5 Probability

The question of what probability really is does not have a totally satisfactory answer. The mathematical approach is to regard it as a function which satisfies certain 'axioms'.

Definition Let S be a sample space. A *probability for* S is a function $\mathbb{P}$ which assigns to each event $A \subseteq S$ a real number $\mathbb{P}(A)$ and satisfies the following axioms.

Axiom 1. For every event $A \subseteq S$ we have $\mathbb{P}(A) \geq 0$

Axiom 2. $\mathbb{P}(S) = 1$

Axiom 3. If $A_1, A_2, \ldots, A_n$ are events and $A_i \cap A_j = \emptyset$ for all $i \neq j$ then

$$\mathbb{P}(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} \mathbb{P}(A_i).$$

If $A_1, A_2, \ldots$ is a countably infinite sequence of events and $A_i \cap A_j = \emptyset$ for all $i \neq j$ then

$$\mathbb{P}(A_1 \cup A_2 \cup \ldots) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

This definition was first suggested by the Russian mathematician A.N. Kolmogorov in 1933. We refer to Axioms 1, 2 and 3 as *Kolmogorov's Axioms for Probability.*

We say that events $A_1, A_2, A_3, \ldots$ satisfying $A_i \cap A_j = \emptyset$ for $i \neq j$ are *pairwise disjoint* or *mutually exclusive.*

Example Suppose S is a finite sample space. We showed in lectures that setting $\mathbb{P}(A) = \frac{|A|}{|S|}$ for each $A \subseteq S$ gives a probability for S. This is the case when every outcome in the sample space is equally likely.

Warning Do not assume that every outcome is equally likely without good reason.

Starting from the axioms we can deduce various properties. Hopefully, these will agree with our intuition about probability. The proofs that all of these properties hold are simple deductions from the axioms.

Proposition 5.1. *If A is an event then*

$$\mathbb{P}(A^c) = 1 - \mathbb{P}(A).$$

Proof We have $S = A \cup A^c$ and $A \cap A^c = \emptyset$ so, by Axiom 3, $\mathbb{P}(S) = \mathbb{P}(A) + \mathbb{P}(A^c)$. Since $\mathbb{P}(S) = 1$ by Axiom 2, we have $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$. •

Corollary 5.2.

$$\mathbb{P}(\emptyset) = 0.$$

Proof We have $S^c = \emptyset$. Proposition 5.1 and Axiom 2 now give $\mathbb{P}(\emptyset) = 1 - \mathbb{P}(S) = 0$. •

Proposition 5.3. *If A and B are events and $A \subseteq B$ then*

$$\mathbb{P}(A) \leq \mathbb{P}(B).$$

Proof Since $A \subseteq B$ we have $B = A \cup (B \setminus A)$ and $A \cap (B \setminus A) = \emptyset$. By Axiom 3, $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A)$. Since $\mathbb{P}(B \setminus A) \geq 0$ by Axiom 1, we have $\mathbb{P}(B) \geq \mathbb{P}(A)$. •

Corollary 5.4. *If A is an event then $\mathbb{P}(A) \leq 1$.*

Proof We have $A \subseteq S$. Proposition 5.3 and Axiom 2 now give $\mathbb{P}(A) \leq \mathbb{P}(S) = 1$. •

Notation If x is an outcome of the experiment then $x \in S$ and $\{x\} \subseteq S$. Hence $\{x\}$ is the event that the outcome of the experiment is x, and $\mathbb{P}(\{x\})$ is the probability that this event occurs. We will usually write $\mathbb{P}(x)$ as shorthand for $\mathbb{P}(\{x\})$, although technically the latter is more correct. Similarly we will often refer to $\mathbb{P}(x)$ as the probability of the outcome x when we should really refer to it as the probability of the simple event $\{x\}$.

