Students' Solutions Manual

Probability and Statistics

This manual contains solutions to odd-numbered exercises from the book *Probability and Statistics* by Miroslav Lovrić, published by Nelson Publishing.

Keep in mind that the solutions provided represent *one* way of answering a question or solving an exercise. In many cases there are alternatives, so make sure that you don't dismiss your solution just because it does not look like the solution in this manual.

This solutions manual is not meant to be read! Think, try to solve an exercise on your own, investigate different approaches, experiment, see how far you get. If you get stuck and don't know how to proceed, try to understand why you are having difficulties before looking up the solution in this manual. If you just read a solution you might fail to recognize the hard part(s); even worse, you might completely miss the point of the exercise.

I accept full responsibility for errors in this text and will be grateful to anybody who brings them to my attention. Your comments and suggestions will be greatly appreciated.

Miroslav Lovrić
Department of Mathematics and Statistics
McMaster University
e-mail: lovric@mcmaster.ca

September 2011

Section 2 Stochastic Models

1. (a) The deterministic part $p_{t+1} = p_t$ models a population which does not change in size (a dead lion is immediately replaced by another lion). The stochastic term I_t represents a possible influx of 6 new lions per year. There is a 50% chance that the influx (and thus an increase in population) occurs in any given year. To make a prediction, it is reasonable to assume that in a period of 10 years, an influx of 6 new lions will occur in 5 years. Thus, the most likely value for p_{10} is $100 + 5 \cdot 6 = 130$. The most likely *values* are those close to the 50-50 split: 5 or 7 years with an influx of 6 new lions per year. So, the three most likely values for p_{10} are 125, 130 and 135.

(b) Assume that heads (H) means influx of 6 new lions, and tails (T) represents no influx.

First simulation: HTHTHHHTTT; the corresponding values of p_t, starting with $p_0 = 100$ are 100, 106, 106, 112, 112, 118, 124, 130, 130, 130, 130.

Second simulation: HTTTHTHHHT; the corresponding values of p_t, starting with $p_0 = 100$ are $100, 106, 106, 106, 106, 112, 112, 118, 124, 130, 130$.

Third simulation: TTTHHTTTHT; the corresponding values of p_t, starting with $p_0 = 100$ are $100, 100, 100, 100, 106, 112, 112, 112, 112, 118, 118$.

(c) The two extreme cases are: no immigration in any of the 10 years (in which case $p_{10} = 100$) and immigration in every year (in which case $p_{10} = 160$). In-between are the cases of immigration occurring anywhere from once in 10 years to nine times in 10 years. Thus, the values of p_{10} (and thus the sample space) are $100, 106, 112, 118, 124, 130, 136, 142, 148, 152, 156, 160$.

3. (a) There is a 50% chance that $m_1 = 2$ and a 50% chance that $m_1 = -1$. If $m_1 = 2$, then there is a 50% chance that $m_2 = 4$ and a 50% chance that $m_1 = -2$. If $m_1 = -1$, then there is a 50% chance that $m_2 = -2$ and a 50% chance that $m_1 = 1$. Thus, there are three outcomes for m_2: -4, -2 and 1. The value $m_2 = -4$ can happen in only one way; $m_2 = -2$ can happen in two ways; $m_2 = 1$ can happen in one way. Thus, the chance that $m_2 = 1$ is 1/4. (For the record: the chance that $m_2 = -4$ is 1/4, and the chance that $m_2 = -2$ is $2/4 = 1/2$.)

(b) To get m_4, we have to multiply $m_0 = 1$ the total of four times by a combination of the two factors 2 or -1. If we multiply 1 by 2 four times, we get $m_4 = 16$. If we multiply 1 by 2 three times, then the fourth multiplication is by -1; we get $m_4 = -8$. If we multiply 1 by 2 two times, the remaining two multiplications are by -1; we get $m_4 = 4$. If we multiply 1 by 2 once, the remaining three factors are -1 and we get $m_4 = -2$. Finally, if we multiply 1 by -1 four times, we get $m_4 = 1$. Thus, the sample space for m_4 is the set $\{1, -2, 4, -8, 16\}$.

5. (a) The deterministic part $p_{t+1} = p_t$ models a population which does not change in size (a dead leopard is immediately replaced by another leopard). The stochastic term I_t represents the change in the number of leopards. There is a 75% chance that the influx (i.e., an increase in the population by 3 leopards) occurs in any given year. With a 25% chance, 3 leopards leave in any given year.

(b) Take a four-year interval. In three of the four years, we expect an influx of 3 leopards per year. In one of the four years, we expect that 3 leopards will leave. Thus, the total change in population in the four years is $3 \cdot 3 - 3 = 6$ leopards; equivalently, the increase in population is, on average, $6/4 = 1.5$ leopards per year. Thus, we predict that in 10 years the population will increase by 15 leopards.

In the long term, the population of leopards will increase (at an average of 1.5 leopards per year).

(c) We declare that diamonds (D for decrease) represent 3 leopards leaving in a given year, and the remaining three suits (spades, hearts, and clubs; call them I for increase) represent an influx of 3 leopards in a given year. Assuming that the deck of cards is complete and fair, the chance of picking a diamonds card is $1/4 = 25\%$.

First simulation: DIIIDIIIII; the corresponding values of p_t, starting with $p_0 = 100$ are 100, 97, 100, 103, 106, 103, 106, 109, 112, 115, 118.

Second simulation: DIDDIIIDII; the corresponding values of p_t, starting with $p_0 = 100$ are 100, 97, 100, 97, 94, 97, 100, 103, 100, 103, 106.

Thurd simulation: IIIDDDDDID; the corresponding values of p_t, starting with $p_0 = 100$ are 100, 103, 106, 109, 106, 103, 100, 97, 94, 97, 94.

7. (a) We use a deck of cards and declare that one suit (say, diamonds) represents the no-immigration year, and the remaining three suits (spades, hearts, and clubs) represent immigration of 12 new lions in a year. Assuming that the four suits are equally likely to be drawn, the chance of one suit (say, diamonds) to be picked is $1/4 = 25\%$.

An alternative is to use a mechanism capable of randomly generating numbers between 0 and 99 (there are 100 outcomes). We declare any number between 0 and 24 (total of 25 numbers) to represent no-immigration, and the remaining 75 numbers (from 25 to 99) to represent immigration. (This mechanism could be software or home-made: we could write the numbers on pieces of paper, place them in a bowl and randomly pick a number, keeping in mind that we have to return the number back into the bowl before picking another number.)

(b) In our simulation, we obtained the following: $\diamondsuit$ $\clubsuit$ $\heartsuit$ $\spadesuit$ $\diamondsuit$ $\heartsuit$. The corresponding number of lions is, starting with $p_0 = 160$ (we perform calculations using decimal numbers, and round off when we are done):

$$\begin{aligned}
p_1 &= 0.95p_0 + I_0 = 0.95(160) + 0 = 152 \\
p_2 &= 0.95p_1 + I_1 = 0.95(152) + 12 = 156.40 \\
p_3 &= 0.95p_2 + I_2 = 0.95(156.40) + 12 = 160.58 \\
p_4 &= 0.95p_3 + I_3 = 0.95(160.58) + 12 = 164.55 \\
p_5 &= 0.95p_4 + I_4 = 0.95(164.55) + 0 = 156.32 \\
p_6 &= 0.95p_5 + I_5 = 0.95(156.32) + 12 = 160.50
\end{aligned}$$

Thus, $p_6 = 160$ (or $p_6 = 161$). We expect p_6 to be larger than the values in Figure 2.1, since the chance of immigration is higher (75%, compared to 50%).

9. (a) The distribution of genotypes among the first generation is: 1/4 of all offspring are AA, 1/2 of all offspring are AB, and 1/4 of all offspring are BB.

(b) The ratio of genotype BB offspring in the second generation is: 1/4 (since all offspring of a genotype BB plant are of genotype BB) + $(1/4) \cdot (1/2)$ (since one quarter of offspring of genotype AB parents are of genotype BB). Thus, in the second generation: $1/4 + (1/4) \cdot (1/2) = 3/8$ of all offspring are BB. For AA offspring, we use exactly the same reasoning; the ratio is 3/8 as well. The ratio of AB offspring is 1 minus the sum of the ratios of AA and BB offspring, which is $1 - 3/8 - 3/8 = 2/8 = 1/4$.

(c) We continue in the same way: All BB plants and 1/4 of AB plants from the second generation will produce BB offspring. Thus, the ratio of BB offspring in the third generation is $3/8 + (1/4)(1/4) = 7/16$.

11. (a) All offspring of AA and BB parents are of genotype AB, and so have long ears. Thus the chance that an offspring of AA and BB parents has short ears is 0%.

(b) Making all possible combinations, we get AB, AB, BB, BB. Thus, an offspring of AB and BB parents is of genotype AB (with a chance of 50%) or of genotype BB (with a chance of 50%). Thus, the chance that an offspring of AB and BB parents is BB, i.e., has short ears, is 50%.

13. Denote by p_t the chance that the molecule is still inside the region during the time interval t. Thus, $p_0 = 1$ (initially, the molecule is inside the region). After one hour, the molecule is still inside the region with a chance of 75%. Thus, $p_1 = 0.75$. After two hours, the molecule is still inside the region if it was inside the region during the first hour and during the second hour. Thus, $p_2 = 0.75 \cdot 0.75 = 0.75p_1$ Continuing in the same way, we obtain the dynamical system $p_{t+1} = 0.75p_t$ whose solution is $p_t = 0.75^t$. From

$$\begin{aligned}
0.75^t &< 0.1 \\
t \ln 0.75 &< \ln 0.1 \\
t &> \frac{\ln 0.1}{\ln 0.75} \\
t &> 8.0039
\end{aligned}$$

we conclude that the chance the molecule is still inside the region falls below 10% after 8 hours.

15. The chance that the molecule is inside the region after 2 minutes is $0.25 \cdot 0.25 = 0.625$ (the molecule needs to be inside the region during the first minute and during the second minute). The the chance that the molecule is inside the region after 3 minutes is $0.25 \cdot 0.25 \cdot 0.25 = 0.015625$, i.e., about 1.56%.

17. (a) By adding 1 and -1 to all elements of the sample space at time t we obtain the sample space at time $t+1$. When $t = 0$, the sample space is $\{0\}$. When $t = 1$, the sample space is $\{-1, 1\}$. When $t = 2$, the sample space is $\{-2, 0, 2\}$. When $t = 3$, the sample space is $\{-3, -1, 1, 3\}$. When $t = 4$, the sample space is $\{-4, -2, 0, 2, 4\}$. When $t = 5$, the sample space is $\{-5, -3, -1, 1, 3, 5\}$.

(b) Continuing part (a), we find the sample space at time $t = 6$ to be $\{-6, -4, -2, 0, 2, 4, 6\}$.

(c) Looking at the pattern in (a) and (b), we see that the sample space at time t (i.e., after t steps have been completed) is the set $\{-t, -t+2, -t+4, \ldots, t-4, t-2, t\}$.

Section 3 Basics of Probability Theory

1. Examples of experiments whose sample space consists of three simple events that are not equally likely: (1) Modify the random walk routine: assume that a particle moves from its present position to the left for one unit of distance with a 70% chance, to the right for one unit of distance with a 20% chance, and remains where it is with a 10% chance. Declare the outcome of the experiment to be the location of the particle starting at $x = 0$ after one step of this modified random walk. The sample set is $S = \{-1, 0, 1\}$, and the three simple events occur with different probability. (2) Only one of three molecules diffuses out of a cell. Molecule A diffuses out with probability 0.4, molecule B diffuses out with probability 0.1, and molecule C diffuses out with probability 0.5. (3) A wolf is hunting for food. It catches a rabbit with probability 0.4, a mouse with probability 0.5 and does not catch anything with probability 0.1.

Examples of experiments whose sample space consists of three simple events that are equally likely: (1) Modify the random walk routine: with equal probability (1/3) the particle moves to the right, to the left, or stays where it is. Declare the outcome of the experiment to be the location of the particle starting at $x = 0$ after one step of this modified random walk. The sample set is $S = \{-1, 0, 1\}$, and the three outcomes have equal chance of occurring. (2) Only one of three molecules diffuses out of a cell. Each molecule diffuses with the probability of 1/3. (3) A person randomly picks one of the three flights available from Toronto to Vancouver.

3. The sample space S consists of all mutual products of numbers $1, 2, 3, 4, 5$ and 6: Multiplying this sequence by 1, we get $1, 2, 3, 4, 5, 6$; multiplying by 2, we get $2, 4, 6, 8, 10, 12$; multiplying by 3, we get $3, 6, 9, 12, 15, 18$; multiplying by 4, we get $4, 8, 12, 16, 20, 24$; multiplying by 5, we get $5, 10, 15, 20, 25, 30$; multiplying by 6, we get $6, 12, 18, 24, 30, 36$. Listing each number once, we write the sample space as $S = \{1, 2, 3, 4, 5, 6, 8, 9, 10, 12, 15, 16, 18, 20, 24, 25, 30, 36\}$. Counting the number of elements in S, we see that $|S| = 18$.

5. The sample space S consists of all numbers from 0 to 8. Its size is $|S| = 9$.

7. Let's look at a small value for n first, say $n = 3$. To construct the sample space, we think of forming three-letter sequences where each letter is either H or T; for instance, HTH, TTH, and so on. We have two choices for the first letter (H or T), two choices for the second letter, and two choices for the third letter. Thus, the total number of three-letter sequences is $2 \cdot 2 \cdot 2 = 2^3 = 8$. By reasoning in the same way, we conclude that the sample space for the experiment of tossing a coin n times consists of 2^n elements.

9. Four years: the sample space consists of all four-letter sequences of letters, where each letter is either I or N. Because we have two choices for each letter, the total number of sequences is $2 \cdot 2 \cdot 2 \cdot 2 = 2^4 = 16$. Its elements are: IIII, IIIN, IINI, INII, IINN, INNI, ININ, INNN, NIII, NIIN, NINI, NNII, NINN, NNNI, NNIN, NNNN. (This, and the next part of the question are exercises in organizing the list: note that the first eight elements have I as the first letter, and the remaining eight elements were obtained from those by changing that first letter from I to N.)

Five years: the sample space consists of all five-letter sequences of letters, where each letter is either I or N. Since we have two choices for each letter, the total number of sequences is $2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 = 2^5 = 32$. To obtain a list of all elements, we use the list for the four years, and append I as the first letter, and then use the same list and append N as the first letter: IIIII, IIIIN, IIINI, IINII, IIINN, IINNI, IININ, IINNN, INIII, INIIN, ININI, INNII, ININN, INNNI, INNIN, INNNN, NIIII, NIIIN, NIINI, NINII, NIINN, NINNI, NININ, NINNN, NNIII, NNIIN, NNINI, NNNII, NNINN, NNNNI, NNNIN, NNNNN.

Reasoning in the same way, we conclude that the sample space for n years of the immigration/ no-immigration dynamics contains 2^n elements.

11. We find

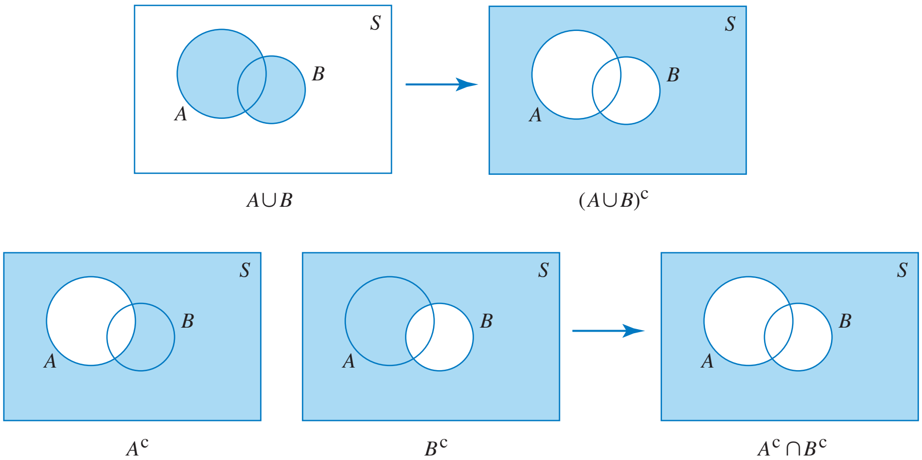
$$\begin{aligned}
A \cup B &= \{1,3,5,6,7,8,9\} \cup \{1,3,4\} = \{1,3,4,5,6,7,8,9\} \\
A \cap B &= \{1,3,5,6,7,8,9\} \cap \{1,3,4\} = \{1,3\} \\
A^{\mathrm{c}} &= \{1,3,5,6,7,8,9\}^{\mathrm{c}} = \{2,4\} \\
A \cap B^{\mathrm{c}} &= \{1,3,5,6,7,8,9\} \cap \{1,3,4\}^{\mathrm{c}} = \{1,3,5,6,7,8,9\} \cap \{2,5,6,7,8,9\} = \{5,6,7,8,9\}
\end{aligned}$$

13. Since all numbers divisible by 4 are even (i.e., $B \subset A$), it follows that $A \cup B = A$ and $A \cap B = B$. The complement of A consists of all odd (non-negative) numbers, $A^{\mathrm{c}} = \{1,3,5,7,9,\ldots\}$. Finally,

$$\begin{aligned}
A \cap B^{\mathrm{c}} &= \{0,2,4,6,8,10,\ldots\} \cap \{0,4,8,12,16,20,\ldots\}^{\mathrm{c}} \\
&= \{0,2,4,6,8,10,\ldots\} \cap \{1,2,3,5,6,7,9,10,11,\ldots\} = \{2,6,10,14,18,22,\ldots\}
\end{aligned}$$

So $A \cap B^{\mathrm{c}}$ is the set of all (non-negative) even numbers which are not divisible by 4.

15. See below.



17. Looking at De Morgan's law $(A \cup B)^{\mathrm{c}} = A^{\mathrm{c}} \cap B^{\mathrm{c}}$ we realize that we can find $P(A^{\mathrm{c}} \cap B^{\mathrm{c}})$ if we can find $P(A \cup B)$; recall that $P((A \cup B)^{\mathrm{c}}) = 1 - P(A \cup B)$. Thus

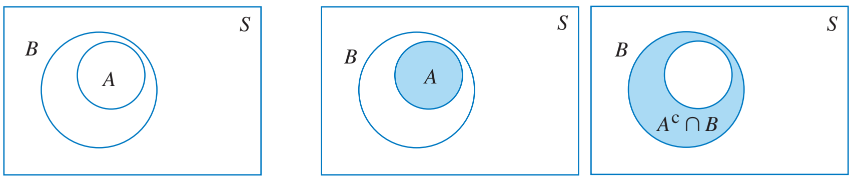
$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.4 + 0.2 - 0.1 = 0.5$$

$P((A \cup B)^{\mathrm{c}}) = 1 - 0.5 = 0.5$ and so $P(A^{\mathrm{c}} \cap B^{\mathrm{c}}) = P((A \cup B)^{\mathrm{c}}) = 0.5$.

19. See the figure below for a proof that $B = A \cup (A^{\mathrm{c}} \cap B)$. Because A and $A^{\mathrm{c}} \cap B$ are disjoint ($A^{\mathrm{c}} \cap B$ is a subset of A^{c}), we conclude that

$$P(B) = P(A \cup (A^{\mathrm{c}} \cap B)) = P(A) + P(A^{\mathrm{c}} \cap B)$$

Since A is a proper subset of B, the set $A^{\mathrm{c}} \cap B$ is non-empty, and therefore $P(A^{\mathrm{c}} \cap B) > 0$. Thus, $P(A) < P(B)$.



21. Interpreting probability as area (or using the argument presented in Exercise 19), we realize that $A \cap B \subseteq B$ implies that $P(A \cap B) \leq P(B)$. The data given ($P(A \cap B) = 0.4$ and $P(B) = 0.2$)

contradict this formula.

23. (a) The probabilities add up to 1. Thus

$$P(4) = 1 - P(1) - P(2) - P(3) - P(5) = 1 - 0.4 - 0.15 - 0.2 - 0.1 = 0.15$$

(Because the meaning is clear, we drop the curly braces from the notation for the probability of an event which consists of a single element, and write $P(1)$ instead of $P(\{1\})$, $P(2)$ instead of $P(\{2\})$, and so on.)

(b) We compute

$$\begin{aligned} P(A) &= P(\{1,2\}) = P(1) + P(2) = 0.4 + 0.15 = 0.55 \\ P(B) &= P(\{2,3,4\}) = 0.15 + 0.2 + 0.15 = 0.5 \\ P(A \cup B) &= P(\{1,2,3,4\}) = P(1) + P(2) + P(3) + P(4) = 0.4 + 0.15 + 0.2 + 0.15 = 0.9 \end{aligned}$$

(Or, $P(A \cup B) = P(\{1,2,3,4\}) = 1 - P(5) = 1 - 0.1 = 0.9$.)

(c) A and B are not disjoint, and therefore $P(A \cup B) \neq P(A) + P(B)$. To verify, we use the probabilities we found in (b): $P(A) + P(B) = 1.05$ whereas $P(A \cup B) = 0.8$.

25. (a) Since the sum of all probabilities is 1, we get

$$P(2) = 1 - P(1) - P(3) - P(4) - P(5) = 1 - 0.2 - 0.4 - 0.3 - 0.1 = 0$$

(Because the meaning is clear, we drop the curly braces from the notation for the probability of an event which consists of a single element, and write $P(1)$ instead of $P(\{1\})$, $P(2)$ instead of $P(\{2\})$, and so on.)

(b) We compute

$$\begin{aligned} P(A) &= P(2) = 0 \quad \text{[Thus, } A = \{2\} \text{ is an impossible event.]} \\ P(A^{\mathrm{C}}) &= 1 - P(A) = 1 - 0 = 1 \\ P(B) &= P(\{1,3,4,5\}) = 0.2 + 0.4 + 0.3 + 0.1 = 1 \\ P(B^{\mathrm{C}}) &= 1 - P(B) = 1 - 1 = 0 \end{aligned}$$

(c) Consider the formula $P(A \cup C) = P(A) + P(C) - P(A \cap C)$. Since $A \cap C \subset A$ it follows that $P(A \cap C) \leq P(A)$. But since $P(A) = 0$, the probability $P(A \cap C) = 0$ and therefore $P(A \cup C) = P(A) + P(C)$ is true.

27. (a) The sample space S is

$$S = \{\mathrm{HHH}, \mathrm{HHT}, \mathrm{HTH}, \mathrm{HTT}, \mathrm{THH}, \mathrm{THT}, \mathrm{TTH}, \mathrm{TTT}\}$$

Let A = "exactly two heads in a row occurred". Then $A = \{\mathrm{HHT}, \mathrm{THH}\}$ and $P(A) = |A|/|S| = 2/8 = 1/4$.

(b) The sample space consists of four-letter sequences in which each letter is either H or T. Since there are two choices for each of the four locations in the sequence, there is a total of $2^4 = 16$ distinct sequences. Thus, $|S| = 16$. Let A = "exactly two heads in a row occurred". Then $A = \{\mathrm{HHTT}, \mathrm{HHTH}, \mathrm{THHT}, \mathrm{TTHH}, \mathrm{HTHH}\}$ and $P(A) = |A|/|S| = 5/16$.

29. The sample space consists of 36 simple events (see Example 3.12 and Table 3.2). A simple event is an ordered pair (m, n) where m is the number that came up on the first die and n is the number that came up on the second die ($1 \leq m, n \leq 6$). Let A = "maximum of the two numbers is 4". A simple event (an ordered pair (m, n)) belongs to A if neither of its entries is larger than 4. There are 4 choices for m, and 4 choices for n, and so $|A| = 16$. [For the record: the following ordered pairs belong to A: $(1,1)$, $(1,2)$, $(1,3)$, $(1,4)$, $(2,1)$, $(2,2)$, $(2,3)$, $(2,4)$, $(3,1)$, $(3,2)$, $(3,3)$, $(3,4)$, $(4,1)$, $(4,2)$, $(4,3)$, $(4,4)$.] We conclude that $P(A) = 16/36 = 4/9$.

31. Let A = "at least one child is a girl". Then A^{C} = "all children are boys". The sample space consists of eight equally likely events (G for a girl and B for a boy): S = {GGG, BGG, GBG, GGB, BBG, BGB, GBB, BBB}. Thus $P(A^{\mathrm{C}}) = P(\mathrm{BBB}) = 1/8$, and so $P(A) = 1 - 1/8 = 7/8$.

