DATA ANALYTICS

Unit 2:Logistic Regression

Mamatha.H.R
Department of Computer Science and Engineering

DATA ANALYTICS

Unit 2:Logistic Regression

Mamatha H R, Gowri Srinivasa

Department of Computer Science and Engineering

Contents

estimates the probability

Supervisied ML method for classification.

$odd_1 = \frac{P_1}{1-P}$ $odd_2 =$ $\frac{odd_2}{odd_1}$

Oddratio → interpret Log Reg

Classification Problems

- Classification is an important category of problems in which the decision maker would like to classify the case/entity/customers into two or more groups

- Examples of Classification Problems:

  - Customer profiling (customer segmentation)
  - Customer Churn
  - Credit Classification (low, high and medium risk)
  - Employee attrition
  - Fraud (classification of transaction to fraud/no-fraud)
  - Stress levels
  - Text Classification (Sentiment Analysis)
  - Outcome of any binomial and multinomial experiment

1. Logistic Regression

- Logistic regression predicts the probability of an outcome that can only have two values (i.e. a dichotomy).
- The prediction is based on the use of one or several predictors (numerical and categorical).

  x -> numerical

- A linear regression is not appropriate for predicting the value of a binary variable for two reasons:

  1. A linear regression will predict values outside the acceptable range (e.g. predicting probabilities outside the range 0 to 1)

  2. Since the dichotomous experiments can only have one of two possible values for each experiment, the residuals will not be normally distributed about the predicted line.

1.Logistic Regression

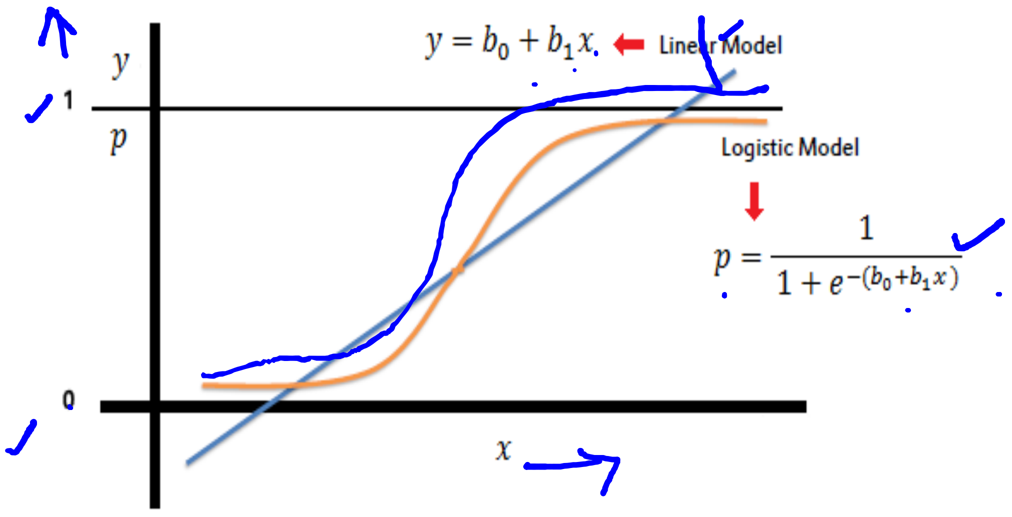
— Sigmoid function

On the other hand, a logistic regression produces a logistic curve, which is limited to values between 0 and 1.

Logistic regression is similar to a linear regression, but the curve is constructed using the natural logarithm of the "odds" of the target variable, rather than the probability.

Moreover, the predictors do not have to be normally distributed or have equal variance in each group.

$$\frac{p}{1-p} = \exp\left(b_0 + b_1 x\right)$$



In the logistic regression the constant (b_0) moves the curve left and right and the slope (b_1) defines the steepness of the curve. By simple transformation, the logistic regression equation can be written in terms of an odds ratio.

2. ODDS and ODDS Ratio

I) Probability $P = \frac{\text{Outcome of interest}}{\text{all possible outcomes}}$

Example

1. Fair Coin flip $P(\text{heads}) = \frac{1}{2} = 0.5$
2. Fair die roll $P(1 \text{ or } 2) = \frac{2}{6} = \frac{1}{3} = 0.333$

II Odds

$$\text{odds} = \frac{P(\text{occurring})}{P(\text{not occurring})} = \frac{P}{1-P}$$

1) $\text{odds}(\text{heads}) = \frac{0.5}{1-0.5} = \frac{0.5}{0.5} = 1$

2) $\text{odds}(1 \text{ or } 2) = \frac{0.333}{1-0.333} = \frac{0.333}{0.666} = \frac{1}{2} = 0.5$

Odds Ratio

It is a ratio of two odds

1) Fair coin flip

$P(\text{heads}) = 0.5$

$\text{Odds}(\text{heads}) = 1$

2) Loaded coin flip

$P(\text{heads}) = \frac{7}{10} = 0.7$

$\text{Odd}(\text{heads}) = \frac{0.7}{1-0.7} = \frac{0.7}{0.3} = 2.333$

Odds ratio $\rightarrow$ important logistic regression.

between loaded coin : fair coin

$$\text{oddsratio} = \frac{\cancel{0.7}\ \text{odds(loaded coin)}}{\text{odd(fair coin)}}$$

$$\text{Oddratio} = \frac{\text{Odds}(1)}{\text{odds}(0)}$$

$$= \frac{\frac{0.7}{1-0.7}}{\frac{0.5}{1-0.5}} = \frac{0.7}{0.3} \times \frac{0.5}{0.5} = 2.333$$

"The odds of getting heads on the loaded coin are 2.333 times greater than the fair coin".

PES UNIVERSITY ONLINE

Odds ratio :

ODDS and ODDS RATIO

We will consider a data-set that tells us about depending on the gender, whether a customer will purchase a product or not

Gender	Purchase		Total
	Yes	No	
Female	159	106	265
Male	121	125	246

Odds, which describes the ratio of success to ratio of failure

Considering females group,

- we see that probability that a female will purchase (success) the product is = 159/265 (yes/total number of females).
- Probability of failure (no purchase) for female is 106/265.
- In this case the odds is defined as (159/265)/(106/265) = 1.5.
- Higher the odds, better is the chance for success. Range of odds can be any number between [0 , ∞].

