Chapter 3

3.1 Take as sample space the set consisting of the four outcomes $(0,0)$, $(1,0)$, $(0,1)$ and $(1,1)$, where the first (second) component is 0 if the first (second) student picked has not done homework and is 1 otherwise. The probabilities $\frac{5}{20} \times \frac{4}{19} = \frac{1}{19}$, $\frac{15}{20} \times \frac{5}{19} = \frac{15}{76}$, $\frac{5}{20} \times \frac{15}{19} = \frac{15}{76}$, and $\frac{15}{20} \times \frac{14}{19} = \frac{21}{38}$ are assigned to these four outcomes. The random variable X takes on the value 0 for the outcome $(1,1)$, the value 1 for the outcomes $(0,1)$ and $(1,0)$, and the value 2 for the outcome $(0,0)$. Thus $P(X=0) = \frac{21}{38}$, $P(X=1) = \frac{15}{76} + \frac{15}{76} = \frac{15}{38}$, and $P(X=2) = \frac{1}{19}$.

3.2 The probability mass function of X can be calculated by using conditional probabilities. Let A_i be the event that the ith person entering the room is the first person matching a birthday. Then $P(X=i) = P(A_i)$ for $i = 2, 3, \ldots, 366$. Using the chain rule for conditional probabilities, it follows that $P(X=2) = \frac{1}{365}$ and

$$P(X=i) = \frac{364}{365} \times \cdots \frac{365-i+2}{365} \times \frac{i-1}{365} \quad \text{for } i \geq 3.$$

3.3 The random variable X can take on the values 0, 1, and 2. By the law of conditional probability, $P(X=0) = \frac{1}{3} \times 0 + \frac{2}{3} \times \frac{1}{4} = \frac{1}{6}$, $P(X=1) = \frac{1}{3} \times 0 + \frac{2}{3} \times \frac{1}{2} = \frac{1}{3}$, and $P(X=2) = \frac{1}{3} \times 1 + \frac{2}{3} \times \frac{1}{4} = \frac{1}{2}$.

3.4 Denote by the random variable X the number of prize winners. The random variable X takes on the value 1 if all three digits of the lottery number drawn are the same, the value 3 if exactly two of the three digits of the lottery number drawn are the same, and the value 6 if all three digits of the lottery number drawn are different. There are 10 lottery numbers for which all three digits are the same and so $P(X=1) = \frac{10}{1{,}000}$. There are $10 \times 9 \times 8 = 720$ lottery numbers with three different digits and so $P(X=6) = \frac{720}{1{,}000}$. The probability $P(X=3) = 1 - P(X=1) - P(X=6) = \frac{270}{1{,}000}$.

3.5 The random variable X can take on the values 2, 3, and 4. Two tests are needed if the first two tests give the depleted batteries, while three tests are needed if the first three batteries tested are not depleted or if a second depleted battery is found at the third test. Thus, by the chain rule for conditional probabilities, $P(X=2) = \frac{2}{5} \times \frac{1}{4} = \frac{1}{10}$, $P(X=3) = \frac{3}{5} \times \frac{2}{4} \times \frac{1}{3} + + \frac{2}{5} \times \frac{3}{4} \times \frac{1}{3} + \frac{3}{5} \times \frac{2}{4} \times \frac{1}{3} = \frac{3}{10}$. The probability $P(X=4)$ is calculated as $1 - P(X=2) - P(X=3) = \frac{6}{10}$.

3.6 For $1 \leq k \leq 4$, you get $2^{k-1} \times 10$ dollars if the first k tosses are heads and the $(k+1)$th toss is tails. You get 160 dollars if the first five tosses are heads. The probability mass function of X is $P(X = 0) = \frac{1}{2}$, $P(X = 10) = \frac{1}{4}$, $P(X = 20) = \frac{1}{8}$, $P(X = 40) = \frac{1}{16}$, $P(X = 80) = \frac{1}{32}$, and $P(X = 160) = \frac{1}{32}$.

3.7 The random variable X can take on the values 0, 5, 10, 15, 20, and 25. Using the chain rule for conditional probabilities, $P(X = 0) = \frac{4}{7}$, $P(X = 5) = \frac{1}{7} \times \frac{4}{6}$, $P(X = 10) = \frac{2}{7} \times \frac{4}{6}$, $P(X = 15) = \frac{2}{7} \times \frac{1}{6} \times \frac{4}{5} + \frac{1}{7} \times \frac{2}{6} \times \frac{4}{5}$, $P(X = 20) = \frac{2}{7} \times \frac{1}{6} \times \frac{4}{5}$, and $P(X = 25) = \frac{2}{7} \times \frac{1}{6} \times \frac{1}{5} + \frac{2}{7} \times \frac{1}{6} \times \frac{1}{5} + \frac{1}{7} \times \frac{2}{6} \times \frac{1}{5}$.

3.8 The sample space is the set $\{1, \ldots, 10\}$, where the outcome i means that the card with number i is picked. Let the X be your payoff. The random variable X can take on the values 0.5, 5, 6, 7, 8, 9, and 10. We have $P(X = 0.5) = \frac{4}{10}$ and $P(X = k) = \frac{1}{10}$ for $5 \leq k \leq 10$. Therefore $E(X) = 0.5 \times \frac{4}{10} + \sum_{k=5}^{10} k \times \frac{1}{10} = 4.70$ dollars.

3.9 The probability that a randomly chosen student belongs to a particular class gets larger when the class has more students. Therefore $E(Y) > E(X)$. We have $E(X) = 15 \times \frac{1}{4} + 20 \times \frac{1}{4} + 70 \times \frac{1}{4} + 125 \times \frac{1}{4} = 57.5$ and $E(Y) = 15 \times \frac{15}{230} + 20 \times \frac{20}{230} + 70 \times \frac{70}{230} + 125 \times \frac{125}{230} = 88.125$.

3.10 Let the random variable X be the net cost of the risk reduction in dollars. The random variable X takes on the values 2,000, 2,000−5,000 and 2,000−10,000 with respective probabilities 0.75, 0.15 and 0.10. We have $E(X) = 2{,}000 \times 0.75 - 3{,}000 \times 0.15 - 8{,}000 \times 0.10 = 250$ dollars.

3.11 Let the random variable X be the amount of money you win. Then $P(X = m) = \binom{10}{m} / \binom{12}{m}$ and $P(X = 0) = 1 - P(X = m)$. This gives

$$E(X) = m\binom{10}{m} / \binom{12}{m}.$$

This expression is maximal for $m = 4$ with $E(X) = \frac{56}{33}$.

3.12 Using conditional probabilities, your probability of winning is $\frac{4}{6} \times \frac{3}{5} = \frac{2}{5}$. Hence your expected winnings is $\frac{2}{5} \times 1.25 - \frac{3}{5} \times 1 = -0.10$ dollars. This is not a fair bet.

3.13 Put for abbreviation $c_k = \binom{52}{k-1}$. Using the second argument from Example 3.2, we get $P(X_2 = k) = \frac{1}{c_k}\binom{48}{k-2}\binom{4}{1} \times \frac{3}{52-(k-1)}$, $P(X_3 = k) =$

$\frac{1}{c_k}\binom{48}{k-3}\binom{4}{2} \times \frac{2}{52-(k-1)}$, and $P(X_4 = k) = \frac{1}{c_k}\binom{48}{k-4}\binom{4}{3} \times \frac{1}{52-(k-1)}$. This leads to $E(X_2) = 21.2$, $E(X_3) = 31.8$, and $E(X_4) = 42.4$. Intuitively, the result $E(X_j) = \frac{52+1}{5} \times j$ for $1 \leq j \leq 4$ can be explained by a symmetry argument.

Note: More generally, suppose that balls are removed from a bowl containing r red and b blue balls, one at a time and at random. Then the expected number of picks until a red ball is removed for the jth time is $\frac{r+b+1}{r+1} \times j$ for $1 \leq j \leq r$.

3.14 Let the random variable X denote the number of chips you get back in any given round. The possible values of X are 0, 2, and 5. The random variable X is defined on the sample space consisting of the 36 equiprobable outcomes $(1,1), (1,2), \ldots, (6,6)$. Outcome (i,j) means that i points turn up on the first die and j points on the second die. The total of the two dice is 7 for the six outcomes $(1,6)$, $(6,1)$, $(2,5)$, $(5,2)$, $(3,4)$, and $(4,3)$. Thus $P(X = 5) = \frac{6}{36}$. Similarly, $P(X = 0) = \frac{15}{36}$ and $P(X = 2) = \frac{15}{36}$. This gives

$$E(X) = 0 \times \frac{15}{36} + 2 \times \frac{15}{36} + 5 \times \frac{6}{36} = 1\frac{2}{3}.$$

You bet two chips each round. Thus, your average loss is $2 - 1\frac{2}{3} = \frac{1}{3}$ chip per round when you play the game over and over.

3.15 Let the random variable X be the total amount staked and the random variable Y be the amount won. The probability p_k that k bets will be placed is $\left(\frac{19}{37}\right)^{k-1}\frac{18}{37}$ for $k = 1, \ldots, 10$ and $\left(\frac{19}{37}\right)^{10}$ for $k = 11$. Thus,

$$\begin{aligned} E(X) &= \sum_{k=1}^{10}(1 + 2 + \ldots + 2^{k-1})p_k + (1 + 2 + \ldots + 2^9 + 1{,}000)p_{11} \\ &= 12.583 \text{ dollars.} \end{aligned}$$

If the round goes to 11 bets, the player's loss is \$23 if the 11th bet is won and is \$2,023 if the 11th bet is lost. Thus

$$\begin{aligned} E(Y) &= 1 \times (1 - p_{11}) - 23 \times p_{11} \times \frac{18}{37} - 2{,}023 \times p_{11} \times \frac{19}{37} \\ &= -0.3401 \text{ dollars.} \end{aligned}$$

The ratio of 0.3401 and 12.583 is in line with the house advantage of 2.70% of the casino.

