Probability and Statistics for SIC
Exercises

Chapter 1

Solution 1 If the set of distinct characters is $\mathcal{C}$, then the set of passwords is the Cartesian product $\mathcal{C} \times \mathcal{C} \times \mathcal{C} \times \mathcal{C} \times \mathcal{C} \times \mathcal{C} \times \mathcal{C} \times \mathcal{C}$, and this has size $|\mathcal{C}^8| = |\mathcal{C}|^8 = 66^8$.

Solution 2 With sets $\mathcal{A} = \{A, \dots, Z\}$ and $\mathcal{D} = \{0, \dots, 9\}$, using the multiplication rule and Cartesian products, the set of possibilities is $\mathcal{A}^2 \times \mathcal{D}^3$, which has size $|\mathcal{A}|^2 \times |\mathcal{D}|^3 = 26^2 \times 10^3$.

Solution 3 There are 23! distinct ways to order the maths books, and 9! distinct ways to order the physics books, and 2! distinct orders for the types of books, so the answer is $23! \times 9! \times 2!$.

Solution 4 Using the logic of the previous exercise, the answer is $4! \times 3! \times 5! \times 3!$.

Solution 5 We must take into account the permutations of A and B: $2! = 2$, and that the k people between A and B and $n-k-2$ people around A and B are not to be taken independently. Thus, there are $(n-2)$ permutations to place everyone except A and B, and $1 + (n-k-2) = n-k-1$ ways of placing the block "A ... B" of length $k+2$ in the queue. Therefore, there are $2(n-2)!(n-k-1)$ ways of having k people between A and B.

Solution 6 A committee is an unordered 4-set of persons, and there are C_{23}^4 distinct 4-sets that can be made from 23 persons.

Solution 7 The logic of the previous solution gives C_{52}^5.

Solution 8 a) There are $C_{10}^3 \times C_8^3 = 6720$ ways of choosing a committee of 3 men and 3 women from a group of 10 women and 8 men: the choices of men and women are independent, so the numbers of possibilities can be multiplied, and within each group an unordered selection is made without replacement.

b) The number of committees of 3 men containing both men who refuse to be together is $C_6^1 = 6$ (there are six ways to choose the other man). So, the number of committees of 3 men without these 2 men is $C_8^3 - 6 = 50$. Therefore the answer sought is

$$C_{10}^3 \times (C_8^3 - 6) = 120 \times 50 = 6000.$$

c) The reasoning here is similar to that in b), giving

$$(C_{10}^3 - C_8^1) \times C_8^3 = (120 - 8) \times 56 = 6272.$$

d) The number of committees with the man and the woman who refuse to work together is $C_9^2 C_7^2 = 756$. So the answer is

$$C_{10}^3 \times C_8^3 - C_9^2 \times C_7^2 = 6720 - 756 = 5964.$$

Solution 9 There are $4 \times 3 \times 2 = 24$ ways of placing the three pairs of twins in the 4 rooms. Furthermore, for each allocation of twins to rooms there are $2^3 = 8$ ways of distributing the three pairs of twins in the beds. There are therefore $24 \times 2^3 = 192$ ways of setting up the experiment.

Solution 10 A direct calculation gives

$$C_{n-1}^{k-1} + C_{n-1}^{k} = \frac{(n-1)!}{(n-k)!(k-1)!} + \frac{(n-1)!}{(n-k-1)!k!} = \frac{(n-1)!(k+n-k)}{(n-k)!k!} = C_n^k.$$

Chapter 2

Solution 11 a) The sample space for this experiment is $\Omega = \Omega_1 \cup \Omega_2$, where

$$\Omega_1 = \{(D_1, D_2, \ldots) : D_i \in \{1, 2, 3, 4, 5\}, i = 1, 2, \ldots\}$$

represents the event 'no 6 is cast' and

$$\Omega_2 = \{6\} \cup \bigcup_{n=1}^{\infty} \{(D_1, D_2, \ldots, D_n, 6) : D_i \in \{1, 2, 3, 4, 5\}, i = 1, 2, \ldots\}$$

represents the event 'the experiment stops'.

It may seem puzzling why Ω_1 is needed. Two reasons are: (i) no probabilities have been given for the outcomes, and perhaps the probability of getting a '6' is zero because the die is biased, in which case only Ω_1 could occur; (ii) any event space must contain infinite unions and complements, so if the possible events include $E_n \equiv$ 'the first six occurs on throw n' ($n = 1, 2, \ldots$), then the event $\bigcup_{n=1}^{\infty} E_n \equiv$ 'the first six occurs for some $n \in \mathbb{N}$' is in the event space, and so is its complement, Ω_1.

b) The points of the sample space which are contained in E_n are of the form $(D_1, \ldots, D_{n-1}, 6)$. The set $(\cup_1^{\infty} E_n)^c$ corresponds to the event "no 6 is cast", that is Ω_1.

Solution 12 a) The sample space Ω comprises all possible ordered arrangements of n people, so $|\Omega| = n!$. If we assume that the event space contains all subsets of Ω, then the number of events is $2^{|\Omega|} = 2^{n!}$.

b) Let the event E_k be "there are k people between A and B". Let us determine $\#E_k =$"number of favourable cases of E_k":

Consider first the block "$A \overset{\leftarrow k \rightarrow}{\ldots} B$" of length $k+2$. There are $1 + (n-k-2) = (n-k-1)$ ways to place it in the line.

Then take into account the permutations

- of A and B: $2! = 2$ permutations,
- of the $(n-2)$ people different to A and B: $(n-2)!$ different permutations.

Thus

$$\#E_k = 2 \times (n-k-1) \times (n-2)!$$

and

$$\Pr(E_k) = \frac{\#E_k}{\#\Omega} = \frac{2(n-k-1)(n-2)!}{n!} = \frac{2(n-k-1)}{n(n-1)}.$$

NB: One can check that $\sum_{k=0}^{n-2} 2(n-k-1)/\{n(n-1)\} = 1$.

c) It is easy to establish the list of the $3! = 6$ possible cases. We then obtain

$$\Pr(k=0) = \frac{4}{6} = \frac{2}{3}, \quad \Pr(k=1) = \frac{2}{6} = \frac{1}{3},$$

which corresponds to the expression found under b).

NB: We have $\Pr(k=0) + \Pr(k=1) = 1$.

Solution 13 a) Two signals S_1 and S_2 reach the receptor in the interval $(0, t)$. Let X_1 be the arrival time of S_1 and let X_2 be the arrival time of S_2. The sample space of possible results is therefore

$$\Omega = \left\{(x_1, x_2) \in R^2 : 0 \le x_1, x_2 \le t\right\}.$$

One element of this set (a pair (x_1, x_2)) represents a possible outcome of this experiment (that is an elementary event of the sample space): "the signal S_1 arrives at time x_1 and the signal S_2 arrives at time x_2".

b) The event A which we are interested in (i.e. ,"the receptor blocks") is a subset of Ω defined as

$$A = \{(x_1, x_2) \in \Omega : |x_1 - x_2| < \theta\};$$

for the next step it is helpful to sketch A, and note that its area is the area of the square Ω minus the area of two triangles, each of which has area $(t-\theta)^2/2$, The fact that both signals "arrive independently of each other and at random"' tells us we are dealing with an equiprobable model. We can then obtain

$$\Pr(A) = \frac{\text{area of } A}{\text{area of } \Omega} = \frac{t^2 - 2 \times (t-\theta)^2/2}{t^2} = \frac{2\theta t - \theta^2}{t^2}.$$

c) If $\theta \ll t$, we can neglect $(\theta/t)^2$ relative to $2\theta/t$, giving $\Pr(A) \simeq 2\theta/t$.

Solution 14 a) The sample space of this experiment is $\Omega = \{(y_1, y_2, y_3) : y_1, y_2, y_3 \in \{1, \ldots, 6\}\}$.

b) Let X_1 be the die which shows the smallest number, X_3 the die which shows the greatest number and X_2 the die which shows a number between X_1 and X_3 (that is $X_1 \leq X_2 \leq X_3$). We want to calculate the probability of the event "$X_1 + X_2 + X_3 \geq 15$". We will call this event E. The table below shows the values X_1, X_2 and X_3 which make their sum greater than or equal to 15.

X_1	X_2	X_3	sum	weight	probability
3	6	6	15	3	1/216
4	5	6	15	3!	1/216
4	6	6	16	3	1/216
5	5	5	15	1	1/216
5	5	6	16	3	1/216
5	6	6	17	3	1/216
6	6	6	18	1	1/216

The fifth column represents the number of ways there are to obtain the given values of X_1, X_2 and X_3. The sixth column represents the probability of each configuration (that is $1/6^3 = 1/216$). Thus, the probability sought is

$$\Pr(E) = \frac{(3+6+3+1+3+3+1)}{216} = \frac{20}{216} \simeq 0.0926.$$

Solution 15 a) and b) The answers are the same:

$$P = \frac{15 \times 14 \times 13 \times 10 \times 9}{25 \times 24 \times 23 \times 22 \times 21} = 0.03854.$$

c) We obtain

$$P = \frac{15 \times 14 \times 10 \times 9 \times 8}{25 \times 24 \times 23 \times 22 \times 21} = 0.0237.$$

Solution 16 First note that if Xavier and both his parents have brown eyes and his sister has blue eyes, that must mean that both parents each have one gene for blue eyes, and the other for brown.

a) Xavier can have the genes $(Br, Br), (Br, B)$ and (B, Br), so the probability that he has one gene for blue eyes is 2/3.

b) Let G_X, G_W and G_C respectively denote the eye genes belonging to Xavier, his wife, and his child. So,

$$\Pr(G_C = (B, B)) = \Pr(G_X \in \{(Br, B), (B, Br)\}) \times \tfrac{1}{2} = \tfrac{2}{3} \times \tfrac{1}{2} = \tfrac{1}{3}.$$

Solution 17 Here $\Omega = \{(a, b) : a, b \in \{1, \ldots, 6\}\}$ is the collection of ordered pairs (because the dice can be distinguished) of numbers between 1 and 6.

a) $E \cap F = \{(1, 2), (1, 4), (1, 6), (2, 1), (4, 1), (6, 1)\}$.

b) $E \cup F$ is the event "the sum of the dice is odd or one of the dice shows a 1".

c) $F \cap G = \{(1, 4), (4, 1)\}$.

d) $E \cap F^c$ is equal to the event "the sum of the dice is odd and each die shows a number greater than or equal to 2".

e) $E \cap F \cap G = F \cap G$.

Solution 18 a) $000\ldots$ represents the event "no-one wins". The other sequences correspond to the events in which A, B or C wins.

b) (i) Those with the 1 in position $3n + 1$ for $n \in \{0, 1, \ldots\}$ correspond to the event 'A wins'.

(ii) Those with the 1 in position $3n + 2$ for $n \in \{0, 1, \ldots\}$ correspond to the event 'B wins'.

(iii) Those with the 1 in position $3n$ for $n \in \{0, 1, \ldots\}$ correspond to the event 'C wins'.

Finally, $(A \cup B)^c$ is event in which either C wins or no-one wins.

Note: in i), for example, $3n + 1$ represents the sequence $(3n)$ + the position the winner is in $(+1)$.

