

# 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).
3. (b) If one simulates a sufficiently large number of rolls, one should be able to conclude that the gamblers were correct.
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.
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.
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.
15. In about 25 percent of the games the player will have a streak of five.

## 1.2 Discrete Probability Distributions

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

$$3. \text{ (b), (d)}$$

$$5. \text{ (a) } 1/2$$

$$\text{(b) } 1/4$$

$$\text{(c) } 3/8$$

$$\text{(d) } 7/8$$

$$7. 11/12$$

$$9. 3/4, 1$$

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

$$13. 11:4$$

15. Let the sample space be:

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

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

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

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

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

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

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

- (c) Probably, since  $6 \log 2 \approx 4.159$  and  $36 \log 2 \approx 24.953$ .

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.

21.  $7/2^{12}$

23. We have

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

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

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

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.

27.

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

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

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  $r$ ,  $P(r) \approx .15$ .

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

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

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

(c)

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

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

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

13.  $2 \log 2 - 1$ .

15. Yes.

## Chapter 3

# Combinatorics

### 3.1 Permutations

1. 24

3.  $2^{32}$

5. 9, 6.

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

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

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

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

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

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

21. They are the same.

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 second 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
3.  $\binom{9}{7} = 36$
5. .998, .965, .729
- 7.

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

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

9.  $n = 15, r = 7$
11. Eight pieces of each kind of pie.
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.

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

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

$$\binom{n}{a} \binom{n-a}{b} = \frac{n!}{a!b!(n-a-b)!}$$

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

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

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

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

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

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

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

$$1 - \sum_{k \geq 27} b(42, .75, k) = .042.$$

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  $p = m/n$ .

31. .999996.

33. By Stirling's formula,

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

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.

39.

$$\begin{aligned} b\left(2n, \frac{1}{2}, n\right) &= 2^{-2n} \frac{2n!}{n!n!} = \frac{2n(2n-1) \cdots 2 \cdot 1}{2n \cdot 2(n-1) \cdots 2 \cdot 2n \cdot 2(n-1) \cdots 2} \\ &= \frac{(2n-1)(2n-3) \cdots 1}{2n(2n-2) \cdots 2} . \end{aligned}$$

### 3.3 Card Shuffling

3. (a) 96.99%

(b) 55.16%



## Chapter 4

# Conditional Probability

### 4.1 Discrete Conditional Probability

3. (a)  $1/2$

(b)  $2/3$

(c)  $0$

(d)  $1/4$

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

(b) (1)

7. (a) We have

$$\begin{aligned}P(A \cap B) &= P(A \cap C) = P(B \cap C) = \frac{1}{4}, \\P(A)P(B) &= P(A)P(C) = P(B)P(C) = \frac{1}{4}, \\P(A \cap B \cap C) &= \frac{1}{4} \neq P(A)P(B)P(C) = \frac{1}{8}.\end{aligned}$$

(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.

9. (a)  $1/3$

(b)  $1/2$

13.  $1/2$

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

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

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  $A$ .

19. .273.

21. No.

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.

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  $k/n$ .

27.

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

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

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

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.

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

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

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

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.

37.

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

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

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

39. (a)  $1/9$

(b) 1/4

(c) No

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

43. .710.

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):

$$W, LLW, LLLLW, \dots$$

The probability that this happens is equal to

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

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.

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

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

so  $Y_1$  and  $Y_2$  are independent.

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

$$P(\text{both coins turn up heads using (b)}) = p_1p_2.$$

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.

51.

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

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 the card on which the registrar sees Mathematics 23 and Government 35 is Mary's.

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

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

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

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

$P(\text{player 1 has a flush and player 2 has a full house}) =$

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

Taking the ratio of these last two quantities gives:

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

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

57.

$$\begin{aligned} &P(B|A) \leq P(B) \text{ and } P(B|A) \geq P(A) \\ \Leftrightarrow &P(B \cap A) \leq P(A)P(B) \text{ and } P(B \cap A) \geq P(A)P(B) \\ \Leftrightarrow &P(A \cap B) = P(A)P(B). \end{aligned}$$

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  $\tilde{B}$ .

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

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

Then

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

which is a contradiction.

## 4.2 Continuous Conditional Probability

1. (a) 2/3

(b) 1/3

(c) 1/2

(d) 1/2

3. (a) .01

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

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

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

5. (a) 1

(b) 1

(c) 1/2

(d)  $\pi/8$

(e) 1/2

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.

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 uniformly 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.)

### 4.3 Paradoxes

1. 2/3

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

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

(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.



## Chapter 5

# Important Distributions and Densities

### 5.1 Important Distributions

1. (a), (c), (d)
3. Assume that  $X$  is uniformly distributed, and let the countable set of values be  $\{\omega_1, \omega_2, \dots\}$ . 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

$$\sum_{i=1}^{\infty} f(\omega_i) = 1 .$$

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

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

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

9. (a) 1000

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

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

13. .7408, .2222, .0370

17. 649741

19. The probability of at least one call in a given day with  $n$  hands of bridge can be estimated by  $1 - e^{-n \cdot (6.3 \times 10^{-12})}$ . 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.

