Chapter 1

Discrete Probability Distributions

1.1 Simulation of Discrete Probabilities

1. As n increases, the *proportion* of heads gets closer to 1/2, but the *difference* between the number of heads and half the number of flips tends to increase (although it will occasionally be 0).

2. n must be approximately 100.

3. (b) If one simulates a sufficiently large number of rolls, one should be able to conclude that the gamblers were correct.

4. Player one has a probability of about .83 of winning.

5. The smallest n should be about 150.

7. The graph of winnings for betting on a color is much smoother (i.e. has smaller fluctuations) than the graph for betting on a number.

8. For two tosses both probabilities are 1/2. For four tosses they are both 6/16. They are, in fact, the same for any even number of tosses. (This is not at all obvious; see Chapter 12 for a discussion of this and related topics.)

9. Each time you win, you either win an amount that you have already lost or one of the original numbers 1, 2, 3, 4, and hence your net winning is just the sum of these four numbers. This is not a foolproof system, since you may reach a point where you have to bet more money than you have. If you and the bank had unlimited resources it would be foolproof.

10. You are very likely to win 5 dollars, but we shall see that this is still an unfair game, so we might say that Thackeray was right.

11. For two tosses, the probabilities that Peter wins 0 and 2 are 1/2 and 1/4, respectively. For four tosses, the probabilities that Peter wins 0, 2, and 4 are 3/8, 1/4, and 1/16, respectively.

13. Your simulation should result in about 25 days in a year having more than 60 percent boys in the large hospital and about 55 days in a year having more than 60 percent boys in the small hospital.

14. About 1/2 the time you win 2, 1/4 of the time you win 4, 1/8 of the time you win 8, etc. If you add up all of these potential winnings, weighted by their probabilities, you get ∞, so it would seem that you should be willing to pay quite a lot to play this game. Few are willing to pay more than \$10.

15. In about 25 percent of the games the player will have a streak of five.

16. In the case of having children until they have a boy, they should have about 200,000 children. In the case that they have children until they have both a boy and a girl, they should have about 300,000 children, or about 100,000 more.

1.2 Discrete Probability Distributions

1. $P(\{a,b,c\}) = 1 \qquad P(\{a\}) = 1/2$

$$\begin{array}{ll} P(\{a,b\}) = 5/6 & P(\{b\}) = 1/3 \\ P(\{b,c\}) = 1/2 & P(\{c\}) = 1/6 \\ P(\{a,c\}) = 2/3 & P(\phi) = 0 \end{array}$$

2. (a) $\Omega =$ $\{A$ elected, B elected$\}$.

(b) $\Omega =$ {Head,Tail}.

(c) $\Omega =$ {(Jan., Mon.), (Jan., Tue.),...,(Jan., Sun.),...,(Dec., Sun.)}.

(d) $\Omega =$ {Student 1,...,Student 10}.

(e) $\Omega = \{A, B, C, D, F\}$.

3. (b), (d)

4. (a) In three tosses of a coin the first outcome is a head.

(b) In three tosses of a coin the same side turns up on each toss.

(c) In three tosses of a coin exactly one tail turns up.

(d) In three tosses of a coin at least one tail turns up.

5. (a) 1/2

(b) 1/4

(c) 3/8

(d) 7/8

6. $\frac{4}{7}$.

7. 11/12

8. Art $\frac{1}{4}$, Psychology $\frac{1}{2}$, Geology $\frac{1}{4}$.

9. 3/4, 1

10. $\frac{1}{2}$.

11. 1 : 12, 1 : 3, 1 : 35

12. $\frac{3}{4}$.

13. 11:4

14. (a) $m_Y(2) = 1/5,\ m_Y(3) = 1/5,\ m_Y(4) = 2/5,\ m_Y(5) = 1/5$

(b) $m_Z(0) = 1/5,\ m_Z(1) = 3/5,\ m_Z(4) = 1/5$

15. Let the sample space be:

$\omega_1 = \{A, A\}$ $\omega_4 = \{B, A\}$ $\omega_7 = \{C, A\}$

$\omega_2 = \{A, B\}$ $\omega_5 = \{B, B\}$ $\omega_8 = \{C, B\}$

$\omega_3 = \{A, C\}$ $\omega_6 = \{B, C\}$ $\omega_9 = \{C, C\}$

where the first grade is John's and the second is Mary's. You are given that

$$P(\omega_4) + P(\omega_5) + P(\omega_6) = .3,$$

$$P(\omega_2) + P(\omega_5) + P(\omega_8) = .4,$$

$$P(\omega_5) + P(\omega_6) + P(\omega_8) = .1.$$

Adding the first two equations and subtracting the third, we obtain the desired probability as

$$P(\omega_2) + P(\omega_4) + P(\omega_5) = .6.$$

16. 10 per cent. An example: 10 lost eye, ear, hand, and leg; 15 eye, ear, and hand; 20 eye, ear, and leg; 25 eye, hand, and leg; 30 ear, hand, and leg.

17. The sample space for a sequence of m experiments is the set of m-tuples of S's and F's, where S represents a success and F a failure. The probability assigned to a sample point with k successes and $m - k$ failures is

$$\left(\frac{1}{n}\right)^k \left(\frac{n-1}{n}\right)^{m-k}.$$

(a) Let $k = 0$ in the above expression.

(b) If $m = n \log 2$, then

$$\begin{aligned}
\lim_{n\to\infty}\left(1-\frac{1}{n}\right)^m &= \lim_{n\to\infty}\left(\left(1-\frac{1}{n}\right)^n\right)^{\log 2} \\
&= \left(\lim_{n\to\infty}(\left(1-\frac{1}{n}\right)^n\right)^{\log 2} \\
&= \left(e^{-1}\right)^{\log 2} \\
&= \frac{1}{2}.
\end{aligned}$$

(c) Peter Doyle provided the following answer. To achieve more accuracy than what was obtained in part b), we can use power series in the following way. We begin with the equation

$$\left(1-\frac{1}{n}\right)^m = \frac{1}{2},$$

and as before, we solve for m. We obtain

$$m = -\frac{\log 2}{\log(1 - 1/n)}.$$

Thus, for example, if $n = 36$, then $m = 24.6051$. DeMoivre would probably have said that this means 24 rolls is insufficient to make it a favorable bet that at least one success will occur. In fact, the probability of at least one success in 24 rolls is about 0.4914, while the corresponding probability for 25 rolls is about 0.5055, so DeMoivre would have gotten it right.

We can get a more accurate approximation for m in terms of n by expanding the function $1/\log(1-1/n)$ in a power series. We obtain the formula

$$m = (\log 2)\left(n - \frac{1}{2} - \frac{1}{12n} - \frac{1}{24n^2} - \frac{19}{720n^3} + \dots\right).$$

Note that if we let $n = 6$ and use just the first summand in the above series, we obtain the approximation 4.159. So if DeMoivre had claimed that this shows that 4 rolls is not enough to make a favorable bet, he would have been incorrect.

It can easily be shown that if one uses the formula

$$m = \lceil (\log 2) n \rceil ,$$

then this formula will be wrong for about 34.7% of the values of n. However, if one uses the formula

$$m = \left\lceil (\log 2)\left(n - \frac{1}{2}\right)\right\rceil ,$$

then the first failure of this formula occurs at the value

$$n = 1121626023352384 .$$

The above formula gives the value

$$m = 777451915729369 ,$$

but if we use one less than this value, we obtain a probability that is still greater than 1/2:

$$0.500000000000000000000000000000000163773529998 .$$

There is a heuristic argument which suggests if we use the approximation

$$m = \left\lceil (\log 2)\left(n - \frac{1}{2} - \frac{1}{12n}\right)\right\rceil ,$$

then this formula will be incorrect for only finitely many (and perhaps no) values of n.

18. (a) The right-hand side is the sum of the probabilities of all outcomes occurring in the left-hand side plus some more because of duplication.

(b)

$$1 \geq P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

19. The left-side is the sum of the probabilities of all elements in one of the three sets. For the right side, if an outcome is in all three sets its probability is added three times, then subtracted three times, then added once, so in the final sum it is counted just once. An element that is in exactly two sets is added twice, then subtracted once, and so it is counted correctly. Finally, an element in exactly one set is counted only once by the right side.

20. We would have to have the same probability assigned to all outcomes. If this probability is 0, the sum of the probabilities would be 0 so that $P(\Omega) = 0$ instead of 1 as it should be. If this common probability is $a > 0$, then the sum of all the probabilities of the first n outcomes would be na and for large enough n this would be greater than 1, contradicting the requirement that the sum of the probabilities for all possible outcomes should be 1.

21. $7/2^{12}$

22. $\Omega = \{1, 2, 3, \ldots\}$ and the distribution is $m(n) = (5/6)^{n-1}(1/6)$. Now if $0 < x < 1$, then

$$\sum_{n=0}^{\infty} x^j = \frac{1}{1-x} \ .$$

Hence

$$(1/6)\sum_{n=1}^{\infty}(5/6)^{n-1} = (1/6) \cdot \frac{1}{1-5/6} = 1.$$

23. We have

$$\sum_{n=0}^{\infty} m(\omega_n) = \sum_{n=0}^{\infty} r(1-r)^n = \frac{r}{1-(1-r)} = 1 \ .$$

24. He just meant that if you pick a month at random within a complete 400-year cycle of the calendar the thirteenth of the month is more likely to fall on Friday than on any other day.

25. They call it a fallacy because if the subjects are thinking about probabilities they should realize that

P(Linda is bank teller and in feminist movement) $\leq P$(Linda is bank teller).

One explanation is that the subjects are not thinking about probability as a measure of likelihood. For another explanation see Exercise 53 of Section 4.1.

26. The probability that the two cards are of the same rank is $\frac{52\cdot 3}{52\cdot 51} = \frac{1}{17}$. Thus $2x + \frac{1}{17} = 1$ giving $x = \frac{8}{17}$.

27.

$$P_x = P(\text{male lives to age } x) \ = \ \frac{\text{number of male survivors at age } x}{100,000} \ .$$

$$Q_x = P(\text{female lives to age } x) \ = \ \frac{\text{number of female survivors at age } x}{100,000} \ .$$

28. (a) $\frac{1}{3}$

(b) $P_3(N) = \frac{[\frac{N}{3}]}{N}$, where $[\frac{N}{3}]$ is the greatest integer in $\frac{N}{3}$. Note that

$$\frac{N}{3} - 1 \le \left[\frac{N}{3}\right] \le \frac{N}{3} \ .$$

From this we see that

$$P_3 = \lim_{N\to\infty} P_3(N) = \frac{1}{3} \ .$$

(c) If A is a finite set with K elements then

$$\frac{A(N)}{N} \le \frac{K}{N} \ ,$$

so

$$\lim_{N\to\infty} \frac{A(N)}{N} = 0 \ .$$

On the other hand, if A is the set of all positive integers, then

$$\lim_{N\to\infty} \frac{A(N)}{N} = \lim_{N\to\infty} \frac{N}{N} = 1 \ .$$

(d) Let $N_k = 10^k - 1$. Then the integers between $N_{k-1}+1$ and N_k have exactly k digits. Thus, if k is odd, then

$$A(N_k) = (N_k - N_{k-1}) + (N_{k-2} - N_{k-3}) + \dots \ ,$$

while if k is even, then

$$A(N_k) = (N_{k-1} - N_{k-2}) + (N_{k-3} - N_{k-4}) + \dots \ .$$

Thus, if k is odd, then

$$A(N_k) \ge (N_k - N_{k-1}) = 9 \cdot 10^{k-1} \ ,$$

while if k is even, then

$$A(N_k) \le N_{k-1} < 10^{k-1} \ .$$

So, if k is odd, then

$$\frac{A(N_k)}{N_k} \geq \frac{9 \cdot 10^{k-1}}{10^k - 1} > \frac{9}{10} \ ,$$

while if k is even, then

$$\frac{A(N_k)}{N_k} < \frac{10^{k-1}}{10^k - 1} < \frac{2}{10} \ .$$

Therefore,

$$\lim_{N \to \infty} \frac{A(N)}{N}$$

does not exist.

29. (Solution by Richard Beigel)

(a) In order to emerge from the interchange going west, the car must go straight at the first point of decision, then make $4n + 1$ right turns, and finally go straight a second time. The probability $P(r)$ of this occurring is

$$P(r) = \sum_{n=0}^{\infty} (1-r)^2 r^{4n+1} = \frac{r(1-r)^2}{1-r^4} = \frac{1}{1+r^2} - \frac{1}{1+r} \ ,$$

if $0 \leq r < 1$, but $P(1) = 0$. So $P(1/2) = 2/15$.

(b) Using standard methods from calculus, one can show that $P(r)$ attains a maximum at the value

$$r = \frac{1+\sqrt{5}}{2} - \sqrt{\frac{1+\sqrt{5}}{2}} \approx .346 \ .$$

At this value of r, $P(r) \approx .15$.

30. In order to depart to the east, one must make $4n + 3$ right-hand turns in succession, and then go straight. The probability is

$$P(r) = \sum_{n=0}^{\infty} (1-r) r^{4n+3} = \frac{r^3}{(1+r)(1+r^2)} \ ,$$

if $0 \leq r < 1$. This function increases on the interval $[0, 1)$, so the maximum value of $P(r)$, if a maximum exists, must occur at $r = 1$. Unfortunately, if $r = 1$, then the car never leaves the interchange, so no maximum exists.

31. (a) Assuming that the students did not actually have a flat tire and that each student gives any given tire as an answer with probability $1/4$, then

probability that they both give the same answer is 1/4. If the students actually had a flat tire, then the probability is 1 that they both give the same answer. So, if the probability that they actually had a flat tire is p, then the probability that they both give the same answer is

$$\frac{1}{4}(1-p)+p=\frac{1}{4}+\frac{3}{4}p\,.$$

(b) In this case, they will both answer 'right front' with probability $(.58)^2$, etc. Thus, the probability that they both give the same answer is 39.8%.

Chapter 2

Continuous Probability Distributions

2.1 Simulation of Continuous Probabilities

The problems in this section are all computer programs.

2.2 Continuous Density Functions

1. (a) $f(\omega) = 1/8$ on $[2, 10]$

(b) $P([a, b]) = \frac{b-a}{8}$.

2. (a) $c = 1/48$.

   (b) $P(E) = \frac{1}{96}(b^2 - a^2)$.

   (c) $P(X > 5) = \frac{75}{96},\ P(x < 7) = \frac{45}{96}$.

   (d)

$$\begin{aligned} P(x^2 - 12x + 35 > 0) &= P(x - 5 > 0, x - 7 > 0) + P(x - 5 < 0, x - 7 < 0) \\ &= P(x > 7) + P(x < 5) = \frac{3}{4}\ . \end{aligned}$$

3. (a) $C = \frac{1}{\log 5} \approx .621$

(b) $P([a,b]) = (.621)\log(b/a)$

(c)

$$\begin{aligned} P(x > 5) &= \frac{\log 2}{\log 5} \approx .431 \\ P(x < 7) &= \frac{\log(7/2)}{\log 5} \approx .778 \\ P(x^2 - 12x + 35 > 0) &= \frac{\log(25/7)}{\log 5} \approx .791 \ . \end{aligned}$$

4. (a) .04, (b) .36, (c) .25, (d) .09.

5. (a) $1 - \frac{1}{e^1} \approx .632$

(b) $1 - \frac{1}{e^3} \approx .950$

(c) $1 - \frac{1}{e^1} \approx .632$

(d) 1

6. (a) $e^{-.01T}$, (b) $T = 100\log(2) = 69.3$.

7. (a) 1/3, (b) 1/2, (c) 1/2, (d) 1/3

8. (a) 1/8, (b) $\frac{1}{2}(1 + \log(2))$, (c) .75, (d) .25,

(e) 3/4, (f) 1/4, (g) 1/8, (h) $\pi/8$, (i) $\pi/4$.

12. 1/4.

13. $2\log 2 - 1$.

14. (a) 13/24, (b) 1/48.

15. Yes.

16. Consider the circumference to be the interval $[0,1]$, as in the hint. Let $A = 0$. There are two cases to consider; $0 < B < 1/2$, and $1/2 < B < 1$. In the first case, C must lie between 1/2 and $B + 1/2$, for otherwise there would be a gap of length greater than 1/2, corresponding to a semicircle containing none of the points. Similarly, if $1/2 < B < 1$, it can be seen that C must lie between $B - 1/2$ and 1/2. The probability of one of these two cases occurring is 1/4.

Chapter 3

Combinatorics

3.1 Permutations

1. 24

2. 1/12

3. 2^{32}

4. At this writing, 37 Presidents have died. The probability that no two people from a group of 37 (all of whom are dead) died on the same day is about .15. Thus, the probability that at least two died on the same day is .85. Yes; Jefferson, Adams, and Monroe (all signers of the Declaration of Independence) died on July 4.

5. 9, 6.

6. Since we do not get a different situation if we rotate the table we can consider one person's position as fixed, and then there are $(n-1)!$ possible arrangements for the other $n-1$ people.

7. $\frac{5!}{5^5}$.

8. Each subset S corresponds to a unique r-tuple of 0's and 1's, where a 1 in the i'th location means that i is an element of S. Since each location has two possibilities, there are 2^r r-tuples, and hence there are 2^r subsets.

10. 1/13

11. $\frac{3n-2}{n^3}$, $\frac{7}{27}$, $\frac{28}{1000}$.

12. (a) $30 \cdot 15 \cdot 9 = 4050$

(b) $4050 \cdot (3 \cdot 2 \cdot 1) = 24300$

(c) 148824

13. (a) $26^3 \times 10^3$

(b) $\binom{6}{3} \times 26^3 \times 10^3$

14. (a) $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$

(b) 60

15. $\dfrac{\binom{3}{1} \times (2^n - 2)}{3^n}$.

16. Note that we have changed the city from Atlanta to Philadelphia. The number of possible sets of initials is

$$26^4 + 26^3 + 26^2 + 26 = 475254\,,$$

but there are more than this number of people in Philadelphia.

17. $1 - \dfrac{12 \cdot 11 \cdot \ldots \cdot (12 - n + 1)}{12^n}$, if $n \leq 12$, and 1, if $n > 12$.

18. 36

20. Think of the person on your right at lunch and at dinner as determining a permutation. Do the same for the person on your left at lunch and at dinner. We have two examples of the problem of a random permutation having no fixed point. The probability of no match for large n for each random permutation would be approximately e^{-1} and if they were independent the probability of no match in either would be e^{-2}. They are not quite independent but for large n they are close enough to being independent to make this a good estimate.

21. They are the same.

22. The sample space is the set of all permutations of size 16 from the set $\{1, 2, \ldots, N\}$, where N is the number of counterfeits. If $x_1, x_2, \ldots, x_{15}, x_{16}$ are the numbers observed with a maximum of 56, then there are 16! sample points that would give rise to this observation. So

$$P(x_1, x_2, \ldots, x_{15}, x_{16}) = \frac{16!}{(N)_{16}}$$

for any $N \geq 56$ and 0 for any $N < 56$. Thus, this probability is greatest when $N = 56$. Your program should verify that Watson's guess is much better.

23. (a) $\frac{1}{n}$, $\frac{1}{n}$

(b) She will get the best candidate if the second best candidate is in the first half and the best candidate is in the secon half. The probability that this happens is greater than 1/4.

3.2 Combinatorics

1. (a) 20

(b) .0064

(c) 21

(d) 1

(e) .0256

(f) 15

(g) 10

(h) 0.04668

2. $\binom{10}{5} = 252$

3. $\binom{9}{7} = 36$

5. .998, .965, .729

6. If Charles has the ability, the probability that he wins is

$$b(10, .75, 7) + b(10, .75, 8) + b(10, .75, 9) + b(10, .75, 10) = .776.$$

If Charles is guessing, the probability that Ruth wins is

$$1 - b(10, .5, 7) - b(10, .5, 8) - b(10, .5, 9) - b(10, .5, 10) = .828.$$

7.

$$\frac{b(n,p,j)}{b(n,p,j-1)} = \frac{\binom{n}{j} p^j q^{n-j}}{\binom{n}{j-1} p^{j-1} q^{n-j+1}} = \frac{n!}{j!(n-j)!} \frac{(n-j+1)!(j-1)!}{n!} \frac{p}{q}$$

$$= \frac{(n-j+1)}{j}\frac{p}{q}$$

But $\frac{(n-j+1)}{j}\frac{p}{q} \geq 1$ if and only if $j \leq p(n+1)$, and so $j = [p(n+1)]$ gives $b(n, p, j)$ its largest value. If $p(n+1)$ is an integer there will be two possible values of j, namely $j = p(n+1)$ and $j = p(n+1) - 1$.

8. $b(30, \frac{1}{6}, 5) = \binom{30}{5}\left(\frac{1}{6}\right)^5\left(\frac{5}{6}\right)^{25} = .1921$. The most probable number of times is 5.

9. $n = 15,\ r = 7$

10. $\frac{11}{64} \approx .172$

11. Eight pieces of each kind of pie.

12. (a) $4/\binom{52}{5} \approx .0000015$

(b) $36/\binom{52}{5} \approx .000014$

(c) $624/\binom{52}{5} \approx .00024$

(d) $3744/\binom{52}{5} \approx .0014$

(e) $5108/\binom{52}{5} \approx .0020$

(f) $10200/\binom{52}{5} \approx .0039$

13. The number of subsets of $2n$ objects of size j is $\binom{2n}{j}$.

$$\frac{\binom{2n}{i}}{\binom{2n}{i-1}} = \frac{2n-i+1}{i} \geq 1 \Rightarrow i \leq n + \frac{1}{2}.$$

Thus $i = n$ makes $\binom{2n}{i}$ maximum.

14. By Stirling's formula, $n! \sim \sqrt{2\pi n}(n^n)e^{-n}$. Thus,

$$\begin{aligned} b(2n, \frac{1}{2}, n) &= \binom{20}{n}\frac{1}{2^{2n}} \\ &= \frac{2n!}{(n!)^2}\cdot\frac{1}{2^{2n}} \end{aligned}$$

$$\sim \frac{1}{2^{2n}} \frac{\sqrt{2\pi 2n}(2n)^{2n}e^{-2n}}{2\pi n(n^{2n})e^{-2n}} = \frac{1}{\sqrt{\pi n}}.$$

15. .343, .441, .189, .027.

16. There are $\binom{8}{3}$ ways of winning three games. After winning, there are $\binom{5}{3}$ ways of losing three games. After losing, there is only one way of tying two games. Thus the the total number of ways to win three games, lose three games, and tie two is $\binom{8}{3}\binom{5}{3} = 560$.

17. There are $\binom{n}{a}$ ways of putting a different objects into the 1st box, and then $\binom{n-a}{b}$ ways of putting b different objects into the 2nd and then one way to put the remaining objects into the 3rd box. Thus the total number of ways is
$$\binom{n}{a}\binom{n-a}{b} = \frac{n!}{a!b!(n-a-b)!}.$$

18. P(no student gets 2 or fewer correct) $= b(340, 7/128, 0) \approx 4.96 \cdot 10^{-9}$; P(no student gets 0 correct) $= b(340, 1/1024, 0) \approx .717$. So Prosser is right to expect at least one student with 2 or fewer correct, but Crowell is wrong to expect at least one student with none correct.

19. (a) $$\frac{\binom{4}{1}\binom{13}{10}}{\binom{52}{10}} = 7.23 \times 10^{-8}.$$

(b) $$\frac{\binom{4}{1}\binom{3}{2}\binom{13}{4}\binom{13}{3}\binom{13}{3}}{\binom{52}{10}} = .044.$$

(c) $$\frac{4!\binom{13}{4}\binom{13}{3}\binom{13}{2}\binom{13}{1}}{\binom{52}{10}} = .315.$$

20. (a) $\binom{13}{6}/\binom{52}{6} \approx .000084$

(b) $\binom{4}{3}\binom{4}{2}\binom{4}{1}/\binom{52}{6} \approx .0000047$

(c) $\binom{4}{2}\binom{13}{3}\binom{13}{3}/\binom{52}{6}$

21. $3(2^5) - 3 = 93$ (we subtract 3 because the three pure colors are each counted twice).

22. $\binom{8}{2} = 28$

23. To make the boxes, you need $n + 1$ bars, 2 on the ends and $n - 1$ for the divisions. The $n - 1$ bars and the r objects occupy $n - 1 + r$ places. You can choose any $n - 1$ of these $n - 1 + r$ places for the bars and use the remaining r places for the objects. Thus the number of ways this can be done is
$$\binom{n-1+r}{n-1} = \binom{n-1+r}{r}.$$

24. $\binom{19}{10} / \binom{29}{20} \approx .009$

25. (a) $6!\binom{10}{6}/10^6 \approx .1512$

(b) $\binom{10}{6}/\binom{15}{6} \approx .042$

26. (a) $pq,\ qp,\ p^2,\ q^2$

27. Ask John to make 42 trials and if he gets 27 or more correct accept his claim. Then the probability of a type I error is
$$\sum_{k\geq 27} b(42, .5, k) = .044,$$
and the probability of a type II error is
$$1 - \sum_{k\geq 27} b(42, .75, k) = .042.$$

28. $n = 114,\ m = 81$

29. $b(n, p, m) = \binom{n}{m} p^m(1 - p)^{n-m}$. Taking the derivative with respect to p and setting this equal to 0 we obtain $m(1-p) = p(n-m)$ and so $p = m/n$.

30. (a) $p(.5) = .5,\ p(.6) = .71,\ p(.7) = .87$

(b) Mets have a 95.2% chance of winning in a 7-game series.

31. .999996.

32. If $u = 1$, you only need to be sure to send at least one to each side. If $u = 0$, it doesn't matter what you do. Let $v = 1 - u$ and $q = 1 - p$. If $0 < v < 1$, let x be the nearest integer to

$$\frac{n}{2} - \frac{1}{2}\frac{\log(p/q)}{\log v} \ .$$

33. By Stirling's formula,

$$\frac{\binom{2n}{n}^2}{\binom{4n}{2n}} = \frac{(2n!)^2(2n!)^2}{n!^4(4n)!} \sim \frac{(\sqrt{4\pi n}(2n)^{2n}e^{-2n})^4}{(\sqrt{2\pi n}(n^n)e^{-n})^4\sqrt{2\pi(4n)}(4n)^{4n}e^{-4n}} = \sqrt{\frac{2}{\pi n}}.$$

34. Let E_i be the event that you do not get the ith player's picture. Then for any k of these events

$$P(E_{i1} \cap E_{i2} \cap \ldots \cap E_{ik}) = \left(\frac{n-k}{n}\right)^m.$$

You have $\binom{n}{k}$ ways of choosing k different E_i's. Thus the result follows from Theorem 9.

35. Consider an urn with n red balls and n blue balls inside. The left side of the identity

$$\binom{2n}{n} = \sum_{j=0}^{n}\binom{n}{j}^2 = \sum_{j=0}^{n}\binom{n}{j}\binom{n}{n-j}$$

counts the number of ways to choose n balls out of the $2n$ balls in the urn. The right hand counts the same thing but breaks the counting into the sum of the cases where there are exactly j red balls and $n - j$ blue balls.

36. (a) $\binom{n}{j}$

(b) $1 - \binom{n-j}{j}/\binom{n}{j}$

38. Consider the Pascal triangle (mod 3) for example.

0	$\underline{1}$
1	1 1
2	1 2 1
3	$\underline{1}$ 0 0 $\underline{1}$
4	1 1 0 1 1
5	1 2 1 1 2 1
6	$\underline{1}$ 0 0 $\underline{2}$ 0 0 $\underline{1}$
7	1 1 0 2 2 0 1 1
8	1 2 1 2 1 2 1 2 1
9	$\underline{1}$ 0 0 $\underline{0}$ 0 0 $\underline{0}$ 0 0 $\underline{1}$
10	1 1 0 0 0 0 0 0 0 1 1
11	1 2 1 0 0 0 0 0 0 1 2 1
12	$\underline{1}$ 0 0 $\underline{1}$ 0 0 $\underline{0}$ 0 0 $\underline{1}$ 0 0 $\underline{1}$
13	1 1 0 1 1 0 0 0 0 1 1 0 1 1
14	1 2 1 1 2 1 0 0 0 1 2 1 1 2 1
15	$\underline{1}$ 0 0 $\underline{2}$ 0 0 $\underline{1}$ 0 0 $\underline{1}$ 0 0 $\underline{2}$ 0 0 $\underline{1}$
16	1 1 0 2 2 0 1 1 0 1 1 0 2 2 0 1 1
17	1 2 1 2 1 2 1 2 1 1 2 1 2 1 2 1 2 1
18	$\underline{1}$ 0 0 $\underline{0}$ 0 0 $\underline{0}$ 0 0 2 0 0 $\underline{0}$ 0 0 $\underline{0}$ 0 0 $\underline{1}$

Note first that the entries in the third row are 0 for $0 < j < 3$. Lucas notes that this will be true for any p. To see this assume that $0 < j < p$. Note that

$$\binom{p}{j} = \frac{p(p-1)\cdots p-j+1}{j(j-1)\cdots 1}$$

is an integer. Since p is prime and $0 < j < p$, p is not divisible by any of the terms of $j!$, and so $(p-1)!$ must be divisible by $j!$. Thus for $0 < j < p$ we have $\binom{p}{j} = 0 \bmod p$. Let us call the triangle of the first three rows a **basic triangle**. The fact that the third row is

$$1\ 0\ 0\ 1$$

produces two more basic triangles in the next three rows and an inverted triangle of 0's between these two basic triangles. This leads to the 6'th row

1 0 0 2 0 0 1 .

