6

Probability

We use the word "chance" frequently in our everyday conversation. For example, we meet somebody "by chance," we "chance upon" the solution to a problem, one team has a "better chance" of reaching the Super Bowl than another, the Weather Channel announces a "30% chance" of precipitation.

There are two ideas here. In every case there is the feature of unpredictability—a chance occurrence is one for which we cannot be certain of the outcome. The other feature is that sometimes chance is *quantitative*—either it can be measured exactly (the weather forecast) or else we can at least say one "chance" is greater than another (the football teams). We shall refer to these two aspects of chance as *randomness* and *probability*, respectively.

For example, say you flip a coin. There is no way to tell whether it will fall heads or tails, so we would say this is a *random* occurrence. If the coin is made uniformly, so that it is equally likely to show heads or a tails, we normally call it a *fair* coin, and we say the probability of a heads is $\frac{1}{2}$, and so is the probability of a tails. (Sometimes we say "50%" instead of "$\frac{1}{2}$.")

Suppose the coin in our example was not uniform; for example, say it was made as a sandwich of disks of metal, like a quarter, but the different disks were of different densities. It might be that, if we flipped it enough times, heads would come up 70% of the time and tails only 30%. Then we would say the probability of a heads is $\frac{7}{10}$. However, this still qualifies as random because no one knows beforehand whether a particular flip will be one of the (more common) ones that results in a heads or one that gives a tails. The probabilities do not have to be equal in order for an event to be random.

To discuss randomness formally, we shall use the idea of an "event" just as we did in Section 5.1. The ideas and notation of set theory—and in particular the number of elements—will be very useful.

W.D. Wallis, *A Beginner's Guide to Discrete Mathematics*,
DOI 10.1007/978-0-8176-8286-6_6,

6.1 Probability Measures

Probability Distributions

Consider a random event: for example, the fall of a heads when a fair coin is flipped. We do not know whether a heads will show on any particular flip, but we expect to see a heads half the time when a large number of attempts are made. We use the word "probability" for this numerical measure of the likelihood of an event. The probabilities 1 and 0 indicate absolute certainty and impossibility, respectively.

We shall write $P(E)$ for the probability that the event E occurs. For example, if H means "a heads shows" and T means "a tails shows" when a fair coin is flipped, then

$$P(H) = P(T) = \frac{1}{2}.$$

A list of the probabilities of all outcomes of an experiment is called the *probability distribution* of the experiment.

One important case is where each of the possible outcomes is equally likely, as in the case of the fair coin. In this case we say the experiment has *equally likely outcomes*, and set their probabilities equal. In this case the probability distribution is called *uniform*, and we also refer to the experiment as a *uniform experiment*. If a uniform experiment has a sample space with n elements, then we shall assign probability $\frac{1}{n}$ to each outcome.

If a uniform experiment has sample space S, then an event E will occur if the outcome is one of the $|E|$ outcomes in E. We expect that E will occur in fraction $|E|/|S|$ of cases, if the experiment is repeated. So we define

$$P(E) = \frac{|E|}{|S|}.$$

Sample Problem 6.1. *A fair die is rolled. S_i is the event that the number i is rolled for $i = 1, 2, 3, 4, 5, 6$. E means the roll of an odd number, and F the roll of a number less than* 3. *Find the probabilities of these events.*

Solution. Since the die is fair, the outcomes are equally likely:

$$P(S_1) = P(S_2) = P(S_3) = P(S_4) = P(S_5) = P(S_6) = \frac{1}{6}.$$

E and F contain three and two outcomes, respectively, so $P(E) = \frac{3}{6} = \frac{1}{2}$ and $P(F) = \frac{2}{6} = \frac{1}{3}$.

Practice Exercise. A single die is rolled. What are the probabilities of the following events?

E: A 4 or a 5 is rolled.

F: An even number is rolled.

G: An odd number greater than 2 is rolled.

If E is any event in a uniform experiment, then $P(E)$ will be a proper fraction. If E and F are two disjoint events in such an experiment, then the event $E \cup F$ ("either E or F occurs") has $|E| + |F|$ elements, so

$$P(E \cup F) = \frac{|E| + |F|}{|S|} = \frac{|E|}{|S|} + \frac{|F|}{|S|} = P(E) + P(F).$$

We shall now apply these ideas to the general situation to make a general definition of probability.

Consider an experiment with sample space $S = \{s_1, s_2, \ldots, s_m\}$. A *probability distribution* for the experiment is a function P with the following properties:

1. For each s_i, $1 \le i \le n$, $P(s_i)$ is a real number and $0 \le P(s_i) \le 1$;
2. $P(s_1) + P(s_2) + \cdots + P(s_m) = 1$.

Sample Problem 6.2. *A black die and a white die are thrown simultaneously. What is the probability that the sum of the numbers shown is 8 given that the dice are fair?*

Solution. Write (x, y) to mean that x shows on the black die and y shows on the white die. Then there are 36 possible outcomes, namely $(1, 1), (1, 2), \ldots, (1, 6), (2, 1), \ldots, (6, 6)$; they are equally likely. Five of them—$(6, 2)$, $(5, 3)$, $(4, 4)$, $(3, 5)$, and $(2, 6)$—have a sum of 8. So

$$P(E) = \frac{|E|}{|S|} = \frac{5}{36}.$$

Practice Exercise. A quarter and a nickel are flipped simultaneously. What is the probability that exactly one heads shows, assuming that both coins are fair?

Sample Problem 6.3. *A deck of cards is shuffled and one card is dealt. What is the probability that it is a spade?*

Solution. There are 52 cards in a deck of which 13 are spades. So

$$P(E) = \frac{|E|}{|S|} = \frac{13}{52} = \frac{1}{4}.$$

Practice Exercise. A deck is shuffled and one card dealt face up. What is the probability that it is a picture card (king, queen, or jack)?

$E_2 = \{11\}$	$\lvert E_2\rvert = 1$	$P(E_2) = \frac{1}{36}$
$E_3 = \{12, 21\}$	$\lvert E_3\rvert = 2$	$P(E_3) = \frac{2}{36} = \frac{1}{18}$
$E_4 = \{14, 22, 31\}$	$\lvert E_4\rvert = 3$	$P(E_4) = \frac{3}{36} = \frac{1}{12}$
$E_5 = \{14, 23, 32, 41\}$	$\lvert E_5\rvert = 4$	$P(E_5) = \frac{4}{36} = \frac{1}{9}$
$E_6 = \{15, 24, 33, 42, 51\}$	$\lvert E_6\rvert = 5$	$P(E_6) = \frac{5}{36}$
$E_7 = \{16, 25, 34, 43, 52, 61\}$	$\lvert E_7\rvert = 6$	$P(E_7) = \frac{6}{36} = \frac{1}{6}$
$E_8 = \{26, 35, 44, 53, 62\}$	$\lvert E_8\rvert = 5$	$P(E_8) = \frac{5}{36}$
$E_9 = \{36, 45, 54, 63\}$	$\lvert E_9\rvert = 4$	$P(E_9) = \frac{4}{36} = \frac{1}{9}$
$E_{10} = \{46, 55, 64\}$	$\lvert E_{10}\rvert = 3$	$P(E_{10}) = \frac{3}{36} = \frac{1}{12}$
$E_{11} = \{56, 65\}$	$\lvert E_{11}\rvert = 2$	$P(E_{11}) = \frac{2}{36} = \frac{1}{18}$
$E_{12} = \{66\}$	$\lvert E_{12}\rvert = 1$	$P(E_{12}) = \frac{1}{36}$

Table 6.1. Probabilities when rolling two fair dice

Nonuniform Experiments

Not all experiments are uniform. But in many cases, given an experiment A, we can find a uniform experiment B such that the outcomes of A are events (not necessarily simple) of B.

For example, consider an experiment in which two fair dice are rolled, and the total of the points on them is recorded. This experiment has 11 possible outcomes $s_2, s_3, \ldots, s_{12}$; s_i means the total is i. To calculate the probability of s_i, we shall look at a slightly different experiment. In this one, two fair dice are rolled and the result is recorded as an ordered pair of digits—(1, 3) means "1 on die 1, 3 on die 2." (For brevity, one commonly writes "13" instead of "(1, 3)".) There are $6^2 = 36$ outcomes, and each has probability $\frac{1}{36}$. In this experiment we write E_i for the event that the two numbers showing add to i. Then

$$P(E) = \frac{|E_i|}{36},$$

and we can calculate all the probabilities; see Table 6.1. But in each case the probability of the outcome s_i in the first experiment is obviously equal to the probability of E_i in the second experiment. So $P(s_2) = \frac{1}{36}$, $P(s_3) = \frac{2}{36}$, and so on.

Sample Problem 6.4. *Three coins are flipped and the number of heads is recorded. What are the possible outcomes and what is the probability distribution?*

Solution. The outcomes are the four numbers 0, 1, 2, and 3. To calculate probabilities, suppose three coins were flipped and all results recorded. E_i is the event "there are i heads showing." Then $|S| = 8$ and

$$E_0 = \{TTT\}, \qquad P(E_0) = \frac{1}{8},$$
$$E_1 = \{HTT, THT, TTH\}, \qquad P(E_1) = \frac{3}{8},$$
$$E_2 = \{HHT, HTH, THH\}, \qquad P(E_2) = \frac{3}{8},$$
$$E_3 = \{HHH\}, \qquad P(E_3) = \frac{1}{8}.$$

Practice Exercise. A game is played using two dice each of which has the numbers 1, 2, 3 on its faces (so each number appears twice per die). An experiment consists of rolling the dice and adding the numbers showing. What is the sample space? What is the probability distribution?

Sample Problem 6.5. *There are five marbles—two blue, three red—in a box. One is selected at random. What is the probability that it is blue?*

Solution. If the marbles were marked v, w, x, y, z, where v and w are blue and the others are red, then the event "blue is chosen" is $\{v, w\}$ and its probability is

$$P(E) = \frac{|E|}{|S|} = \frac{2}{5}.$$

Practice Exercise. There are two red, three white, and four blue marbles in a box. One is drawn at random. What is the probability that it is *not* blue?

Nonuniform Probabilities

Even when the outcomes are not equally likely, we calculate the probability of an event from the formula

$$P(E) = \sum_{s \in E} P(s).$$

Several important facts can be deduced from this; in particular, if E is any event, $0 \le P(E) \le 1$; $P(E) = 0$ if and only if E is impossible, and $P(E) = 1$ if and only if E is certain. If E and F cannot both happen—they are *mutually exclusive events*, as defined in Section 5.1—then

$$P(E \cup F) = P(E) + P(F),$$

and for general E and F,

$$P(E \cup F) = P(E) + P(F) - P(E \cap F).$$

The probability of the complement $\overline{E}$ of E is

$$P(\overline{E}) = 1 - P(E).$$

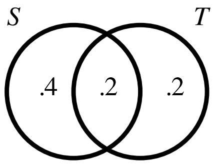
Sample Problem 6.6. *E and F are events in a sample space with $P(E) = 0.6$, $P(F) = 0.4$, and $P(E \cap F) = 0.2$. What is $P(E \cup F)$? What is $P(\overline{E})$?*

Solution.

$$\begin{aligned} P(E \cup F) &= P(E) + P(F) - P(E \cap F) \\ &= 0.6 + 0.4 + 0.2 \\ &= 0.8, \\ P(\overline{E}) &= 1 - P(E) \\ &= 1 - 0.6 \\ &= 0.4. \end{aligned}$$

Practice Exercise. E and F are events in a sample space with $P(E) = 0.7$, $P(F) = 0.3$, and $P(E \cap F) = 0.1$. What are $P(E \cup F)$, $P(\overline{E})$?

These probabilities can be represented in a Venn diagram—in each area of the diagram write the probability of the corresponding event. Then the answer can be obtained by addition. The Venn diagram for the preceding Sample Problem is



Sample Problem 6.7. *A discrete mathematics class is restricted to business and math majors, all freshmen or sophomores. The percentage makeup of the class is*

Freshman business majors	30%,
Freshman math majors	35%,
Sophomore business majors	21%,
Sophomore math majors	14%.

