b) Show that there are either at least three freshmen, at least 19 sophomores, or at least five juniors in the class.

20. Find an increasing subsequence of maximal length and a decreasing subsequence of maximal length in the sequence 22, 5, 7, 2, 23, 10, 15, 21, 3, 17.

21. Construct a sequence of 16 positive integers that has no increasing or decreasing subsequence of five terms.

22. Show that if there are 101 people of different heights standing in a line, it is possible to find 11 people in the order they are standing in the line with heights that are either increasing or decreasing.

*23. Show that whenever 25 girls and 25 boys are seated around a circular table there is always a person both of whose neighbors are boys.

**24. Suppose that 21 girls and 21 boys enter a mathematics competition. Furthermore, suppose that each entrant solves at most six questions, and for every boy-girl pair, there is at least one question that they both solved. Show that there is a question that was solved by at least three girls and at least three boys.

*25. Describe an algorithm in pseudocode for producing the largest increasing or decreasing subsequence of a sequence of distinct integers.

26. Show that in a group of five people (where any two people are either friends or enemies), there are not necessarily three mutual friends or three mutual enemies.

27. Show that in a group of 10 people (where any two people are either friends or enemies), there are either three mutual friends or four mutual enemies, and there are either three mutual enemies or four mutual friends.

28. Use Exercise 27 to show that among any group of 20 people (where any two people are either friends or enemies), there are either four mutual friends or four mutual enemies.

29. Show that if n is an integer with $n \geq 2$, then the Ramsey number $R(2, n)$ equals n. (Recall that Ramsey numbers were discussed after Example 13 in Section 6.2.)

30. Show that if m and n are integers with $m \geq 2$ and $n \geq 2$, then the Ramsey numbers $R(m, n)$ and $R(n, m)$ are equal. (Recall that Ramsey numbers were discussed after Example 13 in Section 6.2.)

31. Show that there are at least six people in California (population: 37 million) with the same three initials who were born on the same day of the year (but not necessarily in the same year). Assume that everyone has three initials.

32. Show that if there are 100,000,000 wage earners in the United States who earn less than 1,000,000 dollars (but at least a penny), then there are two who earned exactly the same amount of money, to the penny, last year.

33. In the 17th century, there were more than 800,000 inhabitants of Paris. At the time, it was believed that no one had more than 200,000 hairs on their head. Assuming these numbers are correct and that everyone has at least one hair on their head (that is, no one is completely bald), use the pigeonhole principle to show, as the French writer Pierre

Nicole did, that there had to be two Parisians with the same number of hairs on their heads. Then use the generalized pigeonhole principle to show that there had to be at least five Parisians at that time with the same number of hairs on their heads.

34. Assuming that no one has more than 1,000,000 hairs on the head of any person and that the population of New York City was 8,008,278 in 2010, show there had to be at least nine people in New York City in 2010 with the same number of hairs on their heads.

35. There are 38 different time periods during which classes at a university can be scheduled. If there are 677 different classes, how many different rooms will be needed?

36. A computer network consists of six computers. Each computer is directly connected to at least one of the other computers. Show that there are at least two computers in the network that are directly connected to the same number of other computers.

37. A computer network consists of six computers. Each computer is directly connected to zero or more of the other computers. Show that there are at least two computers in the network that are directly connected to the same number of other computers. [*Hint:* It is impossible to have a computer linked to none of the others and a computer linked to all the others.]

38. Find the least number of cables required to connect eight computers to four printers to guarantee that for every choice of four of the eight computers, these four computers can directly access four different printers. Justify your answer.

39. Find the least number of cables required to connect 100 computers to 20 printers to guarantee that 2every subset of 20 computers can directly access 20 different printers. (Here, the assumptions about cables and computers are the same as in Example 9.) Justify your answer.

*40. Prove that at a party where there are at least two people, there are two people who know the same number of other people there.

41. An arm wrestler is the champion for a period of 75 hours. (Here, by an hour, we mean a period starting from an exact hour, such as 1 P.M., until the next hour.) The arm wrestler had at least one match an hour, but no more than 125 total matches. Show that there is a period of consecutive hours during which the arm wrestler had exactly 24 matches.

*42. Is the statement in Exercise 41 true if 24 is replaced by
a) 2? **b)** 23? **c)** 25? **d)** 30?

43. Show that if f is a function from S to T, where S and T are nonempty finite sets and $m = \lceil |S| / |T| \rceil$, then there are at least m elements of S mapped to the same value of T. That is, show that there are distinct elements $s_1, s_2, \ldots, s_m$ of S such that $f(s_1) = f(s_2) = \cdots = f(s_m)$.

44. There are 51 houses on a street. Each house has an address between 1000 and 1099, inclusive. Show that at least two houses have addresses that are consecutive integers.

EXAMPLE 14   How many bit strings of length n contain exactly r 1s?

Solution: The positions of r 1s in a bit string of length n form an r-combination of the set $\{1, 2, 3, \ldots, n\}$. Hence, there are $C(n, r)$ bit strings of length n that contain exactly r 1s.   ◀

EXAMPLE 15   Suppose that there are 9 faculty members in the mathematics department and 11 in the computer science department. How many ways are there to select a committee to develop a discrete mathematics course at a school if the committee is to consist of three faculty members from the mathematics department and four from the computer science department?

Solution: By the product rule, the answer is the product of the number of 3-combinations of a set with nine elements and the number of 4-combinations of a set with 11 elements. By Theorem 2, the number of ways to select the committee is

$$C(9, 3) \cdot C(11, 4) = \frac{9!}{3!6!} \cdot \frac{11!}{4!7!} = 84 \cdot 330 = 27{,}720.$$

◀

Exercises

1. List all the permutations of $\{a, b, c\}$.

2. How many different permutations are there of the set $\{a, b, c, d, e, f, g\}$?

3. How many permutations of $\{a, b, c, d, e, f, g\}$ end with a?

4. Let $S = \{1, 2, 3, 4, 5\}$.
  a) List all the 3-permutations of S.
  b) List all the 3-combinations of S.

5. Find the value of each of these quantities.
  a) $P(6, 3)$ **b)** $P(6, 5)$
  c) $P(8, 1)$ **d)** $P(8, 5)$
  e) $P(8, 8)$ **f)** $P(10, 9)$

6. Find the value of each of these quantities.
  a) $C(5, 1)$ **b)** $C(5, 3)$
  c) $C(8, 4)$ **d)** $C(8, 8)$
  e) $C(8, 0)$ **f)** $C(12, 6)$

7. Find the number of 5-permutations of a set with nine elements.

8. In how many different orders can five runners finish a race if no ties are allowed?

9. How many possibilities are there for the win, place, and show (first, second, and third) positions in a horse race with 12 horses if all orders of finish are possible?

