1
Introduction

I shot an arrow into the air
It fell to earth, I knew not where.
H.W. Longfellow

O! many a shaft at random sent
Finds mark the archer little meant.
W. Scott

1.1 PREVIEW

This chapter introduces probability as a measure of likelihood, which can be placed on a numerical scale running from 0 to 1. Examples are given to show the range and scope of problems that need probability to describe them. We examine some simple interpretations of probability that are important in its development, and we briefly show how the well-known principles of mathematical modelling enable us to progress. Note that in this chapter exercises and problems are chosen to motivate interest and discussion; they are therefore non-technical, and mathematical answers are not expected.

Prerequisites. This chapter contains next to no mathematics, so there are no prerequisites. Impatient readers keen to get to an equation could proceed directly to chapter 2.

1.2 PROBABILITY

We all know what light is, but it is not easy to tell what it is.
Samuel Johnson

From the moment we first roll a die in a children's board game, or pick a card (*any* card), we start to learn what probability is. But even as adults, it is not easy to *tell* what it is, in the general way.

For mathematicians things are simpler, at least to begin with. We have the following:

Probability is a number between zero and one, inclusive.

This may seem a trifle arbitrary and abrupt, but there are many excellent and plausible reasons for this convention, as we shall show. Consider the following eventualities.

(i) You run a mile in less than 10 seconds.
(ii) You roll two ordinary dice and they show a double six.
(iii) You flip an ordinary coin and it shows heads.
(iv) Your weight is less than 10 tons.

If you think about the relative likelihood (or chance or probability) of these eventualities, you will surely agree that we can compare them as follows.

The chance of running a mile in 10 seconds is *less* than the chance of a double six, which in turn is *less* than the chance of a head, which in turn is *less* than the chance of your weighing under 10 tons. We may write

$$\begin{aligned}\text{chance of 10 second mile} &< \text{chance of a double six}\\ &< \text{chance of a head}\\ &< \text{chance of weighing under 10 tons.}\end{aligned}$$

(Obviously it is assumed that you are reading this on the planet Earth, not on some asteroid, or Jupiter, that you are human, and that the dice are not crooked.)

It is easy to see that we can very often compare probabilities in this way, and so it is natural to represent them on a numerical scale, just as we do with weights, temperatures, earthquakes, and many other natural phenomena. Essentially, this is what numbers are *for.*

Of course, the two extreme eventualities are special cases. It is quite certain that you weigh less than 10 tons; nothing could be more certain. If we represent certainty by unity, then no probabilities exceed this. Likewise it is quite impossible for you to run a mile in 10 seconds or less; nothing could be less likely. If we represent impossibility by zero, then no probability can be less than this. Thus we can, if we wish, present this on a scale, as shown in figure 1.1.

The idea is that any chance eventuality can be represented by a point somewhere on this scale. Everything that is impossible is placed at zero – that the moon is made of

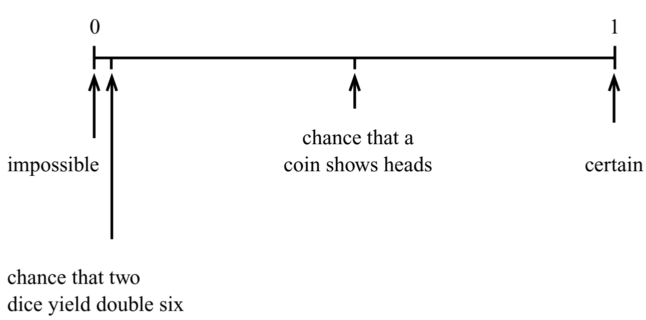

Figure 1.1. A probability scale.

cheese, formation flying by pigs, and so on. Everything that is certain is placed at unity – the moon is not made of cheese, Socrates is mortal, and so forth. Everything else is somewhere in [0, 1], i.e. in the interval between 0 and 1, the more likely things being closer to 1 and the more unlikely things being closer to 0.

Of course, if two things have the same chance of happening, then they are at the same point on the scale. That is what we mean by 'equally likely'. And in everyday discourse everyone, including mathematicians, has used and will use words such as very likely, likely, improbable, and so on. However, any detailed or precise look at probability requires the use of the numerical scale. To see this, you should ponder on just how you would describe a chance that is more than very likely, but less than very very likely.