A name is chosen at random from the class list. What are the probabilities of the following events?

(i) *The student is a freshman.*

(ii) *The student is not a freshman.*

(iii) *The student is either a freshman or a math major.*

Solution. We write F for the event that the student is a freshman, S for sophomore, B for business major, and M for math major. Then the data mean

$$P(F \cap B) = 0.3, \qquad P(F \cap M) = 0.35,$$
$$P(S \cap B) = 0.21, \qquad P(S \cap M) = 0.14.$$

(i) Since $M = \overline{B}$ in this problem (all students are math or business majors and none can be both or else the total would add to more than 100%), the usual equation

$$F = (F \cap B) \cup (F \cap \overline{B}),$$

becomes

$$F = (F \cap B) \cup (F \cap M),$$

and since this is a disjoint union,

$$\begin{aligned} P(F) &= P(F \cap B) + P(F \cap M) \\ &= 0.3 + 0.35 \\ &= 0.65. \end{aligned}$$

(ii) $P(\overline{F}) = 1 - P(F) = 1 - 0.65 = 0.35.$

(iii) $F \cup M = F \cup (S \cap M)$, and this is a disjoint union, so

$$\begin{aligned} P(F \cup M) &= P(F) + P(S \cap M) \\ &= 0.65 + 0.14 \\ &= 0.79. \end{aligned}$$

The data can also be represented as

	B	M
F	0.30	0.35
S	0.21	0.14

The questions can then be answered by adding the probability in the appropriate cells of the diagram. For example,

$$P(F \cup M) = 0.30 + 0.15 + 0.14 = 0.79.$$

Practice Exercise. An entomologist has found that the butterflies in a certain area can be classified as follows:

striped males	19%,
striped females	22%,
unstriped males	28%,
unstriped females	31%.

Represent the data in a diagram. Assuming the butterflies appear at random, what is the probability that the next butterfly to be sighted will be (i) striped, (ii) female, (iii) either striped or female?

Exercises 6.1

1. A fair coin is flipped four times. Find the probabilities that

(i) at least three heads appear,

(ii) an even number of heads appear,

(iii) the first result is a heads.

In Exercises 2 to 7 two fair dice are rolled. What is the probability of the indicated event?

2. The total is 9.

3. The total is 6.

4. The total is even.

5. The total is odd.

6. Both scores are even.

7. One odd and one even number are shown.

8. A box contains 12 cards, one for each month of the year. A card is drawn at random.

(i) What is the probability that the selected card is March?

(ii) What is the probability that the selected card is from a month with an r in its name?

9. A box contains six red and four white balls. One ball is drawn at random. What is the probability that it is white?

10. A box contains seven red and five white balls. One ball is drawn at random. What is the probability that it is white?

11. In the game of roulette, a wheel is divided into 38 equal parts, labeled with the numbers from 1 to 36, 0 and 00. The spin of the wheel causes the ball to be randomly placed in one of the parts; the chances of the ball landing on any part are equal. Half of the parts numbered from 1 to 36 are red and half are black; the 0 and 00 are green. If the chances of the ball landing in any one part are equal, what are the probabilities of the following events?

(i) The ball lands on a red number.

(ii) The ball lands on a black number.

(iii) The ball lands on a green number.

(iv) The ball lands on 17.

(v) The ball lands on a number from 25 to 36 inclusive.

12. In a lottery there are 90 losing tickets and 10 winning tickets. You draw one ticket at random. What is the probability that it is a winner?

13. The moose population in a Canadian park are 45% plain brown, 35% mottled, and 20% spotted. One moose is captured at random for the Bronx Zoo. What are the probabilities that

(i) it is spotted?

(ii) it is not spotted?

(iii) it is not mottled?

14. A random number from 0 to 99 is chosen by a computer. What are the probabilities of the following events?

(i) The number is even.

(ii) The number ends in 5.

(iii) The number is divisible by 11.

15. Box A contains tickets numbered 1, 2, 3, 4. Box B contains tickets numbered 2, 3, 4, 5. One ticket is selected from each box.

(i) List all the elements of the sample space of this experiment.

(ii) Find the probability that the tickets have the same number.

(iii) Find the probability that the sum of the two selected numbers is even.

16. A box contains one red, two blue, and two white marbles. One marble is selected at random. What are the probabilities that it is

(i) red?

(ii) blue?

(iii) white?

17. A box contains four red, two blue, and three white marbles. One is selected at random. What is the probability that it is

(i) blue?

(ii) not blue?

18. There are 20 students in a class. There are 12 men (eight physics and four chemistry majors) and eight women (five physics and three chemistry majors). One student's name is selected at random. What is the probability that

(i) the student is a chemistry major?

(ii) the student is male?

19. A sample space contains two events, E and F, and

$$P(E) = 0.70, \qquad P(F) = 0.25, \qquad P(E \cap F) = 0.15.$$

Determine

$$P(\overline{E}),\ P(E \cup F),\ P(\overline{E \cup F}).$$

20. A sample space contains two events, E and F, and

$$P(E) = 0.5, \qquad P(F) = 0.3, \qquad P(E \cap F) = 0.2.$$

Determine

$$P(\overline{F}),\ P(E \cup F),\ P(\overline{E} \cap F).$$

21. The events E, F, and G satisfy

$$P(E) = 0.6, \qquad P(F) = 0.6, \qquad P(G) = 0.8,$$
$$P(E \cup F) = 0.8, \qquad P(E \cap G) = 0.5, \qquad P(F \cap G) = 0.5.$$

Determine

$$P(E \cap F),\ P(E \cup G),\ P(F \cup G).$$

22. The events E, F, and G satisfy

$$P(E) = 0.6, \qquad P(F) = 0.5, \qquad P(G) = 0.5,$$
$$P(E \cup F) = 0.8, \qquad P(E \cup G) = 0.8, \qquad P(F \cap G) = 0.2.$$

Determine

$$P(E \cap F),\ P(E \cap G),\ P(F \cup G).$$

23. An examination has two questions. Of 100 students, 75 do Question 1 correctly and 72 do Question 2 correctly. Sixty-four do both questions correctly.

(i) Represent the data in a Venn diagram.

(ii) A student's answer book is chosen at random. What is the probability that

(a) Question 1 contains an error?

(b) exactly one question contains an error?

(c) at least one question contains an error?

24. Of 1000 researchers at Microsoft, 375 have a degree in mathematics and 450 have a degree in computer science. Of the researchers, 150 have degrees in both fields. One researcher's name is selected at random.

(i) What is the probability that the researcher has a degree in mathematics, but not in computer science?

(ii) What is the probability that the researcher has no degree in either mathematics or computer science?

25. In a market survey, 50% of those polled said that they usually buy medications at a pharmacy and the others that they buy them at a supermarket. 80% buy meat at the supermarket, 20% at the butcher's. 40% buy both meat and medications at the supermarket. One shopper is selected at random from the survey group. What is the probability that he

(i) buys meat at the supermarket but buys medications at the pharmacy?

(ii) does not buy either meat or medications at the supermarket?

26. During the 1992 presidential elections, 200 voters were surveyed in the St. Louis area. Voters were classified according to whether they lived in Missouri (MO), Illinois (IL), or other states (OS), and whether they intended to vote for Bush (REP), Clinton (DEM), or Perot (IND). The results were as follows:

	REP	DEM	IND
MO	44	33	21
IL	20	22	14
OS	14	20	12

One of the voter's responses is selected at random. What are the probabilities of the following events?

(i) The voter is an Illinois democrat.

(ii) The voter is from Missouri.

(iii) The voter intends to vote for Perot.

(iv) The voter is from outside Illinois.

(v) The voter is either a Democrat or is from Missouri.

27. Two hundred new automobile buyers were surveyed. Their purchases were as follows:

	Sedan	Pickup	Van
Ford	50	10	10
G M	35	10	15
Other	40	10	20

One survey form is chosen at random. What is the probability that the vehicle was

(i) a sedan?

(ii) a Ford?

(iii) a General Motors van or pickup?

28. Five hundred researchers at the National Security Agency were surveyed. It was found that 300 have at least one degree in computer science, 173 have a Bachelor's degree in computer science, 123 have a Master's degree in computer science, and 99 have a Ph.D. in computer science. Additionally, it was found that 63 have both a Bachelor's and Master's degree in computer science, 22 have both a Master's and Ph.D. degree in computer science, and 19 have both Bachelor's and Ph.D. degrees in computer science. One employee is selected at random. What are the probabilities of the following events?

(i) She has all three degrees in computer science.

(ii) She has the Bachelor's and Ph.D. degrees, but not the Master's.

29. Last year, there were 200 students enrolled in Calculus II. Of these students. 140 also enrolled in Linear Algebra. Their majors were distributed as follows:

Enrolment	Major		
	Math	Science	Business
Calculus only	20	35	5
Calculus and Linear Algebra	90	45	5

Suppose the same trend is followed this year. When the next student comes to enroll in Calculus II, what is the probability that this student

(i) is a science major?

(ii) is not enrolling in Linear Algebra this year?

(iii) is not a mathematics major, but will take Linear Algebra?

30. One hundred fifty Business students tried to sign up at registration for required Marketing, Accounting, and Finance courses. Seventy registered for Marketing, 60 for Accounting, and 50 for Finance. Twenty-five registered for both Marketing and Accounting, 15 for Finance and Marketing, 15 for Finance and Accounting, and 5 signed up for all three courses. One student's registration form is chosen at random. What is the probability that the student

(i) was unable to sign up for any of the courses?

(ii) signed up for exactly one of the courses?

(iii) signed up for Marketing and Accounting, but not for Finance?

31. Jerry's Gas sells three grades of gasoline at both their full and self-service pumps. One day they kept track of their first 100 customers and their purchases in the following table:

	Regular Leaded	Regular Unleaded	Premium Unleaded
Full service	15	15	20
Self service	20	25	5

If this trend continues, what is the probability that the next customer who comes in

(i) goes to the self-service pumps?

(ii) buys regular unleaded at the self-service pumps?

(iii) buys premium unleaded at the full-service pumps?

6.2 Repeated Experiments

Stochastic Processes

Sometimes an experiment can be viewed as a sequence of smaller experiments. The outcome consists of a sequence: "outcome of subexperiment 1" followed by "outcome of subexperiment 2," and so on. An experiment of this kind is called a *stochastic process*.

For example, suppose a jar contains three red and two blue marbles. An experiment consists of drawing a marble from the jar, noting its color, drawing another marble, and noting the second marble's color. (The first marble is not replaced.) The possible outcomes are the four ordered pairs of colors: RR, RB, BR, and BB.

To analyze the experiment, we first observe that the first marble drawn is either red (three-fifths of the cases) or blue (two-fifths). If the first marble is red, then the remaining marbles are two red and two blue, so in half of these cases ($\frac{1}{2} \times \frac{3}{5} = \frac{3}{10}$ of the original cases) the second marble drawn is red and in the other half ($\frac{3}{10}$ of the original) it is blue. So, in the obvious notation,

$$P(RR) = \frac{3}{10},$$
$$P(RB) = \frac{3}{10}.$$

In the same way, if the first marble is blue, then the second marble drawn is red in $\frac{3}{4}$ and blue in $\frac{1}{4}$ of the cases. So

$$P(BR) = \frac{2}{5} \times \frac{3}{4} = \frac{3}{10},$$
$$P(BB) = \frac{2}{5} \times \frac{1}{4} = \frac{1}{10}.$$

So the four outcomes are RR, RB, BR, and BB, and

$$P(RR) = P(RB) = P(BR) = \frac{3}{10}, \qquad P(BB) = \frac{1}{10}.$$

Sample Problem 6.8. *There are two bags A and B. Bag A contains two red balls and one green ball; bag B contains two red and three green balls. First a bag is chosen at random, then a ball is chosen at random from it. What are the possible outcomes, and what are their probabilities? What is the probability that a red ball is selected?*

Solution. The outcomes are AR, AG, BR, and BG where A and B are the bags and R and G denote the color of the ball selected. The possible outcomes of the first subexperiment—the selection of bag—are A and B, with probabilities each $\frac{1}{2}$. If A is selected, the probability of R is $\frac{2}{3}$ and the probability of G is $\frac{1}{3}$. If B is selected, R has probability $\frac{2}{5}$ and G has probability $\frac{3}{5}$. So

$$\begin{aligned} P(AR) &= \frac{1}{2} \times \frac{2}{3} = \frac{1}{3}, \\ P(AG) &= \frac{1}{2} \times \frac{1}{3} = \frac{1}{6}, \\ P(BR) &= \frac{1}{2} \times \frac{2}{5} = \frac{1}{5}, \\ P(BG) &= \frac{1}{2} \times \frac{3}{5} = \frac{3}{10}. \end{aligned}$$

So the probability of red is

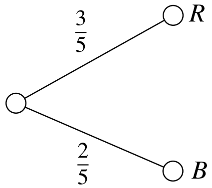
$$P(AR) + P(BR) = \frac{1}{3} + \frac{1}{5} = \frac{8}{15}.$$

Notice that there are four red and four green balls. If the two bags were emptied and one ball chosen, the probability of red would be $\frac{1}{2}$. The process of dividing the balls into two bags, whose contents are not identical, changes the probabilities.

