Probability Theory

3

3

1 INTRODUCTION

We often hear statements of the following kind: "It is likely to rain today," "I have a fair chance of passing this course," "There is an even chance that a coin will come up heads," etc. In each case our statement refers to a situation in which we are not certain of the outcome, but we express some degree of confidence that our prediction will be verified. The theory of probability provides a mathematical framework for such assertions.

Consider an experiment whose outcome is not known. Suppose that someone makes an assertion p about the outcome of the experiment, and we want to assign a probability to p. When statement p is considered in isolation, we usually find no natural assignment of probabilities. Rather, we look for a method of assigning probabilities to all conceivable statements concerning the outcome of the experiment. At first this might seem to be a hopeless task, since there is no end to the statements we can make about the experiment. However, we are aided by a basic principle:

Fundamental Assumption Any two equivalent statements will be assigned the same probability.

As long as there are a finite number of logical possibilities, there are only a finite number of truth sets, and hence the process of assigning probabilities is a finite one. We proceed in three steps: (1) we first determine $\mathcal{U}$, the possibility set, that is, the set of all logical possibilities; (2) to each subset X of $\mathcal{U}$ we assign a number called the measure $m(X)$; (3) to each statement p we assign $m(P)$, the measure of its truth set, as a probability. The probability of statement p is denoted by $\Pr[p]$.

The first step, that of determining the set of logical possibilities, is one that we considered in the previous chapters. It is important to recall that there is no unique method for analyzing logical possibilities. In a given problem we may arrive at a very fine or a very rough analysis of possibilities, causing $\mathcal{U}$ to have many or few elements.

Having chosen $\mathcal{U}$, the next step is to assign a number to each subset X of $\mathcal{U}$, which will in turn be taken to be the probability of any statement having truth set X. We do this in the following way.

Assignment of a Measure Assign a positive number (weight) to each element of $\mathcal{U}$, so that the sum of the weights assigned is 1. Then the measure of a set is the sum of the weights of its elements. The measure of the set $\mathcal{E}$ is 0.

EXAMPLE 1 An ordinary die is thrown. What is the probability that the number which turns up is less than four? Here the possibility set is $\mathcal{U} = \{1, 2, 3, 4, 5, 6\}$. The symmetry of the die suggests that each face should have the same probability of turning up. To make this so, we assign weight $\frac{1}{6}$ to each of the outcomes.

In applications of probability to scientific problems, the analysis of the logical possibilities and the assignment of measures may depend upon factual information and hence can best be done by the scientist making the application.

Once the weights are assigned, to find the probability of a particular statement we must find its truth set and find the sum of the weights assigned to elements of the truth set. This problem, which might seem easy, can often involve considerable mathematical difficulty. The development of techniques to solve this kind of problem is the main task of probability theory.

EXAMPLE 1 (continued) For the case of throwing an ordinary die we have already assigned equal weights to each outcome. Let us consider statements relative to $\mathcal{U} = \{1, 2, 3, 4, 5, 6\}$. The truth set of the statement "The number that turns up is less than four" is $\{1, 2, 3\}$. Hence the probability of this statement is $\frac{3}{6} = \frac{1}{2}$, the sum of the weights of the elements in its truth set. Similarly, the truth set of the statement "The number that turns up is odd or is less than four" is $\{1, 2, 3, 5\}$. Hence the probability of this statement is $\frac{4}{6} = \frac{2}{3}$, which again is the sum of the weights assigned to elements in its truth set.

EXAMPLE 2 A man attends a race involving three horses A, B, and C. He feels that A and B have the same chance of winning but that A (and hence also B) is twice as likely to win as C is. What is the probability that A or C wins? We take as $\mathcal{U}$ the set {A, B, C}. If we were to assign weight a to the outcome C, then we would assign weight $2a$ to each of the outcomes A and B. Since the sum of the weights must be 1, we have $2a + 2a + a = 1$, or $a = \frac{1}{5}$. Hence we assign weights $\frac{2}{5}, \frac{2}{5}, \frac{1}{5}$ to the outcomes A, B, and C, respectively. The truth set of the statement "Horse A or C wins" is {A, C}. The sum of the weights of the elements of this set is $\frac{2}{5} + \frac{1}{5} = \frac{3}{5}$. Hence the probability that A or C wins is $\frac{3}{5}$.

EXERCISES

1. Briefly explain the difference between the terms "weight," "measure," and "probability".
2. Let $\mathcal{U} = \{a, b, c\}$. Assign weights to the three elements so that no two have the same weight, and find the measures of the eight subsets of $\mathcal{U}$.
3. Give the possibility set $\mathcal{U}$ for each of the following experiments:
   (a) A number from 1 to 7 is chosen at random.
   (b) A and B play a game of chess.
   (c) A student is asked for the month in which his birthday falls.
   (d) A die with all faces having the number six is thrown.
4. For which of the cases in Exercise 3 might it be appropriate to assign the same weight to each outcome?
5. In an election Jones has probability $\frac{3}{8}$ of winning, Smith has probability $\frac{1}{8}$, and Black has probability $\frac{1}{2}$.
   (a) Construct $\mathcal{U}$.
   (b) Assign weights.
   (c) Find the measures of the eight subsets.
   (d) Give a pair of nonequivalent predictions which have the same probability.
6. A die is loaded in such a way that the probability of each face is proportional to the number of dots on that face. (For instance, a six is 3 times as probable as a two.) What is the probability of getting an even number in one throw? [*Ans.* $\frac{4}{7}$.]
7. The owner of a certain hardware store places a sign stating "Back in 15 minutes" on his door at noon when he goes to lunch. Customers find that when this sign is posted the probabilities are .4 that he is back within 10 minutes, .45 that he returns in more than 10 but less than 20 minutes, and .145 that he returns after 20 minutes or more have elapsed.
   (a) What is the probability that he returns within 20 minutes? [*Ans.* .85.]
   (b) What is the probability that he takes the rest of the day off and does not return at all?
   (c) Is it possible to determine the probability that he returns within 5 minutes? [*Ans.* No.]
8. If a coin is thrown three times, list the eight possibilities for the outcomes of the three successive throws. A typical outcome can be written (HTH). Determine a probability measure by assigning an equal weight to each outcome. Find the probabilities of the following statements:
   r: The number of heads that occur is greater than the number of tails. [*Ans.* $\frac{1}{2}$.]

s: Exactly two heads occur. [*Ans.* $\frac{3}{8}$.]
t: The same side turns up on every throw. [*Ans.* $\frac{1}{4}$.]

9. For the statements given in Exercise 8, which of the following equalities are true?
(a) $\Pr[r \vee s] = \Pr[r] + \Pr[s]$.
(b) $\Pr[s \vee t] = \Pr[s] + \Pr[t]$.
(c) $\Pr[r \vee \sim r] = \Pr[r] + \Pr[\sim r]$.
(d) $\Pr[r \vee t] = \Pr[r] + \Pr[t]$.

10. Which of the following pairs of statements (see Exercise 8) are inconsistent? (Recall that two statements are inconsistent if their truth sets have no element in common.)
(a) r, s. **(b)** s, t.
(c) $r, \sim r$. **(d)** r, t. [*Ans.* (b) and (c).]

11. State a property which is suggested by Exercises 9 and 10.

12. A number is chosen from the set $\{1, 2, 3\}$. If weights have been assigned to the three outcomes such that Pr [a 1 or 2 is chosen] $= \frac{3}{5}$ and Pr [a 2 or 3 is chosen] $= \frac{2}{3}$, find the weights. [*Ans.* $\frac{1}{3}, \frac{4}{15}, \frac{2}{5}$.]

13. Repeat Exercise 12 for each of the following cases
(a) Pr [a 1 or 2 is chosen] $= \frac{2}{5}$ and
Pr [a 2 or 3 is chosen] $= \frac{2}{5}$.
(b) Pr [a 2 is chosen] $= \frac{1}{3}$, and
Pr [a 1 or 2 is chosen] $= \frac{1}{2} \cdot$ Pr [a 2 or 3 is chosen].

2 PROPERTIES OF A PROBABILITY MEASURE

Before studying special probability measures, we shall consider some general properties of such measures which are useful in computations and in the general understanding of probability theory.

Three basic properties of a probability measure are

(A) $m(X) = 0$ if and only if $X = \mathcal{E}$.
(B) $0 \le m(X) \le 1$ for any set X.
(C) For two sets X and Y,

$$m(X \cup Y) = m(X) + m(Y)$$

if and only if X and Y are disjoint, i.e., have no elements in common.

The proofs of properties (A) and (B) are left as an exercise (see Exercise 16). We shall prove (C).

We observe first that $m(X) + m(Y)$ is the sum of the weights of the elements of X added to the sum of the weights of Y. If X and Y are disjoint, then the weight of every element of $X \cup Y$ is added once and only once, and hence $m(X) + m(Y) = m(X \cup Y)$.

Assume now that X and Y are not disjoint. Here the weight of every element contained in both X and Y—i.e., in $X \cap Y$—is added twice in the sum $m(X) + m(Y)$. Thus this sum is greater than $m(X \cup Y)$ by an amount

$m(X \cap Y)$. By (A) and (B), if $X \cap Y$ is not the empty set, then $m(X \cap Y) > 0$. Hence in this case we have $m(X) + m(Y) > m(X \cup Y)$. Thus if X and Y are not disjoint, the equality in (C) does not hold. Our proof shows that in general we have

(C′) For any two sets X and Y,

$$m(X \cup Y) = m(X) + m(Y) - m(X \cap Y).$$

Since the probabilities for statements are obtained directly from the probability measure $m(X)$, any property of $m(X)$ can be translated into a property about the probability of statements. For example, the above properties become, when expressed in terms of statements,

(a) $\Pr[p] = 0$ if and only if p is logically false.
(b) $0 \leq \Pr[p] \leq 1$ for any statement p.
(c) The equality

$$\Pr[p \vee q] = \Pr[p] + \Pr[q]$$

holds if and only if p and q are inconsistent.
(c′) For any two statements p and q,

$$\Pr[p \vee q] = \Pr[p] + \Pr[q] - \Pr[p \wedge q].$$

Another property of a probability measure which is often useful in computation is

(D) $m(\tilde{X}) = 1 - m(X)$,

or, in the language of statements,

(d) $\Pr[\sim p] = 1 - \Pr[p]$.

The proofs of (D) and (d) are left as an exercise (see Exercise 17).

It is important to observe that our probability measure assigns probability 0 only to statements which are logically false, i.e., which are false for every logical possibility. Hence, a prediction that such a statement will be true is certain to be wrong. Similarly, a statement is assigned probability 1 if and only if it is true in every case, i.e., logically true. Thus the prediction that a statement of this type will be true is certain to be correct. (While these properties of a probability measure seem quite natural, it is necessary, when dealing with infinite possibility sets, to weaken them slightly. We consider in this book only finite possibility sets.)

We shall now discuss the interpretation of probabilities that are not 0 or 1. We shall give only some intuitive ideas that are commonly held concerning probabilities. While these ideas can be made mathematically more precise, we offer them here only as a guide to intuitive thinking.

Suppose that, relative to a given experiment, a statement has been assigned probability p. From this it is often inferred that if a sequence of such experiments is performed under identical conditions, the fraction of experiments which yield outcomes making the statement true would be approximately p. The mathematical version of this is the "law of large numbers" of probability theory (which will be treated in Section 9). In cases where there is no natural way to assign a probability measure, the probability of a statement is estimated experimentally. A sequence of experiments is performed and the fraction of the experiments which make the statement true is taken as the approximate probability for the statement.

A second and related interpretation of probabilities is concerned with betting. Suppose that a certain statement has been assigned probability p. We wish to offer a bet that the statement will in fact turn out to be true. We agree to give r dollars if the statement does not turn out to be true, provided that we receive s dollars if it does turn out to be true. What should r and s be to make the bet fair? If it were true that in a large number of such bets we would win s a fraction p of the times and lose r a fraction $1 - p$ of the time, then our average winning per bet would be $sp - r(1 - p)$. To make the bet fair we should make this average winning 0. This will be the case if $sp = r(1 - p)$ or if $r/s = p/(1 - p)$. Notice that this determines only the ratio of r to s. Such a ratio, written $r{:}s$, is said to give *odds* in favor of the statement.

Definition The *odds* in favor of an outcome are $r{:}s$ (r to s), if the probability of the outcome is p, and $r/s = p/(1 - p)$. Any two numbers having the required ratio may be used in place of r and s. Thus 6:4 odds are the same as 3:2 odds.

EXAMPLE Assume that a probability of $\frac{3}{4}$ has been assigned to a certain horse winning a race. Then the odds for a fair bet would be $\frac{3}{4}{:}\frac{1}{4}$. These odds could be equally well written as 3:1, 6:2 or 12:4, etc. A fair bet would be to agree to pay \$3 if the horse loses and receive \$1 if the horse wins. Another fair bet would be to pay \$6 if the horse loses and win \$2 if the horse wins.

EXERCISES

1. Let p and q be statements such that $\Pr[p \vee q] = \frac{3}{4}$, $\Pr[p] = \frac{2}{3}$, and $\Pr[\sim q] = \frac{3}{4}$. Find $\Pr[p \wedge q]$. [*Ans.* $\frac{1}{6}$.]
2. Using the results of Exercise 1, find $\Pr[\sim p \vee \sim q]$.
3. Let p and q be statements such that $\Pr[p] = \frac{1}{2}$ and $\Pr[q] = \frac{2}{3}$. Are p and q consistent? [*Ans.* Yes.]
4. Show that, if $\Pr[p] + \Pr[q] > 1$, then p and q are consistent.
5. A student is worried about his grades in English and Art. He estimates that the probability of passing English is .4, that he will pass at least one course with probability .6, but that he has only probability .1 of

passing both courses. What is the probability that he will pass Art? [*Ans.* .3.]

6. Given that a school has grades A, B, C, D, and F, and that a student has probability .9 of passing a course, and .6 of getting a grade lower than B, what is the probability that he will get a C or D? [*Ans.* $\frac{1}{2}$.]

7. State what odds a person should give on the following events:

(a) That a card chosen at random from a 52-card deck is on ace.

(b) That a four turns up when a dice is thrown.

(c) That a coin which is flipped twice comes up heads both times.

8. Prove that if the odds in favor of a given statement are $r:s$, then the probability that the statement will be true is $r/(r + s)$.

9. Using the result of Exercise 8 and the definition of "odds," show that if the odds are $r:s$ that a statement is true, then the odds are $s:r$ that it is false.

10. A man is willing to give 3:1 odds that the Democratic candidate will win the next presidential election. What must the probability of a Democratic victory be to make this a fair bet?

11. An American roulette wheel contains 38 slots (18 red, 18 black, and 2 green). What are the odds that red will turn up on a given spin?

12. A man offers 3:2 odds that A will occur, and 1:2 odds that B will occur. If he knows that A and B cannot both occur, what odds should he give that A or B will occur? [*Ans.* 14:1.]

13. Suppose now the man offers 2:3 odds that A will occur, and 2:1 odds that B will occur. Again, he knows that A and B cannot both occur. What odds should he give that A or B will occur?

14. A man offers to bet "dollars to doughnuts" that a certain event will take place. Assuming that a doughnut costs a dime, what must the probability of the event be for this to be a fair bet? [*Ans.* $\frac{10}{11}$.]

15. If X and Y are two sets such that X is a subset of Y, prove that $m(X) \leq m(Y)$. Use this to prove that if p implies q then $\Pr[p] \leq \Pr[q]$.

16. Show from the definition of a probability measure that properties (A) and (B) of the text are true.

17. Prove property (D) of the text. Why does property (d) follow from this property?

18. Let X, Y, and Z be any three sets, the let m be any probability measure. Prove in two ways that $m(X \cup Y \cup Z) = m(X) + m(Y) + m(Z) - m(X \cap Y) - m(X \cap Z) - m(Y \cap Z) + m(X \cap Y \cap Z)$. Use a Venn diagram for the first proof. For the second, notice that $X \cup Y \cup Z = (X \cup Y) \cup Z$ and use property (C′) of the text.

19. Suppose we assume that X, Y, and Z are *pairwise disjoint*—i.e., that $X \cap Y = X \cap Z = Y \cap Z = \mathcal{E}$. Show that $m(X \cup Y \cup Z) = m(X) + m(Y) + m(Z)$.

20. Suppose that we make the assumption (weaker than that in Exercise 19) that $X \cap Y \cap Z = \mathcal{E}$. Show by example that it is not necessarily the case that $m(X \cup Y \cup Z) = m(X) + m(Y) + m(Z)$.

21. Show that $m(X \cup Y \cup Z) \leq m(X) + m(Y) + m(Z)$.
22. Translate the result of Exercise 18 into a result concerning three statements p, q, and r.
23. Suppose $\Pr[p \wedge q] = \Pr[p \wedge r] = \Pr[q \cap r] = 0$ and $\Pr[p \vee q \vee r] = 1$. What can be said about the statements p, q, and r?
24. Suppose a card is drawn from a deck of playing cards. Let p be the statements "The card is an honor card. [i.e., an ace, king, queen, jack, or ten]" let q be the statement "The card is a spade," and let r be the statement "The card is either a heart or the king of clubs." Then $\Pr[p] = \frac{5}{13}$, $\Pr[q] = \frac{1}{4}$, $\Pr[r] = \frac{7}{26}$, $\Pr[p \wedge q] = \frac{5}{52}$, and $\Pr[p \wedge r] = \frac{3}{26}$. Find $\Pr[p \vee q \vee r]$. What is the probability that the card is neither a spade, a heart, nor an honor card?
25. The following is an alternative proof of property (C′) of the text. Give a reason for each step.
    (a) $X \cup Y = (X \cap \tilde{Y}) \cup (X \cap Y) \cup (Y \cap \tilde{X})$.
    (b) $m(X \cup Y) = m(X \cap \tilde{Y}) + m(X \cap Y) + m(\tilde{X} \cap Y)$.
    (c) $m(X \cup Y) = m(X) + m(Y) - m(X \cap Y)$.
26. Two women, A and B, go out to lunch. If the probability that A's check is exactly \$3 is .25, the probability that B's check is exactly \$3 is .35, and the probability that the larger of the two checks is exactly \$3 is .05, what is the probability that the smaller check is exactly \$3? [*Hint:* Enumerate the logical possibilities for the checks, and see which ones correspond to the quantities given above.] [*Ans.* .55.]

3 THE EQUIPROBABLE MEASURE

We have already seen several examples where it was natural to assign the same weight to all possibilities in determining the appropriate probability measure. The probability measure determined in this manner is called the *equiprobable measure.* The measure of sets in the case of the equiprobable measure has a very simple form. In fact, if $\mathcal{U}$ has n elements and if the equiprobable measure has been assigned, then for any set X, $m(X)$ is r/n, where r is the number of elements in the set X. This is true since the weight of each element in X is $1/n$, and hence the sum of the weights of elements of X is r/n.

The particularly simple form of the equiprobable measure makes it easy to work with. In view of this, it is important to observe that a particular choice for the set of possibilities in a given situation may lead to the equiprobable measure, while some other choice will not. For example, consider the case of two throws of an ordinary coin. Suppose that we are interested in statements about the number of heads which occur. If we take for the possibility set the set $\mathcal{U} = \{\text{HH, HT, TH, TT}\}$ then it is reasonable to assign the same weight to each outcome, and we are led to the equiprobable measure. If, on the other hand, we were to take as possible outcomes the set $\mathcal{U} = \{\text{no H, one H, two H}\}$, it would not be natural to assign the same weight to each outcome, since one head can occur in two different ways, while each of the other possibilities can occur in only one way.

EXAMPLE 1 Suppose that we throw two ordinary dice. Each die can turn up a number from 1 to 6; hence there are $6 \cdot 6$ possibilities. We assign weight $\frac{1}{36}$ to each possibility. A prediction that is true in j cases will then have probability $j/36$. For example, "The sum of the dice is 5" will be true if we get $1 + 4$, $2 + 3, 3 + 2$, or $4 + 1$, that is, the sum can be 5 in four different ways. Hence the probability that the sum of the dice is 5 is $\frac{4}{36} = \frac{1}{9}$. The sum can be 12 in only one way, $6 + 6$. Hence the probability that the sum is 12 is $\frac{1}{36}$.

