Gerolamo Cardano

CHAPTER 1

MEASUREMENT

One way to understand the roots of a subject is to examine how its originators thought about it. Some basic philosophical issues are already evident at the very beginning. The first great idea is simply that chance can be measured. It emerged during the sixteenth and seventeenth centuries, and it is something of a mystery why it took so long. The Greeks had a goddess of chance, Tyche. Democritus and his followers postulated a physical chance affecting all the atoms that made up the universe. This is the "swerve" of atoms in Lucretius' *De Rerum Natura*. Games of chance, using knucklebones or dice, were known to Egyptians and Babylonians and were popular in Rome. Soldiers cast lots for Christ's cloak. Greek Skeptics of the later Academy postulated probability (eikos) as the guide to life.[1] Nevertheless, it appears that there was no quantitative theory of chance in these times.[2]

Figure 1.1. Determination of the lawful rood

How do you measure anything?[3] Consider length. You find a standard of equal length, apply it repeatedly, and count. The standard might be your foot, as you pace off a distance. Different feet may not lead to the same result. One refinement, proposed in 1522 for determining a lawful rood (rod), was to line up the feet of 16 people as they emerged from church, as shown in figure 1.1.[4] As the illustration shows, the various folks have very different foot lengths, but an implicit averaging effect was accepted by a group—even though the explicit notion of an average seems to not have existed at the time.

It is worth mentioning a certain philosophical objection at this point. There is a kind of circularity involved in the procedure. We are defining length, but we are already assuming that our standard remains the same length as we step off the distance.

No sensible person would let this objection stop her from stepping off distance. That is how we start. Eventually we refine our notion of length. Your foot may change length; so may the rod; so may the standard meter stick, at a fine-enough precision. Using physics, we refine the measurement of length.[5] So the circularity is real, but it indicates a path for refinement rather than a fatal objection.*

So it is with chance. To measure probability, we first find—or make—equally probable cases. Then we count them. The probability of an event A, denoted by $P(A)$, is then

$$P(A) = \frac{\text{no. of cases in which } A \text{ occurs}}{\text{total no. of cases}}.$$

Note that it follows that

1. Probability is never negative,
2. If A occurs in all cases, $P(A) = 1$,
3. If A and B never occur in the same case,

$$P(A \text{ or } B) = P(A) + P(B).$$

In particular, the probability of an event not occurring is 1 less the probability of its occurring:

$$P(\text{not } A) = 1 - P(A).$$

It is surprising how much can be done by ingenious application of this simple idea. Consider the birthday problem. What is the probability that at least two people in a room share the same birthday, neglecting leap years, assuming birthdates are equiprobable and birthdays of individuals in the room are independent (no twins)? If you have not seen it before, the results are a bit surprising.

The probability of a shared birthday in the group is 1 minus the probability that they are all different. The probability that the second person has a different birthday from the first is $(\frac{364}{365})$. If they are different, the probability that the third is different from them is $(\frac{363}{365})$, and so on, for all in the room. So the probability of a shared birthday among N people is

$$1 - \left(\frac{364}{365} \cdot \frac{363}{365} \cdot \ldots \cdot \frac{365 - N + 1}{365} \right).$$

* What are the paths open for refinement of the notion of equiprobable? They will unfold as we move through the book.

If you are interested in an even-money bet, this formula can be used to find a value of N such that the product is close to $\frac{1}{2}$. If there are 23 people in the room, the probability of a shared birthday is slightly greater than $\frac{1}{2}$. If there are 50 people, it is close to 97%.

There are many variations on the birthday problem. These are used for thinking about surprising coincidences. For instance, it is overwhelmingly likely that there are two people in the United States who share a birthday, whose fathers share the same birthday, whose fathers' fathers share this birthday, and so on, four generations back. Useful approximations for working with these variations may be found in an appendix to this chapter. These approximations are, in turn, used to prove de Finetti's representation theorem in an appendix at the end of this book. The point for now is that the basic "equally likely cases" structure has real breadth and strength.

BEGINNINGS

Nothing provides us better candidates for equiprobable cases than vigorous throws of symmetric dice or draws from a well-shuffled deck of cards. This is where the measurement of probability began. We cannot say who was there first, but the idea was clearly there in the sixteenth-century work on gambling by the algebraist, physician, and astrologer Gerolamo Cardano.[6] Cardano, who sometimes made a living at gambling, was quite sensitive to the equiprobability assumption. He knew about shaved dice and dirty deals: ". . . the die may be dishonest either because it has been rounded off, or because it is too narrow (a fault which is easily visible), or because it has been extended in one direction by pressure on the opposite faces. . . . There are even worse ways of being cheated at cards."[7]

In the early seventeenth century Galileo composed a short note on dice to answer a question posed to him (by his patron, the Grand Duke of Tuscany). The Duke believed that counting possible cases seemed to give the wrong answer. Three dice are thrown. Counting combinations of numbers, 10 and 11 can be made in 6 ways, as can 9 and 12. ". . . yet it is known that long observation has made dice-players consider 10 and 11 to be more advantageous than 9 and 12."* How can this be?