10. There are six different candidates for governor of a state. In how many different orders can the names of the candidates be printed on a ballot?

11. How many bit strings of length 10 contain
  a) exactly four 1s?
  b) at most four 1s?
  c) at least four 1s?
  d) an equal number of 0s and 1s?

12. How many bit strings of length 12 contain
  a) exactly three 1s?
  b) at most three 1s?
  c) at least three 1s?
  d) an equal number of 0s and 1s?

13. A group contains n men and n women. How many ways are there to arrange these people in a row if the men and women alternate?

14. In how many ways can a set of two positive integers less than 100 be chosen?

15. In how many ways can a set of five letters be selected from the English alphabet?

16. How many subsets with an odd number of elements does a set with 10 elements have?

17. How many subsets with more than two elements does a set with 100 elements have?

18. A coin is flipped eight times where each flip comes up either heads or tails. How many possible outcomes
  a) are there in total?
  b) contain exactly three heads?
  c) contain at least three heads?
  d) contain the same number of heads and tails?

19. A coin is flipped 10 times where each flip comes up either heads or tails. How many possible outcomes
  a) are there in total?
  b) contain exactly two heads?
  c) contain at most three tails?
  d) contain the same number of heads and tails?

20. How many bit strings of length 10 have
  a) exactly three 0s?
  b) more 0s than 1s?
  c) at least seven 1s?
  d) at least three 1s?

We can prove combinatorial identities by counting bit strings with different properties, as the proof of Theorem 4 will demonstrate.

THEOREM 4   Let n and r be nonnegative integers with $r \le n$. Then

$$\binom{n+1}{r+1} = \sum_{j=r}^{n} \binom{j}{r}.$$

Proof: We use a combinatorial proof. By Example 14 in Section 6.3, the left-hand side, $\binom{n+1}{r+1}$, counts the bit strings of length $n + 1$ containing $r + 1$ ones.

We show that the right-hand side counts the same objects by considering the cases corresponding to the possible locations of the final 1 in a string with $r + 1$ ones. This final one must occur at position $r + 1, r + 2, \ldots,$ or $n + 1$. Furthermore, if the last one is the kth bit there must be r ones among the first $k - 1$ positions. Consequently, by Example 14 in Section 6.3, there are $\binom{k-1}{r}$ such bit strings. Summing over k with $r + 1 \le k \le n + 1$, we find that there are

$$\sum_{k=r+1}^{n+1} \binom{k-1}{r} = \sum_{j=r}^{n} \binom{j}{r}$$

bit strings of length n containing exactly $r + 1$ ones. (Note that the last step follows from the change of variables $j = k - 1$.) Because the left-hand side and the right-hand side count the same objects, they are equal. This completes the proof.   ◁

Exercises

1. Find the expansion of $(x + y)^4$
 a) using combinatorial reasoning, as in Example 1.
 b) using the binomial theorem.

2. Find the expansion of $(x + y)^5$
 a) using combinatorial reasoning, as in Example 1.
 b) using the binomial theorem.

3. Find the expansion of $(x + y)^6$.

4. Find the coefficient of $x^5 y^8$ in $(x + y)^{13}$.

5. How many terms are there in the expansion of $(x + y)^{100}$ after like terms are collected?

6. What is the coefficient of x^7 in $(1 + x)^{11}$?

7. What is the coefficient of x^9 in $(2 - x)^{19}$?

8. What is the coefficient of $x^8 y^9$ in the expansion of $(3x + 2y)^{17}$?

9. What is the coefficient of $x^{101} y^{99}$ in the expansion of $(2x - 3y)^{200}$?

*10. Give a formula for the coefficient of x^k in the expansion of $(x + 1/x)^{100}$, where k is an integer.

*11. Give a formula for the coefficient of x^k in the expansion of $(x^2 - 1/x)^{100}$, where k is an integer.

12. The row of Pascal's triangle containing the binomial coefficients $\binom{10}{k}$, $0 \le k \le 10$, is:

   1 10 45 120 210 252 210 120 45 10 1

 Use Pascal's identity to produce the row immediately following this row in Pascal's triangle.

13. What is the row of Pascal's triangle containing the binomial coefficients $\binom{9}{k}$, $0 \le k \le 9$?

14. Show that if n is a positive integer, then $1 = \binom{n}{0} < \binom{n}{1} < \cdots < \binom{n}{\lfloor n/2 \rfloor} = \binom{n}{\lceil n/2 \rceil} > \cdots > \binom{n}{n-1} > \binom{n}{n} = 1$.

15. Show that $\binom{n}{k} \le 2^n$ for all positive integers n and all integers k with $0 \le k \le n$.

16. a) Use Exercise 14 and Corollary 1 to show that if n is an integer greater than 1, then $\binom{n}{\lfloor n/2 \rfloor} \ge 2^n/n$.
 b) Conclude from part (a) that if n is a positive integer, then $\binom{2n}{n} \ge 4^n/2n$.

☞ 17. Show that if n and k are integers with $1 \le k \le n$, then $\binom{n}{k} \le n^k/2^{k-1}$.

18. Suppose that b is an integer with $b \ge 7$. Use the binomial theorem and the appropriate row of Pascal's triangle to find the base-b expansion of $(11)_b^4$ [that is, the fourth power of the number $(11)_b$ in base-b notation].

19. Prove Pascal's identity, using the formula for $\binom{n}{r}$.

20. Suppose that k and n are integers with $1 \le k < n$. Prove the **hexagon identity**

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

 which relates terms in Pascal's triangle that form a hexagon.

25. How many positive integers less than 1,000,000 have the sum of their digits equal to 19?

26. How many positive integers less than 1,000,000 have exactly one digit equal to 9 and have a sum of digits equal to 13?

