Lesson 1 - Probability Concepts

We face uncertainty all the time and probability is the way in which we try to measure that uncertainty. For example, how likely is it that it will rain today? What is the probability that our cricket team will win? What is the likelihood of getting heads when you toss a coin?

This week, you will learn about the concept of probability including the laws of probability and how to apply these rules to solve simple problems. This will lay the foundation for many of the concepts that you will be learning in the upcoming lessons.

Introduction

We face uncertainty all the time and probability is the way in which we try to measure that uncertainty. For example, how likely is it that it will rain today? What is the probability that our cricket team will win? What is the likelihood of getting heads when you toss a coin?

This week, you will learn all about the concept of probability including the laws of probability and how to apply these rules to solve simple problems. This will lay the foundation for many of the concepts that you will be learning in the upcoming lessons.

Learning Outcomes:

After completing this lesson, you will be able to explain the concept of uncertainty and probability. Further, you will be able to describe the laws of probability and apply them for solving problems.

Basic Probability Concepts

What is probability?

A probability is a numeric value representing the chance, likelihood or possibility that a particular event will occur, such as the day being rainy or your favourite cricket team winning.

It is a proportion or fraction whose value lies between 0 and 1 (including the values 0 and 1). An event that has no chance of occurring (an impossible event) has a probability of 0 and an event that is sure to occur (a certain event) has a probability of 1.

Elements of probability theory

You will need the following definitions to understand probability theory:

- ***Experiment***: The observation of some activity or the act of taking some measurement is called an experiment.
- ***Outcome***: The result of an experiment is called an outcome.
- ***Sample space***: The collection of all possible outcomes of a random experiment is called the sample space.
- ***Event***: A collection of one or more outcomes of an experiment is an event.
- ***Simple event***: An individual outcome which can't be broken down further is called a simple event.
- ***Complement***: The complement of an event, A, includes all events in the sample space that are not part of event A. It is denoted as A'.
- ***Joint event***: A joint event takes places when more than one event takes place at the same time.
- ***Mutually exclusive events***: Two events are said to be mutually exclusive if they can't both occur together when the experiment is performed.
- ***Collectively exhaustive events***: A set of events is collectively exhaustive if one of the events must occur.

Example 1: *Consider the act of rolling a six-sided die once.*

Experiment*: rolling a die*

Outcome*: the number appearing on the uppermost face*

Sample space*: all the numbers that could appear when you roll the die, {1,2,3,4,5,6}. Note that you denote the events from an experiment as a set using curly brackets {}.*

Simple event*: There are six possible simple events from rolling a die.*

E.g. Let event A = rolling the number two. This is a set consisting of just one outcome, {2}

Event*: Any of the simple events, or a set or subset of the simple events constitute an event.*

E.g. Let event B = rolling an even number. This involves the set of outcomes, {2,4,6}.

Complement*:*

The complement of event A, A' = rolling a number other than 2. This is the set {1,3,4,5,6}

The complement of event B, B' = rolling a non-even, or odd, number. This is the set {1,3,5}

Joint event*: A joint event involves the combination of more than 1 event.*

E.g. Let event C = rolling a number other than 2 AND rolling an even number. This is a joint event, which consists of the set of outcomes {4,6}

Mutually exclusive events*: A and A' are mutually exclusive as you cannot both roll the number 2 and not roll the number 2. Similarly, events A and B' are mutually exclusive as you cannot roll the number 2 and, at the same time, roll an odd number.*

However, the events A and B are not mutually exclusive, as rolling the number 2 is a subset of the event rolling an even number.

***Collectively exhaustive events:** A and A' are collectively exhaustive as together they make up the full sample space. Similarly, B and B' are collectively exhaustive.*

Question 1

Suppose you and a friend each toss a coin. Let H and T denote Heads and Tails respectively. Match the following with the associated probability concept.

Good job!

Tossing the two coins is an experiment *with four possible outcomes:*

{HH, HT, TH, TT} - this is the sample sample *of all possible outcomes*

Each outcome or combination of outcomes are also called events.

For example, ***obtaining two heads {HH} is an outcome*** *but is also an event.*

The complement of an event is includes all events in the sample space that are not part of that particular event.

For example, the event of getting heads at least once includes the following set of outcomes: {HH, HT, TH}.

The complement of this event includes all other outcomes of the experiment. In this case, the only remaining outcome is {TT}. Therefore, ***the complement of getting heads at least once is getting two tails or {TT}****.*

Score: 4

The responses are incorrect. Try reviewing the lesson again.

