Proposition 3.6 Let $\lambda_i = p_i/\sum_{j=i}^{m} p_j$, $i = 1, \ldots, m$. Then

$$E[T] = k + (k-1)\sum_{i=1}^{m-1} \lambda_i$$

Proof To begin, suppose that the observed random variables $X_1\ X_2, \ldots$ take on one of the values $i, i+1, \ldots, m$ with respective probabilities

$$P\{X = j\} = \frac{p_j}{p_i + \cdots + p_m}, \qquad j = i, \ldots, m$$

Let T_i denote the first k-record index when the observed data have the preceding mass function, and note that since the each data value is at least i it follows that the k-record value will equal i, and T_i will equal k, if $X_k = i$. As a result,

$$E[T_i|X_k = i] = k$$

On the other hand, if $X_k > i$ then the k-record value will exceed i, and so all data values equal to i can be disregarded when searching for the k-record value. In addition, since each data value greater than i will have probability mass function

$$P\{X = j|X > i\} = \frac{p_j}{p_{i+1} + \cdots + p_m}, \qquad j = i+1, \ldots, m$$

it follows that the total number of data values greater than i that need be observed until a k-record value appears has the same distribution as T_{i+1}. Hence,

$$E[T_i|X_k > i] = E[T_{i+1} + N_i|X_k > i]$$

where T_{i+1} is the total number of variables greater than i that we need observe to obtain a k-record, and N_i is the number of values equal to i that are observed in that time. Now, given that $X_k > i$ and that $T_{i+1} = n (n \geqslant k)$ it follows that the time to observe T_{i+1} values greater than i has the same distribution as the number of trials to obtain n successes given that trial k is a success and that each trial is independently a success with probability $1 - p_i/\sum_{j \geqslant i} p_j = 1 - \lambda_i$ Thus, since the number of trials needed to obtain a success is a geometric random variable with mean $1/(1-\lambda_i)$, we see that

$$E[T_i|T_{i+1}, X_k > i] = 1 + \frac{T_{i+1} - 1}{1 - \lambda_i} = \frac{T_{i+1} - \lambda_i}{1 - \lambda_i}$$

Taking expectations gives that

$$E[T_i|X_k > i] = E\left[\frac{T_{i+1} - \lambda_i}{1 - \lambda_i}\middle| X_k > i\right] = \frac{E[T_{i+1}] - \lambda_i}{1 - \lambda_i}$$

Thus, upon conditioning on whether $X_k = i$, we obtain

$$\begin{aligned} E[T_i] &= E[T_i|X_k = i]\lambda_i + E[T_i|X_k > i](1 - \lambda_i) \\ &= (k-1)\lambda_i + E[T_{i+1}] \end{aligned}$$

Starting with $E[T_m] = k$, the preceding gives that

$$\begin{aligned} E[T_{m-1}] &= (k-1)\lambda_{m-1} + k \\ E[T_{m-2}] &= (k-1)\lambda_{m-2} + (k-1)\lambda_{m-1} + k \\ &= (k-1)\sum_{j=m-2}^{m-1} \lambda_j + k \\ E[T_{m-3}] &= (k-1)\lambda_{m-3} + (k-1)\sum_{j=m-2}^{m-1} \lambda_j + k \\ &= (k-1)\sum_{j=m-3}^{m-1} \lambda_j + k \end{aligned}$$

In general,

$$E[T_i] = (k-1)\sum_{j=i}^{m-1} \lambda_j + k$$

and the result follows since $T = T_1$. ■

Exercises

1. If X and Y are both discrete, show that $\sum_x p_{X|Y}(x|y) = 1$ for all y such that $p_Y(y) > 0$.

*2. Let X_1 and X_2 be independent geometric random variables having the same parameter p. Guess the value of

$$P\{X_1 = i|X_1 + X_2 = n\}$$

Hint: Suppose a coin having probability p of coming up heads is continually flipped. If the second head occurs on flip number n, what is the conditional probability that the first head was on flip number i, $i = 1, \ldots, n-1$?

Verify your guess analytically.

. The joint probability mass function of X and Y, $p(x, y)$, is given by

$$p(1,1) = \tfrac{1}{9}, \quad p(2,1) = \tfrac{1}{3}, \quad p(3,1) = \tfrac{1}{9}$$
$$p(1,2) = \tfrac{1}{9}, \quad p(2,2) = 0, \quad p(3,2) = \tfrac{1}{18}$$
$$p(1,3) = 0, \quad p(2,3) = \tfrac{1}{6}, \quad p(3,3) = \tfrac{1}{9}$$

ompute $E[X|Y = i]$ for $i = 1, 2, 3$.

. In Exercise 3, are the random variables X and Y independent? ???

. An urn contains three white, six red, and five black balls. Six of these balls are andomly selected from the urn. Let X and Y denote respectively the number of hite and black balls selected. Compute the conditional probability mass function f X given that $Y = 3$. Also compute $E[X|Y = 1]$.

6. Repeat Exercise 5 but under the assumption that when a ball is selected its olor is noted, and it is then replaced in the urn before the next selection is made.