Proposition 5.5. *(a) If $A = \{a_1, a_2, \ldots, a_n\}$ is a finite event then*

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(a_i).$$

(b) If $A = \{a_1, a_2, a_3, \ldots\}$ is a countably infinite event then

$$\mathbb{P}(A) = \sum_{i=1}^{\infty} \mathbb{P}(a_i).$$

Proof We prove (b). (The proof of (a) is almost identical.) Let $A_i = \{a_i\}$ for all integers $i \geq 1$. Then $A = A_1 \cup A_2 \cup A_3 \cup \ldots$ and $A_i \cap A_j = \emptyset$ for all $i \neq j$, so by Axiom 3

$$\mathbb{P}(A) = \sum_{i=1}^{\infty} \mathbb{P}(A_i) = \sum_{i=1}^{\infty} \mathbb{P}(a_i).$$

•

Proposition 5.5 tells us that we can calculate the probability of a finite or countably infinite event by adding together the probabilities of the outcomes which belong to it. In particular, if the whole sample space S is finite or countably infinite, then the probabilities of *all* events are uniquely determined by the probabilities of the outcomes. We will see an example later which shows that this statement is not true when S is uncountable.

Proposition 5.6. *For any two events A and B we have*

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

Proof The events $A \setminus B$, $B \setminus A$ and $A \cap B$ are pairwise disjoint and their union is $A \cup B$. By Axiom 3

$$\mathbb{P}(A \cup B) = \mathbb{P}(A \setminus B) + \mathbb{P}(B \setminus A) + \mathbb{P}(A \cap B) \tag{1}$$

Axiom 3 also implies that $\mathbb{P}(A) = \mathbb{P}(A \setminus B) + \mathbb{P}(A \cap B)$ so

$$\mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(A \cap B), \tag{2}$$

and $\mathbb{P}(B) = \mathbb{P}(B \setminus A) + \mathbb{P}(A \cap B)$ so

$$\mathbb{P}(B) = \mathbb{P}(B \setminus A) + \mathbb{P}(A \cap B). \tag{3}$$

We can now substitute equations (2) and (3) into (1) to obtain

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

•

Proposition 5.7. *For any three events A, B and C we have*

$$\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C).$$

Proof We use Proposition 5.6. Let $D = B \cup C$. By Proposition 5.6

$$\begin{aligned}
\mathbb{P}(A \cup B \cup C) &= \mathbb{P}(A \cup D) \\
&= \mathbb{P}(A) + \mathbb{P}(D) - \mathbb{P}(A \cap D) \\
&= \mathbb{P}(A) + \mathbb{P}(B \cup C) - \mathbb{P}(A \cap (B \cup C) \\
&= \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(B \cap C) - \mathbb{P}(A \cap (B \cup C). \qquad (4)
\end{aligned}$$

The distributive law, Lemma 2.1(c), gives

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

We can apply Proposition 5.6 to the right hand side of this equation to deduce that

$$\mathbb{P}(A \cap (B \cup C)) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap C) - \mathbb{P}((A \cap B) \cap (A \cap C)). \qquad (5)$$

Since $(A \cap B) \cap (A \cap C) = A \cap B \cap C$, we can substitute equation (5) into (4) to obtain

$$\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C).$$

•

Propositions 5.6 and 5.7 can be generalized to give a formula for calculating the probability of the union of any number of events. This general result is called the *Principle of Inclusion and Exclusion*. You should try to work out what the formula is for four events. (Hint: there are 15 terms on the right hand side.)

6 Proofs and Elementary Logic

This is not the place for a formal or philosophical discussion of what proof is. However, a willingness to think logically and to justify our assertions is a must for studying mathematics at university level. In most cases the language and ideas we use are the same as in normal writing but must be absolutely precise. In this short digression there are a few brief remarks on how to understand and write proofs. You will learn more about different sorts of proof by seeing examples as the module progresses.

We consider mathematical statements which can be either true or false. For example the statement "x is an integer and $x \geq 2$" is true when $x = 3$ and false when

$x = 1$ or $x = 22/7$. Statements which are always true are called *tautologies*. Statements which are always false are called *contradictions*. The *negation* of a statement p is the statement *not* p which is false whenever p is true and true whenever p is false. Working out the negation of a statement can be tricky so we give a few examples. Make sure you understand all of these.