This produces a basic triangle, a basic triangle multiplied by 2 (mod 3), and then another basic triangle in the next three rows. Again these triangles are separated by inverted 0 triangles. We can continue this way to construct the entire Pascal triangle as a bunch of multiples of basic triangles separated by inverted 0 triangles. We need only know what the mutiples are. The multiples in row np occur at positions $0, p, 2p, ..., np$. Looking at the triangle we see that the multiple at position $(mp,\, jp)$ is the sum of the multiples at positions $(j-1)p$ and jp in the $(m-1)p$'th row. Thus these multiples satisfy the same recursion relation

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

that determined the Pascal triangle. Therefore the multiple at position $(mp,\, jp)$ in the triangle is $\binom{m}{j}$. Suppose we want to determine the value in the Pascal triangle mod p at the position $(n,\, j)$. Let $n = sp + s_0$ and $j = rp + r_0$, where s_0 and r_0 are $< p$. Then the point $(n,\, j)$ is at position (s_0, r_0) in a basic triangle multiplied by $\binom{s}{r}$. Thus

$$\binom{n}{j} = \binom{s}{r}\binom{s_0}{r_0} .$$

But now we can repeat this process with the pair (s, r) and continue until $s < p$. This gives us the result:

$$\binom{n}{j} = \prod_{i=0}^{k} \binom{s_i}{r_j} (\text{mod } p) ,$$

where

$$\begin{aligned} s &= s_0 + s_1 p^1 + s_2 p^2 + \cdots + s_k p^k , \\ j &= r_0 + r_1 p^1 + r_2 p^2 + \cdots + r_k p^k . \end{aligned}$$

If $r_j > s_j$ for some j then the result is 0 since, in this case, the pair (s_j, r_j) lies in one of the inverted 0 triangles. If we consider the row $p^k - 1$ then for all k, $s_k = p-1$ and $r_k \leq p-1$ so the product will be positive resulting in no zeros in the rows $p^k - 1$. In particular for $p = 2$ the rows $p^k - 1$ will consist of all 1's.

39.

$$b(2n, \frac{1}{2}, n) = 2^{-2n} \frac{2n!}{n!n!} \quad = \quad \frac{2n(2n-1)\cdots 2 \cdot 1}{2n \cdot 2(n-1) \cdots 2 \cdot 2n \cdot 2(n-1) \cdots 2}$$

$$= \frac{(2n-1)(2n-3)\cdots 1}{2n(2n-2)\cdots 2} \ .$$

3.3 Card Shuffling

3. (a) 96.99%

(b) 55.16%

Chapter 4

Conditional Probability

4.1 Discrete Conditional Probability

2. (a) 1/2

(b) 1/4

(c) 1/2

(d) 0

(e) 1/2

3. (a) 1/2

(b) 2/3

(c) 0

(d) 1/4

4. (a) 1/2

(b) 4/13

(c) 1/13

5. (a) (1) and (2)

(b) (1)

6. 3/10

7. (a) We have

$$P(A \cap B) = P(A \cap C) = P(B \cap C) = \frac{1}{4},$$

$$P(A)P(B) = P(A)P(C) = P(B)P(C) = \frac{1}{4},$$

$$P(A \cap B \cap C) = \frac{1}{4} \neq P(A)P(B)P(C) = \frac{1}{8}.$$

(b) We have

$$P(A \cap C) = P(A)P(C) = \frac{1}{4},$$

so C and A are independent,

$$P(C \cap B) = P(B)P(C) = \frac{1}{4},$$

so C and B are independent,

$$P(C \cap (A \cap B)) = \frac{1}{4} \neq P(C)P(A \cap B) = \frac{1}{8},$$

so C and $A \cap B$ are not independent.

8. We have

$$P(A \cap B \cap C) = P(\{a\}) = \frac{1}{8},$$

and

$$P(A) = P(B) = P(C) = \frac{1}{2}.$$

Thus while

$$P(A \cap B \cap C) = P(A)P(B)P(C) = \frac{1}{8},$$

we have

$$P(A \cap B) = P(A \cap C) = P(B \cap C) = \frac{5}{16},$$

and

$$P(A)P(B) = P(A)P(C) = P(B)P(C) = \frac{1}{4}.$$

Therefore no two of these events are independent.

9. (a) 1/3

(b) 1/2

10. It is probably a reasonable estimate. One might refer to earlier life tables to see how much this number has changed over the last four or five censuses.

12. .0481

13. 1/2

14. 1/8

15. (a) $\dfrac{\binom{48}{11}\binom{4}{2}}{\binom{52}{13}-\binom{48}{13}} \approx .307$.

(b) $\dfrac{\binom{48}{11}\binom{3}{1}}{\binom{51}{12}} \approx .427$.

16.

$$\begin{aligned} P(A)P(B|A)P(C|A\cap B) &= P(A)\cdot\frac{P(A\cap B)}{P(A)}\cdot\frac{P(A\cap B\cap C)}{P(A\cap B)} \\ &= P(A\cap B\cap C)\,. \end{aligned}$$

17.

(a)

$$\begin{aligned} P(A\cap\tilde{B}) &= P(A)-P(A\cap B) \\ &= P(A)-P(A)P(B) \\ &= P(A)(1-P(B)) \\ &= P(A)P(\tilde{B})\ . \end{aligned}$$

(b) Use (a), replacing A by $\tilde{B}$ and B by A.

18.

$$\begin{gathered} P(D1|+) = 4/9 \\ P(D2|+) = 1/3 \\ P(D3|+) = 2/9\,. \end{gathered}$$

19. .273.

20. It can be shown that after n draws, $P(k$ white balls, $n+1-k$ black balls in the urn$) = 1/(n+1)$ for any $0 \le k \le n$. Thus you are equally likely to have any proportion of white balls after n draws. In fact, the fraction of white balls will tend to a limit but this limit is a random number. This rather suprising fact is a consequence of the fact that the Polya Urn

model is mathematicallly exactly the same as the following apparently very different model. You have a coin where the probability of heads p is chosen by **rnd**. Once this random p is chosen, the coin is tosses n times. If a head turns up you say you have a white ball, and if a tail turns up you have a black ball. Then the probability for any particular sequence of colors for the balls is exactly the same as the probability of this sequence occurring in the Polya urn model. In the coin model it is obvious that the proportion of heads will tend to a limit which is again a random number since it just depends upon what kind of a coin was chosen by **rnd**. This random coin model will be discussed more in the next section.

21. No.

22. 1/2

23. Put one white ball in one urn and all the rest in the other urn. This gives a probability of nearly 3/4, in particular greater than 1/2, for obtaining a white ball which is what you would have with an equal number of balls in each urn. Thus the best choice must have more white balls in one urn than the other. In the urn with more white balls, the best we can do is to have probability 1 of getting a white ball if this urn is chosen. In the urn with less white balls than black, the best we can do is to have one less white ball than black and then to have as many white balls as possible. Our solution is thus best for the urn with more white balls than black and also for the urn with more black balls than white. Therefore our solution is the best we can do.

24. $P(A \text{ head on the } j\text{th trial and a total of } k \text{ heads in } n \text{ trials}) = \left(\frac{1}{2}\right)^n \binom{n-1}{k-1}$.

$P(\text{Exactly } k \text{ heads in } n \text{ trials}) = \left(\frac{1}{2}\right)^n \binom{n}{k}$.

Thus

$$P(\text{Head on } j\text{'th trial } \mid k \text{ heads in } n \text{ trials}) = \frac{\binom{n-1}{k-1}}{\binom{n}{k}} = \frac{k}{n} .$$

25. We must have

$$p\binom{n}{j}p^k q^{n-k} = p\binom{n-1}{k-1}p^{k-1}q^{n-k} .$$

This will be true if and only if $np = k$. Thus p must equal k/n.

26. $P(A|B) = P(B|A)$ implies that $P(A) = P(B)$.

Thus, since since $P(A \cap B) > 0$,

$$1 = P(A \cup B) = P(A) + P(B) - P(A \cap B) < 2P(A) \ ,$$

and $P(A) > \frac{1}{2}$.

27.

(a) P(Pickwick has no umbrella, given that it rains)$= \dfrac{2}{9}$.

(b) P(Pickwick brings his umbrella, given that it doesn't rain)$= \dfrac{5}{9}$.

Note that the statement in part (b) of this problem was changed in the errata list for the book.

28. The most obvious objection is the assumption that all of the events in question are independent. A more subtle objection is that, since Los Angeles is so large, it is reasonable to ask for the probability that there is a second couple with the same discription, given that there is one such couple. This probability is not so small. (See Exercise 23 of Section 9.3.)

29.

$$P(\text{Accepted by Dartmouth} \mid \text{Accepted by Harvard}) = \frac{2}{3} \, .$$

The events 'Accepted by Dartmouth' and 'Accepted by Harvard' are not independent.

30. Neither has a convincing argument based upon comparing grouped data only. You need more information. For example, suppose that each defective bulb cost \$10 whether a regular or a softglow bulb. Then in making 3000 bulbs the loss to A is \$130 and to B is \$110 so B has a smaller loss than A. But suppose that a defective regular bulb results in a loss of \$20 and a defective softglow bulb in a loss of \$10. Now making 3000 bulbs A has a loss of $2 \times \$20 + \$11 \times \$10 = \$150$ while B has a loss of $5 \times \$20 + 6 \times \$10 = \$160$. Thus A has a smaller loss than B. Paradoxes caused by comparing percentages when data are grouped are called **Simpson paradoxes**. An interesting real life example can be found in Parsani, and Purvis, W.W. Norton l978. In a study of the admission to graduate school at the University of California, Berkeley in 1973 it was found that about 44% of the men who applied were admitted, but 35% of the women who applied were admitted. Suspecting sex bias, an attempt was made to locate where it occurred by examining the individual departments. But within individual departments there did not seem to be any bias. Indeed the only department with a significant difference was one that favored women. Again more information is needed. In this case the explantion lay

in the fact that the women applied to majors that were difficult to get into (lower acceptance rate) and men applied generally to the majors that were easy to get in (high acceptance rate). In each of these examples you are interested in comparing one trait in the presence of a second confounding trait. In the first example it was good or bad bulb confounded by the type of bulb and in the second example it was sex confounded by the major. Freedman et al. discuss how to control the confounding trait to make a more valid comparison.

31. The probability of a 60 year old male living to 80 is .41, and for a female it is .62.

32. (a) pq

(b) $1-(1-p)(1-q)$

(c) .958

33. You have to make a lot of calculations, all of which are like this:

$$\begin{aligned} P(\tilde{A}_1 \cap A_2 \cap A_3) &= P(A_2)P(A_3) - P(A_1)P(A_2)P(A_3) \\ &= P(A_2)P(A_3)(1 - P(A_1)) \\ &= P(\tilde{A}_1)P(A_2)P(A_3). \end{aligned}$$

34. $P_{X_j} = \begin{pmatrix} 1 & 0 \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix}.$

They are not independent. For example, if we know that $X_1 = 1, X_2 = 1, \text{and } X_3 = 1$, then it must be the case that $X_4 = 1$.

35. The random variables X_1 and X_2 have the same distributions, and in each case the range values are the integers between 1 and 10. The probability for each value is 1/10. They are independent. If the first number is not replaced, the two distributions are the same as before but the two random variables are not independent.

36. $p = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{11}{36} & \frac{9}{36} & \frac{7}{36} & \frac{5}{36} & \frac{3}{36} & \frac{1}{36} \end{pmatrix}.$

37.

$$\begin{aligned} P(\max(X,Y) = a) &= P(X = a, Y \le a) + P(X \le a, Y = a) - P(X = a, Y = a). \\ P(\min(X,Y) = a) &= P(X = a, Y > a) + P(X > a, Y = a) + P(X = a, Y = a). \end{aligned}$$

Thus $P(\max(X,Y) = a) + P(\min(X,Y) = a) = P(X = a) + P(Y = a)$

and so $u = t + s - r$.

38. (a) $p_X = \begin{pmatrix} 0 & 1 & 2 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}, \quad p_Y = \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$

(b) $p_Z = \begin{pmatrix} 0 & 1 & 2 & 3 \\ \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8} \end{pmatrix}.$

(c) $p_W = \begin{pmatrix} -1 & 0 & 1 & 2 \\ \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8} \end{pmatrix}.$

39. (a) 1/9

(b) 1/4

(c) No

(d) $p_Z = \begin{pmatrix} -2 & -1 & 0 & 1 & 2 & 4 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}$

40. $p = 1/2 \quad p_X = \begin{pmatrix} 0 & 1 \\ 1/2 & 1/2 \end{pmatrix} \quad p_Y = \begin{pmatrix} 3 & 4 & 5 \\ 1/4 & 3/8 & 3/8 \end{pmatrix}$ Independent

$p = 2/3 \quad p_X = \begin{pmatrix} 0 & 1 \\ 17/81 & 64/81 \end{pmatrix} \quad p_Y = \begin{pmatrix} 3 & 4 & 5 \\ 1/3 & 10/27 & 8/27 \end{pmatrix}$ Not independent

42. Let $u = N - r$ and $v = N - s$ be the number of games that A and B, respectively, must win to win the series. Then the series will surely be over in $u+v-1$ games, so Fermat extended the game to assure this many plays. The player with the most points in the extended game wins. Therefore,

$$P(r,s) = P(u,v) = \sum_{j=u}^{u+v-1} \binom{u+v-1}{j} p^j q^{u+v-1-j}.$$

An alternative formula can be derived without extending the game. To win the series in $u+j$ games, A must win $u-1$ games among the first $u+j-1$ games and then win the $(u+j)$th game with $j \le v-1$. Thus,

$$P(r,s) = P(u,v) = \sum_{j=0}^{v-1} \binom{u+j-1}{j} p^u q^j.$$

43. .710.

44. (a) **First convention.** If you serve $N+1$ times, then your opponent must serve N times. The total number of points played is $2N+1$, so one of you must have won at least $N+1$ points. That is a contradition, since the game is over when a player has won N points.

Second convention. If you serve $N+1$ times, then except for the first time, before each time you serve, you have won a point. Thus at the $(N+1)$st time you serve you have already won N points. The game should have already ended. Therefore, you serve at most N times. Before each serve of your apponent he won the previous point. Thus, as he serves for the Nth time he has already won N points. Therefore, your opponent serves at most $N-1$ times.

(b) Since the total number of points for the two players is $2N-1$, and one player has already got N points before the game is extended, the other can get at most $N-1$ points in the extended game and hence not change the winner.

(c) For the extended game, probabilistically, the two methods are the same: in one, we have N Bernoulli trials with probability p for success and in the other we have $N-1$ trials with probability $\bar{p}$ for success, and in either method you win if you win the most points.

45.

(a) The probability that the first player wins under either service convention is equal to the probability that if a coin has probability p of coming up heads, and the coin is tossed $2N+1$ times, then it comes up heads more often than tails. This probability is clearly greater than .5 if and only if $p > .5$.

(b) If the first team is serving on a given play, it will win the next point if and only if one of the following sequences of plays occurs (where 'W' means that the team that is serving wins the play, and 'L' means that the team that is serving loses the play):

$$W,\ LLW,\ LLLLW,\ \ldots\ .$$

The probability that this happens is equal to

$$p + q^2p + q^4p + \ldots\ ,$$

which equals

$$\frac{p}{1-q^2} = \frac{1}{1+q}\ .$$

Now, consider the game where a 'new play' is defined to be a sequence of plays that ends with a point being scored. Then the service convention is that at the beginning of a new play, the team that won the last new play serves. This is the same convention as the second convention in the preceding problem.

From part a), we know that the first team to serve under the second service convention will win the game more than half the time if and only if $p > .5$.

In the present case, we use the new value of p, which is $1/(1+q)$. This is easily seen to be greater than .5 as long as $q < 1$. Thus, as long as $p > 0$, the first team to serve will win the game more than half the time.

46. $P(X = i) = P(Y = i) = \dfrac{\binom{4}{i}\binom{5-i}{48}}{\binom{52}{5}}$,

$P(X = i,\ Y = j) = \dfrac{\binom{4}{i}\binom{4}{j}\binom{44}{5-i-j}}{\binom{52}{5}}$, if $i \le 4,\ j \le 4$, and $i + j \le 5$,

$P(X = i,\ Y = j) = 0$, otherwise.

47. (a)

$$\begin{aligned} P(Y_1 = r, Y_2 = s) &= P(\Phi_1(X_1) = r, \Phi_2(X_2) = s) \\ &= \sum_{\substack{\Phi_1(a)=r \\ \Phi_2(b)=s}} P(X_1 = a, X_2 = b)\ . \end{aligned}$$

(b) If X_1, X_2 are independent, then

$$\begin{aligned} P(Y_1 = r, Y_2 = s) &= \sum_{\substack{\Phi_1(a)=r \\ \Phi_2(b)=s}} P(X_1 = a, X_2 = b) \\ &= \sum_{\substack{\Phi_1(a)=r \\ \Phi_2(b)=s}} P(X_1 = a)P(X_2 = b) \\ &= \Big(\sum_{\Phi_1(a)=r} P(X_1 = a)\Big)\Big(\sum_{\Phi_2(b)=s} P(X_2 = b)\Big) \\ &= P(\Phi_1(X_1) = r)P(\Phi_2(X_2) = s) \\ &= P(Y_1 = r)P(Y_2 = s)\ , \end{aligned}$$

so Y_1 and Y_2 are independent.

48.

$$\begin{aligned} \sum_{\omega\in\Omega} m_E(\omega) &= \frac{1}{P(E)} \sum_{\omega\in\Omega} P(\omega \cap E) \\ &= \frac{1}{P(E)} P(E) = 1\ . \end{aligned}$$

49. $P(\text{both coins turn up using (a)}) = \frac{1}{2}p_1^2 + \frac{1}{2}p_2^2$.

$P(\text{both coins turn up heads using (b)}) = p_1 p_2$.

Since $(p_1 - p_2)^2 = p_1^2 - 2p_1p_2 + p_2^2 > 0$, we see that $p_1p_2 < \frac{1}{2}p_1^2 + \frac{1}{2}p_2^2$, and so (a) is better.

50. For any sequence $B_1, \ldots, B_n$ with $B_k = A_k$ or $\tilde{A}_k$

$$P(B_1 \cap \cdots \cap B_n) = P(B_1)P(B_2)\cdots P(B_n) > 0 \ .$$

Thus there is at least one sample point ω in each of the sets

$$B_1 \cap B_2 \cap \cdots \cap B_n \ .$$

Since there are 2^n such subsets, there must be at least this many sample points.

51.

$$\begin{aligned} P(A) &= P(A|C)P(C) + P(A|\tilde{C})P(\tilde{C}) \\ &\geq P(B|C)P(C) + P(B|\tilde{C})P(\tilde{C}) = P(B) \ . \end{aligned}$$

52.

$$\begin{aligned} P(\text{coin not found in the } i'\text{th box}) &= P(\text{coin not in } i'\text{th box}) + \\ &\quad\ P(\text{coin in } i'\text{th box but not found}) \\ &= 1 - p_i + (1 - a_i)p_i \\ &= 1 - a_ip_i \ . \end{aligned}$$

Thus

$$P(\text{coin is in } \ j'\text{th } \ \text{box} \mid \text{not found in } \ i'\text{th } \ \text{box}) = p_j/(1 - a_ip_i) \ \text{ if } j \neq i.$$

$$\begin{aligned} &\text{P(coin is in the } \ i'\text{th } \ \text{box} \mid \text{not found in the } \ i'\text{th } \ \text{box)} \\ &\quad = \ 1 - \sum_{j \neq i} \text{P(coin is in } \ j'\text{th } \ \text{box} \mid \text{not found in } \ i'\text{th } \ \text{box)} \\ &\quad = \ 1 - \frac{1 - p_i}{(1 - a_ip_i)} \\ &\quad = \ \frac{(1 - a_i)p_i.}{1 - a_ip_i}. \end{aligned}$$

53. We assume that John and Mary sign up for two courses. Their cards are dropped, one of the cards gets stepped on, and only one course can be read on this card. Call card I the card that was not stepped on and on which the registrar can read government 35 and mathematics 23; call card II the card that was stepped on and on which he can just read mathematics 23. There are four possibilities for these two cards. They are:

Card I	Card II	Prob.	Cond. Prob.
Mary(gov,math)	John(gov, math)	.0015	.224
Mary(gov,math)	John(other,math)	.0025	.373
John(gov,math)	Mary(gov,math)	.0015	.224
John(gov,math)	Mary(other,math)	.0012	.179

In the third column we have written the probability that each case will occur. For example, for the first one we compute the probability that the students will take the appropriate courses: $.5 \times .1 \times .3 \times .2 = .0030$ and then we multiply by 1/2, the probability that it was John's card that was stepped on. Now to get the conditional probabilities we must renormalize these probabilities so that they add up to one. In this way we obtain the results in the last column. From this we see that the probability that card I is Mary's is .597 and that card I is John's is .403, so it is more likely that that the card on which the registrar sees Mathematics 23 and Government 35 is Mary's.

54.

(a) Let

$$A = \{(a, \bullet) : a \in \text{a subset } \mathcal{A} \text{ of } \{\clubsuit, \diamondsuit, \heartsuit, \spadesuit\}, \bullet \in \{2, 3, \ldots, J, Q, K, A\}\}$$

be a suit event and

$$B = \{(\bullet, b) : b \in \text{a subset } \mathcal{B} \text{ of } \{2, 3, \ldots, \text{J}, \text{Q}, \text{K}, \text{A}\}, \bullet \in \{\clubsuit, \diamondsuit, \heartsuit, \spadesuit\}\}$$

be a rank event. Then

$$\begin{aligned} P(A) &= \frac{\text{size of } \mathcal{A}}{4} \\ P(B) &= \frac{\text{size of } \mathcal{B}}{13} \\ P(A \cap B) &= P\{(a, b) : a \in \mathcal{A}, b \in \mathcal{B}\} = \frac{(\text{size of } \mathcal{A})(\text{size of } \mathcal{B})}{52} = P(A)P(B)\ . \end{aligned}$$

(b) The possible sizes of a rank event are $(4i + 3j)$, where $i = 0, \ldots, 12$ and $j = 0$ or 1, and the possible sizes of a suit event are $(13m + 12n)$, where $m = 0,1,2,3$, and $n = 0,1$. By the hint we must have

$$\frac{(4i + 3j)(13m + 12n)}{51}$$

an integer. This means that either $4i + 3$ or $m + 12n$ must be divisible by 17. We show that this is not possible. Assume for example that $4i + 3j = 17k$ where k is 1 or 2. Since $17k = 16k + k$, when $17k$ is divided by 4 we would get a remainder of k, that is, 1 or 2. But when $4i + 3j$ is divided by 4 we get a remainder of $3j$, which is 0 or 3 depending on the value of j. Therefore we cannot have $4i + 3j = 17k$, for k= 1 or 2 and j = 0 or 1. Similarly, assume that $13m + 12n = 17k$ with k = 1 or 2. Since $17k = 13k + 4k$, when $17k$ is divided by 13 we get a remainder of 4 or 8 depending on the value of k, but when $13m$ + $12n$ is divided by 13 we get a remainder of 0 or 12 depending on the value of n. Thus we cannot have $17k = 13m + 12n$ for k = 1 or 2 and n = 0 or 1.

(c)

$$\begin{aligned} A &= \{(\spadesuit, 2), (\spadesuit, 3), (\spadesuit, 4)\} \\ B &= \{(\spadesuit, 4), \ldots, (\spadesuit, 7), (\heartsuit, 1), \ldots, (\heartsuit, K)\}. \end{aligned}$$

(d) Let a be the size of A, b the size of B, and c the size of $A \cap B$. Then A and B independent implies that

$$\frac{c}{53} = \frac{a}{53} \cdot \frac{b}{53}.$$

Thus $ab = 53c$. But, since 53 is prime, this means that either a or b must be 53, which means that either A or B must be trivial.

55.

$$P(R_1) = \frac{4}{\binom{52}{5}} = 1.54 \times 10^{-6}.$$

$$P(R_2 \cap R_1) = \frac{4 \cdot 3}{\binom{52}{5}\binom{47}{5}}.$$

Thus

$$P(R_2 \mid R_1) = \frac{3}{\binom{47}{5}} = 1.96 \times 10^{-6}.$$

Since $P(R_2|R_1) > P(R_1)$, a royal flush is attractive.

$$P(\text{player 2 has a full house}) = \frac{13 \cdot 12\binom{4}{3}\binom{4}{2}}{\binom{52}{5}}.$$

P(player 1 has a flush and player 2 has a full house) =

$$\frac{4\cdot 8\cdot 7\binom{4}{3}\binom{4}{2}+4\cdot 8\cdot 5\binom{4}{3}\cdot\binom{3}{2}+4\cdot 5\cdot 8\cdot\binom{3}{3}\binom{4}{2}+4\cdot 5\cdot 4\binom{3}{3}\binom{3}{2}}{\binom{52}{5}\binom{47}{5}} .$$

Taking the ratio of these last two quantities gives:

$$\text{P(player 1 has a royal flush} \mid \text{player 2 has a full house)} = 1.479 \times 10^{-6}.$$

Since this probability is less than the probability that player 1 has a royal flush (1.54×10^{-6}), a full house repels a royal flush.

56.

$$\begin{aligned} P(B|A) > P(B) \quad &\Leftrightarrow P(B\cap A) > P(A)P(B) \\ &\Leftrightarrow P(A|B) = \tfrac{P(A\cap B)}{P(B)} > P(A) . \end{aligned}$$

57.

$$\begin{gathered} P(B|A) \le P(B) \text{ and } P(B|A) \ge P(A) \\ \Leftrightarrow P(B\cap A) \le P(A)P(B) \text{ and } P(B\cap A) \ge P(A)P(B) \\ \Leftrightarrow P(A\cap B) = P(A)P(B) . \end{gathered}$$

58.

$$\begin{aligned} P(A|B) > P(A) \quad &\Leftrightarrow P(A\cap B) > P(A)P(B) \\ &\Leftrightarrow P(A\cap B) - P(A)P(A\cap B) > P(A)P(B) - P(A)P(B\cap A) \\ &\Leftrightarrow P(A\cap B)P(\tilde{A}) > P(A)P(B\cap\tilde{A}) \\ &\Leftrightarrow \tfrac{P(A\cap B)}{P(A)} > \tfrac{P(B\cap\tilde{A})}{P(\tilde{A})} \\ &\Leftrightarrow P(B|A) > P(B|\tilde{A}) . \end{aligned}$$

59. Since A attracts B, $P(B|A) > P(A)$ and

$$P(B\cap A) > P(A)P(B) ,$$

and so

$$P(A) - P(B\cap A) < P(A) - P(A)P(B) .$$

Therefore,

$$P(\tilde{B}\cap A) < P(A)P(\tilde{B}) ,$$

$$P(\tilde{B}|A) < P(\tilde{B}) ,$$

and A repels $\tilde{B}$.

60. If A attracts B and C, and A repels $B \cap C$ then we have

$$P(A \cap B) > P(A)P(B) \ ,$$
$$P(A \cap C) > P(A)P(C) \ ,$$
$$P(B \cap C \cap A) < P(B \cap C)P(A) \ .$$

Thus

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

Therefore

$$P(B \cup C|A) > P(B \cup C) \ .$$

Here is an example in which A attracts B and C and repels $B \cup C$. Let

$$\Omega = \{a, b, c, d\} \ ,$$

$$p(a) = .2,\ p(b) = .25,\ p(c) = .25,\ p(d) = .3 \ .$$

Let

$$A = \{a, d\}, B = \{b, d\}, C = \{c, d\} \ .$$

Then

$$P(B|A) = .6 > P(B) = .55 \ ,$$
$$P(C|A) = .6 > P(C) = .55 \ ,$$

and

$$P(B \cup C|A) = .6 < P(B \cup C) = .8 \ .$$

61. Assume that A attracts B_1, but A does not repel any of the B_j's. Then

$$P(A \cap B_1) > P(A)P(B_1),$$

and

$$P(A \cap B_j) \geq P(A)P(B_j), \qquad 1 \leq j \leq n.$$

Then

$$\begin{aligned} P(A) &= P(A \cap \Omega) \\ &= P(A \cap (B_1 \cup \ldots \cup B_n)) \\ &= P(A \cap B_1) + \cdots + P(A \cap B_n) \\ &> P(A)P(B_1) + \cdots + P(A)P(B_n) \\ &= P(A)\Big(P(B_1) + \cdots + P(B_n)\Big) \\ &= P(A) \ , \end{aligned}$$

which is a contradiction.

4.2 Continuous Conditional Probability

1. (a) 2/3

(b) 1/3

(c) 1/2

(d) 1/2

2. (a) $1 - e^{-.9} = .593$

(b) $1 - e^{-.5} = .393$

(c) 1

(d) $(1 - e^{-1})/(1 - e^{-2}) = .731$

3. (a) .01

(b) $e^{-.01\,T}$ where T is the time after 20 hours.

(c) $e^{-.2} \approx .819$

(d) $1 - e^{-.01} \approx .010$

4. (a) 1/2

(b) 1/4

(c) 3/4

(d) 1/2

5. (a) 1

(b) 1

(c) 1/2

(d) $\pi/8$

(e) 1/2

6. (a) $$f(x) = \begin{cases} 4/3, & \text{if } 1/4 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

(b) $$f(t) = \begin{cases} e^{-t}/(e^{-1} - e^{-10}), & \text{if } 1 < t < 10, \\ 0, & \text{otherwise.} \end{cases}$$

(c) $$f(x, y) = \begin{cases} \pi/50, & \text{if } (x, y) \text{ is in upper half of the target,} \\ 0, & \text{otherwise.} \end{cases}$$

(d) $$f(x,y) = \begin{cases} 2, & \text{if } x > y, \\ 0, & \text{otherwise.} \end{cases}$$

7. $P(X > \frac{1}{3}, Y > \frac{2}{3}) = \int_{\frac{1}{3}}^{1} \int_{\frac{2}{3}}^{1} dydx = \frac{2}{9}$.

But $P(X > \frac{1}{3})P(Y > \frac{2}{3}) = \frac{2}{3} \cdot \frac{1}{3}$, so X and Y are independent.

8. a and b; c and d.