EXAMPLE 2 Suppose that two cards are drawn successively from a deck of cards. What is the probability that both are hearts? There are 52 possibilities for the first card, and for each of these there are 51 possibilities for the second. Hence there are $52 \cdot 51$ possibilities for the result of the two draws. We assign the equiprobable measure. The statement "The two cards are hearts" is true in $13 \cdot 12$ of the $52 \cdot 51$ possibilities. Hence the probability of this statement is $13 \cdot 12/52 \cdot 51 = \frac{1}{17}$.

EXAMPLE 3 Assume that, on the basis of a predictive index applied to students A, B, and C when entering college, it is predicted that after four years of college the scholastic record of A will be the highest, C the second highest, and B the lowest of the three. Suppose, in fact, that these predictions turn out to be exactly correct. If the predictive index has no merit at all and hence the predictions amount simply to guessing, what is the probability that such a prediction will be correct? There are $3! = 6$ orders in which the men might finish. If the predictions were really just guessing, then we would assign an equal weight to each of the six outcomes. In this case the probability is reasonably large, we would hesitate to conclude that the predictive index is in fact useful on the basis of this one experiment. Suppose, on the other hand, it predicted the order of six men correctly. Then a similar analysis would show that, by guessing, the probability is $1/6! = 1/720$ that such a prediction would be correct. Hence, we might conclude here that there is strong evidence that the index has some merit.

EXERCISES

1. A letter is chosen at random from the word "probability." What is the probability that it is a b? That it is a vowel? [*Ans.* $\frac{2}{11}$; $\frac{4}{11}$.]
2. A card is drawn at random from a deck of playing cards.
   (a) What is the probability that it is either a heart or a king but not both? [*Ans.* $\frac{15}{52}$.]
   (b) What is the probability that it is an honor card (ten, jack, queen, king, ace) and either a club or a spade?

3. An office building with ten floors has a broken elevator which lets people off at random floors. If a man starting on the first floor wants to go to the fourth floor,
   (a) What is the probability that he ends up on the floor he wants? [*Ans.* $\frac{1}{10}$.]
   (b) What is the probability that he ends up no closer to the fourth floor than when he started?
4. A word is chosen at random from the set of words $\mathcal{U} = \{$men, bird, ball, field, book$\}$. Let p, q, and r be the statements:
   p: The word has two vowels.
   q: The first letter of the word is b.
   r: The word rhymes with cook.
   Find the probability of the following statements:
   (a) p.
   (b) q.
   (c) r.
   (d) $p \vee q$.
   (e) $\sim(p \wedge q) \wedge r$.
   (f) $p \rightarrow q$. [*Ans.* $\frac{4}{5}$.]
   (g) $\sim p \leftrightarrow q$.
5. A single die is thrown. Find the probability that
   (a) An odd number turns up.
   (b) The number which turns up is greater than two.
   (c) A seven turns up.
6. A single die is thrown twice. What value for the sum of the two outcomes has the highest probability? What value or values of the sum has the lowest probability of occurring?
7. In Exercise 6, what value or values for the product of the two outcomes has the highest probability of occurring?
8. A certain college has 500 students and it is known that

   250 read French.
   200 read German.
   100 read Russian.
   55 read French and Russian.
   35 read German and Russian.
   60 read German and French.
   20 read all three languages.

   If a student is chosen at random from the school, what is the probability that the student
   (a) Reads two and only two languages? [*Ans.* $\frac{9}{50}$.]
   (b) Reads at least one language?
9. The letters of the word "connect" are scrambled and placed in a random order.

(a) What is the probability that they still spell "connect"?
[*Ans.* $\frac{1}{1260}$.]

(b) What is the probability that they are arranged in alphabetical order?
[*Ans.* $\frac{1}{1260}$.]

(c) What is the probability that the first and last letters are the same?
[*Ans.* $\frac{2}{21}$.]

10. Suppose that three people enter a restaurant which has a row of six seats. If they choose their seats at random, what is the probability that they sit with no seats between them? What is the probability that none of them is sitting in a seat next to somebody else?

11. Find the probability that a bridge hand will have suits of

(a) 5, 4, 3, and 1 cards.

$$\textit{Ans.}\ \frac{4!\binom{13}{5}\binom{13}{4}\binom{13}{3}\binom{13}{1}}{\binom{52}{13}} \cong .129.$$

(b) 6, 4, 2, and 1 cards. [*Ans.* .047.]

(c) 4, 4, 3, and 2 cards. [*Ans.* .216.]

(d) 4, 3, 3, and 3 cards. [*Ans.* .105.]

12. There are $\binom{52}{13} = 6.35 \times 10^{11}$ possible bridge hands. Find the probability that a bridge hand dealt at random will be all of one suit. Estimate roughly the number of bridge hands dealt in the entire country in a year. Is it likely that a hand of all one suit will occur sometime during the year in the United States?

13. If ten people are seated at a circular table, what are the probabilities that

(a) A particular pair of people are seated next to each other?
[*Ans.* $\frac{2}{9}$.]

(b) Three particular people are sitting together?

14. A contestant on a TV quiz show is shown four pieces of merchandise and is given a list of four prices. She wins the grand prize if she matches each piece of merchandise with its correct pricetag. Assume she knows little about the current prices and assigns the prices randomly.

(a) What is the probability that she wins the grand prize?
[*Ans.* $\frac{1}{24}$.]

(b) What is the probability that she prices none of the items correctly?

15. A room contains a group of n people who are wearing badges numbered from 1 to n. If two people are selected at random, what is the probability that the larger badge number is 4? Answer this problem assuming that $n = 3, 4, 5, 6$. [*Ans.* 0; $\frac{1}{2}$; $\frac{3}{10}$; $\frac{1}{5}$.]

16. Find the probability of obtaining each of the following poker hands. (A poker hand is a set of five cards chosen at random from a deck of 52 cards.)

(a) Royal flush (ten, jack, queen, king, ace in a single suit).

$$[Ans.\ \frac{4}{\binom{52}{5}} = .0000015.]$$

(b) Straight flush (five in a sequence in a single suit, but not a royal flush).

$$[Ans.\ \frac{(40-4)}{\binom{52}{5}} = .000014.]$$

(c) Four of a kind (four cards of the same face value).

$$[Ans.\ \frac{624}{\binom{52}{5}} = .00024.]$$

(d) Full house (one pair and one triple of the same face value).

$$[Ans.\ \frac{3744}{\binom{52}{5}} = .0014.]$$

(e) Flush (five cards in a single suit but not a straight or royal flush).

$$[Ans.\ \frac{(5148-40)}{\binom{52}{5}} = .0020.]$$

(f) Straight (five cards in a row, not all of the same suit).

$$[Ans.\ \frac{(10{,}240-40)}{\binom{52}{5}} = .0039.]$$

(g) Straight or better. [*Ans.* 0076.]

17. Find the probability of not having a pair in a hand of poker.

18. Find the probability of a "bust" hand in poker. (A hand is a "bust" if there is no pair and it is neither a straight nor a flush.) [*Ans.* .5012.]

19. In a survey, 100,000 people were interviewed. It was found that 65,832 of them had checking accounts, 43,971 of them had at least one credit card, and 32,348 of them had neither. What is the probability that a person selected at random from this sample has both a checking account and a credit card?

20. A certain French professor announces that he will select three out of eight pages of text to put on an examination and that each student can choose one of these three pages to translate. What is the minimum number of pages that a student should prepare in order to be certain of being able to translate a page that he has studied?

Smith decides to study only five of the eight pages. What is the probability that one of these five pages will appear on the examination? What is the smallest number of pages that Smith can study and still have probability greater than $\frac{3}{4}$ of being able to translate one page?

*4 TWO NONINTUITIVE EXAMPLES

There are occasions in probability theory when one finds a problem for which the answer, based on probability theory, is not at all in agreement with one's intuition. It is usually possible to arrange a few wagers that will bring one's intuition into line with the mathematical theory. A particularly good example of this is provided by the matching birthdays problem.

Assume that we have a room with r people in it and we propose the bet that there are at least two people in the room having the same birthday, i.e., the same month and day of the year. We ask for the value of r which will make this a fair bet. Few people will be willing to bet even money on this wager unless there were at least 100 people in the room. Most people would suggest 150 as a reasonable number. However, we shall see that with 150 people the odds are approximately, 4,100,000,000,000,000, to 1 in favor of two people having the same birthday, and that one should be willing to bet even money with as few as 23 people in the room.

Let us first find the probability that in a room with r people, no two have the same birthday. There are 365 possibilities for each person's birthday (neglecting February 29). There are, then, 365^r possibilities for the birthdays of r people. We assume that all these possibilities are equally likely. To find the probability that no two have the same birthday we must find the number of possibilities for the birthdays which have *no* day represented twice. The first person can have any of 365 days for his birthday. For each of these, if the second person is to have a different birthday, there are only 364 possibilities for his birthday. For the third man, there are 363 possibilities if he is to have a different birthday than the first two, etc. Thus the probability that no two people have the same birthday in a group of r people is

$$q_r = \frac{365 \cdot 364 \cdot \ldots \cdot (365 - r + 1)}{365^r}.$$

The probability that at least two people have the same birthday is then $p_r = 1 - q_r$. In Figure 1 the values of p_r and the odds for a fair bet, $p_r : (1 - p_r)$, are given for several values of r.

We consider now a second problem in which intuition does not lead to the correct answer. A hat-check girl has checked n hats, but they have become hopelessly scrambled. She hands back the hats at random. What is the probability that at least one man gets his own hat? For this problem some people's intuition would lead them to guess that for a large number of hats this probability should be small, while others guess that it should be large. Few people guess that the probability is neither large nor small and essentially independent of the number of hats involved.

Let p_j be the statement "the jth man gets his own hat back." We wish to find $\Pr[p_1 \vee p_2 \vee \ldots \vee p_n]$. A probability of this form can be found from the inclusion-exclusion formula as follows. We first add all proba-

Number of people in the room	Probability of at least two with same birthday	Approximate odds for a fair bet
5	.027	
10	.117	
15	.253	
20	.411	70:100
21	.444	80:100
22	.476	91:100
23	.507	103:100
24	.538	117:100
25	.569	132:100
30	.706	241:100
40	.891	819:100
50	.970	33:1
60	.994	170:1
70		1,200:1
80		12,000:1
90		160,000:1
100		3,300,000:1
125		31,000,000,000:1
150		4,100,000,000,000,000:1

Figure 1

bilities of the form Pr $[p_i]$, then subtract the sum of all probabilities of the form Pr $[p_i \wedge p_j]$, then add the sum of all probabilities of the form Pr $[p_i \wedge p_j \wedge p_k]$, etc. However, each of these probabilities represents the probability that a particular set of men get their own hats back. These probabilities are very easy to compute.

Let us find the probability that out of n men some particular m of them get back their own hats. There are $n!$ ways that the hats can be returned. If a particular m of them are to get their own hats there are only $(n - m)!$ ways that it can be done. Hence the probability that a particular m men get their own hats back is

$$\frac{(n-m)!}{n!}$$

There are $\binom{n}{m}$ different ways we can choose m men out of n. Hence the mth group of terms contributes

$$\binom{n}{m} \cdot \frac{(n-m)!}{n!} = \frac{1}{m!}$$

Number of hats	Probability p_n that at least one man gets his own hat
2	.500000
3	.666667
4	.625000
5	.633333
6	.631944
7	.632143
8	.632118

Figure 2

to the alternating sum. Thus

$$\Pr[p_1 \vee p_2 \vee \ldots \vee p_n] = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} \cdots \pm \frac{1}{n!},$$

where the + sign is chosen if n is odd and the − sign if n is even. In Figure 2, these numbers are given for the first few values of n.

It can be shown that, as the number of hats increases, the probabilities approach a number $1 - (1/e) = .632121\ldots$, where the number $e = 2.71828\ldots$ is a number that plays an important role in many branches of mathematics.

EXERCISES

1. What odds should you be willing to give on a bet that at least two of the presidents of the United States have had the same birthday? Would you win the bet?
   [*Ans.* More than 4:1. Yes, Polk and Harding were both born on November 2.]
2. What odds should you be willing to give on the bet that at least two of the presidents of the United States have died on the same day of the year? Would you win the bet?
   [*Ans.* More than 2.7:1. Yes; Jefferson, Adams and Monroe all died on July 4.]
3. What odds should you be willing to give on a bet that at least two people in the United States Senate have the same birthday?
4. What is the probability that at least two members of the House of Representatives have the same birthday?
5. Find the probability that, in a group of r people, at least one pair has the same birthmonth. How large does r have to be for this probability to be greater than $\frac{1}{2}$? (Assume that the probability of being born in any month is the same.)
6. Show that the probability that, in a group of r people, *exactly* one pair

has the same birthday is

$$t_r = \binom{r}{2} \frac{365 \cdot 364 \cdot \ldots \cdot (365 - r + 2)}{365^r}.$$

7. Show that $t_r = \binom{r}{2} \frac{q_r}{366 - r}$, where t_r is defined in Exercise 6 and q_r is the probability that no pair has the same birthday.
8. Find a formula for the probability of having more than one coincidence of birthdays among r people, i.e., of having at least two pairs of identical birthdays, or of three or more people having the same birthday. [*Hint:* Express the answer in terms of t_r.]
9. Is it very surprising that there was more than one coincidence of the dates on which presidents died (see Exercise 2)?
10. A contest requires entrants to match the stage names of four movie stars with their real names. Assuming a contestant guesses at random, what is his probability of getting none right? Of getting exactly four right? Exactly three? Two? One?
11. In how many ways can 8 rooks be placed on a chessboard so that none can attack any of the others? What is the probability that, in such an arrangement of rooks, the black diagonal has no rooks on it? (Do not carry out the arithmetical details.) Would the probability change if we used 16 rooks and a 16×16 chess board?
12. A teacher has her class of 75 students correct their own homework. She collects the papers, shuffles them, and passes one to each student. What is the approximate probability that no student receives his own paper back?
13. The clubs are removed from a deck of playing cards, shuffled, and dealt face up on a table. The position of each of the thirteen cards is noted, and then they are picked up, shuffled, and again dealt face up on the table. What is the approximate probability that no card occupies the same position in both deals?
14. The integers 1, 2, and 3 are written down in an arbitrary order. What is the probability that no two adjacent integers are consecutive (i.e., that the patterns 12 and 23 do not occur)? Do the problem for 1, 2, 3, and 4.

5 CONDITIONAL PROBABILITY

Suppose that we have a given $\mathcal{U}$ and that measures have been assigned to all subsets of $\mathcal{U}$. A statement p will have probability $\Pr[p] = m(P)$. Suppose we now receive some additional information, say that statement q is true. How does this additional information alter the probability of p?

EXAMPLE 1 Suppose we throw an ordinary die, and we are interested in statement p, "A 3 turns up." By our usual analysis the probability of this statement is

$\frac{1}{6}$. Suppose now that someone looks at the die and tells us statement q, "An odd number turned up." How does knowing q change our probability for statement p? Clearly the possibility set has been reduced from $\{1, 2, 3, 4, 5, 6\}$ to $\{1, 3, 5\}$. Assigning equal weights to the new set gives the new probability that a 3 turns up as $\frac{1}{3}$.

The probability of p after the receipt of the information q is called its *conditional probability,* and it is denoted by $\Pr[p \mid q]$, which is read "the probability of p given q." In this section we shall construct a method of finding this conditional probability in terms of the measure m.

If we know that q is true, then the original possibility set $\mathcal{U}$ has been reduced to Q and therefore we must define our measure on the subsets of Q instead of on the subsets of $\mathcal{U}$. Of course, every subset X of Q is a subset of $\mathcal{U}$, and hence we know $m(X)$, its measure before q was discovered. Since q cuts down on the number of possibilities, its new measure $m'(X)$ should be larger.

The basic idea on which the definition of m' is based is that, while we know that the possibility set has been reduced to Q, we have no new information about subsets of Q. If X and Y are subsets of Q, and $m(X) = 2 \cdot m(Y)$, then we will want $m'(X) = 2 \cdot m'(Y)$. This will be the case if the measures of subsets of Q are simply increased by a proportionality factor $m'(X) = k \cdot m(X)$, and all that remains is to determine k. Since we know that $1 = m'(Q) = k \cdot m(Q)$, we see that $k = l/m(Q)$ and our new measure on subsets of $\mathcal{U}$ is determined by the formula

$$m'(X) = \frac{m(X)}{m(Q)}. \tag{1}$$

How does this affect the probability of p? First of all, the truth set of p has been reduced. Because all elements of $\tilde{Q}$ have been eliminated, the new truth set of p is $P \cap Q$ and therefore

$$\Pr[p \mid q] = m'(P \cap Q) = \frac{m(P \cap Q)}{m(Q)} = \frac{\Pr[p \wedge q]}{\Pr[q]}. \tag{2}$$

Note that if the original measure m is the equiprobable measure, then the new measure m' will also be the equiprobable measure on the set Q.

We must take care that the denominators in (1) and (2) be different from zero. Observe that $m(Q)$ will be zero if Q is the empty set, which happens only if q is self-contradictory. This is also the only case in which $\Pr[q] = 0$, and hence we make the obvious assumption that our information q is not self-contradictory.

EXAMPLE 2 In an election, candidate A has .4 chance of winning, B has .3 chance, C has .2 chance, and D has .1 chance. Just before the election, C withdraws. Now what are the chances of the other three candidates? Let q be the statement that C will not win, i.e., that A or B or D will win. Observe that $\Pr[q] = .8$, hence all the other probabilities are increased by a factor of $1/.8 = 1.25$. Candidate A now has .5 chance of winning, B has .375, and D has .125.

EXAMPLE 3 A family is chosen at random from the set of all families having exactly two children (not twins). What is the probability that the family has two boys, if it is known that there is a boy in the family? Without any information being given, we would assign the equiprobable measure on the set $\mathcal{U} = \{BB, BG, GB, GG\}$, where the first letter of the pair indicates the sex of the older child and the second that of the younger. The information that there is a boy causes $\mathcal{U}$ to change to $\{BB, BG, GB\}$, but the new measure is still the equiprobable measure. Thus the conditional probability that there are two boys given that there is a boy is $\frac{1}{3}$. If, on the other hand, we know that the first child is a boy, then the possibilities are reduced to $\{BB, BG\}$ and the conditional probability is $\frac{1}{2}$.

A particularly interesting case of conditional probability is that in which $\Pr[p \mid q] = \Pr[p]$. That is, the information that q is true has no effect on our prediction for p. If this is the case, we note that

$$\Pr[p \wedge q] = \Pr[p]\Pr[q]. \tag{3}$$

And the case $\Pr[q \mid p] = \Pr[q]$ leads to the same equation. Whenever equation (3) holds, we say that p and q are *independent*. Thus if q is not a self-contradiction, p and q are independent if and only if $\Pr[p \mid q] = \Pr[p]$.

EXAMPLE 4 Consider three throws of an ordinary coin, where we consider the eight possibilities to be equally likely. Let p be the statement "A head turns up on the first throw" and q be the statement "A tail turns up on the second throw." Then $\Pr[p] = \Pr[q] = \frac{1}{2}$ and $\Pr[p \wedge q] = \frac{1}{4}$ and therefore p and q are independent statements.

While we have an intuitive notion of independence, it can happen that two statements that may not seem to be independent are in fact independent. For example, let r be the statement "The same side turns up all three times." Let s be the statement "At most one head occurs." Then r and s are independent statements (see Exercise 10).

An important use of conditional probabilities arises in the following manner. A set of statements $q_1, q_2, \ldots, q_n$ is said to be a *complete set of alternatives* if one and only one statement can be true. We wish to find the probability of a statement p, given a complete set of alternatives $q_1, q_2, \ldots, q_n$ such that the probability $\Pr[q_i]$ as well as the conditional probabilities $\Pr[p|q_i]$ can be found for every i. Then in terms of these we can find $\Pr[p]$ by

$$\Pr[p] = \Pr[q_1]\Pr[p \mid q_1] + \Pr[q_2]\Pr[p \mid q_2] + \ldots + \Pr[q_n]\Pr[p \mid q_n].$$

The proof of this assertion is left as an exercise (see Exercise 13).