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

(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

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

23.

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

Thus when  $j+1 \leq m$ , the probability is increasing, and when  $j+1 \geq m$  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)

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

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



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

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

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  $P(\text{at least one accident}) = 1 - e^{-.1} = .0952$ .

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.

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  $280 \cdot \frac{(.70)^x e^{-(.70)}}{x!}$ .

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.

33. Poisson with mean 3.

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

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

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

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.

## 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  $[0, 1]$ .

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

(b)  $F(x) = 2\sqrt{x}$ ,  $f(x) = \frac{1}{\sqrt{x}}$  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)  $F(y) = \begin{cases} \frac{y^2}{2}, & 0 \leq y \leq 1; \\ 1 - \frac{(2-y)^2}{2}, & 1 \leq y \leq 2, \end{cases} \quad f(y) = \begin{cases} y, & 0 \leq y \leq 1; \\ 2 - y & 1 \leq y \leq 2. \end{cases}$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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



## Chapter 6

# Expected Value and Variance

### 6.1 Expected Value of Discrete Random Variables

1.  $-1/9$
3.  $5'10.1''$
5.  $-1/19$
7. Since  $X$  and  $Y$  each take on only two values, we may choose  $a, b, c, d$  so that

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

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.

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  $-7/498 = -.0141414\dots$ . 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  $-.0136$ .

13. 45

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

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

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

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

19.

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

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.

27. (a) 4

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

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.

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

(b)  $\frac{N}{k} \cdot \left( (k + 1)(1 - (1 - p)^k) + (1 - p)^k \right) .$

(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  $k = 1/\sqrt{p}$ .

33. We begin by noting that

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

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

$$\begin{aligned} a_1 &= t_1 \\ a_2 &= t_1 + t_2 \\ a_3 &= t_1 + t_2 + t_3 \\ &\vdots \\ a_j &= t_1 + t_2 + \dots + t_j . \end{aligned}$$

The sequence  $a_1, a_2, \dots, a_j$  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  $t_1, t_2, \dots, t_j$ . 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  $a_1, a_2, \dots, a_j$ . On the other hand, in order to have  $t_1 + t_2 + \dots + t_j \leq n$  the values of  $a_1, a_2, \dots, a_j$  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

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

$$\begin{aligned} E(X) &= P(X = 1) + P(X = 2) + P(X = 3) \dots \\ &\quad + P(X = 2) + P(X = 3) \dots \\ &\quad + P(X = 3) \dots . \end{aligned}$$

If we sum this by rows we see that

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

Thus,

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

The limit of this last expression as  $n \rightarrow \infty$  is  $e = 2.718\dots$  .

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  $t_1 + t_2 + \dots + t_j$  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  $n$ ,

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



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.

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

$$\begin{aligned} E &= \sum_{n=0}^{\infty} (-1)^{n+1} \frac{2^n}{n} \cdot \frac{1}{2^n} \\ &= \sum_{n=0}^{\infty} (-1)^{n+1} \frac{1}{n} = \log 2 . \end{aligned}$$

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.

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)

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

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

Continuing, we find that

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

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

$$\begin{aligned} p_1 &= 1 - q, \\ p_2 &= q - q^2, \end{aligned}$$

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.

## 6.2 Variance of Discrete Random Variables

1.  $E(X) = 0$ ,  $V(X) = \frac{2}{3}$ ,  $\sigma = D(X) = \sqrt{\frac{2}{3}}$ .
3.  $E(X) = \frac{-1}{19}$ ,  $E(Y) = \frac{-1}{19}$ ,  $V(X) = 33.21$ ,  $V(Y) = .99$ .
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.  $V(X) = \frac{3}{4}$ ,  $D(X) = \frac{\sqrt{3}}{2}$ .
9.  $V(X) = \frac{20}{9}$ ,  $D(X) = \frac{2\sqrt{5}}{3}$ .
11.  $E(X) = (1 + 2 + \cdots + n)/n = (n + 1)/2$ .

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

13. Let  $X_1, \dots, X_n$  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  $W_n = \sum_{j=1}^n X_j$ . Therefore  $E(W_n) = \sum_{i=1}^n E(X_i) = 0$ , and  $V(W_n) = \sum_{i=1}^n V(X_i) = n$ .

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

Therefore,  $E(X_i X_j) = \frac{1}{n(n-1)}$ .

(c)

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

(d)

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

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

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

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.

23. If  $X$  and  $Y$  are independent, then

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

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  $\text{Cov}(X, Y) = 0$ . However, since

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

$X$  and  $Y$  are dependent.

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

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

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

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

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)  $\mu = 0, \sigma^2 = 3/5$
3.  $\mu = 40, \sigma^2 = 800$
5. (d)  $a = -3/2, b = 0, c = 1$
- (e)  $a = \frac{45}{48}, b = 0, c = \frac{3}{16}$
7.  $f(a) = E(X - a)^2 = \int (x - a)^2 f(x) dx$  . Thus

$$f'(a) = - \int 2(x - a) f(x) dx$$

$$\begin{aligned}
&= -2 \int x f(x) dx + 2a \int f(x) dx \\
&= -2\mu(X) + 2a .
\end{aligned}$$

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

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

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

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

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

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

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  $b = 50\sqrt{2}$ .