Solution 19 No, because the configurations are not equiprobable. If we take into account the order in which they appear, we must then take into account the possible permutations of each configuration. Thus (3,3,3) "counts only once" whereas (5,2,2) "counts three times" and (5,3,1) "counts six times". We obtain $\Pr(S = 9) = \frac{25}{6^3}$ and $\Pr(S = 10) = \frac{27}{6^3}$.

Solution 20 The probability of obtaining at least one 6 equals unity minus the probability of casting no 6s: Pr(at least one 6) = 1 − Pr(no 6s). The probability of obtaining no 6s is $(5/6)^4 \simeq 0.4822$, so the probability of casting at least one 6 is $1 - (5/6)^4 \simeq 0.5177$.

Solution 21 The probability that the 6 appears at least once is $1-(5/6)^{2n}$. In order for this probability to reach 1/2, we must have $1-(5/6)^{2n} \geq 1/2$, or equivalently

$$1-1/2 \geq (5/6)^{2n} \Leftrightarrow \log(1/2) \geq 2n\log(5/6) \Leftrightarrow -\log 2 \geq 2n\times\{-\log(6/5)\} \Leftrightarrow n \geq \tfrac{1}{2}\log 2/\log(6/5) \approx 1.9.$$

Thus we must have $n \geq 2$.

Solution 22 The probability that the birthdays of n people are in different months is

$$\frac{12\times 11\times\cdots\times(12-n+1)}{12^n}$$

So the probability that at least two of the n people have their birthdays the same month is

$$1-\frac{12\times 11\times\cdots\times(12-n+1)}{12^n}$$

You need at least 5 people for this probability to exceed 1/2.

Solution 23 The probability that neither of the two coins shows tails is the probability of the outcome (H,H), i.e., 1/4, so the answer is $1-1/4=3/4$.

Solution 24 We use the fact that the event "have at least one success" complements the event "have no success". The probability of obtaining at least one 6 with 4 dice is therefore

$$1-(1-\tfrac{1}{6})^4 \simeq 0.518,$$

and the probability of obtaining at least one double 6 in 24 casts of two dice is

$$1-(1-\tfrac{1}{36})^{24} \simeq 0.491.$$

Solution 25 $\Pr\{\text{"At least one ball} \geq 17\text{"}\} = 1-\Pr\{\text{"All balls} < 17\text{"}\} = 1-\frac{16\times 15\times 14}{20\times 19\times 18}$.

Solution 26 The sample space of this experiment is the set of four ordered pairs $\Omega=\{(E_1,E_2)\}$ where E_1 is the sex of the first child and E_2 is the sex of the second. The event "both children are girls" is $A=\{(G,G)\}$, and "the eldest is a girl" is $B=\{(G,G),(G,B)\}$. The required probability is

$$\Pr(A\mid B)=\frac{\Pr(A\cap B)}{\Pr(B)}=\frac{\Pr(A)}{\Pr(B)}=\frac{1/4}{1/2}=\tfrac{1}{2}.$$

Solution 27 The sample space of this experiment is $\Omega=\{(D_1,D_2): 1\leq D_i\leq 6, i=1,2\}$. Let us call A the event "at least one die shows a 6", and B the event "both results are different". We want to calculate

$$\Pr(A\mid B)=\frac{\Pr(A\cap B)}{\Pr(B)}.$$

But

$$\Pr(B)=\tfrac{6\times 5}{6\times 6}=\tfrac{5}{6},\quad \Pr(A\cap B)=\Pr(B)-\Pr(A^c\cap B)=\tfrac{5}{6}-\tfrac{5\times 4}{6\times 6}=\tfrac{5}{6}-\tfrac{5}{9},$$

so the required probability is

$$\Pr(A\mid B)=1-\tfrac{6}{9}=1/3.$$

Solution 28 Let A denote the event "the first card is a spade" and B the event "the other two are spades". Since

$$\Pr(A\cap B)=\tfrac{13}{52}\tfrac{12}{51}\tfrac{11}{50},\quad \Pr(B)=\Pr(A\cap B)+\Pr(A^c\cap B)=\tfrac{13}{52}\tfrac{12}{51}\tfrac{11}{50}+\tfrac{39}{52}\tfrac{13}{51}\tfrac{12}{50},$$

the required probability is

$$\Pr(A\mid B)=\frac{\Pr(A\cap B)}{\Pr(B)}=\frac{13\times 12\times 11}{13\times 12\times 11+39\times 13\times 12}=\tfrac{11}{50}.$$

Solution 29 Let S_n denote the event "the n^{th} day is sunny" and N_n the event "the n^{th} day is cloudy". Then

$$\begin{aligned} s_n &= \Pr(S_n) = \Pr(S_n \mid S_{n-1})\Pr(S_{n-1}) + \Pr(S_n \mid N_{n-1})\Pr(N_{n-1}) \\ &= ps_{n-1} + q(1 - s_{n-1}) \\ &= (p - q)s_{n-1} + q \end{aligned}$$

We show that $s_n = \frac{1}{2}(1 + (p - q)^n)$, $n \geq 0$, by induction on n. If $n = 0$ this is clearly true. Let us now suppose that this formula is valid for m. Then, since $p + q = 1$,

$$\begin{aligned} s_{m+1} &= (p - q)s_m + q \\ &= (p - q)\left(\tfrac{1}{2} + \tfrac{1}{2}(p - q)^m\right) + q \\ &= \frac{p + q}{2} + \tfrac{1}{2}(p - q)^{m+1} \\ &= \tfrac{1}{2}\left\{1 + (p - q)^{m+1}\right\}, \end{aligned}$$

so the result is true for $m + 1$. Thus $s_n = \frac{1}{2}(1 + (p - q)^n)$ for all $n \geq 0$.

Solution 30 a) Let S_n be the outcome (either H or T, for 'heads' or 'tails') for toss n and let E_n be the event "two successive tails don't appear". Write $P_n = \Pr(E_n)$ and note that

$$\Pr(E_n) = \Pr(E_n \cap \{S_n = T\}) + \Pr(E_n \cap \{S_n = H\}).$$

Now

$$\Pr(E_n \cap \{S_n = H\}) = \Pr(E_{n-1} \cap \{S_n = H\}) = \tfrac{1}{2}\Pr(E_{n-1})$$

as if the last throw is an H, then the sequence ends with HH or TH and thus the last toss is independent of E_{n-1}. By a similar argument,

$$\Pr(E_n \cap \{S_n = T\}) = \Pr(E_{n-2} \cap \{S_{n-1} = H\} \cap \{S_n = T\}) = \tfrac{1}{4}\Pr(E_{n-2}),$$

because $\Pr(E_n \cap \{S_{n-1} = T\} \cap \{S_n = T\}) = 0$. Hence

$$P_n = \tfrac{1}{2}P_{n-1} + \tfrac{1}{4}P_{n-2}. \qquad (1)$$

b) Since P_n is a decreasing function of n, i.e., $0 \leq P_{n+1} \leq P_n$, the limit $\lim_{n\to\infty} P_n$ exists. Let us call it $P_\infty \geq 0$. In taking the limit $n \to \infty$ in (1) we deduce that

$$P_\infty = \tfrac{3}{4}P_\infty,$$

which is possible only if $P_\infty = 0$.

c) Let $G_{n,i}$ denote the event "$\{S_{8i} = T, S_{8i+1} = H, S_{8i+2} = T, S_{8i+3} = T, S_{8i+4} = T, S_{8i+5} = H, S_{8i+6} = T, S_{8(i+1)-1} = T\}$", and let $R_{n,i} = G_{n,i}^c$ (that is the event "the series $S_j, 8i \leq j \leq 8(i+1) - 1$, is different from T, H, T, T, T, H, H, T"). Let $R_n = \cap_{i=1}^{n/8-8} R_{n,i}$. Through the independence of the events $R_{n,i}$, for $i = 1, \ldots, , n/8$, we have

$$\Pr(R_n) = \prod_{i=1}^{n/8} \Pr(R_{n,i}) = \left(1 - \frac{1}{2^8}\right)^{n/8} \qquad (2)$$

where in the last equality we have used the fact that $\Pr(R_{n,i}) = 1 - 1/2^8$. From (2), we deduce that $\lim_{n\to\infty} \Pr(R_n) = 0$. But clearly $\Pr(Q_n) \leq \Pr(R_n)$ (because $Q_n \subset R_n$). Thus, $\lim_{n\to\infty} \Pr(Q_n) = 0$.

Solution 31 Let RR, NN and RN respectively represent the events "the chosen card is entirely red", "entirely black" and "bicoloured". Let R denote the event "the visible side of the chosen card is red". Then

$$\begin{aligned} \Pr(RN \mid R) &= \frac{\Pr(R \mid RN)\Pr(RN)}{\Pr(R \mid RR)\Pr(RR) + \Pr(R \mid RN)\Pr(RN) + \Pr(R \mid NN)\Pr(NN)} \\ &= \frac{\frac{1}{2}\frac{1}{3}}{1 \times \frac{1}{3} + \frac{1}{2}\frac{1}{3} + 0 \times \frac{1}{3}} = \tfrac{1}{3}. \end{aligned}$$

Solution 32 Let C denote the event "the selected person is colour-blind", M the event "that person is a man" and F the event "that person is a woman". Bayes' formula gives

$$\Pr(M \mid C) = \frac{\Pr(C \mid M)\Pr(M)}{\Pr(C \mid M)\Pr(M) + \Pr(C \mid F)\Pr(F)}$$

If we say that there are as many men as women, then $\Pr(M) = \Pr(F) = 1/2$ and

$$\Pr(M \mid C) = \frac{\frac{5}{100} \times \frac{1}{2}}{\frac{5}{100} \times \frac{1}{2} + \frac{0.25}{100} \times \frac{1}{2}} = \frac{5}{5.25} = \frac{20}{21}.$$

If on the other hand there were twice as many woman as men, we would have

$$\Pr(M \mid C) = \frac{\frac{5}{100} \times \frac{1}{3}}{\frac{5}{100} \times \frac{1}{3} + \frac{0.25}{100} \times \frac{2}{3}} = \frac{5}{5.5} = \frac{10}{11}.$$

Solution 33 Let M denote the event "the patient is infected", B the event "the patient is healthy", and $+$ the event "the result of the test is positive". Bayes' formula gives

$$\begin{aligned}\Pr(M \mid +) &= \frac{\Pr(+ \mid M)\Pr(M)}{\Pr(+ \mid M)\Pr(M) + \Pr(+ \mid B)\Pr(B)} \\ &= \frac{\frac{99}{100}\frac{1}{1000}}{\frac{99}{100}\frac{1}{1000} + \frac{2}{100}\frac{999}{1000}} \\ &= \frac{99}{2097} \simeq 0.0472.\end{aligned}$$

This isn't very helpful. For a better result, the test would have to be repeated on the same individual.