$P(\text{Success}_1) = \frac{159}{265}$

$P(\text{failure}_{0}) = \frac{106}{265}$

$$\text{Odds} = \frac{p}{1-p} = \frac{159}{265-159} = \frac{159}{106} = 1.5$$

ODDS and ODDS RATIO

Odds ratio, is the ratio of odds.

Considering the example, Odds ratio, represents which group (male/female) has better odds of success, and it's given by calculating the ratio of odds for each group.

$$odds = \frac{\pi}{1-\pi} = \frac{p}{1-p}$$

So odds ratio for females = odds of successful purchase by female / odds of successful purchase by male = (159/106)/(121/125).

$$OR = \frac{\pi(1)/1-\pi(1)}{\pi(0)/1-\pi(0)}$$

Odds ratio for males will be the reciprocal of the above number.

Odds ratio can vary between 0 to positive infinity, log (odds ratio) will vary between [-∞, ∞]

Example: Consider a hypothetical data

$Y \to$ qualitative

Binary $y = 0$ or 1

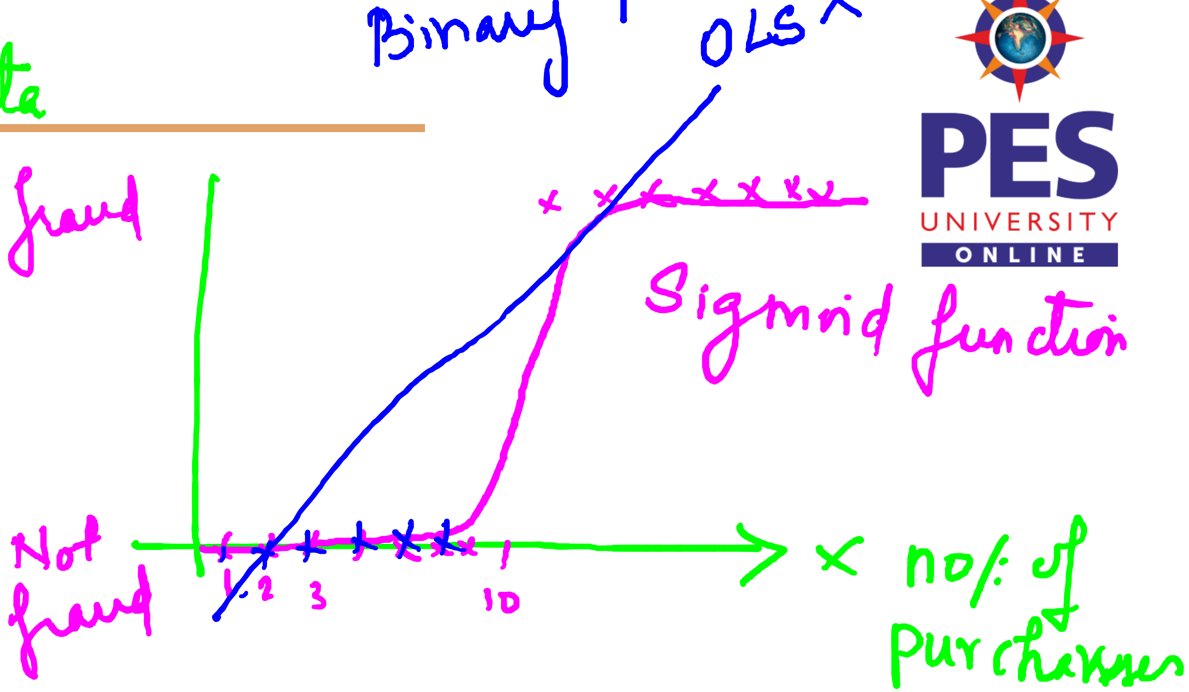
Two categories

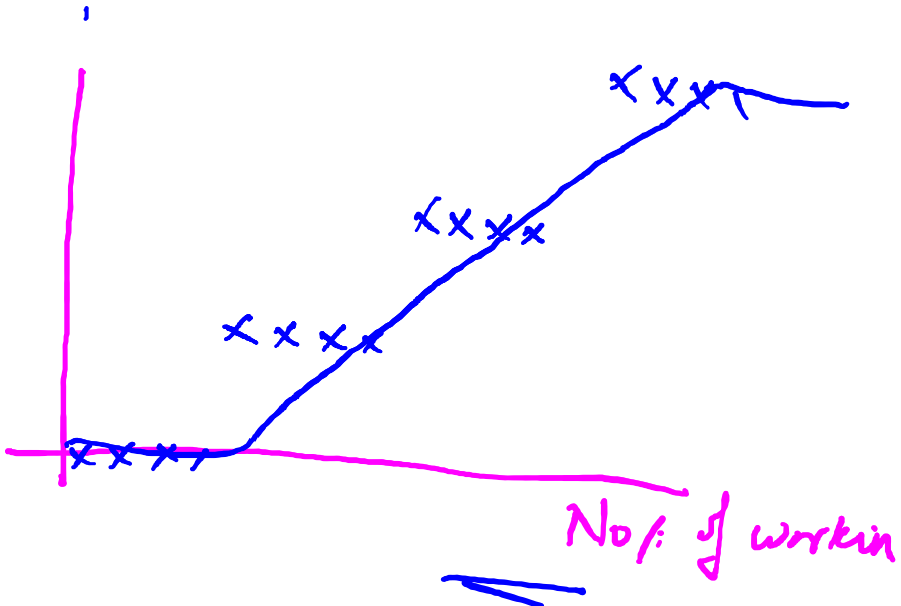
1. Fraud
2. Not fraudulent

} Y binary

Odds $\to \dfrac{p}{1-p}$

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x$$





logit(p) / logit function

$Y \in B$
0/1

$\rightarrow x \in R$

PES UNIVERSITY ONLINE

We know that $\text{Odds} = \frac{P}{1-P} \rightarrow (1)$

Taking natural logarithm, $\rightarrow (2)$

$$\text{logit}(p) = \ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 x_1$$

linear function of the independent variable

Antilog of the logit function (allows to estimate the Log Reg equation)