13. Integrating by parts, we have

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

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

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

Letting  $a$  tend to infinity, we have that

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

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)

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

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

(c)

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

19. (a) 0

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

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

(d) 0

21. We have

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

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

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

27. Let  $Z$  represent the payment. Then

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

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

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

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

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  $n$ .

9. Let  $p_1, \dots, p_6$  be the probabilities for one die and  $q_1, \dots, q_6$  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.

## 7.2 Sums of Continuous Random Variables

3. (a)

$$f_z(x) = \begin{cases} x^3/24, & \text{if } 0 \leq x \leq 2; \\ x - x^3/24 - 4/3, & \text{if } 2 \leq x \leq 4. \end{cases}$$

- (b)

$$f_z(x) = \begin{cases} (x^3 - 18x^2 + 108x - 216)/24, & \text{if } 6 \leq x \leq 8; \\ (-x^3 + 18x^2 - 84x + 40)/24, & \text{if } 8 \leq x \leq 10. \end{cases}$$

- (c)

$$f_z(x) = \begin{cases} x^2/8, & \text{if } 0 \leq x \leq 2; \\ 1/2 - (x - 2)^2/8, & \text{if } 2 \leq x \leq 4. \end{cases}$$

5. (a)

$$f_z(x) = \begin{cases} \frac{\lambda\mu}{\mu+\lambda} e^{\lambda x}, & x < 0; \\ \frac{\lambda\mu}{\mu+\lambda} e^{-\mu x}, & x \geq 0. \end{cases}$$

(b)

$$f_z(x) = \begin{cases} 1 - e^{-\lambda x}, & 0 < x < 1; \\ (e^\lambda - 1)e^{-\lambda x}, & x \geq 1. \end{cases}$$

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

$$f_x^2(x) = \frac{1}{\sigma\sqrt{2\pi}} \frac{1}{2\sqrt{x}} \exp(-x/2\sigma^2 - \mu^2/2\sigma^2) \left( \exp(\sqrt{x}\mu/\sigma^2) + \exp(-\sqrt{x}\mu/\sigma^2) \right) .$$

Replacing the last two exponentials by their series representation, we have

$$f_x^2(x) = e^{-\mu/2\sigma^2} \sum_{r=0}^{\infty} \left( \frac{\mu}{2\sigma^2} \right)^r \frac{1}{r!} G(1/2\sigma^2, r + 1/2, x) ,$$

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  $\mu_1, \sigma_1$  and  $\mu_2, \sigma_2$ . 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,

$$f_{Z_1^2} = e^{-c^2/2} \sum_{r=0}^{\infty} \left( \frac{c^2}{2} \right)^r \frac{1}{r!} G(1/2, r + 1/2, x) ,$$

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

$$f_{Z_1^2+Z_2^2} = f_{X_1^2+X_2^2} = e^{-c^2/2} \sum_{r=0}^{\infty} \left(\frac{c^2}{2}\right)^r \frac{1}{r!} G(1/2, r+1, x) .$$

(This derivation is adapted from that of C.R. Rao in his book *Advanced Statistical Methods in Biometric Research*, Wiley, 1952.)

9.  $P(X_{10} > 22) = .341$  by numerical integration. This also could be estimated by simulation.
11. 10 hours
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

$$f_{Z_n}(x) = \frac{1}{(n-1)!} \left(\log \frac{1}{x}\right)^{n-1} .$$

19. The support of  $X + Y$  is  $[a + c, b + d]$ .
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.

5.  $k = 10$

7.

$$\begin{aligned} p(1-p) &= \frac{1}{4} - \left(\frac{1}{4} - p + p^2\right) \\ &= \frac{1}{4} - \left(\frac{1}{2} - p\right)^2 \leq \frac{1}{4}. \end{aligned}$$

Thus,  $\max_{0 \leq p \leq 1} p(1-p) = \frac{1}{4}$ . From Exercise 6 we have that

$$P\left(\left|\frac{S_n}{n} - p\right| \geq \epsilon\right) \leq \frac{p(1-p)}{n\epsilon^2} \leq \frac{1}{4n\epsilon^2}.$$

9.

$$\begin{aligned} P(S_n \geq 11) &= P(S_n - E(S_n) \geq 11 - E(S_n)) \\ &= P(S_n - E(S_n) \geq 10) \\ &\leq \frac{V(S_n)}{10^2} = .01. \end{aligned}$$

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.
15. Take as  $\Omega$  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.)
17. For  $x \in [0,1]$ , let us toss a biased coin that comes up heads with probability  $x$ . Then

$$E\left(\frac{f(S_n)}{n}\right) \rightarrow f(x).$$

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

$$\sum_{k=0}^n f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k} \rightarrow f(x).$$

This shows that we can 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

8.2. LAW OF LARGE NUMBERS FOR CONTINUOUS RANDOM VARIABLES 47

- (b) 1
- (c) 100/243
- (d) 1/12

3.

