Introduction to Probability

Agenda

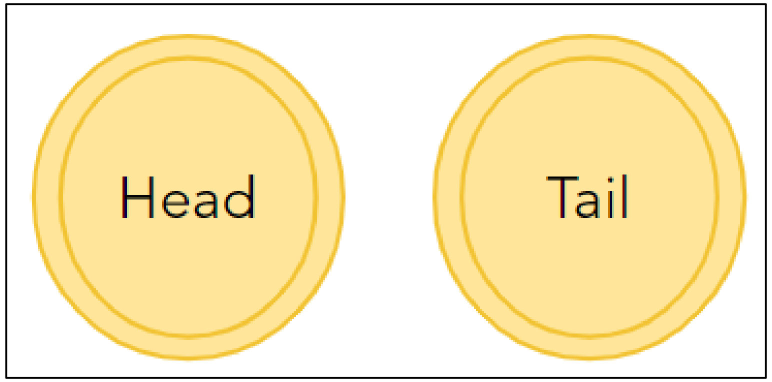
- Introduction to Probability
    - Probability
    - Odds
    - Independence of Events
    - Conditional Probability
    - Bayes Theorem

Experiment

An experiment is procedure that can be repeated umpteenth times, having a set of possible outcomes.

For example:

- The result of rolling a die is either 1 or 2 or 3 or 4 or 5 or 6
- Measuring height of a person
- The result of tossing a coin is either head or tail



Trial of an experiment

- A trial of an experiment is performing the experiment once

- ‘n’ trials imply the experiment is performed ‘n’ times

- Trials can either be sequential or simultaneous

Experiment

Experiment

Based on the number of outcomes of an experiment

Deterministic Experiment

Random Experiment

Only one possible outcome and can be predicted in advance
For instance,
Experiment: Heating water to 100˚C
Outcome: Water turns to vapour

More than one possible outcomes and can not be predicted in advance
For instance,
Experiment: Returns on an investment
Outcomes: More, less, no change

Probability

Probability

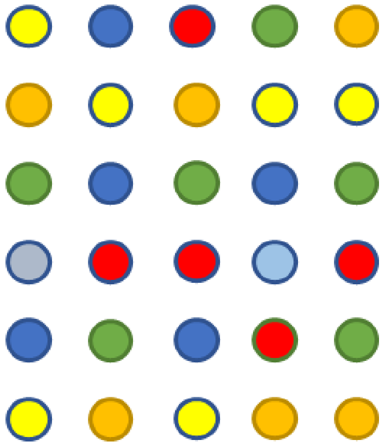
- Probability is a measure for the likelihood of occurrence of an event

$$\text{Probability of an event} = \frac{\text{Number of favourable outcomes}}{\text{Number of all possible outcomes}}$$

- The probability of any event lies in between 0 to 1
- The sum of all possible events of an experiment is 1
- Probability of an event A is denoted by P(A)

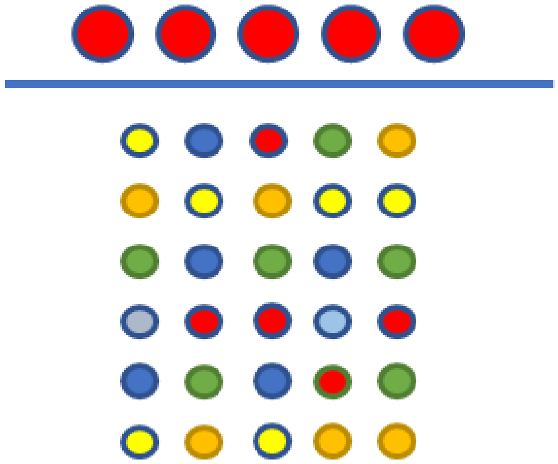
Probability

$$\text{Probability of an event } P(A) = \frac{\text{Number of outcomes in A}}{\text{Number of all possible outcomes}}$$



All possible outcomes

Probability (Ball is red) =



$$= \frac{4}{30}$$

- Probability of complement of an event A is 1 - P(A)
- Probability of an impossible event is 0, i.e. $P(\Phi) = 0$
- For any event B, $0 \leq P(B) \leq 1$
- If event A is a subset of event B ($A \subset B$), then $P(A) \leq P(B)$

Probability

Question:

The new vaccine is to be tested on several patients. There are 5 diabetic patients with same type of diabetes, 9 non-diabetic patients with same heart condition and 11 non-diabetic patients with same liver condition. One patient is randomly chosen. What is the probability that the patient is not diabetic?