. Suppose $p(x, y, z)$, the joint probability mass function of the random variables , Y, and Z, is given by

$$p(1,1,1) = \tfrac{1}{8}, \quad p(2,1,1) = \tfrac{1}{4},$$
$$p(1,1,2) = \tfrac{1}{8}, \quad p(2,1,2) = \tfrac{3}{16},$$
$$p(1,2,1) = \tfrac{1}{16}, \quad p(2,2,1) = 0,$$
$$p(1,2,2) = 0, \quad p(2,2,2) = \tfrac{1}{4}$$

What is $E[X|Y = 2]$? What is $E[X|Y = 2, Z = 1]$?

3. An unbiased die is successively rolled. Let X and Y denote, respectively, the number of rolls necessary to obtain a six and a five. Find (a) $E[X]$, (b) $E[X|Y = 1]$, c) $E[X|Y = 5]$.

9. Show in the discrete case that if X and Y are independent, then

$$E[X|Y = y] = E[X] \qquad \text{for all } y$$

10. Suppose X and Y are independent continuous random variables. Show that

$$E[X|Y = y] = E[X] \qquad \text{for all } y$$

11. The joint density of X and Y is

$$f(x, y) = \frac{(y^2 - x^2)}{8} e^{-y}, \qquad 0 < y < \infty, \quad -y \leqslant x \leqslant y$$

Show that $E[X|Y = y] = 0$.

12. The joint density of X and Y is given by

$$f(x, y) = \frac{e^{-x/y} e^{-y}}{y}, \qquad 0 < x < \infty, \quad 0 < y < \infty$$

Show $E[X|Y = y] = y$.

*13. Let X be exponential with mean $1/\lambda$; that is,

$$f_X(x) = \lambda e^{-\lambda x}, \qquad 0 < x < \infty$$

Find $E[X|X > 1]$.

14. Let X be uniform over $(0, 1)$. Find $E[X|X < \frac{1}{2}]$.

15. The joint density of X and Y is given by

$$f(x, y) = \frac{e^{-y}}{y}, \qquad 0 < x < y, \quad 0 < y < \infty$$

Compute $E[X^2|Y = y]$. $f_{X,Y}$ $\quad f_Y = \int f_{XY}\,dx$

16. The random variables X and Y are said to have a bivariate normal distribution if their joint density function is given by $f_{XYZ} = f_{XY} = \int f_{XYZ}\,dz$

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)} \times \left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]\right\}$$

for $-\infty < x < \infty$, $-\infty < y < \infty$, where σ_x, σ_y, μ_x, μ_y, and ρ are constants such that $-1 < \rho < 1$, $\sigma_x > 0$, $\sigma_y > 0$, $-\infty < \mu_x < \infty$, $-\infty < \mu_y < \infty$.

(a) Show that X is normally distributed with mean μ_x and variance σ_x^2, and Y is normally distributed with mean μ_y and variance σ_y^2.
(b) Show that the conditional density of X given that $Y = y$ is normal with mean $\mu_x + (\rho\sigma_x/\sigma_y)(y - \mu_y)$ and variance $\sigma_x^2(1 - \rho^2)$.

The quantity ρ is called the correlation between X and Y. It can be shown that

$$\rho = \frac{E[(X-\mu_x)(Y-\mu_y)]}{\sigma_x\sigma_y} = \frac{\text{Cov}(X, Y)}{\sigma_x\sigma_y}$$

17. Let Y be a gamma random variable with parameters (s, α). That is, its density is

$$f_Y(y) = Ce^{-\alpha y} y^{s-1}, \qquad y > 0$$

where C is a constant that does not depend on y. Suppose also that the conditional distribution of X given that $Y = y$ is Poisson with mean y. That is,

$$P\{X = i | Y = y\} = e^{-y} y^i / i!, \qquad i \geqslant 0$$

Show that the conditional distribution of Y given that $X = i$ is the gamma distribution with parameters $(s + i, \alpha + 1)$.

18. Let $X_1, \ldots, X_n$ be independent random variables having a common distribution function that is specified up to an unknown parameter θ. Let $T = T(\mathbf{X})$ be a function of the data $\mathbf{X} = (X_1, \ldots, X_n)$. If the conditional distribution of $X_1, \ldots, X_n$ given $T(\mathbf{X})$ does not depend on θ then $T(\mathbf{X})$ is said to be a *sufficient statistic* for θ. In the following cases, show that $T(\mathbf{X}) = \sum_{i=1}^{n} X_i$ is a sufficient statistic for θ.

(a) The X_i are normal with mean θ and variance 1.
(b) The density of X_i is $f(x) = \theta e^{-\theta x}$, $x > 0$.
(c) The mass function of X_i is $p(x) = \theta^x (1 - \theta)^{1-x}$, $x = 0, 1, 0 < \theta < 1$.
(d) The X_i are Poisson random variables with mean θ.

*19. Prove that if X and Y are jointly continuous, then

$$E[X] = \int_{-\infty}^{\infty} E[X|Y = y] f_Y(y)\, dy$$

20. An individual whose level of exposure to a certain pathogen is x will contract the disease caused by this pathogen with probability $P(x)$. If the exposure level of a randomly chosen member of the population has probability density function f, determine the conditional probability density of the exposure level of that member given that he or she

(a) has the disease.
(b) does not have the disease.
(c) Show that when $P(x)$ increases in x, then the ratio of the density of part (a) to that of part (b) also increases in x.