3.16 Let the random variable X be the payoff of the game. Then $P(X = 0) = (\frac{1}{2})^m$ and $P(X = 2^k) = (\frac{1}{2})^{k-1} \times \frac{1}{2}$ for $k = 1, \ldots, m$. Therefore $E(X) = m$.

3.17 Take as sample space the set $0 \cup \{(x, y) : x^2 + y^2 \leq 25\}$, where 0 means that the dart has missed the target. The score is a random variable X with $P(X = 0) = 0.25$, $P(X = 5) = 0.75 \times \frac{25-9}{25} = 0.48$, $P(X = 8) = 0.75 \times \frac{9-1}{25} = 0.24$, and $P(X = 15) = 0.75 \times \frac{1}{25} = 0.03$. The expected value of the score is $E(X) = 0.48 \times 5 + 0.24 \times 8 + 0.03 \times 15 = 4.77$.

3.18 Let the random variable X be your net winnings. The random variable X takes on the values -1, 0 and 10. We have $P(X = 0) = \binom{6}{1}\binom{4}{2}/\binom{10}{3} = \frac{3}{10}$, $P(X = 10) = \binom{4}{3}/\binom{10}{3} = \frac{1}{30}$, and $P(X = -1) = 1 - P(X = 0) - P(X = 10) = \frac{2}{3}$. Thus

$$E(X) = -1 \times \frac{2}{3} + 0 \times \frac{3}{10} + 10 \times \frac{1}{30} = -\frac{1}{3}.$$

3.19 Let the random variable X be the payoff when you go for a second spin given that the first spin showed a score of a points. Then, $P(X = a + k) = 1/1{,}000$ for $1 \leq k \leq 1{,}000 - a$ and $P(X = 0) = a/1{,}000$. Thus

$$E(X) = \frac{1}{1{,}000} \sum_{k=1}^{1{,}000-a} (a + k) = \frac{1}{2{,}000}(1{,}000 - a)(1{,}000 + a + 1).$$

The largest value of a for which $E(X) > a$ is $a^* = 414$. The optimal strategy is to stop after the first spin if this spin gives a score of more than 414 points.

3.20 The expected value of the number of fiches you leave the casino is $U(a, b) = b \times P(a, b) + 0 \times (1 - P(a, b))$. Since $P(a,\, b) \approx (q/p)^{-b}$ for a large, we get

$$U(a,\, b) \approx b \times (q/p)^{-b}.$$

It is matter of simple algebra to verify that $b \times (q/p)^{-b}$ is maximal for $b \approx 1/\ln(q/p)$. This gives

$$P\Big(a,\, \frac{1}{\ln(q/p)}\Big) \approx \frac{1}{e} \;\text{ and }\; U\Big(a,\, \frac{1}{\ln(q/p)}\Big) \approx \frac{b}{e}.$$

3.21 Let the random variable X be the payoff of the game. Then, using conditional probabilities, $P(X = 0) = (\frac{5}{6})^3$, $P(X = 2) = \binom{3}{1}\frac{1}{6}(\frac{5}{6})^2 \times \frac{4}{5}$,

$P(X = 2.5) = \binom{3}{1}\frac{1}{6}(\frac{5}{6})^2 \times \frac{1}{5}$, $P(X = 3) = \binom{3}{2}(\frac{1}{6})^2\frac{5}{6}$, and $P(X = 4) = (\frac{1}{6})^3$. This gives $E(X) = 0\times\frac{125}{216}+2\times\frac{60}{216}+2.5\times\frac{15}{216}+3\times\frac{15}{216}+4\times\frac{1}{216} = 0.956$.

3.22 Let the random variable X be the payoff in the dice game. The random variable X can take on the values 0, 10, and 100. We have $P(X = 100) = 6 \times (\frac{1}{6})^4 = \frac{1}{216}$, $P(X = 10) = 3 \times 3 \times \binom{4}{2} \times (\frac{1}{6})^4 = \frac{1}{24}$, and $P(X = 0) = 1 - P(X = 100) - P(X = 10) = \frac{206}{216}$. Therefore $E(X) = \frac{100}{216} + \frac{10}{24} = 0.8796$. The dice game is unfavorable to you. Let Y be the payoff in the coin-tossing game. Then $P(Y = 0) = \frac{1}{2}$, $P(Y = i) = (\frac{1}{2})^{i+1}$ for $1 \leq i \leq 4$, and $P(Y = 30) = (\frac{1}{2})^5$. Therefore $E(Y) = \sum_{i=1}^{4} i \times (\frac{1}{2})^{i+1} + 30 \times (\frac{1}{2})^5 = 1.75$. The coin-tossing game is also unfavorable to you.

3.23 Let the random variable X be the payoff of the game. The probability that X takes on the value k with $k < m$ is the same as the probability that a randomly chosen number from the interval $(0, 1)$ falls into the subinterval $(\frac{1}{k+2}, \frac{1}{k+1})$ or into the subinterval $(1 - \frac{1}{k+1}, 1 - \frac{1}{k+2})$. Thus $P(X = k) = 2\big(\frac{1}{k+1} - \frac{1}{k+2}\big)$ for $1 \leq k \leq m - 1$ and $P(X = m) = \frac{2}{m+1}$. The stake should be $E(X) = 2\big(\frac{1}{2} + \cdots + \frac{1}{m+1}\big)$.

3.24 Suppose that your current total is i points. If you decide to roll again the die and then to stop, the expected value of the change of your total is

$$\frac{1}{6}\sum_{k=2}^{6} k - \frac{1}{6} \times i = \frac{20}{6} - \frac{i}{6}.$$

Therefore the one-stage-look-ahead rule prescribes to stop as soon as your total is 20 points or more. This stopping rule is optimal.

3.25 Suppose you have rolled a total of $i < 10$ points so far. The expected value of the change of your current total is

$$\sum_{k=1}^{10-i} k \times \frac{1}{6} - i \times \frac{i-4}{6} = \frac{1}{12}(10-i)(10-i+1) - \frac{1}{6}\,i(i-4)$$

if you decide to continue for one more roll. This expression is positive for $i \leq 5$ and negative for $i \geq 6$. Thus, the one-stage-look-ahead rule prescribes to stop as soon as the total number of points rolled is 6 or more. This rule maximizes the expected value of your reward.

3.26 Suppose that at a given moment there are i_0 empty bins and i_1 bins with exactly one ball (and $25 - i_0 - i_1$ bins with two or more balls). If you decide to drop one more ball before stopping rather than stopping immediately, the expected value of the change of your net winnings is $\frac{i_0}{25} \times 1 - \frac{i_1}{25} \times 1.50$. The one-stage-look-ahead rule prescribes to stop in the states (i_0, i_1) with $i_0 - 1.5i_1 \leq 0$ and to continue otherwise. For the case that you lose $\frac{1}{2}k$ dollars for every containing $k \geq 2$ balls, the expected value of the change of your net winnings is

$$\frac{i_0}{25} \times 1 - \frac{i_1}{25} \times 2 - \frac{25 - i_0 - i_1}{25} \times 0.50$$

when you decide to drop one more ball before stopping rather than stopping immediately. Therefore, the one-stage-look-ahead rule prescribes to stop in the states (i_0, i_1) with $3i_0 - 3i_1 - 25 \leq 0$ and to continue otherwise.

3.27 Suppose that you have gathered so far a dollars. Then there are still $w - a$ white balls in the bowl. If you decide to pick one more ball, then the expected change of your bankroll is

$$\frac{1}{r + w - a}(w - a) - \frac{r}{r + w - a}a.$$

This expression is less than or equal to zero for $a \geq \frac{w}{r+1}$. It is optimal to stop as soon as you have gathered at least $\frac{w}{r+1}$ dollars.

3.28 Suppose your current total is i points. If you decide to continue for one more roll, then the expected value of the change of your dollar value is

$$3 \times \frac{2}{36} + 4 \times \frac{2}{36} + 5 \times \frac{4}{36} + 6 \times \frac{4}{36} + 7 \times \frac{6}{36} + 8 \times \frac{4}{36} + 9 \times \frac{4}{36}$$
$$+ 10 \times \frac{2}{36} + 11 \times \frac{2}{36} - \frac{1}{6} \times i = \frac{210}{36} - \frac{i}{6}.$$

The smallest value of i for which the expected change is less than or equal to 0 is $i = 35$. The one-stage-look-ahead rule prescribes to stop as soon you have gathered 35 or more points. This rule is optimal among all conceivable stopping rules.
Note: The maximal expected reward is 14.22 dollars, see Problem 7.52.

3.29 (**a**) Writing $\sum_{k=0}^{\infty} P(X > k) = \sum_{k=0}^{\infty} \sum_{j=k+1}^{\infty} P(X = j)$ and interchanging the order of summation give

$$\sum_{k=0}^{\infty} P(X > k) = \sum_{j=0}^{\infty} \sum_{k=0}^{j-1} P(X = j) = \sum_{j=0}^{\infty} jP(X = j)$$

and so $\sum_{k=0}^{\infty} P(X > k) = E(X)$. The interchange of the order of summation is justified by the nonnegativity of the terms involved.
(**b**) Let X be the largest among the 10 random numbers. Then, $P(X > 0) = 1$ and $P(X > k) = 1 - \left(\frac{k}{100}\right)^{10}$ for $1 \leq k \leq 99$. Thus, $E(X) = \sum_{k=0}^{99} P(X > k) = 91.4008$.

3.30 Define the random variable X as the number of floors on which the elevator will stop. Let the random variable $X_j = 1$ if the elevator does not stop on floor j and $X_j = 0$ otherwise. Then $X = r - \sum_{j=1}^{r} X_j$. We have $P(X_j = 1) = \left(\frac{r-1}{r}\right)^m$ and so $E(X_j) = \left(\frac{r-1}{r}\right)^m$ for all for $j = 1, 2, \ldots, r$. Hence

$$E(X) = r - \sum_{j=1}^{r} E(X_j) = r\Big(1 - \Big(\frac{r-1}{r}\Big)^m\Big).$$

Note: This problem is an instance of the balls-and-bins model from Example 3.8.

