Probability

Contents

2 Probability – Foundations

Probability is the area of mathematics which deals with non-deterministic phenomena or occurrences.

There are several ideas central to the topic, these will be defined first; everything we come to study in this area can only be properly understood if these fundamental definitions are kept in mind.

2.1 First Definition of Probability

We will start by defining events, experiments and probability. We will introduce these ideas, with the example of the dice, in an intuitive framework and then move on to the formal mathematical definitions.

2.1.1 Definition: An Event

An *event* is a well-defined occurrence that may or may not happen.

The event is *well-defined* so we know whether or not it has happened.

2.1.2 First Definition: Probability

The *probability* of an event is a measure of how likely it is to occur. It is a number between 0 and 1, where a probability of 0 for an event means it will not occur and a probability of 1 means it definitely will occur. The higher the probability of an event, the more likely it is to occur.

2.1.3 Notation for probability

Mathematically, an event can be denoted by any symbol. The probability of event E is written mathematically as

$$P[E].$$

Note the use of a functional-style notation; $P[.]$.

The argument may be a description of the event if it is mathematical; if X is a variable, the probability it produces a value above 4 is:

$$P[X > 4].$$

More of this in due course. Here is a simple example of the notation.

“Met Eireann has stated that there is a 35%
chance of rain in Dublin today.”

Set R to be the event that there is rain in Dublin today, however it is actually defined.

The casual ‘35% chance’ becomes the probability of 0.35 and we write:

$$P[R] = 0.35.$$

2.2 Second Definition of Probability

The first definition given for the probability of an event invoked a number. To allow us to define this number for the definition of probability, we need another concept, that of an experiment.

2.2.1 Definition: Experiment

An action, with a well-defined set of outcomes, that can be repeated a large number of times, is called an *experiment*.

In the case of the dice, throwing the dice is the experiment, with the numbers which come up as the outcomes.

The events such as ‘getting an even number’ or ‘getting a 5’ are now defined as made up of the outcomes.

We will now give a definition of probability with these concepts which will give us a number.

2.2.2 Second Definition: Probability

Let E denote an event, which is the possible outcome of an experiment. Let N be the number of times the experiment is repeated. Let N_E be the number of times event E turns up. Then the probability of event E is defined as:

$$P[E] = \lim_{N \to \infty} \frac{N_E}{N}.$$

This is the *frequentist* interpretation of probability: objective, empirical and data-driven.

The is a hidden assumption in this definition, that the limit

$$P[E] = \lim_{N \to \infty} \frac{N_E}{N}.$$

does indeed converge.

The Law of Large Numbers is a statement that this will happen; it can be proved or alternatively regarded as an axiom to establish probability science.

2.2.3 The Single Dice as an Experiment

In the throw of one fair dice, the set of possible outcomes is

$$\{1, 2, 3, 4, 5, 6\}$$

If the dice is thrown N times, N a large number, each face comes up approximately the same number of times n, and

$$6n \approx N.$$

The proportion of times each number comes up is then approximately

$$\frac{n}{N} \cong \frac{1}{6}.$$

As the number of throws N increases without limit, the ratio becomes closer to the value 1/6, then

$$P[E] = \frac{1}{6}.$$

2.2.4 Empirical Data and Probabilities

Now let is look at some empirical examples, where we rely on data. The first two involve full data, coming from the Census of 2011 and maternity hospital data. The third involves a *sample*.

Empirical Data	*Probability*
Clinical data suggests that 8% of the population born in Ireland has O Negative blood-type.	The probability that a person chosen at random and born in Ireland has this blood-type is 0.08.
A recent census established that 6.2% of the population of Ireland have no religion.	The probability that a person in Ireland, chosen at random, has no religion is 0.062.

A company manufactures a certain type of component, by the thousands. A random sample of 200 components is taken and 5 are found to be faulty.	The company management concludes that the probability of an individual component being faulty is 0.025.