21. Consider Example 3.12 which refers to a miner trapped in a mine. Let N denote the total number of doors selected before the miner reaches safety. Also, let T_i denote the travel time corresponding to the ith choice, $i \geqslant 1$. Again let X denote the time when the miner reaches safety.

(a) Give an identity that relates X to N and the T_i.
(b) What is $E[N]$?
(c) What is $E[T_N]$?
(d) What is $E[\sum_{i=1}^{N} T_i | N = n]$?
(e) Using the preceding, what is $E[X]$?

22. Suppose that independent trials, each of which is equally likely to have any of m possible outcomes, are performed until the same outcome occurs k consecutive times. If N denotes the number of trials, show that

$$E[N] = \frac{m^k - 1}{m - 1}$$

Some people believe that the successive digits in the expansion of $\pi = 3.14159\ldots$ are "uniformly" distributed. That is, they believe that these digits have all the appearance of being independent choices from a distribution that is equally likely to be any of the digits from 0 through 9. Possible evidence against this hypothesis is the fact that starting with the 24,658,601st digit there is a run of nine successive 7s. Is this information consistent with the hypothesis of a uniform distribution?

To answer this, we note from the preceding that if the uniform hypothesis were correct, then the expected number of digits until a run of nine of the same value occurs is

$$(10^9 - 1)/9 = 111{,}111{,}111$$

Thus, the actual value of approximately 25 million is roughly 22 percent of the theoretical mean. However, it can be shown that under the uniformity assumption the standard deviation of N will be approximately equal to the mean. As a result, the observed value is approximately 0.78 standard deviations less than its theoretical mean and is thus quite consistent with the uniformity assumption.

*23. A coin having probability p of coming up heads is successively flipped until two of the most recent three flips are heads. Let N denote the number of flips. (Note that if the first two flips are heads, then $N = 2$.) Find $E[N]$.

24. A coin, having probability p of landing heads, is continually flipped until at least one head and one tail have been flipped.

(a) Find the expected number of flips needed.
(b) Find the expected number of flips that land on heads.
(c) Find the expected number of flips that land on tails.
(d) Repeat part (a) in the case where flipping is continued until a total of at least two heads and one tail have been flipped.

25. A gambler wins each game with probability p. In each of the following cases, determine the expected total number of wins.

(a) The gambler will play n games; if he wins X of these games, then he will play an additional X games before stopping.
(b) The gambler will play until he wins; if it takes him Y games to get this win, then he will play an additional Y games.

26. You have two opponents with whom you alternate play. Whenever you play A, you win with probability p_A; whenever you play B, you win with probability p_B, where $p_B > p_A$. If your objective is to minimize the number of games you need to play to win two in a row, should you start with A or with B?

Hint: Let $E[N_i]$ denote the mean number of games needed if you initially play i. Derive an expression for $E[N_A]$ that involves $E[N_B]$; write down the equivalent expression for $E[N_B]$ and then subtract.

27. A coin that comes up heads with probability p is continually flipped until the pattern T, T, H appears. (That is, you stop flipping when the most recent flip lands heads, and the two immediately preceding it lands tails.) Let X denote the number of flips made, and find $E[X]$.

28. The random variables X and Y have the following joint probability mass function:

$$P\{X = i, Y = j\} = e^{-(a+bi)} \frac{(bi)^j}{j!} \frac{a^i}{i!}, \qquad i \geqslant 0,\ j \geqslant 0$$

(a) What is the conditional distribution of Y given that $X = i$?
(b) Find $\mathrm{Cov}(X, Y)$.

29. Two players take turns shooting at a target, with each shot by player i hitting the target with probability p_i, $i = 1, 2$. Shooting ends when two consecutive shots hit the target. Let μ_i denote the mean number of shots taken when player i shoots first, $i = 1, 2$.

(a) Find μ_1 and μ_2.
(b) Let h_i denote the mean number of times that the target is hit when player i shoots first, $i = 1, 2$. Find h_1 and h_2.

30. Let X_i, $i \geqslant 0$ be independent and identically distributed random variables with probability mass function

$$p(j) = P\{X_i = j\}, \qquad j = 1, \ldots, m \qquad \sum_{j=1}^{m} P(j) = 1$$

Find $E[N]$, where $N = \min\{n > 0 : X_n = X_0\}$.

31. Each element in a sequence of binary data is either 1 with probability p or 0 with probability $1 - p$. A maximal subsequence of consecutive values having identical outcomes is called a run. For instance, if the outcome sequence is 1, 1, 0, 1, 1, 1, 0, the first run is of length 2, the second is of length 1, and the third is of length 3.

(a) Find the expected length of the first run.
(b) Find the expected length of the second run.

32. If each trial is independently a success with probability p, find the expected number of trials needed for there to have been both at least n successes and at least m failures.

Hint: Is it useful to know the results of the first $n + m$ trials?