27. There are 10 questions on a discrete mathematics final exam. How many ways are there to assign scores to the problems if the sum of the scores is 100 and each question is worth at least 5 points?

28. Show that there are $C(n + r - q_1 - q_2 - \cdots - q_r - 1, n - q_1 - q_2 - \cdots - q_r)$ different unordered selections of n objects of r different types that include at least q_1 objects of type one, q_2 objects of type two, $\ldots$, and q_r objects of type r.

29. How many different bit strings can be transmitted if the string must begin with a 1 bit, must include three additional 1 bits (so that a total of four 1 bits is sent), must include a total of 12 0 bits, and must have at least two 0 bits following each 1 bit?

30. How many different strings can be made from the letters in *MISSISSIPPI*, using all the letters?

31. How many different strings can be made from the letters in *ABRACADABRA*, using all the letters?

32. How many different strings can be made from the letters in *AARDVARK*, using all the letters, if all three *A*s must be consecutive?

33. How many different strings can be made from the letters in *ORONO*, using some or all of the letters?

34. How many strings with five or more characters can be formed from the letters in *SEERESS*?

35. How many strings with seven or more characters can be formed from the letters in *EVERGREEN*?

36. How many different bit strings can be formed using six 1s and eight 0s?

37. A student has three mangos, two papayas, and two kiwi fruits. If the student eats one piece of fruit each day, and only the type of fruit matters, in how many different ways can these fruits be consumed?

38. A professor packs her collection of 40 issues of a mathematics journal in four boxes with 10 issues per box. How many ways can she distribute the journals if

a) each box is numbered, so that they are distinguishable?

b) the boxes are identical, so that they cannot be distinguished?

39. How many ways are there to travel in xyz space from the origin $(0, 0, 0)$ to the point $(4, 3, 5)$ by taking steps one unit in the positive x direction, one unit in the positive y direction, or one unit in the positive z direction? (Moving in the negative x, y, or z direction is prohibited, so that no backtracking is allowed.)

40. How many ways are there to travel in $xyzw$ space from the origin $(0, 0, 0, 0)$ to the point $(4, 3, 5, 4)$ by taking steps one unit in the positive x, positive y, positive z, or positive w direction?

41. How many ways are there to deal hands of seven cards to each of five players from a standard deck of 52 cards?

42. In bridge, the 52 cards of a standard deck are dealt to four players. How many different ways are there to deal bridge hands to four players?

43. How many ways are there to deal hands of five cards to each of six players from a deck containing 48 different cards?

44. In how many ways can a dozen books be placed on four distinguishable shelves

a) if the books are indistinguishable copies of the same title?

b) if no two books are the same, and the positions of the books on the shelves matter? [*Hint:* Break this into 12 tasks, placing each book separately. Start with the sequence 1, 2, 3, 4 to represent the shelves. Represent the books by b_i, $i = 1, 2, \ldots, 12$. Place b_1 to the right of one of the terms in 1, 2, 3, 4. Then successively place $b_2, b_3, \ldots,$ and b_{12}.]

45. How many ways can n books be placed on k distinguishable shelves

a) if the books are indistinguishable copies of the same title?

b) if no two books are the same, and the positions of the books on the shelves matter?

46. A shelf holds 12 books in a row. How many ways are there to choose five books so that no two adjacent books are chosen? [*Hint:* Represent the books that are chosen by bars and the books not chosen by stars. Count the number of sequences of five bars and seven stars so that no two bars are adjacent.]

∗47. Use the product rule to prove Theorem 4, by first placing objects in the first box, then placing objects in the second box, and so on.

∗48. Prove Theorem 4 by first setting up a one-to-one correspondence between permutations of n objects with n_i indistinguishable objects of type i, $i = 1, 2, 3, \ldots, k$, and the distributions of n objects in k boxes such that n_i objects are placed in box i, $i = 1, 2, 3, \ldots, k$ and then applying Theorem 3.

∗49. In this exercise we will prove Theorem 2 by setting up a one-to-one correspondence between the set of r-combinations with repetition allowed of $S = \{1, 2, 3, \ldots, n\}$ and the set of r-combinations of the set $T = \{1, 2, 3, \ldots, n + r - 1\}$.

a) Arrange the elements in an r-combination, with repetition allowed, of S into an increasing sequence $x_1 \le x_2 \le \cdots \le x_r$. Show that the sequence formed by adding $k - 1$ to the kth term is strictly increasing. Conclude that this sequence is made up of r distinct elements from T.

b) Show that the procedure described in (a) defines a one-to-one correspondence between the set of r-combinations, with repetition allowed, of S and the r-combinations of T. [*Hint:* Show the correspondence can be reversed by associating to the r-combination $\{x_1, x_2, \ldots, x_r\}$ of T, with $1 \le x_1 < x_2 < \cdots < x_r \le n + r - 1$, the r-combination with

31. Find the probability that a family with five children does not have a boy, if the sexes of children are independent and if
   a) a boy and a girl are equally likely.
   b) the probability of a boy is 0.51.
   c) the probability that the ith child is a boy is $0.51 - (i/100)$.

32. Find the probability that a randomly generated bit string of length 10 begins with a 1 or ends with a 00 for the same conditions as in parts (a), (b), and (c) of Exercise 30, if bits are generated independently.

33. Find the probability that the first child of a family with five children is a boy or that the last two children of the family are girls, for the same conditions as in parts (a), (b), and (c) of Exercise 31.

34. Find each of the following probabilities when n independent Bernoulli trials are carried out with probability of success p.
   a) the probability of no successes
   b) the probability of at least one success
   c) the probability of at most one success
   d) the probability of at least two successes

35. Find each of the following probabilities when n independent Bernoulli trials are carried out with probability of success p.
   a) the probability of no failures
   b) the probability of at least one failure
   c) the probability of at most one failure
   d) the probability of at least two failures

36. Use mathematical induction to prove that if $E_1, E_2, \ldots, E_n$ is a sequence of n pairwise disjoint events in a sample space S, where n is a positive integer, then $p(\bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} p(E_i)$.