10. (b) Let Z be a number chosen with uniform density on the interval $[0, a]$. We find the density for $\max(Z, a - Z)$. This density is nonzero only on the interval $[a/2,a]$. For x in this interval:

$$\begin{aligned} P(\max(Z, a - Z)) \le x) &= P(a - Z \ge Z, a - Z \le x) + \\ &\quad P(Z > a - Z, Z \le x) \\ &= P(Z \ge a - x, Z \le \frac{a}{2}) + P(Z > \frac{a}{2}, Z \le x) \\ &= \frac{2x - a}{a} \ . \end{aligned}$$

Taking $a = 1$, we see that the density for the length of the largest stick from the first cut is uniform on the interval $[\frac{1}{2}, 1]$. Assume that the length of this longest piece is X. Let Y be position on the interval $[0, X]$ of the second cut. Then we obtain a triangle if $\max(Y, X - Y) \le \frac{1}{2}$. But by our first computation with $a = X$ and $x = \frac{1}{2}$, we see that

$$P(\max(Y, X - Y)) \le \frac{1}{2} = \Bigl(\frac{1 - X}{X}\Bigr) \ .$$

Thus

$$P(\text{triangle}) = 2 \int_{\frac{1}{2}}^{1} \Bigl(\frac{1 - x}{x}\Bigr) dx = 2 \log 2 - 1 \ .$$

11. If you have drawn n times (total number of balls in the urn is now $n + 2$) and gotten j black balls, (total number of black balls is now $j + 1$), then the probability of getting a black ball next time is $(j + 1)/(n + 2)$. Thus at each time the conditional probability for the next outcome is the same in the two models. This means that the models are determined by the same probability distribution, so either model can be used in making predictions. Now in the coin model, it is clear that the proportion of heads will tend to the unknown bias p in the long run. Since the value of p was assumed to be unformly distributed, this limiting value has a random value between 0 and 1. Since this is true in the coin model, it is also true in the Polya Urn model for the proportion of black balls.(See Exercise 20 of Section 4.1.)

12. A new beta density with $\alpha = 6$ and $\beta = 9$. It will be successful next time with probability .4.

4.3 Paradoxes

1. 2/3

2. Let M be the event that the hand has an ace. Let N be the event that the hand has at least two aces. Then

$$\begin{aligned} P(N|M) &= \frac{P(N \cap M)}{P(M)} = \frac{P(N)}{P(M)} = \frac{\binom{4}{2}\binom{48}{11} + \binom{4}{3}\binom{48}{10} + \binom{4}{4}\binom{48}{9}}{\binom{52}{13} - \binom{48}{13}} \\ &\approx .3696\,. \end{aligned}$$

Let S be the event that the hand has the ace of hearts. Then

$$\begin{aligned} P(N|S) &= \frac{P(S \cap T)}{P(S)} = \frac{\binom{3}{1}\binom{48}{11} + \binom{3}{2}\binom{48}{10} + \binom{3}{3}\binom{48}{9}}{\binom{52}{13} - \binom{51}{13}} \\ &\approx .5612\,. \end{aligned}$$

3. (a) Consider a tree where the first branching corresponds to the number of aces held by the player, and the second branching corresponds to whether the player answers 'ace of hearts' or anything else, when asked to name an ace in his hand. Then there are four branches, corresponding to the numbers 1, 2, 3, and 4, and each of these except the first splits into two branches. Thus, there are seven paths in this tree, four of which correspond to the answer 'ace of hearts.' The conditional probability that he has a second ace, given that he has answered 'ace of hearts,' is therefore

$$\frac{\left(\left(\binom{48}{12} + \frac{1}{2}\binom{3}{1}\binom{48}{11} + \frac{1}{3}\binom{3}{2}\binom{48}{10} + \frac{1}{4}\binom{3}{3}\binom{48}{9}\right)\Big/\binom{52}{13}\right)}{\left(\binom{51}{12}\Big/\binom{52}{13}\right)} \approx .6962\,.$$

(b) This answer is the same as the second answer in Exercise 2, namely .5612.

5. Let $x = 2^k$. It is easy to check that if $k \geq 1$, then
$$\frac{p_{x/2}}{p_{x/2} + p_x} = \frac{3}{4} \ .$$
If $x = 1$, then
$$\frac{p_{x/2}}{p_{x/2} + p_x} = 0 \ .$$
Thus, you should switch if and only if your envelope contains 1.

6. The sample space consists of all hands of size 13 (for the left-hand opponent) that can be drawn from the 26 hands not held by the declarer or the dummy (our two hands). If A_k denotes the event that our left-hand opponent (LHO) has exactly k cards in the suit in question (say the suit is hearts), then
$$P(A_k) = \frac{\binom{4}{k}\binom{22}{13-k}}{\binom{26}{13}} \ .$$
At the point in the problem where we need to make a decision, we know that LHO either has only the queen or the queen and jack. The probabilities of these two events are 143/2300 and 156/2300. Thus, it is more likely that LHO has the jack, so we should play the king.

As an interesting follow-up to this question, suppose we have played with LHO for many years, and from experience, we know that in this situation, he plays the queen on the first trick half the time and the jack on the first trick half the time. Now, one can compute, using a decision tree, that if he plays the queen on the first trick, the probability that he has the jack is 143/455. So, in this case, we should finesse on the second trick (i.e. assume that he doesn't have the jack).

More generally, if we know that in this situation he plays the queen with probability p and the jack with probability $1 - p$, then the probability he has the jack, assuming he plays the queen on the first trick, is
$$\frac{p(143/299)}{p(143/299) + (156/299)} \ .$$

Chapter 5

Important Distributions and Densities

5.1 Important Distributions

1. (a), (c), (d)

2. Yes.

3. Assume that X is uniformly distributed, and let the countable set of values be $\{\omega_1, \omega_2, \ldots\}$. Let p be the probability assigned to each outcome by the distribution function f of X. If $p > 0$, then
$$\sum_{i=1}^{\infty} f(\omega_i) = \sum_{i=1}^{\infty} p \ ,$$
and this last sum does not converge. If $p = 0$, then
$$\sum_{i=1}^{\infty} f(\omega_i) = 0 \ .$$
So, in both cases, we arrive at a contradiction, since for a distribution function, we must have
$$\sum_{i=1}^{\infty} f(\omega_i) = 1 \ .$$

4. One way to help decide whether the given experiment will result in a random subset is to ask whether one might reasonably expect any of the possible subsets to occur if the experiment is performed. Since some close friends almost always eat lunch together, the first experiment will almost never give a subset which has as a member exactly one of a pair of close

friends. The second experiment will always give the same subset when performed on the same student body. In addition, Social Security numbers are assigned based upon geography, among other things. Thus, depending upon the use we are going to make of this subset, this experiment may or may not give an adequate subset. Finally, the third experiment will probably return each subset with approximately the same probability. (A tenth-floor window would be much better.)

5. (b) Ask the Registrar to sort by using the sixth, seventh, and ninth digits in the Social Security numbers.

(c) Shuffle the cards 20 times and then take the top 100 cards. (Can you think of a method of shuffling 3000 cards?

6. The distribution function of Y is given by

$$f(x) = \frac{(k-x+1)^n - (k-x)^n}{k^n} ,$$

for $1 \le x \le k$. The numerator counts the number of n-tuples, all of whose entries are at least x, and subtracts the number of n-tuples, all of whose entries are at least $x+1$.

7. (a) $p(n) = \frac{1}{6}\left(\frac{5}{6}\right)^{n-1}$

(b) $P(T > 3) = (\frac{5}{6})^3 = \frac{125}{216}$.

(c) $P(T > 6 \mid T > 3) = (\frac{5}{6})^3 = \frac{125}{216}$.

8. $\frac{1}{8}$.

9. (a) 1000

(b) $\dfrac{\binom{100}{10}\binom{N-100}{90}}{\binom{N}{100}}$

(c) $N = 999$ or $N = 1000$

10. (a) $\dfrac{\binom{N}{k}\binom{N-k}{n_1-k}\binom{N-n_1}{n_2-k}}{\binom{N}{n_1}\binom{N}{n_2}}$

(b) Let p_N denote the probability that $X = n_{12}$, given that the population size is N. Then p_N equals the expression in the answer to part (a), with k replaced by n_{12}. After some gruesome algebra, one obtains

$$\frac{p_{N+1}}{p_N} = \frac{N^2 - (n_1+n_2-2)N + (n_1-1)(n_2-1)}{N^2 - (n_1+n_2-n_{12}-2)N + (n_{12}+1-n_1-n_2} .$$

Let a_N and b_N denote the numerator and denominator of this expression. We want the smallest value of N for which the expression is less than or equal to 1, or equivalently, we want the smallest value of N for which $a_N \le b_N$. If we solve this inequality for N, we obtain

$$N = \left\lceil \frac{n_1 n_2}{n_{12}} \right\rceil - 1 \ .$$

13. .7408, .2222, .0370

14. (a) $e^{-10} \approx 4.54 \times 10^{-5}$

(b) We need $e^{-n/1000} = 1/2$, or $n = 1000 \cdot \log 2 \approx 694$.

16. P(miss 0 calls) + P(miss 1 call) = .0498+.1494 = .1992.

17. 649741

18. (a) $m = 600 \times \frac{1}{500}$ so $P(\text{no raisins}) = e^{-m} = .301$.

(b) $m = 400 \times \frac{1}{500}$, so $P(\text{exactly two chocolate chips}) = \frac{m^2 e^{-m}}{2!} = .144$.

(c) $m = 1000 \times \frac{1}{500}$, so

$$P(\text{at least two bits}) = 1 - P(0 \text{ bits}) - P(1 \text{ bit}) = 1 - .1353 - .2707 = .594.$$

19. The probability of at least one call in a given day with n hands of bridge can be estimated by $1 - e^{-n\cdot(6.3\times10^{-12})}$. To have an average of one per year we would want this to be equal to $\frac{1}{365}$. This would require that n be about 400,000,000 and that the players play on the average 8,700 hands a day. Very unlikely! It's much more likely that someone is playing a practical joke.

20. $e^{-5} \approx .00674$

21. (a) $b(32, j, 1/80) = \binom{32}{j}\left(\frac{1}{80}\right)^j\left(\frac{79}{80}\right)^{32-j}$

(b) Use $\lambda = 32/80 = 2/5$. The approximate probability that a given student is called on j times is $e^{-2/5}(2/5)^j/j!$. Thus, the approximate probability that a given student is called on more than twice is

$$1 - e^{-2/5}\left(\frac{(2/5)^0}{0!} + \frac{(2/5)^1}{1!} + \frac{(2/5)^2}{2!}\right) \approx .0079 \ .$$

22. .0077

23.

$$P(\text{outcome is } j+1)/\text{P}(\text{outcome is } j) \;=\; \frac{m^{j+1}e^{-m}}{(j+1)!}\Big/\frac{m^j e^{-m}}{j!} = \frac{m}{j+1}.$$

Thus when $j+1 \le m$, the probability is increasing, and when $j+1 \ge m$ it is decreasing. Therefore, $j = m$ is a maximum value. If m is an integer, then the ratio will be one for $j = m-1$, and so both $j = m-1$ and $j = m$ will be maximum values. (cf. Exercise 7 of Chapter 3, Section 2)

24. The probability that Kemeny receives no mail on a given weekday can be estimated by $e^{-10} = 4.54 \times 10^{-5}$. Thus in ten years, the probability that at least one day brings no mail can be estimated by $1 - e^{-3000 \cdot 4.54 \times 10^{-5}} =$ .127. Thus he finds that the probability is .127, which is inconclusive.

25. Using the Poisson approximation, we find that without paying the meter Prosser pays

$$2\frac{5^2e^{-5}}{2!} + (2+5)\frac{5^3e^{-5}}{3!} + \cdots + (2+5*98)\frac{5^{100}e^{-5}}{100!} = \$17.155.$$

If one computes the exact value, using the binomial distribution, one finds the sum to be finds the answer to be

$$2\binom{100}{2}(.05)^2(.95)^{98}+7\binom{100}{3}(.05)^3(.95)^{97}+\ldots+(2+5*98)\binom{100}{100}(.05)^{100}(.95)^0 = \$17.141\,.$$

He is better off putting a dime in the meter each time for a total cost of \$10.

26.

number	observed	expected
0	229	227
1	211	211
2	93	99
3	35	31
4	7	9
5	1	1

27. $m = 100 \times (.001) = .1$. Thus $P(\text{at least one accident}) = 1 - e^{-.1} = .0952$.

28. .9084

29. Here $m = 500 \times (1/500) = 1$, and so $P(\text{at least one fake}) = 1 - e^{-1} = .632$. If the king tests two coins from each of 250 boxes, then $m = 250 \times \frac{2}{500} = 1$, and so the answer is again .632.

30. $P(\text{win } \geq 3 \text{ times}) \approx .5071$, expected winnings ≈ -2.703

31.The expected number of deaths per corps per year is

$$1 \cdot \frac{91}{280} + 2 \cdot \frac{32}{280} + 3 \cdot \frac{11}{280} + 4 \cdot \frac{2}{280} = .70.$$

The expected number of corps with x deaths would then be $280 \cdot \dfrac{(.70)^x e^{-(.70)}}{x!}$. From this we obtain the following comparison:

Number of deaths	Corps with x deaths	Expected number of corps
0	144	139.0
1	91	97.3

The fit is quite good.

32.

$$\begin{aligned} P(X+Y=j) &= \sum_{k=0}^{j} P(X=k)P(Y=j-k) = \sum_{k=0}^{j} \frac{m^k e^{-m}}{k!} \cdot \frac{\bar{m}^{j-k} e^{-\bar{m}}}{(j-k)!} \\ &= e^{-(m+\bar{m})} \Big(\sum_{k=0}^{j} \frac{j! m^k \bar{m}^{j-k}}{k!(j-k)!} \Big) \frac{1}{j!} \\ &= e^{-(m+\bar{m})} \frac{1}{j!} \Big(\sum_{k=0}^{j} \binom{j}{k} m^k \bar{m}^{j-k} \Big) = \frac{(m+\bar{m})^j}{j!} e^{-(m+\bar{m})}. \end{aligned}$$

Thus, $X+Y$ has a Poisson density with mean $m+\bar{m}$.

33. Poisson with mean 3.

34. .168

35.

(a) In order to have d defective items in s items, you must choose d items out of D defective ones and the rest from $S-D$ good ones. The total number of sample points is the number of ways to choose s out of S.

(b) Since

$$\sum_{j=0}^{\min(D,s)} P(X=j) = 1,$$

we get

$$\sum_{j=0}^{\min(D,s)} \binom{D}{j} \binom{s-D}{s-j} = \binom{S}{s}.$$

36. $D = 20$. This illustrates the general fact that the maximum probability is achieved when
$$\frac{d}{D} = \frac{s}{S} .$$

37. The maximum likelihood principle gives an estimate of 1250 moose.

38. With replacement: $P(X = 1) \approx .396$

Without replacement: $P(X = 1) \approx .440$

40. (a) $$\frac{\frac{4!}{2}\binom{13}{4}\binom{13}{4}\binom{13}{3}\binom{13}{2}}{\binom{52}{13}} = .2155.$$

(b) $$\frac{\frac{4!}{2}\binom{13}{5}\binom{13}{3}\binom{13}{3}\binom{13}{2}}{\binom{52}{13}} = .1552.$$

42. $$p_{X_1} = \begin{pmatrix} 0 & 4 \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}, \quad p_{X_2} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \quad p_{X_3} = \begin{pmatrix} 2 & 6 \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}, \quad p_{X_4} = \begin{pmatrix} 1 & 5 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

If your friend chooses die 1, choose die 4; if she chooses die 2, choose die 1; if she chooses die 3, choose die 2; if she chooses die 4, choose die 3. Then
$$P(X_1 < X_4) = P(X_2 < X_1) = P(X_3 < X_2) = P(X_4 < X_3) = \frac{2}{3}.$$

Thus you are assured of winning with probability 2/3.

43. If the traits were independent, then the probability that we would obtain a data set that differs from the expected data set by as much as the actual data set differs is approximately .00151. Thus, we should reject the hypothesis that the two traits are independent.

44. The value of χ^2 corresponding to the data is $v = 9931.6$, which is much greater than v_0, so the hypothesis that the chosen numbers are uniformly distributed should be rejected.

5.2 Important Densities

1. (a) $f(x) = 1$ on $[2, 3]$; $F(x) = x - 2$ on $[2, 3]$.

(b) $f(x) = \frac{1}{3}x^{-2/3}$ on $[0,1]$; $F(x) = x^{1/3}$ on $[0,1]$.

2. (a) $F(x) = 2 - \frac{1}{x}\,, \quad f(x) = \frac{1}{x^2}$ on $[\frac{1}{2}, 1]\,.$

(b) $F(x) = e^x - 1\,, \quad f(x) = e^x$ on $[0, \log 2]\,.$

5. (a) $F(x) = 2x\,, \quad f(x) = 2$ on $[0,1]\,.$

(b) $F(x) = 2\sqrt{x}\,, \quad f(x) = \frac{1}{\sqrt{x}}$ on $[0, \frac{1}{2}]\,.$

7. Using Corollary 5.2, we see that the expression $\sqrt{rnd}$ will simulate the given random variable.

9. (a) $F(y) = \begin{cases} \frac{y^2}{2}, & 0 \le y \le 1; \\ 1 - \frac{(2-y)^2}{2}, & 1 \le y \le 2, \end{cases} \quad f(y) = \begin{cases} y, & 0 \le y \le 1; \\ 2 - y & 1 \le y \le 2. \end{cases}$

(b) $F(y) = 2y - y^2, \qquad f(y) = 2 - 2y, \quad 0 \le y \le 1.$

10. (a) $F(x) = x^2$ and $f(x) = 2x$ on $[0,1]$.

(b) $F(x) = 2x - x^2$ and $f(x) = 2 - 2x$ on $[0,1]$.

12. (a) 1/2

(b) 1

(c) .2

13. (a) $F(r) = \sqrt{r}\,, \quad f(r) = \frac{1}{2\sqrt{r}}\,,$ on $[0,1]\,.$

(b) $F(s) = 1 - \sqrt{1-4s}\,, \quad f(s) = \frac{2}{\sqrt{1-4x}}\,,$ on $[0, 1/4]\,.$

(c) $F(t) = \frac{t}{1+t}\,, \quad f(t) = \frac{1}{(1+t)^2}\,,$ on $[0, \infty)\,.$

14. (a) 3/4

(b) $\pi/16$

15. $F(d) = 1 - (1-2d)^2, \qquad f(d) = 4(1-2d) \qquad$ on $\quad [0, \frac{1}{2}]$.

16. (a) c = 6

(b) $F(x) = 3x^2 - 2x^3$

(c) .156

17. (a) $f(x) = \begin{cases} \frac{\pi}{2}\sin(\pi x), & 0 \le x \le 1; \\ 0, & \text{otherwise.} \end{cases}$

(b) $\sin^2(\frac{\pi}{8}) = .146$.

18. $F_W(w) =$

$a > 0: \quad F_X(\frac{w-b}{a}); \quad a = 0: \quad \begin{cases} 1, & w \ge b; \\ 0, & \text{otherwise;} \end{cases} \quad a < 0: \quad 1 - F_X(\frac{w-b}{a}).$

19. a $\neq 0: f_W(w) = \frac{1}{|a|} f_X(\frac{w-b}{a})$, a = 0: $f_W(w) = 0$ if $w \neq 0$.

20. $a = \dfrac{1}{d-c}$ and $b = \dfrac{c}{c-d}$

21. $P(Y \le y) = P(F(X) \le y) = P(X \le F^{-1}(y)) = F(F^{-1}(y)) = y \quad \text{on} \quad [0,1]$.

22. (a) $\dfrac{a+b}{2}$

(b) μ

(c) $\dfrac{1}{\lambda}\log 2$

23. The mean of the uniform density is $(a+b)/2$. The mean of the normal density is μ. The mean of the exponential density is $1/\lambda$.

24. The mode of the uniform density is any number in $[0,1]$. The mode of the normal is μ. The mode of the exponential is 0.

25. (a) .9773, (b) .159, (c) .0228, (d) .6827.

26. 13.4% are likely to be rejected. For 1% rejection rate, let $\sigma = .0012$.

27. A: 15.9%, B: 34.13%, C: 34.13%, D: 13.59%, F: 2.28%.

28. 2.4%

29. e^{-2}, e^{-2}.

30. The car will last for 4 years with probability $1/e \approx .368$.

31. $\frac{1}{2}$.

34.
$$P(X < Y) = \int_{x=0}^{\infty}\int_{y=x}^{\infty} f(x)g(y)\,dx\,dy = \int_{x=0}^{\infty} f(x)(1 - G(x))\,dx.$$

Thus

$$P(X < Y) = \int_0^\infty \lambda e^{-\lambda x} \cdot e^{-\mu x} dx = \frac{\lambda}{\lambda + \mu}.$$

Therefore, the probability that a 100 watt bulb will outlast a 60 watt bulb is $\frac{(1/200)}{1/200+1/100} = 1/3$.

35. $P(\text{size increases}) = P(X_j < Y_j) = \lambda/(\lambda + \mu)$.

$P(\text{size decreases}) = 1 - P(\text{size increases}) = \mu/(\lambda + \mu)$.

36. Exponential with parameter λ/r.

37. $F_Y(y) = \dfrac{1}{\sqrt{2\pi y}} e^{-\frac{\log^2(y)}{2}}$, for $y > 0$.

38.

$$\begin{aligned} P(Y_1 \le y_1, Y_2 \le y_2) &= P(X_1 \le \Phi_1^{-1}(y_1), X_2 \le \Phi_2^{-1}(y_2)) \\ &= P(X_1 \le \Phi_1^{-1}(y_1)) P(X_2 \le \Phi_2^{-1}(y_2)) \\ &= P(Y_1 \le y_1) P(Y_2 \le y_2), \end{aligned}$$

so Y_1 and Y_2 are independent.

Chapter 6

Expected Value and Variance

6.1 Expected Value of Discrete Random Variables

1. -1/9

2. -1/2

3. 5′ 10.1”

4. -1/19

5. -1/19

6. Let U and V be independent identically distributed random variables with the density:
$$p_U = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}.$$
Then
$$XY = (U+V)(U-V) = U^2 - V^2,$$
so
$$E(XY) = E(U^2) - E(V^2) = 0.$$
Since
$$E(Y) = E(U) - E(V) = 0,$$
we have
$$E(XY) = E(X)E(Y) = 0.$$
But X and Y are not independent. For example, if we know that $X = 12$, then we know that $Y = 0$.

7. Since X and Y each take on only two values, we may choose a, b, c, d so that

$$U = \frac{X+a}{b},\ V = \frac{Y+c}{d}$$

take only values 0 and 1. If $E(XY) = E(X)E(Y)$ then $E(UV) = E(U)E(V)$. If U and V are independent, so are X and Y. Thus it is sufficient to prove independence for U and V taking on values 0 and 1 with $E(UV) = E(U)E(V)$. Now

$$E(UV) = P(U=1, V=1) = E(U)E(V) = P(U=1)P(V=1),$$

and

$$\begin{aligned} P(U=1, V=0) &= P(U=1) - P(U=1, V=1) \\ &= P(U=1)(1 - P(V=1)) = P(U=1)P(V=0). \end{aligned}$$

Similarly,

$$\begin{aligned} P(U=0, V=1) &= P(U=0)P(V=1) \\ P(U=0, V=0) &= P(U=0)P(V=0). \end{aligned}$$

Thus U and V are independent, and hence X and Y are also.

8. The expected number of boys and the expected number of girls are both $\frac{7}{8}$.

9. The second bet is a fair bet so has expected winning 0. Thus your expected winning for the two bets is the same as the original bet which was $-7/498 = -.0141414...$. On the other hand, you bet 1 dollar with probability 1/3 and 2 dollars with probability 2/3. Thus the expected amount you bet is $1\frac{2}{3}$ dollars and your expected winning per dollar bet is -.0141414/1.666667 = -.0085 which makes this option a better bet in terms of the amount won per dollar bet. However, the amount of time to make the second bet is negligible, so in terms of the expected winning per time to make one play the answer would still be -.0141414.

11. The roller has expected winning -.0141; the pass bettor has expected winning -.0136.

12. 0

13. 45

14. $E(X_j) = \dfrac{1}{N}$. For $j \neq k,\ E(X_j E_k) = \dfrac{1}{N(N-1)}$. Thus X_j and X_k are not independent.

15. $E(X) = \frac{1}{5}$, so this is a favorable game.

16. (a)

$$E(X) = \frac{(1+2+3+4+5+6)}{6} = 3\frac{1}{2}.$$

(b) The large sums are much less likely to occur than small sums. For example

$$P(\text{total} = 21) = (1/6)^6 = 2.14 \times 10^{-5}$$

and

$$P(\text{total} = 0) = (5/6)^6 = .335.$$

17. $p_k = p(\overbrace{S \cdots S}^{k-1 \text{ times}}\ F) = p^{k-1}(1-p) = p^{k-1}q,\ k = 1, 2, 3, \ldots$.

$$\sum_{k=1}^{\infty} p_k = q\sum_{k=0}^{\infty} p^k = q\frac{1}{1-p} = 1\ .$$

$$E(X) = q\sum_{k=1}^{\infty} kp^{k-1} = \frac{q}{(1-p)^2} = \frac{1}{q}\ .\ \text{(See Example 6.4.)}$$

18. 7/2

19.

$$\begin{aligned} E(X) &= \frac{\binom{4}{4}}{\binom{4}{4}}(3-3) + \frac{\binom{3}{2}}{\binom{4}{3}}(3-2) + \frac{\binom{3}{3}}{\binom{4}{3}}(0-3) + \frac{\binom{3}{1}}{\binom{4}{2}}(3-1) \\ &\quad + \frac{\binom{3}{2}}{\binom{4}{2}}(0-2) + \frac{\binom{3}{0}}{\binom{4}{1}}(3-0) + \frac{\binom{3}{1}}{\binom{4}{1}}(0-1) = 0\,. \end{aligned}$$

20. (a)

$$E(X) = \frac{1}{2}\cdot 2 + \left(\frac{1}{2}\right)^2 \cdot 2^2 + \left(\frac{1}{2}\right)^3 \cdot 2^3 \cdots = 1 + 1 + 1 + \cdots = \infty\,,$$

and so $E(X)$ does not exist. This means that if we could play the game, it would be favorable now matter how much we pay to play it. However, we cannot realize this game, since it requires arbitrarly large amounts of money.

(b)

$$\begin{aligned} E(X) &= \frac{1}{2}\cdot 2 + \left(\frac{1}{2}\right)^2 \cdot 2^2 + \cdots + \left(\frac{1}{2}\right)^{10} \cdot 2^{10} + 2^{10}\left(\frac{1}{2^{11}} + \frac{1}{2^{12}} + \cdots\right) \\ &= 10 + \frac{1}{2} + \frac{1}{2^2} + \cdots = 11\ . \end{aligned}$$

(d) If the utility of n dollars is $\sqrt{n}$, then the expected utility of the payment is given by

$$\sum_{i=1}^{\infty} \frac{1}{2^i}\sqrt{2^i} = \frac{1}{\sqrt{2}-1} \ .$$

If the utility of n dollars is $\log n$, then the expected utility of the payment is given by

$$\sum_{i=1}^{\infty} \frac{1}{2^i}\log(2^i) = 2\log 2 \ .$$

22. The expected number of days in a year with more than 60 percent boys for the large hospital is

$$365 \cdot \sum_{k=28}^{k=45} b(45, .5, k) = 24.67 \ .$$

For the small hospital it is

$$365 \cdot \sum_{k=10}^{k=15} b(15, .5, k) = 55.1 \ .$$

23. 10

25.

(b) Let S be the number of stars and C the number of circles left in the deck. Guess star if $S > C$ and guess circle if $S < C$. If $S = C$ toss a coin.

(d) Consider the recursion relation:

$$h(S, C) = \frac{\max(S, C)}{S + C} + \frac{S}{S + C}h(S - 1, C) + \frac{C}{S + C}h(S, C - 1)$$

and $h(0, 0) = h(-1, 0) = h(0, -1) = 0$. In this equation the first term represents your expected winning on the current guess and the next two terms represent your expected total winning on the remaining guesses. The value of $h(10, 10)$ is 12.34.

26. (a) Let L be the horizontal line passing through $S - C$. If the random walk is below L, then there are more stars than circles in the remaining deck, and so, using the optimal strategy, you guess star. If you are right, the graph goes up. If the walk is above L, then there are more circles than stars, and you guess circle. If you are right, the graph goes down. Since $S \geq C$, the graph ends at $(S + C, S - C)$. Let a be the number of times the graph goes up under L, b the number of times it goes down under L, c the number of times it goes down above L, and d the number of times it

goes up above L. Then $a + b + c + d = S + C, a - b = S - C, c - d = 0$. Thus $2a - S + C + 2c = S + C$, and this implies $a + c = S$, i.e., we have S correct guesses.

(b) We arrive at (x, x) if $S - x$ stars turn up and $C - x$ circles turn up in $S + C - 2x$ guesses. The probability of this happening is

$$\frac{\binom{S}{S-x}\binom{C}{C-x}}{\binom{S+C}{S+C-2x}} = \frac{\binom{S}{x}\binom{C}{x}}{\binom{S+C}{2x}} .$$

(c) The number of correct guesses equals the number of correct guesses when the graph is under or above L plus the number of correct guesses when the graph hits L. Thus the expected number of correct guesses is:

$$S + \sum_{x=1}^{C} \frac{\binom{S}{x}\binom{C}{x}}{\binom{S+C}{2x}} \cdot \frac{1}{2} .$$

27. (a) 4

(b) $4 + \sum_{x=1}^{4} \frac{\binom{4}{x}\binom{4}{x}}{\binom{8}{x}} = 5.79$.

28. (a) Assume that $n = 2^k - 1$. Choose the middle number of the numbers from 1 to $2^k - 1$, and then continue to choose the middle number until you guess the number correctly. If you have not yet succeeded after $k - 1$ guesses you will be down to a single number and will be sure to get it on the kth question. This strategy obviously works just as well if $n < 2^{k-1}$.

(b) Whenever you make a guess and are wrong the search is narrowed to a new and smaller interval $[a, b]$. The probability that you guess correctly on a question when the interval is $[a, b]$ is

$$P(\text{correct}) = \frac{(b-a)}{n} \cdot \frac{1}{(b-a)} = \frac{1}{n} .$$

Thus the probabililty that you guess the number on the kth question is $a(k)/n$ where $a(k)$ is the number of possible subintervals for the kth question. The probability of guessing the number correctly for a strategy with at most k guesses is

$$\frac{\sum_k a(k)}{n} .$$

Thus any strategy that makes this sum as large as possible is optimal. We can at most double the number of intervals on each question.Thus any strategy that achieves this is optimal.The resulting probability of guessing the number in k questions is

$$\frac{\sum_{j=0}^{k-1} 2^k}{n} = \frac{2^k - 1}{n} .$$

If $n \geq 2^k - 1$ we can achieve this optimal strategy by continuing to bisect the numbers between 1 and $2^k - 1$.