33. If R_i denotes the random amount that is earned in period i, then $\sum_{i=1}^{\infty} \beta^{i-1} R_i$, where $0 < \beta < 1$ is a specified constant, is called the total discounted reward with discount factor β. Let T be a geometric random variable with parameter $1 - \beta$ that is independent of the R_i. Show that the expected total discounted reward is equal to the expected total (undiscounted) reward earned by time T. That is, show that

$$E\left[\sum_{i=1}^{\infty} \beta^{i-1} R_i\right] = E\left[\sum_{i=1}^{T} R_i\right]$$

34. A set of n dice is thrown. All those that land on six are put aside, and the others are again thrown. This is repeated until all the dice have landed on six. Let N denote the number of throws needed. (For instance, suppose that $n = 3$ and that on the initial throw exactly two of the dice land on six. Then the other die will be thrown, and if it lands on six, then $N = 2$.) Let $m_n = E[N]$.

(a) Derive a recursive formula for m_n and use it to calculate m_i, $i = 2, 3, 4$ and to show that $m_5 \approx 13.024$.
(b) Let X_i denote the number of dice rolled on the ith throw. Find $E[\sum_{i=1}^{N} X_i]$.

35. If $E[X|Y = y] = c$ for all y, show that $\mathrm{Cov}(X, Y) = 0$.

36. Suppose that conditional on $Y = y$, the random variables X_1 and X_2 are independent with mean y. Show that

$$\mathrm{Cov}(X_1, X_2) = \mathrm{Var}(Y)$$

37. A manuscript is sent to a typing firm consisting of typists A, B, and C. If it is typed by A, then the number of errors made is a Poisson random variable with mean 2.6; if typed by B, then the number of errors is a Poisson random variable with mean 3; and if typed by C, then it is a Poisson random variable with mean 3.4. Let X denote the number of errors in the typed manuscript. Assume that each typist is equally likely to do the work.

(a) Find $E[X]$.
(b) Find $\mathrm{Var}(X)$.

38. Let U be a uniform $(0, 1)$ random variable. Suppose that n trials are to be performed and that conditional on $U = u$ these trials will be independent with a common success probability u. Compute the mean and variance of the number of successes that occur in these trials.

39. A deck of n cards, numbered 1 through n, is randomly shuffled so that all $n!$ possible permutations are equally likely. The cards are then turned over one at a time until card number 1 appears. These upturned cards constitute the first cycle. We now determine (by looking at the upturned cards) the lowest numbered card that has not yet appeared, and we continue to turn the cards face up until that card appears. This new set of cards represents the second cycle. We again determine the lowest numbered of the remaining cards and turn the cards until it appears, and so on until all cards have been turned over. Let m_n denote the mean number of cycles.

(a) Derive a recursive formula for m_n in terms of m_k, $k = 1, \dots, n-1$.
(b) Starting with $m_0 = 0$, use the recursion to find m_1, m_2, m_3, and m_4.
(c) Conjecture a general formula for m_n.
(d) Prove your formula by induction on n. That is, show it is valid for $n = 1$, then assume it is true for any of the values $1, \dots, n-1$ and show that this implies it is true for n.
(e) Let X_i equal 1 if one of the cycles ends with card i, and let it equal 0 otherwise, $i = 1, \dots, n$. Express the number of cycles in terms of these X_i.
(f) Use the representation in part (e) to determine m_n.
(g) Are the random variables $X_1, \dots, X_n$ independent? Explain.
(h) Find the variance of the number of cycles.

40. A prisoner is trapped in a cell containing three doors. The first door leads to a tunnel that returns him to his cell after two days of travel. The second leads to a tunnel that returns him to his cell after three days of travel. The third door leads immediately to freedom.

(a) Assuming that the prisoner will always select doors 1, 2, and 3 with probabilities 0.5, 0.3, 0.2, what is the expected number of days until he reaches freedom?
(b) Assuming that the prisoner is always equally likely to choose among those doors that he has not used, what is the expected number of days until he reaches freedom? (In this version, for instance, if the prisoner initially tries door 1, then when he returns to the cell, he will now select only from doors 2 and 3.)
(c) For parts (a) and (b) find the variance of the number of days until the prisoner reaches freedom.

41. A rat is trapped in a maze. Initially it has to choose one of two directions. If it goes to the right, then it will wander around in the maze for three minutes and will then return to its initial position. If it goes to the left, then with probability $\frac{1}{3}$ it will depart the maze after two minutes of traveling, and with probability $\frac{2}{3}$ it will return to its initial position after five minutes of traveling. Assuming that the rat is at all times equally likely to go to the left or the right, what is the expected number of minutes that it will be trapped in the maze?

42. A total of 11 people, including you, are invited to a party. The times at which people arrive at the party are independent uniform $(0, 1)$ random variables.

(a) Find the expected number of people who arrive before you.
(b) Find the variance of the number of people who arrive before you.

43. The number of claims received at an insurance company during a week is a random variable with mean μ_1 and variance σ_1^2. The amount paid in each claim is a random variable with mean μ_2 and variance σ_2^2. Find the mean and variance of the amount of money paid by the insurance company each week. What independence assumptions are you making? Are these assumptions reasonable?

44. The number of customers entering a store on a given day is Poisson distributed with mean $\lambda = 10$. The amount of money spent by a customer is uniformly distributed over $(0, 100)$. Find the mean and variance of the amount of money that the store takes in on a given day.