Solution 34 Let A denote the event that a piece is accepted, and B denote the event that it is good. Then $\Pr(A \mid B) = 0.9$ and $\Pr(A^c \mid B^c) = 0.8$.

a) All 4 pieces are accepted, therefore 3 good pieces are checked correctly and there is an error in the checking of the defective piece. The probability of this event is:

$$\Pr(A \mid B)^3 \times \Pr(A \mid B^c) = (0.9)^3 \times 0.2 \simeq 0.146.$$

b) Let the event $E =$ denote "there is an error during the checking of a piece". Then

$$\Pr(E) = \Pr(A^c \mid B) \times \Pr(B) + \Pr(A \mid B^c) \times \Pr(B^c),$$

so, since $\Pr(B^c) = 0.2$, $\Pr(E) = 0.1 \times 0.8 + 0.2 \times 0.2 = 0.12$.

c) Bayes' theorem gives

$$\Pr(B^c \mid A) = \frac{\Pr(A \mid B^c) \times \Pr(B^c)}{\Pr(A)} = \frac{\Pr(A \mid B^c) \times \Pr(B^c)}{\Pr(A \mid B) \times \Pr(B) + \Pr(A \mid B^c) \times \Pr(B^c)} \simeq 0.053.$$

Chapter 3

Solution 35 X_n can only take the values 0 and 1. There cannot be more than one broken-down machine at the start of a day's work. Let B_n be the random variable equal to the number of machines that fail on the n^{th} day. Then

$$\begin{aligned}\Pr(X_{n+1} = 0 \mid X_n = 0) &= \Pr(B_n = 0 \cup B_n = 1) \\ &= \Pr(B_n = 0 \mid X_n = 0) + \Pr(B_n = 1 | X_n = 0) \\ &= p^2 + p(1-p) + p(1-p) = p(2-p) \\ &\text{(neither machine fails,} \\ &\text{one machine fails and is repaired,} \\ &\text{the other machine fails and is repaired)} \\ \Pr(X_{n+1} = 0 \mid X_n = 1) &= \Pr(B_n = 0 \mid X_n = 1) \\ &= \Pr(\text{the only machine that is not broken doesn't fail}) = p, \\ \Pr(X_{n+1} = 1 \mid X_n = 0) &= \Pr(B_n = 2 | X_n = 0) = (1-p)^2, \\ \Pr(X_{n+1} = 1 \mid X_n = 1) &= \Pr(\text{the machine that is not yet broken fails}) = 1 - p.\end{aligned}$$

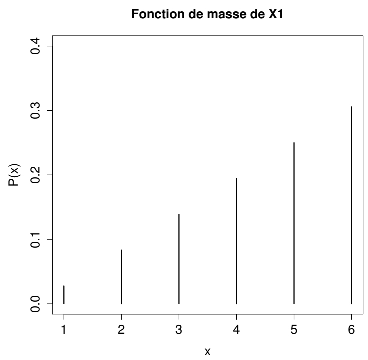
Solution 36 The space of the possible values taken by X_1 is $S = \{1, 2, 3, 4, 5, 6\}$. The enumeration of different possible cases gives us:

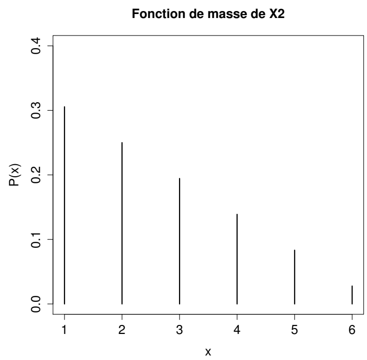
$$\begin{aligned} \Pr(X_1 = 1) &= 1/36, & \Pr(X_1 = 2) &= 1/12, \\ \Pr(X_1 = 3) &= 5/36, & \Pr(X_1 = 4) &= 7/36, \\ \Pr(X_1 = 5) &= 9/36, & \Pr(X_1 = 6) &= 11/36. \end{aligned}$$

Similarly, for X_2,

$$\begin{aligned} \Pr(X_2 = 1) &= 11/36, & \Pr(X_2 = 2) &= 9/36, \\ \Pr(X_2 = 3) &= 7/36, & \Pr(X_2 = 4) &= 5/36, \\ \Pr(X_2 = 5) &= 3/36, & \Pr(X_2 = 6) &= 1/36. \end{aligned}$$

We get the following graphs:





Solution 37 We suppose that birthdays are evenly spread over the year and, to simplify, that a year comprises 365 days.

a) The probability that a random person was born on a January 1st is therefore 1/365. The probability that both partners were born on a January 1st is, supposing independence, $1/365^2$. Among the 42800 couples married in 2010, the number of those in which both partners were born on a January 1st, X, follows binomial distribution of parameters $n = 42800$ and $p = 1/365^2$. Thus

$$\Pr(X = 2) = C_{42800}^2 \frac{1}{365^4} \left(1 - \frac{1}{365^2}\right)^{42798} = 0.0374.$$

b) The probability that both partners were born on the same day is 365 times higher than the probability that they were both born on a January 1st. The number Y of couples married having their birthdays the same day therefore follows a binomial distribution of parameters $n = 42800$ and $p = 1/365$. We then have

$$\Pr(X = 2) = C_{42800}^2 \frac{1}{365^2} \left(1 - \frac{1}{365}\right)^{42798} \simeq 0.$$

Solution 38 a) Here we are dealing with a draw with replacement, since the same animal can be seen twice. The probability that any given sighting is of a lion is $L/(L + T)$. The number of lions noted in the report follows a binomial distribution of parameters n and $p = L/(L + T)$. The probability that k lions have been noted is therefore

$$\binom{n}{k} \left(\frac{L}{L+T}\right)^k \left(\frac{T}{L+T}\right)^{n-k}, \quad k = 0, \dots, n.$$

b) Here we have a draw without replacement. The number of lions captured follows a hypergeometric distribution of parameters L, T and n. The probability that k lions have been captured is

$$\frac{\binom{L}{k}\binom{T}{n-k}}{\binom{L+T}{n}}, \quad k = \max(0, n - T), \dots, \min(L, n).$$

Solution 39 a) Arnaud bets the amount $100 \times 2^{n-1}$ CHF at turn n.

b) Since each turn is a Bernoulli experiment of parameter p, the number of turns N until the first win follows a geometric distribution of parameter p. Its mass function is therefore

$$\Pr(N = n) = p(1 - p)^{n-1}, \quad n = 1, 2, \dots.$$

The event $\{N = n\}$ happens if the $(n-1)$ first turns have been lost and the nth has been won.

c) Let Y denote the amount invested during the last turn, that is to say when Arnaud wins for the first time. We have $Y = 100 \times 2^{N-1}$ where N is the random variable of the number of turns until the first win. We have

$$\begin{aligned} \mathrm{E}(Y) &= \mathrm{E}(100 \times 2^{N-1}) = \sum_{n=1}^{\infty}(100 \times 2^{n-1})\Pr(N=n) = \sum_{n=1}^{\infty}(100 \times 2^{n-1})p(1-p)^{n-1} \\ &= 100p\sum_{n=1}^{\infty}\{2(1-p)\}^{n-1} = \begin{cases} \frac{100p}{1-2(1-p)}, & p > 1/2, \\ \infty, & p \le 1/2. \end{cases} \end{aligned}$$

If $p \le 1/2$, it is best to be infinitely rich to be able to follow this strategy! If $p > 1/2$, a finite fortune should be sufficient, at least on average—but if a casino sees that you are playing this strategy, they will throw you out!

Solution 40 a) Let X_i denote the sum of two dice at the ith throw: $X_i \in \{2, \ldots, 12\}$. We have $\Pr(X_i = 5) = 1/9$ and $\Pr(X_i = 7) = 1/6$. The time of end of play is called τ, $\tau = 1, 2, \ldots$, and $F_j = \{\tau = j\}$ is the event in which the game ends at the jth throw. Note that we can write the event F_j as

$$\{X_j = 5 \text{ or } X_j = 7\} \cap \left[\cap_{i=1}^{j-1} \{X_i \neq 5 \text{ and } \mathrm{X_i} \neq 7\}\right].$$

By using this and the independence of the variables X_i we get that for every j

$$\Pr(X_j = 5|F_j) = P\left(\{X_j = 5\}|\{X_j = 5\} \cup \{X_j = 7\}\right) = \frac{1/9}{1/9 + 1/6} = \frac{2}{5}.$$

b) Since the game is almost sure to end (that is $\sum_{j=1}^{\infty} \Pr(F_j) = 1$), by the formula of total probability we reach

$$\Pr(\text{the game ends with a 5}) = \sum_{j=1}^{\infty} \Pr(X_j = 5|F_j)\Pr(F_j) = \left(\frac{2}{5}\right)\sum_{j=1}^{\infty}\Pr(F_j) = \frac{2}{5}.$$

c) By definition of τ and F_j we have

$$\mathrm{E}(\tau) = \sum_{j=1}^{\infty} j\Pr(F_j).$$

Since $\Pr(X_i = 5 \text{ or } X_i = 7) = 5/18$,

$$\Pr(F_j) = \left(1 - \frac{5}{18}\right)^{j-1}\left(\frac{5}{18}\right).$$

We are therefore trying to calculate the expectation of geometric distribution of parameter 5/18, i.e.,

$$\mathrm{E}(\tau) = \frac{18}{5} = 3.6.$$

Solution 41 a) Number the white balls $i = 1, \ldots, N$. Let $X_i = 1$ if the white ball numbered i has been chosen, and otherwise let $X_i = 0$. Then

$$\Pr(X_i = 0) = \frac{M+N-1}{M+N} \times \frac{M+N-2}{M+N-1} \times \ldots \times \frac{M+N-n}{M+N-n+1} = \frac{M+N-n}{M+N}, \quad \Pr(X_i = 1) = \frac{n}{M+N}.$$

Since $X = \sum_{i=1}^{N} X_i$, we deduce that

$$\mathrm{E}(X) = \sum_{i=1}^{N} \mathrm{E}(X_i) = N\Pr(X_i = 1) = \frac{Nn}{M+N}.$$

b) Number the black balls $j = 1, \ldots, M$ and let $Y_j = 1$ if the black ball numbered j is chosen before the first white ball and $Y_j = 0$ otherwise. Then consider any possible sequence (of length $N + M$) in which all the balls might be drawn, such as

$$B \cdots BW_2B \cdots BB_jB \cdots BW_NW_1B \cdots B,$$

where B denotes any black ball except B_j and $W_1, \dots, W_N$ denote the N white balls. Suppose we condition on the configuration, meaning the length of the sequence and the positioning of all the Bs except B_j. Then the only thing that can vary is the positioning of the balls $B_j, W_1, \dots, W_N$. Clearly the probability that B_j appears first in these $N+1$ positions is $1/(N+1)$, by symmetry, and since this does not depend on the configuration, then it must be true that $\Pr(Y_j = 1) = 1/(N+1)$. (If you are nervous about this, note that

$$\Pr(Y_j = 1) = \sum_{C \in \text{all possible configurations}} \Pr(Y_j = 1 \mid C)\Pr(C) = \sum_{C \in \text{all possible configurations}} \frac{1}{N+1}\Pr(C) = \frac{1}{N+1},$$

as required.) Then, since $X = 1 + \sum_{j=1}^{M} Y_j$, we have

$$\mathrm{E}(X) = 1 + \sum_{j=1}^{M} \mathrm{E}(Y_j) = 1 + \sum_{j=1}^{M} \Pr(Y_j = 1) = 1 + \frac{M}{N+1}.$$

For an alternative solution, suppose for greater generality that we remove the balls one by one until w whites have appeared, and let X be the total number of balls then drawn. Then

$$\Pr(X = r) = \frac{\binom{N}{w-1}\binom{M}{r-1-(w-1)}}{\binom{N+M}{r-1}} \times \frac{N-w+1}{N+M-(r-1)}, \quad r = w, \dots, w+M,$$

where to avoid trivial cases we assume that $w \in \{1, \dots, N\}$. The argument for this expression is that just before we take the wth white ball out on the rth trial, we have drawn $w-1$ white balls from among N and $(r-1)-(w-1)$ black balls from M in $r-1$ trials (corresponding to the hypergeometric probability given), and then we must choose a white ball with probability $(N-w+1)/(N+M-r+1)$. It is then possible to check algebraically that

$$\Pr(X = r) = \frac{N}{N+M} \frac{\binom{N-1}{w-1}\binom{M}{r-w}}{\binom{N+M-1}{r-1}} = \frac{w}{r} \frac{\binom{N}{w}\binom{M}{r-w}}{\binom{N+M}{r}}, \quad r = w, \dots, w+M,$$

or to argue by symmetry. For example, to obtain the first expression, note that the probability of first getting a white and then any configuration of $w-1$ whites among the remaining $N-1$ and $r-w$ blacks among the remaining M is as given, but by symmetry this probability must be the same as that of getting the sequence in the opposite order, which is the sequence we want. Note also that the sum of any of the three expressions for the probabilities must equal unity. This is called the negative hypergeometric distribution, by analogy with the negative binomial distribution.

