Finite Mathematics

TSILB[1]

Version 4.0A0, 5 October 1998

[1]This Space Intentionally Left Blank. Contributors include: John G. Kemeny, J. Laurie Snell, and Gerald L. Thompson. Additional work by: Peter Doyle.

Chapter 2

Sets and subsets

2.1 Introduction

A well-defined collection of objects is known as a *set*. This concept, in its complete generality, is of great importance in mathematics since all of mathematics can be developed by starting from it.

The various pieces of furniture in a given room form a set. So do the books in a given library, or the integers between 1 and 1,000,000, or all the ideas that mankind has had, or the human beings alive between one billion B.C. and ten billion A.D. These examples are all examples of finite sets, that is, sets having a finite number of elements. All the sets discussed in this book will be finite sets.

There are two essentially different ways of specifying a set. One can give a rule by which it can be determined whether or not a given object is a member of the set, or one can give a complete list of the elements in the set. We shall say that the former is a *description* of the set and the latter is a *listing* of the set. For example, we can define a set of four people as (a) the members of the string quartet which played in town last night, or (b) four particular persons whose names are Jones, Smith, Brown, and Green. It is customary to use braces to, surround the listing of a set; thus the set above should be listed $\{\text{Jones}, \text{Smith}, \text{Brown}, \text{Green}\}$.

We shall frequently be interested in sets of logical possibilities, since the analysis of such sets is very often a major task in the solving of a problem. Suppose, for example, that we were interested in the successes of three candidates who enter the presidential primaries (we assume there are no other entries). Suppose that the key primaries will be held in New Hampshire, Minnesota, Wisconsin, and California. Assume that candidate A enters all the primaries, that B does not contest in New Hampshire's primary, and C does not contest in Wisconsin's. A list of the logical possibilities is given in Figure 2.1. Since the New Hampshire and Wisconsin primaries can each end in two ways, and the Minnesota and California primaries can each end in three ways, there are in all $2 \cdot 2 \cdot 3 \cdot 3 = 36$ different logical possibilities as listed in Figure 2.1.

A set that consists of some members of another set is called a *subset* of that set. For example, the set of those logical possibilities in Figure 2.1 for which the statement "Candidate A wins at least three primaries" is true, is a subset of the set of all logical possibilities. This subset can also be defined by listing its members: $\{\mathrm{P1}, \mathrm{P2}, \mathrm{P3}, \mathrm{P4}, \mathrm{P7}, \mathrm{P13}, \mathrm{P19}\}$.

In order to discuss all the subsets of a given set, let us introduce the following terminology. We shall call the original set the *universal set*, one-element subsets will be called *unit* sets, and the set which contains no members the *empty set*. We do not introduce special names for other kinds of subsets of the universal set. As an example, let the universal set $\mathcal{U}$ consist of the three elements $\{a, b, c\}$. The proper subsets of $\mathcal{U}$ are those sets containing some but not all of the elements of $\mathcal{U}$. The proper subsets consist of three two-element sets namely, $\{a, b\}$, $\{a, c\}$, and $\{b, c\}$ and three unit sets, namely, $\{a\}$, $\{b\}$, and $\{c\}$. To complete the picture, we also consider the universal set a subset (but not a proper subset) of itself, and we consider the empty set $\emptyset$, that contains no elements of $\mathcal{U}$, as a subset of $\mathcal{U}$. At first it may seem strange that we should include the sets $\mathcal{U}$ and $\emptyset$ as subsets of $\mathcal{U}$, but the reasons for their inclusion will become clear later.

We saw that the three-element set above had $8 = 2^3$ subsets. In general, a set with n elements has 2^n subsets, as can be seen in the following manner. We form subsets P of $\mathcal{U}$ by considering each of the elements of $\mathcal{U}$ in turn and deciding whether or not to include it in the subset P. If we decide to put every element of $\mathcal{U}$ into P, we get the universal set, and if we decide to put no element of $\mathcal{U}$ into P, we get the empty set. In most cases we will put some but not all the elements into P and thus obtain a proper subset of $\mathcal{U}$. We have to make n decisions, one for each element of the set, and for each decision we have to choose between two alternatives. We can make these decisions in $2 \cdot 2 \cdot \ldots \cdot 2 = 2^n$ ways, and hence this is the number of different subsets of $\mathcal{U}$ that can be formed. Observe that our formula would not have been so simple if we had not included the universal set and the empty set as subsets of $\mathcal{U}$.

Possibility Number	Winner in New Hampshire	Winner in Minnesota	Winner in Wisconsin	Winner in California
P1	A	A	A	A
P2	A	A	A	B
P3	A	A	A	C
P4	A	A	B	A
P5	A	A	B	B
P6	A	A	B	C
P7	A	B	A	A
P8	A	B	A	B
P9	A	B	A	C
P10	A	B	B	A
P11	A	B	B	B
P12	A	B	B	C
P13	A	C	A	A
P14	A	C	A	B
P15	A	C	A	C
P16	A	C	B	A
P17	A	C	B	B
P18	A	C	B	C
P19	C	A	A	A
P20	C	A	A	B
P21	C	A	A	C
P22	C	A	B	A
P23	C	A	B	B
P24	C	A	B	C
P25	C	B	A	A
P26	C	B	A	B
P27	C	B	A	C
P28	C	B	B	A
P29	C	B	B	B
P30	C	B	B	C
P31	C	C	A	A
P32	C	C	A	B
P33	C	C	A	C
P34	C	C	B	A
P35	C	C	B	B
P36	C	C	B	C

Figure 2.1: ♢

In the example of the voting primaries above there are 2^{36} or about 70 billion subsets. Of course, we cannot deal with this many subsets in a practical problem, but fortunately we are usually interested in only a few of the subsets. The most interesting subsets are those which can be defined by means of a simple rule such as "the set of all logical possibilities in which C loses at least two primaries". It would be difficult to give a simple description for the subset containing the elements $\{\mathrm{P1}, \mathrm{P4}, \mathrm{P14}, \mathrm{P30}, \mathrm{P34}\}$. On the other hand, we shall see in the next section how to define new subsets in terms of subsets already defined.

Example 2.1 We illustrate the two different ways of specifying sets in terms of the primary voting example. Let the universal set $\mathcal{U}$ be the logical possibilities given in Figure 2.1.

1. What is the subset of $\mathcal{U}$ in which candidate B wins more primaries than either of the other candidates?

   [Ans. $\{\mathrm{P11}, \mathrm{P12}, \mathrm{P17}, \mathrm{P23}, \mathrm{P26}, \mathrm{P28}, \mathrm{P29}\}$.]

2. What is the subset in which the primaries are split two and two?

   [Ans. $\{\mathrm{P5}, \mathrm{P8}, \mathrm{P10}, \mathrm{P15}, \mathrm{P21}, \mathrm{P30}, \mathrm{P31}, \mathrm{P35}\}$.]

3. Describe the set $\{\mathrm{P1}, \mathrm{P4}, \mathrm{P19}, \mathrm{P22}\}$.

   [Ans. The set of possibilities for which A wins in Minnesota and California.]

4. How can we describe the set $\{\mathrm{P18}, \mathrm{P24}, \mathrm{P27}\}$

   [Ans. The set of possibilities for which C wins in California, and the other primaries are split three ways.]

$\diamondsuit$

Exercises

1. In the primary example, give a listing for each of the following sets.

   (a) The set in which C wins at least two primaries.

(b) The set in which the first three primaries are won by the same candidate.

(c) The set in which B wins all four primaries.

2. The primaries are considered decisive if a candidate can win three primaries, or if he or she wins two primaries including California. List the set in which the primaries are decisive.

3. Give simple descriptions for the following sets (referring to the primary example).

   (a) $\{P33, P36\}$.

   (b) $\{P10, P11, P12, P28, P29, P30\}$.

   (c) $\{P6, P20, P22\}$.

4. Joe, Jim, Pete, Mary, and Peg are to be photographed. They want to line up so that boys and girls alternate. List the set of all possibilities.

5. In Exercise 4, list the following subsets.

   (a) The set in which Pete and Mary are next to each other.

   (b) The set in which Peg is between Joe and Jim.

   (c) The set in which Jim is in the middle.

   (d) The set in which Mary is in the middle.

   (e) The set in which a boy is at each end.

6. Pick out all pairs in Exercise 5 in which one set is a subset of the other.

7. A TV producer is planning a half-hour show. He or she wants to have a combination of comedy, music, and commercials. If each is allotted a multiple of five minutes, construct the set of possible distributions of time. (Consider only the total time allotted to each.)

8. In Exercise 7, list the following subsets.

   (a) The set in which more time is devoted to comedy than to music.

(b) The set in which no more time is devoted to commercials than to either music or comedy.

(c) The set in which exactly five minutes is devoted to music.

(d) The set in which all three of the above conditions are satisfied.

9. In Exercise 8, find two sets, each of which is a proper subset of the set in 8a and also of the set in 8c.

10. Let $\mathcal{U}$ be the set of paths in Figure **??**. Find the subset in which

    (a) Two balls of the same color are drawn.

    (b) Two different color balls are drawn.

11. A set has 101 elements. How many subsets does it have? How many of the subsets have an odd number of elements?

[Ans. $2^{101}; 2^{100}$.]

12. Do Exercise 11 for the case of a set with 102 elements.

2.2 Operations on subsets

In Chapter **??** we considered the ways in which one could form new statements from given statements. Now we shall consider an analogous procedure, the formation of new sets from given sets. We shall assume that each of the sets that we use in the combination is a subset of some universal set, and we shall also want the newly formed set to be a subset of the same universal set. As usual, we can specify a newly formed set either by a description or by a listing.

If P and Q are two sets, we shall define a new set $P \cap Q$, called the *intersection* of P and Q, as follows: $P \cap Q$ is the set which contains those and only those elements which belong to both P and Q. As an example, consider the logical possibilities listed in Figure 2.1. Let P be the subset in which candidate A wins at least three primaries, i.e., the set $\{\mathrm{P1}, \mathrm{P2}, \mathrm{P3}, \mathrm{P4}, \mathrm{P7}, \mathrm{P13}, \mathrm{P19}\}$; let Q be the subset in which A wins the first two primaries, i.e., the set $\{\mathrm{P1}, \mathrm{P2}, \mathrm{P3}, \mathrm{P4}, \mathrm{P5}, \mathrm{P6}\}$. Then the intersection $P \cap Q$ is the set in which both events take place, i.e., where A wins the first two primaries and wins at least three primaries. Thus $P \cap Q$ is the set $\{\mathrm{P1}, \mathrm{P2}, \mathrm{P3}, \mathrm{P4}\}$.

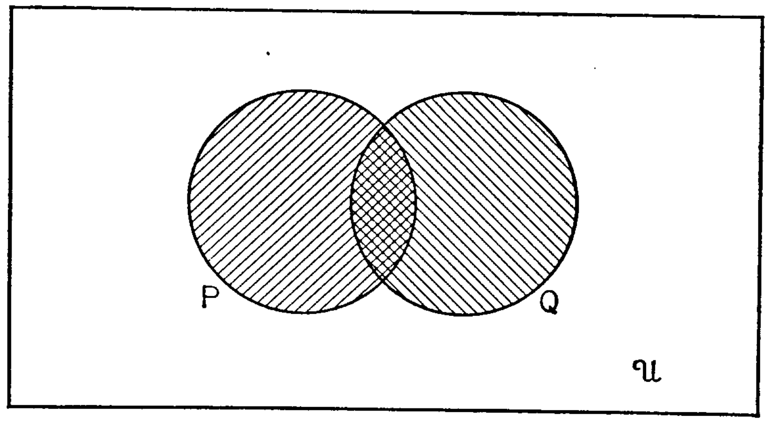


Figure 2.2: $\diamondsuit$

If P and Q are two sets, we shall define a new set $P \cup Q$ called the *union* of P and Q as follows: $P \cup Q$ is the set that contains those and only those elements that belong either to P or to Q (or to both). In the example in the paragraph above, the union $P \cup Q$ is the set of possibilities for which either A wins the first two primaries or wins at least three primaries, i.e., the set $\{\mathrm{P1, P2, P3, P4, P5, P6, P7, P13, P19}\}$.

To help in visualizing these operations we shall draw diagrams, called *Venn diagrams*, which illustrate them. We let the universal set be a rectangle and let subsets be circles drawn inside the rectangle. In Figure 2.2 we show two sets P and Q as shaded circles. Then the doubly crosshatched area is the intersection $P \cap Q$ and the total shaded area is the union $P \cup Q$.

If P is a given subset of the universal set $\mathcal{U}$, we can define a new set $\tilde{P}$ called the *complement* of P as follows: P is the set of all elements of $\mathcal{U}$ that are not contained in P. For example, if, as above, Q is the set in which candidate A wins the first two primaries, then $\tilde{Q}$ is the set $\{\mathrm{P7, P8}, \ldots, \mathrm{P36}\}$. The shaded area in Figure 2.3 is the complement of the set P. Observe that the complement of the empty set $\emptyset$ is the universal set $\mathcal{U}$, and also that the complement of the universal set is the empty set.

Sometimes we shall be interested in only part of the complement of a set. For example, we might wish to consider the part of the complement of the set Q that is contained in P, i.e., the set $P \cap \tilde{Q}$. The shaded area in Figure 2.4 is $P \cap \tilde{Q}$.

A somewhat more suggestive definition of this set can be given as

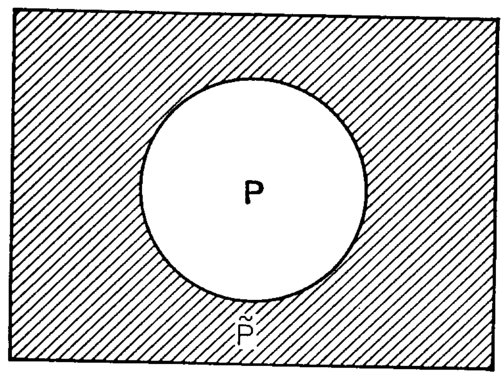


Figure 2.3: $\diamondsuit$

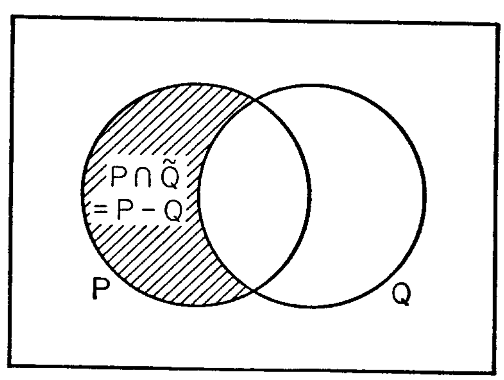


Figure 2.4: $\diamondsuit$

follows: Let $P - Q$ be the *difference* of P and Q, that is, the set that contains those elements of P that do not belong to Q. Figure 2.4 shows that $P \cap \tilde{Q}$ and $P - Q$ are the same set. In the primary voting example above, the set $P - Q$ can be listed as $\{\mathrm{P7}, \mathrm{P13}, \mathrm{P19}\}$.

The complement of a subset is a special case of a difference set, since we can write $\tilde{Q} = \mathcal{U} - Q$. If P and Q are nonempty subsets whose intersection is the empty set, i.e., $P \cap Q = \emptyset$, then we say that they are *disjoint* subsets.

Example 2.2 In the primary voting example let R be the set in which A wins the first three primaries, i e., the set $\{\mathrm{P1}, \mathrm{P2}, \mathrm{P3}\}$; let S be the set in which A wins the last two primaries, i.e., the set $\{\mathrm{P1}, \mathrm{P7}, \mathrm{P13}, \mathrm{P19}, \mathrm{P25}, \mathrm{P31}\}$. Then $R \cap S = \{\mathrm{P1}\}$ is the set in which A wins the first three primaries and also the last two, that is, he or she wins all the primaries. We also have

$$R \cup S = \{\mathrm{P1}, \mathrm{P2}, \mathrm{P3}, \mathrm{P7}, \mathrm{P13}, \mathrm{P19}, \mathrm{P25}, \mathrm{P31}\},$$

which can be described as the set in which A wins the first three primaries or the last two. The set in which A does not win the first three primaries is $\tilde{R} = \{\mathrm{P4}, \mathrm{P5}, \ldots, \mathrm{P36}\}$. Finally, we see that the difference set $R - S$ is the set in which A wins the first three primaries but not both of the last two. This set can be found by taking from R the element P1 which it has in common with S, so that $R - S = \{\mathrm{P2}, \mathrm{P3}\}$. $\diamondsuit$

Exercises

1. Draw Venn diagrams for $P \cap Q$, $P \cap \tilde{Q}$, $\tilde{P} \cap Q$, $\tilde{P} \cap \tilde{Q}$.

2. Give a step-by-step construction of the diagram for $(\tilde{P} - Q) \cup (P \cap \tilde{Q})$.

3. Venn diagrams are also useful when three subsets are given. Construct such a diagram, given the subsets P. Q. and R. Identify each of the eight resulting areas in terms of P, Q, and R.

4. In testing blood, three types of antigens are looked for: A, B, and Rh. Every person is classified doubly. He or she is Rh positive if he or she has the Rh antigen, and Rh negative otherwise. He or she is type AB, A, or B depending on which of the other antigens he or she has, with type O having neither A nor B. Draw a Venn diagram, and identify each of the eight areas.

	Liked Very Much	Liked Slightly	Disliked Slightly	Disliked Very Much
Men	1	3	5	10
Women	6	8	3	1
Boys	5	5	3	2
Girls	8	5	1	1

Figure 2.5: ◊

5. Considering only two subsets, the set X of people having antigen A, and the set Y of people having antigen B. define (symbolically) the types AB, A, B and O.

6. A person can receive blood from another person if he or she has all the antigens of the donor. Describe in terms of X and Y the sets of people who can give to each of the four types. Identify these sets in terms of blood types.

7. The tabulation in Figure 2.5 records the reaction of a number of spectators to a television show. A11 the categories can be defined in terms of the following four: M (male), G (grown-up), L (liked), V (very much). How many people fall into each of the following categories?

   (a) M.

   [Ans. 34.]

   (b) L.

   (c) V.

   (d) $M \cap \tilde{G} \cap \tilde{L} \cap V$.

   [Ans. 2.]

   (e) $\tilde{M} \cap G \cap L$.

   (f) $(M \cap G) \cup (L \cap V)$.

(g) $\widetilde{M \cap G}$.

[Ans. 48.]

(h) $\tilde{M} \cup \tilde{G}$.

(i) $M - G$.

(j) $[\tilde{M} - (G \cap L \cap \tilde{V})]$.

8. In a survey of 100 students, the numbers studying various languages were found to be: Spanish, 28; German, 30; French, 42; Spanish and German, 8; Spanish and French, 10; German and French, 5; all three languages, 3.

   (a) How many students were studying no language?

   [Ans. 20.]

   (b) How many students had French as their only language?

   [Ans. 30.]

   (c) How many students studied German if and only if they studied French?

   [Ans. 38.]

   [Hint: Draw a Venn diagram with three circles, for French, German, and Spanish students. Fill in the numbers in each of the eight areas, using the data given above. Start from the end of the list and work back.]

9. In a later survey of the 100 students (see Exercise 8) the numbers studying the various languages were found to be: German only, 18; German but not Spanish, 23; German and French, 8; German, 26; French, 48; French and Spanish, 8; no language, 24.

   (a) How many students took Spanish?

   [Ans. 18.]

   (b) How many took German and Spanish but not French?

   [Ans. None.]

   (c) How many took French if and only if they did not take Spanish?

[Ans. 50.]

10. The report of one survey of the 100 students (see Exercise 8) stated that the numbers studying the various languages were: all three languages, 5; German and Spanish, 10; French and Spanish, 8; German and French, 20; Spanish, 30; German, 23; French, 50. The surveyor who turned in this report was fired. Why?

2.3 The relationship between sets and compound statements

The reader may have observed several times in the preceding sections that there was a close connection between sets and statements, and between set operations and compounding operations. In this section we shall formalize these relationships.

If we have a number of statements relative to a set of logical possibilities, there is a natural way of assigning a set to each statement. First of all, we take the set of logical possibilities as our universal set. Then to each statement we assign the subset of logical possibilities of the universal set for which that statement is true. This idea is so important that we embody it in a formal definition.

Definition. Let $\mathcal{U}$ be a set of logical possibilities, let p be a statement relative to it, and let P be that subset of the possibilities for which p is true; then we call P the *truth set* of p.

If p and q are statements, then $p \vee q$ and $p \wedge q$ are also statements and hence must have truth sets. To find the truth set of $p \vee q$, we observe that it is true whenever p is true or q is true (or both). Therefore we must assign to $p \vee q$ the logical possibilities which are in P or in Q (or both); that is, we must assign to $p \vee q$ the set $P \cup Q$. On the other hand, the statement $p \wedge q$ is true only when both p and q are true, so that we must assign to $p \wedge q$ the set $P \cap Q$.

Thus we see that there is a close connection between the logical operation of disjunction and the set operation of union, and also between conjunction and intersection. A careful examination of the definitions of union and intersection shows that the word "or" occurs in the definition of union and the word "and" occurs in the definition of intersection. Thus the connection between the two theories is not surprising.

Since the connective "not" occurs in the definition of the complement of a set, it is not surprising that the truth set of $\neg p$ is $\tilde{P}$. This

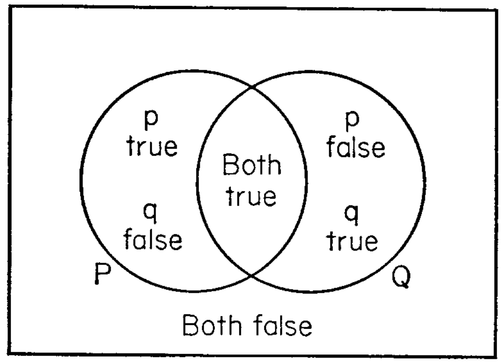


Figure 2.6: $\diamondsuit$

follows since $\neg p$ is true when p is false, so that the truth set of $\neg p$ contains all logical possibilities for which p is false, that is, the truth set of $\neg p$ is $\tilde{P}$.

The truth sets of two propositions p and q are shown in Figure 2.6. Also marked on the diagram are the various logical possibilities for these two statements. The reader should pick out in this diagram the truth sets of the statements $p \vee q$, $p \wedge q$, $\neg p$, and $\neg q$.

The connection between a statement and its truth set makes it possible to "translate" a problem about compound statements into a problem about sets. It is also possible to go in the reverse direction. Given a problem about sets, think of the universal set as being a set of logical possibilities and think of a subset as being the truth set of a statement. Hence we can "translate" a problem about sets into a problem about compound statements.

So far we have discussed only the truth sets assigned to compound statements involving $\vee$, $\wedge$, and $\neg$. All the other connectives can be defined in terms of these three basic ones, so that we can deduce what truth sets should be assigned to them. For example, we know that $p \to q$ is equivalent to $\neg p \vee q$ (see Figure **??**). Hence the truth set of $p \to q$ is the same as the truth set of $\neg p \vee q$, that is, it is $\tilde{P} \cup Q$. The Venn diagram for $p \to q$ is shown in Figure 2.7, where the shaded area is the truth set for the statement. Observe that the unshaded area in Figure 2.7 is the set $P - Q = P \cap \tilde{Q}$, which is the truth set of the statement $p \wedge \neg q$. Thus the shaded area is the set $\widetilde{P - Q} = \widetilde{P \cap \tilde{Q}}$, which is the truth set of the statement $\neg(p \wedge \neg q)$. We have thus discovered the fact that $p \to q$, $\neg p \vee q$, and $\neg(p \wedge \neg q)$ are equivalent. It is always the case that two compound statements are equivalent if and only if they

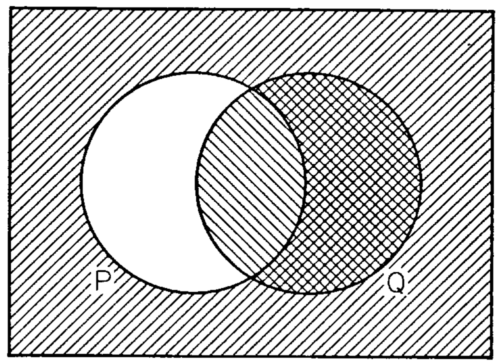


Figure 2.7: $\diamondsuit$

have the same truth sets. Thus we can test for equivalence by checking whether they have the same Venn diagram.

Suppose that p is a statement that is logically true. What is its truth set? Now p is logically true if and only if it is true in every logically possible case, so that the truth set of p must be $\mathcal{U}$. Similarly, if p is logically false, then it is false for every logically possible case, so that its truth set is the empty set $\emptyset$.

Finally, let us consider the implication relation. Recall that p implies p if and only if the conditional $p \to q$ is logically true. But $p \to q$ is logically true if and only if its truth set is $\mathcal{U}$, that is, $\widetilde{(P - Q)} = \mathcal{U}$, or $(P - Q) = \emptyset$. From Figure 2.4 we see that if $P - Q$ is empty, then P is contained in Q. We shall symbolize the containing relation as follows: $P \subset Q$ means "P is a subset of Q". We conclude that $p \to q$ is logically true if and only if $P \subset Q$.

Figure 2.8 supplies a "dictionary" for translating from statement language to set language, and back. To each statement relative to a set of possibilities $\mathcal{U}$ there corresponds a subset of $\mathcal{U}$, namely the truth set of the statement. This is shown in lines 1 and 2 of the figure. To each connective there corresponds an operation on sets, as illustrated in the next four lines. And to each relation between statements there corresponds a relation between sets, examples of which are shown in the last two lines of the figure.

Example 2.3 Prove by means of a Venn diagram that the statement $[p \vee (\neg p \vee q)]$ is logically true. The assigned set of this statement is $[P \cup (\tilde{P} \cup Q)]$, and its Venn diagram is shown in Figure 2.9. In that figure the set P is shaded vertically, and the set $\tilde{P} \cup Q$ is shaded

Statement Language	Set Language
r	R
s	S
$\sim r$	$\tilde{R}$
$r \vee s$	$R \cup S$
$r \wedge s$	$R \cap S$
$r \rightarrow s$	$\widetilde{(R - S)}$
r implies s	$R \subset S$
r is equivalent to s	$R = S$

Figure 2.8: ♢

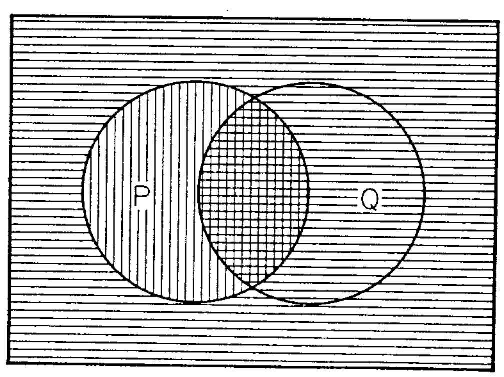
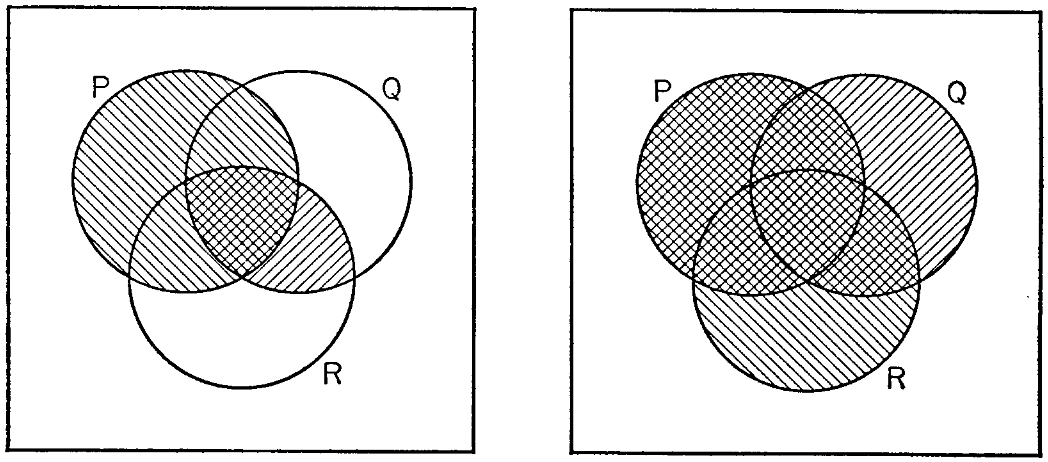


Figure 2.9: ♢



(a) (b)

Figure 2.10: $\diamondsuit$

horizontally. Their union is the entire shaded area, which is $\mathcal{U}$, so that the compound statement is logically true. $\diamondsuit$

Example 2.4 Prove by means of Venn diagrams that $p \vee (q \wedge r)$ is equivalent to $(p \vee q) \wedge (p \vee r)$. The truth set of $p \vee (q \wedge r)$ is the entire shaded area in diagram (a) of Figure 2.10, and the truth set of $(p \vee q) \wedge (p \vee r)$ is the doubly shaded area in diagram (b). Since these two sets are equal, we see that the two statements are equivalent. $\diamondsuit$

Example 2.5 Show by means of a Venn diagram that q implies $p \to q$. The truth set of $p \to q$ is the shaded area in Figure 2.7. Since this shaded area includes the set Q. we see that q implies $p \to q$. $\diamondsuit$

Exercises

Note. In Exercises 1, 2, and 3, find first the truth set of each statement.

1. Use Venn diagrams to test which of the following statements are logically true or logically false.

   (a) $p \vee \neg p$.

   [Ans. logically true.]

(b) $p \wedge \neg p$.

[Ans. logically false.]

(c) $p \vee (\neg p \wedge q)$.

(d) $p \to (q \to p)$.

[Ans. logically true.]

(e) $p \wedge \neg(q \to p)$.

[Ans. logically false.]

2. Use Venn diagrams to test the following statements for equivalences.

   (a) $p \vee \neg q$.

   (b) $\neg(p \wedge q)$.

   (c) $\neg(q \wedge \neg q)$.

   (d) $p \to \neg q$.

   (e) $\neg p \vee \neg q$.

   [Ans. 2a and 2c equivalent; 2b and 2d and 2e equivalent.]

3. Use Venn diagrams for the following pairs of statements to test whether one implies the other.

   (a) $p; p \wedge q$.

   (b) $p \wedge \neg q; \neg p \to \neg q$.

   (c) $p \to q; q \to p$.

   (d) $p \wedge q; p \wedge \neg q$.

4. Devise a test for inconsistency of p and q, using Venn diagrams.

5. Three or more statements are said to be inconsistent if they cannot all be true. What does this state about their truth sets?

6. Consider these three statements.

   If this is a good course, then I will work hard in it.

   If this is not a good course, then I shall get a bad grade in it.

I will not work hard, but I will get a good grade in this course.

(a) Assign variables to the components of each of these statements.