3.31 Let $I_k = I(A_k)$. Then $P(A_1^c \cap \cdots \cap A_n^c) = E[(1 - I_1) \cdots (1 - I_n)]$. We have that

$$(1 - I_1) \cdots (1 - I_n) = 1 - \sum_{j=1}^{n} I_j + \sum_{j<k} I_j I_k + \cdots + (-1)^n I_1 \cdots I_n.$$

Taking the expected value of both sides, noting that $E(I_{i_1} \cdots I_{i_r})$ is given by $P(A_{i_1} \cdots A_{i_r})$ and using $P\big(\bigcup_{k=1}^{n} A_k\big) = 1 - P\big(\bigcap_{k=1}^{n} A_k^c\big)$, we get the inclusion-exclusion formula.

3.32 Let the random variable X be the number of times that two adjacent letters are the same in a random permutation of the word Mississippi. Then, X can be represented as $X = \sum_{j=1}^{10} X_j$, where the random variable X_j equals 1 if the letters j and $j + 1$ are the same in the random permutation and equals 0 otherwise. By numbering the eleven letters of the word Mississippi as $1, 2, \ldots, 11$, it is easily seen that

$$P(X_j = 1) = \frac{4 \times 3 \times 9! + 4 \times 3 \times 9! + 2 \times 1 \times 9!}{11!} = \frac{26}{110}.$$

for all j (alternatively, using conditional probabilities, $P(X_j = 1)$ can be calculated as $\frac{4}{11} \times \frac{3}{10} + \frac{4}{11} \times \frac{3}{10} + \frac{2}{11} \times \frac{1}{10} = \frac{26}{110}$). Hence, using the linearity of the expectation operator,

$$E(X) = \sum_{j=1}^{10} E(X_j) = 10 \times \frac{26}{110} = 2.364.$$

3.33 Let the indicator variable I_k be equal to 1 if the kth team has a married couple and zero otherwise. Then $P(I_k = 1) = (12 \times 22)/\binom{24}{3} = \frac{3}{23}$ for any k. The expected number of teams with a married couple is $\sum_{k=1}^{8} E(I_k) = \frac{24}{23}$.
Note: As a sanity check, the probability that a given team has no married couple is $\frac{22}{23} \times \frac{20}{22} = \frac{20}{23}$, by using the chain rule for conditional probabilities.

3.34 It is no restriction to assume that n is even. Let the indicator variable I_{2j} be 1 if the random walk returns to the zero level at time $2j$ and 0 otherwise. Then, by the linearity of the expectation operator, the expected number of returns to the zero level during the first n time units is $\sum_{j=1}^{n/2} E(I_{2j})$. Let $p_j = P(I_{2j} = 1)$, then $E(I_{2j}) = p_j$. In Example 1.4, it is shown that $p_j \approx 1/\sqrt{\pi j}$ for j large. Next, we get $\sum_{j=1}^{n/2} E(I_{2j}) \approx \sqrt{2/\pi}\sqrt{n}$ for n large.

3.35 Let I_k be 1 if two balls of the same color are removed on the kth pick and 0 otherwise. The expected number of times that you pick two balls of the same color is $\sum_{k=1}^{r+b} E(I_k)$. By a symmetry argument, each I_k has the same distribution as I_1 (the order in which you draw the pairs of balls does not matter, that is, for all practical purposes the kth pair can be considered as the first pair). Since

$$P(I_1 = 1) = \frac{2r}{2r+2b} \times \frac{2r-1}{2r+2b-1} + \frac{2b}{2r+2b} \times \frac{2b-1}{2r+2b-1},$$

the expected number of times that you pick two balls of the same color is $[r(2r-1) + b(2b-1)]/(2r+2b-1)$.

3.36 (**a**) Let X_i be equal to 1 if there is a birthday on day i and 0 otherwise. For each i, $P(X_i = 0) = (364/365)^{100}$ and $P(X_i = 1) = 1 - P(X_i = 0)$. The expected number of distinct birthdays is

$$E\Big(\sum_{i=1}^{365} X_i\Big) = 365 \times [1 - (364/365)^{100}] = 87.6.$$

(**b**) Let X_i be equal to 1 if some child in the second class shares the birthday of the ith child in the first class and X_i is zero otherwise. Then $P(X_i = 1) = 1 - (364/365)^s$ for all i. The expected number of children in the first class sharing a birthday with some child in the other class is

$$E\Big(\sum_{i=1}^{r} X_i\Big) = r \times [1 - (364/365)^s].$$

3.37 Let the indicator variable I_s be 1 if item s belongs to T after n steps and 0 otherwise. Then $P(I_s = 0) = \left(\frac{n-1}{n}\right)^n$ and so $E(I_s) = 1 - \left(1 - \frac{1}{n}\right)^n$. Thus the expected value of the number of distinct items in T after n steps is $n\left[1 - \left(1 - \frac{1}{n}\right)^n\right]$. Note that the expected value is about $n\left(1 - \frac{1}{e}\right)$ for n large.

3.38 Let the random variable X_i be 1 if the ith person survives the first round and X_i be zero otherwise. The random variable X_i takes on the value 1 if and only if nobody of the other $n - 1$ persons shoots at person i. Hence

$$P(X_i = 1) = \frac{n-2}{n-1} \times \cdots \times \frac{n-2}{n-1} = \left(\frac{n-2}{n-1}\right)^{n-1}$$

for all $1 \leq i \leq n$. The expected value of the number of people who survive the first round is

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

This expected value can be approximated by $\frac{n}{e}$ for n large, where $e = 2.71828\ldots$.

3.39 Label the white balls as $1, \ldots, w$. Let the indicator variable I_k be equal to 1 if the white ball with label k remains in the bag when you stop and 0 otherwise. To find $P(I_k = 1)$, you can discard the other white balls. Therefore $P(I_k) = \frac{1}{r+1}$. The expected number of remaining white balls is $\sum_{k=1}^{w} E(I_k) = \frac{w}{r+1}$.

3.40 Number the 25 persons as $1, 2, \ldots, 25$ and let person 1 be the originator of the rumor. Let the random variable X_k be equal to 1 if person k hears about the rumor for $2 \leq k \leq 25$. For fixed k with $2 \leq k \leq 25$, let A_j be the event that person k hears about the rumor for the first time when the rumor is told the jth time. Then, $E(X_k) = P(X_k = 1) = \sum_{j=1}^{10} P(A_j)$, where

$$P(A_1) = \frac{1}{24} \quad \text{and} \quad P(A_j) = \left(1 - \frac{1}{24}\right)\left(1 - \frac{1}{23}\right)^{j-2}\frac{1}{23}, \quad 2 \leq j \leq 10.$$

This gives

$$E(X_k) = \frac{1}{24} + \sum_{j=2}^{10}\left(1 - \frac{1}{24}\right)\left(1 - \frac{1}{23}\right)^{j-2}\frac{1}{23} = 0.35765$$

for $2 \leq k \leq 25$. Hence the expected value of the number of persons who know about the rumor is

$$1 + \sum_{k=2}^{25} E(X_k) = 1 + 24 \times 0.35765 = 9.584.$$

3.41 Let the indicator variable I_k be equal to 1 if the numbers k and $k+1$ appear in the lotto drawing and 0 otherwise. Then, $P(I_k = 1) = \binom{43}{4}/\binom{45}{6}$. The expected number of consecutive numbers is $\sum_{k=1}^{44} E(I_k) = \frac{2}{3}$.

3.42 For any $k \geq 2$, let X_k be the amount you get at the kth game. Then the total amount you will get is $\sum_{k=2}^{s} X_k$. By the linearity of the expectation operator, the expected value of the total amount you will get is $\sum_{k=2}^{s} E(X_k)$. Since $P(X_k = 1) = \frac{1}{k(k-1)}$ and $P(X_k = 0) = 1 - P(X_k = 1)$, it follows that $E(X_k) = \frac{1}{k(k-1)}$ for $k = 2, \ldots, s$. Hence the expected value of the total amount you will get is equal to

$$\sum_{k=2}^{s} \frac{1}{k(k-1)} = \frac{s-1}{s}.$$

The fact that the sum equals $\frac{s-1}{s}$ is easily verified by induction.

3.43 Let the random variable X_i be equal to 1 if box i contains more than 3 apples and X_i be equal to 0 otherwise. Then,

$$P(X_i = 1) = \sum_{k=4}^{25} \binom{25}{k} \left(\frac{1}{10}\right)^k \left(\frac{9}{10}\right)^{25-k} = 0.2364.$$

and so $E(X_i) = 0.2364$. Thus the expected value of the number of boxes containing more than 3 apples is given by $E\big(\sum_{k=1}^{10} X_k\big) = 10 \times 0.2364 = 2.364$.

3.44 In Problem 1.73 the reader was asked to prove that the probability that a particular number r belongs to a cycle of length k is $\frac{1}{n}$, regardless of the value of k (this result can also be proved by using conditional probabilities: $\frac{n-1}{n} \times \frac{n-2)}{n-1} \times \cdots \times \frac{n-k+2}{n-k+1} \times \frac{1}{n-k+1} = \frac{1}{n}$). Thus, for fixed k, $E(X_r) = \frac{1}{n}$ for $r = 1, \ldots, n$. Therefore the expected number of cycles of length k is

$$\frac{1}{k} E\Big(\sum_{r=1}^{n} X_r\Big) = \frac{1}{k} \sum_{r=1}^{n} E(X_r) = \frac{1}{k}$$

for any $1 \leq k \leq n$. This shows that the expected value of the total number of cycles is

$$\sum_{k=1}^{n} \frac{1}{k}.$$

This sum can be approximated by $\ln(n) + \gamma$ for n sufficiently large, where $\gamma = 0.57722\ldots$ is Euler's constant.