29. If you have no ten-cards and the dealer has an ace, then in the remaining 49 cards there are 16 ten cards. Thus the expected payoff of your insurance bet is:

$$2 \cdot \frac{16}{49} - 1 \cdot \frac{33}{49} = -\frac{1}{49} \; .$$

If you are playing two hands and do not have any ten-cards then there are 16 ten-cards in the remaining 47 cards and your expected payoff on an insurance bet is:

$$2 \cdot \frac{16}{47} - 1 \cdot \frac{31}{47} = \frac{1}{47} \; .$$

Thus in the first case the insurance bet is unfavorable and in the second it is favorable.

30.

(a) $P(X_k = j) = P(j-1$ boxes have old pictures and the jth box has a new picture)

$$= \left(\frac{k-1}{n}\right)^j \left(\frac{n-k+1}{n}\right) \; ,$$

and so X_k has a geometric distribution with $p = (n-k+1)/n$.

(c) The expected time for getting the first half of the players is

$$\begin{aligned} E(X_1) + \cdots + E(X_n) &= \frac{2n}{2n-1+1} + \frac{2n}{2n-2+1} + \cdots + \frac{2n}{2n-n+1} \\ &= 2n\left(\frac{1}{2n} + \frac{1}{2n-1} + \cdots \frac{1}{n+1}\right) \; . \end{aligned}$$

The expected time for getting the second half of the players is:

$$\begin{aligned} E(X_{n+1}) + \cdots + E(X_{2n}) &= \frac{2n}{2n-(n+1)-1} + \cdots + \frac{2n}{2n-2n+1} \\ &= 2n\left(\frac{1}{n} + \frac{1}{n-1} + \cdots + \frac{1}{1}\right) \; . \end{aligned}$$

(d)

$$\begin{aligned} 1 + \frac{1}{2} + \cdots + \frac{1}{n} &\sim \log n + .5772 + \frac{1}{2n} \; . \\ 1 + \frac{1}{2} + \cdots + \frac{1}{n} + \frac{1}{n+1} + \cdots + \frac{1}{2n} &\sim \log 2n + .5772 + \frac{1}{4n} \; . \\ 2n\left(\frac{1}{2n} + \cdots + \frac{1}{n+1}\right) &\sim 2n\left(\log 2n + \frac{1}{4n} - \log n - \frac{1}{2n}\right) \; . \\ &= 2n\left(\log 2 - \frac{1}{4n}\right) \\ &= 2n\log 2 - \frac{1}{2} \end{aligned}$$

$$2n\Big(1 + \frac{1}{2} + \cdots + \frac{1}{n}\Big) \quad \sim \quad 2n\Big(\log n + .5772 + \frac{1}{2n}\Big) \ .$$

31. (a) $1 - (1-p)^k$.

(b) $\dfrac{N}{k} \cdot \Big((k+1)(1-(1-p)^k) + (1-p)^k\Big)$.

(c) If p is small, then $(1-p)^k \sim 1 - kp$, so the expected number in (b) is $\sim N[kp + \frac{1}{k}]$, which will be minimized when $k = 1/\sqrt{p}$.

32. Your estimate should be near $e = 2.718...$

33. We begin by noting that

$$P(X \geq j+1) = P((t_1 + t_2 + \cdots + t_j) \leq n) \ .$$

Now consider the j numbers $a_1, a_2, \cdots, a_j$ defined by

$$\begin{array}{lll} a_1 & = t_1 & \\ a_2 & = & t_1 + t_2 \\ a_3 & = & t_1 + t_2 + t_3 \\ \vdots & & \vdots \quad \vdots \\ a_j & = & t_1 + t_2 + \cdots + t_j \ . \end{array}$$

The sequence $a_1, a_2, \cdots, a_j$ is a monotone increasing sequence with distinct values and with successive differences between 1 and n. There is a one-to-one correspondence between the set of all such sequences and the set of possible sequences $t_1, t_2, \cdots, t_j$. Each such possible sequence occurs with probability $1/n^j$. In fact, there are n possible values for t_1 and hence for a_1. For each of these there are n possible values for a_2 corresponding to the n possible values of t_2. Continuing in this way we see that there are n^j possible values for the sequence $a_1, a_2, \cdots, a_j$. On the other hand, in order to have $t_1 + t_2 + \cdots + t_j \leq n$ the values of $a_1, a_2, \cdots, a_j$ must be distinct numbers lying between 1 to n and arranged in order. The number of ways that we can do this is $\binom{n}{j}$. Thus we have

$$P(t_1 + t_2 + \cdots + t_j \leq n) = P(X \geq j+1) = \binom{n}{j} \frac{1}{n^j} \ .$$

$$\begin{array}{rcl} E(X) & = & P(X=1) + P(X=2) + P(X=3) \cdots \\ & & \qquad\qquad\quad + P(X=2) + P(X=3) \cdots \\ & & \qquad\qquad\qquad\qquad\qquad\quad + P(X=3) \cdots \ . \end{array}$$

If we sum this by rows we see that

$$E(X) = \sum_{j=0}^{n-1} P(X \geq j+1) \ .$$

Thus,

$$E(X) = \sum_{j=1}^{n} \binom{n}{j} \Big(\frac{1}{n}\Big)^j = \Big(1 + \frac{1}{n}\Big)^n \ .$$

The limit of this last expression as $n \to \infty$ is $e = 2.718...$.
There is an interesting connection between this problem and the exponential density discussed in Section 2.2 (Example 2.17). Assume that the experiment starts at time 1 and the time between occurrences is equally likely to be any value between 1 and n. You start observing at time n. Let T be the length of time that you wait. This is the amount by which $t_1 + t_2 + \cdots + t_j$ is greater than n. Now imagine a sequence of plays of a game in which you pay $n/2$ dollars for each play and for the j'th play you receive the reward t_j. You play until the first time your total reward is greater than n. Then X is the number of times you play and your total reward is $n+T$. This is a perfectly fair game and your expected net winning should be 0. But the expected total reward is $n + E(T)$. Your expected payment for play is $\frac{n}{2}E(X)$. Thus by fairness, we have

$$n + E(T) = (n/2)E(X) \ .$$

Therefore,

$$E(T) = \frac{n}{2}E(X) - n \ .$$

We have seen that for large n, $E(X) \sim e$. Thus for large n,

$$E(\text{waiting time}) = E(T) \sim n(\frac{e}{2} - 1) = .718n \ .$$

Since the average time between occurrences is $n/2$ we have another example of the paradox where we have to wait on the average longer than $1/2$ the average time time between occurrences.

34. (a) We prove first that for Bernoulli trials the probability that the kth failure precedes the rth success is

$$f(k, p, r) = \binom{r+k-1}{k} p^{r-1} q^k \cdot p \ .$$

To prove this, we note that for the kth failure to precede the rth success we must have $r-1$ successes and k failures in the first $r+k-1$ trials and then have a success. The probability that this happens is $f(k, p, r)$. Now consider a Bernoulli trials process where success is getting a match from the right pocket. In order to have r matches in the left pocket when the right pocket has none we must have $N-r$ failures before the $(N+1)$st success. Thus the probability that there are r matches in the left pocket when the right pocket has none is

$$f(N - r, \frac{1}{2}, N + 1) \ .$$

The same argument applies for the probability that there are r matches in the right pocket when the left pocket has none. Thus

$$p_r = 2f(N-r, \frac{1}{2}, N+1) = \binom{2N-r}{N}\left(\frac{1}{2}\right)^{2N-r} .$$

(c)

$$\begin{aligned}
(N-\frac{r}{2})p_{r+1} &= (N-\frac{r}{2})\binom{2N-r-1}{N}\left(\frac{1}{2}\right)^{2N-r-1} \\
&= (2N-r)\binom{2N-r-1}{N}\left(\frac{1}{2}\right)^{2N-r} \\
&= (N-r)\binom{2N-r}{N}\left(\frac{1}{2}\right)^{2N-r} \\
&= (N-r)p_r .
\end{aligned}$$

(d) $\displaystyle\sum_{r=0}^{N} p_r = 1$.

(e)

$$\sum_{r=0}^{N}(N-r)p_r = \sum_{r=0}^{N}\frac{1}{2}(2N+1)p_{r+1} - \sum_{r=0}^{N}\frac{1}{2}(r+1)p_{r+1} .$$

Thus

$$N - E = \frac{1}{2}(2N+1)(1-p_0) - \frac{1}{2}E ,$$

and

$$E = p_0(2N+1) - 1 .$$

But

$$p_0 = \binom{2N}{N}\left(\frac{1}{2}\right)^{2N} \sim \frac{1}{\sqrt{\pi N}} ,$$

so

$$E \sim 2\sqrt{\frac{N}{\pi}} .$$

Using this asymptotic expression leads to an estimate of 133 for the number of matches needed in each pocket to make $E = 13$. It is easy to make an exact calculation with the computer, and this gives 153 matches.

35. One can make a conditionally convergent series like the alternating harmonic series sum to anything one pleases by properly rearranging the series. For example, for the order given we have

$$E = \sum_{n=0}^{\infty}(-1)^{n+1}\frac{2^n}{n}\cdot\frac{1}{2^n}$$

$$= \sum_{n=0}^{\infty} (-1)^{n+1} \frac{1}{n} = \log 2 \ .$$

But we can rearrange the terms to add up to a negative value by choosing negative terms until they add up to more than the first positive term, then choosing this positive term, then more negative terms until they add up to more than the second positive term, then choosing this positive term, etc.

36. $c \dfrac{k}{c+d}$

37. (a) Under option (a), if red turns up, you win 1 franc, if black turns up, you lose 1 franc, and if 0 turns up, you lose 1/2 franc. Thus, the expected winnings are

$$1\left(\frac{18}{37}\right) + (-1)\left(\frac{18}{37}\right) + \left(\frac{-1}{2}\right)\left(\frac{1}{37}\right) \approx -.0135 \ .$$

(b) Under option (b), if red turns up, you win 1 franc, if black turns up, you lose 1 franc, and if 0 comes up, followed by black or 0, you lose 1 franc. Thus, the expected winnings are

$$1\left(\frac{18}{37}\right) + (-1)\left(\frac{18}{37}\right) + (-1)\left(\frac{1}{37}\right)\left(\frac{19}{37}\right) \approx -.0139 \ .$$

(c)

38. (Solution by Victor Hernández) Let p_{ij} be the probability that book i is above book j. Then the average depth of book j is

$$d_j = \sum_{i \neq j} p_{ij} \ ,$$

where the top book is considered to be at depth 0. Now if book i is above book j, then the relative order of books i and j is changed after a call if and only if book j is consulted. Hence,

$$\begin{aligned} p_{ij} &= p_{ij}(1 - p_j) + p_{ji}p_i \\ &= p_{ij}(1 - p_j) + (1 - p_{ij})p_i \\ &= p_i + p_{ij}(1 - p_i - p_j) \ . \end{aligned}$$

Thus, we have

$$p_{ij} = \frac{p_i}{p_i + p_j} \ ,$$

and

$$d_j = \sum_{k \neq j} \frac{p_k}{p_k + p_j} \ .$$

If $p_i \geq p_j$, then

$$\frac{p_k}{p_k + p_i} \leq \frac{p_k}{p_k + p_j}$$

for $k \neq i$, j, and

$$\frac{p_j}{p_i + p_j} \leq \frac{p_i}{p_i + p_j} \ .$$

Since each term in the sum for d_i is less than or equal to the corresponding term in the sum for d_j, we have $d_i \leq d_j$.

39. (Solution by Peter Montgomery) The probability that book 1 is in the right place is the probability that the last phone call referenced book 1, namely p_1. The probability that book 2 is in the right place, given that book 1 is in the right place, is

$$p_2 + p_2 p_1 + p_2 p_1^2 + \ldots = \frac{p_2}{(1 - p_1)} \ .$$

Continuing, we find that

$$P = p_1 \frac{p_2}{(1 - p_1)} \frac{p_3}{(1 - p_1 - p_2)} \cdots \frac{p_n}{(1 - p_1 - p_2 - \ldots - p_{n-1})} \ .$$

Now let q be a real number between 0 and 1, let

$$p_1 = 1 - q \ ,$$

$$p_2 = q - q^2 \ ,$$

and so on, and finally let

$$p_n = q^{n-1} \ .$$

Then

$$P = (1 - q)^{n-1} \ ,$$

so P can be made arbitrarily close to 1.

40. If $a_1, \ a_2, \ \ldots$ is the sequence, then the event $\{a_1 < a_2 < \ldots < a_k\}$ occurs with probability $1/k!$, since there are $k!$ different orderings of k real numbers, and all of them are equally likely to occur in this experiment. Therefore,

$$P(X > k) = \frac{1}{k!} \ .$$

Now let

$$p_k = P(X = k) \ .$$

Then

$$\begin{aligned} E(X) &= p_1 + 2p_2 + 3p_3 + \ldots \\ &= (p_1 + p_2 + \ldots) + (p_2 + p_3 + \ldots) + \ldots \\ &= P(X > 0) + P(X > 1) + \ldots \\ &= 1 + \frac{1}{1!} + \frac{1}{2!} + \ldots \\ &= e \ . \end{aligned}$$

6.2 Variance of Discrete Random Variables

1. $E(X) = 0,\ V(X) = \frac{2}{3},\ \ \sigma = D(X) = \sqrt{\frac{2}{3}}$.

2. $E(X) = \frac{4}{3},\ \ V(X) = \frac{17}{9},\ \ \sigma = D(X) = \frac{\sqrt{17}}{3}$.

3. $E(X) = \frac{-1}{19},\ \ E(Y) = \frac{-1}{19},\ \ V(X) = 33.21,\ \ V(Y) = .99$.

4. (a) 10015, (b) 310, (c) -100, (d) 15, (e) $\sqrt{15}$.

5. (a) $E(F) = 62,\ \ V(F) = 1.2$.

(b) $E(T) = 0,\ \ V(T) = 1.2$.

(c) $E(C) = \frac{50}{3}\ ,\ \ V(C) = \frac{10}{27}$.

7. $V(X) = \frac{3}{4}\ ,\ \ D(X) = \frac{\sqrt{3}}{2}$.

8. $E(S_{2400}) = 960,\quad V(S_{2400}) = 576,\quad \sigma = D(S_{2400}) = 24.$

9. $V(X) = \frac{20}{9}\ ,\ \ D(X) = \frac{2\sqrt{5}}{3}$.

10. (a) $V(X + c) = V(X)$, so $D(X + c) = D(X)$.

(b) $V(cX) = c^2 V(X),$ so $D(cX) = |c|X.$

11. $E(X) = (1 + 2 + \cdots + n)/n = (n + 1)/2.$

$$\begin{aligned} V(X) &= (1^2 + 2^2 + \cdots + n^2)/n - (E(X))^2 \\ &= (n + 1)(2n + 1)/6 - (n + 1)^2/4 = (n + 1)(n - 1)/12. \end{aligned}$$

12. $E\Big(X - \mu/\sigma\Big) = (1/\sigma)(E(X) - \mu) = 0,$

$V\Big(X - \mu/\sigma\Big)^2 = (1/\sigma^2)E(X - \mu)^2 = \sigma^2/\sigma^2 = 1.$

13. Let $X_1, \ldots, X_n$ be identically distributed random variables such that

$$P(X_i = 1) = P(X_i = -1) = \frac{1}{2}.$$

Then $E(X_i) = 0$, and $V(X_i) = 1$. Thus $W_n = \sum_{j=1}^n X_i$. Therefore $E(W_n) = \sum_{i=1}^n E(X_i) = 0$, and $V(W_n) = \sum_{i=1}^n V(X_i) = n$.

14. Let X be the number of boys and Y be the number of girls. Then

$$E(X) = E(Y) = \frac{7}{8}\ ,$$

and

$$V(X) = \frac{7}{64}\ , \qquad V(Y) = \frac{71}{64}\ .$$

15. (a) $P_{X_i} = \begin{pmatrix} 0 & 1 \\ \frac{n-1}{n} & \frac{1}{n} \end{pmatrix}$. Therefore, $E(X_i)^2 = 1/n$ for $i \neq j$.

(b) $P_{X_iX_j} = \begin{pmatrix} 0 & 1 \\ 1 - \frac{1}{n(n-1)} & \frac{1}{n(n-1)} \end{pmatrix}$ for $i \neq j$.

Therefore, $E(X_iX_j) = \dfrac{1}{n(n-1)}$.

(c)

$$\begin{aligned} E(S_n)^2 &= \sum_i E(X_i)^2 + \sum_i \sum_{j \neq i} E(X_iX_j) \\ &= n \cdot \frac{1}{n} + n(n-1) \cdot \frac{1}{n(n-1)} = 2\ . \end{aligned}$$

(d)

$$\begin{aligned} V(S_n) &= E(S_n)^2 - E(S_n)^2 \\ &= 2 - (n \cdot (1/n))^2 = 1\ . \end{aligned}$$

16. (a) For $p = .5$:

		k		
		1	2	3
	10	.656	.979	.998
N	30	.638	.957	.999
	50	.678	.967	.997

For $p = .2$:

		k		
		1	2	3
	10	.772	.967	.994
N	30	.749	.964	.997
	50	.629	.951	.997

(b) Use Exercise 12 and the fact that $E(S_n) = np$ and $V(S_n) = npq$. The two examples in (a) suggests that the probability that the outcome is within k standard deviations is approximately the same for different values of p. We shall see in Chapter 9 that the Central Limit Theorem explains why this is true.

18. (a) $E(\bar{x}) = \frac{1}{n}\sum_{i=1}^{n} E(x_i) = \frac{1}{n} \cdot n\mu = \mu$.

(b) We have

$$E\Big((\bar{x}-\mu)^2\Big) = V(\bar{x}) ,$$

which was shown to equal σ^2/n in Theorem 6.9.

(c) We have from the hint:

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n}(x_i-\mu)^2 - n(\bar{x}-\mu)^2 .$$

Thus,

$$\begin{aligned} E(s^2) &= \frac{1}{n}E\left(\sum_{i=1}^{n}(x_i-\bar{x})^2\right) \\ &= \frac{1}{n}\left(E\left(\sum_{i=1}^{n}(x_i-\mu)^2\right) - nE(\bar{x}-\mu)^2\right) \\ &= \frac{1}{n}(n\sigma^2 - n\frac{\sigma^2}{n}) = \frac{n-1}{n}\sigma^2 , \end{aligned}$$

where we have used the definition of the variance and part (b) to obtain the penultimate expression.

(d) Since the expectation operator is linear, and the 'new' s^2 is $n/(n-1)$ times the 'old' s^2, the new s^2 has expectation

$$\frac{n}{n-1}\frac{n-1}{n}\sigma^2 = \sigma^2 .$$

19. Let X_1, X_2 be independent random variables with

$$p_{X_1} = p_{X_2} = \begin{pmatrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} .$$

Then

$$p_{X_1+X_2} = \begin{pmatrix} -2 & 0 & 2 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix} .$$

Then

$$\bar{\sigma}_{X_1} = \bar{\sigma}_{X_2} = 1,\ \bar{\sigma}_{X_1+X_2} = 1 .$$

Therefore

$$V(X_1 + X_2) = 1 \neq V(X_1) + V(X_2) = 2 ,$$

and

$$\bar{\sigma}_{X_1+X_2} = 1 \neq \bar{\sigma}_{X_1} + \bar{\sigma}_{X_2} = 2 .$$

20. (a) $E(\bar{\mu}) = \mu$

(b) $w = \dfrac{V(X_2)}{V(X_1) + V(X_2)}$

21.

$$\begin{aligned} f'(x) &= -\sum_\omega 2(X(\omega) - x)p(\omega) \\ &= -2\sum_\omega X(\omega)p(\omega) + 2x\sum_\omega p(\omega) \\ &= -2\mu + 2x\ . \end{aligned}$$

Thus $x = \mu$ is a critical point. Since $f''(x) \equiv 2$, we see that $x = \mu$ is the minimum point.

22. $X, Y, X+Y$, and $X-Y$ have the same distribution, so they have the same mean and variance. Thus $E(X) = E(Y) = E(X) + E(Y)$. This implies that

$$E(X) = E(Y) = 0\ .$$

It also implies that

$$E(X+Y)^2 = E(X-Y)^2 = E(X^2) = E(Y^2)\ .$$

Thus $E(XY) = 0$ and $E(Y^2) = E(X^2) = 0$. Therefore, $P(X = Y = 0) = 1$.

23. If X and Y are independent, then

$$\mathrm{Cov}(X, Y) = E(X - E(X)) \cdot E(Y - E(Y)) = 0\ .$$

Let U have distribution

$$p_U = \begin{pmatrix} 0 & \pi/2 & \pi & 3\pi/2 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}\ .$$

Then let $X = \cos(U)$ and $Y = \sin(U)$. X and Y have distributions

$$p_X = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}\ ,$$

$$p_Y = \begin{pmatrix} 0 & 1 & 0 & -1 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}\ .$$

Thus $E(X) = E(Y) = 0$ and $E(XY) = 0$, so $\mathrm{Cov}(X, Y) = 0$. However, since

$$\sin^2(x) + \cos^2(x) = 1\,,$$

X and Y are dependent.

24. Consider the variance of S_n:

$$V(S_n) = \sum_{i=1}^n p_i(1-p_i) = \sum_{i=1}^n p_i - \sum_{i=1}^n p_i^2 \ ,$$

with the constraint

$$\sum_{i=1}^n p_i = np \ .$$

Assume that we have values of p_i that satisfy the constraint and that m of the values of p_i are equal to x and one is equal to y with $x > y$. By rearranging the terms if necessary we can assume that the first m are equal to x and the $(m+1)$st is equal to y. We shall show that we can increase the variance by making these $m+1$ values equal. To do this we define

$$\bar{p}_i = p_i - \frac{\epsilon}{m}, \quad \text{for } i = 1 \text{ to } m$$

and

$$\bar{p}_{m+1} = p_{m+1} + \epsilon \ ,$$

where

$$\epsilon = \frac{m}{m+1}(x-y) \ .$$

Then the new $\bar{p}_i$s satisfy the constraint, and the difference between the new variance $\bar{V}$ and the old variance V is

$$\bar{V} - V = m(x - \frac{\epsilon}{m})^2 + (y+\epsilon)^2 - mx^2 - y^2 \ .$$

After simplifying and substituting the value for ϵ this becomes

$$\bar{V} - V = \frac{m}{m+1}(x-y)^2 \ .$$

Since this value is positive we have increased the variance by making the first $m+1$ values equal. The same argument applies in case $y > x$. By induction we see that the variance is maximized by making all the values equal. (Note: Students who know about the technique of Lagrange multipliers will find this easier to prove using that method.)

25. (a) The expected value of X is

$$\mu = E(X) = \sum_{i=1}^{5000} iP(X=i) \ .$$

The probability that a white ball is drawn is

$$P(\text{white ball is drawn}) = \sum_{i=1}^n P(X=i)\frac{i}{5000} \ .$$

Thus

$$P(\text{white ball is drawn}) = \frac{\mu}{5000} \ .$$

(b) To have $P(\text{white,white}) = P(\text{white})^2$ we must have

$$\sum_{i=1}^{5000} (\frac{i}{5000})^2 P(X = i) = (\sum_{i=1}^{n} \frac{i}{5000} P(X = i))^2 \ .$$

But this would mean that $E(X^2) = E(X)^2$, or $V(X) = 0$. Thus we will have independence only if X takes on a specific value with probability 1.

(c) From (b) we see that

$$P(\text{white,white}) = \frac{1}{5000^2} E(X^2) \ .$$

Thus

$$V(X) = \frac{(\sigma^2 + \mu^2)}{5000^2} \ .$$

26. (a) $P(X = k) = pq^{k-1}, k = 1, 2, \ldots$ Thus by Example 3 we have

$$E(X_j) = \frac{1}{p},\ V(X_j) = \frac{q}{p^2} \ .$$

(b) $E(T_n) = n/p,\ V(T_n) = nq/p^2$.

(c) $E(T_n) = 2n,\ V(T_n) = 2n$.

27. The number of boxes needed to get the j'th picture has a geometric distribution with

$$p = \frac{(2n - k + 1)}{2n} \ .$$

Thus

$$V(X_j) = \frac{2n(k-1)}{(2n - k + 1)^2} \ .$$

Therefore, for a team of 26 players the variance for the number of boxes needed to get the first half of the pictures would be

$$\sum_{k=1}^{13} \frac{26(k-1)}{(26 - k + 1)^2} = 7.01 \ ,$$

and to get the second half would be

$$\sum_{k=14}^{26} \frac{26(k-1)}{(26 - k + 1)^2} = 979.23 \ .$$

Note that the variance for the second half is much larger than that for the first half.

6.3 Continuous Random Variables

1. (a) $\mu = 0,\ \sigma^2 = 1/3$

(b) $\mu = 0,\ \sigma^2 = 1/2$

(c) $\mu = 0,\ \sigma^2 = 3/5$

(d) $\mu = 0,\ \sigma^2 = 3/5$

2. (a) $\mu = 0,\ \sigma^2 = 1/5$

(b) $\mu = 0,\ \sigma^2 = \dfrac{\pi^2 - 8}{\pi^2}$

(c) $\mu = 1/3,\ \sigma^2 = 2/9$

(d) $\mu = 1/2,\ \sigma^2 = 3/20$

3. $\mu = 40,\ \sigma^2 = 800$

4. (a) $\int_{-1}^{1}(ax+b)dx = 2b = 1$, so $b = \frac{1}{2}$.

(b) $ax + \frac{1}{2} \geq 0$, so when $x = 1, a \geq -\frac{1}{2}$, and when $x = -1, a \leq \frac{1}{2}$. Thus $-\frac{1}{2} \leq a \leq \frac{1}{2}$.

(c) $\mu = \int_{-1}^{1}(ax^2 + bx)dx = \frac{2}{3}a$.

(d) $E(X^2) = \int_{-1}^{1}(ax^3 + bx^2)dx = \frac{2b}{3} = 1/3 - (4/9)a^2$. Thus $\sigma^2(X) = \frac{2}{3}b - \frac{4}{9}a^2$.

5. (d) $a = -3/2,\ b = 0,\ c = 1$

(e) $a = \dfrac{45}{48},\ b = 0,\ c = \dfrac{3}{16}$

6. (a) $\mu_T = \dfrac{1}{3}, \qquad \sigma_T^2 = \dfrac{1}{9}$.

(b) $\mu_T = \dfrac{1}{3}, \qquad \sigma_T^2 = \dfrac{2}{9}$.

(c) $\mu_T = \dfrac{1}{2}, \qquad \sigma_T^2 = \dfrac{3}{4}$.

7. $f(a) = E(X-a)^2 = \displaystyle\int (x-a)^2 f(x)dx$. Thus

$$
\begin{aligned}
f'(a) &= -\int 2(x-a)f(x)dx \\
&= -2\int xf(x)dx + 2a\int f(x)dx
\end{aligned}
$$

$$= \quad -2\mu(X) + 2a \ .$$

Since $f''(a) = 2$, $f(a)$ achieves its minimum when $a = \mu(X)$.

8. $a(\sigma^2 + \mu^2) + b\mu + c$.

9. (a) $3\mu,\ 3\sigma^2$

(b) $E(A) = \mu,\ V(A) = \dfrac{\sigma^2}{3}$

(c) $E(S^2) = 3\sigma^2 + 9\mu^2,\ E(A^2) = \dfrac{\sigma^2}{3} + \mu^2$

10. (a) $\dfrac{1}{3}$

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

(c) $\dfrac{1}{3}$

(d) $\dfrac{2}{3}$

(e) $\dfrac{7}{6}$

11. In the case that X is uniformly distributed on $[0, 100]$, one finds that

$$E(|X - b|) = \frac{1}{200}\Big(b^2 + (100 - b)^2\Big) \ ,$$

which is minimized when $b = 50$.

When $f_X(x) = 2x/10{,}000$, one finds that

$$E(|X - b|) = \frac{200}{3} - b + \frac{b^3}{15000} \ ,$$

which is minimized when $b = 50\sqrt{2}$.

12. $\displaystyle\int_0^1 \int_0^1 x^y\,dxdy = \log(2) \approx .693$.

13. Integrating by parts, we have

$$\begin{aligned} E(X) &= \int_0^\infty x dF(x) \\ &= -x(1 - F(x))\Big|_0^\infty + \int_0^\infty (1 - F(x))dx \end{aligned}$$

$$= \int_0^\infty (1 - F(x))dx \ .$$

To justify this argment we have to show that $a(1 - F(a))$ approaches 0 as a tends to infinity. To see this, we note that

$$\begin{aligned}
\int_0^\infty xf(x)dx &= \int_0^a xf(x)dx + \int_a^\infty xf(x)dx \\
&\geq \int_0^a xf(x)dx + \int_0^a af(x)dx \\
&= \int_0^a xf(x)dx + a(1 - F(a)) \ .
\end{aligned}$$

Letting a tend to infinity, we have that

$$E(X) \geq E(X) + \lim_{a \to \infty} a(1 - F(a)) \ .$$

Since both terms are non-negative, the only way this can happen is for the inequality to be an equality and the limit to be 0.

To illustrate this with the exponential density, we have

$$\int_0^\infty (1 - F(x))dx = \int_0^\infty e^{-\lambda x}dx = \frac{1}{\lambda} = E(X) \ .$$

15. $E(Y) = 9.5,\ E(Z) = 10,\ E(|X - Y|) = 1/2,\ E(|X - Z|) = 1/4$.

Z is better, since the expected value of the error committed by rounding using this method is one-half of that using the other method.