Galileo replies that his patron is counting the wrong thing. He counts three 3s as one possibility for making a 9 and two 3s and a 4 as one possibility for making a 10. Galileo points out the latter covers three possibilities, depending on which die exhibits the 4:

$$<4, 3, 3>, <3, 4, 3>, <3, 3, 4>.$$

For the former, there is only $<3, 3, 3>$. Galileo has a complete grasp of permutations and combinations and does not seem to regard it as anything new.

In constructing equiprobable cases, both Galileo and Cardano appear to make implicit use of *independence*. They suppose that for each die, all 6 faces are equally probable and that for throws of 3 dice, all 216 possible outcomes are also equally probable. When we treated the birthday problem earlier, we assumed that different people had independent chances for their birthdays.

With this basic machinery well understood, Pascal and Fermat in their famous correspondence attacked more subtle problems with a different conceptual flavor.

PASCAL AND FERMAT (1654)

The first substantial work in the mathematics of probability appears to be the correspondence between Pascal and Fermat, which began in 1654. We include a discussion for three reasons: (1) It *is* the first; (2) it shows how seemingly complex problems can be reduced to straightforward calculations with equally likely cases; and (3) it introduces the crucial notion of expectation—a mainstay of the subject.

* One strange aspect of the statement of the problem is the comment about long observation. The observation would have had to be long indeed. From Galileo's calculations, the chance of a 9 is $\frac{25}{216}$, about 0.116; the chance of a 10 is $\frac{27}{216}$, about 0.125. The difference between these is 0.009, or about $\frac{1}{100}$. As an exercise, you could calculate how many observations would be required.

Pascal and Fermat addressed problems with a different conceptual flavor from those solved by Cardano and Galileo, defining fairness and focusing on expectation.

There are two problems given to Pascal by his sometime gambling friend, the Chevalier de Méré. Pascal communicated these, together with his thoughts on them, to Fermat, with whom he had a connection through the academy of Father Mersenne. This was the *Académie Parisienne* that Mersenne formed in 1635 where the work of leading mathematicians, scientists, and philosophers—including Galileo, Descartes, and Leibniz—was shared.

The *problem of dice*: A player has undertaken to throw a 6 in 8 throws of a die. The stakes have been settled, and the 3 throws have been taken without obtaining a 6. What proportion of the stake would be fair to give the player to forego his fourth throw (just the fourth).*

The *problem of points*: Two players of equal skill† are playing a series of games. The one to win a round gets a point. They have agreed that the first to reach a certain number of points wins the game and collects the stakes. A certain number of rounds have been played, and the game is interrupted. What is a fair division of the stakes?

Both of these problems are stated in terms of fairness. But what is fairness in the theory of probability? We will see that Pascal and Fermat implicitly employ the concept of *expectation* to answer that question.

The expected value of a gamble that pays off $V(x)$ in outcome x is the probability weighted average:

$$\text{expectation } (V) = V(x_1)\, p(x_1) + V(x_2)\, p(x_2) + \cdots .$$

A transaction that leaves the players' expected values unchanged is assumed to be fair. For example, consider flipping a fair coin. If it comes up heads you win 1; if tails you lose 1. Then the expected value is $(+1)(\frac{1}{2}) + (-1)(\frac{1}{2}) = 0$.

Let's apply this idea to the *problem of dice*. The stakes, s, are still on the table. If the player does not forego his fourth throw, he has 5 throws remaining. His expectation is

* You might try to think about this directly before you proceed. Suppose that before starting, bets are laid so that there is \$10 on the table that you will get a 6 in eight throws. Would you take \$5 to forego just trial 4? Would this be fair?
† We may as well think of them as flipping a fair coin.

$\frac{1}{6}s$	+	$\frac{5}{6}\left(1-\left(\frac{5}{6}\right)^4\right)s$
(win in fourth throw)		(lose on fourth but win on 1 of the remaining 4 throws).*

Suppose that that player foregoes his fourth throw for $\frac{1}{6}$ of the stakes, as Fermat suggests in the correspondence.[8] Then his expectation is

$\frac{1}{6}s$	+	$\left(1-\left(\frac{5}{6}\right)^4\right)\left(\frac{5}{6}\right)s$
(the amount he is paid to forego that throw)		(probability of winning in the remaining 4 throws times the diminished stakes).