3.45 For any $i \neq j$, let $X_{ij} = 1$ if the integers i and j are switched in the random permutation and $X_{ij} = 0$ otherwise. Then $P(X_{ij} = 1) = \frac{(n-2)!}{n!}$. The expected number of interchanges is

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E(X_{ij}) = \binom{n}{2} \frac{1}{n(n-1)} = \frac{1}{2}.$$

3.46 Let X be the payoff in dollars for investment A and Y be the payoff in dollars for investment B . Then

$$E(X) = 0.20 \times 1{,}000 + 0.40 \times 2{,}000 + 0.40 \times 3{,}000 = 2{,}200.$$

Also, we have

$$E(X^2) = 0.20 \times 1{,}000^2 + 0.40 \times 2{,}000^2 + 0.40 \times 3{,}000^2 = 5{,}400{,}000.$$

Thus $\sigma^2(X) = 5{,}400{,}000 - 2{,}200^2 = 560{,}000$ and so $\sigma(X) = 748.33$. In the same way, we get $E(Y) = 2{,}200$, $E(Y^2) = 5{,}065{,}000$, $\sigma^2(Y) = 225{,}000$, and $\sigma(Y) = 474.34$.

3.47 Using the basic sums $\sum_{k=1}^{n} k = \frac{1}{2}n(n+1)$ and $\sum_{k=1}^{n} k^2 = \frac{1}{6}n(n+1)(2n+1)$, it follows that $E(X)$ and $E(X^2)$ are $\frac{1}{2}(a+b)$ and $\frac{1}{6}(2a^2 + 2ab - a + 2b^2 + b)$, which leads to $\text{var}(X) = \frac{1}{12}(a^2 - 2ab - 2a + b^2 + 2b)$.

3.48 Since $E(X) = pa + (1-p)b$ and $E(X^2) = pa^2 + (1-p)b^2$, we have $\text{var}(X) = pa^2 + (1-p)b^2 - \big(pa + (1-p)b\big)^2$. It is a matter of simple algebra to get $\text{var}(X) = p(1-p)(a-b)^2$. Then, by noting that $p(1-p)$ is maximal for $p = \frac{1}{2}$, the desired result follows.
Note: We have $E(X) \leq \frac{1}{4}(a-b)^2$ for any discrete random variable X that is concentrated on the integers $a, a+1, \ldots, b$.

3.49 Let I_k be 1 if the kth team has a married couple and 0 otherwise. Let $X = \sum_{k=1}^{8} I_k$. Then $E(X^2) = \sum_{k=1}^{8} E(I_k^2) + 2\sum_{j=1}^{7}\sum_{k=j+1}^{8} E(I_j I_k)$.

We have

$$E(I_k^2) = P(I_k = 1) = \frac{12 \times 22}{\binom{24}{3}} \quad \text{for all } k$$

$$E(I_j I_k) = P(I_j = 1,\, I_k = 1) = \frac{12 \times 11 \times 20 \times 19}{\binom{24}{3}\binom{21}{3}} \quad \text{for all } j \neq k.$$

This gives $E(X) = \frac{24}{23}$, $E(X^2) = \frac{48}{23}$, and $\sigma(X) = 0.9981$.

3.50 Define X as the number of integers that do not show up in the 15 lotto drawings. Let $X_i = 1$ if the number i does not show up in the 15 lotto drawings and $X_i = 0$ otherwise. Then, $X = \sum_{i=1}^{45} X_i$. The probability that a specified number does not show up in any given drawing is $\binom{44}{6}/\binom{45}{6} = 39/45$. Hence $E(X_i) = P(X_i = 1) = (39/45)^{15}$ and so

$$E(X) = 45 \times \left(\frac{39}{45}\right)^{!15} = 5.2601.$$

The probability that two specified numbers i and j with $i \neq j$ do not show up in any given drawing is $\binom{43}{6}/\binom{45}{6} = (39/45) \times (38/44)$. Hence $E(X_i X_j) = P(X_i = 1, X_j = 1) = [(39 \times 38)/(45 \times 44)]^{15}$ and so

$$E(X^2) = 45 \times \left(\frac{39}{45}\right)^{15} + 2\binom{45}{2} \times \left(\frac{39 \times 38}{45 \times 44}\right)^{15} = 30.9292.$$

This leads to $\sigma(X) = 1.8057$.

3.51 By the substitution rule, we have

$$E(X) = \sum_{k=1}^{10} k\, \frac{11-k}{55} = 4 \text{ and } E(X^2) = \sum_{k=1}^{10} k^2\, \frac{11-k}{55} = 22,$$

and so $\sigma(X) = \sqrt{22 - 16} = 2.449$. The number of reimbursed treatments is $Y = g(X)$, where $g(x) = \min(x, 5)$. Then, by applying the substitution rule,

$$E(Y) = \sum_{k=1}^{4} k\, \frac{11-k}{55} + \sum_{k=5}^{10} 5\, \frac{11-k}{55} = \frac{37}{11}$$

$$E(Y^2) = \sum_{k=1}^{4} k^2\, \frac{11-k}{55} + \sum_{k=5}^{10} 25\, \frac{11-k}{55} = \frac{151}{11}.$$

The standard deviation $\sigma(Y) = \sqrt{151/11 - (37/11)^2} = 1.553$.

3.52 Let V be the stock left over and W be the amount of unsatisfied demand. We have $V = g_1(X)$ and $W = g_2(X)$, where the functions $g_1(x)$ and $g_2(x)$ are given by $g_1(x) = Q - x$ for $x < Q$ and $g_1(x) = 0$ otherwise, and $g_2(x) = x - Q$ for $x > Q$ and $g_1(x) = 0$ otherwise. By the substitution rule,

$$E(V) = \sum_{k=0}^{Q-1}(Q-k)p_k \quad \text{and} \quad E(W) = \sum_{k=Q+1}^{\infty}(k-Q)p_k.$$

Note: Writing $\sum_{k=Q+1}^{\infty}(k-Q)p_k$ as $\sum_{k=0}^{\infty}(k-Q)p_k - \sum_{k=0}^{Q}(k-Q)p_k$, it follows that $E(W) = \mu - Q + E(V)$, where μ is the expected demand.

3.53 Let the random variable X be the number of repairs that will be necessary in the coming year and Y be the maintenance costs in excess of the prepaid costs. Then $Y = g(X)$, where the function $g(x) = 100(x - 155)$ for $x > 155$ and $g(x) = 0$ otherwise. By the substitution rule,

$$E(Y) = \sum_{x=156}^{\infty} 100(x-155)e^{-150}\frac{150^x}{x!} = 280.995$$

$$E(Y^2) = \sum_{x=156}^{\infty} 100^2(x-155)^2 e^{-150}\frac{150^x}{x!} = 387{,}929.$$

The standard deviation of Y is $\sqrt{E(Y^2) - E^2(Y)} = 555.85$.

3.54 Let the random variable X be the monthly demand for the medicine and Y be the net profit in any given month. Then $Y = g(X)$, where the function $g(x)$ is given by

$$g(x) = \begin{cases} 400x - 800 & \text{for } 3 \leq x \leq 8 \\ 400x - 800 - 350(x-8) & \text{for } x > 8. \end{cases}$$

By the substitution rule,

$$E(Y) = \sum_{x=3}^{10} g(x)\, P(X = x) \quad \text{and} \quad E(Y^2) = \sum_{x=3}^{10}[g(x)]^2\, P(X = x).$$

The standard deviation of Y is $\sqrt{E(Y^2) - E^2(Y)}$. Substituting the values of $g(x)$ and $P(X = x)$, it follows after some calculations that the expected value and the standard deviation of the monthly net profit $Y = g(X)$ are given by \$1227.50 and \$711.24.

3.55 Your probability of winning the contest is

$$\sum_{k=0}^{n} \frac{1}{k+1} P(X = k) = E\Big(\frac{1}{1+X}\Big),$$

by the law of conditional probability. The random variable X can be written as $X = X_1 + \cdots + X_n$, where X_i is equal to 1 if the ith person in the first round survives this round and 0 otherwise. Since $E(X_i) = \frac{1}{n}$ for all i, we have $E(X) = 1$. Thus, by Jensen's inequality,

$$E\Big(\frac{1}{1+X}\Big) \geq \frac{1}{1+E(X)} = \frac{1}{2}.$$

3.56 By the linearity of the expectation operator, $E[(X_1+\cdots+X_n)^2+(Y_1+\cdots+Y_n)^2]$ is equal to $E[(X_1+\cdots+X_n)^2]+E[(Y_1+\cdots+Y_n)^2]$. Using the algebraic formula, $(a_1+\cdots+a_n)^2 = \sum_{i=1}^n a_i^2 + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^n a_i a_j$ and again the linearity of the expectation operator, it follows that

$$\begin{aligned} E[(X_1 + \cdots + X_n)^2] &= \sum_{i=1}^{n} E(X_i^2) + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} E(X_i X_j) \\ &= \sum_{i=1}^{n} E(X_i^2) + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} E(X_i)E(X_j), \end{aligned}$$

where the last equality uses the fact that $E(X_i X_j) = E(X_i)E(X_j)$ by the independence of X_i and X_j for $i \neq j$. For each i, $E(X_i) = 1 \times \frac{1}{4} + (-1) \times \frac{1}{4} = 0$ and $E(X_i^2) = 1^2 \times \frac{1}{4} + (-1)^2 \times \frac{1}{4} = \frac{1}{2}$. This gives

$$E[(X_1 + \cdots + X_n)^2] = \frac{1}{2}n.$$

In the same way, $E[(Y_1+\cdots+Y_n)^2] = \frac{1}{2}n$. Hence we find the interesting result that the expected value of the squared distance between the drunkard's position after n steps and his starting position is equal to n for any value of n.

Note: It is not true that the expected value of the distance between the drunkard's position after n steps and his starting position equals $\sqrt{n}$. Otherwise, the variance of the distance would be zero and so the distance would exhibit no variability, but this cannot be true.

