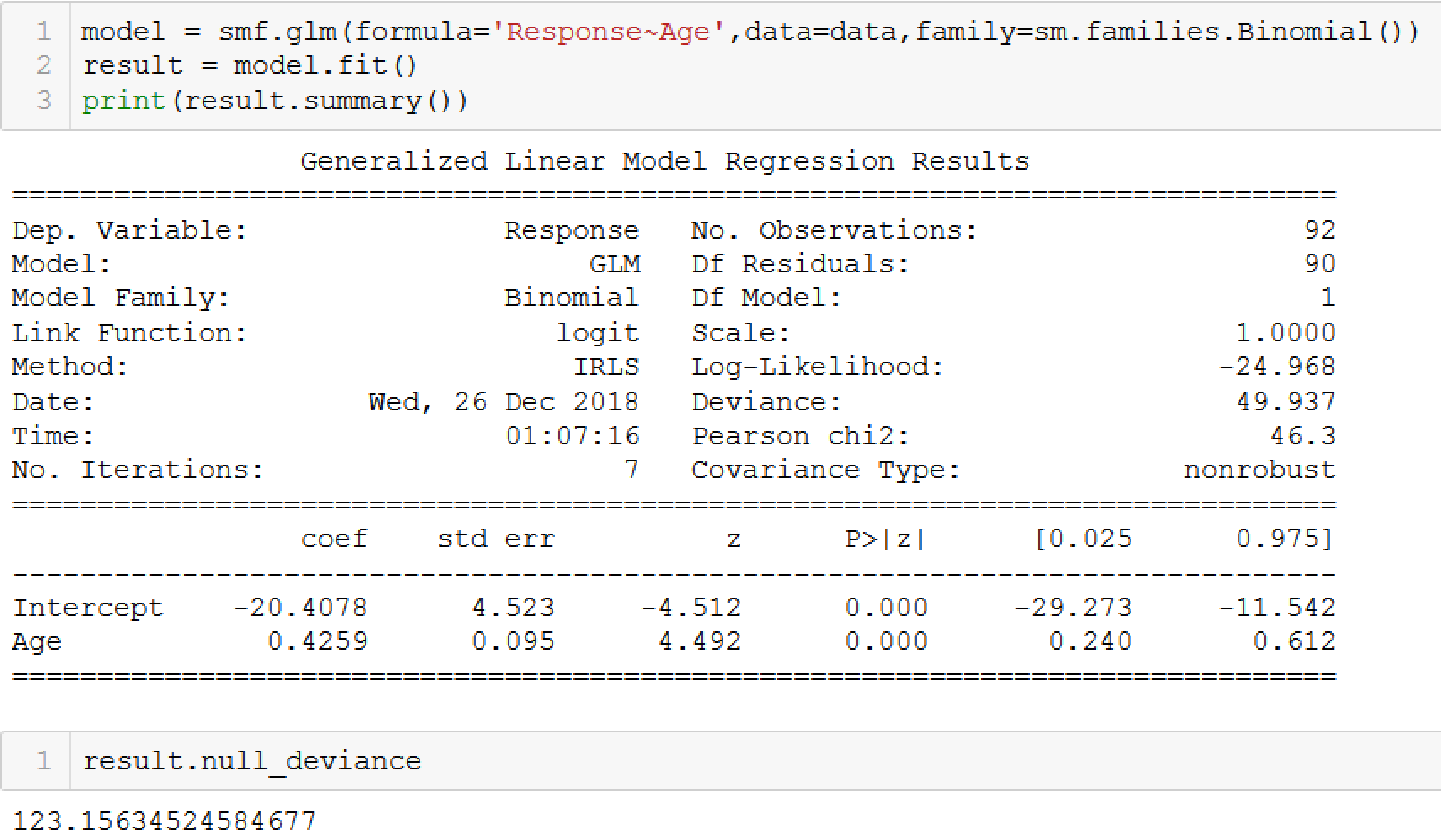
$$s = \frac{\left(\frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n}}{1+e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n}}\right)}{1-\left(\frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n}}{1+e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n}}\right)}$$

$$\therefore, S = e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n}$$

$$\ln(S) = \ln\left(e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

Logistic Model

The log of the odds ratio is called logit, and the transformed model is linear in βs.



```
model = smf.glm(formula='Response~Age',data=data,family=sm.families.Binomial())
result = model.fit()
print(result.summary())
```

```
                 Generalized Linear Model Regression Results
==============================================================================
Dep. Variable:               Response   No. Observations:                   92
Model:                            GLM   Df Residuals:                       90
Model Family:                Binomial   Df Model:                            1
Link Function:                  logit   Scale:                          1.0000
Method:                          IRLS   Log-Likelihood:                -24.968
Date:                Wed, 26 Dec 2018   Deviance:                       49.937
Time:                        01:07:16   Pearson chi2:                     46.3
No. Iterations:                     7   Covariance Type:             nonrobust
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept    -20.4078      4.523     -4.512      0.000     -29.273     -11.542
Age            0.4259      0.095      4.492      0.000       0.240       0.612
==============================================================================
```

```
result.null_deviance
```

```
123.15634524584677
```

What is logit equation ?

$$\ln(s) = -20.40782 + 0.42592 * Age$$

Determining Logistic Regression Model

Suppose we want a probability that a 50-year old club member will return the form.

ln S = -20.40782 + 0.42592 * 50 = 0.89

S = $e^{0.89} = 2.435$

The odds that a 50-year old returns the form are 2.435 to 1.

Determining Logistic Regression Model

$$\hat{p} = \frac{s}{s+1} = \frac{2.435}{2.435+1} = 0.709$$

Using a probability of 0.50 as a cut-off between predicting a 0 or a 1, this member would be classified as a 1.

The output of the logistic regression forecast is a probability value. One needs to decide on a threshold value before a class is assigned.

Reference

Head First Statistics