These are the same, so $\frac{1}{6}$ of the stakes is the fair price for foregoing the throw.[9]

The *problem of points* is also an expected-value problem. It had baffled many previous thinkers. In 1494, Fra Luca Pacioli considered a problem of points where the play is complete with 6 points; one player has won 5 and the other, 3. Pacioli—perhaps under the influence of Aristotle's proportional theory of justice—argues that the fair division is in proportion to the rounds already won, 5 to 3. About 50 years later, Tartaglia objected that, according to this rule, if the game were stopped after 1 round, 1 player would be awarded the whole stake. This consequence looks worse and worse as the number of points necessary to win is increased. Tartaglia tried to modify Pacioli's rule to take this into account, but in the end he doubted that a definitive answer was possible. The problem puzzled all who thought about it, including Cardano and the Chevalier de Méré.

Fermat had the key insight. Suppose that one player needs r points to win and the other, s. Then the game will surely be decided in $r+s-1$ rounds. It may be decided earlier, but there is no harm in considering all sequences of $r+s-1$ coin flips, since the outcome is well defined for each. This reduces the problem to one of equiprobable cases and we can calculate probabilities by counting.

So for Pacioli's problem, where player 1 has 5 points and player 2 has 3, since 6 points concludes the game, 3 more rounds will suffice. Of

* Probability of winning on one of the remaining 4 throws $= 1 - P(\text{lose on all } 4) = (1 - (\frac{5}{6})^4)$.

the 8 equiprobable cases, player 2 will win the game only if he wins all 3 rounds. His expectation is $\frac{1}{8}$ of the stakes, while player 1 has an expectation of $\frac{7}{8}$ on the stakes. It is fair, then, to divide the stakes in this proportion.

Expectation, computed by counting equiprobable cases, solves the problem. But there may be a large number of equiprobable cases to count. Consider Tartaglia's example. Six points win, and one player has no points and the other, 1 point. Then play must be complete after 10 more rounds. It would be tedious to write out the 1024 possible outcomes. But Pascal had a better way of counting.

To count the cases in which the first player wins, one adds the number of cases in which she gets 6 wins in 10 trials [called 10 choose 6] + the number where she gets 7 wins in 10 trials [10 choose 7] + · · · + the number where she gets 10 wins in 10 trials [10 choose 10]. These numbers are conveniently to be found on the tenth row of Pascal's arithmetical triangle (or Tartaglia's triangle, or Omar Khayyam's triangle[10]), which we show in figure 1.2. The row tells us the number of ways we can choose from a group of 10 objects. Reading from the left, there are 1 way of choosing nothing, 10 ways of choosing 1 object, 45 ways of choosing 2 objects, 120 ways of choosing 3, and so on, to only 1 way of choosing 10.

We want the number of ways of getting 6 wins in 10 trials + the number of ways to get 7 wins in 10 trials + · · · + the number where she gets 10 wins in 10 trials. From row 10 we get

$$210 + 120 + 45 + 10 + 1 = 386$$

for a probability of winning of

$$\frac{386}{1024} \quad \text{(about 38\%)}.$$

Thus a fair division of the stakes gives player 1 (who had no points) $\frac{386}{1024}$ of the stakes and player 2 the rest.

After Pascal and Fermat, the basic elements of measuring probability by counting equiprobable cases, calculating by combinatorial principles, and using expected value are all on the table.

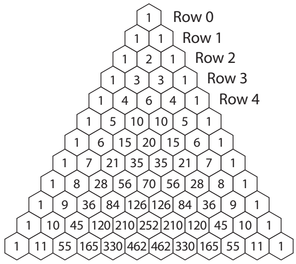

Figure 1.2. Pascal's triangle

HUYGENS (1657)

The ideas in the Pascal-Fermat correspondence were taken up and developed by the great Dutch scientist Christiaan Huygens[11] after he heard about the correspondence on a visit to Paris. He then worked them out by himself and wrote the first book on the subject in 1656. It was translated into English by John Arbuthnot in 1692 as *Of the Laws of Chance*.[12]

Huygens begins his book with a fundamental principle:

> *Postulat*
>
> As a Foundation to the following Proposition, I shall take Leave to lay down this Self-evident Truth: That any one Chance or Expectation to win any thing is worth just such a Sum, as wou'd procure in the same Chance and Expectation at a fair Lay. As for Example, if any one shou'd put 3 Shillings in one Hand, without letting me know which, and 7 in the other, and give me Choice of either of them; I say, it is the same thing as if he shou'd give me 5 Shillings; because with 5 Shillings I can, at a fair Lay, procure the same even Chance or Expectation to win 3 or 7 Shillings.