33. We follow the strategy of Exercise 31. The sample space consists of six-letter sequences, where each letter is either G or B. Since there are 2 choices for the first letter, two choices for the second letter, and so on, the total number of these six-letter sequences is $2^6 = 64$. Let $A =$ "at least one child is a girl" and $A^{\mathrm{C}} =$ "all children are boys". Since A^{C} consists of one event (BBBBBB), it follows that $P(A^{\mathrm{C}}) = 1/64$, and so $P(A) = 1 - 1/64 = 63/64$.

35. From $p/(1-p) = 2/100$ we get $100p = 2 - 2p$ and $102p = 2$. Thus, the corresponding probability is $p = 2/102 = 1/51$.

37. (a) The sample space is $\{-1, 0, 1\}$. By assumption, the three simple events are equally likely: $P(-1) = P(0) = P(1) = 1/3$.

(b) Assume that the particle is at -1 at $t = 1$. At $t = 2$, with equal probability, it is located at -2, or -1, or 0. We proceed by listing the remaining cases: a particle which is at 0 at $t = 1$ will be at -1, 0 or 1 at $t = 2$ (with equal probability). A particle which is at 1 at $t = 1$ will be at 0, 1 or 2 at $t = 2$ (with equal probability).

Summarizing the above information: 1 path leads to -2, 2 paths lead to -1, 3 paths lead to 0, 2 paths lead to 1 and 1 path leads to 2 (note the symmetry for the locations x and $-x$). There are $1 + 2 + 3 + 2 + 1 = 9$ equally likely paths, and so $P(-2) = P(2) = 1/9$, $P(-1) = P(1) = 2/9$, and $P(0) = 3/9$.

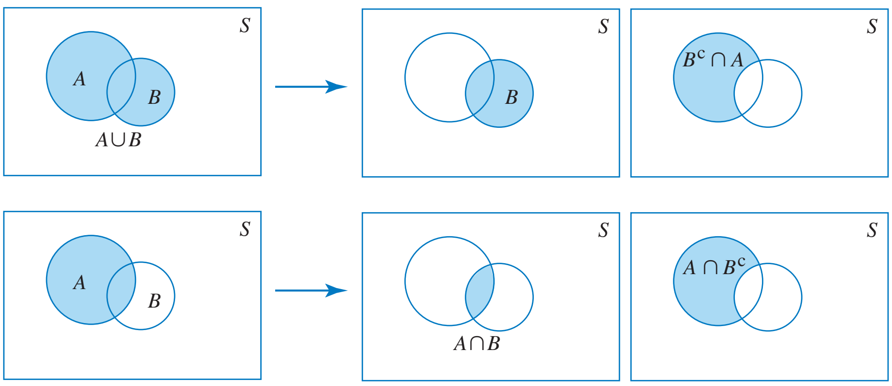
(c) We proceed as in (b). A particle located at -2 when $t = 2$ is at -3, -2, or -1 when $t = 3$. A particle located at -1 when $t = 2$ is at -2, -1, or 0 when $t = 3$. A particle located at 0 when $t = 2$ is at -1, -0, or 1 when $t = 3$. A particle located at 1 when $t = 2$ is at 0, 1, or 2 when $t = 3$. A particle located at 2 when $t = 2$ is at 1, 2, or 3 when $t = 3$. This time, there is a total of $3^3 = 27$ paths (for each t, there are three choices (move left, don't move, and move right), so there is a total of $3 \cdot 3 \cdot 3 = 3^3$ choices).

One path leads to -3. How many paths lead to -2? One from -2 (to which a particle can arrive along one path; see (b)) and one from -1 (to which a particle can arrive along two paths; see (b)). Thus, there are three paths that end at -2 at $t = 3$.

How many paths lead to -1? One from -2 (to which a particle can arrive along one path) one from -1 (to which a particle can arrive along two paths) and one from 0 (to which a particle can arrive along three paths) Thus, there are six paths that end at -1 at $t = 3$. Due to symmetry, we get that one path leads to 3, three paths lead to 2 and six paths lead to 1.

To count the paths that lead to 0 we can proceed as above, or subtract from 27 the number of paths that lead to all other locations; thus, there are $27 - (1 + 3 + 6 + 6 + 3 + 1) = 7$ paths that lead to 0. It follows that $P(-3) = P(3) = 1/27$, $P(-2) = P(2) = 3/27$, $P(-1) = P(1) = 6/27$, and $P(0) = 7/27$.

39. (a) and (b) See below.



(c) Since $A \cup B$ is a disjoint union of B and $B^{\mathrm{C}} \cap A$, it follows that

$$P(A \cup B) = P(B) + P(B^{\mathrm{C}} \cap A)$$

Likewise, A is a disjoint union of $A \cap B$ and $A \cap B^{\mathrm{c}}$, and thus

$$P(A) = P(A \cap B) + P(A \cap B^{\mathrm{c}})$$

Eliminating $P(A \cap B^{\mathrm{c}})$ we get

$$\begin{aligned} P(A \cup B) &= P(B) + P(B^{\mathrm{c}} \cap A) \\ &= P(B) + (P(A) - P(A \cap B)) \\ &= P(A) + P(B) - P(A \cap B) \end{aligned}$$

Section 4 Conditional Probability and the Law of Total Probability

1. Take C to be a subset of A, so that $A \cap C = C$; in that case, $P(A \mid C) = P(A \cap C)/P(C) = P(C)/P(C) = 1$. (Since we are asked to supply a specific example, we pick $A = \{2,3,4\}$ and $C = \{2,4\}$.) If we take disjoint sets B and D, then $P(B \mid D) = P(B \cap D)/P(D) = P(\emptyset)/P(D) = 0$. (For example, $B = \{2,3\}$ and $D = \{4,5\}$.)

3. We compute $P(A \cap B) = P(1) = 0.2$, $P(A) = P(\{1,2,3\}) = 0.2 + 0.1 + 0.15 = 0.45$, and $P(B) = P(\{1,4,5\}) = 0.2 + 0.45 + 0.1 = 0.75$. Thus, $P(A \mid B) = P(A \cap B)/P(B) = 0.2/0.75 = 2/75$ and $P(B \mid A) = P(B \cap A)/P(A) = 0.2/0.45 = 2/45$.

5. We compute $P(A \cap B) = P(\{4,5\}) = 0.4$, $P(A) = P(\{1,2,4,5\}) = 0.8$, and $P(B) = P(\{4,5\}) = 0.4$. Thus, $P(A \mid B) = P(A \cap B)/P(B) = 0.4/0.4 = 1$ (not a surprize, since $B \subset A$) and $P(B \mid A) = P(B \cap A)/P(A) = 0.4/0.8 = 1/2$.

7. Looking at the formulas

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \quad \text{and} \quad P(B \mid A) = \frac{P(B \cap A)}{P(A)}$$

we notice that $P(A \mid B)$ and $P(B \mid A)$ have equal numerators. Thus, it is the denominators that we need to think about.

(a) Since all simple events are equally likely, to make $P(A) \neq P(B)$ we pick A and B to be of different sizes (with a non-empty intersection). For instance, if $A = \{1,2,3\}$ and $B = \{3,4\}$ then

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(3)}{P(\{3,4\})} = 1/2$$
$$P(B \mid A) = \frac{P(B \cap A)}{P(A)} = \frac{P(3)}{P(\{1,2,3\})} = 1/3$$

(b) To make $P(A) = P(B)$ we pick sets of the same size. For instance, if $A = \{2,3,4\}$ and $B = \{3,4,5\}$ then

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\{3,4\})}{P(\{3,4,5\})} = 2/3$$
$$P(B \mid A) = \frac{P(B \cap A)}{P(A)} = \frac{P(\{3,4\})}{P(\{2,3,4\})} = 2/3$$

Alternatively, we can pick two disjoint sets for A and B (not necessarily of the same size), in which case both conditional probabilities are zero.

9. Define A = "two children are girls" and B = "third child is a boy". The probability that the third child is a boy given that two children are girls is $P(B \mid A)$.

If "two children are girls" means "*exactly* two children are girls" then $P(B \mid A) = 1$.

If "two children are girls" means "*at least* two children are girls" then we proceed as follows. The sample space consists of eight equally likely events (G for a girl and B for a boy): S = {GGG, BGG, GBG, GGB, BBG, BGB, GBB, BBB}. Thus

$$P(A) = P(\{\text{GGG}, \text{BGG}, \text{GBG}, \text{GGB}\}) = 4/8$$

and

$$P(A \cap B) = P(\text{BGG, GBG, GGB }\}) = 3/8$$

and therefore $P(B \mid A) = P(A \cap B)/P(A) = (3/8)/(4/8) = 3/4$.

11. Define A = "three children are of the same sex" and B = "fourth child is a girl". The probability that the fourth child is a girl given that three children are of the same sex is $P(B \mid A)$. The sample space consists of four-letter sequences, where each letter is either G (girl) or B (boy). Since there are two choices for the first letter, two choices for the second letter, and so on, the total number of these four-letter sequences is $2^4 = 16$.

If A = "three children are of the same sex" means A = "*exactly* three children are of the same sex" then

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{P(\{\text{BBBG}, \text{BBGB}, \text{BGBB}, \text{GBBB}\})}{P(\{\text{BBBG}, \text{BBGB}, \text{BGBB}, \text{GBBB}, \text{GGGB}, \text{GGBG}, \text{GBGG}, \text{BGGG}\})} = \frac{1}{2}$$

If A = "three children are of the same sex" means A = "*at least* three children are of the same sex" then

$$\begin{aligned} P(B \mid A) &= \frac{P(A \cap B)}{P(A)} \\ &= \frac{P(\{\text{BBBG}, \text{BBGB}, \text{BGBB}, \text{GBBB}, \text{GGGG}\})}{P(\{\text{BBBB}, \text{BBBG}, \text{BBGB}, \text{BGBB}, \text{GBBB}, \text{GGGB}, \text{GGBG}, \text{GBGG}, \text{BGGG}, \text{GGGG}\})} \\ &= \frac{1}{2} \end{aligned}$$

13. Define A = "one toss is H" and B = "at least two H". The probability we are looking for is $P(B \mid A)$. The sample space (tossing a coin three times) is

$$S = \{\text{HHH}, \text{HHT}, \text{HTH}, \text{THH}, \text{HTT}, \text{THT}, \text{TTH}, \text{TTT}\}$$

Thus

$$P(A) = 1 - P(\text{all tosses are T}) = 1 - 1/8 = 7/8$$

and

$$P(A \cap B) = P(\text{as least two H}) = P(\{\text{HHH}, \text{HHT}, \text{HTH}, \text{THH}\}) = 4/8$$

and therefore $P(B \mid A) = P(A \cap B)/P(A) = (4/8)/(7/8) = 4/7$.

15. Define A = "one die shows a number larger than 3" and B = "sum is equal to 7"; we are looking for $P(B \mid A)$. The sample space consists of 36 elements (see Example 3.12 and Table 3.2 in Section 3).

To find $P(A)$ we can list all ordered pairs (m, n) such that one or both numbers are equal to 4, 5, or 6. Alternatively, we look at the complementary event A^{c} = "both dice show 1, or 2, or 3". Since there are 3 choices for each of the two numbers, $|A^{\text{c}}| = 3^2 = 9$, and $P(A^{\text{c}}) = 9/36$; thus, $P(A) = 1 - 9/36 = 27/36$. Now

$$\begin{aligned} P(A \cap B) &= P(\text{sum is 7 and one die shows a number larger than 3}) \\ &= P(\{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}) = 6/36 \end{aligned}$$

and therefore $P(B \mid A) = P(A \cap B)/P(A) = (6/36)/(27/36) = 6/27 = 2/9$.

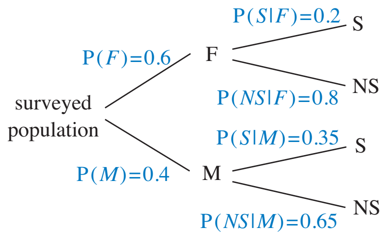
17. Define A = "baby tiger has one T allele" and B = "baby tiger has a striped tail"; we are looking for $P(B \mid A)$. The sample space of genotypes consists of three events $\{PP, PT, TT\}$ with probabilities $P(PP) = 1/4$, $P(PT) = 1/2$, and $P(TT) = 1/4$. Thus $P(A) = P(\{PT, TT\}) = 1/2 + 1/4 = 3/4$, $P(A \cap B) = P(TT) = 1/4$ and therefore $P(B \mid A) = P(A \cap B)/P(A) = (1/4)/(3/4) = 1/3$.

19. (a) We have to pick at least one pair of non-disjoint subsets (i.e., they have to have a non-empty intersection). For instance, take $A = \{1, 2, 3\}$, $B = \{3, 4\}$, and $C = \{5\}$. The union of these sets is S, but A and B are not mutually exclusive.

(b) Let $A = \{1\}$, $B = \{2\}$, and $C = \{3\}$. The three sets are mutually exclusive, but their union $\{1, 2, 3\}$ is not equal to S.

(c) Let $A = \{1, 4\}$, $B = \{2, 5\}$, and $C = \{3\}$. The three sets are mutually exclusive and their union is equal to the universal set S.

21. The diagram below will help us calculate the probabilities. We use the following: F = "female", M = "male", S = "smoker", and NS = "non-smoker". The subsets F and M form a partition of the surveyed population.



(a) By the law of total probability,

$$P(S) = P(S \mid F)P(F) + P(S \mid M)P(M) = (0.2)(0.6) + (0.35)(0.4) = 0.26$$

(b) Using Bayes' formula, we get

$$P(M \mid S) = \frac{P(S \mid M)P(M)}{P(S \mid F)P(F) + P(S \mid M)P(M)} = \frac{(0.35)(0.4)}{0.26} = \frac{14}{26} = \frac{7}{13}$$

23. We use the following: C = "child", Y = "adolescent", A = "adult", and F = "flu". The subsets C, Y, and A form a partition of the given population.
(a) By the law of total probability,

$$\begin{aligned} P(F) &= P(F \mid C)P(C) + P(F \mid Y)P(Y) + P(F \mid A)P(A) \\ &= (0.45)(0.2) + (0.2)(0.3) + (0.15)(0.5) = 0.225 \end{aligned}$$

(b) Using Bayes' formula, we get

$$P(A \mid F) = \frac{P(F \mid A)P(A)}{P(F \mid C)P(C) + P(F \mid Y)P(Y) + P(F \mid A)P(A)} = \frac{(0.15)(0.5)}{0.225} = \frac{0.075}{0.225} = \frac{1}{3}$$

25. Let F = "female", M = "male", and A = "asthma". The subsets F and M form a partition of the population of young adults. It is given that $P(F) = P(M) = 0.5$, $P(A \mid F) = 0.064$ and $P(A \mid M) = 0.045$.
(a) By the law of total probability,

$$P(A) = P(A \mid F)P(F) + P(A \mid M)P(M) = (0.064)(0.5) + (0.045)(0.5) = 0.0545$$

(b) Using Bayes' formula, we get

$$P(F \mid A) = \frac{P(A \mid F)P(F)}{P(A \mid F)P(F) + P(A \mid M)P(M)} = \frac{(0.064)(0.5)}{0.0545} = \frac{0.032}{0.0545} \approx 0.587$$

27. Let R = "rain", NR = "no rain", and C = "car available". The subsets R and NR form a partition of the set of possible weather conditions tomorrow. It is given that $P(R) = 0.6$ (thus $P(NR) = 0.4$), $P(C \mid R) = 0.3$, and $P(C \mid NR) = 0.9$. By the law of total probability,

$$P(C) = P(C \mid R)P(R) + P(C \mid NR)P(NR) = (0.3)(0.6) + (0.9)(0.4) = 0.54$$

29. Let M = "person has meningitis", NM = "person does not have meningitis", and A = "test for meningitis is positive". The subsets M and NM form a partition of the population of Canada. It is given that $P(M) = 3.4/100,000$ (thus $P(NM) = 1 - 3.4/100,000 = 99,996.6/100,000$), $P(A \mid M) = 0.85$, and $P(A \mid NM) = 0.07$.

(a) By the law of total probability,

$$\begin{aligned} P(A) &= P(A \mid M)P(M) + P(A \mid NM)P(NM) \\ &= 0.85 \cdot \frac{3.4}{100,000} + 0.07 \cdot \frac{99,996.6}{100,000} = \frac{7,002.652}{100,000} \approx 0.0700 \end{aligned}$$

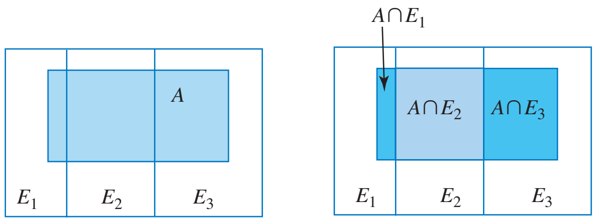
So, the probability that a randomly selected person tests positive for meningitis is about 7%.

(b) Using Bayes' formula, we get

$$P(M \mid A) = \frac{P(A \mid M)P(M)}{P(A \mid M)P(M) + P(A \mid NM)P(NM)} = \frac{0.85 \cdot \frac{3.4}{100,000}}{\frac{7,002.652}{100,000}} = \frac{2.890}{7002.62} \approx 0.00041270$$

So if a person tests positive for bacterial meningitis, the probability that they have it is very small, about 0.04%.

31. (a) See below. Because E_1, E_2, and E_3 form a partition, they are disjoint, and thus the sets $A \cap E_1$, $A \cap E_2$, and $A \cap E_3$ are disjoint as well (because $A \cap E_1$ is a subset of E_1, $A \cap E_2$ is a subset of E_2, and $A \cap E_3$ is a subset of E_3). Therefore, $P(A) = P(A \cap E_1) + P(A \cap E_2) + P(A \cap E_3)$ is true.



(b) From $P(A \mid E_1) = P(A \cap E_1)/P(E_1)$ it follows that $P(A \cap E_1) = P(A \mid E_1)P(E_1)$. Likewise, $P(A \cap E_2) = P(A \mid E_2)P(E_2)$ and $P(A \cap E_3) = P(A \mid E_3)P(E_3)$. So, the equation

$$P(A) = P(A \cap E_1) + P(A \cap E_2) + P(A \cap E_3)$$

implies that

$$P(A) = P(A \mid E_1)P(E_1) + P(A \mid E_2)P(E_2) + P(A \mid E_3)P(E_3)$$

(c) Assume that $E_1, E_2, \ldots, E_n$ form a partition of S. As in (a), a Venn diagram shows that A can be written as a union of disjoint sets

$$A = (A \cap E_1) \cup (A \cap E_2) \cup \cdots \cup (A \cap E_n)$$

Thus

$$P(A) = P(A \cap E_1) + P(A \cap E_2) + \cdots + P(A \cap E_n)$$

Repeating the calculation in (b), we show that $P(A \cap E_i) = P(A \mid E_i)P(E_i)$ for all $i = 1, 2, \ldots, n$. Therefore,

$$P(A) = P(A \mid E_1)P(E_1) + P(A \mid E_2)P(E_2) + \cdots + P(A \mid E_n)P(E_n)$$

and we are done.

Section 5 Independence

1. No. If A and B are disjoint, then $P(A \cap B) = P(\emptyset) = 0$, and the condition for independence $P(A \cap B) = P(A)P(B)$ reads $0 = P(A)P(B)$. This equation implies that either $P(A) = 0$ and $P(B) = 0$, which contradicts the assumption that $P(A) > 0$ and $P(B) > 0$.

3. We compute

$$\begin{aligned} P(B \mid A) &= \frac{P(B \cap A)}{P(A)} = \frac{P(5)}{P(\{4,5\})} = \frac{0.1}{0.4+0.1} = \frac{0.1}{0.5} = 0.2 \\ P(B) &= P(\{2,5\}) = 0.1 + 0.1 = 0.2 \end{aligned}$$

Since $P(B \mid A) = P(B)$, it follows that B and A are independent. (Because the meaning is clear, we drop the curly braces from the notation for the probability of an event which consists of a single element, and write $P(5)$ instead of $P(\{5\})$.)

5. We compute

$$\begin{aligned} P(A \mid B) &= \frac{P(A \cap B)}{P(B)} = \frac{P(\{1,5\})}{P(\{1,2,5\})} = \frac{0.2+0.1}{0.2+0.1+0.1} = \frac{0.3}{0.4} = 0.75 \\ P(A) &= P(\{1,5\}) = 0.2 + 0.1 = 0.3 \end{aligned}$$

Since $P(A \mid B) \neq P(A)$, it follows that A and B are not independent.

7. We compute

$$\begin{aligned} P(A \cap B) &= P(\{1\}) = 0.2 \\ P(A) &= P(\{1,3\}) = 0.2 + 0.2 = 0.4 \\ P(B) &= P(\{1,4\}) = 0.2 + 0.3 = 0.5 \end{aligned}$$

Since $P(A)P(B) = (0.4)(0.5) = 0.2 = P(A \cap B)$, it follows that A and B are independent.

9. Because $A = \{1,3\}$ and $B = \{2,3\}$ are independent,

$$\begin{aligned} P(A \cap B) &= P(A)P(B) \\ P(\{3\}) &= (0.5)(0.4) \end{aligned}$$

and so $P(3) = 0.2$. (Since the meaning is clear, we drop the curly brace notation for the probability of an event which consists of a single element and write $P(3)$ instead of $P(\{3\})$, $P(1)$ instead of $P(\{1\})$, and so on.) From $P(A) = P(\{1,3\}) = P(1) + P(3)$ we get $0.5 = P(1) + 0.2$ and $P(1) = 0.3$. Likewise, $P(B) = P(\{2,3\}) = P(2) + P(3)$ implies $0.4 = P(2) + 0.2$ and $P(2) = 0.2$. Finally, $P(4) = 1 - P(1) - P(2) - P(3) = 0.3$.

11. We find $P(A) = 0.2 + 0.3 = 0.5$. The relation for independence $P(A \cap B) = P(A)P(B)$ tells us to look for a two-element set B (thus $P(B) > 0$) so that $P(A \cap B) = 0.5P(B)$.

Note that B cannot be disjoint from A (since in that case $P(A \cap B) = 0$ and the relation $P(A \cap B) = 0.5P(B)$ does not hold). If $B = A$, then $P(A \cap B) = P(B)$, and again $P(A \cap B) = 0.5P(B)$ does not hold. This analysis shows that B must have one element in common with A.

Now, it's a matter of trial-and-error. Assume that $A \cap B = \{2\}$ and take $B = \{2,1\}$. Then $P(A \cap B) = P(2) = 0.2$ and $P(B) = P(2) + P(1) = 0.5$, so $P(A \cap B) = 0.5P(B)$ does not hold. In the same way, we learn that the choice $B = \{2,3\}$ does not work either.

Thus, it must be that $A \cap B = \{4\}$. Take $B = \{4,1\}$. Then $P(A \cap B) = P(4) = 0.3$ and $P(B) = P(4) + P(1) = 0.6$ and $P(A \cap B) = 0.5P(B)$ is satisfied. Thus, $B = \{4,1\}$. By showing that $B = \{4,3\}$ does not satisfy $P(A \cap B) = 0.5P(B)$, we show that the answer for B is unique.

13. Let Q_i = "student answers the ith question correctly", where $i = 1, 2, \ldots, 10$. The context implies that Q_i are independent events and $P(Q_i) = 1/2$ for all i. The probability of complementary events Q_i^{c} = "student answers the ith question incorrectly" is $P(Q_i^{\mathrm{c}}) = 1/2$, $i = 1, 2, \ldots, 10$.

(a) As usual, the phrase "at least" suggests that we use a complementary event. If A = "student answers at least one question correctly" then A^{C} = "student answers all questions incorrectly." From

$$A^{\mathrm{C}} = Q_1^{\mathrm{c}} \cap Q_2^{\mathrm{c}} \cap \cdots \cap Q_{10}^{\mathrm{c}}$$

we get (by independence)

$$P(A^{\mathrm{C}}) = P(Q_1^{\mathrm{c}})P(Q_2^{\mathrm{c}}) \cdots P(Q_{10}^{\mathrm{c}}) = \left(\frac{1}{2}\right)^{10} = \frac{1}{1024}$$

Thus,

$$P(A) = 1 - P(A^{\mathrm{C}}) = 1 - \frac{1}{1024} = \frac{1023}{1024} \approx 0.99902$$

(b) Let B = "student answers all questions correctly." Then from

$$B = Q_1 \cap Q_2 \cap \cdots \cap Q_{10}$$

we compute

$$P(B) = \left(\frac{1}{2}\right)^{10} = \frac{1}{1024} \approx 0.00098$$

15. Let G_i = "ith child is a girl" and B_i = "ith child is a boy". It is given that $P(G_i) = 0.45$ and $P(B_i) = 0.55$. In part (a), $i = 1, 2, 3$; in part (b), $i = 1, 2, 3, 4$.