$$f(x) = \begin{cases} 1 - x/10, & \text{if } 0 \leq x \leq 10; \\ 0 & \text{otherwise.} \end{cases}$$

$$g(x) = \frac{100}{3x^2} .$$

5. (a) 1, 1/4, 1/9

(b) 1 vs. .3173, .25 vs. .0455, .11 vs. .0027

7. (b) 1, 1, 100/243, 1/12

9. (a) 0

(b) 7/12

(c) 11/12

11. (a) 0

(b) 7/12

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

$$E(X) \geq \int_{x \geq a} xp(x)dx \geq aP(X \geq a) .$$





## 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
3. (a)  $P(\text{June passes}) \approx .985$   
(b)  $P(\text{April passes}) \approx .056$
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).
7. .322
9. (a) 0  
(b) 1 (Law of Large Numbers)  
(c) .977 (Central Limit Theorem)

(d) 1 (Law of Large Numbers)

$$13. \quad P(S_{1900} \geq 115) = P\left(S_{1900}^* \geq \frac{115 - 95}{\sqrt{1900 \cdot .05 \cdot .95}}\right) = P(S_{1900}^* \geq 2.105) = .0176.$$

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

3. (a) .5

(b) .9987

5. (a)  $P(S_{210} < 700) \approx .0757$ .

(b)  $P(S_{189} \geq 700) \approx .0528$

(c)

$$\begin{aligned} P(S_{179} < 700, S_{210} \geq 700) &= P(S_{179} < 700) - P(S_{179} < 700, S_{210} < 700) \\ &= P(S_{179} < 700) - P(S_{210} < 700) \\ &\approx .9993 - .0757 = .9236. \end{aligned}$$

7. (a) Expected value = 200, variance = 2

(b) .9973

$$9. \quad P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) = P\left(|S_n - n\mu| \geq n\epsilon\right) = P\left(\left|\frac{S_n - n\mu}{\sqrt{n\sigma^2}}\right| \geq \frac{n\epsilon}{\sqrt{n\sigma^2}}\right).$$

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.

9.3. CENTRAL LIMIT THEOREM FOR CONTINUOUS INDEPENDENT TRIALS 51

11. Her expected loss is 60 dollars. The probability that she lost no money is about .0013.
13.  $p = .0056$

### 9.3 Central Limit Theorem for Continuous Independent Trials

1.

$$\begin{aligned} E(X^*) &= \frac{1}{\sigma}(E(X) - \mu) = \frac{1}{\sigma}(\mu - \mu) = 0, \\ \sigma^2(X^*) &= E\left(\frac{X - \mu}{\sigma}\right)^2 = \frac{1}{\sigma^2}\sigma^2 = 1. \end{aligned}$$

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  $T_n^* = \frac{T_n}{\sqrt{n}} = \frac{S_n - n\mu}{\sigma\sqrt{n}} = S_n^*$ .

11. (a) .5
- (b) .148
- (c) .018
13. .0013



## 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)  $h(z) = \frac{2}{3-z}$

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)

$$g(t) = \frac{1}{4} + \frac{1}{2} \left( \sum_{k=0}^{\infty} \frac{t^k}{k!} \right) + \frac{1}{4} \left( \sum_{k=0}^{\infty} \frac{2^k}{k!} t^k \right)$$

$$= 1 + \sum_{k=1}^{\infty} \left( \frac{1}{2k!} + \frac{2^{k-2}}{k!} \right) t^k = 1 + \sum_{k=1}^{\infty} \frac{\mu_k}{k!} t^k.$$

Thus  $\mu_0 = 1$ , and  $\mu_k = \frac{1}{2} + 2^{k-2}$  for  $k \geq 1$ .

$$(d) \quad p_0 = \frac{1}{4}, \quad p_1 = \frac{1}{2}, \quad p_2 = \frac{1}{4}.$$

$$5. \quad (a) \quad \mu_1(p) = \mu_1(p') = 3, \quad \mu_2(p) = \mu_2(p') = 11$$

$$\mu_3(p) = 43, \quad \mu_3(p') = 47$$

$$\mu_4(p) = 171, \quad \mu_4(p') = 219$$

$$7. \quad (a) \quad g_{-X}(t) = g(-t)$$

$$(b) \quad g_{X+1}(t) = e^t g(t)$$

$$(c) \quad g_{3X}(t) = g(3t)$$

$$(d) \quad g_{aX+b} = e^{bt} g(at)$$

$$9. \quad (a) \quad h_X(z) = \sum_{j=1}^6 a_j z^j, \quad h_Y(z) = \sum_{j=1}^6 b_j z^j.$$

$$(b) \quad h_Z(z) = \left( \sum_{j=1}^6 a_j z^j \right) \left( \sum_{j=1}^6 b_j z^j \right).$$

(c) Assume that  $h_Z(z) = (z^2 + \dots + z^{12})/11$ . Then

$$\left( \sum_{j=1}^6 a_j z^{j-1} \right) \left( \sum_{j=1}^6 b_j z^{j-1} \right) = \frac{1 + z + \dots + z^{10}}{11} = \frac{z^{11} - 1}{11(z - 1)}.$$

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).