Huygens assumes that he could, in effect, flip a fair coin to choose which hand to pick.* Then $(\frac{1}{2})3 + (\frac{1}{2})7 = 5$. He then says that the value

* A point made much later by Howard Raiffa against the so-called Ellsberg paradox, which we will visit in our chapter on psychology of chance (chapter 3).

of the wager is the same as the value of 5 for sure. Thus he makes explicit (a special case of) the principle that is implicit in Pascal and Fermat: *expectation is the correct measure of value*.

He then goes on to justify this measure by a fairness argument. Suppose I bet 10 shillings with someone on the flip of a fair coin. This is fair by reasons of symmetry. Now suppose we modify this by an agreement that whoever wins shall give 3 to the loser. This preserves symmetry, so the modified arrangement is also fair. But now the loser nets 3 and the winner retains 7. Any such agreement preserves fairness, including where the winner gives the loser 5, and each has 5 for sure. Huygens then shows how the argument generalizes to arbitrary finite numbers of outcomes and arbitrary rational-valued probabilities of outcomes. It will be a recurring theme that an equality is justified by a symmetry.

NEWTONIAN CONSIDERATIONS

In the preface to the translation of Huygens, Arbuthnot, who was a follower of Newton,[13] makes the following noteworthy remark (L. Todhunter, *A History of the Mathematical Theory of Probability* (Cambridge: Macmillan, 1865); reprinted by Chelsea (New York, 1965), p. 51):

> It is impossible for a Die, with such determin'd force and direction, not to fall on such determin'd side, only I don't know the force and direction which makes it fall on such determin'd side, and therefore I call it Chance, which is nothing but the want of art.

Arbuthnot thus introduces the fundamental question of the proper conception of chance in a deterministic setting. His answer is that chance is an artifact of our ignorance.

Consider tossing a coin just once. The thumb hits the coin; the coin spins upward and is caught in the hand. It is clear that if the thumb hits the coin in the same place with the same force, the coin will land with the same side up. Coin tossing is physics, not random! To demonstrate this, we had the physics department build us a coin-tossing machine. The coin starts out on a spring, the spring is released, the coin spins upward and lands in a cup, as shown in figure 1.3. Because the forces are controlled, the coin always lands with the same side up.

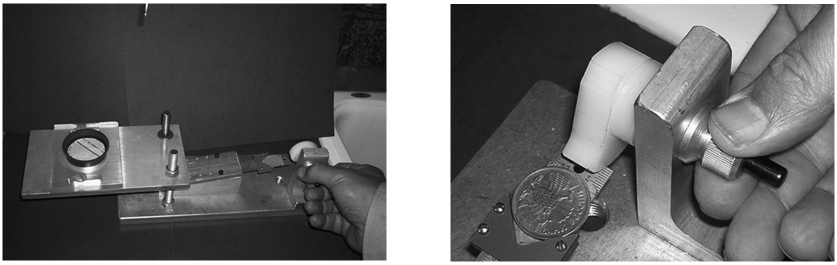

Figure 1.3. A deterministic coin-tossing machine

This is viscerally quite disturbing (even to the two of us). Magicians and crooked gamblers (including one of your authors) have the same ability.

How then is the probabilistic treatment of coin flips so widespread and so successful? The basic answer is due to Poincaré. If the coin is flipped vigorously, with sufficient vertical and angular velocity, there is sensitive dependence on initial conditions. Then a little uncertainty as to initial conditions is amplified to a large uncertainty about the outcome, where equiprobability of outcomes is not such a bad assumption. But the provisos are important. See appendix 2 for a little more on this. We will return to the question in more detail in our chapter on physical chance (chapter 9).

BERNOULLI 1713

In 1713 Jacob Bernoulli's *Ars Conjectandi*[14] was published, 8 years after his death. Bernoulli made explicit the practice of his predecessors. The first part is a reprint, with commentary, of Huygens. The probability of an event is now explicitly defined as the ratio of the number of (equiprobable) cases in which the event happens to the total number of (equiprobable) cases. The probability of being dealt a club from a deck of cards is $\frac{13}{52}$. He also defines the *conditional probability* of a second event (B) conditional on a first (A) as the ratio of the number of cases both happen to the number of cases the first happens:

$$\text{Probability}\,(B \text{ conditional on } A) = \frac{\text{no. of cases in which } A \text{ and } B \text{ occur}}{\text{no. of cases in which } A \text{ occurs}}.$$

The probability of being dealt a queen given that one is dealt a club is $\frac{1}{13}$.

On the basis of these definitions, he shows that the probabilities of mutually exclusive events add and that probabilities satisfy the multiplicative law, $P(A \text{ and } B) = P(A)P(B \text{ conditional on } A)$. These simple rules form the heart of all calculations of probability.