Practice Exercise. A die is chosen from a pair of dice and rolled. One die (die A) is standard; the other (die B) has three faces marked 1 and three marked 6. What are the outcomes? Assuming that the die is chosen at random, what is the probability of a 6? Of a 3?

Tree Diagrams

The use of tree diagrams to represent experiments (see Section 5.1) is particularly useful for stochastic processes. Let us go back to the example of three red and two blue marbles. The first stage of the experiment can be represented as a branching into two parts, labeled “red” and “blue,” and the branches can be also be labeled with the probabilities $\frac{3}{5}$ (on red) and $\frac{2}{5}$ (on blue):



In each case the second stage can be represented similarly. The tree diagram for the whole experiment is shown in Figure 6.1.

Finally, the outcomes are sequences of the results of the subexperiments. Each endpoint on the right of the tree represents an outcome, and the sequence can be read

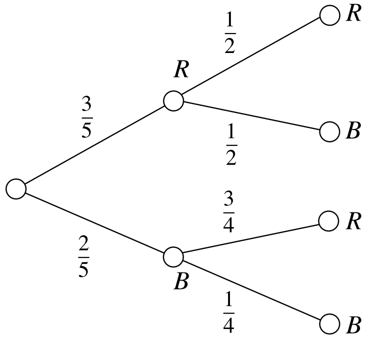


Fig. 6.1. Tree diagram for the whole experiment

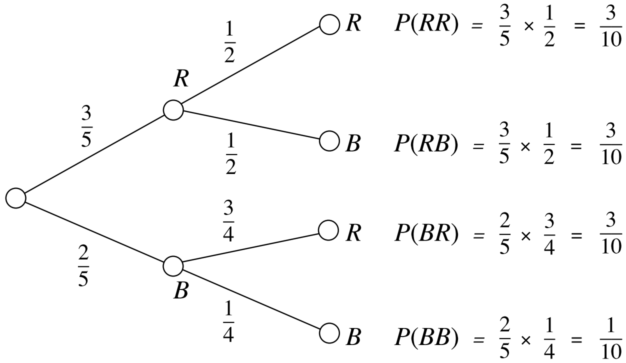


Fig. 6.2. Tree diagram with branch probabilities

by tracing back through the branch to the root. The probability of an outcome can be calculated by multiplying the probabilities along the branch. So the experiment under discussion yields the diagram Figure 6.2.

Bernoulli Trials and Binomial Experiments

Consider the experiment "flip a coin three times and count how many heads occur." This can be viewed as repeating one basic experiment that has exactly two outcomes ("flip a coin") three times and then counting the results.

We shall define a *Bernoulli trial* to be an experiment in which there are exactly two possible outcomes. In many applications it makes sense to think of one of these events as "success" and the other as "failure" (abbreviated to S and F). We shall often denote the probability of success by p. A *binomial experiment* is one in which a Bernoulli trial is repeated a certain number of times, and the outcome is the number of successes in total. So our initial example was a binomial experiment in which the trial is repeated three times. Since the number of heads is to be counted, we would probably refer to a heads as a "success." If the coin was fair, p will equal $\frac{1}{2}$.

In a binomial experiment, if the trial is repeated n times, there are $n + 1$ possible outcomes, from "0 successes" to "n successes."

Sample Problem 6.9. *A trial consists of throwing a die; a result of 1 or 2 is a success, other throws are failures. What is the probability of two successes in three trials?*

Solution. There are three ways in which two successes can occur: the sequences SSF, SFS, and FSS. Now

$$P(SSF) = \frac{1}{3} \times \frac{1}{3} \times \frac{2}{3} = \frac{2}{27},$$
$$P(SFS) = \frac{1}{3} \times \frac{2}{3} \times \frac{1}{3} = \frac{2}{27},$$
$$P(FSS) = \frac{2}{3} \times \frac{1}{3} \times \frac{1}{3} = \frac{2}{27}.$$

So the probability of two successes is

$$\frac{2}{27} + \frac{2}{27} + \frac{2}{27} = \frac{6}{27}.$$

Practice Exercise. In the experiment described in this sample problem, what is the probability of exactly one success in three trials?

More generally, the number of sequences of n trials that contain k successes and $n - k$ failures is $\binom{n}{k}$. The probability of any such sequence is $p^k(1 - p)^{n-k}$. So the probability of k successes is

$$\binom{n}{k} p^k(1 - p)^{n-k}.$$

Sample Problem 6.10. *A sales representative estimates that a sale results from one in four of his calls to companies. If he makes five calls today, what is the probability of at least two sales?*

Solution. Each call is a Bernoulli trial with $p = \frac{1}{4}$, so the five calls in a day can be thought of as a binomial experiment with $p = \frac{1}{4}$, $n = 5$. So

$$P(2 \text{ successes}) = \binom{5}{2}\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^3 = 10 \times \frac{3^3}{4^5} = \frac{270}{1024},$$
$$P(3 \text{ successes}) = \binom{5}{3}\left(\frac{1}{4}\right)^3\left(\frac{3}{4}\right)^2 = 10 \times \frac{3^2}{4^5} = \frac{90}{1024},$$

$$P(4\text{ successes}) = \binom{5}{4}\left(\frac{1}{4}\right)^4\left(\frac{3}{4}\right)^1 = 5 \times \frac{3}{4^5} = \frac{15}{1024},$$
$$P(5\text{ successes}) = \binom{5}{5}\left(\frac{1}{4}\right)^5\left(\frac{3}{4}\right)^0 = 1 \times \frac{1}{4^5} = \frac{1}{1024},$$

and the probability of at least two successes is the sum of these:

$$\frac{270 + 90 + 15 + 1}{1024} = \frac{376}{1024} = \frac{47}{128},$$

or about 37%. Alternatively, we could have noticed that "he makes *at least* two sales per day" is the complement of "he makes 0 or 1 sales per day." The probability of 0 or 1 sales is

$$\begin{aligned} P(0\text{ sales}) + P(1\text{ sale}) &= \binom{5}{0}\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^5 + \binom{5}{1}\left(\frac{1}{4}\right)^1\left(\frac{3}{4}\right)^4 \\ &= 1 \times \frac{3^5}{4^5} + 5 \times \frac{3^4}{4^5} \\ &= \frac{243 + 5 \times 81}{1024} = \frac{243 + 405}{1024} \\ &= \frac{648}{1024} = \frac{81}{128}, \end{aligned}$$

so the probability of at least two sales is

$$1 - \frac{81}{128} = \frac{47}{128}.$$

Practice Exercise. If the salesman could improve his record to give him a 50% chance of a sale from each call, what is his probability of at least three successes in the day?

Note that, for practical purposes, the exact value $\frac{47}{128}$ in the preceding sample problem is not important. It would normally be enough to say "his chance of two or more sales is a little better than one in three." In probability problems, it is important to have the correct formula, and to have an approximate idea of the numerical value. If, for any reason, a precise value is needed, it can be calculated from the formula.

Sample Problem 6.11. *The salesman in the preceding sample problem would like to have at least a* 60% *chance of two successes per day. How many calls should he make per day?*

Solution. Suppose he makes n calls. The probability of 0 or 1 success is

$$\binom{n}{0}\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^n + \binom{n}{1}\left(\frac{1}{4}\right)^1\left(\frac{3}{4}\right)^{n-1} = \frac{(1 \times 3^n)}{4^n} + \frac{n(1 \times 3^{n-1})}{4^n}.$$

So what he wants is

$$\frac{3^{n-1}(3+n)}{4^n} < \frac{4}{10}.$$

We do some experimental arithmetic:

$$n=5: \frac{3^{n-1}(3+n)}{4^n} = \frac{81 \times 8}{1024} = 0.63\ldots,$$
$$n=6: \frac{3^{n-1}(3+n)}{4^n} = \frac{243 \times 9}{4096} = 0.53\ldots,$$
$$n=7: \frac{3^{n-1}(3+n)}{4^n} = \frac{729 \times 10}{16384} = 0.44\ldots,$$
$$n=8: \frac{3^{n-1}(3+n)}{4^n} = \frac{2187 \times 11}{65536} = 0.36\ldots.$$

So at least eight calls are needed.

Practice Exercise. How many calls are needed to give the salesman at least a 70% chance of two successes per day?

The Birthday Coincidence

Sometimes probabilities can be very surprising. One well-known example is the "birthday coincidence" problem. Suppose there are 30 people in a room. What is the probability that two of them share the same birthday (day and month)?

For simplicity let us ignore leap years. (The answer will still be approximately correct.) What is the probability of the complementary event—that no two have the same birthday? (The probability of a birthday coincidence will be found by subtracting this probability from 1.) If we assume the people are ordered in some way (alphabetical order, for example), then there are 365 choices for the first person's birthday, 364 for the second, 363 for the third, and so on. So the total number of possible lists of birthdays (with no repeats) is

$$P(365, 30) = 365 \times 364 \times \cdots \times 336.$$

If there is no restriction, each person has 365 possible birthdays. So there are 365^{30} possibilities. Therefore the probability of the event "no two have the same birthday" is

$$P(E) = \frac{|E|}{|S|} = \frac{P(365, 30)}{365^{30}},$$

which is approximately 0.294. So the probability of the "birthday coincidence" is about 0.706.

Similar calculations can, of course, be carried out for any number of people. Table 6.2 shows the probability of a "birthday coincidence" for various numbers of people. It is about a 50% chance when there are 23 people in the room.

n	P_n
5	0.027
10	0.117
15	0.253
20	0.411
23	0.507
25	0.569
30	0.706
40	0.891
50	0.970

Table 6.2. P_n is the probability of a "birthday coincidence" among n people

Most students find this very surprising, and would guess that many more people would be needed to bring the probability up to 50%.

Exercises 6.2

1. A jar contains six marbles, four red and two white. In stage 1 of an experiment, one marble is selected at random and its color noted; it is not replaced. In stage 2, another marble is drawn and its color noted.

(i) Draw a tree diagram for this experiment.

(ii) What are the probabilities of the four outcomes (RR, RW, WR, WW)?

(iii) What is the probability that the two marbles selected are different colors?

2. A cage contains six mice: three white females, one white male, and two gray males. In an experiment, two mice are selected one after the other, without replacement, and sex and color are noted.

(i) Draw a tree diagram for this experiment.

(ii) Find the probabilities of all the outcomes.

(iii) What is the probability that two males are selected?

(iv) What is the probability that both a white and a gray mouse are selected?

(v) What is the probability that the second mouse is gray?

3. Three cards are dealt in order from a standard deck. In each case it is only recorded whether the card is a face card (ace, king, queen, jack) or a minor card (10 through 2).

(i) Draw a tree diagram for this experiment. Find the probabilities of the different outcomes.

(ii) What is the probability that at least two face cards are dealt?

4. Three cards are dealt in order from a standard deck. Only the color (red or black) of the card is recorded. The cards are not replaced.

(i) Draw a tree diagram for this experiment.

(ii) What is the probability that the three cards are red?