Score: 0

Tossing two coins

Score: 0

Obtaining two heads

Score: 0

{HH, HT, TH, TT}

Score: 0

{TT}

Score: 0

Approaches to Probability

How do we determine the probability of an outcome occurring? There are three approaches to determining the probability of an event:

1. *A priori* or classical approach
2. Empirical or relative frequency approach
3. Subjective approach

In *a priori* probability, the probability of an event is calculated based on the assumption that the outcomes of the experiment are equally likely.

So if an experiment has T different possible outcomes, the probability of a single outcome occurring is given as:

$$Probability \ of \ occurrence \ of \ a \ single \ outcome =\ \frac{1}{T}\ $$

The probability of a particular event occurring is given as:

$$Probability \ of \ A \ occurring = \ \frac{Number \ of \ ways \ that \ A \ can \ occur}{T} \ $$

Example 2: *Let's go back to the example of rolling a six-sided die once. Suppose we want to calculate the probability of event A (rolling the number 2) and event B (rolling an even number).*

There are six outcomes in total that can take place, so T=6, and we assume that the die is fair so that each outcome has an equal chance of occurring.

Event A is a simple event, i.e. it consists of a single outcome. Therefore,

$$Probability \ of \ rolling \ the \ number \ 2 =\ P(A) = \ \frac{1}{6}\ $$

Event B consists of three possible outcomes, {2,4,6}. Therefore,

$$Probability \ of \ rolling \ an \ even \ number =\ P(B) = \ \frac{3}{6} = \frac{1}{2} \ $$

Under the empirical approach, probabilities are based on observed past frequency of occurrence, for example, through surveys. So if a random experiment is repeated n times and a particular outcome is observed x times in these trials in the past, the probability of that outcome is given as:

$$Probability \ of \ occurrence =\ \frac{Number \ of \ times \ the \ event \ occurred \ in \ the \ past}{Total \ number \ of \ trials} = \frac{x}{n} $$

Example 3: *Suppose that the results of a survey indicate that 60% of students have part-time jobs, then we take 0.6 to be the probability that an individual student has a part-time job.*

The subjective approach is usually used when it is not possible to use a priori or empirical probabilities. A subjective probability differs from person to person and will be based on a combination of an individual’s past experience, personal opinion and analysis of a particular situation.

Example 4: *A development team for a new product may assign a probability of 0.6 to the chance of product success whereas the CEO of the company may be less optimistic and assign a probability of 0.3.*

Probability in real life

Tossing a coin has two outcomes, which means that the probability of obtaining heads is 0.5. Does this mean that if we tossed a coin ten times, each time recording the outcome, we would obtain heads for exactly five of those tosses? No, because we cannot say for certain what will happen each of those ten times. However, if we were to keep on tossing the coin over and over, in the long run, the proportion of heads obtained will approach 0.5.

Question 2

Which approach to probability do the following scenarios use?

Excellent!
The probability of selecting a red ball is based on the assumption that each outcome is equally likely and is simply calculated as the number of red balls divided by the total number of balls: 2/5. Therefore, this is the classical approach.

Using historical or survey data to measure probability is the empirical approach. Therefore, the probability of mobile phone ownership among population aged 15-30 and the probability of rain in April fall under the empirical approach.

The political scientist is setting the probability based on a hunch. Another person may set a different probability based on his/her views. This is an example of subjective probability.

Score: 4

Some of your responses are incorrect. Review the lesson and try again.

Score: 0

You do a survey and find that 90% of the population aged 15-30 owns a mobile phone.

Score: 0

The probability of choosing a red ball from a bag containing 3 blue balls and 2 red balls is 2/5.

Score: 0

A political scientist has a hunch that the current government has a 40% chance of winning the next election.

Score: 0

A study of the rainfall data for April reveals the probability of rain on any given day in April is 35%.

Score: 0

Probability Notation and Venn Diagrams

Some probability concepts are denoted using mathematical symbols. These include

$P(A)$: the probability of event A occurring

$P(A \cup B)$: P(A or B) or the probability that A or B or both occur (the symbol $\cup$ denotes a union of events)

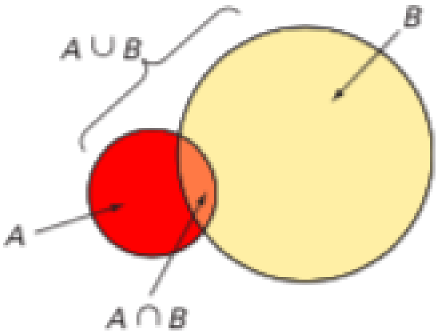
$P(A \cap B)$: P(A and B) or the probability that A and B both occur (the symbol $\cap$ denotes an intersection of events)