(b) Bring the statements into symbolic form.

(c) Find the truth sets of the statements.

(d) Rest for consistency.

[Ans. Inconsistent.]

Note. In Exercises 7, 8, and 9, assign to each set a statement having it as a truth set.

7. Use truth tables to find which of the following sets are empty.

(a) $(P \cup Q) \cap (\tilde{P} \cup \tilde{Q})$.

(b) $(P \cap Q) \cap (\tilde{Q} \cap R)$.

(c) $(P \cap Q) - P$.

(d) $(P \cup R) \cap (\tilde{P} \cup \tilde{Q})$

[Ans. 7b and 7c.]

8. Use truth tables to find out whether the following sets are all different.

(a) $P \cap (Q \cup R)$.

(b) $(R - Q) \cup (Q - R)$.

(c) $(R \cup Q) \cap (\widetilde{R \cap Q})$.

(d) $(P \cap Q) \cup (P \cap R)$.

(e) $(P \cap Q \cap \tilde{R}) \cup (P \cap \tilde{Q} \cap R) \cup (\tilde{P} \cap Q \cap \tilde{R}) \cup (\tilde{P} \cap \tilde{Q} \cap R)$.

9. Use truth tables for the following pairs of sets to test whether one is a subset of the other.

(a) $P; P \cap Q$.

(b) $P \cap \tilde{Q}; Q \cap \tilde{P}$.

(c) $P - Q; Q - P$.

(d) $P \cap \tilde{Q}; P \cup Q$.

10. Show, both by the use of truth tables and by the use of Venn diagrams, that $p \wedge (q \vee r)$ is equivalent to $(p \wedge q) \vee (p \wedge r)$.

11. The *symmetric difference* of P and Q is defined to be $(P - Q) \cup (Q - P)$. What connective corresponds to this set operation?

12. Let p, q, r be a complete set of alternatives (see Section **??**). What can we say about the truth sets P, Q, R?

2.4 The abstract laws of set operations

The set operations which we have introduced obey some very simple abstract laws, which we shall list in this section. These laws can be proved by means of Venn diagrams or they can be translated into statements and checked by means of truth tables.

The abstract laws given below bear a close resemblance to the elementary algebraic laws with which the student is already familiar. The resemblance can be made even more striking by replacing $\cup$ by $+$ and $\cap$ by $\times$. For this reason, a set, its subsets, and the laws of combination of subsets are considered an algebraic system, called a *Boolean algebra*—after the British mathematician George Boole who was the first person to study them from the algebraic point of view. Any other system obeying these laws, for example, the system of compound statements studied in Chapter **??**, is also known as a Boolean algebra. We can study any of these systems from either the algebraic or the logical point of view.

Below are the basic laws of Boolean algebras. The proofs of these laws will be left as exercises.

The laws governing union and intersection:

A1. $A \cup A = A$.
A2. $A \cap A = A$.
A3. $A \cup B = B \cup A$.
A4. $A \cap B = B \cap A$.
A5. $A \cup (B \cup C) = (A \cup B) \cup C$.
A6. $A \cap (B \cap C) = (A \cap B) \cap C$.
A7. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
A8. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.
A9. $A \cup \mathcal{U} = \mathcal{U}$.
A10. $A \cap \emptyset = \emptyset$.
A11. $A \cup \emptyset = A$.
A12. $A \cap \mathcal{U} = A$.

The laws governing complements:

B1. $\tilde{\tilde{A}} = A$.
B2. $A \cup \tilde{A} = \mathcal{U}$.
B3. $A \cap \tilde{A} = \emptyset$.
B4. $\widetilde{A \cup B} = \tilde{A} \cap \tilde{B}$.
B5. $\widetilde{A \cap B} = \tilde{A} \cup \tilde{B}$.
B6. $\tilde{\mathcal{U}} = \emptyset$.

The laws governing set-differences:

C1. $A - B = A \cap \tilde{B}$.
C2. $\mathcal{U} - A = \tilde{A}$.
C3. $A - \mathcal{U} = \emptyset$.
C4. $A - \emptyset = A$.
C5. $\emptyset - A = \emptyset$.
C6. $A - A = \emptyset$.
C7. $(A - B) - C = A - (B \cup C)$.
C8. $A - (B - C) = (A - B) \cup (A \cap C)$.
C9. $A \cup (B - C) = (A \cup B) - (C - A)$.
C10. $A \cap (B - C) = (A \cap B) - (A \cap C)$.

Exercises

1. Test laws in the group A1–A12 by means of Venn diagrams.

2. "Translate" the A-laws into laws about compound statements. Test these by truth tables.

3. Test the laws in groups B and C by Venn diagrams.

4. "Translate" the B- and C-laws into laws about compound statements. Test these by means of truth tables.

5. Derive the following results from the 28 basic laws.

   (a) $A = (A \cap B) \cup (A \cap \tilde{B})$.

   (b) $A \cup B = (A \cap B) \cup (A \cap \tilde{B}) \cup (\tilde{A} \cap B)$.

   (c) $A \cap (A \cup B) = A$.

   (d) $A \cup (\tilde{A} \cap B) = A \cup B$.

6. From the A- and B-laws and from C1, derive C2–C6.

7. Use A1–A12 and C2–C10 to derive B1, B2, B3, and B6.

Supplementary exercises.

Note. Use the following definitions in these exercises: Let $+$ be symmetric difference (see Exercise 11), $\times$ be intersection, let 0 be $\emptyset$ and 1 be $\mathcal{U}$.

8. From A2, A4, and A6 derive the properties of multiplication.

9. Find corresponding properties for addition.

10. Set up addition and multiplication tables for 0 and 1.

11. What do $A \times 0$, $A \times 1$, $A + 0$, and $A + 1$ equal?

    [Ans. 0; A; A; $\tilde{A}$.]

12. Show that

$$A \times (B + C) = (A \times B) + (A \times C).$$

13. Show that the following equation is not always true.

$$A + (B \times C) = (A + B) \times (A + C).$$

+	0	1
0	0	1
1	1	?

•	0	1
0	0	0
1	0	1

Figure 2.11: $\diamondsuit$

2.5 Two-digit number systems

In the decimal number system one can write any number by using only the ten digits, $0, 1, 2, \ldots, 9$. Other number systems can be constructed which use either fewer or more digits. Probably the simplest number system is the binary number system which uses only the digits 0 and 1. We shall consider all the possible ways of forming number systems using only these two digits.

The two basic arithmetical operations are addition and multiplication. To understand any arithmetic system, it is necessary to know how to add or multiply any two digits together. Thus to understand the decimal system, we had to learn a multiplication table and an addition table, each of which had 100 entries. To understand the binary system, we have to learn a multiplication and an addition table, each of which has only four entries. These are shown in Figure 2.11. The multiplication table given there is completely determined by the two familiar rules that multiplying a number by zero gives zero, and multiplying a number by one leaves it unchanged. For addition, we have only the rule that the addition of zero to a number does not change that number. The latter rule is sufficient to determine all but one of the entries in the addition table in Figure 2.11. We must still decide what shall be the sum $1 + 1$.

What are the possible ways in which we can complete the addition table? The only one-digit numbers that we can use are 0 and l, and these lead to interesting systems. Of the possible two-digit numbers, we see that 00 and 01 are the same as 0 and l and so do not give anything new. The number 11 or any greater number would introduce a "jump" in the table, hence the only other possibility is 10. The addition tables of these three different number systems are shown in Figure 2.12, and they all have the multiplication table shown in Figure 2.11. Each of these systems is interesting in itself as the interpretations below show.

Let us say that the *parity* of a positive integer is the fact of its being

+	0	1
0	0	1
1	1	0

(a)

+	0	1
0	0	1
1	1	1

(b)

+	0	1
0	0	1
1	1	10

(c)

Figure 2.12: $\diamondsuit$

odd or even. Consider now the number system having the addition table (a) in Figure 2.12 and let 0 represent "even" and 1 represent "odd". The tables above now tell how the parity of a combination of two positive integers is related to the parity of each. Thus $0 \cdot 1 = 0$ tells us that the product of an even number and an odd number is even, while $1 + 1 = 0$ tells us that the sum of two odd numbers is even, etc. Thus the first number system is that which we get from the arithmetic of the positive integers if we consider only the parity of numbers.

The second number system, which has the addition table (b) in Figure 2.12, has an interpretation in terms of sets. Let 0 correspond to the empty set $\emptyset$ and 1 correspond to the universal set $\mathcal{U}$. Let the addition of numbers correspond to the union of sets and let the multiplication of sets correspond to the intersection of sets. Then $0 \cdot 1 = 0$ tells us that $\emptyset \cap \mathcal{U} = \emptyset$ and $1 + 1 = 1$ tells us that $\mathcal{U} \cup \mathcal{U} = \mathcal{U}$. The student should give the interpretations for the other arithmetic computations possible for this number system.

Finally, the third number system, which has the addition table in (c) of Figure 2.12, is the so-called *binary number system.* Every ordinary integer can be written as a binary integer. Thus the binary 0 corresponds to the ordinary 0, and the binary unit 1 to the ordinary single unit. The binary number 10 means a "unit of higher order" and corresponds to the ordinary number two (not to ten). The binary number 100 then means two times two or four. In general, if $b_n b_{n-1} \ldots b_2 b_1 b_0$ is a binary number, where each digit is either 0 or 1, then the corresponding ordinary integer I is given by the formula

$$I = b_n \cdot 2^n + b_{n-1} \cdot 2^{n-1} + \ldots + b_2 \cdot 2^2 + b_1 \cdot 2 + b_0.$$

Thus the binary number 11001 corresponds to $2^4 + 2^3 + 1 = 16 + 8 + 1 = 25$. The table in Figure 2.13 shows some binary numbers and their decimal equivalents.

Binary number	1	10	11	100	101	110	111	1000	10000	100000
Decimal equiv.	1	2	3	4	5	6	7	8	16	32

Figure 2.13: $\diamondsuit$

Because electronic circuits are particularly well adapted to performing computations in the binary system, modern high-speed electronic computers are frequently constructed to work in the binary system.

Example 2.6 As an example of a computation, let us multiply 5 by 5 in the binary system. Since the binary equivalent of 5 is the number 101, the multiplication is done as follows.

$$\begin{array}{ccccc}
 & & 1 & 0 & 1 \\
 & & 1 & 0 & 1 \\
\hline
 & & 1 & 0 & 1 \\
 & 0 & 0 & 0 & \\
1 & 0 & 1 & & \\
\hline
1 & 1 & 0 & 0 & 1
\end{array}$$

The answer is the binary number 11001, which we saw above was equivalent to the decimal integer 25, the answer we expected to get. $\diamondsuit$

Exercises

1. Complete the interpretations of the addition and multiplication tables for the number systems representing
   (a) parity,
   (b) the sets $\mathcal{U}$ and $\emptyset$.
2. (a) What are the binary numbers corresponding to the integers 11, 52, 64, 98, 128, 144?
   [Partial Ans. 1100010 corresponds to 98.]
   (b) What decimal integers correspond to the binary numbers 1111, 1010101, 1000000, 11011011?

[Partial Ans. 1010101 corresponds to 85.]

3. Carry out the following operations in the binary system. Check your answer.

   (a) $29 + 20$.

   (b) $9 \cdot 7$.

4. Of the laws listed below, which apply to each of the three systems?

   (a) $x + y = y + x$.

   (b) $x + x = x$.

   (c) $x + x + x = x$.

5. Interpret $a + b$ to be the larger of the two numbers a and b, and $a \cdot b$ to be the smaller of the two. Write tables of "addition" and "multiplication" for the digits 0 and 1. Compare the result with the three systems given above.

   [Ans. Same as the $\mathcal{U}$, $\emptyset$ system.]

6. What do the laws A1–A10 of Section 2.4 tell us about the second number system established above?

7. The first number system above (about parity) can be interpreted to deal with the remainders of integers when divided by 2. An even number leaves 0, an odd number leaves 1. Construct tables of addition and multiplication for remainders of integers when divided by 3. [Hint: These will be 3 by 3 tables.]

8. Given a set of four elements, suppose that we want to number its subsets. For a given subset, write down a binary number as follows: The first digit is 1 if and only if the first element is in the subset, the second digit is 1 if and only if the second element is in the subset, etc. Prove that this assigns a unique number, from 0 to 15, to each subset.

9. In a multiple choice test the answers were numbered 1, 2, 4, and 8. The students were told that there might be no correct answer, or that one or more answers might be correct. They were told to add together the numbers of the correct answers (or to write 0 if no answer was correct).

(a) By using the result of Exercise 8, show that the resulting number gives the instructor all the information he or she wants.

(b) On a given question the correct sum was 7. Three students put down 4, 8, and 15, respectively. Which answer was most nearly correct? Which answer was worst?

[Ans. 15 best, 8 worst.]

10. In the ternary number system, numbers are expressed to the base 3, so that 201 in this system stands for $2 \cdot 3^2 + 0 \cdot 3 + 1 \cdot 1 = 19$.

    (a) Write the numbers from 1 through 30 in this notation.

    (b) Construct a table of addition and multiplication for the digits 0, 1, 2.

    (c) Carry out the multiplication of $5 \cdot 5$ in this system. Check your answer.

11. Explain the meaning of the numeral "2907" in our ordinary (base 10) notation, in analogy to the formula given for the binary system.

12. Show that the addition and multiplication tables set up in Exercise 10 correspond to one of our three systems.

2.6 Voting coalitions

As an application of our set concepts, we shall consider the significance of voting coalitions in voting bodies. Here the universal set is a set of human beings which form a decision-making body. For example, the universal set might be the members of a committee, or of a city council, or of a convention, or of the House of Representatives, etc. Each member can cast a certain number of votes. The decision as to whether or not a measure is passed can be decided by a simple majority rule, or two-thirds majority, etc.

Suppose now that a subset of the members of the body forms a coalition in order to pass a measure. The question is whether or not they have enough votes to guarantee passage of the measure. If they have enough votes to carry the measure, then we say they form a winning coalition. If the members not in the coalition can pass a measure of their own, then we say that the original coalition is a losing coalition. Finally, if the members of the coalition cannot carry their measure, and if the members not in the coalition cannot carry their measure, then the coalition is called a blocking coalition.

Let us restate these definitions in set-theoretic terms. A coalition C is winning if they have enough votes to carry an issue; coalition C is losing if the coalition $\tilde{C}$ is winning; and coalition C is blocking if neither C nor $\tilde{C}$ is a winning coalition.

The following facts are immediate consequences of these definitions. The complement of a winning coalition is a losing coalition. The complement of a losing coalition is a winning coalition. The complement of a blocking coalition is a blocking coalition.

Example 2.7 A committee consists of six members each having one vote. A simple majority vote will carry an issue. Then any coalition of four or more members is winning, any coalition with one or two members is losing, and any three-person coalition is blocking. $\diamondsuit$

Example 2.8 Suppose in Example 2.7 one of the six members (say the chair) is given the additional power to break ties. Then any three-person coalition of which the chair is a member is winning, while the other three-person coalitions are losing; hence there are no blocking coalitions. The other coalitions are as in Example 2.7. $\diamondsuit$

Example 2.9 Let the universal set $\mathcal{U}$ be the set $\{\mathrm{x}, \mathrm{y}, \mathrm{w}, \mathrm{z}\}$, where x and y each has one vote, w has two votes, and z has three votes. Suppose it takes five votes to carry a measure. Then the winning coalitions are: $\{\mathrm{z}, \mathrm{w}\}$, $\{\mathrm{z}, \mathrm{x}, \mathrm{y}\}$, $\{\mathrm{z}, \mathrm{w}, \mathrm{x}\}$, $\{\mathrm{z}, \mathrm{w}, \mathrm{y}\}$, and $\mathcal{U}$. The losing coalitions are the complements of these sets. Blocking coalitions are: $\{\mathrm{z}\}$, $\{\mathrm{z}, \mathrm{x}\}$, $\{\mathrm{z}, \mathrm{y}\}$, $\{\mathrm{w}, \mathrm{x}\}$, $\{\mathrm{w}, \mathrm{y}\}$, and $\{\mathrm{w}, \mathrm{x}, \mathrm{y}\}$. $\diamondsuit$

The last example shows that it is not always necessary to list all members of a winning coalition. For example, if the coalition $\{\mathrm{z}, \mathrm{w}\}$ is winning, then it is obvious that the coalition $\{\mathrm{z}, \mathrm{w}, \mathrm{y}\}$ is also winning. In general, if a coalition C is winning, then any other set that has C as a subset will also be winning. Thus we are led to the notion of a minimal winning coalition. A minimal winning coalition is a winning coalition which contains no smaller winning coalition as a subset. Another way of stating this is that a minimal winning coalition is a winning coalition such that, if any member is lost from the coalition, then it ceases to be a winning coalition.

If we know the minimal winning coalitions, then we know everything that we need to know about the voting problem. The winning coalitions are all those sets that contain a minimal winning coalition, and the losing coalitions are the complements of the winning coalitions. All other sets are blocking coalitions.

In Example 2.7 the minimal winning coalitions are the sets containing four members. In Example 2.8 the minimal winning coalitions are the three-member coalitions that contain the tie-breaking member and the four-member coalitions that do not contain the tie-breaking member. The minimal winning coalitions in the third example are the sets $\{\mathrm{z}, \mathrm{w}\}$ and $\{\mathrm{z}, \mathrm{x}, \mathrm{y}\}$.

Sometimes there are committee members who have special powers or lack of power. If a member can pass any measure he or she wishes without needing anyone else to vote with him or her, then we call him or her a dictator. Thus member x is a dictator if and only if $\{x\}$ is a winning coalition. A somewhat weaker but still very powerful member is one who can by himself or herself block any measure. If x is such a member, then we say that x has veto power. Thus x has veto power if and only if $\{x\}$ is a blocking coalition. Finally if x is not a member of any minimal winning coalition, we shall call him or her a powerless member. Thus x is powerless if and only if any winning coalition of which x is a member is a winning coalition without him or her.

Example 2.10 An interesting example of a decision-making body is the Security Council of the United Nations. (We discuss the rules prior to 1966.) The Security Council has eleven members consisting of the five permanent large-nation members called the Big Five, and six small-nation members. In order that a measure be passed by the Council, seven members including all of the Big Five must vote for the measure. Thus the seven-member sets made up of the Big Five plus two small nations are the minimal winning coalitions. Then the losing coalitions are the sets that contain at most four small nations. The blocking coalitions are the sets that are neither winning nor losing. In particular, a unit set that contains one of the Big Five as a member is a blocking coalition. This is the sense in which a Big Five member has a veto. [The possibility of "abstaining" is immaterial in the above discussion.]

In 1966 the number of small-nation members was increased to 10.

A measure now requires the vote of nine members, including all of the Big Five. (See Exercise 11.) ♢

Exercises

1. A committee has w, x, y, and z as members. Member w has two votes, the others have one vote each. List the winning, losing, and blocking coalitions.

2. A committee has n members, each with one vote. It takes a majority vote to carry an issue. What are the winning, losing, and blocking coalitions?

3. Rhe Board of Estimate of New York City consists (that is, consisted at one time) of eight members with voting strength as follows:

s.	Mayor	4 votes
t.	Controller	4
u.	Council President	4
v.	Brooklyn Borough President	2
w.	Manhattan Borough President	2
x.	Bronx Borough President	2
y.	Richmond Borough President	2
z.	Queens Borough President	2

   A simple majority is needed to carry an issue. List the minimal winning coalitions. List the blocking coalitions. Do the same if we give the mayor the additional power to break ties.

4. A company has issued 100,000 shares of common stock and each share has one vote. How many shares must a stockholder have to be a dictator? How many to have a veto?

   [Ans. 50,001; 50,000.]

5. In Exercise 4, if the company requires a two-thirds majority vote to carry an issue, how many shares must a stockholder have to be a dictator or to have a veto?

   [Ans. At least 66,667; at least 33,334.]

6. Prove that if a committee has a dictator as a member, then the remaining members are powerless.

7. We can define a maximal losing coalition in analogy to the minimal winning coalitions. What is the relation between the maximal losing and minimal winning coalitions? Do the maximal losing coalitions provide all relevant information?

8. Prove that any two minimal winning coalitions have at least one member in common.

9. Find all the blocking coalitions in the Security Council example (Example 2.10).

10. Prove that if a member has veto power and if he or she together with any one other member can carry a measure, then the distribution of the remaining votes is irrelevant.

11. Find the winning, losing, and blocking coalitions in the Security Council, using the revised (1966) structure.

Suggested reading.

Birkhoff, G., and S. MacLane, *A Survey of Modern Algebra*, 1953, Chapter XI.

Tarski, A., *Introduction to Logic*, 2d rev. ed., 1946, Chapter IV.

Allendoerfer, C. B., and C. O. Oakley, *Principles of Mathematics*, 1955, Chapter V.

Johnstone, H. W., Jr., *Elementary Deductive Logic*, 1954, Part Three.

Breuer, Joseph, *Introduction to the Theory of Sets*, 1958.

Fraenkel, A. A., *Abstract Set Theory*, 1953.

Kemeny, John G., Hazleton Mirkil, J. Laurie Snell, and Gerald L. Thompson, *Finite Mathematical Structures*, 1959, Chapter 2.

Chapter 3

Partitions and counting

3.1 Partitions

The problems to be studied in this chapter can be most conveniently described in terms of partitions of a set. A *partition* of a set $\mathcal{U}$ is a subdivision of the set into subsets that are disjoint and exhaustive, i.e., every element of $\mathcal{U}$ must belong to one and only one of the subsets. The subsets A_i in the partition are called *cells*. Thus $[A_1, A_2, \ldots, A_r]$ is a partition of $\mathcal{U}$ if two conditions are satisfied: (1) $A_i \cap A_j = \emptyset$ if $i \neq j$ (the cells are disjoint) and (2) $A_1 \cup A_2 \cup \ldots \cup A_r = \mathcal{U}$ (the cells are exhaustive).

Example 3.1 If $\mathcal{U} = \{a, b, c, d, e\}$, then $[\{a, b\}, \{c, d, e\}]$ and $[\{b, c, e\}, \{a\}, \{d\}]$ and $[\{a\}, \{b\}, \{c\}, \{d\}, \{e\}]$ are three different partitions of $\mathcal{U}$. The last is a partition into unit sets. $\diamondsuit$

The process of going from a fine to a less fine analysis of a set of logical possibilities is actually carried out by means of a partition. For example, let us consider the logical possibilities for the first three games of the World Series if the Yankees play the Dodgers. We can list the possibilities in terms cf the winner of each game as

$$\{\text{YYY}, \text{YYD}, \text{YDY}, \text{DYY}, \text{DDY}, \text{DYD}, \text{YDD}, \text{DDD}\}.$$

We form a partition by putting all the possibilities with the same number of wins for the Yankees in a single cell,

$$[\{\text{YYY}\}, \{\text{YYD}, \text{YDY}, \text{DYY}\}, \{\text{DDY}, \text{DYD}, \text{YDD}\}, \{\text{DDD}\}].$$

Thus, if we wish the possibilities to be Yankees win three games, win two, win one, win zero, then we are considering a less detailed analysis obtained from the former analysis by identifying the possibilities in each cell of the partition.

If $[A_1, A_2, \ldots, A_r]$ and $[B_1, B_2, \ldots, B_s]$ are two partitions of the same set $\mathcal{U}$, we can obtain a new partition by considering the collection of all subsets of $\mathcal{U}$ of the form $A_i \cap B_j$ (see Exercise 7). This new partition is called the *cross-partition* of the original two partitions.

Example 3.2 A common use of cross-partitions is in the problem of classification. For example, from the set $\mathcal{U}$ of all life forms we can form the partition $[P, A]$ where P is the set of all plants and A is the set of all animals. We may also form the partition $[E, F]$ where E is the set of extinct life forms and F is the set of all existing life forms. The cross-partition

$$[P \cap E, P \cap F, A \cap E, A \cap F]$$

gives a complete classification according to the two separate classifications. $\diamondsuit$

Many of the examples with which we shall deal in the future will relate to processes which take place in stages. It will be convenient to use partitions and cross-partitions to represent the stages of the process. The graphical representation of such a process is, of course, a tree. For example, suppose that the process is such that we learn in succession the truth values of a series of statements relative to a given situation. If $\mathcal{U}$ is the set of logical possibilities for the situation, and p is a statement relative to $\mathcal{U}$, then the knowledge of the truth value of p amounts to knowing which cell of the partition $[P, \tilde{P}]$ contains the actual possibility. Recall that P is the truth set of p, and $\tilde{P}$ is the truth set of $\neg p$. Suppose now we discover the truth value of a second statement q. This information can again be described by a partition, namely, $[Q, \tilde{Q}]$. The two statements together give us information which can be represented by the cross-partition of these two partitions,

$$[P \cap Q, P \cap \tilde{Q}, \tilde{P} \cap Q, \tilde{P} \cap \tilde{Q}].$$

That is, if we know the truth values of p and q, we also know which of the cells of this cross-partition contains the particular logical possibility describing the given situation. Conversely, if we knew which cell contained the possibility, we would know the truth values for the statements p and q.

The information obtained by the additional knowledge of the truth value of a third statement r, having a truth set R, can be represented by the cross-partition of the three partitions $[P, \tilde{P}]$, $[Q, \tilde{Q}]$, $[R, \tilde{R}]$ This cross-partition is

$$[P \cap Q \cap R, P \cap Q \cap \tilde{R}, P \cap \tilde{Q} \cap R, P \cap \tilde{Q} \cap \tilde{R}, \tilde{P} \cap Q \cap R, \tilde{P} \cap Q \cap \tilde{R}, \tilde{P} \cap \tilde{Q} \cap R, \tilde{P} \cap \tilde{Q} \cap \tilde{R}].$$

Notice that now we have the possibility narrowed down to being in one of $8 = 2^3$ possible cells. Similarly, if we knew the truth values of n statements, our partition would have 2^n cells.

If the set $\mathcal{U}$ were to contain 2^{20} (approximately one million) logical possibilities, and if we were able to ask yes-no questions in such a way that the knowledge of the truth value of each question would cut the number of possibilities in half each time, then we could determine in 20 questions any given possibility in the set $\mathcal{U}$. We could accomplish this kind of questioning, for example, if we had a list of all the possibilities and were allowed to ask "Is it in the first half?" and, if the answer is yes, then "Is it in the first one-fourth?", etc. In practice we ordinarily do not have such a list, and we can only approximate this procedure.

Example 3.3 In the familiar radio game of twenty questions it is not unusual for a contestant to try to carry out a partitioning of the above kind. For example, he or she may know that he or she is trying to guess a city. He or she might ask, "Is the city in North America?" and if the answer is yes, "Is it in the United States?" and if yes, "Is it west of the Mississippi?" and if no, "Is it in the New England states?", etc. Of course, the above procedure does not actually divide the possibilities exactly in half each time. The more nearly the answer to each question comes to dividing the possibilities in half, the more certain one can be of getting the answer in twenty questions, if there are at most a million possibilities. $\diamondsuit$

Exercises

1. If $\mathcal{U}$ is the set of integers from 1 to 6, find the cross-partitions of the following pairs of partitions

   (a) $[\{1, 2, 3\}, \{4, 5, 6\}]$ and $[\{1, 4\}, \{2, 3, 5, 6\}]$.

   [Ans. $[\{1\}, \{2, 3\}, \{4\}, \{5, 6\}]$.]

   (b) $[\{1, 2, 3, 4, 5\}, \{6\}]$ and $[\{1, 3, 5\}, \{2, 6\}, \{4\}]$.

2. A coin is thrown three times. List the possibilities according to which side turns up each time. Give the partition formed by

putting in the same cell all those possibilities for which the same number of heads occur.

3. Let p and q be two statements with truth set P and Q. What can be said about the cross-partition of $[P, \tilde{P}]$ and $[Q, \tilde{Q}]$ in the case that

   (a) p implies q.

   [Ans. $P \cap \tilde{Q} = \emptyset$.]

   (b) p is equivalent to q.

   (c) p and q are inconsistent.

4. Consider the set of eight states consisting of Illinois, Colorado, Michigan, New York, Vermont, Texas, Alabama, and California.

   (a) Show that in three "yes" or "no" questions one can identify any one of the eight states.

   (b) Design a set of three "yes" or "no" questions which can be answered independently of each other and which will serve to identify any one of the states.

5. An unabridged dictionary contains about 600,000 words and 3000 pages. If a person chooses a word from such a dictionary, is it possible to identify this word by twenty "yes" or "no" questions? If so, describe the procedure that you would use and discuss the feasibility of the procedure. (One approach is the following. Use 12 questions to locate the page, but then you may need 9 questions to locate the word.)

6. Jones has two parents, each of his or her parents had two parents, each of these had two parents, etc. Tracing a person's family tree back 40 generations (about 1000 years) gives Jones 2^{40} ancestors, which is more people than have been on the earth in the last 1000 years. What is wrong with this argument?

7. Let $[A_1, A_2, A_3]$ and $[B_1, B_2]$ be two partitions. Prove that the cross-partition of the two given partitions really is a partition, that is, it satisfies requirements (1) and (2) for partitions.

8. The cross-partition formed from the truth sets of n statements has 2^n cells. As seen in Chapter **??**, the truth table of a statement compounded from n statements has 2^n rows. What is the relationship between these two facts?

9. Let p and q be statements with truth sets P and Q. Form the partition $[P \cap Q, P \cap \tilde{Q}, \tilde{P} \cap Q, \tilde{P} \cap \tilde{Q}]$. State in each case below which of the cells must be empty in order to make the given statement a logically true statement.

   (a) $p \to q$

   (b) $p \leftrightarrow q$

   (c) $p \vee \neg p$

   (d) p

10. A partition $[A_1, A_2, \ldots, A_n]$ is said to be a *refinement* of the partition $[B_1, B_2, \ldots, B_m]$ if every A_j is a subset of some B_k. Show that a cross-partition of two partitions is a refinement of each of the partitions from which the cross-partition is formed.

11. Consider the partition of the people in the United States determined by classification according to states. The classification according to county determines a second partition. Show that this is a refinement of the first partition. Give a third partition which is different from each of these and is a refinement of both.

12. What can be said concerning the cross-partition of two partitions, one of which is a refinement of the other?

13. Given nine objects, of which it is known that eight have the same weight and one is heavier, show how, in two weighings with a pan balance, the heavy one can be identified.

14. Suppose that you are given thirteen objects, twelve of which are the same, but one is either heavier or lighter than the others. Show that, with three weighings using a pan balance, it is possible to identify the odd object. (A complete solution to this problem is given on page 42 of *Mathematical Snapshots*, second edition, by H. Steinhaus.)

15. A subject can be completely classified by introducing several simple subdivisions and taking their cross-partition. Thus, courses

in college may be partitioned according to subject, level of advancement, number of students, hours per week, interests, etc. For each of the following subjects, introduce five or more partitions. How many cells are there in the complete classification (cross-partition) in each case?