5. Eight evenly matched horses run a race. Three are bay colts, one is a bay filly, two are brown colts, and two are brown fillies. The color and sex of the first and second finishers are recorded.

(i) Draw a tree diagram for this experiment.

(ii) What is the probability that both the horses recorded are colts?

(iii) What is the probability that the two horses are of different colors?

(iv) What is the probability that the second horse is a filly?

6. A jar contains four red, three white, and two blue marbles. One marble is drawn and placed aside; then another marble is drawn. Only the colors are recorded. Draw a tree diagram for this experiment. What is the probability that the two marbles are of different colors?

7. A coin is weighted so that a heads is twice as likely to occur as a tails. It is tossed four times, and the result is noted.

(i) What is the probability that exactly one heads occurs?

(ii) What is the probability that at least three heads occur?

8. A die is rolled six times. What is the probability of

(i) exactly one six?

(ii) at most one six?

9. A fair coin is tossed six times. Which is more probable—that heads and tails occur three times each, or that they divide 4 and 2 (either 4 heads or 4 tails)?

10. A multiple choice test contains five questions, and each question has three possible answers. A passing grade is three or more correct answers. What is the probability of passing if you guess the answers at random?

11. In a 20-question, true-false test, what is the probability of getting exactly 16 answers correct by random guessing?

12. A pair of fair dice are rolled and the total is recorded. If this is done ten times, what is the probability that 7 is recorded at least three times?

13. On average, 1 light bulb in 20 is defective. What is the probability that there is more than one defective bulb in a box of ten?

14. A drug is effective in nine cases out of ten. If it is tested on 12 patients, what is the probability that it will be ineffective in more than two cases?

15. A baseball pitcher estimates that he can throw his fast ball for a strike seven times in every eight.

(i) What is the probability that he throws exactly seven strikes in eight pitches?

(ii) What is his probability of throwing at least seven strikes in eight pitches?

16. A children's game uses a die that has 1 on three of its faces, 2 on two of its faces, and a 3 on the last face. If it is rolled six times, what is the probability of rolling 2 exactly three times?

17. The probability that a new tire will last 35,000 miles is 0.9. If you replace all four tires now, what is the probability that you will need to replace at least one within the next 35,000 miles?

18. There are five people in a room. What is the probability that two were born in the same month (but not necessarily the same year)?

6.3 Counting and Probability

All of the elementary counting principles that we studied in Chapter 5 are useful in calculating probabilities. We shall show this using several examples.

Choosing Marbles

Many elementary problems involve random drawing of marbles from a jar (or bag or box). These obviously arise in lotteries and other games of chance, but this is not the only reason for studying such problems. The mathematics is the same if we consider many other types of random selection. As an example, suppose there are 500 Democrat supporters and 400 Republicans in your area. Working out the probabilities of various combinations being contacted in an opinion poll involves exactly the same calculations as the ones made in studying random drawings from a jar of 500 black and 400 white marbles. There are also examples in physical science—for example, molecules escaping from a container of heated gas behave in the same way.

Sample Problem 6.12. *There are* 12 *red marbles and* 3 *blue marbles in a jar. Three are selected at random. What is the probability that all are red?*

Solution. There are 15 marbles, so the number of ways of selecting three is $\binom{15}{3}$. So $|S| = \binom{15}{3}$. The number of ways of selecting three red balls is $|E| = \binom{12}{3}$. So

$$\begin{aligned} P(E) &= \frac{\binom{12}{3}}{\binom{15}{3}} = \frac{12!}{9!3!} \times \frac{12!3!}{15!} \\ &= \frac{12 \cdot 11 \cdot 10}{15 \cdot 14 \cdot 13} = \frac{44}{91}. \end{aligned}$$

Practice Exercise. An urn contains five red, four blue, and two green marbles. Three marbles are chosen at random. What is the probability that the three are all different colors?

Sample Problem 6.13. *A box contains four red, three white, and two blue balls. Two balls are chosen at random. What is the probability that they are the same color?*

Solution. There are $\binom{9}{2}$ selections available. There are $\binom{4}{2} = 6$ ways to choose two red balls, $\binom{3}{2} = 3$ ways to choose two white balls, and $\binom{2}{2} = 1$ ways to choose two blue balls. So

$$P(\text{two are the same}) = \frac{12+6+1}{\binom{9}{2}} = \frac{19}{36}.$$

Practice Exercise. A jar contains six white and four blue balls. Two balls are chosen at random. What is the probability of selecting at least one white ball?

Card Problems

Remember that in a standard deck there are 52 cards, 13 in each of the four suits. So there are four cards in each of the 13 denominations: Ace, King, ..., 4, 3, 2. In most card games, the hands of cards are dealt face down to the players, so the order in which cards are received does not matter.

Sample Problem 6.14. *A poker hand of five cards is dealt from a standard deck. What are the probabilities of the following hands?*

(i) *The four kings and the ace of spades.*

(ii) *A full house* (*three cards of one denomination, two of another*).

(iii) *A hand with no pair* (*all different denominations*).

Solution. In each case, the hand dealt is one of the $\binom{52}{5}$ possible selections of five cards from the full deck of 52. So, in each case, $|S| = \binom{52}{5}$.

(i) The event E: "the hand consists of four kings and the ace of spades" contains only one outcome. So

$$\begin{aligned} |E| &= 1, \\ P(E) &= \frac{|E|}{|S|} = \frac{1}{\binom{52}{5}} \\ &= \frac{5\cdot 4\cdot 3\cdot 2\cdot 1}{52\cdot 51\cdot 50\cdot 49\cdot 48} = \frac{1}{2598960}. \end{aligned}$$

(ii) Let's say a full house is of "type (x, y)" if it contains three x's and two y's—one could have "type (7,5)," "type (king, 4)," and so on. There are 13 ways to choose the denomination of the three, and for each such choice there are 12 ways to select the two. So there are $13 \times 12 = 156$ types. Once the type is known, we can calculate the number of hands of that type. For example, if there are three kings and two fours, then there are $\binom{4}{3}$ ways of choosing which kings are to be used (we must choose three out of the four possible kings), and $\binom{4}{2}$ ways of choosing which fours. So there are $\binom{4}{3} \times \binom{4}{2} = 4 \times 6 = 24$ hands of each type. So the number of full houses is

$$\begin{aligned} 13 \times 12 \times \binom{4}{3} \times \binom{4}{2} \\ = 13 \times 12 \times 24, \end{aligned}$$

and the probability is

$$\frac{13 \cdot 12 \cdot 24}{\binom{52}{5}} = \frac{6}{4165}.$$

(iii) If there is no pair, the hands have one card of each of five denominations. This collection of denominations can be chosen in $\binom{13}{5}$ ways. The card in each denomination can be selected in four ways. So the number of hands is

$$\binom{13}{5} \times 4^5,$$

and the probability is

$$\frac{\binom{13}{5} \cdot 4^5}{\binom{52}{5}} = \frac{2112}{4165},$$

which is a little greater than 50%.

Practice Exercise. A poker hand of five cards is dealt from a standard deck. What is the probability that it contains a flush (all five cards of the same suit)?

Choosing a Committee

Sample Problem 6.15. *A committee of four people is to be chosen at random from a club with* 12 *members, four men and eight women.*

(i) *What is the probability that at least one man is chosen?*

(ii) *What is the probability that one particular member, Jack Smith, is chosen?*

Solution. The number of possible outcomes—all possible committees—is $\binom{12}{4}$. This is $|S|$. They are equally likely.

(i) Let E be the event that at least one man is chosen. Then $\overline{E}$ is the event that no man is chosen. $|\overline{E}| = \binom{8}{4}$. So $|E| = \binom{12}{4} - \binom{8}{4}$ and

$$P(E) = 1 - \frac{\binom{8}{4}}{\binom{12}{4}} = 1 - \frac{14}{99} = 1 - \frac{85}{99}.$$

(ii) The number of committees containing a given individual is $\binom{11}{3}$. (Jack Smith must be a member; the remaining three are chose from the other 11 members and that can be done in $\binom{11}{3}$ ways.) Therefore,

$$P(\text{Jack Smith is chosen}) = \frac{\binom{11}{3}}{\binom{12}{4}} = \frac{1}{3}.$$

Practice Exercise. A class has 18 members—six math majors, five in economics, and seven in computer science. A class committee of three is chosen at random. What is the probability that all three have the same major?

Derangement Problems

The number of derangements of n objects is

$$D_n = n!\left[1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!}\right].$$

The number of ways to arrange n objects is $n!$, so the probability that a randomly selected arrangement is a derangement is

$$\frac{D_n}{n!} = 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!}.$$

Sample Problem 6.16. *Four students have identical backpacks. As they leave the library, each chooses a backpack at random. What is the chance that none gets the correct backpack?*

Solution. The probability is

$$\frac{D_4}{4!} = 1 - \frac{1}{1} + \frac{1}{2} - \frac{1}{6} + \frac{1}{24} = \frac{3}{8}.$$

Practice Exercise. What is the probability that some, but not all, get their correct backpacks?

Exercises 6.3

1. There are three red, three blue, and four white marbles in a jar. Two marbles are chosen at random. What are the probabilities that

(i) both are blue?

(ii) the two are different colors?

2. There are ten red, three white, and four blue marbles in a jar. Two are chosen at random. What are the probabilities that

(i) neither are red?

(ii) exactly one is red?

(iii) both are red?

3. From the jar in the preceding exercise, three marbles are chosen. What are the probabilities that

(i) the three are red?

(ii) the three are of different colors?

4. A jar contains four red, five blue, and three white marbles. Three marbles are chosen at random. What are the probabilities that

(i) all three are red?

(ii) none are red?

5. A student must select two courses from a list of two humanities and three science courses. He selects at random. What is the probability that both are science courses?

6. A box of 50 matches contains three broken matches. Two matches are drawn at random. What is the probability that neither is broken?

7. Your midterm test contains ten questions with true-false answers. If you select your answers at random, what are the probabilities of getting

(i) exactly four correct answers?

(ii) exactly two correct answers?

(iii) at most two correct answers?

8. A committee of six people (four men and two women) need to select a chairman and secretary. They do so by drawing names from a hat. What are the probabilities that

(i) both are men?

(ii) exactly one is a man?

9. In the game Yahtzee, five fair dice are rolled at once. What are the probabilities of getting

(i) a Yahtzee, that is, all five dice the same?

(ii) four of a kind (that is four of one denomination, with the other different)?

(iii) a full house (that is three of one denomination, and two of another)?

10. An electrician knows that two switches are faulty out of a batch of five. He tests them one at a time. What is the probability that he finds the two faulty switches in the first two trials?

11. A five-card poker hand is dealt at random from a standard deck. What is the probability that it contains a straight (A2345, 23456, 34567, ..., 10JKQA)? (In a straight, suits do not matter.)

12. Suppose license plates consist of two letters followed by four numbers. Suppose all possible combinations have been used. If you select a car at random, what is the probability that the plate has

(i) both letters the same?

(ii) no repeated symbol (letter or number)?

(iii) an even number?

13. A congressional committee contains five Democrats and four Republicans. To choose a subcommittee they list all groups of four that contain at least one Republican and one Democrat, and then select one at random. What is the probability that the selected subcommittee consists of two Democrats and two Republicans?

14. A secretary types three letters and the three corresponding envelopes. She is in a hurry to quit work for the day, and forgets to check that she puts the correct letters in the correct envelopes, so it is as though she puts the letters in the envelopes at random. What are the probabilities that

(i) every envelope gets the correct letter?

(ii) no envelope gets the correct letter?

15. A committee of six men and four women select a steering committee of three people. If all combinations are equally likely, what is the probability that all three will be of the same gender?

16. Four people apply for two jobs. Among them are a husband and wife. Assume that the four are equally well qualified and choices are made at random.

(i) What is the probability that both husband and wife are appointed?

(ii) What is the probability that at least one is appointed?

17. In a lottery, you must select three different numbers from 1, 2, 3, 4, 5, 6, 7, 8, 9. If your three are the numbers drawn, you win first prize; if you have two right out of three, you win second prize. (The order of the numbers does not count.)