This still leaves some questions to be answered. For example, the choice of 0 and 1 as the ends of the scale may appear arbitrary, and, in particular, we have not said exactly which numbers represent the chance of a double six, or the chance of a head. We have not even justified the claim that a head is more likely than double six. We discuss all this later in the chapter; it will turn out that if we regard probability as an extension of the idea of proportion, then we can indeed place many probabilities accurately and confidently on this scale.

We conclude with an important point, namely that the chance of a head (or a double six) is just a *chance*. The whole point of probability is to discuss uncertain eventualities *before* they occur. *After* this event, things are completely different. As the simplest illustration of this, note that even though we agree that if we flip a coin and roll two dice then the chance of a head is greater than the chance of a double six, nevertheless it may turn out that the coin shows a tail when the dice show a double six. Likewise, when the weather forecast gives a 90% chance of rain, or even a 99% chance, it may in fact not rain. The chance of a slip on the San Andreas fault this week is very small indeed, nevertheless it may occur today. The antibiotic is overwhelmingly likely to cure your illness, but it may not; and so on.

Exercises for section 1.2

1. Formulate your own definition of probability. Having done so, compare and contrast it with those in appendix I of this chapter.

2. (a) Suppose you flip a coin; there are two possible outcomes, head or tail. Do you agree that the probability of a head is $\frac{1}{2}$? If so, explain why.
(b) Suppose you take a test; there are two possible outcomes, pass or fail. Do you agree that the probability of a pass is $\frac{1}{2}$? If not, explain why not.

3. In the above discussion we claimed that it was intuitively reasonable to say that you are more likely to get a head when flipping a coin than a double six when rolling two dice. Do you agree? If so, explain why.

1.3 THE SCOPE OF PROBABILITY

> . . . nothing between humans is 1 to 3. In fact, I long ago come to the conclusion that all life is 6 to 5 against.
>
> Damon Runyon, *A Nice Price*

Life is a gamble at terrible odds; if it was a bet you wouldn't take it.
Tom Stoppard, *Rosencrantz and Guildenstern are Dead,* Faber and Faber

In the next few sections we are going to spend a lot of time flipping coins, rolling dice, and buying lottery tickets. There are very good reasons for this narrow focus (to begin with), as we shall see, but it is important to stress that probability is of great use and importance in many other circumstances. For example, today seems to be a fairly typical day, and the newspapers contain articles on the following topics (in random order).

1. How are the chances of a child's suffering a genetic disorder affected by a grand-parent's having this disorder? And what difference does the sex of child or ancestor make?
2. Does the latest opinion poll reveal the true state of affairs?
3. The lottery result.
4. DNA profiling evidence in a trial.
5. Increased annuity payments possible for heavy smokers.
6. An extremely valuable picture (a Van Gogh) might be a fake.
7. There was a photograph taken using a scanning tunnelling electron microscope.
8. Should risky surgical procedures be permitted?
9. Malaria has a significant chance of causing death; prophylaxis against it carries a risk of dizziness and panic attacks. What do you do?
10. A commodities futures trader lost a huge sum of money.
11. An earthquake occurred, which had not been predicted.
12. Some analysts expected inflation to fall; some expected it to rise.
13. Football pools.
14. Racing results, and tips for the day's races.
15. There is a 10% chance of snow tomorrow.
16. Profits from gambling in the USA are growing faster than any other sector of the economy. (In connection with this item, it should be carefully noted that profits are made by the casino, not the customers.)
17. In the preceding year, British postmen had sustained 5975 dogbites, which was around 16 per day on average, or roughly one every 20 minutes during the time when mail is actually delivered. One postman had sustained 200 bites in 39 years of service.

Now, this list is by no means exhaustive; I could have made it longer. And such a list could be compiled every day (see the exercise at the end of this section). The subjects reported touch on an astonishingly wide range of aspects of life, society, and the natural world. And they all have the common property that chance, uncertainty, likelihood, randomness – call it what you will – is an inescapable component of the story. Conversely, there are few features of life, the universe, or anything, in which chance is not in some way crucial.