17. (a)

$$\begin{aligned}
\mathrm{Cov}(X, Y) &= E(XY) - \mu(X)E(Y) - E(X)\mu(Y) + \mu(X)\mu(Y) \\
&= E(XY) - \mu(X)\mu(Y) = E(XY) - E(X)E(Y) \ .
\end{aligned}$$

(b) If X and Y are independent, then $E(XY) = E(X)E(Y)$, and so $\mathrm{Cov}(X, Y) = 0$.

(c)

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

18. (a)

$$\begin{aligned} 0 &\leq V\Big(\frac{X}{\sigma(X)} + \frac{Y}{\sigma(Y)}\Big) \\ &= V\Big(\frac{X}{\sigma(X)}\Big) + V\Big(\frac{Y}{\sigma(Y)}\Big) + 2\mathrm{Cov}\Big(\frac{X}{\sigma(X))}, \frac{Y}{\sigma(Y)}\Big) \\ &= 1 + 1 + 2\frac{\mathrm{cov}(X,Y)}{\sqrt{V(X)V(Y)}} = 2(1+\rho(X,Y)) \ . \end{aligned}$$

(b)

$$\begin{aligned} 0 &\leq V\Big(\frac{X}{\sigma(X)} - \frac{Y}{\sigma(Y)}\Big) \\ &= V\Big(\frac{X}{\sigma(X)}) + V(\frac{Y}{\sigma(Y)}\Big) - 2\mathrm{Cov}\Big(\frac{X}{\sigma(X))}, -\frac{Y}{\sigma(Y)}\Big) \\ &= 1 + 1 - 2\frac{\mathrm{Cov}(X,Y)}{\sigma(X)\sigma(Y)} = 2(1-\rho(X,Y)) \ . \end{aligned}$$

(c) From (a),

$$1 + \rho(X,Y) \geq 0\,,$$

so

$$\rho(X,Y) \geq -1\,.$$

From (b),

$$1 - \rho(X,Y) \geq 0\,,$$

so

$$\rho(X,Y) \leq 1\,.$$

Thus

$$-1 \leq \rho(X,Y) \leq 1\,.$$

19. (a) 0

(b) $\frac{1}{\sqrt{2}}$

(c) $-\frac{1}{\sqrt{2}}$

(d) 0

20. (a)

$$\begin{aligned} f_X(x) &= \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \exp\Big(\frac{-(x^2 - 2\rho xy + y^2)}{2(1-\rho^2)}\Big) dy \\ &= \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \exp(-(y-\rho x)^2) \cdot \exp(-\frac{1}{2}x^2) dy \end{aligned}$$

$$= \frac{1}{\sqrt{2\pi}} \cdot \exp(-\frac{1}{2}x^2) \ .$$

Thus X has a standard normal density. By symmetry, Y also has a standard normal density.

(b)

$$\begin{aligned}
E(XY) &= \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy \cdot \exp\Big(\frac{-(x^2-2\rho xy+y^2)}{2(1-\rho^2)}\Big) dxdy \\
&= \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty}\Big(\int_{-\infty}^{\infty} y \cdot \exp\Big(\frac{-(y-\rho x)^2}{(2(1-\rho^2)}\Big) dy\Big) x \cdot \exp(-\frac{1}{2}x^2) dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \rho x^2 \cdot \exp(-\frac{1}{2}x^2) = \rho \ .
\end{aligned}$$

Now

$$\mathrm{Cov}(X,Y) = \frac{E(XY) - E(X)E(Y)}{\sqrt{V(X)V(Y)}} \ .$$

Since $E(X) = E(Y) = 0$ and $V(X) = V(Y) = 1$,

$$\mathrm{Cov}(X,Y) = E(XY) = \rho \, .$$

21. We have

$$\begin{aligned}
\frac{f_{XY}(x,y)}{f_Y(y)} &= \frac{\frac{1}{2\pi\sqrt{1-\rho^2}} \cdot \exp\Big(\frac{-(x^2-2\rho xy+y^2)}{2(1-\rho^2)}\Big)}{\sqrt{2\pi} \cdot \exp(-\frac{y^2}{2})} \\
&= \frac{1}{\sqrt{2\pi(1-\rho^2)}} \cdot \exp\left(\frac{-(x-\rho y)^2}{2(1-\rho^2)}\right)
\end{aligned}$$

which is a normal density with mean ρy and variance $1-\rho^2$. Thus,

$$\begin{aligned}
E(X|Y=y) &= \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi(1-\rho^2)}} \cdot \exp\Big(\frac{-(x-\rho y)^2}{2(1-\rho^2)}\Big)\, dx \\
&= \rho y \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi(1-\rho^2)}} \cdot \exp(-(x-\rho y)^2) \\
&= \begin{cases} \rho y < y, & \text{if } 0 < \rho < 1; \\ y, & \text{if } \rho = 1. \end{cases}
\end{aligned}$$

22. We have

$$\begin{aligned}
f_X(x) &= \int_0^1 f_{X,Y}(x,y)dy \\
&= \int_0^1 f_{X|Y}(x|y) f_Y(y) dy
\end{aligned}$$

$$
\begin{aligned}
&= \int_x^1 \frac{1}{y} dy \\
&= -\log x \ ,
\end{aligned}
$$

if $0 < x \le 1$.

24. (a) Since the father's height is 72 inches, $Y = (72 - 68)/2.7 = 1.48$. Therefore the density for X given Y is normal with mean $.5 \cdot 1.48 = .74$ and variance $1 - .5^2 = .75$. Thus the density for the son's height, given that the father's height is 72, is normal with mean $2.7 \cdot .74 + 68 = 70$ and variance $(2.7)^2 \cdot .75 = 5.47$.

26. (a) Let θ denote the angle that our path makes with the river bank, and assume without loss of generality that $0 \le \theta \le \pi/2$. Let X denote the distance from P to the river. Then $X = \sin(\theta)$. Thus, the cumulative distribution function of X is given by

$$
\begin{aligned}
F_X(x) &= P(X \le x) \\
&= P\big(\sin(\theta) \le x\big) \\
&= P(\theta \le \arcsin x) \\
&= \frac{2}{\pi} \arcsin x \ .
\end{aligned}
$$

So,

$$
f_X(x) = \frac{2}{\pi} \frac{1}{\sqrt{1 - x^2}} \ .
$$

Therefore,

$$
\begin{aligned}
E(X) &= \int_0^1 \frac{2}{\pi} \frac{x}{\sqrt{1 - x^2}} \, dx \\
&= \frac{2}{\pi} \Big[(1 - x^2)^{1/2} \Big]_0^1 \\
&= \frac{2}{\pi} \ .
\end{aligned}
$$

(b) For a fixed θ between 0 and $\pi/2$, let A_θ denote the set of angles α that you can choose at P and get back to the river by walking at most 1 mile in the direction α. If $\alpha = 0$ corresponds to the direction directly towards the river from P, then

$$
A_\theta = \Big[\theta - \frac{\pi}{2}, \frac{\pi}{2} - \theta \Big] \ .
$$

So the probability that you choose a good angle α, given that you are at P, is

$$
\frac{|A_\theta|}{2\pi} = \frac{\pi - 2\theta}{2\pi} \ .
$$

This must be averaged over all $\theta \in [0, \pi/2]$ to obtain the final answer:

$$\begin{aligned} \frac{2}{\pi}\int_0^{\pi/2} \frac{1}{2} - \frac{1}{\pi}\theta \, d\theta &= \frac{2}{\pi}\left[\frac{\theta}{2} - \frac{1}{2\pi}\theta^2\right]_0^{\pi/2} \\ &= \frac{1}{4}\ . \end{aligned}$$

27. Let Z represent the payment. Then

$$\begin{aligned} P(Z = k | X = x) &= P(Y_1 \le x, Y_2 \le x, \dots, Y_k \le x, Y_{k+1} > x) \\ &= x^k(1 - x)\ . \end{aligned}$$

Therefore,

$$\begin{aligned} P(Z = k) &= \int_0^1 x^k(1 - x)\, dx \\ &= \left[\frac{1}{k+1}x^{k+1} - \frac{1}{k+2}x^{k+2}\right]_0^1 \\ &= \frac{1}{k+1} - \frac{1}{k+2} \\ &= \frac{1}{(k+1)(k+2)}\ . \end{aligned}$$

Thus,

$$E(Z) = \sum_{k=0}^{\infty} k\left(\frac{1}{(k+1)(k+2)}\right)\ ,$$

which diverges. Thus, you should be willing to pay any amount to play this game.

Chapter 7

Sums of Independent Random Variables

7.1 Sums of Discrete Random Variables

1. (a) .625

(b) .5

2. $\begin{pmatrix} -2 & -1 & 0 & 1 & 2 & 3 & 4 \\ \frac{1}{16} & \frac{1}{4} & \frac{5}{16} & \frac{3}{16} & \frac{9}{64} & \frac{1}{32} & \frac{1}{64} \end{pmatrix}$.

3. $\begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ \frac{1}{64} & \frac{3}{32} & \frac{17}{64} & \frac{3}{8} & \frac{1}{4} \end{pmatrix}$

5. (a) $\begin{pmatrix} 3 & 4 & 5 & 6 \\ \frac{1}{12} & \frac{4}{12} & \frac{4}{12} & \frac{3}{12} \end{pmatrix}$

(b) $\begin{pmatrix} 1 & 2 & 3 & 4 \\ \frac{1}{12} & \frac{4}{12} & \frac{4}{12} & \frac{3}{12} \end{pmatrix}$

6. (a) $P(T_r = k) = \binom{r+k-1}{k} p^r q^k$.

(b) $P(C_r = k) = b(r, p, k) = \binom{r}{k} p^k q^{r-k}$.

(c) $E(C_r) = rp, \quad V(C_r) = rpq$.

7. (a) $P(Y_3 \le j) = P(X_1 \le j, X_2 \le j, X_3 \le j) = P(X_1 \le j)^3$.
Thus

$$p_{Y_3} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{216} & \frac{7}{216} & \frac{19}{216} & \frac{37}{216} & \frac{61}{216} & \frac{91}{216} \end{pmatrix} .$$

This distribution is not bell-shaped.

(b) In general,

$$P(Y_n \leq j) = P(X_1 \leq j)^3 = \left(\frac{j}{6}\right)^n .$$

Therefore,

$$P(Y_n = j) = \left(\frac{j}{6}\right)^n - \left(\frac{j-1}{6}\right)^n .$$

This distribution is not bell-shaped for large n.

8. (b) .304

(c) .325

9. Let $p_1, \ldots, p_6$ be the probabilities for one die and $q_1, \ldots, q_6$ be the probabilities for the other die. Assume first that all probabilities are positive. Then $p_1 q_1 > p_1 q_6$, since there is only one way to get a 2 and several ways to get a 7. Thus $q_1 > q_6$. In the same way $q_6 q_6 > q_1 p_6$ and so $q_6 > q_1$. This is a contradiction. If any of the sides has probability 0, then we can renumber them so that it is side 1. But then the probability of a 2 is 0 and so all sums would have to have probability 0, which is impossible.

   Here's a fancy way to prove it. Define the polynomials

$$p(x) = \sum_{k=0}^{5} p_{(k+1)} x^k$$

   and

$$q(x) = \sum_{k=0}^{5} q_{(k+1)} x^k .$$

   Then we must have

$$p(x)q(x) = \sum_{k=0}^{10} \frac{x^k}{11} = \frac{(1-x^{11})}{(1-x)} .$$

   The left side is the product of two fifth degree polynomials. A fifth degree polynomial must have a real root which will not be 0 if $p_1 > 0$. Consider the right side as a polynomial. For x to be a non-zero root of this polynomial it would have to be a real eleventh root of unity other than 1, and there are no such roots. Hence again we have a contradiction.

10. Let $n = rs$. Then consider the following two distributions

$$p_X = \begin{pmatrix} 0 & 1 & 2 & \ldots & r-1 \\ 1/r & 1/r & 1/r & \ldots & 1/r \end{pmatrix} ,$$

$$p_Y = \begin{pmatrix} 0 & r & 2r & \dots & (s-1)r \\ 1/s & 1/s & 1/s & \dots & 1/s \end{pmatrix} .$$

If X and Y are independent, then $X+Y$ takes on all possible values from 0 to $n-1$. Further, there is only one choice of X and Y that gives $X+Y$ a particular value and the probability for this choice is $1/rs$. Thus $X+Y$ has a uniform distribution on the values from 0 to $n-1$.

7.2 Sums of Continuous Random Variables

2. (a) $f_Z(x) = \dfrac{x+2}{4}$ on $[-2, 0]$ and $\dfrac{2-x}{4}$ on $[0, 2]$.

3. (a)

$$f_Z(x) = \begin{cases} x^3/24, & \text{if } 0 \le x \le 2; \\ x - x^3/24 - 4/3, & \text{if } 2 \le x \le 4. \end{cases}$$

(b)

$$f_Z(x) = \begin{cases} (x^3 - 18x^2 + 108x - 216)/24, & \text{if } 6 \le x \le 8; \\ (-x^3 + 18x^2 - 84x + 40)/24, & \text{if } 8 \le x \le 10. \end{cases}$$

(c)

$$f_Z(x) = \begin{cases} x^2/8, & \text{if } 0 \le x \le 2; \\ 1/2 - (x-2)^2/8, & \text{if } 2 \le x \le 4. \end{cases}$$

4.

$$f_Z(x) = \begin{cases} x^2/2, & \text{if } 0 \le x \le 1; \\ (-2x^2 + 6x - 3)/2, & \text{if } \ 1 \le x \le 2; \\ (x-3)^2/2, & \text{if } \ 2 \le x \le 3. \end{cases}$$

5. (a)

$$f_Z(x) = \begin{cases} \frac{\lambda\mu}{\mu+\lambda} e^{\lambda x}, & x < 0; \\ \frac{\lambda\mu}{\mu+\lambda} e^{-\mu x}, & x \ge 0. \end{cases}$$

(b)

$$f_Z(x) = \begin{cases} 1 - e^{-\lambda x}, & 0 < x < 1; \\ (e^{\lambda} - 1) e^{-\lambda x}, & x \ge 1. \end{cases}$$

6. Z is normally distributed with mean $\mu = \mu_1 + \mu_2$ and variance $\sigma^2 = \sigma_1^2 + \sigma_2^2$.

7. We first find the density for X^2 when X has a general normal density

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} dx \ .$$

Then (see Theorem 1 of Chapter 5, Section 5.2 and the discussion following) we have

$$f_X^2(x) = \frac{1}{\sigma\sqrt{2\pi}} \frac{1}{2\sqrt{x}} \exp(-x/2\sigma^2 - \mu^2/2\sigma^2)\Big(\exp(\sqrt{x}\mu/\sigma^2) + \exp(-\sqrt{x}\mu/\sigma^2)\Big) \ .$$

Replacing the last two exponentials by their series representation, we have

$$f_X^2(x) = e^{-\mu/2\sigma^2} \sum_{r=0}^{\infty} \Big(\frac{\mu}{2\sigma^2}\Big)^r \frac{1}{r!} G(1/2\sigma^2, r+1/2, x) \ ,$$

where

$$G(a,p,x) = \frac{a^p}{\Gamma(p)} e^{-ax} x^{p-1}$$

is the gamma density. We now consider the original problem with X_1 and X_2 two random variables with normal density with parameters μ_1, σ_1 and μ_2, σ_2. This is too much generality for us, and we shall assume that the variances are equal, and then for simplicity we shall assume they are 1. Let

$$c = \sqrt{\mu_1^2 + \mu_2^2} \ .$$

We introduce the new random variables

$$Z_1 = \frac{1}{c}(\mu_1 X_1 + \mu_2 X_2) \ ,$$

$$Z_2 = \frac{1}{c}(\mu_2 X_1 - \mu_1 X_2) \ .$$

Then Z_1 is normal with mean c and variance 1 and Z_2 is normal with mean 0 and variance 1. Thus,

$$f_{Z_1^2} = e^{-c^2/2} \sum_{r=0}^{\infty} \Big(\frac{c^2}{2}\Big)^r \frac{1}{r!} G(1/2, r+1/2, x) \ ,$$

and

$$f_{Z_2^2} = G(1/2, 1/2, x) \ .$$

Convoluting these two densities and using the fact that the convolution of a gamma density $G(a,p,x)$ and $G(a,q,x)$ is a gamma density $G(a,p+q,x)$ we finally obtain

$$f_{Z_1^2+Z_2^2} = f_{X_1^2+X_2^2} = e^{-c^2/2} \sum_{r=0}^{\infty} \Big(\frac{c^2}{2}\Big)^r \frac{1}{r!} G\Big(1/2, r+1, x\Big) \ .$$

(This derivation is adapted from that of C.R. Rao in his book *Advanced Statistical Methods in Biometric Research*, Wiley, l952.)

8.

$$f_{R^2} = \begin{cases} \pi/4, & \text{if } 0 \le x \le 1; \\ (1/2)\arcsin\Big((2-x)/x\Big), & \text{if } 1 \le x \le 2. \end{cases}$$

$$f_R = \begin{cases} (\pi/2)x, & \text{if } 0 \le x \le 1; \\ x\arcsin\Big((2-x^2)/x^2\Big), & \text{if } 1 \le x \le \sqrt{2}. \end{cases}$$

9. $P(X_{10} > 22) = .341$ by numerical integration. This also could be estimated by simulation.

10. $P(\min(X_1, \ldots, X_n > x) = (P(X_1 > x))^n = (e^{-x/\mu})^n = e^{-(n/\mu)x}$. Thus

$$f_{\min(X_1,\ldots,X_n)} = \frac{n}{\mu} e^{-(n/\mu)x} .$$

This is the exponential density with mean μ/n.

11. 10 hours

12. By Exercise 10 the first claim has the mean of $\mu/50$. If μ is about 30 years, then $\mu/50$ is about 7 months, which is practical. Once we have estimated $\mu/50$, we have an estimate for μ.

13. $Y_1 = -\log(X_1)$ has an exponential density $f_{Y_1}(x) = e^{-x}$. Thus S_n has the gamma density

$$f_{S_n}(x) = \frac{x^{n-1}e^{-x}}{(n-1)!} .$$

Therefore

$$f_{Z_n}(x) = \frac{1}{(n-1)!}\Big(\log\frac{1}{x}\Big)^{n-1} .$$

14. $X_3 = -X_2$ has density

$$f_{-X_2}(x) = \begin{cases} e^{\lambda x}, & -\infty < x \le 0; \\ 0, & \text{otherwise.} \end{cases}$$

Thus $Z = X_1 + X_3$ has density

$$\begin{aligned} f_Z(x) &= \int_0^\infty e^{\lambda(x-2y)}dy = \frac{1}{2\lambda}e^{\lambda x}, & x < 0; \\ &= \int_x^\infty e^{\lambda(x-2y)}dy = \frac{1}{2\lambda}e^{\lambda x}\Big(e^{-2\lambda x}\Big) = \frac{1}{2\lambda}e^{-\lambda x}, & x \ge 0. \end{aligned}$$

19. The support of $X + Y$ is $[a + c, b + d]$.

20. We prove it by induction. It is true for $n = 1$. Suppose that f_{S_k} has support on $[kc, kb]$. Then $f_{S_{k+1}} = f_{S_k} * f_X$ has support on $[ka+a, kb+b] = [(k+1)a, (k+1)b]$. (See Exercise 19 above.)

21. (a)

$$f_A(x) = \frac{1}{\sqrt{2\pi n}} e^{-x^2/(2n)} \ .$$

(b)

$$f_A(x) = n^n x^n e^{-nx}/(n-1)! \ .$$

Chapter 8

Law of Large Numbers

8.1 Law of Large Numbers for Discrete Random Variables

1. 1/9

3. We shall see that $S_n - n/2$ tends to infinity as n tends to infinity. While the difference will be small compared to $n/2$, it will not tend to 0. On the other hand the difference $S_n/n - 1/2$ does tend to 0.

4. You will lose on the average 1.41 percent of the money that you bet. Thus if you play a long time, you will lose a lot. The law of large numbers tells you that the probability that you will be ahead in the long run tends to 0.

5. $k = 10$

6. $V\left(\frac{S_n}{n} - p\right) = V\left(\frac{S_n}{n}\right) = \frac{p(1-p)}{n}$. Thus $P\left(\left|\frac{S_n}{n} - p\right| \geq \epsilon\right) \leq \frac{p(1-p)}{n\epsilon^2}$.

7.

$$\begin{aligned} p(1-p) &= \frac{1}{4} - \left(\frac{1}{4} - p + p^2\right) \\ &= \frac{1}{4} - (\frac{1}{2} - p)^2 \leq \frac{1}{4} \ . \end{aligned}$$

Thus, $\max_{0 \leq p \leq 1} p(1-p) = \frac{1}{4}$. From Exercise 6 we have that

$$P\left(|\frac{S_n}{n} - p| \geq \epsilon\right) \leq \frac{p(1-p)}{n\epsilon^2} \leq \frac{1}{4n\epsilon^2} \ .$$

8. No.

9.

$$\begin{aligned}
P(S_n \geq 11) &= P(S_n - E(S_n) \geq 11 - E(S_n)) \\
&= P(S_n - E(S_n) \geq 10) \\
&\leq \frac{V(S_n)}{10^2} = .01.
\end{aligned}$$

10.

$$\begin{aligned}
P(X \geq k+1) &= P(X - E(X) \geq k + 1 - E(X)) \\
&= P(X - E(X) \geq k) \\
&\leq \tfrac{V(X)}{k^2} = \frac{1}{k^2}.
\end{aligned}$$

11. No, we cannot predict the proportion of heads that should turn up in the long run, since this will depend upon which of the two coins we pick. If you have observed a large number of trials then, by the Law of Large Numbers, the proportion of heads should be near the probability for the coin that you chose. Thus, in the long run, you will be able to tell which coin you have from the proportion of heads in your observations. To be 95 percent sure, if the proportion of heads is less than .625, predict $p = 1/2$; if it is greater than .625, predict $p = 3/4$. Then you will get the correct coin if the proportion of heads does not deviate from the probability of heads by more than .125. By Exercise 7, the probability of a deviation of this much is less than or equal to $1/(4n(.125)^2)$. This will be less than or equal to .05 if $n > 320$. Thus with 321 tosses we can be 95 percent sure which coin we have.

12.

$$\begin{aligned}
P\left(\left|\frac{S_n}{n} - \frac{M_n}{n}\right| > \epsilon\right) &= P\left(\frac{1}{n}\left|\sum_{i=1}^{n}(X_i - m_i)\right| > \epsilon\right) \\
&= P\left(\left|\sum_{i=1}^{n}(X_i - m_i)\right| > n\epsilon\right) \\
&\leq \frac{1}{n^2\epsilon^2}\sum_{i=1}^{n}\sigma_k^2 < \frac{nR}{n^2\epsilon^2} \\
&= \frac{R}{n\epsilon^2}\ .
\end{aligned}$$

This last expression approaches 0 as n goes to ∞.

14.

$$E(|X - E(X)|) = \sum_{\omega} |X(\omega) - E(X)|$$

$$\begin{aligned} &\geq \sum_{\{\omega:|X(\omega)-E(X)|\geq\epsilon\}} |X(\omega)-E(X)| \\ &\geq \epsilon P(|X-E(X)|\geq\epsilon). \end{aligned}$$

15. Take as Ω the set of all sequences of 0's and 1's, with 1's indicating heads and 0's indicating tails. We cannot determine a probability distribution by simply assigning equal weights to all infinite sequences, since these weights would have to be 0. Instead, we assign probabilities to finite sequences in the usual way, and then probabilities of events that depend on infinite sequences can be obtained as limits of these finite sequences. (See Exercise 28 of Chapter 1, Section 1.2.)

16. The exercise as stated in the text is incorrect. The following replacement exercise, sent to us by David Maslen, is correct: In this exercise, we shall construct an example of a sequence of random variables that satisfies the weak law of large numbers, but not the strong law. The distribution of X_i will have to depend on i, because otherwise both laws would be satisfied. As a preliminary, we need to prove a lemma, which is one of the Borel-Cantelli lemmas. Suppose we have an infinite sequence of mutually independent events $A_1, A_2, \ldots$. Let $a_i = \text{Prob}(A_i)$, and let r be a positive integer.

(a) Find an expression of the probability that none of the A_i with $i > r$ occur.

(b) Use the fact that $x - 1 \leq e^{-x}$ to show that

$$\text{Prob(No } A_i \text{ with } i > r \text{ occurs)} \leq e^{-\sum_{i=r}^{\infty} a_i} \ .$$

(c) Prove that if $\sum_{i=1}^{\infty} a_i$ diverges, then

$$\text{Prob(infinitely many } A_i \text{ occur)} = 1 \ .$$

Now, let X_i be a sequence of mutually independent random variables such that for each positive integer $i \geq 2$,

$$\text{Prob}(X_i = i) = \frac{1}{2i\log i}\ , \quad \text{Prob}(X_i = -i) = \frac{1}{2i\log i}\ , \quad \text{Prob}(X_i = 0) = 1 - \frac{1}{i\log i}\ .$$

When $i = 1$ we let $X_i = 0$ with probability 1. As usual we let $S_n = X_1 + \cdots + X_n$. Note that the mean of each X_i is 0.

(d) Find the variance of S_n.

(e) Show that the sequence $\{X_i\}$ satisfies the weak law of large numbers, i.e. prove that for any $\epsilon > 0$

$$\text{Prob}\left(\left|\frac{S_n}{n}\right| \geq \epsilon\right) \to 0\ ,$$

as n tends to infinity. We now show that $\{X_i\}$ does not satisfy the strong law of large numbers. Suppose that $S_n/n \to 0$. Then because

$$\frac{X_n}{n} = \frac{S_n}{n} - \frac{n-1}{n}\frac{S_{n-1}}{n-1}\ ,$$

we know that $X_n/n \to 0$. From the definition of limits, we conclude that the inequality $|X_i| \geq i/2$ can only be true for finitely many i.

(f) Let A_i be the event $|X_i| \geq i/2$. Find $\text{Prob}(A_i)$. Show that

$$\sum_{i=1}^{\infty} \text{Prob}(A_i)$$

diverges (think Integral Test).

(g) Prove that A_i occurs for infinitely many i.

(h) Prove that

$$\text{Prob}\left(\frac{S_n}{n} \to 0\right) = 0\ ,$$

and hence that the Strong Law of Large Numbers fails for the sequence $\{X_i\}$.

17. For $x \in [0,1]$, let us toss a biased coin that comes up heads with probability x. Then

$$E\left(\frac{f(S_n)}{n}\right) \to f(x).$$

But

$$E\left(\frac{f(S_n)}{n}\right) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-x)^{n-k}.$$

The right side is a polynomial, and the left side tends to f(x). Hence

$$\sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-x)^{n-k} \to f(x).$$

This shows that we can obtain obtain any continuous function $f(x)$ on [0,1] as a limit of polynomial functions.

8.2 Law of Large Numbers for Continuous Random Variables

1. (a) 1

(b) 1

(c) 100/243

(d) 1/12

2. (a) $E(X) = 10,\ V(X) = 100/3.$

(b) $P(|X - 10| \geq 2) = 4/5, \quad P(|X - 10| \geq 5) = 1/2.$

$P(|X - 10| \geq 9) = 1/10, \quad P(|X - 10| \geq 20) = 0.$

3.

$$f(x) = \begin{cases} 1 - x/10, & \text{if } 0 \leq x \leq 10; \\ 0 & \text{otherwise.} \end{cases}$$

$$g(x) = \frac{100}{3x^2}\ .$$

4. (a) E$(X) = 1/\lambda = 10$, V$(X) = (1/\lambda)^2 = 100$.

(b) For the first three probabilities Chebyshev's estimate is greater than 1, and so the best estimate is 1. For the last one Chebyshev's estimate gives $P(|X - 10| \geq 20) \leq .25$.

(c) Comparing these Chebyshev's estimates with the exact values, we have:

$$(1, .852),\ (1, .617),\ (1, .245),\ (.25, .0498).$$

5. (a) 1, 1/4, 1/9

(b) 1 vs. .3173, .25 vs. .0455, .11 vs. .0027

6. (a) 1, (b) 1/4, (c) 1/ 9, (d) 1/ 16.

The exact values are (a) .3173, (b) .0455, (c) .0027, (d) 0.

7. (b) 1, 1, 100/243, 1/12

8. (a)

$$\begin{aligned} P(|X^*| \geq a) &= P\left(\left|\frac{X - \mu}{\sigma}\right| \geq a\right) \\ &= P(|X - \mu| \geq a\sigma) \\ &\leq \tfrac{\sigma^2}{\sigma^2 a^2} = \tfrac{1}{a^2}. \end{aligned}$$

(b) $P(|X^*| \geq 2) = 1/4, \quad P(|X^*| \geq 5) = 1/25.$

$P(|X^*| \geq 9) = 1/81.$

9. (a) 0

(b) 7/12

(c) 11/12

10. (a)

$$\begin{aligned} P(65 \leq X \leq 75) &= P(65 - 70 \leq X - 70 \leq 75 - 70) \\ &= P(-5 \leq X - 70 \leq 5) \\ &= 1 - P(|X - 70| \geq 5) \\ &\geq 1 - 25/25 = 0. \end{aligned}$$

Thus Chebyshev's estimate gives us a useless lower bound in this case.

(b) $E(\bar{X}) = 70,\ V(\bar{X}) = 25/100 = .25$. Thus

$$\begin{aligned} P(65 \leq \bar{X} \leq 75) &= 1 - P(|\bar{X} - 70| \geq 5) \\ &\geq 1 - \frac{.25}{25} = .99. \end{aligned}$$

Therefore, Chebyshev's estimate gives a lower bound of .99.

11. (a) 0

(b) 7/12

12. (a) $E(Y_2) = 30, \quad V(Y_2) = \frac{1}{4}. \quad$ Thus $P(25 \leq Y_2 \leq 35) \geq .99.$

(b) $E(Y_{11}) = 30, \quad V(Y_{11}) = \frac{10}{4}. \quad$ Thus $P(25 \leq Y_{11} \leq 35) \geq .9.$

(c) $E(Y_{101}) = 30, \quad V(Y_{101}) = \frac{100}{4}. \quad$ Thus $P(25 \leq Y_{101} \leq 35) \geq 0.$

13. (a) 2/3

(b) 2/3

(c) 2/3

17. $E(X) = \int_{-\infty}^{\infty} xp(x)dx$. Since X is non-negative, we have

$$E(X) \geq \int_{x \geq a} xp(x)dx \geq aP(X \geq a)\ .$$

18. Since $E(X) = 20$ and X is non-negative, we have:

$$20 = \int_0^\infty xp(x)dx \geq \int_a^\infty xp(x)dx \geq aP(X \geq a)\ .$$

Therefore,

$$P(X \geq a) \leq \frac{20}{a}\ .$$

This is interesting only for $a \geq 20$.