$P(A')$ or $P(A^c)$: the probability that A does not occur or the probability that the complement of A occurs

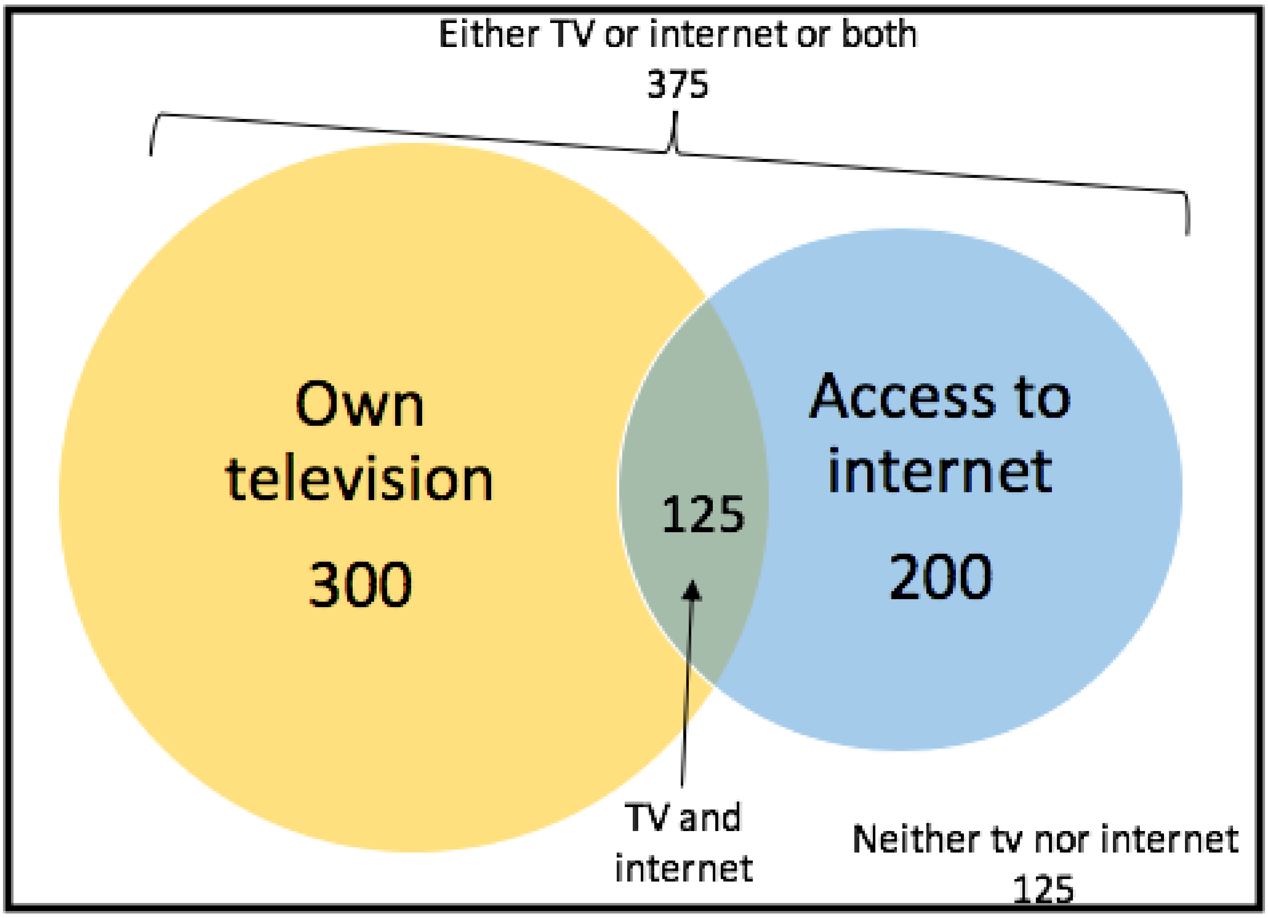
Venn diagrams are commonly used to illustrate the sample space of an experiment, with various events represented as the “unions” or “intersections” of circles.

In the diagram below:

- the red circle denotes event A and the yellow circle denotes event B
- the area common to both circles (the orange area) is the intersection of A and B ($P(A \cap B)$)
- the total area of the two circles is the union of A and B ($P(A \cup B)$)
- the area outside of the two circle contains outcomes that are neither part of A nor B



Example 5: *Suppose from a survey of 500 households, we find that 300 own televisions and 200 have access to internet.125 households both own a television and have access to internet. This means that 300-125=175 households only own a television and 200-125=75 households only have access to the internet. The remaining 125 households have neither a television nor access to internet. This can be shown diagrammatically using a Venn diagram.*



The information represented in the Venn diagram can also be displayed in a tabular format where frequencies across categories are cross tabulated. We call this kind of table a contingency table.