Nor is this merely some abstruse academic point; assessing risks and taking chances are inescapable facets of everyday existence. It is a trite maxim to say that life is a lottery; it would be more true to say that life offers a collection of lotteries that we can all, to some extent, choose to enter or avoid. And as the information at our disposal increases, it does not reduce the range of choices but in fact increases them. It is, for example, increasingly difficult successfully to run a business, practise medicine, deal in finance, or engineer things without having a keen appreciation of chance and probability. Of course you can make the attempt, by relying entirely on luck and uninformed guesswork, but in the long run you will probably do worse than someone who plays the odds in an informed way. This is amply confirmed by observation and experience, as well as by mathematics.

Thus, probability is important for all these severely practical reasons. And we have the bonus that it is also entertaining and amusing, as the existence of all those lotteries, casinos, and racecourses more than sufficiently testifies.

Finally, a glance at this and other section headings shows that chance is so powerful and emotive a concept that it is employed by poets, playwrights, and novelists. They clearly expect their readers to grasp jokes, metaphors, and allusions that entail a shared understanding of probability. (This feat has not been accomplished by algebraic structures, or calculus, and is all the more remarkable when one recalls that the *literati* are not otherwise celebrated for their keen numeracy.) Furthermore, such allusions are of very long standing; we may note the comment attributed by Plutarch to Julius Caesar on crossing the Rubicon: ‘Iacta alea est’ (commonly rendered as ‘The die is cast’). And the passage from Ecclesiastes: ‘The race is not always to the swift, or the battle to the strong, but time and chance happen to them all’. The Romans even had deities dedicated to chance, Fors and Fortuna, echoed in Shakespeare’s *Hamlet*: ‘. . . the slings and arrows of outrageous fortune . . .’.

Many other cultures have had such deities, but it is notable that deification has not occurred for any other branch of mathematics. There is no god of algebra.

One recent stanza (by W.H. Henley) is of particular relevance to students of probability, who are often soothed and helped by murmuring it during difficult moments in lectures and textbooks:

> In the fell clutch of circumstance
> I have not winced or cried aloud:
> Under the bludgeonings of chance
> My head is bloody, but unbowed.

Exercise for section 1.3

1. Look at today’s newspapers and mark the articles in which chance is explicitly or implicitly an important feature of the report.

1.4 BASIC IDEAS: THE CLASSICAL CASE

> The perfect die does not lose its usefulness or justification by the fact that real dice fail to live up to it.
>
> W. Feller

Our first task was mentioned above; we need to supply reasons for the use of the standard probability scale, and methods for deciding where various chances should lie on this scale. It is natural that in doing this, and in seeking to understand the concept of probability, we will pay particular attention to the experience and intuition yielded by flipping coins and rolling dice. Of course this is not a very bold or controversial decision;

any theory of probability that failed to describe the behaviour of coins and dice would be widely regarded as useless. And so it would be. For several centuries that we know of, and probably for many centuries before that, flipping a coin (or rolling a die) has been the epitome of probability, the paradigm of randomness. You flip the coin (or roll the die), and nobody can accurately predict how it will fall. Nor can the most powerful computer predict correctly how it will fall, if it is flipped energetically enough.

This is why cards, dice, and other gambling aids crop up so often in literature both directly and as metaphors. No doubt it is also the reason for the (perhaps excessive) popularity of gambling as entertainment. If anyone had any idea what numbers the lottery would show, or where the roulette ball will land, the whole industry would be a dead duck.

At any rate, these long-standing and simple gaming aids do supply intuitively convincing ways of characterizing probability. We discuss several ideas in detail.

I Probability as proportion

Figure 1.2 gives the layout of an American roulette wheel. Suppose such a wheel is spun once; what is the probability that the resulting number has a 7 in it? That is to say, what is the probability that the ball hits 7, 17, or 27? These three numbers comprise a proportion $\frac{3}{38}$ of the available compartments, and so the essential symmetry of the wheel (assuming it is well made) suggests that the required probability ought to be $\frac{3}{38}$. Likewise the