But Bernoulli's major contribution was to establish a rigorous connection between probability and frequency that had heretofore only been conjectured. He called this his golden theorem.

As an illustration he considers an urn containing 3000 white pebbles and 2000 black pebbles and postulates independent draws with replacement of the pebble drawn. He asks whether one can find a number of draws so that it becomes "morally certain" that the ratio of white pebbles to black ones becomes approximately 3:2. He then chooses a high probability as moral certainty and establishes a number of draws sufficient to provide a positive answer. Then he shows the weak law of large numbers:

> Given any interval around the probability (here $\frac{3}{5}$) as small as you please and any approximation to certainty, $1-e$, as close as you please, there is a number of trials, N, such that in N trials the probability that the relative frequency of draws of white falls within the specified interval is at least $1-e$.

This is a story to which we will return in our chapter on frequency (chapter 4).

SUMMING UP

Probability, like length, can be measured by dividing things into equally likely cases, counting the number of successful cases and dividing by the total number of cases. This definition satisfies the following:

1. Probability is a number between 0 and 1.
2. If A never occurs, $P(A) = 0$. If A occurs in all cases, $P(A) = 1$.
3. If A and B never occur in the same case, then $P(A \text{ or } B) = P(A) + P(B)$.

4. Conditional probability for B given A is defined by counting all the cases in which B and A occur together and dividing by the number of cases in which A occurs. Then, $P(A \text{ and } B) = P(A)P(B$ conditional on $A)$. If A and B are independent, that is, if $P(B$ conditional on $A)$ just equals $P(B)$, then $P(A \text{ and } B) = P(A)P(B)$.

Finding the cases and doing the counting leads to math problems such as the probability of winning a complicated wager or the birthday problem.

Expectation, weighting the costs and benefits of various outcomes by their chances, is useful for calculations *and* is a measure of fairness and value.

The law of large numbers, to which we will return in chapter 4 (and again in chapter 6), proves that chances can be approximated (with high probability) by frequencies in repeated independent trials.

APPENDICES

These three appendices give, respectively, a more detailed look at the correspondence between Pascal and Fermat, a development of the physics of coin tossing, and a more detailed analysis of the connection between the mathematics of probability and the real-world occurrence of chance events. (For those who might find it useful, there is a probability-refresher appendix at the end of this book.)

APPENDIX 1. PASCAL AND FERMAT

THE PROBLEM OF DICE

Pascal's first letter to Fermat is lost, but it must state the problem of the dice.

Fermat's reply points out that Pascal has made an error ("Pascal and Fermat on Probability," tr. by Vera Sanford in *A Sourcebook in Mathematics*, ed. David Eugene Smith (New York: McGraw Hill, 1929), 546–65. Dover reprint in 1969 available online at https://www.york.ac.uk/depts/maths/histstat/pascal.pdf):

> If I undertake to make a point with a single die in eight throws, and if we agree after the money is put at stake, that I shall not cast the first throw, it is necessary by my theory that I take $\frac{1}{6}$ of the total sum to [be] impartial because of the aforesaid first throw.
>
> And if we agree after that, that I shall not play the second throw, I should, for my share, take the sixth of the remainder that is $\frac{5}{36}$ of the total.
>
> If, after that, we agree that I shall not play the third throw, I should to recoup myself, take $\frac{1}{6}$ of the remainder, which is $\frac{25}{216}$ of the total.
>
> And if subsequently, we agree again that I shall not cast the fourth throw, I should take $\frac{1}{6}$ of the remainder or $\frac{125}{1296}$ of the total, and I agree with you that that is the value of the fourth throw supposing that one has already made the preceding plays.
>
> But you proposed in the last example in your letter (I quote your very terms) that if I undertake to find the six in eight throws and if I have thrown three times without getting it, and if my opponent proposes that I should not play the fourth time, and if he wishes me to be justly treated, it is proper that I have $\frac{125}{1296}$ of the entire sum of our wagers.
>
> This, however, is not true by my theory. For in this case, the three first throws having gained nothing for the player who holds the die, the total sum thus remaining at stake, he who holds the die and who agrees to not play his fourth throw should take $\frac{1}{6}$ as his reward. And if he has played four throws without finding the desired point and if they agree that he shall not play the fifth time, he will, nevertheless, have $\frac{1}{6}$ of the total for his share. Since the whole sum stays in play it not only follows from the theory, but it is indeed common sense that each throw should be of equal value.

It is clear that the central issue here is that of *expected value*. The combination of foregoing a round and receiving a proportion of the stake is fair if it leaves the expected value of the game unchanged.