Example 6: ____________

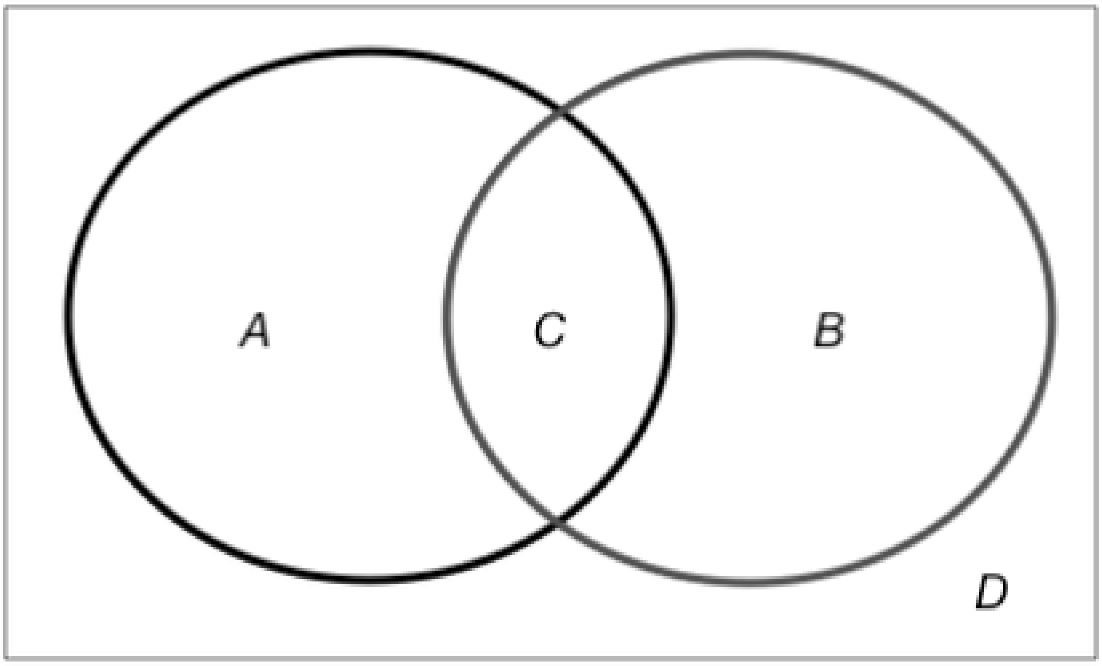
Owning a television	Access to internet		
	Yes	No	Total
Yes	125	175	300
No	75	125	200
Total	200	300	500

Question 3

Consider a sample of 10 students. 3 of them like only Maths and 2 like only Science. 2 of them like both Maths and Science.

See the Venn diagram below: the black circle (on the left) corresponds to the set of students who like Maths and the grey circle (on the right) corresponds to the set of students who like Science.

Match the number of students to the corresponding areas A, B, C, and D.



Good job!

A corresponds to the set of students who like only maths, B is the set of the students who likely only science, C is the set that like both maths and science and D is the set of students who don't like either subject.

A = 3,

B=2,

C=2,

D=10-3-2-2=3

Score: 4

Your responses are incorrect. Review the notes and try again.
Score: 0

A

Score: 0

B

Score: 0

C

Score: 0

D

Score: 0

Conditional Probability

Up to now we calculated probabilities of events sampled from the entire sampling space. However, given certain additional information about the events involved, we might be able to update the probability of a particular event.

Conditional probability refers to the probability of event A, given information about the occurrence of another event, B. We denote this conditional probability as $P(A|B)$, the probability of event A given event B occurs.

The conditional probability of event A given the occurrence of event B is calculated as:

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

where $P(B)>0$.