Now

$$\mathrm{E}(X) = \sum_{r=w}^{w+M} r\Pr(X+r) = \sum_{r=w}^{w+M} r \frac{w}{r} \frac{\binom{N}{w}\binom{M}{r-w}}{\binom{N+M}{r}} = \frac{w(M+N+1)}{N+1} \sum_{r=w}^{w+M} \frac{N+1}{M+N+1} \frac{\binom{N}{w}\binom{M}{r-w}}{\binom{N+M}{r}}.$$

By setting $r = w + s$ we can write the sum on the right as

$$\sum_{s=0}^{M} \frac{N+1}{M+N+1} \frac{\binom{N}{w}\binom{M}{s}}{\binom{N+M}{w+s}} = \sum_{s=0}^{M} \frac{N'}{M+N'} \frac{\binom{N'-1}{w'-1}\binom{M}{s}}{\binom{N'+M-1}{w'+s-1}} = \sum_{r'=w'}^{w'+M} \frac{N'}{M+N'} \frac{\binom{N'-1}{w'-1}\binom{M}{r'-w'}}{\binom{N'+M-1}{r'-1}} = 1,$$

where we put $N' = N+1$, $w' = w+1$ and $r' = w' + s$, and note that the sum contains the probabilities $\Pr(X = r)$ for the same problem, stopping after having selected w' white balls from a bag with N' white balls and M black ones. Hence

$$\mathrm{E}(X) = \frac{w(M+N+1)}{N+1} = w + \frac{wM}{N+1},$$

which gives the result above when $w = 1$.

Yet another solution is as follows. Suppose all $N+M$ balls are taken successively at random, and let Z_0 be the number of black balls before the first white, Z_1 the number of black balls between the first and second whites, and so on, with Z_N the number of black balls after the last white one. Obviously $Z_0 + \cdots + Z_N = M$, and, almost as obviously, $\mathrm{E}(Z_0) = \cdots = \mathrm{E}(Z_N)$ by symmetry. Hence $\mathrm{E}(Z_i) = M/(N+1)$, and therefore the expected number of balls drawn (including whites) up to and including the wth white one is $w + w\mathrm{E}(Z_i) = w + wM/(N+1)$, as above.

Solution 42 a) X has the binomial distribution with parameters n and p, so $\Pr\{X = k\} = \binom{n}{k} p^k (1 - p)^{n-k}$, for $k \in \{0, \ldots, n\}$.

b) See the lecture notes: $\mathrm{E}(X) = np$ and $\mathrm{var}(X) = np(1-p)$.

Note: To show that X has as expectation np and variance $np(1-p)$, we can write $X = \sum_{i=1}^{n} Y_i$ where Y_i are independent Bernouilli random variables with parameter p. Then

$$\mathrm{E}(Y_i) = 0 \times \Pr(Y_i = 0) + 1 \times \Pr(Y_i = 1) = 0 \times (1-p) + 1 \times p = p,$$

and

$$\mathrm{var}(Y_i) = \mathrm{E}(Y_i^2) - E(Y_i)^2 = \{0^2 \times \Pr(Y_i = 0) + 1^2 \times \Pr(Y_i = 1)\} - p^2 = p - p^2 = p(1-p),$$

so by the independence of the Y_i, we have

$$\begin{aligned} \mathrm{E}(X) &= \mathrm{E}\left(\sum_{i=1}^{n} Y_i\right) = \sum_{i=1}^{n} \mathrm{E}(Y_i) = \sum_{i=1}^{n} p = np, \\ \mathrm{var}(X) &= \mathrm{var}\left(\sum_{i=1}^{n} Y_i\right) = \sum_{i=1}^{n} \mathrm{var}(Y_i) = \sum_{i=1}^{n} \{p(1-p)\} = np(1-p). \end{aligned}$$

NB: Here $\mathrm{var}(\sum_{i=1}^{n} Y_i) = \sum_{i=1}^{n} \mathrm{var}(Y_i)$ because the Y_i are independent, but this is not true in general. It is however always true that $\mathrm{E}(\sum_{i=1}^{n} Y_i) = \sum_{i=1}^{n} \mathrm{E}(Y_i)$.

Solution 43 T is a geometric random variable, and

$$\Pr(T = n) = p(1-p)^{n-1} \quad n \geq 1.$$

We have

$$\mathrm{E}(T) = \sum_{n=1}^{\infty} n \Pr(T = n) = \sum_{n=1}^{\infty} n\, p(1-p)^{n-1} = p \sum_{n=1}^{\infty} n(1-p)^{n-1} = -p \tfrac{d}{dp} \sum_{n=1}^{\infty} (1-p)^n = -p \tfrac{d}{dp} \tfrac{1}{p} = p \times \tfrac{1}{p^2} = \tfrac{1}{p}.$$

Furthermore,

$$\begin{aligned} \mathrm{E}(T^2) &= \sum_{n=1}^{\infty} n^2 \Pr(T = n) = \sum_{n=1}^{\infty} n^2\, p(1-p)^{n-1} = p \sum_{n=1}^{\infty} n^2 (1-p)^{n-1} = -p \tfrac{d}{dp} \left\{ \sum_{n=1}^{\infty} n(1-p)^n \right\} \\ &= -p \tfrac{d}{dp} \left\{ (1-p) \sum_{n=1}^{\infty} n(1-p)^{n-1} \right\} = -p \tfrac{d}{dp} \left\{ (1-p) \tfrac{1}{p^2} \right\} = -p \left(-\tfrac{2}{p^3} + \tfrac{1}{p^2} \right) = \tfrac{2}{p^2} - \tfrac{1}{p}. \end{aligned}$$

T's variance is therefore:

$$\mathrm{var}(T) = \mathrm{E}(T^2) - \mathrm{E}(T)^2 = \tfrac{2}{p^2} - \tfrac{1}{p} - \tfrac{1}{p^2} = \tfrac{1}{p^2} - \tfrac{1}{p}, \;\; p \in (0, 1].$$

Solution 44 a) Since

$$\sum_{i=0}^{\infty} \Pr\{X = i\} = 1 = c \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = c e^{\lambda},$$

we must have $c = e^{-\lambda}$. This is a Poisson distribution of parameter λ.

b) $\Pr\{X = 0\} = c \frac{\lambda^o}{0!} = e^{-\lambda}$.

c) $\Pr\{X > 2\} = 1 - (\Pr(X = 0) + \Pr(X = 1) + \Pr(X = 2)) = 1 - e^{-\lambda}(1 + \lambda + \frac{\lambda^2}{2}) = e^{-\lambda} \sum_{i=3}^{\infty} \frac{\lambda^i}{i!}$.

d) See lecture notes. $\mathrm{E}(X) = \mathrm{var}(X) = \lambda$.

Solution 45 Consider the first page. A given error will appear on this page with a probability of 1/350, since the errors are uniformly distributed (i.e., distributed at random) and there are 350 pages in total. The number of printing errors X on the first page therefore has a Binomial distribution $B(n = 450, p = 1/350)$, so

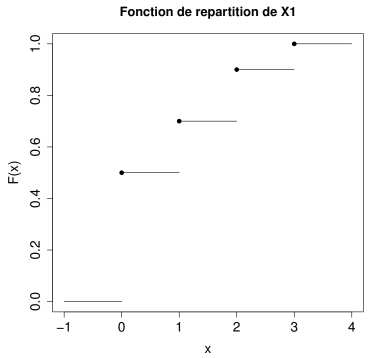
$$\Pr(X \geq 3) = 1 - \Pr(X \leq 2) = 1 - \sum_{i=0}^{2} C_{450}^{i} \left(\frac{1}{350}\right)^i \left(\frac{349}{350}\right)^{450-i} \simeq 0.14.$$

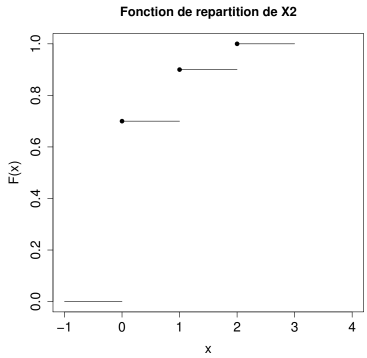
Since n is large and p is small, we can also approximate the binomial variable X by a Poisson variable Y of parameter $\lambda = np \simeq 1.29$ and we then obtain

$$\Pr(Y \geq 3) = 1 - \Pr(Y \leq 2) = 1 - e^{-\lambda} \left(1 + \lambda + \lambda^2/2\right) \simeq 0.14.$$

Solution 46 a) We check that the sum of the lines is 1.