2.2.5 Rolling Two Dice

We will now look at the throw of two dice. For this case, we will calculate the probability of the result adding up to 5.

Firstly, let E be this event. Theoretically, to find the probability, we need the number N and N_E, for

$$P[E] = \frac{N_E}{N}\text{, evaluated as } N \text{ gets bigger.}$$

There are 36 possible outcomes in all. These outcomes are all equally likely to come up as the experiment is repeated a large number of times. Only 4 have the sum of 5:

(1, 4), (2, 3), (3, 2) and (4, 1)

This list of outcomes is the event ‘the sum equals 5’.

The proportion of times the sum of 5 comes up is then

$$P[E] = \lim_{N \to \infty} \frac{N_E}{N} = \frac{4}{36}, \text{ so } P[E] = \frac{1}{9}.$$

2.2.6 The Sample Space

In the previous example, there was an important distinction between the outcome of the experiment and the event, which is that they add up to 5. More specifically, the outcomes were the pairs

$$(1,1), (1,2), \ldots, (1,6), (2,1), (2,2), \ldots, (6,1), (6,2), \ldots, (6,6).$$

The event was the list of 4 outcomes:

$$(1, 4), (2, 3), (3, 2) \text{ and } (4, 1).$$

Therefore we see clearly that the event is a subset of this list of outcomes.

2.2.7 Definition of Sample Space

The sample space is the set of all possible outcomes of an experiment. Any event which arises as a result of that experiment is a subset of the sample space.

2.2.8 An Example of the Sample Space

For the case of two dice, we will calculate the probability that the sum is below 6. We calculate the sum for each possible outcome:

	1	*2*	*3*	*4*	*5*	*6*
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

This is equivalent to the sample space.

To find the probability of this event, count those outcomes where the sum is below 6; this is 10, so the probability is 10/36.

The full information is:

((1, 1), 2), ((1, 2), 3), …, ((1,6),7), ((2,1), 3), ((2,2), 4), …, ((6,1), 7), ((6,2),8) …, ((6,6), 12).

We are now getting on to the idea of a random variable, the ‘sum’, is the criteria which determines whether the event has happened or not.

2.2.9 The Difference

In the same experiment, rolling two dice, we calculate the probability that the difference of the two numbers is 2.

This is essentially the same list of outcomes, but now with an alternative label, the difference, which is drawn from the list of numbers:

$$\{0, 1, 2, 3, 4, 5\}.$$

Here is the event space for 'the difference':

	1	2	3	4	5	6
1	0	1	2	3	4	5
2	1	0	1	2	3	4
3	2	1	0	1	2	3
4	3	2	1	0	1	2
5	4	3	2	1	0	1
6	5	4	3	2	1	0

There are 8 results where the difference is 2, so the probability is:

$$8/36 = 2/9.$$

In the case of rolling two dice and noting the sum,

- The experiment was rolling the two dice.
- The sample space was the set of all pairs of numbers, with the sum/difference noted. These are the outcomes of the experiment.

2.2.10 Complementary Events

Now calculate the probability that the sum is equal to or greater than 6. From our existing information, the number of outcomes equal to or greater than 6 is 36 – 10 = 26. The probability is then 26/36.

This leads us to a particularly important rule.

In an experiment with a number of possible outcomes, let A be an event and let B be the exactly opposite event. The event B is called the complement of A. Then

$$P[A] + P[B] = 1.$$

In other words, when two distinct events between them cover all possibilities eventualities, the sum of their probabilities is 1.

2.2.11 The Birthday Enigma

Here is an example which gives what seems like an unusual result, using the same ratio of outcomes from a sample space.

> "In a group of 25 people, calculate the chances that two or more share the same birth date."

This event is more likely than not; consider a slightly more abstract, but simpler, version of the same question.

> "20 people are asked to choose a number from 1 to 100. Calculate the probability that two or more people pick the same number."