∗37. (*Requires calculus*) Show that if $E_1, E_2, \ldots$ is an infinite sequence of pairwise disjoint events in a sample space S, then $p(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} p(E_i)$. [*Hint:* Use Exercise 36 and take limits.]

38. A pair of dice is rolled in a remote location and when you ask an honest observer whether at least one die came up six, this honest observer answers in the affirmative.
   a) What is the probability that the sum of the numbers that came up on the two dice is seven, given the information provided by the honest observer?

   b) Suppose that the honest observer tells us that at least one die came up five. What is the probability the sum of the numbers that came up on the dice is seven, given this information?

∗∗39. This exercise employs the probabilistic method to prove a result about round-robin tournaments. In a **round-robin tournament** with m players, every two players play one game in which one player wins and the other loses.

   We want to find conditions on positive integers m and k with $k < m$ such that it is possible for the outcomes of the tournament to have the property that for every set of k players, there is a player who beats every member in this set. So that we can use probabilistic reasoning to draw conclusions about round-robin tournaments, we assume that when two players compete it is equally likely that either player wins the game and we assume that the outcomes of different games are independent. Let E be the event that for every set S with k players, where k is a positive integer less than m, there is a player who has beaten all k players in S.
   a) Show that $p(\overline{E}) \leq \sum_{j=1}^{\binom{m}{k}} p(F_j)$, where F_j is the event that there is no player who beats all k players from the jth set in a list of the $\binom{m}{k}$ sets of k players.
   b) Show that the probability of F_j is $(1 - 2^{-k})^{m-k}$.
   c) Conclude from parts (a) and (b) that $p(\overline{E}) \leq \binom{m}{k}(1 - 2^{-k})^{m-k}$ and, therefore, that there must be a tournament with the described property if $\binom{m}{k}(1 - 2^{-k})^{m-k} < 1$.
   d) Use part (c) to find values of m such that there is a tournament with m players such that for every set S of two players, there is a player who has beaten both players in S. Repeat for sets of three players.

∗40. Devise a Monte Carlo algorithm that determines whether a permutation of the integers 1 through n has already been sorted (that is, it is in increasing order), or instead, is a random permutation. A step of the algorithm should answer "true" if it determines the list is not sorted and "unknown" otherwise. After k steps, the algorithm decides that the integers are sorted if the answer is "unknown" in each step. Show that as the number of steps increases, the probability that the algorithm produces an incorrect answer is extremely small. [*Hint:* For each step, test whether certain elements are in the correct order. Make sure these tests are independent.]

41. Use pseudocode to write out the probabilistic primality test described in Example 16.

7.3 Bayes' Theorem

Introduction

There are many times when we want to assess the probability that a particular event occurs on the basis of partial evidence. For example, suppose we know the percentage of people who have a particular disease for which there is a very accurate diagnostic test. People who test positive for

event that the message contains word w_i, assuming that the number of incoming spam messages is approximately the same as the number of incoming messages that are not spam, and that the events $E_i \mid S$ are independent, then by Bayes' theorem the probability that a message containing all the words $w_1, w_2, \ldots, w_k$ is spam is

$$p(S \mid \bigcap_{i=1}^{k} E_i) = \frac{\prod_{i=1}^{k} p(E_i \mid S)}{\prod_{i=1}^{k} p(E_i \mid S) + \prod_{i=1}^{k} p(E_i \mid \overline{S})}.$$

We can estimate this probability by

$$r(w_1, w_2, \ldots, w_k) = \frac{\prod_{i=1}^{k} p(w_i)}{\prod_{i=1}^{k} p(w_i) + \prod_{i=1}^{k} q(w_i)}.$$

For the most effective spam filter, we choose words for which the probability that each of these words appears in spam is either very high or very low. When we compute this value for a particular message, we reject the message as spam if $r(w_1, w_2, \ldots, w_k)$ exceeds a preset threshold, such as 0.9.

Bayesian poisoning, the insertion of extra words to defeat spam filters, can use random or purposefully selected words.

Another way to improve the performance of a Bayesian spam filter is to look at the probabilities that particular pairs of words appear in spam and in messages that are not spam. We then treat appearances of these pairs of words as appearance of a single block, rather than as the appearance of two separate words. For example, the pair of words "enhance performance" most likely indicates spam, while "operatic performance" indicates a message that is not spam. Similarly, we can assess the likelihood that a message is spam by examining the structure of a message to determine where words appear in it. Also, spam filters look at appearances of certain types of strings of characters rather than just words. For example, a message with the valid e-mail address of one of your friends is less likely to be spam (if not sent by a worm) than one containing an e-mail address that came from a country known to originate a lot of spam. There is an ongoing war between people who create spam and those trying to filter their messages out. This leads to the introduction of many new techniques to defeat spam filters, including inserting into spam messages long strings of words that appear in messages that are not spam, as well as including words inside pictures. The techniques we have discussed here are only the first steps in fighting this war on spam.

Exercises

1. Suppose that E and F are events in a sample space and $p(E) = 1/3$, $p(F) = 1/2$, and $p(E \mid F) = 2/5$. Find $p(F \mid E)$.

2. Suppose that E and F are events in a sample space and $p(E) = 2/3$, $p(F) = 3/4$, and $p(F \mid E) = 5/8$. Find $p(E \mid F)$.

3. Suppose that Frida selects a ball by first picking one of two boxes at random and then selecting a ball from this box at random. The first box contains two white balls and three blue balls, and the second box contains four white balls and one blue ball. What is the probability that Frida picked a ball from the first box if she has selected a blue ball?

4. Suppose that Ann selects a ball by first picking one of two boxes at random and then selecting a ball from this box. The first box contains three orange balls and four black balls, and the second box contains five orange balls and six black balls. What is the probability that Ann picked a ball from the second box if she has selected an orange ball?

5. Suppose that 8% of all bicycle racers use steroids, that a bicyclist who uses steroids tests positive for steroids 96% of the time, and that a bicyclist who does not use steroids tests positive for steroids 9% of the time. What is the probability that a randomly selected bicyclist who tests positive for steroids actually uses steroids?

6. When a test for steroids is given to soccer players, 98% of the players taking steroids test positive and 12% of the players not taking steroids test positive. Suppose that 5% of soccer players take steroids. What is the probability that a soccer player who tests positive takes steroids?