Antilog of (2)

$$\Rightarrow \frac{P}{1-P} = e^{(\beta_0 + \beta_1 x)}$$

$$\Rightarrow p = (1-p)e^{(\beta_0 + \beta_1 x)}$$

$$p = e^{(\beta_0+\beta_1 x)}(1-p)$$

$$\Rightarrow p = e^{(\beta_0+\beta_1 x)} - p.e^{(\beta_0+\beta_1 x)}$$

$$\Rightarrow p + p e^{\beta_0+\beta_1 x} = e^{\beta_0+\beta_1 x}$$

$$\Rightarrow \boxed{\hat{p} = \frac{e^{\beta_0+\beta_1 x}}{1+e^{\beta_0+\beta_1 x}}} \checkmark \rightarrow (4)$$

estimated probability

Text book

Further $\hat{p}$ can be simplified $\times$ & $\div$ by $e^{-(\beta_0+\beta_1 x)}$ in (4)

$$\boxed{\hat{p} = \frac{1}{1+e^{-(\beta_0+\beta_1 x)}}}$$

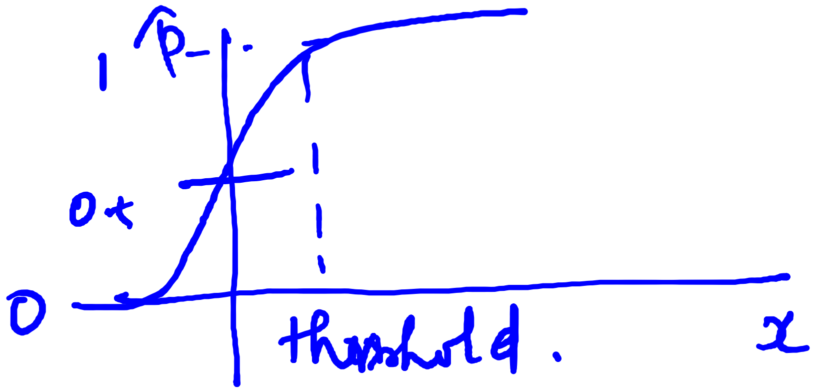
OR

estimate the parameters.

β ?

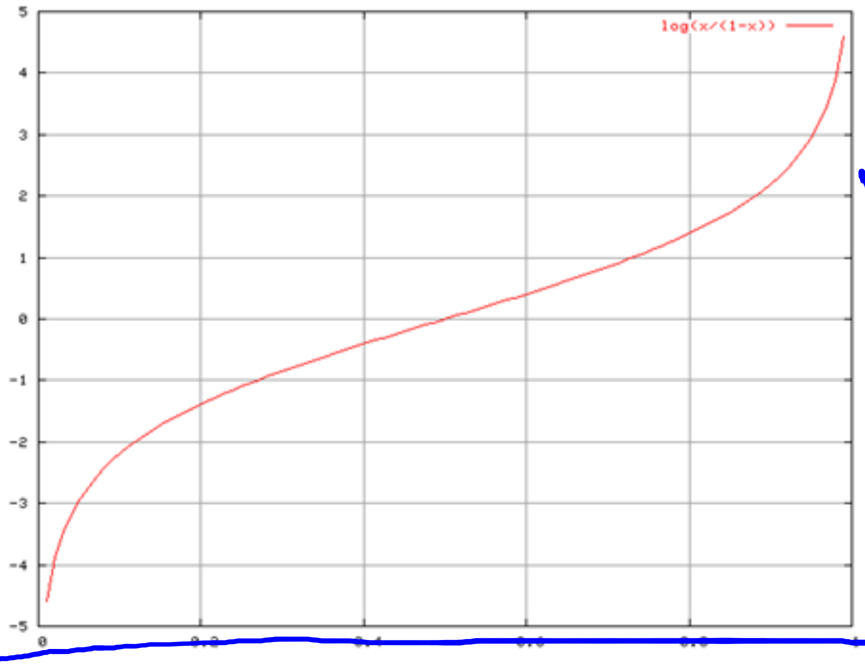
We need to estimate 'β'

$$\boxed{\hat{p} = \frac{1}{1+e^{-(\beta_0+\beta_1 x)}}}$$



3.Logit Function

- The logit function is the logarithmic transformation of the logistic function. It is defined as the natural logarithm of odds.
- Logit of a variable π (with value between 0 and 1) is given by:



P / π

$$\boxed{Logit(\pi) = \ln\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 x}$$

4.Logistic Transformation

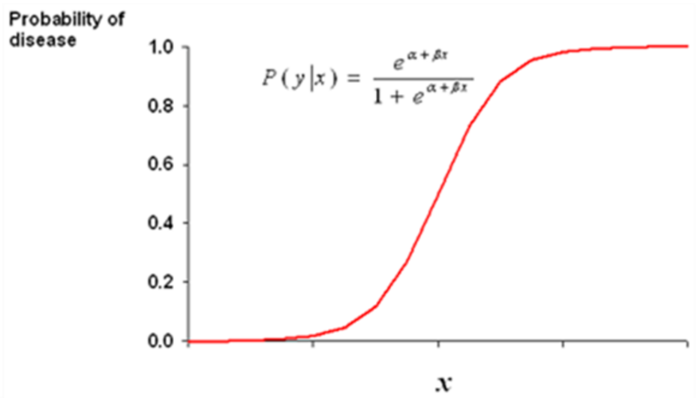
- The logistic regression model is given by:

$\hat{p}$

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

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

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



$$\ln\left(\frac{\pi_i}{1 - \pi_i}\right) = \beta_0 + \beta X_{1i}$$

Function with linear properties

β = 0 implies that P(Y|x) is same for each value of x

β > 0 implies that P(Y|x) is increases as the value of x increases

β < 0 implies that P(Y|x) is decreases as the value of x increases

$\beta_0 + \beta_1 x$ . . $\dot{x}$. $Y = \beta_0 + \beta_1 x$

6.Parameter Estimation in Logistic Regression (Maximum Likelihood Estimate)

Likelihood function for Binary Logistic Function

- Probability density function for binary logistic regression is given by:

$$P(Y=1 \mid Z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_m X_m) = \pi(Z) = \frac{e^Z}{(1+e^Z)} \tag{11.7}$$

The probability (likelihood) function of binary logistic regression for specific observation Y_i (Y_i = 0 or 1) is given by

$$P(Y_i) = \pi(Z)^{Y_i} (1 - \pi(Z))^{1-Y_i} \tag{11.8}$$

Bernoulli

$Y = 0 \mid 1$

Success/failure

MLE $\rightarrow \beta_0 + \beta_1$

Estimation of parameters

MLE

PES UNIVERSITY ONLINE

Assume that the data set has n observations, $Y_1, Y_2, \ldots, Y_n$. The likelihood function, which is a joint probability, $L(Y_1, Y_2, \ldots, Y_n)$ for a specific Z_i $(= \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_m X_{mi})$ is given by

$$L = P(Y_1, Y_2, \ldots, Y_n) = \prod_{i=1}^{n} \pi(Z_i)^{Y_i}[1 - \pi(Z_i)]^{1-Y_i} \tag{11.9}$$