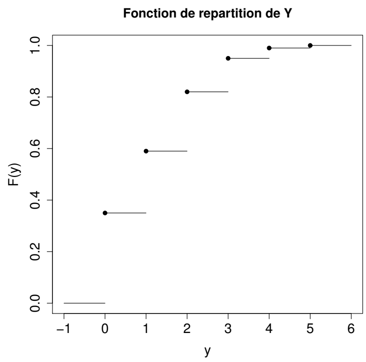
b) Addition gives $F_{X_1}(0) = 0.5$, $F_{X_1}(1) = 0.7$, $F_{X_1}(2) = 0.9$, $F_{X_1}(3) = 1$, and $F_{X_2}(0) = 0.7$, $F_{X_2}(1) = 0.9$, $F_{X_2}(2) = 1$.





c) We find that $\Pr(Y = 0) = \Pr(X_1 = 0)\Pr(X_2 = 0) = 0.35$, $\Pr(Y = 1) = \Pr(X_1 = 0)\Pr(X_2 = 1) + \Pr(X_1 = 1)\Pr(X_2 = 0) = 0.24$, and by similar calculations, $\Pr(Y = 2) = 0.23$, $\Pr(Y = 3) = 0.13$, $\Pr(Y = 4) = 0.04$, and $\Pr(Y = 5) = 0.01$. As a check, the total mass is $0.35+0.24+0.23+0.13+0.04+0.01 = 1$.

d) In c) we found that $F_Y(0) = 0.35$, $F_Y(1) = 0.59$, $F_Y(2) = 0.82$, $F_Y(3) = 0.95$, $F_Y(4) = 0.99$, $F_Y(5) = 1$, so the sketch is



Chapter 4

Solution 47 First calculate c. For f to be a density we must have $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1$, that is to say

$$c = \frac{1}{\int_0^\infty x\exp(-x/2)\mathrm{d}x}.$$

Integration by parts gives:

$$\int_0^\infty x\exp(-x/2)\mathrm{d}x = -2x\exp(-x/2)\big|_0^\infty + 2\int_0^\infty \exp(-x/2)\mathrm{d}x = 4,$$

so c=1/4. Consequently the probability that the system functions for at least 5 months is

$$\int_5^\infty \frac{x}{4}\exp(-x/2)\mathrm{d}x = -\exp(-x/2)\big|_5^\infty - \frac{1}{2}x\exp(-x/2)\big|_5^\infty = e^{-5/2} + \frac{5}{2}e^{-5/2} \simeq 0.287.$$

Solution 48 a) Integration by parts yields

$$\Gamma(\alpha) = \int_0^\infty e^{-y}y^{\alpha-1}dx = -e^{-y}y^{\alpha-1}\mid_0^\infty + (\alpha-1)\Gamma(\alpha-1) = (\alpha-1)\Gamma(\alpha-1), \quad (\alpha > 1).$$

Since $\Gamma(1) = 1$, we deduce that, for n a positive integer, $\Gamma(n) = (n-1)!$.

b) We have

$$\mathrm{E}(X) = \tfrac{\lambda^\alpha}{\Gamma(\alpha)}\int_0^\infty x^\alpha e^{-\lambda x}dx = \tfrac{\lambda^\alpha}{\Gamma(\alpha)}\int_0^\infty \tfrac{y^\alpha}{\lambda^\alpha}e^{-y}\tfrac{1}{\lambda}dy = \tfrac{\Gamma(\alpha+1)}{\lambda\Gamma(\alpha)} = \tfrac{\alpha}{\lambda}.$$

Solution 49 Since

$$E(|X|) = \int_{-\infty}^{\infty} \frac{|x|}{\pi(1+x^2)}\,dx = 2\int_0^{\infty} \frac{x}{\pi(1+x^2)} dx = \frac{1}{\pi}\left[\ln\left(1+x^2\right)\right]_0^{\infty} = \infty,$$

E(X) is undefined.

Solution 50 a) We have $f(t) = c(10000t^2 - 200t^3 + t^4)$ in the set $0 < t < 100$. By integrating we get $10^{10} \times c \times \frac{1}{30} = 1$ or $c = 3 \times 10^{-9}$.

b) The expected lifetime in years is

$$\mathrm{E}(T) = \int_0^{100} tf(t)dt = c[10000t^4/4 - 40t^5 + t^6/6]_0^{100} = 3 \times 10^{-9} \times 10^{12} \times (1/4 - 2/5 + 1/6) = 50.$$

c) We have

$$\Pr(50 \le T \le 80) = \int_{50}^{80} f(t)dt = c[10000t^3/3 - 50t^4 + t^5/5]_{50}^{80} \simeq 0.4421.$$

Solution 51 a) f being a density, we must have $\int_{-\infty}^{\infty} f(x)\,dx = 1$. But

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_1^{\infty} f(x)\,dx = \int_1^{\infty} \frac{k}{x^4}\,dx = \left[-\frac{k}{3x^3}\right]_1^{\infty} = \frac{k}{3},$$

so $k = 3$.

b) To find the distribution function, we have to integrate the density function $f(x)$. We find

$$F(x) = \int_{-\infty}^{x} f(x)\,dx = \begin{cases} 1 - 1/x^3, & x \ge 1, \\ 0, & x < 1. \end{cases}$$

c) The required probability is $\Pr(X > 3) = 1 - F(3) = 0.0370$.

d) The average lifetime is equal to the expectation of X, that is

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} xf(x)\,dx = \int_1^{\infty} \frac{3x}{x^4}\,dx = \int_1^{\infty} \frac{3}{x^3}\,dx = \left[-\frac{3}{2x^2}\right]_1^{\infty} = 3/2,$$

or a year and a half.

e) First calculate $\mathrm{E}(X^2)$:

$$\mathrm{E}(X^2) = \int_1^{\infty} \frac{3x^2}{x^4}\,dx = \int_1^{\infty} \frac{3}{x^2}\,dx = \left[-\frac{3}{x}\right]_1^{\infty} = 3,$$

therefore $\mathrm{var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = 0.75$; the standard deviation is $\sqrt{\mathrm{var}(X)} = 0.87$ years.

Solution 52 (a) The distribution function of U is $F(u) = u/2$, for $0 < u < 2$, so we seek

$$\Pr(U \le u \mid U < 1) = \frac{\Pr(U \le u, U \le 1)}{\Pr(U \le 1)} = \frac{u/2}{1/2} = u, \quad 0 < u < 1.$$

Thus the required distribution is $U(0,1)$.

(b) Write $\mathcal{A} = \{u : |u-1| \ge 1/2\} = \{u : u < 1/2\} \cup \{u : u > 3/2\}$ and note that the density of U is $f(u) = \frac{1}{2}I(0 < u < 2)$. Hence the density function of U conditional on $U \in \mathcal{A}$ is

$$f(u \mid U \in \mathcal{A}) = \frac{f(u)}{\Pr(U \in \mathcal{A})}, \quad u \in \mathcal{A},$$

and zero elsewhere. Clearly $\Pr(U \in \mathcal{A}) = 1/2$, so

$$f(u \mid U \in \mathcal{A}) = \begin{cases} 1, & u \in \mathcal{A} \cap (0,2), \\ 0, & \text{elsewhere.} \end{cases}$$

This density is uniform on $\mathcal{A} \cap (0,2) = (0,1/2) \cup (3/2,2)$.

Solution 53 Note that since X takes positive real values, Y takes values in $\mathbb{R}$. Since the logarithm function is strictly increasing, we can write

$$F_Y(y) = \Pr(Y \le y) = \Pr(\log X \le y) = \Pr\{X \le \exp(y)\} = \int_0^{\exp(y)} \exp(-x)\mathrm{d}x, \quad y \in \mathbb{R}.$$

Differentiation of this expression with respect to y yields

$$f_Y(y) = \frac{dF_Y(y)}{dy} = e^y \times \exp(-e^y), \quad y \in \mathbb{R}.$$

Solution 54 Let $Y = g(X)$ with $g(x) = 1/x$. Since g is monotonic with inverse $g^{-1}(y) = 1/y$ on $(0, 1]$, then

$$f_Y(y) = f_X\{g^{-1}(y)\} \times \left|\tfrac{\mathrm{d}g^{-1}(y)}{\mathrm{d}y}\right| = (1/y)^{-2} \times \tfrac{1}{y^2} = 1, \quad 0 < y \le 1.$$

Alternatively,

$$F_X(x) = \int_1^x t^{-2}\,\mathrm{d}t = \left[-\tfrac{1}{t}\right]_1^x = 1 - x^{-1}, \quad x \ge 1.$$

So, for $0 < y \le 1$ we have

$$\begin{aligned} F_Y(y) &= \Pr(Y \le y) = \Pr\left(\tfrac{1}{X} \le y\right) = \Pr\left(\tfrac{1}{y} \le X\right) = 1 - \Pr\left(X < \tfrac{1}{y}\right) \\ &= 1 - F_X(1/y) = 1 - (1 - y) = y, \end{aligned}$$

which implies that $Y \sim U(0, 1)$; Y has the standard uniform distribution.

Solution 55 If $Y = g(X)$ and g is a monotone function, then

$$f_Y(y) = \left|\frac{dg^{-1}(y)}{dy}\right| f_X\{g^{-1}(y)\},$$

defined for those values of y for which $g^{-1}(y)$ lies in the support of X.

In this case the function $y = g(x) = e^x$ is monotone and $g^{-1}(y) = \log y$; therefore $dg^{-1}(y)/dy = 1/y$, defined only for $y > 0$. Therefore since $f_X(x) = (2\pi)^{-1/2}\exp(-x^2/2)$, for $x \in \mathbb{R}$, we have

$$f_Y(y) = \frac{1}{y}\frac{1}{\sqrt{2\pi}}e^{-(\log y)^2/2}, \quad y > 0,$$

and $f_Y(y) = 0$ if $y \le 0$.

Solution 56 Let $X \sim N(175, 6^2)$ denote height of a 25-year-old man. Then

$$\Pr(X > 185) = 1 - \Phi\{(185 - 175)/6\} \simeq 4.78\,\%$$

and two similar calculations give

$$\Pr(X > 192 \mid X > 180) = \tfrac{\Pr(X>192)}{\Pr(X>180)} \simeq 1.14\,\%.$$

Solution 57 The quantity a must satisfy $\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1$, so

$$a = \frac{1}{\int_0^\infty x^2\exp(-bx^2)\mathrm{d}x}.$$

If we make the change of variables $y/\sqrt{2b} = x$, we obtain

$$\int_0^\infty x^2\exp(-bx^2)\mathrm{d}x = \left(\frac{1}{\sqrt{2b}}\right)^3 \int_0^\infty y^2\exp(-y^2/2)\mathrm{d}y = \left(\frac{1}{\sqrt{2b}}\right)^3 \frac{\sqrt{2\pi}}{2} \times \int_{-\infty}^{+\infty} \frac{y^2}{\sqrt{2\pi}}\exp(-y^2/2)\mathrm{d}y.$$

The last integral is equal to 1 (the second moment of a standard normal variable) and therefore

$$a = \frac{4}{\sqrt{\pi}}b^{3/2}.$$

Chapter 5

Solution 58 For all real z we have

$$\begin{aligned} F_Z(z) &= F_X(z)F_Y(z), \\ F_{\tilde{Z}}(z) &= 1-\{1-F_X(z)\}\{1-F_Y(z)\} = F_X(z)+F_Y(z)-F_X(z)F_Y(z), \end{aligned}$$

differentiation of which yields

$$\begin{aligned} f_Z(z) &= f_X(z)F_Y(z)+F_X(z)f_Y(z), \\ f_{\tilde{Z}}(z) &= f_X(z)\{1-F_Y(z)\}+f_Y(z)\{1-F_X(z)\}. \end{aligned}$$

Solution 59 Write $W=\min(X_1,\ldots,X_n)$ and note that

$$\Pr(W\ge t)=\Pr(X_1\ge t,\ldots,X_n\ge t)=\Pr(X_1\ge t)^n$$

But $\Pr(X_1\ge t)=e^{-t\lambda}$, so $\Pr(W\ge t)=e^{-tn\lambda}$, and therefore the distribution function of W is

$$\Pr(W\le t)=\begin{cases}1-e^{-tn\lambda}, & t>0,\\ 0, & \text{otherwise.}\end{cases}$$

Thus $W\sim\exp(n\lambda)$.