(b) Now assume $E(X) = 20$ and $V(X) = 25$. Then

$$E(X^2) = V(X) + E(X)^2 = 425\ .$$

Thus

$$425 = \int_0^\infty x^2 p(x)dx \geq \int_a^\infty x^2 dx \geq a^2 P(X \geq a)\ .$$

Therefore

$$P(X \geq a) \leq \frac{425}{a^2}\ .$$

From part (a) we also have that $P(X \geq a) \leq \dfrac{20}{a}$. Thus our best upper bounds are:

$$P(X \geq a) \leq \frac{20}{a} \quad \text{if } 20 \leq a \leq 21.25\ ,$$

and

$$P(X \geq a) \leq \frac{425}{a^2} \quad \text{if } a \geq 21.25\ .$$

(c) Since X is non-negative and the density is symmetric with mean 20, we must have $p(x)$ positive only on the interval [0,40]. Again by symmetry we have

$$10 = \int_{20}^{40} xp(x)\, dx\ .$$

Thus for $a \geq 20$,

$$10 = \int_{20}^{40} xp(x)\, dx \geq \int_a^{40} xp(x)\, dx \geq aP(X \geq a)\ .$$

Therefore,

$$P(X \geq a) \leq \frac{10}{a}\ .$$

Again by symmetry we have

$$\int_{20}^{40} (x-20)^2 p(x)dx = 12.5\ .$$

Then

$$\int_{20}^{40} x^2 p(x)dx - 40 \int_{20}^{40} xp(x)dx + 400 \cdot \frac{1}{2} = 12.5 \ .$$

From this we obtain

$$\int_{20}^{40} x^2 p(x)dx = 12.5 + 40 \cdot 10 - 200 = 212.5 \ .$$

Therefore, for $a \geq 20$ we have

$$212.5 = \int_{20}^{40} x^2 p(x)dx \geq \int_{a}^{40} xp(x)dx \geq a^2 P(X \geq a) \ .$$

Thus for $a \geq 20$ we have

$$P(X \geq a) \leq \frac{212.5}{a^2} \ .$$

Combining our two estimates we have:

$$P(X \geq a) \leq \frac{10}{a} \quad \text{if } 20 \leq a \leq 21.25 \ ,$$

and

$$P(X \geq a) \leq \frac{212.5}{a^2} \quad \text{if } 21.25 \leq a \leq 40 \ .$$

Chapter 9

Central Limit Theorem

9.1 Central Limit Theorem for Discrete Independent Trials

(The answers to the problems in this chapter do not use the '1/2 correction' mentioned in Section 9.1.)

1. (a) .158655

(b) .6318

(c) .0035

(d) .9032

2. (a) .0564

(b) .0208

(c) 1.033×10^{-3}

3. (a) $P(\text{June passes}) \approx .985$

(b) $P(\text{April passes}) \approx .056$

4. (a) $P(499,500 < S_{1,000,000} < 500,500) \geq 0$ by Chebyschev.

(b) $P(499,500 < S_{1,000,000} < 500,500) \approx .6826$ by the Central Limit Theorem.

(a) $P(499,000 < S_{1,000,000} < 501,000) \geq .75$ by Chebyschev.

(b) $P(499,000 < S_{1,000,000} < 501,000) \approx .9545$ by the Central Limit Theorem.

(a) $P(498,500 < S_{1,000,000} < 501,500) \geq .8889$ by Chebyschev.

(b) $P(498,500 < S_{1,000,000} < 501,500) \approx .9973$ by the Central Limit Theorem.

5. Since his batting average was .267, he must have had 80 hits. The probability that one would obtain 80 or fewer successes in 300 Bernoulli trials, with individual probability of success .3, is approximately .115. Thus, the low average is probably not due to bad luck (but a statistician would not reject the hypothesis that the player has a probability of success equal to .3).

6. We need to choose k so that $P(S_{1000} \leq k) \geq .99$. This is the same as

$$P\left(S^*_{1000} \leq \frac{k-500}{15.81}\right) \geq .99 \ .$$

Thus we want

$$\frac{k-500}{15.81} = 2.33 \ .$$

This will be true if $k = 537$.

7. .322

8. We want $np + 2\sqrt{npq} = 108$ and $np - 2\sqrt{npq} = 72$. Adding and subtracting gives $2np = 180$ and $4\sqrt{npq} = 36$. Solving these two equations for n and p gives
$p = .1$ and $n = 900$.

9. (a) 0

(b) 1 (Law of Large Numbers)

(c) .977 (Central Limit Theorem)

(d) 1 (Law of Large Numbers)

10. We want

$$P(S_{10,000} \leq 931) = P(S^*_{10,000} \leq \frac{931-1000}{30}) = P(S^*_{10,000} \leq -2.3) \approx .011.$$

12. 13

13. $P\left(S_{1900} \geq 115\right) = P\left(S^*_{1900} \geq \frac{115-95}{\sqrt{1900 \cdot .05 \cdot .95}}\right) = P\left(S^*_{1900} \geq 2.105\right) =$ .0176.

14. (a) 64 to 96

(b) 6400

16. $n = 108, \ m = 77$

17. We want $\frac{2\sqrt{pq}}{\sqrt{n}} = .01$. Replacing $\sqrt{pq}$ by its upper bound $\frac{1}{2}$, we have $\frac{1}{\sqrt{n}} = .01$. Thus we would need $n = 10{,}000$. Recall that by Chebyshev's inequality we would need 50,000.

9.2 Central Limit Theorem for Discrete Independent Trials

1. (a) .4762

(b) .0477

2. .3174

3. (a) .5

(b) .9987

5. (a) $P(S_{210} < 700) \approx .0757$.

(b) $P(S_{189} \geq 700) \approx .0528$

(c)

$$\begin{aligned} P(S_{179} < 700, S_{210} \geq 700) &= P(S_{179} < 700) - P(S_{179} < 700, S_{210} < 700) \\ &= P(S_{179} < 700) - P(S_{210} < 700) \\ &\approx .9993 - .0757 = .9236\ . \end{aligned}$$

6. (a) The expected loss is .2 cents and the variance of this loss is .36.

(b) .2024

(c) .047

(d) .9994

(e) 54

7. (a) Expected value = 200, variance = 2

(b) .9973

8. $P(S_{30} = 0) \approx \frac{N(0)}{\sqrt{30 \cdot 1.5}} = .0595$.

9. $P\Big(\Big|\frac{S_n}{n} - \mu\Big| \geq \epsilon\Big) = P\Big(\Big|S_n - n\mu\Big| \geq n\epsilon\Big) = P\Big(\Big|\frac{S_n - n\mu}{\sqrt{n\sigma^2}}\Big| \geq \frac{n\epsilon}{\sqrt{n\sigma^2}}\Big).$

   By the Central Limit Theorem, this probability is approximated by the area under the normal curve between $\frac{\sqrt{n}\epsilon}{\sigma}$ and infinity, and this area approaches 0 as n tends to infinity.

10. (a) The law of large numbers states that the average of Peter's fortune will be close to 0.

(b) The Central Limit Theorem states, for example, that with probability .95 Peter will not have won or lost more than \$2 after the 10,000 plays.

11. Her expected loss is 60 dollars. The probability that she lost no money is about .0013.

12. Betting 1 dollar on red gives $E(X) = -\frac{1}{74}$ and Var(X)= .980. Betting 1 dollar on 17 gives $E(X) = -\frac{1}{37}$ and Var$(X) = 34.08$. Thus, by the Central Limit Theorem, if we bet 1 dollar on red for 100 plays,

$$P(S_{100} > 0) = P(S^*_{100} > 0.137) \approx .446.$$

If we bet 1 dollar on 17 for 100 plays, we have

$$P(S_{100} > 0) = P(S^*_{100} > .046) \approx .482.$$

Note that for this game, the Central Limit Theorem must give an approximation that you are ahead which is less than .5, because the mean of each play is negative. Nonetheless, the actual probability is greater than .5. Thus, if we were to bet that we would be ahead after 100 plays, this would be a favorable bet, although the Central Limit Theorem approximation makes it seem that the bet is unfavorable.

13. p = .0056

9.3 Central Limit Theorem for Continuous Independent Trials

1.

$$\begin{aligned} E(X^*) &= \frac{1}{\sigma}(E(X) - \mu) = \frac{1}{\sigma}(\mu - \mu) = 0 \ , \\ \sigma^2(X^*) &= E\Big(\frac{X - \mu}{\sigma}\Big)^2 = \frac{1}{\sigma^2}\sigma^2 = 1 \ . \end{aligned}$$

2. $S_n^* = \frac{S_n - n\mu}{\sqrt{n}\sigma} = \frac{S_n}{\sqrt{n}} = \frac{nA_n}{\sqrt{n}} = \sqrt{n}A_n$.

3. $T_n = Y_1 + Y_2 + \cdots + Y_n = \frac{S_n - n\mu}{\sigma}$. Since each Y_j has mean 0 and variance 1, $E(T_n) = 0$ and $V(T_n) = n$. Thus $T_n^* = \frac{T_n}{\sqrt{n}} = \frac{S_n - n\mu}{\sigma\sqrt{n}} = S_n^*$.

4. For one uniform random variable on [0,20] the mean is 10 and the variance is 400/12. For the sum of 25 such random variables the mean is 250 and the standard deviation is $\sqrt{25(400/12)} = 28.87$. Thus the normal density used to approximate the sum is:

$$f(x) = \frac{1}{28.87}\frac{1}{\sqrt{2\pi}}\ \exp\left(-\frac{1}{2}\left(\frac{x-250}{28.87}\right)^2\right)\ .$$

For the standardized sum S^* the density for the normal approximation is the density with mean 0 and standard deviation 1:

$$f(x) = \frac{1}{\sqrt{2\pi}}\ e^{-x^2/2}\ .$$

The average of 25 numbers A_{25} has expected value 10 and standard deviation $\sqrt{(400/12)/25} = 1.155$. Thus the normal density used to approximate the average is:

$$f(x) = \frac{1}{1.155}\frac{1}{\sqrt{2\pi}}\ exp\left(-\frac{1}{2}\left(\frac{x-10}{1.155}\right)^2\right).$$

10. By Chebyshev's inequality we would need $10/n \leq .05$ or $n \geq 200$. By the Central Limit Theorem we would need $2\sqrt{10/n} \approx 1$, or $n \approx 40$. To find the variance necessary for 10 measurements to suffice using Chebyshev's inequality, we would need $\sigma^2/10 \approx .05$, or $\sigma^2 \approx .5$. Using the Central Limit Theorem we would need $2\sqrt{\sigma^2/10} \approx 1$, or $\sigma^2 \approx 2.5$.(A larger variance is easier to obtain.)

11. (a) .5

(b) .148

(c) .018

13. 7.6×10^{-24}

14. (b) (20.53, 25.87)

Chapter 10

Generating Functions

10.1 Generating Functions for Discrete Distributions

1. In each case, to get $g(t)$ just replace z by e^t in $h(z)$.

(a) $h(z) = \frac{1}{2}(1+z)$

(b) $h(z) = \frac{1}{6}\sum_{j=1}^{6} z^j$

(c) $h(z) = z^3$

(d) $h(z) = \frac{1}{k+1} z^n \sum_{j=0}^{k} z^j$

(e) $h(z) = z^n(pz+q)^k$

(f) $h(z) = \frac{2}{3-z}$

2. (a) $\mu_1 = 1/2$, $\mu_2 = 1/2$, $h'(1) = 1/2 = \mu_1$, $h''(1) = 0 = \mu_2 - \mu_1$.

(b) $\mu_1 = 7/2$, $\mu_2 = 91/6$, $h'(1) = 7/2$, $h''(1) = 70/6 = \mu_2 - \mu_1$.

(c) $\mu_1 = 3$, $\mu_2 = 9$, $h'(1) = 3$, $h''(1) = 6 = \mu_2 - \mu_1$.

(d) $\mu_1 = n + k/2$, $\mu_2 = n^2 + nk + k(1+2k)/6$,

$$h'(1) = n + k/2 = \mu_1, \qquad h''(1) = n^2 + (k-1)n + k(k-1)/3 = \mu_2 - \mu_1.$$

3. (a) $h(z) = \frac{1}{4} + \frac{1}{2}z + \frac{1}{4}z^2$.

(b) $g(t) = h(e^t) = \frac{1}{4} + \frac{1}{2}e^t + \frac{1}{4}e^{2t}$.

(c)

$$\begin{aligned} g(t) &= \frac{1}{4} + \frac{1}{2}\Big(\Big(\sum_{k=0}^{\infty} \frac{t^k}{k!}\Big) + \frac{1}{4}\Big(\sum_{k=0}^{\infty} \frac{2^k}{k!} t^k\Big) \\ &= 1 + \sum_{k=1}^{\infty}\Big(\frac{1}{2k!} + \frac{2^{k-2}}{k!}\Big)t^k = 1 + \sum_{k=1}^{\infty} \frac{\mu_k}{k!} t^k \ . \end{aligned}$$

Thus $\mu_0 = 1$, and $\mu_k = \frac{1}{2} + 2^{k-2}$ for $k \geq 1$.

(d) $p_0 = \frac{1}{4}, \quad p_1 = \frac{1}{2}, \quad p_2 = \frac{1}{4}$.

4. $h(z) = \left(1 - \frac{3}{2}\mu_1 + \frac{1}{2}\mu_2\right) + (2\mu_1 - \mu_2)z + \frac{(\mu_2 - \mu_1)}{2} z^2$. Thus $p_0 = 1 - \frac{3}{2}\mu_1 + \frac{1}{2}\mu_2, \quad p_1 = 2\mu_1 - \mu$

5. (a) $\mu_1(p) = \mu_1(p') = 3, \ \mu_2(p) = \mu_2(p') = 11$

$\mu_3(p) = 43, \ \mu_3(p') = 47$

$\mu_4(p) = 171, \ \mu_4(p') = 219$

6. (a) $p_2 = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1/4 & 4/9 & 4/9 \end{pmatrix}$.

(b) $h(z) = \frac{z}{3}(1 + 2z), \qquad h_2(z) = \frac{z^2}{9}(1 + 4z + 4z^2) = \big(h(z)\big)^2$.

(c) $h_n(z) = \Big(\frac{z}{3}(1 + 2z)\Big)^n$.

(d) $\mu_1 = \frac{5}{3}n, \qquad \mu_2 = \frac{25}{9}n^2 + \frac{2}{9}n$.

Thus the mean of $p_n = \frac{5}{3}n$, and the variance of $p_n = \frac{2}{9}n$. Since p has mean $\frac{5}{3}$, we see that the mean of p_n is n times the mean of p.

(e) $p_n(j) > 0$ for $j = n, \cdots, 2n$.

7. (a) $g_{-X}(t) = g(-t)$

(b) $g_{X+1}(t) = e^t g(t)$

(c) $g_{3X}(t) = g(3t)$

(d) $g_{aX+b} = e^{bt} g(at)$

8.

$$\begin{aligned}
(a) \quad h(z) &= \frac{1}{3}(z+2)\ , \\
(b) \quad h_2(z) &= \left(\frac{1}{3}(z+2)\right)^2\ , \\
(c) \quad h_n(z) &= \left(\frac{1}{3}(z+2)\right)^n\ .
\end{aligned}$$

Note: $\mu = 1/3$ and $\sigma = \sqrt{2}/3$.

9. (a) $h_X(z) = \sum_{j=1}^{6} a_j z^j, \qquad h_Y(z) = \sum_{j=1}^{6} b_j z^j.$

(b) $h_Z(z) = \left(\sum_{j=1}^{6} a_j z^j\right)\left(\sum_{j=1}^{6} b_j z^j\right)\ .$

(c) Assume that $h_Z(z) = (z^2 + \cdots + z^{12})/11$. Then

$$\left(\sum_{j=1}^{6} a_j z^{j-1}\right)\left(\sum_{j=1}^{6} b_j z^{j-1}\right) = \frac{1 + z + \cdots z^{10}}{11} = \frac{z^{11}-1}{11(z-1)}\ .$$

Either $\sum_{j=1}^{6} a_j z^{j-1}$or $\sum_{j=1}^{6} b_j z^{j-1}$ is a polynomial of degree 5 (i.e., either $a_6 \neq 0$ or $b_6 \neq 0$). Suppose that $\sum_{j=1}^{6} a_j z^{j-1}$ is a polynomial of degree 5. Then it must have a real root, which is a real root of $(z^{11}-1)/(z-1)$. However $(z^{11}-1)/(z-1)$ has no real roots. This is because the only real root of $z^{11}-1$ is 1, which cannot be a real root of $(z^{11}-1)/(z-1)$. Thus, we have a contradiction. This means that you cannot load two dice in such a way that the probabilities for any sum from 2 to 12 are the same. (cf. Exercise 11 of Section 7.1).

10.

$$\begin{aligned}
h(1) &= \frac{1-\sqrt{1-4pq}}{2q} = \frac{1-\sqrt{1-4p+4p^2}}{2q} = \frac{1-|2p-1|}{2q} \\
&= \begin{cases} q/p, & \text{if } p \leq q, \\ 1, & \text{if } p \geq q. \end{cases}
\end{aligned}$$

$$h'(z) = \begin{cases} \dfrac{4pqz}{2qz\sqrt{1-4pqz^2}} - \dfrac{1}{2qz^2}(1-\sqrt{1-4pqz})\ , & \text{if } p > q, \\ -\dfrac{1}{z^2}\left(1-\sqrt{1-z^2}\right) + \dfrac{1}{\sqrt{1-z^2}}\ . & \text{if } p = q. \end{cases}$$

Therefore,

$$h'(1) = \begin{cases} 1/(p-q), & \text{if } p > q, \\ \infty, & \text{if } p = q. \end{cases}$$

14.

$$g_{X^*}(t) = E(e^{X^* t}) = E(e^{\frac{X-\mu}{\sigma} t}) = e^{-\frac{\mu}{\sigma} t} E(e^{\frac{X}{\sigma} t}) = e^{-\frac{\mu}{\sigma} t} g_X\left(\frac{t}{\sigma}\right) .$$

10.2 Branching Processes

1. (a) $d = 1$

(b) $d = 1$

(c) $d = 1$

(d) $d = 1$

(e) $d = 1/2$

(f) $d \approx .203$

2. (a) .618, (b) .414, (c) 0 if $t = 0$, $\dfrac{-1+\sqrt{5}}{2}$ =.618 if $t \neq 0$

3. (a) 0

(b) 276.26

4.

$$\begin{aligned} h(z) = E\Big(Z^{S_n}\Big) &= \sum_k E(Z^{S_k}|N = k)P(N = k) \\ &= \sum_k \Big(E(Z^{X_1})\Big)^k P(N = k) \\ &= \sum_k (f(z))^k P(N = k) \\ &= g(f(z)) \ . \end{aligned}$$

5. Let Z be the number of offspring of a single parent. Then the number of offspring after two generations is

$$S_N = X_1 + \cdots + X_N \ ,$$

where $N = Z$ and X_i are independent with generating function f. Thus by Exercise 4, the generating function after two generations is $h(z) = f(f(z))$.

6. (a) $f(z) = p + qz,\ g(z) = \dfrac{rz}{1-(1-r)z}$.

(b) Let N be the time she needs to be served. Then the number of customers arriving during this time is $X_1 + \cdots + X_N$, where X_i are identically distributed independent of N. $P(X_0 = 0) = p,\ P(X_i = 1) = q$. Thus by Exercise 4, $h(z) = g(f(z))$.

(c) The server ultimately has a time when he is not busy if the branching process dies out. For this we need $m \leq 1$ or $h'(1) = 1$. But $h(z) = g(f(z))$, so we need

$$h'(1) = g'(1)f'(1) = \text{ mean arrival rate} \cdot \text{ mean service time } \leq 1\ .$$

This means that we need $q/r \leq 1$.

7. If there are k offspring in the first generation, then the expected total number of offspring will be kN, where N is the expected total numer for a single offspring. Thus we can compute the expected total number by counting the first offspring and then the expected number after the first generation. This gives the formula

$$N = 1 + \Bigl(\sum_k kp_k\Bigr) = 1 + mN\ .$$

From this it follows that N is finite if and only if $m < 1$, in which case

$$N = 1/(1-m).$$

8. You can work this by passing to the limit in the expressions given in Example 4, but it is easier to do it directly as follows: The generating function $h_1(z) = h(z)$ for the population after one generation is

$$h(z) = \sum_{k=0}^{\infty} (1/2)^{j+1} z^j = \frac{1/2}{1-(1/2)z} = \frac{1}{(2-z)}\ .$$

Then we can get the generating functions for future generations by using the relation $h_{n+1}(z) = h_n(h(z))$. For example,

$$h_2(z) = \frac{1}{(2 - \frac{1}{(2-z)})} = \frac{2-x}{3-2x}\ .$$

Continuing in this way, we get

$$h_3(z) = \frac{3-2x}{4-3x}\ .$$

These results suggest that the general case is

$$h_n(z) = \frac{n - x(n-1)}{(n+1) - nx}\ .$$

It is easy to check that this satisfies the equation $h_{n+1} = h_n(h(z))$, so by induction we see that our guess for $h_n(z)$ is correct. Then

$$\begin{aligned} h_n(z) &= \frac{n - x(n-1)}{n+1}\Big(\frac{1}{1 - \frac{n}{n+1}x}\Big) \\ &= \frac{n - z(n-1)}{n+1}\Big(\sum_{j=0}^{\infty}\Big(\frac{n}{n+1}\Big)^j z^j \ . \end{aligned}$$

The constant term is $p_n(0) = n/(n+1)$. Collecting the coefficients of z^j and simplifying gives $p^{(n)}(j) = \dfrac{1}{n(n+1)}\Big(\dfrac{n}{n+1}\Big)^j$.

(b) The probability that the population dies out at the nth generation is equal to the difference between the probability that it has died out by the nth generation and the probability that it has died out by the $(n-1)$st generation. This is:

$$p_0^n - p_0^{n-1} = \frac{n}{n+1} - \frac{n-1}{n} = \frac{1}{n(n+1)} \ .$$

(c) The expected lifetime is

$$\begin{aligned} \sum_n n \cdot \mathrm{P}(\text{population dies out on the } n\text{th generation}) &= \sum_{n=1}^{\infty} \frac{n}{n(n+1)} \\ &= \sum_{n=1}^{\infty} \frac{1}{n+1} = \infty \ . \end{aligned}$$

10.3 Generating Functions for Continuous Densities

1. (a) $g(t) = \dfrac{1}{2t}(e^{2t} - 1)$

(b) $g(t) = \dfrac{e^{2t}(2t-1)+1}{2t^2}$

(c) $g(t) = \dfrac{e^{2t} - 2t - 1}{2t^2}$

(d) $g(t) = \dfrac{e^{2t}(ty-1) + 2e^t - t - 1}{t^2}$

(e) $(3/8)\left(\frac{e^{2t}(4t^2-4t+2)-2}{t^3}\right)$

2. (a) $\mu_1 = 1 = g'(0), \qquad \mu_2 = \frac{4}{3} = g''(0)$.

(b) $\mu_1 = \frac{4}{3} = g'(0), \qquad \mu_2 = 2 = g''(0)$.

(c) $\mu_1 = \frac{2}{3} = g'(0), \qquad \mu_2 = \frac{2}{3} = g''(0)$.

(d) $\mu_1 = 1 = g'(0), \qquad \mu_2 = \frac{3}{2} = g''(0)$.

(e) $\mu_1 = \frac{3}{2} = g'(0), \qquad \mu_2 = \frac{12}{5} = g''(0)$.

3. (a) $g(t) = \frac{2}{2-t}$

(b) $g(t) = \frac{4-3t}{2(1-t)(2-t)}$

(c) $g(t) = \frac{4}{(2-t)^2}$ (d) $g(t) = \left(\frac{\lambda}{\lambda+t}\right), \quad t < \lambda$.

4. (a) $\mu_1 = \frac{1}{2} = g'(0), \qquad \mu_2 = \frac{1}{2} = g''(0)$.

(b) $\mu_1 = \frac{3}{4} = g'(0), \qquad \mu_2 = \frac{5}{4} = g''(0)$.

(c) $\mu_1 = 1 = g'(0), \qquad \mu_2 = \frac{3}{2} = g''(0)$.

(d) $\mu_1 = \frac{n}{\lambda} = g'(0), \qquad \mu_2 = \frac{n(n+1)}{\lambda^2} = g''(0)$.

5. (a) $k(\tau) = \frac{1}{2i\tau}(e^{2i\tau}-1)$

(b) $k(\tau) = \frac{e^{2i\tau}(2i\tau-1)+1}{-2\tau^2}$

(c) $k(\tau) = \frac{e^{2i\tau}-2i\tau-1}{-2\tau^2}$

(d) $k(\tau) = \frac{e^{2i\tau}(i\tau-1)+2e^{i\tau}-i\tau-1}{-\tau^2}$

(e) $k(\tau) = (3/8)\left(\frac{e^{2i\tau}(-4\tau^2-4i\tau+2}{-i\tau^3}\right)$

6.

$$\begin{aligned} f_X(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\tau}e^{|\tau|}d\tau = \frac{1}{2\pi}\int_0^{\infty} e^{-i\tau}e^{-\tau}d\tau + \frac{1}{2\pi}\int_{-\infty}^{0} e^{-i\tau}e^{\tau}d\tau \\ &= \frac{1}{2\pi}\int_0^{\infty}\Bigl(e^{-i\tau x}+e^{i\tau x}\Bigr)e^{-\tau}d\tau = \frac{1}{\pi}\int_0^{\infty}\cos(\tau x)e^{-\tau}d\tau\ . \end{aligned}$$

Now to calculate this last integral:

$$\begin{aligned} \frac{1}{\pi}\int_0^{\infty}\cos(\tau x)e^{-\tau}d\tau &= \frac{1}{\pi}\Bigl[-e^{-\tau}\cos(\tau x)\Big|_0^{\infty} - x\int_0^{\infty} e^{-\tau}\sin(\tau x)d\tau\Bigr] \\ &= \frac{1}{\pi}\Bigl[1 - x(-e^{-\tau}\sin(\tau x))\Big|_0^{\infty} + x^2\int_0^{\infty} e^{-\tau}\cos(\tau x)\Bigr] \\ &= \frac{1}{\pi}\Bigl[1 - x^2\int_0^{\infty} e^{-\tau}\cos(\tau x)d\tau\Bigr]\ . \end{aligned}$$

Solving this for the integral we obtain:

$$\frac{1}{\pi}\int_0^{\infty}\cos(\tau x)e^{-\tau}d\tau = \frac{1}{\pi(1+x^2)}\ .$$

Thus,

$$f_X(x) = \frac{1}{\pi}\int_0^{\infty}\cos(\tau x)e^{-\tau}d\tau = \frac{1}{\pi(1+x^2)}\ .$$

7. (a) $g(-t) = \dfrac{1-e^{-t}}{t}$

(b) $e^t g(t) = \dfrac{e^{2t}-e^t}{t}$

(c) $g(et) = \dfrac{e^{3t}-1}{3t}$

(d) $e^b g(at) = \dfrac{e^b(e^{at}-1)}{at}$

8. (a) $g(t) = \dfrac{e^{at}-e^{bt}}{t(b-a)}$.

(b) $g(t) = \Bigl(\dfrac{e^{at}-e^{bt}}{t(b-a)}\Bigr)^2$.

(c) $g(t) = \Bigl(\dfrac{e^{at}-e^{bt}}{t(b-a)}\Bigr)^n$.

(d) $g(t) = \left(\frac{e^{at/n} - e^{bt/n}}{t(b-a)}\right)^n$.

(e) $g(t) = e^{-\sqrt{n}\mu/\sigma}\left(\frac{e^{at/\sqrt{n}\sigma} - e^{bt/\sqrt{n}\sigma}}{t(b-a)}\right)^n$.

9. (a) $g(t) = e^{t^2+t}$

(b) $\Big(g(t)\Big)^2$

(c) $\Big(g(t)\Big)^n$

(d) $\Big(g(t/n)\Big)^n$

(e) $e^{t^2/2}$

10. (a) $m = 0, \qquad \sigma^2 = 2$.

(b) $g_{X_1}(t) = \frac{1}{1-t^2}, \qquad g_{S_n}(t) = \left(\frac{1}{1-t^2}\right)^n, \quad t < 1$,

$$g_{A_n}(t) = \left(\frac{1}{1-(\frac{t}{n})^2}\right)^n \qquad g_{S_n^*}(t) = \left(\frac{1}{1-\frac{t^2}{2n}}\right)^n .$$

(c) $g_{S_n^*}(t) \to e^{-\frac{t^2}{2}}$ as $n \to \infty$.

(d) $g_{A_n}(t) \to 1$ as $n \to \infty$.

Chapter 11

Markov Chains

11.1 Introduction

1. $\mathbf{w}(1) = (.5,\ .25,\ .25)$

$\mathbf{w}(2) = (.4375,\ .1875,\ .375)$

$\mathbf{w}(3) = (.40625,\ .203125,\ .390625)$

2. $\mathbf{P} = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad \mathbf{P}^2 = \begin{pmatrix} 1 & 0 \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix}, \quad \mathbf{P}^3 = \begin{pmatrix} 1 & 0 \\ \frac{7}{8} & \frac{1}{8} \end{pmatrix}.$

$\mathbf{P}^n = \begin{pmatrix} 1 & 0 \\ \frac{2^n-1}{2^n} & \frac{1}{2^n} \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}.$

Whatever the President's decision, in the long run each person will be told that he or she is going to run.

3. $\mathbf{P}^n = \mathbf{P}$ for all n.