Fermat sees clearly that the analysis is the same at any point in the game. Suppose that after the round in question, there will be $n+1$ rounds remaining; give the stakes at this point value 1. Then the value of taking the play is $\frac{1}{6}$ for winning now and $(\frac{5}{6})(1-\left(\frac{5}{6}\right)^n)$ for failing on this throw but possibly eventually winning. The value of taking $\frac{1}{6}$ of the stakes and proceeding with the rest of the game for the diminished stakes is $\frac{1}{6}$ for the cash in hand plus $1-\left(\frac{5}{6}\right)^n$, the probability of eventually, winning times $\frac{5}{6}$ of the the diminished stakes. Pascal immediately agrees with Fermat's analysis.

THE PROBLEM OF POINTS

There is another aspect of Pascal's discussion that is of interest. He starts with the example of a game where two players play for 3 points, where each has staked 32 pistoles ("Pascal and Fermat on Probability," tr. by Vera Sanford in *A Sourcebook in Mathematics*, ed. David Eugene Smith (New York: McGraw Hill, 1929), 546–65. Dover reprint in 1969 available online at https://www.york.ac.uk/depts/maths/histstat/pascal.pdf):

> Let us suppose that the first of them has two (points) and the other one. They now play one throw of which the chances are such that if the first wins, he will win the entire wager that is at stake, that is to say 64 pistoles. If the other wins, they will be two to two and in consequence, if they wish to separate, it follows that each will take back his wager that is to say 32 pistoles.
>
> Consider then, Monsieur, that if the first wins, 64 will belong to him. If he loses, 32 will belong to him. Then if they do not wish to play this point, and separate without doing it, the first should say "I am sure of 32 pistoles, for even a loss gives them to me. As for the 32 others, perhaps I will have them and perhaps you will have them, the risk is equal. Therefore let us divide the 32 pistoles in half, and give me the 32 of which I am certain besides." He will then have 48 pistoles and the other will have 16.

This is not just a calculation of expected value but also a justification of the *fairness* of using it, in terms that are hard for anyone to reject. What you have for sure is yours. For what is uncertain, equal probabilities match equal division. It is a definitive answer to Fra Pacioli's line of thought.

Pascal goes on to show how this reasoning can be further iterated:

> Now let us suppose that the first has *two* points and the other *none*, and that they are beginning to play for a point. The chances are such that if the first wins, he will win all of the wager, 64 pistoles. If the other wins, behold they have come back to the preceding case in which the first has *two* points and the other *one*.
>
> But we have already shown that in this case 48 pistoles will belong to the one who has *two* points. Therefore if they do not wish to play this point, he should say, "If I win, I shall gain all, that is 64. If I lose, 48 will legitimately belong to me. Therefore give me the 48 that are certain to be mine, even if I lose, and let us divide the other 16 in half because there is as much chance that you will gain them as that I will." Thus he will have 48 and 8, which is 56 pistoles.
>
> Let us now suppose that the first has but *one* point and the other *none*. You see, Monsieur, that if they begin a new throw, the chances are such that if the first wins, he will have *two* points to *none*, and dividing by the preceding case, 56 will belong to him. If he loses, they will [be] point for point, and 32 pistoles will belong to him. He should therefore say, "If you do not wish to play, give me the 32 pistoles of which I am certain, and let us divide the rest of the 56 in half. From 56 take 32, and 24 remains. Then divide 24 in half, you take 12 and I take 12 which with 32 will make 44.

This gives us a recursive procedure for fair division. Pascal then projects to games with larger numbers of points, and comes to a general solution of the problem.

APPENDIX 2. PHYSICS OF COIN TOSSING

Drawing balls from an urn, flipping coins, rolling dice, and shuffling cards are basic probability models. How are they connected to their parallels in the real world? Going further afield, these basic models are often used to calculate chances in much more complicated setups; Bernoulli considered the successive scores of two tennis players. Gilovitch, Tversky, and Valone[15] considered the successive hits and misses of basketball players. Shouldn't physics and psychology come into these analyses?

Each of the foregoing examples has its own literature. To give a flavor of this, we consider a single flip of a coin. Afterward, pointers to the analysis of other examples will be given.

Let's take a brief look at a simple version of the physics.[16] When the coin leaves the hand, it has an initial velocity upward v (feet/second) and a rate of spin ω (revolutions/second). If v and ω are known, Newton tells us how much time the coin will take before landing and thus heads or tails are determined. The phase space of a coin in this model is thus as shown in figure 1.4.