Solution 60 We first compute the mass function of X. For non-negative integer k,

$$\begin{aligned} \Pr\{X=k\} &= \sum_{n\ge k}\Pr\{X=k\mid Z=n\}\Pr\{Z=n\} \\ &= \sum_{n\ge k}\binom{n}{k}p^k(1-p)^{n-k}e^{-\lambda}\tfrac{\lambda^n}{n!} \\ &= e^{-\lambda}\tfrac{(p\lambda)^k}{k!}\sum_{n\ge k}\binom{n}{k}(1-p)^{n-k}\tfrac{\lambda^{n-k}}{n!}\tfrac{k!(n-k)!}{(n-k)!} \\ &= e^{-\lambda}\tfrac{(p\lambda)^k}{k!}\sum_{n\ge k}\tfrac{(1-p)^{n-k}\lambda^{n-k}}{(n-k)!} \\ &= e^{-\lambda}\tfrac{(p\lambda)^k}{k!}e^{(1-p)\lambda} \\ &= e^{-p\lambda}\tfrac{(p\lambda)^k}{k!}\quad, \end{aligned}$$

so X is Poisson with parameter $p\lambda$.

Then, for non-negative integer l,

$$\begin{aligned} \Pr\{Y=l\} &= \sum_{n\ge l}\Pr\{Y=l\mid Z=n\}\Pr\{Z=n\} \\ &= \sum_{n\ge l}\Pr\{X=n-l\mid Z=n\}\Pr\{Z=n\} \\ &= \sum_{n\ge l}\binom{n}{l}p^{n-l}(1-p)^l e^{-\lambda}\tfrac{\lambda^n}{n!} \\ &= e^{-(1-p)\lambda}\tfrac{((1-p)\lambda)^l}{l!}\quad, \end{aligned}$$

so Y is also a Poisson variable, of parameter $(1-p)\lambda$. Finally, for all $k,l\in\{0,1,2,\ldots\}$, since $n-l=k$,

$$\begin{aligned} \Pr\{X=k,Y=l\} &= \sum_{n\ge 0}\Pr\{X=k,Y=l\mid Z=n\}\Pr\{Z=n\} \\ &= \Pr\{X=k,Y=l\mid Z=k+l\}\Pr\{Z=k+l\} \\ &= \binom{k+l}{k}p^k(1-p)^l e^{-\lambda}\tfrac{\lambda^{k+l}}{(k+l)!} \\ &= e^{-p\lambda}\tfrac{(p\lambda)^k}{k!}\times e^{-(1-p)\lambda}\tfrac{((1-p)\lambda)^l}{l!}, \end{aligned}$$

which is $\Pr\{X=k\}\Pr\{Y=l\}$, so X and Y are indeed independent.

Solution 61 a) Since

$$\int_0^2 \int_0^1 f_{X,Y}(x,y)\,dx\,dy = 1,$$

we have $c \times 7/6 = 1$, thus $c = 6/7$.

b) The marginal density of X is

$$f_X(x) = \int_0^2 f_{X,Y}(x,y)\,dy = \frac{6}{7}\int_0^1 \left(x^2 + xy/2\right)\,dy = \frac{6}{7}\left[x^2y + xy^2/4\right]_0^2 = \frac{6}{7}(2x^2 + x), \quad 0 \le x \le 1.$$

The marginal density of Y is

$$f_Y(y) = \int_0^1 f_{X,Y}(x,y)\,dx = \frac{6}{7}\int_0^1 (x^2 + xy/2)\,dx = \frac{6}{7}\left[x^3/3 + x^2y/4\right]_0^1 = \frac{6}{7}(1/3 + y/4), \quad 0 \le y \le 2.$$

The variables X and Y are dependent, because the joint density is not the product of the marginal densities.

c) We have

$$\Pr(X > Y) = \frac{6}{7}\int_0^1 \left(\int_0^x x^2 + xy/2dy\right)dx = \frac{6}{7}\int_0^1 (x^3 + x^3/4)dx = \frac{6}{7} \times \frac{5}{4} \times \frac{1}{4} = \frac{30}{112} \simeq 0.2679.$$

Solution 62 a) Let f_X denote the density of X and f_Y that of Y. Note that

$$f_Y(y) = \int_0^\infty f(x,y)dx = \left.\frac{-xe^{-x(1+y)}}{1+y}\right|_0^\infty + \left.\frac{-e^{-x(1+y)}}{(1+y)^2}\right|_0^\infty = \frac{1}{(1+y)^2}, \quad y > 0,$$

so

$$F_Y(y) = \begin{cases} y/(1+y), & y > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Also,

$$f_X(x) = \int_0^\infty f(x,y)dy = e^{-x}, \quad x > 0$$

so $X \sim \exp(1)$. X and Y are not independent because $f(x,y) \neq f_X(x)f_Y(y)$ for all x, y.

b) Let f_X denote the density of X and f_Y the density of Y. Note that

$$f_X(x) = \int_0^{1-x} f(x,y)dy = 60x\int_0^{1-x} y^2dy = 20x(1-x)^3, \quad 0 < x < 1,$$

and that

$$f_Y(y) = \int_0^{1-y} f(x,y)dx = 60y^2\int_0^{1-y} xdx = 30y^2(1-y)^2, \quad 0 < y < 1.$$

X and Y are not independent because $f(x,y) \neq f_X(x)f_Y(y)$ for all x, y. In this case the lack of independence is obvious without any need for calculation because the support of the joint density is not a Cartesian product.

Solution 63 The first part is obvious, since if $X_1 > X_2$ then the right-hand side equals $X_1 + X_2$ and similarly if $X_2 \ge X_1$.

For the second part, note that $\min(X_1, X_2) > x$ if and only if $X_1, X_2 > x$, and since they are independent this has probability $e^{-\lambda_1 x} \times e^{-\lambda_2 x}$ if $x > 0$. Therefore

$$\Pr\{\min(X_1, x_2) \le x\} = 1 - \exp\{-(\lambda_1 + \lambda_2)x\}, \quad x > 0,$$

and thus the minimum has the $\exp(\lambda_1 + \lambda_2)$ distribution. By the first part,

$$\mathrm{E}(X_1 + X_2) = \mathrm{E}\{\max(X_1, X_2) + \min(X_1, X_2)\},$$

and the linearity of expectation implies that $\mathrm{E}\{\max(X_1, X_2)\} = \mathrm{E}(X_1) + \mathrm{E}(X_2) - \mathrm{E}\{\min(X_1, X_2)\}$, which gives the result.

Solution 64 a) The conditional density of X given Y is defined as $f_{X|Y}(x \mid y) = f_{X,Y}(x, y)/f_Y(y)$, and

$$f_{X|Y}(x \mid y) = \frac{1}{\sqrt{2\pi}\sqrt{1/(2y)}} \exp\left\{x^2/(2 \times 1/2y)\right\}, \quad f_Y(y) = (\pi y)^{-1/2} e^{-y}, \quad y > 0, x \in \mathbb{R},$$

so

$$f_{X,Y}(x, y) = f_Y(y) f_{X|Y}(x \mid y) = \begin{cases} \frac{1}{\pi} e^{-y(1+x^2)}, & y > 0, x \in \mathbb{R}, \\ 0, & y \leq 0, x \in \mathbb{R}. \end{cases}$$

b) We have

$$f_X(x) = \int_0^\infty f_{X,Y}(x, y) dy = \frac{1}{\pi(1 + x^2)}, \quad x \in \mathbb{R},$$

which is the standard Cauchy density.

c) Now $f_{Y|X}(x, y) = f_{X,Y}(x, y)/f_X(x)$, so

$$f_{Y|X}(x, y) = \begin{cases} (1 + x^2) e^{-y(1+x^2)}, & y > 0, \\ 0, & y \leq 0. \end{cases}$$

So, conditional on $X = x$, Y follows an exponential distribution of parameter $1 + x^2 > 0$.

Solution 65 a) We have

$$f(x) = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi} dy = \frac{2}{\pi}\sqrt{1 - x^2}, \quad -1 \leq x \leq 1,$$

and by symmetry we must have

$$g(y) = \frac{2}{\pi}\sqrt{1 - y^2}, \quad -1 \leq y \leq 1.$$

Clearly X and Y are not independent because $h(x, y) \neq f(x)g(y)$ for all x, y; it is quicker to note that since the support of the joint density is the unit disk, and this is not a Cartesian product, they cannot be independent.

b) Now $\text{cov}(X, Y) = \text{E}(XY) - \text{E}(X)\text{E}(Y)$, and evidently $\text{E}(X) = \text{E}(Y) = 0$. Now $\text{E}(XY)$ consists of integrals in the four quadrants, corresponding to $x, y > 0$, $x < 0 < y$, $x, y < 0$, $y < 0 < x$ (the integrals on the axes themselves equal zero). Let D denote the unit disk, and let $Q_1, \ldots, Q_4$ denote the quadrants, starting with the upper right and proceeding anticlockwise. Then we need to compute

$$\text{E}(XY) \propto \int_{D \cap Q_1} xy \, dxdy + \int_{D \cap Q_2} xy \, dxdy + \int_{D \cap Q_3} xy \, dxdy + \int_{D \cap Q_4} xy \, dxdy,$$

and this clearly equals zero, since

$$\int_{D \cap Q_2} xy \, dxdy = \int_{D \cap Q_1} (-x)y \, dxdy = -\int_{D \cap Q_1} xy \, dxdy, \quad \int_{D \cap Q_4} xy \, dxdy = \int_{D \cap Q_3} (-x)y \, dxdy = -\int_{D \cap Q_3} xy \, dxdy.$$

Solution 66 $\text{var}(X_1 - X_2) = (+1)^2\text{var}(X_1) + (-1)^2\text{var}(X_2) - 2\text{cov}(X_1, x_2) = 2n\pi(1 - \pi)$, because the covariance is zero (by independence) and the variances both equal $n\pi(1 - \pi)$. Since independence implies that $\text{cov}(X_i, X_k) = 0$ if $i \neq k$,

$$\text{var}(S) = \sum_{i=1}^n \text{var}(X_i) + \sum_{i \neq k} \text{cov}(X_i, X_k) = \sum_{i=1}^n \text{var}(X_i) = n^2\pi(1 - \pi).$$

Solution 67 The variance may be decomposed as follows:

$$\text{var}(3X - 2Y + 1) = 9\text{var}(X) - 12\text{cov}(X, Y) + 4\text{var}(Y).$$

We have $\text{cov}(X, Y) = \text{corr}(X, Y)\sqrt{\text{var}(X)\text{var}(Y)} = 2$. Hence $\text{var}(3X - 2Y + 1) = 48$. By bilinearity of the covariance,

$$\text{cov}(X + 2Y, X - Y) = \text{var}(X) + \text{cov}(X, Y) - 2\text{var}(Y) = -12.$$

By linearity of expectation, $\text{E}(3X - 2Y + Z) = 0 \iff \text{E}(Z) = 2\text{E}(Y) - 3\text{E}(X)$, giving $\text{E}(Z) = 1$.

Solution 68 a) By linearity of expectation, $\mathrm{E}(5X - 3Y + 9) = 5\mathrm{E}(X) - 3\mathrm{E}(Y) + 9 = 7$. By the independence of X and Y and properties of the covariance,

$$\mathrm{var}(3Y - 2X) = 3^2\mathrm{var}(Y) + (-2)^2\mathrm{var}(X) + 2 \times 3 \times (-2)\mathrm{cov}(X, Y) = 57.$$

b) We have

$$\mathrm{var}(X + Y) = \mathrm{cov}(X + Y, X + Y)) = \mathrm{var}(X) + \mathrm{var}(Y) + 2\mathrm{cov}(X, Y),$$

and likewise

$$\mathrm{var}(X - Y) = \mathrm{cov}(X - Y, X - Y) = \mathrm{var}(X) + \mathrm{var}(Y) - 2\mathrm{cov}(X, Y).$$

From these we get the system of equations:

$$\begin{cases} \mathrm{cov}(X, Y) = 48 - \mathrm{var}(X) \\ \mathrm{cov}(X, Y) = -32 + \mathrm{var}(X), \end{cases}$$

from which we can deduce $\mathrm{var}(X) = \mathrm{var}(Y) = 40$ and $\mathrm{cov}(X, Y) = 8$. The correlation is $\frac{8}{\sqrt{40\times 40}} = 0.2$.

