

Chapter 1

Discrete Probability Distributions

1.1 Simulation of Discrete Probabilities

Probability

In this chapter, we shall first consider chance experiments with a finite number of possible outcomes $\omega_1, \omega_2, \dots, \omega_n$. For example, we roll a die and the possible outcomes are 1, 2, 3, 4, 5, 6 corresponding to the side that turns up. We toss a coin with possible outcomes H (heads) and T (tails).

It is frequently useful to be able to refer to an outcome of an experiment. For example, we might want to write the mathematical expression which gives the sum of four rolls of a die. To do this, we could let $X_i, i = 1, 2, 3, 4$, represent the values of the outcomes of the four rolls, and then we could write the expression

$$X_1 + X_2 + X_3 + X_4$$

for the sum of the four rolls. The X_i 's are called *random variables*. A random variable is simply an expression whose value is the outcome of a particular experiment. Just as in the case of other types of variables in mathematics, random variables can take on different values.

Let X be the random variable which represents the roll of one die. We shall assign probabilities to the possible outcomes of this experiment. We do this by assigning to each outcome ω_j a nonnegative number $m(\omega_j)$ in such a way that

$$m(\omega_1) + m(\omega_2) + \dots + m(\omega_6) = 1 .$$

The function $m(\omega_j)$ is called the *distribution function* of the random variable X . For the case of the roll of the die we would assign equal probabilities or probabilities $1/6$ to each of the outcomes. With this assignment of probabilities, one could write

$$P(X \leq 4) = \frac{2}{3}$$

to mean that the probability is $2/3$ that a roll of a die will have a value which does not exceed 4.

Let Y be the random variable which represents the toss of a coin. In this case, there are two possible outcomes, which we can label as H and T. Unless we have reason to suspect that the coin comes up one way more often than the other way, it is natural to assign the probability of $1/2$ to each of the two outcomes.

In both of the above experiments, each outcome is assigned an equal probability. This would certainly not be the case in general. For example, if a drug is found to be effective 30 percent of the time it is used, we might assign a probability .3 that the drug is effective the next time it is used and .7 that it is not effective. This last example illustrates the intuitive *frequency concept of probability*. That is, if we have a probability p that an experiment will result in outcome A , then if we repeat this experiment a large number of times we should expect that the fraction of times that A will occur is about p . To check intuitive ideas like this, we shall find it helpful to look at some of these problems experimentally. We could, for example, toss a coin a large number of times and see if the fraction of times heads turns up is about $1/2$. We could also simulate this experiment on a computer.

Simulation

We want to be able to perform an experiment that corresponds to a given set of probabilities; for example, $m(\omega_1) = 1/2$, $m(\omega_2) = 1/3$, and $m(\omega_3) = 1/6$. In this case, one could mark three faces of a six-sided die with an ω_1 , two faces with an ω_2 , and one face with an ω_3 .

In the general case we assume that $m(\omega_1), m(\omega_2), \dots, m(\omega_n)$ are all rational numbers, with least common denominator n . If $n > 2$, we can imagine a long cylindrical die with a cross-section that is a regular n -gon. If $m(\omega_j) = n_j/n$, then we can label n_j of the long faces of the cylinder with an ω_j , and if one of the end faces comes up, we can just roll the die again. If $n = 2$, a coin could be used to perform the experiment.

We will be particularly interested in repeating a chance experiment a large number of times. Although the cylindrical die would be a convenient way to carry out a few repetitions, it would be difficult to carry out a large number of experiments. Since the modern computer can do a large number of operations in a very short time, it is natural to turn to the computer for this task.

Random Numbers

We must first find a computer analog of rolling a die. This is done on the computer by means of a *random number generator*. Depending upon the particular software package, the computer can be asked for a real number between 0 and 1, or an integer in a given set of consecutive integers. In the first case, the real numbers are chosen in such a way that the probability that the number lies in any particular subinterval of this unit interval is equal to the length of the subinterval. In the second case, each integer has the same probability of being chosen.