3.57 For $x, y \in \{-1, 1\}$, we have $P(X = x, Y = y) = P(X = x \mid Y = y)P(Y = y)$ and $P(X = x \mid Y = y) = P(Z = x/y \mid Y = y)$. Since Y

and Z are independent, $P(Z = x/y \mid Y = y) = P(Z = x/y) = 0.5$. This gives

$$P(X = x,\, Y = y) = 0.5 \times P(Y = y) \quad \text{for all } x, y \in \{-1, 1\}.$$

Also, $P(X = 1) = P(Y = 1,\, Z = 1) + P(Y = -1,\, Z = -1)$. Thus, by the independence of Y and Z, we get $P(X = 1) = 0.25 + 0.25 = 0.5$ and so $P(X = -1) = 0.5$. Therefore, the result $P(X = x,\, Y = y) = 0.5 \times P(Y = y)$ implies $P(X = x,\, Y = y) = P(X = x)P(Y = y)$, proving that X and Y are independent. However, X is not independent of $Y + Z$. To see this, note that $P(X = 1,\, Y + Z = 0) = 0$ and $P(X = 1)P(Y + Z = 0) > 0$.

3.59 Noting that $X_i = X_{i-1} + R_i$, we get $X_i = R_2 + \cdots + R_i$ for $2 \le i \le 10$. This implies $\sum_{i=2}^{10} X_i = \sum_{k=2}^{10}(11-k)R_k$. Since $P(R_k = 0) = P(R_k = 1) = \frac{1}{2}$, we have $E(R_k) = \frac{1}{2}$ and $\sigma^2(R_k) = \frac{1}{4}$. The random variables R_k are independent and so, by the Rules 3.1 and 3.9,

$$E\Big(\sum_{i=2}^{10} X_i\Big) = \sum_{k=2}^{10}(11-k)E(R_k) = 22.5$$
$$\sigma^2\Big(\sum_{i=1}^{10} X_i\Big) = \sum_{k=2}^{10}(11-k)^2\sigma^2(R_k) = 71.25.$$

3.60 We have $P(X+Y = k) = \sum_{j=1}^{k-1}(1-p)^j p\,(1-p)^{k-j-1}p$ for $k = 2, 3, \ldots$, by the convolution formula. This leads to

$$P(X + Y = k) = (k-1)p^2(1-p)^{k-2} \quad \text{for } k = 2, 3, \ldots .$$

3.61 We have $E\big(\sum_{k=1}^{\infty} X_k I_k\big) = \sum_{k=1}^{\infty} E(X_k I_k)$, since it is always allowed to interchange expectation and summation for nonnegative random variables. Since X_k and I_k are independent, $E(X_k I_k) = E(X_k)E(I_k)$ for any $k \ge 1$. Also, $E(I_k) = P(N \ge k)$. Thus,

$$E\Big(\sum_{k=1}^{\infty} X_k I_k\Big) = E(X_1)\sum_{k=1}^{\infty} P(N \ge k) = E(X_1)E(N),$$

using the fact that $E(N) = \sum_{n=0}^{\infty} P(N > n)$, see Problem 3.29.
Note: If the X_k are not nonnegative, the proof still applies in view of the fact that $E\big(\sum_{k=1}^{\infty} Y_k\big) = \sum_{k=1}^{\infty} E(Y_k)$ when $\sum_{k=1}^{\infty} E(|Y_k|) < \infty$. The proof of this so-called dominated convergence result can be found in advanced texts.

3.62 Let the random variable X be the number of passengers showing up. Then X is binomially distributed with parameters $n = 160$ and $p = 0.9$. The overbooking probability is

$$\sum_{k=151}^{160} \binom{160}{k} 0.9^k (0.1)^{160-k} = 0.0359.$$

Denote by the random variable R the daily return. Using the substitution rule, the expected value of R is calculated as

$$\sum_{k=0}^{150} \big[75k + 37.5(160-k)\big] \binom{160}{k} 0.9^k (0.1)^{160-k} + \sum_{k=151}^{160} \big[75 \times 150$$
$$+37.5(160-k) - 425(k-150)\big] \binom{160}{k} 0.9^k (0.1)^{160-k}.$$

This gives $E(R) = \$11367.63$. Also, by the substitution rule, $E(R^2)$ is calculated as

$$\sum_{k=0}^{150} \big[75k + 37.5(160-k)\big]^2 \binom{160}{k} 0.9^k (0.1)^{160-k} \sum_{k=151}^{160} \big[75 \times 150$$
$$+37.5(160-k) - 425(k-150)\big]^2 \binom{160}{k} 0.9^k (0.1)^{160-k}.$$

Hence the standard deviation of R equals $\sqrt{E(R^2) - E^2(R)}$ =\$194.71.

3.63 The probability

$$P_r = \sum_{k=r}^{6r} \binom{6r}{k} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{6r-k}$$

of getting at least r sixes in one throw of $6r$ dice has the values 0.6651, 0.6187, and 0.5963 for $r = 1$, 2, and 3. Thus it is best to throw 6 dice. Pepys believed that it was best to throw 18 dice.
Note: The probability P_r is decreasing in r and tends to $\frac{1}{2}$ as $r \to \infty$.

3.64 If the competition would have been continued, the Yankees would have won with probability $\sum_{k=2}^{7} \binom{7}{k} 0.5^k \, 0.5^{7-k} = \frac{11}{16}$. The prize money should be divided between the Yankees and the Mets according to the proportion 11:5.

3.65 This question can be translated into the question what the probability is of getting 57 *or less* heads in 199 tosses of a fair coin. A binomial random variable with parameters $n = 199$ and $p = 0.5$ has expected value $199 \times 0.5 = 99.5$ and standard deviation $0.5\sqrt{199} = 7.053$. Thus the observed number of polio cases in the treatment group is more than six standard deviations below the expected number. Without doing any further calculations, we can say that the probability of this occurring is extremely small (the precise value of the probability is 7.4×10^{-10}). This makes clear that there is overwhelming evidence that the vaccine does work.

3.66 Let X be the number of beans that will come up white and Y be the number of point gained by the bean thrower. Then, $P(X = k) = \binom{8}{k} 0.5^8$ for $k = 0, 1, \ldots, 8$. We have $P(Y = 1) = \sum_{k=0}^{3} P(X = 2k + 1) = 0.5$, $P(Y = 2) = P(X = 0) + P(X = 8) = \frac{1}{128}$, and $P(Y = -1) = 1 - P(Y = 1) - P(Y = 2) = \frac{63}{128}$. This gives $E(Y) = \frac{3}{128}$. Thus the bean thrower has a slight advantage.

3.67 Using the law of conditional probability, we have that the sought probability is given by

$$\sum_{k=0}^{n} \frac{k}{n} \binom{n}{k} p^k (1-p)^{n-k}.$$

This probability is nothing else than $\frac{1}{n}E(X)$, where X is binomially distributed with parameters n and p. Thus the probability that you will be admitted to the program is $\frac{np}{n} = p$.

3.68 Let X and Y be the numbers of successful penalty kicks of the two teams. The independent random variables X and Y are binomially distributed with parameters $n = 5$ and $p = 0.7$. The probability of a tie is $\sum_{k=0}^{5} P(X = k, Y = k$. Using the independence of X and Y, this probability can be evaluated as

$$\sum_{k=0}^{5} \binom{5}{k} 0.7^k \, 0.3^{5-k} \times \binom{5}{k} 0.7^k \, 0.3^{5-k} = 0.2716.$$

3.69 Let the random variable X be the number of coins that will be set aside. Then the random variable $100 - X$ is binomially distributed with parameters $n = 100$ and $p = \frac{1}{8}$. Therefore $P(X = k) = \binom{100}{100-k} (\frac{1}{8})^{100-k} (\frac{7}{8})^k$ for $k = 0, 1, \ldots, 100$.

3.70 Using only the expected value and the standard deviation of the binomial distribution, you can see that is highly unlikely that the medium

has to be paid out. The expected value and the standard deviation of the binomial distribution with parameters $n = 250$ and $p = \frac{1}{5}$ are 50 and $\sqrt{250 \times \frac{1}{5} \times \frac{4}{5}} = 6.32$. Thus the requirement of 82 or more correct answers means more than five standard deviations above the expected value, which has a negligible probability.
Note: The psychic who took the challenge of the famous scientific skeptic James Randi was able to get only fifty predictions correct.

3.71 Let the random variable X be the number of rounds you get cards with an ace. If the cards are well-shuffled each time, then X is binomially distributed with parameters $n = 10$ and $p = 1 - \binom{48}{13}/\binom{52}{13}$. The answer to the question should be based on

$$P(X \leq 2) = \sum_{k=0}^{2} \binom{10}{k} p^k (1-p)^{10-k} = 0.0017.$$

This small probability is a strong indication that the cards were not well-shuffled.

3.72 By the same argument as used for the binomial distribution, $P(X_1 = x_1, X_2 = x_2, \ldots, X_r = x_r)$ is given by

$$\binom{n}{x_1}\binom{n-x_1}{x_2}\cdots\binom{n-x_1-\cdots-x_{r-1}}{x_r} p_1^{x_1}\cdots p_r^{x_r}.$$

Using this expression, the result follows.

3.73 Let X_i be the number of times that image i shows up in the roll of the five poker dice. Then $(X_1, X_2, \ldots, X_6)$ has a multinomial distribution with parameters $n = 5$ and $p_1 = p_2 = \cdots = p_6 = \frac{1}{6}$. Let X be the payoff to the player for each unit staked. Then $E(X) = 3 \times P(X_1 \geq 1,\ X_2 \geq 1)$. This gives

$$\begin{aligned} E(X) &= 3 \sum_{x_1=1}^{4} \sum_{x_2=1}^{5-x_1} \frac{5!}{x_1! x_2! (5-x_1-x_2)!} p_1^{x_1} p_2^{x_2} (1-p_1-p_2)^{5-x_1-x_2} \\ &= 0.9838. \end{aligned}$$

3.74 The number of winners in any month has a binomial distribution with parameters $n = 200$ and $p = \frac{50}{450{,}000}$. This distribution can be very well approximated by a Poisson distribution with an expected value of $\lambda = 200 \times \frac{50}{450{,}000} = \frac{1}{45}$. The monthly amount the corporation will have to give away is \$0 with probability $e^{-\lambda} = 0.9780$, \$25,000 with probability $e^{-\lambda}\lambda = 0.0217$ and \$50,000 with probability $e^{-\lambda}\frac{\lambda^2}{2!} = 2.4 \times 10^{-4}$.