Solution 69 a) X_1 and X_2 are binomial variables of parameters $(n, 1/6)$. A binomial variable may be written as a sum of independent Bernoulli variables, each with the same success probability.

b) We have $\mathrm{var}(X_1) = \mathrm{var}(X_2) = n \times \frac{1}{6} \times (1 - \frac{1}{6}) = n \times \frac{5}{36}$.

c) The variable U represents the total number of 1s and 2s obtained during n throws. Its distribution is $\mathrm{Bin}(n, 1/3)$, so its variance is $n \times \frac{2}{9}$. Since $\mathrm{var}(X_1 + x_2) = \mathrm{var}(X_1) + 2\mathrm{cov}(X_1, X_2) + \mathrm{var}(X_2)$, it follows that $\mathrm{cov}(X_1, X_2) = -\frac{1}{2}[\mathrm{var}(X_1) + \mathrm{var}(X_2) - \mathrm{var}(X_1 + X_2)] = -n/36$, so $\mathrm{corr}(X_1, X_2) = -\frac{n}{36}/\frac{5n}{36} = -1/5$.

d) We have $\mathrm{var}(V) = \mathrm{var}(X_1) + \mathrm{var}(X_2) - 2\mathrm{cov}(X_1, X_2) = \frac{12n}{36}$, and $\mathrm{corr}(U, V) = \mathrm{cov}(U, V)/\sqrt{\mathrm{var}(U)\mathrm{var}(V)}$, where $\mathrm{cov}(U, V) = \mathrm{var}(X_1) - \mathrm{var}(X_2) = 0$. Thus, $\mathrm{corr}(U, V) = 0$.

Solution 70 By enumerating the possible events, we observe that $\Pr(Z = 1) = \Pr(Z = 7) = 1/12$ and that $\Pr(Z = 2) = \cdots = \Pr(Z = 6) = 1/6$. Moreover

$$\mathrm{E}(Z \mid X = x) = \mathrm{E}(X + Y \mid X = x) = x + \mathrm{E}(Y) = x + 3.5.$$

Likewise $\mathrm{var}(Z \mid X = x) = \mathrm{var}(X+Y \mid X = x) = \mathrm{var}(x+Y \mid X = x) = \mathrm{var}(Y) = 35/12$, since conditional on $X = x$, the only random variable in $x + Y$ is Y.

It is easy to check that Theorem 168 is satisfied.

Solution 71 If $z = 1$, then $\Pr(Z = z) = \Pr(X = 0)\Pr(Y = 1) = (1 - p)p$. For $z \geq 2$,

$$\begin{aligned} \Pr(Z = z) &= \Pr(X = 0)\Pr(Y = z) + \Pr(X = 1)\Pr(Y = z - 1) \\ &= (1 - p) \times p(1 - p)^{z-1} + p \times p(1 - p)^{z-2} \\ &= p(1 - p)^{z-2}\{p + (1 - p)^2\}. \end{aligned}$$

We have $\mathrm{E}(Z \mid X = x) = \mathrm{E}(X + Y \mid X = x) = \mathrm{E}(Y) + x$ and also $\mathrm{E}(Z \mid Y = y) = \mathrm{E}(X) + y$.

Solution 72 The moment generating function is $M_X(t) = \mathrm{E}(e^{tX})$. Thus

$$\begin{aligned} \mathrm{E}(Y) &= \mathrm{E}(e^X) = M_X(1) = \exp(1/2), \\ \mathrm{E}(Y^2) &= \mathrm{E}(e^{2X}) = M_X(2) = \exp(2), \\ \mathrm{var}(Y) &= \mathrm{E}(Y^2) - \mathrm{E}(Y)^2 = \exp(2) - \exp(1). \end{aligned}$$

Solution 73 a) For any positive t and by independence of X and Y we have

$$\begin{aligned} \Pr\{Z \leq t\} &= \iint_{\{x+y\leq z\}} f_{X,Y}(x, y)dxdy = \iint_{\{x+y\leq z\}} f_X(x)f_Y(y)dxdy \\ &= \lambda_1\lambda_2 \int_0^t e^{-\lambda_1 u} \left\{ \int_0^{t-u} e^{-\lambda_2 v}dv \right\} du \\ &= \lambda_1 \int_0^t e^{-\lambda_1 u}(1 - e^{-\lambda_2(t-u)})du \\ &= (1 - e^{-\lambda_1 t}) - \lambda_1 e^{-\lambda_2 t} \int_0^t e^{(\lambda_2-\lambda_1)u}du \\ &= \begin{cases} (1 - e^{-\lambda_1 t}) + \frac{\lambda_1}{\lambda_1-\lambda_2}(e^{-\lambda_1 t} - e^{-\lambda_2 t}), & \lambda_1 \neq \lambda_2, \\ 1 - (1 + \lambda t)e^{-\lambda t}, & \lambda_1 = \lambda_2 = \lambda. \end{cases} \end{aligned}$$

b) Recall that independence implies that for all real t,

$$M_Z(t) = \mathrm{E}[e^{tZ}] = \mathrm{E}[e^{tX}e^{tY}] = \mathrm{E}[e^{tX}]\mathrm{E}[e^{tY}].$$

In the case of an exponential variable,

$$M_X(t) = \int_0^{+\infty} \lambda_1 e^{-\lambda_1 u} e^{tu} du = \tfrac{\lambda_1}{\lambda_1 - t}, \quad t < \lambda_1,$$

so

$$M_Z(t) = \tfrac{\lambda_1}{\lambda_1 - t}\tfrac{\lambda_2}{\lambda_2 - t}, \quad t < \min(\lambda_1, \lambda_2).$$

c) We deduce from b) that if $\lambda_1 = \lambda_2$, then Z is a gamma variable of parameters $\alpha = 2$ and λ.

Solution 74 The mass function of X is

$$\Pr(X = n) = p(1-p)^{n-1}, \quad n = 1, 2, \ldots, 0 < p < 1,$$

and its MGF is

$$M_X(t) = \sum_{n=1}^{\infty} e^{tn} p(1-p)^{n-1} = pe^t \sum_{n=1}^{\infty} [e^t(1-p)]^{n-1} = \frac{pe^t}{1 - e^t(1-p)}, \quad |e^t(1-p)| < 1.$$

The moments are well defined for values of p for which $e^t(1-p)\Big|_{t=0} < 1$ or equivalently such that $p > 0$.

The first moment is $M_X'(0)$, i.e.,

$$\mathrm{E}(X) = \frac{pe^t}{1 - e^t(1-p)} + pe^t \frac{e^t(1-p)}{1 - e^t(1-p)}\Big|_{t=0} = M_X(t) + M_X(t)^2\Big|_{t=0} \tfrac{(1-p)}{p} = \tfrac{1}{p},$$

where we used the fact that $M_X(0) = 1$.

The second moment is given by the second derivative of $M_X(t)$ evaluated at $t = 0$ and is

$$\mathrm{E}(X^2) = \tfrac{d[M_X(t)']}{dt}\Big|_{t=0} = \tfrac{dM_X(t)}{dt}\Big|_{t=0} + \tfrac{dM_X(t)^2}{dt}\Big|_{t=0} \tfrac{1-p}{p} = \tfrac{1}{p} + \tfrac{1-p}{p} 2M_X(t)\tfrac{dM_X(t)}{dt}\Big|_{t=0} = \tfrac{2}{p^2} - \tfrac{1}{p}.$$

Solution 75 a) By direct calculation: for $k \in \{0, 1, \ldots\}$:

$$\begin{aligned}
\Pr\{Z = k\} &= \textstyle\sum_{l=0}^{k} \Pr\{X = l, Y = k - l\} \\
&= \textstyle\sum_{l=0}^{k} \Pr\{X = l\}\Pr\{Y = k - l\} \\
&= e^{-(\lambda_1 + \lambda_2)} \textstyle\sum_{l=0}^{k} \tfrac{\lambda_1{}^l}{l!} \tfrac{\lambda_2{}^{(k-l)}}{(k-l)!} \\
&= \tfrac{e^{-(\lambda_1+\lambda_2)}}{k!} \textstyle\sum_{l=0}^{k} k! \times \tfrac{\lambda_1{}^l}{l!} \tfrac{\lambda_2{}^{(k-l)}}{(k-l)!} \\
&= e^{-(\lambda_1+\lambda_2)} \tfrac{(\lambda_1+\lambda_2)^k}{k!}
\end{aligned}$$

by the binomial theorem, so Z is Poissonian, of parameter $\lambda_1 + \lambda_2$.

b) Remember that for all real t, independence of X and Y gives

$$M_Z(t) = E[e^{tZ}] = E[e^{tX}e^{tY}] = E[e^{tX}]E[e^{tY}]$$

and that

$$M_X(t) = e^{-\lambda_1} \sum_{k=0}^{\infty} e^{tk} \tfrac{\lambda_1{}^k}{k!} = e^{-\lambda_1} e^{\lambda_1 e^t} = \exp\{\lambda_1(e^t - 1)\}.$$

Therefore

$$M_Z(t) = \exp(\lambda_1(e^t - 1)) \exp(\lambda_2(e^t - 1)) = \exp\{(\lambda_1 + \lambda_2)(e^t - 1)\} :$$

and we recognise this as the MGF of a Poisson variable of parameter $\lambda_1 + \lambda_2$.

Solution 76 a) The variables X and Y are centred at 0, and ϵ and X are independent, so

$$\mathrm{cov}(X, Y) = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y) = \mathrm{E}(XY) = \mathrm{E}(\epsilon X^2) = \mathrm{E}(\epsilon)\mathrm{E}(X^2) = 0.$$

b) The vector (X, Y) cannot be jointly Gaussian because $X + Y$ isn't Gaussian: $X + Y$ takes the value 0 with probability 1/2. For a more formal argument, we can compute the joint MGF of (X, Y), which is

$$\mathrm{E}\{\exp(sX + tY)\} = \mathrm{E}_\epsilon[\mathrm{E}\{\exp(sX + tY) \mid \epsilon\}] = \tfrac{1}{2}\mathrm{E}\{\exp(sX + tX)\}] + \tfrac{1}{2}\mathrm{E}\{\exp(sX - tX)\}],$$

where $\epsilon = +1$ in the first summand and equals -1 in the second summand. Therefore

$$\mathrm{E}\{\exp(sX + tY)\} = \tfrac{1}{2}\exp\{(s+t)^2/2\} + \tfrac{1}{2}\exp\{(s-t)^2/2\},$$

and this is clearly not of the form $\exp\{(s,t)\mu + (s,t)\Omega(s,t)^{\mathrm{T}}/2\}$ for some μ and Ω, which would be necessary for the joint density to be Gaussian.

c) No, because Y is a function of X!