Let S be the event that two or more of the numbers chosen are the same and let T be the complementary event that none of the numbers are the same.

The probability $P[T]$ can be found by seeing how many ways in which the 20 people can pick their numbers with no restrictions, the quantity N. Then find the number of outcomes which satisfy the criteria for the event, the quantity N_E.

We are implicitly assuming that each outcome is equally likely to come up. The probability is then the ratio

$$P[T] = \frac{N_T}{N}.$$

The number of unrestricted possible combinations of the 20 numbers is:

$$100 \text{ x } 100 \text{ x } \dots \text{ x } 100 = 100^{20} = 10^{40}.$$

For the number of ways numbers can be chosen so none are the same:

- The first person has a choice of the full 100.

- To ensure the first number is not picked again, the second person has the choice of the other 99 numbers.
- To ensure the first or second number is not picked again, the third person has the choice of the other 98 numbers.
- Continuing like this, the last person will have a choice of 81 numbers, so the full number of combinations will be 100 x 99 x 98 x …x 81

This number is

$$100 \times 99 \times 98 \times \ldots \times 81 = {}^{100}P_{20} = 1.3\text{x}10^{39}.$$

If the 20 numbers are chosen at random, the probability that none are the same is

$$P[T] = \frac{^{100}P_{20}}{100^{20}} = 0.13$$

and so $P[S] = 1 - 0.13 = 0.87$.

2.2.12 Birthdays

Now return to the original question of birthdays. In a group of 25, calculate the probability that two or more people share a birthday.

Let S be the event that two or more of the people share the same birthday and let N be the complementary event that none are the same. Then

$$P[N] = \frac{{}^{365}P_{25}}{365^{25}}.$$

The figures are:

$$365^{25} = 1.14\text{x}10^{64}, \text{ and } {}^{365}P_{25} = 4.92\text{x}10^{63}.$$

The probability of nobody having a birthday in common is then

$$P[N] = 0.43, \text{ so } P[S] = 0.57.$$

2.2.13 Definition – Mutually Exclusive Events

Two events A and B, the possible results of the same experiment, are said to be *mutually exclusive* if it is impossible for them to happen together. If two events A and B are not mutually exclusive, that is, A and B can occur together, they are said to be mutually *non-exclusive* events.

This allows us to discuss the following scenario.

2.2.14 Three or More Complementary Events

In the case of throwing two dice and noting the sum, we calculate the probability of the following events; the sum is less than 7, the sum is equal to 7, and then the sum is greater than 7. Counting the outcomes which give the correct respective sum, there are

15, 6, and 15 possibilities respectively.

The probabilities are then

15/36, 6/36 and 15/36.

Note that these all add up to 1, as they should from the counting process.

Say an experiment has n mutually exclusive possible events:

$$E_1, E_2, E_3, \ldots, E_n,$$

and no others. Then

$$P[E_1] + P[E_2] + \ldots + P[E_n] = 1.$$

In other words, when all distinct possibilities have been taken care of, the sum of the probabilities is 1.

A generalisation of the rule for complementary events.

2.3 The Laws of Probability

We have already seen the rule for complementary events. If A is an event and B is the exact opposite event, then

$$P[A] + P[B] = 1.$$

For n distinct possible events:

$$E_1, E_2, E_3, \ldots, E_n,$$

and no others, then

$$P[E_1] + P[E_2] + \ldots + P[E_n] = 1.$$

The 'addition law' concerns how we calculate the probability of the event 'A or B'.

2.3.1 The Addition Law

If two events A and B are mutually exclusive, then the following law holds:

$$P[A \text{ or } B] = P[A] + P[B].$$

If the two events are not mutually exclusive, the law above is broadened so that the probability of A or B occurring is given by:

$$P[A \text{ or } B] = P[A] + P[B] - P[A \text{ and } B].$$

2.3.2 An Example of the Addition Law

If a single dice is thrown, we determine the probability of getting a multiple of 2 or a multiple of 3, and then the probability of one or the other. Let A be the event of getting a multiple of 2, and B be the event of getting a multiple of 3.