The log-likelihood function is given by

$$\ln(L) = LL = \sum_{i=1}^{n} Y_i \ln[\pi(Z_i)] + \sum_{i=1}^{n}(1 - Y_i)[\ln(1 - \pi(Z_i))] \tag{11.10}$$

For mathematical simplicity, assume that $Z_i = \beta_0 + \beta_1 X_i$. Equation (11.10) can be written as

$$LL(\beta_0, \beta_1) = \sum_{i=1}^{n} Y_i(\beta_0 + \beta_1 X_i) - \sum_{i=1}^{n} \ln[1 + \exp(\beta_0 + \beta_1 X_i)] \tag{11.11}$$

Taking partial derivatives with respect to β_0 and β_1 and equating them to zero, we get the following first-order conditions (Hosmer and Lemeshow, 2000; Kleinbaum and Klein, 2011):

$$\frac{\partial \ln(L(\beta_0, \beta_1))}{\partial \beta_0} = \sum_{i=1}^{n} Y_i - \sum_{i=1}^{n} \frac{\exp(\beta_0 + \beta_1 X_i)}{1 + \exp(\beta_0 + \beta_1 X_i)} = 0 \tag{11.12}$$

$$\frac{\partial \ln(L(\beta_0, \beta_1))}{\partial \beta_1} = \sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} \frac{X_i \exp(\beta_0 + \beta_1 X_i)}{1 + \exp(\beta_0 + \beta_1 X_i)} = 0 \tag{11.13}$$

closed form

Numerical methods

"Newton Raphson" method.

β_0 & β_1

Limitations of MLE

- Maximum likelihood estimator may not be unique or may not exist.

- No closed form solution ⇒ use iterative procedure to estimate the parameter values.

odd

odd ratio

Logit function

$p = \frac{1}{1+e^{-(\beta_0+\beta_1 x)}}$

$\beta_0 = ?$ $\beta_1 = ?$

Exercise :Space Shuttle Challenger Data

Space shuttle orbiter Challenger (Mission STS-51-L) was the 25th shuttle launched by NASA on January 28, 1986 (Smith, 1986; Feynman 1988). The Challenger crashed 73 seconds into its flight due to the erosion of O-rings which were part of the solid rocket boosters of the shuttle. Before the launch, the engineers at NASA were concerned about the outside temperature which was very low (the actual launch occurred at 36°F). Data in Table 11.1 shows the O-ring erosion and the launch temperature of the previous shuttle launches, where 'damage to O-ring = 1' implies there was a damage to O-ring and 'damage to O-ring = 0' implies there was no damage to O-ring during that launch. In this case, the outcome is binary − either there is a damage to O-ring or there is no damage to O-ring. **We can develop a logistic regression model to predict the probability of erosion of O-ring based on the launch temperature.**

Exercise :Space Shuttle Challenger Data

n=24

Prob(y)

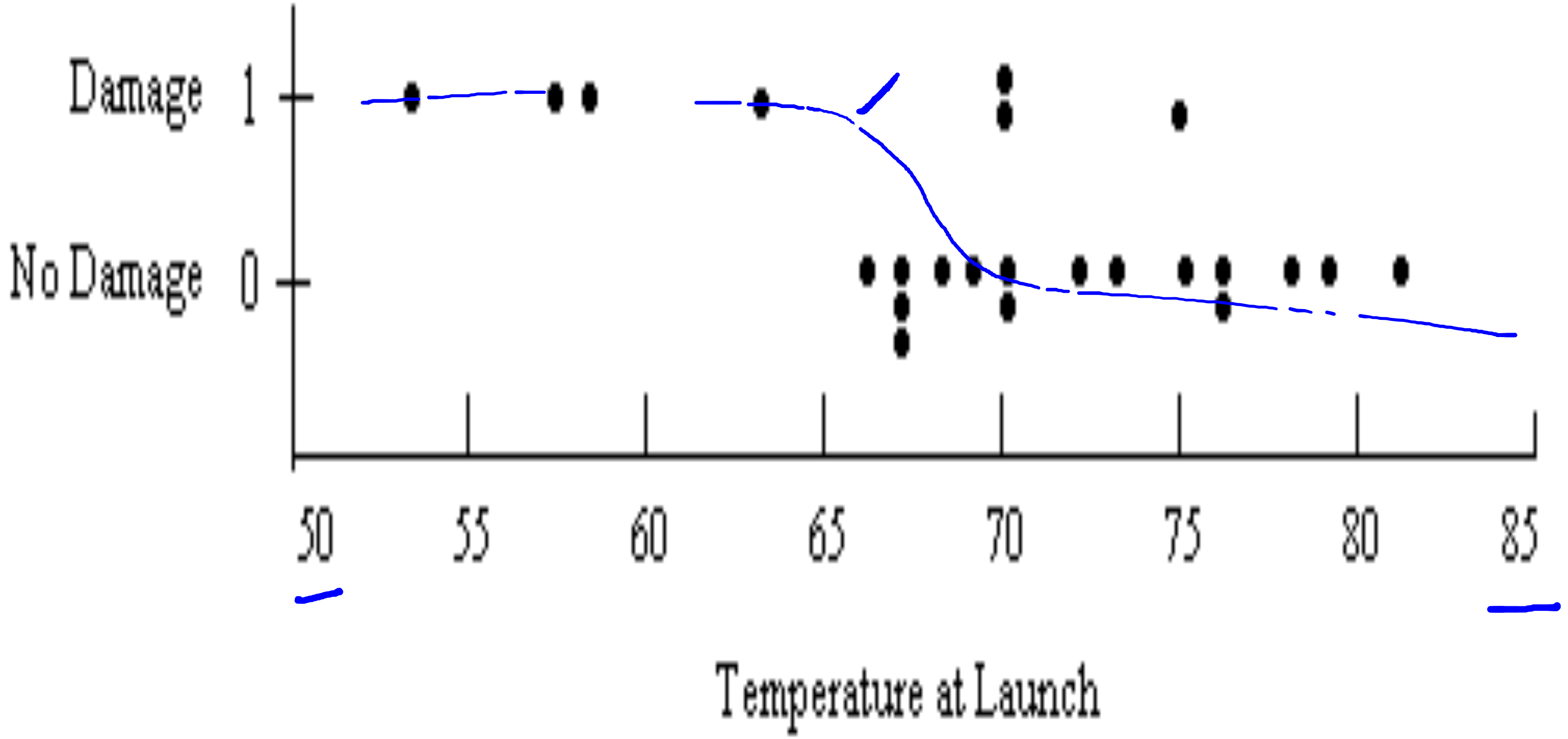
x   Y   'O' ring

Flt	Temp	Damage
STS-1	66 ✓	No 0
STS-2	70	Yes 1
STS-3	69	No 0
STS-4	80	No 0
STS-5	68	No 0
STS-6	67	No 0
STS-7	72	No
STS-8	73	No
STS-9	70	No
STS-41B	57	Yes
STS-41C	63	Yes
STS-41D	70	Yes