EXAMPLE 5 A psychology student once studied the way mathematicians solve problems and contended that at times they try too hard to look for symmetry in a problem. To illustrate this she asked a number of mathematicians the following problem: Fifty balls (25 white and 25 black) are to be put in two urns, not necessarily the same number of balls in each. How should the balls be placed in the urns so as to maximize the chance of drawing a black ball, if an urn is chosen at random and a ball drawn from this urn? A quite surprising number of mathematicians answered that you could not do any better than $\frac{1}{2}$, by the symmetry of the problem. In fact one can do a good deal better by putting one black ball in urn 1, and all the 49 other balls in urn 2. To find the probability in this case let p be the statement "A black ball is drawn," q_1 the statement "Urn 1 is drawn" and q_2 the statement "Urn 2 is drawn." Then q_1 and q_2 are a complete set of alternatives, so

$$\Pr[p] = \Pr[q_1]\Pr[p \mid q_1] + \Pr[q_2]\Pr[p \mid q_2].$$

But $\Pr[q_1] = \Pr[q_2] = \frac{1}{2}$ and $\Pr[p \mid q_1] = 1$, $\Pr[p \mid q_2] = \frac{24}{49}$. Thus

$$\Pr[p] = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \frac{24}{49} = \frac{73}{98} \approx .745.$$

When told the answer, a number of the mathematicians that had said $\frac{1}{2}$ replied that they thought there had to be the same number of balls in each urn. However, since this had been carefully stated not to be necessary, they also had fallen into the trap of assuming too much symmetry.

EXERCISES

1. A card is drawn at random from a pack of playing cards. What is the probability that it is a 6 or a king, given that it is between 5 and 9 inclusive?
2. A die is loaded in such a way that the probability of a given number turning up is proportional to that number (e.g., a six is 3 times as likely to turn up as a two).
   (a) What is the probability of rolling an odd number, given that a six does not turn up? [*Ans.* $\frac{3}{5}$.]
   (b) What is the probability of rolling a six, given that an even number turns up? [*Ans.* $\frac{1}{2}$.]
3. Suppose we arrange the letters of the word "random" in a random order.
   (a) Find the probability that the letters are in alphabetical order given that the new arrangement begins with *a* and ends with *r*. [*Ans.* $\frac{1}{24}$.]
   (b) Which is greater, the probability that the two vowels are not together or the probability that the two vowels are not together and the new arrangement begins with *d*?
4. Referring to Exercise 8 in Section 3, what is the probability that the man selected studies German if
   (a) He studies French?

(b) He studies French and Russian?
(c) He studies neither French nor Russian?

5. A student takes a five-question true-false exam. What is the probability that he will get all answers correct if
   (a) He is only guessing?
   (b) He knows that the instructor puts more true than false questions on his exams?
   (c) He knows, in addition to (b), that the instructor never puts three questions in a row with the same answer?
   (d) He knows, in addition to (b) and (c), that the first and last questions must have the opposite answer?
   (e) He knows, in addition to (b), (c), and (d), that the answer to the second problem is "false"?
6. A die is thrown twice. What is the probability that the sum of the faces which turn up is 7, 8, or 9, given that one of them is a 4? Given that the first throw is a 4? [*Ans.* $\frac{5}{11}$; $\frac{1}{2}$.]
7. If $\Pr[q] = \frac{2}{5}$ and $\Pr[\sim p | \sim q] = \frac{1}{3}$, find $\Pr[p \vee q]$. [*Ans.* $\frac{4}{5}$.]
8. A certain motorist knows that before he reaches his destination the road forks four times, giving 16 possible paths. However, he does not remember which way he should turn at each of the forks. He decides that at each fork he will pick randomly which direction to go; thus each of the 16 possible patterns is equally likely. Unfortunately, after the four turns he realizes that he is in the wrong place.
   (a) What is the probability that he made a wrong turn at the first fork?
   (b) Given that he made the correct first turn, what is the probability that his second turn was incorrect? [*Ans.* $\frac{4}{7}$.]
   (c) Given that he made at least two correct turns, what is the probability that his first turn was correct?
9. Three persons, A, B, and C, are placed at random in a straight line. Let r be the statement "B is on the left" and let s be the statement "C is on the right."
   (a) What is $\Pr[r \wedge s]$?
   (b) Are r and s independent? [*Ans.* No.]
10. Prove that statements r and s in Example 4 are independent.
11. Let a deck of cards consist of the jacks and queens chosen from a bridge deck, and let two cards be drawn from the new deck. Find
   (a) The probability that the cards are both jacks, given that one is a jack. [*Ans.* $\frac{3}{11} = .27$.]
   (b) The probability that the cards are both jacks, given that one is a red jack. [*Ans.* $\frac{5}{13} = .38$.]
   (c) The probability that the cards are both jacks, given that one is the jack of hearts. [*Ans.* $\frac{3}{7} = .43$.]
12. Which is greater, Pr [a bridge hand contains 4 aces | it contains 1 ace] or Pr [a bridge hand contains 4 aces | it contains the ace of spades]?
13. Let p be any statement and q_1, q_2, q_3 be a complete set of alternatives.

Prove that

$$\Pr[p] = \Pr[q_1]\Pr[p|q_1] + \Pr[q_2]\Pr[p|q_2] + \Pr[q_3]\Pr[p|q_3].$$

14. The following example shows that r may be independent of p and q without being independent of $p \wedge q$ and $p \vee q$. We throw a coin twice. Let p be "The first toss comes out heads," q be "The second toss comes out heads," and r be "The two tosses come out different." Compute $\Pr[r]$, $\Pr[r|p]$, $\Pr[r|q]$, $\Pr[r|p \wedge q]$, $\Pr[r|p \vee q]$. [*Ans.* $\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, 0, \frac{2}{3}$.]

15. Assume that p and q are independent statements relative to a given measure. Prove that each of the following pairs of statements are independent relative to this same measure.

(a) p and $\sim q$.

(b) $\sim q$ and p.

(c) $\sim p$ and $\sim q$.

16. Prove that for any three statements p, q and r,

$$\Pr[p \wedge q \wedge r] = \Pr[p] \cdot \Pr[q|p] \cdot \Pr[r|p \wedge q].$$

17. **(a)** What is true about $\Pr[p|q]$ and $\Pr[q|p]$ if p and q are inconsistent?

(b) Under what other circumstances will it be true that $\Pr[p|q] = \Pr[q|p]$?

18. A card is drawn at random from a deck of playing cards. Are the following pairs of statements independent?

(a) p: A jack is drawn.
q: A black card is drawn.

(b) p: A black jack, queen, or king is drawn.
q: A spade which is not a 2, 3, or 4 is drawn.

19. A multiple-choice-test question lists four alternative answers, of which just one is correct. If a student has done his homework, then he is certain to identify the correct answer; otherwise he chooses an answer at random. Let p be the statement "A student does his homework" and q the statement "He answers the question correctly." Let $\Pr[p] = a$.

(a) Find a formula for $\Pr[p|q]$ in terms of a.

(b) Show that $\Pr[p|q] \geq \Pr[p]$ for all values of a. When does the equality hold?

20. A simple genetic model for the color of a person's eyes is the following: There are two kinds of color-determining genes, B and b, and each person has two color-determining genes. If both are b, he has blue eyes; otherwise he has brown eyes. Assume that one-quarter of the people have two B genes, one-quarter of the people have two b genes, and the rest have one B gene and one b gene.

(a) If a man has brown eyes, what is the probability that he has two B genes?

Assume that a man has brown eyes and that his wife has brown eyes. A child born to this couple will get one gene from the man and one from his wife, the selection in each case being a random selection from the parent's two genes.

(b) What is the probability that the child will have blue eyes?

(c) If the child has brown eyes, what is the probability that both of the parents have two B genes? [*Ans.* $\frac{1}{8}$.]

21. Two unfair coins, labeled A and B, are tossed. If Pr [A and B come up heads] = $\frac{1}{8}$ and Pr [A and B come up heads | at least one of them comes up heads] = $\frac{3}{14}$, find Pr [A comes up heads] and Pr [B comes up heads]. Assume that A has the greater probability of coming up heads and that the statements "A comes up heads" and "B comes up heads" are independent.

22. Three red, three green, and three blue balls are to be put into three urns, with at least two balls in each urn. Then an urn is selected at random and two balls withdrawn.

(a) How should the balls be put in the urns in order to maximize the probability of drawing two balls of different color? What is the probability? [*Partial Ans.* 1.]

(b) How should the balls be put in the urns in order to maximize the probability of withdrawing a red and a green ball? What is the maximum probability? [*Partial Ans.* $\frac{7}{10}$.]

23. A man who is extremely worried about having the plane in which he is flying blown up nevertheless always carries a bomb with him when he flies, because he has read that the probability of *two* people on the same plane having bombs is very low. Is his reasoning correct?

6 FINITE STOCHASTIC PROCESSES

We consider here a very general situation which we shall specialize in later sections. We deal with a sequence of experiments where the outcome on each particular experiment depends on some chance element. Any such sequence is called a *stochastic process*. (The Greek word *stochos* means "guess.") We shall assume a finite number of experiments and a finite number of possibilities for each experiment. We assume that, if all the outcomes of the experiments which precede a given experiment were known, then both the possibilities for this experiment and the probability that any particular possibility will occur would be known. We wish to make predictions about the process as a whole. For example, in the case of repeated throws of an ordinary coin we would assume that on any particular experiment we have two outcomes, and the probabilities for each of these outcomes is one-half regardless of any other outcomes. We might be interested, however, in the probabilities of statements of the form, "More than two-thirds of the throws result in heads," or "The number of heads and tails which occur is the same," etc. These are questions which can be answered only when a probability measure has been assigned to the process as a whole. In this section we show how probability measure can be assigned, using the given information. In the case of coin tossing, the probabilities

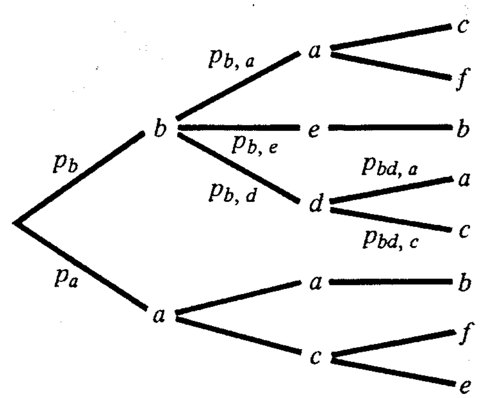


Figure 3

(hence also the possibilities) on any given experiment do not depend upon the previous results. We shall not make any such restriction here since the assumption is not true in general.

We shall show how the probability measure is constructed for a particular example, and the procedure in the general case is similar.

We assume that we have a sequence of three experiments, the possibilities for which are indicated in Figure 3. The set of all possible outcomes which might occur on any of the experiments is represented by the set $\{a, b, c, d, e, f\}$. Note that if we know that outcome b occurred on the first experiment, then we know that the possibilities on experiment two are $\{a, e, d\}$. Similarly, if we know that b occurred on the first experiment and a on the second, then the only possibilities for the third are $\{c, f\}$. We denote by p_a the probability that the first experiment results in outcome a, and by p_b the probability that outcome b occurs in the first experiment. We denote by $p_{b,d}$ the probability that outcome d occurs on the second experiment, which is the probability computed on the assumption that outcome b occurred on the first experiment. Similarly for $p_{b,a}, p_{b,e}, p_{a,a}, p_{a,c}$. We denote by $p_{bd,c}$ the probability that outcome c occurs on the third experiment, the latter probability being computed on the assumption that outcome b occurred on the first experiment and d on the second. Similarly for $p_{ba,c}, p_{ba,f}$, etc. We have assumed that these numbers are given and the fact that they are probabilities assigned to possible outcomes would mean that they are positive and that

$$p_a + p_b = 1, p_{b,a} + p_{b,e} + p_{b,d} = 1, \quad \text{and} \quad p_{bd,a} + p_{bd,c} = 1, \text{ etc.}$$

It is convenient to associate each probability with the branch of the tree that connects the branch point representing the predicted outcome. We have done this in Figure 3 for several branches. The sum of the numbers assigned to branches from a particular branch point is one, e.g.,

$$p_{b,a} + p_{b,e} + p_{b,d} = 1.$$

A possibility for the sequence of three experiments is indicated by a path through the tree. We define now a probability measure on the set of all paths. We call this a *tree measure*. To the path corresponding to outcome b on the first experiment, d on the second, and c on the third, we assign the weight $p_b \cdot p_{b,d} \cdot p_{bd,c}$. That is the *product* of the probabilities associated with each branch along the path being considered. We find the probability for each path through the tree.

Before showing the reason for this choice, we must first show that it determines a probability measure—in other words, that the weights are positive and the sum of the weights is one. The weights are products of positive numbers and hence positive. To see that their sum is one we first find the sum of the weights of all paths corresponding to a particular outcome, say b, on the first experiment and a particular outcome, say d, on the second. We have

$$p_b \cdot p_{b,d} \cdot p_{bd,a} + p_b \cdot p_{b,d} \cdot p_{bd,c} = p_b \cdot p_{b,d}[p_{bd,a} + p_{bd,c}] = p_b \cdot p_{b,d}.$$

For any other first two outcomes we would obtain a similar result. For example, the sum of the weights assigned to paths corresponding to outcome a on the first experiment and c on the second is $p_a \cdot p_{a,c}$. Notice that when we have verified that we have a probability measure, this will be the probability that the first outcome results in a and the second experiment results in c.

Next we find the sum of the weights assigned to all the paths corresponding to the cases where the outcome of the first experiment is b. We find this by adding the sums corresponding to the different possibilities for the second experiment. But by our preceding calculation this is

$$p_b \cdot p_{b,a} + p_b \cdot p_{b,e} + p_b \cdot p_{b,d} = p_b\,[p_{b,a} + p_{b,e} + p_{b,d}] = p_b.$$

Similary, the sum of the weights assigned to paths corresponding to the outcome a on the first experiment is p_a. Thus the sum of all weights is $p_a + p_b = 1$. Therefore we do have a probability measure. Note that we have also shown that the probability that the outcome of the first experiment is a has been assigned probability p_a in agreement with our given probability.

To see the complete connection of our new measure with the given probabilities, let $X_j = z$ be the statement "The outcome of the jth experiment was z." Then the statement $[X_1 = b \wedge X_2 = d \wedge X_3 = c]$ is a compound statement that has been assigned probability $p_b \cdot p_{b,d} \cdot p_{bd,c}$. The statement $[X_1 = b \wedge X_2 = d]$ we have noted has been assigned probability $p_b \cdot p_{b,d}$ and the statement $[X_1 = b]$ has been assigned probability p_b. Thus

$$\Pr[X_3 = c \mid X_2 = d \wedge X_1 = b] = \frac{p_b \cdot p_{b,d} \cdot p_{bd,c}}{p_b \cdot p_{b,d}} = p_{bd,c}$$

$$\Pr[X_2 = d \mid X_1 = b] = \frac{p_b \cdot p_{b,d}}{p_b} = p_{b,d}.$$

Thus we see that our probabilities, computed under the assumption that previous results were known, become the corresponding conditional probabilities when computed with respect to the tree measure. It can be shown that the tree measure which we have assigned is the only one which will lead to this agreement. We can now find the probability of any statement concerning the stochastic process from our tree measure.

EXAMPLE 1 Suppose that we have two urns. Urn 1 contains two black balls and three white balls. Urn 2 contains two black balls and one white ball. An urn is chosen at random and a ball chosen from this urn at random. What is the probability that a white ball is chosen? A hasty answer might be $\frac{1}{2}$, since there are an equal number of black and white balls involved and everything is done at random. However, it is hasty answers like this (which is wrong) which show the need for a more careful analysis.

We are considering two experiments. The first consists in choosing the urn and the second in choosing the ball. There are two possibilities for the first experiment, and we assign $p_1 = p_2 = \frac{1}{2}$ for the probabilities of choosing the first and the second urn, respectively. We then assign $p_{1,W} = \frac{3}{5}$ for the probability that a white ball is chosen, under the assumption that urn 1 is chosen. Similarly we assign $p_{1,B} = \frac{2}{5}, p_{2,W} = \frac{1}{3}, p_{2,B} = \frac{2}{3}$. We indicate these probabilities on the possibility tree in Figure 4. The probability that a white

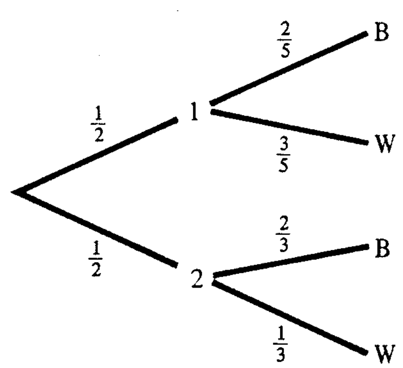


Figure 4

ball is drawn is then found from the tree measure as the sum of the weights assigned to paths which lead to a choice of a white ball. This is $\frac{1}{2} \cdot \frac{3}{5} + \frac{1}{2} \cdot \frac{1}{3} = \frac{7}{15}$.

EXAMPLE 2 Suppose that a man leaves a bar which is on a corner which he knows to be one block from his home. He is unable to remember which street leads to his home. He proceeds to try each of the streets at random without ever choosing the same street twice until he goes on the one which leads to his home. What possibilities are there for his trip home, and what is the probability for each of these possible trips? We label the streets A, B, C, and Home. The possibilities together with typical probabilities are given in Figure 5. The probability for any particular trip, or path, is found by taking the product of the branch probabilities.

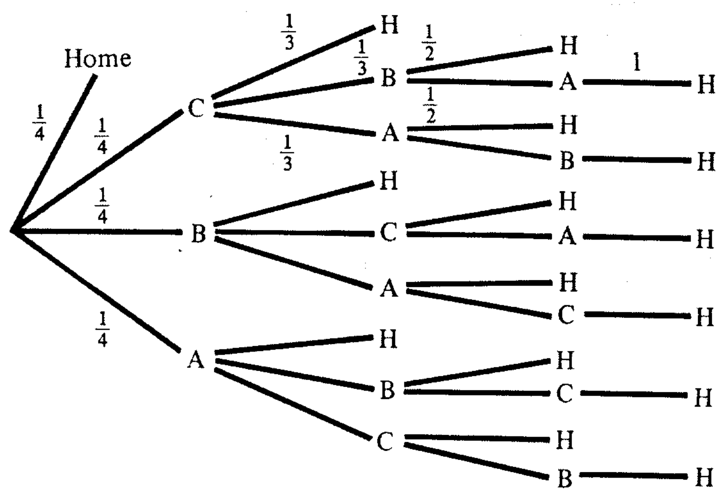


Figure 5

EXAMPLE 3 Assume that you are presented with two slot machines, A and B. Each machine pays the same fixed amount when it pays off. Machine A pays off each time with probability $\frac{1}{2}$, and machine B with probability $\frac{1}{4}$. You are not told which machine is A. Suppose that you choose a machine at random and win. What is the probability that you chose machine A? We first construct the tree (Figure 6) to show the possibilities and assign branch

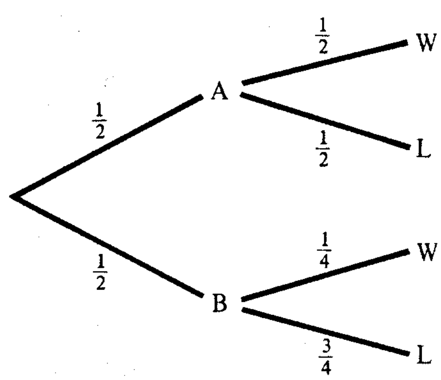


Figure 6

probabilities to determine a tree measure. Let p be the statement "Machine A was chosen" and q be the statement "The machine chosen paid off." Then we are asked for

$$\Pr[p \mid q] = \frac{\Pr[p \wedge q]}{\Pr[q]}.$$

The truth set of the statement $p \wedge q$ consists of a single path which has been assigned weight $\frac{1}{4}$. The truth set of the statement q consists of two paths, and the sum of the weights of these paths is $\frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{4} = \frac{3}{8}$.

Thus $\Pr[p|q] = \frac{2}{3}$. Thus if we win, it is more likely that we have machine A than B and this suggests that next time we should play the same machine. If we lose, however, it is more likely that we have machine B than A, and hence we would switch machines before the next play. (See Exercise 9.)

EXERCISES

1. Construct a tree measure to represent the possibilities for four throws of an ordinary coin. Assume that the probability of a head on any toss is $\frac{1}{2}$ regardless of any information about other throws.

2. Using the tree constructed in Exercise 1, find the probability of the following events:

(a) Two heads and two tails occur. [*Ans.* $\frac{3}{8}$.]

(b) The third toss is heads, given that the first two were tails. [*Ans.* $\frac{1}{2}$.]