(a) Detective stories.

(b) Diseases.

16. Assume that in a given generation x men are Republicans and y are Democrats and that the total number of men remains at 50 million in each generation. Assume further that it is known that 20 per cent of the sons of Republicans are Democrats and 30 per cent of the sons of Democrats are Republicans in any generation. What conditions must x and y satisfy if there are to be the same number of Republicans in each generation? Is there more than one choice for x and y? If not, what must x and y be?

[Partial Ans. There are 30 million Republicans.]

17. Assume that there are 30 million Democratic and 20 million Republican men in the country. It is known that p per cent of the sons of Democrats are Republicans, and q per cent of the sons of Republicans are Democrats. If the total number of men remains 50 million, what condition must p and q satisfy so that the number in each party remains the same? Is there more than one choice of p and q?

3.2 The number of elements in a set

The remainder of this chapter will be devoted to certain counting problems. For any set X we shall denote by $n(X)$ the number of elements in the set.

Suppose we know the number of elements in certain given sets and wish to know the number in other sets related to these by the operations of unions, intersections, and complementations. As an example, consider the following problem.

Suppose that we are told that 100 students take mathematics, and 150 students take economics. Can we then tell how many take either mathematics or economics? The answer is no, since clearly we would

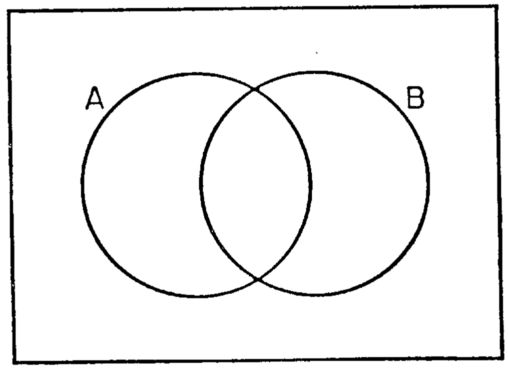


Figure 3.1: $\diamondsuit$

also need to know how many students take both courses. If we know that no student takes both courses, i.e., if we know that the two sets of students are disjoint, then the answer would be the sum of the two numbers or 250 students.

In general, if we are given disjoint sets A and B, then it is true that $n(A \cup B) = n(A) + n(B)$. Suppose now that A and B are not disjoint as shown in Figure 3.1. We can divide the set A into disjoint sets $A \cap \tilde{B}$ and $A \cap B$. Similarly, we can divide B into the disjoint sets $\tilde{A} \cap B$ and $A \cap B$. Thus,

$$n(A) = n(A \cap \tilde{B}) + n(A \cap B),$$
$$n(B) = n(\tilde{A} \cap B) + n(A \cap B).$$

Adding these two equations, we obtain

$$n(A) + (B) = n(A \cap B) + n(A \cap \tilde{B}) + n(\tilde{A} \cap B) + 2n(A \cap B).$$

Since the sets $A \cap B$, $A \cap \tilde{B}$, and $\tilde{A} \cap B$ are disjoint sets whose union is $A \cup B$, we obtain the formula

$$n(A \cup B) = n(A) + n(B) - n(A \cap B),$$

which is valid for any two sets A and B.

Example 3.4 Let p and q be statements relative to a set $\mathcal{U}$ of logical possibilities. Denote by P and Q the truth sets of these statements. The truth set of $p \vee q$ is $P \cup Q$ and the truth set of $p \wedge q$ is $P \cap Q$. Thus the above formula enables us to find the number of cases where $p \vee q$ is true if we know the number of cases for which p, q, and $p \wedge q$ are true. $\diamondsuit$

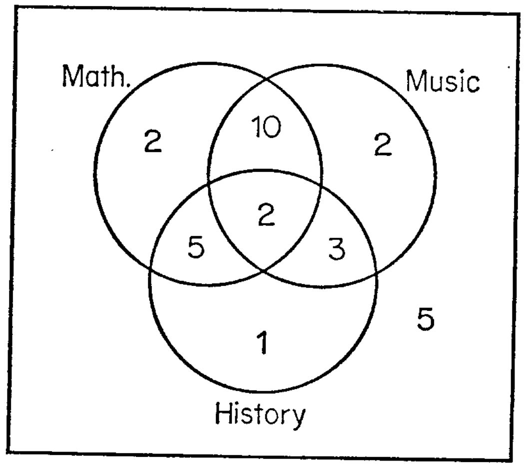


Figure 3.2: $\diamondsuit$

Example 3.5 More than two sets. It is possible to derive formulas for the number of elements in a set which is the union of more than two sets (see Exercise 6), but usually it is easier to work with Venn diagrams. For example, suppose that the registrar of a school reports the following statistics about a group of 30 students: 19 take mathematics. 17 take music. 11 take history. 12 take mathematics and music. 7 take history and mathematics. 5 take music and history. 2 take mathematics, history, and music. We draw the Venn diagram in Figure 3.2 and fill in the numbers for the number of elements in each subset working from the bottom of our list to the top. That is, since 2 students take all three courses, and 5 take music and history, then 3 take history and music but not mathematics, etc. Once the diagram is completed we can read off the number which take any combination of the courses. For example, the number which take history but not mathematics is $3 + 1 = 4$. $\diamondsuit$

Example 3.6 Cancer studies. The following reasoning is often found in statistical studies on the effect of smoking on the incidence of lung cancer. Suppose a study has shown that the fraction of smokers among those who have lung cancer is greater than the fraction of smokers among those who do not have lung cancer. It is then asserted that the fraction of smokers who have lung cancer is greater than the fraction of nonsmokers who have lung cancer. Let us examine this argument.

Let S be the set of all smokers in the population, and C be the

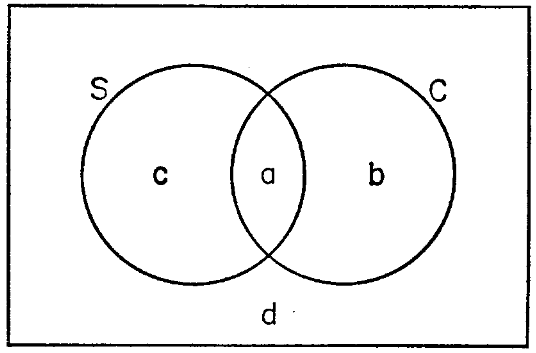


Figure 3.3: $\diamondsuit$

set of all people with lung cancer. Let $a = n(S \cap C)$, $b = n(\tilde{S} \cap C)$, $c = n(S \cap \tilde{C})$ and $d = n(\tilde{S} \cap \tilde{C})$, as indicated in Figure 3.3. The fractions in which we are interested are

$$p_1 = \frac{a}{a+b}, p_2 = \frac{c}{c+d}, p_3 = \frac{a}{a+c}, p_4 = \frac{b}{b+d},$$

where p_1 is the fraction of those with lung cancer that smoke, p_2 the fraction of those without lung cancer that smoke, p_3 the fraction of smokers who have lung cancer, and p_4 the fraction of nonsmokers who have cancer.

The argument above states that if $p_1 > p_2$, then $p_3 > p_4$. The hypothesis,

$$\frac{a}{a+b} > \frac{c}{c+d}$$

is true if and only if $ac + ad > ac + bc$, that is, if and only if $ad > bc$. The conclusion

$$\frac{a}{a+c} > \frac{b}{b+d}$$

is true if and only if $ab + ad > ab + bc$, that is, if and only if $ad > bc$. Thus the two statements $p_1 > p_2$ and $p_3 > p_4$ are in fact equivalent statements, so that the argument is valid. $\diamondsuit$

Exercises

1. In Example 3.5, find

   (a) The number of students that take mathematics but do not take history.

[Ans. 12.]

(b) The number that take exactly two of the three courses.

(c) The number that take one or none of the courses.

2. In a chemistry class there are 20 students, and in a psychology class there are 30 students. Find the number in either the psychology class or the chemistry class if

   (a) The two classes meet at the same hour.

   [Ans. 50.]

   (b) The two classes meet at different hours and 10 students are enrolled in both courses.

   [Ans. 40.]

3. If the truth set of a statement p has 10 elements, and the truth set of a statement q has 20 elements, find the number of elements in the truth set of $p \vee q$ if

   (a) p and q are inconsistent.

   (b) p and q are consistent and there are two elements in the truth set of $p \wedge q$.

4. If p is a statement that is true in ten cases, and q is a statement that is true in five cases, find the number of cases in which both p and q are true if $p \vee q$ is true in ten cases. What relation holds between p and q?

5. Assume that the incidence of lung cancer is 16 per 100,000, and that it is estimated that 75 per cent of those with lung cancer smoke and 60 per cent of those without lung cancer smoke. (These numbers are fictitious.) Estimate the fraction of smokers with lung cancer, and the fraction of nonsmokers with lung cancer.

   [Ans. 20 and 10 per 100,000.]

6. Let A, B, and C be any three sets of a universal set $\mathcal{U}$. Draw a Venn diagram and show that

$$\begin{aligned} n(A \cup B \cup C) &= n(A) + n(B) + n(C) \\ &\quad - n(A \cap B) - n(A \cap C) - n(B \cap C) \\ &\quad + n(A \cap B \cap C). \end{aligned}$$

7. Analyze the data given below and draw a Venn diagram like that in Figure 3.2. Assuming that every student in the school takes one of the courses, find the total number of students in the school.

   (a) First case:

   28 students take English.
   23 students take French.
   23 students take German.
   12 students take English and French.
   11 students take English and German.
   8 students take French and German.
   5 students take all three courses.

   (b) Second case:

   36 students take English.
   23 students take French.
   13 students take German.
   6 students take English and French.
   11 students take English and German.
   4 students take French and German.
   1 students take all three courses.

   (Comment on this result.)

8. Suppose that in a survey concerning the reading habits of students it is found that:

   60 per cent read magazine A; 50 per cent read magazine B; 50 per cent read magazine C; 30 per cent read magazines A and B; 20 per cent read magazines B and C; 30 per cent read magazines A and C; 10 per cent read all three magazines.

   (a) What per cent read exactly two magazines?

   [Ans. 50.]

   (b) What per cent do not read any of the magazines?

   [Ans. 10.]

9. If p and q are equivalent statements and $n(P) = 10$, what is $n(P \cup Q)$?

10. If p implies q, prove that $n(P \cup \tilde{Q}) = n(P) + n(\tilde{Q})$.

11. On a transcontinental airliner, there are 9 boys, 5 American children, 9 men, 7 foreign boys, 14 Americans, 6 American males, and 7 foreign females. What is the number of people on the plane?

[Ans. 33.]

Supplementary exercises.

12. Prove that $n(\tilde{A}) = n(\mathcal{U}) - n(A)$.

13. Show that $n(\tilde{A} \cap \tilde{B}) = n(A \,\tilde{\cup}\, B) = n(\mathcal{U}) - n(A \cup B)$.

14. In a collection of baseball players there are ten who can play only outfield positions, five who can play only infield positions but cannot pitch, three who can pitch, four who can play any position but pitcher, and two who can play any position at all. How many players are there in all?

[Ans. 22.]

15. Ivyten College awarded 38 varsity letters in football, 15 in basketball, and 20 in baseball. If these letters went to a total of 58 men and only three of these men lettered in all three sports, how many men received letters in exactly two of the three sports?

[Ans. 9.]

16. Let $\mathcal{U}$ be a finite set. For any two sets A and B define the "distance" from A to B to be $d(A, B) = n(A \cap \tilde{B}) + n(\tilde{A} \cap B)$.

    (a) Show that $d(A, B) \geq 0$. When is $d(A, B) = 0$?

    (b) If A, B, and C are nonintersecting sets, show that

    $$d(A, C) \leq d(A, B) + d(B, C).$$

    (c) Show that for any three sets A, B, and C

    $$d(A, C) \leq d(A, B) + d(B, C).$$

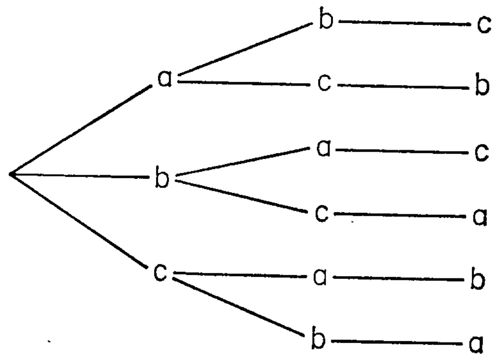


Figure 3.4: ♢

3.3 Permutations

We wish to consider here the number of ways in which a group of n different objects can be arranged. A listing of n different objects in a certain order is called a *permutation* of the n objects. We consider first the case of three objects, a, b, and c. We can exhibit all possible permutations of these three objects as paths of a tree, as shown in Figure 3.4. Each path exhibits a possible permutation, and there are six such paths. We could also list these permutations as follows: abc, bca, acb, cab, bac, cba. If we were to construct a similar tree for n objects, we would find that the number of paths could be found by multiplying together the numbers n, $n-1$, $n-2$, continuing down to the number 1. The number obtained in this way occurs so often that we give it a symbol, namely $n!$, which is read "n factorial". Thus, for example, $3! = 3 \cdot 2 \cdot 1 = 6$, $4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$, etc. For reasons which will be clear later, we define $0! = 1$. Thus we can say *there are* $n!$ *different permutations of* n *distinct objects.*

Example 3.7 In the game of Scrabble, suppose there are seven lettered blocks from which we try to form a seven-letter word. If the seven letters are all different, we must consider $7! = 5040$ different orders. ♢

Example 3.8 A quarterback has a sequence of ten plays. Suppose his or her coach instructs him or her to run through the ten-play sequence without repetition. How much freedom is left to the quarterback? He or she may choose any one of $10! = 3,628,800$ orders in which to call the plays. ♢

Example 3.9 How many ways can n people be seated around a circular table? When this question is asked, it is usually understood that two arrangements are different only if at least one person has a different person on the right in the two arrangements. Consider then one person in a fixed position. There are $(n-1)!$ ways in which the other people may be seated. We have now counted all the arrangements we wish to consider different. $\diamondsuit$

A general principle. There are many counting problems for which it is not possible to give a simple formula for the number of possible cases. In many of these the only way to find the number of cases is to draw a tree and count them (see Exercise 4). In some problems, the following general principle is useful.

If one thing can be done in exactly r different ways, for each of these a second thing can be done in exactly s different ways, for each of the first two, a third can be done in exactly t ways, etc., then the sequence of things can be done in the product of the numbers of ways in which the individual things can be done, i.e., $r \cdot s \cdot t$ ways.

The validity of the above general principle can be established by thinking of a tree representing all the ways in which the sequence of things can be done. There would be r branches from the starting position. From the ends of each of these r branches there would be s new branches, and from each of these t new branches, etc. The number of paths through the tree would be given by the product $r \cdot s \cdot t$.

Example 3.10 The number of permutations of n distinct objects is a special case of this principle. If we were to list all the possible permutations, there would be n possibilities for the first, for each of these $n-1$ for the second, etc., until we came to the last object, and for which there is only one possibility. Thus there are $n(n-1)\ldots 1 = n!$ possibilities in all. $\diamondsuit$

Example 3.11 If there are three roads from city x to city y and two roads from city y to city z, then there are $3 \cdot 2 = 6$ ways that a person can drive from city x to city z passing through city y. $\diamondsuit$

Example 3.12 Suppose there are n applicants for a certain job. Three interviewers are asked independently to rank the applicants according to their suitability for the job. It is decided that an applicant will be hired if he or she is ranked first by at least two of the three interviewers.

What fraction of the possible reports would lead to the acceptance of some candidate? We shall solve this problem by finding the fraction of the reports which do not lead to an acceptance and subtract this answer from 1. Frequently, an indirect attack of this kind on a problem is easier than the direct approach. The total number of reports possible is $(n!)^3$ since each interviewer can rank the men in $n!$ different ways. If a particular report does not lead to the acceptance of a candidate, it must be true that each interviewer has put a different applicant in first place. This can be done in $n(n-1)(n-2)$ different ways by our general principle. For each possible first choices, there are $[(n-1)!]^3$ ways in which the remaining men can be ranked by the interviewers. Thus the number of reports which do not lead to acceptance is $n(n-1)(n-2)[(n-1)!]^3$. Dividing this number by $(n!)^3$ we obtain

$$\frac{(n-1)(n-2)}{n^2}$$

as the fraction of reports which fail to accept a candidate. The fraction which leads to acceptance is found by subtracting this fraction from 1 which gives

$$\frac{3n-2}{n^2}$$

For the case of three applicants, we see that $\frac{7}{9}$ of the possibilities lead to acceptance. Here the procedure might be criticized on the grounds that even if the interviewers are completely ineffective and are essentially guessing there is a good chance that a candidate will be accepted on the basis of the reports. For n equal to ten, the fraction of acceptances is only .28, so that it is possible to attach more significance to the interviewers ratings, if they reach a decision. $\diamondsuit$

Exercises

1. In how many ways can five people be lined up in a row for a group picture? In how many ways if it is desired to have three in the front row and two in the back row?

   [Ans. 120;120.]

2. Assuming that a baseball team is determined by the players and the position each is playing, how many teams can be made from 13 players if

   (a) Each player can play any position?
   (b) Two of the players can be used only as pitchers?

3. Grades of A, B, C, D, or F are assigned to a class of five students.

   (a) How many ways may this be done, if no two students receive the same grade?

   [Ans. 120.]

   (b) Two of the students are named Smith and Jones. How many ways can grades be assigned if no two students receive the same grade and Smith must receive a higher grade than Jones?

   [Ans. 60.]

   (c) How many ways may grades be assigned if only grades of A and F are assigned?

   [Ans. 32.]

4. A certain club wishes to admit seven new members, four of whom are Republicans and three of whom are Democrats. Suppose the club wishes to admit them one at a time and in such a way that there are always more Republicans among the new members than there are Democrats. Draw a tree to represent all possible ways in which new members can be admitted, distinguishing members by their party only.

5. There are three different routes connecting city A to city B. How many ways can a round trip be made from A to B and back? How many ways if it is desired to take a different route on the way back?

   [Ans. 9;6.]

6. How many different ways can a ten-question multiple-choice exam be answered if each question has three possibilities, a, b, and c? How many if no two consecutive answers are the same?

7. Modify Example 3.12 so that, to be accepted, an applicant must be first in two of the interviewers' ratings and must be either first or second in the third interviewers' rating. What fraction of the possible reports lead to acceptance in the case of three applicants? In the case of n?

[Ans. $\frac{4}{9}$; $\frac{4}{n^2}$.]

8. A town has 1240 registered Republicans. It is desired to contact each of these by phone to announce a meeting. A committee of r people devise a method of phoning s people each and asking each of these to call t new people. If the method is such that no person is called twice,

   (a) How many people know about the meeting after the phoning?

   (b) If the committee has 40 members and it is desired that all 1240 Republicans be informed of the meeting and that s and t should be the same, what should they be?

9. In the Scrabble example (Example 3.7), suppose the letters are Q, Q, U, F, F, F, A. How many distinguishable arrangements are there for these seven letters?

   [Ans. 420.]

10. How many different necklaces can be made

    (a) If seven different sized beads are available?

    [Ans. 360.]

    (b) If six of the beads are the same size and one is larger?

    [Ans. 1.]

    (c) If the beads are of two sizes, five of the smaller size and two of the larger size?

    [Ans. 3.]

11. Prove that two people in Columbus, Ohio, have the same initials.

12. Find the number of distinguishable arrangements for each of the following collections of five symbols. (The same letters with different subscripts indicate distinguishable objects.)

    (a) A_1, A_2, B_1, B_2, B_3.

    [Ans. 120.]

(b) A, A, B_1, B_2, B_3.

[Ans. 60.]

(c) A, A, B, B, B.

13. Show that the number of distinguishable arrangements possible for n objects, n_1 of type 1, n_2 of type 2, etc., for r different types is

$$\frac{n!}{n_1!n_2!\cdots n_r!}.$$

Supplementary exercises.

14. (a) How many four digit numbers can be formed from the digits 1, 2, 3, 4, using each digit only once?

    (b) How many of these numbers are less than 3000?

[Ans. 12.]

15. How many license plates can be made if they are to contain five symbols, the first two being letters and the last three digits?

16. How many signals can a ship show if it has seven flags and a signal consists of five flags hoisted vertically on a rope?

[Ans. 2520.]

17. We must arrange three green, two red, and four blue books on a single shelf.

    (a) In how many ways can this be done if there are no restrictions?

    (b) In how many ways if books of the same color must be grouped together?

    (c) In how many ways if, in addition to the restriction in 17b, the red books must be to the left of the blue books?

    (d) In how many ways if, in addition to the restrictions in 17b and 17c, the red and blue books must not be next to each other?

[Ans. 288.]

18. A youngster has three shades of nail polish with which to paint his or her fingernails. In how many ways can he or she do this (each nail being one solid color) if there are no more than two different shades on each hand?

[Ans. 8649.]

3.4 Counting partitions

Up to now we have not had occasion to consider the partitions $[\{1,2\},\{3,4\}]$ and $[\{3,4\},\{1,2\}]$ of the integers from 1 to 4 as being different partitions. Here it will be convenient to do so, and to indicate this distinction we shall use the term *ordered partition.* An *ordered partition with r cells* is a partition with r cells (some of which may be empty), with a particular order specified for the cells.

We are interested in counting the number of possible ordered partitions with r cells that can be formed from a set of n objects having a prescribed number of elements in each cell. We consider first a special case to illustrate the general procedure.

Suppose that we have eight students, A, B, C, D, E, F, G, and H, and we wish to assign these to three rooms, Room 1, which is a triple room, Room 2, a triple room, and Room 3, a double room. In how many different ways can the assignment be made? One way to assign the students is to put them in the rooms in the order in which they arrive, putting the first three in Room 1, the next three in Room 2, and the last two in Room 3. There are 8! ways in which the students can arrive, but not all of these lead to different assignments. We can represent the assignment corresponding to a particular order of arrival as follows,

$$|BCA|DFE|HG|.$$

In this case, B, C, and A are assigned to Room 1, D, F, and E to Room 2, and H and G to Room 3. Notice that orders of arrival which simply change the order within the rooms lead to the same assignment. The number of different orders of arrival which lead to the same assignment as the one above is the number of arrangements which differ from the given one only in that the arrangement within the rooms is different. There are $3! \cdot 3! \cdot 2!$ such orders of arrival, since we can arrange the three in Room 1 in 3! different ways, for each of these the ones in Room 2 in 3! different ways, and for each of these, the ones in Room

3 in 2! ways. Thus we can divide the 8! different orders of arrival into groups of $3! \cdot 3! \cdot 2!$ different orders such that all the orders of arrival in a single group lead to the same room assignment. Since there are $3! \cdot 3! \cdot 2!$ elements in each group and 8! elements altogether, there are $\frac{8!}{3!3!2!}$ groups, or this many different room assignments.

The same argument could be carried out for n elements and r rooms, with n_1 in the first, n_2 in the second, etc. This would lead to the following result. Let $n_1, n_2, \ldots, n_r$ be nonnegative integers with

$$n_1 + n_2 + \ldots + n_r = n.$$

Then:

The number of ordered partitions with r cells $[A_1, A_2, \ldots, A_r]$ of a set of n elements with n_1 in the first cell, n_2 in the second, etc. is

$$\frac{n!}{n_1!n_2!\ldots n_r!}$$

We shall denote this number by the symbol

$$\binom{n!}{n_1!, n_2!, \ldots, n_r!}.$$

Note that this symbol is defined only if $n_1 + n_2 + \ldots + n_r = n$.

The special case of two cells is particularly important. Here the problem can be stated equivalently as the problem of finding the number of subsets with r elements that can be chosen from a set of n elements. This is true because any choice defines a partition $[A, \tilde{A}]$, where A is the set of elements chosen and $\tilde{A}$ is the set of remaining elements.

The number of such partitions is $\frac{n!}{r!(n-r)!}$ and hence this is also the number of subsets with r elements. Our notation $\binom{n}{r, n-r}$ for this case is shortened to $\binom{n}{r}$.

Notice that $\binom{n}{n-r}$ is the number of subsets with $n-r$ elements which can be chosen from n, which is the number of partitions of the form $[\tilde{A}, A]$ above. Clearly, this is the same as the number of $[A, \tilde{A}]$ partitions. Hence $\binom{n}{r} = \binom{n}{n-r}$.

Example 3.13 A college has scheduled six football games during a season. How many ways can the season end in two wins, three losses, and one tie? From each possible outcome of the season, we form a partition, with three cells, of the opposing teams. In the first cell we put the teams which our college defeats, in the second the teams to which our college loses, and in the third cell the teams which our college ties. There are $\binom{6}{2,3,1} = 60$ such partitions, and hence 60 ways in which the season can end with two wins, three losses, and one tie. $\diamondsuit$

Example 3.14 In the game of bridge, the hands N, E, S, and W determine a partition of the 52 cards having four cells, each with 13 elements. Thus there are $\frac{52!}{13!13!13!13!}$ different bridge deals. This number is about $5.3645 \cdot 10^{28}$, or approximately 54 billion billion billion deals. $\diamondsuit$

Example 3.15 The following example will be important in probability theory, which we take up in the next chapter. If a coin is thrown six times, there are 2^6 possibilities for the outcome of the six throws, since each throw can result in either a head or a tail. How many of these possibilities result in four heads and two tails? Each sequence of six heads and tails determines a two-cell partition of the numbers from one to six as follows: In the first cell put the numbers corresponding to throws which resulted in a head, and in the second put the numbers corresponding to throws which resulted in tails. We require that the first cell should contain four elements and the second two elements. Hence the number of the 2^6 possibilities which lead to four heads and two tails is the number of two-cell partitions of six elements which have four elements in the first cell and two in the second cell. The answer is $\binom{6}{4} = 15$. For n throws of a coin, a similar analysis shows that there are $\binom{n}{r}$ different sequences of H's and T's of length n which have exactly r heads and $n - r$ tails. $\diamondsuit$

Exercises

1. Compute the following numbers.

   (a) $\binom{7}{5}$

   [Ans. 21.]

   (b) $\binom{3}{2}$

   (c) $\binom{7}{2}$

   (d) $\binom{250}{249}$

[Ans. 250.]

(e) $\binom{5}{0}$

(f) $\binom{5}{1,2,2}$

(g) $\binom{4}{2,0,2}$

[Ans. 6.]

(h) $\binom{2}{1,1,1}$

2. Give an interpretation for $\binom{n}{0}$ and also for $\binom{n}{n}$. Can you now give a reason for making $0! = 1$?

3. How many ways can nine students be assigned to three triple rooms? How many ways if one particular pair of students refuse to room together?

[Ans. 1680; 1260.]

4. A group of seven boys and ten girls attends a dance. If all the boys dance in a particular dance, how many choices are there for the girls who dance? For the girls who do not dance? How many choices are there for the girls who do not dance, if three of the girls are sure to be asked to dance?

5. Suppose that a course is given at three different hours. If fifteen students sign up for the course,

   (a) How many possibilities are there for the ways the students could distribute themselves in the classes?

   [Ans. 3^{15}.]

   (b) How many of the ways would give the same number of students in each class?

   [Ans. 756,756.]

6. A college professor anticipates teaching the same course for the next 35 years. So as not to become bored with his or her jokes, he or she decides to tell exactly three jokes every year and in no two years to tell exactly the same three jokes. What is the minimum number of jokes that will accomplish this? What is the minimum number if he or she determines never to tell the same joke twice?

7. How many ways can you answer a ten-question true-false exam, marking the same number of answers true as you do false? How many if it is desired to have no two consecutive answers the same?

8. From three Republicans and three Democrats, find the number of committees of three which can be formed

   (a) With no restrictions.

   [Ans. 20.]

   (b) With three Republicans and no Democrats.

   [Ans. 1.]

   (c) With two Republicans and one Democrat.

   [Ans. 9.]

   (d) With one Republican and two Democrats.

   [Ans. 9.]

   (e) With no Republicans and three Democrats.

   [Ans. 1.]

   What is the relation between your answers to the five parts of this question?

9. Exercise 8 suggests that the following should be true.

$$\binom{2n}{n} = \binom{n}{0}\binom{n}{n} + \binom{n}{1}\binom{n}{n-1} + \binom{n}{2}\binom{n}{n-2} + \ldots + \binom{n}{n}\binom{n}{0}.$$

   Show that it is true.

10. A student needs to choose two electives from six possible courses.

   (a) How many ways can he or she make his or her choice?

   [Ans. 15.]

   (b) How many ways can he or she choose if two of the courses meet at the same time?

   [Ans. 14.]

(c) How many ways can he or she choose if two of the courses meet at 10 o'clock, two at 11 o'clock, and there are no other conflicts among the courses?

[Ans. 13.]

Supplementary exercises.

11. Consider a town in which there are three plumbers, A, B, and C. On a certain day six residents of the town telephone for a plumber. If each resident selects a plumber from the telephone directory, in how many ways can it happen that

    (a) Three residents call A, two residents call B, and one resident calls C?

    [Ans. 60.]

    (b) The distribution of calls to the plumbers is three, two, and one?

    [Ans. 360.]

12. Two committees (a labor relations committee and a quality control committee) are to be selected from a board of nine men. The only rules are (1) the two committees must have no members in common, and (2) each committee must have at least four men. In how many ways can the two committees be appointed?

13. A group of ten people is to be divided into three committees of three, three, and six members, respectively. The chair of the group is to serve on all three committees and is the only member of the group who serves on more than one committee. In how many ways can the committee assignments be made?

    [Ans. 756.]

14. In a class of 20 students, grades of A, B, C, D, and F are to be assigned. Omit arithmetic details in answering the following.

    (a) In how many ways can this be done if there are no restrictions?

    [Ans. 520.]

(b) In how many ways can this be done if the grades are assigned as follows: 2 A's, 3 B's, 10 C's, 3 D's, and 2 F's?

(c) In how many ways can this be done if the following rules are to be satisfied: exactly 10 C's; the same number of A's as F's; the same number of B's as D's; always more B's than A's?

[Ans. $\binom{20}{5,10,5} + \binom{20}{1,4,10,4,1} + \binom{20}{2,3,10,3,2}$.]

15. Establish the identity

$$\binom{n}{r}\binom{r}{k} = \binom{n}{k}\binom{n-k}{r-k}$$

for $n \geq r \geq k$ in two ways, as follows:

(a) Replace each expression by a ratio of factorials and show that the two sides are equal.

(b) Consider the following problem: From a set of n people a committee of r is to be chosen, and from these r people a steering subcommittee of k people is to be selected. Show that the two sides of the identity give two different ways of counting the possibilities for this problem.

3.5 Some properties of the numbers $\binom{n}{j}$.

The numbers $\binom{n}{j}$ introduced in the previous section will play an important role in our future work. We give here some of the more important properties of these numbers.