(a) Let A = "two girls". Then

$$A = (G_1 \cap G_2 \cap B_3) \cup (G_1 \cap B_2 \cap G_3) \cup (B_1 \cap G_2 \cap G_3)$$

Note that A is a union of three disjoint sets. By the mutual exclusivity property and then by the independence, we get

$$\begin{aligned}
P(A) &= P(G_1 \cap G_2 \cap B_3) + P(G_1 \cap B_2 \cap G_3) + P(B_1 \cap G_2 \cap G_3) \\
&= P(G_1)P(G_2)P(B_3) + P(G_1)P(B_2)P(G_3) + P(B_1)P(G_2)P(G_3) \\
&= (0.45)(0.45)(0.55) + (0.45)(0.55)(0.45) + (0.55)(0.45)(0.45) \\
&= 3(0.45)^2(0.55) \approx 0.334
\end{aligned}$$

(b) Let C = "at least two children are boys." The event C includes all combinations involving 2, 3, and 4 boys. To reduce the number of combinations, we consider C^{C} = "no boys or one boy." Since

$$\begin{aligned}
C^{\mathrm{C}} = (G_1 \cap G_2 \cap G_3 \cap G_4) &\cup (B_1 \cap G_2 \cap G_3 \cap G_4) \cup (G_1 \cap B_2 \cap G_3 \cap G_4) \\
&\cup (G_1 \cap G_2 \cap B_3 \cap G_4) \cup (G_1 \cap G_2 \cap G_3 \cap B_4)
\end{aligned}$$

it follows that (again, by the mutual exclusivity and the independence of the events)

$$\begin{aligned}
P(C^{\mathrm{C}}) &= P(G_1)P(G_2)P(G_3)P(G_4) + P(B_1)P(G_2)P(G_3)P(G_4) + P(G_1)P(B_2)P(G_3)P(G_4) \\
&\quad + P(G_1)P(G_2)P(B_3)P(G_4) + P(G_1)P(G_2)P(G_3)P(B_4) \\
&= (0.45)^4 + (0.45)^3(0.55) + (0.45)^3(0.55) + (0.45)^3(0.55) + (0.45)^3(0.55) \\
&= (0.45)^4 + 4(0.45)^3(0.55) \approx 0.241
\end{aligned}$$

The probability that at least two children are boys is

$$P(C) = 1 - P(C^{\mathrm{C}}) \approx 1 - 0.241 = 0.759$$

17. Let H_i = "ith house dust mite survives in laundry washed at 60^{o}C", where $i = 1, 2, \ldots, 100$. It is given that $P(H_i) = 0.01$ for all i; thus, $P(H_i^{\mathrm{C}}) = 0.99$ for all i. We are looking for the probability of A = "at least one house dust mite survives". Consider the complementary event A^{C} = "none of the 100 house dust mites survives", i.e.,

$$A^{\mathrm{C}} = H_1^{\mathrm{C}} \cap H_2^{\mathrm{C}} \cap \cdots \cap H_{100}^{\mathrm{C}}$$

Then (using the independence)

$$P(A^{\mathrm{C}}) = P(H_1^{\mathrm{C}})P(H_2^{\mathrm{C}}) \cdots P(H_{100}^{\mathrm{C}}) = (0.99)^{100}$$

and so $P(A) = 1 - (0.99)^{100} \approx 0.634$.

19. Let F_i = "test result for the ith person is false-negative", where $i = 1, 2, \ldots, 50$. It is given that $P(F_i) = 0.012$ for all i; thus, F_i^{C} = "test result for the ith person is not false-negative" and $P(F_i^{\mathrm{C}}) = 0.988$ for all i. We are looking for the probability of F = "at least one false-negative test result in a group of 50 people." Consider the complementary event F^{C} = "no one in a group of 50 people receives a false-negative test result"

$$F^{\mathrm{C}} = F_1^{\mathrm{C}} \cap F_2^{\mathrm{C}} \cap \cdots \cap F_{50}^{\mathrm{C}}$$

Assuming the independence of testing,

$$P(F^{\mathrm{C}}) = P(F_1^{\mathrm{C}})P(F_2^{\mathrm{C}}) \cdots P(F_{50}^{\mathrm{C}}) = (0.988)^{50}$$

and $P(F) = 1 - (0.988)^{50} \approx 0.453$. Thus, the probability of at least one false-negative test result is quite high, about 45.3%.

21. Let C_i = "use of a condom prevents pregnancy in year i", where $i = 1, 2, 3, 4, 5$. It is given that $P(C_i) = 0.86$ for all i. We are looking for the probability of A = "sexually active woman who uses condoms regularly gets pregnant at least once in a 5-year period." Consider the complementary event A^{C} = "sexually active woman who uses condoms regularly does not get pregnant in a 5-year period." We write

$$A^{\mathrm{C}} = C_1 \cap C_2 \cap C_3 \cap C_4 \cap C_5$$

Assuming the independence of events,

$$P(A^{\mathrm{C}}) = P(C_1)P(C_2)P(C_3)P(C_4)P(C_5) = (0.86)^5$$

and $P(A) = 1 - (0.86)^5 \approx 0.530$. The probability at least one pregnancy in five years is about 53%.

23. We use abbreviated symbols to represent independence conditions: we write XY for $P(X \cap Y) = P(X)P(Y)$, XYZ for $P(X \cap Y \cap Z) = P(X)P(Y)P(Z)$, and so on.

(a) To prove that the four events A, B, C, and D are independent, we need to check: pairs of events AB, AC, AD, BC, BD, CD; triples of events ABC, ABD, ACD, BCD; and the quadruple of events $ABCD$. Thus, we need to check the total of $6 + 4 + 1 = 11$ conditions.

(b) To prove that the five events A, B, C, D, and E are independent, we need to check: pairs of events AB, AC, AD, AE, BC, BD, BE, CD, CE, DE; triples of events ABC, ABD, ABE, ACD, ACE, ADE, BCD, BCE, BDE, CDE; quadruples of events $ABCD$, $ABCE$, $ABDE$, $ACDE$, $BCDE$; and the quintuplet of events $ABCDE$. Thus, we need to check the total of $10 + 10 + 5 + 1 = 26$ conditions.

(c) (Section 10 reasoning.) Assume that there are n events. The number of conditions involving 2 events is the number of ways we can pick a group of 2 symbols out of the group of n symbols, which is $\binom{n}{2}$; the number of conditions involving 3 events is the number of ways we can pick a group of 3 symbols out of the group of n symbols, which is $\binom{n}{3}$; and so on. (So, the sum in (b) is $\binom{5}{2} + \binom{5}{3} + \binom{5}{4} + \binom{5}{5}$.)

25. (a) Using $g_{t+1} = ag_t$, we compute $g_1 = ag_0$, $g_2 = ag_1 = a(ag_0) = a^2 g_0$, $g_3 = ag_2 = a(a^2 g_0) = a^3 g_0$, and so on. Thus, $g_t = a^t g_0$. When $a = 1 - m$, we get $g_t = g_0(1 - m)^t$.

(b) We check that the left side r_{t+1} is equal to the right side $(1 - m)r_t + m$ when we substitute $r_t = (r_0 - 1)(1 - m)^t + 1$:

$$\begin{aligned} r_{t+1} &= (r_0 - 1)(1 - m)^{t+1} + 1 \\ (1 - m)r_t + m &= (1 - m)[(r_0 - 1)(1 - m)^t + 1] + m \\ &= [(r_0 - 1)(1 - m)^{t+1} + 1 - m] + m \\ &= (r_0 - 1)(1 - m)^{t+1} + 1 \end{aligned}$$

Section 6 Discrete Random Variables

1. The range of X is the set $\{5, 6, 7, 8, 9, 10, \ldots\}$; it is an infinite, countable set (since its elements can be listed in a sequence). Thus, X is a discrete random variable.

3. Since $p(1) + p(2) + p(3) = 0.16 + 0.54 + 0.29 = 0.99 \neq 1$, p cannot be a probability mass function of a random variable.

5. Because $F(x) = 0.32$ if $0 \leq x < 1$ and $F(x) = 0.31$ if $1 \leq x < 2$ it follows that $F(x)$ is decreasing on a part of its domain. Thus one of the properties of a cumulative distribution function ($F(x)$ is non-decreasing for all x) fails to hold.

7. The sample space S consists of four-letter sequences, where each letter is either H or T. Thus, S contains $2 \cdot 2 \cdot 2 \cdot 2 = 2^2 = 16$ elements:

$$S = \{\mathrm{THHH}, \mathrm{TTHH}, \mathrm{THTH}, \mathrm{THHT}, \mathrm{TTTH}, \mathrm{TTHT}, \mathrm{THTT}, \mathrm{TTTT},$$
$$\mathrm{HHHH}, \mathrm{HTHH}, \mathrm{HHTH}, \mathrm{HHHT}, \mathrm{HTTH}, \mathrm{HTHT}, \mathrm{HHTT}, \mathrm{HTTT}\}$$

Since all 16 events are equally likely, the probability of any one occurring is $1/16$.

The range of X is $\{0, 1, 2, 3, 4\}$. The probabilities are:

$$P(X = 0) = P(\{\mathrm{HHHH}\}) = 1/16$$
$$P(X = 1) = P(\{\mathrm{THHH}, \mathrm{TTHH}, \mathrm{THTH}, \mathrm{THHT}, \mathrm{TTTH}, \mathrm{TTHT}, \mathrm{THTT}, \mathrm{TTTT}\}) = 8/16$$
$$P(X = 2) = P(\{\mathrm{HTHH}, \mathrm{HTTH}, \mathrm{HTHT}, \mathrm{HTTT}\}) = 4/16$$
$$P(X = 3) = P(\{\mathrm{HHTH}, \mathrm{HHTT}\}) = 2/16$$
$$P(X = 4) = P(\{\mathrm{HHHT}\}) = 1/16$$

The probability mass function of X is given in the table below.

x	$P(X = x)$
0	$1/16$
1	$1/2$
2	$1/4$
3	$1/8$
4	$1/16$

9. The sample space of the experiment consists of 36 simple events

$$S = \{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6), (2, 1), \ldots, (6, 5), (6, 6)\}$$

(where (m, n) means that the number m came up on the first die and n came up on the second die). Since all simple events are equally likely, the probability that any one occurs is $1/36$.

The range of X is $\{1, 2, 3, 4, 5, 6\}$. The probabilities are:

$$P(X = 1) = P(\{(1, 1)\}) = 1/36$$
$$P(X = 2) = P(\{(1, 2), (2, 1), (2, 2)\}) = 3/36$$
$$P(X = 3) = P(\{(1, 3), (3, 1), (2, 3), (3, 2), (3, 3)\}) = 5/36$$
$$P(X = 4) = P(\{(1, 4), (4, 1), (2, 4), (4, 2), (3, 4), (4, 3), (4, 4)\}) = 7/36$$
$$P(X = 5) = P(\{(1, 5), (5, 1), (2, 5), (5, 2), (3, 5), (5, 3), (4, 5), (5, 4), (5, 5)\}) = 9/36$$
$$P(X = 6) = P(\{(1, 6), (6, 1), (2, 6), (6, 2), (3, 6), (6, 3), (4, 6), (6, 4), (5, 6), (6, 5), (6, 6)\}) = 11/36$$

The probability mass function of X is given in the table below.

x	$P(X = x)$
1	1/36
2	3/36
3	5/36
4	7/36
5	9/36
6	11/36

11. We use the mutual exclusivity of events to calculate the probabilities. The probability distribution for Y: $P(Y = 0) = 1/8$, $P(Y = 1) = 4/8$, $P(Y = 2) = 2/8$, $P(Y = 3) = 1/8$. The probability distribution for Z: $P(Z = 0) = 1/8$, $P(Z = 1) = 4/8$, $P(Z = 2) = 2/8$, $P(Z = 3) = 1/8$. The probability distribution for W: $P(W = -6) = 1/8$, $P(W = -1) = 3/8$, $P(W = 4) = 3/8$, $P(W = 9) = 1/8$.

13. Denote the presence of a virus by V and its absence by N. The sample space is

$$S = \{\mathrm{VV}, \mathrm{NV}, \mathrm{VN}, \mathrm{NN}\}$$

For instance, NV describes the event that the population (which starts virus-free, by assumption) is virus-free for a month; then the virus appears.

The random variable X counts the number of virus-free months in the 2-month period; thus $X(\mathrm{VV}) = 0$, $X(\mathrm{NV}) = 1$, $X(\mathrm{VN}) = 1$, and $X(\mathrm{NN}) = 2$.

Using independence, we compute

$$\begin{aligned}
P(\mathrm{VV}) &= P(\text{V during the first month and V during the second month} \\
&= P(\text{V during the first month})P(\text{V during the second month}) \\
&= (0.3)(0.4) = 0.12 \\
P(\mathrm{NV}) &= P(\mathrm{N})P(\mathrm{V}) = (0.7)(0.3) = 0.21 \\
P(\mathrm{VN}) &= P(\mathrm{V})P(\mathrm{N}) = (0.3)(0.6) = 0.18 \\
P(\mathrm{NN}) &= P(\mathrm{N})P(\mathrm{N}) = (0.7)(0.7) = 0.49
\end{aligned}$$

Thus, the probability mass function for X is given by $P(X = 0) = 0.12$, $P(X = 1) = 0.21+0.18 = 0.39$, and $P(X = 2) = 0.49$.

15. Denote a dark brown baby monkey by D and a light brown baby monkey by L. It is given that, in every year, $P(D) = 0.65$ and $P(L) = 0.35$. The sample space is

$$S = \{\mathrm{DDD}, \mathrm{LDD}, \mathrm{DLD}, \mathrm{DDL}, \mathrm{LLD}, \mathrm{LDL}, \mathrm{DLL}, \mathrm{LLL}\}$$

and the probabilities of the simple events in S are (using independence):

$$\begin{aligned}
P(\mathrm{DDD}) &= P(D)P(D)P(D) = (0.65)^3 = 0.274625 \\
P(\mathrm{LDD}) = P(\mathrm{DLD}) = P(\mathrm{DDL}) &= P(D)P(D)P(L) = (0.65)^2(0.35) = 0.147875 \\
P(\mathrm{LLD}) = P(\mathrm{LDL}) = P(\mathrm{DLL}) &= P(D)P(L)P(L) = (0.65)(0.35)^2 = 0.079625 \\
P(\mathrm{LLL}) &= P(L)P(L)P(L) = (0.35)^3 = 0.042875
\end{aligned}$$

The range of R is $\{0, 1, 2, 3\}$. Using mutual exclusivity, we compute

$$\begin{aligned}
P(R = 0) &= 0.042875 \\
P(R = 1) &= 3(0.079625) = 0.238875 \\
P(R = 2) &= 3(0.147875) = 0.443625 \\
P(R = 3) &= 0.274625
\end{aligned}$$

17. Denote a red-eyed baby monkey by D, a blue-eyed baby monkey by E, a brown-eyed baby monkey by N (for lack of good notation, we use the last letter of the word for the colour). It is given that, in every year, $P(D) = 0.15$, $P(E) = 0.05$, and $P(N) = 0.8$. The sample space is

$$S = \{\mathrm{DD}, \mathrm{EE}, \mathrm{NN}, \mathrm{DE}, \mathrm{ED}, \mathrm{NE}, \mathrm{EN}, \mathrm{DN}, \mathrm{ND}\}$$

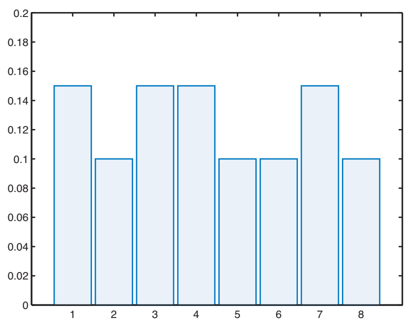
and the probabilities of the events in S are (using independence):

$$\begin{aligned}
P(\mathrm{DD}) &= P(D)P(D) = (0.15)^2 = 0.0225 \\
P(\mathrm{EE}) &= P(E)P(E) = (0.05)^2 = 0.0025 \\
P(\mathrm{NN}) &= P(D)P(D) = (0.8)^2 = 0.64 \\
P(\mathrm{DE}) &= P(\mathrm{ED}) = P(D)P(E) = (0.15)(0.05) = 0.0075 \\
P(\mathrm{NE}) &= P(\mathrm{EN}) = P(N)P(E) = (0.8)(0.05) = 0.04 \\
P(\mathrm{DN}) &= P(\mathrm{ND}) = P(D)P(N) = (0.15)(0.8) = 0.12
\end{aligned}$$

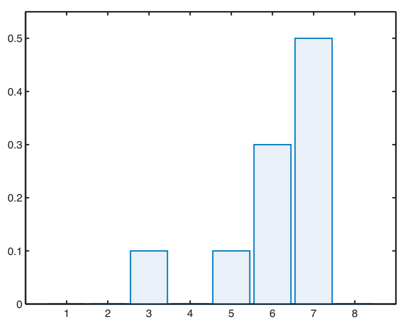
The probability mass function of B = "number of blue-eyed baby monkeys born to the couple in a 2-year period" is given in the table below.

x	$P(B = x)$
0	$P(\mathrm{DD}, \mathrm{NN}, \mathrm{DN}, \mathrm{ND}) = 0.0225 + 0.64 + 2(0.12) = 0.9025$
1	$P(\mathrm{DE}, \mathrm{ED}, \mathrm{NE}, \mathrm{EN}) = 2(0.0075) + 2(0.04) = 0.095$
2	$P(\mathrm{EE}) = 0.0025$

19. See below. The word that best describes the histogram is "uniform".



21. See below. The word that best describes the histogram is "skewed right".



23. The discontinuities of $F(x)$ occur at $x = 0.7$, 1, and 1.2. The sizes of the jumps determine the non-zero probabilities, and hence the probability mass function of X: $P(X = 0.7) = 0.3 - 0 = 0.3$, $P(X = 1) = 0.7 - 0.3 = 0.4$, and $P(X = 1.2) = 1 - 0.7 = 0.3$.

25. The discontinuities of $F(x)$ occur at $x = 1/2$, 1, 3/2, and 3. The sizes of the jumps determine the non-zero probabilities, and hence the probability mass function of X: $P(X = 1/2) = 0.1 - 0 = 0.1$, $P(X = 1) = 0.5 - 0.1 = 0.4$, $P(X = 3/2) = 0.8 - 0.5 = 0.3$, and $P(X = 3) = 1 - 0.8 = 0.2$.

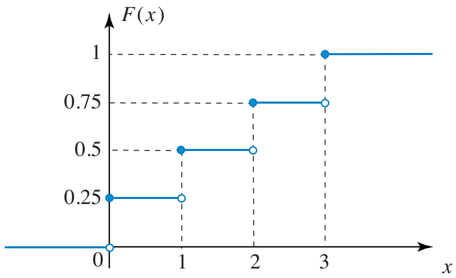
27. The values of $F(x)$ are zero until x (moving from $-\infty$ to ∞) reaches the smallest value in the range of X, which is $x = 0$. There, $F(0) = P(X \le 0) = P(X = 0) = 0.25$. Then, $F(x)$ remains constant until x reaches the next value in the range of X, which is $x = 1$. The value of F is

$$F(1) = P(X \le 1) = P(X = 0) + P(X = 1) = 0.25 + 0.25 = 0.5$$

Continuing in this way, we obtain the following:

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.25 & 0 \le x < 1 \\ 0.5 & 1 \le x < 2 \\ 0.75 & 2 \le x < 3 \\ 1 & x \ge 3 \end{cases}$$

See the graph below.



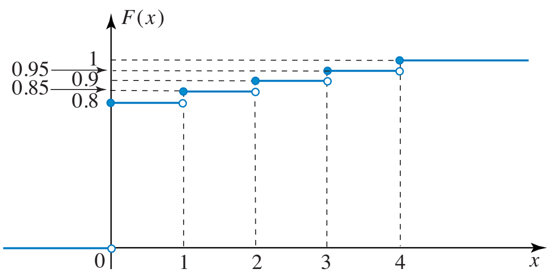
29. The values of $F(x)$ are zero until x (moving from $-\infty$ to ∞) reaches the smallest value in the range of X, which is $x = 0$. There, $F(0) = P(X \le 0) = P(X = 0) = 0.8$. Then, $F(x)$ remains constant until x reaches the next value in the range of X, which is $x = 1$. The value of F is

$$F(1) = P(X \le 1) = P(X = 0) + P(X = 1) = 0.8 + 0.05 = 0.85$$

Continuing in the same way, we obtain the following:

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.8 & 0 \le x < 1 \\ 0.85 & 1 \le x < 2 \\ 0.9 & 2 \le x < 3 \\ 0.95 & 3 \le x < 4 \\ 1 & x \ge 4 \end{cases}$$

See the graph below.



31. (a) Initial location: 0; locations after 1 step: -1 and 1; locations after 2 steps: -2, 0 and 2; locations after 3 steps: -3, -1, 1, and 3; locations after 4 steps: -4, -2, 0, 2, and 4; locations after 5 steps: -5, -3, -1, 1, 3, and 5. To move one step ahead, we add 1 to the locations in the previous step or subtract 1 from the locations in the previous step. Adding 1 to an even number (or subtracting 1 from an even number) makes it odd, and vice versa. A particle starts at an even numbered location ($x = 0$). Thus, after an even (odd) number of steps, the particle arrives at an even-numbered (odd-numbered) location.

(b) To be absorbed, the particle needs to reach 3 or -3, which are odd numbers. The particle can reach 3 or -3 in 3 steps (in which case $X = 3$). If it does not, it means that it ended at -1 or 1 after 3 steps (since after an odd number of steps a particle can only be in an odd-numbered location). Thus, the particle needs 2 more steps to reach 3 or -3 (in which case $X = 5$); if it does not, it means that it ended at -1 or 1; repeating this routine, we see that X can assume only odd-numbered values.

33. We read the values from the table. The probability mass function of X is given by $P(X = 1) = 0.3$, $P(X = 2) = 0.1$, $P(X = 3) = 0.2$, $P(X = 4) = 0.1$, and $P(X = 5) = 0.3$. The discontinuities of the cumulative distribution function $F(x)$ of X occur at $x = 1, 2, 3, 4$ and 5. We find that

$$F(x) = \begin{cases} 0 & x < 1 \\ 0.3 & 1 \le x < 2 \\ 0.4 & 2 \le x < 3 \\ 0.6 & 3 \le x < 4 \\ 0.7 & 4 \le x < 5 \\ 1 & x \ge 5 \end{cases}$$

Section 7 The Mean, the Median, and the Mode

1. Ordering S_1, we get $S_1 = \{2, 3, 4, 5, 6, 7, 10\}$; the median is 5. Ordering S_2, we get $S_2 = \{2, 3, 4, 5, 6, 700{,}000, 1{,}000{,}000\}$; the median is 5 as well. The median fails to capture large difference in the values at the right ends of the two distributions.

3. Adding up the values of all elements in S_1 and dividing by the number of elements in S_1 we get the mean of S_1. To calculate the mean of S_2, the numerator doubles whereas the denominator remains the same. Thus, the mean of S_2 is double the mean of S_1.

The location of the midpoint of the two distributions does not change, since multiplication by 2 does not change the order (assume that S_1 and S_2 are ordered; if a is before b in the list for S_1, then $2a$ is before $2b$ in the list for S_2). Thus, the median of S_2 is double the median of S_1.

If a is the value that appears most often in S_1, then the value $2a$ appears most often in S_2. So, the mode of S_2 is double the mode of S_1.