(c) The first and third tosses are the same, given that the second and third tosses are the same. [*Ans.* $\frac{1}{2}$.]

3. A man has found through long experience with a certain soda machine that after depositing his money he will receive a soda with probability .8, his money will be returned with probability .1, and the machine will take the money with probability .1. If his money is returned, he deposits it again. If the machine takes his money, he kicks it, so that when he deposits more money the possible outcomes and their probabilities are soda, .85; return, 0; and take money, .15. After the second try he gives up. Construct a tree measure to represent the possible outcomes of the man's encounter with the soda machine.

4. In Exercise 3 find the probability of the following events:

(a) The man gets a soda.

(b) The man loses some money.

(c) The man gets a soda on his second try.

(d) The man gets a soda, given that he tried twice.

5. A man wins a certain tournament if he can win two consecutive games out of three played alternately with two opponents A and B. A is a better player than B. The probability of winning a game when B is the opponent is $\frac{2}{3}$. The probability of winning a game when A is his opponent is only $\frac{1}{3}$. Construct a tree measure for the possibilities for three games, assuming that he plays alternately but plays A first. Do the same assuming that he plays B first. In each case find the probability that he will win two consecutive games. Is it better to play two games against the stronger player or against the weaker player? [*Ans.* $\frac{10}{27}$; $\frac{8}{27}$; better to play strong player twice.]

6. A manufacturing plant makes a certain part on two different machines. Of the parts made by machine A, 80 percent are good and 20 percent defective; machine B is older and produces good parts only 75 percent of the time. Construct a tree measure for the experiment of picking a machine at random, then choosing two pieces of its output and inspecting them. What is the probability that both pieces are good? What is the probability that the pieces are good, given that they came from machine A? What is the probability that the pieces came from machine B, given that both are defective?

7. A cancer researcher has observed that 60 percent of young men begin to smoke cigarettes. Once someone begins to smoke, he quits with probability .25. Smokers develop lung cancer with probability .15, while one-time smokers who have quit get lung cancer with probability .1. Those men who have never smoked get lung cancer with probability .025. Construct a tree measure which illustrates this data. What is the probability that a man gets lung cancer? Given that he gets lung cancer, what is the probability that he once smoked cigarettes?

8. An urn contains three coins. One coin is fair, one falls heads with probability .6, and the other falls heads with probability .4. Construct a tree measure, and find the probability that a coin chosen from the urn at random and flipped will come up heads. Find the probability that the coin chosen was the fair one, given that it came up tails. [*Ans.* $\frac{1}{2}$; $\frac{1}{3}$.]

9. In Example 3, assume that the player makes two plays. Find the probability that he wins at least once under the assumption that

(a) He plays the same machine twice. [*Ans.* $\frac{19}{32}$.]

(b) He plays the same machine the second time if and only if he won the first time. [*Ans.* $\frac{20}{32}$.]

10. An urn initially contains two red and two blue balls. A ball is drawn from the urn, and it and two more balls of the same color are replaced in the urn. This process is carried out again, and finally a single ball is drawn. Construct a tree measure for the possible outcomes of the experiment. What is the probability that all three balls drawn are of the same color? What is the probability that the third ball is blue, given that the first two are red? What is the probability that the first two balls have the same color? [*Partial Ans.* $\frac{1}{2}$.]

11. A chess player plays three successive games of chess. His psychological makeup is such that the probability of his winning a given game is $(\frac{1}{2})^{k+1}$, where k is the number of games he has won so far. (For instance, the probability of his winning the first game is $\frac{1}{2}$, the probability of his winning the second game *if he has already won the first game* is $\frac{1}{4}$, etc.) What is the probability that he will win at least two of the three games. [*Ans.* $\frac{9}{32}$.]

12. Two defective lightbulbs have become mixed with three good bulbs. The bulbs are chosen one by one and tested until it is discovered which bulbs are defective. What is the least possible number of draws necessary? What is the greatest possible number of draws necessary? What is the probability that at most three draws are needed? Exactly three draws? Given that four draws are needed, what is the probability that the second and fourth bulbs are defective?

13. A composer of aleatory (random) music writes his works in three-note sections. The first note of each section is randomly chosen from A, C, and F. If the first note is A, the second is F with probability $\frac{1}{2}$ and B with probability $\frac{1}{2}$. If the first note is C, the second is B with probability $\frac{1}{4}$ and D with probability $\frac{3}{4}$. If the first note is F, the second is E with probability $\frac{1}{3}$ and A with probability $\frac{2}{3}$. The third note is the same as the first with probability $\frac{3}{8}$ and is one note higher than the first with probability $\frac{5}{8}$ (ignore sharps and flats). What is the probability that a given 3-note section contains a B? What is the probability that it contains a B, given that it contains no note twice? (Musical notes are arranged in ascending alphabetical order; thus B is one note higher than A, etc.)

14. Before a political convention, a political expert has assigned the following probabilities. The probability that the President will be willing to run again is $\frac{1}{2}$. If he is willing to run, he and his Vice-President are sure to be nominated and have probability $\frac{3}{5}$ of being elected again. If the President does not run, the present Vice-President has probability $\frac{1}{10}$ of being nominated, and any other presidential candidate has probability $\frac{1}{2}$ of being elected. What is the probability that the present Vice-President will be re-elected as either Vice-President or President? [*Ans.* $\frac{13}{40}$.]

15. A and B, finalists in a table tennis tournament, agree to play a best-of-three series for the championship. A has probability .6 of winning each game. What is the probability that A wins the championship? What is the probability that exactly three games are needed? What is the probability that the player who wins the first game goes on to win the championship?

16. In a room there are three chests, each chest contains two drawers, and each drawer contains one coin. In one chest each drawer contains a gold coin; in the second chest each drawer contains a silver coin; and in the last chest one drawer contains a gold coin and the other contains a silver coin. A chest is picked at random and then a drawer is picked at random from that chest. When the drawer is opened, it is found to contain a gold coin. What is the probability that the other drawer of that same chest will also contain a gold coin? [*Ans.* $\frac{2}{3}$.]

17. Four slips of paper, marked with the integers 1 through 4, are placed in a hat. What is the probability that the numbers on two slips drawn at random from the hat are in ascending (not necessarily consecutive) order?

18. A survey revealed that 75 percent of all mathematicians are eldest sons. Given that 90 percent of mathematicians are male, and the average family has three children, are the results surprising? (Assume that male and female children are equally likely.)

19. A student claims to be able to distinguish beer from ale. He is given a series of three tests. In each test he is given two glasses of beer and one of ale and asked to pick out the ale. If he gets two or more correct we shall admit his claim. Draw a tree to represent the possibilities (either he guesses right or he guesses wrong) for his answers. Construct the tree measure corresponding to his guessing and find the probability that his claim will be established if he guesses on every trial.

20. Urn A contains two red balls and one black ball; urn B contains one ball of each color. An urn is selected at random and a ball drawn from it. If the ball is black, it is returned to the urn; if it is red, it is placed in the other urn. Then another ball is drawn, this one from the other urn. Find the probability that the second ball drawn is black. What is the probability that both balls are the same color? Given that both balls drawn are red, what is the probability that the first urn chosen was A?

7 BAYES'S PROBABILITIES

The following situation often occurs. Measures have been assigned in a possibility space $\mathcal{U}$. A complete set of alternatives, $p_1, p_2, \ldots, p_n$ has been singled out. Their probabilities are determined by the assigned measure. (Recall that a complete set of alternatives is a set of statements such that for any possible outcome one and only one of the statements is true.) We are now given that a statement q is true. We wish to compute the new probabilities for the alternatives relative to this information. That is, we wish the conditional probabilities $\Pr[p_j|q]$ for each p_j. We shall give two different methods for obtaining these probabilities.

The first is by a general formula. We illustrate this formula for the case of four alternatives: p_1, p_2, p_3, p_4. Consider $\Pr[p_2|q]$. From the definition of conditional probability,

$$\Pr[p_2|q] = \frac{\Pr[p_2 \wedge q]}{\Pr[q]}.$$

But since p_1, p_2, p_3, p_4, are a complete set of alternatives,

$$\Pr[q] = \Pr[p_1 \wedge q] + \Pr[p_2 \wedge q] + \Pr[p_3 \wedge q] + \Pr[p_4 \wedge q].$$

Thus

$$\Pr[p_2|q] = \frac{\Pr[p_2 \wedge q]}{\Pr[p_1 \wedge q] + \Pr[p_2 \wedge q] + \Pr[p_3 \wedge q] + \Pr[p_4 \wedge q]}.$$

Since $\Pr[p_j \wedge q] = \Pr[p_j]\Pr[q|p_j]$, we have the desired formula

$$\Pr[p_2|q] = \frac{\Pr[p_2]\cdot\Pr[q|p_2]}{\Pr[p_1]\cdot\Pr[q|p_1]+\Pr[p_2]\cdot\Pr[q|p_2]+\Pr[p_3]\cdot\Pr[q|p_3]+\Pr[p_4]\cdot\Pr[q|p_4]}.$$

Similar formulas apply for the other alternatives, and the formula generalizes in an obvious way to any number of alternatives. In its most general form it is called *Bayes's theorem.*

EXAMPLE 1 Suppose that a freshman must choose among mathematics, physics, chemistry, and botany as his science course. On the basis of the interest he expressed, his adviser assigns probabilities of .4, .3, .2, and .1 to his choosing each of the four courses, respectively. His adviser does not hear which course he actually chose, but at the end of the term the adviser hears that he received A in the course chosen. On the basis of the difficulties of these courses the adviser estimates the probability of the student getting an A in mathematics to be .1, in physics .2, in chemistry .3, and in botany .9. How can the adviser revise his original estimates as to the probabilities of the student taking the various courses? Using Bayes's theorem we get

$$\Pr[\text{He took math} \mid \text{He got an A}] = \frac{(.4)(.1)}{(.4)(.1) + (.3)(.2) + (.2)(.3) + (.1)(.9)} = \frac{4}{25} = .16.$$

Similar computations assign probabilities of .24, .24, and .36 to the other three courses. Thus the new information, that he received an A, had little effect on the probability of his having taken physics or chemistry, but it has made it much less likely that he took mathematics, and much more likely that he took botany.

It is important to note that knowing the conditional probabilities of q relative to the alternatives is not enough. Unless we also know the probabilities of the alternatives at the start, we cannot apply Bayes's theorem. However, in some situations it is reasonable to assume that the alternatives are equally probable at the start. In this case the factors $\Pr[p_1], \ldots, \Pr[p_4]$ cancel from our basic formula, and we get the special form of the theorem:

$$\text{If} \quad \Pr[p_1] = \Pr[p_2] = \Pr[p_3] = \Pr[p_4], \quad \text{then}$$

$$\Pr[p_2 \mid q] = \frac{\Pr[q \mid p_2]}{\Pr[q \mid p_1] + \Pr[q \mid p_2] + \Pr[q \mid p_3] + \Pr[q \mid p_4]}.$$

EXAMPLE 2 In a sociological experiment the subjects are handed four sealed envelopes, each containing a problem. They are told to open one envelope and try to solve the problem in ten minutes. From past experience, the experimenter knows that the probability of their being able to solve the hardest problem is .1. With the other problems, they have probabilities of .3, .5, and .8. Assume the group succeeds within the allotted time. What is the probability that they selected the hardest problem? Since they have no way of knowing which problem is in which envelope, they choose at random, and we assign equal probabilities to the selection of the various problems. Hence the above simple formula applies. The probability of their having selected the hardest problem is

$$\frac{.1}{.1 + .3 + .5 + .8} = \frac{1}{17}.$$

The second method of computing Bayes's probabilities is to draw a tree, and then to redraw the tree in a different order. This is illustrated in the following example.

EXAMPLE 3 There are three urns. Each urn contains one white ball. In addition, urn I contains one black ball, urn II contains two, and urn III contains three. An urn is selected and one ball is drawn. The probability for selecting the three urns is $\frac{1}{6}$, $\frac{1}{2}$, and $\frac{1}{3}$, respectively. If we know that a white ball is drawn, how does this alter the probability that a given urn was selected?

First we construct the ordinary tree and tree measure (Figure 7).

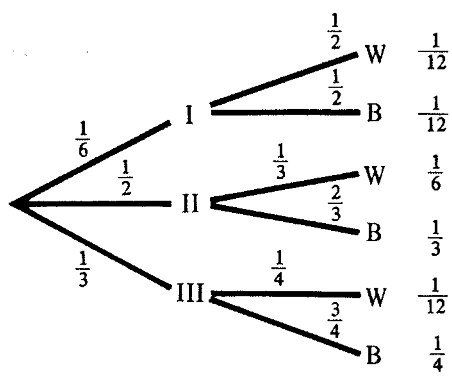


Figure 7

Next we redraw the tree, using the ball drawn as stage 1, and the urn selected as stage 2. We have the same paths as before, but in a different order. So the path weights are read off from the previous tree. The probability of drawing a white ball is

$$\tfrac{1}{12} + \tfrac{1}{6} + \tfrac{1}{12} = \tfrac{1}{3}.$$

This leaves the branch weights of the second stage to be computed (see Figure 8). But this is simply a matter of division. For example, the branch

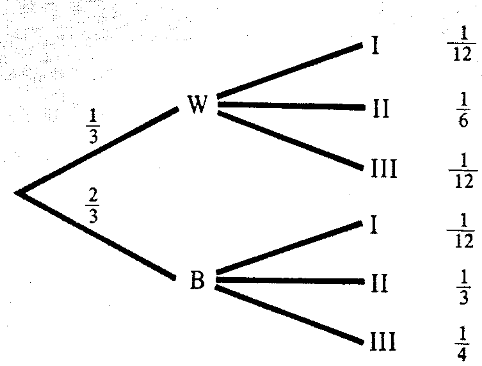


Figure 8

weights for the branches starting at "W" must be $\frac{1}{4}$, $\frac{1}{2}$, $\frac{1}{4}$ to yield the correct path weights. Thus, if a white ball is drawn, the probability of having selected urn I has increased to $\frac{1}{4}$, the probability of having picked urn III has fallen to $\frac{1}{4}$, while the probability of having chosen urn II is unchanged (see Figure 9).

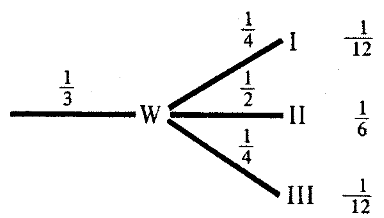


Figure 9

This method is particularly useful when we wish to compute all the conditional probabilities. We shall apply the method next to Example 1. The tree and tree measure for this example in the natural order is shown in Figure 10. In that figure the letters M, P, C, and B stand for mathematics, physics, chemistry, and botany, respectively.

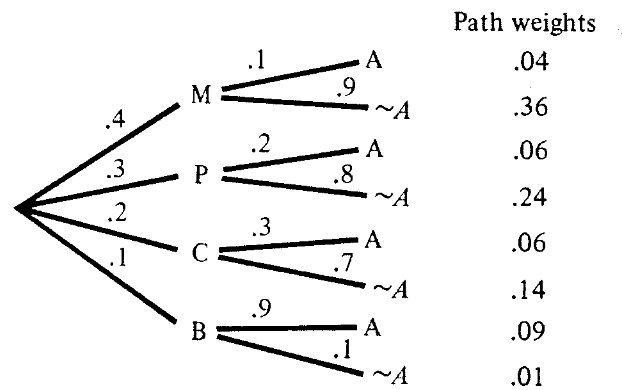


Figure 10

The tree drawn in reverse order is shown in Figure 11.

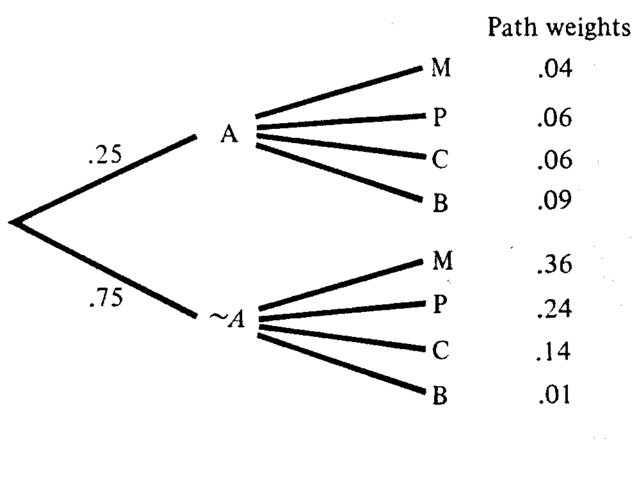


Figure 11

Each path in this tree corresponds to one of the paths in the original tree. Therefore the path weights for this new tree are the same as the weights assigned to the corresponding paths in the first tree. The two branch weights at the first level represent the probability that the student receives an A or that he does not receive an A. These probabilities are also easily obtained from the first tree. In fact,

$$\Pr[A] = .04 + .06 + .06 + .09 = .25$$

and

$$\Pr[\sim A] = 1 - .25 = .75.$$

We have now enough information to obtain the branch weights at the second level, since the product of the branch weights must be the path weights. For example, to obtain $p_{A,M}$ we have

$$.25 \cdot p_{A,M} = .04 \quad \text{or} \quad p_{A,M} = .16.$$

But $p_{A,M}$ is also the conditional probability that the student took math given that he got an A. Hence this is one of the new probabilities for the alternatives in the event that the student received an A. The other branch probabilities are found in the same way and represent the probabilities for the other alternatives. By this method we obtain the new probabilities for all alternatives under the hypothesis that the student receives an A as well as the hypothesis that the student does not receive an A. The results are shown in the completed tree in Figure 12.

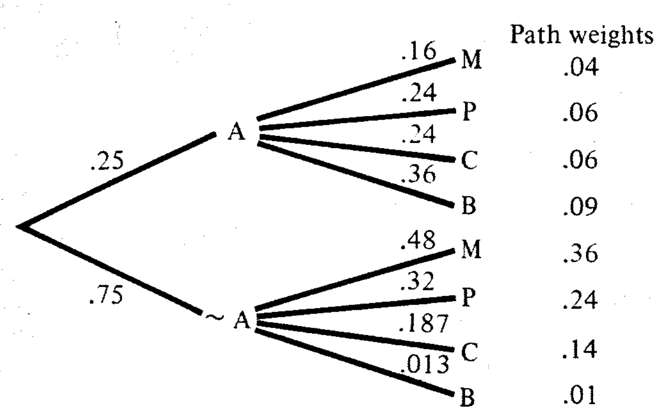


Figure 12

EXERCISES

1. A certain New England state has fair weather 20 percent of the time and foul weather 80 percent of the time. If a given day is fair, the probability that the next day is fair is .25; if a given day is foul, the next day is also foul with probability .75. If it is fair today, what is the probability that it was fair yesterday?

2. A survey showed that 25 percent of American cars are compacts, 40 percent are intermediates, and 35 percent are standard size. If a compact car is involved in an accident, the probability that its occupants are seriously injured is .6. For an intermediate car the probability is .5, and for standard size car it is .4. Given that the occupants of a car were seriously injured in an accident, what is the probability that the car was a compact? What is the probability that the car was an intermediate? A standard?
3. For Exercise 2, construct the tree measure and the tree drawn in reverse order.
4. On a multiple-choice exam there are four possible answers for each question. Therefore, if a student knows the right answer, he has probability 1 of choosing correctly; if he is guessing, he has probability $\frac{1}{4}$ of choosing correctly. Let us further assume that a good student will know 90 percent of the answers, a poor student only 50 percent. If a good student chooses the right answer, what is the probability that he was only guessing? Answer the same question about a poor student, if the poor student chooses the right answer. [*Ans.* $\frac{1}{37}$, $\frac{1}{5}$.]
5. At a small coeducational college, 50 percent of the students are majoring in liberal arts, 10 percent in nursing, 10 percent in performing arts, and 30 percent in education. The proportion of women in the various majors is 40, 90, 60, and 50 percent respectively. Find the probability that a given male student is enrolled in each of the majors.
6. Of 200 people attending an office picnic, 150 eat one helping of potato salad, 30 eat two helpings, and 20 eat three helpings. Later many of those who attended the picnic became sick, and it is discovered that the potato salad was the cause. A doctor estimates that the probability of becoming sick is .3 times the number of servings of potato salad eaten. Find the probability that a person who became sick ate one, two, or three helpings. Do the same for a person who did not get sick.
7. Three men, A, B, and C, are in jail, and one of them is to be hanged the next day. The jailor knows which man will hang, but must not announce it. Man A says to the jailor, "Tell me the name of one of the two who will not hang. If both are to go free, just toss a coin to decide which to say. Since I already know that at least one of them will go free, you are not giving away the secret." The jailor thinks a moment and then says, "No, this would not be fair to you. Right now you think the probability that you will hang is $\frac{1}{3}$; but if I tell you the name of one of the others who is to go free, your probability of hanging increases to $\frac{1}{2}$. You would not sleep as well tonight." Was the jailor's reasoning correct? [*Ans.* No.]
8. A machine for testing radio tubes will detect a defective tube with probability .95, but will show that a good tube is defective with probability .1. A technician knows that one tube in a radio with ten tubes is defective. He selects a tube at random, tests it, and finds that the machine indicates the tube is defective. What is the probability that

the tube actually is defective? Suppose the machine says the tube is good. What is now the probability that the tube is in fact defective?