7. Suppose that a test for opium use has a 2% false positive rate and a 5% false negative rate. That is, 2% of people who do not use opium test positive for opium, and

5% of opium users test negative for opium. Furthermore, suppose that 1% of people actually use opium.

a) Find the probability that someone who tests negative for opium use does not use opium.

b) Find the probability that someone who tests positive for opium use actually uses opium.

8. Suppose that one person in 10,000 people has a rare genetic disease. There is an excellent test for the disease; 99.9% of people with the disease test positive and only 0.02% who do not have the disease test positive.

a) What is the probability that someone who tests positive has the genetic disease?

b) What is the probability that someone who tests negative does not have the disease?

9. Suppose that 8% of the patients tested in a clinic are infected with HIV. Furthermore, suppose that when a blood test for HIV is given, 98% of the patients infected with HIV test positive and that 3% of the patients not infected with HIV test positive. What is the probability that

a) a patient testing positive for HIV with this test is infected with it?

b) a patient testing positive for HIV with this test is not infected with it?

c) a patient testing negative for HIV with this test is infected with it?

d) a patient testing negative for HIV with this test is not infected with it?

10. Suppose that 4% of the patients tested in a clinic are infected with avian influenza. Furthermore, suppose that when a blood test for avian influenza is given, 97% of the patients infected with avian influenza test positive and that 2% of the patients not infected with avian influenza test positive. What is the probability that

a) a patient testing positive for avian influenza with this test is infected with it?

b) a patient testing positive for avian influenza with this test is not infected with it?

c) a patient testing negative for avian influenza with this test is infected with it?

d) a patient testing negative for avian influenza with this test is not infected with it?

11. An electronics company is planning to introduce a new camera phone. The company commissions a marketing report for each new product that predicts either the success or the failure of the product. Of new products introduced by the company, 60% have been successes. Furthermore, 70% of their successful products were predicted to be successes, while 40% of failed products were predicted to be successes. Find the probability that this new camera phone will be successful if its success has been predicted.

✱12. A space probe near Neptune communicates with Earth using bit strings. Suppose that in its transmissions it sends a 1 one-third of the time and a 0 two-thirds of the time. When a 0 is sent, the probability that it is received correctly is 0.9, and the probability that it is received incorrectly (as a 1) is 0.1. When a 1 is sent, the probability that it is received correctly is 0.8, and the probability that it is received incorrectly (as a 0) is 0.2.

a) Find the probability that a 0 is received.

b) Use Bayes' theorem to find the probability that a 0 was transmitted, given that a 0 was received.

13. Suppose that E, F_1, F_2, and F_3 are events from a sample space S and that F_1, F_2, and F_3 are pairwise disjoint and their union is S. Find $p(F_1 \mid E)$ if $p(E \mid F_1) = 1/8$, $p(E \mid F_2) = 1/4$, $p(E \mid F_3) = 1/6$, $p(F_1) = 1/4$, $p(F_2) = 1/4$, and $p(F_3) = 1/2$.

14. Suppose that E, F_1, F_2, and F_3 are events from a sample space S and that F_1, F_2, and F_3 are pairwise disjoint and their union is S. Find $p(F_2 \mid E)$ if $p(E \mid F_1) = 2/7$, $p(E \mid F_2) = 3/8$, $p(E \mid F_3) = 1/2$, $p(F_1) = 1/6$, $p(F_2) = 1/2$, and $p(F_3) = 1/3$.

15. In this exercise we will use Bayes' theorem to solve the Monty Hall puzzle (Example 10 in Section 7.1). Recall that in this puzzle you are asked to select one of three doors to open. There is a large prize behind one of the three doors and the other two doors are losers. After you select a door, Monty Hall opens one of the two doors you did not select that he knows is a losing door, selecting at random if both are losing doors. Monty asks you whether you would like to switch doors. Suppose that the three doors in the puzzle are labeled 1, 2, and 3. Let W be the random variable whose value is the number of the winning door; assume that $p(W = k) = 1/3$ for $k = 1, 2, 3$. Let M denote the random variable whose value is the number of the door that Monty opens. Suppose you choose door i.

a) What is the probability that you will win the prize if the game ends without Monty asking you whether you want to change doors?

b) Find $p(M = j \mid W = k)$ for $j = 1, 2, 3$ and $k = 1, 2, 3$.

c) Use Bayes' theorem to find $p(W = j \mid M = k)$ where i and j and k are distinct values.

d) Explain why the answer to part (c) tells you whether you should change doors when Monty gives you the chance to do so.

16. Ramesh can get to work in three different ways: by bicycle, by car, or by bus. Because of commuter traffic, there is a 50% chance that he will be late when he drives his car. When he takes the bus, which uses a special lane reserved for buses, there is a 20% chance that he will be late. The probability that he is late when he rides his bicycle is only 5%. Ramesh arrives late one day. His boss wants to estimate the probability that he drove his car to work that day.

a) Suppose the boss assumes that there is a 1/3 chance that Ramesh takes each of the three ways he can get to work. What estimate for the probability that Ramesh drove his car does the boss obtain from Bayes' theorem under this assumption?

b) Suppose the boss knows that Ramesh drives 30% of the time, takes the bus only 10% of the time, and takes his bicycle 60% of the time. What estimate for the probability that Ramesh drove his car does the boss obtain from Bayes' theorem using this information?

*17. Prove Theorem 2, the extended form of Bayes' theorem. That is, suppose that E is an event from a sample space S and that $F_1, F_2, \ldots, F_n$ are mutually exclusive events such that $\bigcup_{i=1}^{n} F_i = S$. Assume that $p(E) \neq 0$ and $p(F_i) \neq 0$ for $i = 1, 2, \ldots, n$. Show that

$$p(F_j \mid E) = \frac{p(E \mid F_j)p(F_j)}{\sum_{i=1}^{n} p(E \mid F_i)p(F_i)}.$$

[*Hint:* Use the fact that $E = \bigcup_{i=1}^{n}(E \cap F_i)$.]

18. Suppose that a Bayesian spam filter is trained on a set of 500 spam messages and 200 messages that are not spam. The word "exciting" appears in 40 spam messages and in 25 messages that are not spam. Would an incoming message be rejected as spam if it contains the word "exciting" and the threshold for rejecting spam is 0.9?