The probability of scoring a multiple of 2 is that of getting a 2, 4 or 6 so that

$$P[A] = 3/6 = ½.$$

Getting a multiple of 3 means a 3 or a 6 so the probability of this is:

$$P[B] = 2/6 = 1/3.$$

Now both of these events could happen at the same time; 6. It then follows that A and B are non-exclusive events, so

$$P[A \text{ or } B] = P[A] + P[B] - P[A \text{ and } B].$$

The event 'A and B' is equivalent to getting a 6. Thus

$$P[A \text{ and } B] = 1/6.$$

The final probability is then

$$P[A \text{ or } B] = 2/6 + 3/6 - 1/6 = 4/6 = 2/3.$$

2.4 Independent and Dependent Events

Consider the rolling of a dice on two occasions. The outcome of the first throw will not affect the probabilities for the second. In this context we say that each throw of the dice has *no memory*.

Now consider the event of drawing a lemon from a bag of lemons and limes without replacing it and the second event of drawing a lime after this. Clearly the probability of the second event depends on whether the first event *does* or *does not* happen.

2.4.1 Definition – Independent Events

Two events are *independent* when the occurrence of one event does not affect the probability of the occurrence of the second event. If the outcome of one event does affect the probability of the second event, they are said to be *dependent*.

Consider again the experiment of picking fruit from a bag of 5 lemons and 10 limes. Let A be the event of picking a lemon from the bag, without replacement and let B be the event of picking a lime after this.

From a counting process:

$$P[A] = 5/15 = 1/3.$$

The fruit is not replaced. If the first fruit *was* a lemon, there are 4 lemons among the 14 remaining fruit, so then:

$$P[B] = 10/14.$$

If the first fruit *was not* a lemon, there are 5 lemons in the 14 remaining fruit. Therefore:

$$P[B] = 9/14.$$

The event B is dependent on A.

2.4.2 The Multiplication Law of Probability (1)

For two *independent events* A and B, the probability of the occurrence of both events A and B, is given by

$$P[A \text{ and } B] = P[A]P[B]$$

The probability of the occurrence of both events is the product of the two individual probabilities.

2.4.3 An Example of Independent Events

A single fair dice is thrown 4 times. Calculate the probability of getting the same number four times in a row.

The event of rolling the die each time is an independent event. Clearly the probability of getting a given number four times in a row is

$$(1/6)^4 = 1/6^4.$$

For the second problem, let A_i be the event of getting four i's in a row. Then the second event of the same number four times in a row is

$$A_1 \text{ or } A_2 \text{ or } \ldots \text{ or } A_6$$

So it follows that

$$P[A_1 \text{ or } A_2 \text{ or } \ldots \text{ or } A_6] = P[A_1] + P[A_2] + \ldots + P[\ A_6]$$

Each of these events has the same probability, so therefore

$$P[A_1 \text{ or } A_2 \text{ or } \ldots \text{ or } A_6] = 6\text{x}1/6^4 = 1/6^3.$$

2.4.4 The Binomial Distribution

Given it is known that 11% of the Irish population is left-handed, find the probability of finding 3 people who are left-handed in a randomly selected group of 12 Irish people.

Look at one case of the first 3 of the 12 being left-handed.

If the group have been randomly selected, the probability that each individual being left-handed is 0.11.

If there is no connection between them, the multiplicative law applies so the probability of the first 3 being left-handed and therefore the remaining 9 being right-handed, is

$$0.11^3 0.89^9.$$

Any selection of 3 from the 12 could be the relevant left-handed people and the probability of that selection is the same since multiplication is commutative.

These possible outcomes are all mutually exclusive, so to find the overall probability, add on this probability for each combination.