9. (This problem should be done both with and without using Bayes's theorem.) A deck contains three cards. One is black on both sides, another is red on both sides, and the third has one red side and one black side. A card is selected from the deck at random and dealt onto a table; the face showing is black. What is the probability that the other face is also black?

10. One coin in a collection of 8 million coins has two heads. The rest are fair coins. A coin chosen at random from the collection is tossed ten times and comes up heads every time. What is the probability that it is the two-headed coin?

11. Referring to Exercise 10, assume that the coin is tossed n times and comes up heads every time. How large does n have to be to make the probability approximately $\frac{1}{2}$ that you have the two-headed coin? [*Ans.* 23.]

12. A musicologist is attempting to determine the composer of a newly discovered baroque ditty. He thinks it equally likely to be Archangelo Spumani or his lesser-known brother Pistachio. Unfortunately both composed only in the keys of A major and F minor; Archangelo used the former 60 percent of the time, while Pistachio used the latter in 80 percent of his compositions. If the musicologist discovers that the work is in F minor, what is the probability that it was written by Archangelo? By Pistachio?

13. One-third of the subjects in a test of cold remedies are given vitamin C, $\frac{1}{2}$ are given antibiotics, and $\frac{1}{6}$ are given a placebo. The colds of $\frac{1}{4}$ of the vitamin-C group, $\frac{1}{2}$ of the antibiotic group, and $\frac{3}{5}$ of the placebo group are cured. What is the probability that a subject whose cold was *not* cured was given vitamin C? What is the probability that a subject whose cold was cured was given a placebo?

14. The manager of an office employing 15 women and 5 men discovers that the men are equally likely to use a paper clip as a nail cleaner, as a paper fastener, or as ammunition for a rubber-band slingshot. The women never shoot paper clips, but use them as nail cleaners with probability .75 and as paper fasteners with probability .25. If a paper clip is used as a nail cleaner, what is the probability that it was used by a woman? What is the probability that a clip shot across the office was shot by a man?

8 INDEPENDENT TRIALS WITH TWO OUTCOMES

In the preceding section we developed a way to determine a probability measure for any sequence of chance experiments where there are only a finite number of possibilities for each experiment. While this provides the framework for the general study of stochastic processes, it is too general to be studied in complete detail. Therefore, in probability theory we look for simplifying assumptions which will make our probability measure easier to work with. It is desired also that these assumptions be such as to apply to a variety of experiments which would occur in practice. In this book we shall limit ourselves to the study of two types of processes. The first, the independent trials process, will be considered in the present section. This process was the first one to be studied extensively in probability theory. The second, the Markov chain process, is a process that is finding increasing application, particularly in the social and biological sciences, and will be considered in Section 12.

A process of independent trials applies to the following situation. Assume that there is a sequence of chance experiments, each of which consists of a repetition of a single experiment, carried out in such a way that the results of any one experiment in no way affect the results in any other experiment. We label the possible outcome of a single experiment by $a_1, \ldots, a_r$. We assume that we are also given probabilities $p_1, \ldots, p_r$ for each of these outcomes occurring on any single experiment, the probabilities being independent of previous results. The tree representing the possibilities for the sequence of experiments will have the same outcomes from each branch point, and the branch probabilities will be assigned by assigning probability p_j to any branch leading to outcome a_j. The tree measure determined in this way is the measure of an *independent trials process*. In this section we shall consider the important case of two outcomes for each experiment. The more general case is studied in Section 10.

In the case of two outcomes we arbitrarily label one outcome "success" and the other "failure." For example, in repeated throws of a coin we might call heads success, and tails failure. We assume there is given a probability p for success and a probability $q = 1 - p$ for failure. The tree measure for a sequence of three such experiments is shown in Figure 13. The weights assigned to each path are indicated at the end of the path.

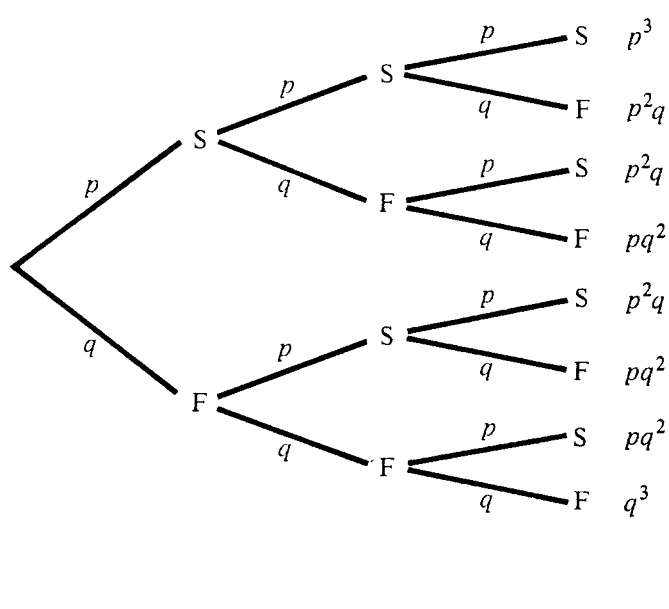


Figure 13

The question which we now ask is the following. Given an independent trials process with two outcomes, what is the probability of *exactly* x successes in n experiments? We denote this probability by $f(n, x; p)$ to indicate that it depends upon n, x, and p.

Assume that we had a tree for this general situation, similar to the tree in Figure 13 for three experiments, with the branch points labeled S for success and F for failure. Then the truth set of the statement "Exactly x successes occur" consists of all paths which go through x branch points labeled S and $n - x$ labeled F. For instance, in Figure 13 suppose $x = 2$ so that we are interested in the probability that "exactly two successes" occur. We look for all the paths that go through two branch points labeled S and $(3 - 2)$ or one branch point labeled F. (There are three paths of this type.) To find the probability of this statement we must add the weights for all such paths. We are helped first by the fact that our tree measure assigns the same weight to any such path, namely $p^x q^{n-x}$. The reason for this is that every branch leading to an S is assigned probability p, and every branch leading to F is assigned probability q, and in the product there will be x p's and $(n - x)$ q's. To find the desired probability we need only find the number of paths in the truth set of the statement "Exactly x successes occur." But that is just the number of ways we can label x branch points with S and $n - x$ branch points with F. We found in Chapter 2 that this labeling could be done in $\binom{n}{x}$ ways. Thus we have proved:

In an independent trials process with two outcomes the probability of exactly x successes in n experiments is given by

$$f(n, x; p) = \binom{n}{x} p^x q^{n-x}.$$

EXAMPLE 1 Consider n throws of an ordinary coin. We label heads "success" and tails "failure," and we assume that the probability is $\frac{1}{2}$ for heads on any one throw independently of the outcome of any other throw. Then the probability that exactly x heads will turn up is

$$f(n, x; \tfrac{1}{2}) = \binom{n}{x}\left(\frac{1}{2}\right)^n.$$

For example, in 100 throws the probability that exactly 50 heads will turn up is

$$f(100, 50; \tfrac{1}{2}) = \binom{100}{50}\left(\frac{1}{2}\right)^{100},$$

which is approximately .08. Thus we see that it is quite unlikely that exactly one-half of the tosses will result in heads. On the other hand, suppose that we ask for the probability that nearly one-half of the tosses will be heads. To be more precise, let us ask for the probability that the number of heads which occur does not deviate by more than 10 from 50. To find this we must add

$$f(100, x; \tfrac{1}{2}) \quad \text{for} \quad x = 40, 41, \ldots, 60.$$

If this is done, we obtain a probability of approximately .96. Thus, while it is unlikely that exactly 50 heads will occur, it is very likely that the number of heads which occur will not deviate from 50 by more than 10.

EXAMPLE 2 Assume that we have a machine which, on the basis of data given, is to predict the outcome of an election as either a Republican victory or a Democratic victory. If two identical machines are given the same data, they should predict the same result. We assume, however, that any such machine has a certain probability q of reversing the prediction that it would ordinarily make, because of a mechanical or electrical failure. To improve the accuracy of our prediction we give the same data to r identical machines, and choose the answer which the majority of the machines give. To avoid ties we assume that r is odd. Let us see how this decreases the probability of an error due to a faulty machine.

Consider r experiments, where the jth experiment results in success if the jth machine produces the prediction which it would make when operating without any failure of parts. The probability of success is then $p = 1 - q$. The majority decision will agree with that of a perfectly operating machine if we have more than $r/2$ successes. Suppose, for example, that we have five machines, each of which has a probability of .1 of reversing the prediction because of a parts failure. Then the probability for success is .9, and the probability that the majority decision will be the desired one is

$$f(5, 3; 0.9) + f(5, 4; 0.9) + f(5, 5; 0.9),$$

which is found to be approximately .991 (see Exercise 3).

Thus the above procedure decreases the probability of error due to machine failure from .1 in the case of one machine to .009 for the case of five machines.

EXERCISES

1. Compute for $n = 4$, $n = 6$, and $n = 8$ the probability of obtaining heads exactly half the time when an ordinary coin is thrown.
2. Do Exercise 1 for a loaded coin which has probability $\frac{3}{4}$ of coming up heads. How do the answers compare with those in Exercise 1? [*Ans.* .211, .132, .087.]
3. Verify that the probability .991 given in Example 2 is correct.
4. A machine produces light bulbs that are good with probability .95 and defective with probability .05. What is the probability that a sample of ten bulbs selected at random from the machine's output contains at most one defective bulb? (Do not carry out the computation.)

5. Suppose an unprepared student takes a five-question multiple-choice exam. Each question has four possible answers, only one of which is correct. What is the probability that he can attain a passing grade of 80 percent by guessing?
6. A coin is to be thrown eight times. What is the most probable number of heads that will occur? What is the number having the highest probability, given that the first four throws result in heads?
7. A die is made by marking the faces of a regular dodecahedron with the numbers 1 through 12. What is the probability that on exactly three out of six throws of the die, a number larger than 8 turns up?
8. Suppose a coin is flipped six times. What is the probability that more than half of the tosses come up tails? Answer the same for seven throws, and for 17,219 throws.
9. Suppose seven coins are flipped eleven times each. What is the probability that more than half of the coins come up heads more than one-half of the time?
   [*Ans.* $\frac{1}{2}$.] [*Hint:* Use the result of Exercise 8 twice.]
10. A small factory has ten workers. The workers eat their lunch at one of two diners, and they are just as likely to eat at one as in the other. If the proprietors want to be more than .95 sure of having enough seats, how many seats must each of the diners have? [*Ans.* Eight seats.]
11. Suppose we have a computer routine to produce random digits. What is the probability that in ten trials of the routine, more than two zeros are output?
12. A trapper has found through experience that he can expect a given trap to catch an animal once every three weeks. How many traps should he set to have probability at least .7 of catching at least two animals a week?
13. In a certain board game players move around the board, and each turn consists of a player's rolling a pair of dice. If a player is on the square marked "Park Bench," he must roll a seven or doubles before he is allowed to move out.
    (a) What is the probability that a player stuck on "Park Bench" will be allowed to move out on his next turn?
    (b) A player stuck on "Park Bench" has probability greater than $\frac{3}{4}$ of getting out after how many rolls? [*Ans.* (a) $\frac{1}{3}$; (b) 4.]
14. A machine produces small electrical parts which are perfect with probability .8, defective but usable with probability .15, and useless with probability .05. Find the probability that a sample of ten parts made by the machine contains exactly eight perfect parts. (Do not carry out the computation.)
15. Find the probability that the sample in Exercise 14 contains seven perfect parts, two defective but usable parts, and one useless part.
16. Show that $f(n, x; p) = \dfrac{(n - x + 1)p}{x \cdot q} f(n, x - 1; p)$.

17. For given n and p, find the k such that $f(n, k; p)$ is the largest. [*Hint:* We want $f(n, k; p) \geq f(n, k - 1; p)$ and $f(n, k; p) \geq f(n, k + 1; p)$; use the result of Exercise 16.]

18. Without actually computing the probabilities, find the value of x for which $f(20, x; .3)$ is largest.

19. A restaurant orders five pieces of apple pie and five pieces of cherry pie. Assume that the restaurant has ten customers, and the probability that a customer will ask for apple pie is $\frac{3}{4}$ and for cherry pie is $\frac{1}{4}$.

(a) What is the probability that the ten customers will all be able to have their first choice? (Do not carry out the computation.)

(b) What number of each kind of pie should the restaurant order if it wishes to order ten pieces of pie and wants to maximize the probability that the ten customers will all have their first choice?

20. Suppose a computer routine for generating random digits is operated 1000 times. What is the most likely number of times the digit 7 appears?

9 THE LAW OF LARGE NUMBERS

In this section we shall study some further properties of the independent trials process with two outcomes. In Section 8 we saw that the probability for x successes in n trials is given by

$$f(n, x; p) = \binom{n}{x} p^x q^{n-x}.$$

In Figure 14 we show these probabilities graphically for $n = 8$ and $p = \frac{3}{4}$.

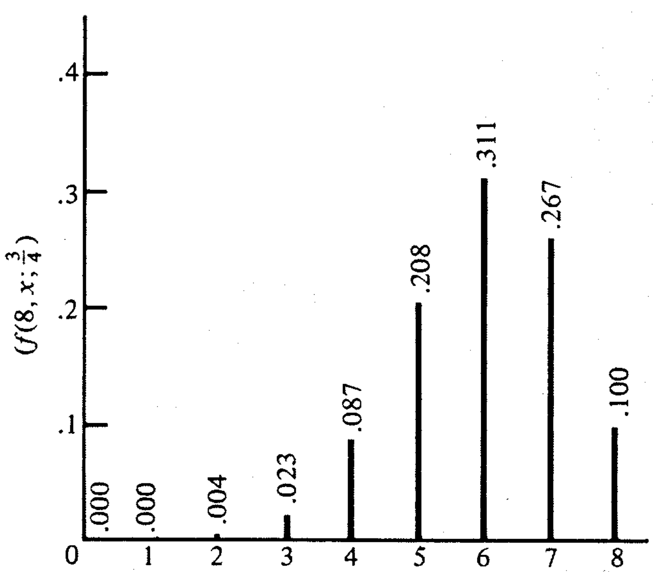


Figure 14

In Figure 15 we have done similarly for the case of $n = 7$ and $p = \frac{3}{4}$.

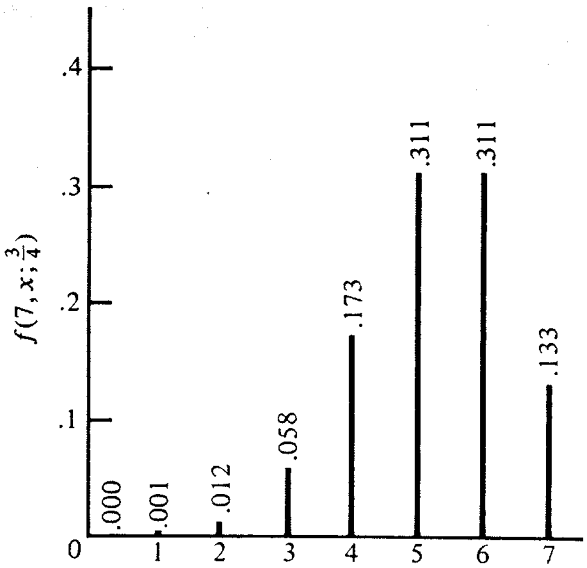


Figure 15

We see in the first case that the values increase up to a maximum value at $x = 6$ and then decrease. In the second case the values increase up to a maximum value at $x = 5$, have the same value for $x = 6$, and then decrease. These two cases are typical of what can happen in general.

Consider the ratio of the probability of $x + 1$ successes in n trials to the probability of x successes in n trials, which is

$$\frac{\binom{n}{x+1}p^{x+1}q^{n-x-1}}{\binom{n}{x}p^x q^{n-x}} = \frac{n-x}{x+1}\cdot\frac{p}{q}.$$

This ratio will be greater than one as long as $(n - x)p > (x + 1)q$ or as long as $x < np - q$. If $np - q$ is not an integer, the values $\binom{n}{x}p^x q^{n-x}$ increase up to a maximum value, which occurs at the first integer greater than $np - q$, and then decrease. In case $np - q$ is an integer, the values $\binom{n}{x}p^x q^{n-x}$ increase up to $x = np - q$, are the same for $x = np - q$ and $x = np - q + 1$, and then decrease.

Thus we see that, in general, values near np will occur with the largest probability. It is not true that one particular value near np is highly likely to occur, but only that it is relatively more likely than a value further from np. For example, in 100 throws of a coin, $np = 1000 \cdot \frac{1}{2} = 50$. The probability of exactly 50 heads is approximately .08. The probability of exactly 30 is approximately .00002.

More information is obtained by studying the probability of a given deviation of the proportion of successes x/n from the number p; that is, by studying for $\epsilon > 0$,

$$\Pr\left[\left|\frac{x}{n} - p\right| < \epsilon\right].$$

For any fixed n, p, and ϵ, the latter probability can be found by adding all the values of $f(n, x; p)$ for values of x for which the inequality $p - \epsilon < x/n < p + \epsilon$ is true. In Figure 16 we have given these probabilities for the case $p = .3$ with various values for ϵ and n. In the first column we have the case $\epsilon = .1$. We observe that as n increases, the probability that the fraction of successes deviates from .3 by less than .1 tends to the value 1. In fact, to four decimal places the answer is 1.0000 after $n = 400$. In the second column we have the same probabilities for the smaller value of $\epsilon = .05$. Again the probabilities are tending to 1 but not so fast. In the third column we have given these probabilities for the case $\epsilon = .02$. We see now that even after 1000 trials there is still a reasonable chance that the fraction x/n is not within .02 of the value of $p = .3$. It is natural to ask if we can expect these probabilities also to tend to 1 if we increase n sufficiently. The answer is yes and this is assured by one of the fundamental theorems of probability called the *law of large numbers*. This theorem asserts that, for any $\epsilon > 0$,

$$\Pr\left[\left|\frac{x}{n} - p\right| < \epsilon\right]$$

tends to 1 as n increases indefinitely.

$\Pr\left[\left|\frac{x}{n} - p\right| < \epsilon\right]$ for $p = .3$ and $\epsilon = .1, .05, .02.$

n	$\Pr\left[\left\|\frac{x}{n} - .3\right\| < .10\right]$	$\Pr\left[\left\|\frac{x}{n} - .3\right\| < .05\right]$	$\Pr\left[\left\|\frac{x}{n} - .3\right\| < .02\right]$
20	.5348	.1916	.1916
40	.7738	.3945	.1366
60	.8800	.5184	.3269
80	.9337	.6068	.2853
100	.9626	.6740	.2563
200	.9974	.8577	.4107
300	.9998	.9326	.5116
400	1.0000	.9668	.5868
500	1.0000	.9833	.6461
600	1.0000	.9915	.6944
700	1.0000	.9956	.7345
800	1.0000	.9977	.7683
900	1.0000	.9988	.7970
1000	1.0000	.9994	.8216

Figure 16

It is important to understand what this theorem says and what it does not say. Let us illustrate its meaning in the case of coin tossing.

We are going to toss a coin n times and we want the probability to be very high, say greater than .99, that the fraction of heads which turn up will be very close, say within .001, of the value .5. The law of large numbers assures us that we can have this if we simply choose n large enough. The theorem itself gives us no information about how large n must be. Let us, however, consider this question.