.203309	.762057	.151121	.623868
.932052	.415178	.716719	.967412
.069664	.670982	.352320	.049723
.750216	.784810	.089734	.966730
.946708	.380365	.027381	.900794

Table 1.1: Sample output of the program **RandomNumbers**.

Let X be a random variable with distribution function $m(\omega)$, where ω is in the set $\{\omega_1, \omega_2, \omega_3\}$, and $m(\omega_1) = 1/2$, $m(\omega_2) = 1/3$, and $m(\omega_3) = 1/6$. If our computer package can return a random integer in the set $\{1, 2, \dots, 6\}$, then we simply ask it to do so, and make 1, 2, and 3 correspond to ω_1 , 4 and 5 correspond to ω_2 , and 6 correspond to ω_3 . If our computer package returns a random real number r in the interval $(0, 1)$, then the expression

$$\lfloor 6r \rfloor + 1$$

will be a random integer between 1 and 6. (The notation $\lfloor x \rfloor$ means the greatest integer not exceeding x , and is read “floor of x .”)

The method by which random real numbers are generated on a computer is described in the historical discussion at the end of this section. The following example gives sample output of the program **RandomNumbers**.

Example 1.1 (Random Number Generation) The program **RandomNumbers** generates n random real numbers in the interval $[0, 1]$, where n is chosen by the user. When we ran the program with $n = 20$, we obtained the data shown in Table 1.1. \square

Example 1.2 (Coin Tossing) As we have noted, our intuition suggests that the probability of obtaining a head on a single toss of a coin is $1/2$. To have the computer toss a coin, we can ask it to pick a random real number in the interval $[0, 1]$ and test to see if this number is less than $1/2$. If so, we shall call the outcome *heads*; if not we call it *tails*. Another way to proceed would be to ask the computer to pick a random integer from the set $\{0, 1\}$. The program **CoinTosses** carries out the experiment of tossing a coin n times. Running this program, with $n = 20$, resulted in:

THTTTHTTTTHTTTTTHHTT.

Note that in 20 tosses, we obtained 5 heads and 15 tails. Let us toss a coin n times, where n is much larger than 20, and see if we obtain a proportion of heads closer to our intuitive guess of $1/2$. The program **CoinTosses** keeps track of the number of heads. When we ran this program with $n = 1000$, we obtained 494 heads. When we ran it with $n = 10000$, we obtained 5039 heads.

We notice that when we tossed the coin 10,000 times, the proportion of heads was close to the “true value” .5 for obtaining a head when a coin is tossed. A mathematical model for this experiment is called Bernoulli Trials (see Chapter 3). The *Law of Large Numbers*, which we shall study later (see Chapter 8), will show that in the Bernoulli Trials model, the proportion of heads should be near .5, consistent with our intuitive idea of the frequency interpretation of probability.

Of course, our program could be easily modified to simulate coins for which the probability of a head is p , where p is a real number between 0 and 1. \square

In the case of coin tossing, we already knew the probability of the event occurring on each experiment. The real power of simulation comes from the ability to estimate probabilities when they are not known ahead of time. This method has been used in the recent discoveries of strategies that make the casino game of blackjack favorable to the player. We illustrate this idea in a simple situation in which we can compute the true probability and see how effective the simulation is.

Example 1.3 (Dice Rolling) We consider a dice game that played an important role in the historical development of probability. The famous letters between Pascal and Fermat, which many believe started a serious study of probability, were instigated by a request for help from a French nobleman and gambler, Chevalier de Méré. It is said that de Méré had been betting that, in four rolls of a die, at least one six would turn up. He was winning consistently and, to get more people to play, he changed the game to bet that, in 24 rolls of two dice, a pair of sixes would turn up. It is claimed that de Méré lost with 24 and felt that 25 rolls were necessary to make the game favorable. It was *un grand scandale* that mathematics was wrong.