5. Intutively: since all outcomes are equally likely, the mean is $(1+2+3+\cdots+10)/10 = 55/10 = 5.5$. Formally,

$$E(X) = \sum_{k=1}^{10} k \cdot P(X = k) = \sum_{k=1}^{10} k \cdot \frac{1}{10} = \frac{1}{10}(1 + 2 + 3 + \cdots + 10) = \frac{1}{10} \cdot \frac{10 \cdot 11}{2} = 5.5$$

(Recall the formula: $\sum_{k=1}^{n} k = 1 + 2 + 3 + \cdots + n = n(n + 1)/2$.)

7. No. Consider the random variable X given by $P(X = 0) = 0.5$ and $P(X = 6) = 0.5$ Then $E(X) = 0 \cdot 0.5 + 6 \cdot 0.5 = 3$. The distribution of X^2 is $P(X^2 = 0) = 0.5$ and $P(X^2 = 36) = 0.5$ and so $E(X^2) = 0 \cdot 0.5 + 36 \cdot 0.5 = 18$. (This is just one of many counterexamples.)

9. Using properties of the expected value,

$$E(2X^2 - 4X + 1) = E(2X^2) - E(4X) + E(1) = 2E(X^2) - 4E(X) + E(1)$$

Since

$$E(1) = \sum_x 1 \cdot P(X = x) = \sum_x P(X = x) = 1$$

we get $E(2X^2 - 4X + 1) = 2(3) - 4(2) + 1 = -1$.

11. Using the properties of the expected value,

$$E(Y) = E\left(\frac{1}{\sigma}(X - \mu)\right) = \frac{1}{\sigma}E\,(X - \mu) = \frac{1}{\sigma}(E(X) - \mu) = 0,$$

since, by assumption, $E(X) = \mu$. (Recall that $E(X + b) = E(X) + b$ for a real number b; replacing b by $-\mu$, we get $E(X - \mu) = E(X) - \mu$, which is how the second last equality was obtained.)

13. We compute

$$E(X) = \sum_{x=0}^{3} x \cdot P(X = x) = 0 \cdot 0.25 + 1 \cdot 0.25 + 2 \cdot 0.25 + 3 \cdot 0.25 = 6(0.25) = 1.5$$

$$E(X^2) = \sum_{x=0}^{3} x^2 \cdot P(X = x) = 0 \cdot 0.25 + 1 \cdot 0.25 + 4 \cdot 0.25 + 9 \cdot 0.25 = 14(0.25) = 3.5$$

$$E(X(X-1)) = \sum_{x=0}^{3} x(x-1) \cdot P(X = x) = 0 \cdot 0.25 + 0 \cdot 0.25 + 2 \cdot 0.25 + 6 \cdot 0.25 = 8(0.25) = 2$$

To check:

$$E(X(X - 1)) = E(X^2 - X) = E(X^2) - E(X) = 3.5 - 1.5 = 2$$

15. We compute

$$E(X) = \sum_{x=0}^{4} x \cdot P(X = x) = 0 \cdot 0.8 + 1 \cdot 0.05 + 2 \cdot 0.05 + 3 \cdot 0.05 + 4 \cdot 0.05 = 10(0.05) = 0.5$$

$$E(X^2) = \sum_{x=0}^{4} x^2 \cdot P(X = x) = 0 \cdot 0.8 + 1 \cdot 0.05 + 4 \cdot 0.05 + 9 \cdot 0.05 + 16 \cdot 0.05 = 30(0.05) = 1.5$$

$$E(X(X-1)) = \sum_{x=0}^{4} x(x-1) \cdot P(X = x) = 0 \cdot 0.8 + 0 \cdot 0.05 + 2 \cdot 0.05 + 6 \cdot 0.05 + 12 \cdot 0.05$$
$$= 20(0.05) = 1$$

To check:

$$E(X(X-1)) = E(X^2 - X) = E(X^2) - E(X) = 1.5 - 0.5 = 1$$

17. Instead of ordering the list (call it S), we record the frequencies:

Value	14	18	19	20	22	25	27	29	30
Frequency	1	15	3	5	1	1	3	2	5

Clearly, the mode is 18.

The data set S contains 36 elements. In identifying the median, we calculate the mean of the 18th and the 19th entries. Since both are equal to 19, the median of S is 19. The mean is

$$\overline{S} = \frac{1}{36}(1 \cdot 14 + 15 \cdot 18 + 3 \cdot 19 + 5 \cdot 20 + 1 \cdot 22 + 1 \cdot 25 + 3 \cdot 27 + 2 \cdot 29 + 5 \cdot 30)$$
$$= \frac{777}{36} \approx 21.58$$

19. From

$$E(X) = \sum_{x=1}^{4} x \cdot P(X = x) = 1 \cdot 0.2 + 2 \cdot 0.4 + 3 \cdot 0.3 + 4 \cdot 0.1 = 2.3$$

$$E(\sin(X)) = \sum_{x=1}^{4} \sin x \cdot P(X = x) = \sin 1 \cdot 0.2 + \sin 2 \cdot 0.4 + \sin 3 \cdot 0.3 + \sin 4 \cdot 0.1 \approx 0.49867$$

we compute $E(\sin X) - \sin(E(X)) = 0.49867 - \sin 2.3 \approx -0.24704$.

21. From

$$E(\ln(X)) = \sum_{x=1}^{4} \ln x \cdot P(X = x) = \ln 1 \cdot 0.2 + \ln 2 \cdot 0.4 + \ln 3 \cdot 0.3 + \ln 4 \cdot 0.1 \approx 0.74547$$

we compute $e^{E(\ln X)} = e^{0.74547} \approx 2.10743$.

23. We compute

$$E(1/X) = \sum_{x=1}^{4} \frac{1}{x} \cdot P(X = x) = \frac{1}{1} \cdot 0.2 + \frac{1}{2} \cdot 0.4 + \frac{1}{3} \cdot 0.3 + \frac{1}{4} \cdot 0.1 = 0.525$$

25. Let R represent the per capita production rate of the fish population. Its probability mass function is $P(R = 1.25) = 0.7$ and $P(R = 0.1) = 0.3$. From

$$\begin{aligned} E(\ln(R)) &= \ln 1.25 \cdot P(R = 1.25) + \ln 0.1 \cdot P(R = 0.1) \\ &= (\ln 1.25)(0.7) + (\ln 0.1)(0.3) \\ &\approx -0.53458 \end{aligned}$$

we get the geometric mean $e^{E(\ln R)} = e^{-0.53458} \approx 0.58591$. The geometric mean predicts a decline in the population at the rate of $1 - 0.58591 = 0.41409$ per year.

27. The mode consists of three values: 2, 6, and 7. The probability mass function is given below.

x	$P(X = x)$
1	0.15
2	0.2
5	0.1
6	0.2
7	0.2
8	0.15

The mean is

$$E(X) = (1)(0.15) + 2(0.2) + 5(0.1) + 6(0.2) + 7(0.2) + 8(0.15) = 4.85$$

To find the median, we keep calculating the values of the cumulative distribution function until we reach 0.5: $F(1) = 0.15$, $F(2) = 0.35$, $F(5) = 0.45$, $F(6) = 0.65$. The median is $(5 + 6)/2 = 5.5$.

29. The mode consists of two values: 3 and 6. The probability mass function is given in the table below.

x	$P(X = x)$
1	0.2
3	0.3
6	0.3
8	0.2

The mean is

$$E(X) = (1)(0.2) + 3(0.3) + 6(0.3) + 8(0.2) = 4.5$$

To find the median, we keep calculating the values of the cumulative distribution function until we reach 0.5: $F(1) = 0.2$, $F(3) = 0.5$, $F(6) = 0.8$. The median is $(3 + 6)/2 = 4.5$.

Section 8 The Spread of a Distribution

1. All three samples share the same mean: $\mu_A = \mu_B = \mu_C = 3$. The sample B is least spread out (the two values which differ from the mean are one unit away from it). The sample A is less spread out than C: the four values in A which differ from 3 are closer to the mean than the four values in C which differ from 3. Thus, B has the smallest standard deviation, followed by A; the sample C has the largest standard deviation of the three samples.

We confirm our reasoning by calculating the three standard deviations:

$$\begin{aligned}\operatorname{var}(A) &= \sum_a (a-\mu_A)^2 P(A=a) = \frac{1}{5}\sum_a (a-3)^2\\ &= \frac{1}{5}\left((2-3)^2+(2-3)^2+(3-3)^2+(4-3)^2+(4-3)^2\right) = \frac{4}{5}\end{aligned}$$

and $\sigma_A = \sqrt{\operatorname{var}(A)} = 2/\sqrt{5}$. Likewise,

$$\begin{aligned}\operatorname{var}(B) &= \sum_b (b-\mu_B)^2 P(B=b) = \frac{1}{5}\sum_b (b-3)^2\\ &= \frac{1}{5}\left((2-3)^2+(3-3)^2+(3-3)^2+(3-3)^2+(4-3)^2\right) = \frac{2}{5}\end{aligned}$$

and $\sigma_B = \sqrt{\operatorname{var}(B)} = \sqrt{2}/\sqrt{5}$. Finally,

$$\begin{aligned}\operatorname{var}(C) &= \sum_c (c-\mu_C)^2 P(C=c) = \frac{1}{5}\sum_c (c-3)^2\\ &= \frac{1}{5}\left((1-3)^2+(1-3)^2+(3-3)^2+(5-3)^2+(5-3)^2\right) = \frac{16}{5}\end{aligned}$$

and $\sigma_C = \sqrt{\operatorname{var}(C)} = 4/\sqrt{5}$. Thus, $\sigma_B < \sigma_A < \sigma_C$.

3. Consider multiplying X by a real number a. The formula $\operatorname{var}(aX) = a^2\operatorname{var}(X)$ gives $2 = a^2(22)$, so $a^2 = 1/22$ and $a = \pm 1/\sqrt{11}$. Define $Y = \pm(1/\sqrt{11})X$; the variance of Y is 2. To check:

$$\operatorname{var}(Y) = \operatorname{var}\left(\pm\frac{1}{\sqrt{11}}X\right) = \left(\pm\frac{1}{\sqrt{11}}\right)^2 \operatorname{var}(X) = \frac{1}{11}\cdot 22 = 2$$

(Note that adding a real number to X does not change its variance; that's why we considered the multiplication by a real number).

5. The expected value of X is zero:

$$E(X) = \sum_{k=-4}^{4} kP(X=k) = \frac{1}{9}\sum_{k=-4}^{4} k = \frac{1}{9}(-4-3-2-1+0+1+2+3+4) = 0$$

Therefore

$$\operatorname{var}(X) = \sum_{k=-4}^{4}(k-E(X))^2 P(X=k) = \frac{1}{9}\sum_{k=-4}^{4} k^2 = \frac{1}{9}(16+9+4+1+0+1+4+9+16) = \frac{60}{9}$$

7. From

$$E(X) = \sum_x xP(X=x) = (0)(0.15)+(1)(0.15)+(2)(0.15)+(4)(0.15) = (7)(0.15) = 1.05$$

and

$$E(X^2) = \sum_x x^2P(X=x) = (0)(0.15)+(1)(0.15)+(4)(0.15)+(16)(0.15) = (21)(0.15) = 3.15$$

we compute

$$\operatorname{var}(X) = E(X^2) - [E(X)]^2 = 3.15 - (1.05)^2 = 2.0475$$

and $\sigma_X = \sqrt{2.0475} \approx 1.43091$.

9. From

$$E(X) = \sum_x xP(X = x) = (-2)(0.25) + (-1)(0.2) + (0)(0.1) + (1)(0.2) + (2)(0.25) = 0$$

and

$$E(X^2) = \sum_x x^2 P(X = x) = (4)(0.25) + (1)(0.2) + (0)(0.1) + (1)(0.2) + (4)(0.25) = 2.4$$

we compute

$$\text{var}(X) = E(X^2) - [E(X)]^2 = 2.4 - (0)^2 = 2.4$$

and $\sigma_X = \sqrt{2.4} \approx 1.54919$.

11. Let $E(X) = \mu$ and $X_1 = X - E(X) = X - \mu$.
Direct proof:

$$\begin{aligned} E(X_1) &= \sum_{x_1} x_1 P(X_1 = x_1) \\ &= \sum_x (x - \mu)P(X - \mu = x - \mu) \\ &= \sum_x (x - \mu)P(X = x) \\ &= \sum_x xP(X = x) - \sum_x \mu P(X = x) \\ &= E(X) - \mu \sum_x P(X = x) \\ &= \mu - \mu \cdot 1 = 0 \end{aligned}$$

Using Theorem 7:

$$E(X_1) = E(X - \mu) = E(X) - \mu = \mu - \mu = 0$$

13. Replacing X in $\text{var}(X) = E(X^2) - [E(X)]^2$ by aX, we get

$$\begin{aligned} \text{var}(aX) &= E[(aX)^2] - [E(aX)]^2 \\ &= E[a^2 X^2] - [aE(X)]^2 \\ &= a^2 E[X^2] - a^2[E(X)]^2 \\ &= a^2(E(X^2) - [E(X)]^2) = a^2 \text{var}(X) \end{aligned}$$

15. The sample of 12 healthy adults, sorted:

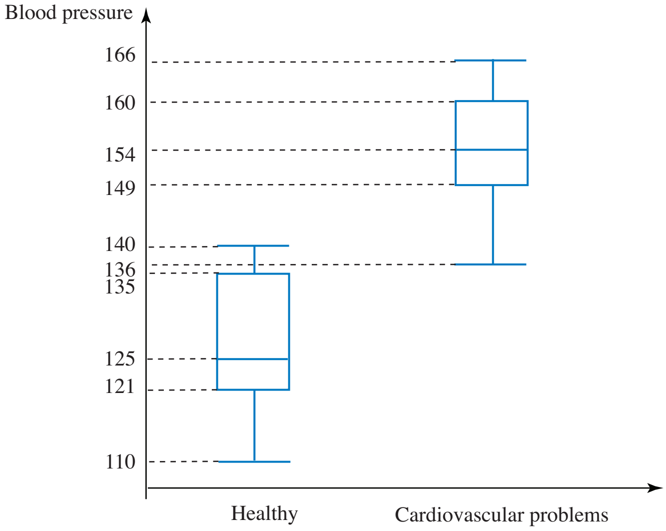
$$110, 116, 120, 122, 123, 125, 125, 128, 132, 138, 138, 140$$

The minimum is 110, and the maximum is 140. The median is the mean of the sixth and the seventh numbers: 125. The lower quartile is the mean of the third and the fourth numbers: $Q_1 = (120 + 122)/2 = 121$, and the upper quartile is the mean of the ninth and the tenth numbers: $Q_3 = (132 + 138)/2 = 135$.

The sample of 12 adults with a history of cardiovascular problems, sorted:

$$136, 142, 148, 150, 154, 154, 154, 158, 160, 160, 162, 166$$

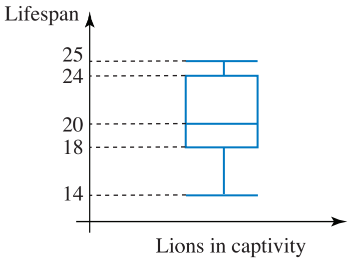
The minimum is 136, and the maximum is 166. The median is the mean of the sixth and the seventh numbers: 154. The lower quartile is the mean of the third and the fourth numbers: $Q_1 = (148+150)/2 = 149$, and the upper quartile is the mean of the ninth and the tenth numbers: $Q_3 = 160$. See the figure below for the boxplots.



17. The sample, sorted:

$$14, 16, 17, 18, 19, 20, 20, 20, 22, 22, 24, 24, 24, 25$$

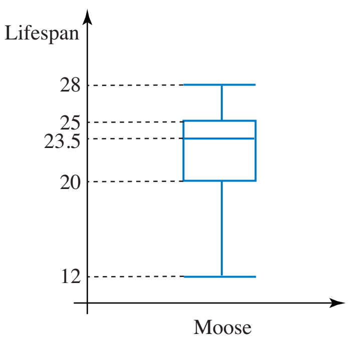
The sample contains 14 numbers. The minimum is 14, and the maximum is 25. The median is the mean of the seventh and the eighth numbers: 20. The lower quartile is the fourth number: $Q_1 = 18$, and the upper quartile is the eleventh number: $Q_3 = 24$.



19. The sample, sorted:

$$12, 20, 20, 20, 21, 23, 23, 24, 24, 25, 25, 26, 27, 28$$

The sample contains 14 numbers. The minimum is 12, and the maximum is 28. The median is the mean of the seventh and the eighth numbers: 23.5. The lower quartile is the fourth number: $Q_1 = 20$, and the upper quartile is the eleventh number: $Q_3 = 25$.



21. We extract the probability mass function from the histogram.

x	$P(X = x)$
1	0.15
2	0.1
3	0.05
4	0.15
5	0.1
6	0.2
7	0.05
8	0.2

From

$$\begin{aligned} E(X) &= \sum_x xP(X = x) \\ &= (1)(0.15) + (2)(0.1) + (3)(0.05) + (4)(0.15) + (5)(0.1) + (6)(0.2) + (7)(0.05) + (8)(0.2) \\ &= 4.75 \end{aligned}$$

and

$$\begin{aligned} E(X^2) &= \sum_x x^2 P(X = x) \\ &= (1)(0.15) + (4)(0.1) + (9)(0.05) + (16)(0.15) + (25)(0.1) + (36)(0.2) + (49)(0.05) + (64)(0.2) \\ &= 28.35 \end{aligned}$$

we compute

$$\text{var}(X) = E(X^2) - [E(X)]^2 = 28.35 - (4.75)^2 = 5.7875$$

and $\sigma_X = \sqrt{5.7875} \approx 2.40572$.

23. We extract the probability mass function from the histogram.

x	$P(X = x)$
1	0.05
2	0.05
3	0.1
4	0.1
5	0.15
6	0.15
7	0.2
8	0.2

From

$$\begin{aligned} E(X) &= \sum_x xP(X = x) \\ &= (1)(0.15) + (2)(0.15) + (3)(0.1) + (4)(0.1) + (5)(0.15) + (6)(0.15) + (7)(0.2) + (8)(0.2) = 5.8 \end{aligned}$$

and

$$\begin{aligned}
E(X^2) &= \sum_x x^2 P(X = x) \\
&= (1)(0.05) + (4)(0.05) + (9)(0.1) + (16)(0.1) + (25)(0.15) + (36)(0.15) + (49)(0.2) + (64)(0.2) \\
&= 34.5
\end{aligned}$$

we compute

$$\mathrm{var}(X) = E(X^2) - [E(X)]^2 = 34.5 - (5.8)^2 = 0.86$$

and $\sigma_X = \sqrt{0.86} \approx 0.92736$.

25. The mean of all three distributions is 24.5. For the Milky Way Farm,

$$\begin{aligned}
\mathrm{MAD}(X_1) &= E(|X_1 - E(X_1)|) = E(|X_1 - 24.5|) \\
&= |18 - 24.5| \cdot \frac{6}{30} + |20 - 24.5| \cdot \frac{5}{30} + |22 - 24.5| \cdot \frac{2}{30} + |24 - 24.5| \cdot \frac{1}{30} \\
&\quad + |25 - 24.5| \cdot \frac{1}{30} + |27 - 24.5| \cdot \frac{4}{30} + |29 - 24.5| \cdot \frac{4}{30} + |30 - 24.5| \cdot \frac{7}{30} \\
&= \frac{134}{30}
\end{aligned}$$

For the Milkshake Farm,

$$\begin{aligned}
\mathrm{MAD}(X_2) &= E(|X_2 - 24.5|) \\
&= |22 - 24.5| \cdot \frac{4}{30} + |23 - 24.5| \cdot \frac{6}{30} + |24 - 24.5| \cdot \frac{3}{30} + |25 - 24.5| \cdot \frac{8}{30} \\
&\quad + |26 - 24.5| \cdot \frac{6}{30} + |27 - 24.5| \cdot \frac{3}{30} \\
&= \frac{41}{30}
\end{aligned}$$

For the Butterscotch Farm,

$$\begin{aligned}
\mathrm{MAD}(X_3) &= E(|X_3 - 24.5|) \\
&= |17 - 24.5| \cdot \frac{2}{30} + |18 - 24.5| \cdot \frac{7}{30} + |19 - 24.5| \cdot \frac{4}{30} + |20 - 24.5| \cdot \frac{3}{30} \\
&\quad + |30 - 24.5| \cdot \frac{2}{30} + |31 - 24.5| \cdot \frac{5}{30} + |32 - 24.5| \cdot \frac{7}{30} \\
&= \frac{192}{30}
\end{aligned}$$

Thus, the MAD is able to detect the differences in the spreads of the three distributions. Note that the order of the three distributions from the smallest to the largest standard deviation is the same as the order of the three distributions from the smallest to the largest mean absolute deviation.

Section 9 Joint Distributions

1. Using independence, we find

$$\begin{aligned} P(X=1, Y=1) &= P(X=1)P(Y=1) = (0.2)(0.7) = 0.14 \\ P(X=1, Y=2) &= P(X=1)P(Y=2) = (0.2)(0.3) = 0.06 \\ P(X=2, Y=1) &= P(X=2)P(Y=1) = (0.8)(0.7) = 0.56 \\ P(X=2, Y=2) &= P(X=2)P(Y=2) = (0.8)(0.3) = 0.24 \end{aligned}$$

These four probabilities form the joint probability distribution of X and Y. See below.

	$X=1$	$X=2$
$Y=1$	0.14	0.56
$Y=2$	0.06	0.24

3. Denote the missing entries by a and b and expand the table to include the horizontal and the vertical totals:

	$X=0$	$X=1$	
$Y=0$	0.1	0.3	$P(Y=0)=0.4$
$Y=1$	a	b	$P(Y=1)=a+b=0.6$
	$P(X=0)=0.1+a$	$P(X=1)=0.4+b$	

By independence

$$\begin{aligned} P(X=0, Y=0) &= P(X=0)P(Y=0) \\ 0.1 &= (0.1+a)(0.4) \\ 0.25 &= 0.1+a \end{aligned}$$

and thus $a=0.15$. From $P(Y=1)=a+b=0.6$ we get $b=0.45$.

5. Using independence, we find

$$\begin{aligned} P(X=1, Y=1) &= P(X=1)P(Y=1) = (0.2)(0.9) = 0.18 \\ P(X=1, Y=2) &= P(X=1)P(Y=2) = (0.2)(0.1) = 0.02 \\ P(X=2, Y=1) &= P(X=2)P(Y=1) = (0.8)(0.9) = 0.72 \\ P(X=2, Y=2) &= P(X=2)P(Y=2) = (0.8)(0.1) = 0.08 \end{aligned}$$

The four probabilities form the joint probability distribution of X and Y, shown in the table below (expanded, to include horizontal and vertical totals):

	$X=1$	$X=2$	
$Y=1$	0.18	0.72	$P(Y=1)=0.9$
$Y=2$	0.02	0.08	$P(Y=2)=0.1$
	$P(X=1)=0.2$	$P(X=2)=0.8$	

We find

$$\begin{aligned} P(X=1 \mid Y=1) &= \frac{P(X=1, Y=1)}{P(Y=1)} = \frac{0.18}{0.9} = 0.2 \\ P(X=1 \mid Y=2) &= \frac{P(X=1, Y=2)}{P(Y=2)} = \frac{0.02}{0.1} = 0.2 \end{aligned}$$

Recall the law of total probability: If A is an event, and E_1 and E_2 form a partition, then

$$P(A) = P(A \mid E_1)P(E_1) + P(A \mid E_2)P(E_2)$$

Substituting $A = \{X = 1\}$, $E_1 = \{Y = 1\}$, and $E_2 = \{Y = 2\}$, we obtain the desired relation

$$P(X = 1) = P(X = 1 \mid Y = 1)P(Y = 1) + P(X = 1 \mid Y = 2)P(Y = 2)$$

between $P(X = 1 \mid Y = 1)$, $P(X = 1 \mid Y = 2)$, and $P(X = 1)$. We illustrate it by substituting the probabilities we calculated:

$$0.2 = (0.2)(0.9) + (0.2)(0.1)$$

7. We need to find $P(R = + \text{ and } G = \mathrm{B} \mid G = \mathrm{B})$. From $P(G = \mathrm{B}) = 0.076 + 0.014 = 0.09$ we get

$$P(R = + \text{ and } G = \mathrm{B} \mid G = \mathrm{B}) = \frac{P(R = +, G = \mathrm{B})}{P(G = \mathrm{B})} = \frac{0.076}{0.09} \approx 0.844$$

9. The two probabilities

$$P(R = + \mid G = \mathrm{B}) = \frac{P(R = +, G = \mathrm{B})}{P(G = \mathrm{B})} = \frac{0.076}{0.09} = \frac{76}{90}$$
$$P(R = - \mid G = \mathrm{B}) = \frac{P(R = -, G = \mathrm{B})}{P(G = \mathrm{B})} = \frac{0.014}{0.09} = \frac{14}{90}$$

define the distribution of R conditional on $G = \mathrm{B}$. (Note that $P(G = \mathrm{B}) = 0.076 + 0.014 = 0.09$.)

