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. ◇

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 $\frac{1}{2}$. ◇

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. ◇

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. $\diamond$

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?

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

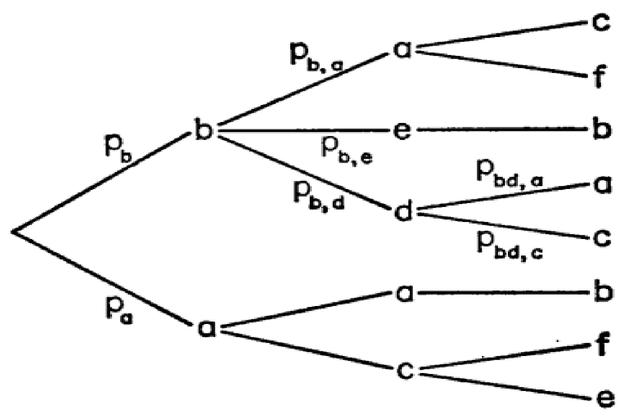

Figure 4.3: ◇

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

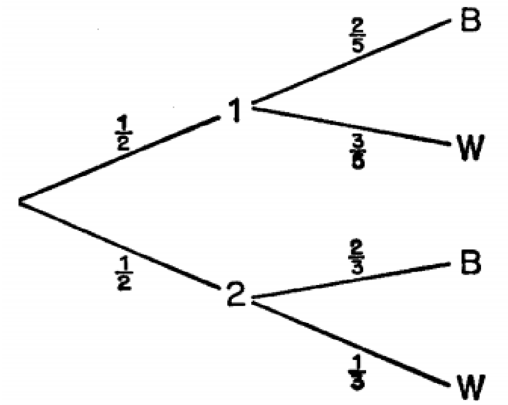

Figure 4.4: ♢

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

Example 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. ♢

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

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

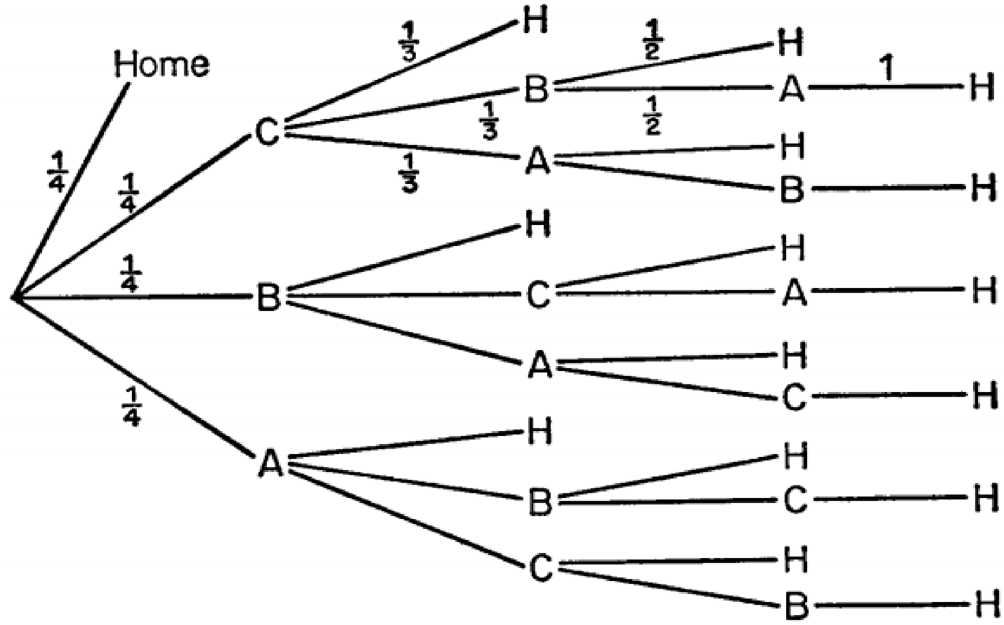

Figure 4.5: ◇

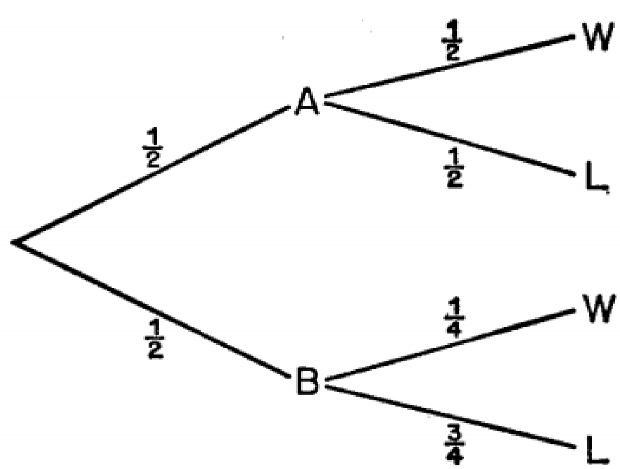

Figure 4.6: ◇

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.) $\diamond$

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. $\frac{20}{32}$.]

Supplementary exercises.

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.

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.]