## 10.2 Branching Processes

1. (a)  $d = 1$
- (b)  $d = 1$
- (c)  $d = 1$
- (d)  $d = 1$
- (e)  $d = 1/2$
- (f)  $d \approx .203$
3. (a) 0
- (b) 276.26
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  $h(z) = f(f(z))$ .

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 number 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

$$N = 1 + \left( \sum_k kp_k \right) = 1 + mN .$$

From this it follows that  $N$  is finite if and only if  $m < 1$ , in which case

$$N = 1/(1 - m).$$

## 10.3 Generating Functions for Continuous Densities

1. (a)  $g(t) = \frac{1}{2t}(e^{2t} - 1)$

$$(b) \quad g(t) = \frac{e^{2t}(2t-1) + 1}{2t^2}$$

$$(c) \quad g(t) = \frac{e^{2t} - 2t - 1}{2t^2}$$

$$(d) \quad g(t) = \frac{e^{2t}(ty-1) + 2e^t - t - 1}{t^2}$$

$$(e) \quad (3/8) \left( \frac{e^{2t}(4t^2 - 4t + 2) - 2}{t^3} \right)$$

$$3. \quad (a) \quad g(t) = \frac{2}{2-t}$$

$$(b) \quad g(t) = \frac{4-3t}{2(1-t)(2-t)}$$

$$(c) \quad g(t) = \frac{4}{(2-t)^2} \quad (d) \quad g(t) = \left( \frac{\lambda}{\lambda+t} \right), \quad t < \lambda.$$

$$5. \quad (a) \quad k(\tau) = \frac{1}{2i\tau}(e^{2i\tau} - 1)$$

$$(b) \quad k(\tau) = \frac{e^{2i\tau}(2i\tau - 1) + 1}{-2\tau^2}$$

$$(c) \quad k(\tau) = \frac{e^{2i\tau} - 2i\tau - 1}{-2\tau^2}$$

$$(d) \quad k(\tau) = \frac{e^{2i\tau}(i\tau - 1) + 2e^{i\tau} - i\tau - 1}{-\tau^2}$$

$$(e) \quad k(\tau) = (3/8) \left( \frac{e^{2i\tau}(-4\tau^2 - 4i\tau + 2)}{-i\tau^3} \right)$$

$$7. \quad (a) \quad g(-t) = \frac{1 - e^{-t}}{t}$$

$$(b) \quad e^t g(t) = \frac{e^{2t} - e^t}{t}$$

$$(c) \quad g(et) = \frac{e^{3t} - 1}{3t}$$

$$(d) \quad e^b g(at) = \frac{e^b(e^{at} - 1)}{at}$$

$$9. \quad (a) \quad g(t) = e^{t^2+t}$$

$$(b) \quad (g(t))^2$$



(c)  $(g(t))^n$

(d)  $(g(t/n))^n$

(e)  $e^{t^2/2}$



# Chapter 11

## Markov Chains

### 11.1 Introduction

1.  $\mathbf{w}(1) = (.5, .25, .25)$

$\mathbf{w}(2) = (.4375, .1875, .375)$

$\mathbf{w}(3) = (.40625, .203125, .390625)$

3.  $\mathbf{P}^n = \mathbf{P}$  for all  $n$ .

5. 1

7. (a)  $\mathbf{P}^n = \mathbf{P}$

(b)  $\mathbf{P}^n = \begin{cases} \mathbf{P}, & \text{if } n \text{ is odd,} \\ \mathbf{I}, & \text{if } n \text{ is even.} \end{cases}$

9.  $p^2 + q^2, q^2, 1 \begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix}$

11. .375

19. (a) 5/6.

(b) The 'transition matrix' is

$$\mathbf{P} = \begin{matrix} & \begin{matrix} H & T \end{matrix} \\ \begin{matrix} H \\ T \end{matrix} & \begin{pmatrix} 5/6 & 1/6 \\ 1/2 & 1/2 \end{pmatrix} \end{matrix}.$$

(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  $b = 0$
3. Examples 11.10 and 11.11
5. The transition matrix in canonical form is

$$P = \begin{array}{c} \begin{array}{ccccc} & GG, Gg & GG, gg & Gg, Gg & Gg, gg & GG, GG & gg, gg \end{array} \\ \begin{array}{c} GG, Gg \\ GG, gg \\ Gg, Gg \\ Gg, gg \\ GG, GG \\ gg, gg \end{array} \end{array} \begin{pmatrix} 1/2 & 0 & 1/4 & 0 & 1/4 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1/4 & 1/8 & 1/4 & 1/4 & 1/16 & 1/16 \\ 0 & 0 & 1/4 & 1/2 & 0 & 1/4 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Thus