11. By adding up the entries horizontally, we obtain the marginal distribution for A:

$$\begin{aligned} P(A = \text{allergy}) &= P(A = \text{allergy}, T = \text{positive}) \\ &\quad + P(A = \text{allergy}, T = \text{negative}) + P(A = \text{allergy}, T = \text{inconclusive}) \\ &= 0.3 + 0.07 + 0.1 = 0.47 \\ P(A = \text{no allergy}) &= P(A = \text{no allergy}, T = \text{positive}) \\ &\quad + P(A = \text{no allergy}, T = \text{negative}) + P(A = \text{no allergy}, T = \text{inconclusive}) \\ &= 0.03 + 0.45 + 0.05 = 0.53 \end{aligned}$$

Thus, there is a 47% chance that a randomly selected person from the group has allergy.

By adding up the entries vertically, we obtain the marginal distribution for T:

$$\begin{aligned} P(T = \text{positive}) &= P(A = \text{allergy}, T = \text{positive}) + P(A = \text{no allergy}, T = \text{positive}) \\ &= 0.3 + 0.03 = 0.33 \\ P(T = \text{negative}) &= P(A = \text{allergy}, T = \text{negative}) + P(A = \text{no allergy}, T = \text{negative}) \\ &= 0.07 + 0.45 = 0.52 \\ P(T = \text{inconclusive}) &= P(A = \text{allergy}, T = \text{inconclusive}) + P(A = \text{no allergy}, T = \text{inconclusive}) \\ &= 0.1 + 0.05 = 0.15 \end{aligned}$$

Thus, for 33% of the population the test is positive, and for 52% it is negative; for 15% of the population the test is inconclusive.

13. We compute

$$P(A = \text{allergy} \mid T = \text{negative}) = \frac{P(A = \text{allergy}, T = \text{negative})}{P(T = \text{negative})} = \frac{0.07}{0.07 + 0.45} = \frac{7}{52} \approx 0.13461$$

15. We need to find the probabilities a, b, c and d which define the joint distribution:

	$X = 1$	$X = 2$
$Y = 3$	a	b
$Y = 4$	c	d

From the given information, we get the following equations:

$$\begin{aligned} P(X=1) = 0.4 &\quad \text{implies that} \quad a + c = 0.4 \\ P(X=2) = 0.6 &\quad \text{implies that} \quad b + d = 0.6 \\ P(Y=3 \,|\, X=1) = 0.7 &\quad \text{implies that} \quad \frac{P(Y=3, X=1)}{P(X=1)} = \frac{a}{a+c} = 0.7 \\ P(Y=3 \,|\, X=2) = 0.1 &\quad \text{implies that} \quad \frac{P(Y=3, X=2)}{P(X=2)} = \frac{b}{b+d} = 0.1 \end{aligned}$$

Combining the first and the third equation we get $a/0.4 = 0.7$ and thus $a = 0.28$. From $a + c = 0.4$ it follows that $c = 0.12$. Combining the second and the fourth equation we get $b/0.6 = 0.1$ and thus $b = 0.06$. From $b + d = 0.6$ it follows that $d = 0.54$. The joint distribution is

	$X = 1$	$X = 2$
$Y = 3$	0.28	0.06
$Y = 4$	0.12	0.54

17. By adding up the entries along the rows we obtain the distribution for F:

$$\begin{aligned} P(F = \text{fish}) &= P(F = \text{fish}, P = \text{brown bear}) \\ &\qquad + P(F = \text{fish}, P = \text{wolf}) + P(F = \text{fish}, P = \text{fox}) \\ &= 0.2 + 0.02 + 0.03 = 0.25 \\ P(F = \text{insects}) &= P(F = \text{insects}, P = \text{brown bear}) \\ &\qquad + P(F = \text{insects}, P = \text{wolf}) + P(F = \text{insects}, P = \text{fox}) \\ &= 0.1 + 0.05 + 0.05 = 0.2 \\ P(F = \text{small mammals}) &= P(F = \text{small mammals}, P = \text{brown bear}) \\ &\qquad + P(F = \text{small mammals}, P = \text{wolf}) + P(F = \text{small mammals}, P = \text{fox}) \\ &= 0.2 + 0.25 + 0.1 = 0.55 \end{aligned}$$

By adding up the entries vertically we obtain the distribution for P:

$$\begin{aligned} P(P = \text{brown bear}) &= P(P = \text{brown bear}, F = \text{fish}) \\ &\qquad + P(P = \text{brown bear}, F = \text{insects}) \\ &\qquad + P(P = \text{brown bear}, F = \text{small mammals}) \\ &= 0.2 + 0.1 + 0.2 = 0.5 \\ P(P = \text{wolf}) &= P(P = \text{wolf}, F = \text{fish}) \\ &\qquad + P(P = \text{wolf}, F = \text{insects}) + P(P = \text{wolf}, F = \text{small mammals}) \\ &= 0.02 + 0.05 + 0.25 = 0.32 \\ P(P = \text{fox}) &= P(P = \text{fox}, F = \text{fish}) \\ &\qquad + P(P = \text{fox}, F = \text{insects}) + P(P = \text{fox}, F = \text{small mammals}) \\ &= 0.03 + 0.05 + 0.1 = 0.18 \end{aligned}$$

19. The conditional probabilities are:

$$P(F = \text{fish} \,|\, P = \text{wolf}) = \frac{P(F = \text{fish}, P = \text{wolf})}{P(P = \text{wolf})} = \frac{0.02}{0.02 + 0.05 + 0.25} = \frac{0.02}{0.32}$$
$$P(F = \text{insects} \,|\, P = \text{wolf}) = \frac{P(F = \text{insects}, P = \text{wolf})}{P(P = \text{wolf})} = \frac{0.05}{0.02 + 0.05 + 0.25} = \frac{0.05}{0.32}$$
$$P(F = \text{small mammals} \,|\, P = \text{wolf}) = \frac{P(F = \text{small mammals}, P = \text{wolf})}{P(P = \text{wolf})} = \frac{0.25}{0.02 + 0.05 + 0.25} = \frac{0.25}{0.32}$$

The probabilities add up to 1, because a wolf will have one of the three for food.

21. The probability that a bear will prey on a small mammal is

$$P(F = \text{small mammals} \,|\, P = \text{bear}) = \frac{P(F = \text{small mammals}, P = \text{bear})}{P(P = \text{bear})} = \frac{0.2}{0.2 + 0.1 + 0.2} = \frac{2}{5}$$

or 40%.

23. (a) The marginal distribution for X is given by

$$P(X = 0) = P(X = 0, Y = 0) + P(X = 0, Y = 1) = 0.05 + 0.45 = 0.5$$
$$P(X = 1) = P(X = 1, Y = 0) + P(X = 1, Y = 1) = 0.1 + 0.4 = 0.5$$

The marginal distribution for Y is given by

$$P(Y = 0) = P(X = 0, Y = 0) + P(X = 1, Y = 0) = 0.05 + 0.1 = 0.15$$
$$P(Y = 1) = P(X = 0, Y = 1) + P(X = 1, Y = 1) = 0.45 + 0.4 = 0.85$$

(b) The random variables X and Y are not independent; for instance, $P(X = 0, Y = 0) = 0.05$ is not equal to $P(X = 0)P(Y = 0) = (0.5)(0.15) = 0.075$.

25. (a) The marginal distribution for X is given by

$$P(X = 0) = P(X = 0, Y = 0) + P(X = 0, Y = 1) + P(X = 0, Y = 2) = 0.12 + 0.22 + 0.02 = 0.36$$
$$P(X = 1) = P(X = 1, Y = 0) + P(X = 1, Y = 1) + P(X = 1, Y = 2) = 0.18 + 0.28 + 0.18 = 0.64$$

The marginal distribution for Y is given by

$$P(Y = 0) = P(X = 0, Y = 0) + P(X = 1, Y = 0) = 0.12 + 0.18 = 0.3$$
$$P(Y = 1) = P(X = 0, Y = 1) + P(X = 1, Y = 1) = 0.22 + 0.28 = 0.5$$
$$P(Y = 2) = P(X = 0, Y = 2) + P(X = 1, Y = 2) = 0.02 + 0.18 = 0.2$$

(b) The random variables X and Y are not independent; for instance, $P(X = 0, Y = 1) = 0.22$ is not equal to $P(X = 0)P(Y = 1) = (0.36)(0.5) = 0.18$.

27. We find

$$P(Y = 0 \,|\, X = 0) = \frac{P(Y = 0, X = 0)}{P(X = 0)} = \frac{0.2}{0.2 + 0.08 + 0.12} = \frac{0.2}{0.4} = 0.5$$
$$P(Y = 0 \,|\, X = 1) = \frac{P(Y = 0, X = 1)}{P(X = 1)} = \frac{0.3}{0.3 + 0.12 + 0.18} = \frac{0.3}{0.6} = 0.5$$

We see that $P(Y = 0 \,|\, X = 0) + P(Y = 0 \,|\, X = 1) = 0/5 + 0.5 = 1$. From the joint distribution table we compute $P(Y = 0) = 0.2 + 0.3 = 0.5$. By examining the joint distribution closer, we realize that X and Y are independent. Thus $P(Y = 0 \,|\, X = 0) + P(Y = 0 \,|\, X = 1) = P(Y = 0) + P(Y = 0) = 2P(Y = 0)$, which is illustrated by their numeric values above.

29. The probabilities $P(X = 0) = 0.05 + 0.1 + 0.4 = 0.55$ and $P(X = 1) = 0.1 + 0.1 + 0.25 = 0.45$ define the marginal probability distribution of X.

31. The probabilities

$$P(X=0\,|\,Y=2)=\frac{P(X=0,Y=2)}{P(Y=2)}=\frac{0.4}{0.4+0.25}=\frac{0.4}{0.65}=\frac{8}{13}$$
$$P(X=1\,|\,Y=2)=\frac{P(X=1,Y=2)}{P(Y=2)}=\frac{0.25}{0.4+0.25}=\frac{0.25}{0.65}=\frac{5}{13}$$

define the distribution of X conditional on $Y=2$.

33. The marginal probability distributions of X and Y are given in the last row and the last column in the table below.

	$Y=1$	$Y=2$	
$X=-2$	0	0.12	$P(X=-2)=0.12$
$X=-1$	0.1	0.38	$P(X=-1)=0.48$
$X=0$	0.26	0.14	$P(X=0)=0.4$
	$P(Y=1)=0.36$	$P(Y=2)=0.64$	

35. Assume that $P(X=1)=p_1$ and $P(X=2)=p_2$ is the probability distribution of X and $P(Y=3)=q_1$, $P(Y=4)=q_2$, and $P(Y=5)=q_3$ is the probability distribution of Y. Then $E(X)=p_1+2p_2$ and $E(Y)=3q_1+4q_2+5q_3$. The range of XY is $\{3,4,5,6,8,10\}$ and its distribution is given by (here we use independence)

$$\begin{aligned}
P(XY=3)&=P(X=1\text{ and }Y=3)=P(X=1)P(Y=3)=p_1q_1\\
P(XY=4)&=P(X=1\text{ and }Y=4)=P(X=1)P(Y=4)=p_1q_2\\
P(XY=5)&=P(X=1\text{ and }Y=5)=P(X=1)P(Y=5)=p_1q_3\\
P(XY=6)&=P(X=2\text{ and }Y=3)=P(X=2)P(Y=3)=p_2q_1\\
P(XY=8)&=P(X=2\text{ and }Y=4)=P(X=2)P(Y=4)=p_2q_2\\
P(XY=10)&=P(X=2\text{ and }Y=5)=P(X=2)P(Y=5)=p_2q_3
\end{aligned}$$

It follows that

$$E(XY)=3p_1q_1+4p_1q_2+5p_1q_3+6p_2q_1+8p_2q_2+10p_2q_3$$

Since

$$E(X)E(Y)=(p_1+2p_2)(3q_1+4q_2+5q_3)=3p_1q_1+4p_1q_2+5p_1q_3+6p_2q_1+8p_2q_2+10p_2q_3$$

we conclude that $E(XY)=E(X)E(Y)$.

In general: the range of X is $\{x_1,x_2,\dots,x_m\}$; assume that its distribution is given by $P(X=x_i)=p_i$. The range of Y is $\{y_1,y_2,\dots,y_n\}$; assume that its distribution is given by $P(Y=y_i)=q_i$. Then $E(X)=p_1x_1+p_2x_2+\cdots+p_mx_m$ and $E(Y)=q_1y_1+q_2y_2+\cdots+q_ny_n$. The range of XY consists of all products x_iy_j, where $i=1,2,\dots,m$ and $j=1,2,\dots,n$. The probability distribution is (by independence)

$$P(XY=x_iy_j)=P(X=x_i\text{ and }Y=y_j)=P(X=x_i)P(Y=y_j)=p_iq_j$$

and

$$\begin{aligned}
E(XY)=&\,x_1y_1p_1q_1+x_1y_2p_1q_2+\cdots+x_1y_np_1q_n\\
&+x_2y_1p_2q_1+x_2y_2p_2q_2+\cdots+x_2y_np_2q_n\\
&+\cdots\\
&+x_my_1p_mq_1+x_my_2p_mq_2+\cdots+x_my_np_mq_n
\end{aligned}$$

Computing the product

$$E(X)E(Y)=(p_1x_1+p_2x_2+\cdots+p_mx_m)(q_1y_1+q_2y_2+\cdots+q_ny_n)$$

we see that $E(XY)=E(X)E(Y)$.

Section 10 The Binomial Distribution

1. No. The binomial distribution requires that the *same* experiment (with the same probability of success) be repeated. In this case, the probability of success (a male is interviewed) *changes:* initially, the probability that a male is selected for an interview is 1/2. Assuming independence, the probability that the second interviewee is a male is 10/19 (if the first interviewee was a woman) or 9/19 (if the first interviewee was a man); however, neither is equal to 50%.

3. Define

$$B = \begin{cases} 1 & \text{goshawk catches a small mammal (success)} \\ 0 & \text{goshawk does not catch a small mammal} \end{cases}$$

B is a Bernoulli random variable with the probability of success equal to 0.6. Repeat the experiment 10 times; by assumption, the outcomes are independent. The random variable X = "number of small mammals captured" counts the number of successes in ten independent repetitions of the same experiment. Thus, X is a binomial variable.

5. Let S represent success and F represent a no-success (failure). Exactly two successes in four trials occur in the following six cases: SSFF, SFSF, SFFS, FSSF, FSFS, and FFSS. They are all equally likely, with the probability

$$P(\text{SSFF}) = P(\text{S})P(\text{S})P(\text{F})P(\text{F}) = (0.3)(0.3)(0.7)(0.7) = (0.3)^2(0.7)^2$$

Thus, probability of obtaining exactly two successes in four trials is $6P(\text{SSFF}) = 6 \cdot (0.3)^2(0.7)^2 = 0.2646$.

Now the binomial distribution approach: the probability of success in a single experiment is 0.3. The probability of 2 successes in 4 independent repetitions of the experiment is given by

$$b(2, 4; 0.3) = \binom{4}{2} (0.3)^2(0.7)^2 = 6 \cdot (0.3)^2(0.7)^2$$

Clearly, the two answers match.

7. $\binom{12}{3}$ represents the number of ways to choose a group of three objects from a group of 12 objects (say, the number of ways of picking three shirts from a collection of 12 shirts in different colours). We compute

$$\binom{12}{3} = \frac{12 \cdot 11 \cdot 10}{1 \cdot 2 \cdot 3} = 2 \cdot 11 \cdot 10 = 220$$

9. We compute

$$C(8, 0) = \binom{8}{0} = \frac{8!}{0! \cdot (8-0)!} = 1$$

since $0! = 1$. In theory, $C(8, 0)$ represents the number of ways to choose zero objects from a group of eight objects (we can *define* that there is one way of not picking any object from a group of 8 objects).

11. The number $b(1, 4; 0.6)$ represents the probability of one success in 4 independent repetitions of the same Bernoulli experiment with the probability of success equal to 0.6. We compute

$$b(1, 4; 0.6) = \binom{4}{1} (0.6)^1(1 - 0.6)^{4-1} = 4\,(0.6)(0.4)^3 = 0.1536.$$

13. The number $b(1,7;0.2)$ represents the probability of one success in 7 independent repetitions of the same Bernoulli experiment whose probability of success is 0.2. We compute

$$b(1,7;0.2) = \binom{7}{1}(0.2)^1(1-0.2)^{7-1} = 7\,(0.2)(0.8)^6 \approx 0.3670.$$

15. Label the tosses by numbers $S = \{1,2,3,4,5,6,7,8\}$. Picking a group of three numbers from S corresponds to one event in which 3 tails occurred in 8 tosses (for instance, picking 4, 5 and 8 describes the event in which tails occurred on the 4th, 5th and 8th tosses). Thus, the number of ways of getting three tails in eight tosses of a coin is equal to the number of ways of selecting a group of 3 numbers from the set S of 8 numbers, which is

$$\binom{8}{3} = \frac{8\cdot 7\cdot 6}{1\cdot 2\cdot 3} = 56$$

17. The number of ways of selecting a team of 4 students from a group of 20 students is

$$\binom{20}{4} = \frac{20\cdot 19\cdot 18\cdot 17}{1\cdot 2\cdot 3\cdot 4} = 5\cdot 19\cdot 3\cdot 17 = 4845$$

19. "At least three successes" means 3, 4, or 5 successes. Thus, the probability of at least three successes in five trials is given by $b(3,5;0.6) + b(4,5;0.6) + b(5,5;0.6)$.

21. The number of successes could be 5, 6, 7, 8, or 9. The probability is given by $b(5,25;0.6) + b(6,25;0.6) + b(7,25;0.6) + b(8,25;0.6) + b(9,25;0.6)$.

23. Formula (10.3) states that

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

Replacing k by $n-k$, we get

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

Thus,

$$\binom{22}{20} = \binom{22}{22-20} = \binom{22}{2} = \frac{22\cdot 21}{1\cdot 2} = 231$$

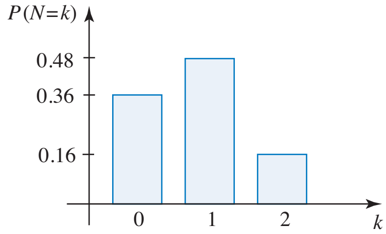
25. (a) The probability distribution of X is given by:

$$P(X=0) = b(0,2;0.4) = \binom{2}{0}(0.4)^0(1-0.4)^2 = (0.6)^2 = 0.36$$

$$P(X=1) = b(1,2;0.4) = \binom{2}{1}(0.4)^1(1-0.4)^1 = 2\,(0.4)(0.6) = 0.48$$

$$P(X=2) = b(2,2;0.4) = \binom{2}{2}(0.4)^2(1-0.4)^0 = (0.4)^2 = 0.16$$

(b) See below.



(c) The mean is

$$E(X) = 0 \cdot 0.36 + 1 \cdot 0.48 + 2 \cdot 0.16 = 0.8$$

From

$$E(X^2) = 0 \cdot 0.36 + 1 \cdot 0.48 + 4 \cdot 0.16 = 1.12$$

we compute the variance

$$\mathrm{var}(X) = E(X^2) - (E(X))^2 = 1.12 - 0.8^2 = 0.48$$

(d) Using (10.4), $E(X) = np = 2 \cdot 0.4 = 0.8$; using (10.5), $\mathrm{var}(X) = np(1-p) = 2 \cdot 0.4 \cdot 0.6 = 0.48$.

27. (a) The probability distribution of X is given by:

$$P(X = 0) = b(0, 4; 0.4) = \binom{4}{0}(0.4)^0(0.6)^4 = (0.6)^4 = 0.1296$$

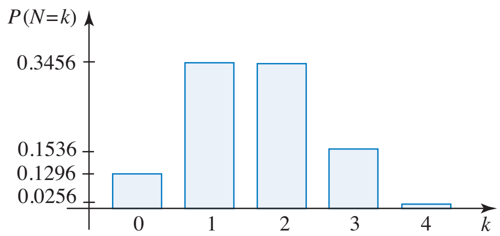
$$P(X = 1) = b(1, 4; 0.4) = \binom{4}{1}(0.4)^1(0.6)^3 = 4\,(0.4)(0.6)^3 = 0.3456$$

$$P(X = 2) = b(2, 4; 0.4) = \binom{4}{2}(0.4)^2(0.6)^2 = 6\,(0.4)^2(0.6)^2 = 0.3456$$

$$P(X = 3) = b(3, 4; 0.4) = \binom{4}{3}(0.4)^3(0.6)^1 = 4\,(0.4)^3(0.6) = 0.1536$$

$$P(X = 4) = b(4, 4; 0.4) = \binom{4}{4}(0.4)^4(0.6)^0 = (0.4)^4 = 0.0256$$

(b) See below.



(c) The mean is

$$E(X) = 0 \cdot 0.1296 + 1 \cdot 0.3456 + 2 \cdot 0.3456 + 3 \cdot 0.1536 + 4 \cdot 0.0256 = 1.6$$

From

$$E(X^2) = 0 \cdot 0.1296 + 1 \cdot 0.3456 + 4 \cdot 0.3456 + 9 \cdot 0.1536 + 16 \cdot 0.0256 = 3.52$$

we compute the variance

$$\mathrm{var}(X) = E(X^2) - (E(X))^2 = 3.52 - 1.6^2 = 0.96$$

(d) Using (10.4), $E(X) = np = 4 \cdot 0.4 = 1.6$; using (10.5), $\mathrm{var}(X) = np(1-p) = 4 \cdot 0.4 \cdot 0.6 = 0.96$.

29. (a) Define

$$H = \begin{cases} 1 & \text{a chocolate has a hazelnut (success)} \\ 0 & \text{a chocolate has no hazelnut} \end{cases}$$

H is a Bernoulli random variable with probability of success equal to 0.03. Let N = "number of chocolates with a hazelnut in a box of 20 chocolates". Since N counts the number of successes in repetitions of H (assumed independent), N is a binomially distributed random variable with $n = 20$ and $p = 0.03$.

The expected number of chocolates with a hazelnut per box is $E(N) = np = 20(0.03) = 0.6$.

(b) The probability that there are no chocolates with a hazelnut in one box of 20 chocolates is the probability of no successes in 20 repetitions:

$$b(0, 20; 0.03) = \binom{20}{0} (0.03)^0 (0.97)^{20} = 0.54379$$

i.e., a bit over 54%.

(c) Define

$$B = \begin{cases} 1 & \text{a box of chocolates has no chocolates with a hazelnut (success)} \\ 0 & \text{a box of chocolates has a chocolate with a hazelnut} \end{cases}$$

B is a Bernoulli random variable with probability of success equal to 0.54379. Let M = "number of boxes of chocolates which do not contain a chocolate with a hazelnut". Since M counts the number of successes in 15 independent repetitions of B, M is a binomially distributed random variable with $n = 15$ and $p = 0.54379$.

The expected number of of boxes that contain no chocolates with a hazelnut is $E(M) = np = 15(0.54379) = 8.15685$; i.e., 8 boxes.

31. Define

$$T = \begin{cases} 1 & \text{a tomato plant has been infested with hornworms (success)} \\ 0 & \text{a tomato plant has not been infested with hornworms} \end{cases}$$

T is a Bernoulli random variable with probability of success equal to 0.15. Let N = "number of tomato plants which have been infested with hornworms". Since N counts the number of successes in independent repetitions of T, N is a binomially distributed random variable; it is given that $n = 10$ and $p = 0.15$. The probability that none of the ten randomly picked tomato plants have been infested with hornworms is (zero successes in ten trials)

$$b(0, 10; 0.15) = \binom{10}{0} (0.15)^0 (0.85)^{10} \approx 0.19687$$

33. The probability distribution of the genotype of a puppy of SC parents is $P(\text{SS}) = 0.25$, $P(\text{SC}) = 0.5$, and $P(\text{CC}) = 0.25$. Thus,

$$P(\text{puppy has straight hair}) = P(\text{SS}) + P(\text{SC}) = 0.75$$
$$P(\text{puppy has curly hair}) = P(\text{CC}) = 0.25$$

Define

$$H = \begin{cases} 1 & \text{a puppy has curly hair (success)} \\ 0 & \text{a puppy does not have curly hair} \end{cases}$$

H is a Bernoulli random variable with probability of success equal to 0.25. Let N = "number of puppies which have curly hair". Since N counts the number of successes in independent repetitions of H, N is a binomially distributed random variable; it is given that $n = 8$ and $p = 0.25$.