A single flip corresponds to a point in this plane. Consider the point in figure 1.4. The velocity is large (so the coin goes up rapidly), but the rate of spin is low. Thus the coin goes up like a pizza tossed in the air, hardly turning. Similarly, a point with v small and ω large may be turning like crazy but never goes high enough to turn over once. From these considerations, it follows that there is a region of initial conditions, close to the two axes, where the coin never turns.

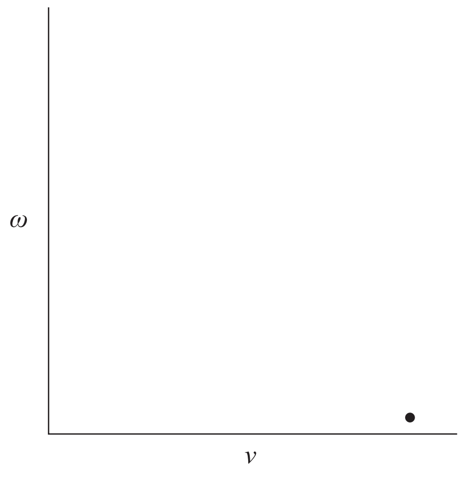

Figure 1.4. The $v\omega$-plane with a single flip

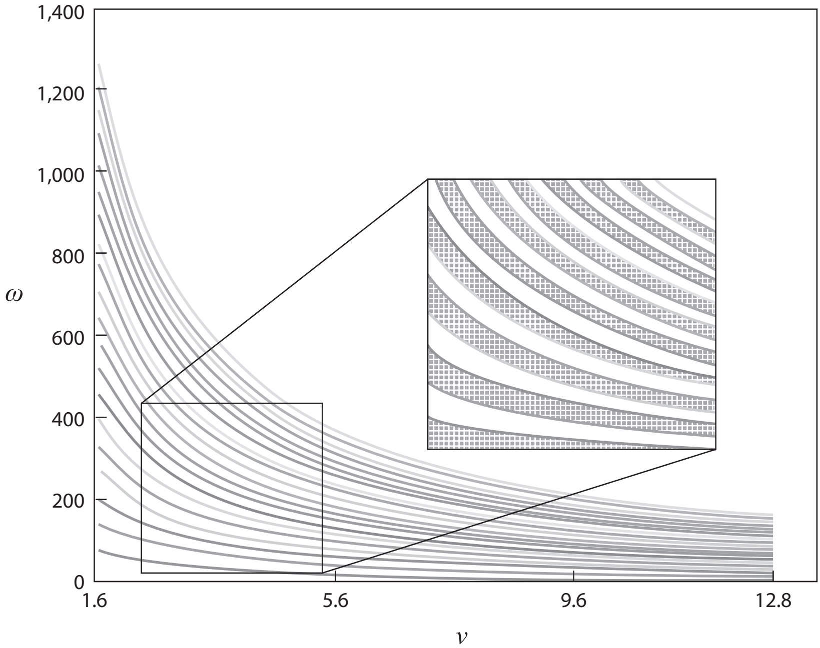

Figure 1.5. The hyperbolas separating heads from tails in part of phase space. Initial conditions leading to heads are hatched, tails are left white, and ω is measured in s^{-1}.

There is an adjoining region where the coin turns once, then a region for two turns, and so on. The full picture is shown in figure 1.5.

Inspection of the picture (and some easy mathematics) shows that regions far from 0 get closer together. So small changes in initial conditions make for the difference between heads or tails.

To go further, one must know the answer to the following question: When real people flip real coins, where are the points on the picture? We have carried out experiments and a normal flip takes about $\frac{1}{2}$ second and turns at about 40 revolutions/second. Look at figure 1.5. In the units of the picture, velocity is about $\frac{1}{5}$, very close to zero. The rate of spin, ω, is 40 units up, however, way off the picture. The math behind the picture says how close the regions are. This coupled with experimental work shows that coin tossing is fair to two decimal places but not to three.

The preceding analysis is in a simple model, which assumes that the coin flips about an axis through the coin. In fact, real coins are more complicated. They precess in amazing ways. A full analysis, with many details, caveats, and full references is in "Dynamical Bias in the Coin Toss,"[17] which concludes that vigorous tosses of ordinary coins are *slightly biased*. The chance of the coin landing the same way it started is about 0.51.