45. An individual traveling on the real line is trying to reach the origin. However, the larger the desired step, the greater is the variance in the result of that step. Specifically, whenever the person is at location x, he next moves to a location having mean 0 and variance βx^2. Let X_n denote the position of the individual after having taken n steps. Supposing that $X_0 = x_0$, find

(a) $E[X_n]$
(b) $\text{Var}(X_n)$

46. (a) Show that

$$\text{Cov}(X, Y) = \text{Cov}(X, E[Y \mid X])$$

(b) Suppose, that, for constants a and b,

$$E[Y \mid X] = a + bX$$

Show that

$$b = \text{Cov}(X, Y)/\text{Var}(X)$$

***47.** If $E[Y \mid X] = 1$, show that

$$\text{Var}(X\,Y) \geqslant \text{Var}(X)$$

48. Give another proof of the result of Example 3.17 by computing the moment generating function of $\sum_{i=1}^{N} X_i$ and then differentiating to obtain its moments.

Hint: Let

$$\begin{aligned}\phi(t) &= E\left[\exp\left(t\sum_{i=1}^{N} X_i\right)\right] \\ &= E\left[E\left[\exp\left(t\sum_{i=1}^{N} X_i\right)\Big|N\right]\right]\end{aligned}$$

Now,

$$E\left[\exp\left(t\sum_{i=1}^{N} X_i\right)\Big|N=n\right] = E\left[\exp\left(t\sum_{i=1}^{n} X_i\right)\right] = (\phi_X(t))^n$$

since N is independent of the Xs where $\phi_X(t) = E[e^{tX}]$ is the moment generating function for the Xs Therefore,

$$\phi(t) = E[(\phi_X(t))^N]$$

Differentiation yields

$$\begin{aligned}\phi'(t) &= E[N(\phi_X(t))^{N-1}\phi'_X(t)], \\ \phi''(t) &= E[N(N-1)(\phi_X(t))^{N-2}(\phi'_X(t))^2 + N(\phi_X(t))^{N-1}\phi''_X(t)]\end{aligned}$$

Evaluate at $t = 0$ to get the desired result.

49. A and B play a series of games with A winning each game with probability p. The overall winner is the first player to have won two more games than the other.

(a) Find the probability that A is the overall winner.
(b) Find the expected number of games played.

50. There are three coins in a barrel. These coins, when flipped, will come up heads with respective probabilities 0.3, 0.5, 0.7. A coin is randomly selected from among these three and is then flipped ten times. Let N be the number of heads obtained on the ten flips. Find

(a) $P\{N = 0\}$.
(b) $P\{N = n\}, n = 0, 1, \ldots, 10$.
(c) Does N have a binomial distribution?
(d) If you win \$1 each time a head appears and you lose \$1 each time a tail appears, is this a fair game? Explain.

51. Do Exercise 50 under the assumption that each time a coin is flipped, it is then put back in the barrel and another coin is randomly selected. Does N have a binomial distribution now?

52. Suppose that X and Y are independent random variables with probability density functions f_X and f_Y. Determine a one-dimensional integral expression for $P\{X + Y < x\}$.

***53.** Suppose X is a Poisson random variable with mean λ. The parameter λ is itself a random variable whose distribution is exponential with mean 1. Show that $P\{X = n\} = (\frac{1}{2})^{n+1}$.

54. A coin is randomly selected from a group of ten coins, the nth coin having a probability $n/10$ of coming up heads. The coin is then repeatedly flipped until a head appears. Let N denote the number of flips necessary. What is the probability distribution of N? Is N a geometric random variable? When would N be a geometric random variable; that is, what would have to be done differently?

55. Suppose in Exercise 42 that, aside from yourself, the number of other people who are invited is a Poisson random variable with mean 10.

(a) Find the expected number of people who arrive before you.
(b) Find the probability that you are the n^{th} person to arrive.

56. Data indicate that the number of traffic accidents in Berkeley on a rainy day is a Poisson random variable with mean 9, whereas on a dry day it is a Poisson random variable with mean 3. Let X denote the number of traffic accidents tomorrow. If it will rain tomorrow with probability .6, find

(a) $E[X]$;
(b) $P\{X = 0\}$;
(c) $\mathrm{Var}(X)$.

57. The number of storms in the upcoming rainy season is Poisson distributed but with a parameter value that is uniformly distributed over $(0, 5)$. That is, Λ is uniformly distributed over $(0, 5)$, and given that $\Lambda = \lambda$, the number of storms is Poisson with mean λ. Find the probability there are at least three storms this season.

58. A collection of n coins is flipped. The outcomes are independent, and the ith coin comes up heads with probability $\alpha_i, i = 1, \ldots, n$. Suppose that for some value of j, $1 \leqslant j \leqslant n$, $\alpha_j = \frac{1}{2}$. Find the probability that the total number of heads to appear on the n coins is an even number.

59. Let A and B be mutually exclusive events of an experiment. If independent replications of the experiment are continually performed, what is the probability that A occurs before B?

60. Two players alternate flipping a coin that comes up heads with probability . The first one to obtain a head is declared the winner. We are interested in e probability that the first player to flip is the winner. Before determining this robability, which we will call $f(p)$, answer the following questions.

(a) Do you think that $f(p)$ is a monotone function of p? If so, is it increasing or decreasing?
(b) What do you think is the value of $\lim_{p \to 1} f(p)$?
(c) What do you think is the value of $\lim_{p \to 0} f(p)$?
(d) Find $f(p)$.