The expected number of puppies with curly hair is $E(N) = np = 8(0.25) = 2$. The probability that exactly 2 puppies have curly hair is

$$b(2, 8; 0.25) = \binom{8}{2} (0.25)^2 (0.75)^6 \approx 0.31146$$

35. The probability distribution of the genotype of an offspring of LS parents is $P(\text{LL}) = 0.25 = P(\text{long})$, $P(\text{LS}) = 0.5 = P(\text{medium-sized})$, and $P(\text{SS}) = 0.25 = P(\text{short})$. Define

$$T = \begin{cases} 1 & \text{an offspring is medium-sized (success)} \\ 0 & \text{an offspring is not medium-sized} \end{cases}$$

T is a Bernoulli random variable with probability of success equal to 0.5. Let N = "number of medium-sized offspring". Since N counts the number of successes in repetitions of T, N is a binomially distributed random variable; it is given that $n = 12$ and $p = 0.5$.

(a) The expected number of medium-sized offspring is $E(N) = np = 12(0.5) = 6$.

(b) The probability that that there are at most two medium-sized offspring is

$$\begin{aligned} &b(0, 12; 0.55) + b(1, 12; 0.55) + b(2, 12; 0.55) \\ &\qquad = \binom{12}{0}(0.5)^0(0.5)^{12} + \binom{12}{1}(0.5)^1(0.5)^{11} + \binom{12}{2}(0.5)^2(0.5)^{10} \\ &\qquad = (1 + 12 + 66)(0.5)^{12} \approx 0.01929 \end{aligned}$$

37. (a) We approximate

$$50! \approx \sqrt{2\pi 50}\left(\frac{50}{e}\right)^{50} = 10\sqrt{\pi}\left(\frac{50}{e}\right)^{50} \approx 3.036344619 \cdot 10^{64}$$

The true value is

$$50! = 30414093201713378043612608166064768844377641568960512000000000000$$

(b) We get

$$\begin{aligned} \log_{10}\left(\sqrt{2\pi n}\left(\frac{n}{e}\right)^n\right) &= \log_{10}\sqrt{2\pi n} + \log_{10}\left(\frac{n}{e}\right)^n \\ &= \frac{1}{2}(\log_{10}(2\pi) + \log_{10} n) + n(\log_{10} n - \log_{10} e) \\ &= \frac{1}{2}\log_{10}(2\pi) + \left(n + \frac{1}{2}\right)\log_{10} n - n\log_{10} e \end{aligned}$$

When $n = 120$,

$$\log_{10} 120! \approx \frac{1}{2}\log_{10}(2\pi) + 120.5\log_{10} 120 - 120\log_{10} e \approx 198.8250922$$

(c) Using (b), we get

$$\begin{aligned} \log_{10}\binom{120}{36} &= \log_{10}\frac{120!}{36!84!} \\ &= \log_{10} 120! - [\log_{10} 36! + \log_{10} 84!] \\ &\approx \frac{1}{2}\log_{10}(2\pi) + 120.5\log_{10} 120 - 120\log_{10} e \\ &\qquad - \left[\frac{1}{2}\log_{10}(2\pi) + 36.5\log_{10} 36 - 36\log_{10} e + \frac{1}{2}\log_{10}(2\pi) + 84.5\log_{10} 84 - 84\log_{10} e\right] \\ &= -\frac{1}{2}\log_{10}(2\pi) + 120.5\log_{10} 120 - 36.5\log_{10} 36 - 84.5\log_{10} 84 \\ &\approx 30.7356092 \end{aligned}$$

Section 11 The Multinomial and the Geometric Distributions

1. (a) We can do it in $\frac{4!}{1!\,3!} = \frac{24}{6} = 4$ ways: $\{1\,|\,2,3,4\}$, $\{2\,|\,1,3,4\}$, $\{3\,|\,1,2,4\}$, and $\{4\,|\,1,2,3\}$.
(b) We can do it in $\frac{4!}{2!\,2!} = \frac{24}{4} = 6$ ways: $\{1,2\,|\,3,4\}$, $\{1,3\,|\,2,4\}$, $\{1,4\,|\,2,3\}$, $\{2,3\,|\,1,4\}$, $\{2,4\,|\,1,3\}$, and $\{3,4\,|\,1,2\}$.
(c) We can do it in $\frac{4!}{1!\,1!\,2!} = \frac{24}{2} = 12$ ways: $\{1\,|\,2\,|\,3,4\}$, $\{2\,|\,1\,|\,3,4\}$, $\{1\,|\,3\,|\,2,4\}$, $\{3\,|\,1\,|\,2,4\}$, $\{1\,|\,4\,|\,2,3\}$, $\{4\,|\,1\,|\,2,3\}$, $\{2\,|\,3\,|\,1,4\}$, $\{3\,|\,2\,|\,1,4\}$, $\{2\,|\,4\,|\,1,3\}$, $\{4\,|\,2\,|\,1,3\}$, $\{3\,|\,4\,|\,1,2\}$, and $\{4\,|\,3\,|\,1,2\}$.

3. (a) The probability that the 80 wolves will prey on 10 deer, 70 beavers, no moose, and no animals from the "other" group is

$$\begin{aligned}P(N_1=10, N_2=70, N_3=0, N_4=0) &= \frac{80!}{10!\cdot 70!\cdot 0!\cdot 0!}\,0.33^{10}\cdot 0.55^{70}\cdot 0.05^{0}\cdot 0.07^{0}\\ &= \frac{80!}{10!\cdot 70!}\,0.33^{10}\cdot 0.55^{70}\end{aligned}$$

(b) Is it given that $N_2 = 60$, $N_4 = 16$ and $N_1 + N_3 = 4$. The probability is (we go through all combinations of N_1 and N_3 whose sum is 4):

$$\begin{aligned}&P(N_1=0, N_2=60, N_3=4, N_4=16) + P(N_1=1, N_2=60, N_3=3, N_4=16)\\ &\quad + P(N_2=2, N_2=60, N_3=2, N_4=16) + P(N_1=3, N_2=60, N_3=1, N_4=16)\\ &\quad + P(N_1=4, N_2=60, N_3=0, N_4=16)\\ &= \frac{80!}{0!\cdot 60!\cdot 4!\cdot 16!}\,0.33^{0}\cdot 0.55^{60}\cdot 0.05^{4}\cdot 0.07^{16} + \frac{80!}{1!\cdot 60!\cdot 3!\cdot 16!}\,0.33^{1}\cdot 0.55^{60}\cdot 0.05^{3}\cdot 0.07^{16}\\ &\quad + \frac{80!}{2!\cdot 60!\cdot 2!\cdot 16!}\,0.33^{2}\cdot 0.55^{60}\cdot 0.05^{2}\cdot 0.07^{16} + \frac{80!}{3!\cdot 60!\cdot 1!\cdot 16!}\,0.33^{3}\cdot 0.55^{60}\cdot 0.05^{1}\cdot 0.07^{16}\\ &\quad + \frac{80!}{4!\cdot 60!\cdot 0!\cdot 16!}\,0.33^{4}\cdot 0.55^{60}\cdot 0.05^{0}\cdot 0.07^{16}\end{aligned}$$

5. The probability distribution of the genotype of an offspring of AB parents is $P(\text{AA}) = 0.25$, $P(\text{AB}) = 0.5$, and $P(\text{BB}) = 0.25$. There is a total of 9 offspring. The probability is

$$P(\text{three AA, two AB, four BB}) = \frac{9!}{3!\cdot 2!\cdot 4!}\,0.25^3\cdot 0.5^2\cdot 0.25^4 = \frac{9!}{6\cdot 2\cdot 24}\,0.25^8 \approx 0.01923$$

i.e., close to 2%.

7. The probability distribution of the genotype of an offspring of LS parents is $P(\text{LL}) = 0.25 = P(\text{long})$, $P(\text{LS}) = 0.5 = P(\text{medium length})$, and $P(\text{SS}) = 0.25 = P(\text{short})$.
(a) The probability is

$$P(\text{two LL, two LS, two SS}) = \frac{6!}{2!\cdot 2!\cdot 2!}\,0.25^2\cdot 0.5^2\cdot 0.25^2 = \frac{6!}{8}\,0.25^5 \approx 0.08789$$

(b) The probability is

$$\begin{aligned}&P(\text{two LL, zero LS, four SS}) + P(\text{two LL, one LS, three SS})\\ &= \frac{6!}{2!\cdot 0!\cdot 4!}\,0.25^2\cdot 0.5^0\cdot 0.25^4 + \frac{6!}{2!\cdot 1!\cdot 3!}\,0.25^2\cdot 0.5^1\cdot 0.25^3\\ &\approx 0.00366 + 0.02930 = 0.03296\end{aligned}$$

9. The probability distribution of the genotype of an offspring of AB parents is

$$\begin{aligned}P(\text{AA}) &= 0.25 = P(\text{neither carrier nor has the trait})\\ P(\text{AB}) &= 0.5 = P(\text{carrier})\\ P(\text{BB}) &= 0.25 = P(\text{has the trait})\end{aligned}$$

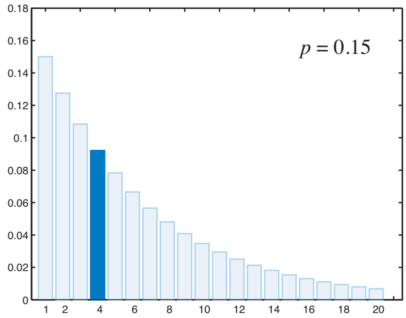
The probability that one child will have attached earlobes, two will be carriers, and one will neither be a carrier nor have attached earlobes is

$$P(\text{one AA, two AB, one BB}) = \frac{4!}{1! \cdot 2! \cdot 1!}(0.25)^1 \cdot (0.5)^2 \cdot (0.25)^1 = 12\,(0.25)^3 = 0.1875$$

11. (a) Consider the geometric distribution with probability of success $p = 0.15$. The probability of the first success occurring on the fourth trial is

$$P(X = 4) = (1 - 0.15)^3(0.15) = (0.85)^3(0.15) \approx 0.092$$

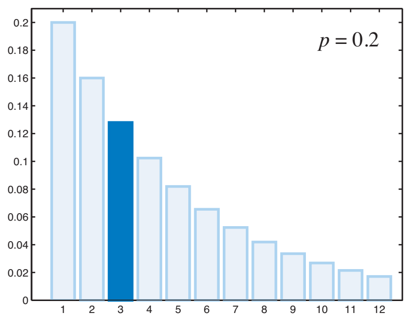
(b) See below.



13. (a) Consider the geometric distribution with probability of success $p = 0.2$. The probability of the first success occurring on the third trial is

$$P(X = 3) = (1 - 0.2)^2(0.2) = (0.8)^2(0.2) = 0.128$$

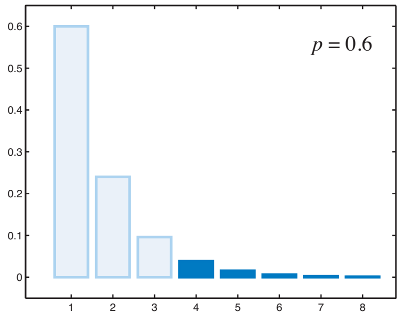
(b) See below.



15. (a) Consider the geometric distribution with probability of success $p = 0.6$. The probability that the first success occurs on or after the fourth trial is (use the complementary event "success occurs before the fourth trial")

$$\begin{aligned} P(X \geq 4) &= 1 - P(X < 4) \\ &= 1 - [P(X = 1) + P(X = 2) + P(X = 3)] \\ &= 1 - [0.6 + (1 - 0.6)(0.6) + (1 - 0.6)^2(0.6)] = 0.064 \end{aligned}$$

(b) See below.

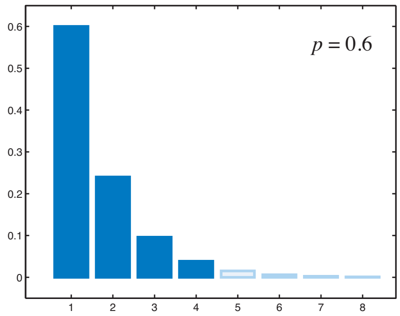


17. (a) Consider the geometric distribution with probability of success $p = 0.6$. The probability that the first success occurs on or before the fourth trial is

$$\begin{aligned} P(X \leq 4) &= P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) \\ &= 0.6 + (1 - 0.6)(0.6) + (1 - 0.6)^2(0.6) + (1 - 0.6)^3(0.6) \\ &= 0.6[1 + 0.4 + 0.4^2 + 0.4^3] \\ &= 0.6 \cdot \frac{1 - 0.4^4}{1 - 0.4} = 1 - 0.4^4 = 0.9744 \end{aligned}$$

(In calculating the sum in the end, we used the formula $1 + q + q^2 + q^3 + \cdots + q^n = (1 - q^{n+1})/(1 - q)$ with $q = 0.4$ and $n = 3$.)

(b) See below.



19. A geometric distribution will larger p is less spread out that the one with smaller p (look at histograms in Figure 11.1). Thus, the geometric distribution with $p_2 = p/2$ is more spread out than the one with $p_1 = p$.

Formally: the variances are $\mathrm{var}_1 = (1 - p)/p^2$ and

$$\mathrm{var}_2 = \frac{1 - \frac{p}{2}}{\left(\frac{p}{2}\right)^2} = \frac{1 - \frac{p}{2}}{\frac{p^2}{4}} = \frac{4 - 2p}{p^2}$$

From $4 - 2p > 1 - p$ (which is true whenever $p < 3$) we conclude

$$\mathrm{var}_1 = \frac{1 - p}{p^2} < \frac{4 - 2p}{p^2} = \mathrm{var}_2$$

So, $p_2 = p/2$ yields larger variance (thus, wider spread) than $p_1 = p$.

21. From $E(X) = 1/p = 5$ we get $p = 0.2$ The variance is $\mathrm{var}(X) = (1 - p)/p^2 = 0.8/0.04 = 20$ and the standard deviation is $\sqrt{20} \approx 4.472$.

23. From $\mathrm{var} = (1-p)/p^2 = 2$ we get $2p^2 = 1-p$ and $2p^2+p-1 = (2p-1)(p+1) = 0$. Thus $p = 1/2$ (the remaining solution $p = -1$ makes no sense) and so the mean is $1/(1/2) = 2$.

25. Let X = "number of trials until gene mutates". X is a geometrically distributed random variable with the probability of success $p = 0.001$.

The probability that a gene will mutate during the 20th cell division is

$$P(X = 20) = (1 - 0.001)^{19}(0.001) \approx 0.00098$$

The probability that the gene will mutate before or during the 20th cell division is

$$\begin{aligned} P(X \le 20) &= P(X = 1) + P(X = 2) + P(X = 3) + \cdots + P(X = 20) \\ &= 0.001 + (0.999)(0.001) + (0.999)^2(0.001) + \cdots + (0.999)^{19}(0.001) \\ &= 0.001[1 + 0.999 + (0.999)^2 + \cdots + (0.999)^{19}] \\ &= 0.001 \cdot \frac{1 - 0.999^{20}}{1 - 0.999} = 1 - 0.999^{20} \approx 0.0198 \end{aligned}$$

(In calculating the sum we used the formula $1 + q + q^2 + q^3 + \cdots + q^n = (1 - q^{n+1})/(1 - q)$ with $q = 0.999$ and $n = 19$.)

27. (a) We compute

$$\begin{aligned} s_n - qs_n &= 1 + q + q^2 + q^3 + \cdots + q^n - q(1 + q + q^2 + q^3 + \cdots + q^n) \\ s_n(1 - q) &= 1 + q + q^2 + q^3 + \cdots + q^n - q - q^2 - q^3 - \cdots - q^n - q^{n+1} \\ s_n(1 - q) &= 1 - q^{n+1} \\ s_n &= \frac{1 - q^{n+1}}{1 - q} \end{aligned}$$

(b) Since $|q| < 1$, it follows that the limit of q^{n+1} as $n \to \infty$ is zero. Thus,

$$\lim_{n\to\infty} s_n = \lim_{n\to\infty} \frac{1 - q^{n+1}}{1 - q} = \frac{1}{1 - q}$$

i.e.,

$$1 + q + q^2 + q^3 + \cdots = \frac{1}{1 - q}$$

29. (a) Differentiating

$$1 + q + q^2 + q^3 + \cdots = \frac{1}{1 - q}$$

with respect to q, we get

$$1 + 2q + 3q^2 + 4q^3 + \cdots = -(1 - q)^{-2}(-1) = \frac{1}{(1 - q)^2}$$

replacing q by $1 - p$ yields

$$\begin{aligned} 1 + 2(1 - p) + 3(1 - p)^2 + 4(1 - p)^3 + \cdots &= \frac{1}{(1 - (1 - p))^2} \\ \sum_{k=1}^{\infty} k(1 - p)^{k-1} &= \frac{1}{p^2} \end{aligned}$$

(b) Differentiating

$$1 + q + q^2 + q^3 + \cdots = \frac{1}{1 - q}$$

with respect to q, then multiplying by q and differentiating with respect to q again, we obtain

$$\begin{aligned} 1 + 2q + 3q^2 + 4q^3 + \cdots &= \frac{1}{(1 - q)^2} \\ q + 2q^2 + 3q^3 + 4q^4 + \cdots &= \frac{q}{(1 - q)^2} \end{aligned}$$

$$1 + 2^2 q + 3^2 q^2 + 4^2 q^3 + \cdots = \frac{(1-q)^2 - q \cdot 2(1-q)(-1)}{(1-q)^4} = \frac{(1-q) + 2q}{(1-q)^3}$$
$$1 + 2^2 q + 3^2 q^2 + 4^2 q^3 + \cdots = \frac{q+1}{(1-q)^3}$$

Replacing q by $1-p$ and then multiplying by p yields

$$1 + 2^2(1-p) + 3^2(1-p)^2 + 4^2(1-p)^3 + \cdots = \frac{(1-p)+1}{(1-(1-p))^3}$$
$$\sum_{k=1}^{\infty} k^2(1-p)^{k-1} = \frac{2-p}{p^3}$$
$$p\sum_{k=1}^{\infty} k^2(1-p)^{k-1} = \frac{2-p}{p^2}$$

Section 12 The Poisson Distribution

1. It is given that $X \sim \text{Po}\,(2.5)$. Using

$$P(X=k) = \frac{e^{-\lambda}\lambda^k}{k!} = \frac{e^{-2.5}(2.5)^k}{k!}$$

we obtain

$$\begin{aligned}
P(X=0) &= \frac{e^{-2.5}(2.5)^0}{0!} = e^{-2.5} \approx 0.0820850 \\
P(X=1) &= \frac{e^{-2.5}(2.5)^1}{1!} \approx 0.205212 \\
P(X=2) &= \frac{e^{-2.5}(2.5)^2}{2!} \approx 0.256516 \\
P(X=3) &= \frac{e^{-2.5}(2.5)^3}{3!} \approx 0.213763 \\
P(X=4) &= \frac{e^{-2.5}(2.5)^4}{4!} \approx 0.133602
\end{aligned}$$

3. It is given that $X \sim \text{Po}\,(12)$. We find

$$\begin{aligned}
P(4 \le X \le 7) &= P(X=4) + P(X=5) + P(X=6) + P(X=7) \\
&= \frac{e^{-12}12^4}{4!} + \frac{e^{-12}12^5}{5!} + \frac{e^{-12}12^6}{6!} + \frac{e^{-12}12^7}{7!} \\
&= e^{-12}\left(\frac{12^4}{4!} + \frac{12^5}{5!} + \frac{12^6}{6!} + \frac{12^7}{7!}\right) \approx 0.087213
\end{aligned}$$

5. It is given that $X \sim \text{Po}\,(4)$. We find

$$\begin{aligned}
P(0 \le X \le 3) &= P(X=0) + P(X=1) + P(X=2) + P(X=3) \\
&= \frac{e^{-4}4^0}{0!} + \frac{e^{-4}4^1}{1!} + \frac{e^{-4}4^2}{2!} + \frac{e^{-4}4^3}{3!} \\
&= e^{-4}\left(1 + 4 + 8 + \frac{32}{3}\right) \approx 0.433470
\end{aligned}$$

7. Look at Figure 12.1. There are two identical probabilities, corresponding to the values $P(X=\lambda-1)$ and $P(X=\lambda)$. Thus, the given graph represents the Poisson distribution with $\lambda = 4$.

We now prove that the observation we made is indeed true. Assume that λ is an integer, $\lambda \ge 1$. and that $X \sim \text{Po}\,(\lambda)$. Then

$$P(X=\lambda-1) = \frac{e^{-\lambda}\lambda^{\lambda-1}}{(\lambda-1)!} = \frac{e^{-\lambda}\lambda^{\lambda-1}}{(\lambda-1)!}\cdot\frac{\lambda}{\lambda} = \frac{e^{-\lambda}\lambda^{\lambda}}{\lambda!} = P(X=\lambda)$$

(In the above, we used the fact that $(\lambda-1)!\lambda = \lambda!$.)

9. Define X = "number of people with a respiratory infection in a group of 5000 people." The occurrence of 3 out of 2,000 translates to (multiply by 2.5) 7.5 out of 5,000. Thus, X is a Poisson distribution with parameter $\lambda = 7.5$. The probability that 12 out of 5,000 people are diagnosed with the infection is

$$P(X=12) = \frac{e^{-7.5}(7.5)^{12}}{12!} \approx 0.036575$$

11. Let X = "number of more serious traffic accidents per week." Then X is a Poisson distribution with parameter $\lambda = 4$. The probability that at least two more serious accidents happen in a week is

$$\begin{aligned} P(X \geq 2) &= 1 - P(X < 2) = 1 - (P(X = 0) + P(X = 1)) \\ &= 1 - \left(\frac{e^{-4}(4)^0}{0!} + \frac{e^{-4}(4)^1}{1!} \right) = 1 - 5e^{-4} \approx 0.908422 \end{aligned}$$

13. Define X = "number of spoiled apples in a bag of 15 apples." The occurrence of 2 spoiled apples in a bag of 30 apples translates to 1 spoiled apple in a bag of 15 apples. Thus, X is a Poisson distribution with parameter $\lambda = 1$. The probability that there are no more than two spoiled apples in the bag of 15 apples is

$$\begin{aligned} P(X \leq 2) &= P(X = 0) + P(X = 1) + P(X = 2) \\ &= \frac{e^{-1}(1)^0}{0!} + \frac{e^{-1}(1)^1}{1!} + \frac{e^{-1}(1)^2}{2!} = e^{-1}(1 + 1 + 0.5) = 2.5e^{-1} \approx 0.919699 \end{aligned}$$

15. Define X = "number of heavy metal particles in a half-litre glass of tap water." The occurrence of six heavy metal particles in 1 L of tap water translates to three heavy metal particles in $1/2$ L of tap water. Thus, X is a Poisson distribution with parameter $\lambda = 3$. The probability that there are no heavy metal particles in a half-litre glass of tap water is

$$P(X = 0) = \frac{e^{-3}(3)^0}{0!} = e^{-3} \approx 0.049787$$

i.e., a bit less than 5%.

