PROBABILITY - BASICS

Shah Ayub Quadri

Ayub.quadri89@gmail.com

Content

- Probability
- Event
  - Types of Events
  - Independent, Dependent, Mutually Exclusive events
- Probability
  - Types of probabilities
  - Conditional probability
- Bayes Theorem
- Confusion Matrix
- Introduction to R Programming

Probability:

- Probability is the measure of how much likelihood that an event will occur.
- Probability is qualified between 0 & 1
- Higher the probability of an event, more certain that the event will occur
- 0 – Impossibility of an event
- 1 – highly certain that even will occur

Eg: Weather condition: *“80% chances that it would rain today”*

Perspectives on Probability

- Classical (A priori or Theoretical)
- Empirical (Posterior or Frequentist)
- Subjective Probability

Classical Probability (A priori):

- Probability can be determined theoretically prior to conducting any experiment

- $P(E) = \frac{No\ of\ Outcomes}{Total\ possible\ Outcomes}$

- Eg: Tossing a fair dice to get 4 on a die
  - $P(4) = 1/6$

- When to use Classical method? When possible outcomes are finite or known

Empirical Probability (A posterior or Frequentist):

- It defines the probability via conducting experiments

- $P(E) = \frac{No\ of\ times\ event\ occured}{Total\ no\ of\ time\ exp\ is\ carried\ out}$

- Eg: Out of 5 fair coin tosses, Heads appeared 2 times
  - $P(H) = 2/5$

Subjective Probability:

- Its an individual person's measure of belief that an event will occur based on the feeling, insights, knowledge etc of a person.
- It differs from person to person

Events:

- Events are nothing but one or more outcome of an experiment
- Eg: Getting heads when a fair coin is tossed
- $P(H) = 1/2$

Types of Events

- Independent
- Dependent
- Mutually Exclusive

Independent Event:

- An event which is not effected by any other events

- Eg: Toss 3 fair coins, event of getting heads at one coin does not effect any other coin tosses

- $P(A \& B) = P(A) * P(B)$

Dependent Event:

- An event that is affected by previous event

- $P(A \text{ and } B) = P(A).P(A|B)$

Mutually exclusive events:

- Consider two events, If an event 'A' happened than event 'B' cannot happen.

- $P(A \text{ or } B) = P(A) + P(B) - P(A \& B)$

Contingency table summarizing 2 variables, Loan Defaulters vs Age.

		Age			
		Young	**Middle-aged**	**Old**	**Total**
Loan Default	**No**	10,503	27,368	259	38,130
	Yes	3,586	4,851	120	8,557
	Total	14,089	32,219	379	46,687

		Age			
		Young	**Middle-aged**	**Old**	**Total**
Loan Default	**No**	0.225	0.586	0.005	0.816
	Yes	0.077	0.104	0.003	0.184
	Total	0.302	0.690	0.008	1.000

Types of Probabilities

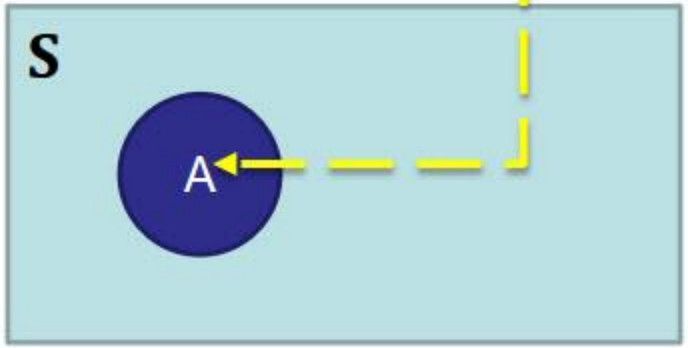
Marginal Probability

		Age			
		Young	Middle-aged	Old	Total
Loan Default	No	0.225	0.586	0.005	0.816
	Yes	0.077	0.104	0.003	0.184
	Total	0.302	0.690	0.008	1.000

Probability describing a single attribute.

P(No) = 0.816

P(Old) = 0.008

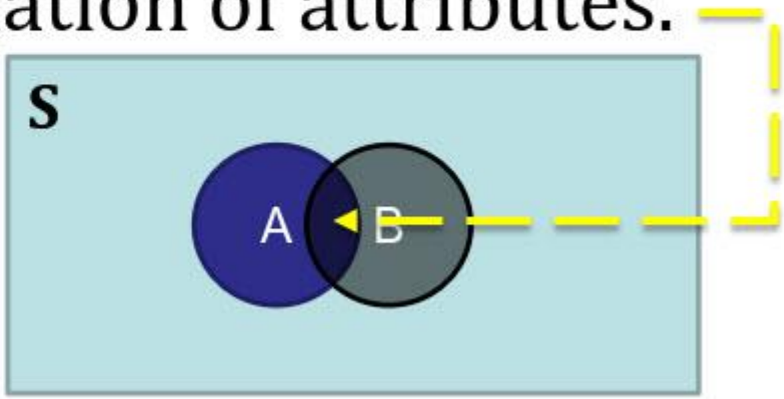


Joint Probability

		Age			
		Young	Middle-aged	Old	Total
Loan Default	No	0.225	0.586	0.005	0.816
	Yes	0.077	0.104	0.003	0.184
	Total	0.302	0.690	0.008	1.000

Probability describing a combination of attributes.