We can also write, the conditional probability of event B taking place, given that event A has taken place as:

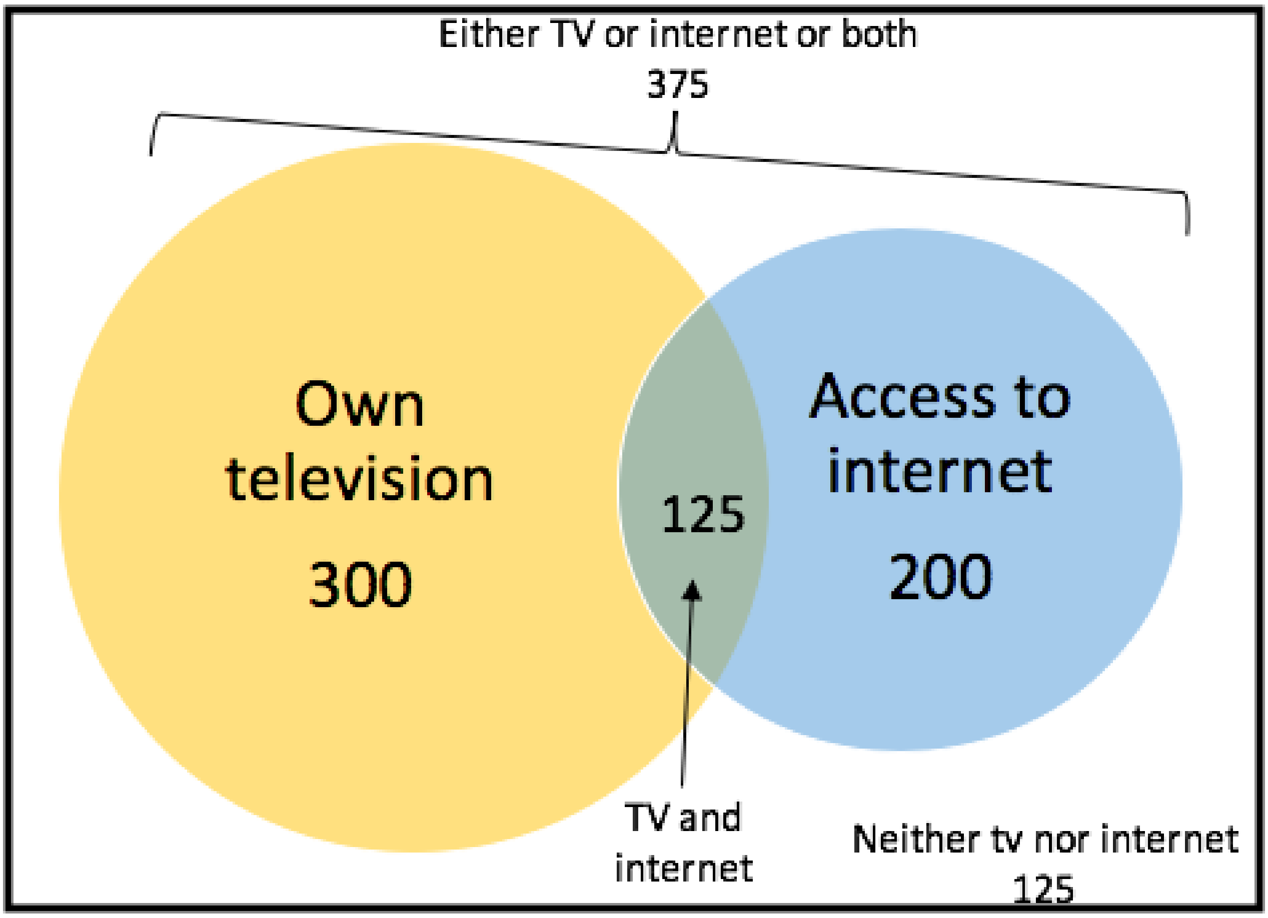
$$P(B|A)=\frac{P(A \cap B)}{P(A)}$$

If knowing that event B occurred had no bearing on the probability of event A, we say that event A and B are independent. In that case, we have:

$$P(A|B)=P(A)$$

Example 7: *Going back to the previous example where we have information on households owning televisions and having access to internet, suppose we are told that a particular household owns a television. Now, what is the probability that the household also has access to internet?*

Since we are told that the household has a television, the sample space is no longer all 500 households covered by the survey. It is only the 300 households that own a television. We know that out of these households 125 also have access to internet.



Therefore, the probability of having access to internet, given that a household owns a television is:

$$P(Internet|TV)= \frac{125}{300} = 0.42$$

We can also use the equation given above:

$$P(Internet|TV)=\frac{P(Internet \ AND \ TV)}{P(TV)} = \frac{125/500}{300/500} = \frac{125}{300} = 0.42$$

Question 4

Let us go back to the previous example where out of 10 students: 5 students liked Maths (3 liked only Maths and 2 liked both Maths and Science) and 4 students liked Science and 2 students liked both Maths and Science.

What is the probability that a student chosen randomly likes Maths, given that he also likes Science?