17. Define X = "number of molecules leaving the region by the end of the second hour." The rate of 0.4 molecules per hour translates to the rate of 0.8 molecules per two hours. Thus, X is a Poisson distribution with parameter $\lambda = 0.8$. The probability that three or fewer molecules leave by the end of the second hour is

$$\begin{aligned} P(X \leq 3) &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) \\ &= \frac{e^{-0.8}(0.8)^0}{0!} + \frac{e^{-0.8}(0.8)^1}{1!} + \frac{e^{-0.8}(0.8)^2}{2!} + \frac{e^{-0.8}(0.8)^3}{3!} \approx 0.990920 \end{aligned}$$

19. Define X = "number of hits by cosmic rays in an eight-hour interval." The rate of one cosmic ray per day translates to the rate of $1/3$ cosmic rays per eight hours. Thus, X is a Poisson distribution with parameter $\lambda = 1/3$. The probability that we will be hit at least once in an eight-hour interval is

$$P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0) = 1 - \frac{e^{-1/3}(1/3)^0}{0!} = 1 - e^{-1/3} \approx 0.283469$$

21. Let X = "number of text messages received in an hour." The context implies that X is a Poisson distribution with parameter $\lambda = 3$. The probability that the student receives more than five messages in an hour is

$$\begin{aligned} P(X > 5) &= 1 - P(X \leq 5) \\ &= 1 - (P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5)) \\ &= 1 - \left(\frac{e^{-3}(3)^0}{0!} + \frac{e^{-3}(3)^1}{1!} + \frac{e^{-3}(3)^2}{2!} + \frac{e^{-3}(3)^3}{3!} + \frac{e^{-3}(3)^4}{4!} + \frac{e^{-3}(3)^5}{5!} \right) \\ &= 1 - \frac{92}{5}e^{-3} \approx 0.083918 \end{aligned}$$

23. Since $X \sim \text{Po}(1)$ and $Y \sim \text{Po}(9)$, it follows that (assuming independence) $X + Y \sim \text{Po}(1 + 9)$, i.e., $X + Y \sim \text{Po}(10)$. Thus

$$P(X + Y = 2) = \frac{e^{-10}(10)^2}{2!} = 50e^{-10} \approx 0.002270$$

and

$$\begin{aligned} P(Y = 2 \mid X + Y = 2) &= \frac{P(Y = 2 \text{ and } X + Y = 2)}{P(X + Y = 2)} \\ &= \frac{P(Y = 2 \text{ and } X = 0)}{P(X + Y = 2)} \\ &= \frac{P(Y = 2)P(X = 0)}{P(X + Y = 2)} \\ &= \frac{\frac{e^{-9}(9)^2}{2!}\frac{e^{-1}(1)^0}{0!}}{50e^{-10}} = \frac{(9)^2}{50 \cdot 2!} = 0.81 \end{aligned}$$

25. Define T = "number of text messages in an hour" and C = "number of phone calls in an hour." It is given that $T \sim \text{Po}(4)$ and $C \sim \text{Po}(2)$. Let $I = T + C$ = "number of interruptions in an hour." Assuming independence, $I \sim \text{Po}(6)$. The probability that the student will experience no interruptions in 1 hour is

$$P(I = 0) = \frac{e^{-6}(6)^0}{0!} = e^{-6} \approx 0.002479$$

Let J = "number of interruptions in a ten-minute interval." Then $J \sim \text{Po}(1)$, and the probability that the student will experience one interruption every 10 minutes is

$$P(J = 1) = \frac{e^{-1}(1)^1}{1!} = e^{-1} \approx 0.367879$$

27. Define

$$A = \begin{cases} 1 & \text{a person experiences serious side effects from allergy medication (success)} \\ 0 & \text{a person does not experience serious side effects from allergy medication} \end{cases}$$

A is a Bernoulli random variable with probability of success equal to 0.003. Let N = "number of people experiencing serious side effects from allergy medication in a group of 200 people". Since N counts the number of successes in 200 independent repetitions of the event A, it follows that N is a binomially distributed random variable with $n = 200$ and $p = 0.003$. Thus, the probability that in a group of 200 people nobody experiences serious side effects is

$$b(0, 200; 0.003) = \binom{200}{0}(0.003)^0(0.997)^{200} = (0.997)^{200} \approx 0.548317$$

Using Poisson approximation (recall that $b(k, n; p) \approx P(X = k)$ if $X \sim \text{Po}(np)$), we get

$$b(0, 200; 0.003) \approx P(X = 0)$$

where $X \sim \text{Po}(200 \cdot 0.003 = 0.6)$. Thus

$$P(X = 0) = \frac{e^{-0.6}(0.6)^0}{0!} = e^{-0.6} \approx 0.548812$$

29. Define

$$L = \begin{cases} 1 & \text{a person has serious consequences from lactose intolerance (success)} \\ 0 & \text{a person does not have serious consequences from lactose intolerance} \end{cases}$$

L is a Bernoulli random variable with probability of success equal to 0.002. Let N = "number of people who have serious consequences from lactose intolerance in a group of 500 people". Since N counts the number of successes in 500 independent repetitions of the event L, it follows that N is a binomially distributed random variable with $n = 500$ and $p = 0.002$. Thus, the probability that in a group of 500 people one person experiences serious consequences from lactose intolerance is

$$b(1, 500; 0.002) = \binom{500}{1}(0.002)^1(0.998)^{499} = (0.998)^{499} \approx 0.368248$$

Using Poisson approximation (recall that $b(k, n; p) \approx P(X = k)$ if $X \sim \text{Po}\,(np)$), we get

$$b(1, 500; 0.002) \approx P(X = 1)$$

where $X \sim \text{Po}\,(500 \cdot 0.002 = 1)$. Thus

$$P(X = 1) = \frac{e^{-1}(1)^1}{1!} = e^{-1} \approx 0.367879$$

Section 13 Continuous Random Variables

1. The function $f(x) = 1 - x^2$, $x \in [0, 2]$, cannot be a probability density function because $f(x) \geq 0$ does not hold on $[0, 2]$. For instance, $f(1.5) = 1 - (1.5)^2 = -1.25$.

3. To satisfy $f(x) \geq 0$, we need $a \geq 0$ (actually, we need $a > 0$; if $a = 0$, then f is identically zero and cannot be a probability density function). As well, the integral of f has to be equal to 1:

$$\int_1^{10} \frac{a}{x}\,dx = a \ln|x|\bigg|_1^{10} = a\ln 10 - a\ln 1 = a\ln 10 = 1$$

Thus, $a = 1/\ln 10$.

5. To satisfy $f(x) \geq 0$, we need $a \geq 0$ (actually we need $a > 0$; if $a = 0$, then f is identically zero and cannot be a probability density function). As well, the integral of f has to be 1:

$$\int_0^{\infty} \frac{a}{1+x^2}\,dx = a \arctan x\bigg|_0^{\infty} = a\arctan(\infty) - a\arctan 0 = a\frac{\pi}{2} = 1$$

(since $\arctan 0 = 0$). Thus, $a = 2/\pi$.

In the above, we abbreviated the calculation of the improper integral. Without skipping steps:

$$\begin{aligned}\int_0^{\infty} \frac{a}{1+x^2}\,dx &= a \lim_{T\to\infty} \int_0^{T} \frac{a}{1+x^2}\,dx \\ &= a \lim_{T\to\infty} \arctan x\bigg|_0^{T} \\ &= a \lim_{T\to\infty} (\arctan T - \arctan 0) = a\frac{\pi}{2}\end{aligned}$$

7. Clearly, $f(x) = 2/x^3$ is positive for $x \in [1, \infty)$. As well,

$$\begin{aligned}\int_1^{\infty} \frac{2}{x^3}\,dx &= \lim_{T\to\infty} \int_1^{T} \frac{2}{x^3}\,dx \\ &= \lim_{T\to\infty} 2\frac{x^{-2}}{-2}\bigg|_1^{T} \\ &= \lim_{T\to\infty} -\frac{1}{x^2}\bigg|_1^{T} \\ &= \lim_{T\to\infty} \left(-\frac{1}{T^2} + \frac{1}{1^2}\right) = 0 + 1 = 1\end{aligned}$$

The mean is

$$\begin{aligned}\mu = \int_1^{\infty} x\frac{2}{x^3}\,dx &= \lim_{T\to\infty} \int_1^{T} \frac{2}{x^2}\,dx \\ &= \lim_{T\to\infty} 2\frac{x^{-1}}{-1}\bigg|_1^{T} \\ &= \lim_{T\to\infty} -\frac{2}{x}\bigg|_1^{T} \\ &= \lim_{T\to\infty} \left(-\frac{2}{T} + \frac{2}{1}\right) = 0 + 2 = 2\end{aligned}$$

9. No. Let $f(x) = a$ for $x \in [0, \infty)$, where $a > 0$ is a constant. Since

$$\int_0^{\infty} a\,dx = \lim_{T\to\infty} \int_0^{T} a\,dx = \lim_{T\to\infty} ax\bigg|_0^{T} = \lim_{T\to\infty} (aT) = \infty$$

the integral of f cannot be equal to 1, no matter what value of a is used.

11. Using the probability density function, we compute

$$P(0.5 \le X \le 2) = \int_{0.5}^{2} (0.3 + 0.2x)\,dx = (0.3x + 0.1x^2)\Big|_{0.5}^{2} = (0.6 + 0.4) - (0.15 + 0.025) = 0.825$$

The cumulative distribution function of $f(x)$ is

$$F(x) = \int_0^x (0.3 + 0.2t)\,dt = (0.3t + 0.1t^2)\Big|_0^x = (0.3x + 0.1x^2) - 0 = 0.3x + 0.1x^2$$

for x in $[0, 2]$. Thus,

$$P(0.5 \le X \le 2) = F(2) - F(0.5) = [(0.3)(2) + (0.1)(2)^2] - [(0.3)(0.5) + (0.1)(0.5)^2] = 0.825$$

13. Using the probability density function, we compute

$$P(1 \le X \le 2) = \int_1^2 \frac{1}{x}\,dx = \ln|x|\Big|_1^2 = \ln 2 - \ln 1 = \ln 2$$

The cumulative distribution function of $f(x)$ is

$$F(x) = \int_1^x \frac{1}{t}\,dt = \ln|t|\Big|_1^x = \ln x - \ln 1 = \ln x$$

for x in $[1, e]$. Thus,

$$P(1 \le X \le 2) = F(2) - F(1) = \ln 2 - \ln 1 = \ln 2$$

15. We check the properties listed in Theorem 13:

(a) Since $e^{-2x} \le 1$ for $x \ge 0$, it follows that $F(x) = 1 - e^{-2x} \ge 0$ for all $x \in [0, \infty)$. As well, $e^{-2x} > 0$, and thus $F(x) = 1 - e^{-2x} \le 1$ for all $x \in [0, \infty)$.

(b) The function $F(x)$ is continuous for all x, as the difference of two continuous functions. The fact that $F'(x) = -e^{-2x}(-2) = 2e^{-2x} > 0$ implies that $F(x)$ is increasing (thus, it is non-decreasing) for all $x \in [0, \infty)$.

(c) The limits:

$$\lim_{x\to 0} F(x) = \lim_{x\to 0}(1 - e^{-2x}) = 1 - e^0 = 0$$

and

$$\lim_{x\to\infty} F(x) = \lim_{x\to\infty}(1 - e^{-2x}) = 1 - e^{-\infty} = 1$$

Thus $F(x) = 1 - e^{-2x}$, $x \in [0, \infty)$, is indeed a cumulative distribution function. The corresponding probability density function is $f(x) = F'(x) = 2e^{-2x}$.

The expected value is given by

$$\mu = \int_0^\infty x(2e^{-2x})\,dx = 2\int_0^\infty xe^{-2x}\,dx$$

First we calculate the indefinite integral (using integration by parts): let $u = x$ and $v' = e^{-2x}$. Then $u' = 1$, $v = -e^{-2x}/2$, and

$$\begin{aligned}\int xe^{-2x}\,dx &= uv - \int vu'\,dx \\ &= -\frac{1}{2}xe^{-2x} + \frac{1}{2}\int e^{-2x}\,dx \\ &= -\frac{1}{2}xe^{-2x} - \frac{1}{4}e^{-2x} = -\frac{1}{4}(2x+1)e^{-2x}\end{aligned}$$

Thus

$$\mu = 2\int_0^\infty xe^{-2x}\,dx = 2\lim_{T\to\infty}\int_0^T xe^{-2x}\,dx$$

$$= 2 \lim_{T\to\infty} \left(-\frac{1}{4}(2x+1)e^{-2x} \right) \Big|_0^T$$
$$= 2 \left[\lim_{T\to\infty} \left(-\frac{1}{4}(2T+1)e^{-2T} \right) - \left(-\frac{1}{4} \right) \right] = 2\left(0 + \frac{1}{4}\right) = \frac{1}{2}$$

Recall that

$$\lim_{T\to\infty} e^{-2T} = 0$$

and, by L'Hôpital's rule,

$$\lim_{T\to\infty} Te^{-2T} = \lim_{T\to\infty} \frac{T}{e^{2T}} = \lim_{T\to\infty} \frac{1}{2e^{2T}} = 0$$

17. (a) Clearly, $f(x) = 2x \geq 0$ for $x \in [0,1]$. As well,

$$\int_0^1 f(x)\,dx = \int_0^1 2x\,dx = x^2 \Big|_0^1 = 1 - 0 = 1$$

(b) The cumulative distribution function is

$$F(x) = \int_0^x 2t\,dt = t^2 \Big|_0^x = x^2$$

for $x \in [0,1]$.
(c) The expected value of X is

$$\mu = E(X) = \int_0^1 x(2x)\,dx = \frac{2x^3}{3}\Big|_0^1 = \frac{2}{3}$$

(d) We find

$$P(X \leq \mu) = P(X \leq 2/3) = F(2/3) = \left(\frac{2}{3}\right)^2 = \frac{4}{9}$$

19. (a) Clearly, $f(x) = 3x^2 \geq 0$ for $x \in [0,1]$. As well,

$$\int_0^1 f(x)\,dx = \int_0^1 3x^2\,dx = x^3 \Big|_0^1 = 1 - 0 = 1$$

(b) The cumulative distribution function is

$$F(x) = \int_0^x 3t^2\,dt = t^3 \Big|_0^x = x^3$$

for $x \in [0,1]$.
(c) The expected value of X is

$$\mu = E(X) = \int_0^1 x(3x^2)\,dx = \frac{3x^4}{4}\Big|_0^1 = \frac{3}{4}$$

(d) We find

$$P(X \leq \mu) = P(X \leq 3/4) = F(3/4) = \left(\frac{3}{4}\right)^3 = \frac{27}{64}$$

21. (a) From $0 \leq x \leq 3$ we get (after multiplying by $2/9$) $0 \leq 2x/9 \leq 2/3$. Thus, $2/3 - 2x/9 \geq 0$ for $x \in [0,3]$. As well,

$$\int_0^3 f(x)\,dx = \int_0^3 \left(\frac{2}{3} - \frac{2x}{9}\right) dx = \left(\frac{2x}{3} - \frac{x^2}{9}\right)\Big|_0^3 = (2-1) - 0 = 1$$

(b) The cumulative distribution function is

$$F(x) = \int_0^x \left(\frac{2}{3} - \frac{2t}{9}\right) dt = \left(\frac{2t}{3} - \frac{t^2}{9}\right)\Bigg|_0^x = \frac{2x}{3} - \frac{x^2}{9}$$

for $x \in [0, 3]$.

(c) The expected value of X is

$$\mu = E(X) = \int_0^3 x\left(\frac{2}{3} - \frac{2x}{9}\right) dx = \int_0^3 \left(\frac{2x}{3} - \frac{2x^2}{9}\right) dx = \left(\frac{x^2}{3} - \frac{2x^3}{27}\right)\Bigg|_0^3 = (3-2) - 0 = 1$$

(d) We find

$$P(X \le \mu) = P(X \le 1) = F(1) = \left(\frac{2}{3} - \frac{1}{9}\right) = \frac{5}{9}$$

23. We compute

$$\begin{aligned}
\mu = E(X) &= \int_0^1 x(3x^2)\, dx = \frac{3x^4}{4}\Bigg|_0^1 = \frac{3}{4} = 0.75 \\
E(X^2) &= \int_0^1 x^2(3x^2)\, dx = \frac{3x^5}{5}\Bigg|_0^1 = \frac{3}{5} \\
\text{var}(X) &= E(X^2) - (E(X))^2 = \frac{3}{5} - \frac{9}{16} = \frac{3}{80}
\end{aligned}$$

and $\sigma = \sqrt{\text{var}(X)} = \sqrt{3/80} \approx 0.19365$. The probability that the values of X are at most one standard deviation away from the mean is

$$\begin{aligned}
P(\mu - \sigma \le X \le \mu + \sigma) &= \int_{\mu-\sigma}^{\mu+\sigma} 3x^2\, dx \\
&= x^3\Big|_{\mu-\sigma}^{\mu+\sigma} \\
&= (\mu+\sigma)^3 - (\mu-\sigma)^3 \\
&= (0.75 + 0.19365)^3 - (0.75 - 0.19365)^3 = 0.668093
\end{aligned}$$

25. The cumulative distribution function is

$$F(x) = \int_0^x 3t^2\, dt = t^3\Big|_0^x = x^3$$

for $x \in [0, 1]$. The median is the value x where $F(x) = 1/2$, i.e., where $x^3 = 1/2$. Thus, the median is $\sqrt[3]{1/2}$.

27. We are looking for a number Q_3 such that $P(X \le Q_3) = 0.75$.

$$\begin{aligned}
\int_0^{Q_3} \left(\frac{2}{3} - \frac{2x}{9}\right) dx &= 0.75 \\
\left(\frac{2x}{3} - \frac{x^2}{9}\right)\Bigg|_0^{Q_3} &= 0.75 \\
\frac{2Q_3}{3} - \frac{Q_3^2}{9} &= 0.75
\end{aligned}$$

Multiplying by 9 and using the quadratic formula, we get

$$\begin{aligned}
Q_3^2 - 6Q_3 + 6.75 &= 0 \\
Q_3 &= \frac{6 \pm \sqrt{36 - 27}}{2}
\end{aligned}$$

So $Q_3 = 1.5$ or $Q_3 = 4.5$. Since the probability density function and the cumulative distribution function are defied on $0 \le x \le 3$, the upper quartile is 1.5.

29. The average lifetime of the tree is given by the integral

$$\int_0^\infty t\, f(t)\, dt = \int_0^\infty t\, 0.01 e^{-0.01t}\, dt = \lim_{T\to\infty} \int_0^T 0.01 t e^{-0.01t}\, dt$$

To calculate the indefinite integral, we use the integration by parts with $u = t$ and $v' = e^{-0.01t}$. Then $u' = 1$, $v = -e^{-0.01t}/0.01 = -100e^{-0.01t}$, and

$$\begin{aligned}\int 0.01 t e^{-0.01t}\, dt &= 0.01\left(uv - \int vu'\, dt\right)\\ &= 0.01\left(-100te^{-0.01t} + \int 100e^{-0.01t}\, dt\right)\\ &= 0.01\left(-100te^{-0.01t} - 10000e^{-0.01t}\right)\\ &= -te^{-0.01t} - 100e^{-0.01t}\end{aligned}$$

Thus,

$$\begin{aligned}\lim_{T\to\infty} \int_0^T 0.01te^{-0.01t}\, dt &= \lim_{T\to\infty} \left(-te^{-0.01t} - 100e^{-0.01t}\right)\Big|_0^T\\ &= \lim_{T\to\infty} \left[\left(-Te^{-0.01T} - 100e^{-0.01T}\right) - (0 - 100)\right]\\ &= 100\end{aligned}$$

because

$$\lim_{T\to\infty} e^{-0.01T} = 0$$

and, by L'Hôpital's rule,

$$\lim_{T\to\infty} Te^{-0.01T} = \lim_{T\to\infty} \frac{T}{e^{0.01T}} = \lim_{T\to\infty} \frac{1}{0.01e^{0.01T}} = 0$$

Thus, the average lifetime of a tree is 100 years.

The probability that a tree will live longer than 70 years is

$$\begin{aligned}P &= \int_{70}^\infty f(t)\, dt = \int_{70}^\infty 0.01e^{-0.01t}\, dt\\ &= \lim_{T\to\infty} \int_{70}^T 0.01e^{-0.01t}\, dt\\ &= \lim_{T\to\infty} -e^{-0.01t}\Big|_{70}^T\\ &= \lim_{T\to\infty} \left(-e^{-0.01T} + e^{-0.01(70)}\right)\\ &= e^{-0.7} \approx 0.49659\end{aligned}$$

i.e., about 50%.

31. The probability is

$$\begin{aligned}P(\text{distance} \le 10) &= \int_0^{10} \frac{2}{\pi(1+x^2)}\, dx\\ &= \frac{2}{\pi} \arctan x\Big|_0^{10}\\ &= \frac{2}{\pi}\arctan 10 - 0 \approx 0.936550\end{aligned}$$

33. (a) When $x \geq 0$, $f(x) = 1 - |x| = 1 - x$; thus,

$$\begin{aligned} P(1/2 \leq X \leq 3/4) &= \int_{1/2}^{3/4} (1 - |x|)\,dx = \int_{1/2}^{3/4} (1 - x)\,dx \\ &= \left(x - \frac{x^2}{2} \right) \Big|_{1/2}^{3/4} \\ &= \left(\frac{3}{4} - \frac{9}{32} \right) - \left(\frac{1}{2} - \frac{1}{8} \right) = \frac{3}{32} \end{aligned}$$

When $x < 0$, $f(x) = 1 - |x| = 1 - (-x) = 1 + x$; so

$$\begin{aligned} P(-1/2 \leq X \leq 0) &= \int_{-1/2}^{0} (1 - |x|)\,dx = \int_{-1/2}^{0} (1 + x)\,dx \\ &= \left(x + \frac{x^2}{2} \right) \Big|_{-1/2}^{0} \\ &= (0) - \left(-\frac{1}{2} + \frac{1}{8} \right) = \frac{3}{8} \end{aligned}$$

(b) To find the expected value, we need to find the integral

$$E(X) = \int_{-1}^{1} x(1 - |x|)\,dx = \int_{-1}^{0} x(1 + x)\,dx + \int_{0}^{1} x(1 - x)\,dx$$

We can proceed as usual, calculating antiderivatives and evaluating. But, there is a shortcut: the function $x(1 - |x|)$ is odd, and therefore its integral from -1 to 1 is zero. Thus, $E(X) = 0$. We need to find

$$\begin{aligned} E(X^2) &= \int_{-1}^{1} x^2(1 - |x|)\,dx = \int_{-1}^{0} x^2(1 + x)\,dx + \int_{0}^{1} x^2(1 - x)\,dx \\ &= \int_{-1}^{0} (x^2 + x^3)\,dx + \int_{0}^{1} (x^2 - x^3)\,dx \\ &= \left(\frac{x^3}{3} + \frac{x^4}{4} \right) \Big|_{-1}^{0} + \left(\frac{x^3}{3} - \frac{x^4}{4} \right) \Big|_{0}^{1} \\ &= (0) - \left(-\frac{1}{3} + \frac{1}{4} \right) + \left(\frac{1}{3} - \frac{1}{4} \right) - (0) = \frac{1}{6} \end{aligned}$$

Thus, the variance is $\operatorname{var}(X) = E(X^2) - (E(X))^2 = 1/6$.

35. The Intermediate Value Theorem states that a continuous function defined on a closed interval $[a, b]$ assumes all values between $f(a)$ and $f(b)$. The cumulative distribution function is continuous, and by assumption in this exercise, it is defined on a closed interval $[a, b]$ (and not on an interval that includes $-\infty$ or ∞). Any cumulative distribution function $F(x)$ satisfies $F(a) = 0$ and $F(b) = 1$. Thus, the Intermediate Value Theorem implies that F assumes all values between 0 and 1, in particular the value $1/2$. In other words, there is a number x where $F(x) = 1/2$; this number is the median of X.

Section 14 The Normal Distribution

1. Assume that X is normally distributed with mean μ and variance σ^2. The z-score of a number a is the number $(a-\mu)/\sigma$; it is used to convert a probability related to a normal distribution to a probability related to the *standard* normal distribution.

If $X \sim N(3, 16)$ then $\mu = 3$ and $\sigma = 4$. To calculate $P(0 \leq X \leq 7)$ we convert the numbers to their z-scores:

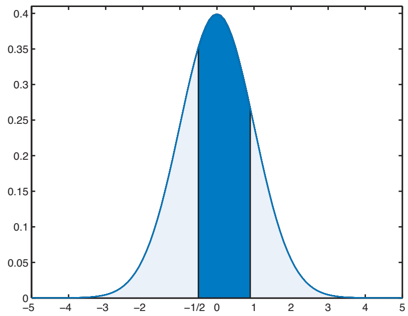
$$P(0 \leq X \leq 7) = P\left(\frac{0-3}{4} \leq \frac{X-3}{4} \leq \frac{7-3}{4}\right) = P(-3/4 \leq Z \leq 1)$$

The random variable $Z = (X-3)/4$ has the standard normal distribution.