4. .7.

5. 1

6. $\mathbf{w}^{(1)} = \mathbf{w}^{(2)} = \mathbf{w}^{(3)} = \mathbf{w}^{(n)} = (.25, .5, .25)$.

7. (a) $\mathbf{P}^n = \mathbf{P}$

(b) $\mathbf{P}^n = \begin{cases} \mathbf{P}, & \text{if } n \text{ is odd,} \\ \mathbf{I}, & \text{if } n \text{ is even.} \end{cases}$

8. $\mathbf{P} = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix} \end{array}.$

9. $p^2+q^2,\ q^2,\ \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & p & q \\ 1 & q & p \end{array}$

$$\begin{array}{c} \\ 0 \\ 1 \end{array}\begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} p & q \\ q & p \end{pmatrix} \end{array}$$

11. .375

12. (a)
$$\mathbf{P} = \begin{array}{c} \\ P \\ SL \\ UL \\ NS \end{array}\begin{array}{c} \begin{array}{cccc} P & SL & UL & NS \end{array} \\ \begin{pmatrix} .64 & .08 & .08 & .2 \\ .16 & .48 & .16 & .2 \\ .2 & .2 & .4 & .2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}.$$

(b) .24.

19. (a) 5/6.

(b) The ‘transition matrix’ is

$$\mathbf{P} = \begin{array}{c} \\ H \\ T \end{array}\begin{array}{c} \begin{array}{cc} H & T \end{array} \\ \begin{pmatrix} 5/6 & 1/6 \\ 1/2 & 1/2 \end{pmatrix} \end{array}.$$

(c) 9/10.

(d) No. If it were a Markov chain, then the answer to (c) would be the same as the answer to (a).

11.2 Absorbing Markov Chains

1. $a = 0$ or $b = 0$

2. H is the absorbing state. Y and D are transient states. It is possible to go from each of these states to the absorbing state, in fact in one step.

3. Examples 11.10 and 11.11

4.
$$N = \begin{array}{c} \\ GG \\ gg \end{array}\begin{array}{c} \begin{array}{cc} Gg & gg \end{array} \\ \begin{pmatrix} 2 & 0 \\ 2 & 1 \end{pmatrix} \end{array}.$$

5. The transition matrix in canonical form is

$$P = \begin{array}{c} \\ GG,Gg \\ GG,gg \\ Gg,Gg \\ Gg,gg \\ GG,GG \\ gg,gg \end{array} \begin{array}{c} \begin{array}{cccccc} GG,Gg & GG,gg & Gg,Gg & Gg,gg & GG,GG & gg,gg \end{array} \\ \begin{pmatrix} 1/2 & 0 & 1/4 & 0 & 1/4 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1/4 & 1/8 & 1/4 & 1/4 & 1/16 & 1/16 \\ 0 & 0 & 1/4 & 1/2 & 0 & 1/4 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}.$$

Thus

$$Q = \begin{array}{c} \\ GG,Gg \\ GG,gg \\ Gg,Gg \\ Gg,gg \\ . \end{array} \begin{array}{c} \begin{array}{cccc} GG,Gg & GG,gg & Gg,Gg & Gg,gg \end{array} \\ \begin{pmatrix} 1/2 & 0 & 1/4 & 0 \\ 0 & 0 & 1 & 0 \\ 1/4 & 1/8 & 1/4 & 1/4 \\ 0 & 0 & 1/4 & 1/2 \\ & & & \end{pmatrix} \end{array},$$

and

$$N = (I - Q)^{-1} = \begin{array}{c} \\ GG,Gg \\ GG,gg \\ Gg,Gg \\ Gg,gg \end{array} \begin{array}{c} \begin{array}{cccc} GG,Gg & GG,gg & Gg,Gg & Gg,gg \end{array} \\ \begin{pmatrix} 8/3 & 1/6 & 4/3 & 2/3 \\ 4/3 & 4/3 & 8/3 & 4/3 \\ 4/3 & 1/3 & 8/3 & 4/3 \\ 2/3 & 1/6 & 4/3 & 8/3 \end{pmatrix} \end{array}.$$

From this we obtain

$$t = Nc = \begin{array}{c} GG,Gg \\ GG,gg \\ Gg,Gg \\ Gg,gg \end{array} \begin{pmatrix} 29/6 \\ 20/3 \\ 17/3 \\ 29/6 \end{pmatrix},$$

and

$$\mathbf{B} = \mathbf{NR} = \begin{array}{c} \\ GG,Gg \\ GG,gg \\ Gg,Gg \\ Gg,gg \end{array} \begin{array}{c} \begin{array}{cc} GG,GG & gg,gg \end{array} \\ \begin{pmatrix} 3/4 & 1/4 \\ 1/2 & 1/2 \\ 1/2 & 1/2 \\ 1/4 & 3/4 \end{pmatrix} \end{array}.$$

6. The canonical form of the transition matrix is

$$\mathbf{P} = \begin{array}{c} \\ N \\ S \\ R \end{array} \begin{array}{c} \begin{array}{ccc} N & S & R \end{array} \\ \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/4 & 1/2 & 1/4 \\ 0 & 0 & 1 \end{pmatrix} \end{array},$$

$$\mathbf{N} = \begin{array}{c} \\ N \\ S \end{array} \!\!\begin{array}{c} \begin{array}{cc} N & S \end{array} \\ \begin{pmatrix} 4/3 & 4/3 \\ 2/3 & 8/3 \end{pmatrix} \end{array},$$

$$\mathbf{t} = \mathbf{Nc} = \begin{array}{c} N \\ S \end{array} \begin{pmatrix} 8/3 \\ 10/3 \end{pmatrix},$$

$$\mathbf{B} = \mathbf{NR} = \begin{array}{c} N \\ S \end{array} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Here is a typical interpretation for an entry of **N**. If it is snowing today, the expected number of nice days before the first rainy day is 2/3. The entries of **t** give the expected number of days until the next rainy day. Starting with a nice day this is 8/3, and starting with a snowy day it is 10/3. The entries of **B** reflect the fact that we are certain to reach the absorbing state (rainy day) starting in either state N or state S.

7. $$\mathbf{N} = \begin{pmatrix} 2.5 & 3 & 1.5 \\ 2 & 4 & 2 \\ 1.5 & 3 & 2.5 \end{pmatrix}$$

$$\mathbf{Nc} = \begin{pmatrix} 7 \\ 8 \\ 7 \end{pmatrix}$$

$$\mathbf{B} = \begin{pmatrix} 5/8 & 3/8 \\ 1/2 & 1/2 \\ 3/8 & 5/8 \end{pmatrix}$$

8. The transition matrix in canonical form is

$$\mathbf{P} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 0 \\ 4 \end{array} \!\!\begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 0 & 4 \end{array} \\ \begin{pmatrix} 0 & 2/3 & 0 & 1/3 & 0 \\ 1/3 & 0 & 2/3 & 0 & 0 \\ 0 & 1/3 & 0 & 0 & 2/3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1\ cr \end{pmatrix} \end{array},$$

$$\mathbf{N} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \!\!\begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 7/5 & 6/5 & 4/5 \\ 3/5 & 9/5 & 6/5 \\ 1/5 & 3/5 & 7/5 \end{pmatrix} \end{array},$$

$$\mathbf{B} = \mathbf{NR} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \!\!\begin{array}{c} \begin{array}{cc} 0 & 4 \end{array} \\ \begin{pmatrix} 7/15 & 8/15 \\ 3/15 & 12/15 \\ 1/15 & 14/15 \end{pmatrix} \end{array},$$

$$\mathbf{t} = \mathbf{NC} = \begin{array}{c} 1 \\ 2 \\ 3 \end{array} \begin{pmatrix} 17/5 \\ 18/5 \\ 11/5 \end{pmatrix}.$$

9. 2.08

12.

$$\mathbf{P} = \begin{array}{c} \\ ABC \\ AC \\ BC \\ A \\ B \\ C \\ none \end{array} \begin{array}{c} \begin{array}{ccccccc} ABC & AC & BC & A & B & C & none \end{array} \\ \begin{pmatrix} 5/18 & 5/18 & 4/18 & 0 & 0 & 4/18 & 0 \\ 0 & 5/12 & 0 & 5/2 & 0 & 1/12 & 1/12 \\ 0 & 0 & 10/18 & 0 & 5/18 & 2/18 & 1/18 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

$$\mathbf{N} = \begin{pmatrix} 1.385 & .659 & .692 \\ 0 & 1.714 & 0 \\ 0 & 0 & 2.25 \end{pmatrix}$$

$$\mathbf{Nc} = \begin{pmatrix} 2.736 \\ 1.714 \\ 2.25 \end{pmatrix}$$

$$\mathbf{B} = \begin{array}{c} \\ ABC \\ AC \\ BC \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & none \end{array} \\ \begin{pmatrix} .275 & .192 & .440 & .093 \\ .714 & 0 & .143 & .143 \\ 0 & .625 & .25 & .125 \end{pmatrix} \end{array}$$

13. Using timid play, Smith's fortune is a Markov chain with transition matrix

$$\mathbf{P} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 0 \\ 8 \end{array} \begin{array}{c} \begin{array}{ccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 0 & 8 \end{array} \\ \begin{pmatrix} 0 & .4 & 0 & 0 & 0 & 0 & 0 & .6 & 0 \\ .6 & 0 & .4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .6 & 0 & .4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & .6 & 0 & .4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & .6 & 0 & .4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & .6 & 0 & .4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & .6 & 0 & 0 & .4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}.$$

For this matrix we have

$$\mathbf{B} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 8 \end{array} \\ \begin{pmatrix} .98 & .02 \\ .95 & .05 \\ .9 & .1 \\ .84 & .16 \\ .73 & .27 \\ .58 & .42 \\ .35 & .65 \end{pmatrix} \end{array} .$$

For bold strategy, Smith's fortune is governed instead by the transition matrix

$$\mathbf{P} = \begin{array}{c} \\ 3 \\ 4 \\ 6 \\ 0 \\ 8 \end{array} \begin{array}{c} \begin{array}{ccccc} 3 & 4 & 6 & 0 & 8 \end{array} \\ \begin{pmatrix} 0 & 0 & .4 & .6 & 0 \\ 0 & 0 & 0 & .6 & .4 \\ 0 & .6 & 0 & 0 & .4 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array} ,$$

with

$$\mathbf{B} = \begin{array}{c} \\ 3 \\ 4 \\ 6 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 8 \end{array} \\ \begin{pmatrix} .744 & .256 \\ .6 & .4 \\ .36 & .64 \end{pmatrix} \end{array} .$$

From this we see that the bold strategy gives him a probability .256 of getting out of jail while the timid strategy gives him a smaller probability .1. Be bold!

14. It is the same.

15. (a)

$$\mathbf{P} = \begin{array}{c} \\ 3 \\ 4 \\ 5 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccccc} 3 & 4 & 5 & 1 & 2 \end{array} \\ \begin{pmatrix} 0 & 2/3 & 0 & 1/3 & 0 \\ 1/3 & 0 & 2/3 & 0 & 0 \\ 0 & 2/3 & 0 & 0 & 1/3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array} .$$

(b)

$$\mathbf{N} = \begin{array}{c} \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccc} 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 5/3 & 2 & 4/3 \\ 1 & 3 & 2 \\ 2/3 & 2 & 7/3 \end{pmatrix} \end{array} ,$$

$$\mathbf{t} = \begin{array}{c} 3 \\ 4 \\ 5 \end{array} \begin{pmatrix} 5 \\ 6 \\ 5 \end{pmatrix} ,$$

$$\mathbf{B} = \begin{array}{c} \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{cc} 1 & 2 \end{array} \\ \begin{pmatrix} 5/9 & 4/9 \\ 1/3 & 2/3 \\ 2/9 & 7/9 \end{pmatrix} \end{array} .$$

(c) Thus when the score is deuce (state 4), the expected number of points to be played is 6, and the probability that B wins (ends in state 2) is 2/3.

16.

$$\begin{array}{c} \\ g, GG \\ G, Gg \\ g, Gg \\ G, gg \\ G, GG \\ g, gg \end{array} \begin{array}{c} \begin{array}{cccccc} g, GG & G, Gg & g, Gg & G, gg & G, GG & g, gg \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ .25 & .25 & .25 & 0 & .25 & 0 \\ 0 & .25 & .25 & .25 & 0 & .25 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array} ,$$

$$\mathbf{N} = \begin{pmatrix} 1.667 & 2.667 & 1.333 & .333 \\ .667 & 2.667 & 1.333 & .333 \\ .333 & 1.333 & 2.667 & .667 \\ .333 & 1.333 & 2.667 & 1.667 \end{pmatrix} ,$$

$$\mathbf{Nc} = \begin{pmatrix} 6 \\ 5 \\ 5 \\ 6 \\ , \end{pmatrix}$$

$$\mathbf{B} = \begin{pmatrix} .667 & .333 \\ .667 & .333 \\ .333 & .667 \\ .333 & .667 \end{pmatrix} .$$

17. For the color-blindness example, we have

$$\mathbf{B} = \begin{array}{c} \\ g, GG \\ G, Gg \\ g, Gg \\ G, gg \end{array} \begin{array}{c} \begin{array}{cc} G, GG & g, gg \end{array} \\ \begin{pmatrix} 2/3 & 1/3 \\ 2/3 & 1/3 \\ 1/3 & 2/3 \\ 1/3 & 2/3 \end{pmatrix} \end{array} ,$$

and for Example 9 of Section 11.1, we have

$$\mathbf{B} = \begin{array}{c} \\ GG, Gg \\ GG, gg \\ Gg, Gg \\ Gg, gg \end{array} \begin{array}{c} \begin{array}{cc} GG, GG & gg, gg \end{array} \\ \begin{pmatrix} 3/4 & 1/4 \\ 1/2 & 1/2 \\ 1/2 & 1/2 \\ 1/4 & 3/4 \end{pmatrix} \end{array} .$$

In each case the probability of ending up in a state with all G's is proportional to the number of G's in the starting state. The transition matrix for Example 9 is

$$\mathbf{P} = \begin{array}{c} \\ GG,GG \\ GG,Gg \\ GG,gg \\ Gg,Gg \\ Gg,gg \\ gg,gg \end{array} \begin{array}{c} \begin{array}{cccccc} GG,GG & GG,Gg & GG,gg & Gg,Gg & Gg,gg & gg,gg \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1/4 & 1/2 & 0 & 1/4 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1/16 & 1/4 & 1/8 & 1/4 & 1/4 & 1/16 \\ 0 & 0 & 0 & 1/4 & 1/2 & 1/4 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array} .$$

Imagine a game in which your fortune is the number of G's in the state that you are in. This is a fair game. For example, when you are in state Gg,gg your fortune is 1. On the next step it becomes 2 with probability 1/4, 1 with probability 1/2, and 0 with probability 1/4. Thus, your expected fortune after the next step is equal to 1, which is equal to your current fortune. You can check that the same is true no matter what state you are in. Thus if you start in state Gg,gg, your expected final fortune will be 1. But this means that your final fortune must also have expected value 1. Since your final fortune is either 4 if you end in GG,GG or 0 if you end in gg,gg, we see that the probability of your ending in GG,GG must be 1/4.

18. (a)
$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ F \\ G \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & F & G \end{array} \\ \begin{pmatrix} r & p & 0 & q & 0 \\ 0 & r & p & q & 0 \\ 0 & 0 & r & q & p \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

(b) Expected time in second year = 1.09.

Expected time in med school = 3.3 years.

(c) Probability of an incoming student graduating = .67.

19. (a)
$$\mathbf{P} = \begin{array}{c} \\ 1 \\ 2 \\ 0 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 0 & 3 \end{array} \\ \begin{pmatrix} 0 & 2/3 & 1/3 & 0 \\ 2/3 & 0 & 0 & 1/3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array} .$$

(b)
$$\mathbf{N} = \begin{array}{c} \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{cc} 1 & 2 \end{array} \\ \begin{pmatrix} 9/5 & 6/5 \\ 6/5 & 9/5 \end{pmatrix} \end{array} ,$$

$$\mathbf{B} = \begin{array}{c c} & \begin{array}{cc} 0 & 3 \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{pmatrix} 3/5 & 2/5 \\ 2/5 & 3/5 \end{pmatrix} \end{array},$$

$$\mathbf{t} = \begin{array}{c} 1 \\ 2 \end{array} \begin{pmatrix} 3 \\ 3 \end{pmatrix}.$$

(c) The game will last on the average 3 moves.

(d) If Mary deals, the probability that John wins the game is 3/5.

20. Consider the Markov chain with state i (for $1 \leq i < k$) the length of the current run, and k an absorbing state. Then when in state $i < k$, the chain goes to $i + 1$ with probability $1/m$ or to 1 with probability $(m - 1)/m$. Thus, starting in state 1, in order to get to state $j + 1$ the chain must be in state j and then move to $j + 1$. This means that

$$N_{1,j+1} = N_{1,j}(1/m)\ ,$$

or

$$N_{1,j} = mN_{1,j+1}\ .$$

This will be true also for $j + 1 = k$ if we interpret $N_{1,k}$ as the number of times that the chain enters the state k, namely, 1. Thus, starting with $N_{1,k} = 1$ and working backwards, we see that $N_{1,j} = m^{k-j}$ for $j = 1, \cdots, k$. Therefore, the expected number of experiments until a run of k occurs is

$$1 + m + m^2 + \cdots + m^{k-1} = \frac{m^k - 1}{m - 1}\ .$$

(The initial 1 is to start the process off.) Putting $m = 10$ and $k = 9$ we see that the expected number of digits in the decimal expansion of π until the first run of length 7 would be about 111 million if the expansion were random. Thus we should not be surprised to find such a run in the first 100,000,000 digits of π and indeed there are runs of length 9 among these digits.

21. The problem should assume that a fraction

$$q_i = 1 - \sum_j q_{ij} > 0$$

of the pollution goes into the atmosphere and escapes.

(a) We note that $\mathbf{u}$ gives the amount of pollution in each city from today's emission, $\mathbf{uQ}$ the amount that comes from yesterday's emission, $\mathbf{uQ}^2$ from two days ago, etc. Thus

$$\mathbf{w}^n = \mathbf{u} + \mathbf{uQ} + \cdots \mathbf{uQ}^{n-1}\ .$$

(b) Form a Markov chain with $\mathbf{Q}$-matrix $\mathbf{Q}$ and with one absorbing state to which the process moves with probability q_i when in state i. Then

$$\mathbf{I} + \mathbf{Q} + \mathbf{Q}^2 + \cdots + \mathbf{Q}^{n-1} \to \mathbf{N} \ ,$$

so

$$\mathbf{w}^{(n)} \to \mathbf{w} = \mathbf{uN} \ .$$

(c) If we are given $\mathbf{w}$ as a goal, then we can achieve this by solving $\mathbf{w} = \mathbf{Nu}$ for $\mathbf{u}$, obtaining

$$\mathbf{u} = \mathbf{w}(\mathbf{I} - \mathbf{Q}) \ .$$

22.

(a) The total amount of goods that the ith industry needs to produce \$1 worth of goods is

$$x_1 q_{1i} + x_2 q_{2i} + \cdots + x_n q_{ni} \ .$$

This is the i'th component of the vector $\mathbf{xQ}$.

(b) By part (a) the amounts the industries need to meet their internal demands is $\mathbf{xQ}$. Thus to meet both internal and external demands, the companies must produce amounts given by a vector $\mathbf{x}$ satifying the equation

$$\mathbf{x} = \mathbf{xQ} + \mathbf{d} \ .$$

(c) From Markov chain theory we can always solve the equation

$$\mathbf{x} = \mathbf{xQ} + \mathbf{d}$$

by writing it as

$$\mathbf{x}(\mathbf{I} - \mathbf{Q}) = \mathbf{d}$$

and then using the fact that $(\mathbf{I} - \mathbf{Q})\mathbf{N} = \mathbf{I}$ to obtain

$$\mathbf{x} = \mathbf{dN} \ .$$

(d) If the row sums of $\mathbf{Q}$ are all less than 1, this means that every industry makes a profit. A company can rely directly or indirectly on a profit-making company. If for any value of $n, q_{ij}^n > 0$, then i depends at least indirectly on j. Thus *depending upon* is equivalent in the Markov chain interpretation to *being able to reach.* Thus the demands can be met if every company is either profit-making or depends upon a profit-making industry.

(e) Since $\mathbf{x} = \mathbf{dN}$, we see that

$$\mathbf{xc} = \mathbf{dNc} = \mathbf{dt} \ .$$

24. When the walk is in state i, it goes to $i+1$ with probability p and $i-1$ with probability q. Condition (a) just equates the probability of winning in terms of the current state to the probability after the next step. Clearly, if our fortune is 0, then the probability of winning is 0, and if it is T, then the probability is 1. Here is an instructive way (not the simplest way) to see that the values of $\mathbf{w}$ are uniquely determined by (a), (b), and (c). Let $\mathbf{P}$ be the transition matrix for our absorbing chain. Then these equations state that

$$\mathbf{Pw} = \mathbf{w} .$$

That is, the column vector $\mathbf{w}$ is a fixed vector for $\mathbf{P}$. Consider the transition matrix for an arbitrary Markov chain in canonical form and assume that we have a vector $\mathbf{w}$ such that $\mathbf{w} = \mathbf{Pw}$. Multiplying through by $\mathbf{P}$, we see that $\mathbf{P}^2\mathbf{w} = \mathbf{w}$, and in general $\mathbf{P}^n\mathbf{w} = \mathbf{w}$. But

$$\mathbf{P}^n \to \begin{pmatrix} \mathbf{0} & \mathbf{B} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} .$$

Thus

$$\mathbf{w} = \begin{pmatrix} \mathbf{0} & \mathbf{B} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \mathbf{w} .$$

If we write

$$\mathbf{w} = \begin{pmatrix} \mathbf{w_T} \\ w_A \end{pmatrix} ,$$

where T is the set of transient states and A the set of absorbing states, then by the argument above we have

$$\mathbf{w} = \begin{pmatrix} \mathbf{w}_T \\ w_A \end{pmatrix} = \begin{pmatrix} \mathbf{Bw_A} \\ w_A \end{pmatrix} .$$

Thus for an absorbing Markov chain, a fixed column vector $\mathbf{w}$ is determined by its values on the absorbing states. Since in our example we know these values are (0,1), we know that $\mathbf{w}$ is completely determined. The solutions given clearly satisfy (b) and (c), and a direct calculation shows that they also satisfy (a).

26. Again, it is easy to check that the proposed solution $f(x) = x(n-x)$ satisfies conditions (a) and (b). The hard part is to prove that these equations have a unique solution. As in the case of Exercise 23, it is most instructive to consider this problem more generally. We have a special case of the following situation. Consider an absorbing Markov chain with transition matrix $\mathbf{P}$ in canonical form and with transient states T and absorbing states A. Let $\mathbf{f}$ and $\mathbf{g}$ be column vectors that satisfy the following system of equations

$$\begin{pmatrix} \mathbf{Q} & \mathbf{R} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{f}_A \\ \mathbf{0} \end{pmatrix} + \begin{pmatrix} \mathbf{g}_A \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{f}_A \\ \mathbf{0} \end{pmatrix} ,$$

where $\mathbf{g}_A$ is given and it is desired to determine $\mathbf{f}_A$. In our example, $\mathbf{g}_A$ has all components equal to 1. To solve for $\mathbf{f}_A$ we note that these equations are equivalent to

$$\mathbf{Q}\mathbf{f}_A + \mathbf{g}_A = \mathbf{f}_A \ ,$$

or

$$(\mathbf{I} - \mathbf{Q})\mathbf{f}_A = \mathbf{g}_A \ .$$

Solving for $\mathbf{f}_A$, we obtain

$$\mathbf{f}_A = \mathbf{N}\mathbf{g}_A \ .$$

Thus $\mathbf{f}_A$ is uniquely determined by $\mathbf{g}_A$.

27. Use the solution to Exercise 24 with $\mathbf{w} = \mathbf{f}$.

28. Using the program **Absorbing Chain** for the transition matrix corresponding to the pattern HTH, we find that

$$\mathbf{t} = \begin{array}{c} HT \\ H \\ \emptyset \end{array} \begin{pmatrix} 6 \\ 8 \\ 10 \end{pmatrix} .$$

Thus $E(T) = 10$. For the pattern HHH the transition matrix is

$$\mathbf{P} = \begin{array}{c} \\ HHH \\ HH \\ H \\ \emptyset \end{array} \begin{array}{c} \begin{array}{cccc} HHH & HH & H & \emptyset \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ .5 & 0 & 0 & .5 \\ 0 & .5 & 0 & .5 \\ 0 & 0 & .5 & .5 \end{pmatrix} \end{array} .$$

Solving for $\mathbf{t}$ for this matrix gives

$$\mathbf{t} = \begin{array}{c} HH \\ H \\ \emptyset \end{array} \begin{pmatrix} 8 \\ 12 \\ 14 \end{pmatrix} .$$

Thus for this pattern E(T) = 14.

29. For the chain with pattern HTH we have already verified that the conjecture is correct starting in HT. Assume that we start in H. Then the first player will win 8 with probability 1/4, so his expected winning is 2. Thus $E(T|H) = 10 - 2 = 8$, which is correct according to the results given in the solution to Exercise 28. The conjecture can be verified similarly for the chain HHH by comparing the results given by the conjecture with those given by the solution to Exercise 28.

30. T must be at least 3. Thus when you sum the terms

$$P(T > n) = 2P(T = n + 1) + 8P(T = n + 3),$$

the coefficients of the 2 and the 8 just add up to 1 since they are all possible probabilies for T. Let T be an integer-valued random variable. We write

$$\begin{aligned} E(T) \quad = \quad & P(T = 1) + P(T = 2) + P(T = 3) + \cdots \\ & \qquad\qquad\quad + P(T = 2) + P(T = 3) + \cdots \\ & \qquad\qquad\qquad\qquad\quad + P(T = 3) + \cdots \end{aligned}$$

If we add the terms by columns, we get the usual definition of expected value; if we add them by rows, we get the result that

$$E(T) = \sum_{n=0}^{\infty} P(T > n) \ .$$

That the order does not matter follows from the fact that all the terms in the sum are positive.

31. You can easily check that the proportion of G's in the state provides a harmonic function. Then by Exercise 27 the proportion at the starting state is equal to the expected value of the proportion in the final aborbing state. But the proportion of 1s in the absorbing state GG, GG is 1. In the other absorbing state gg, gg it is 0. Thus the expected final proportion is just the probability of ending up in state GG, GG. Therefore, the probability of ending up in GG, GG is the proportion of G genes in the starting state.(See Exercise 17.)

32. The states with all squares the same color are absorbing states. From any non-absorbing state it is possible to reach any absorbing state corresponding to a color still represented in the state. To see that the game is fair, consider the following argument. In order to decrease your fortune by 1 you must choose a red square and then choose a neighbor that is not red. With the same probability you could have chosen the neighbor and then the red square and your fortune would have been increased by 1. Since it is a fair game, if at any time a proportion p of the squares are red, for example, then p is also the probability that we end up with all red squares.

33. In each case Exercise 27 shows that

$$f(i) = b_{iN} f(N) + (1 - b_{iN}) f(0) \ .$$

Thus

$$b_{iN} = \frac{f(i) - f(0)}{f(N) - f(0)} \ .$$

Substituting the values of f in the two cases gives the desired results.

11.3 Ergodic Markov Chains

1. (a), (f)

2. (a) $\mathbf{P}^3 = \begin{pmatrix} .5 & .333 & .167 \\ .562 & .250 & .187 \\ .375 & .500 & .125 \end{pmatrix}$. Since $\mathbf{P}^3$ has no zero entries, $\mathbf{P}$ is regular.

(b) 1/6.

(c) $\mathbf{w} = (1/2,\ 1/3,\ 1/6)$.

3. (a) $a = 0$ or $b = 0$

(b) $a = b = 1$

(c) $(0 < a < 1$ and $0 < b < 1)$ or $(a = 1$ and $0 < b < 1)$ or $(0 < a < 1$ and $b = 1)$.

4. $\mathbf{w} = (b/(b + a),\ a/(b + a))$.

5. (a) $(2/3, 1/3)$

(b) $(1/2, 1/2)$

(c) $(2/7, 3/7, 2/7)$

6. Let $\mathbf{P} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Then $\mathbf{P}^{2n+1} = \mathbf{P}$ and $\mathbf{P}^{2n} = \mathbf{I}$. Thus $\mathbf{P}$ is not regular. However, the average $\mathbf{A}_n = \frac{1}{n}\left(\mathbf{P} + \mathbf{P}^2 + \cdots + \mathbf{P}^n\right)$ of these matrices converges to
$$\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}.$$
The vector $\mathbf{w} = (1/2, 1/2)$ is a fixed vector for $\mathbf{P}$. Its components represent the average number of times the process will be in each state in the long run.

7. The fixed vector is $(1, 0)$ and the entries of this vector are not strictly positive, as required for the fixed vector of an ergodic chain.

8. The vectors (1,0,0) and (0,0,1) are fixed vectors, and so is any vector of the form a(1,0,0) + $(1 - a)$(0,0,1) for $0 < a < 1$.
$$\mathbf{P}^n \to \begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 0 & 1 \end{pmatrix}.$$

9. Let

$$\mathbf{P} = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix},$$

with column sums equal to 1. Then

$$\begin{aligned} (1/3,\ 1/3,\ 1/3)P &= (1/3\sum_{j=1}^{3} p_{j1},\ 1/3\sum_{j=1}^{3} p_{j2},\ 1/3\sum_{j=1}^{3} p_{j3}) \\ &= (1/3, 1/3, 1/3)\ . \end{aligned}$$

The same argument shows that if $\mathbf{P}$ is an $n \times n$ transition matrix with columns that add to 1 then

$$\mathbf{w} = (1/n, \cdots, 1/n)$$

is a fixed probability vector. For an ergodic chain this means the the average number of times in each state is $1/n$.

10. In Example 11.10 of Section 11.1, the state GG is an absorbing state, and it is impossible to go from this state to either of the other two states.

11. In Example 11.11 of Section 11.1, the state (GG, GG) is absorbing, and the same reasoning as in the immediately preceding answer applies to show that this chain is not ergodic.