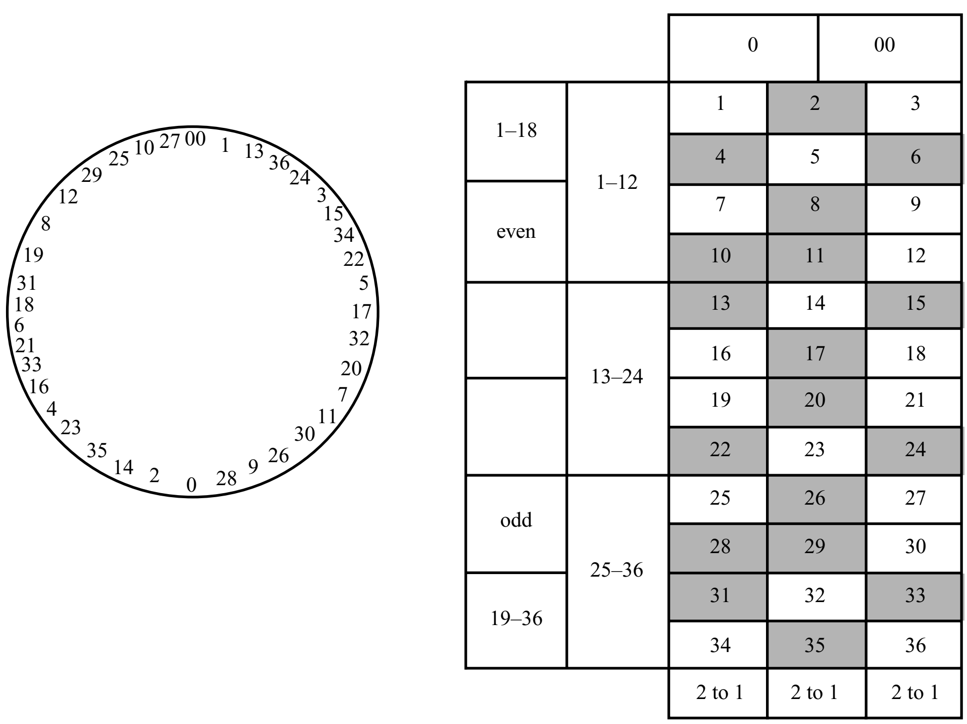

Figure 1.2. American roulette. Shaded numbers are black; the others are red except for the zeros.

probability of an odd compartment is suggested to be $\frac{18}{38} = \frac{9}{19}$, because the proportion of odd numbers on the wheel is $\frac{18}{38}$.

Most people find this proposition intuitively acceptable; it clearly relies on the fundamental symmetry of the wheel, that is, that all numbers are regarded equally by the ball. But this property of symmetry is shared by a great many simple chance activities; it is the same as saying that all possible outcomes of a game or activity are equally likely. For example:

- The ball is equally likely to land in any compartment.
- You are equally likely to select either of two cards.
- The six faces of a die are equally likely to be face up.

With these examples in mind it seems reasonable to adopt the following convention or rule. Suppose some game has n equally likely outcomes, and r of these outcomes correspond to your winning. Then the probability p that you win is r/n. We write

(1) $$p = \frac{r}{n} = \frac{\text{number of ways of winning the game}}{\text{number of possible outcomes of the game}}.$$

This formula looks very simple. Of course, it *is* very simple but it has many useful and important consequences. First note that we always have $0 \leqslant r \leqslant n$, and so it follows that

(2) $$0 \leqslant p \leqslant 1.$$

If $r = 0$, so that it is impossible for you to win, then $p = 0$. Likewise if $r = n$, so that you are certain to win, then $p = 1$. This is all consistent with the probability scale introduced in section 1.2, and supplies some motivation for using it. Furthermore, this interpretation of probability as defined by proportion enables us to place many simple but important chances on the scale.

Example 1.4.1. Flip a coin and choose 'heads'. Then $r = 1$, because you win on the outcome 'heads', and $n = 2$, because the coin shows a head or a tail. Hence the probability that you win, which is also the probability of a head, is $p = \frac{1}{2}$. ○

Example 1.4.2. Roll a die. There are six outcomes, which is to say that $n = 6$. If you win on an even number then $r = 3$, so the probability that an even number is shown is

$$p = \tfrac{3}{6} = \tfrac{1}{2}.$$

Likewise the chance that the die shows a 6 is $\frac{1}{6}$, and so on. ○

Example 1.4.3. Pick a card at random from a pack of 52 cards. What is the probability of an ace? Clearly $n = 52$ and $r = 4$, so that

$$p = \tfrac{4}{52} = \tfrac{1}{13}.$$ ○

Example 1.4.4. A town contains x women and y men; an opinion pollster chooses an adult at random for questioning about toothpaste. What is the chance that the adult is male? Here

$$n = x + y \quad \text{and} \quad r = y.$$

Hence the probability is

$$p = y/(x + y).$$ ○

It may be objected that these results depend on an arbitrary imposition of the ideas of symmetry and proportion, which are clearly not always relevant. Nevertheless, such results and ideas are immensely appealing to our intuition; in fact the first probability calculations in Renaissance Italy take this framework more or less for granted. Thus Cardano (writing around 1520), says of a well-made die: 'One half of the total number of faces always represents equality . . . I can as easily throw 1, 3, or 5 as 2, 4, or 6'.