3.75 Let the random variable X be the number of king's rolls. Then X has a binomial distribution with parameters $n = 4 \times 6^{r-1}$ and $p = \frac{1}{6^r}$. This distribution converges to a Poisson distribution with expected value $np = \frac{2}{3}$ as $r \to \infty$. The binomial probability $1 - (1-p)^n$ tends very fast to the Poisson probability $1 - e^{-2/3} = 0.48658$ as $r \to \infty$. The binomial probability has the values 0.51775, 0.49140, 0.48738, 0.48660, and 0.48658 for $r = 1$, 2, 3, 5, and 7.

3.76 An appropriate model is the Poisson model. We have the situation of 500 independent trials each having a success probability of $\frac{1}{365}$. The number of marriages having the feature that both partners are born on the same day is approximately distributed as a Poisson random variable with expected value $\frac{500}{365} = 1.3699$. The approximate Poisson probabilities are 0.2541, 0.3481, 0.2385, 0.1089, 0.0373, and 0.0102 for 0, 1, 2, 3, 4, and 5 matches, while the exact binomial probabilities are 0.2537, 0.3484, 0.2388, 0.1089, 0.0372, and 0.0101

3.77 The Poisson model is an appropriate model. Using the fact that the sum of two independent Poisson random variables is again Poisson distributed, the sought probability is $1 - \sum_{k=0}^{10} e^{-6.2}\, 6.2^k/k! = 0.0514$.

3.78 It is reasonable to model the number goals scored per team per game by a Poisson random variable X with expected value $\frac{128}{2 \times 48} = 1.3125$. To see how good this fit is, we calculate $P(X = k)$ for $k = 0, 1, 2, 3$ and $P(X > 3)$. These probabilities have the values 0.2691, 0.3533, 0.2318, 0.1014, and 0.0417. These values are close to the empirical probabilities 0.2708, 0.3542, 0.2500, 0.0833, and 0.0444.

3.79 Using the Poisson model, an estimate is $1 - \sum_{k=0}^{7} e^{-4.2} 4.2^k/k! = 0.064$.

3.80 The probability that you will the jackpot in any given week by submitting 5 six-number sequences in the lottery 6/42 is $5/\binom{42}{6} = 9.531 \times 10^{-7}$. The number of times that you will win the jackpot in the next 312 drawings of the lottery can be modeled by a Poisson distribution with expected value $\lambda_0 = 312 \times 5/\binom{42}{6} = 2.9738 \times 10^{-4}$. Therefore

$$P(\text{you win the jackpot two or more times in the next 312 drawings}) \\ = 1 - e^{-\lambda_0} - \lambda_0 e^{-\lambda_0} = 4.421 \times 10^{-8}.$$

Thus the number of people under the 100 million players who will win the jackpot two or more times in the coming three years can be modeled by a Poisson distribution with expected value

$$\lambda = 100{,}000{,}000 \times 4.421 \times 10^{-8} = 4.421.$$

The sought probability is $1 - e^{-\lambda} = 0.9880$.

3.81 Let X be the number of weekly winners. An appropriate model for X is the Poisson distribution with expected value 0.25. The standard deviation of this distribution is $\sqrt{0.25} = 0.5$. The observed number of winners lies $\frac{3-0.25}{0.5} = 5.5$ standard deviations above the expected value. Without doing any further calculations, we can say that the probability of three or more winners is quite small ($P(X \geq 3) = 2.2 \times 10^{-3}$).

3.82 (**a**) Suppose that X has a Poisson distribution. By the substitution rule,

$$\begin{aligned} & E[\lambda g(X+1) - Xg(X)] \\ & = \sum_{k=0}^{\infty} \lambda g(k+1) e^{-\lambda} \frac{\lambda^k}{k!} - \sum_{k=0}^{\infty} k g(k) e^{-\lambda} \frac{\lambda^k}{k!} \\ & = \sum_{k=0}^{\infty} \lambda g(k+1) e^{-\lambda} \frac{\lambda^k}{k!} - \lambda \sum_{l=0}^{\infty} g(l+1) e^{-\lambda} \frac{\lambda^l}{l!} = 0. \end{aligned}$$

(**b**) Let $p_j = P(X = j)$ for $j = 0, 1, \ldots$. For fixed $i \geq 1$, define the indicator function $g(x)$ by $g(k) = 1$ for $k = i$ and $g(k) = 0$ for $k \neq i$. Then the relation $E[\lambda g(X+1) - Xg(X)] = 0$ reduces to

$$\lambda p_{i-1} - i p_i = 0.$$

This gives $p_i = \frac{\lambda}{i} p_{i-1}$ for $i \geq 0$. By repeated application of this equation, it next follows that $p_i = \frac{\lambda^i}{i!} p_0$ for $i \geq 0$. Using the fact that $\sum_{i=0}^{\infty} p_i = 1$, we get $p_0 = e^{-\lambda}$. This gives

$$P(X = i) = e^{-\lambda} \frac{\lambda^i}{i!} \quad \text{for } i = 0, 1, \ldots,$$

proving the desired result.

3.83 Translate the problem into a chance experiment with $\binom{25}{3} = 2{,}300$ trials. There is a trial for each possible combination of three people. Three given people have the same birthday with probability $(\frac{1}{365})^2$ and have birthdays falling within one day of each other with probability $7 \times (\frac{1}{365})^2$. The Poisson heuristic gives the approximations $1 - e^{-2{,}300 \times (1/365)^2} = 0.0171$ and $1 - e^{-2{,}300 \times 7/(365)^2} = 0.1138$ for the sought probabilities. Simulation shows that the approximate values are close to the exact values. In a simulation study we found the values 0.016 and 0.103.

3.84 Label the four suits as $i = 1, \ldots, 4$. Translate the problem into a chance experiment with 4 trials. The ith trial is said to be successful if suit i is missing in the bridge hand. The success probability of each trial is $p = \binom{39}{13}/\binom{52}{13}$. The Poisson heuristic gives the approximation $1 - e^{-4p}$ to the probability that some suit will be missing in the bridge hand. The approximate value is 0.0499. The exact value can be obtained by the inclusion-exclusion formula and is 0.0511.

3.85 Translate the problem into a chance experiment with n trials. The ith trial is said to be successful if couple i is paired as bridge partners. The success probability of each trial is $p = \frac{1}{2n-1}$. Letting $\lambda = np = \frac{n}{2n-1}$, the Poisson heuristic gives the approximation $1 - e^{-n/(2n-1)}$ to the probability that no couple will be paired as bridge partners. For $n = 10$, the approximate value is 0.4092. This approximate value is quite close to the exact value 0.4088, which is obtained from the inclusion-exclusion method.

3.86 Imagine a chance experiment with 51 trials. In the ith trial the face values of the cards in the positions i and $i+1$ are compared. The trial is said to be successful if the face values are the same. The success probability is $\frac{3}{51}$. The sought probability is equal to the probability of no successful trial. The latter probability can be approximated by the Poisson probability $e^{-51 \times (3/51)} = e^{-3} = 0.950$. In a simulation study we found the value 0.955.

3.87 Translate the problem into a chance experiment with 365 trials. The ith trial is said to be successful if seven or more people have their birthdays on day i. The success probability is $p = [1 - (\frac{354}{365})^{75} - 75 \times \frac{1}{365} \times (\frac{364}{365})^{74}]$. Letting $\lambda = 365p = 6.6603$, the Poisson heuristic gives the approximation $1 - \sum_{k=0}^{6} e^{-\lambda} \lambda^k / k! = 0.499$ to the probability that there are seven or more days so that on each of these days two or more people have their birthday. In a simulation study we found the value 0.516.

3.88 There are $\binom{n}{2}$ combinations of two different integers from 1 to n. Take a random permutation of the integers 1 to n. Imagine a chance experiment with $\binom{n}{2}$ trials, where the ith trial is said to be successful if the two integers involved in the trial have interchanged positions in the random permutation. The success probability of each trial is $\frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$. The number of successful trials can be approximated by a Poisson distribution with expected value $\binom{n}{2}\frac{1}{n(n-1)} = \frac{1}{2}$. In particular, the probability of no successful trial is approximated by $e^{-\frac{1}{2}}$.

3.89 Think of a sequence of $2n$ trials with $n = 5$, where in each trial a person draws at random a card from the hat. The trial is said to be successful if the person draws the card with their own number, or the card with the number of their spouse. The success probability of each trial is $p = \frac{2}{2n}$. The number of successes can be approximated by a Poisson distribution with expected value $\lambda = 2n \times p = 2$. In particular, the probability of no success can be approximated by $e^{-2} = 0.1353$. The exact value is 0.1213 when $n = 5$. This value is obtained from the exact formula

$$\frac{1}{(2n)!}\int_0^\infty (x^2 - 4x + 2)^n e^{-x}\, dx,$$

which is stated without proof.

Note: Using the exact formula, it can be experimentally verified that $e^{-2}(1 - \frac{1}{2n})$ is a better approximation than e^{-2}. In a generalization of the Las Vegas card game, you have two thoroughly shuffled decks of cards, where each deck has $r(= 13)$ types of cards and $s(= 4)$ cards of each type. A match occurs when two cards of the same type occupy the same position in their respective decks. Then the probability of no match can be approximated by e^{-s} for r large, while the exact value of this probability can be calculated from

$$(-1)^{rs}\frac{(s!)^r}{(rs)!}\int_0^\infty [L_s(x)]^r e^{-x}\, dx$$

with $L_s(x) = \sum_{j=0}^{s}(-1)^j\binom{s}{j}\frac{x^j}{j!}$ is the Laguerre polynomial of degree s.

3.90 Imagine a chance experiment with b trials, where the ith trial is said to be successful if the ith bin receives no ball. The success probability of each trial is $p = \left(\frac{b-1}{b}\right)^m$. The trials are weakly dependent when b

is large. Then the probability mass function of the number of empty bins can be approximated by a Poisson distribution with expected value $b\left(\frac{b-1}{b}\right)^m$.