Probability

Solution:

Let A: event that a patient is chosen

Total number of patients = 5 + 9 + 11 = 25

Number of diabetic patients = 5

Probability that a diabetic patients is chosen, i.e. P(A) = $\dfrac{\text{Number of diabetic patients}}{\text{Total number of patients}}$

Probability

Solution:

$$P(A) = \frac{\text{Number of diabetic patients}}{\text{Total number of patients}} = \frac{5}{25} = 0.2$$

The required probability is the a diabetic patients is not chosen, i.e P(A’)

P(A’) = 1 - P(A) = 1- 0.2 = 0.8

Odds

Odds

- Odds of an event is the ratio of the number of observations in favour of an event to the number of observations not in favour of the event

$$\text{odds} = \frac{\text{number of observations in favour of the event}}{\text{number of observations not in favour of the event}}$$

- Probability is sometimes expressed in terms of odds

Odds

Odds in favour of an event A are a:b

$$P(A) = \frac{a}{a+b}$$

$$P(A') = \frac{b}{a+b}$$

$$P(A') = 1 - P(A)$$

Odds

Question:

The odds that a New Yorker picked at random will be either overweight or obese are 14:11. What is the probability that the person is fit (is not overweight or obese)?

Odds

Solution:

Let A: a New Yorker is obese

The odds in favour of A are 14:11

Here, a = 14 and b = 11

The probability that the person is fit (is not overweight or obese) is 1-P(A) = P(A')

$$P(A') = \frac{b}{a+b} = \frac{11}{11+14} = \frac{11}{25} = 0.44$$

Summary

Probability of an event is the ratio of number of observations in favour of an event to all possible observations

$$\text{probability} = \frac{\text{number of observations in favour of the event}}{\text{number of observations}}$$

Odds of an event is the ratio of number of observations in favour of an event to number of observations not in favour of the event

$$\text{odds} = \frac{\text{number of observations in favour of the event}}{\text{number of observations not in favour of the event}}$$

Independence of Events

Independence of Events

- Let A and B be two events, they are said to be independent if occurrence or non-occurrence of one event does not affect the occurence of others
- Events A and B, are said to be independent if $P(A \cap B) = P(A) . P(B)$
- The results hold vice versa. For two events A and B if $P(A \cap B) = P(A) . P(B)$, then the events A and B are independent

Independent event

Question:

Consider rolling a fair die twice :

A: First trial - Occurrence of a number greater than 3

B: Second trail: Occurence of a number greater than 3

Are the events A and B independent? Why.

Independent event

Solution:

Here the events are:

A: Occurrence of a number greater than 3

B: Occurrence of a number greater than 3

i.e., A = {4, 5,6} and B = {4, 5, 6}

$P(A) = ⅓$ and $P(B) = ½$

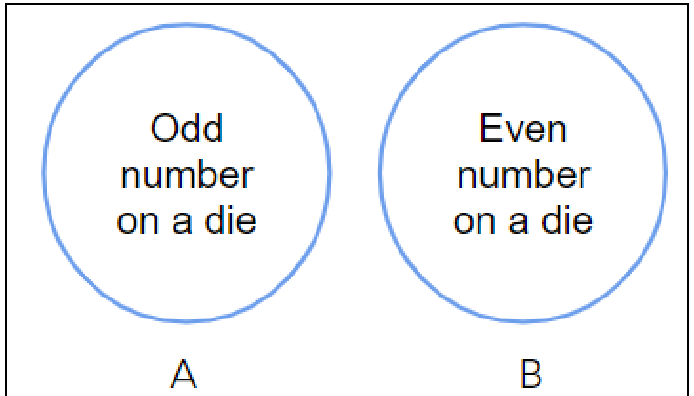
$P(A \cap B) = ⅙ = P(A).P(B)$

Thus, events A and B are independent.

Independence and mutually exclusive events

- If A and B are independent events with P(A) and P(B) both non-zero, then A and B cannot be mutually exclusive

- If A and B are mutually exclusive with P(A) and P(B) both non-zero, then A and B cannot be independent i.e A and B are dependent



Mutually exclusive

Question:

A card is drawn from a pack of 52 cards. Let events A and B be:

A: Occurrence of a club card

B: Occurrence of a red ace card

Are the events A and B mutually exclusive? Why.

Solution:

The club card is a black card and event B is the occurrence of a red card. Implies, $A \cap B$ is empty. Thus the events A and B mutually exclusive.