We shall try to see if de Méré is correct by simulating his various bets. The program **DeMere1** simulates a large number of experiments, seeing, in each one, if a six turns up in four rolls of a die. When we ran this program for 1000 plays, a six came up in the first four rolls 48.6 percent of the time. When we ran it for 10,000 plays this happened 51.98 percent of the time.

We note that the result of the second run suggests that de Méré was correct in believing that his bet with one die was favorable; however, if we had based our conclusion on the first run, we would have decided that he was wrong. *Accurate results by simulation require a large number of experiments.* \square

The program **DeMere2** simulates de Méré’s second bet that a pair of sixes will occur in n rolls of a pair of dice. The previous simulation shows that it is important to know how many trials we should simulate in order to expect a certain degree of accuracy in our approximation. We shall see later that in these types of experiments, a rough rule of thumb is that, at least 95% of the time, the error does not exceed the reciprocal of the square root of the number of trials. Fortunately, for this dice game, it will be easy to compute the exact probabilities. We shall show in the next section that for the first bet the probability that de Méré wins is $1 - (5/6)^4 = .518$.

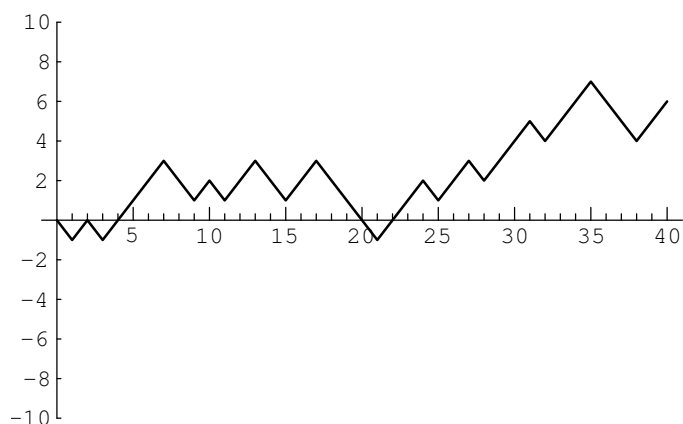


Figure 1.1: Peter's winnings in 40 plays of heads or tails.

One can understand this calculation as follows: The probability that no 6 turns up on the first toss is $(5/6)$. The probability that no 6 turns up on either of the first two tosses is $(5/6)^2$. Reasoning in the same way, the probability that no 6 turns up on any of the first four tosses is $(5/6)^4$. Thus, the probability of at least one 6 in the first four tosses is $1 - (5/6)^4$. Similarly, for the second bet, with 24 rolls, the probability that de Méré wins is $1 - (35/36)^{24} = .491$, and for 25 rolls it is $1 - (35/36)^{25} = .506$.

Using the rule of thumb mentioned above, it would require 27,000 rolls to have a reasonable chance to determine these probabilities with sufficient accuracy to assert that they lie on opposite sides of .5. It is interesting to ponder whether a gambler can detect such probabilities with the required accuracy from gambling experience. Some writers on the history of probability suggest that de Méré was, in fact, just interested in these problems as intriguing probability problems.

Example 1.4 (Heads or Tails) For our next example, we consider a problem where the exact answer is difficult to obtain but for which simulation easily gives the qualitative results. Peter and Paul play a game called *heads or tails*. In this game, a fair coin is tossed a sequence of times—we choose 40. Each time a head comes up Peter wins 1 penny from Paul, and each time a tail comes up Peter loses 1 penny to Paul. For example, if the results of the 40 tosses are

THTHHHTTHTTHHTTHTTTTTHHHHTHHHTHHHTHHHTTTHH.

Peter's winnings may be graphed as in Figure 1.1.

Peter has won 6 pennies in this particular game. It is natural to ask for the probability that he will win j pennies; here j could be any even number from -40 to 40. It is reasonable to guess that the value of j with the highest probability is $j = 0$, since this occurs when the number of heads equals the number of tails. Similarly, we would guess that the values of j with the lowest probabilities are $j = \pm 40$.