Where does all this analysis leave us? The standard model is a very good approximation. It would take about 250,000 flips to detect the difference between 0.50 and 0.51 (in the sense of giving second-digit accurately). We wish some of the other instances of the standard model were as solidly useful. Similar statements hold for Galileo's dice, but roulette or shuffling cards is another story![18]

If an honest analysis of a simple coin flip leads us into such complications, how much more would be required for an analysis of chances in games of skill or for the application of probability to medicine and law, as envisioned by Leibniz and Bernoulli? Bernoulli appreciated the point (Jacob Bernoulli, *The Art of Conjecturing*, tr. with an introduction and notes by Edith Dudley Sylla (Baltimore: Johns Hopkins University Press, 2006), 327):

> But what mortal, I ask, may determine, for example, the number of diseases, as if they were just as many cases, which may invade at any age the innumerable parts of the human body and which imply our death? And who can determine how much more easily one disease may kill than another—the plague compared to dropsy, dropsy compared to fever? Who, then, can form conjectures on the future state of life and death on this basis? Likewise who will count the innumerable cases of the changes to which the air is subject every day and on this basis conjecture its future constitution after a month, not to say after a year?
>
> Again, who has a sufficient perspective on the nature of the human mind or on the wonderful structure of the body so that they would dare to determine the cases in which this or that player may win or lose in games that depend in whole or in part on the shrewdness or the agility of the players? In these and similar situations, since they may depend on causes that are entirely hidden and that would forever mock our diligence by an

innumerable variety of combinations, it would clearly be mad to want to learn anything in this way.[19]

Bernoulli thought he had an answer to these problems in his law of large numbers. We return to this issue in chapter 4. There, and in subsequent chapters, we will assess the adequacy of the answer and discuss the possible alternatives.

APPENDIX 3. COINCIDENCES AND THE BIRTHDAY PROBLEM

Coincidences occur to all of us. Should we be surprised, or worried? The simple birthday problem (and its variations) has emerged as a useful tool to enable standards of surprise. While most people *are* surprised when there is a birthday match within a group of 23 people, the easy calculation in the introduction to this chapter shows that it is not surprising at all. Let us abstract and extend this calculation.

Consider what we will call the "watch" problem. Old-fashioned watches—watches with second hands—are coming back into current fashion. We believe that the second hands are "random"—completely out of sync and equally likely to show anything from 1 to 60—independently from watch to watch. Consider a group of N people, each having a watch with a second hand. What is the probability that two or more of these match—say, right at this second?

This is the birthday problem with 60 categories. The original birthday problem has 365 categories. Abstracting, consider C categories (so $C = 60$ for watches but $C = 365$ for birthdays.) There are N people, each independently and uniformly distributed in $\{1, 2, 3, \dots, C\}$. What is the chance that all these numbers are distinct? Of course, this depends on C and N; the chance is zero if $N = C + 1$.

Call this chance $P(C, N)$. By our earlier reasoning,

$$P(C,N) = 1 \cdot \left(1 - \frac{1}{C}\right) \cdot \left(1 - \frac{2}{C}\right) \cdot \cdots \cdot \left(1 - \frac{(N-1)}{C}\right).$$

This is a neat formula. You can use it with a pocket calculator to give an exact answer for any fixed C and N.

But this is not particularly useful for understanding. For later use, we can compute a simple approximation, which shows that when $N = 1.2\sqrt{C}$, the chance of success is close to $\frac{1}{2}$. For the watch problem, $1.2\sqrt{60} = 9.3$, so a match has at least even odds with 10 people. Intuitively, a match would seem a striking coincidence. (For the original birthday problem, $1.2\sqrt{365} = 22.9$.)

We state our approximation as a proposition.

Proposition: With N people and C possibilities, N and C large, the chance of no match is

$$P(C, N) \sim e^{-N(N-1)/2C}.$$

Proof: The argument uses simple properties of the logarithm: $\log(1-x) \sim -x$ when x is small. Then

$$\begin{aligned} P(C, N) &= \left(1 - \frac{1}{C}\right) \cdot \left(1 - \frac{2}{C}\right) \cdot \cdots \cdot \left(1 - \frac{(N-1)}{C}\right) \\ &= e^{\log(1-1/C) + \log(1-2/C) + \cdots + \log(1-(N-1)/C)} \\ &\sim e^{-1/C - 2/C \cdots -(N-1)/C} \\ &= e^{-N(N-1)/2C}. \end{aligned}$$

The approximations are accurate provided that N and C are large, with $N^{2/3}/C$ small.

Diaconis and Mosteller[20] use the birthday problem more generally in studying coincidences. They use these ideas to study multiple coincidences. For instance, how large should N be to have approximately even odds of a triple birthday match? (Answer: about 81.)

As a counterpoint to a philosophy that tries to make much of coincidences, we have provided a simple chance model for comparison. It seems useful and believable for studying things like a birthday match in a classroom. But one might consider instead a group of people in a very fancy restaurant. Since people are often taken out to dinner on their birthdays, it is quite likely that there may be several matches on a given night. The assumptions of our chance model don't hold, so the conclusions aren't relevant. The caution applies to all the simple chance models of this section. For more, see Diaconis and Holmes (2002).[21]