Conditional Probability

Conditional Probability

- The conditional probability of an event A is the probability that the event will occur given the knowledge that an event B has already occurred

- It is denoted by,

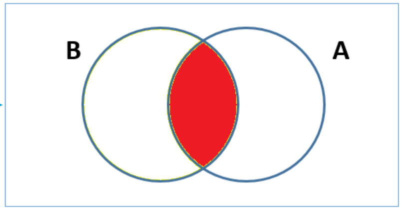
$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

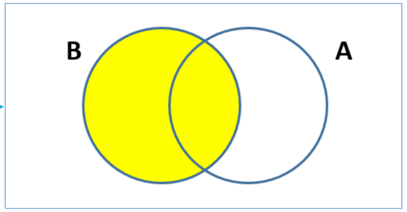
Conditional Probability

- The conditional probability of an event A is the probability that the event will occur given the knowledge that an event B has already occurred

- It is denoted by,

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$





Conditional probability

Question:

A pair of fair dice is rolled. If the product of numbers that appear is 6, find the probability that the second die shows an even number?

Conditional probability

Solution:

Let A: the event of getting the product as 6

The ways A can occur: {(1,6), (2,3), (3,2), (6,1)}

Let B: the event that the second die shows an even number

The ways B can occur: {(1,2), (1,4), (1,6), (2,2), (2,4), (2,6), (3,2), (3,4), (3,6), (4,2), (4,4), (4,6), (5,2), (5,4), (5,6), (6,2), (6,4), (6,6)}

Thus, the event that product of the die is 6 and number 2 appears on the dice is A ∩ B.

A ∩ B = {(1,6), (3,2)}

The total number of samples is 36.

First Dice

Second Dice	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Conditional probability

Solution:

The required probability is

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)}$$

$$P(\text{number on the second die is even} \mid \text{the product is 6}) = \frac{\text{P (product is 6 and number on the second die is even)}}{\text{P(the product is 6)}}$$

$$P(\text{number on the second die is even} \mid \text{the product is 6}) = \frac{\frac{2}{36}}{\frac{4}{36}} = \frac{2}{4} = 0.5$$

PLEASE!! NOTE

Let A and B be two events:

Addition theorem: $$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Multiplication theorem:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

$$\Rightarrow P(A \cap B) = P(A \mid B).\,P(B)$$

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}$$

$$\Rightarrow P(A \cap B) = P(B \mid A).\,P(A)$$

Bayes’ Theorem

Bayes' theorem

- Conditional probability is the likelihood of an event given that another event has occurred
- Bayes theorem provides a way to update the probability based on the new information
- It is completely based on the conditional probability
- Also known as Bayes' Rule or Bayes' law

Bayes' theorem

Suppose L_1, L_2 and L_3 are the production lines of an industry and D is the event that the article manufactured is defective. Given the article is defective, we need to find the probability that it was produced on which line.