To say that the fraction of the times success results is near p is the same as saying that the actual number of successes x does not deviate too much from the expected number np. To see the kind of deviations which might be expected we can study the value of $\Pr[|x - np| \geq d]$. A table of these values for $p = .3$ and various values of n and d are given in Figure 17. Let us ask how large d must be before a deviation as large as d could be considered surprising. For example, let us see for each n the value of d which makes $\Pr[|x - np| \geq d]$ about .04. From the table, we see that d should be 7 for $n = 50$, 9 for $n = 80$, 10 for $n = 100$, etc. To see deviations which might be considered more typical we look for the values of d which make $\Pr[|x - np| \geq d]$ approximately $\frac{1}{3}$. Again from the table, we see that d should be 3 or 4 for $n = 50$, 4 or 5 for $n = 80$, 5 for $n = 100$, etc. The answers to these two questions are given in the last two columns of the table. An examination of these numbers shows us that deviations which we would consider surprising are approximately $\sqrt{n}$ while those which are more typical are about one half as large or $\sqrt{n}/2$.

This suggests that $\sqrt{n}$, or a suitable multiple of it, might be taken as a unit of measurement for deviations. Of course, we would also have to study how $\Pr[|x - np| \geq d]$ depends on p. When this is done, one finds that $\sqrt{npq}$ is a natural unit; it is called a *standard deviation*. It can be shown that for large n the following approximations hold:

$$\Pr[|x - np| \geq \sqrt{npq}] \approx .3174$$
$$\Pr[|x - np| \geq 2\sqrt{npq}] \approx .0455$$
$$\Pr[|x - np| \geq 3\sqrt{npq}] \approx .0027.$$

That is, a deviation from the expected value of one standard deviation is rather typical, while a deviation of as much as two standard deviations is quite surprising and three very surprising. For values of p not too near 0 or 1, the value of $\sqrt{pq}$ is approximately $\frac{1}{2}$. Thus these approximations are consistent with the results we observed from our table.

For large n, $\Pr[x - np \geq k\sqrt{npq}]$ or $\Pr[x - np \leq -k\sqrt{npq}]$ can be shown to be approximately the same. Hence these probabilities can be estimated for $k = 1$, 2, and 3 by taking $\frac{1}{2}$ the values given above.

EXAMPLE 1 In throwing an ordinary coin 10,000 times, the expected number of heads is 5000, and the standard deviation for the number of heads is $\sqrt{10{,}000(\frac{1}{2})(\frac{1}{2})} = 50$. Thus the probability that the number of heads which

$p = 3$; $\Pr[|x - np| \geq d]$.

n \ d	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	Pr near to .04	Pr near to $\frac{1}{3}$
50	.878	.644	.441	.280	.164	.088	.043	.020	.008									7	3–4
80	.903	.715	.542	.393	.272	.179	.112	.066	.037	.020	.010							9	4–5
100	.913	.744	.586	.445	.326	.230	.155	.101	.063	.037	.021	.012						10	5
120	.921	.765	.619	.486	.370	.273	.195	.135	.090	.058	.036	.022	.012					11	5–6
140	.927	.782	.645	.519	.407	.310	.230	.166	.116	.079	.052	.033	.021	.012				12	6
170	.933	.802	.676	.558	.451	.357	.276	.209	.154	.111	.078	.054	.036	.024	.015	.009		13	6
200	.939	.817	.700	.589	.488	.396	.316	.247	.189	.142	.105	.076	.053	.037	.025	.017	.011	14	7

Figure 17

turn up deviates from 5000 by as much as one standard deviation, or 50, is approximately .317. The probability of a deviation of as much as two standard deviations, or 100, is approximately .046. The probability of a deviation of as much as three standard deviations, or 150, is approximately .003.

EXAMPLE 2 Assume that in a certain large city, 900 people are chosen at random and asked if they favor a certain proposal. Of the 900 asked, 550 say they favor the proposal and 350 are opposed. If, in fact, the people in the city are equally divided on the issue, would it be unlikely that such a large majority would be obtained in a sample of 900 of the citizens? If the people were equally divided, we would assume that the 900 people asked would form an independent trials process with probability $\frac{1}{2}$ for a "yes" answer and $\frac{1}{2}$ for a "no" answer. Then the standard deviation for the number of "yes" answers in 900 trials is $\sqrt{900(\frac{1}{2})(\frac{1}{2})} = 15$. Then it would be very unlikely that we would obtain a deviation of more than 45 from the expected number of 450. The fact that the deviation in the sample from the expected number was 100, then, is evidence that the hypothesis that the voters were equally divided is incorrect. The assumption that the true proportion is any value less than $\frac{1}{2}$ would also lead to the fact that a number as large as 550 favoring in a sample of 900 is very unlikely. Thus we are led to suspect that the true proportion is greater than $\frac{1}{2}$. On the other hand, if the number who favored the proposal in the sample of 900 were 465, we would have only a deviation of one standard deviation, under the assumption of an equal division of opinion. Since such a deviation is not unlikely, we could not rule out this possibility on the evidence of the sample.

EXAMPLE 3 A certain Ivy League college would like to admit 800 students in their freshman class. Experience has shown that if they accept 1250 students they will have acceptances from approximately 800. If they admit as many as 50 too many students they will have to provide additional dormitory space. Let us find the probability that this will happen assuming that the acceptances of the students can be considered to be an independent trials process. We take as our estimate for the probability of an acceptance $p = \frac{800}{1250} = .64$. Then the expected number of acceptances is 800 and the standard deviation for the number of acceptances is $\sqrt{1250 \times .64 \times .36} \approx 17$. The probability that the number accepted is three standard deviations or 51 from the mean is approximately .0027. This probability takes into account a deviation above the mean or below the mean. Since in this case we are only interested in a deviation above the mean, the probability we desire is half of this or approximately .0013. Thus we see that it is highly unlikely that the college will have to have new dormitory space under the assumptions we have made.

We finish this discussion of the law of large numbers with some final remarks about the interpretation of this important theorem.

Of course no matter how large n is we cannot prevent the coin from coming up heads every time. If this were the case we would observe a fraction of heads equal to 1. However, this is not inconsistent with the theorem, since the probability of this happening is $(\frac{1}{2})^n$ which tends to 0 as n increases. Thus a fraction of 1 is always possible, but becomes increasingly unlikely.

The law of large numbers is often misinterpreted in the following manner. Suppose that we plan to toss the coin 1000 times and after 500 tosses we have already obtained 400 heads. Then we must obtain less than one-half heads in the remaining 500 tosses to have the fraction come out near $\frac{1}{2}$. It is tempting to argue that the coin therefore owes us some tails and it is more likely that tails will occur in the last 500 tosses. Of course this is nonsense, since the coin has no memory. The point is that something very unlikely has already happened in the first 500 tosses. The final result can therefore also be expected to be a result not predicted before the tossing began.

We could also argue that perhaps the coin is a biased coin, but this would make us predict more heads than tails in the future. Thus the law of averages, or the law of large numbers, should not give you great comfort if you have had a series of very bad hands dealt you in your last 100 poker hands. If the dealing is fair, you have the same chance as ever of getting a good hand.

Early attempts to define the probability p that success occurs on a single experiment sounded like this. If the experiment is repeated indefinitely, the fraction of successes obtained will tend to a number p, and this number p is called the probability of success on a single experiment. While this fails to be satisfactory as a definition of probability, the law of large numbers captures the spirit of this frequency concept of probability.

EXERCISES

1. In 64 tosses of an ordinary coin, what is the expected number of heads that turn up? What is the standard deviation for the number of heads that occur? [*Ans.* 32, 4.]
2. A die is loaded so that the probability of any face turning up is proportional to the number on that face. If the die is rolled 150 times, what is the expected number of times a three will turn up? What is the standard deviation for the number of threes that turn up? [*Ans.* $\frac{150}{7}$, $\frac{30}{7}$.]
3. Suppose the die in Exercise 2 is tossed 75 times. What is the expected number of times a three or a six will turn up? What is the standard deviation for the number of such throws?
4. An unknown coin is tossed 10,000 times and comes up heads 5100 times? Is it likely that the coin is fair?
5. In a large number of independent trials with probability p for success, what is the approximate probability that the number of successes will

deviate from the expected number by more than one standard deviation but less than two standard deviations? [*Ans.* .272.]

6. A farmer has found that, on the average, 1 percent of his 1000 apple trees die and must be replaced each year. The year after an atomic power plant begins operating nearby, 19 of the farmer's trees die. Should he suspect that his trees are dying due to other than natural causes?
7. Consider n independent trials with probability p for success. Let r and s be numbers such that $p < r < s$. What does the law of large numbers say about

$$\Pr\left[r < \frac{x}{n} < s\right]$$

as we increase n indefinitely? Answer the same question in the case that $r < p < s$.
8. Although 10 percent of those receiving Ph.D's in mathematics each year are women, all ten people hired by the mathematics department of a small college over the past five years have been men. Is there reason to suspect that the department is discriminating against women?
9. A researcher studying the effects of diet on heart disease notes that 15 percent of all men over 55 have heart disease, and 47 men out of a sample of 500 men over 55 on a low-cholesterol diet have heart disease. Is it reasonable for him to hypothesize that a low-cholesterol diet may reduce the incidence of heart disease?
10. What is the approximate probability that, in 10,000 throws of an ordinary coin, the number of heads which turn up lies between 4850 and 5150? What is the probability that the number of heads lies in the same interval, given that in the first 1900 throws there were 1600 heads?
11. Suppose we want to be 95 percent sure that the fraction of heads that turn up when a fair coin is tossed n times does not differ from $\frac{1}{2}$ by more than .01. How large should n be?

    [*Ans.* Approximately 10,000.]
12. A small college found that, while it was all female, 2 percent of its students majored in mathematics. After it became coeducational, the college had 10 math majors in its first mixed graduating class of 400 students. Is it likely that coeducation had some effect on the number of students electing to major in mathematics?
13. A preelection poll indicates that 55 percent of the voters will choose candidate A. When the election is over, candidate A has received 5200 of the 9900 votes cast. How accurate was the poll?
14. Suppose that for each roll of a fair die you lose \$1 when an odd number comes up and win \$1 when an even number comes up. Then after 40,000 rolls you can, with approximately 84 percent confidence, expect to have lost not more than how many dollars?
15. The Dartmouth computer, having nothing better to do, flipped a coin 1,000,000 times. It obtained 499,452 heads. Is this number reasonable?

*10 INDEPENDENT TRIALS WITH MORE THAN TWO OUTCOMES

By extending the results of Section 8, we shall study the case of independent trials in which we allow more than two outcomes. We assume that we have an independent trials process where the possible outcomes are $a_1, a_2, \ldots, a_k$, occurring with probabilities $p_1, p_2, \ldots, p_k$, respectively. We denote by

$$f(r_1, r_2, \ldots, r_k; p_1, p_2, \ldots, p_k)$$

the probability that, in

$$n = r_1 + r_2 + \ldots + r_k$$

such trials, there will be r_1 occurences of a_1, r_2 or a_2, etc. In the case of two outcomes this notation would be $f(r_1, r_2; p_1, p_2)$. In Section 8 we wrote this as $f(n, r; p) = f(n, r_1; p_1)$ since r_2 and p_2 are determined from n, r_1, and p_1. We shall indicate how this probability is found in general, but carry out the details only for a special case. We choose $k = 3$, and $n = 5$ for purposes of illustration. We shall find $f(1, 2, 2; p_1, p_2, p_3)$.

We show in Figure 18 enough of the tree for this process to indicate the branch probabilities for a path (heavy lines) corresponding to the outcomes a_2, a_3, a_1, a_2, a_3. The tree measure assigns weight $p_2 \cdot p_3 \cdot p_1 \cdot p_2 \cdot p_3 = p_1 \cdot p_2^2 \cdot p_3^2$ to this path.

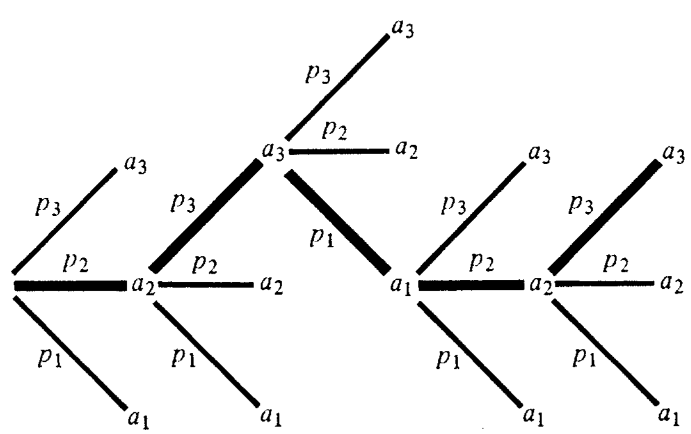


Figure 18

There are, of course, other paths through the tree corresponding to one occurrence of a_1, two of a_2, and two of a_3. However, they would all be assigned the same weight $p_1 \cdot p_2^2 \cdot p_3^2$, by the tree measure. Hence to find $f(1, 2, 2; p_1, p_2, p_3)$, we must multiply this weight by the number of paths having the specified number of occurrences of each outcome.

We note that the path a_2, a_3, a_1, a_2, a_3 can be specified by labeling the numbers 1 to 5 with the outcomes a_1, a_2, a_3. Thus trial 3 is labeled with outcome a_1, trials 1 and 4 are labeled with outcome a_2, and trials 2 and 5 are labeled with outcome a_3. Conversely, any such labeling of the numbers 1 to 5 which uses label a_1 once and labels a_2 and a_3 twice each corresponds to a unique path of the desired kind. Hence the number of such paths is the number of such labelings. But this is

$$\binom{5}{1,2,2} = \frac{5!}{1!\,2!\,2!}$$

(see Chapter 2, Section 5), so that the probability of one occurrence of a_1, two of a_2, and two of a_3 is

$$\binom{5}{1,2,2} \cdot p_1 \cdot p_2^2 \cdot p_3^2.$$

The above argument carried out in general leads, for the case of independent trials with outcomes $a_1, a_2, \ldots, a_k$ occurring with probabilities $p_1, p_2, \ldots, p_k$, to the following.

The probability for r_1 occurrences of a_1, r_2 occurrences of a_2, etc., is given by

$$f(r_1, r_2, \ldots, r_k; p_1, p_2, \ldots, p_k) = \binom{n}{r_1, r_2, \ldots, r_k} p_1^{r_1} \cdot p_2^{r_2} \ldots p_k^{r_k}.$$

EXAMPLE 1 A die is thrown 12 times. What is the probability that each number will come up twice? Here there are six outcomes, 1, 2, 3, 4, 5, 6 corresponding to the six sides of the die. We assign each outcome probability $\frac{1}{6}$. We are then asked for

$$f(2, 2, 2, 2, 2, 2; \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}),$$

which is

$$\binom{12}{2,2,2,2,2,2}\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2 = .0034....$$

EXAMPLE 2 Suppose that we have an independent trials process with four outcomes a_1, a_2, a_3, a_4 occurring with probability p_1, p_2, p_3, p_4, respectively. It might be that we are interested only in the probability that r_1 occurrences of a_1 and r_2 occurrences of a_2 will take place with no specification about the number of each of the other possible outcomes. To answer this question we simply consider a new experiment where the outcomes are $a_1, a_2, \bar{a}_3$. Here $\bar{a}_3$ corresponds to an occurrence of either a_3 or a_4 in our original experiment. The corresponding probabilities would be p_1, p_2, and $\bar{p}_3$ with $\bar{p}_3 = p_3 + p_4$. Let $\bar{r}_3 = n - (r_1 + r_2)$. Then our question is answered by finding the probability in our new experiment for r_1 occurrences of a_1, r_2 or a_2, and $\bar{r}_3$ of $\bar{a}_3$, which is

$$\binom{n}{r_1, r_2, \bar{r}_3} p_1^{r_1} p_2^{r_2} \bar{p}_3^{\bar{r}_3}.$$

The same procedure can be carried out for experiments with any number of outcomes where we specify the number of occurrences of such particular outcomes. For example, if a die is thrown ten times the probability that a one will occur exactly twice and a three exactly three times is given by

$$\binom{10}{2, 3, 5}\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^3\left(\frac{4}{6}\right)^5 = .043\ldots.$$

EXERCISES

1. Gypsies sometimes toss a thick coin for which heads and tails are equally likely, but which also has probability $\frac{1}{5}$ of standing on edge (i.e., neither heads nor tails). What is the probability of exactly two heads and three tails in five tosses of a gypsy coin? [*Ans.* $\frac{64}{625}$.]
2. Three horses, A, B, and C, compete in four races. Assuming that each horse has an equal chance in each race, what is the probability that A wins two races and B and C win one each? What is the probability that the same horse wins all four races? [*Ans.* $\frac{4}{27}$, $\frac{1}{27}$.]
3. Three children go into a restaurant where each gets either an ice-cream cone, a sundae, or a milkshake. Assuming that each gets an ice-cream cone twice as often as a milkshake and a sundae twice as often as an ice-cream cone, what is the probability that at least two of them order the same thing?
4. Assume that in a certain large college 40 percent of the students are freshmen, 30 percent are sophomores, 20 percent are juniors, and 10 percent are seniors. A committee of eight is chosen at random from the student body. What is the probability that there are equal numbers from each class on the committee?
5. If four dice are thrown, find the probability that there are 2 twos and 2 threes, given that all the outcomes are less than four. [*Ans.* $\frac{2}{27}$.]
6. Let us assume that when a batter comes to bat, he has probability .6 of being put out, .1 of getting a walk, .2 of getting a single, .1 of getting an extra-base hit. If he comes to bat five times in a game, what is the probability that
   (a) He gets a walk and a single (and three outs)? [*Ans.* $\frac{54}{625}$.]
   (b) He has a perfect day (no outs)? [*Ans.* $\frac{32}{3125}$.]
   (c) He gets a single, two extra base hits, and a walk (and one out)?
7. Assume that a single torpedo has a probability $\frac{1}{2}$ of sinking a ship, probability $\frac{1}{4}$ of damaging it, and probability $\frac{1}{4}$ of missing. Assume further that two damaging shots are sufficient to sink a ship. What is the probability that four torpedoes will succeed in sinking a ship? [*Ans.* $\frac{251}{256}$.]
8. A hiker is planning to make a three-day trip to the mountains. He estimates that on a given day it is clear with probability $\frac{1}{3}$, is cloudy with probability $\frac{1}{2}$, and rains with probability $\frac{1}{6}$. He will consider the trip enjoyable if he does not get rained on and if he has at least two clear days. Assuming the weather on a given day is independent of previous weather,
   (a) Find the probability that he enjoys the trip.

(b) Given that at least one day was not clear, what is probability that he enjoyed the trip?

9. Let us assume that in a World Series game a batter has probability $\frac{1}{4}$ of getting no hits, $\frac{1}{2}$ of getting one hit, and $\frac{1}{4}$ of getting two hits, assuming that the probability of getting more than two hits is negligible. In a four-game World Series, find the probability that the batter gets
   (a) Exactly two hits.
   (b) Exactly three hits.
   (c) Exactly four hits.
   (d) Exactly five hits.
   (e) Fewer than two hits or more than five.

   [*Ans.* $\frac{7}{64}$; $\frac{7}{32}$; $\frac{35}{128}$; $\frac{7}{32}$; $\frac{23}{128}$.]

10. Jones, Smith, and Green live in the same house. The mailman has observed that on the average Jones receives twice as much mail as Green and three times as much as Smith. If he has four letters for this house, what is the probability that each man receives at least one letter?

11. Assume that in a certain course the probability that a student chosen at random will get an A is .1, that he will get a B is .2, that he will get a C is .4, that he will get a D is .2, and that he will get an E is .1. What distribution of grades is most likely in the case of four students? [*Ans.* One B, two C's, one D.]

12. A professor decides that he will fail any student who misses more than two classes in a given week or is absent or walks in late every day in a given week. The class meets four times a week. A particular student's sleeping habits are such that he misses class $\frac{1}{5}$ of the time, walks in late $\frac{3}{10}$ of the time, and arrives before the start of class $\frac{1}{2}$ of the time.
   (a) What is the probability that he is failed during the first week of the course?
   (b) What is the probability that he is failed during the first two weeks of the course? (Note that when he begins the second week, his performance the first week does not matter any more.)

   [*Ans.* .151.]

11 EXPECTED VALUE

In this section we shall discuss the concept of expected value. Although it originated in the study of gambling games, it enters into almost any detailed probabilistic discussion.