0.5

Score: 1

0.4

Score: 0

0.36

Score: 0

We cannot say

Score: 0

Question 5

You calculated the probability that a student likes maths (event A), given that he likes science (event B) as 0.5. Which of the following statements about A and B is correct?

A and B are mutually exclusive

Score: 0

A and B are mutually exhaustive

Score: 0

A and B are independent

Score: 0

A and B are not independent.

Score: 1

Joint and Marginal Probability

Joint probability measures the likelihood that two or more events will happen at the same time. In other words, it is the probability of a joint event. The probability of events A and B taking place at once is denoted by the probability of the intersection of those events, $P(A \cap B)$.

Marginal or simple probabilities refer to the probability of occurrence of simple events. For instance, the probability of event A taking place, P(A). We are interested in the probability of A taking place, regardless of any other events. The marginal probability can also be written as the sum of joint probabilities. This is explained in the Rule of Total Probability described in the following section.

Contingency tables can be useful for identifying joint and marginal probabilities.

Example 8: *We can use the contingency table constructed in the previous example to calculate joint and marginal probabilities of our choice.*

Owning a television	Access to internet		
	Yes	No	Total
Yes	125	175	300
No	75	125	200
Total	200	300	500

Let A=Event that a selected household owns a television

B=Event that a selected household has access to internet

Then, $A \cap B$ = The joint event that a selected household owns a television and has access to internet, with $P(A \cap B)$ the joint probability.

$$P(A \cap B) = \frac{125}{500} = 0.25$$

We can also calculate the marginal probability of owning a television by considering all households owning televisions regardless of internet access (this is the sum of the first-row frequencies) and dividing by the total number of households.

$$P(A) = \frac{(125+175)}{500} = 0.6$$

Question 6

Consider the following contingency table based on a sample of 15 individuals:

	Right-handed	Left-handed
Male	6	2
Female	3	4

Match the following events to the correct probabilities.

Great job!

P(Male) = (6+2)/15 = 0.53

P(Right handed AND female) = 3/15 = 0.2

P(Left handed) = (2+4)/15 = 0.4

Score: 3

Some of your responses are incorrect. Review the material and try again!

Score: 0

P(Male)

Score: 0

P(Right handed AND Female)

Score: 0

P(Left handed)

Score: 0

Rules of Probability

Probability is governed by some mathematical rules. These are explained below.

Assume that E_i is a simple event in a sample space consisting of n outcomes so $S=\{E_1, E_2, \dots , E_n\}$

$$0 \leq P(E_i) \leq 1$$

$$\sum\limits_{i=1}^n P(E_i) = 1$$

Complement rule

For any event A,

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

For any two events A and B,

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

And if A and B are mutually exclusive:

$$P(A \cup B) = P(A) + P(B)$$

Multiplication rule

For any two events A and B, the formula for conditional probability can be rearranged so that:

$$P(A \cap B) = P(A).P(B|A) = P(B).P(A|B)$$

When events A and B are independent, $P(B|A)=P(B)$ and $P(A|B)=P(A)$ so that the above equation simplifies to:

$$P(A \cap B) = P(A).P(B)$$

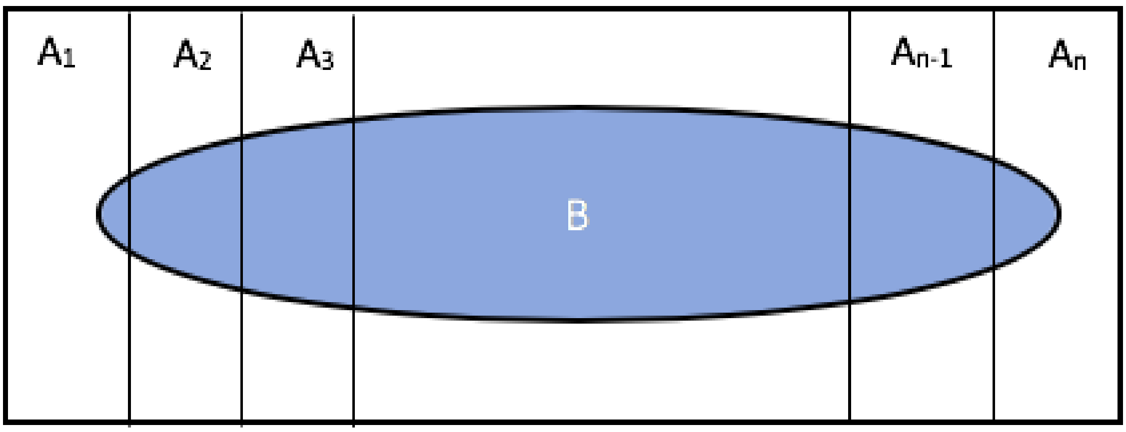
Rule of total probability

Suppose events $A_1, A_2, ..., A_n$ are mutually exclusive and exhaustive, then for any event B,