A convenient way to obtain these numbers is given by the famous Pascal triangle, shown in Figure 3.5. To obtain the triangle we first write the 1's down the sides. Any of the other numbers in the triangle has the property that it is the sum of the two adjacent numbers in the row just above. Thus the next row in the triangle is 1, 6, 15, 20, 15, 6, 1. To find the number $\binom{n}{j}$ we look in the row corresponding to the number n and see where the diagonal line corresponding to the value of j intersects this row. For example, $\binom{4}{2} = 6$ is in the row marked $n = 4$ and on the diagonal marked $j = 2$.

The property of the numbers $\binom{n}{j}$ upon which the triangle is based is

$$\binom{n+1}{j} = \binom{n}{j-1} + \binom{n}{j}.$$

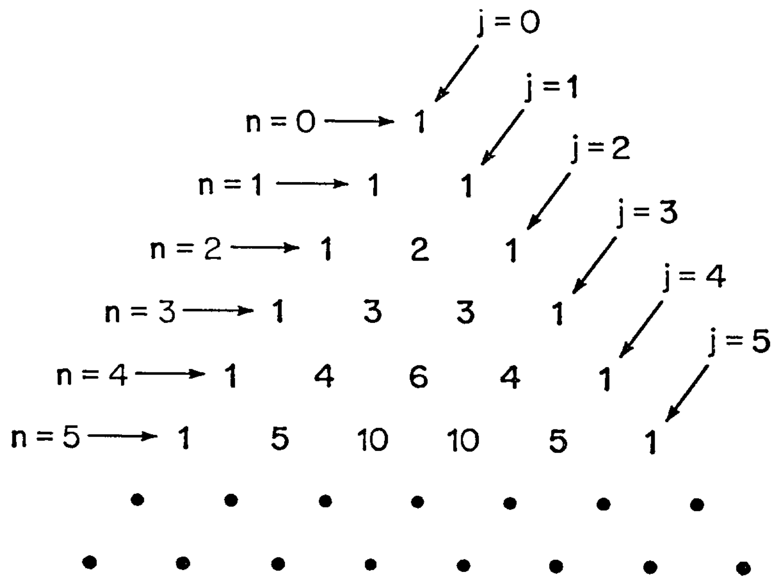


Figure 3.5: $\diamondsuit$

This fact can be verified directly (see Exercise 6), but the following argument is interesting in itself. The number $\binom{n+1}{j}$ is the number of subsets with j elements that can be formed from a set of $n+1$ elements. Select one of the $n+1$ elements, x. The $\binom{n+1}{j}$ subsets can be partitioned into those that contain x and those that do not. The latter are subsets of j elements formed from n objects, and hence there are $\binom{n}{j}$ such subsets. The former are constructed by adding x to a subset of $j-1$ elements formed from n elements, and hence there are $\binom{n}{j-1}$ of them. Thus

$$\binom{n+1}{j} = \binom{n}{j-1} + \binom{n}{j}.$$

If we look again at the Pascal triangle, we observe that the numbers in a given row increase for a while, and then decrease. We can prove this fact in general by considering the ratio of two successive terms,

$$\frac{\binom{n}{j+1}}{\binom{n}{j}} = \frac{n!}{(j+1)!(n-j-1)!} \cdot \frac{j!(n-j)!}{n!} = \frac{n-j}{j+1}.$$

The numbers increase as long as the ratio is greater than 1, i.e., $n-j > j+1$. This means that $j < \frac{1}{2}(n-1)$. We must distinguish the case of an even n from an odd n. For example, if $n = 10$, j must be less than $\frac{1}{2}(10-1) = 4.5$. Hence for j up to 4 the terms are increasing, from $j = 5$ on, the terms decrease. For $n = 11$, j must be less than $\frac{1}{2}(11-1) = 5$. For $j = 5$, $(11-j)/(j+1) = 1$. Hence, up to $j = 5$ the terms increase, then $\binom{11}{5} = \binom{11}{6}$, and then the terms decrease.

Exercises

1. Extend the Pascal triangle to $n = 16$. Save the result for later use.

2. Prove that

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \ldots + \binom{n}{n} = 2^n,$$

   using the fact that a set with n elements has 2^n subsets.

3. For a set of ten elements prove that there are more subsets with five elements than there are subsets with any other fixed number of elements.

4. Using the fact that $\binom{n}{r+1} = \frac{n-r}{r+1} \cdot \binom{n}{r}$, compute $\binom{30}{s}$ for $s = 1, 2, 3, 4$ from the fact that $\binom{30}{0} = 1$.

   [Ans. 30; 435; 4060; 27,405.]

5. There are $\binom{52}{13}$ different possible bridge hands. Assume that a list is made showing all these hands, and that in this list the first card in every hand is crossed out. This leaves us with a list of twelve-card hands. Prove that at least two hands in the latter list contain exactly the same cards.

6. Prove that

$$\binom{n+1}{j} = \binom{n}{j-1} + \binom{n}{j}$$

   using only the fact that

$$\binom{n}{j} = \frac{n!}{j!(n-j)!}.$$

7. Construct a triangle in the same way that the Pascal triangle was constructed, except that whenever you add two numbers, use parity addition (table (a) in Figure 2.12). Construct the triangle for 16 rows. What does this triangle tell you about the numbers in the Pascal triangle? Use this result to check your triangle in Exercise 1.

8. In the triangle obtained in Exercise 7, what property do the rows 1, 2, 4, 8, and 16 have in common? What does this say about the numbers in the corresponding rows of the Pascal triangle? What would you predict for the terms in the 32nd row of the Pascal triangle?

9. For the following table state how one row is obtained from the preceding row and give the relation of this table to the Pascal triangle.

1	1	1	1	1	1	1
1	2	3	4	5	6	7
1	3	6	10	15	21	28
1	4	10	20	35	56	84
1	5	15	35	70	126	210
1	6	21	56	126	252	462
1	7	28	84	210	462	924

10. Referring to the table in Exercise 9, number the columns starting with $0, 1, 2, \ldots$ and number the rows starting with $1, 2, 3, \ldots$. Let $f(n, r)$ be the element in the nth column and the rth row. The table was constructed by the rule

$$f(n, r) = f(n-1, r) + f(n, r-1)$$

for $n > 0$ and $r > 1$, and $f(n, 1) = f(0, r) = 1$ for all n and r. Verify that

$$f(n, r) = \binom{n+r-1}{n}$$

satisfies these conditions and is in fact the only choice for $f(n, r)$ which will satisfy the conditions.

11. Consider a set $\{a_1, a_2, a_3\}$ of three objects which cannot be distinguished from one another. Then the ordered partitions with two cells which could be distinguished are: $[\{a_1, a_2, a_3\}, \emptyset]$, $[\{a_1, a_2\}, \{a_3\}]$, $[\{a_1\}, \{a_2, a_3\}]$, $[\emptyset, \{a_1, a_2, a_3\}]$. List all such ordered partitions with three cells. How many are there?

[Ans. 10.]

12. Let $f(n, r)$ be the number of distinguishable ordered partitions with r cells which can be formed from a set of n indistinguishable objects. Show that $f(n, r)$ satisfies the conditions

$$f(n, r) = f(n-1, r) + f(n, r-1)$$

for $n > 0$ and $r > 1$, and $f(n, 1) = f(0, r) = 1$ for all n and r. [Hint: Show that $f(n, r-1)$ is the number of partitions which have the last cell empty and $f(n-1, r)$ is the number which have at least one element in the last cell.]

13. Using the results of Exercises 10 and 12, show that the number of distinguishable ordered partitions with r cells which can be formed from a set of n indistinguishable objects is

$$\binom{n+r-1}{n}.$$

14. Assume that a mail carrier has seven letters to put in three mail boxes. How many ways can this be done if the letters are not distinguished?

[Ans. 36.]

15. For $n \geq r \geq k \geq s$ show that the identity

$$\binom{n}{r}\binom{r}{k}\binom{k}{s} = \binom{n}{s}\binom{n-s}{k-s}\binom{n-k}{r-k}.$$

holds by replacing each binomial coefficient by a ratio of factorials.

16. Establish the identity in Exercise 15 in another way by showing that the two sides of the expression are simply two different ways of counting the number of solutions to the following problem: From a set of n people a subset of r is to be chosen; from the set of r people a subset of k is to be chosen; and from the set of k people a subset of s people is to be chosen.

17. Generalize the identity in Exercises 15 and 16 to solve the problem of finding the number of ways of selecting a t-element subset from an s-element subset from a k-element subset from an r-element subset of an n-element set, where $n \geq r \geq k \geq s \geq t$.

3.6 Binomial and multinomial theorems

It is sometimes necessary to expand products of the form $(x+y)^3$, $(x+2y+11z)^5$, etc. In this section we shall consider systematic ways of carrying out such expansions.

Consider first the special case $(x+y)^3$. We write this as

$$(x+y)^3 = (x+y)(x+y)(x+y).$$

To perform the multiplication, we choose either an x or a y from each of the three factors and multiply our choices together; we do this for all possible choices and add the results. We represent a particular set of choices by a two-cell partition of the numbers 1, 2, 3. In the first cell we put the numbers which correspond to factors from which we chose an x. In the second cell we put the numbers which correspond to factors from which we chose a y. For example, the partition $[\{1,3\},\{2\}]$ corresponds to a choice of x from the first and third factors and y from the second. The product so obtained is $xyx = x^2y$. The coefficient of x^2y in the expansion of $(x+y)^3$ will be the number of partitions which lead to a choice of two x's and one y, that is, the number of two-cell partitions of three elements with two elements in the first cell and one in the second, which is $\binom{3}{2} = 3$. More generally, the coefficient of the term of the form $x^j y^{3-j}$ will be $\binom{3}{j}$ for $j = 0, 1, 2, 3$. Thus we can write the desired expansion as

$$\begin{aligned}(x+y)^3 &= \binom{3}{3}x^3 + \binom{3}{2}x^2y + \binom{3}{1}xy^2 + \binom{3}{0}y^3 \\ &= x^3 + 3x^2y + 3xy^2 + y^3.\end{aligned}$$

The same argument carried out for the expansion $(x+y)^n$ leads to the binomial theorem of algebra.

Binomial theorem. The expansion of $(x+y)^n$ is given by

$$\binom{n}{n}x^n + \binom{n}{n-1}x^{n-1}y + \binom{n}{n-2}x^{n-2}y^2 + \ldots + \binom{n}{1}xy^{n-1} + \binom{n}{0}y^n.$$

Example 3.16 Let us find the expansion for $(a-2b)^3$. To fit this into the binomial theorem, we think of x as being a and y as being $-2b$. Then we have

$$\begin{aligned}(a-2b)^3 &= a^3 + 3a^2(-2b) + 3a(-2b)^2 + (-2b)^3 \\ &= a^3 - 6a^2b + 12ab^2 - 8b^3.\end{aligned}$$

$\diamondsuit$

We turn now to the problem of expanding the trinomial $(x+y+z)^3$. Again we write

$$(x+y+z)^3 = (x+y+z)(x+y+z)(x+y+z).$$

This time we choose either an x or y or z from each of the three factors. Our choice is now represented by a three-cell partition of the set of numbers $\{1,2,3\}$. The first cell has the numbers corresponding to factors from which we choose an x, the second cell the numbers corresponding to factors from which we choose a y, and the third those from which we choose a z. For example, the partition $[\{1,3\},\emptyset,\{2\}]$ corresponds to a choice of x from the first and third factors, no y's, and a z from the second factor. The term obtained is $xzx = x^2z$. The coefficient of the term x^2z in the expansion is thus the number of three-cell partitions with two elements in the first cell, none in the second, and one in the third. There are $\binom{3}{2,0,1} = 3$ such partitions. In general, the coefficient of the term of the form $x^ay^bz^c$ in the expansion of $(x+y+z)^3$ will be

$$\binom{3}{a,b,c} = \frac{3!}{a!b!c!}.$$

Finding this way the coefficient for each possible a, b, and c, we obtain

$$(x+y+z)^3 = x^3+y^3+x^3+3x^2y+3xy^2+3yz^2+3y^2z+3xz^2+3x^2z+6xyz.$$

The same method can be carried out in general for finding the expansion of $(x_1+x_2+\ldots+x_r)^n$. From each factor we choose either an x_1, or x_2, $\ldots$, or x_r, form the product and add these products for all possible choices. We will have r^n products, but many will be equal. A particular choice of one term from each factor determines an r-cell partition of the numbers from 1 to n. In the first cell we put the numbers of the factors from which we choose an x_1, in the second cell those from which we choose x_2, etc. A particular choice gives us a term of the form $x_1^{n_1}x_2^{n_2}\ldots x_r^{n_r}$ with $n_1+n_2+\ldots+n_r = n$. The corresponding partition has n_1 elements in the first cell, n_2 in the second, etc. For each such partition we obtain one such term. Hence the number of these terms which we obtain is the number of such partitions, which is

$$\binom{n}{n_1,n_2,\ldots,n_r} = \frac{n!}{n_1!n_2!\ldots n_r!}.$$

Thus we have the multinomial theorem.

Multinomial theorem. The expansion of $(x_1 + x_2 + \ldots + x_r)^n$ is found by adding all terms of the form

$$\binom{n}{n_1, n_2, \ldots, n_r} x_1^{n_1} x_2^{n_2} \ldots x_r^{n_r}$$

where $n_1 + n_2 + \ldots + n_r = n$.

Exercises

1. Expand by the binomial theorem

   (a) $(x + y)^4$.

   (b) $(1 + x)^5$.

   (c) $(x - y)^3$.

   (d) $(2x + a)^4$.

   (e) $(2x - 3y)^3$.

   (f) $(100 - 1)^5$.

2. Expand

   (a) $(x + y + x)^4$.

   (b) $(2x + y - z)^3$.

   (c) $(2 + 2 + 1)^3$. (Evaluate two ways.)

3. (a) Find the coefficient of the term $x^2y^3z^2$ in the expansion of $(x + y + z)^7$.

   [Ans. 210.]

   (b) Find the coefficient of the term $x^6y^3z^2$ in the expression $(x - 2y + 5z)^{11}$

   [Ans. -924,000.]

4. Using the binomial theorem prove that

   (a)

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \ldots + \binom{n}{n} = 2^n.$$

(b)

$$\binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \binom{n}{3} + \ldots \pm \binom{n}{n} = 0$$

for $n > 0$.

5. Using an argument similar to the one in Section 3.6, prove that

$$\binom{n+1}{i,j,k} = \binom{n}{i-1,j,k} + \binom{n}{i,j-1,k} + \binom{n}{i,j,k-1}.$$

6. Let $f(n,r)$ be the number of terms in the multinomial expansion of

$$(x_1 + x_2 + \ldots + x_r)^n$$

and show that

$$f(n,r) = \binom{n+r-1}{n}.$$

[Hint: Show that the conditions of Exercise 10 are satisfied by showing that $f(n,r-1)$ is the number of terms which do not have x_r and $f(n-1,r)$ is the number which do.]

7. How many terms are there in each of the expansions:

(a) $(x+y+z)^6$?

[Ans. 28.]

(b) $(a+2b+5c+d)^4$?

[Ans. 35.]

(c) $(r+s+t+u+v)^6$?

[Ans. 210.]

8. Prove that k^n is the sum of the numbers $\binom{n}{r_1,r_2,\ldots,r_k}$ for all choices of $r_1, r_2, \ldots, r_k$ such that

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

Supplementary exercises.

9. Show that the problem given in Exercise 15b can also be solved by a multinomial coefficient, and hence show that

$$\binom{n}{n-r, r-k, k} = \binom{n}{r}\binom{r}{k} = \binom{n}{k}\binom{n-k}{r-k}.$$

10. Show that the problem given in Exercise 16 can also be solved by a multinomial coefficient, and hence show that

$$\binom{n}{n-r, r-k, k-s, s} = \binom{n}{r}\binom{r}{k}\binom{k}{s} = \binom{n}{s}\binom{n-s}{k-s}\binom{n-k}{r-k}.$$

11. If $a+b+c=n$, show that

$$\binom{n}{a, b, c} = \binom{n}{a}\binom{n-a}{b}.$$

12. If $a+b+c+d=n$, show that

$$\binom{n}{a, b, c, d} = \binom{n}{a}\binom{n-a}{b}\binom{n-a-b}{c}.$$

13. If $n_1+n_2+\ldots+n_r=n$, guess a formula that relates the multinomial coefficient to a product of binomial coefficients. [Hint: Use the formulas in Exercises 11 and 12 to guide you.]

14. Use Exercises 11, 12, and 13 to show that the multinomial coefficients can always be obtained by taking products of suitable numbers in the first n rows of the Pascal triangle.

3.7 Voting power

We return to the problem raised in Section 2.6. Now we are interested not only in coalitions, but also in the power of individual members. We will develop a numerical measure of voting power that was suggested by L. S. Shapley and M. Shubik. While the measure will be explained in detail below, for the reasons for choosing this particular measure the reader is referred to the original paper.

First of all we must realize that the number of votes a member controls is not in itself a good measure of his or her power. If x has three votes and y has one vote, it does not necessarily follow that x has three times the power that y has. Thus if the committee has just three members $\{x, y, z\}$ and z also has only one vote, then x is a dictator and y is powerless.

The basic idea of the power index is found in considering various alignments of the committee members on a number of issues. The n members are ordered $x_1, x_2, \ldots, x_n$ according to how likely they are to vote for the measure. If the measure is to carry, we must persuade x_1 and x_2 up to x_i to vote for it until we have a winning coalition. If $\{x_1, x_2, \ldots, x_i\}$ is a winning coalition but $\{x_1, x_2, \ldots, x_{i-1}\}$ is not winning, then x_i is the crucial member of the coalition. We must persuade him or her to vote for the measure, and he or she is the one hardest to persuade of the i necessary members. We call x_i the pivot.

For a purely mathematical measure of the power of a member we do not consider the views of the members. Rather we consider all possible ways that the members could be aligned on an issue, and see how often a given member would be the pivot. That means considering all permutations, and there will be $n!$ of them. In each permutation one member will be the pivot. The frequency with which a member is the pivot of an alignment is a good measure of his or her voting power.

Definition. The *voting power* of a member of a committee is the number of alignments in which he or she is pivotal divided by the total number of alignments. (The total number of alignments, of course, is $n!$ for a committee of n members.)

Example 3.17 If all n members have one vote each, and it takes a majority vote to carry a measure, it is easy to see (by symmetry) that each member is pivot in $1/n$ of the alignments. Hence each member has power equal to $1/n$. Let us illustrate this for $n = 3$. There are $3! = 6$ alignments. It takes two votes to carry a measure; hence the second member is always the pivot. The alignments are: 1**2**3, 1**3**2, 2**1**3, 2**3**1, 3**1**2, 3**2**1. The pivots are emphasized. Each member is pivot twice, hence has power $\frac{2}{6} = \frac{1}{3}$. ♢

Example 3.18 Reconsider Example 2.9 of Section 2.6 from this point of view. There are 24 permutations of the four members. We will list them, with the pivot emphasized:

wxy**z**	wx**z**y	wyx**z**	wy**z**x	w**z**xy	w**z**yx
xwy**z**	xw**z**y	xyw**z**	xy**z**w	xz**w**y	xz**y**w
yxw**z**	yx**z**w	ywx**z**	yw**z**x	yz**x**w	yz**w**x
zx**y**w	zx**w**y	zy**x**w	zy**w**x	z**w**xy	z**w**yx

We see that z has power of $\frac{14}{24}$, w has $\frac{6}{24}$, x and y have $\frac{2}{24}$ each. (Or, simplified, they have $\frac{7}{12}$, $\frac{3}{12}$, $\frac{1}{12}$, $\frac{1}{12}$ power, respectively.) We note that these ratios are much further apart than the ratio of votes which is $3:2:1:1$. Here three votes are worth seven times as much as the single vote and more than twice as much as two votes. $\diamondsuit$

Example 3.19 Reconsider Example 2.10 of Section 2.6. By an analysis similar to the ones used so far it can be shown that in the Security Council of the United Nations before 1966, each of the Big Five had $\frac{76}{385}$ or approximately .197 power, while each of the small nations had approximately .002 power. (See Exercise 12.) This reproduces our intuitive feeling that, while the small nations in the Security Council are not powerless, nearly all the power is in the hands of the Big Five.

The voting powers according to the 1966 revision will be worked out in Exercise 13. $\diamondsuit$

Example 3.20 In a committee of five each member has one vote, but the chair has veto power. Hence the minimal winning coalitions are three-member coalitions including the chair. There are $5! = 120$ permutations. The pivot cannot come before the chair, since without the chair we do not have a winning coalition. Hence, when the chair is in place number 3, 4, or 5, he or she is the pivot. This happens in $\frac{3}{5}$ of the permutations. When he or she is in position 1 or 2, then the number 3 member is pivot. The number of permutations in which the chair is in one of the first two posltions and a given member is third is $2 \cdot 3! = 12$. Hence the chair has power $\frac{3}{5}$, and each of the others has power $\frac{1}{10}$. $\diamondsuit$

Exercises

1. A committee of three makes decisions by majority vote. Write out all permutations, and calculate the voting powers if the three members have

   (a) One vote each.

   [Ans. $\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$.]

   (b) One vote for two of them, two votes for the third.

   [Ans. $\frac{1}{6}, \frac{1}{6}, \frac{2}{3}$.]

   (c) One vote for two of them, three votes for the third.

[Ans. $0, 0, 1$.]

(d) One, two, and three votes, respectively.

[Ans. $\frac{1}{6}, \frac{1}{6}, \frac{2}{3}$.]

(e) Two votes each for two of them, and three votes for the third.

[Ans. $\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$.]

2. Prove that in any decision-making body the sum of the powers of the members is 1.

3. What is the power of a dictator? What is the power of a "powerless" member? Prove that your answers are correct.

4. A large company issued 100,000 shares. These are held by three stockholders, who have 50,000, 49,999, and one share, respectively. Calculate the powers of the three members.

[Ans. $\frac{2}{3}, \frac{1}{6}, \frac{1}{6}$.]

5. A committee consists of 100 members having one vote each, plus a chairman who can break ties. Calculate the power distribution. (Do not try to write out all permutations!)

6. In Exercise 5, give the chairman a veto instead of the power to break ties. How does this change the power distribution?

[Ans. The chairman has power $\frac{50}{101}$.]

7. How are the powers in Exercise 1 changed if the committee requires a $\frac{3}{4}$ vote to carry a measure?

8. If in a committee of five, requiring majority decisions, each member has one vote, then each has power $\frac{1}{5}$. Now let us suppose that two members team up, and always vote the same way. Does this increase their power? (The best way to represent this situation is by allowing only those permutations in which these two members are next to each other.)

[Ans. Yes, the pair's power increases from .4 to .5.]

9. If the minimal winning coalitions are known, show that the power of each member can be determined without knowing anything about the number of votes that each member controls.

10. Answer the following questions for a three-man committee.

    (a) Find all possible sets of minimal winning coalitions.

    (b) For each set of minimal winning coalitions find the distribution of voting power.

    (c) Verify that the various distributions of power found in Exercises 1 and 7 are the only ones possible.

11. In Exercise 1 parts 1a and 1e have the same answer, and parts 1b and 1d and Exercise 4 also have the same answer. Use the results of Exercise 9 to find a reason for these coincidences.

12. Compute the voting power of one of the Big Five in the Security Council of the United Nations as follows:

    (a) Show that for the nation to be pivotal it must be in the number 7 spot or later.

    (b) Show that there are $\binom{6}{2}6!4!$ permutations in which the nation is pivotal in the number 7 spot.

    (c) Find similar formulas for the number of permutations in which it is pivotal in the number 8, 9, 10, or 11 spot.

    (d) Use this information to find the total number of permutations in which it is pivotal, and from this compute the power of the nation.

13. Apply the method of Exercise 12 to the revised voting scheme in the Security Council (10 small-nation members, and 9 votes required to carry a measure). What is the power of a large nation? Has the power of one of the small nations increased or decreased?

    [Ans. $\frac{421}{2145}$ (nearly the same as before); decreased.]

3.8 Techniques for counting

We know that there is no single method or formula for solving all counting problems. There are, however, some useful techniques that can be learned. In this section we shall discuss two problems that illustrate important techniques.

The first problem illustrates the importance of looking for a general pattern in the examination of special cases. We have seen in Section 3.2 and Exercise 6 of that section, that the following formulas hold for the number of elements in the union of two and three sets, respectively.

$$n(A_1 \cup A_2) = n(A_1) + n(A_2) - n(A_1 \cap A_2),$$

$$\begin{aligned} n(A_1 \cup A_2 \cup A_3) \quad = \quad & n(A_1) + n(A_2) + n(A_3) \\ & -n(A_1 \cap A_2) - n(A_1 \cap A_3) - n(A_2 \cap A_3) \\ & +n(A_1 \cap A_2 \cap A_3). \end{aligned}$$

On the basis of these formulas we might conjecture that the number of elements in the union of any finite number of sets could be obtained by adding the numbers in each of the sets, then subtracting the numbers in each possible intersection of two sets, then adding the numbers in each possible intersection of three sets, etc. If this is correct, the formula for the intersection of four sets should be

$$\begin{aligned} n(A_1 \cup A_2 \cup A_3 \cup A_4) \quad = \quad & n(A_1) + n(A_2) + n(A_3) + n(A_4) \qquad (3.1) \\ & - n(A_1 \cap A_2) - n(A_1 \cap A_3) - n(A_1 \cap A_4) \\ & - n(A_2 \cap A_3) - n(A_2 \cap A_4) - n(A_3 \cap A_4) \\ & + n(A_1 \cap A_2 \cap A_3) + n(A_1 \cap A_2 \cap A_4) \\ & + n(A_1 \cap A_3 \cap A_4) + n(A_2 \cap A_3 \cap A_4) \\ & - n(A_1 \cap A_2 \cap A_3 \cap A_4) \end{aligned}$$

Let us try to establish this formula. We must show that if u is an element of at least one of the four sets, then it is counted exactly once on the right-hand side of 3.1. We consider separately the cases where u is in exactly 1 of the sets, exactly 2 of the sets, etc.

For instance, if u is in exactly two of the sets it will be counted twice in the terms of the right-hand side of 3.1 that involve single sets, once in the terms that involve the intersection of two sets, and not at all in the terms that involve the intersections of three or four sets. Again, if u is in exactly three of the sets it will be counted three times in the terms involving single sets, twice in the terms involving intersections of two sets, once in the terms involving the intersections of three sets, and not at all in the last term involving the intersection of all four sets. Considering each possibility we have the following table.

Number of sets that contain u	Number of times it is counted
1	1
2	$2-1$
3	$3-3+1$
4	$4-6+4-1$

We see from this that, in every case, u is counted exactly once on the right-hand side of 3.1. Furthermore, if we look closely, we detect a pattern in the numbers in the righthand column of the above table. If we put a -1 in front of these numbers we have

1	$-1+1$
2	$-1+2-1$
3	$-1+3-3+1$
4	$-1+4-6+4-1$

We now recognize that these numbers are simply the numbers in the first four rows of the Pascal triangle, but with alternating $+$ and $-$ signs. Since we put a -1 in each row of the table, we want to show that the sum of each row is 0. If that is true, it should be a general property of the Pascal triangle. That is, if we put alternating signs in the jth row of the Pascal triangle, we should get a sum of 0. But this is indeed the case, since, by the binomial theorem, for $j > 0$,

$$\begin{aligned}
0 &= \pm(1-1)^j \\
&= 1 - \binom{j}{1} + \binom{j}{2} - \binom{j}{3} + \ldots \pm 1 \\
&= -1 + \binom{j}{1} - \binom{j}{2} + \binom{j}{3} - \ldots \mp 1.
\end{aligned}$$

Thus we have not only seen why the formula works for the case of four sets, but we have also found the method for proving the formula for the general case. That is, suppose we wish to establish that the number of elements in the union of n sets may be obtained as an alternating sum by adding the numbers of elements in each of the sets, subtracting the numbers of elements in each pairwise intersection of the sets, adding the numbers of elements in each intersection of three sets, etc. Consider an element u that is in exactly j of the sets. Let us see how many times u will be counted in the alternating sum. If it is in j of the sets, it will first be counted j times in the sum of the elements in the sets by themselves. For u to be in the intersection of two sets, we must choose two of the j sets to which it belongs. This can be done in $\binom{j}{2}$

different ways. Hence an amount $\binom{j}{2}$ will be subtracted from the sum. To be in the intersection of three sets, we must choose three of the j sets containing u. This can be done in $\binom{j}{3}$ different ways. Thus, an amount $\binom{j}{3}$ will be added to the sum, etc. Hence the total number of times u will be counted by the alternating sum is

$$\binom{j}{1} - \binom{j}{2} + \binom{j}{3} - \ldots \pm 1$$

since we have just seen that, if we add -1 to the sum, we obtain 0. Hence the sum itself must always be 1. That is, no matter how many sets u is in, it will be counted exactly once by the alternating sum, and this is true for every element u in the union. We have thus established the general formula

$$\begin{aligned} n(A_1 \cup A_2 \cup \ldots \cup A_n) \quad = \quad & n(A_1) + n(A_2) + \ldots + n(A_n) \\ & -n(A_1 \cap A_2) - n(A_1 \cap A_3) - \ldots \\ & +n(A_1 \cap A_2 \cap A_3) + n(A_1 \cap A_2 \cap A_4) + \ldots \\ & - \ldots \\ & \pm n(A_1 \cap A_2 \cap \ldots \cap A_n). \end{aligned} \tag{3.2}$$

This formula is called the *inclusion-exclusion formula* for the number of elements in the union of n sets. It can be extended to formulas for counting the number of elements that occur in two of the sets, three of the sets, etc. See Exercises 21, 25, and 27.

Example 3.21 In a high school the following language enrollments are recorded for the senior class.

English	150
French	75
German	35
Spanish	50

Also, the following overlaps are noted.

Taking English and French	70
Taking English and German	30
Taking English and Spanish	40
Taking French and German	5
Taking English, French and German	2

If every student takes at least one language, how many seniors are there?

Let E, F, G, and S be the sets of students taking English, French, German, and Spanish, respectively. Using formula 3.1 and ignoring empty sets, we have

$$\begin{aligned} n(E \cup F \cup G \cup S) &= n(E) + n(F) + n(G) + n(S) \\ &\quad - n(E \cap F) - n(E \cap G) - n(E \cap S) - n(F \cap G) \\ &\quad + n(E \cap F \cap G) \\ &= 150 + 75 + 35 + 50 - 70 - 30 - 40 - 5 + 2 \\ &= 167. \end{aligned}$$

Since every student takes at least one language, the total number of students is 167. ♢

Example 3.22 The four words

TABLE, BASIN, CLASP, BLUSH

have the following interesting properties. Each word consists of five different letters. Any two words have exactly two letters in common. Any three words have one letter in common. No letter occurs in all four words. How many different letters are there?