Flt	Temp	Damage
STS-41G	78	No
STS-51-A	67	No
STS-51-C	53	Yes
STS-51-D	67	No
STS-51-B	75	No
STS-51-G	70	No
STS-51-F	81	No
STS-51-I	76	No
STS-51-J	79	No
STS-61-A	75	Yes
STS-61-B	76	No
STS-61-C	58	Yes

PES UNIVERSITY ONLINE

Space Shuttle Challenger Data



logit function $= \ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x$

Exercise :Space Shuttle Challenger Data

Solution

The logit function for the example in Table 11.1 is given by

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

$$\hat{p}_i = \frac{1}{1+e^{-(\beta_0+\beta_1 x)}}$$

X_i is the launch temperature of i^{th} launch. Binary logistic regression output from SPSS is shown in Table 11.2. The values of β_0 and β_1 are 15.297 and –0.236, respectively.

TABLE 11.2 Logistic regression coefficient for the challenger data in Table 11.1

	B (beta values)		S.E. (Standard error of estimate)		Wald Statistic	*df*	Sig. (*p*-value)
Launch Temperature	–0.236	β_1	0.107	std β_1	4.832	1	0.028
Constant	15.297	β_0	7.329	std β_0	4.357	1	0.037

Exercise :Space Shuttle Challenger Data

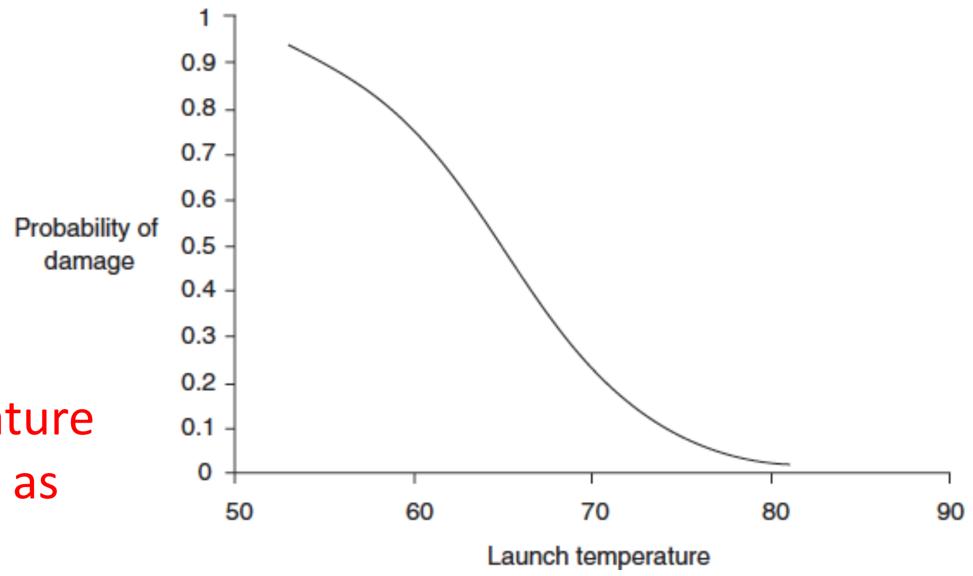
$$p = \frac{e^{(\beta_0+\beta_1 x)}}{1+e^{(\beta_0+\beta_1 x)}} = \frac{1}{1+e^{-(\beta_0+\beta_1 x)}}$$

We know that

$$P(Y=1) = \frac{e^Z}{1+e^Z}$$

$$Z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_m X_m$$

$$P(Y=1) = \pi(Z) = \frac{\exp(15.297 - 0.236 \times X_i)}{1+\exp(15.297 - 0.236 \times X_i)}$$



The plot of probability of damage to O-ring and launch temperature is shown in graph in which the probability of damage decreases as the launch temperature increases.

However, we have to test the validity of the model using diagnostic tests before the model can be accepted.

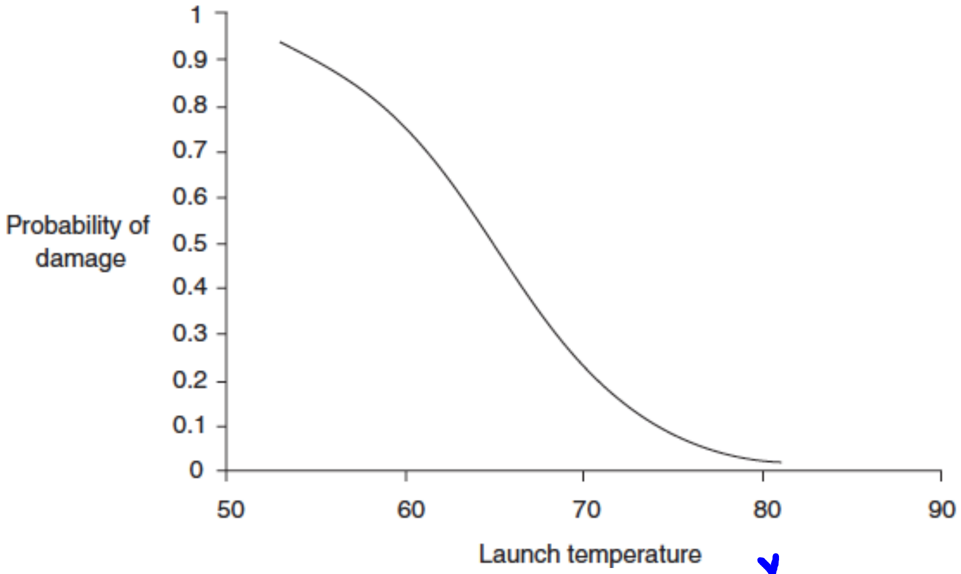
$$x = 66, \quad \hat{p} = \frac{1}{1 + e^{-(15.297 + (-0.236)*66)}}$$

$$= 0.43$$

Exercise :Space Shuttle Challenger Data

TABLE 11.8 Challenger crash data — predicted probability using logistic regression model

