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

$$E = a_1 p_1 + a_2 p_2 + \ldots + a_k p_k.$$

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. $\diamond$

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. $\diamond$

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

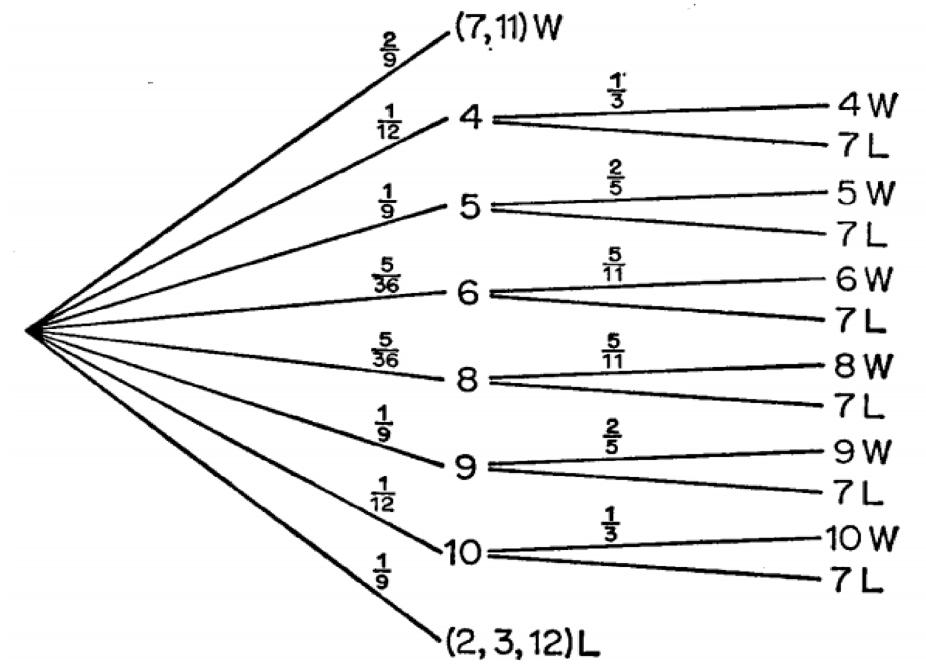


Figure 4.21: ◊

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. ◊

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?

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

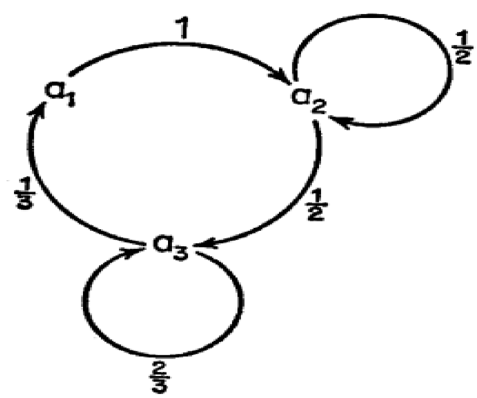


Figure 4.22: ◇

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.

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

$$\begin{array}{c} \\ a_1 \\ a_2 \\ a_3 \end{array} \begin{array}{c} \begin{array}{ccc} a_1 & a_2 & a_3 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{2}{3} \end{pmatrix} \end{array}.$$

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

$$p_{11}^{(3)} = 1 \cdot \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}.$$

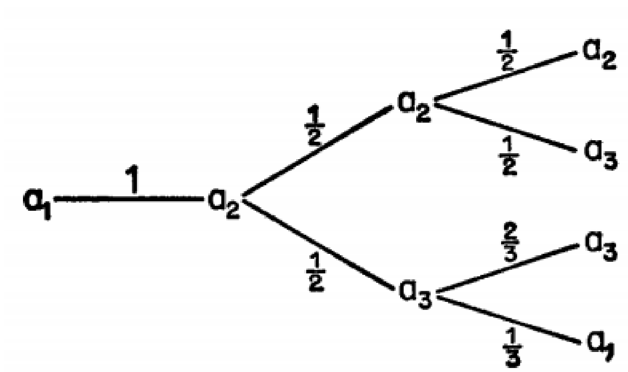


Figure 4.23: ◇

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:

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

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 $P^{(n)}$. ◇

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

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

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,

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

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)