Here we can clearly see the beginnings of the idea of probability as an expression of proportion, an idea so powerful that it held sway for centuries. However, there is at least one unsatisfactory aspect to this interpretation: it seems that we do not need ever to roll a die to say that the chance of a 6 is $\frac{1}{6}$. Surely actual experiments should have a role in our definitions? This leads to another idea.

II Probability as relative frequency

Figure 1.3 shows the proportion of sixes that appeared in a sequence of rolls of a die. The number of rolls is n, for $n = 0, 1, 2, \ldots$; the number of sixes is $r(n)$, for each n, and the proportion of sixes is

(3) $$p(n) = \frac{r(n)}{n}.$$

What has this to do with the probability that the die shows a six? Our idea of probability as a proportion suggests that the proportion of sixes in n rolls should not be too far from the theoretical chance of a six, and figure 1.3 shows that this seems to be true for large values of n. This is intuitively appealing, and the same effect is observed if you record such proportions in a large number of other repeated chance activities.

We therefore make the following general assertion. Suppose some game is repeated a large number n of times, and in $r(n)$ of these games you win. Then the probability p that

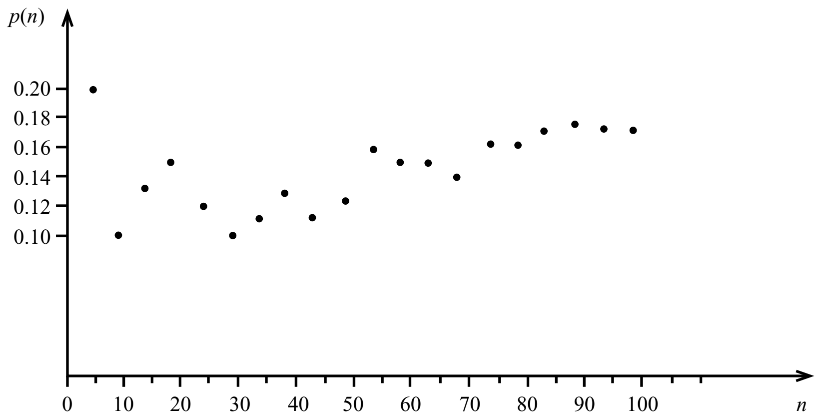

Figure 1.3. The proportion of sixes given in 100 rolls of a die, recorded at intervals of 5 rolls. Figures are from an actual experiment. Of course, $\frac{1}{6} = 0.1\dot{6}$.

Cambridge University Press & Assessment
978-0-521-64445-7 — Probability and Random Variables: A Beginner's Guide
David Stirzaker
Excerpt
More Information

you win some future similar repetition of this game is close to $r(n)/n$. We write

(4) $$p \simeq \frac{r(n)}{n} = \frac{\text{number of wins in } n \text{ games}}{\text{number } n \text{ of games}}.$$

The symbol $\simeq$ is read as 'is approximately equal to'. Once again we note that $0 \leqslant r(n) \leqslant n$ and so we may take it that $0 \leqslant p \leqslant 1$.

Furthermore, if a win is impossible then $r(n) = 0$, and $r(n)/n = 0$. Also, if a win is certain then $r(n) = n$, and $r(n)/n = 1$. This is again consistent with the scale introduced in figure 1.1, which is very pleasant. Notice the important point that this interpretation supplies a way of approximately measuring probabilities rather than calculating them merely by an appeal to symmetry.

Since we can now calculate simple probabilities, and measure them approximately, it is tempting to stop there and get straight on with formulating some rules. That would be a mistake, for the idea of proportion gives another useful insight into probability that will turn out to be just as important as the other two, in later work.