S. No.	Flight Number	Launch Temperature	Damage to O-ring	Predicted Probability
1	STS 1	66.00	0	0.43
2	STS 2	70.00	1	0.23
3	STS 3	69.00	0	0.27
4	STS 4	80.00	0	0.03
5	STS 5	68.00	0	0.32
6	STS 6	67.00	0	0.37
7	STS 7	72.00	0	0.15
8	STS 8	73.00	0	0.13
9	STS 9	70.00	0	0.23
10	STS 41B	57.00	1	0.86
11	STS 41C	63.00	1	0.61
12	STS 41D	70.00	1	0.23
13	STS 41G	78.00	0	0.04
14	STS 51A	67.00	0	0.37
15	STS 51C	53.00	1	0.94
16	STS 51D	67.00	0	0.37
17	STS 51B	75.00	0	0.08
18	STS 51G	70.00	0	0.23
19	STS 51F	81.00	0	0.02
20	STS 51I	76.00	0	0.07
21	STS 51J	79.00	0	0.03
22	STS 61A	75.00	1	0.08
23	STS 61B	76.00	0	0.07
24	STS 61C	58.00	1	0.83



x

7.Interpretation of LR coefficients using odds ratio.

Page 343

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

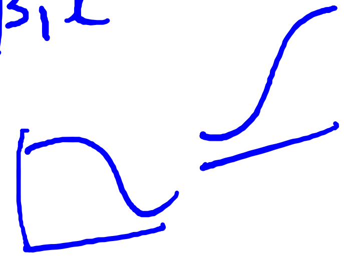
$$p \begin{cases} y=1 \\ y=0 \end{cases}$$

$$\ln\left(\frac{P(Y=1)}{1-P(Y=1)}\right) = \beta_0 + \beta_1 \cdot temp$$

$\rightarrow$ derivation

7.Interpretation of LR coefficients

$\ln(\text{odds}) = \beta_0 + \beta_1 x$



$\beta_1 = \text{slope.}$

- β_1 is the change in log-odds ratio for unit change in the explanatory variable.

- β_1 is the change in odds ratio by a factor $\exp(\beta_1)$.

$$\beta_1 = \ln\left(\frac{P(x+1)\;/(1-P(x+1))}{P(x)\;/(1-P(x))}\right) = \text{Change in ln odds ratio}$$

$\beta_1 = \ln(\text{odds ratio})$

$$e^{\beta_1} = \frac{P(x+1)\,/\,(1-P(x+1))}{P(x)\,/\,(1-P(x))} = \text{Change in odds ratio}$$

Detailed example: https://www.youtube.com/watch?v=Rj3pVkjAwY0

Exercise :Space Shuttle Challenger Data

Solution

The logit function for the example in Table 11.1 is given by

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

X_i is the launch temperature of i^{th} launch. Binary logistic regression output from SPSS is shown in Table 11.2. The values of β_0 and β_1 are 15.297 and –0.236, respectively.

$\beta_1 = -0.236$

$\beta_1 = \ln(\text{odds ratio})$

$\text{odds ratio} = e^{\beta_1}$

$= e^{-0.236}$

$= 0.789$

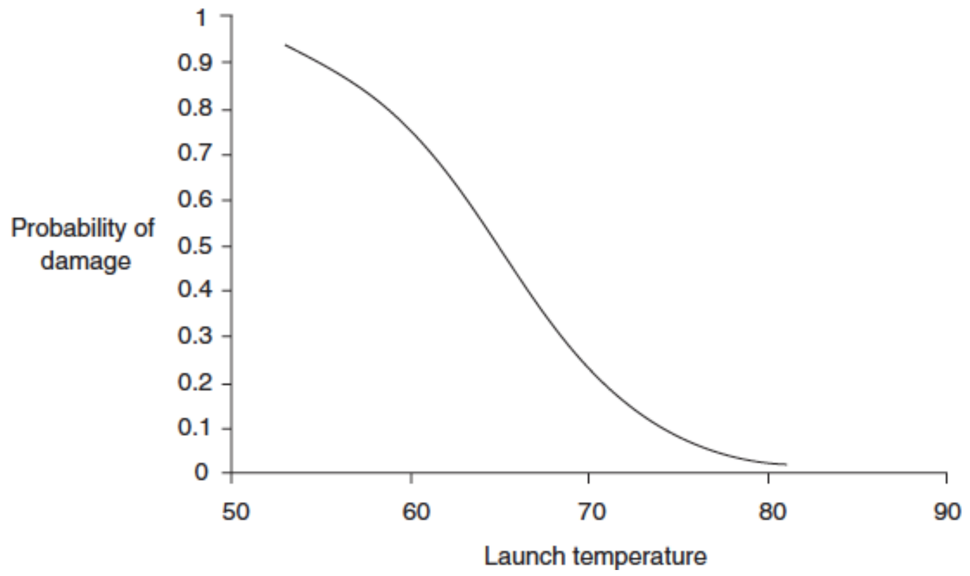
TABLE 11.2 Logistic regression coefficient for the challenger data in Table 11.1

	B (beta values)	S.E. (Standard error of estimate)	Wald Statistic	*df*	Sig. (*p*-value)
Launch Temperature	–0.236 ✓	0.107	4.832	1	0.028
Constant	15.297	7.329	4.357	1	0.037

Exercise :Space Shuttle Challenger Data

TABLE 11.8 Challenger crash data — predicted probability using logistic regression model