$$Q = \begin{array}{c} \begin{array}{cccc} & GG, Gg & GG, gg & Gg, Gg & Gg, gg \end{array} \\ \begin{array}{c} GG, Gg \\ GG, gg \\ Gg, Gg \\ Gg, gg \end{array} \end{array} \begin{pmatrix} 1/2 & 0 & 1/4 & 0 \\ 0 & 0 & 1 & 0 \\ 1/4 & 1/8 & 1/4 & 1/4 \\ 0 & 0 & 1/4 & 1/2 \end{pmatrix},$$

and

$$N = (I - Q)^{-1} = \begin{array}{c} \begin{array}{cccc} & GG, Gg & GG, gg & Gg, Gg & Gg, gg \end{array} \\ \begin{array}{c} GG, Gg \\ GG, gg \\ Gg, Gg \\ Gg, gg \end{array} \end{array} \begin{pmatrix} 8/3 & 1/6 & 4/3 & 2/3 \\ 4/3 & 4/3 & 8/3 & 4/3 \\ 4/3 & 1/3 & 8/3 & 4/3 \\ 2/3 & 1/6 & 4/3 & 8/3 \end{pmatrix}.$$

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

$$\mathbf{B} = \mathbf{NR} = \begin{matrix} & \begin{matrix} GG, GG & gg, gg \end{matrix} \\ \begin{matrix} GG, Gg \\ GG, gg \\ Gg, Gg \\ Gg, gg \end{matrix} & \begin{pmatrix} 3/4 & 1/4 \\ 1/2 & 1/2 \\ 1/2 & 1/2 \\ 1/4 & 3/4 \end{pmatrix} \end{matrix}.$$

$$7. \quad \mathbf{N} = \begin{pmatrix} 2.5 & 3 & 1.5 \\ 2 & 4 & 2 \\ 1.5 & 3 & 2.5 \end{pmatrix}$$

$$\mathbf{Nc} = \begin{pmatrix} 7 \\ 8 \\ 7 \end{pmatrix}$$

$$\mathbf{B} = \begin{pmatrix} 5/8 & 3/8 \\ 1/2 & 1/2 \\ 3/8 & 5/8 \end{pmatrix}$$

9. 2.08

13. Using timid play, Smith's fortune is a Markov chain with transition matrix

$$\mathbf{P} = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 0 & 8 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 0 \\ 8 \end{matrix} & \begin{pmatrix} 0 & .4 & 0 & 0 & 0 & 0 & 0 & .6 & 0 \\ .6 & 0 & .4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .6 & 0 & .4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & .6 & 0 & .4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & .6 & 0 & .4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & .6 & 0 & .4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & .6 & 0 & 0 & .4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}.$$

For this matrix we have

$$\mathbf{B} = \begin{matrix} & \begin{matrix} 0 & 8 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{matrix} & \begin{pmatrix} .98 & .02 \\ .95 & .05 \\ .9 & .1 \\ .84 & .16 \\ .73 & .27 \\ .58 & .42 \\ .35 & .65 \end{pmatrix} \end{matrix}.$$

For bold strategy, Smith's fortune is governed instead by the transition matrix

$$\mathbf{P} = \begin{matrix} & \begin{matrix} 3 & 4 & 6 & 0 & 8 \end{matrix} \\ \begin{matrix} 3 \\ 4 \\ 6 \\ 0 \\ 8 \end{matrix} & \begin{pmatrix} 0 & 0 & .4 & .6 & 0 \\ 0 & 0 & 0 & .6 & .4 \\ 0 & .6 & 0 & 0 & .4 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix},$$

with

$$\mathbf{B} = \begin{matrix} & \begin{matrix} 0 & 8 \end{matrix} \\ \begin{matrix} 3 \\ 4 \\ 6 \end{matrix} & \begin{pmatrix} .744 & .256 \\ .6 & .4 \\ .36 & .64 \end{pmatrix} \end{matrix}.$$

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!

15. (a)

$$\mathbf{P} = \begin{matrix} & \begin{matrix} 3 & 4 & 5 & 1 & 2 \end{matrix} \\ \begin{matrix} 3 \\ 4 \\ 5 \\ 1 \\ 2 \end{matrix} & \begin{pmatrix} 0 & 2/3 & 0 & 1/3 & 0 \\ 1/3 & 0 & 2/3 & 0 & 0 \\ 0 & 2/3 & 0 & 0 & 1/3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}.$$

(b)

$$\mathbf{N} = \begin{matrix} & \begin{matrix} 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 3 \\ 4 \\ 5 \end{matrix} & \begin{pmatrix} 5/3 & 2 & 4/3 \\ 1 & 3 & 2 \\ 2/3 & 2 & 7/3 \end{pmatrix} \end{matrix},$$

$$\mathbf{t} = \begin{matrix} & \begin{matrix} 3 \\ 4 \\ 5 \end{matrix} \\ \begin{matrix} 3 \\ 4 \\ 5 \end{matrix} & \begin{pmatrix} 5 \\ 6 \\ 5 \end{pmatrix} \end{matrix},$$

$$\mathbf{B} = \begin{matrix} & \begin{matrix} 1 & 2 \end{matrix} \\ \begin{matrix} 3 \\ 4 \\ 5 \end{matrix} & \begin{pmatrix} 5/9 & 4/9 \\ 1/3 & 2/3 \\ 2/9 & 7/9 \end{pmatrix} \end{matrix}.$$

(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  $2/3$ .