1. Suppose in Exercise 29 that the shooting ends when the target has been hit vice. Let m_i denote the mean number of shots needed for the first hit when player shoots first, $i = 1, 2$. Also, let P_i, $i = 1, 2$, denote the probability that the first it is by player 1, when player i shoots first.

(a) Find m_1 and m_2.
(b) Find P_1 and P_2.

or the remainder of the problem, assume that player 1 shoots first.

(c) Find the probability that the final hit was by 1.
(d) Find the probability that both hits were by 1.
(e) Find the probability that both hits were by 2.
(f) Find the mean number of shots taken.

2. A, B, and C are evenly matched tennis players. Initially A and B play a et, and the winner then plays C. This continues, with the winner always playing e waiting player, until one of the players has won two sets in a row. That player then declared the overall winner. Find the probability that A is the overall inner.

3. Let X_1 and X_2 be independent geometric random variables with respective arameters p_1 and p_2. Find $P\{|X_1 - X_2| \leqslant 1\}$.

4. A and B roll a pair of dice in turn, with A rolling first. A's objective is to btain a sum of 6, and B's is to obtain a sum of 7. The game ends when either layer reaches his or her objective, and that player is declared the winner.

(a) Find the probability that A is the winner.
(b) Find the expected number of rolls of the dice.
(c) Find the variance of the number of rolls of the dice.

5. The number of red balls in an urn that contains n balls is a random variable at is equally likely to be any of the values $0, 1, \ldots, n$. That is,

$$P\{i \text{ red}, n - i \text{ non-red}\} = \frac{1}{n+1}, \qquad i = 0, \ldots, n$$

The n balls are then randomly removed one at a time. Let Y_k denote the number of red balls in the first k selections, $k = 1, \ldots, n$.

(a) Find $P\{Y_n = j\}$, $j = 0, \ldots, n$.
(b) Find $P\{Y_{n-1} = j\}$, $j = 0, \ldots, n$.
(c) What do you think is the value of $P\{Y_k = j\}$, $j = 0, \ldots, n$?
(d) Verify your answer to part (c) by a backwards induction argument. That is, check that your answer is correct when $k = n$, and then show that whenever it is true for k it is also true for $k - 1$, $k = 1, \ldots, n$.

66. The opponents of soccer team A are of two types: either they are a class 1 or a class 2 team. The number of goals team A scores against a class i opponent is a Poisson random variable with mean λ_i, where $\lambda_1 = 2$, $\lambda_2 = 3$. This weekend the team has two games against teams they are not very familiar with. Assuming that the first team they play is a class 1 team with probability 0.6 and the second is, independently of the class of the first team, a class 1 team with probability 0.3, determine

(a) the expected number of goals team A will score this weekend.
(b) the probability that team A will score a total of five goals.

*67. A coin having probability p of coming up heads is continually flipped. Let $P_j(n)$ denote the probability that a run of j successive heads occurs within the first n flips.

(a) Argue that

$$P_j(n) = P_j(n-1) + p^j(1-p)[1 - P_j(n-j-1)]$$

(b) By conditioning on the first non-head to appear, derive another equation relating $P_j(n)$ to the quantities $P_j(n-k)$, $k = 1, \ldots, j$.

68. In a knockout tennis tournament of 2^n contestants, the players are paired and play a match. The losers depart, the remaining 2^{n-1} players are paired, and they play a match. This continues for n rounds, after which a single player remains unbeaten and is declared the winner. Suppose that the contestants are numbered 1 through 2^n, and that whenever two players contest a match, the lower numbered one wins with probability p. Also suppose that the pairings of the remaining players are always done at random so that all possible pairings for that round are equally likely.

(a) What is the probability that player 1 wins the tournament?
(b) What is the probability that player 2 wins the tournament?

Hint: Imagine that the random pairings are done in advance of the tournament. That is, the first-round pairings are randomly determined; the 2^{n-1} first-round pairs are then themselves randomly paired, with the winners of each pair to play in round 2; these 2^{n-2} groupings (of four players each) are then randomly paired, with the winners of each grouping to play in round 3, and so on. Say that players i and j are scheduled to meet in round k if, provided they both win their first $k-1$ matches, they will meet in round k. Now condition on the round in which players 1 and 2 are scheduled to meet.

9. In the match problem, say that (i, j), $i < j$, is a pair if i chooses j's hat and chooses i's hat.

(a) Find the expected number of pairs.
(b) Let Q_n denote the probability that there are no pairs, and derive a recursive formula for Q_n in terms of Q_j, $j < n$.

Hint: Use the cycle concept.

(c) Use the recursion of part (b) to find Q_8.

0. Let N denote the number of cycles that result in the match problem.

(a) Let $M_n = E[N]$, and derive an equation for M_n in terms of $M_1, \ldots, M_{n-1}$.
(b) Let C_j denote the size of the cycle that contains person j. Argue that

$$N = \sum_{j=1}^{n} 1/C_j$$

and use the preceding to determine $E[N]$.
(c) Find the probability that persons $1, 2, \ldots, k$ are all in the same cycle.
(d) Find the probability that $1, 2, \ldots, k$ is a cycle.