12. The chain is ergodic but not regular. Note that it is impossible to reach states 1 and 3 from state 0 in an even number of steps, and it is impossible to reach states 0, 2, and 4 from state 0 in an odd number of steps.

13. The fixed vector is $\mathbf{w} = (a/(b + a), b/(b + a))$. Thus in the long run a proportion $b/(b + a)$ of the people will be told that the President will run. The fact that this is independent of the starting state means it is independent of the decision that the President actually makes. (See Exercise 2 of Section 11.1)

14. The fixed vector is the common row of $\mathbf{P}$.

The chain is regular if and only if the entries of this vector are strictly positive.

15. It is clearly possible to go between any two states, so the chain is ergodic. From 0 it is possible to go to states 0, 2, and 4 only in an even number of steps, so the chain is not regular. For the general Erhrenfest Urn model the fixed vector must statisfy the following equations:

$$\frac{1}{n} w_1 = w_0\ ,$$

$$w_{j+1}\frac{j+1}{n} + w_{j-1}\frac{n-j+1}{n} = w_j, \qquad \text{if } 0 < j < n,$$

$$\frac{1}{n} w_{n-1} = w_n \ .$$

It is easy to check that the binomial coefficients satisfy these conditions.

16. $\mathbf{P}^2$ is strictly positive. The fixed vector is $\mathbf{w} = (1/4, 1/2, 1/4)$.

17. Consider the Markov chain whose state is the value of S_n mod(7), that is, the remainder when S_n is divided by 7. Then the transition matrix for this chain is

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \!\!\begin{array}{c} \begin{array}{ccccccc} \ \ 0 \ \ & \ \ 1 \ \ & \ \ 2 \ \ & \ \ 3 \ \ & \ \ 4 \ \ & \ \ 5 \ \ & \ \ 6 \ \ \end{array} \\ \begin{pmatrix} 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 0 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 0 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 0 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 \end{pmatrix} \end{array} .$$

Since the column sums of this matrix are 1, the fixed vector is

$$\mathbf{w} = (1/7, 1/7, 1/7, 1/7, 1/7, 1/7, 1/7) \ .$$

18. $2^r + 1$ (by this power there must have been a repetition of the pattern of positive numbers in the matrix so nothing new can occur). $N(3) = 5$. See Exercise 19.

19.

(a) For the general chain it is possible to go from any state i to any other state j in $r^2 - 2r + 2$ steps. We show how this can be done starting in state 1. To return to 1, circle $(1, 2, .., r-1, 1)$ $r - 2$ times ($r^2 - 3r + 2$ steps) and $(1, ..., r, 1)$ once (r steps). For $k = 1, ..., r - 1$ to reach state $k + 1$, circle $(1, 2, \ldots, r, 1)$ $r - k$ times ($r^2 - rk$ steps) then $(1, 2, \ldots, r-1, 1)$ $k - 2$ times ($rk - 2r - k + 2$ steps) and then move to $k + 1$ in k steps.You have taken $r^2 - 2r + 2$ steps in all. The argument is the same for any other starting state with everything translated the appropriate amount.

(b)

$$\mathbf{P} = \begin{pmatrix} 0 & * & 0 \\ * & 0 & * \\ * & 0 & 0 \end{pmatrix} , \ \mathbf{P}^2 = \begin{pmatrix} * & 0 & * \\ * & * & 0 \\ 0 & * & 0 \end{pmatrix} , \ \mathbf{P}^3 = \begin{pmatrix} * & * & 0 \\ * & * & * \\ * & 0 & * \end{pmatrix} ,$$

$$\mathbf{P}^4 = \begin{pmatrix} * & * & * \\ * & * & * \\ * & * & 0 \end{pmatrix} , \ \mathbf{P}^5 = \begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \end{pmatrix} .$$

20. The transition matrix is

$$
\begin{array}{c} \\ 0 \\ 1 \\ 2 \\ \vdots \\ n \end{array}
\begin{array}{c}
\begin{array}{cccccc} 0 & 1 & 2 & \dots & n-1 & n \end{array} \\
\begin{pmatrix}
1-p & p & 0 & \dots & 0 & 0 \\
r(1-p) & pr+(1-p)(1-r) & p(1-r) & \dots & 0 & 0 \\
0 & r(1-p) & pr+(1-p)(1-r) & \dots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & & \\
0 & 0 & \dots & \dots & r & 1-r
\end{pmatrix}
\end{array} .
$$

This transition matrix has a property called *reversibility* which will be discussed in Section 11.5. For such a chain the fixed vector $\mathbf{w}$ satisfies the condition

$$w_i p_{ij} = w_j p_{ji} \ .$$

When this condition is satisfied, it is easy to determine the fixed vector. For this example, reversibility in this chain means that

$$w_i p(1-r) = w_{i+1} r(1-p) \ , \quad 0 < i < n \ ,$$

or

$$\frac{w_{i+1}}{w_i} = \frac{p(1-r)}{r(1-p)} \ , \quad 0 < i < n \ .$$

Thus

$$w_i = c\rho^i, \qquad 0 < i < n,$$

where

$$\rho = \frac{p(1-r)}{r(1-p)} \ .$$

The values for w_0 and w_n are obtained from $w_0 p = w_1 r(1-p)$ and $w_n r = w_{n-1} p(1-r)$. The constant c is then chosen to make the probabilities add to 1. If the traffic intensity $s = p/r$ is greater than 1 then ρ is greater than 1, and if it is less than 1 then ρ is less than one. Thus when the traffic intensity is greater than 1 the fixed vector is large for large values of i, and when it is less than 1 the fixed vector is small for large values of n. This means that the queue size will build up when the traffic intensity is greater than or equal to 1. The case ρ eaual to 1 is a border-line case, and in this case the equilibrium vector is constant for $0 < i < n$.

22. (a)

$$
\begin{aligned}
P(S=j) &= P(\text{customer does not finish in } j-1 \text{ sec.,but does in } j\text{th sec. }) \\
&= (1-r)^{j-1} r \ ,
\end{aligned}
$$

$$
\begin{aligned}
P(T=j) &= P(\text{no arrival in } j-1 \text{ seconds,but arrival in } j\text{th second }) \\
&= (1-p)^{j-1} p \ .
\end{aligned}
$$

(b) S and T have geometric distributions, and so $E(S) = 1/r$ and $E(T) = 1/p$.

(c) The traffic intensity s greater than 1 means p is greater than r, or $E(S)$ is greater than $E(T)$. This means that the arrival rate is faster than the service rate, so the queue size builds up; s equal to 1 means p is equal to r or that $E(S)$ is equal to $E(T)$; s less than 1 means that p is less than r, or $E(S)$ is less than $E(T)$ and the queue size does not build up.

24. Fixed probability vector is (1/5, 4/5). Thus $\mathbf{wP} = \mathbf{w}$ implies

$$\frac{1}{5} \times .5 + \frac{4}{5}p = \frac{1}{5} \ ,$$

so $p = .125$.

25. To each Markov chain we can associate a *directed graph*, whose vertices are the states i of the chain, and whose edges are determined by the transition matrix: the graph has an edge from i to j if and only if $p_{ij} > 0$. Then to say that $\mathbf{P}$ is ergodic means that from any state i you can find a path following the arrows until you reach any state j. If you cut out all the loops in this path you will then have a path that never interesects itself, but still leads from i to j. This path will have at most $r - 1$ edges, since each edge leads to a different state and none leads to i. Following this path requires at most $r - 1$ steps.

26. Assume that $\mathbf{P}$ is ergodic. Let M be the maximum of the steps required to go between two states. Then it is possible to go from any state i to any state j in M steps. To see this, assume that it is possible to go from i to j in m steps with $m < M$. Then just stay in i for $M - m$ steps before starting on your route to M. Thus $\mathbf{P}$ is regular.

27. If $\mathbf{P}$ is ergodic it is possible to go between any two states. The same will be true for the chain with transition matrix $\frac{1}{2}(\mathbf{I}+\mathbf{P})$. But for this chain it is possible to remain in any state; therefore, by Exercise 26, this chain is regular.

28. Assume that $\mathbf{wP} = \mathbf{w}$. Then

$$\mathbf{w}(\mathbf{I} + \mathbf{P})/2 = (\mathbf{w} + \mathbf{w})/2 = \mathbf{w} \ .$$

Conversely, if

$$\mathbf{w}(\mathbf{I} + \mathbf{P})/2 = \mathbf{w}$$

then $\mathbf{wP}/2 = \mathbf{w}/2$, and $\mathbf{wP} = \mathbf{w}$.

29.

(b) Since $\mathbf{P}$ has rational transition probabilities, when you solve for the fixed vector you will get a vector $\mathbf{a}$ with rational components. We can multiply through by a sufficiently large integer to obtain a fixed vector $\mathbf{u}$ with integer components such that each component of $\mathbf{u}$ is an integer multiple of the corresponding component of $\mathbf{a}$. Let $\mathbf{a}^{(n)}$ be the vector resulting from the nth iteration. Let $\mathbf{b}^{(n)} = \mathbf{a}^{(n)}\mathbf{P}$. Then $\mathbf{a}^{(n+1)}$ is obtained by adding chips to $\mathbf{b}^{(n+1)}$. We want to prove that $\mathbf{a}^{(n+1)} \geq \mathbf{a}^{(n)}$. This is true for $n = 0$ by construction. Assume that it is true for n. Then multiplying the inequality by $\mathbf{P}$ gives that $\mathbf{b}^{(n+1)} \geq \mathbf{b}^{(n)}$. Consider the component $a_j^{(n+1)}$. This is obtained by adding chips to $b_j^{(n+1)}$ until we get a multiple of a_j. Since $b_j^{(n)} \leq b_j^{(n+1)}$, any multiple of a_j that could be obtained in this manner to define $a_j^{(n+1)}$ could also have been obtained to define $a_j^{(n)}$ by adding more chips if necessary. Since we take the smallest possible multiple a_j, we must have $a_j^{(n)} \leq a_j^{(n+1)}$. Thus the results after each iteration are monotone increasing. On the other hand, they are always less than or equal to $\mathbf{u}$. Since there are only a finite number of integers between components of $\mathbf{a}$ and $\mathbf{u}$, the iteration will have to stop after a finite number of steps.

30. Assume that the tape can hold at most n units.Then the transition matrix is

$$\mathbf{P} = \begin{array}{c c} & \begin{array}{ccccccc} 0 & 1 & 2 & 3 & \dots & n-1 & n \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \\ n \end{array} & \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & p & q & 0 & \dots & 0 & 0 \\ 0 & p^2 & \binom{2}{1}pq & q^2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ p^n & \binom{n}{1}p^{n-1}q & \binom{n}{2}p^{n-2}q^2 & \binom{n}{3}p^{n-3}q^3 & \dots & \binom{n}{n-1}pq^{n-1} & q^n \end{pmatrix} \end{array}.$$

(Note that we assume that the request has no effect when the tape is full). When the chain is in state i the expected value of the next position is $1 - ip$. No matter how small p is, for large enough i this will be negative. Thus, no matter how small p is, we see that if the tape is big enough, there will be a tendancy to free up space when a large number of spaces are occupied.

31. If the maximum of a set of numbers is an average of other elements of the set, then each of the elements with positive weight in this average must also be maximum. By assumption, $\mathbf{P}\ \mathbf{x} = \mathbf{x}$. This implies $\mathbf{P}^n\mathbf{x} = \mathbf{x}$ for all n. Assume that $x_i = M$, where M is the maximum value for the x_k's, and let j be any other state. Then there is an n such that $p_{ij}^n > 0$. The ith row of the equation $\mathbf{P}^n\mathbf{x} = \mathbf{x}$ presents x_i as an average of values of x_k with positive weight,one of which is x_j. Thus $x_j = M$, and $\mathbf{x}$ is constant.

32. If 0 is the average of non-negative numbers, then any element of this average occurring with positive weight must be 0. Assume that $\mathbf{w}$ is a fixed probability vector for an ergodic chain $\mathbf{P}$. Then $\mathbf{wP} = \mathbf{w}$ and

$\mathbf{w}\mathbf{P}^n = \mathbf{w}$. Assume that $w_i = 0$. For any j there is an n such that $p_{ij}^n > 0$. Thus the ith column of $\mathbf{w}\mathbf{P}^n$ presents w_i as an average of other values of w_k, with w_j occurring with positive weight. Hence $w_j = 0$.

11.4 Fundamental Limit Theorem for Regular Chains

1. $\begin{pmatrix} 1/3 \\ 1/3 \end{pmatrix}$

2. $\mathbf{P}^n \to \mathbf{c}\mathbf{w}$. Thus $\mathbf{P}^n\mathbf{y} \to \mathbf{c}\mathbf{w}\mathbf{y} = \mathbf{w}\mathbf{y}\mathbf{c}$, since $\mathbf{w}\mathbf{y}$ is a number.

3. For regular chains, only the constant vectors are fixed column vectors.

4. All vectors of the form $a\mathbf{y} + b\mathbf{z}$ for a and b constants are fixed vectors for the matrix $\mathbf{P}$ of Exercise 3. There are no other fixed vectors.

6. Let $\mathbf{y}^{(n)} = \mathbf{P}^n\mathbf{y}$. Then $y_i^{(n+1)} = \sum_j p_{ij} y_j^{(n)}$. Thus $y_i^{(n+1)} \leq \sum_j p_{ij} M_n = M_n$. This means that M_n is a upper bound for $\mathbf{y}_i^{(n+1)}$. Therefore, $M_{n+1} \leq M_n$. A similar argument shows that $m_{n+1} \geq m_n$. Hence $M_{n+1} - m_{n+1} \leq M_n - m_n$ for any $n \geq 1$. Thus if we can show that a subsequence of differences $M_n - m_n$ tends to 0, then the same will be true for the entire sequence of differences, since this sequence is monotone decreasing.

8. This game has the flavor of Doeblin's coupling idea. Once you and your friend happen to be looking at the same number, from that time on you will continue together. If this happens you will end up together and you can successfully state that she stopped the same place you did. To see how successful you will be, you will have to estimate the probability that the coupling takes place. It is easy to write a program using random permutations of 52 objects to simulate this process. If you do, you will find that about 75 percent of the time the coupling is successful.

11.5 Mean First Passage Time for Ergodic Chains

1.
$$\mathbf{Z} = \begin{pmatrix} 11/9 & -2/9 \\ -1/9 & 10/9 \end{pmatrix} .$$

and

$$\mathbf{M} = \begin{pmatrix} 0 & 2 \\ 4 & 0 \end{pmatrix} .$$

2. $$\mathbf{P} = \begin{array}{c} \\ S \\ A \\ W \end{array} \begin{array}{c} \begin{array}{ccc} S & A & W \end{array} \\ \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 1/3 & 2/3 \end{pmatrix} \end{array} .$$

$$\mathbf{Z} = \begin{array}{c} \\ S \\ A \\ W \end{array} \begin{array}{c} \begin{array}{ccc} S & A & W \end{array} \\ \begin{pmatrix} 1.333 & 0 & -1 \\ -.222 & .889 & -.333 \\ -.889 & -.444 & 1.667 \end{pmatrix} \end{array} , \mathbf{w} = (2/9,\ 4/9,\ 3/9).$$

(a) Mean recurrence time for $S = 1/w_1 = 4.5$.

(b) $m_{31} = 10$.

3. 2

4. Mean recurrence time for Yes is $1 + a/b$, and for No it is $1 + b/a$. For $a = .5$ and $b = .75$ this gives a mean recurrence time of 5/3 for Yes and 5/2 for No.

5. The fixed vector is $\mathbf{w} = (1/6,1/6,1/6,1/6,1/6,1/6)$, so the mean recurrence time is 6 for each state.

6. $$\mathbf{P} = \begin{array}{c} \\ R \\ N \\ S \end{array} \begin{array}{c} \begin{array}{ccc} R & N & S \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 0 & 1/2 \\ 1/4 & 1/4 & 1/2 \end{pmatrix} \end{array} ,$$

$$\mathbf{N} = \begin{array}{c} \\ N \\ S \end{array} \begin{array}{c} \begin{array}{cc} N & S \end{array} \\ \begin{pmatrix} 4/3 & 4/3 \\ 2/3 & 8/3 \end{pmatrix} \end{array} ,$$

$$\mathbf{t} = \mathbf{Nc} = \begin{array}{c} N \\ S \end{array} \begin{pmatrix} 8/3 \\ 10/3 \end{pmatrix} .$$

This tells us that if it is nice today, then the expected number of days until rain is 8/3. If it is snowy today, then the expected number of days until rain is 10/3.

7. (a)

$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array}\begin{array}{c} \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\ \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1/4 & 1/4 & 0 & 1/4 & 1/4 & 0 \\ 0 & 0 & 1/2 & 0 & 0 & 1/2 \\ 0 & 0 & 1/2 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 \end{pmatrix} \end{array}$$

(b) The rat alternates between the sets $\{1, 2, 4, 5\}$ and $\{3, 6\}$.

(c) $\mathbf{w} = (1/12, 1/12, 4/12, 2/12, 2/12, 2/12)$.

(d) $m_{1,5} = 7$

8. The mean recurrence time for state 0 is the average time between times that the server is busy.

9. (a) if n is odd, $\mathbf{P}$ is regular. If n is even, $\mathbf{P}$ is ergodic but not regular.

(b) $\mathbf{w} = (1/n, \cdots, 1/n)$.

(c) From the program **Ergodic** we obtain

$$\mathbf{M} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array}\begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 4 & 6 & 6 & 4 \\ 4 & 0 & 4 & 6 & 6 \\ 6 & 4 & 0 & 4 & 6 \\ 6 & 6 & 4 & 0 & 4 \\ 4 & 6 & 6 & 4 & 0 \end{pmatrix} \end{array}.$$

This is consistent with the conjecture that $m_{ij} = d(n - d)$, where d is the clockwise distance from i to j.

10. The transition matrix is:

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array}\begin{array}{c} \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array},$$

$$\mathbf{w} = (.25, .125, .125, .125, .125, .25),$$

and

$$\mathbf{M} = \begin{array}{c c} & \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} & \begin{pmatrix} 0 & 1 & 4 & 9 & 16 & 25 \\ 9 & 0 & 3 & 8 & 15 & 24 \\ 16 & 7 & 0 & 5 & 12 & 21 \\ 21 & 12 & 5 & 0 & 7 & 16 \\ 24 & 15 & 8 & 3 & 0 & 9 \\ 25 & 16 & 9 & 4 & 1 & 0 \end{pmatrix} \end{array} .$$

Note that the entries of the first passage matrix are all integers. They also form arithmetic progressions going down diagonals. General formulas for basic quantities for random walks can be found in *Finite Markov Chains* by John G. Kemeny and J. Laurie Snell (New York: Springer-Verlag, l976).

11. Yes, the reverse transition matrix is the same matrix.

12. $\mathbf{P} = \begin{pmatrix} 1/2 & 0 & 1/2 \\ 2/3 & 1/3 & 0 \\ 0 & 2/3 & 1/3 \end{pmatrix}$,

$\mathbf{w} = (4/10, 3/10, 3/10), \qquad w_1 p_{12} = 0 \neq w_2 p_{21}.$

13. Assume that $\mathbf{w}$ is a fixed vector for $\mathbf{P}$. Then

$$\sum_i w_i p^*_{ij} = \sum_i \frac{w_i w_j p_{ji}}{w_i} = \sum_i w_j p_{ji} = w_j ,$$

so $\mathbf{w}$ is a fixed vector for $\mathbf{P}$ *. Thus if $\mathbf{w}^*$ is the unique fixed vector for $\mathbf{P}$ * we must have $\mathbf{w} = \mathbf{w}^*$.

14. No. In the Land of Oz example we found the mean first-passage matrix to be

$$\mathbf{M} = \begin{array}{c c} & \begin{array}{ccc} R & N & S \end{array} \\ \begin{array}{c} R \\ N \\ S \end{array} & \begin{pmatrix} 0 & 4 & 10/3 \\ 8/3 & 0 & 8/3 \\ 10/3 & 4 & 0 \end{pmatrix} \end{array} .$$

Note, for example, that $m_{NR} = 8/3 \neq m_{RN} = 4$.

15. If $p_{ij} = p_{ji}$ then $\mathbf{P}$ has column sums 1. We have seen (Exercise 9 of Section 11.3) that in this case the fixed vector is a constant vector. Thus for any two states s_i and s_j, $w_i = w_j$ and $p_{ij} = p_{ji}$. Thus $w_i p_{ij} = w_j p_{ji}$, and the chain is reversible.

16. We first show that

$$(\mathbf{I} + \mathbf{P} + \cdots + \mathbf{P}^{n-1})(\mathbf{I} - \mathbf{P} + \mathbf{W}) = \mathbf{I} - \mathbf{P}^n + n\mathbf{W} .$$

Recall that $\mathbf{PW} = \mathbf{W}$ so also $\mathbf{P^k W} = \mathbf{W}$. Thus, just multiplying out the left side gives this equality. Now if we divide through by n and pass to the limit we have

$$\frac{\mathbf{I} + \mathbf{P} + \ldots + \mathbf{P}^{n-1}}{n}(\mathbf{I} - \mathbf{P} + \mathbf{W}) \to \mathbf{W}$$

Multiplying both sides on the right by $\mathbf{Z}$ and recalling that $\mathbf{WZ} = \mathbf{W}$ we see that

$$\frac{\mathbf{I} + \mathbf{P} + \ldots + \mathbf{P}^{n-1}}{n} \to \mathbf{W}$$

17. We know that $\mathbf{wZ} = \mathbf{w}$. We also know that $m_{ki} = (z_{ii} - z_{ki})/w_i$ and $w_i = 1/r_i$. Putting these in the relation

$$\bar{m}_i = \sum_k w_k m_{ki} + w_i r_i \ ,$$

we see that

$$\begin{aligned}
\bar{m}_i &= \sum_k w_k \frac{z_{ii} - z_{ki}}{w_i} + 1 \\
&= \frac{z_{ii}}{w_i} \sum_k w_k - \frac{1}{w_i} \sum_k w_k z_{ki} + 1 \\
&= \frac{z_{ii}}{w_i} - 1 + 1 = \frac{z_{ii}}{w_i} \ .
\end{aligned}$$

18. Form a Markov chain whose states are the possible outcomes of a roll. After 100 rolls we may assume that the chain is in equilibrium. We want to find the mean time in equilibrium to obtain snake eyes for the first time. For this chain m_{ki} is the same as r_i, since the starting state does not effect the time to reach another state for the first time. The fixed vector has all entries equal to 1/36, so $r_i = 36$. Using this fact, we obtain

$$\bar{m}_i = \sum_k w_k m_{ki} + w_i r_i = 35 + 1 = 36.$$

We see that the expected time to obtain snake eyes is 36, so the second argument is correct.

19. Recall that

$$m_{ij} = \sum_j \frac{z_{jj} - z_{ij}}{w_j}.$$

Multiplying through by w_j summing on j and, using the fact that $\mathbf{Z}$ has row sums 1, we obtain

$$m_{ij} = \sum_j z_{jj} - \sum_j z_{ij} = \sum_j z_{jj} - 1 = K,$$

which is independent of i.

20. Assume that you start in state a. Then the expected amount you win on the nth step is $\sum_j P^n_{a,j} f_j$. From this it follows that your expected winning on the first n steps can be represented by the column vector $\mathbf{g}^{(n)}$, with

$$\mathbf{g}^{(n)} = (\mathbf{I} + \mathbf{P} + \mathbf{P}^2 + \cdots + \mathbf{P}^n)\mathbf{f}.$$

But since $\mathbf{wf} = \mathbf{0}$ also $\mathbf{Wf} = \mathbf{0}$. Thus we have

$$\mathbf{g}^{(n)} = (\mathbf{I} + (\mathbf{P} - \mathbf{W}) + (\mathbf{P}^2 - \mathbf{W}) + \cdots + (\mathbf{P}^n - \mathbf{W}))\mathbf{f}$$

Letting $n \to \infty$ we obtain,

$$\mathbf{g}^{(n)} \to \mathbf{Zf}.$$

We used here that if $\mathbf{P}$ is the transition matrix for a regular chain then Z is equal to the infinite series:

$$\mathbf{Z} = \mathbf{I} + (\mathbf{P} - \mathbf{W}) + (\mathbf{P^2} - \mathbf{W}) + (\mathbf{P^3} - \mathbf{W}) \cdots .$$

21. The transition matrix is

$$\mathbf{P} = \begin{array}{c} \\ GO \\ A \\ B \\ C \end{array} \!\!\! \begin{array}{c} \begin{array}{cccc} GO & A & B & C \end{array} \\ \begin{pmatrix} 1/6 & 1/3 & 1/3 & 1/6 \\ 1/6 & 1/6 & 1/3 & 1/3 \\ 1/3 & 1/6 & 1/6 & 1/3 \\ 1/3 & 1/3 & 1/6 & 1/6 \end{pmatrix} \end{array} .$$

Since the column sums are 1, the fixed vector is

$$\mathbf{w} = (1/4,\ 1/4,\ 1/4,\ 1/4)\ .$$

From this we see that $\mathbf{wf} = \mathbf{0}$. From the result of Exercise 20 we see that your expected winning starting in GO is the first component of the vector $\mathbf{Zf}$ where

$$\mathbf{f} = \begin{pmatrix} 15 \\ -30 \\ -5 \\ 20 \end{pmatrix} .$$

Using the program **ergodic** we find that the long run expected winning starting in GO is 10.4.

22. $\mathbf{PW} = \mathbf{W}$ follows from the fact the columns of $\mathbf{W}$ are constant and the row sums of $\mathbf{P}$ are 1. Similarly $\mathbf{WW} = \mathbf{W}$ follows from the fact that $\mathbf{W}$ has row sums 1 and constant columns. Thus $\mathbf{W^2} = \mathbf{W}$. Multiplying both sides by $\mathbf{W}$ gives $\mathbf{W^3} = \mathbf{W^2} = \mathbf{W}$. Continuing in this way we obtain $\mathbf{W^k} = \mathbf{W}$.

23. Assume that the chain is started in state s_i. Let $X_j^{(n)}$ equal 1 if the chain is in state s_i on the nth step and 0 otherwise. Then

$$S_j^{(n)} = X_j^{(0)} + X_j^{(1)} + X_j^{(2)} + \dots X_j^{(n)}$$

and

$$E(X_j^{(n)}) = P_{ij}^n.$$

Thus

$$E(S_j^{(n)}) = \sum_{h=0}^{n} p_{ij}^{(h)}.$$

If now follows from Exercise 16 that

$$\lim_{n\to\infty} \frac{E(S_j^{(n)})}{n} = w_j.$$

24. He got it!

Chapter 12

Random Walks

12.1 Random Walks in Euclidean Space

1. Let p_n = probability that the gambler is ruined at play n. Then

$$
\begin{aligned}
p_n &= 0, \qquad \text{if } n \text{ is even}, \\
p_1 &= q, \\
p_n &= p(p_1 p_{n-2} + p_3 p_{n-4} + \cdots + p_{n-2} p_1), \qquad \text{if } n > 1 \text{ is odd}.
\end{aligned}
$$

Thus

$$h(z) = qz + pz\Big(h(x)\Big)^2 ,$$

so

$$h(z) = \frac{1 - \sqrt{1 - 4pqz^2}}{2pz} .$$

By Exercise 10 we have

$$h(1) = \begin{cases} q/p, & \text{if } q \leq p, \\ 1, & \text{if } q \geq p, \end{cases}$$

$$h'(1) = \begin{cases} 1/(q-p), & \text{if } q > p, \\ \infty, & \text{if } q = p. \end{cases}$$

This says that when $q > p$, the gambler must be ruined, and the expected number of plays before ruin is $1/(q-p)$. When $p > q$, the gambler has a probability q/p of being ruined. When $p = q$, the gambler must be ruined eventually, but the expected number of plays until ruin is not finite.

2.

$$\frac{d^n}{dx^n}\Big((1-x)^{1/2}\Big)\Big|_{x=0} = (-1)^n \frac{1}{2}\Big(\frac{1}{2} - 1\Big) \cdots \Big(\frac{1}{2} - n + 1\Big) .$$

Thus,

$$(1-x)^{1/2} \quad = \quad \sum_{n=0}^{\infty} (-1)^n \frac{\frac{1}{2}\Big(\frac{1}{2}-1\Big)\cdots\Big(\frac{1}{2}-n+1\Big)}{n!}(-x)^n$$

$$
\begin{aligned}
&= \sum_{n=0}^{\infty} \frac{\frac{1}{2}\left(\frac{1}{2}-1\right)\cdots\left(\frac{1}{2}-n+1\right)}{n!} x^n \\
&= \sum_{n=0}^{\infty} \binom{1/2}{n} x^n .
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
h(z) &= \frac{1-\sqrt{1-4pqz^2}}{2pz} = \frac{1-\sum_{n=0}^{\infty}\binom{1/2}{n}(-4pqz^2)^n}{2pz} \\
&= \sum_{n=1}^{\infty} \frac{1}{2p}(-1)^n \binom{1/2}{n} (4pq)^n z^{2n-1} \ ,
\end{aligned}
$$

and

$$
p_T(n) = \begin{cases} \frac{1}{2p}(-1)^k \binom{1/2}{k} (4pq)^k, & \text{if } n = 2k-1 = \text{odd}, \\ 0, & \text{if } n = \text{even}. \end{cases}
$$

3. (a) From the hint:

$$
h_k(z) = h_{U_1}(z) \cdots h_{U_k}(z) = \Big(h(z)\Big)^k \ .
$$

(b)

$$
h_k(1) = \Big(h(1)\Big)^k = \begin{cases} (q/p)^k & \text{if } q \le p, \\ 1 & \text{if } q \ge p. \end{cases}
$$

$$
h'(1) = \begin{cases} k/(q-p) & \text{if } q > p, \\ \infty & \text{if } q = p. \end{cases}
$$

Thus the gambler must be ruined if $q \ge p$. The expected number of plays in this case is $k/(q-p)$ if $q > p$ and ∞ if $q = p$. When $q < p$ he is ruined with probability $(q/p)^k$.