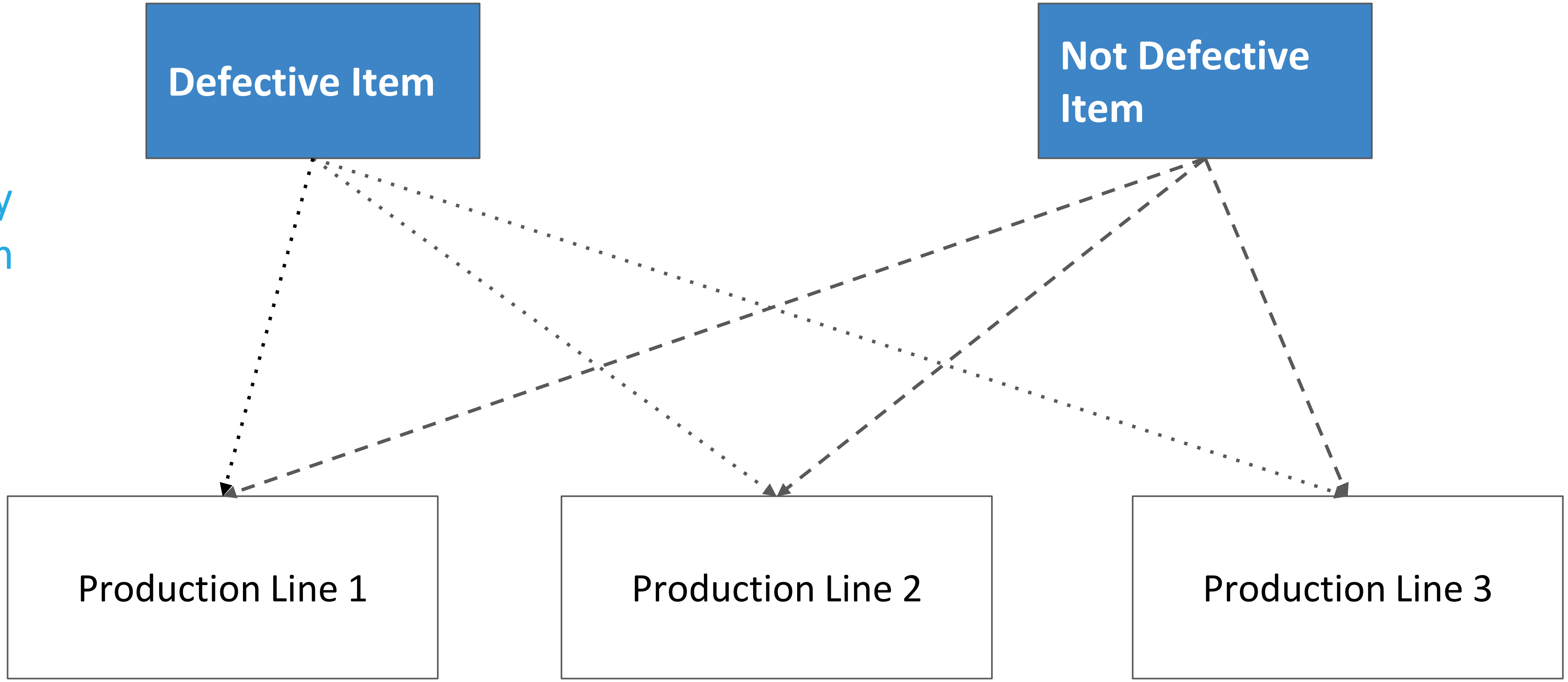
$$P(L_i \mid D) = \frac{P(L_i).P(D|L_i)}{\sum_{j=1}^{3} P(L_j).P(D|L_j)} \quad \text{for all } i = 1, 2,3$$

Which production line do I belong to?

Defective Item

Bayes' theorem

What is the probability that the defective item chosen was from the production line 1?



$P(L_1) = 0.3$ $P(L_2) = 0.4$ $P(L_3) = 0.3$

Bayes' theorem

Suppose events $A_1, A_2, \ldots, A_n$ form a partition on a sample space Ω of a random experiment and B is any other event such that $P(B) > 0$ defined on Ω, then

Prior Probabilities

Likelihood

$$P(A_i \mid B) = \frac{P(A_i).P(B|A_i)}{\sum_{j=1}^{n} P(A_j).P(B|A_j)} \quad \text{for all } i = 1, 2, \ldots, n$$

Posterior Probabilities

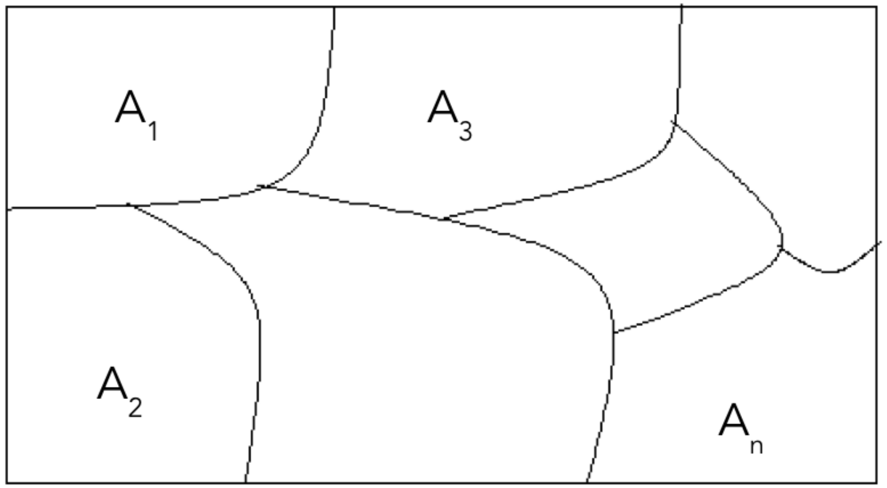
Evidence

Partition of a sample space

The events $A_1, A_2, \ldots, A_n$ defined on a sample space Ω are said to form a partition if and only if