17. For the color-blindness example, we have

$$\mathbf{B} = \begin{array}{c} G, GG \quad g, gg \\ g, GG \\ G, Gg \\ g, Gg \\ G, gg \end{array} \begin{pmatrix} 2/3 & 1/3 \\ 2/3 & 1/3 \\ 1/3 & 2/3 \\ 1/3 & 2/3 \end{pmatrix},$$

and for Example 9 of Section 11.1, we have

$$\mathbf{B} = \begin{array}{c} GG, GG \quad gg, gg \\ GG, Gg \\ GG, gg \\ Gg, Gg \\ Gg, gg \end{array} \begin{pmatrix} 3/4 & 1/4 \\ 1/2 & 1/2 \\ 1/2 & 1/2 \\ 1/4 & 3/4 \end{pmatrix}.$$

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

$$\mathbf{P} = \begin{array}{c} GG, GG \\ GG, Gg \\ GG, gg \\ Gg, Gg \\ Gg, gg \\ gg, gg \end{array} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1/4 & 1/2 & 0 & 1/4 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1/16 & 1/4 & 1/8 & 1/4 & 1/4 & 1/16 \\ 0 & 0 & 0 & 1/4 & 1/2 & 1/4 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

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.

19. (a)

$$\mathbf{P} = \begin{array}{c} 1 \\ 2 \\ 0 \\ 3 \end{array} \begin{pmatrix} 1 & 2 & 0 & 3 \\ 0 & 2/3 & 1/3 & 0 \\ 2/3 & 0 & 0 & 1/3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

(b)

$$\mathbf{N} = \begin{array}{c} 1 \quad 2 \\ 1 \quad \begin{pmatrix} 9/5 & 6/5 \\ 6/5 & 9/5 \end{pmatrix} \\ 2 \end{array},$$

$$\mathbf{B} = \begin{array}{c} 0 \quad 3 \\ 1 \quad \begin{pmatrix} 3/5 & 2/5 \\ 2/5 & 3/5 \end{pmatrix} \\ 2 \end{array},$$

$$\mathbf{t} = \begin{array}{c} 1 \quad \begin{pmatrix} 3 \\ 3 \end{pmatrix} \\ 2 \end{array}.$$

(c) The game will last on the average 3 moves.

(d) If Mary deals, the probability that John wins the game is  $3/5$ .

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

$$\mathbf{I} + \mathbf{Q} + \mathbf{Q}^2 + \cdots + \mathbf{Q}^{n-1} \rightarrow \mathbf{N},$$

so

$$\mathbf{w}^{(n)} \rightarrow \mathbf{w} = \mathbf{uN}.$$

(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

$$\mathbf{u} = \mathbf{w}(\mathbf{I} - \mathbf{Q}).$$

27. Use the solution to Exercise 24 with  $\mathbf{w} = \mathbf{f}$ .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.



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 absorbing 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.)
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)
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  $b = 1)$ .
5. (a)  $(2/3, 1/3)$   
(b)  $(1/2, 1/2)$   
(c)  $(2/7, 3/7, 2/7)$
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.
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

$$\mathbf{w} = (1/n, \dots, 1/n)$$

is a fixed probability vector. For an ergodic chain this means the the average number of times in each state is  $1/n$ .

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.
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)
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 Ehrenfest Urn model the fixed vector must satisfy the following equations:

$$\begin{aligned} \frac{1}{n}w_1 &= w_0 , \\ w_{j+1} \frac{j+1}{n} + w_{j-1} \frac{n-j+1}{n} &= w_j, \quad \text{if } 0 < j < n, \\ \frac{1}{n}w_{n-1} &= w_n . \end{aligned}$$

It is easy to check that the binomial coefficients satisfy these conditions.

17. Consider the Markov chain whose state is the value of  $S_n \bmod(7)$ , that is, the remainder when  $S_n$  is divided by 7. Then the transition matrix for this chain is

$$\mathbf{P} = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{pmatrix} 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 0 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 0 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 0 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 \end{pmatrix} \end{matrix}.$$

Since the column sums of this matrix are 1, the fixed vector is

$$\mathbf{w} = (1/7, 1/7, 1/7, 1/7, 1/7, 1/7, 1/7).$$

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  $(1, 2, \dots, r-1, 1)$   $r-2$  times ( $r^2 - 3r + 2$  steps) and  $(1, \dots, r, 1)$  once ( $r$  steps). For  $k = 1, \dots, r-1$  to reach state  $k+1$ , circle  $(1, 2, \dots, r, 1)$   $r-k$  times ( $r^2 - rk$  steps) then  $(1, 2, \dots, r-1, 1)$   $k-2$  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)

$$\mathbf{P} = \begin{pmatrix} 0 & * & 0 \\ * & 0 & * \\ * & 0 & 0 \end{pmatrix}, \quad \mathbf{P}^2 = \begin{pmatrix} * & 0 & * \\ * & * & 0 \\ 0 & * & 0 \end{pmatrix}, \quad \mathbf{P}^3 = \begin{pmatrix} * & * & 0 \\ * & * & * \\ * & 0 & * \end{pmatrix},$$

$$\mathbf{P}^4 = \begin{pmatrix} * & * & * \\ * & * & * \\ * & * & 0 \end{pmatrix}, \quad \mathbf{P}^5 = \begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \end{pmatrix}.$$

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 intersects 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.
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.