Letting the words be sets of letters, there are $\binom{4}{1}$ ways of taking these sets one at a time, $\binom{4}{2}$ ways of taking them two at a time, etc. Hence formula 3.2 gives

$$\binom{4}{1} \cdot 5 - \binom{4}{2} \cdot 2 + \binom{4}{3} \cdot 1 - \binom{4}{4} \cdot 0 = 12$$

as the number of distinct letters. The reader should verify this answer by direct count. ♢

It often happens that a counting problem can be formulated in a number of different ways that sound quite different but that are in fact equivalent. And in one of these ways the answer may suggest itself readily. To illustrate how a reformulation can make a hard sounding problem easy, we give an alternate method for solving the problem considered in Exercise 13.

The problem is to count the number of ways that n indistinguishable objects can be put into r cells. For instance, if there are three objects and three cells, the number of different ways can be enumerated as follows (Using $\circ$ for object and bars to indicate the sides of the cells:

```
| ○ ○ ○ |       |       |
| ○○    | ○     |       |
| ○○    |       | ○     |
| ○     | ○○    |       |
| ○     | ○     | ○     |
|       | ○ ○ ○ |       |
|       | ○○    | ○     |
|       | ○     | ○○    |
|       |       | ○ ○ ○ |
```

We see that in this case there are ten ways the task can be accomplished. But the answer for the general case is not clear.

If we look at the problem in a slightly different manner, the answer suggests itself. Instead of putting the objects in the cells, we imagine putting the cells around the objects. In the above case we see that three cells are constructed from four bars. Two of these bars must be placed at the ends. The two other bars together with the three objects we regard as occupying five intermediate positions. Of these five intermediate positions we must choose two of them for bars and three for the objects. Hence the total number of ways we can accomplish the task is $\binom{5}{2} = \binom{5}{3} = 10$, which is the answer we got by counting all the ways.

For the general case we can argue in the same manner. We have r cells and n objects. We need $r+1$ bars to form the r cells, but two of these must be fixed on the ends. The remaining $r-1$ bars together with the n objects occupy $r-1+n$ intermediate positions. And we must choose $r-1$ of these for the bars and the remaining n for the objects. Hence our task can be accomplished in

$$\binom{n+r-1}{r-1} = \binom{n+r-1}{n}$$

different ways.

Example 3.23 Seven people enter an elevator that will stop at five floors. In how many different ways can the people leave the elevator if we are interested only in the number that depart at each floor, and do not distinguish among the people? According to our general formula, the answer is

$$\binom{7+5-1}{7} = \binom{11}{7} = 330.$$

Suppose we are interested in finding the number of such possibilities in which at least one person gets off at each floor. We can then arbitrarily assign one person to get off at each floor, and the remaining two can get off at any floor. They can get off the elevator in

$$\binom{2+5-1}{2} = \binom{6}{2} = 15$$

different ways. ♢

Exercises

1. The survey discussed in Exercise 8 has been enlarged to include a fourth magazine D. It was found that no one who reads either magazine A or magazine B reads magazine D. However, 10 per cent of the people read magazine D and 5 per cent read both C and D. What per cent of the people do not read any magazine?

   [Ans. 5 per cent.]

2. A certain college administers three qualifying tests. They announce the following results: "Of the students taking the tests, 2 per cent failed all three tests, 6 per cent failed tests A and B, 5 per cent failed A and C, 8 per cent failed B and C, 29 per cent failed test A, 32 per cent failed B, and 16 per cent failed C." How many students passed all three qualifying tests?

3. Four partners in a game require a score of exactly 20 points to win. In how many ways can they accomplish this?

   [Ans. $\binom{23}{3}$.]

4. In how many ways can eight apples be distributed among four boys? In how many ways can this be done if each boy is to get at least one apple?

5. Suppose we have n balls and r boxes with $n \geq r$. Show that the number of different ways that the balls can be put into the boxes which insures that there is at least one ball in every box is $\binom{n-1}{r-1}$.

6. Identical prizes are to be distributed among five boys. It is observed that there are 15 ways that this can be done if each boy is to get at least one prize. How many prizes are there?

[Ans. 7.]

7. Let $p_1, p_2, \ldots, p_n$ be n statements relative to a possibility space $\mathcal{U}$. Show that the inclusion-exclusion formula gives a formula for the number of elements in the truth set of the disjunction $p_1 \vee p_2 \vee \ldots \vee p_n$ in terms of the numbers of elements in the truth sets of conjunctions formed from subsets of these statements.

8. A boss asks his or her secretary to put letters written to seven different persons into addressed envelopes. Find the number of ways that this can be done so that at least one person gets his or her own letter. [Hint: Use the result of Exercise 7, letting p_i be the statement "The ith person gets his or her own letter".]

[Ans. 3186.]

9. Consider the numbers from 2 to 10 inclusive. Let A_2 be the set of numbers divisible by 2 and A_3 the set of numbers divisible by 3. Find $n(A_2 \cup A_3)$ by using the inclusion-exclusion formula. From this result find the number of prime numbers between 2 and 10 (where a prime number is a number divisible only by itself and by 1). [Hint: Be sure to count the numbers 2 and 3 among the primes.]

10. Use the method of Exercise 9 to find the number of prime numbers between 2 and 100 inclusive. [Hint: Consider first the sets A_2, A_3, A_5, and A_7.]

[Ans. 25.]

11. Verify that the following formula gives the number of elements in the intersection of three sets.

$$\begin{aligned} n(A_1 \cap A_2 \cap A_3) \quad = \quad & n(A_1) + n(A_2) + n(A_3) \\ & - n(A_1 \cup A_2) - n(A_1 \cup A_3) - n(A_2 \cup A_3) \\ & + n(A_1 \cup A_2 \cup A_3). \end{aligned}$$

12. Show that if we replace $\cap$ by $\cup$ and $\cup$ by $\cap$ in formula 3.2, we get a valid formula for the number of elements in the intersection of n sets. [Hint: Apply the inclusion-exclusion formula to the left-hand side of
$$n(\tilde{A}_1 \cup \tilde{A}_2 \cup \ldots \cup \tilde{A}_n) = n(\mathcal{U}) - n(A_1 \cap A_2 \cap \ldots \cap A_n.]$$

13. For $n \le m$ prove that
$$\binom{m}{0}\binom{n}{0} + \binom{m}{1}\binom{n}{1} + \binom{m}{2}\binom{n}{2} + \ldots + \binom{m}{n}\binom{n}{n} = \binom{m+n}{n}$$
by carrying out the following two steps:
    (a) Show that the left-hand side counts the number of ways of choosing equal numbers of men and women from sets of m men and n women.
    (b) Show that the right-hand side also counts the same number by showing that we can select equal numbers of men and women by selecting any subset of n persons from the whole set, and then combining the men selected with the women not selected.

14. By an ordered partition of n with r elements we mean a sequence of r nonnegative integers, possibly some 0, written in a definite order, and having n as their sum. For instance, $[1, 0, 3]$ and $[3, 0, 1]$ are two different ordered partitions of 4 with three elements. Show that the number of ordered partitions of n with r elements is $\binom{n+r-1}{n}$.

15. Show that the number of different possibilities for the outcomes of rolling n dice is $\binom{n+5}{n}$.

    Note. The next few exercises illustrate an important counting technique called the *reflection principle*. In Figure 3.6 we show a path from the point $(0, 2)$ to the point $(7, 1)$. We shall be interested in counting the number of paths of this type where at each step the path moves one unit to the right, and either one unit up or one unit down. We shall see that this model is useful for analyzing voting outcomes.

16. Show that the number of different paths leading from the point $(0, 2)$ to $(7, 1)$ is $\binom{7}{3}$. [Hint: Seven decisions must be made, of which three moves are up and the rest down.]

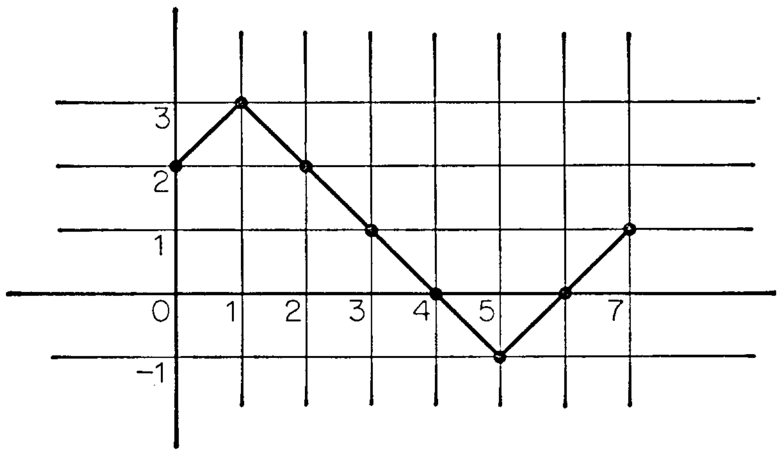


Figure 3.6: ♢

17. Show that the number of different paths from $(0, 2)$ to $(7, 1)$ which touch the x-axis at least once is the same as the total number of paths from the point $(0, -2)$ to the point $(7, 1)$. [Hint: Show that for every path to be counted from $(0, 2)$ that touches the x-axis, there corresponds a path from $(0, -2)$ to $(7,\ 1)$ obtained by reflecting the part of the path to the first touching point through the x-axis. A specific example is shown in Figure 3.7.]

18. Use the results of Exercises 16 and 17 to find the number of paths from $(0, 2)$ to $(7, 1)$ that never touch the x-axis.

[Ans. 14.]

19. Nine votes are cast in a race between two candidates A and B. Candidate A wins by one vote. Find the number of ways the ballots can be counted so that candidate A is leading throughout the entire count. [Hint: The first vote counted must be for A. Counting the remaining eight votes corresponds to a path from $(1, 1)$ to $(9, 1)$. We want the number of paths that never touch the x-axis.]

[Ans. 14.]

20. Let the symbol $n_r^{(k)}$ stand for "the number of elements that are in k or more of the r sets $A_1, A_2, \ldots, A_r$". Show that $n_3^{(1)} = n(A_1 \cup A_2 \cup A_3)$.

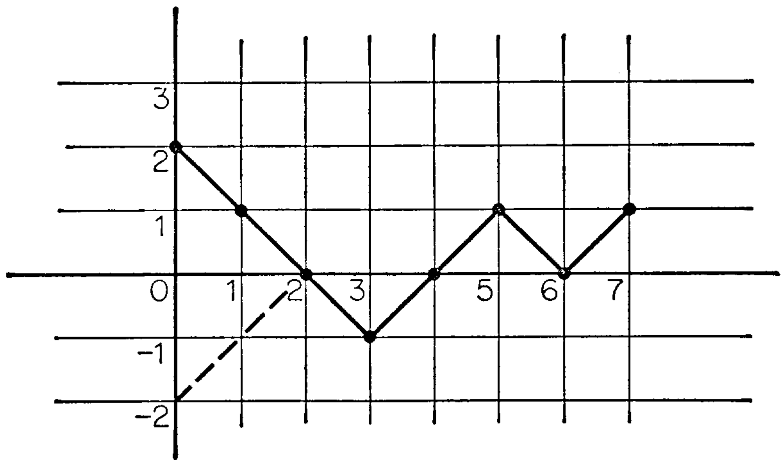


Figure 3.7: ♢

21. Show that

$$\begin{aligned} n_3^{(2)} &= n((A_1 \cap A_2) \cup (A_1 \cap A_3) \cup (A_2 \cap A_3)) \\ &= n(A_1 \cap A_2) + n(A_1 \cap A_3) + n(A_2 \cap A_3) - 2n(A_1 \cap A_2 \cap A_3) \end{aligned}$$

by using the inclusion-exclusion formula. Also develop an independent argument for the last formula.

22. Use Exercise 21 to find the number of letters that appear two or more times in the three words TABLE, BASIN, and CLASP.

23. Give an interpretation for $n_3^{(1)} - n_3^{(2)}$.

24. Use Exercise 23 to find the number of letters that occur exactly once in the three words of Exercise 22.

25. Develop a general argument like that in Exercise 21 to show that

$$\begin{aligned} n_4^{(2)} = \ & n(A_1 \cap A_2) + n(A_1 \cap A_3) + n(A_1 \cap A_4) \\ & + n(A_2 \cap A_3) + n(A_2 \cap A_4) + n(A_3 \cap A_4) \\ & - 2[n(A_1 \cap A_2 \cap A_3) + n(A_1 \cap A_2 \cap A_4) \\ & \quad + n(A_1 \cap A_3 \cap A_4) + n(A_2 \cap A_3 \cap A_4)] \\ & + 3n(A_1 \cap A_2 \cap A_3 \cap A_4) \end{aligned}$$

26. Use Exercise 25 to find the number of letters used two or more times in the four words of Example 3.22.

27. From the formulas in Exercises 21 and 25 guess the general formula for $n_r^{(2)}$ and develop a general argument to establish its correctness.

Suggested reading.

Shapley, L. S., and M. Shubik, "A Method for Evaluating the Distribution of Power in a Committee System", The American Political Science Review 48 (1954), pp. 787–792.

Whitworth, W. A., *Choice and Chance, with 1000 Exercises*, 1934.

Goldberg, S., *Probability: An Introduction*, 1960.

Parzen, E., *Modern Probability Theory and Its Applications*, 1960.

Chapter 4

Probability theory

4.1 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: (l) 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 $\emptyset$ is 0.

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 4.1 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. The truth set of the statement "The number which 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. ♢

Example 4.2 A gambler attends a race involving three horses A, B, and C. He or she 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 $\{\text{A}, \text{B}, \text{C}\}$. 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 l, 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 $\{\mathrm{A}, \mathrm{C}\}$. 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. Assume that there are n possibilities for the outcome of a given experiment. How should the weights be assigned if it is desired that all outcomes be assigned the same weight?

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. In an election Jones has probability $\frac{1}{2}$ of winning, Smith has probability $\frac{1}{3}$, and Black has probability $\frac{1}{6}$.

   (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.

4. Give the possibility set $\mathcal{U}$ for each of the following experiments.

   (a) An election between candidates A and B is to take place.

   (b) A number from 1 to 5 is chosen at random.

   (c) A two-headed coin is thrown.

   (d) A student is asked for the day of the year on which his or her birthday falls.

5. For which of the cases in Exercise 4 might it be appropriate to assign the same weight to each outcome?

6. Suppose that the following probabilities have been assigned to the possible results of putting a penny in a certain defective peanut-vending machine: The probability that nothing comes out is $\frac{1}{2}$. The probability that either you get your money back or you get peanuts (but not both) is $\frac{1}{3}$.

(a) What is the probability that you get your money back and also get peanuts?

[Ans. $\frac{1}{6}$.]

(b) From the information given, is it possible to find the probability that you get peanuts?

[Ans. No.]

7. 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 6 is three times as probable as a 2.) What is the probability of getting an even number in one throw?

[Ans. $\frac{4}{7}$.]

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.

(a) r: The number of heads that occur is greater than the number of tails.

[Ans. $\frac{1}{2}$.]

(b) s: Exactly two heads occur.

[Ans. $\frac{5}{8}$.]

(c) 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) $\Pr[r \vee \neg r] = \Pr[r] + \Pr[\neg r]$

(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

    [Ans. consistent.]

    (b) s, t

    [Ans. inconsistent.]

    (c) $r, \neg r$

    [Ans. inconsistent.]

    (d) r, t

    [Ans. consistent.]

11. State a theorem suggested by Exercises 9 and 10.

12. An experiment has three possible outcomes, a, b, and c. Let p be the statement "the outcome is a or b", and q be the statement "the outcome is b or c". Assume that weights have been assigned to the three outcomes so that $\Pr[p] = \frac{2}{3}$ and $\Pr[q] = \frac{5}{6}$. Find the weights.

    [Ans. $\frac{1}{6}, \frac{1}{2}, \frac{1}{3}$.]

13. Repeat Exercise 12 if $\Pr[p] = \frac{1}{2}$ and $\Pr[q] = \frac{3}{8}$.

4.2 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 $X = \emptyset$.

(B) $0 < m(X) < 1$ 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 $m(X) + m(Y) > m(X \cap Y)$. 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) $0 < \Pr[p] < 1$ for any statement p.

(c) The equality

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

holds 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) $\Pr[\neg p] = 1 - \Pr[p]$.

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 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 the 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 4.10). 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 and 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 4.3 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 $\Pr[p \wedge q] = \frac{1}{4}$, $\Pr[\neg p] = \frac{1}{3}$, and $\Pr[q] = \frac{1}{2}$. What is $\Pr[p \vee q]$?

   [Ans. $\frac{11}{12}$.]

2. Using the result of Exercise 1, find $\Pr[\neg p \wedge \neg q]$.

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 or her grades in English and Art. The student estimates that the probability of passing English is .4, the probability of passing at least one course with probability .6, but that the probability of 1 of passing both courses is only .1? What is the probability that the student 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 the student will get a C or D?

   [Ans. $\frac{1}{2}$.]

7. What odds should a person give on a bet that a six will turn up when a die is thrown?

8. Referring to Example 4.2, what odds should the man be willing to give for a bet that either A or B will come in first?

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

10. Using the result of Exercise 9 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.

11. A gambler is willing to give $5 : 4$ odds that the Dodgers will win the World Series. What must the probability of a Dodger victory be for this to be a fair bet?

    [Ans. $\frac{5}{9}$.]

12. A statistician has found through long experience that if he or she washes the car, it rains the next day 85 per cent of the time. What odds should the statistician give that this will occur next time?

13. A gambler offers $1 : 3$ odds that A will occur, $1 : 2$ odds that B will occur. The gambler knows that A and B cannot both occur. What odds should he or she give that A or B will occur?

    [Ans. $7 : 5$.]

14. A gambler offers $3 : 1$ odds that A will occur, $2 : 1$ odds that B will occur. The gambler knows that A and B cannot both occur. What odds should he or she give that A or B will occur?

15. Show from the definition of a probability measure that $m(X) = 1$ if and only if $X = \mathcal{U}$.

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

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

18. Prove that if R, S, and T are three sets that have no element in common,

$$m(R \cup S \cup T) = m(R) + m(S) + m(T).$$

19. If X and Y are two sets such that X is a subset of Y, prove that $m(X) \leq m(Y)$.

20. If p and q are two statements such that p implies q, prove that $\Pr[p] \leq \Pr[q]$.

21. Suppose that you are given n statements and each has been assigned a probability less than or equal to r. Prove that the probability of the disjunction of these statements is less than or equal to nr.

22. The following is an alternative proof of property (C′) of the text. Give a reason for each step.

    (a) $X \cup Y = (X \cap \tilde{Y}) \cup (X \cap Y) \cup (\tilde{X} \cap Y)$.

    (b) $m(X \cup Y) = m(X \cap \tilde{Y}) + m(X \cap Y) + m(\tilde{X} \cap Y)$.

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

23. If X, Y, and Z are any three sets, prove that, for any probability measure,

$$\begin{aligned} m(X \cup Y \cup Z) &= m(X) + m(Y) + m(Z) \\ &\quad -m(X \cap Y) - m(Y \cap Z) - m(X \cap Z) \\ &\quad +m(X \cap Y \cap Z). \end{aligned}$$

24. Translate the result of Exercise 23 into a result concerning three statements p, q, and r.

25. A man offers to bet "dollars to doughnuts" that a certain event will take place. Assuming that a doughnut costs a nickel, what must the probability of the event be for this to be a fair bet?

    [Ans. $\frac{20}{21}$.]

26. Show that the inclusion-exclusion formula 3.2 is true is n is replaced by m. Apply this result to

$$\Pr[p_1 \vee p_2 \vee \ldots \vee p_n].$$

4.3 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 $\mathcal{U} = \{\mathrm{HH}, \mathrm{HT}, \mathrm{TH}, \mathrm{TT}\}$ 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 $\mathcal{U} = \{\text{no H}, \text{one H}, \text{two H}\}$, 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 4.4 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 $4 + 1$. 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 $\frac{1}{36}$. $\diamondsuit$

Example 4.5 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 $13 \cdot 12/(52 \cdot 51) = \frac{1}{17}$. $\diamondsuit$

Example 4.6 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 that a particular prediction is true is $\frac{1}{6}$. Since this 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 $\frac{1}{6!} = \frac{1}{720}$ that such a prediction would be correct. Hence, we might conclude here that there is strong evidence that the index has some merit. $\diamondsuit$

Exercises

1. A letter is chosen at random from the word "random". What is the probability that it is an n? That it is a vowel?

   [Ans. $\frac{1}{6}$; $\frac{1}{3}$.]

2. An integer between 3 and 12 inclusive is chosen at random. What is the probability that it is an even number? That it is even and divisible by three?

3. A card is drawn at random from a pack of playing cards.

   (a) What is the probability that it is either a heart or the king of clubs?

   [Ans. $\frac{7}{26}$.]

   (b) What is the probability that it is either the queen of hearts or an honor card (i.e., ten, jack, queen, king, or ace)?

[Ans. $\frac{5}{13}$.]

4. A word is chosen at random from the set of words

$$\mathcal{U} = \{\text{men}, \text{bird}, \text{ball}, \text{field}, \text{book}\}.$$

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) $p \wedge q$.
(e) $(p \vee q) \wedge \neg r$.
(f) $p \to q$.

[Ans. $\frac{4}{5}$.]

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. In the Primary voting example of Section 2.1, assume that all 36 possibilities in the elections are equally likely. Find

(a) The probability that candidate A wins more states than either B or C.

[Ans. $\frac{7}{18}$.]

(b) That all the states are won by the same candidate.

[Ans. $\frac{1}{36}$.]

(c) That every state is won by a different candidate.

[Ans. 0.]

7. 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?

8. Two boys and two girls are placed at random in a row for a picture. What is the probability that the boys and girls alternate in the picture?

[Ans. $\frac{1}{3}$.]

9. A certain college has 500 students and it is known that

300	read French.
200	read German.
50	read Russian.
20	read French and Russian.
30	read German and Russian.
20	read German and French.
10	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?

(b) Reads at least one language?

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 there is at least one empty seat between any two of them?

11. 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. $4/\binom{52}{5} = .0000015$.]

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

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

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

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

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

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

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

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

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

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

(g) Straight or better.

[Ans. .0076.]

12. If ten people are seated at a circular table at random, what is the probability that a particular pair of people are seated next to each other?

[Ans. $\frac{2}{9}$.]

13. 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 a 3? Answer this problem assuming that $n = 5, 4, 3, 2$.

[Ans. $\frac{1}{5}; \frac{1}{3}; \frac{2}{3}; 0$.]

14. In Exercise 13, suppose that we observe two men leaving the room and that the larger of their badge numbers is 3. What might we guess as to the number of people in the room?

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

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

[Ans. $\frac{4!\binom{13}{5}\binom{13}{4}\binom{13}{3}\binom{13}{1}}{\binom{52}{13}} = .129$.]

(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.]

16. There are $\binom{52}{13} = 6.5 \times 10^{11}$ 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?

Supplementary exercises.

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

18. Find the probability of a "bust" hand in poker. [Hint: A hand is a "bust" if there is no pair, and it is neither a straight nor a flush.]

[Ans. .5012.]

19. In poker, find the probability of having

(a) Exactly one pair.

[Ans. .4226.]

(b) Two pairs.

[Ans. .0475.]

(c) Three of a kind.

[Ans. .0211.]

20. Verify from Exercises 11, 18, 19 that the probabilities for all possible poker hands add up to one (within a rounding error).

21. A certain French professor announces that he or she 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.

(a) What is the maximum number of pages that a student should prepare in order to be certain of being able to translate a page that he or she has studied?

(b) Smith decides to study only four of the eight pages. What is the probability that one of these four pages will appear on the examination?

4.4 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 would 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 or her birthday. For each of these, if the second person is to have a different birthday, there are only 364 possibilities for his or her birthday. For the third person, there are 363 possibilities if he or she 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

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

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 4.1: ♢

The probability that at least two people have the same birthday is then $p_r = 1 - q_r$. In Figure 4.1 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 clerk has checked n hats, but they have become hopelessly scrambled. The clerk hands back the hats at random. What is the probability that at least one head gets its 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 head gets its own hat back". We wish to find $\Pr[p_1 \vee p_2 \vee \ldots \vee p_n]$. We know from Exercise 26 that a probability of this form can be found from the inclusion-exclusion formula. We must add all probabilities 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 heads get their own hats back. These probabilities are very easy to compute. Let us find the probability that out of n heads 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 heads get their own hats back is

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

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

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

to the alternating sum. Thus

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

where the $+$ sign is chosen if n is odd and the $-$ sign if n is even. In Figure 4.2, 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 people in the United States Senate have the same birthday?

[Ans. $3,300,000 : 1$.]

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 4.2: ♢

2. What is the probability that in the House of Representatives at least two men have the same birthday?

3. What odds should you be willing to give on a bet that at 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 born on Nov. 2.]

4. 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.]

5. Four men check their hats. Assuming that the hats are returned at random, what is the probability that *exactly* four men get their own hats? Calculate the answer for 3, 2, 1, 0 men.

   [Ans. $\frac{1}{24}$; 0; $\frac{1}{4}$; $\frac{1}{3}$; $\frac{3}{8}$.]

6. A group of 50 knives and forks a dance. The partners for a dance are chosen by lot (knives dance with forks). What is the approximate probability that no knife dances with its own fork?

7. Show that the probability that, in a group of r people, exactly one pair has the same birthday is

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

8. Show that $t_r = \binom{r}{2} \frac{q_r}{366-r}$, where t_r is defined in Exercise 7, and q_r is the probability that no pair has the same birthday.

9. Using the result of Exercise 8 and the results given in Figure 4.1, find the probability of exactly one pair of people with the same birthday in a group of r people, for $r = 15, 20, 25, 30, 40, 50$.

[Ans. .22; .32; .38; .38; .26; .12.]

10. What is the approximate probability that there has been exactly one pair of Presidents with the same birthday?

Supplementary exercises.

11. Find a formula for the probability of having more than one coincidence of birthdays among n people, i.e., of having at least two pairs of identical birthdays, or of three or more people having the same birthday. [Hint: Take the probability of at least one coincidence, and subtract the probability of having exactly one pair.]

12. Compute the probability of having more than one coincidence of birthdays when there are 20, 25, 30, 40, or 50 people in the room.

13. What is the smallest number of people you need in order to have a better than even chance of finding more than one coincidence of birthdays?

[Ans. 36.]

14. Is it very surprising that there was more than one coincidence of birthdays among the dates on which Presidents died?

15. A game of solitaire is played as follows: A deck of cards is shuffled, and then the player turns the cards up one at a time. As the player turns the cards, he or she calls out the names of the cards

in a standard order—say "two of clubs", "three of clubs", etc. The object of the game is to go through the entire deck without once calling out the name of the card one turns up. What is the probability of winning? How does the probability change if one uses a single suit in place of a whole deck?

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?

The probability of p after the receipt of the information q is called its *conditional probability*, and it is denoted by $\Pr[p|q]$, which is read "the probability of p given q". In this section we will 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 $k = 1/m(Q)$ and our new measure on subsets of $\mathcal{U}$ is determined by the formula

$$m'(X) = \frac{m(X)}{m(Q)}. \tag{4.1}$$

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

$$\Pr[p|q] = m'(P \cap Q) = \frac{m(P \cap Q)}{m(Q)} = \frac{\Pr[p \wedge q]}{\Pr[q]}. \tag{4.2}$$

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 4.1 and 4.2 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 4.7 In an election, candidate A has a .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. $\diamondsuit$

Example 4.8 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 $\mathcal{U} = \{\mathrm{BB}, \mathrm{BG}, \mathrm{GB}, \mathrm{GG}\}$, where the first letter of the pair indicates the sex of the younger child and the second that of the older. The information that there is a boy causes $\mathcal{U}$ to change to $\{\mathrm{BB}, \mathrm{BG}, \mathrm{GB}\}$, 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 $\{\mathrm{BB}, \mathrm{BG}\}$ and the conditional probability is $\frac{1}{2}$. $\diamondsuit$

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

$$\Pr[p \wedge q] = \Pr[p]\Pr[q]. \tag{4.3}$$

And the case $\Pr[q|p] = q$ leads to the same equation. Whenever equation 4.3 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 $\Pr[p|q] = \Pr[p]$.

Example 4.9 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. $\diamondsuit$

While we have an intuitive notion of independence, it can happen that two statements, which 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. We wish to find the probability of a statement p. We observe that there is 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

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

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

Example 4.10 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

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

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

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

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 5, given that it is between 2 and 7 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 6 is three times as likely to turn up as a 2).

   (a) What is the probability of rolling a 3 given that an odd number turns up?

   [Ans. $\frac{1}{3}$.]

   (b) What is the probability of rolling an even number given that a number greater than three turns up?

   [Ans. $\frac{2}{3}$.]

3. A die is thrown twice. What is the probability that the sum of the faces which turn up is greater than 10, given that one of them is a 6? Given that the first throw is a 6?

   [Ans. $\frac{3}{11}$; $\frac{1}{3}$.]

4. Referring to Exercise 9, what is the probability that the students selected studies German if

   (a) He or she studies French?

   (b) He or she studies French and Russian?

   (c) He or she studies neither French nor Russian?

5. In the primary voting example of Section 2.1, assuming that the equiprobable measure has been assigned, find the probability that A wins at least two primaries, given that B drops out of the Wisconsin primary.

[Ans. $\frac{7}{9}$.]

6. If $\Pr[\neg p] = \frac{1}{4}$ and $\Pr[q|p] = \frac{1}{2}$, what is $\Pr[p \wedge q]$?

[Ans. $\frac{3}{8}$.]

7. A student takes a five-question true-false exam. What is the probability that the student will get all answers correct if

   (a) The student is only guessing?

   (b) The student knows that the instructor puts more true than false questions on his or her exams?

   (c) The student also knows that the instructor never puts three questions in a row with the same answer?

   (d) The student also knows that the first and last questions must have the opposite answer?

   (e) The student also knows that the answer to the second problem is "false"?

8. Three persons, A, B, and C, are placed at random in a straight line. Let r be the statement "B is to the right of A" and let s be the statement "C is to the right of A".

   (a) What is the $\Pr[r \wedge s]$?

   [Ans. $\frac{1}{3}$.]

   (b) Are r and s independent?

   [Ans. No.]

9. 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. $\frac{5}{13} = .38$.]

The probability that the cards are both jacks, given that one is the jack of hearts.

[Ans. $\frac{3}{7} = .43$.]

10. Prove that statements r and s in Example 4.9 are independent.

11. 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 the same". Compute $\Pr[r], \Pr[r|p], \Pr[r|q], \Pr[r|p \wedge q], \Pr[r|p \vee q]$.