A second interesting question about this game is the following: How many times in the 40 tosses will Peter be in the lead? Looking at the graph of his winnings (Figure 1.1), we see that Peter is in the lead when his winnings are positive, but we have to make some convention when his winnings are 0 if we want all tosses to contribute to the number of times in the lead. We adopt the convention that, when Peter's winnings are 0, he is in the lead if he was ahead at the previous toss and not if he was behind at the previous toss. With this convention, Peter is in the lead 34 times in our example. Again, our intuition might suggest that the most likely number of times to be in the lead is $1/2$ of 40, or 20, and the least likely numbers are the extreme cases of 40 or 0.

It is easy to settle this by simulating the game a large number of times and keeping track of the number of times that Peter's final winnings are j , and the number of times that Peter ends up being in the lead by k . The proportions over all games then give estimates for the corresponding probabilities. The program **HTSimulation** carries out this simulation. Note that when there are an even number of tosses in the game, it is possible to be in the lead only an even number of times. We have simulated this game 10,000 times. The results are shown in Figures 1.2 and 1.3. These graphs, which we call spike graphs, were generated using the program **Spikegraph**. The vertical line, or spike, at position x on the horizontal axis, has a height equal to the proportion of outcomes which equal x . Our intuition about Peter's final winnings was quite correct, but our intuition about the number of times Peter was in the lead was completely wrong. The simulation suggests that the least likely number of times in the lead is 20 and the most likely is 0 or 40. This is indeed correct, and the explanation for it is suggested by playing the game of heads or tails with a large number of tosses and looking at a graph of Peter's winnings. In Figure 1.4 we show the results of a simulation of the game, for 1000 tosses and in Figure 1.5 for 10,000 tosses.

In the second example Peter was ahead most of the time. It is a remarkable fact, however, that, if play is continued long enough, Peter's winnings will continue to come back to 0, but there will be very long times between the times that this happens. These and related results will be discussed in Chapter 12. \square

In all of our examples so far, we have simulated equiprobable outcomes. We illustrate next an example where the outcomes are not equiprobable.

Example 1.5 (Horse Races) Four horses (Acorn, Balky, Chestnut, and Dolby) have raced many times. It is estimated that Acorn wins 30 percent of the time, Balky 40 percent of the time, Chestnut 20 percent of the time, and Dolby 10 percent of the time.

We can have our computer carry out one race as follows: Choose a random number x . If $x < .3$ then we say that Acorn won. If $.3 \leq x < .7$ then Balky wins. If $.7 \leq x < .9$ then Chestnut wins. Finally, if $.9 \leq x$ then Dolby wins.

The program **HorseRace** uses this method to simulate the outcomes of n races. Running this program for $n = 10$ we found that Acorn won 40 percent of the time, Balky 20 percent of the time, Chestnut 10 percent of the time, and Dolby 30 percent

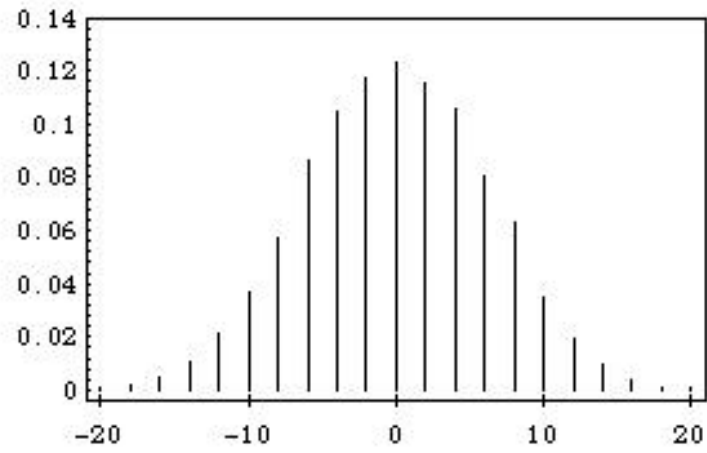


Figure 1.2: Distribution of winnings.