(i) What is the probability of winning first prize?

(ii) What is the probability of winning second prize?

18. A club has 12 members, including Mr. and Mrs. Smith. A committee of three is to be chosen at random.

(i) What is the probability that both are chosen?

(ii) What is the probability that neither are chosen?

19. Your discussion group chooses its chairman at each monthly meeting by drawing names from a hat. If there are 12 members, what is the probability that you will not be chosen at any meeting this year?

20. A pinball machine selects a number from 0 to 9 at random; you are awarded a free game if that number equals the last digit in your score.

(i) What is the probability that you win a game in this way on your first attempt?

(ii) What is the probability that you do not win a game on your first attempt, but you win a game on your second attempt?

(iii) What is the probability that you will get no "match" in five games?

6.4 Conditional Probabilities

Conditional Probability Defined

Suppose you draw a marble at random from a jar containing two red and two blue marbles. You look at this marble, then discard it, and draw another marble. What is the probability of drawing a blue marble the second time? There are two answers. If the first marble is red, then the probability is $\frac{2}{3}$; if it is blue, the probability is $\frac{1}{3}$. We write

$$P(\text{second blue} \mid \text{first red}) = \frac{2}{3},$$
$$P(\text{second blue} \mid \text{first blue}) = \frac{1}{3}.$$

The sign "|" is read as the word "given," so the first probability above is "probability that the second is blue given that the first is red." These probabilities are called *conditional* because they express the probability of a result in the second draw if a certain condition (the result of the first draw) is satisfied.

In general, suppose the fact that event A occurs gives extra information about the probability that B occurs. The *probability that B occurs, given that A occurs*, which is the *conditional probability* $P(B \mid A)$, is the best measure to use in deciding whether or not B will happen.

Sample Problem 6.17. *Two cards are dealt from a standard deck. The first card is not replaced before the second is dealt. What is the probability that the second card is a king, given that the first card is an ace? What is the probability that the second card is a king, given that the first card is a king?*

Solution. Given that the first card is an ace, the second card is a random selection from 51 equally likely possibilities. Four of these outcomes are kings. So

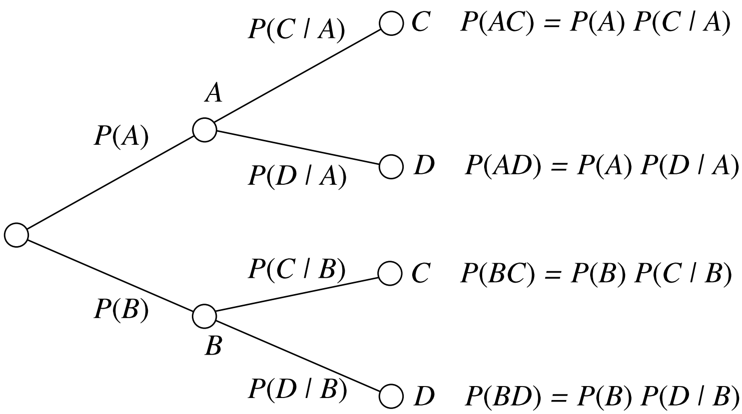


Fig. 6.3. Tree diagram with branch probabilities

$$P(K \mid A) = \frac{n(E)}{n(S)} = \frac{4}{51}.$$

If the first card is a king, there are only three kings among the cards for the second trial, so

$$P(K \mid K) = \frac{3}{51}.$$

Practice Exercise. Two cards are dealt, without replacement, from a standard deck. If the first is a spade, what is the probability that the second is (i) a spade? (ii) a heart?

The probabilities written on the later branches of a tree diagram are all conditional probabilities. If an experiment consists of two stages in which the first stage has possible outcomes A and B and the second stage has possible outcomes C and D, the tree diagram is shown in Figure 6.3. Of course, this all works in the same way when there are more than two stages, or when there are more than two outcomes at a stage.

Sample Problem 6.18. *Two cards are dealt from a standard deck without replacement, and it is noted whether or not the cards are kings. What are the outcomes of this experiment, and their probabilities? Represent the information in a tree diagram.*

Solution. We shall write K and N for "a king" and "a card other than a king." By a similar argument to Sample Problem 6.17 we see that the probability of a king on the second draw, given a king on the first draw, is

$$P(K \mid K) = \frac{3}{51},$$

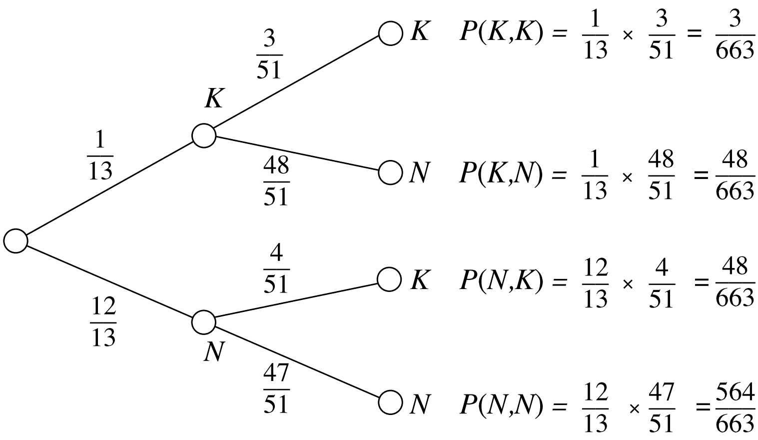


Fig. 6.4. Tree diagram for Sample Problem 6.18

and similarly

$$P(K \mid N) = \frac{4}{51}.$$

Given that the first card is a king, the second card is either a king or not, so

$$P(K \mid K) + P(N \mid K) = 1;$$

therefore,

$$P(N \mid K) = \frac{48}{51},$$

and similarly,

$$P(N \mid N) = \frac{47}{51}.$$

So the diagram is the one shown in Figure 6.4 (where, for example, "K, N" means "king first, non-king second").

Practice Exercise. Repeat this sample Problem, in the case where the information noted is whether or not the card is a spade.

The Conditional Probability Formula

The probability of the sequence "A followed by B" is

$$P(AB) = P(A)P(B|A). \tag{6.1}$$

This formula can be used more generally, even when A does not happen before B. Whenever we get more information about whether an event B occurs by knowing whether or not event A occurs, we define the conditional probability of B given A by

$$P(B|A) = \frac{P(A \cap B)}{P(A)}. \tag{6.2}$$

Sample Problem 6.19. *A car pool contains both ten red and ten white Fords and 15 red and five white Buicks. A car is chosen at random, by picking up the keys.*

1. *What is the probability that a red car is chosen?*
2. *You notice that the keys belong to a Buick. What is the probability that a red car is chosen?*

Solution. 1. There are 40 cars, of which 25 are red. So

$$P(R) = \frac{|R|}{|S|} = \frac{25}{40} = \frac{5}{8}.$$

2. There are 20 Buicks of which 15 are red. So

$$\begin{aligned} P(B) &= \frac{20}{40} = \frac{1}{2}, \\ P(R \cap B) &= \frac{15}{40} = \frac{3}{8}, \\ P(R \mid B) &= \frac{P(R \cap B)}{P(B)} = \frac{3}{8} \times \frac{2}{1} = \frac{3}{4}. \end{aligned}$$

Practice Exercise. There are five Republicans and three Democrats on a committee. A subcommittee of two is chosen by a random drawing.

1. What is the probability that both are Democrats?
2. You are told that the committee contains at least one Democrat. What is the probability that both are Democrats?

Sample Problem 6.20. *A carpenter uses screws made by two companies, X and Y; 40% of his stock comes from X and the rest from Y. About 2% of the screws from Y are faulty. If he chooses a screw at random, what is the probability that it is a faulty screw from company Y?*

Solution. We have

$$P(Y) = 0.6, \qquad P(F \mid Y) = 0.02.$$

So

$$\begin{aligned} P(Y \cap F) &= P(Y)P(F \mid Y) \\ &= 0.6 \times 0.02 = 0.012, \end{aligned}$$

and the probability is 0.012 or 1.2%.

Practice Exercise. Records show that two March days out of five are dull, and it rains on one out of every three dull days. What is the probability that March 12th next year will be dull and rainy?

Independence

We say two events are independent in everyday English if they have no effect on each other. In probability theory, this idea is formalized as follows: Events A and B are *independent* when

$$P(A \mid B) = P(A).$$

Provided neither A nor B is impossible, the relation

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)},$$

leads us to the equivalent definition: A and B are independent when

$$P(A \cap B) = P(A) \times P(B).$$

Sample Problem 6.21. *Two dice are rolled and their sum is recorded. Consider the following events.*

A: Sum is 2, 8, *or* 11.

B: Sum is even.

C: Sum is 4, 7, *or* 10.

Which of these events are independent?

Solution. We write E_i for the outcome "the sum is i." These outcomes are the events shown with their probabilities in Table 6.1. Then

$$\begin{aligned}
A &= \{E_2, E_8, E_{11}\},\\
B &= \{E_2, E_4, E_6, E_8, E_{10}, E_{12}\},\\
C &= \{E_4, E_7, E_{10}\},\\
A \cap B &= \{E_2, E_8\},\\
A \cap C &= \emptyset,\\
B \cap C &= \{E_4, E_{10}\},
\end{aligned}$$

and from the table,

$$\begin{aligned}
P(A) &= (1+5+2)/36 = 8/36,\\
P(B) &= (1+3+5+5+3+1)/36 = 18/36,\\
P(C) &= (3+6+3)/36 = 12/36,
\end{aligned}$$

$$P(A \cap B) = \frac{(1+5)}{36} = 6/36,$$
$$P(A \cap C) = 0,$$
$$P(B \cap C) = (3+3)/36 = 6/36.$$

Now

$$P(A) \times P(B) = 8/36 \times 18/36 = 4/36 \neq P(A \cap B),$$
$$P(A) \times P(C) = 8/36 \times 12/36 = 8/108 \neq P(A \cap C),$$
$$P(B) \times P(C) = 18/36 \times 12/36 = 6/36 = P(B \cap C),$$

so B and C are independent, but A and B are dependent, and so are A and C.

Practice Exercise. An urn contains eight balls numbered 1 through 8. Balls 1, 2, and 3 are red; 4, 5, 6, and 7 are white; 8 is blue. One ball is drawn. Consider the following events

A: The ball is red.

B: The ball is blue.

C: The number is odd.

Which of these events are independent?

A Summation Formula

Suppose two events, A and B are considered. Since

$$A = (A \cap B) \cup (A \cap \overline{B}),$$

and the latter events are mutually exclusive, we have

$$P(A) = P(A \cap B) + P(A \cap \overline{B}),$$

so

$$P(A) = P(A \mid B)P(B) + P(A \mid \overline{B})P(\overline{B}).$$

More generally, if the possible outcomes of an experiment are $B_1, B_2, \ldots, B_k$, then

$$P(A) = \sum_{i=1}^{k} P(A \mid B_i)P(B_i). \tag{6.3}$$

In terms of tree diagrams, this formula is a way of stating in symbols the following rule: "to find the probability of A, find the probabilities associated with all branches that contain A, and sum them."