[Ans. $\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, 1, \frac{1}{3}$.]

12. Prove that for any two statements p and q,

$$\Pr[p] = \Pr[p \wedge q] + \Pr[p \wedge \neg q].$$

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. Prove that the procedure given in Example 4.10 does maximize the chance of getting a black ball. [Hint: Show that you can assume that one urn contains more black balls than white balls and then consider what is the best that could be achieved, first in the urn with more black than white balls, and then in the urn with more white than black balls.]

Supplementary exercises.

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 $\neg q$.

    (b) $\neg q$ and p.

    (c) $\neg p$ and $\neg q$

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 coin is thrown twice. Let p be the statement "Heads turns up on the first toss" and q the statement "Heads turns up on the second toss". Show that it is possible to assign a measure to the possibility space $\{\mathrm{HH}, \mathrm{HT}, \mathrm{TH}, \mathrm{TT}\}$ so that these statements are *not* independent.

18. A multiple-choice test question lists five alternative answers, of which just one is correct. If a student has done the homework, then he or she is certain to identify the correct answer; otherwise, the student chooses an answer at random. Let p be the statement "The student does the homework" and q the statement "The student 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?

19. A coin is weighted so that heads has probability .7, tails has probability .2, and it stands on edge with probability .1. What is the probability that it does not come up heads, given that it does not come up tails?

    [Ans. $\frac{1}{8}$.]

20. 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: An even numbered heart is drawn. q: A red card smaller than a five is drawn.

21. 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 or she has blue eyes; otherwise he or she 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 person has brown eyes, what is the probability that he or she has two B genes?

Assume that a child's mother and father have brown eyes and blue eyes, respectively.

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

(c) If the child has brown eyes, what is the probability that the father has two B genes?

[Ans. $\frac{1}{2}$.]

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?

[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?

[Ans. $\frac{7}{10}$.]

4.6 Finite stochastic processes

We consider here a very general situation which we will 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 $\frac{1}{2}$ regardless

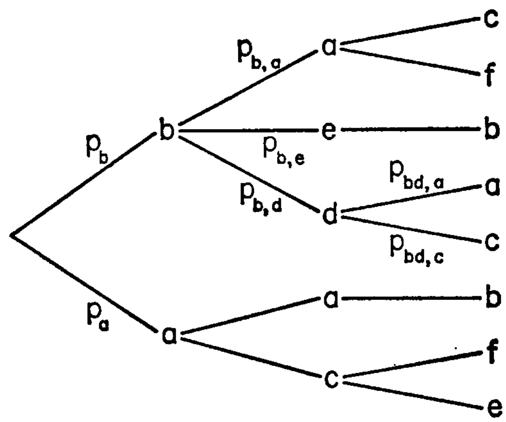


Figure 4.3: $\diamondsuit$

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 a probability measure can be assigned, using the given information. In the case of coin tossing, the probabilities (hence also the possibilities) on any given experiment do not depend upon the previous results. We will 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 4.3. 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 $p_a + p_b = 1$, $p_{b,a} + p_{b,e} + p_{b,d} = 1$, and $p_{bd,a} + p_{bd,c} = 1$, etc.

It is convenient to associate each probability with the branch of the tree that connects to the branch point representing the predicted outcome. We have done this in Figure 4.3 for several branches. The sum of the numbers assigned to branches from a particular branch point is 1, 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 1. The weights are products of positive numbers and hence positive. To see that their sum is 1 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

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

Similarly, 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

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

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

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 4.11 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 their possibility tree in Figure 4.4. The probability that a white ball is drawn is then found

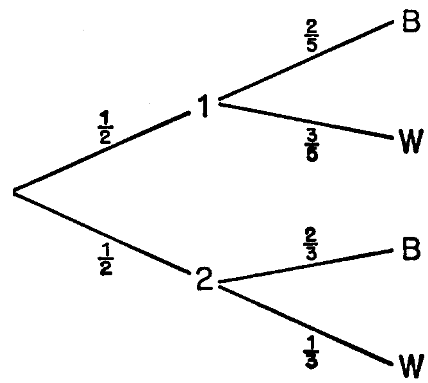


Figure 4.4: $\diamondsuit$

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

Example 4.12 Suppose that a drunkard leaves a bar which is on a corner which he or she knows to be one block from home. He or she is unable to remember which street leads home, and proceeds to try each of the streets at random without ever choosing the same street twice until he or she goes on the one which leads home. What possibilities are there for the 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 4.5. The probability for any particular trip, or path, is found by taking the product of the branch probabilities. $\diamondsuit$

Example 4.13 Assume that we 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 $\frac{1}{4}$. We are not told which machine is A. Suppose that we choose a machine at random and win. What is the probability that we chose machine A? We first construct the tree (Figure 4.6) to show the possibilities and assign branch 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

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

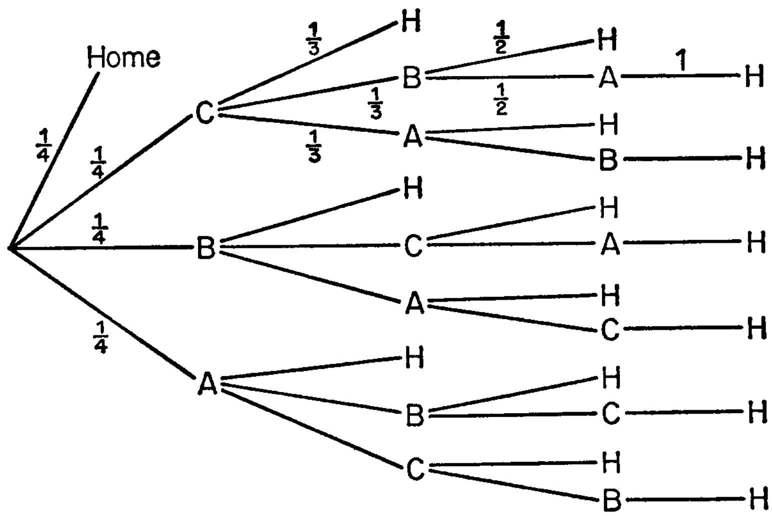


Figure 4.5: ♢

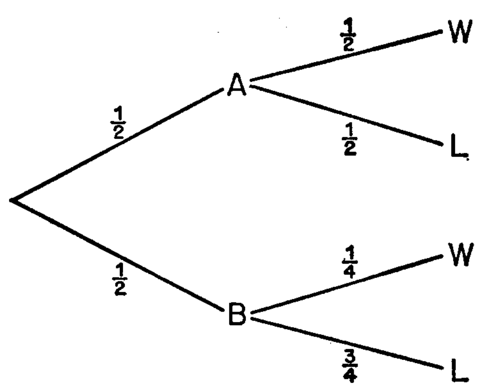


Figure 4.6: ♢

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.) $\diamondsuit$

Exercises

1. The fractions of Republicans, Democrats, and Independent voters in cities A and B are

   City A: .30 Republican, .40 Democratic, .30 Independent;

   City B: .40 Republican, .50 Democratic, .10 Independent.

   A city is chosen at random and two voters are chosen successively and at random from the voters of this city. Construct a tree measure and find the probability that two Democrats are chosen. Find the probability that the second voter chosen is an Independent voter.

   [Ans. .205; .2.]

2. A coin is thrown. If a head turns up, a die is rolled. If a tail turns up, the coin is thrown again. Construct a tree measure to represent the two experiments and find the probability that the die is thrown and a six turns up.

3. An athlete wins a certain tournament if he or she 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 $\frac{2}{3}$. The probability of winning a game when A is the opponent is only $\frac{1}{3}$. Construct a tree measure for the possibilities for three games, assuming that he or she plays alternately but plays A first. Do the same assuming that he or she plays B first. In each case find the probability that he or she will win two consecutive games. Is it better to play two games against the strong player or against the weaker player?

[Ans. $\frac{10}{27}$; $\frac{8}{27}$; better to play strong player twice.]

4. 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.

5. A student claims to be able to distinguish beer from ale. The student is given a series of three tests. In each test, the student is given two cans of beer and one of ale and asked to pick out the ale. If the student gets two or more correct we will admit the claim. Draw a tree to represent the possibilities (either right or wrong) for the student's answers. Construct the tree measure which would correspond to guessing and find the probability that the claim will be established if the student guesses on every trial.

6. A box contains three defective light bulbs and seven good ones. Construct a tree to show the possibilities if three consecutive bulbs are drawn at random from the box (they are not replaced after being drawn). Assign a tree measure and find the probability that at least one good bulb is drawn out. Find the probability that all three are good if the first bulb is good.

   [Ans. $\frac{119}{120}$; $\frac{5}{12}$.]

7. In Example 4.12, find the probability that the drunkard reaches home after trying at most one wrong street.

8. In Example 4.13, find the probability that machine A was chosen, given that we lost.

9. In Example 4.13, assume that we make two plays. Find the probability that we win at least once under the assumption

   (a) That we play the same machine twice.

   [Ans. $\frac{19}{32}$.]

   (b) That we play the same machine the second time if and only if we won the first time.

   [Ans. $\frac{20}{32}$.]

10. A chess player plays three successive games of chess. The player's psychological makeup is such that the probability of winning a given game is $(\frac{1}{2})^{k+1}$, where k is the number of games won so far. (For instance, the probability of winning the first game is $\frac{1}{2}$, the probability of winning the second game if the player has already won the first game is $\frac{1}{4}$, etc.) What is the probability that the player will win at least two of the three games?

11. 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 the President is willing to run, the President and his or her 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?

[Ans. $\frac{13}{40}$.]

12. There are two urns, A and B. Urn A contains one black and one red ball. Urn B contains two black and three red balls. A ball is chosen at random from urn A and put into urn B. A ball is then drawn at random from urn B.

    (a) What is the probability that both balls drawn are of the same color?

    [Ans. $\frac{7}{12}$.]

    (b) What is the probability that the first ball drawn was red, given that the second ball drawn was black?

    [Ans. $\frac{2}{5}$.]

Supplementary exercises.

13. Assume that in the World Series each team has probability $\frac{1}{2}$ of winning each game, independently of the outcomes of any other game. Assign a tree measure. (See Section **??** for the tree.) Find the probability that the series ends in four, five, six, and seven games, respectively.

14. Assume that in the World Series one team is stronger than the other and has probability .6 for winning each of the games. Assign a tree measure and find the following probabilities.

    (a) The probability that the stronger team wins in 4, 5, 6, and 7 games, respectively.

    (b) The probability that the weaker team wins in 4, 5, 6, and 7 games, respectively.

    (c) The probability that the series ends in 4, 5, 6, and 7 games, respectively.

    [Ans. .16; .27; .30; .28.]

    (d) The probability that the strong team wins the series.

    [Ans. .71.]

15. Redo Exercise 14 for the case of two poorly matched teams, where the better team has probability .9 of winning a game.

    [Ans. (c).66;.26;.07;.01; (d).997.]

16. In the World Series from 1905 through 1965 (excluding series of more than seven games) there were 11 four-game, 14 five-game, 13 six-game, and 20 seven-game series. Which of the assumptions in Exercises 13, 14, 15 comes closest to predicting these results? Is it a good fit?

    [Ans. .6; No.]

17. Consider the following assumption concerning World Series: Ninety per cent of the time the two teams are evenly matched, while 10 per cent of the time they are poorly matched, with the better team having probability .9 of winning a game. Show that this assumption comes closer to predicting the actual outcomes than those considered in Exercise 16.

18. We are given three coins. Coin A is fair while coins B and C are loaded: B has probability .6 of heads and C has probability .4 of heads. A game is played by tossing a coin twice starting with coin B. If a head is obtained, B is tossed again, otherwise the second coin to be tossed is chosen at random from A and C.

(a) Draw the tree for this game, assigning branch and path weights.

(b) Let p be the statement "The first toss results in heads" and let q be the statement "The second toss results in heads". Find $\Pr[p]$, $\Pr[q]$, $\Pr[q|p]$.

[Ans. .6; .54; .6.]

19. A and B play a series of games for which they are evenly matched. A player wins the series either by winning two games in a row, or by winning a total of three games. Construct the tree and the tree measure.

    (a) What is the probability that A wins the series?

    (b) What is the probability that more than three games need to be played?

20. 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. $\frac{2}{3}$.]

4.7 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

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

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 4.14 Suppose that a freshman must choose among mathematics, physics, chemistry, and botany as his or her science course. On the basis of the interest he or she expressed, his or her adviser assigns probabilities of .4, .3, .2 and .1 to the student's choosing each of the four courses, respectively. The adviser does not hear which course the student actually chose, but at the end of the term the adviser hears that he or she received an 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 the original estimates as to the probabilities of the student taking the various courses? Using Bayes's theorem we get

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

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

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:

If $\Pr[p_1] = \Pr[p_2] = \Pr[p_3] = \Pr[p_4]$ then

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

Example 4.15 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

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

$\diamondsuit$

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 4.16 There are three urns. Each urn contains one white ball. ln addition, urn I contains one black ball, urn II contains two, and urn III contains 3. 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, in Figure 4.7.

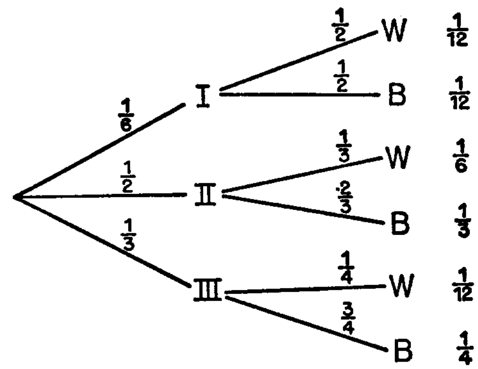


Figure 4.7: $\diamondsuit$

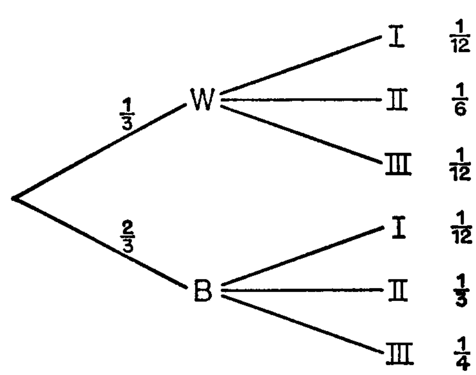


Figure 4.8: $\diamondsuit$

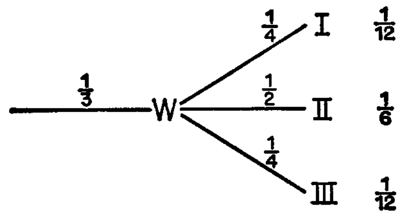


Figure 4.9: $\diamondsuit$

Next we redraw the tree, using the ball drawn as stage 1, and the urn selected as stage 2. (See Figure 4.8.) 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

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

This leaves the branch weights of the second stage to be computed. But this is simply a matter of division. For example, the branch 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 4.9). $\diamondsuit$

This method is particularly useful when we wish to compute all the conditional probabilities. We will apply the method next to Example 4.14. The tree and tree measure for this example in the natural order is shown in Figure 4.10. In that figure the letters M, P, C, and B stand for mathematics, physics, chemistry, and botany, respectively.

The tree drawn in reverse order is shown in Figure 4.11. 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 or she does not receive an A. These probabilities are also easily obtained from the first tree. In fact,

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

and

$$\Pr[\neg\mathrm{A}] = 1-\Pr[\mathrm{A}] = .75.$$

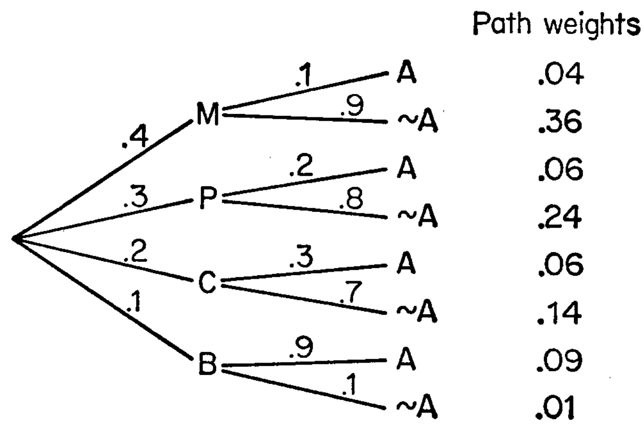


Figure 4.10: ♢

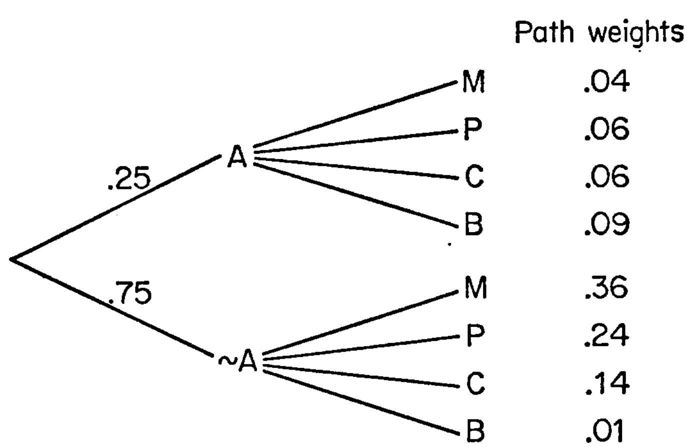


Figure 4.11: ♢

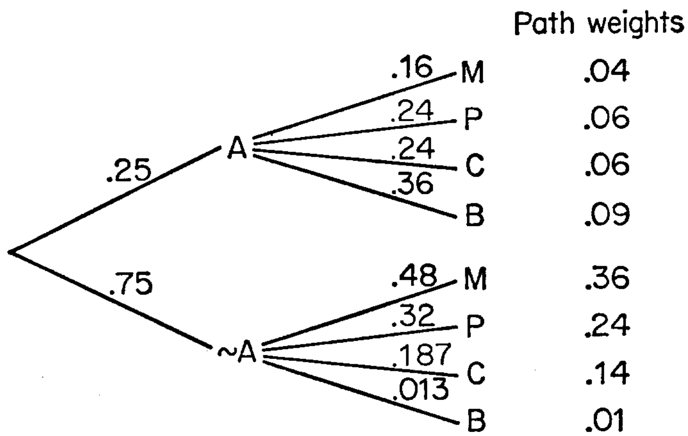


Figure 4.12: $\diamondsuit$

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 $p_{\mathrm{A,M}}$ we have

$$.25 \cdot p_{\mathrm{A,M}} = .04;\ p_{\mathrm{A,M}} = .16.$$

But $p_{\mathrm{A,M}}$ is also the conditional probability that the student took math given that he or she 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 4.12.

Exercises

1. Urn I contains 7 red and 3 black balls and urn II contains 6 red and 4 black balls. An urn is chosen at random and two balls are drawn from it in succession without replacement. The first ball is red and the second black. Show that it is more probable that urn II was chosen than urn I.

2. A gambler is told that one of three slot machines pays off with probability $\frac{1}{2}$, while each of the other two pays off with probability $\frac{1}{3}$.

(a) If the gambler selects one at random and plays it twice, what is the probability that he or she will lose the first time and win the second?

[Ans. $\frac{25}{108}$.]

(b) If the gambler loses the first time and wins the second, what is the probability he or she chose the favorable machine?

[Ans. $\frac{9}{25}$.]

3. During the month of May the probability of a rainy day is .2. The Dodgers win on a clear day with probability .7, but on a rainy day only with probability .4. If we know that they won a certain game in May, what is the probability that it rained on that day?

[Ans. $\frac{1}{8}$.]

4. Construct a diagram to represent the truth sets of various statements occurring in the previous exercise.

5. On a multiple-choice exam there are four possible answers for each question. Therefore, if a student knows the right answer, he or she has probability 1 of choosing correctly; if the student is guessing, he or she has probability $\frac{1}{4}$ of choosing correctly. Let us further assume that a good student will know 90 per cent of the answers, a poor student only 50 per cent. If a good student has the right answer, what is the probability that he or she was only guessing? Answer the same question about a poor student, if the poor student has the right answer.

[Ans. $\frac{1}{37}$; $\frac{1}{5}$.]

6. Three economic theories are proposed at a given time, which appear to be equally likely on the basis of existing evidence. The state of the American economy is observed the following year, and it turns out that its actual development had probability .6 of happening according to the first theory; and probabilities .4 and .2 according to the others. How does this modify the probabilities of correctness of the three theories?

7. Let p_1, p_2, p_3, and p_4 be a set of equally likely alternatives. Let $\Pr[q|p_1] = a$, $\Pr[q|p_2] = b$, $\Pr[q|p_3] = c$, $\Pr[q|p_4] = d$. Show that if $a+b+c+d = 1$, then the revised probabilities of the alternatives relative to q are a, b, c, and d, respectively.

8. In poker, Smith holds a very strong hand and bets a considerable amount. The probability that Smith's opponent, Jones, has a better hand is .05. With a better hand Jones would raise the bet with probability .9, but with a poorer hand Jones would raise only with probability .2. Suppose that Jones raises, what is the new probability that he or she has a winning hand?

   [Ans. $\frac{9}{47}$.]

9. A rat is allowed to choose one of five mazes at random. If we know that the probabilities of his or her getting through the various mazes in three minutes are .6, .3, .2, .1, .1, and we find that the rat escapes in three minutes, how probable is it that he or she chose the first maze? The second maze?

   [Ans. $\frac{6}{13}$; $\frac{3}{13}$.]

10. Three men, A, B, and C, are in jail, and one of them is to be hanged the next day. The jailer knows which man will hang, but must not announce it. Man A says to the jailer, "Tell me the name of one of the other 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 jailer 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 jailer's reasoning correct?

    [Ans. No.]

11. 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?

12. Referring to Exercise 11, 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.]

13. A statistician will accept job a with probability $\frac{1}{2}$, job b with probability $\frac{1}{3}$, and job c with probability $\frac{1}{6}$. In each case he or she must decide whether to rent or buy a house. The probabilities of buying are $\frac{1}{3}$ if he or she takes job a, $\frac{2}{3}$ if he or she takes job b, and 1 if he or she takes job c. Given that the statistician buys a house, what are the probabilities of having taken each job?

[Ans. .3; .4; .3.]

14. Assume that chest X-rays for detecting tuberculosis have the following properties. For people having tuberculosis the test will detect the disease 90 out of every 100 times. For people not having the disease the test will in 1 out of every 100 cases diagnose the patient incorrectly as having the disease. Assume that the incidence of tuberculosis is 5 persons per 10,000. A person is selected at random, given the X-ray test, and the radiologist reports the presence of tuberculosis. What is the probability that the person in fact has the disease?

4.8 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 outselves 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 4.13.

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 4.11.

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 4.13. The weights assigned to each path are indicated at the end of the path. 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 4.13 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. 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

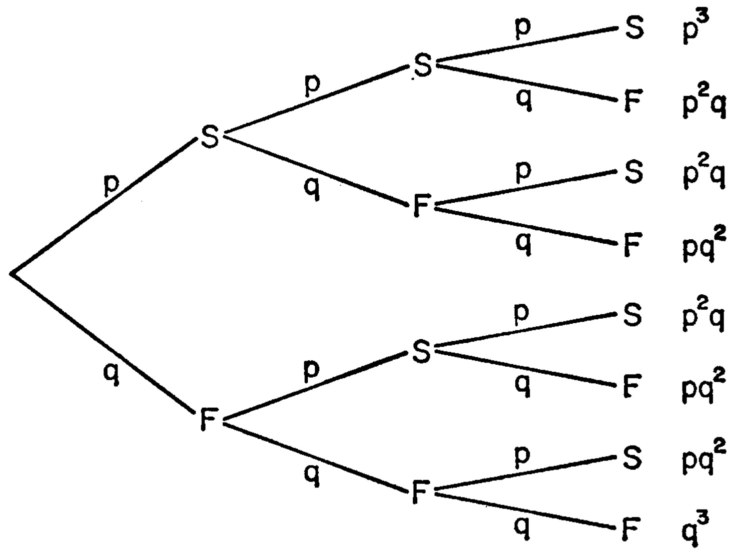


Figure 4.13: $\diamondsuit$

successes occur". To each such path we make correspond an ordered partition of the integers from 1 to n which has two cells, x elements in the first and $n-x$ in the second. We do this by putting the numbers of the experiments on which success occurred in the first cell and those for which failure occurred in the second cell. Since there are $\binom{n}{x}$ such partitions there are also this number of paths in the truth set of the statement considered. 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 4.17 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

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

For example, in l00 throws the probability that exactly 50 heads will turn up is

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

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 l0 from 50. To find this we must add $f(100, x; \frac{1}{2})$ for $x = 40, 41, \ldots, 60$. 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 l0. ♢

Example 4.18 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 $n = 4$, $n = 8$, $n = 12$, and $n = 16$ the probability of obtaining exactly $\frac{1}{2}$ heads when an ordinary coin is thrown.

   [Ans. .375; .273; .226; .196.]

2. Compute for $n = 4$, $n = 8$, $n = 12$, and $n = 16$ the probability that the fraction of heads deviates from $\frac{1}{2}$ by less than $\frac{1}{5}$.

   [Ans. .375; .711, .854; .923.]

3. Verify that the probability .991 given in Example 4.18 is correct.

4. Assume that Peter and Paul match pennies four times. (In matching pennies, Peter wins a penny with probability $\frac{1}{2}$, and Paul wins a penny with probability $\frac{1}{2}$.) What is the probability that Peter wins more than Paul? Answer the same for five throws. For the case of 12,917 throws.

   [Ans. $\frac{5}{16}$; $\frac{1}{2}$; $\frac{1}{2}$.]

5. If an ordinary die is thrown four times, what is the probability that exactly two sixes will occur?

6. In a ten-question true-false exam, what is the probability of getting 70 per cent or better by guessing?

   [Ans. $\frac{11}{64}$.]

7. Assume that, every time a batter comes to bat, he or she has probability .3 for getting a hit. Assuming that hits form an independent trials process and that the batter comes to bat four times, what fraction of the games would he or she expect to get at least two hits? At least three hits? Four hits?

   [Ans. .348; .084; .008.]

8. 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 resulted in heads?

9. A small factory has ten workers. The workers eat their lunch at one of two diners, and they are just as likely to eat in 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.]

10. Suppose that five people are chosen at random and asked if they favor a certain proposal. If only 30 per cent of the people favor the proposal, what is the probability that a majority of the five people chosen will favor the proposal?

11. In Example 4.18, if the probability for a machine reversing its answer due to a parts failure is .2, how many machines would have to be used to make the probability greater than .89 that the answer obtained would be that which a machine with no failure would give?

    [Ans. Three machines.]

12. Assume that it is estimated that a torpedo will hit a ship with probability $\frac{1}{3}$. How many torpedos must be fired if it is desired that the probability for at least one hit should be greater than .9?

13. A student estimates that, if he or she takes four courses, he or she has probability .8 of passing each course. If he or she takes five courses, he or she has probability .7 of passing each course, and if he or she takes six courses he or she has probability .5 for passing each course. The student's only goal is to pass at least four courses. How many courses should he or she take for the best chance of achieving this goal?

    [Ans. 5.]

Supplementary exercises.

14. 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 Park Bench, he or she must roll a seven or doubles before being allowed to move out.

(a) What is the probability that a player stuck on Park Bench will be allowed to move out on the next turn?

[Ans. $\frac{1}{3}$.]

(b) How many times must a player stuck on Park Bench roll before the chances of getting out exceed $\frac{3}{4}$.

[Ans. 4.]

15. 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?

    (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?

16. Show that it is more probable to get at least one ace with 4 dice than at least one double ace in 24 throws of two dice.

17. A thick coin, when tossed, will land "heads" with a probability of $\frac{5}{12}$, "tails" with a probability of $\frac{5}{12}$, and will land on edge with a probability of $\frac{1}{6}$. If it is tossed six times, what is the probability that it lands on edge exactly two times?

[Ans. .2009.]

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

19. A certain team has probability $\frac{2}{3}$ of winning whenever it plays.

    (a) What is the probability the team will win exactly four out of five games?

    [Ans. $\frac{80}{243}$.]

    (b) What is the probability the team will win at most four out of five games?

[Ans. $\frac{211}{243}$.]

(c) What is the probability the team will win exactly four games out of five if it has already won the first two games of the five?

[Ans. $\frac{4}{9}$.]

4.9 A problem of decision

In the preceding sections we have dealt with the problem of calculating the probability of certain statements based on the assumption of a given probability measure. In a statistics problem, one is often called upon to make a decision in a case where the decision would be relatively easy to make if we could assign probabilities to certain statements, but we do not know how to assign these probabilities. For example, if a vaccine for a certain disease is proposed, we may be called upon to decide whether or not the vaccine should be used. We may decide that we could make the decision if we could compare the probability that a person vaccinated will get the disease with the probability that a person not vaccinated will get the disease. Statistical theory develops methods to obtain from experiments some information which will aid in estimating these probabilities, or will otherwise help in making the required decision. We shall illustrate a typical procedure.

Smith claims to have the ability to distinguish ale from beer and has bet Jones a dollar to that effect. Now Smith does not mean that he or she can distinguish beer from ale every single time, but rather a proportion of the time which is significantly greater than $\frac{1}{2}$.

Assume that it is possible to assign a number p which represents the probability that Smith can pick out the ale from a pair of glasses, one containing ale and one beer. We identify $p = \frac{1}{2}$ with having no ability, $p > \frac{1}{2}$ with having some ability, and $p < \frac{1}{2}$ with being able to distinguish, but having the wrong idea which is the ale. If we knew the value of p, we would award the dollar to Jones if p were $\leq \frac{1}{2}$, and to Smith if p were $> \frac{1}{2}$. As it stands, we have no knowledge of p and thus cannot make a decision. We perform an experiment and make a decision as follows.

Smith is given a pair of glasses, one containing ale and the other beer, and is asked to identify which is the ale. This procedure is repeated ten times, and the number of correct identifications is noted. If the number correct is at least eight, we award the dollar to Smith, and if it is less than eight, we award the dollar to Jones.

We now have a definite procedure and shall examine this procedure both from Jones's and Smith's points of view. We can make two kinds of errors. We may award the dollar to Smith when in fact the appropriate value of p is $\leq \frac{1}{2}$, or we may award the dollar to Jones when the appropriate value for p is $> \frac{1}{2}$ There is no way that these errors can be completely avoided. We hope that our procedure is such that each bettor will be convinced that, if he or she is right, he or she will very likely win the bet.

Jones believes that the true value of p is $\frac{1}{2}$. We shall calculate the probability of Jones winning the bet if this is indeed true. We assume that the individual tests are independent of each other and all have the same probability $\frac{1}{2}$ for success. (This assumption will be unreasonable if the glasses are too large.) We have then an independent trials process with $p = \frac{1}{2}$ to describe the entire experiment. The probability that Jones will win the bet is the probability that Smith gets fewer than eight correct. From the table in Figure 4.14 we compute that this probability is approximately .945. Thus Jones sees that, if he or she is right, it is very likely that he or she will win the bet.