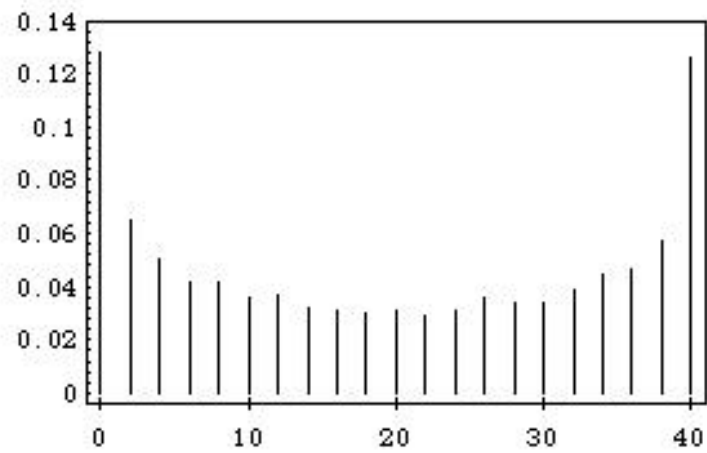


Figure 1.3: Distribution of number of times in the lead.

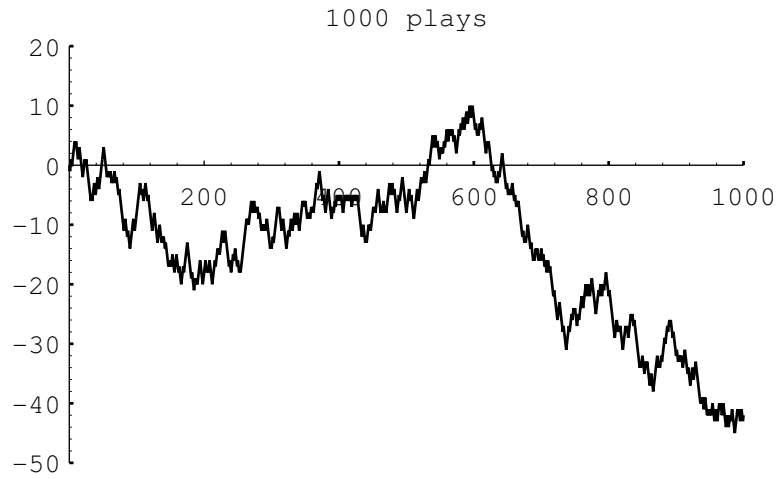


Figure 1.4: Peter's winnings in 1000 plays of heads or tails.

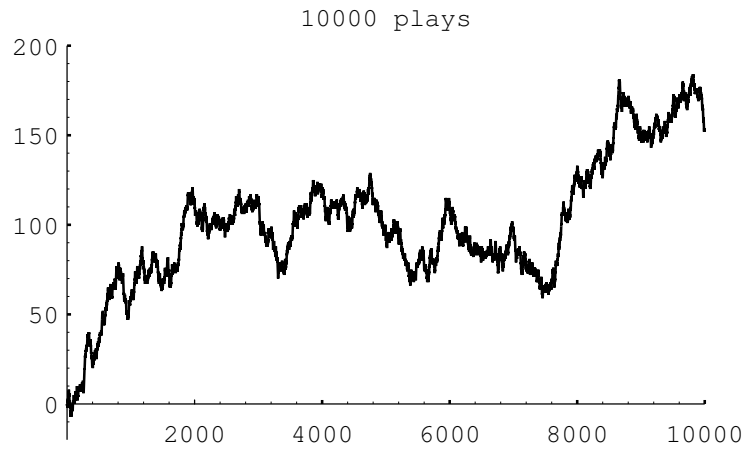


Figure 1.5: Peter's winnings in 10,000 plays of heads or tails.