

Techniques used for Discrete Output

- Logistic Analysis
- Logit Analysis
- Probit Analysis

Slide-114

Regression Techniques – Simple Logistic Regression

Y-Discrete, x – Single & continuous

We apply Simple logistic Regression

Y- Discrete, x – Single & discrete

We create dummy variable for discrete component and

We then apply Simple logistic Regression

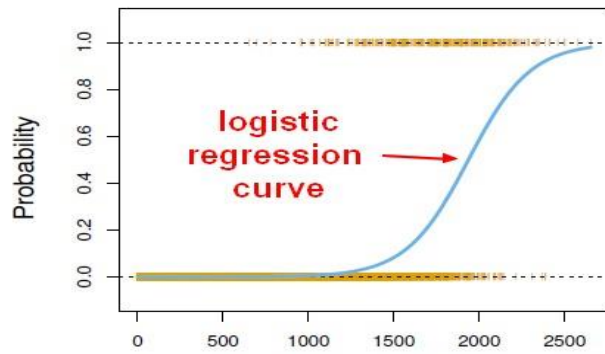
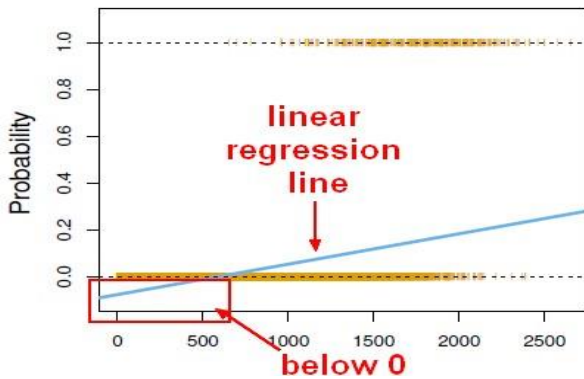
Slide-115

Logistic Regression

- Logistic Regression model predicts the probability associated with each dependent variable Category

How does it do this?

- It finds linear relationship between independent variables and a link function of these probabilities. Then the link function that provides the best goodness-of-fit for the given data is chosen.



Slide-116

- Multiple Logistic Regression Model is quite similar to the Multiple Linear Regression Model, Only β coefficients vary

Multiple Linear Regression model :

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon$$

Where β_0 = Y intercept

β_k -> the model coefficient for the linear effect of variable i on y

e -> the random error

Slide-117

The probability in a logistic regression curve

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

Where e is a real constant, the base of natural logarithm and equals to 2.7183

Y is the response value of the observation

SLIDE-118

Logistic Regression Methods

Method	Description of categorical response variable	Example
binary	2 categories	Presence/absence of disease
Nominal	3 or more categories with no natural ordering to the levels	Crunchy/mushy/crispy
ordinal	Three or more variables with ordering of levels	Strongly disagree/disagree/neutral/Agree/strongly agree

SLIDE-119

Assumptions in Logistic Regression

- Only one outcome per event – Like pass or fail
- The outcomes are statistically independent
- All relevant predictors are in the model
- One category at a time – Mutually exclusive & collectively exhaustive
- Sample sizes are larger than for linear regression

SLIDE-120

Steps in Logistic Regression

- Collect & organize sample data
- Formulate Logistic Regression Model
- Check the model's validity
- Determine Probabilities using Probability equation
- Compile the results

SLIDE-121

Logistic Regression Example

Imagine that you are a Data Scientist at a very large scale integration circuit manufacturing company. You want to know whether or not the time spent inspecting each product impacts the quality assurance department's ability to detect a designing error in the circuit

→ *Step-1: Collect and organize the sample data*

→ *Number of Observations*

→ *Error Identification*

→ *Inspection Time*

Number of Observations: 55 Observations of circuits with errors, and determine whether those errors were detected by QA