3.91 Imagine a trial for each person. The trial is said to be successful if the person involved has a lone birthday. The success probability is $p = (\frac{364}{365})^{m-1}$. The probability that nobody in the group has a lone birthday is the same as the probability of having no successful trial. Thus, by the Poisson heuristic, the probability that nobody in the group has a lone birthday is approximately equal to $e^{-m(364/365)^{m-1}}$. This leads to the approximate value 3,061 for the minimum number of people that are required in order to have a fifty-fifty probability of no lone birthday. The exact value is 3,064, see Problem 1.89.

3.92 Translate the problem into a chance experiment with 8 trials. The ith trial is said to be successful if you have predicted correctly the two teams for the ith match. The success probability is $p = \frac{8\times 2\times 14!}{16!} = \frac{1}{15}$. The Poisson distribution with expected value $\lambda = 8 \times \frac{1}{15} = \frac{8}{15}$ provides a remarkable accurate approximation for the distribution of the number of correctly predicted matches. In a simulation study we found the values 0.587, 0.312, 0.083, and 0.015 for the probability that k matches are correctly predicted for $k = 0$, 1, 2, and 3, while the approximate values are 0.5866, 0.3129, 0.0834, and 0.0148.

3.93 To approximate the probability of drawing two consecutive numbers, translate the problem into a chance experiment with 44 trials, where there is a trial for any two consecutive numbers from 1 to 45. The probability of drawing two *specific* consecutive numbers is $\binom{43}{4}/\binom{45}{6}$. Thus, letting $\lambda_1 = 44 \times \left[\binom{43}{4}/\binom{45}{6}\right]$, the Poisson heuristic gives the approximation $1 - e^{-\lambda_1} = 0.487$ for the probability of drawing two consecutive numbers. In the same way, letting $\lambda_2 = 43 \times \left[\binom{42}{3}/\binom{45}{6}\right]$, we get the approximation $1 - e^{-\lambda_2} = 0.059$ for the probability of drawing three consecutive numbers. In a simulation study we found the values 0.529 and 0.056 for the two probabilities.

Note: An exact expression for the probability of two or more consecutive numbers in a draw of the lottery is given by

$$1 - \binom{40}{6}/\binom{45}{6}.$$

The trick to get this result is as follows. There is a one-to-one correspondence between the *non-adjacent* ways of choosing six distinct numbers from 1 to 45 and *all* ways of choosing six distinct numbers from 1 to 40. To explain this, take a particular non-adjacent draw of 6 from 45, say $3-12-18-27-35-44$. This non-adjacent draw can be converted to a draw of 6 from 40 by subtracting respectively 0, 1, 2, 3, 4, and 5 from the ordered six numbers. This gives the draw $3-11-16-24-31-39$ for the Lotto 6/40. Conversely, take any set of 6 from 40 and add respectively 0, 1, 2, 3, 4, and 5.

3.94 Imagine that the twenty numbers drawn from the numbers $1,\ldots,80$ are identified as $R=20$ red balls in an urn and that the remaining sixty, nonchosen numbers are identified as $W=60$ white balls in the urn. You have ticked ten numbers on your game form. The probability that you have chosen r numbers from the red group is simply the probability that r red balls will come up in the random drawing of $n=10$ balls from the urn when no balls are replaced. Thus

$$P(r \text{ numbers correct out of 10 ticked numbers}) = \frac{\binom{20}{r}\binom{60}{10-r}}{\binom{80}{10}}.$$

This probability has the values 4.58×10^{-2}, 1.80×10^{-1}, 2.95×10^{-1}, 2.67×10^{-1}, 1.47×10^{-1}, 5.14×10^{-2}, 1.15×10^{-2}, 1.61×10^{-3}, 1.35×10^{-4}, 6.12×10^{-6}, and 1.12×10^{-7} for $r=0,1,\ldots 10$.

3.95 Let X denote how many numbers you will correctly guess. Then X has a hypergeometric distribution with parameters $R=5$, $W=34$, and $n=5$. Therefore

$$P(X=k) = \frac{\binom{5}{k}\binom{34}{5-k}}{\binom{39}{5}} \quad \text{for } k=0,\ldots,5.$$

Let E be the expected payoff per dollar staked. Then,

$$\begin{aligned} E = {} & 100{,}000 \times P(X=5) + 500 \times P(X=4) + 25 \times P(X=3) \\ & + E \times P(X=2). \end{aligned}$$

This gives $E=0.631$. The house percentage is 36.9%.

3.96 Let the random variable X be the number of left shoes among the four shoes you have chosen. The random variable X has a hypergeometric distribution with parameters $R=10$, $W=10$ and $n=4$ applies. The desired probability is

$$1-P(X=0)-P(X=4) = 1-\frac{\binom{10}{4}}{\binom{20}{4}}-\frac{\binom{10}{4}}{\binom{20}{4}} = 0.9133.$$

3.97 The hypergeometric model with $R = W = 25$ and $n = 25$ is applicable under the hypothesis that the psychologist blindly guesses which 25 persons are left-handed. Then, the probability of identifying correctly 18 or more of the 25 left-handers is

$$\sum_{k=18}^{25} \frac{\binom{25}{k}\binom{25}{25-k}}{\binom{50}{25}} = 2.1 \times 10^{-3}.$$

This small probability provides evidence against the hypothesis.

3.98 Let the random variables X_1, X_2, and X_3 indicate how many syndicate tickets that match five, four, or three of the winning numbers. Then the expected amount of money won by the syndicate is

$$\$25{,}000 \times E(X_1) + \$925 \times E(X_2) + \$27.50 \times E(X_3).$$

To work out this expression, we need the probability that a particular ticket matches 5, 4, or 3 numbers given that none of the 200,000 tickets matches all six winning numbers. Denote by A_k the event that a particular ticket matches exactly k of the six winning numbers and B be the event that none of the 200,000 tickets matches all six winning numbers. Then, by the hypergeometric model,

$$P(A_k) = \frac{\binom{6}{k}\binom{40}{6-k}}{\binom{46}{6}}.$$

The sought probability $P(A_k \mid B)$ satisfies

$$P(A_k \mid B) = \frac{P(A_k B)}{P(B)} = \frac{P(B \mid A_k)P(A_k)}{P(B)}.$$

The probability $P(B)$ can be calculated as the Poisson probability of zero successes in 200,000 independent trials each having success probability $1/\binom{46}{6}$ and $P(B \mid A_k)$ can be calculated as the Poisson probability of zero successes in 199,999 of such trials. Noting that $P(B \mid A_k)/P(B)$ is 1 for all practical purposes, we get

$$P(A \mid B_k) = P(A_k).$$

We now find that

$$E(X_k) = 200{,}000 \times \frac{\binom{6}{6-k}\binom{40}{k}}{\binom{46}{6}} \quad \text{for } k = 1, 2, 3.$$

This gives the values $E(X_1) = 5.12447$, $E(X_2) = 249.8180$ and $E(X_3) = 4219.149$. Therefore the expected amount of money won by the syndicate is

$$25{,}000 \times 5.12447 + 925 \times 249.8180 + 27.50 \times 4219.149 = 475{,}220$$

dollars. The expected profit is \$75,220. To conclude, we remark that the random variables X_1, X_2, and X_3 are Poisson distributed. These random variables are practically independent of each other and so the standard deviation of the random variable $25{,}000X_1+925X_2+27.50X_3$ can be approximated by

$$\sqrt{25{,}000^2\sigma^2(X_1) + 925^2\sigma^2(X_2) + 27.50^2\sigma^2(X_3)} = 58{,}478.50.$$

Note: The probability distribution of the random variable $25{,}000X_1 + 925X_2 + 27.50X_3$ can be approximated by a normal distribution with expected value \$475,220 and standard deviation \$58,478.50 (see Chapter 4). This leads to the approximation of 9.9% for the probability that the syndicate will lose money on its investment of \$400,000.

3.99 Use the hypergeometric model with $R = 8$, $W = 7$, and $n = 10$. The sought probability is equal to the probability of picking 5 or more red balls and this probability is $\sum_{k=5}^{8} \binom{8}{k}\binom{7}{10-k}/\binom{15}{10} = \frac{9}{11}$.

3.100 Let the random variable X be the largest number drawn and Y the smallest number drawn. Then,

$$P(X = k) = \frac{\binom{k-1}{5}}{\binom{45}{6}} \quad \text{for } 6 \leq k \leq 45$$

$$P(Y = k) = \frac{\binom{45-k}{5}}{\binom{45}{6}} \quad \text{for } 1 \leq k \leq 40.$$

3.101 For a single player, the problem can be translated into the urn model with 24 red and 56 white balls. This leads to $Q_k = 1 - \binom{24}{24}\binom{56}{k-24}/\binom{80}{k}$ for $24 \leq k \leq 79$, where $Q_{23} = 1$ and $Q_{80} = 0$. The probability that more than 70 numbers must be called out before one of the players has achieved a full card is given by $Q_{70}^{36} = 0.4552$. The probability that you will be the first player to achieve a full card while no other player has a full card at the same time as you is equal to

$$\sum_{k=24}^{79} \left(Q_{k-1} - Q_k\right) Q_k^{35} = 0.0228.$$

The probability that you will be among the first players achieving a full card is

$$\sum_{k=24}^{80}\sum_{a=0}^{35}\binom{35}{a}\left(Q_{k-1}-Q_k\right)^{a+1}Q_k^{35-a}=0.0342.$$

3.102 Write $X = X_1 + \cdots + X_r$, where X_i is the ith number picked. Then $E(X) = \sum_{i=1}^{r} E(X_i)$, $E(X^2) = \sum_{i=1}^{r} E(X_i^2) + 2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r} E(X_iX_j)$. Since the X_i are interchangeable random variables, $E(X_i) = E(X_1)$ for all i and $E(X_iX_j) = E(X_1X_2)$ for all $i \neq j$. Obviously, $E(X_1) = \frac{1}{2}(s+1)$ and so

$$E(X)=\frac{1}{2}r(s+1).$$

Since $P(X_1 = k,\, X_2 = l) = \frac{1}{s} \times \frac{1}{s-1}$ for any $k \neq l$, we have

$$E(X_1X_2)=\sum_{k,l:\, l\neq k} kl\frac{1}{s(s-1)}=\sum_{k=1}^{s}k\Big(\sum_{l=1}^{s}l-k\Big).$$

Using the formulas $\sum_{l=1}^{s} l = \frac{1}{2}s(s+1)$ and $\sum_{k=1}^{s} k^2 = \frac{1}{6}s(s+1)(2s+1)$, it follows after some algebra that $E(X_1X_2) = \frac{1}{12}(s+1)(3s+2)$. Also, we have $E(X_1^2) = \frac{1}{s}\sum_{k=1}^{s} k^2 = \frac{1}{6}(s+1)(2s+1)$. Putting the pieces together, we get

$$E(X^2)=\frac{1}{6}r(s+1)(2s+1)+r(r-1)\frac{1}{12}(s+1)(3s+2).$$

Next, it is a matter of simple algebra to get the formula for $\sigma^2(X)$.