19. Suppose that a Bayesian spam filter is trained on a set of 1000 spam messages and 400 messages that are not spam. The word "opportunity" appears in 175 spam messages and 20 messages that are not spam. Would an incoming message be rejected as spam if it contains the word "opportunity" and the threshold for rejecting a message is 0.9?

20. Would we reject a message as spam in Example 4

a) using just the fact that the word "undervalued" occurs in the message?

b) using just the fact that the word "stock" occurs in the message?

21. Suppose that a Bayesian spam filter is trained on a set of 10,000 spam messages and 5000 messages that are not spam. The word "enhancement" appears in 1500 spam

messages and 20 messages that are not spam, while the word "herbal" appears in 800 spam messages and 200 messages that are not spam. Estimate the probability that a received message containing both the words "enhancement" and "herbal" is spam. Will the message be rejected as spam if the threshold for rejecting spam is 0.9?

22. Suppose that we have prior information concerning whether a random incoming message is spam. In particular, suppose that over a time period, we find that s spam messages arrive and h messages arrive that are not spam.

a) Use this information to estimate $p(S)$, the probability that an incoming message is spam, and $p(\overline{S})$, the probability an incoming message is not spam.

b) Use Bayes' theorem and part (a) to estimate the probability that an incoming message containing the word w is spam, where $p(w)$ is the probability that w occurs in a spam message and $q(w)$ is the probability that w occurs in a message that is not spam.

23. Suppose that E_1 and E_2 are the events that an incoming mail message contains the words w_1 and w_2, respectively. Assuming that E_1 and E_2 are independent events and that $E_1 \mid S$ and $E_2 \mid S$ are independent events, where S is the event that an incoming message is spam, and that we have no prior knowledge regarding whether or not the message is spam, show that

$$p(S \mid E_1 \cap E_2)$$

$$= \frac{p(E_1 \mid S)p(E_2 \mid S)}{p(E_1 \mid S)p(E_2 \mid S) + p(E_1 \mid \overline{S})p(E_2 \mid \overline{S})}.$$

7.4   Expected Value and Variance

Introduction

The **expected value** of a random variable is the sum over all elements in a sample space of the product of the probability of the element and the value of the random variable at this element. Consequently, the expected value is a weighted average of the values of a random variable. The expected value of a random variable provides a central point for the distribution of values of this random variable. We can solve many problems using the notion of the expected value of a random variable, such as determining who has an advantage in gambling games and computing the average-case complexity of algorithms. Another useful measure of a random variable is its **variance**, which tells us how spread out the values of this random variable are. We can use the variance of a random variable to help us estimate the probability that a random variable takes values far removed from its expected value.

Expected Values

Links

Many questions can be formulated in terms of the value we expect a random variable to take, or more precisely, the average value of a random variable when an experiment is performed a large number of times. Questions of this kind include: How many heads are expected to appear

Exercises

1. What is the expected number of heads that come up when a fair coin is flipped five times?

2. What is the expected number of heads that come up when a fair coin is flipped 10 times?

3. What is the expected number of times a 6 appears when a fair die is rolled 10 times?

4. A coin is biased so that the probability a head comes up when it is flipped is 0.6. What is the expected number of heads that come up when it is flipped 10 times?

5. What is the expected sum of the numbers that appear on two dice, each biased so that a 3 comes up twice as often as each other number?

6. What is the expected value when a $1 lottery ticket is bought in which the purchaser wins exactly $10 million if the ticket contains the six winning numbers chosen from the set $\{1, 2, 3, \ldots, 50\}$ and the purchaser wins nothing otherwise?

7. The final exam of a discrete mathematics course consists of 50 true/false questions, each worth two points, and 25 multiple-choice questions, each worth four points. The probability that Linda answers a true/false question correctly is 0.9, and the probability that she answers a multiple-choice question correctly is 0.8. What is her expected score on the final?

8. What is the expected sum of the numbers that appear when three fair dice are rolled?

9. Suppose that the probability that x is in a list of n distinct integers is $2/3$ and that it is equally likely that x equals any element in the list. Find the average number of comparisons used by the linear search algorithm to find x or to determine that it is not in the list.

10. Suppose that we flip a fair coin until either it comes up tails twice or we have flipped it six times. What is the expected number of times we flip the coin?

11. Suppose that we roll a fair die until a 6 comes up or we have rolled it 10 times. What is the expected number of times we roll the die?

12. Suppose that we roll a fair die until a 6 comes up.
    a) What is the probability that we roll the die n times?
    b) What is the expected number of times we roll the die?

13. Suppose that we roll a pair of fair dice until the sum of the numbers on the dice is seven. What is the expected number of times we roll the dice?

14. Show that the sum of the probabilities of a random variable with geometric distribution with parameter p, where $0 < p \leq 1$, equals 1.

15. Show that if the random variable X has the geometric distribution with parameter p, and j is a positive integer, then $p(X \geq j) = (1 - p)^{j-1}$.

16. Let X and Y be the random variables that count the number of heads and the number of tails that come up when two fair coins are flipped. Show that X and Y are not independent.

17. Estimate the expected number of integers with 1000 digits that need to be selected at random to find a prime, if the probability a number with 1000 digits is prime is approximately $1/2302$.

18. Suppose that X and Y are random variables and that X and Y are nonnegative for all points in a sample space S. Let Z be the random variable defined by $Z(s) = \max(X(s), Y(s))$ for all elements $s \in S$. Show that $E(Z) \leq E(X) + E(Y)$.

19. Let X be the number appearing on the first die when two fair dice are rolled and let Y be the sum of the numbers appearing on the two dice. Show that $E(X)E(Y) \neq E(XY)$.

*20. Show that if $X_1, X_2, \ldots, X_n$ are mutually independent random variables, then $E(\prod_{i=1}^{n} X_i) = \prod_{i=1}^{n} E(X_i)$.

The **conditional expectation** of the random variable X given the event A from the sample space S is $E(X|A) = \sum_{r \in X(S)} r \cdot P(X = r|A)$.