In other words, multiply by ${}^{12}C_3$. The probability of 3 left-handed people in a group of 12 is therefore

$$ {}^{12}C_3 0.11^3 0.89^9. $$

This is an example of what is called the Binomial distribution; let A be the outcome of a trial, a ‘success’, which can occur with a probability p. The trial is repeated n times. The probability of getting r successes is

$$ {}^{n}C_r p^r (1-p)^{n-r}. $$

We will return to this in greater detail.

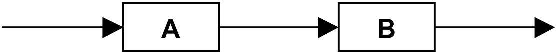
2.5 Some Examples of Complex Systems

In each of the following systems, the probability that each individual component of type A will fail is 0.03, and the probability that each individual component of type B will fail is 0.05.

All components A and components B are independent.

For each system, we will calculate the probability that the system overall will work.

2.5.1 Example 1

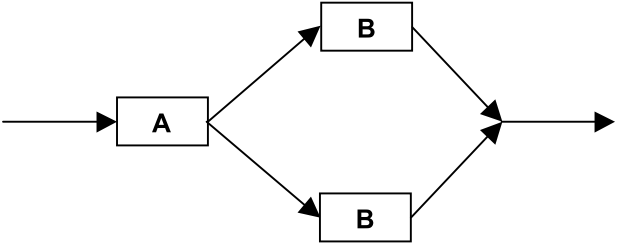


For this example, clearly both components have to work so that the system overall works. The two events, *A* and *B*, are independent, so the probability they both work is:

$$0.97 \times 0.95 = 0.9215.$$

This then is the probability the system works.

2.5.2 Example 2



In this case, going from left to right, the first component A has to work, followed by the second stage, which is *either* one B or the other. For the system overall, both stages must work. Therefore:

$$P[\text{system works}] = P[\text{stage 1 works and stage 2 works}] =$$

$$= P[\text{A}] \text{ x } P[\text{stage 2 works}].$$

$$= P[\text{A}] \text{ x } P[\text{B or B}].$$

In the second stage, either *B* working are not mutually exclusive. So use the second version of the addition law, this is

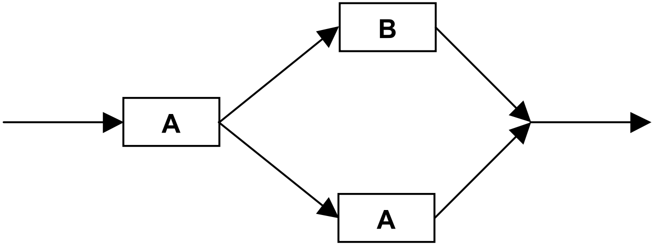
$$P[\text{B or B}] = P[\text{B}] + P[\text{B}] – P[\text{B}]P[\text{B}].$$

The laws then mean that the probability the whole system works is

$$P[\text{system works}] = P[\text{A}] \,(\, P[\text{B}] + P[\text{B}] – P[\text{B}]P[\text{B}] \,) =$$

$$= 0.97\text{x}0.9975 = 0.9676.$$

2.5.3 Example 3



In this case, going from left to right, the first component A has to work, followed by the second stage, which is *either* the B or the second A.

For the system overall, both stages must work.

$$P[\text{system works}] = P[\text{stage 1 works}] \text{ x } P[\text{stage 2 works}]$$

$$= P[\text{A}] \text{ x } P[\text{A or B works}]$$

So to find the probability this part of the system works, we have to use the second version of the addition law, this is just

$$P[\text{A or B}] = P[\text{A}] + P[\text{B}] – P[\text{A}]P[\text{B}].$$

The laws then mean that the probability the whole system works is

$$P[\text{system works}] = 0.97\text{x}(0.97 + 0.95 – 0.97\text{x}0.95) = 0.969.$$

2.6 Conditional Probability and Bayes Rule

To deal with dependent events, some notation is needed to indicate when the probability of one event depends on another.