Sample Problem 6.22. *There are six boxes, two of them round and four square. Each round box contains two green marbles and three blue marbles. Each square box contains one green marble and three blue marbles. A box is chosen at random and a marble is chosen at random from it. What is the probability that the marble is blue? What is the probability that the box was round, given that the marble was blue? Represent the probabilities in a tree diagram.*

Solution. We write R, S, G, B for round, square, green, blue. Then the probabilities of choosing round and square boxes are

$$P(R) = 1/3, \qquad P(S) = 2/3.$$

So

$$\begin{aligned} P(B) &= P(B \mid R)P(R) + P(B \mid S)P(S) \\ &= (3/5)(1/3) + (3/4)(2/3) \\ &= 1/5 + 1/2 = 7/10. \end{aligned}$$

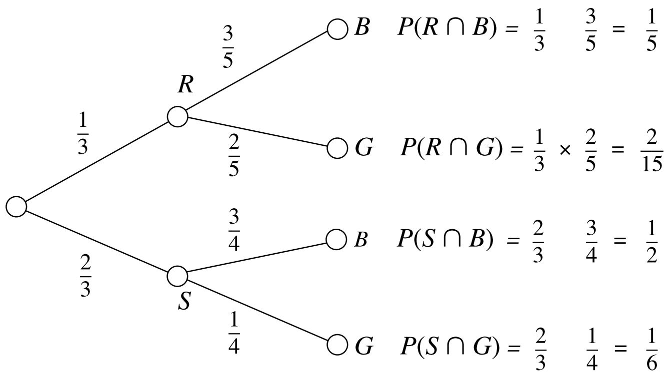
We can calculate $P(R \cap B)$:

$$P(R \cap B) = P(B \mid R)P(R) = (3/5)(1/3) = 1/5.$$

So

$$P(R \mid B) = \frac{P(R \cap B)}{P(B)} = \frac{1/5}{7/10} = \frac{2}{7}.$$

The diagram is



$$P(B) = P(R \cap B) + P(S \cap B) = 1/5 + 1/2 = 7/10.$$

(This last equation is the sum of the probabilities of all branches that contain B.)

Practice Exercise. Consider the builder in Sample Problem 6.20. Suppose further that about 1.5% of the screws from company X are faulty. If he selects a screw at random, what is the probability that it is faulty?

Exercises 6.4

1. Box A contains three red pens and four blue pens. Box B contains two red, one green, and one blue pen. A pen is selected from Box A at random, and placed in Box B. Then a pen is selected at random from Box B. What is the probability that this pen is

(i) red? (ii) blue? (iii) green?

2. A city worker can either go to work by car or by bus. If she goes by car, she uses the tunnel 40% of the time and the bridge 60% of the time. If she takes the bus, it is equally likely that her bus will use the tunnel or the bridge.

(i) Draw a tree diagram to represent her possible routes to work.

(ii) Suppose she drives on three days out of five and takes the bus twice out of every five times. What are the probabilities of the various outcomes? What is the probability, on a given day, that she crosses the bridge on her way to work?

3. Box 1 contains five red balls and three white balls. Box 2 contains two red and two white balls. An experiment consists of selecting two balls at random from Box 1 and placing them in Box 2, then selecting one ball from Box 2 at random.

(i) Draw a tree diagram for this experiment.

(ii) What is the probability that the ball chosen from Box 2 is red?

4. A card is dealt at random from a regular deck.

(i) What is the probability that it is a jack given that it is a picture card (king, queen, or jack)?

(ii) What is the probability that it is a 5, given that it is not a picture card?

5. Two fair dice are rolled. Consider the following events.

A: The sum is 6.

B: Both dice show even numbers.

C: At least one die shows a 4.

What are

$$P(A \mid B),\ P(A \mid C),\ P(B \mid A),\ P(B \mid C),\ P(C \mid A),\ P(C \mid B)?$$

6. A pair of fair dice are rolled. Consider the following events.

D: The sum is 7.

E: The sum is odd.

F: At least one die shows a 4.

What are

$$P(D \mid E),\ P(E \mid D),\ P(D \mid F),\ P(F \mid D)?$$

7. Two cards are selected from a regular deck, without replacement. Consider the following events.

A: Both cards are red.

B: The second card is red.

C: At least one card is red.

D: At least one card is a heart.

What are

$$P(A),\ P(D),\ P(A \mid B),\ P(A \mid C),\ P(D \mid B),\ P(D \mid C)?$$

8. Two fair dice are rolled. Consider the following events.

A: The sum is 8.

B: Both dice show even numbers.

C: At least one die shows a 5.

What are

$$P(A \mid B),\ P(A \mid C),\ P(B \mid A),\ P(B \mid C),\ P(C \mid A),\ P(C \mid B)?$$

9. A jar contains three white and three red marbles. Two are drawn in succession, without replacement, and the colors are noted.

(i) Draw a tree diagram for this experiment.

(ii) Let A, B, and C denote the following events.

A: Both marbles are red.

B: The second marble is red.

C: At least one marble is red.

What are $P(A)$, $P(A \mid B)$, $P(A \mid C)$, $P(B \mid C)$?

10. In Exercise 6.2.2, what are the conditional probabilities of the following events?

(i) Two males are selected, given that at least one male is selected.

(ii) Two females are selected, given that at least one white mouse is selected.

11. In Exercise 1, what are the conditional probabilities of the following events?

(i) The pen selected from Box B is blue, given that the pen selected from Box A was blue.

(ii) The pen selected from Box B is blue, given that the pen selected from Box A was red.

12. In Exercise 6.4.2, what are the conditional probabilities of the following events?

(i) The worker traveled over the bridge on her way to work, given that she traveled by bus.

(ii) She used the tunnel, given that she came by car.

13. In Exercise 6.4.3, what are the probabilities of the following events?

(i) The ball chosen from Box 2 is red, given that both balls chosen from Box 1 were red.

(ii) The ball chosen from Box 2 is red, given that at least one ball chosen from Box 1 was red.

(iii) The ball chosen from Box 2 is red, given that at most one ball chosen from Box 1 was red.

In Exercises 14 to 21, illustrate the data with a Venn diagram and calculate $P(A \mid B)$ *and* $P(B \mid A)$.

14. $P(A) = 0.7$, $P(B) = 0.6$, $P(A \cup B) = 0.9$.

15. $P(A) = 0.7$, $P(B) = 0.4$, $P(A \backslash B) = 0.4$.

16. $P(A) = 0.6$, $P(B) = 0.4$, $P(A \cap B) = 0.1$.

17. $P(A \cup B) = 0.7$, $P(A \cap B) = 0.3$, $P(B) = 0.3$.

18. $P(A) = 0.6$, $P(B) = 0.3$, $P(A \cap B) = 0.2$.

19. $P(A \cup B) = 0.8$, $P(A \cap B) = 0.4$, $P(A) = 0.6$.

20. $P(A) = 0.4$, $P(B) = 0.6$, $P(A \cup B) = 0.9$.

21. $P(A) = 0.6$, $P(B) = 0.6$, and $P(A \backslash B) = 0.4$.

In Exercises 22 to 27, A and B are independent events. Find $P(A \cup B)$ *and* $P(A \cap B)$ *when the probabilities are as given.*

22. $P(A) = 0.6$, $P(B) = 0.4$.

23. $P(A) = 0.5$, $P(B) = 0.8$.

24. $P(A) = 0.8$, $P(B) = 0.8$.

25. $P(A) = 0.7$, $P(B) = 0.5$.

26. $P(A) = 0.4$, $P(B) = 0.4$.

27. $P(A) = 0.9$, $P(B) = 0.7$.

28. Two cards are selected without replacement from a regular deck. It is recorded whether or not each card is a spade. Define the events A and B as follows.

A: The first card is a spade.

B: The second card is a spade.

(i) Draw a tree diagram for this experiment.

(ii) Find $P(A)$, $P(B)$, $P(A \mid B)$, $P(B \mid A)$, and $P(A \cap B)$.

(iii) Are A and B independent?

29. Repeat Problem 28 in the case where the first card is replaced and the deck is shuffled before the second selection.

30. A box contains six red and four blue marbles. Two marbles are drawn from the box (without replacement) and their colors are noted. Define the following events.

E: The first marble is red.

F: The second marble is red.

(i) Draw a tree diagram for this experiment.

(ii) Find $P(E)$, $P(F)$, $P(E \mid F)$, $P(F \mid E)$, and $P(E \cap F)$.

(iii) Are E and F independent?

31. Repeat problem 30 in the case where the first marble is replaced before the second drawing.

32. A weather forecasting station predicts that on December 18 it will snow with a probability of 0.6, there will be a change in the wind direction with a probability of 0.8, and there will be both snow and a change in the wind direction with a probability of 0.4. If the prediction is correct:

(i) What is the probability that there will be neither snow nor a change in the wind direction?

(ii) What is the probability of snow given that the wind direction has changed?

33. A random sample of 100 people were asked their income level and whether they regularly invest in the stock market. The following table gives the results of the survey

Income level	Regularly invest in stock market	
	Yes	No
High	18	6
Medium	21	16
Low	15	24

A person in the survey is chosen at random. Let H be the event that the person has a high income. Let Y be the event that the person regularly invests in the stock market, and let N be the event that the person does not regularly invest in the stock market. Find

(i) $P(H)$,

(ii) $P(Y)$,

(iii) $P(H \cup Y)$,

(iv) $P(N \mid H)$.

34. Professor Jones has taught the same course for the last 20 years. Each time he teaches it he gives a pretest on the first day of the class. At the end of each semester he compares course grades with whether or not students have passed the pretest. Seventy percent of the course grades he gives are C or better. He finds that 80% of those students who got C or better in the course have passed the pretest while 30% of the students who got less than a C in the course also passed the pretest.

(i) Draw and label a tree diagram illustrating the process.

(ii) Find the probability that a student who takes the pretest will *not* pass it.

(iii) If a student has not passed the pretest, what is his probability of getting less than a C in the course?

35. Two dice are rolled and the sum of the top faces is recorded. Let E be the event that the sum is an odd number, let F be the event that the sum is 4, 7, or 10, and let G be the event that the sum is 2, 7, or 11.

(i) Find $P(G \mid E)$.

(ii) Determine whether E and F are independent.

36. Suppose an urn contains two red, four green, and five blue marbles. A marble is selected from the urn and its color noted. If it is red or green it is withdrawn, otherwise it is replaced. Then a second marble is selected from the urn. Determine the probability the first marble was red, given the second was red.

37. Two fair dice, one red and one green, are rolled. Events A, B, and C are defined as follows.

A: The sum of the two dice is 7.

B: The green die shows a 3.

C: The green die shows a 1.

(i) Are A and B independent?

(ii) Are A and C independent?

38. Three jars, labeled A, B, and C, contain red and green jelly beans as follows: A has two red and two green, B has four red and three green, and C has two red and five green. A jar is selected at random, and one jelly bean is selected at random from it. Represent this experiment in a tree diagram, and find the probability that the jelly bean is red.

39. Suppose E, F, and G are any three events. Prove that

$$P(E \cap F \cap G) = P(E) \times P(F \mid G) \times P(G \mid E \cap F).$$

40. Forty percent of those who take drugs also have an alcohol problem and 5% of those who do not take drugs have an alcohol problem. If 32% of the population take drugs, what is the probability that a person who has an alcohol problem also takes drugs?

41. Two fair dice, one red and one blue, are rolled. Events A, B, and C are defined as follows.

A: The sum of the two numbers shown is 6.

B: The number showing on the blue die is 2.

C: The blue die shows a 3.

(i) Are A and B independent?

(ii) Are A and C independent?

42. Two jars, labeled A and B, contain marbles as follows: A has two red, one white, and four blue, and B contains six red, two white, and two blue. A jar is selected at random and a marble is selected at random from it. Represent this experiment in a tree diagram. What is the probability that the marble chosen is blue?

In Exercises 43 to 50, E and F are events in the sample space S. You may assume that $0 < P(E) < 1$ and $0 < P(F) < 1$. Is the statement true, false, or is it impossible to decide?

43. $P(E \mid E) = 1$.

44. $P(E \mid \overline{E}) = 1$.

45. $P(F \mid S) = 1$.

46. $P(S \mid F) = 1$.

47. $P(F \mid S) = P(F)$.

48. $P(E \mid E \cap F) = 0$.

49. $P(E \mid F) = P(F \mid E)$.

50. $P(E \cup F) \geq P(E \cap F)$.