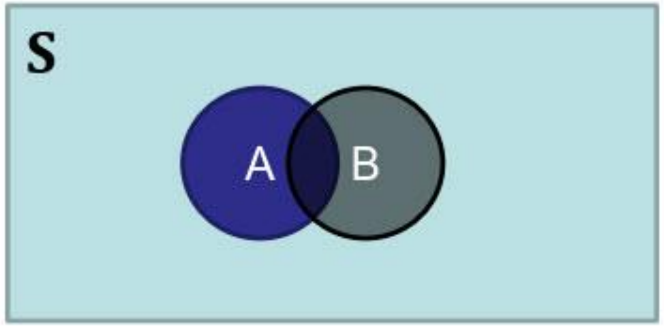
P(Yes **and** Young) = 0.077



Union Probability

		Age			Total
		Young	Middle-aged	Old	Total
Loan Default	No	0.225	0.586	0.005	0.816
	Yes	0.077	0.104	0.003	0.184
	Total	0.302	0.690	0.008	1.000

P(Yes **or** Young) = P(Yes) + P (Young) – P(Yes and Young) = 0.184 + 0.302 – 0.077 = 0.409



Conditional probability:

- Probability that event B will occur given that event A has already occurred is called as conditional probability.

- $P(B|A) = \frac{P(A \ and \ B)}{P(A)}$

Bayes' Theorem

- A way of finding a probability when you know certain other probabilities

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}$$

- Where P(B) != 0
- P(A) – Probability of event 'A'
- P(B) – Probability of event 'B'
- P(B|A) – conditional probability of event 'B' when event 'A' has occurred

Confusion Matrix

		Predicted		
		Positive	Negative	
Actual	Positive	True +ve	False –ve	Recall/Sensitivity/True Positive Rate (Minimize False –ve)
	Negative	False +ve	True –ve	Specificity/True Negative Rate (Minimize False +ve)
		Precision		Accuracy, F_1 score

$$Recall = \frac{True + ve}{Actual\ Total + ve}$$

$$Specificity = \frac{True\ - ve}{predicted\ total\ - ve}$$

$$Precision\ = \frac{True + ve}{Predicted + ve}$$

$$F_1 Score = \frac{2 * Precession\ * Recall}{(Precession + Recall)}$$

$$Accuracy\ = \frac{(True + ve) + (True\ - ve)}{(True + ve) + (False + ve) + (False\ - ve) + (True\ - ve)}$$

Problem 1: Spam Filter

Spam filtering		**Predicted**		**Total**
		Positive	Negative	
Actual	Positive	952	526	1478
	Negative	167	3025	3192
Total		1119	3551	4670

$$Recall\ (Sensitivity) = \frac{952}{1478} = 0.644$$

$$Precision = \frac{952}{1119} = 0.851$$

$$Accuracy = \frac{952 + 3025}{952 + 3025 + 526 + 167} = \frac{3977}{4670} = 0.852$$

$$Specificity = \frac{3025}{3025 + 167} = \frac{3025}{3192} = 0.948$$

$$F_1 = 2 * \frac{Precision * Recall}{Precision + Recall} = \frac{2 * 0.851 * 0.644}{0.851 + 0.644} = \frac{1.096}{1.495} = 0.733$$

Problem 2: Breast Cancer

Breast cancer detection		Predicted		Total
		Positive	Negative	
Actual	Positive	852	126	978
	Negative	67	1025	1092
Total		919	1151	2070

$$Recall\ (Sensitivity) = \frac{852}{978} = 0.871$$

$$Precision = \frac{852}{919} = 0.927$$

$$Accuracy = \frac{852 + 1025}{852 + 1025 + 126 + 67} = \frac{1877}{2070} = 0.907$$

$$Specificity = \frac{1025}{1025 + 67} = \frac{1025}{1092} = 0.939$$

$$F_1 = 2 * \frac{Precision * Recall}{Precision + Recall} = \frac{2 * 0.871 * 0.927}{0.871 + 0.927} = \frac{1.615}{1.798} = 0.898$$

Interview Question:

Q. You have been tasked to build a classifier for cancer diagnosis. It is of high importance that patients with cancer can be diagnosed wrongly as negative but patients without cancer should NEVER be diagnosed as positive.

Which of the following classification models would you prefer?

(Assuming: Positives = Cancer, Negatives = Not cancer)

Options:

- True Positive Rate [which is = True Positive / Actual Positive]
- True Negative Rate [which is = True Negative / Actual Negative]
- Precision [which is = True Positive / Predicted Positive]
- Total Accuracy [which is = (True Positive + True Negative) / Total Population]

Answer – Precision

REFERENCES

Probability & statistics

- http://www.mathsisfun.com/data/index.html#stats

Short definition & terms in probability and statistics

- http://stattrek.com

Conditional probability explained visually

- https://www.khanacademy.org/video/conditional-probability2

Bayes Theorem:

- https://youtu.be/E4rlJ82CUZI

- Confusion Matrix
  - http://www.dataschool.io/simple-guide-to-confusion-matrix-terminology/