Definition If in an experiment the possible outcomes are numbers, $a_1, a_2, \ldots, a_k$, occurring with probability $p_1, p_2, \ldots, p_k$, then the *expected value* is defined to be

$$E = a_1p_1 + a_2p_2 + \ldots + a_kp_k.$$

The term "expected value" is not to be interpreted as the value that will necessarily occur on a single experiment. For example, if a person bets \$1 that a head will turn up when a coin is thrown, he may either win \$1 or lose \$1. His expected value is $(1)(\frac{1}{2}) + (-1)(\frac{1}{2}) = 0$, which is not one of the possible outcomes. The term "expected value" had its origin in the following consideration. If we repeat an experiment with expected value E a large number of times, and if we expect a_1 a fraction p_1 of the time, a_2 a fraction p_2 of the time, etc., then the average that we expect per experiment is E. In particular, in a gambling game E is interpreted as the average winning expected in a large number of plays. Here the expected value is often taken as the value of the game to the player. If the game has a positive expected value, the game is said to be favorable; if the game has expected value zero it is said to be fair; and if it has negative expected value it is described as unfavorable. These terms are not to be taken too literally, since many people are quite happy to play games that, in terms of expected value, are unfavorable. For instance, the buying of life insurance may be considered an unfavorable game which most people choose to play.

EXAMPLE 1 For the first example of the application of expected value we consider the game of roulette as played at Monte Carlo. There are several types of bets which the gambler can make, and we consider two of these.

The wheel has the number 0 and the numbers from 1 to 36 marked on equally spaced slots. The wheel is spun and a ball comes to rest in one of these slots. If the player puts a stake, say \$1, on a given number, and the ball comes to rest in this slot, then he receives from the croupier 36 times his stake, or \$36. The player wins \$35 with probability $\frac{1}{37}$ and loses \$1 with probability $\frac{36}{37}$. Hence his expected winnings are

$$\frac{35}{37} - 1 \cdot \frac{36}{37} = -\frac{1}{37} = -.027.$$

This can be interpreted to mean that in the long run he can expect to lose about 2.7 percent of his stakes.

A second way to play is the following. A player may bet on "red" or "black." The numbers from 1 to 36 are evenly divided between the two colors. If a player bets on "red" and a red number turns up, he receives twice his stake. If a black number turns up, he loses his stake. If 0 turns up, then the wheel is spun until it stops on a number different from 0. If this is black, the player loses; but if it is red, he receives only his original stake, not twice it. For this type of play, the gambler wins \$1 with probability $\frac{18}{37}$, breaks even with probability $\frac{1}{2} \cdot \frac{1}{37} = \frac{1}{74}$, and loses \$1 with probability $\frac{18}{37} + \frac{1}{2} \cdot \frac{1}{37} = \frac{37}{74}$. Hence his expected winning is

$$1 \cdot \frac{18}{37} + 0 \cdot \frac{1}{74} - 1 \cdot \frac{37}{74} = -.0135.$$

In this case the player can expect to lose about 1.35 percent of his stakes in the long run. Thus the expected loss in this case is only half as great as in the previous case.

EXAMPLE 2 A player rolls a die and receives a number of dollars corresponding to the number of dots on the face which turns up. What should the player pay for playing, to make this a fair game? To answer this question, we note that the player wins 1, 2, 3, 4, 5 or 6 dollars, each with probability $\frac{1}{6}$. Hence, his expected winning is

$$1(\tfrac{1}{6}) + 2(\tfrac{1}{6}) + 3(\tfrac{1}{6}) + 4(\tfrac{1}{6}) + 5(\tfrac{1}{6}) + 6(\tfrac{1}{6}) = 3\tfrac{1}{2}.$$

Thus if he pays \$3.50, his expected winnings will be zero.

EXAMPLE 3 What is the expected number of successes in the case of four independent trials with probability $\frac{1}{3}$ for success? We know that the probability of x successes is $\binom{4}{x}\left(\frac{1}{3}\right)^x\left(\frac{2}{3}\right)^{4-x}$. Thus

$$\begin{aligned} E = 0 \cdot \binom{4}{0}\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^4 + 1 \cdot \binom{4}{1}\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^3 + 2 \cdot \binom{4}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^2 \\ + 3 \cdot \binom{4}{3}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^1 + 4 \cdot \binom{4}{4}\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^0 \\ = 0 + \frac{32}{81} + \frac{48}{81} + \frac{24}{81} + \frac{4}{81} = \frac{108}{81} = \frac{4}{3}. \end{aligned}$$

In general, it can be shown that in n trials with probability p for success, the expected number of successes is np.

EXAMPLE 4 In the game of craps a pair of dice is rolled by one of the players. If the sum of the spots shown is 7 or 11, he wins. If it is 2,3, or 12, he loses. If it is another sum, he must continue rolling the dice until he either repeats the same sum or rolls a 7. In the former case he wins, in the latter he loses. Let us suppose that he wins or loses \$1. Then the two possible outcomes are $+1$ and -1. We shall compute the expected value of the game. First we must find the probability that he will win.

We represent the possibilities by a two-stage tree shown in Figure 19. While it is theoretically possible for the game to go on indefinitely, we do not consider this possibility. This means that our analysis applies only to games which actually stop at some time.

The branch probabilities at the first stage are determined by thinking of the 36 possibilities for the throw of the two dice as being equally likely and taking in each case the fraction of the possibilities which correspond to the branch as the branch probability. The probabilities for the branches at the second level are obtained as follows. If, for example, the first outcome was a 4, then when the game ends, a 4 or 7 must have occurred. The possible outcomes for the dice were

$$\{(3,1), (1,3), (2,2), (4,3), (3,4), (2,5), (5,2), (1,6), (6,1)\}.$$

Again we consider these possibilities to be equally likely and assign to the branch considered the fraction of the outcomes which correspond to this

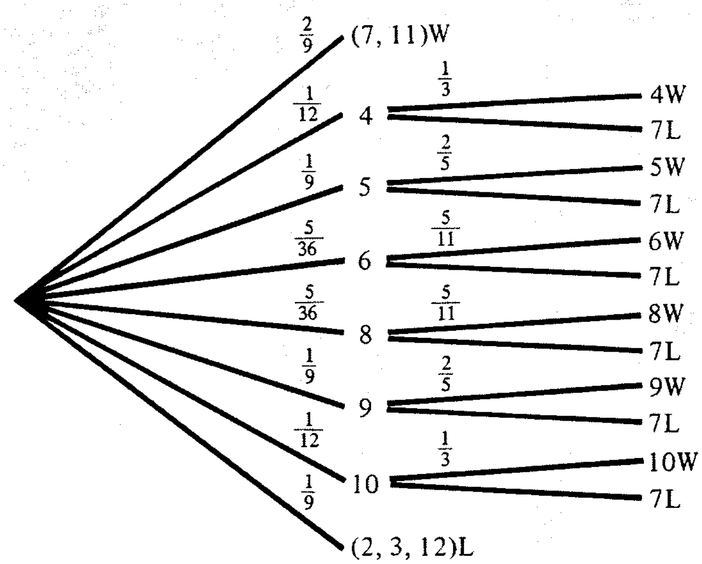


Figure 19

branch. Thus to the 4 branch we assign a probability $\frac{3}{9} = \frac{1}{3}$. The other branch probabilities are determined in a similar way. Having the tree measure assigned, to find the probability of a win we must simply add the weights of all paths leading to a win. If this is done, we obtain $\frac{244}{495}$. Thus the player's expected value is

$$1 \cdot (\tfrac{244}{495}) + (-1) \cdot (\tfrac{251}{495}) = -\tfrac{7}{495} = -.0141.$$

Hence he can expect to lose 1.41 percent of his stakes in the long run. It is interesting to note that this is just slightly less favorable than his losses in betting on "red" in roulette.

EXERCISES

1. If 13 coins are thrown, what is the expected number of heads that will turn up? [*Ans.* $\frac{13}{2}$.]
2. An urn contains two black and three white balls. Balls are successively drawn from the urn without replacement until a white ball is obtained. Find the expected number of draws required. Do the same for the case of four black and six white balls.
3. Suppose that A tosses three coins and receives \$8 if three heads appear, \$4 if two heads appear, \$2 if one head appears, and \$1 if no heads appear. What is the expected value of the game to him? [*Ans.* \$$3\frac{3}{8}$.]
4. Two players, A and B, play the following dice game. A shakes a die that has three 2's and three 3's on the faces, while B shakes a die painted with four 1's and two 6's. Find A's expected winning (or loss) for each of the following sets of rules.
   (a) The player that shakes the lower number pays the other player \$2. [*Ans.* \$$\frac{2}{3}$.]

(b) The player that shakes the lower number pays the other player a number of dollars equal to the difference between the two outcomes.

(c) The player that shakes the lower number pays the other player a number of dollars equal to the number shaken by the other player.

5. A coin is thrown until the second time a head comes up or until three tails in a row occur. Find the expected number of times the coin is thrown. [*Ans.* $\frac{105}{32}$.]

6. A man wishes to purchase a five-cent newspaper. He has in his pocket one dime and five pennies. The newsman offers to let him have the paper in exchange for one coin drawn at random from the customer's pocket.

(a) Is this a fair proposition and, if not, to whom is it favorable? [*Ans.* Favorable to man.]

(b) Answer the same questions as in (a) assuming that the newsman demands two coins drawn at random from the customer's pocket. [*Ans.* Fair proposition.]

7. Referring to Exercise 17 of Chapter 1, Section 5, assuming that each speaker chooses his topic at random from those available to him,

(a) Find the expected number of speeches on brotherhood during a given program.

(b) Find the smallest number of programs that we would have to attend in order that the expected value of the number of speeches on integrity that we hear is to be at least five.

8. Prove that if the expected value of a given experiment is E, and if a constant c is added to each of the outcomes, the expected value of the new experiment is $E + c$.

9. Prove that, if the expected value of a given experiment is E, and if each of the possible outcomes is multiplied by a constant k, the expected value of the new experiment is $k \cdot E$.

10. A bets x cents against B's 78 cents that, if two cards are dealt from a shuffled pack of ordinary playing cards, both cards will be of the same color. What value of x will make this bet fair?

11. Betting on "red" in roulette can be described roughly as follows. We win with probability .49, get our money back with probability .01, and lose with probability .50. Draw the tree for three plays of the game, and compute (to three decimals), the probability of each path. What is the probability that we are ahead at the end of three bets? [*Ans.* .485.]

12. Assume that the odds are $r:s$ that a certain statement will be true. If a man receives s dollars if the statement turns out to be true, and gives r dollars if not, what is his expected winning?

13. In the World Series, we assume that the stronger team has probability .6 of winning each game. In this case the probabilities of the series

lasting 4, 5, 6, or 7 games are .16, .27, .30, and .28, respectively. What is the expected length of the World Series? [*Ans.* 5.75.]

14. An office worker buys root beer from a defective machine which gives him orange soda instead of root beer $\frac{1}{5}$ of the time. He keeps buying soda until he gets his root beer or runs out of dimes. How many dimes must he carry with him every day if he wants to be able to expect to get root beer on at least 99 out of 100 days he uses the machine? [*Ans.* 3.]

15. Suppose that in roulette at Monte Carlo we place 50 cents on "red" and 50 cents on "black." What is the expected value on the game? Is this better or worse than placing $1 on "red"? Which of the two games would it be more desirable to play?

16. Suppose that we modify the game of craps as follows: On a 7 or 11 the player wins $1; on a 2, 3, or 12 he loses $$x$; otherwise the game is as usual, all losses being $$x$. Find the expected value of the new game, and determine the value of x for which it becomes a fair game.

17. A gambler is given the choice of playing one of the following games. Either he pays $10 and throws three dice, receiving in return the number of dollars equal to the sum of the three outcomes, or he pays $12 and throws two dice, receiving in return the number of dollars equal to the product of the two outcomes. Which game should he play?

18. A pair of dice is rolled. Each die has the number 1 on two opposite faces, the number 2 on two opposite faces, and the number 3 on two opposite faces. The "roller" wins a dollar

(i) if the sum of 4 occurs on the first roll; or

(ii) if the sum of 3 or 5 occurs on the first roll and the same sum occurs on a subsequent roll before the sum of 4 occurs.

Otherwise he loses a dollar.

(a) What is the probability that the person rolling the dice wins?

(b) What is the expected value of the game? [*Ans.* (a) $\frac{23}{45}$; (b) $\frac{1}{45}$.]

12 MARKOV CHAINS

In this section we shall study a more general kind of process than the ones considered in the last three sections.

We assume that we have a sequence of experiments with the following properties. The outcome of each experiment is one of a finite number of possible outcomes $a_1, a_2, \ldots, a_r$. It is assumed that the probability of outcome a_j on any given experiment is not necessarily independent of the outcomes of previous experiments but depends at most upon the outcome of the immediately preceding experiment. We assume that there are given numbers p_{ij} which represent the probability of outcome a_j on any given experiment, given that outcome a_i occurred on the preceding experiment. The outcomes $a_1, a_2, \ldots, a_r$ are called *states,* and the numbers p_{ij} are called *transition probabilities*. If we assume that the process begins in some particular state, then we have enough information to determine the tree measure for the process and can calculate probabilities of statements relating to the overall sequence of experiments. A process of the above kind is called a *Markov chain process.*

The transition probabilities can be exhibited in two different ways. The first way is that of a square array. For a Markov chain with states a_1, a_2, and a_3, this array is written as

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix}.$$

Such an array is a special case of a *matrix*. Matrices are of fundamental importance to the study of Markov chains as well as being important in the study of other branches of mathematics. They will be studied in detail in the next chapter.

A second way to show the transition probabilities is by a *transition diagram*. Such a diagram is illustrated for a special case in Figure 20. The arrows from each state indicate the possible states to which a process can move from the given state.

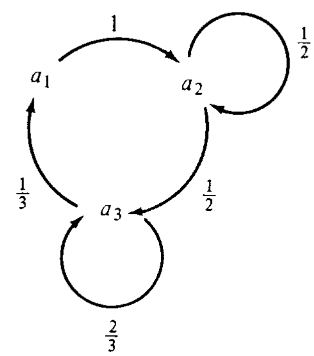


Figure 20

The matrix of transition probabilities which corresponds to this diagram is the matrix

$$P = \begin{matrix} & \begin{matrix} a_1 & a_2 & a_3 \end{matrix} \\ \begin{matrix} a_1 \\ a_2 \\ a_3 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{2}{3} \end{pmatrix} \end{matrix}.$$

An entry of 0 indicates that the transition is impossible.

Notice that in the matrix P the sum of the elements of each row is 1. This must be true in any matrix of transition probabilities, since the elements of the ith row represent the probabilities for all possibilities when the process is in state a_i.

The kind of problem in which we are most interested in the study of Markov chains is the following. Suppose that the process starts in state i.

What is the probability that after n steps it will be in state j? We denote this probability by $p_{ij}^{(n)}$. Notice that we do *not* mean by this the nth power of the number p_{ij}. We are actually interested in this probability for all possible starting positions i and all possible terminal positions j. We can represent these numbers conveniently again by a matrix. For example, for n steps in a three-state Markov chain we write these probabilities as the matrix

$$P^{(n)} = \begin{pmatrix} p_{11}^{(n)} & p_{12}^{(n)} & p_{13}^{(n)} \\ p_{21}^{(n)} & p_{22}^{(n)} & p_{23}^{(n)} \\ p_{31}^{(n)} & p_{32}^{(n)} & p_{33}^{(n)} \end{pmatrix}.$$

EXAMPLE 1 Let us find for a Markov chain with transition probabilities indicated in Figure 20 the probability of being at the various possible states after three steps, assuming that the process starts at state a_1. We find these probabilities by constructing a tree and a tree measure as in Figure 21.

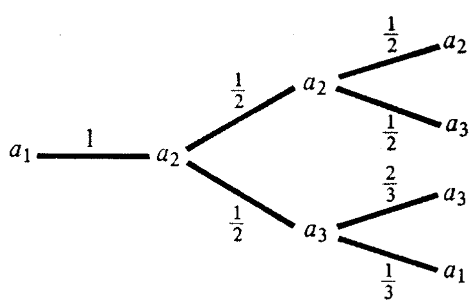


Figure 21

The probability $p_{13}^{(3)}$, for example, is the sum of the weights assigned by the tree measure to all paths through our tree which end at state a_3. That is,

$$1 \cdot \tfrac{1}{2} \cdot \tfrac{1}{2} + 1 \cdot \tfrac{1}{2} \cdot \tfrac{2}{3} = \tfrac{7}{12}.$$

Similarly,

$$p_{12}^{(3)} = 1 \cdot \tfrac{1}{2} \cdot \tfrac{1}{2} = \tfrac{1}{4} \quad \text{and} \quad p_{11}^{(3)} = 1 \cdot \tfrac{1}{2} \cdot \tfrac{1}{3} = \tfrac{1}{6}.$$

By constructing a similar tree measure, assuming that we start at state a_2, we could find $p_{21}^{(3)}$, $p_{22}^{(3)}$, and $p_{23}^{(3)}$. The same is true for $p_{31}^{(3)}$, $p_{32}^{(3)}$, and $p_{33}^{(3)}$. If this is carried out (see Exercise 7) we can write the results in matrix form as follows:

$$P^{(3)} = \begin{matrix} a_1 \\ a_2 \\ a_3 \end{matrix} \overset{\begin{matrix} a_1 & a_2 & a_3 \end{matrix}}{\begin{pmatrix} \frac{1}{6} & \frac{1}{4} & \frac{7}{12} \\ \frac{7}{36} & \frac{7}{24} & \frac{37}{72} \\ \frac{4}{27} & \frac{7}{18} & \frac{25}{54} \end{pmatrix}}.$$

Again the rows add up to 1, corresponding to the fact that if we start at a given state we must reach some state after three steps. Notice now that all the elements of this matrix are positive, showing that it is possible to reach any state from any state in three steps. In the next chapter we shall develop a simple method of computing $P^{(n)}$.

EXAMPLE 2 Suppose that we are interested in studying the way in which a given state votes in a series of national elections. We wish to make long-term predictions and so shall not consider conditions peculiar to a particular election year. We shall base our predictions only on past history of the outcomes of the elections, Republican or Democratic. It is clear that a knowledge of these past results would influence our predictions for the future. As a first approximation, we assume that the knowledge of the past beyond the last election would not cause us to change the probabilities for the outcomes on the next election. With this assumption we obtain a Markov chain with two states R and D and matrix of transition probabilities

$$\begin{array}{c} \\ R \\ D \end{array} \begin{array}{c} \begin{array}{cc} R & D \end{array} \\ \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix} \end{array}.$$

The numbers a and b could be estimated from past results as follows. We could take for a the fraction of the previous years in which the outcome has changed from Republican in one year to Democratic in the next year, and for b the fraction of reverse changes.

We can obtain a better approximation by taking into account the previous two elections. In this case our states are RR, RD, DR, and DD, indicating the outcome of two successive elections. Being in state RR means that the last two elections were Republican victories. If the next election is a Democratic victory, we will be in state RD. If the election outcomes for a series of years is $DDDRDRR$, then our process has moved from state DD to DD to DR to RD to DR, and finally to RR. Notice that the first letter of the state to which we move must agree with the second letter of the state from which we came, since these refer to the same election year. Our matrix of transition probabilities will then have the following form:

$$\begin{array}{c} \\ RR \\ DR \\ RD \\ DD \end{array} \begin{array}{c} \begin{array}{cccc} RR & DR & RD & DD \end{array} \\ \begin{pmatrix} 1-a & 0 & a & 0 \\ b & 0 & 1-b & 0 \\ 0 & 1-c & 0 & c \\ 0 & d & 0 & 1-d \end{pmatrix} \end{array}.$$

Again the numbers a, b, c, and d would have to be estimated. The study of this example is continued in Chapter 4, Section 7.

EXAMPLE 3 The following example of a Markov chain has been used in physics as a simple model for diffusion of gases. We shall see later that a similar model applies to an idealized problem in changing populations.