2.6.1 Notation – Conditional Probability

Let A and B be two events, where event B is dependent on event A. The probability of event B, given that event A has already occurred, is:

$$P[B \mid A].$$

This is ‘the probability of B, given A.’

For two independent events A and B, by definition, the fact that A has already occurred does not affect the probability of event B. It then follows that

$$P[B \mid A] = P[B].$$

If the events are dependent, these probabilities are not the same.

2.6.2 The Multiplication Law of Probability (2)

Now consider two events A and B; with event B dependent on A. The probability of the occurrence of both events is given by

$$P[A \text{ and } B] = P[A]\, P[B \mid A]$$

The probability of both events happening is the probability of A times the probability of B, given that A has occurred.

2.6.3 An Example of Conditional Probabilities

The staff of a factory have put 100 components in a box, of which 8 are defective and 5 components are chosen at random. Identify the experiment and the event space. We will calculate the probability of the following events

(a) None of the five components are defective.

(b) One of the five components is defective.

Let A be the event in part (a) and B for part (b). Intuitively we know that

$$P[A] = \frac{92}{100}\frac{91}{99}\frac{90}{98}\frac{89}{97}\frac{88}{96} = 0.653.$$

More formally, let D_i be the event that component i is defective, N_i be the event it is not. Then the event A is

$$A = N_1, N_2|N_1, N_3|N_1 \,\&\, N_2, \text{etc.}$$

The individual probabilities are calculated with a counting process, then drawn together using the multiplication rule for conditional probabilities, giving the result above.

For event B, the probability that the first component is defective and the rest are not is

$$P[D_1]P[N_2 \mid D_1]P[N_3 \mid D_1 \cup N_2]P[N_4 \mid \ldots]P[N_5 \mid \ldots] = \frac{8}{100}\frac{92}{99}\frac{91}{98}\frac{90}{97}\frac{89}{96}.$$

If the defective component comes up in another place in the list this simply rearranges the numbers on the top line. Therefore:

$$P[A] = 5\frac{8}{100}\frac{92}{99}\frac{91}{98}\frac{90}{97}\frac{89}{96} = 0.297.$$

These probabilities could also be calculated or verified with the counting processes of combinations. For event A:

- The number of ways of picking 5 components from the 100 is ${}^{100}C_5$.
- The number of ways of picking 5 components from the 92 non-defectives is ${}^{92}C_5$.

The probability is then:

$$\frac{{}^{92}C_5}{{}^{100}C_5} = \frac{92 \times 91 \times 90 \times 89 \times 88}{100 \times 99 \times 98 \times 97 \times 96} = 0.653.$$

For event *B*:

- The number of ways of picking 5 components from the 100 is ${}^{100}C_5$.
- The number of ways of picking 5 components with 1 from the 8 defectives and 4 from the 92 non-defectives is ${}^{8}C_1$ x ${}^{92}C_4$.

The probability is then:

$$\frac{{}^{8}C_1\,{}^{92}C_4}{{}^{100}C_5} = \frac{8 \times \dfrac{92 \times 91 \times 90 \times 89}{4 \times 3 \times 2 \times 1}}{\dfrac{100 \times 99 \times 98 \times 97 \times 96}{5 \times 4 \times 3 \times 2 \times 1}} = \frac{5 \times 8 \times 92 \times 91 \times 90 \times 89}{100 \times 99 \times 98 \times 97 \times 96}.$$

2.6.4 Bayes Rule

The second multiplication law, dealing with conditional probabilities, leads to a powerful result in probability known as Bayes Rule or sometimes Bayes Theorem. Recall that

$$P[A \text{ and } B] = P[A]\, P[B \,|\, A]$$

We reorganise this equation as

$$P[B \mid A] = \frac{P[B \cup A]}{P[A]}.$$

The 'U' symbol is used for the concept of 'and' here, a usage which will become more useful soon.