Smith, on the other hand, believes that p is significantly greater than $\frac{1}{2}$. If Smith believes that p is as high as .9, we see from Figure 4.14 that the probability of Smith's getting eight or more correct is .930. Then both parties will be satisfied by the bet.

Suppose, however, that Smith thinks the value of p is only about .75. Then the probability that Smith will get eight or more correct and thus win the bet is .526. There is then only an approximately even chance that the experiment will discover Smith's abilities, and Smith probably will not be satisfied with this. If Smith really thinks his or her ability is represented by a p value of about $\frac{3}{4}$, we would have to devise a different method of awarding the dollar. We might, for example, propose that Smith win the bet if he or she gets seven or more correct. Then, if Smith has probability $\frac{3}{4}$ of being correct on a single trial, the probability that Smith will win the bet is approximately .776. If $p = \frac{1}{2}$ the probability that Jones will win the bet is about .828 under this new arrangement. Jones's chances of winning are thus decreased, but Smith may be able to convince him or her that it is a fairer arrangement than the first procedure.

In the above example, it was possible to make two kinds of errors. The probability of making these errors depended on the way we

Table of Values of $f(10, x; p)$

x \ p	0.1	0.25	0.50	0.75	0.90
0	.349	.056	.001	.000	.000
1	.387	.188	.010	.000	.000
2	.194	.282	.044	.000	.000
3	.057	.250	.117	.003	.000
4	.011	.146	.205	.016	.000
5	.001	.058	.246	.058	.001
6	.000	.016	.205	.146	.011
7	.000	.003	.117	.250	.057
8	.000	.000	.044	.282	.194
9	.000	.000	.010	.188	.387
10	.000	.000	.001	.056	.349

Figure 4.14: ♢

designed the experiment and the method we used for the required decision. In some cases we are not too worried about the errors and can make a relatively simple experiment. In other cases, errors are very important, and the experiment must be designed with that fact in mind. For example, the possibility of error is certainly important in the case that a vaccine for a given disease is proposed, and the statistician is asked to help in deciding whether or not it should be used. In this case it might be assumed that there is a certain probability p that a person will get the disease if not vaccinated, and a probability r that the person will get it if he or she is vaccinated. If we have some knowledge of the approximate value of p, we are then led to construct an experiment to decide whether r is greater than p, equal to p, or less than p. The first case would be interpreted to mean that the vaccine actually tends to produce the disease, the second that it has no effect, and the third that it prevents the disease; so that we can make three kinds of errors. We could recommend acceptance when it is actually harmful, we could recommend acceptance when it has no effect, or finally we could reject it when it actually is effective. The first and third might result in the loss of lives, the second in the loss of time and money of those administering the test. Here it would certainly be important that the probability of the first and third kinds of errors be made small. To see how it is possible to make the probability of both errors small, we return to the case of Smith and Jones.

Suppose that, instead of demanding that Smith make at least eight correct identifications out of ten trials, we insist that Smith make at least 60 correct identifications out of 100 trials. (The glasses must now be very small.) Then, if $p = \frac{1}{2}$, the probability that Jones wins the bet is about .98; so that we are extremely unlikely to give the dollar to Smith when in fact it should go to Jones. (If $p < \frac{1}{2}$ it is even more likely that Jones will win.) If $p > \frac{1}{2}$ we can also calculate the probability that Smith will win the bet. These probabilities are shown in the graph in Figure 4.15. The dashed curve gives for comparison the corresponding probabilities for the test requiring eight out of ten correct. Note that with 100 trials, if p is $\frac{3}{4}$, the probability that Smith wins the bet is nearly 1, while in the case of eight out of ten, it was only about $\frac{1}{2}$. Thus in the case of 100 trials, it would be easy to convince both Smith and Jones that whichever one is correct is very likely to win the bet.

Thus we see that the probability of both types of errors can be made small at the expense of having a large number of experiments.

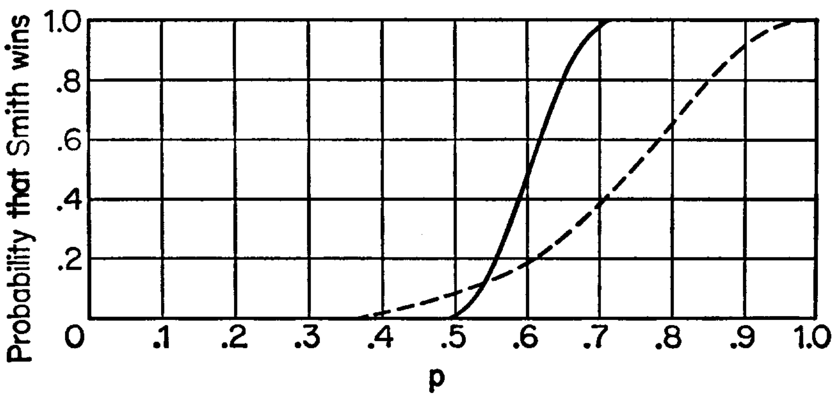


Figure 4.15: $\diamondsuit$

Exercises

1. Assume that in the beer and ale experiment Jones agrees to pay Smith if Smith gets at least nine out of ten correct.

   (a) What is the probability of Jones paying Smith even though Smith cannot distinguish beer and ale, and guesses?

   [Ans. .011.]

   (b) Suppose that Smith can distinguish with probability .9. What is the probability of not collecting from Jones?

   [Ans. .264.]

2. Suppose that in the beer and ale experiment Jones wishes the probability to be less than .1 that Smith will be paid if, in fact, Smith guesses. How many of ten trials must Jones insist that Smith get correct to achieve this?

3. In the analysis of the beer and ale experiment, we assume that the various trials were independent. Discuss several ways that error can enter, because of the nonindependence of the trials, and how this error can be eliminated. (For example, the glasses in which the beer and ale were served might be distinguishable.)

4. Consider the following two procedures for testing Smith's ability to distinguish beer from ale.

(a) Four glasses are given at each trial, three containing beer and one ale, and Smith is asked to pick out the one containing ale. This procedure is repeated ten times. Smith must guess correctly seven or more times. Find the probability that Smith wins by guessing.

[Ans. .003.]

(b) Ten glasses are given to Smith, and Smith is told that five contain beer and five ale, and asked to name the five that contain ale. Smith must choose all five correctly. Find the probability that Smith wins by guessing.

[Ans. .004.]

(c) Is there any reason to prefer one of these two tests over the other?

5. A testing service claims to have a method for predicting the order in which a group of freshmen will finish in their scholastic record at the end of college. The college agrees to try the method on a group of five students, and says that it will adopt the method if, for these five students, the prediction is either exactly correct or can be changed into the correct order by interchanging one pair of adjacent students in the predicted order. If the method is equivalent to simply guessing, what is the probability that it will be accepted?

[Ans. $\frac{1}{24}$.]

6. The standard treatment for a certain disease leads to a cure in $\frac{1}{4}$ of the cases. It is claimed that a new treatment will result in a cure in $\frac{3}{4}$ of the cases. The new treatment is to be tested on ten people having the disease. If seven or more are cured, the new treatment will be adopted. If three or fewer people are cured, the treatment will not be considered further. If the number cured is four, five, or six, the results will be called inconclusive, and a further study will be made. Find the probabilities for each of these three alternatives under the assumption first, that the new treatment has the same effectiveness as the old, and second, under the assumption that the claim made for the treatmnent is correct.

7. Three students debate the intelligence of Springer spaniels. One claims that Springers are mostly (say 90 per cent of them) intelligent. A second claims that very few (say 10 per cent) are intelligent, while a third one claims that a Springer is just as likely to be intelligent as not. They administer an intelligence test to ten Springers, classifying them as intelligent or not. They agree that the first student wins the bet if eight or more are intelligent, the second if two or fewer, the third in all other cases. For each student, calculate the probability that he or she wins the bet, if he or she is right.

   [Ans. .930, .930, .890.]

8. Ten students take a test with ten problems. Each student on each question has probability $\frac{1}{2}$ of being right, if he or she does not cheat. The instructor determines the number of students who get each problem correct. If instructor finds on four or more problems there are fewer than three or more than seven correct, he or she considers this convincing evidence of communication between the students. Give a justification for the procedure. [Hint: The table in Figure 4.14 must be used twice, once for the probability of fewer than three or more than seven correct answers on a given problem, and the second time to find the probability of this happening on four or more problems.]

4.10 The law of large numbers

In this section we shall study some further properties of the independent trials process with two outcomes. In Section 4.8 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 4.16 we show these probabilities graphically for $n = 8$ and $p = \frac{3}{4}$. In Figure 4.17 we have done similarly for the case of $n = 7$ and $p = \frac{3}{4}$.

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

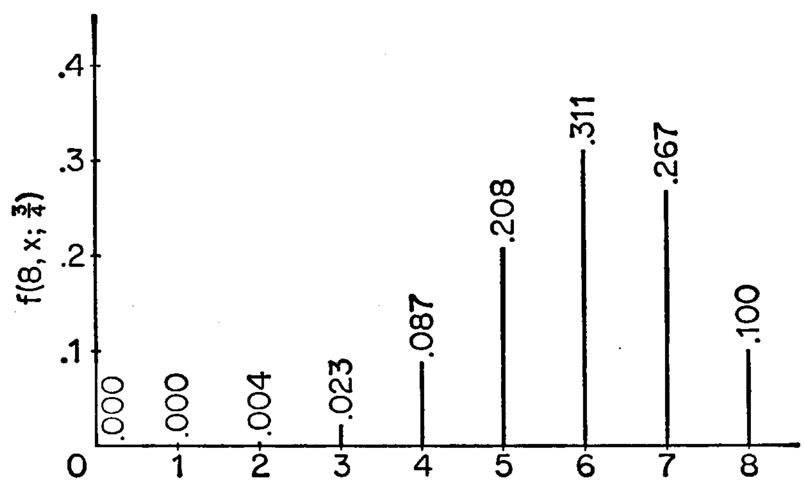


Figure 4.16: ♢

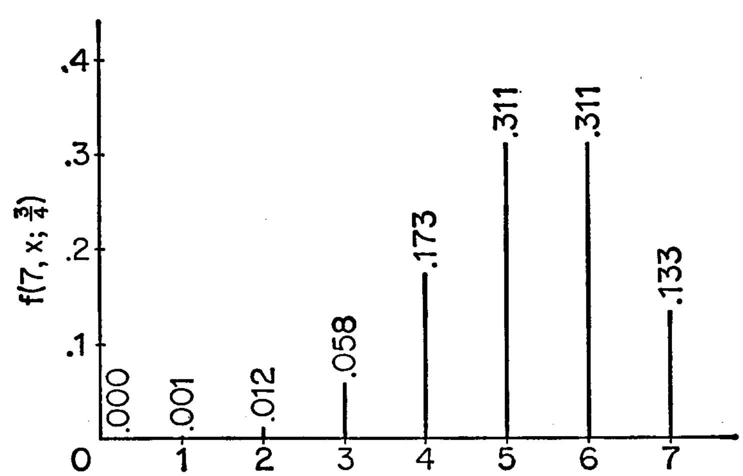


Figure 4.17: ♢

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, $np = 100 \cdot \frac{1}{2} = 50$. 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 $\epsilon > 0$,

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

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 4.18 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 after $n = 400$. In column two 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

$\Pr\left[\left|\frac{x}{n} - p\right| < \epsilon\right]$ for $p = .3$ and $\epsilon = .1, .05, .01$.

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 4.18: ♢

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

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 4.19: $\diamondsuit$

assured by one of the fundamental theorems of probability called the *law of large numbers.* This theorem asserts that, for any $\epsilon > 0$,

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

tends to 1 as n increases indefinitely.

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 .00l 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 4.19. 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 $k = 1, 2, 3$ by taking $\frac{1}{2}$ the values given above.

Example 4.19 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 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. $\diamondsuit$

Example 4.20 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. $\diamondsuit$

Example 4.21 A certain Ivy League college would like to admit 800 students in their freshman class. Experience has shown that if they admit 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 $p = \frac{800}{1250}$. Then the expected number of acceptances is 800 and the standard deviation for the number of acceptances is $\sqrt{1250 \cdot .64 \cdot .36} \approx 17$. 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. $\diamondsuit$

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. If an ordinary die is thrown 20 times, what is the expected number of times that a six will turn up? What is the standard deviation for the number of sixes that turn up?

   [Ans. $\frac{10}{3}$; $\frac{5}{3}$.]

2. Suppose that an ordinary die is thrown 450 times. What is the expected number of throws that result in either a three or a four? What is the standard deviation for the number of such throws?

3. In 16 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. 8;2.]

4. In 16 tosses of a coin, find the exact probability that the number of heads that turn up differs from the expected number by (a) as much as one standard deviation, and (b) by more than one standard deviation. Do the same for the case of two standard deviations, and for the case of three standard deviations. Show that the approximations given for large n lie between the values obtained, but are not very accurate for so small an n.

   [Ans. .454; .210; .077; .021; .004; .001.]

5. 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 $\Pr[r < \frac{x}{n} < s]$ as we increase n indefinitely? Answer the same question in the case that $r < p < s$.

6. A drug is known to be effective in 20 per cent of the cases where it is used. A new agent is introduced, and in the next 900 times the drug is used it is effective 250 times. What can be said about the effectiveness of the drug?

7. 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.]

8. 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?

9. Suppose that it is desired that the probability be approximately .95 that the fraction of sixes that turn up when a die is thrown n times does not deviate by more than .01 from the value $\frac{1}{6}$. How large should n be?

   [Ans. Approximately 5555.]

10. 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 10,000 rolls you can, with approximately 84 per cent confidence, expect to have lost not more than $(how much?).

11. Assume that 10 per cent of the people in a certain city have cancer. If 900 people are selected at random from the city, what is the expected number which will have cancer? What is the standard deviation? What is the approximate probability that more than 108 of the 900 chosen have cancer?

    [Ans. 90;9;.023.]

12. Suppose that in Exercise 11, the 900 people are chosen at random from those people in the city who smoke. Under the hypothesis that smoking has no effect on the incidence of cancer, what is the expected number in the 900 chosen that have cancer? Suppose that more than 120 of the 900 chosen have cancer, what might be said concerning the hypothesis that smoking has no effect on the incidence of cancer?

13. In Example 4.20, we made the assumption in our calculations that, if the true proportion of voters in favor of the proposal were p, then the 900 people chosen at random represented an independent trials process with probability p for a "yes" answer, and $1 - p$ for a "no" answer. Give a method for choosing the 900 people which would make this a reasonable assumption. Criticize the following methods.

    (a) Choose the first 900 people in the list of registered Republicans.

    (b) Choose 900 names at random from the telephone book.

    (c) Choose 900 houses at random and ask one person from each house, the houses being visited in the mid-morning.

14. For n throws of an ordinary coin, let t_n be such that

$$\Pr[-t_n < \frac{x}{n} - \frac{1}{2} < t_n] = .997,$$

    where x is the number of heads that turn up. Find t_n for $n = 10^4$, $n = 10^6$, and $n = 10^{20}$.

[Ans. .015; .0015; .000,000,000,15.]

15. Assume that a calculating machine carries out a million operations to solve a certain problem. In each operation the machine gives the answer 10^{-5} too small, with probability $\frac{1}{2}$, and 10^{-5} too large, with probability $\frac{1}{2}$. Assume that the errors are independent of one another. What is a reasonable accuracy to attach to the answer? What if the machine carries out 10^{10} operations?

[Ans. $\pm.01; \pm 1.$]

16. A computer tosses a coin 1 million times, and obtains 499,588 heads. Is this number reasonable?

4.11 Independent trials with more than two outcomes

By extending the results of Section 4.8, 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 occurrences of a_1, r_2 occurrences of a_2, etc. In the case of two outcomes this notation would be $f(r_1, r_2; p_1, p_2)$. In Section 4.8 we wrote this as $f(n, r+1; p)$ 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 4.20 enough of the tree for this process to indicate the branch probabilities for a path (heavy lined) 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.

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

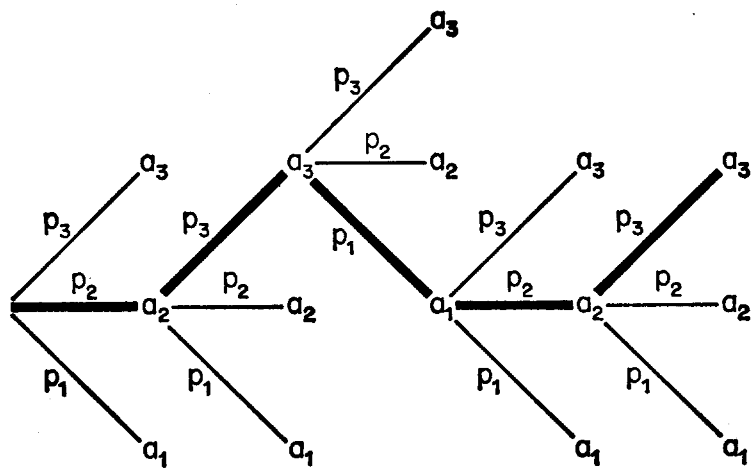


Figure 4.20: $\diamondsuit$

find $f(l, 2, 2; p_1, p_2, p_3)$ 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 the three-cell partition $[\{3\}, \{1, 4\}, \{2, 5\}]$ of the numbers from 1 to 5. Here the first cell shows the experiment which resulted in a_1, the second cell shows the two that resulted in a_2, and the third shows the two that resulted in a_3. Conversely, any such partition of the numbers from 1 to 5 with one element in the first cell, two in the second, and two in the third corresponds to a unique path of the desired kind. Hence the number of paths is the number of such partitions. But this is

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

(see 3.4), 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

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

Example 4.22 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

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

which is

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

$\diamondsuit$

Example 4.23 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 of a_2, and $\bar{r_3}$ of $\bar{a_3}$, which is

$$\binom{n}{r_1, r_2, \bar{r_3}} p_1^{r_1} \cdot p_2^{r_2} \cdot \bar{p_3}^{\bar{r_3}}.$$

$\diamondsuit$

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

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

Exercises

1. Suppose that in a city 60 per cent of the population are Democrats, 30 per cent are Republicans, and 10 per cent are Independents. What is the probability that if three people are chosen at random there will be one Republican, one Democrat, and one Independent voter?

   [Ans. .108.]

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. Assume that in a certain large college 40 per cent of the students are freshmen, 30 per cent are sophomores, 20 per cent are juniors, and 10 per cent 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?

4. Let us assume that when a batter comes to bat, he or she 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 or she comes to bat five times in a game, what is the probability that

   (a) He gets two walks and three singles?

   [Ans. .0008.]

   (b) He gets a walk, a single, an extra base hit (and is out twice)?

   [Ans. .043.]

   (c) He has a perfect day (i.e., never out)?

   [Ans. .010.]

5. 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 sink the ship. What is the probability that four torpedos will succeed in sinking the ship?

[Ans. $\frac{251}{256}$.]

6. Jones, Smith, and Green live in the same house. The mailman has observed that Jones and Smith receive the same amount of mail on the average, but that Green receives twice as much as Jones (and hence also twice as much as Smith). If he or she has four letters for this house, what is the probability that each resident receives at least one letter?

7. If three dice are thrown, find the probability that there is one six and two fives, given that all the outcomes are greater than three.

[Ans. $\frac{1}{9}$.]

8. An athlete plays a tournament consisting of three games. In each game he or she has probability $\frac{1}{2}$ for a win, $\frac{1}{4}$ for a loss, and $\frac{1}{4}$ for a draw, independently of the outcomes of other games. To win the tournament he or she must win more games than he or she loses. What is the probability that he or she wins the tournament?

9. Assume that in a certain course the probability that a student chosen at random will get an A is .1, that he or she will get a B is .2, that he or she will get a C is .4, that he or she will get a D is .2, and that he or she will get an F is .1. What distribution of grades is most likely in the case of four students?

[Ans. One B, two C's, one D.]

10. Let us assume that in a World Series game a batter has probability $\frac{1}{4}$ of getting no hits, $\frac{1}{2}$ for getting one hit, $\frac{1}{4}$ for 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.

    [Ans. $\frac{7}{64}$.]

    (b) Exactly three hits.

    [Ans. $\frac{7}{32}$.]

    (c) Exactly four hits.

[Ans. $\frac{35}{128}$.]

(d) Exactly five hits.

[Ans. $\frac{7}{32}$.]

(e) Fewer than two hits or more than five.

[Ans. $\frac{23}{128}$.]

11. 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 one head and four tails in five tosses of a gypsy coin?

12. A family car is driven by the father, two sons, and the mother. The fenders have been dented four times, three times while the mother was driving. Is it fair to say that the mother is a worse driver than the men?

4.12 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 or she 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 4.24 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 of \$1, on a given number, and the ball comes to rest in this slot, then he or she receives from the croupier 36 times the stake, or \$36. The player wins \$35 with probability $\frac{1}{37}$ and loses \$1 with probability $\frac{36}{37}$. Hence his or her expected winnings are

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

This can be interpreted to mean that in the long run the player can expect to lose about 2.7 per cent of his or her 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, the player receives twice the stake. If a black number turns up, the player loses the 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, the player receives only the original stake, not twice it. For this type of play, the player wins \$1 with probability $\frac{18}{37}$, breaks even with probability $\frac{1}{2} \cdot \frac{1}{37}$, and loses \$1 with probability $\frac{18}{37} + \frac{1}{2} \cdot \frac{1}{37}$. Hence his or her 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 per cent of his or her stakes in the long run. Thus the expected loss in this case is only half as great as in the previous case. ♢

Example 4.25 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, the player's expected winning is

$$1(\frac{1}{6}) + 2(\frac{1}{6}) + 3(\frac{1}{6}) + 4(\frac{1}{6}) + 5(\frac{1}{6}) + 6(\frac{1}{6}) = 3.5.$$

Thus if the player pays \$3.50, the expected winnings will be zero. $\diamondsuit$

Example 4.26 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 $\binom{4}{x}(\frac{1}{3})^x(\frac{2}{3})^{4-x}$. Thus

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

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

Example 4.27 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 or she wins. If it is 2, 3, or 12, he or she loses. If it is another sum, he or she must continue rolling the dice until he or she either repeats the same sum or rolls a 7. In the former case he or she wins, in the latter he or she loses. Let us suppose that he or she wins or loses \$1. Then the two possible outcomes are $+1$ and -1. We will compute the expected value of the game. First we must find the probability that he or she will win.

We represent the possibilities by a two-stage tree shown in Figure 4.21. 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

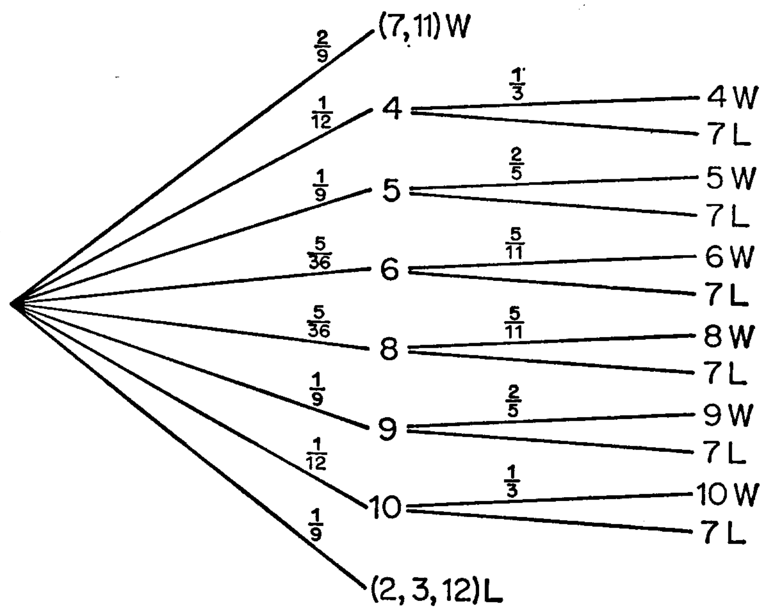


Figure 4.21: $\diamondsuit$

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 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

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

Hence the player can expect to lose 1.41 per cent of his or her stakes in the long run. It is interesting to note that this is just slightly less favorable than the losses in betting on "red" in roulette. $\diamondsuit$

Exercises

1. Suppose that A tosses two coins and receives \$2 if two heads appear, \$1 if one head appears, and nothing if no heads appear. What is the expected value of the game to A?

   [Ans. \$1.]

2. Smith and Jones are matching coins. If the coins match, Smith gets \$1, and if they do not, Jones get \$1.

   (a) If the game consists of matching twice, what is the expected value of the game for Smith?

   (b) Suppose that Smith quits if he or she wins the first round he or she quits, and plays the second round if he or she loses the the first. Jones is not allowed to quit. What is the expected value of the game for Smith?

3. If five coins are thrown, what is the expected number of heads that will turn up?

   [Ans. $\frac{5}{2}$.]

4. A coin is thrown until the first time a head comes up or until three tails in a row occur. Find the expected number of times the coin is thrown.

5. A customer wishes to purchase a five cent newspaper. The customer has in his or her pocket one dime and five pennies. The news agent offers to let the customer 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 customer.]

   (b) Answer the same question assuming that the news agent demands two coins drawn at random from the customer's pocket.

   [Ans. Fair proposition.]

6. A bets 50 cents against B's x 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?

7. 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 $E + c$.

8. 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 $k \cdot E$.

9. A gambler plays the following game: A card is drawn from a bridge deck; if it is an ace, the gambler wins $5; if it is a jack, a queen or a king, he or she wins $2; for any other card he or she loses $1. What is the expected winning per play?

10. An urn contains two black and three white balls. Balls are successively drawn from the urn without replacement until a black ball is obtained. Find the expected number of draws required.

11. Using the result of Exercises 13 and 14 of Section 4.6, find the expected number of games in the World Series (a) under the assumption that each team has probability 1 of winning each game and (b) under the assumption that the stronger team has probability .6 of winning each game.

    [Ans. 5.81; 5.75.]

12. Suppose that we modify the game of craps as follows: On a 7 or 11 the player wins $2, on a 2, 3, or 12 he or she loses $3; otherwise the game is as usual. Find the expected value of the new game, and compare it with the old value.

13. 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"?

14. 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.]

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

16. Referring to Exercise 9 of Section 4.3, find the expected number of languages that a student chosen at random reads.

17. Referring to Exercise 5 of Section 4.4, find the expected number of men who get their own hats.

    [Ans. 1.]

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 if

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

    (ii) the sum of three or five occurs on the first roll and the same sum occurs on a subsequent roll before the sum of four occurs.

    Otherwise he or she loses a dollar.

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

    [Ans. $\frac{23}{45}$.]

    (b) What is the expected value of the game?

    [Ans. $\frac{1}{45}$.]

4.13 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

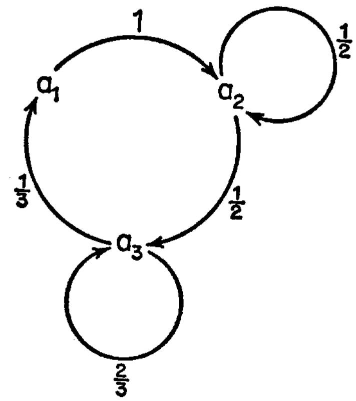


Figure 4.22: $\diamondsuit$

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 over-all 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 **??**.

A second way to show the transition probabilities is by a transition diagram. Such a diagram is illustrated for a special case in Figure 4.22. The arrows from each state indicate the possible states to which a process can move from the given state.

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

$$\begin{array}{cc} & \begin{array}{ccc} a_1 & a_2 & a_3 \end{array} \\ \begin{array}{c} a_1 \\ a_2 \\ a_3 \end{array} & \begin{pmatrix} 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{2}{3} \end{pmatrix} \end{array}.$$

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 4.28 Let us find for a Markov chain with transition probabilities indicated in Figure 4.22 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 4.23.

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,

$$p_{13}^{(3)} = 1 \cdot \frac{1}{2} \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} \cdot \frac{2}{3} = \frac{7}{12}.$$

Similarly

$$p_{12}^{(3)} = 1 \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

and

$$p_{11}^{(3)} = 1 \cdot \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}.$$

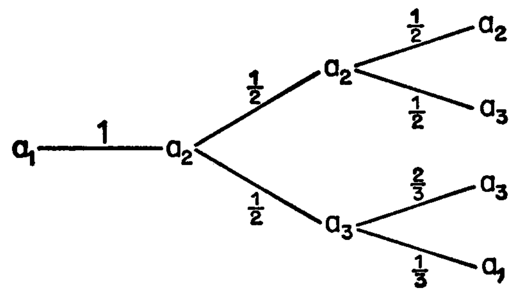


Figure 4.23: $\diamondsuit$

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:

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

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 will develop a simple method of computing $P^{(n)}$. $\diamondsuit$

Example 4.29 example:4.13.2 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 will not consider conditions peculiar to a particular election year. We shall base our predictions only on the 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

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

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 form,

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

Again the numbers a, b, c, and d would have to be estimated. The study of this example is continued in **??**. ♢

Example 4.30 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)

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

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 **??**, that this is indeed the case. $\diamondsuit$

Exercises

1. Draw a state diagram for the Markov chain with transition probabilities given by the following matrices.

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix},$$

$$\begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix},$$

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

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

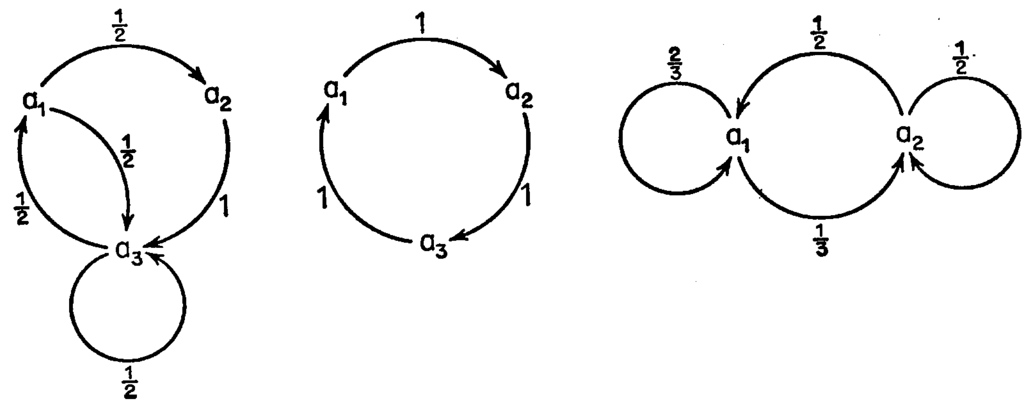


Figure 4.24: $\diamondsuit$

2. Give the matrix of transition probabilities corresponding to the transition diagrams in Figure 4.24.

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

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

[Ans. $\begin{pmatrix} \frac{5}{12} & \frac{7}{12} \\ \frac{7}{18} & \frac{11}{18} \end{pmatrix}$.]