We imagine n black balls and n white balls which are put into two urns so that there are n balls in each urn. A single experiment consists in choosing a ball from each urn at random and putting the ball obtained from the first urn into the second urn, and the ball obtained from the second urn into the first. We take as state the number of black balls in the first urn. If at any time we know this number, then we know the exact composition of each urn. That is, if there are j black balls in urn 1, there must be $n - j$ black balls in urn 2, $n - j$ white balls in urn 1, and j white balls in urn 2. If the process is in state j, then after the next exchange it will be in state $j - 1$, if a black ball is chosen from urn 1 and a white ball from urn 2. It will be in state j if a ball of the same color is drawn from each urn. It will be in state $j + 1$ if a white ball is drawn from urn 1 and a black ball from urn 2. The transition probabilities are then given by (see Exercise 12)

$$\begin{aligned} p_{jj-1} &= \left(\frac{j}{n}\right)^2 && j > 0 \\ p_{jj} &= \frac{2j(n-j)}{n^2} \\ p_{jj+1} &= \left(\frac{n-j}{n}\right)^2 && j < n \\ p_{jk} &= 0 && \text{otherwise.} \end{aligned}$$

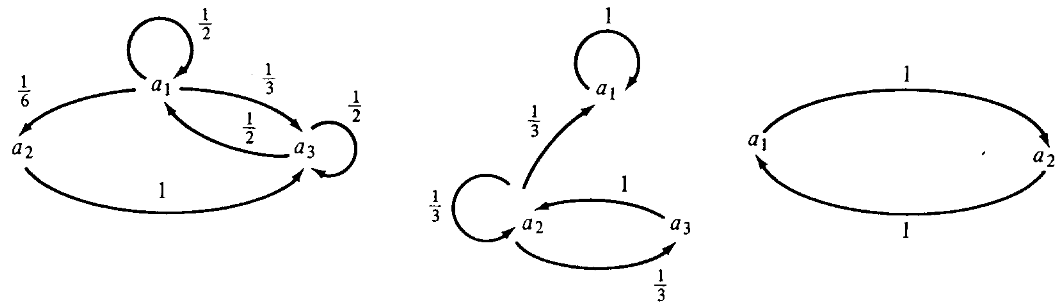
A physicist would be interested, for example, in predicting the composition of the urns after a certain number of exchanges have taken place. Certainly any predictions about the early stages of the process would depend upon the initial composition of the urns. For example, if we started with all black balls in urn 1, we would expect that for some time there would be more black balls in urn 1 than in urn 2. On the other hand, it might be expected that the effect of this initial distribution would wear off after a large number of exchanges. We shall see later, in Chapter 4, Section 7, that this is indeed the case.

EXERCISES

1. Draw a transition diagram for the Markov chains with transition probabilities given by the following matrices:

$$\begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{3} & 0 & \frac{2}{3} \end{pmatrix}, \quad \begin{pmatrix} \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \frac{1}{12} & \frac{1}{6} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}.$$

2. Give the matrix of transition probabilities corresponding to the following transition diagrams:



3. What is the matrix of transition probabilities for the Markov chain in Example 3, for the case of two white balls and two black balls?

4. Find the matrix $P^{(2)}$ for the Markov chain determined by the matrix of transition probabilities

$$P = \begin{pmatrix} \frac{1}{4} & \frac{3}{4} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}. \qquad \left[Ans. \begin{pmatrix} \frac{5}{16} & \frac{11}{16} \\ \frac{11}{36} & \frac{25}{36} \end{pmatrix}.\right]$$

5. Find the matrices $P^{(2)}$, $P^{(3)}$, $P^{(4)}$ for the Markov chain determined by the transition probabilities

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Find the same for the Markov chain determined by the matrix

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

6. (a) What is the relationship between independent trials processes and Markov chains?

[*Ans.* Every independent trials process, given a particular starting state, is a Markov chain.]

(b) Set up a transition diagram and matrix for tossing a fair coin. Find $P^{(2)}$ and $P^{(3)}$.

(c) Repeat part (b) for a coin that comes up heads with probability $\frac{3}{4}$.

7. Referring to the Markov chain with transition probabilities indicated in Figure 21, construct the tree measures and determine the values of

$$p_{21}^{(3)}, p_{22}^{(3)}, p_{23}^{(3)} \quad \text{and} \quad p_{31}^{(3)}, p_{32}^{(3)}, p_{33}^{(3)}.$$

8. Suppose that a Markov chain has two states, a_1 and a_2, and transition probabilities given by the matrix

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix}.$$

By means of a separate chance device we choose a state in which to start the process. This device chooses a_1 with probability $\frac{1}{2}$ and a_2 with probability $\frac{1}{2}$. Find the probability that the process is in state a_1 after the first step. Answer the same question in the case that the device chooses a_1 with probability $\frac{1}{3}$ and a_2 with probability $\frac{2}{3}$. [*Ans.* $\frac{3}{8}$; $\frac{1}{3}$.]

9. A certain calculating machine uses only the digits 0 and 1. It is supposed to transmit one of these digits through several stages. However, at every stage there is a probability p that the digit which enters this stage will be changed when it leaves. We form a Markov chain to represent the process of transmission by taking as states the digits 0 and 1. What is the matrix of transition probabilities?

10. For the Markov chain in Exercise 9, draw a tree and assign a tree measure, assuming that the process begins in state 1 and moves through three stages of transmission. What is the probability that the machine after three stages produces the digit 1, i.e., the correct digit? What is the probability that the machine changed the digit from 1 but ended up with a 1 after three stages?

11. A student has a class that meets on Monday, Wednesday, and Friday. He decides on any one of these days to go to class with a probability that depends only on whether or not he went to the last class. If he did go to class on one day, he goes to the next class with probability $\frac{1}{2}$. If he did not go to one class, he goes to the next class with probability $\frac{3}{4}$. Set up the matrix of transition probabilities and find the probability that if he went to class on Monday, he will also attend the class on Friday of that week.

12. Explain why the transition probabilities given in Example 3 are correct.

13. Assume that a man's profession can be classified as professional, skilled laborer, or unskilled laborer. Assume that of the sons of professional men 80 percent are professionals, 15 percent are skilled laborers, and 5 percent are unskilled laborers. In the case of sons of skilled laborers, 50 percent are skilled laborers, 25 percent are professionals, and 25 percent are unskilled laborers. Finally, in the case of unskilled laborers, 40 percent of the sons are unskilled laborers and 30 percent each are in the other two categories. Assume that every man has a son, and form a Markov chain by following a given family through several generations. Set up the matrix of transition probabilities. Find the probability that the grandson of a skilled laborer is a professional man. [*Ans.* .4.]

14. In Exercise 13 we assumed that every man has a son. Assume instead that the probability a man has a son is .75. Form a Markov chain with four states. The first three states are as in Exercise 13, and the fourth state is such that the process enters it if a man has no son, and that the state cannot be left. This state represents families whose male line has died out. Find the matrix of transition probabilities and find the probability that a skilled laborer has a grandson who is a professional man. [*Ans.* .225.]

15. In another model for diffusion, it is assumed that there are two urns which together contain N balls numbered from 1 to N. Each second a number from 1 to N is chosen at random, and the ball with the corresponding number is moved to the other urn. For $N = 4$ set up a Markov chain by taking as state the number of balls in urn 1. Find the transition matrix.

16. In a two-player game, each player starts out with three chips. Two dice are then tossed. If the sum of the numbers tossed is less than 7, player A gets a chip from player B. If the total is greater than 7, B gets a chip from A. If a 7 is tossed, either the player with less chips gets one from the other player or, if the players are even, they remain even. The game continues until one player runs out of chips. Using as states the number of chips that A has, set up the transition matrix. (Assume that once A gets to the state 0 or 6 he stays there with probability 1 next time.)

17. In Exercise 16, find the following:

(a) The probability that B loses in three turns. $\left[Ans.\ \frac{125}{1728}.\right]$

(b) The probability that B loses in three turns starting with two chips left.

(c) $p_{24}^{(3)}$.

(d) The probability that A loses in five or less turns.

$$\left[Ans.\ \frac{34{,}625}{12^5} \approx .139.\right]$$

*13 GAMBLER'S RUIN

In this section we shall study a particular Markov chain, which is interesting in itself and has far-reaching applications. Its name, "gambler's ruin," derives from one of its many applications. In the text we shall describe the chain from the gambling point of view, but in the exercises we shall present several other applications.

Let us suppose that you are gambling against a professional gambler or gambling house. You have selected a specific game to play, on which you have probability p of winning. The gambler has made sure that the game is favorable to him, so that $p < \frac{1}{2}$. However, in most situations p will be close to $\frac{1}{2}$. (The cases $p = \frac{1}{2}$ and $p > \frac{1}{2}$ are considered in the exercises.)

At the start of the game you have A dollars, and the gambler has B dollars. You bet \$1 on each game, and play until one of you is ruined. What is the probability that you will be ruined? Of course, the answer depends on the exact values of p, A, and B. We shall develop a formula for the ruin-probability in terms of these three given numbers.

First we shall set the problem up as a Markov chain. Let $N = A + B$, the total amount of money in the game. As states for the chain we choose the numbers $0, 1, 2, \ldots, N$. At any one moment the position of the chain is the amount of money *you* have. The initial position is shown in Figure 22.

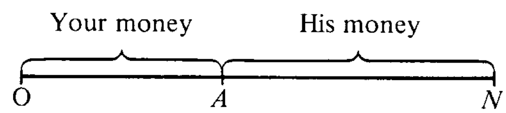


Figure 22

If you win a game, your money increases by \$1, and the gambler's fortune decreases by \$1. Thus the new position is one state to the right of the previous one. If you lose a game, the chain moves one step to the left. Thus at any step there is probability p of moving one step to the right, and probability $q = 1 - p$ of one step to the left. Since the probabilities for the next position are determined by the present position, it is a Markov chain.

If the chain reaches 0 or N, we stop. When 0 is reached, you are ruined. When N is reached, you have all the money, and you have ruined the gambler. We shall be interested in the probability of *your* ruin, i.e., the probability of reaching 0.

Let us suppose that p and N are fixed. We actually want the probability of ruin when we start at A. However, it turns out to be easier to solve a problem that appears much harder: Find the ruin-probability for every possible starting position. For this reason we introduce the notation x_i, to stand for the probability of your ruin if you start in position i (that is, if you have i dollars).

Let us first solve the problem for the case $N = 5$. We have the unknowns x_0, x_1, x_2, x_3, x_4, and x_5. Suppose that we start at position 2. The chain moves to 3, with probability p, or to 1, with probability q. Thus

$$\Pr[\text{ruin} \mid \text{start at } 2] = \Pr[\text{ruin} \mid \text{start at } 3] \cdot p + \Pr[\text{ruin} \mid \text{start at } 1] \cdot q,$$

using the conditional probability formula, with a set of two alternatives. But once it has reached state 3, a Markov chain behaves just as if it had been started there. Thus

$$\Pr[\text{ruin} \mid \text{start at } 3] = x_3.$$

And, similarly,

$$\Pr[\text{ruin} \mid \text{start at } 1] = x_1.$$

We obtain the key relation

$$x_2 = px_3 + qx_1.$$

We can modify this as follows: using $p + q = 1$, we have

$$\begin{aligned} x_2 = (p + q)x_2 &= px_3 + qx_1 \\ p(x_2 - x_3) &= q(x_1 - x_2) \\ x_1 - x_2 &= r(x_2 - x_3), \end{aligned}$$

where $r = p/q$, and hence $r < 1$. When we write such an equation for each of the four "ordinary" positions, we obtain

(1)
$$\begin{aligned} x_0 - x_1 &= r(x_1 - x_2) \\ x_1 - x_2 &= r(x_2 - x_3) \\ x_2 - x_3 &= r(x_3 - x_4) \\ x_3 - x_4 &= r(x_4 - x_5). \end{aligned}$$

We must still consider the two extreme positions. Suppose that the chain reaches 0. Then you are ruined, hence the probability of your ruin is 1. While if the chain reaches $N = 5$, the gambler drops out of the game, and you can't be ruined. Thus

(2)
$$x_0 = 1, \qquad x_5 = 0.$$

If we substitute the value of x_5 in the last equation of (1), we have $x_3 - x_4 = rx_4$. This in turn may be substituted in the previous equation, etc. We thus have the simpler equations

(3)
$$\begin{aligned} x_4 &= 1 \cdot x_4 \\ x_3 - x_4 &= rx_4 \\ x_2 - x_3 &= r^2x_4 \\ x_1 - x_2 &= r^3x_4 \\ x_0 - x_1 &= r^4x_4. \end{aligned}$$

Let us add all the equations. We obtain

$$x_0 = (1 + r + r^2 + r^3 + r^4)x_4.$$

From (2) we have that $x_0 = 1$. We also use the simple identity

$$(1 - r)(1 + r + r^2 + r^3 + r^4) = 1 - r^5,$$

which implies

$$1 + r + r^2 + r^3 + r^4 = \frac{1 - r^5}{1 - r}.$$

And then we solve for x_4:

$$x_4 = \frac{1 - r}{1 - r^5}.$$

If we add the first two equations in (3), we have that $x_3 = (1 + r)x_4$. Similarly, adding the first three equations, we solve for x_2, and adding the

Ruin-probabilities for $p = .45, .48, .49, .495$.

$p = .45$

A \ B	1	5	10	20	50
1	.550	.905	.973	.997	1
5	.260	.732	.910	.988	1
10	.204	.666	.881	.984	1
20	.185	.638	.868	.982	1
50	.182	.633	.866	.982	1

$p = .48$

A \ B	1	5	10	20	50
1	.520	.865	.941	.981	.999
5	.202	.599	.788	.923	.994
10	.131	.472	.690	.878	.990
20	.095	.381	.606	.832	.985
50	.078	.334	.555	.801	.982

$p = .49$

A \ B	1	5	10	20	50
1	.510	.850	.926	.969	.994
5	.184	.550	.731	.871	.972
10	.110	.402	.599	.788	.951
20	.069	.287	.472	.690	.921
50	.045	.204	.363	.586	.881

$p = .495$

A \ B	1	5	10	20	50
1	.505	.842	.918	.961	.989
5	.175	.525	.699	.838	.948
10	.100	.367	.550	.731	.905
20	.058	.242	.402	.599	.839
50	.031	.143	.259	.438	.731

Figure 23

first four equations we obtain x_1. We now have our entire solution:

$$(4) \qquad x_1 = \frac{1 - r^4}{1 - r^5}, \quad x_2 = \frac{1 - r^3}{1 - r^5}, \quad x_3 = \frac{1 - r^2}{1 - r^5}, \quad x_4 = \frac{1 - r}{1 - r^5}.$$

The same method will work for any value of N. And it is easy to guess from (4) what the general solution looks like. If we want x_A, the answer is a fraction like those in (4). In the denominator the exponent of r is always N. In the numerator the exponent is $N - A$, which equals B. Thus the ruin-probability is

$$(5) \qquad x_A = \frac{1 - r^B}{1 - r^N}.$$

We recall that A is the amount of money you have, B is the gambler's stake, $N = A + B$, p is your probability of winning a game, and $r = p/(1 - p)$.

In Figure 23 we show some typical values of the ruin-probability. Some of these are quite startling. If the probability of p is as low as .45 (odds against you on each game 11:9) and the gambler has \$20 to put up, you are almost sure to be ruined. Even in a nearly fair game, say $p = .495$, with each of you having \$50 to start with, there is a .731 chance for your ruin.

It is worth examining the ruin-probability formula, (5), more closely. Since the denominator is always less than 1, your probability of ruin is at least $1 - r^B$. This estimate does not depend on how much money you have, only on p and B. Since r is less than 1, by making B large enough we can make r^B practically 0, and hence make it almost certain that you will be ruined.

Suppose, for example, that a gambler wants to have probability .999 of ruining you. (You can hardly call him a gambler under those circumstances!) He must make sure that $r^B < .001$. For example, if $p = .495$, the gambler needs \$346 to have probability .999 of ruining you, even if you are a millionaire. If $p = .48$, he needs only \$87. And even for the almost fair game with $p = .499$, \$1727 will suffice.

There are two ways that gamblers achieve this goal. Small gambling houses will fix the odds quite a bit in their favor, making r much less than 1. Then even a relatively small bank of B dollars suffices to assure them of winning. Larger houses, with B quite sizable, can afford to let you play nearly fair games.

EXERCISES

1. Suppose you are playing the "shell game" with a gambler. (In this game, the gambler hides a pea under one of three cups and shuffles them around, and you guess which cup it is.) If you guess correctly, you win \$1; if you guess wrong, you lose \$1. Suppose you each start out with \$2 and play until someone is ruined. What is your probability of losing all your money?

2. Verify that the proof of the text is still correct when $p > \frac{1}{2}$. Interpret formula (5) for this case.
3. Show that if $p > \frac{1}{2}$ and both parties have a substantial amount of money, your probability of ruin is approximately $1/r^A$.
4. Suppose that both the gambler and the player in Exercise 1 start with \$10. Find the approximate probability of the *gambler* being ruined. (Use Exercise 3.)
5. Modify the proof in the text to apply to the case $p = \frac{1}{2}$. What is the probability of your ruin? [*Ans.* B/N.]
6. Suppose the game in Exercise 1 is being played with two cups, and you start with \$10, the gambler with \$15. What is your probability of being ruined?
7. A man leaves a bar in a state of intoxication, and starts to walk randomly. Fifty steps to the left of the bar is the subway entrance, and 50 steps to the right of the bar is the police station. What is the probability that the man makes it to the subway before arriving at the police station if:
   (a) He is equally likely to take a step to the right as to the left?
   (b) He has probability .52 of taking a step to the left?
   (c) He has probability .55 of taking a step to the right?
8. A demon operates a gate between two halves of a box. Initially each side of the box contains 20 molecules. The demon attempts to operate the gate in such a way that all the molecules end up on the left side of the box. Since the inside of the box is quite dark, however, he succeeds only 51 percent of the time; on the other occasions when he opens the gate, a molecule escapes from left to right. (The gate shuts so quickly that only one molecule passes through it each time it is opened.) Suppose the demon operates the gate until all the molecules are on one side or the other. What is the probability that all will be on the left side? [*Ans.* .690.]
9. What is the approximate value of x_A if you are rich and the gambler starts with \$1? Assume the game is weighted so the gambler has the advantage.
10. Suppose you are playing the shell game of Exercise 1 with a poverty-stricken gambler who has only \$1, while you are a millionaire. What, approximately, is the probability that you will be ruined?
11. Consider a simple model for evolution. On a small island there is room for 1000 members of a certain species. One year a favorable mutant appears. We assume that in each subsequent generation either the mutants take one place from the regular members of the species, with probability .6, or the reverse happens. Thus for example, the mutation disappears in the very first generation with probability .4. What is the probability that the mutants eventually take over? [*Hint:* See Exercise 9.] [*Ans.* $\frac{1}{3}$.]
12. After a single crystal of the (fictional) substance ice-eight is added to a container of water, there is probability .9 that one molecule of water

changes to ice-eight every minute, and probability .1 that a crystal of ice-eight changes back to water. What is the approximate probability that the entire contents of the container will eventually consist of ice-eight?

13. You are in the following hopeless situation: You are playing the shell game of Exercise 1, in which you have only $\frac{1}{3}$ chance of winning. You have \$1, and your opponent has \$15. What is the probability of your winning all his money if

- **(a)** You bet \$1 each time?
- **(b)** You bet all your money each time? [*Ans.* $\frac{1}{81}$.]

14. Repeat Exercise 13 for the case of a fair game, where you have probability $\frac{1}{2}$ of winning.

NOTE: Exercises 15–18 deal with the following win problem: A and B are playing a game in which A has probability .6 of winning. They play until A wins three games or B wins two games.

15. Set up the process as a Markov chain in which the states are (a, b), where A has won a games and B has won b games.

16. For each state (a, b), find the probability that A wins.

17. What is the probability that A will achieve his goal first? [*Ans.* $\frac{297}{625}$.]

18. Suppose that payments are made as follows: If A wins three games, he receives \$1; if B wins two games then A pays \$1. What is the expected value of A's winnings, to the nearest penny?

SUGGESTED READING

Dwass, M. *Probability: Theory and Applications.* New York: W. A. Benjamin, 1969.

Feller, W. *An Introduction to Probability Theory and its Applications.* New York: John Wiley, 1950.

Goldberg, S. *Probability: An Introduction.* Englewood Cliffs, N.J.: Prentice-Hall, 1960.

Mosteller, F. *Fifty Challenging Problems in Probability with Solutions.* Reading, Mass.: Addison-Wesley, 1960.

Mosteller, F.; Rourke, Robert E. K.; and **Thomas, G. B.** *Probability with Statistical Applications.* Reading, Mass.: Addison-Wesley, 1970.

Parzen, E. *Modern Probability Theory and Its Applications.* New York: John Wiley, 1960.