S. No.	Flight Number	Launch Temperature	Damage to O-ring	Predicted Probability
1	STS 1	66.00	0	0.43
2	STS 2	70.00	1	0.23
3	STS 3	69.00	0	0.27
4	STS 4	80.00	0	0.03
5	STS 5	68.00	0	0.32
6	STS 6	67.00	0	0.37
7	STS 7	72.00	0	0.15
8	STS 8	73.00	0	0.13
9	STS 9	70.00	0	0.23
10	STS 41B	57.00	1	0.86
11	STS 41C	63.00	1	0.61
12	STS 41D	70.00	1	0.23
13	STS 41G	78.00	0	0.04
14	STS 51A	67.00	0	0.37
15	STS 51C	53.00	1	0.94
16	STS 51D	67.00	0	0.37
17	STS 51B	75.00	0	0.08
18	STS 51G	70.00	0	0.23
19	STS 51F	81.00	0	0.02
20	STS 51I	76.00	0	0.07
21	STS 51J	79.00	0	0.03
22	STS 61A	75.00	1	0.08
23	STS 61B	76.00	0	0.07
24	STS 61C	58.00	1	0.83



7.Interpretation of LR coefficients

Flt	Temp	Damage	Predicted Prob
STS-2	70	Yes	0.23
STS-41B	57	Yes	0.86
STS-41C	63	Yes	0.61
STS-41D	70	Yes	0.23
STS-51-C	53	Yes	0.94
STS-61-A	75	Yes	0.08
STS-61-C	58	Yes	0.83

x y=1

Odds ratio:

$x = 57$ $\hat{p} = 0.86$ $odd = \frac{0.86}{1-0.86} = 6.14$

increase

$x+1 = 58$ $\hat{p} = 0.837$ $odd = \frac{0.83}{1-0.83} = 4.88$

$$\text{Oddsratio} = \frac{\frac{P(x+1)}{1-P(x+1)}}{\frac{P(x)}{1-P(x)}} = \frac{4.88}{6.14} = 0.79$$

$\beta_1 =$

$$\boxed{\text{Oddsratio} = e^{\beta_1}}$$

Odds Ratio for Binary Logistic Regression

$x+1$

$$OR = \frac{\pi(1)/1-\pi(1)}{\pi(0)/1-\pi(0)}$$

If OR = 2, then the event is twice likely to occur when X = 1 compared to X = 0.

Odds ratio approximates the relative risk.

$$\boxed{\text{Odds ratio} = e^{\beta 1}}$$

Odd

$\beta_0 \neq \beta_1$

Example

$p = ?$ $\quad odds = \frac{p}{1-p}$

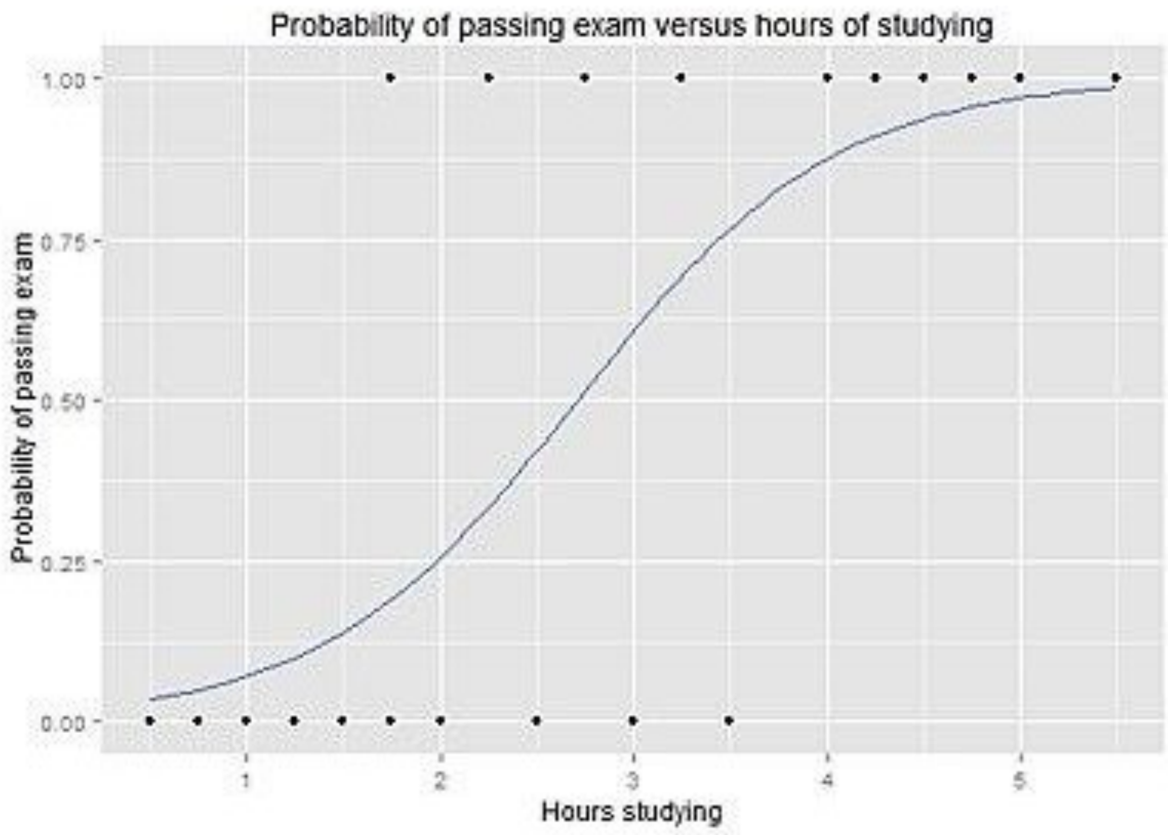
$odds\ ratio = \frac{odds_2}{odd_1} = \frac{0.34}{0.076} = e^{\beta_1}$

$\beta_1 = .$

Hours of study	Passing exam		
	Log-odds	Odds	Probability
1	−2.57	0.076 ≈ 1:13.1	0.07
2	−1.07	0.34 ≈ 1:2.91	0.26
3	0.44	1.55	0.61
4	1.94	6.96	0.87
5	3.45	31.4	0.97