III Probability and expected value

Many problems in chance are inextricably linked with numerical outcomes, especially in gambling and finance (where 'numerical outcome' is a euphemism for money). In these cases probability is inextricably linked to 'value', as we now show.

To aid our thinking let us consider an everyday concrete and practical problem. A plutocrat makes the following offer. She will flip a fair coin; if it shows heads she will give you \$1, if it shows tails she will give Jack \$1. What is this offer worth to you? That is to say, for what fair price $\$p$, should you sell it?

Clearly, whatever price $\$p$ this is worth to you, it is worth the same price $\$p$ to Jack, because the coin is fair, i.e. symmetrical (assuming he needs and values money just as much as you do). So, to the pair of you, this offer is altogether worth $\$2p$. But whatever the outcome, the plutocrat has given away \$1. Hence $\$2p = \$1$, so that $p = \frac{1}{2}$ and the offer is worth $\$\frac{1}{2}$ to you.

It seems natural to regard this value $p = \frac{1}{2}$ as a measure of your chance of winning the money. It is thus intuitively reasonable to make the following general rule.

Suppose you receive \$1 with probability p (and otherwise you receive nothing). Then the *value* or fair price of this offer is $\$p$. More generally, if you receive $\$d$ with probability p (and nothing otherwise) then the fair price or expected value of this offer is given by

(5) $$\text{expected value} = pd.$$

This simple idea turns out to be enormously important later on; for the moment we note only that it is certainly consistent with our probability scale introduced in figure 1.1. For example, if the plutocrat definitely gives you \$1 then this is worth exactly \$1 to you, and $p = 1$. Likewise if you are definitely given nothing, then $p = 0$. And it is easy to see that $0 \leqslant p \leqslant 1$, for any such offers.

In particular, for the specific example above we find that the probability of a head when a fair coin is flipped is $\frac{1}{2}$. Likewise a similar argument shows that the probability of a six when a fair die is rolled is $\frac{1}{6}$. (Simply imagine the plutocrat giving \$1 to one of six people selected by the roll of the die.)

The 'fair price' of such offers is often called the expected value, or *expectation*, to emphasize its chance nature. We meet this concept again, later on.

We conclude this section with another classical and famous manifestation of probability. It is essentially the same as the first we looked at, but is superficially different.

IV Probability as proportion again

Suppose a small meteorite hits the town football pitch. What is the probability that it lands in the central circle?

Obviously meteorites have no special propensity to hit any particular part of a football pitch; they are equally likely to strike any part. It is therefore intuitively clear that the chance of striking the central circle is given by the proportion of the pitch that it occupies. In general, if $|A|$ is the area of the pitch in which the meteorite lands, and $|C|$ is the area of some part of the pitch, then the probability p that C is struck is given by $p = |C|/|A|$.

Once again we formulate a general version of this as follows. Suppose a region A of the plane has area $|A|$, and C is some part of A with area $|C|$. If a point is picked at random in A, then the probability p that it lies in C is given by

$$p = \frac{|C|}{|A|}. \tag{6}$$

As before we can easily see that $0 \leqslant p \leqslant 1$, where $p = 0$ if C is empty and $p = 1$ if $C = A$.

Example 1.4.5. An archery target is a circle of radius 2. The bullseye is a circle of radius 1. A naive archer is equally likely to hit any part of the target (if she hits it at all) and so the probability of a bullseye for an arrow that hits the target is

$$p = \frac{\text{area of bullseye}}{\text{area of target}} = \frac{\pi \times 1^2}{\pi \times 2^2} = \frac{1}{4}. \qquad \bigcirc$$

Exercises for section 1.4

1. Suppose you read in a newspaper that the proportion of \$20 bills that are forgeries is 5%. If you possess what appears to be a \$20 bill, what is its expected value? Could it be more than \$19? Or could it be less? Explain! (Does it make any difference how you acquired the bill?)

2. A point P is picked at random in the square $ABCD$, with sides of length 1. What is the probability that the distance from P to the diagonal AC is less than $\frac{1}{6}$?

1.5 BASIC IDEAS; THE GENERAL CASE

> We must believe in chance, for how else can we account for the successes of those we detest?
>
> Anon.

We noted that a theory of probability would be hailed as useless if it failed to describe the behaviour of coins and dice. But of course it would be equally useless if it failed to