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  $n$ th 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.

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  $i$ th 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.

## 11.4 Fundamental Limit Theorem for Regular Chains

1.  $\begin{pmatrix} 1/3 \\ 1/3 \end{pmatrix}$

3. For regular chains, only the constant vectors are fixed column vectors.

## 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

$$\mathbf{M} = \begin{pmatrix} 0 & 2 \\ 4 & 0 \end{pmatrix}.$$

3. 2

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.

7. (a)

$$\begin{array}{c} \begin{matrix} & 1 & 2 & 3 & 4 & 5 & 6 \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1/4 & 1/4 & 0 & 1/4 & 1/4 & 0 \\ 0 & 0 & 1/2 & 0 & 0 & 1/2 \\ 0 & 0 & 1/2 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 \end{pmatrix} \end{matrix} \end{array}$$

(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)  $m_{1,5} = 7$ 

9. (a) if  $n$  is odd,  $\mathbf{P}$  is regular. If  $n$  is even,  $\mathbf{P}$  is ergodic but not regular.

(b)  $\mathbf{w} = (1/n, \dots, 1/n)$ .(c) From the program **Ergodic** we obtain

$$\mathbf{M} = \begin{array}{c} \begin{matrix} & 0 & 1 & 2 & 3 & 4 \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{matrix} & \begin{pmatrix} 0 & 4 & 6 & 6 & 4 \\ 4 & 0 & 4 & 6 & 6 \\ 6 & 4 & 0 & 4 & 6 \\ 6 & 6 & 4 & 0 & 4 \\ 4 & 6 & 6 & 4 & 0 \end{pmatrix} \end{matrix} \end{array}.$$

This is consistent with the conjecture that  $m_{ij} = d(n - d)$ , where  $d$  is the clockwise distance from  $i$  to  $j$ .

11. Yes, the reverse transition matrix is the same matrix.
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  $\mathbf{w} = \mathbf{w}^*$ .

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.
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

$$\begin{aligned} \bar{m}_i &= \sum_k w_k \frac{z_{ii} - z_{ki}}{w_i} + 1 \\ &= \frac{z_{ii}}{w_i} \sum_k w_k - \frac{1}{w_i} \sum_k w_k z_{ki} + 1 \\ &= \frac{z_{ii}}{w_i} - 1 + 1 = \frac{z_{ii}}{w_i}. \end{aligned}$$

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  $i$ .

21. The transition matrix is

$$\mathbf{P} = \begin{array}{c} \begin{array}{cccc} GO & A & B & C \\ GO & \left( \begin{array}{cccc} 1/6 & 1/3 & 1/3 & 1/6 \\ 1/6 & 1/6 & 1/3 & 1/3 \\ 1/3 & 1/6 & 1/6 & 1/3 \\ 1/3 & 1/3 & 1/6 & 1/6 \end{array} \right) \\ A \\ B \\ C \end{array} \end{array}.$$

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  $\mathbf{w}\mathbf{f} = \mathbf{0}$ . From the result of Exercise 20 we see that your expected winning starting in GO is the first component of the vector  $\mathbf{Z}\mathbf{f}$  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.

23. Assume that the chain is started in state  $s_i$ . Let  $X_j^{(n)}$  equal 1 if the chain is in state  $s_j$  on the  $n$ th 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

$$E(S_j^{(n)}) = \sum_{h=0}^n p_{ij}^{(h)} .$$

It now follows from Exercise 16 that

$$\lim_{n \rightarrow \infty} \frac{E(S_j^{(n)})}{n} = w_j .$$





## 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

$$\begin{aligned} p_n &= 0, & \text{if } n \text{ is even,} \\ p_1 &= q, \\ p_n &= p(p_1 p_{n-2} + p_3 p_{n-4} + \cdots + p_{n-2} p_1), & \text{if } n > 1 \text{ is odd.} \end{aligned}$$

Thus

$$h(z) = qz + pz(h(z))^2,$$

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.

3. (a) From the hint:

$$h_k(z) = h_{U_1}(z) \cdots h_{U_k}(z) = (h(z))^k.$$

(b)

$$h_k(1) = \left(h(1)\right)^k = \begin{cases} (q/p)^k & \text{if } q \leq p, \\ 1 & \text{if } q \geq p. \end{cases}$$
$$h'(1) = \begin{cases} k/(q-p) & \text{if } q > p, \\ \infty & \text{if } q = p. \end{cases}$$

Thus the gambler must be ruined if  $q \geq p$ . The expected number of plays in this case is  $k/(q-p)$  if  $q > p$  and  $\infty$  if  $q = p$ . When  $q < p$  he is ruined with probability  $(q/p)^k$ .