One hr of study increases log odds of passing by 1.5046

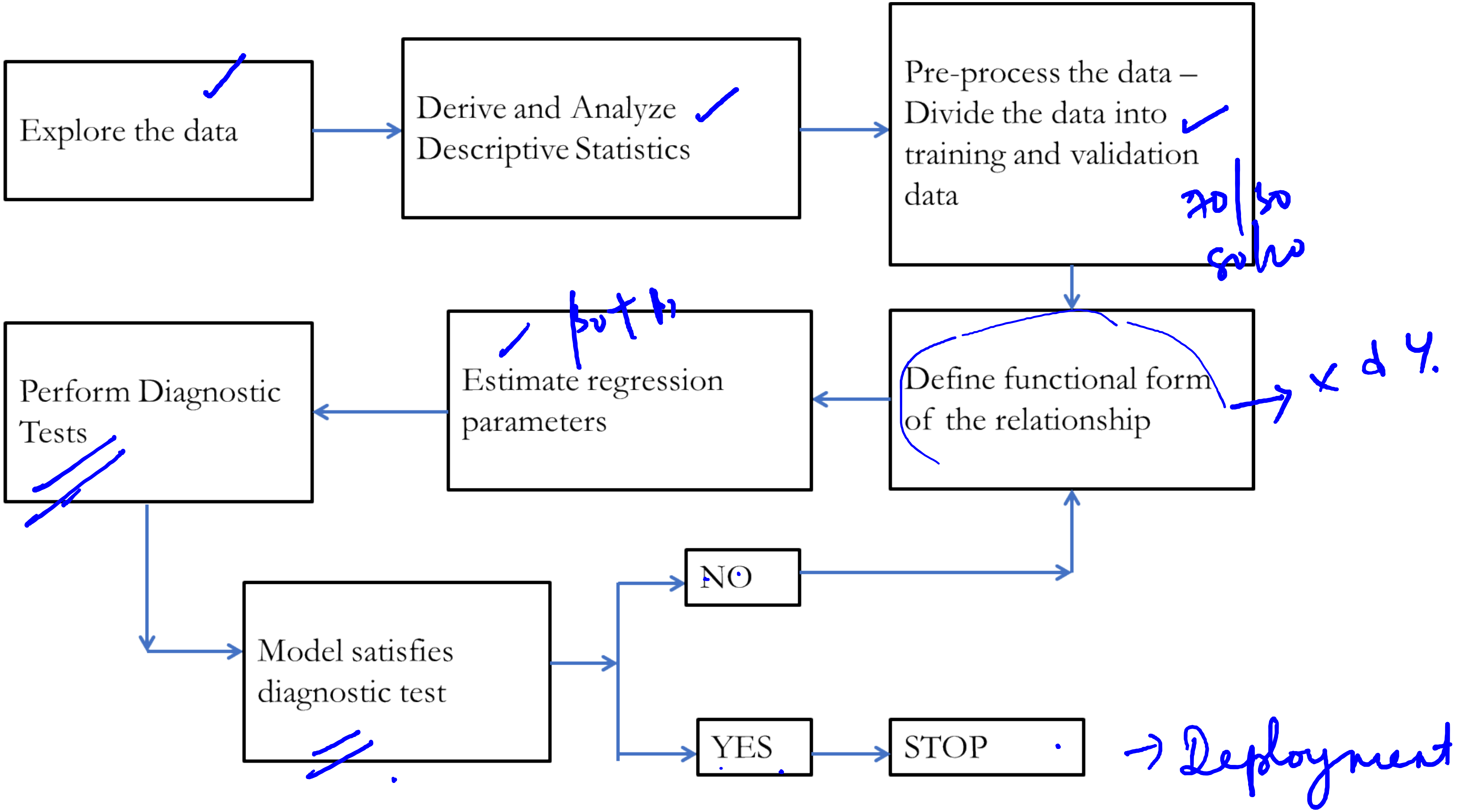
$2\,hr = -1.07 + 2.57 = 1.5$



	Coefficient	Std.Error	P-value (Wald)
Intercept	−4.0777 β_0	1.7610	0.0206
Hours	1.5046 β_1	0.6287	0.0167

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

Logistic Regression Model Development



't' of L.R or MLR

0.01

Wald's test

Wald's test is used for checking statistical significance of individual predictor variables (equivalent to *t*-test in MLR model). The null and alternative hypotheses for Wald's test are:

H_0: $\beta_i = 0$

H_1: $\beta_i \neq 0$

TABLE 11.2 Logistic regression coefficient for the challenger data in Table 11.1					
	B (beta values)	S.E. (Standard error of estimate)	Wald Statistic	*df*	Sig. (*p*-value)
Launch Temperature	−0.236	0.107	4.832	1	0.028
Constant	15.297	7.329	4.357	1	0.037

Wald's test statistic is given by

$$W = \left[\frac{\hat{\beta}_i}{S_e(\hat{\beta}_i)} \right]^2$$

R^2 in Logistic Regression

R^2 / $R^2 \to$ LinearReg

Sigmoid

- In linear regression R^2 is the proportion of variation explained by the regression model.

- It is not possible to develop a R^2 type measure for Logistic Regression since the variance of the error term is not constant.

- Many Pseudo R^2 values are used in Logistic Regression. Pseudo R2 is an indicator of strength of relationship.

R^2 in Logistic Regression

- **R-squared is a measure of improvement from null model to fitted model** - The denominator of the ratio can be thought of as the sum of squared errors from the null model--a model predicting the dependent variable without any independent variables.

- In the null model, each *y* value is predicted to be the mean of the *y* values.

Pseudo R^2

It is not possible to calculate R^2 as in the case of continuous dependent variable in a logistic regression model.

However, many pseudo R^2 values are used which compare the intercept-only model to the model with independent variables.

Cox and Snell R^2

Pseudo R^2 → logistic

PES UNIVERSITY ONLINE

Cox and Snell R^2 is given by

$$R^2 = 1 - \left\{ \frac{L(\text{Intercept only model})}{L(\text{Full Model})} \right\}^{2/N}$$

β_0 $\quad$ ~~$\beta_1 \cdots \beta_n$~~

$\beta_0, \beta_1, \beta_2$

Based on Log-likelihood ratio

$$R^2 = 1 - \left(\frac{LL(\text{Null Model})}{LL(\text{Model})} \right)^{2/n}$$

Null Model : Model without predictors

Full Model: Model with predictors

n is the number of observations

References

Text Book:

"Business Analytics, The Science of Data-Driven Decision Making", U. Dinesh Kumar, Wiley 2017 (Ch. 11.1-11.4, 11.6-11.6.2)

https://towardsdatascience.com/train-validation-and-test-sets-72cb40cba9e7

https://medium.com/hugo-ferreiras-blog/confusion-matrix-and-other-metrics-in-machine-learning

https://online.stat.psu.edu/stat504/node/216/

https://towardsdatascience.com/logit-of-logistic-regression-understanding-the-fundamentals-f384152a33d1

THANK YOU

Dr.Mamatha H R

Professor,Department of Computer Science

mamathahr@pes.edu

+91 80 2672 1983 Extn 834