$$P(B)=\sum\limits_{i=1}^n P(A_i).P(B|A_i)$$

This derives from the multiplication rule and the observation that since $A_1, A_2, ..., A_n$ are mutually exclusive and exhaustive, we have:

$$P(B) = P(A_1 \cap B)+P(A_2 \cap B)+ ... + P(A_n \cap B)$$



Bayes' Theorem

Bayes' theorem is used to revise a previously calculated probability based on new information and combines what you learnt about conditional probability with the rule of total probability to give:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{\sum\limits_{i=1}^n P(A_i).P(B|A_i)}$$

Example 9: *The probability that a person has a certain disease is 0.03. Medical tests are able to determine whether a person actually has the disease. If the person actually has the disease, the probability that the test gives a positive result (indicating the presence of the disease) is 0.9. If the person does not actually have the disease, the probability of a positive test result is 0.02. Suppose that an individual has taken the test and got a positive result. What is the probability that that individual actually has the disease? What is the probability of a positive test result?*

Let

Event D = has disease Event D' = does not have disease

Event T = positive test result Event T' = negative test result

Note that D and D' are mutually exclusive as well as mutually exhaustive. Same goes for T and T'.

So:

$P(D)=0.03$, $P(D')=1-0.03=0.97$

$P(T|D)=0.9$, $P(T|D')=0.02$

Then, using Bayes' Theorem, the probability that an individual actually has the disease given a positive test result is:

$$P(D|T) = \frac{P(T|D)P(D)}{P(T|D)P(D)+P(T|D')P(D')} = \frac{0.9*0.03}{0.9*0.03 + 0.02*0.97} = 0.582$$

And the probability of a positive test result, P(T), is the denominator in the Bayes' formula, 0.046.

Question 7

Suppose we have two events A and B. The calculation of $P(A \cup B)$ requires subtracting $P(A \cap B)$ because $P(A)$ and $P(B)$ both include $P(A \cap B)$.

True

Score: 1

False

Score: 0

Question 8

10% of patients entering a clinic have liver disease.

Among those patients diagnosed with liver disease, 14% are alcoholics. Among patients not diagnosed with liver disease, 7% are alcoholics.

What is the probability of a patient having liver disease given that they are an alcoholic? (give the answer rounded to two decimal places)