21. What is expected value of the sum of the numbers appearing on two fair dice when they are rolled given that the sum of these numbers is at least nine. That is, what is $E(X|A)$ where X is the sum of the numbers appearing on the two dice and A is the event that $X \geq 9$?

The **law of total expectation** states that if the sample space S is the disjoint union of the events $S_1, S_2, \ldots, S_n$ and X is a random variable, then $E(X) = \sum_{j=1}^{n} E(X|S_j)P(S_j)$.

22. Prove the law of total expectations.

23. Use the law of total expectation to find the average weight of a breeding elephant seal, given that 12% of the breeding elephant seals are male and the rest are female, and the expected weights of a breeding elephant seal is 4,200 pounds for a male and 1,100 pounds for a female.

24. Let A be an event. Then I_A, the **indicator random variable** of A, equals 1 if A occurs and equals 0 otherwise. Show that the expectation of the indicator random variable of A equals the probability of A, that is, $E(I_A) = p(A)$.

25. A **run** is a maximal sequence of successes in a sequence of Bernoulli trials. For example, in the sequence $S, S, S, F, S, S, F, F, S$, where S represents success and F represents failure, there are three runs consisting of three successes, two successes, and one success, respectively. Let R denote the random variable on the set of sequences of n independent Bernoulli trials that counts the number of runs in this sequence. Find $E(R)$. [*Hint:* Show that $R = \sum_{j=1}^{n} I_j$, where $I_j = 1$ if a run begins at the jth Bernoulli trial and $I_j = 0$ otherwise. Find $E(I_1)$ and then find $E(I_j)$, where $1 < j \leq n$.]

26. Let $X(s)$ be a random variable, where $X(s)$ is a nonnegative integer for all $s \in S$, and let A_k be the event that $X(s) \geq k$. Show that $E(X) = \sum_{k=1}^{\infty} p(A_k)$.

27. What is the variance of the number of heads that come up when a fair coin is flipped 10 times?

28. What is the variance of the number of times a 6 appears when a fair die is rolled 10 times?

29. Let X_n be the random variable that equals the number of tails minus the number of heads when n fair coins are flipped.

   a) What is the expected value of X_n?

   b) What is the variance of X_n?

30. Show that if X and Y are independent random variables, then $V(XY) = E(X)^2 V(Y) + E(Y)^2 V(X) + V(X)V(Y)$

31. Let $A(X) = E(|X - E(X)|)$, the expected value of the absolute value of the deviation of X, where X is a random variable. Prove or disprove that $A(X + Y) = A(X) + A(Y)$ for all random variables X and Y.

32. Provide an example that shows that the variance of the sum of two random variables is not necessarily equal to the sum of their variances when the random variables are not independent.

33. Suppose that X_1 and X_2 are independent Bernoulli trials each with probability $1/2$, and let $X_3 = (X_1 + X_2) \bmod 2$.

   a) Show that X_1, X_2, and X_3 are pairwise independent, but X_3 and $X_1 + X_2$ are not independent.

   b) Show that $V(X_1 + X_2 + X_3) = V(X_1) + V(X_2) + V(X_3)$.

   c) Explain why a proof by mathematical induction of Theorem 7 does not work by considering the random variables X_1, X_2, and X_3.

*34. Prove the general case of Theorem 7. That is, show that if X_1, $X_2, \ldots, X_n$ are pairwise independent random variables on a sample space S, where n is a positive integer, then $V(X_1 + X_2 + \cdots + X_n) = V(X_1) + V(X_2) + \cdots + V(X_n)$. [*Hint:* Generalize the proof given in Theorem 7 for two random variables. Note that a proof using mathematical induction does not work; see Exercise 33.]

35. Use Chebyshev's inequality to find an upper bound on the probability that the number of tails that come up when a fair coin is tossed n times deviates from the mean by more than $5\sqrt{n}$.

36. Use Chebyshev's inequality to find an upper bound on the probability that the number of tails that come up when a biased coin with probability of heads equal to 0.6 is tossed n times deviates from the mean by more than $\sqrt{n}$.

37. Let X be a random variable on a sample space S such that $X(s) \geq 0$ for all $s \in S$. Show that $p(X(s) \geq a) \leq E(X)/a$ for every positive real number a. This inequality is called **Markov's inequality**.

38. Suppose that the number of cans of soda pop filled in a day at a bottling plant is a random variable with an expected value of 10,000 and a variance of 1000.

   a) Use Markov's inequality (Exercise 37) to obtain an upper bound on the probability that the plant will fill more than 11,000 cans on a particular day.

   b) Use Chebyshev's inequality to obtain a lower bound on the probability that the plant will fill between 9000 and 11,000 cans on a particular day.

39. Suppose that the number of tin cans recycled in a day at a recycling center is a random variable with an expected value of 50,000 and a variance of 10,000.

   a) Use Markov's inequality (Exercise 37) to find an upper bound on the probability that the center will recycle more than 55,000 cans on a particular day.

   b) Use Chebyshev's inequality to provide a lower bound on the probability that the center will recycle 40,000 to 60,000 cans on a certain day.

*40. Suppose the probability that x is the ith element in a list of n distinct integers is $i/[n(n + 1)]$. Find the average number of comparisons used by the linear search algorithm to find x or to determine that it is not in the list.

*41. In this exercise we derive an estimate of the average-case complexity of the variant of the bubble sort algorithm that terminates once a pass has been made with no interchanges. Let X be the random variable on the set of permutations of a set of n distinct integers $\{a_1, a_2, \ldots, a_n\}$ with $a_1 < a_2 < \cdots < a_n$ such that $X(P)$ equals the number of comparisons used by the bubble sort to put these integers into increasing order.

   a) Show that, under the assumption that the input is equally likely to be any of the $n!$ permutations of these integers, the average number of comparisons used by the bubble sort equals $E(X)$.

   b) Use Example 5 in Section 3.3 to show that $E(X) \leq n(n - 1)/2$.

   c) Show that the sort makes at least one comparison for every inversion of two integers in the input.

   d) Let $I(P)$ be the random variable that equals the number of inversions in the permutation P. Show that $E(X) \geq E(I)$.

   e) Let $I_{j,k}$ be the random variable with $I_{j,k}(P) = 1$ if a_k precedes a_j in P and $I_{j,k} = 0$ otherwise. Show that $I(P) = \sum_k \sum_{j<k} I_{j,k}(P)$.

   f) Show that $E(I) = \sum_k \sum_{j<k} E(I_{j,k})$.

   g) Show that $E(I_{j,k}) = 1/2$. [*Hint:* Show that $E(I_{j,k}) = $ probability that a_k precedes a_j in a permutation P. Then show it is equally likely for a_k to precede a_j as it is for a_j to precede a_k in a permutation.]

   h) Use parts (f) and (g) to show that $E(I) = n(n - 1)/4$.

   i) Conclude from parts (b), (d), and (h) that the average number of comparisons used to sort n integers is $\Theta(n^2)$.

