SECTION 6.4 Expected Value and Variance

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

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

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

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

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

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

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

12. If X is the number of times we roll the die, then X has a geometric distribution with $p = 1/6$.
a) $p(X = n) = (1-p)^{n-1}p = (5/6)(1/6) = 5^{n-1}/6^n$
b) $1/(1/6) = 6$ by Theorem 4

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

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

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

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

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

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

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

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

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

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

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

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

From Theorem 6 we have

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

It follows from algebra and linearity of expectation that

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

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

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

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

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

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

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

34. Since

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

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

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

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

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

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

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

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

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

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

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

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

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

This can be done, painfully, by induction.

f) This follows immediately from the previous two parts.

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

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

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

40. We can use the result of Exercise 38. It is easy to see that $E(X) = 7$ and $E(Y) = 7$ (see Example 4). To find the expectation of XY, we construct the following table to show the value of $2i(i+j)$ for the 36 equally-likely outcomes (i is the row label, j the column label):

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

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

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