1. Use the equation following (3.14) to obtain Equation (3.10).

Hint: First multiply both sides of Equation (3.14) by n, then write a new equation by replacing n with $n-1$, and then subtract the former from the latter.

2. In Example 3.25 show that the conditional distribution of N given that $_1 = y$ is the same as the conditional distribution of M given that $U_1 = 1 - y$. lso, show that

$$E[N|U_1 = y] = E[M|U_1 = 1 - y] = 1 + e^y$$

73. Suppose that we continually roll a die until the sum of all throws exceeds 00. What is the most likely value of this total when you stop?

4. There are five components. The components act independently, with comnent i working with probability p_i, $i = 1, 2, 3, 4, 5$. These components form a stem as shown in Figure 3.7.

The system is said to work if a signal originating at the left end of the diagram can ach the right end, where it can pass through a component only if that component

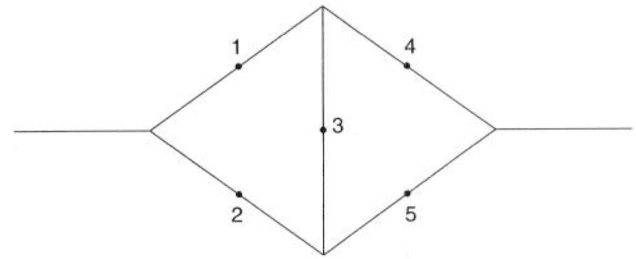

Figure 3.7.

is working. (For instance, if components 1 and 4 both work, then the system also works.) What is the probability that the system works?

75. This problem will present another proof of the ballot problem of Example 3.24.

(a) Argue that

$$P_{n,m} = 1 - P\{A \text{ and } B \text{ are tied at some point}\}$$

(b) Explain why

$$P\{A \text{ receives first vote and they are eventually tied}\}$$
$$= P\{B \text{ receives first vote and they are eventually tied}\}$$

Hint: Any outcome in which they are eventually tied with A receiving the first vote corresponds to an outcome in which they are eventually tied with B receiving the first vote. Explain this correspondence.

(c) Argue that P {eventually tied} = $2m/(n+m)$, and conclude that $P_{n,m} = (n-m)/(n+m)$.

76. Consider a gambler who on each bet either wins 1 with probability 18/38 or loses 1 with probability 20/38. (These are the probabilities if the bet is that a roulette wheel will land on a specified color.) The gambler will quit either when he or she is winning a total of 5 or after 100 plays. What is the probability he or she plays exactly 15 times?

77. Show that

(a) $E[XY|Y = y] = yE[X|Y = y]$
(b) $E[g(X, Y)|Y = y] = E[g(X, y)|Y = y]$
(c) $E[XY] = E[YE[X|Y]]$

78. In the ballot problem (Example 3.24), compute $P\{A \text{ is never behind}\}$.

9. An urn contains n white and m black balls which are removed one at a time. If $> m$, show that the probability that there are always more white than black balls n the urn (until, of course, the urn is empty) equals $(n-m)/(n+m)$. Explain why his probability is equal to the probability that the set of withdrawn balls always ontains more white than black balls. [This latter probability is $(n-m)/(n+m)$ y the ballot problem.]

0. A coin that comes up heads with probability p is flipped n consecutive times. Vhat is the probability that starting with the first flip there are always more heads han tails that have appeared?

1. Let $X_i, i \geqslant 1$, be independent uniform $(0, 1)$ random variables, and efine N by

$$N = \min\{n: X_n < X_{n-1}\}$$

here $X_0 = x$. Let $f(x) = E[N]$.

(a) Derive an integral equation for $f(x)$ by conditioning on X_1.
(b) Differentiate both sides of the equation derived in part (a).
(c) Solve the resulting equation obtained in part (b).
(d) For a second approach to determining $f(x)$ argue that

$$P\{N \geqslant k\} = \frac{(1-x)^{k-1}}{(k-1)!}$$

(e) Use part (d) to obtain $f(x)$.

2. Let $X_1, X_2, \ldots$ be independent continuous random variables with a common stribution function F and density $f = F'$, and for $k \geqslant 1$ let

$$N_k = \min\{n \geqslant k: X_n = k\text{th largest of } X_1, \ldots, X_n\}$$

(a) Show that $P\{N_k = n\} = \frac{k-1}{n(n-1)}, n \geqslant k$.
(b) Argue that

$$f_{X_{N_k}}(x) = f(x)(\bar{F}(x))^{k-1} \sum_{i=0}^{\infty} \binom{i+k-2}{i} (F(x))^i$$

(c) Prove the following identity:

$$a^{1-k} = \sum_{i=0}^{\infty} \binom{i+k-2}{i} (1-a)^i, \qquad 0 < a < 1, k \geqslant 2$$

Hint: Use induction. First prove it when $k = 2$, and then assume it for k. To prove it for $k+1$, use the fact that

$$\sum_{i=1}^{\infty} \binom{i+k-1}{i} (1-a)^i = \sum_{i=1}^{\infty} \binom{i+k-2}{i} (1-a)^i + \sum_{i=1}^{\infty} \binom{i+k-2}{i-1} (1-a)^i$$

where the preceding used the combinatorial identity

$$\binom{m}{i} = \binom{m-1}{i} + \binom{m-1}{i-1}$$

Now, use the induction hypothesis to evaluate the first term on the right side of the preceding equation.