- If p is the statement $x = 2$ then the negation of p is the statement $x \neq 2$
- If p is the statement "$x = 2$ and $y = 3$" then the negation of p is the statement "either $x \neq 2$ or $y \neq 3$" (where or is being used in its usual mathematical sense to include both occurring).
- If p is the statement "either $x = 2$ or $y = 3$" then the negation of p is the statement " $x \neq 2$ and $y \neq 3$".
- If p is the statement "every student at QM is hardworking" the negation of p is "there exists at least student at QM who is not hardworking".
- If p is the statement "for all $x, y \in A$ with $x \neq y$ we have $f(x) \neq f(y)$" then the negation of p is the statement "there exist $x, y \in A$ with $x \neq y$ and $f(x) = f(y)$".

Theorems (and also lemmas, propositions and corollaries) are examples of tautologies. Some tautologies are self evident, e.g. the statement "2 is an integer". Other tautologies require a reasoned argument to establish that they are always true e.g. the statement "$\mathbb{Z}$ is countable" or the statement "if x and y are even integers then $x+y$ is an even integer". To prove that a statement p is always true we have to give a logical argument which starts with something that we know is true and uses a series of deductions to show that p is always true. The first thing you should do when you want to prove that p is always true is *read the definitions* of all the terms used to state p. You have no hope of proving that p is always true if you do not understand the terms used to state p. For example we cannot prove that the statement "if x and y are even integers then $x + y$ is an even integer" is always true without a precise definition of what 'even' means.

Given two statements p and q we can make the compound statement

$$p \Rightarrow q$$

which we read as "p implies q" or equivalently "if p then q".[1] For example the statement "if x and y are even integers then $x+y$ is an even integer" is a combination

[1]The $\Rightarrow$ symbol is much abused. You should avoid using it when you write a proof. What you write is more likely to make sense if you write it out in words.

of the two statements "x and y are even integers" and "$x+y$ is an even integer". The compound statement $p \Rightarrow q$ is *true* if q is true in all situations when p is true. The compound statement $p \Rightarrow q$ is *false* if there exists a situation when p is true and q is false.

Most theorems (and propositions, lemmas and corollaries) are statements of the form $p \Rightarrow q$ is true. In this case we refer to the statement p as the *hypothesis* of the theorem and q as the *conclusion* of the theorem. A proof that $p \Rightarrow q$ is true should look something like the following:

Proof Suppose p is true. Then we have .....

So .....

Hence q is true. •

Where each sentence follows clearly from the previous ones. See for example the proof of Lemma 4.1(a). (We suppose that "X is a finite set and $f : X \to Y$ is injective". We prove "$|X| \leq |Y|$".)

To prove that $p \Rightarrow q$ is false it suffices to give one example when p is true and q is false. We call such an example a *counterexample* to the statement $p \Rightarrow q$. For example, to prove that the statement "$f : \mathbb{N} \to \mathbb{N}$ and f is injective" does not imply that "f is surjective" we only need to give one counterexample i.e. one example of function which is injective but not surjective.

Notice that $p \Rightarrow q$ and $q \Rightarrow p$ are different statements. It is quite possible that one is true and the other is false. For example "$x = 2$" implies that "$x^2 = 4$" but "$x^2 = 4$" does not imply that "$x = 2$" (we could equally well have $x = -2$).

The statement

$$p \Leftrightarrow q$$

means $p \Rightarrow q$ and $q \Rightarrow p$. This is usually read as "p if and only if q" or "p and q are equivalent". Thus $p \Leftrightarrow q$ is true means that p is true whenever q is true and q is true whenever p is true. To prove that $p \Leftrightarrow q$ is true we need to show that both $p \Rightarrow q$ *and* $q \Rightarrow p$ are true. Sometimes it is possible to do both of these at once but it is often clearer to prove them separately (see for example the proof of Theorem 4.2).

The statement $p \Rightarrow q$ is equivalent to the statement (not q) $\Rightarrow$ (not p).[2] So another way to prove that the statement $p \Rightarrow q$ is true is to show that the statement (not q) $\Rightarrow$ (not p) is true. That is we show that whenever q is false, p is also false. The statement (not q) $\Rightarrow$ (not p) is called the *contrapositive* of the statement $p \Rightarrow q$.

[2]Think about why these statements are equivalent.