3.103 Fix $1 \leq r \leq a$. Let the random variable X be the number of picks needed to obtain r red balls. Then X takes on the value k if and only if $r-1$ red balls are obtained in the first $k-1$ picks and another red ball at the kth pick. Thus, for $k = r, \ldots, b+r$,

$$P(X=k)=\frac{\binom{a}{r-1}\binom{b}{k-1-(r-1)}}{\binom{a+b}{k-1}}\times\frac{1}{a+b-(k-1)}.$$

Alternatively, the probability mass function of X can be obtained from the tail probability $P(X > k) = \sum_{j=0}^{r-1}\binom{a}{j}\binom{b}{k-j}/\binom{a+b}{k}$.

3.104 Call your opponents East and West. The probability that East has two spades and West has three spades is

$$\frac{\binom{5}{2}\binom{21}{10}}{\binom{26}{13}} = \frac{39}{115}.$$

Hence the desired probability is $2 \times \frac{39}{115} = 0.6783$.

3.105 For $0 \leq k \leq 4$, let E_k be the event that a diamond has appeared 4 times and a spade k times in the first $4+k$ cards and F_k be the event that the $(4+k+1)$th card is a diamond. The sought probability is $\sum_{k=0}^{4} P(E_k F_k) = \sum_{k=0}^{4} P(E_k)P(F_k \mid E_k)$. Thus, the win probability of player A is

$$\sum_{k=0}^{4} \frac{\binom{8}{4}\binom{7}{k}}{\binom{15}{4+k}} \times \frac{4}{11-k} = 0.6224.$$

3.106 Let A_k be the event that the last drawing has exactly k numbers in common with the second last drawing. Seeing the six numbers from the second last drawing as red balls and the other numbers as white balls, it follows that

$$P(A_k) = \frac{\binom{6}{k}\binom{43}{6-k}}{\binom{49}{6}} \quad \text{for } k = 0, 1, \ldots, 6.$$

Let E be the event that the next drawing will have no numbers common with the last two drawings. Then, by the law of conditional probabilities, $P(E) = \sum_{k=0}^{6} P(E \mid A_k)P(A_k)$ and so

$$P(E) = \sum_{k=0}^{6} \frac{\binom{49-(6+6-k)}{6}}{\binom{49}{6}} \times \frac{\binom{6}{k}\binom{43}{6-k}}{\binom{49}{6}} = 0.1901.$$

3.107 Let the random variable X be the number of tickets you will win. Then X has a hypergeometric distribution with parameters $R = 100$, $W = 124{,}900$, and $n = 2{,}500$. The sought probability is $1 - P(X = 0) = 0.8675$. Since $R + W \gg n$, the hypergeometric distribution can be approximated by the binomial distribution with parameters $n = 2{,}500$ and $p = \frac{R}{R+W} = 0.0008$. The binomial distribution in turn can be approximated by a Poisson distribution with expected value $\lambda = np = 2$.

3.108 The probability of the weaker team winning the final is

$$\sum_{k=4}^{7} \binom{k-1}{3} (0.45)^4 (0.55)^{k-4} = 0.3917.$$

Let the random variable X be the number of games the final will take. Then,

$$P(X = k) = \binom{k-1}{3} (0.45)^4 (0.55)^{k-4} + \binom{k-1}{3} (0.55)^4 (0.45)^{k-4}.$$

This probability has the numerical value 0.1325, 0.2549, 0.3093, and 0.3032 for $k = 4$, 5, 6, and 7. The expected value and the standard deviation of the random variable X are given by 5.783 and 1.020.

3.109 The probability of fifteen successes before four failures is

$$\sum_{n=15}^{19} \binom{n-1}{14} \left(\frac{3}{4}\right)^{15} \left(\frac{1}{4}\right)^{n-15} = 0.4654,$$

using the negative binomial distribution. As a sanity check, imagine that 19 trials are done. Then, the sought probability is the probability of 15 or more successes in 19 trials and is equal to the binomial probability

$$\sum_{k=15}^{19} \binom{19}{k} \left(\frac{3}{4}\right)^{k} \left(\frac{1}{4}\right)^{19-k} = 0.4654.$$

3.110 Let the random variable X have a negative binomial distribution with parameters $r = 15$ and $p = \frac{3}{4}$. The probability that the red bag will be emptied first is

$$\sum_{k=15}^{19} P(X = k) = \sum_{k=15}^{19} \binom{k-1}{14} \left(\frac{3}{4}\right)^{15} \left(\frac{1}{4}\right)^{k-15} = 0.4654.$$

The probability that there still $k \geq 1$ balls in the blue bag when the red bag gets empty is

$$P(X = 20 - k) = \binom{19-k}{14} \left(\frac{3}{4}\right)^{15} \left(\frac{1}{4}\right)^{5-k} \quad \text{for } k = 1, \ldots, 5.$$

3.111 Imagine that each player continues rolling the die until one of the assigned numbers of that player appears. Let X_1 be the number of rolls player A needs to get a 1 or 2 and X_2 be the number of rolls player B needs to get a 4, 5 or 6. Then X_1 and X_2 are independent and geometrically distributed with parameters $p_1 = \frac{1}{3}$ and $p_2 = \frac{1}{2}$. The probability of player A winning is

$$P(X_1 \leq X_2) = \sum_{j=1}^{\infty} p_1(1-p_1)^{j-1}(1-p_2)^{j-1} = \frac{p_1}{p_1 + p_2 - p_1 p_2} = \frac{1}{2}.$$

The length of the game is $X = \min(X_1, X_2)$. Thus

$$P(X > l) = (1-p_1)^l(1-p_2)^l = (1-p)^l \quad \text{for } l = 0, 1, \ldots,$$

where $p = p_1 + p_2 - p_1 p_2$. Therefore the length of the game is geometrically distributed with parameter $p = \frac{2}{3}$.

3.112 The geometric distribution with success probability $p = \frac{1}{37}$ applies to this situation. The probability that the house number 0 will not come up in 25 spins of the roulette wheel is

$$1 - (1-p)^{25} = 0.495897.$$

The expected value of the gambler's net profit per dollar bet is \$0.0082.

3.113 Let P_A be the probability of player A winning and P_d be the probability of a draw. By the law of conditional probability, $P_A = a(1-b) + (1-a)(1-b)P_A$ and $P_d = ab + (1-a)(1-b)P_d$. This gives

$$P_A = \frac{a(1-b)}{a+b-ab} \text{ and } P_d = \frac{ab}{a+b-ab}.$$

The length of the game is geometrically distributed with parameter $p = 1-(1-a)(1-b) = a+b-ab$.

3.114 The random variable X is given by $Y-3$, where the random variable Y has a negative binomial distribution with parameters $n = 3$ and $p = \frac{1}{2}$. Hence

$$P(X = x) = \binom{x+2}{2}\left(\frac{1}{2}\right)^{x+3} \quad \text{for } x = 0, 1, \ldots.$$

The expected value and the standard deviation of X are given $6-3 = 3$ and $\sqrt{6} = 2.449$.

3.115 Suppose the strategy is to stop as soon as you have picked a number larger than or equal to r. The number of trials needed is geometrically distributed with parameter $\frac{25-r+1}{25}$ and the amount you get paid has a discrete uniform distribution on $r, \ldots, 25$. The expected net payoff is given by

$$\frac{1}{25-r+1}\sum_{k=r}^{25} k - \frac{25}{25-r+1} = \frac{1}{2}(25+r) - \frac{25}{25-r+1}.$$

This expression takes on the maximal value \$18.4286 for $r = 19$.

3.116 The probability that both coins simultaneously show the same outcome is $p \times \frac{1}{2} + (1-p) \times \frac{1}{2} = \frac{1}{2}$. The desired probability distribution is the geometric distribution with parameter $\frac{1}{2}$.

3.117 Let X be the number of rounds required for the game. The random variable X is geometrically distributed with parameter $p = \sum_{i=2}^{12} a_i^2 = \frac{73}{648}$, where a_i is the probability of rolling a dice total of i and is given by $a_i = \frac{i-1}{36}$ for $2 \leq i \leq 7$ and $a_{14-i} = a_i$ for $8 \leq i \leq 12$. The probability of John paying for the beer is

$$\sum_{k=1}^{5} a_{2k+1}^2 \Big/ \sum_{i=2}^{12} a_i^2 = \frac{38}{73}.$$

3.118 Let us say that a success occurs each time an ace is drawn that you have not seen before. Denote by X_j be the number of cards drawn between the occurrences of the $(j-1)$th and jth success. The random variable X_j is geometrically distributed with success probability $\frac{4-(j-1)}{52}$. Also, the random variables $X_1, \ldots, X_4$ are independent of each other (the cards are drawn with replacement). A geometrically distributed random variable with parameter p has expected value $1/p$ and variance $(1-p)/p^2$. Hence the expected value and the standard deviation of the number of times you have to draw a card until you have seen all four different aces are

$$E(X_1 + X_2 + X_3 + X_4) = \frac{52}{4} + \frac{52}{3} + \frac{52}{2} + \frac{52}{1} = 108.33$$

$$\sigma(X_1 + X_2 + X_3 + X_4) = \sqrt{\sum_{k=1}^{4} \frac{1-k/52}{(k/52)^2}} = 61.16.$$