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

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. ◇

Exercises

1. Draw a state diagram for the Markov chain with transition probabilities given by the following matrices.

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix},$$

$$\begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix},$$

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

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

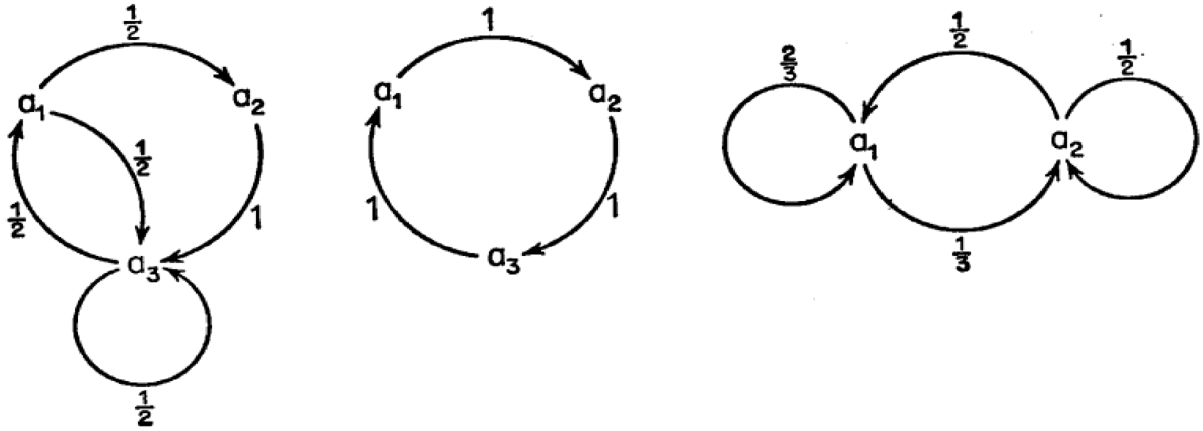


Figure 4.24: ♢