(d) Conclude that X_{N_k} has distribution F.

83. An urn contains n balls, with ball i having weight $w_i, i = 1, \ldots, n$. The balls are withdrawn from the urn one at a time according to the following scheme: When S is the set of balls that remains, ball i, $i \in S$, is the next ball withdrawn with probability $w_i / \sum_{j \in S} w_j$. Find the expected number of balls that are withdrawn before ball i, $i = 1, \ldots, n$.

84. In the list example of Section 3.6.1 suppose that the initial ordering at time $t = 0$ is determined completely at random; that is, initially all $n!$ permutations are equally likely. Following the front-of-the-line rule, compute the expected position of the element requested at time t.

Hint: To compute $P\{e_j \text{ precedes } e_i \text{ at time } t\}$ condition on whether or not either e_i or e_j has ever been requested prior to t.

85. In the list problem, when the P_i are known, show that the best ordering (best in the sense of minimizing the expected position of the element requested) is to place the elements in decreasing order of their probabilities. That is, if $P_1 > P_2 > \cdots > P_n$, show that $1, 2, \ldots, n$ is the best ordering.

86. Consider the random graph of Section 3.6.2 when $n = 5$. Compute the probability distribution of the number of components and verify your solution by using it to compute $E[C]$and then comparing your solution with

$$E[C] = \sum_{k=1}^{5} \binom{5}{k} \frac{(k-1)!}{5^k}$$

7. (a) From the results of Section 3.6.3 we can conclude that there are $\binom{n+m-1}{m-1}$ nonnegative integer valued solutions of the equation $x_1 + \cdots + x_m = n$. Prove this directly.

(b) How many positive integer valued solutions of $x_1 + \cdots + x_m = n$ are there?

Hint: Let $y_i = x_i - 1$.

(c) For the Bose–Einstein distribution, compute the probability that exactly k of the X_i are equal to 0.

8. In Section 3.6.3, we saw that if U is a random variable that is uniform on), 1) and if, conditional on $U = p$, X is binomial with parameters n and p, then

$$P\{X = i\} = \frac{1}{n+1}, \qquad i = 0, 1, \ldots, n$$

or another way of showing this result, let $U, X_1, X_2, \ldots, X_n$ be independent niform (0, 1) random variables. Define X by

$$X = \#i\colon X_i < U$$

hat is, if the $n+1$ variables are ordered from smallest to largest, then U would e in position $X + 1$.

(a) What is $P\{X = i\}$?

(b) Explain how this proves the result of Exercise 88.

9. Let $I_1, \ldots, I_n$ be independent random variables, each of which is equally kely to be either 0 or 1. A well-known nonparametric statistical test (called the gned rank test) is concerned with determining $P_n(k)$ defined by

$$P_n(k) = P\left\{\sum_{j=1}^{n} jI_j \leqslant k\right\}$$

ıstify the following formula:

$$P_n(k) = \tfrac{1}{2}P_{n-1}(k) + \tfrac{1}{2}P_{n-1}(k-n)$$

0. The number of accidents in each period is a Poisson random variable with ean 5. With X_n, $n \geqslant 1$, equal to the number of accidents in period n, find $E[N]$ hen

(a) $N = \min(n\colon X_{n-2} = 2, X_{n-1} = 1, X_n = 0)$

(b) $N = \min(n\colon X_{n-3} = 2, X_{n-2} = 1, X_{n-1} = 0, X_n = 2)$

1. Find the expected number of flips of a coin, which comes up heads with robability p, that are necessary to obtain the pattern h, t, h, h, t, h, t, h.

92. The number of coins that Josh spots when walking to work is a Poisson random variable with mean 6. Each coin is equally likely to be a penny, a nickel, a dime, or a quarter. Josh ignores the pennies but picks up the other coins.

(a) Find the expected amount of money that Josh picks up on his way to work.

(b) Find the variance of the amount of money that Josh picks up on his way to work.

(c) Find the probability that Josh picks up exactly 25 cents on his way to work.

*93. Consider a sequence of independent trials, each of which is equally likely to result in any of the outcomes $0, 1, \ldots, m$. Say that a round begins with the first trial, and that a new round begins each time outcome 0 occurs. Let N denote the number of trials that it takes until all of the outcomes $1, \ldots, m-1$ have occurred in the same round. Also, let T_j denote the number of trials that it takes until j distinct outcomes have occurred, and let I_j denote the j^{th} distinct outcome to occur. (Therefore, outcome I_j first occurs at trial T_j.)

(a) Argue that the random vectors $(I_1, \ldots, I_m)$ and $(T_1, \ldots, T_m)$ are independent.

(b) Define X by letting $X = j$ if outcome 0 is the j^{th} distinct outcome to occur. (Thus, $I_X = 0$.) Derive an equation for $E[N]$ in terms of $E[T_j]$, $j = 1, \ldots, m-1$ by conditioning on X.

(c) Determine $E[T_j]$, $j = 1, \ldots, m-1$.

Hint: See Exercise 42 of Chapter 2.

(d) Find $E[N]$.