3. The notation $X \sim N(0, 2^2)$ says that the mean is $\mu = 0$ and the standard deviation is $\sigma = 2$. Thus,

$$P(-1 \leq X \leq 2) = P\left(\frac{-1-0}{2} \leq \frac{X-0}{2} \leq \frac{2-0}{2}\right) = P(-1/2 \leq Z \leq 1)$$

Below is the graph of the *standard* normal distribution; the area of the shaded region is equal to $P(-1 \leq X \leq 2)$.



5. It is given that $\mu = 5$ and $\sigma = 10$. Thus

$$P(X < 9) = P\left(\frac{X-5}{10} < \frac{9-5}{10}\right) = P(Z < 0.4) = F(0.4) = 0.655422$$

7. It is given that $\mu = 0$ and $\sigma = 10$. Thus

$$\begin{aligned} P(X > 25) &= P\left(\frac{X-0}{10} > \frac{25-0}{10}\right) \\ &= P(Z > 2.5) \\ &= 1 - P(Z \leq 2.5) \\ &= 1 - F(2.5) = 1 - 0.993790 = 0.006210 \end{aligned}$$

9. It is given that $\mu = -5$ and $\sigma = 10$. We find

$$\begin{aligned} P(X < -10) &= P\left(\frac{X-(-5)}{10} < \frac{-10-(-5)}{10}\right) \\ &= P(Z < -0.5) \\ &= F(-0.5) \\ &= 1 - F(0.5) = 1 - 0.691462 = 0.308538 \end{aligned}$$

11. It is given that $\mu = 2$ and $\sigma = 5$. Thus

$$\begin{aligned} P(0 \le X \le 5) &= P\left(\frac{0-2}{5} \le \frac{X-2}{5} \le \frac{5-2}{5}\right) \\ &= P(-0.4 \le Z \le 0.6) \\ &= F(0.6) - F(-0.4) \\ &= F(0.6) - (1 - F(0.4)) = 0.725747 - (1 - 0.655422) = 0.381169 \end{aligned}$$

13. Let W denote the weight of a pink salmon. It is given that $W \sim N(1.7, 0.1^2)$. The ratio of pink salmon which is heavier than 1.9 kg is given by

$$\begin{aligned} P(W > 1.9) &= P\left(\frac{W-1.7}{0.1} > \frac{1.9-1.7}{0.1}\right) \\ &= P(Z > 2) \\ &= 1 - P(Z \le 2) \\ &= 1 - F(2) = 1 - 0.977250 = 0.022750 \end{aligned}$$

So, about 2.3% of salmon is heavier than 1.9 kg.

15. The mean of I is $\mu = 100$ and the standard deviation is $\sigma = 15$. We compute

$$\begin{aligned} P(I > 120) &= P\left(Z > \frac{120-100}{15}\right) \\ &= P(Z > 4/3) \\ &= 1 - P(Z \le 1.33) \\ &= 1 - F(1.33) \\ &\approx 1 - F(1.35) = 1 - 0.911492 = 0.088508 \end{aligned}$$

The probability that someone's IQ is more than 120 is about 8.85%.

17. Given $S \sim N(44, 5^2)$, we compute

$$\begin{aligned} P(S > 50) &= P\left(Z > \frac{50-44}{5}\right) \\ &= P(Z > 1.2) \\ &= 1 - P(Z \le 1.2) \\ &= 1 - F(1.2) = 1 - 0.884930 = 0.115070 \end{aligned}$$

About 11.5% of moose can run faster than 50 km/h.

19. The fraction of the population in the interval $(\mu-\sigma, \mu+\sigma)$ is 0.683. The fraction of the population in the interval $(\mu-2\sigma, \mu+2\sigma)$ which is *outside* of $(\mu-\sigma, \mu+\sigma)$ is $0.955 - 0.683 = 0.272$. The fraction of population in $(\mu-\sigma, \mu+2\sigma)$ is the fraction of the population in $(\mu-\sigma, \mu+\sigma)$ plus (because of symmetry) one half of the population in the interval $(\mu-2\sigma, \mu+2\sigma)$ which is *outside* of $(\mu-\sigma, \mu+\sigma)$. Thus, the required fraction is $0.683 + 0.272/2 = 0.819$.

21. Let X denote the given population. From $P(\mu-\sigma \le X \le \mu+\sigma) = 0.683$ it follows that $P(\mu \le X \le \mu+\sigma) = 0.683/2$ (that's because of the symmetry of the graph). Thus,

$$P(-\infty \le X \le \mu+\sigma) = P(-\infty \le X \le \mu) + P(\mu \le X \le \mu+\sigma) = 0.5 + 0.683/2 = 0.8415$$

23. Let X denote the given population. From $P(\mu-\sigma \le X \le \mu+\sigma) = 0.683$ it follows that $P(\mu-\sigma \le X \le \mu) = 0.683/2$ (because of the symmetry of the graph). Thus,

$$P(-\infty \le X \le \mu-\sigma) = P(-\infty \le X \le \mu) - P(\mu-\sigma \le X \le \mu) = 0.5 - 0.683/2 = 0.1585$$

25. X is normally distributed with mean $E(X) = 2 + 4 = 6$ and variance $\mathrm{var}(X) = 12^2 + 6^2 = 180$ (so the standard deviation of X is $\sigma = \sqrt{180}$).

27. Reducing to z-scores, we obtain

$$P(X \le x) = 0.56$$
$$P\left(Z \le \frac{x-2}{12}\right) = 0.56$$

In Table 14.4 we find

$$P(Z \le 0.15) = 0.559618$$

which is the closest value to 0.56. Thus, $(x - 2)/12 \approx 0.15$, and $x \approx 12(0.15) + 2 = 3.8$.

29. Reducing to z-scores, we obtain

$$\begin{aligned} P(X > x) &= 0.2 \\ P\left(Z > \frac{x-2}{12}\right) &= 0.2 \\ 1 - P\left(Z \le \frac{x-2}{12}\right) &= 0.2 \\ P\left(Z \le \frac{x-2}{12}\right) &= 0.8 \end{aligned}$$

In Table 14.4 we find

$$P(Z \le 0.85) = 0.802337$$

which is the closest value to 0.8. Thus, $(x - 2)/12 \approx 0.85$, and $x \approx 12(0.85) + 2 = 12.2$.

31. Denote by S the grades on the test. It is given that $S \sim N(72, 8^2)$. The ratio of students which scored more than 90% on the test is

$$\begin{aligned} P(S > 90) &= P\left(Z > \frac{90-72}{8}\right) \\ &= P(Z > 18/8 = 2.25) \\ &= 1 - P(Z \le 2.25) \\ &= 1 - F(2.25) \\ &= 1 - 0.987776 = 0.012224 \end{aligned}$$

Thus, about 1.2% of students scored more than 90% on the test.

33. Denote by S the grades on the test. It is given that $S \sim N(72, 8^2)$. We are asked to find s so that $P(S \ge s) = 0.05$. We compute

$$\begin{aligned} P(S \ge s) &= 0.05 \\ P\left(Z > \frac{s-72}{8}\right) &= 0.05 \\ 1 - P\left(Z \le \frac{s-72}{8}\right) &= 0.05 \\ P\left(Z \le \frac{s-72}{8}\right) &= 0.95 \end{aligned}$$

In Table 14.4 we find $P(Z \le 1.65) = 0.950529$, which is the closest value to 0.95. Thus, $(s - 72)/8 = 1.65$, and $s = 8(1.65) + 72 = 85.2$. So the minimum score of the highest 5% of the test scores is 85.2 (of 100).

35. Define the Bernoulli experiment

$$T_i = \begin{cases} 1 & i\text{th tree is infested by canker-rot fungus (success)} \\ 0 & i\text{th tree is not infested by canker-rot fungus} \end{cases}$$

It is given that $p = P(T_i = 1) = 0.014$ and $P(T_i = 0) = 0.986$ for $i = 1, 2, \ldots, 200$ (we find $E(T_i) = p = 0.014$ and $\text{var}(T_i) = p(1-p) = (0.014)(0.986) = 0.013804$ for all i). The random variable

$$M = \sum_{i=1}^{200} T_i$$

counts the number trees infested by canker-rot fungus.

The random variables T_i are identically distributed (and assumed to be) independent. The mean of M is (see Theorem 7 in Section 7) $E(M) = np = (200)(0.014) = 2.8$ and the variance is (see Theorem 9 in Section 9) $\text{var}(M) = np(1-p) = 200(0.014)(0.986) = 2.7608$. Using the Central Limit Theorem, we approximate M by the normal distribution $M \sim N(2.8, 2.7608)$.

The probability that fewer than 25 trees are infested with the fungus is (approximately)

$$\begin{aligned} P(M \leq 25) &= P\left(Z \leq \frac{25 - 2.8}{\sqrt{2.7608}}\right) \\ &\approx P(Z \leq 13.36089) \\ &= F(13.36089) \\ &\approx 0.999999 \end{aligned}$$

(We don't have this value in the tables, but know that it's very close to 1).

37. Consider the random variable B_i = "number of surviving offspring from ith bacterium", where $i = 1, 2, 3, \ldots, 10,000$. It is given that, for all i, $P(B_i = 2) = 0.15$, $P(B_i = 1) = 0.75$, and $P(B_i = 0) = 0.1$. We compute

$$E(B_i) = 2(0.15) + 1(0.75) + 0(0.1) = 1.05$$

From

$$E(B_i^2) = 4(0.15) + 1(0.75) + 0(0.1) = 1.35$$

we compute the variance

$$\text{var}(B_i) = E(B_i^2) - (E(B_i))^2 = 1.35 - 1.05^2 = 0.2475.$$

The random variable

$$M = \sum_{i=1}^{10,000} B_i$$

counts the number surviving offspring.

The random variables B_i are identically distributed (with mean $\mu = 1.05$ and variance $\sigma^2 = 0.2475$) and assumed to be independent. The mean of M is (see Theorem 7 in Section 7) $E(M) = n\mu = (10,000)(1.05) = 10,500$ and the variance is (see Theorem 9 in Section 9) $\text{var}(M) = n\sigma^2 = 10,000(0.2475) = 2,475$. Using the Central Limit Theorem, we approximate M by the normal distribution $M \sim N(10,500, 2,475)$.

The probability that the population will be larger than 10,000 is (approximately)

$$\begin{aligned} P(M > 10,000) &= P\left(Z > \frac{10,000 - 10,500}{\sqrt{2,475}}\right) \\ &= P(Z > -10.05) \\ &= 1 - P(Z \leq -10.05) \\ &= 1 - F(-10.05) \\ &= 1 - (1 - F(10.05)) = F(10.05) \approx 0.999999 \end{aligned}$$

($F(10.05)$ is very close to 1.)

39. Define the Bernoulli experiment

$$V = \begin{cases} 1 & \text{virus is present (success)} \\ 0 & \text{virus is absent} \end{cases}$$

It is given that $p = P(V = 1) = 0.2$ and $P(V = 0) = 0.8$. Repeat the experiment 120 times, and let N count the number of successes (number of months the virus is present). The probability that the virus will be present in between 30 and 36 months during a 10-year period is given by the sum of the probabilities of 30, 31, 32, 33, 34, 35 and 36 successes in 120 repetitions:

$$\begin{aligned} P(30 \le n \le 36) &= b(30, 120; 0.2) + b(31, 120; 0.2) + b(32, 120; 0.2) \\ &\quad + b(33, 120; 0.2) + b(34, 120; 0.2) + b(35, 120; 0.2) + b(36, 120; 0.2) \end{aligned}$$

The major difficulty in evaluating the seven expressions consists of dealing with products of very large numbers (factorials) with very small numbers (coming from the probabilities). For instance,

$$b(33, 120; 0.2) = \binom{120}{33}(0.2)^{33}(0.8)^{87} = \frac{120!}{33!\,87!}(0.2)^{33}(0.8)^{87}$$

However, thus is not a real problem if instead of a pocket calculator we use Maple, Matlab, Mathematica, or similar software.

41. Let $t = -u^2$. Then $dt/du = -2u$, and $udu = -dt/2$; we get

$$\int ue^{-u^2}\,du = \int e^t\left(-\frac{dt}{2}\right)dt = -\frac{1}{2}\int e^t\,dt = -\frac{1}{2}e^t + C = -\frac{1}{2}e^{-u^2} + C.$$

The definite integral is computed to be

$$\begin{aligned} \int_0^\infty ue^{-u^2}\,du &= \lim_{T\to\infty}\int_0^T ue^{-u^2}\,du \\ &= \lim_{T\to\infty}\left(-\frac{1}{2}e^{-u^2}\right)\Bigg|_0^T \\ &= -\frac{1}{2}\lim_{T\to\infty}\left(e^{-T^2} - e^0\right) \\ &= -\frac{1}{2}(0-1) = \frac{1}{2} \end{aligned}$$

Let $f(u) = ue^{-u^2}$. Then $f(-u) = (-u)e^{-(-u)^2} = -ue^{-u^2} = -f(u)$; i.e., $f(u)$ is an odd function. Because $\int_0^\infty ue^{-u^2}\,du$ is a convergent integral (equal to 1/2) it follows that $\int_{-\infty}^0 ue^{-u^2}\,du$ is convergent as well, and equal to $-1/2$. Thus,

$$\int_{-\infty}^\infty ue^{-u^2}\,du = \int_{-\infty}^0 ue^{-u^2}\,du + \int_0^\infty ue^{-u^2}\,du = \frac{1}{2} + \left(-\frac{1}{2}\right) = 0$$

43. (a) The calculation $g(-x) = e^{-(-x)^2} = e^{-x^2} = g(x)$ proves that g is an even function.

(b) We compute $g'(x) = -2xe^{-x^2}$. If $x > 0$, then $g'(x) < 0$ (keep in mind that $e^{-x^2} > 0$ for all x), and so g is decreasing. If $x < 0$, then $g'(x) > 0$ and g is increasing.

The equation $g'(x) = -2xe^{-x^2} = 0$ implies that $x = 0$ is the only critical point of g. Since g changes from increasing to decreasing at $x = 0$, it follows that $g(0) = 1$ is a local maximum.

Because $-x^2 \le 0$ for all x, we conclude that $e^{-x^2} \le e^0 = 1$ for all real numbers x. Thus, $g(0) = 1$ is also a global maximum of g.

(c) Differentiating g', we obtain

$$g''(x) = -2e^{-x^2} - 2xe^{-x^2}(-2x) = -2e^{-x^2}(1 - 2x^2)$$

From $g''(x) = 0$ we get $1 - 2x^2 = 0$, $x^2 = 1/2$ and $x = \pm 1/\sqrt{2}$.

If $x < -1/\sqrt{2}$, then $g''(x) > 0$ and g is concave up. If $-1/\sqrt{2} < x < 1/\sqrt{2}$, then $g''(x) < 0$ and g is concave down. If $x > 1/\sqrt{2}$, then $g''(x) > 0$ and g is concave up. Thus, $x = \pm 1/\sqrt{2}$ are points of inflection of g.

(d) We find

$$\lim_{x\to-\infty} g(x) = \lim_{x\to-\infty} e^{-x^2} = e^{-\infty} = 0$$
$$\lim_{x\to\infty} g(x) = \lim_{x\to\infty} e^{-x^2} = e^{-\infty} = 0$$

45. It is assumed that $X \sim N(\mu, \sigma^2)$. We use z-scores to convert to calculations involving the standard normal distribution:

$$\begin{aligned}
P(\mu-\sigma \le X \le \mu+\sigma) &= P\left(\frac{\mu-\sigma-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{\mu+\sigma-\mu}{\sigma}\right)\\
&= P(-1 \le Z \le 1)\\
&= F(1) - F(-1)\\
&= F(1) - (1 - F(1))\\
&= 2F(1) - 1 = 2(0.841345) - 1 = 0.682690 \approx 0.683
\end{aligned}$$

Likewise,

$$\begin{aligned}
P(\mu-2\sigma \le X \le \mu+2\sigma) &= P\left(\frac{\mu-2\sigma-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{\mu+2\sigma-\mu}{\sigma}\right)\\
&= P(-2 \le Z \le 2)\\
&= F(2) - F(-2)\\
&= 2F(2) - 1 = 2(0.977250) - 1 = 0.9545 \approx 0.955
\end{aligned}$$

and

$$\begin{aligned}
P(\mu-3\sigma \le X \le \mu+3\sigma) &= P\left(\frac{\mu-3\sigma-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{\mu+3\sigma-\mu}{\sigma}\right)\\
&= P(-3 \le Z \le 3)\\
&= F(3) - F(-3)\\
&= 2F(3) - 1 = 2(0.998650) - 1 = 0.9973 \approx 0.997
\end{aligned}$$

Section 15 The Uniform and the Exponential Distributions

1. From $\operatorname{var}(U) = (b-0)^2/12 = 12$ we get $b^2 = 12^2$ and $b = 12$ (since $b > 0$). The mean of U is $E(U) = (0+12)/2 = 6$.

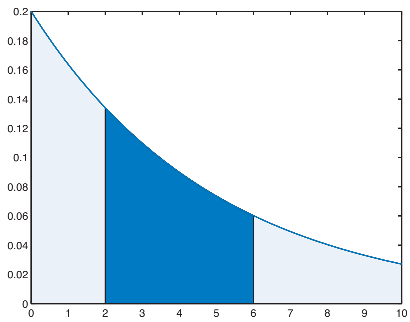
3. (a) The probability density function is $f(t) = 0.2e^{-0.2t}$ and the cumulative distribution function is $F(t) = 1 - e^{-0.2t}$. The probability that the first event occurs between times 2 and 6 is

$$\begin{aligned} P(2 \le T \le 6) &= \int_2^6 0.2e^{-0.2t}\,dt \\ &= \left(-e^{-0.2t}\right)\Big|_2^6 \\ &= -e^{-1.2} + e^{-0.4} \approx 0.369126 \end{aligned}$$

Alternatively, using the cumulative distribution function,

$$P(2 \le T \le 6) = F(6) - F(2) = \left(1 - e^{-0.2(6)}\right) - \left(1 - e^{-0.2(2)}\right) = -e^{-1.2} + e^{-0.4} \approx 0.369126$$

(b) See below.



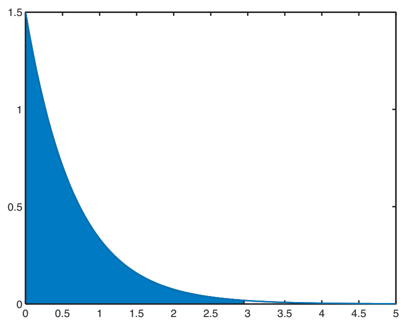
5. (a) The probability density function is $f(t) = 1.5e^{-1.5t}$ and the cumulative distribution function is $F(t) = 1 - e^{-1.5t}$. The probability that the first event occurs before $t = 3$ is

$$\begin{aligned} P(T < 3) &= \int_0^3 1.5e^{-1.5t}\,dt \\ &= \left(-e^{-1.5t}\right)\Big|_0^3 \\ &= -e^{-4.5} + 1 \approx 0.988891 \end{aligned}$$

Alternatively, using the cumulative distribution function, we obtain

$$P(T < 3) = F(3) = 1 - e^{-1.5(3)} = 1 - e^{-4.5} \approx 0.988891$$

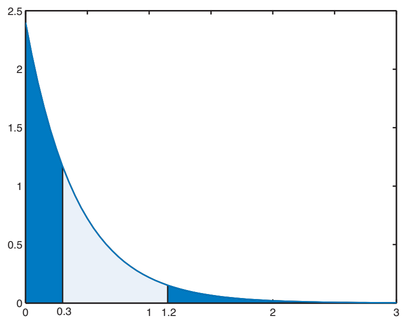
(b) See below.



7. (a) The probability density function is $f(t) = 2.4e^{-2.4t}$ and the cumulative distribution function is $F(t) = 1 - e^{-2.4t}$. The probability that the first event occurs before $t = 0.3$ or after $t = 1.2$ is

$$\begin{aligned} P(T < 0.3) + P(T > 1.2) &= P(T < 0.3) + (1 - P(T \leq 1.2)) \\ &= F(0.3) + 1 - F(1.2) \\ &= (1 - e^{-2.4(0.3)}) + 1 - (1 - e^{-2.4(1.2)}) \approx 0.569383 \end{aligned}$$

(b) See below.



9. Given $s(t) = e^{-0.4t}$, we identify $\lambda = 0.4$/month. The mean lifetime is $1/\lambda = 1/0.4 = 2.5$ months. From $s(3) = e^{-0.4(3)} = e^{-1.2} \approx 0.301194$ we conclude that about 30.1% of insects will survive 3 months.

11. Denote the lifespan of the atom by T. Since the expected lifespan is 4 hours, it follows that $\lambda = 1/4 = 0.25$/hour. The probability density function of T is $f(t) = 0.25e^{-0.25t}$, the cumulative distribution function is $F(t) = 1 - e^{-0.25t}$, and the survivorship function is $s(t) = e^{-0.25t}$.

The probability that the atom will not decay during the first 3 hours is

$$P(T > 3) = 1 - P(T \leq 3) = 1 - F(3) = s(3) = e^{-0.25(3)} = e^{-0.75} \approx 0.472367$$

Repeating this calculation, we obtain the probability that the atom will decay after 6 hours:

$$P(T > 6) = s(6) = e^{-0.25(6)} = e^{-1.5} \approx 0.223130$$

13. (a) The average lifespan of a guinea pig is $1/0.18 \approx 5.56$ years.

(b) The survivorship function for the guinea pig is $s(t) = e^{-0.18t}$. Thus, the chance that a guinea pig will live longer than 6 years is $s(6) = e^{-0.18(6)} \approx 0.236928$.

(c) Let T represent the lifetime of a guinea pig. We find

$$\begin{aligned} P(T > 8 \,|\, T > 2) &= \frac{P((T > 8) \cap (T > 2))}{P(T > 2)} \\ &= \frac{P(T > 8)}{P(T > 2)} \\ &= \frac{s(8)}{s(2)} \\ &= \frac{e^{-0.18(8)}}{e^{-0.18(2)}} = e^{-0.18(6)} = s(6) \end{aligned}$$

The answer is the same as in (b).

15. Young and old organisms are more likely to die, since the survivorship curve is sharply decreasing for them. After the initial sharp drop, the curve continues with a small negative slope. Thus, an adult organism has a good change of living bit longer (until it reaches the age where the survivorship curve drops quickly again).

17. This is not hard to guess: the function $f(x) = 5x$ stretches by a factor of 5: it maps the interval $(0, 1)$ to the interval $(0, 5)$. Now we shift by 3 units, so the answer is $f(x) = 5x + 3$.

(Formally: we are looking for a linear function that maps the initial point of the first interval (0) to the initial point of the second interval (3) and the terminal point of the first interval (1) to the terminal point of the second interval (8). In other words, we are looking for an equation of a line through the points $(0, 3)$ and $(1, 8)$. Using the point-slope equation, we get $y - 3 = \frac{8-3}{1-0}(x - 0)$, i.e., $y = 5x + 3$.)

By generating random numbers in the interval $(0, 1)$ and then applying $f(x)$ to them, we generate random numbers in the interval $(3, 8)$.

The length of the interval (a, b) is $b - a$. Thus $f(x) = (b - a)x$ transforms the interval $(0, 1)$ to $(0, b - a)$. Now we move it so that it starts at a; the function $f(x) = (b - a)x + a$ maps the interval $(0, 1)$ to $(0 + a, b - a + a) = (a, b)$. So, composing a random number generator on the interval $(0, 1)$ with $f(x)$ we obtain a random number generator on the interval (a, b).

19. The half-life of a radioactive substance is the time t_h for which $P(T > t_h) = s(t_h) = 1/2$. From $e^{-\lambda t_h} = 1/2$ we obtain

$$-\lambda t_h = \ln(1/2) = \ln 1 - \ln 2 = -\ln 2$$

and $t_h = \ln 2/\lambda$.

The median is the time t_m such that $F(t_m) = 1 - e^{-\lambda t_m} = 1/2$, i.e., $e^{-\lambda t_m} = 1/2$.

We see that $t_m = t_h$. From $F(t) = 1 - e^{-\lambda t} = 1 - s(t)$ we conclude that $F(t) + s(t) = 1$. So, if one of the $F(t)$ or $s(t)$ is $1/2$, so is the other. Or: the half-life is the time t when the probability of surviving $s(t)$ is the same as the probability of dying $F(t)$.