6.5 Bayes' Formula and Applications

Bayes' Formula

If A and B are any two events, then $A \cap B$ and $B \cap A$ are the same event. However, the conditional probability formula gives two different looking expressions for their probabilities:

$$P(A \cap B) = P(B \mid A)P(A),$$
$$P(B \cap A) = P(A \mid B)P(B).$$

It follows that the two expressions must be equal:

$$P(B \mid A)P(A) = P(A \mid B)P(B).$$

This formula is usually written in the form

$$P(B \mid A) = \frac{P(A \mid B)P(B)}{P(A)}. \tag{6.4}$$

Sample Problem 6.23. *A store gets 40% of its stock of light bulbs from factory X, 35% from Y, and 25% from Z. Some bulbs are faulty: 1% of the output from X, 2% from Y, and 3% from Z. A light bulb is chosen at random from stock and found to be faulty. What is the chance that it comes from factory Z?*

Solution. Write Z for the event "the bulb comes from factory Z" and A for "the bulb is faulty." Then $P(Z \mid A)$ is required. From the data we know $P(Z) = 0.25$ and $P(A \mid Z) = 0.03$. To calculate $P(A)$ we use the formula in (6.3):

$$\begin{aligned}P(A) &= P(A \mid X)P(X) + P(A \mid Y)P(Y) \\ &\quad + P(A \mid Z)P(Z) \\ &= (0.01)(0.40) + (0.02)(0.35) + (0.03)(0.25) \\ &= 0.004 + 0.007 + 0.0075 \\ &= 0.0185.\end{aligned}$$

So

$$P(Z \mid A) = \frac{(0.25)(0.03)}{0.0185} = \frac{75}{185} = \frac{15}{37},$$

which is between 40% and 41%.

Practice Exercise. You are given three coins; one is biased so that it shows heads two-thirds of the time, and the other two are fair (heads and tails are equally likely). You select a coin at random and flip; it shows heads. What is the probability that it is the biased coin?

The formula in (6.4) is called *Bayes' Formula*. It is often combined with (6.3) as follows. Suppose the possible outcomes of an experiment are $B_1, B_2, \ldots, B_k$. Then

$$P(B_i \mid A) = \frac{P(A \mid B_i)P(B_i)}{P(A \mid B_1)P(B_1) + \cdots + P(A \mid B_k)P(B_k)}. \tag{6.5}$$

(The denominator of (6.5) comes directly from (6.3).)

When Bayes' formula is used, very often the outcomes $B_1, B_2, \ldots, B_k$ are the outcomes of an experiment that occurred earlier, and A is an outcome, or set of outcomes, of a later experiment. Bayes' formula answers the question, "given the outcome of the second experiment, what is the probability that the outcome of the first experiment was so-and-so?" This sometimes confuses students because it seems to suggest that the outcome of the later experiment can somehow affect the earlier experiment, and that is impossible—causes must come before effects. What, in fact, happens is that knowledge of the later experiment can increase our (incomplete) knowledge of the earlier experiment.

Sample Problem 6.24. *The car pool contains ten Fords* (*five red and five white*) *and* 15 *Pontiacs* (*five red and ten white*). *You are allocated a car at random. You see from a distance that it is red. What is the probability that you have been given a Ford?*

Solution. We require $P(F \mid R)$. There are 25 cars of which 10 are Fords, so $P(F) = \frac{10}{25} = 0.4$, and ten are red, so $P(R) = 0.4$ also. The probability of a red car, given that it is a Ford, is 0.5, since half the Fords are red. So

$$\begin{aligned}P(R)P(F \mid R) &= P(F)P(R \mid F), \\ 0.4 \times P(F \mid R) &= 0.4 \times 0.5, \\ P(F \mid R) &= 0.5.\end{aligned}$$

Practice Exercise. Box X contains two blue pens and three red pens. Box Y contains two red pens and three blue pens. A box is chosen at random and a pen is chosen at random from it. If the pen is blue, what is the probability that the box was box X?

Tree Diagrams

It is often convenient to use tree diagrams in Bayes' formula problems. When the diagram is completed, the terms to be added for the denominator of (6.5) are on the right-hand side.

The computation of Bayes' formula can be carried out as follows.

1. First construct a tree diagram with the possible outcomes $B_1, B_2, \ldots, B_k$ as branches.
2. To each of these branches add the branches A ("the event A occurs") and $\overline{A}$ ("A does not occur"). The paths ending in A together represent all the circumstances in which A can occur.
3. Now calculate the probability—the product of the conditional probabilities—for each branch that ends in A. The sum of these probabilities will be $P(A)$, the denominator of (6.5).
4. The numerator $P(B_i \mid A)$ will be written next to one of the branches.

Sample Problem 6.25. *Construct a tree diagram for Sample Problem 6.23.*

Solution. In the terminology of experiments, Sample Problem 6.23 can be described as follows. First, a supplier is chosen; then a light bulb is chosen from that supplier. So the outcomes of the first experiment (the $B_1, B_2, \ldots$) are factories X, Y, or Z. The first part of the tree diagram has three branches labeled X, Y, and Z, with probabilities 0.4, 0.35, and 0.25, respectively. Event A is "the bulb is faulty," and the three probabilities are 0.01, 0.02, and 0.03. So we obtain the diagram of Figure 6.5.

From the Figure, the denominator is 0.0185. So

$$P(Z \mid A) = \frac{P(A \mid Z)P(Z)}{0.0185} = \frac{0.0075}{0.0185} = 0.676.$$

Practice Exercise. Repeat the above for Sample Problem 6.24.

Sample Problem 6.26. *The following table shows the proportion of people over* 18 *who are in various age categories, together with the probability that a person in a given category will vote in a given election. A vote is selected at random. What is the probability that the voter was from the* 18–24 *age group?*

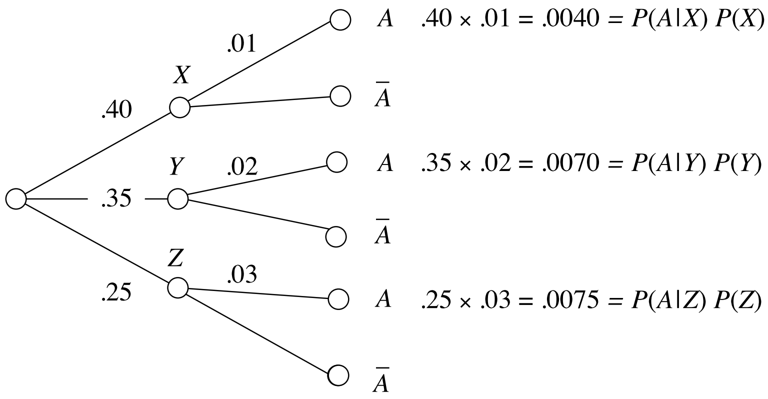


Fig. 6.5. Tree diagram for Sample Problem 6.25

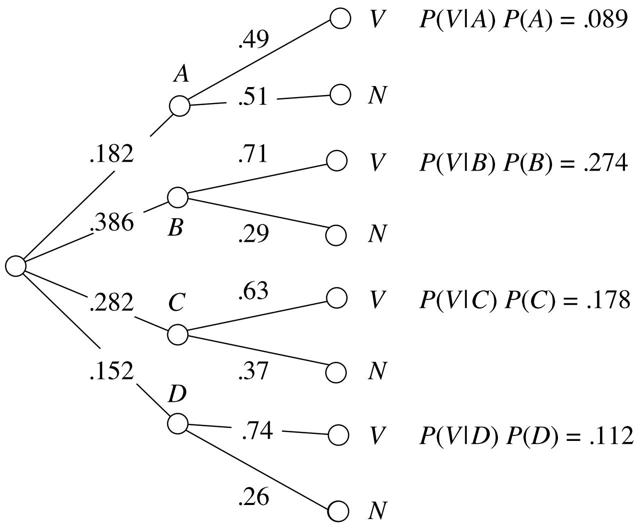


Fig. 6.6. Tree diagram for Sample Problem 6.26

Age	Proportion	Probability
A: 18–24	18.2%	0.49
B: 25–44	38.6%	0.71
C: 45–64	28.2%	0.63
D: 65+	15.2%	0.74

Solution. We use the diagram shown in Figure 6.6. (V and N mean "voter" and "nonvoter," respectively).

$$P(A \mid V) = \frac{0.089}{0.653} = 0.136.$$

	A	B	
X	0.004	0.396	0.4
Y	0.007	0.343	0.35
Z	0.0075	0.2425	0.25
	0.0185	0.9815	

Fig. 6.7. Box diagram of data for Sample Problem 6.23

Practice Exercise. What is the probability that the voter was 65 or older?

Box Diagrams

In many examples it is easiest to represent a two-stage experiment using a box diagram. The outcomes of the first experiment are used as labels for rows, and the outcomes of the second experiment are used as labels for columns. Where row A meets column B, we put a box containing $P(A \cap B)$. The sum of entries in row A, which equals $P(A)$, is written at the end of row A, and so on.

As an example, consider Sample Problem 6.23. The first experiment—determination of factory—has outcomes X, Y, and Z; the second experiment—"check: is it faulty"—has outcomes A (faulty) and B (okay). Then

$$P(A \cap X) = P(A \mid X)P(X) = 0.004,$$

$$P(A \cap Y) = P(A \mid Y)P(Y) = 0.007,$$

$$P(A \cap Z) = P(A \mid Z)P(Z) = 0.0075.$$

(See Figure 6.5 for these calculations.)

We can calculate $P(B \cap X)$ in various ways. For example, A and B are complements, so

$$P(A \cap X) + P(B \cap X) = P(X),$$

$$P(B \cap X) = P(X) - P(A \cap X)$$

$$= 0.4 - 0.004 = 0.396;$$

similarly $P(B \cap Y) = 0.343$, $P(B \cap Z) = 0.2425$. So the box diagram is as shown in Figure 6.7.

Sample Problem 6.27. *Use the box diagram in Figure 6.7 to calculate $P(Z \mid A)$ in Sample Problem 6.23.*

Solution. To find $P(Z \mid A)$, look at column A, which has a total of 0.0185. Then look at the (Z, A) entry 0.0075, and take the ratio

$$P(Z \mid A) = \frac{P(A \cap Z)}{P(Z)} = \frac{0075}{0.0185} = 0.406.$$

Practice Exercise. Produce a box diagram for the Practice Exercise following Sample Problem 6.23 and *use it* to find the probability that you have tossed the biased coin, given that it shows heads.

Medical Testing

Probability theory is often misused in the press. One particular area where probabilistic ideas are mentioned, but not properly analyzed, is in the discussion of tests for disease and drugs. This is a very important topic nowadays, when compulsory drug testing is becoming more common in everyday life and testing for diseases such as AIDS is also very important.

You may see in a newspaper that a certain test is "90% accurate." And you would most likely think this means that, if you test positive, there is a 90% probability that you have the disease in question. But this is not so, as the next example shows.

Sample Problem 6.28. *A test for a venereal disease is* 90% *accurate: if you have the disease, the probability is* 0.9 *that you will test positive; and if you do not have the disease, the probability is* 0.9 *that you will test negative. In the whole population,* 1 *person in* 100 *has the disease. If a given person tests positive, what is the chance that he has the disease?*

Solution. We shall use the abbreviations T (tests positive), N (tests negative), D (has the disease), and H (is healthy). We want to find $P(D \mid T)$. From the data, we know

$$\begin{aligned} P(T \mid D) &= 0.9, & P(N \mid D) &= 0.1, \\ P(T \mid H) &= 0.1, & P(N \mid H) &= 0.9, \\ P(D) &= 0.01, & P(H) &= 0.99. \end{aligned}$$

So

$$\begin{aligned} P(D \mid T) &= \frac{P(T \mid D)P(D)}{P(T \mid H)P(H) + P(T \mid D)P(D)} \\ &= \frac{(0.9)(0.01)}{(0.1)(0.99) + (0.9)(0.01)} \\ &= \frac{0.009}{0.099 + 0.009} = \frac{0.009}{0.108} = 0.083. \end{aligned}$$

So, even if you test positive, it is still most unlikely that you have the disease. This example can be represented by the following box diagram:

	P	N	
D	0.009	0.001	0.01
H	0.099	0.891	0.99
	0.108	0.892	

Practice Exercise. A disease has infected 5% of the population. The test is 95% effective—95% of those with the disease test positive, and 95% of those without the disease test negative. If your test shows a positive result, what is the probability that you have the disease?

The Monty Hall Problem

Marilyn vos Savant's *Ask Marilyn* column in *Parade* Magazine for September 9th, 1990, contained a puzzle that generated a lot of interest. It is, in fact, a version of an older problem called the *Monty Hall Problem*, named for the game show host.