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

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

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

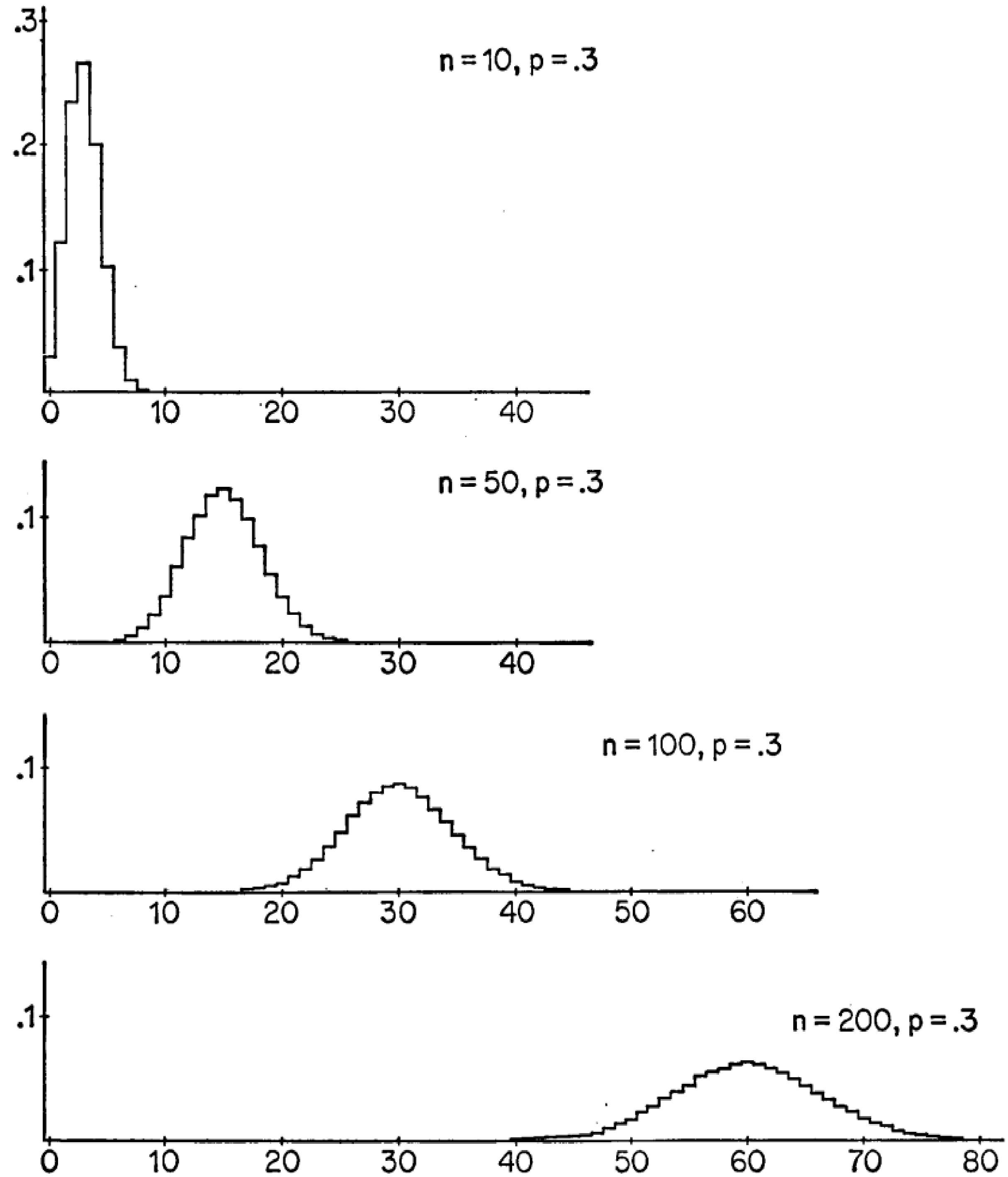


Figure 4.26: $\diamondsuit$

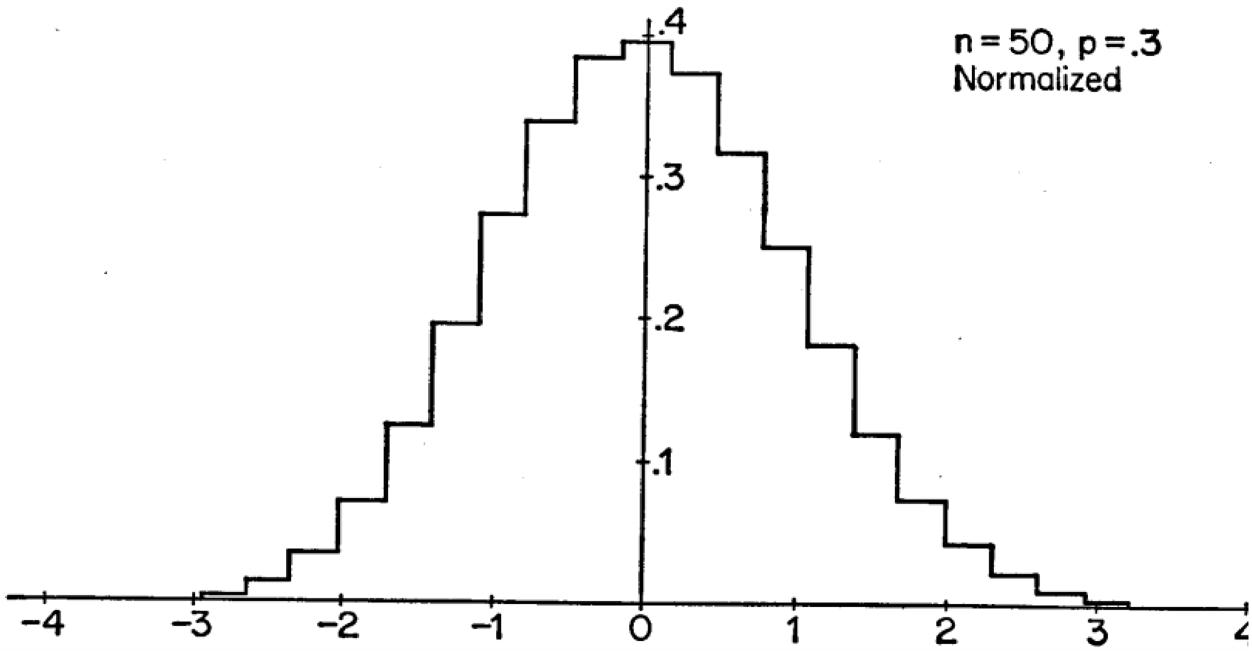


Figure 4.27: ♢

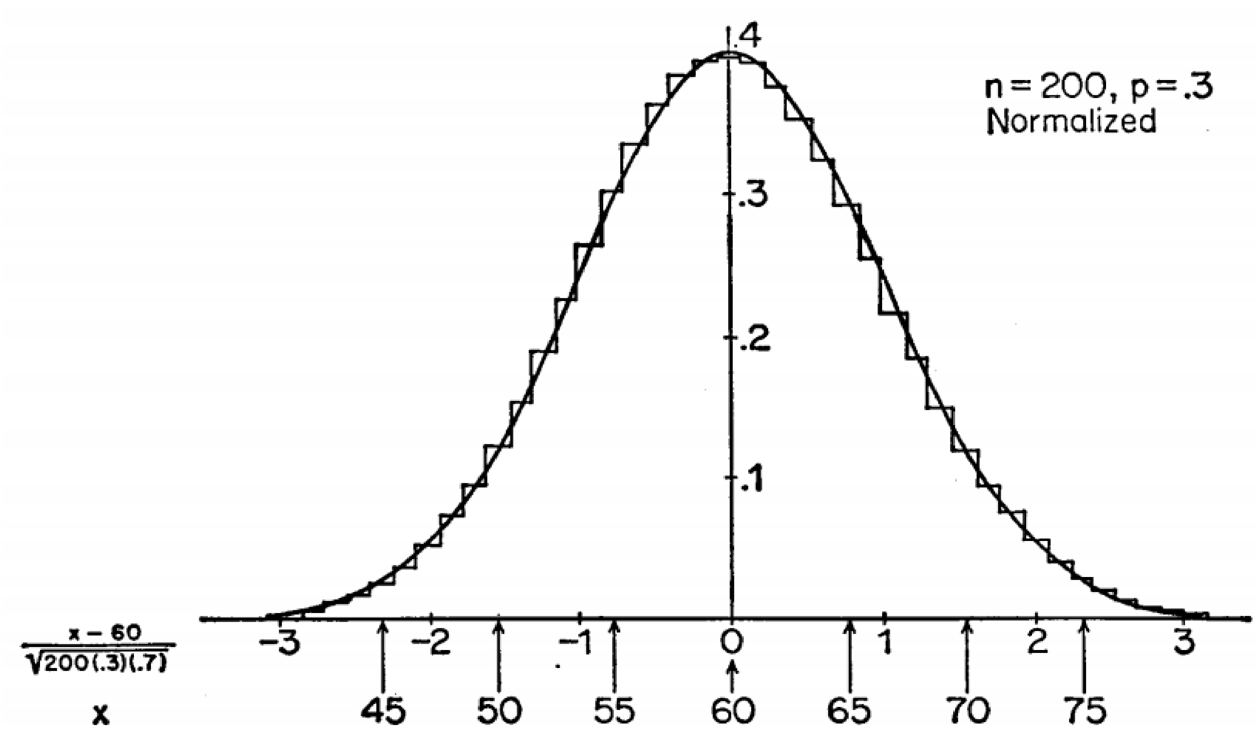
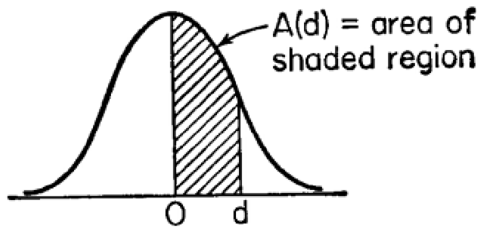


Figure 4.28: ♢



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 $\frac{1}{2}$. ◇

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.]

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.

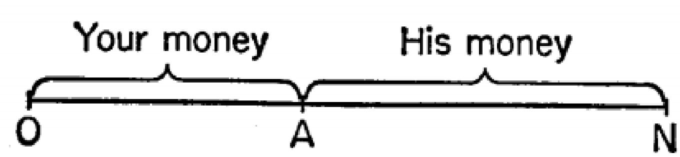


Figure 4.30: ◇

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

$$\begin{aligned} x_4 &= 1 \cdot x_4, \\ x_3 - x_4 &= rx_4. \\ x_2 - x_3 &= r^2 x_4. \\ x_1 - x_2 &= r^3 x_4. \\ x_0 - x_1 &= r^4 x_4 \end{aligned} \tag{4.6}$$

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: $\diamondsuit$

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?

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.