- $A_i \cap A_j = \Phi$ for all i and j; $i \neq j$ (mutually exclusive)

- $\bigcup_{i=1}^{n} A_i = \Omega$ (exhaustive)



Prior probabilities

- Consider three production lines L_1, L_2, L_3 manufacturing the same finished articles
- From the past experience, it is seen that the chance of selecting a production line is 0.3, 0.4, 0.3 respectively
- These probabilities are called the prior probabilities

Posterior probabilities

- Consider three production lines L_1, L_2, L_3 manufacturing the same finished articles

- An article is chosen at random and is found to be defective

- The knowledge of the probability that the defective article is produced by L_1, L_2 or L_3 will help in improving the quality of the product line, i.e. $P(L_1|$ defective article$)$

- These probabilities are called the posterior probabilities

- The probability is computed after the production line manufactures the articles

Likelihood

- Consider three production lines L_1, L_2, L_3 manufacturing the same finished articles

- Likelihood is the conditional probability that the article produced is defective given it is produced on one of the lines L_1, L_2, L_3. i.e. P(defective article | L_1)

- It specifies the chance of producing a defective article of the given production line

Evidence

- Event D is that the article selected is defective, which is not used in the calculation of the prior probability
- In our example, the evidence is the probability of selecting a defective article produced on any of the production lines L_1, L_2 or L_3
- It is also known as marginal probability

Bayes' theorem

Question:

In an armament production station, the explosion can occur due to short circuit, fault in the machinery, negligence of workers. From experience, the chances of these causes are 0.1, 0.3, 0.6 respectively. The chief engineer feels that an explosion can occur with probability:

- 0.3 if there is a short circuit
- 0.2 if there is a fault in the machinery
- 0.25 if the workers are negligent

Given that an explosion has occurred, determine the most likely cause of it?

Bayes' theorem

Solution:

Let the events be defined as

A_1: the explosion can occur due to short circuit

A_2: the explosion can occur due to fault in the machinery

A_3: the explosion can occur due to negligence of workers

The prior probabilities are

$P(A_1) = 0.1,$ $P(A_2) = 0.3,$ $P(A_3) = 0.6$

Bayes' theorem

Solution:

Let B be the event there is an explosion.

Thus the probabilities are

P(there is an explosion| there is a short circuit) = $P(B \mid A_1) = 0.3$

P(there is an explosion| there is a fault in the machinery) = $P(B \mid A_2) = 0.2$

P(there is an explosion| the workers are negligent) = $P(B \mid A_3) = 0.25$

Bayes' theorem

Solution:

To find: the most likely cause of it given that an explosion has occurred

Thus, using the bayes theorem, to compute the probability that there is a short circuit given an explosion has occurred is computed as

$$P(A_1|B) = \frac{P(A_1).P(B|A_1)}{P(A_1).P(B|A_1) + P(A_2).P(B|A_2) + P(A_3).P(B|A_3)}$$

$$P(A_1|B) = \frac{0.1\times0.3}{0.1\times0.3+0.3\times0.2+0.6\times0.25} = \frac{1}{8} = 0.125$$

Bayes' theorem

Solution:

Similarly,

$$P(A_2|B) = \frac{0.3\times0.2}{0.1\times0.3+0.3\times0.2+0.6\times0.25} = \frac{1}{4} = 0.25$$

$$P(A_3|B) = \frac{0.6\times0.25}{0.1\times0.3+0.3\times0.2+0.6\times0.25} = \frac{5}{8} = 0.625$$

Bayes' theorem

Solution:

Thus, we have

P(there is a short circuit | there is an explosion) = $P(A_1 \mid \mathbf{B}) = 0.125$

P(there is a fault in the machinery| there is an explosion) = $P(A_2 \mid \mathbf{B}) = 0.25$

P(the workers are negligent| there is an explosion) = $P(A_3 \mid \mathbf{B}) = 0.625$

Comparing the probabilities, $P(A_3 \mid \mathbf{B})$ has highest probability. It implies that the negligence of workers is the most likely cause of an explosion in the factory.

Summary

Terminologies	Meaning
Prior Probability	Probability of the event A (P(A))
Posterior Probability	Conditional probability of the event A given event B has occurred (P(A|B))
Likelihood	Conditional probability of the event B given event A has occurred (P(B|A))
Evidence	Probability of the event B (P(B))

Thank you!

Happy Learning :)