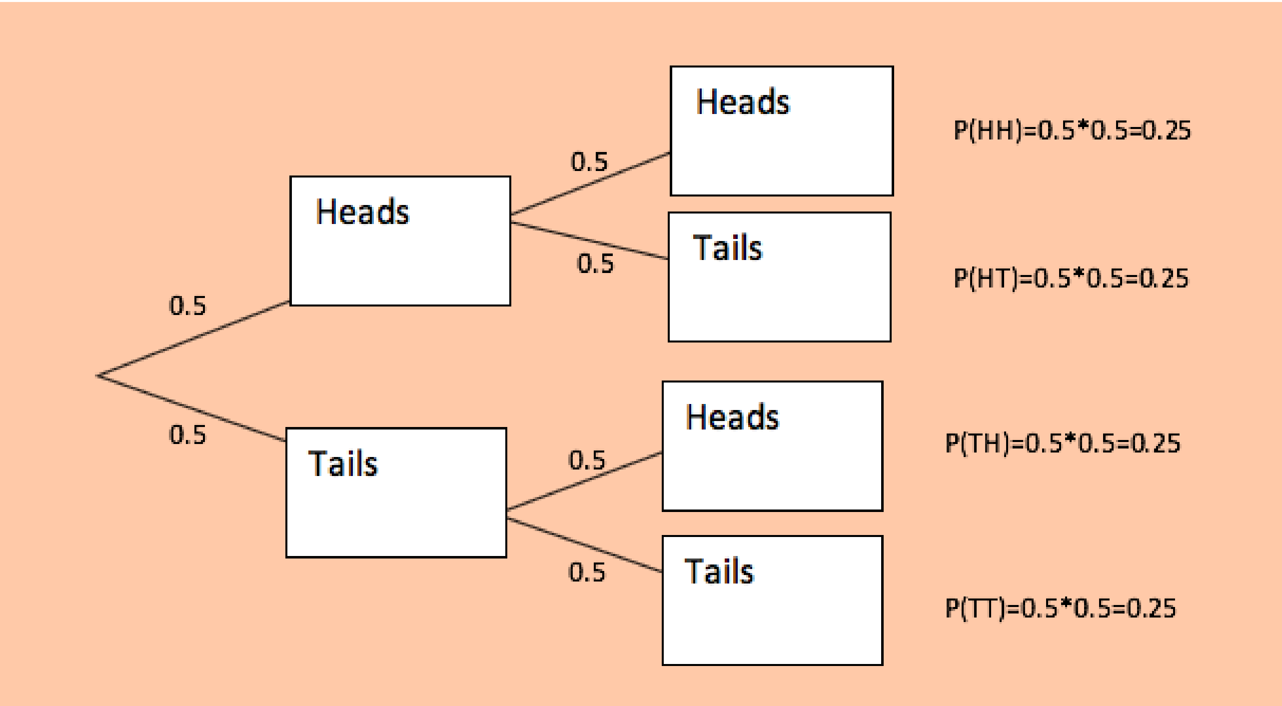
0.18
Score: 1

*
Score: 0

Decision Trees

Sometimes it is easy to understand a problem if we can visualize it. A decision tree (or tree diagram) is a graphical representation of an experiment and its events. The outcomes of an experiment are represented as leaves while the probabilities are the branches. An example will clarify what this means.

Example 10: *Consider an experiment of flipping a coin twice. In the first flip, we get either Heads or Tails, with equal probability, and the same goes for the second flip. The branches of the decision tree denote each of the four outcomes that can result from the experiment: {HH, HT, TH, TT}.*



*In this example, the outcome of the first coin toss is independent of the outcome of the second coin toss. Therefore, we can easily obtain the joint probabilities of each of the four outcomes by multiplying the probabilities along the corresponding branches. For example, the probability of obtaining heads in the first toss and tails in the second toss, P(HT)=0.5*0.5=0.25.*

Counting Rules

We calculate the probability of occurrence of an event as the number of ways that event can occur divided by the total number of all possible outcomes. When there is a large number of possible outcomes, calculating the exact number can be difficult so we have some counting rules that will be useful.

Counting rule 1

If any one of k different mutually exclusive and collectively exhaustive events can occur on each of n trials, the number of possible outcomes is equal to: k^n

Example 11: *If you toss a coin three times, since there are two possible events that can occur for each toss, the number of possible outcomes is: $2^3=8$*

{HHH, HHT, HTT, HTH, TTT, TTH, THT, THH}

Counting rule 2

This is a generalization of the first rule to the case when each trial has a different number of possible outcomes.

Suppose there are k_1 outcomes of the first trial, k_2 on the second trial,…, and k_n events on the n^{th} trial, the number of possible outcomes is equal to: $k_1*k_2*…*k_n$

Example 12: *If you have four shirts, two trousers and two pairs of shoes, the number of outfits that can be chosen is 4*2*2=16*

Counting rule 3

This rule computes the number of ways that a set of items can be arranged in order.

The number of ways that n items can be arranged in order is: $n!=n*(n-1)*(n-2)*…*1$

Example 13: *How many passwords can you form using the letters A, B, and C, where each letter can only be used once? The answer is 3! = 3*2*1=6*

{ABC,ACB,BAC,BCA,CBA,CAB}

Counting rule 4

The fourth rule tells us how many ways you can select a subset from a larger group of items, when the order of selection matters. Each possible arrangement is called a permutation.

The number of ways of arranging x objects selected from n objects, in order, is:

$$ ^n P_x = \frac{n!}{(n-x)!} $$

Example 14: *How many 2 letter passwords can you form using the letters A, B, C and D, where each letter can only be used once?*

$$ ^4 P_3 = \frac{4!}{(4-3)!} = 4! = 24 $$

Counting rule 5

The last counting rule also counts the number of ways that a subset can be selected from a larger group, but here the order of selection does not matter. In this case, each possible selection is called a combination.

The number of ways of selecting x objects from n objects, irrespective of order, is:

$$ ^n C_x = \frac{n!}{x! (n-x)!} $$

Since we do not care about order, the number of combinations will always be less than the number of permutations.

Example 15: *How many 3-member committees can you form out of a group of 6 people?*

Since the order in which the three members are chosen does not matter, we count the combinations:

$$ ^6 C_3 = \frac{6!}{3! (6-3)!} = \frac{6*5*4}{3*2*1} = 20 $$

Note that $0!=1$

Question 9

How many 3 digit numbers can be formed using the digits 1,2,3,4, and 5? (Assume that the digits cannot be repeated)

60
Score: 1

*
Score: 0

Question 10

How many ways can you give 4 textbooks to 8 students?

70
Score: 1

*
Score: 0