*42. In this exercise we find the average-case complexity of the quick sort algorithm, described in the preamble to Exercise 50 in Section 5.4, assuming a uniform distribution on the set of permutations.

   a) Let X be the number of comparisons used by the quick sort algorithm to sort a list of n distinct integers. Show that the average number of comparisons used by the quick sort algorithm is $E(X)$ (where the sample space is the set of all $n!$ permutations of n integers).

b) Let $I_{j,k}$ denote the random variable that equals 1 if the jth smallest element and the kth smallest element of the initial list are ever compared as the quick sort algorithm sorts the list and equals 0 otherwise. Show that $X = \sum_{k=2}^{n} \sum_{j=1}^{k-1} I_{j,k}$.

c) Show that $E(X) = \sum_{k=2}^{n} \sum_{j=1}^{k-1} p$(the jth smallest element and the kth smallest element are compared).

d) Show that p(the jth smallest element and the kth smallest element are compared), where $k > j$, equals $2/(k - j + 1)$.

e) Use parts (c) and (d) to show that $E(X) = 2(n + 1)(\sum_{i=2}^{n} 1/i) - 2(n - 1)$.

f) Conclude from part (e) and the fact that $\sum_{j=1}^{n} 1/j \approx \ln n + \gamma$, where $\gamma = 0.57721\ldots$ is Euler's constant, that the average number of comparisons used by the quick sort algorithm is $\Theta(n \log n)$.

*43. What is the variance of the number of **fixed elements**, that is, elements left in the same position, of a randomly selected permutation of n elements? [*Hint:* Let X denote the number of fixed points of a random permutation. Write $X = X_1 + X_2 + \cdots + X_n$, where $X_i = 1$ if the permutation fixes the ith element and $X_i = 0$ otherwise.]

The **covariance** of two random variables X and Y on a sample space S, denoted by $\text{Cov}(X, Y)$, is defined to be the expected value of the random variable $(X - E(X))(Y - E(Y))$. That is, $\text{Cov}(X, Y) = E((X - E(X))(Y - E(Y)))$.

44. Show that $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$, and use this result to conclude that $\text{Cov}(X, Y) = 0$ if X and Y are independent random variables.

45. Show that $V(X + Y) = V(X) + V(Y) + 2\,\text{Cov}(X, Y)$.

46. Find $\text{Cov}(X, Y)$ if X and Y are the random variables with $X((i, j)) = 2i$ and $Y((i, j)) = i + j$, where i and j are the numbers that appear on the first and second of two dice when they are rolled.

47. When m balls are distributed into n bins uniformly at random, what is the probability that the first bin remains empty?

48. What is the expected number of balls that fall into the first bin when m balls are distributed into n bins uniformly at random?

49. What is the expected number of bins that remain empty when m balls are distributed into n bins uniformly at random?

Key Terms and Results

TERMS

sample space: the set of possible outcomes of an experiment

event: a subset of the sample space of an experiment

probability of an event (Laplace's definition): the number of successful outcomes of this event divided by the number of possible outcomes

probability distribution: a function p from the set of all outcomes of a sample space S for which $0 \le p(x_i) \le 1$ for $i = 1, 2, \ldots, n$ and $\sum_{i=1}^{n} p(x_i) = 1$, where $x_1, \ldots, x_n$ are the possible outcomes

probability of an event E: the sum of the probabilities of the outcomes in E

$p(E|F)$ (conditional probability of E given F): the ratio $p(E \cap F)/p(F)$

independent events: events E and F such that $p(E \cap F) = p(E)p(F)$

pairwise independent events: events $E_1, E_2, \ldots, E_n$ such that $p(E_i \cap E_j) = p(E_i)p(E_j)$ for all pairs of integers i and j with $1 \le j < k \le n$

mutually independent events: events $E_1, E_2, \ldots, E_n$ such that $p(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_m}) = p(E_{i_1})p(E_{i_2}) \cdots p(E_{i_m})$ whenever $i_j, j = 1, 2, \ldots, m$, are integers with $1 \le i_1 < i_2 < \cdots < i_m \le n$ and $m \ge 2$

random variable: a function that assigns a real number to each possible outcome of an experiment

distribution of a random variable X: the set of pairs $(r, p(X = r))$ for $r \in X(S)$

uniform distribution: the assignment of equal probabilities to the elements of a finite set

expected value of a random variable: the weighted average of a random variable, with values of the random variable weighted by the probability of outcomes, that is, $E(X) = \sum_{s \in S} p(s)X(s)$

geometric distribution: the distribution of a random variable X such that $p(X = k) = (1 - p)^{k-1}p$ for $k = 1, 2, \ldots$ for some real number p with $0 \le p \le 1$.

independent random variables: random variables X and Y such that $p(X = r_1$ and $Y = r_2) = p(X = r_1)p(Y = r_2)$ for all real numbers r_1 and r_2

variance of a random variable X: the weighted average of the square of the difference between the value of X and its expected value $E(X)$, with weights given by the probability of outcomes, that is, $V(X) = \sum_{s \in S}(X(s) - E(X))^2 p(s)$

standard deviation of a random variable X: the square root of the variance of X, that is, $\sigma(X) = \sqrt{V(X)}$

Bernoulli trial: an experiment with two possible outcomes

probabilistic (or Monte Carlo) algorithm: an algorithm in which random choices are made at one or more steps

probabilistic method: a technique for proving the existence of objects in a set with certain properties that proceeds by assigning probabilities to objects and showing that the probability that an object has these properties is positive