4. What is the matrix of transition probabilities for the Markov chain in Example 4.30, for the case of two white balls and two black balls?

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

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

Find the same for the Markov chain determined by the matrix

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

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

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

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. $\frac{5}{12}$; $\frac{4}{9}$.]

7. Referring to the Markov chain with transition probabilities indicated in Figure 4.22, construct the tree measures and determine the values of

$$p_{21}^{(3)}, p_{22}^{(3)}, p_{23}^{(3)}$$

and

$$p_{31}^{(3)}, p_{32}^{(3)}, p_{33}^{(3)}.$$

8. 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?

9. For the Markov chain in Exercise 8, draw a tree and assign a tree measure, assuming that the process begins in state 0 and moves through three stages of transmission. What is the probability that the machine after three stages produces the digit 0, i.e., the correct digit? What is the probability that the machine never changed the digit from 0?

10. 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 per cent are professional, 10 per cent are skilled laborers, and 10 per cent are unskilled laborers. In the

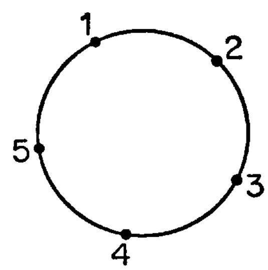


Figure 4.25: ♢

case of sons of skilled laborers, 60 per cent are skilled laborers, 20 per cent are professional, and 20 per cent are unskilled laborers. Finally, in the case of unskilled laborers, 50 per cent of the sons are unskilled laborers, and 25 per cent 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 an unskilled laborer is a professional man.

[Ans. .375.]

11. In Exercise 10 we assumed that every man has a son. Assume instead that the probability a man has a son is .8. Form a Markov chain with four states. The first three states are as in Exercise 10, 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 an unskilled laborer has a grandson who is a professional man.

[Ans. .24.]

12. Explain why the transition probabilities given in Example 4.30 are correct.

Supplementary exercises.

13. Five points are marked on a circle. A process moves clockwise from a given point to its neighbor with probability $\frac{2}{3}$, or counterclockwise to its neighbor with probability $\frac{1}{3}$.

(a) Considering the process to be a Markov chain process, find the matrix of transition probabilities.

(b) Given that the process starts in a state 3, what is the probability that it returns to the same state in two steps?

14. In northern New England, years for apples can be described as good, average, or poor. Suppose that following a good year the probabilities of good, average, or poor years are respectively .4, .4, and .2. Following a poor year the probabilities of good, average, or poor years are .2, .4, and .4 respectively. Following an average year the probabilities that the next year will be good or poor are each .2, and of an average year, .6.

    (a) Set up the transition matrix of this Markov chain.

    (b) 1965 was a good year. Compute the probabilities for 1966, 1967, and 1968.

    [Ans. For 1967: .28, .48, .24.]

15. In Exercise 14 suppose that there is probability $\frac{1}{4}$ for a good year, $\frac{1}{2}$ for an average year, and $\frac{1}{4}$ for a poor year. What are the probabilities for the following year?

16. A teacher in an oversized mathematics class finds, after grading all homework papers for the first two assignments, that it is necessary to reduce the amount of time spent in such grading. He therefore designs the following system: Papers will be marked satisfactory or unsatisfactory. All papers of students receiving a mark of unsatisfactory on any assignment will be read on each of the two succeeding days. Of the remaining papers, the teacher will read one-fifth, chosen at random. Assuming that each paper has a probability of one-fifth of being classified "unsatisfactory",

    (a) Set up a three-state Markov chain to describe the process.

    (b) Suppose that a student has just handed in a satisfactory paper. What are the probabilities for the next two assignments?

17. 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. Set

up a Markov chain by taking as state the number of balls in urn 1. Find the transition matrix.

4.14 The central limit theorem

We continue our discussion of the independent trials process with two outcomes. As usual, let p be the probability of success on a trial, and $f(n,p;x)$ be the probability of exactly x successes in n trials.

In Figure 4.26 we have plotted bar graphs which represent $f(n,.3;x)$ for $n = 10, 50, 100$, and 200. We note first of all that the graphs are drifting off to the right. This is not surprising, since their peaks occur at np, which is steadily increasing. We also note that while the total area is always 1, this area becomes more and more spread out.

We want to redraw these graphs in a manner that prevents the drifting and the spreading out. First of all, we replace x by $x - np$, assuring that our peak always occurs at 0. Next we introduce a new unit for measuring the deviation, which depends on n, and which gives comparable scales. As we saw in Section 4.10, the standard deviation $\sqrt{npq}$ is such a unit.

We must still insure that probabilities are represented by areas in the graph. In Figure 4.26 this is achieved by having a unit base for each rectangle, and having the probability $f(n,p;x)$ as height. Since we are now representing a standard deviation as a single unit on the horizontal axis, we must take $f(n,p;x)\sqrt{npq}$ as the heights of our rectangles. The resulting curves for $n = 50$ and 200 are shown in Figures 4.27 and 4.28, respectively.

We note that the two figures look very much alike. We have also shown in Figure 4.28 that it can be approximated by a bell-shaped curve. This curve represents the function

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

and is known as the normal curve. It is a fundamental theorem of probability theory that as n increases, the appropriately rescaled bar graphs more and more closely approach the normal curve. The theorem is known as the Central Limit Theorem, and we have illustrated it graphically.

More precisely, the theorem states that for any two numbers a and

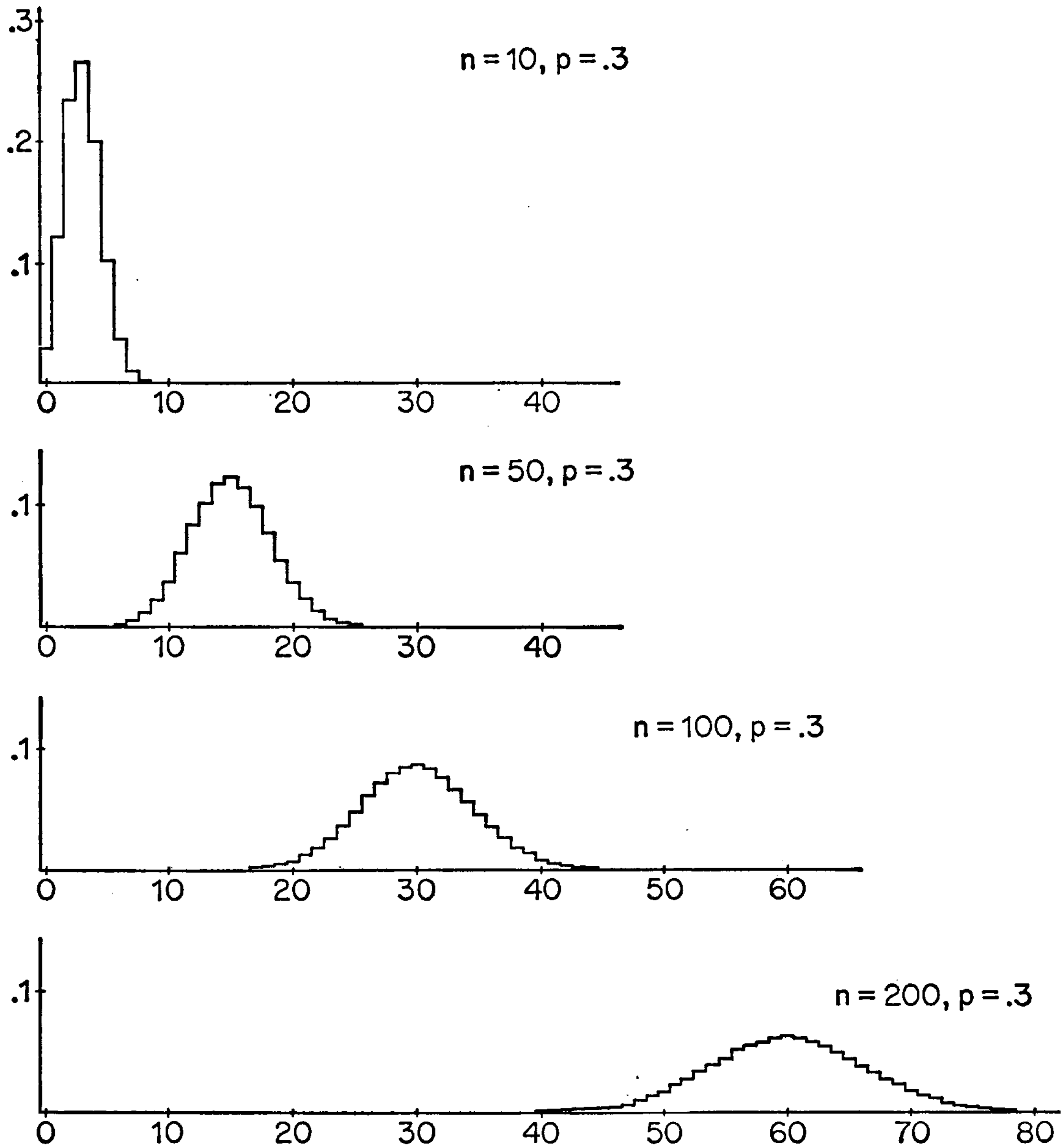


Figure 4.26: $\diamondsuit$

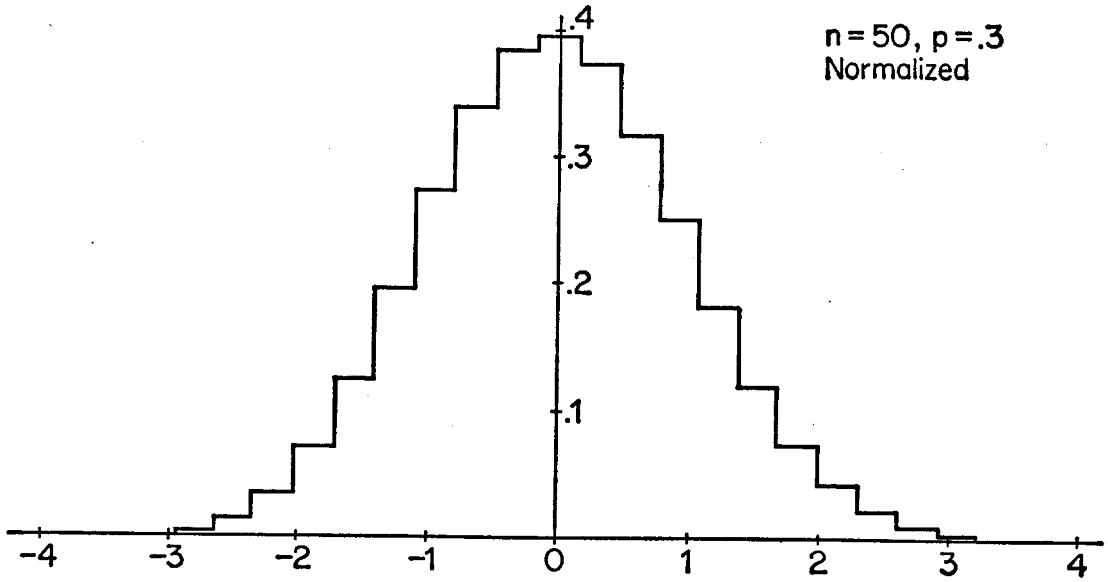


Figure 4.27: ♢

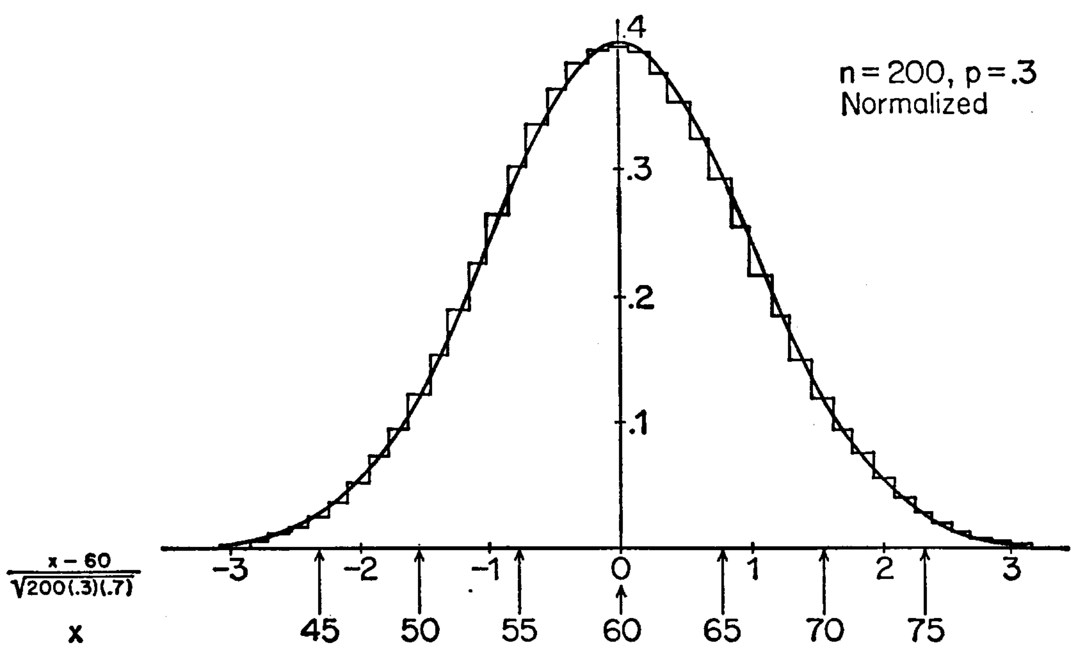
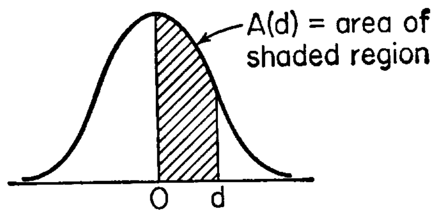


Figure 4.28: ♢



d	A(d)	d	A(d)	d	A(d)	d	A(d)
.0	.000	1.1	.364	2.1	.482	3.1	.4990
.1	.040	1.2	.385	2.2	.486	3.2	.4993
.2	.079	1.3	.403	2.3	.489	3.3	.4995
.3	.118	1.4	.419	2.4	.492	3.4	.4997
.4	.155	1.5	.433	2.5	.494	3.5	.4998
.5	.191	1.6	.445	2.6	.495	3.6	.4998
.6	.226	1.7	.455	2.7	.497	3.7	.4999
.7	.258	1.8	.464	2.8	.497	3.8	.49993
.8	.288	1.9	.471	2.9	.498	3.9	.49995
.9	.316	2.0	.477	3.0	.4987	4.0	.49997
1.0	.341					5.0	.49999997

Figure 4.29: ♢

b, with $a < b$,

$$\Pr[a < \frac{x - np}{\sqrt{npq}} < b]$$

approaches the area under the normal curve between a and b, as n increases. This theorem is particularly interesting in that the normal curve is symmetric about 0, while $f(n, p; x)$ is symmetric about the expected value np only for the case $p = \frac{1}{2}$. It should also be noted that we always arrive at the same normal curve, no matter what the value of p is.

In Figure 4.29 we give a table for the area under the normal curve between 0 and d. Since the total area is 1, and since it is symmetric about the origin, we can compute arbitrary areas from this table. For example, suppose that we wish the area between -1 and $+2$. The area between 0 and 2 is given in the table as .477. The area between -1 and 0 is the same as between 0 and 1, and hence is given as .341. Thus the total area is .818. The area outside the interval $(-1, 2)$ is then

$1 - .818 = .182.$

Example 4.31 Let us find the probability that x differs from the expected value np by as much as d standard deviations.

$$\Pr[|x - np| \geq d\sqrt{npq}] = \Pr[|\frac{x - np}{\sqrt{npq}} \geq d].$$

and hence the approximate answer should be the area outside the interval $(-d, d)$ under the normal curve. For $d = 1, 2, 3$ we obtain $1 - (2 \cdot .341) = .318$, $1 - (2 \cdot .477) = .046$, $1 - (2 \cdot .4987) = .0026$, respectively. These agree with the values given in Section 4.10, to within rounding errors. In fact, the Central Limit Theorem is the basis of those estimates. $\diamondsuit$

Example 4.32 In Example 4.19 we considered the example of throwing a coin 10,000 times. The expected number of heads that turn up is 5000 and the standard deviation is $\sqrt{10,000} \cdot \frac{1}{2} \cdot \frac{1}{2} = 50$. We observed that the probability of a deviation of more than two standard deviations (or 100) was very unlikely. On the other hand, consider the probability of a deviation of less than .1 standard deviation. That is, of a deviation of less than five. The area from 0 to .1 under the normal curve is .040 and hence the probability of a deviation from 5000 of less than five is approximately .08. Thus, while a deviation of 100 is very unlikely, it is also very unlikely that a deviation of less than five will occur. $\diamondsuit$

Example 4.33 The normal approximation can be used to estimate the individual probabilities $f(n, x; p)$ for large n. For example, let us estimate $f(200, 65; .3)$. The graph of the probabilities $f(200, x; .3)$ was given in Figure 4.28 together with the normal approximation. The desired probability is the area of the bar corresponding to $x = 65$. An inspection of the graph suggests that we should take the area under the normal curve between 64.5 and 65.5 as an estimate for this probability. In normalized units this is the area between

$$\frac{4.5}{\sqrt{200}(.3)(.7)}$$

and

$$\frac{5.5}{\sqrt{200}(.3)(.7)}$$

or between .6944 and .8487. Our table is not fine enough to find this area, but from more complete tables, or by machine computation, this area may be found to be .046 to three decimal places. The exact value to three decimal places is .045. This procedure gives us a good estimate.

If we check all of the values of $f(200, x; .3)$ we find in each case that we would make an error of at most .001 by using the normal approximation. There is unfortunately no simple way to estimate the error caused by the use of the Central Limit Theorem. The error will clearly depend upon how large n is, but it also depends upon how near p is to 0 or 1. The greatest accuracy occurs when p is near $\frac{1}{2}$. $\diamondsuit$

Example 4.34 Suppose that a drug has been administered to a number of patients and found to be effective a fraction p of the time. Assuming an independent trials process, it is natural to take p as an estimate for the unknown probability p for success on any one trial. It is useful to have a method of estimating the reliability of this estimate. One method is the following. Let x be the number of successes for the drug given to n patients. Then by the Central Limit Theorem

$$\Pr[|\frac{x - np}{\sqrt{npq}}| \leq 2] \approx .95.$$

This is the same as saying

$$\Pr[|\frac{x/n - p}{\sqrt{pq/n}}| \leq 2] \approx .95.$$

Putting $\bar{p} = x/n$, we have

$$\Pr[|\bar{p} - p| \leq 2\sqrt{pq/n}] \approx .95.$$

Using the fact that $pq < 4$ (see Exercise 12) we have

$$\Pr[|\bar{p} - p| \leq \frac{1}{\sqrt{n}}] \geq .95.$$

This says that no matter what p is, with probability $\geq .95$, the true value will not deviate from the estimate p by more than $\frac{1}{\sqrt{n}}$ It is customary then to say that

$$\bar{p} - \frac{1}{\sqrt{n}} \leq p \leq \bar{p} + \frac{1}{\sqrt{n}}$$

with confidence .95. The interval

$$[\bar{p} - \frac{1}{\sqrt{n}}, \bar{p} + \frac{1}{\sqrt{n}}]$$

is called a 95 per cent confidence interval. Had we started with

$$\Pr[|\frac{x - np}{\sqrt{npq}}| \leq 3] \approx .99,$$

we would have obtained the 99 per cent confidence interval

$$[\bar{p} - \frac{3}{2\sqrt{n}}, \bar{p} + \frac{3}{2\sqrt{n}}]$$

For example, if in 400 trials the drug is found effective 124 times, or .31 of the times, the 95 per cent confidence interval for p is

$$[.31 - \frac{1}{20}, .31 + \frac{1}{20}] = [.26, .36]$$

and the 99 per cent confidence interval is

$$[.31 - \frac{3}{40}, .31 + \frac{3}{40}] = [.235, .385].$$

◊

Exercises

1. Let x be the number of successes in n trials of an independent trials process with probability p for success. Let $x^\star = \frac{x-np}{\sqrt{npq}}$ For large n estimate the following probabilities.

    (a) $\Pr[x^\star < -2.5]$.

    [Ans. .006.]

    (b) $\Pr[x^\star < 2.5]$.

    (c) $\Pr[x^\star \geq -.5]$.

    (d) $\Pr[-1.5 < x^\star < 1]$.

    [Ans. .774.]

2. A coin is biased in such a way that a head comes up with probability .8 on a single toss. Use the normal approximation to estimate the probability that in a million tosses there are more than 800,400 heads.

3. Plot a graph of the probabilities $f(10, x, .5)$. Plot a graph also of the normalized probabilities as in Figures 4.27 and 4.28.

4. An ordinary coin is tossed one million times. Let x be the number of heads which turn up. Estimate the following probabilities.

   (a) $\Pr[499,500 < x < 500,500]$

   [Ans. Approximately .682.]

   (b) $\Pr[499,000 < x < 501,000]$,

   [Ans. Approximately .954.]

   (c) $\Pr[498,500 < x < 501,500]$,

   [Ans. Approximately .997.]

5. Assume that a baseball player has probability .37 of getting a hit each time he or she comes to bat. Find the probability of getting an average of .388 or better if he or she comes to bat 300 times during the season. (In 1957 Ted Williams had a batting average of .388 and Mickey Mantle had an average of .353. If we assume this difference is due to chance, we may estimate the probability of a hit as the combined average, which is about .37.)

   [Ans. .242.]

6. A true-false examination has 48 questions. Assume that the probability that a given student knows the answer to any one question is $\frac{3}{4}$. A passing score is 30 or better. Estimate the probability that the student will fail the exam.

7. In Example 4.21 of Section 4.10, assume that the school decides to admit 1296 students. Estimate the probability that they will have to have additional dormitory space.

   [Ans. Approximately .115.]

8. Peter and Paul each have 20 pennies. They agree to match pennies 400 times, keeping score but not paying until the 400 matches are over. What is the probability that one of the players will not be able to pay? Answer the same question for the case that Peter has 10 pennies and Paul has 30.

9. In tossing a coin 100 times, the probability of getting 50 heads is, to three decimal places, .080. Estimate this same probability using the Central Limit Theorem.

   [Ans. .080.]

10. A standard medicine has been found to be effective in 80 per cent of the cases where it is used. A new medicine for the same purpose is found to be effective in 90 of the first 100 patients on which the medicine is used. Could this be taken as good evidence that the new medication is better than the old?

11. In the Weldon dice experiment, 12 dice were thrown 26,306 times and the appearance of a 5 or a 6 was considered to be a success. The mean number of successes observed was, to four decimal places, 4.0524. Is this result significantly different from the expected average number of 4?

    [Ans. Yes.]

12. Prove that $pq \leq \frac{1}{4}$. [Hint: write $p = \frac{1}{2} + x$.]

13. Suppose that out of 1000 persons interviewed 650 said that they would vote for Mr. Big for mayor. Construct the 99 per cent confidence interval for p, the proportion in the city that would vote for Mr. Big.

14. Opinion pollsters in election years usually poll about 3000 voters. Suppose that in an election year 51 per cent favor candidate A and 49 per cent favor candidate B. Construct 95 per cent confidence limits for candidate A winning.

    [Ans. $[.492, .528]$.]

15. In an experiment with independent trials we are going to estimate p by the fraction p of successes. We wish our estimate to be within .02 of the correct value with probability .95. Show that 2500 observations will always suffice. Show that if it is known that p is approximately .1, then 900 observations would be sufficient.

16. An experimenter has an independent trials process and he or she has a hypothesis that the true value of p is p_0. He decides to carry out a number of trials, and from the observed r calculate the 95 per cent confidence interval for p. He will reject p_0 if it does not fall within these limits. What is the probability that he or she will reject p_0 when in fact it is correct? Should he or she accept p_0 if it does fall within the confidence interval?

17. A coin is tossed 100 times and turns up heads 61 times. Using the method of Exercise 16 test the hypothesis that the coin is a fair coin.

    [Ans. Reject.]

18. Two railroads are competing for the passenger traffic of 1000 passengers by operating similar trains at the same hour. If a given passenger is equally likely to choose one train as the other, how many seats should the railroad provide if it wants to be sure that its seating capacity is sufficient in 99 out of 100 cases?

    [Ans. 537.]

4.15 Gambler's ruin

In this section we will 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 will describe the chain from the gambling point of view, but in the exercises we will 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 in his or her favor, 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 will develop a formula for the ruin-probability in terms of these three given numbers.

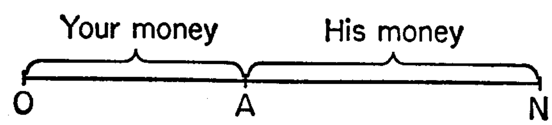


Figure 4.30: $\diamondsuit$

First we will 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 4.30.

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 will 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 $x_0, x_1, x_2, x_3, x_4, x_5$. Suppose that we start at position 2. The chain moves to 3, with probability p, or to 1, with probability q. Thus

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

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

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

And, similarly,

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

We obtain the key relation

$$x_2 = px_3 + qx_1.$$

We can modify this as follows:

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

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

$$\begin{aligned}x_0 - x_1 &= r(x_1 - x_2),\\ x_1 - x_2 &= r(x_2 - x_3),\\ x_2 - x_3 &= r(x_3 - x_4),\\ x_3 - x_4 &= r(x_4 - x_5)\end{aligned} \tag{4.4}$$

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

$$x_0 = 1,\ x_5 = 0. \tag{4.5}$$

If we substitute the value of x_5 in the last equation of 4.4, 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

$$\begin{aligned}x_4 &= 1 \cdot x_4,\\ x_3 - x_4 &= rx_4.\\ x_2 - x_3 &= r^2x_4.\\ x_1 - x_2 &= r^3x_4.\\ x_0 - x_1 &= r^4x_4\end{aligned} \tag{4.6}$$

Let us add all the equations. We obtain

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

From 4.5 we have that $x_0 = 1$. We also use the simple identity

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

And then we solve for x_4:

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

If we add the first two equations in 4.6, we have that $x_3 = (1+r)x_4$. Similarly, adding the first three equations, we solve for x_2, and adding the first four equations we obtain x_1. We now have our entire solution,

$$x_1 = \frac{1-r^4}{1-r^5},\ x_2 = \frac{1-r^3}{1-r^5},\ x_3 = \frac{1-r^2}{1-r^5},\ x_4 = \frac{1-r^1}{1-r^5}. \tag{4.7}$$

The same method will work for any value of N. And it is easy to guess from 4.7 what the general solution looks like. If we want x_A, the answer is a fraction like those in 4.7. In the denominator the exponent of r is always N. In the numerator the exponent is $N-A$, or B. Thus the ruin-probability is

$$x_A = \frac{1-r^B}{1-r^N}. \tag{4.8}$$

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 4.31 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 dollars 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 4.8 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 or her a gambler under those circumstances!) The gambler 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 $p = .48$, the gambler needs only \$87. And even for the almost fair game with $p = .499$, \$1727 will suffice.

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 4.31: ♢

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. An urn has nine white balls and 11 black balls. A ball is drawn, and replaced. If it is white, you win five cents, if black, you lose five cents. You have a dollar to gamble with, and your opponent has fifty cents. If you keep on playing till one of you loses all his or her money, what is the probability that you will lose your dollar?

   [Ans. .868.]

2. Suppose that you are shooting craps, and you always hold the dice. You have \$20, your opponent has \$10, and \$1 is bet on each game; estimate your probability of ruin.

3. Two government agencies, A and B, are competing for the same task. A has 50 positions, and B has 20. Each year one position is taken away from one of the agencies, and given to the other. If 52 per cent of the time the shift is from A to B, what do you predict for the future of the two agencies?

   [Ans. One agency will be abolished. B survives with probability .8, A with probability .2.]

4. What is the approximate value of x_A if you are rich, and the gambler starts with \$1?

5. 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 4.]

[Ans. $\frac{1}{3}$.]

6. Verify that the proof of formula 4.8 in the text is still correct when $p > \frac{1}{2}$. Interpret formula 4.8 for this case.

7. Show that if $p > \frac{1}{2}$, and both parties have a substantial amount of money, your probability of ruin is approximately $1/r^A$.

8. 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.]

9. You are matching pennies. You have 25 pennies to start with, and your opponent has 35. What is the probability that you will win all his or her pennies?

10. Jones lives on a short street, about 100 steps long. At one end of the street is Jones's home, at the other a lake, and in the middle a bar. One evening Jones leaves the bar in a state of intoxication, and starts to walk at random. What is the probability that Jones will fall into the lake if

    (a) Jones is just as likely to take a step to the right as to the left?

    [Ans. $\frac{1}{2}$.]

    (b) Jones has probability .51 of taking a step towards home?

    [Ans. .119.]

11. You are in the following hopeless situation: You are playing a game in which you have only $\frac{1}{3}$ chance of winning. You have \$1, and your opponent has \$7. What is the probability of your winning all his or her money if

    (a) You bet \$1 each time?

    [Ans. $\frac{1}{255}$.]

    (b) You bet all your money each time?

    [Ans. $\frac{1}{27}$.]

12. Repeat Exercise 11 for the case of a fair game, where you have probability $\frac{1}{2}$ of winning.

13. Modify the proof in the text to compute y_i, the probability of reaching state $N = 5$.

14. Verify, in Exercise 13, that $x_i + y_i = 1$ for every state. Interpret.

    Note: The following exercises deal with the following ruin problem: A and B play a game in which A has probability W of winning. They keep playing until either A has won six times or B has won three times.

15. Set up the process as a Markov chain whose states are (a, b), where a is the number of times A won, and b the number of B wins.

16. For each state compute the probability of A winning from that position. [Hint: Work from higher a- and b-values to lower ones.]

17. What is the probability that A reaches his or her goal first?

    [Ans. $\frac{1024}{2187}$.]

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

Suggested reading.

Cramer, Harald, *The Elements of Probability Theory, Part I*, 1955.

Feller, W., *An Introduction to Probability Theory and its Applications*, 1950.

Goldberg, S., *Probability: An Introduction*, 1960.

Mosteller, F., *Fifty Challenging Problems in Probability with Solutions*, 1965.

Neyman, J., *First Course in Probability and Statistics*, 1950.

Parzen, E., *Modern Probability Theory and Its Applications*, 1960.

Whitworth, W. A., *Choice and Chance, with 1000 Exercises*, 1934.