> The climax of a TV game show is run as follows. The contestant is given a choice of three numbered doors. Behind one closed door is a valuable new car; behind each of the others is a nearly worthless goat. The contestant is to choose a door, and wins the prize behind it. However, after the choice is announced but before the door is opened, the host opens one of the other two doors (not the one she chose) and reveals a goat. (Of course the host knows where the car is.) He then asks, "Do you want to stay with your original choice? Or would you rather switch to the third door?"

Well, should the contestant stay or switch? Or does it not matter?

Before analyzing the problem, we need to agree on three points. First, the car is placed behind the doors at random, so that on any given night the chance that it is behind any particular door is $\frac{1}{3}$. Second, the game always proceeds in the same way: the host *always* opens a door to show a goat, then offers the switch. Third, on those nights when the contestant's first choice is the door with the car, there is an equal chance that the host will open either of the other two doors.

Without loss of generality, let us suppose the contestant chooses door 1 and the host opens door 2. We write $C1$, $C2$, and $C3$ as abbreviations for "the car is behind door 1," "the car is behind door 2," and "the car is behind door 3," and $H2$, $H3$ for "the host opens door 2," "the host opens door 3." (He cannot open door 1.) Then what we want to know is

$$P(C3 \mid H2) > P(C1 \mid H2)?$$

If so, the contestant should switch, otherwise not.

We know the probabilities of the host's actions, given the position of the car. If the car is behind door 1, she is equally likely to open either door, so

$$P(H2 \mid C1) = P(H3 \mid C1) = \frac{1}{2}.$$

In the other cases she must open the remaining "goat" door, so

$$P(H2 \mid C2) = P(H3 \mid C3) = 0,$$

$$P(H2 \mid C3) = P(H3 \mid C2) = 1.$$

Moreover we know $P(C1) = P(C2) = P(C3) = \frac{1}{3}$.

Now the calculation of the probabilities is a simple application of Bayes' formula. First

$$\begin{aligned} P(H2) &= P(H2|C1)P(C1) + P(H2|C2)P(C2) + P(H2|C3)P(C3) \\ &= \left(\frac{1}{2}\right)\left(\frac{1}{3}\right) + (0)\left(\frac{1}{3}\right) + (1)\left(\frac{1}{3}\right) = \frac{1}{2}, \end{aligned}$$

and similarly $P(H3) = \frac{1}{2}$. So

$$P(C1 \mid H2) = \frac{P(H2 \mid C1)P(C1)}{P(H2)} = \frac{(\frac{1}{2})(\frac{1}{3})}{\frac{1}{2}} = \frac{1}{3},$$

$$P(C3 \mid H2) = \frac{P(H2 \mid C3)P(C3)}{P(H2)} = \frac{(1)(\frac{1}{3})}{(\frac{1}{2})} = \frac{2}{3}.$$

So the odds are 2 to 1 in favor of switching. You may find this very surprising—intuitively, you might argue that "the car was equally likely to be behind any of the doors, so the probabilities are still equal," but the host's choice of doors has actually given you some information.

The most important assumption in this discussion is that when the contestant has chosen correctly, the host is equally likely to open either door. Suppose this were not true—for example, suppose he always chooses the lowest numbered available door. Then if the contestant chooses door 1 and the host opens door 3, she should always switch; but if he opens door 2, there is no advantage (or disadvantage) in switching.

This problem is particularly well suited to analysis by box diagrams. Assuming that the contestant chooses door 1, then the fact that the car is equally likely to be behind any door means that the diagram looks like

	$C1$	$C2$	$C3$	
$H2$	*	*	*	*
$H3$	*	*	*	*
	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	

(where asterisks represent the numbers that have not yet been determined). Since the host never opens the door that hides the car, the $(H2, C2)$ and $(H3, C3)$ entries must be zero, and to make the column sums come out right we have

	$C1$	$C2$	$C3$	
$H2$	*	0	$\frac{1}{3}$	*
$H3$	*	$\frac{1}{3}$	0	*
	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	

Assuming the host chooses at random when the car is behind door 1, we can make the two missing numbers in column $C1$ equal, and we have

	$C1$	$C2$	$C3$	
$H2$	$\frac{1}{6}$	0	$\frac{1}{3}$	$\frac{1}{2}$
$H3$	$\frac{1}{6}$	$\frac{1}{3}$	0	$\frac{1}{2}$
	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	

When the host opens door 2, we need only look at the $H2$ row to see that the odds are 2 to 1 ($\frac{1}{3}$ to $\frac{1}{6}$) in favor of switching, and similarly if he opens door 3.

Exercises 6.5

1. The events E and F satisfy $P(E) = 0.6$, $P(F \mid E) = 0.5$, and $P(F \mid \overline{E}) = 0.75$. Find $P(E \mid F)$ and $P(\overline{E} \mid F)$.

2. The events E and F satisfy $P(E) = 0.85$, $P(F \mid E) = 0.8$, and $P(F \mid \overline{E}) = 0.8$. Find $P(E \mid F)$ and $P(\overline{E} \mid F)$.

3. One experiment has the possible outcomes A, B, and C; a second experiment has the possible outcomes E and F. Find $P(A \mid E)$, $P(B \mid E)$, $P(C \mid E)$, $P(A \mid F)$, $P(B \mid F)$, and $P(C \mid F)$ in each of the following cases:

(i) $P(A) = 0.4$, $P(B) = 0.4$, $P(E \mid A) = 0.25$, $P(E \mid B) = 0.75$, and $P(E) = 0.6$.

(ii) $P(A) = 0.25$, $P(B) = 0.25$, $P(E \mid A) = 0.4$, $P(E \mid B) = 0.4$, and $P(E) = 0.5$.

(iii) $P(A) = \frac{3}{5}$, $P(B) = \frac{1}{10}$, $P(E \mid A) = \frac{1}{3}$, $P(E \mid B) = 0$, and $P(E) = \frac{1}{2}$.

4. One experiment has the possible outcomes A, B, and C; a second experiment has the possible outcomes E and F. Find $P(A \mid E)$, $P(B \mid E)$, $P(C \mid E)$, $P(A \mid F)$, $P(B \mid F)$, and $P(C \mid F)$ in each of the following cases:

(i) $P(A) = \frac{3}{10}$, $P(B) = \frac{1}{2}$, $P(E \mid A) = \frac{2}{3}$, $P(F \mid B) = \frac{2}{5}$, and $P(F \mid C) = \frac{1}{2}$.

(ii) $P(A) = 0.25$, $P(B) = 0.25$, $P(E \mid A) = 0.6$, $P(F \mid B) = 0.6$, and $P(F \mid C) = 0.5$.

(iii) $P(A) = 0.5$, $P(B) = 0.3$, $P(E \mid A) = 0.6$, $P(F \mid B) = 0$, and $P(F \mid C) = 0.5$.

5. An automobile dealership finds that 4% of their customers default on payments, so that the car must be repossessed. On analyzing the records, it is found that among those who did not default, 40% made a large downpayment ($2000 or more), while only 10% of those who later defaulted made a large downpayment.

(i) Suppose a customer makes a large downpayment. What is the probability that he will default on payments?

(ii) If a customer makes a small downpayment, what is the probability that he will default on payments?

6. A builder buys tiles from two companies, A and B; he gets 80% from A and 20% from B. He finds that 98% of the tiles he gets from A are undamaged, while 96% of those from B are undamaged. If he finds a damaged tile, what is the probability that it came

(i) from A?

(ii) from B?

7. In manufacturing metal office equipment, your company uses nuts supplied by three companies, A, B, and C. Company A supplies 30%, company B supplies 45%, and company C supplies 25%. It is known that on average the following percentages of the nuts are defective: 1% of those from A, 1.5% of those from B, and 0.5% of those from C.

(i) If a nut is selected at random, what is the probability that it is defective?

(ii) If a nut is selected at random and is found to be defective, what is the probability that it was made by company B?

8. Suppose the shipping department of a company has three workers who prepare shipping labels. They prepare 60%, 30%, and 10% of the labels, respectively. The respective percentages of errors are 3%, 5%, and 10%. Find the probability that an incorrect label is due to the first person. Also find the probability that an incorrect label is due to each of the other persons.

9. In a certain city, it is found that equally many people have fair and dark hair. A survey shows that 20% of people with dark hair and 40% of people with fair hair have blue eyes. A person is chosen at random from the population and is found to have blue eyes. What is the probability that this person has fair hair?

10. Box 1 contains three red pens and four blue pens. Box 2 contains four red pens, one green pen, and three black pens. A pen is chosen from Box 1 and then it is placed in Box 2. Then a box is chosen at random and a pen is selected.

(i) What is the probability that the pen is blue?

(ii) If the pen is blue, what is the probability it originally came from Box 1?

(*Hint:* Work this problem as though there were two types of blue pen: light blue initially in Box 1, dark blue initially in Box 2.)

11. In Exercise 10, suppose Box 2 is the one from which a pen in finally selected.

(i) If the pen drawn is blue, what is the probability that the pen drawn earlier from Box 1 was red?

(ii) If the pen drawn is red, what is the probability that the pen drawn earlier from Box 1 was red?

12. In a manufacturing plant, three machines, A, B, and C, produce 50%, 35%, and 15%, respectively, of the total production. The quality control department of the company has determined that 1% of the items produced by Machine A and 2% of the items produced by each of Machines B and C are defective. If an item is selected at random and found to be defective, what is the probability that it was produced by Machine B?

13. In a factory, 30% of the workers smoke. It is found that smokers have three times the absentee rate of other workers. If a worker is absent, what is the probability that he is a smoker?

14. An auto insurance company classifies 20% of its drivers as good risks, 60% as medium risks, and 20% as bad risks (called classes A, B, and C, respectively). The probability of at least one accident in a given year is 1% for class C, 0.5% for class B, and 0.1% for class A. If one of their insureds has an accident this year, what is the probability that he is a class B driver?

15. A class contains 60% women. It is found that 12% of the men students and 7% of the women students are left-handed. A student is chosen at random. If the student is left-handed, what is the probability that the student is male?

16. A school tax proposition is submitted to voters. The voters' registered party affiliation as a percentage of all voters, and the percentage of each group who voted in favor of the proposition, are as follows:

Party	Registration	In Favor
Democrat	40	70
Republican	40	20
Independent	10	80
Other	10	50

A voter is selected at random.

(i) What is the probability that she voted in favor of the proposition?

(ii) If she voted in favor of the proposition, what is the probability that she is a Democrat?

(iii) If she voted against the proposition, what is the probability that she is not a Republican?

17. Among one strain of cattle in Texas, it is estimated that 30% of the bulls suffer from Ruckel's disease. A blood test has been developed to detect the disease. If a bull has the disease, then the test gives a positive response 80% of the time. If a bull does not have the disease, the test is positive in 10% of cases. If a bull is tested for the disease and the test is positive, what is the probability that he really *does not* have the disease?

18. A test is positive for 94% of subjects who have a certain disease and negative for 96% of those who do not have the disease. If 4% of the population suffers from the disease, then

(i) if the test is positive, what is the probability that the patient actually has the disease?

(ii) if the test is negative, what is the probability that the patient does not have the disease?

19. A smallpox vaccine produces total immunity in 95% of cases. If the vaccination does not produce immunity, then it has no effect—the person's probability of catching smallpox is the same as that of an unvaccinated person. Suppose 30% of the population have been vaccinated. If a person contracts smallpox, what is the probability that she had been vaccinated?

20. A laboratory test for a particular disease tests positive 95% of the time when a person has the disease, and tests negative 98% of the time when the person does not have the disease. In general 0.1% of the population has the disease.

(i) What is the probability that a randomly selected person who tested positive did not have the disease?

(ii) What is the probability that a randomly selected person who tested negative actually had the disease?

21. Suppose the "Monty Hall" game involved four doors, not three. What is the probability of winning if you switch? What is the probability of winning if you do not switch?

22. Suppose there are three cards in a hat. One is red on both sides, one black on both sides, and the third has one side red and one side black. One card is withdrawn, and one side is observed to be red. What is the chance that the other side is red?