2.6.5 An Example of Bayes rule

A patient is undergoing a medical check for a disease. Let T be the event the test gives a positive result, let D be the event the patient has the disease. Firstly identify the known probabilities.

- The probability of the disease occurring at random in a patient is known; $P[D]$.
- The probability of getting a positive result given a patient has the disease is known; $P[T|D]$.
- The probability that the patience gives a positive result when they do not have the disease is $P[T|D']$.

What is required is the probability the patient has the disease given a positive result, which is $P[D|T]$, and is given by Bayes rule:

$$P[D \mid T] = \frac{P[D \cup T]}{P[T]}.$$

The probability of the events T and D U T can be calculated by a second application of the multiplicative law.

Firstly, the probability $P[D$ U $T]$ is given by

$$P[D \text{ U } T] = P[T|D]\, P[D].$$

These probabilities are all known.

The probability $P[T]$ is given by

$$P[T] = P[D \cup T] + P[D' \cup T] = P[T|D]\, P[D] + P[T|D']\, P[D'].$$

Then:

$$P[D \mid T] = \frac{P[T \mid D]P[D]}{P[T \mid D]P[D] + P[T \mid D']P[D']}.$$

This version of the equation clearly shows the way in which the rule gives us the probability $P[D|T]$ from the probability $P[T|D]$.

To see how dramatic the results of the use of Bayes rule in this context can be, let us use some data:

- The probability of the disease occurring in a patient at random is $P[D] = 0.001$. This is quite rare.
- The probability of getting a positive result given a patient has the disease is $P[T|D]$ is 0.99. The test is very good.
- The probability a patient gives a positive result when they do not have the disease is low, $P[T|D'] = 0.02$.

Let us calculate the relevant probability $P[D|T]$ with Bayes rule:

$$P[D \mid T] = \frac{P[T \mid D]P[D]}{P[T \mid D]P[D] + P[T \mid D']P[D']}.$$

We need:

$$P[D'] = 1 - P[D] = 0.999.$$

Therefore

$$P[D \mid T] = \frac{0.99 \times 0.001}{0.99 \times 0.001 + 0.02 \times 0.999} = 0.0472.$$

To see why results like this may happen, rearrange the equation:

$$P[D \mid T] = \frac{1}{1 + \dfrac{P[T \mid D']}{P[T \mid D]} \times \dfrac{P[D']}{P[D]}}.$$

The probabilities in the ratios below the line are:

$$\frac{P[T \mid D']}{P[T \mid D]},$$

the ratio of getting a positive test result when the patient does not or does have the disease, and

$$\frac{P[D']}{P[D]},$$

the odds of not having the disease.

2.6.6 Odds form of Bayes Rule

We are familiar with the concept of the odds of an event happening, for an event E:

$$O[E] = \frac{P[E]}{P[E']}, \text{ where } P[E'] \neq 0.$$

By rewriting the current form of Bayes rule for the event A' and then dividing the two identities we find the following relation:

$$\frac{P[A \mid B]}{P[A' \mid B]} = \frac{P[B \mid A]P[A]}{P[B \mid A']P[A']}.$$

In terms of odds,

$$O[A \mid B] = \frac{P[B \mid A]}{P[B \mid A']} O[A].$$

The odds of event A happening, given B, is the original odds of event A happening times the ratio

$$\frac{P[B \mid A]}{P[B \mid A']}.$$

This quantity is known as the relative likelihood of B with respect to A. Essentially we have an equation that starts with the odds of an event A, then updates these odds based on the information from B.

2.6.7 Example – Crime and Punishment

A crime has occurred and only one of two individuals A and B could have committed it. Two pieces of additional evidence have emerged; the perpetrator was of blood type O neg and so is person A. Nothing is known about the blood type of B. The proportion of the population who are O neg is 0.08. We will see how this changes the odds on A's possible guilt.

In this context, we refer to the probabilities or odds before the new evidence as the *a priori* odds and those after the new evidence as the *posterior* odds.

Let H be the hypothesis that person A is the perpetrator. Before the new evidence, the odds of this are 1. Let E be the event of the evidence appearing.

- The probability of the blood found being O neg if A committed the crime is 1; $P[E|H] = 1$.
- If A did not commit the crime, then the probability of the blood being O neg is the same as for the general population; $P[E|H'] = 0.08$.
- The likelihood ratio is therefore 12.5.

The calculation with the Bayes rule then says

$$O[H \mid E] = \frac{P[E \mid H]}{P[E \mid H']} O[H] = 12.5 \text{ x } 1 = 12.5.$$

Therefore the odds of A being guilty have increased considerably.

2.7 Set Theory Formulation of Probability

The definitions and laws of probability we have seen so far will be rephrased here as laws on sets.

2.7.1 The Sample Space and the Universal Set

Let an experiment have a sample space X, from which the set U is constructed, where every event E resulting from the experiment is a subset of U. Thus for any event we can say that

$$E \subset U .$$

The set U is called the universal set, or sometimes the event space, as noted already. It is also known as the universal event; it may be viewed as the union of all possible events. In this context, set Ω to be the set

$$\Omega = \{E \mid E \subset U\}.$$

So the set Ω is the set of all subsets of U. A probability may now be thought of as a function

$$P: \Omega \to [0,1].$$

2.7.2 Complementary Events

Let A be an event resulting from the same experiment, so $A \subset U$. Let B be the complement of A, so

$$B = A' = \{x \mid x \in U, x \notin A\}.$$

Then

$$P[A] + P[B] = 1.$$

We can now rewrite the definition of mutually exclusive events.

2.7.3 Definition – Mutually Exclusive Events

Let two events A and B be the possible results of the same experiment. Let U be the corresponding universal set, so that

$$A, B \subset U .$$

The events A and B are said to be mutually exclusive if

$$A \cap B = \{\}.$$

Now say an experiment has n mutually exclusive possible events:

$$E_1, E_2, E_3, \ldots, E_n,$$

and no others, which also satisfy

$$E_1 \cup E_2 \cup \ldots \cup E_n = U \ ,$$

then

$$P[E_1] + P[E_2] + \ldots + P[E_n] = 1.$$

2.7.4 Formal Definitions

We will now give more formal definitions of the concepts introduced above.

2.7.5 Definition – Sample Space

Consider an experiment that can produce a number of outcomes. The set of all possible outcomes of the experiment is called the sample space. Any event is a subset of the sample space.

2.7.6 Definition – Event Space / Sigma Algebra

For an experiment, the event space, or universal event, is the union of all possible events that can result from the experiment. The sigma algebra is the set of all possible subsets of the sample space.

A probability will now be defined as a map from the sigma algebra to the interval [0,1]

$$P\colon \Omega \rightarrow [0,1],$$

with certain properties.

2.7.7 Third Definition: Probability

Let an experiment have a sample space X, from which we construct the universal set U. Let Ω be the sigma algebra for the experiment, so

$$\Omega = \{E \mid E \subset U\}.$$

A probability is a function P: $\Omega \rightarrow [0,1]$, with:

1. Let E_1, E_2, E_3, ..., E_n, be a list of n mutually exclusive events, then $P\left[\bigcup_i E_i\right] = \sum_i P[E_i]$. In particular, if $\bigcup_i E_i = U$, then $\sum_i P[E_i] = 1$.

2. $P[U] = 1$, $P[\{\}] = 0$.
3. Let A and B be two events, subsets of U, such that $A \subset B$. Then $P[A] < P[B]$. The more the event includes, the higher the probability. This is called monotonicity.
4. Let A and B be two events, subsets of U. Then $P[A \cup B] = P[A] + P[B] - P[A \cap B]$. This is the addition law; in the simpler case where $A \cap B = \{\}$, we have $P[A \cup B] = P[A] + P[B]$.