Solution 77 We use the convolution formula

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx,$$

here with $f_X(x) = f_Y(y) = 1$ if $0 \le x \le 1$, and otherwise zero. We therefore have

$$f_X(x)f_Y(z-x) = 1, \quad 0 \le x \le 1, z-1 \le x \le z, \qquad f_X(x)f_Y(z-x) = 0 \text{ otherwise.}$$

There are several cases to consider:

- If $z > 2$ or $z < 0$, then $\{0 \le x \le 1\} \cap \{z-1 \le x \le z\}$ is empty and $f_Z(z) = 0$.
- If $0 \le z \le 1$, then $\{0 \le x \le 1\} \cap \{z-1 \le x \le z\} = \{0 \le x \le z\}$ and $f_Z(z) = z$.
- If $1 \le z \le 2$, then $\{0 \le x \le 1\} \cap \{z-1 \le x \le z\} = \{z-1 \le x \le 1\}$ and $f_Z(z) = 1-(z-1) = 2-z$.

We find the required density.

Chapter 6

Solution 78 a) $\alpha = 2/\theta^2$.

b) By definition $F_{M_n}(x) = \Pr(\max_{1\le i \le n} X_i \le x)$. So,

$$F_{M_n}(x) = \left(\int_0^x \alpha y dy\right)^n = \alpha^n \left(\frac{x^2}{2}\right)^n = \left(\frac{x}{\theta}\right)^{2n}.$$

Furthermore $f_{M_n}(x) = dF_{M_n}(x)/dx$. So,

$$f_{M_n}(x) = \frac{2n}{\theta^{2n}} x^{2n-1}, \quad 0 < x < \theta,$$

and

$$\begin{aligned}
\mu_1 = \mathrm{E}(M_n) &= \int_0^\theta x \frac{2n}{\theta^{2n}} x^{2n-1} dx = \frac{2n}{\theta^{2n}} \frac{\theta^{2n+1}}{2n+1} = \frac{2n}{2n+1}\theta, \\
\mu_2 = \mathrm{E}(M_n^2) &= \int_0^\theta x^2 \frac{2n}{\theta^{2n}} x^{2n-1} dx = \frac{2n}{\theta^{2n}} \frac{\theta^{2n+2}}{2n+2} = \frac{n}{n+1}\theta^2.
\end{aligned}$$

c) We must show that for each $\epsilon > 0$, $\lim_{n\to\infty} \Pr(\mid M_n - \theta \mid \ge \epsilon) = 0$. But, by Chebyschov's inequality,

$$\begin{aligned}
\Pr(\mid M_n - \theta \mid \ge \epsilon) &\le \frac{1}{\epsilon^2}\mathrm{E}\{(M_n-\theta)^2\} \\
&= \frac{1}{\epsilon^2}\mathrm{E}(M_n^2 - 2M_n\theta + \theta^2) \\
&= \frac{\theta^2}{\epsilon^2}\left(\frac{n}{n+1} - \frac{4n}{2n+1} + 1\right) \\
&= \frac{\theta^2}{\epsilon^2}\frac{1}{(2n+1)(n+1)},
\end{aligned}$$

which gives the result.

Solution 79 a) Since X is a positive random variable, by Markov's inequality:

$$\Pr\{X > 85\} \leq \mathrm{E}\left(\tfrac{X}{85}\right) = \tfrac{75}{85} \simeq 0,88.$$

b) Knowing the second moment of X allows us to use Chebyshov's inequality,

$$\Pr\{X > 85\} = \Pr\{X^2 > 85^2\} \leq \tfrac{1}{85^2}\mathrm{E}(X^2) = \tfrac{\sigma^2+\mathrm{E}(X)^2}{85^2} = \tfrac{25+75^2}{85^2} \simeq 0,78.$$

We also have:

$$\Pr\{65 \leq X \leq 85\} = \Pr\{\mid X - \mathrm{E}(X) \mid \leq 10\} = 1 - \Pr\{\mid X - \mathrm{E}(X) \mid > 10\} \geq 1 - \tfrac{\sigma^2}{10^2} = 3/4.$$

c) Since the marks of each of the n students are independent variables,

$$\overline{X}_n = \tfrac{1}{n}\sum_{k=1}^{n} X_k$$

is also of expectation 75, and its variance is $(n25)/n^2 = 25/n$. Hence

$$\Pr\{\mid \overline{X}_n - \mathrm{E}(\overline{X}_n) \mid \leq 5\} = 1 - \Pr\{\mid \overline{X}_n - \mathrm{E}(\overline{X}_n) \mid > 5\} \geq 1 - \tfrac{25/n}{25} = \tfrac{n-1}{n} :$$

with 10 students our probability is at least 0.9 .

d) Let's see if it is pertinent to use the CLT. If it can be applied, we have:

$$\Pr\{\mid \overline{X}_n - \mathrm{E}(\overline{X}_n) \mid \leq 5\} \simeq \Pr(-5 \leq Z_n \leq 5) = \Pr(\tfrac{-5}{\sqrt{25/n}} \leq Z \leq \tfrac{5}{\sqrt{25/n}}),$$

where Z_n is normal and centred of variance $25/n$ and Z is standard normal. The minimum number of students obtained this time is $n = 3$, quite a different result! With $n = 10$, we are not in an asymptotic case and the CLT therefore doesn't apply.

Solution 80 The number Y of sixes obtained in 120 throws follows a binomial distribution with $n = 120$ and $p = 1/6$, with mean $\mu = np = 120 \times 1/6 = 20$ and variance $\sigma^2 = np(1-p) = 120 \times 1/6 \times 5/6 \simeq 16.7$.

By the CLT, we can approximate the binomial Y by $X \sim \mathcal{N}(20, 16.7)$. The required probability is thus approximately (using the continuity correction)

$$\Pr(X < 15.5) = \Pr\left(Z < \tfrac{15.5-\mu}{\sigma}\right) = \Pr\left(Z < \tfrac{15.5-20}{\sqrt{16.7}}\right) = \Pr(Z < -1.1) \simeq 0.135,$$

where $Z \sim N(0, 1)$.

Solution 81 The number of heads obtained, X, is a sum of 500 independent random Bernoulli variables of parameter 1/2, so it is binomial with parameters $n = 500$ and $p = 1/2$. Its expectation and variance are therefore $\mu = np = 500 \times \frac{1}{2} = 250$ and $\sigma^2 = np(1-p) = 500 \times \frac{1}{2} \times \frac{1}{2} = 125$. The CLT gives

$$\Pr(250 - 10 \leq X \leq 250 + 10) = \Pr(240 \leq X \leq 260) \simeq \Pr\left(\tfrac{240-\mu}{\sigma} \leq Z \leq \tfrac{260-\mu}{\sigma}\right)$$

i.e.,

$$\Pr(-0.894 \leq Z \leq 0.894) = 2 \times \Pr(0 \leq Z \leq 0.894) = 2 \times 0.314 = 0.628,$$

where $Z \sim N(0, 1)$.

Solution 82 Let $X_1, \ldots, X_{50}$ denote the numbers, $A_1, \ldots, A_{50}$ their rounded versions and $U_1, \ldots, U_{50}$ the corresponding errors. Thus $X_k = A_k + U_k$. The sum obtained by rounding is $\sum_{k=1}^{50} A_k$ and the exact sum is $\sum_{i=k}^{50} X_k$, so the error is

$$\sum_{k=1}^{50} X_k - \sum_{k=1}^{50} A_k = \sum_{k=1}^{50} U_k.$$

The variable $\sum_{k=1}^{50} U_k$ has zero expectation and variance $50 \times 1/12$ (by independence). We therefore have, by application of the CLT:

$$\Pr\left(\left|\sum_{k=1}^{50} U_k\right| > 3\right) \simeq \Pr\left(|Z| > \frac{3}{\sqrt{50/12}}\right) = 2 \times \Pr\left(Z > \frac{3}{\sqrt{50/12}}\right) \approx 2\left\{1 - \Phi(1.47)\right\} \simeq 0.142.$$

Here Z is standard normal.

Solution 83 Let $T = X_1 + \cdots + X_{100}$ denote the random variable for the lifespan of the system. Then since T is a sum of many independent identically distributed random variables, the central limit theorem implies that T has an approximate normal distribution of expectation $100 \times 5 = 500$ and variance $\sigma^2 = 100 \times 25 = 2500 = 50^2$. Hence

$$\Pr(T > 525) = \Pr\left(\frac{T-500}{50} > \frac{525-500}{50}\right) \simeq \Pr(Z > 0.5) = 1 - \Phi(0.5) \simeq 0.31,$$

where $Z \sim N(0,1)$.

Solution 84 In this problem, a first light-bulb of lifetime X_1 is put in place, then, when this one is at the end of its life, it is replaced by another light-bulb of lifetime X_2, etc.

Let X_A denote the total lifetime of the light-bulbs of type A. Since $40 \geq 25$ and the bulbs should be independent, we can apply the central limit theorem, which tells us that the sum $X_A \dot{\sim} N(\mu_A, \sigma_A^2)$, where $\mu_A = 40/\lambda_A = 4000$ [hours] and $\sigma_A^2 = 40/\lambda_A^2 = 400000$ [hours^2].

Similarly, with X_B the total lifetime of the B-type light-bulbs, $X_B \dot{\sim} N(\mu_B = 3000, \sigma_B^2 = 150000)$.

Using properties of the normal distribution, the total lifetime of all the light-bulbs, $X = X_A + X_B$ is approximately a normal variable of mean $\mu = \mu_A + \mu_B = 7000$ hours and variance $\sigma^2 = \sigma_A^2 + \sigma_B^2 = 550000$ hours^2. Having obtained the distribution of X we can now calculate our probability to be

$$\Pr(X \geq 6500) = 1 - \Pr(X \leq 6500) = 1 - \Phi\left(\frac{6500-7000}{\sqrt{550000}}\right) = 1 - \Phi(-0.67) = 0.75.$$

Solution 85 a) Since $Y = \ln X$,

$$\mathrm{E}(Y) = \int_0^1 \ln x \, \mathrm{d}x = x \ln x - x|_0^1 = -1,$$

and

$$\mathrm{var}(Y) = \mathrm{E}\left(Y^2\right) - \mathrm{E}(Y)^2 = \left[\int_0^1 (\ln x)^2 \mathrm{d}x\right] - 1 = x(\ln x)^2 - 2x\ln x + 2x|_0^1 - 1 = 1.$$

b) We observe that the function ln is strictly increasing and therefore

$$\Pr\left(Z < 10^{-40}\right) = \Pr\left(\ln Z < -40 \ln 10\right),$$

and that, if we set $Y_i = \ln X_i$, we obtain

$$\ln Z = \sum_{i=1}^{100} Y_i,$$

that is to say $\ln Z$ is the sum of 100 independent variables with the same distribution, and whose expectation and variance were obtained in a). We use the CLT and replace $\ln Z$ by $W \sim N(-100, 100)$, giving

$$\Pr\left(Z < 10^{-40}\right) \simeq \Pr\left(W < -40 \ln 10\right) = \Pr\left(\frac{W - E(W)}{\sqrt{\mathrm{var}(W)}} < \frac{-40 \ln 10 + 100}{10}\right) \simeq \Phi(0.790) \simeq 0.78.$$

Solution 86 Let A, B, C denote the random variables for the calculation times of each of the three sections.

a) In an obvious notation, $\mathrm{cov}(A, C) = \mathrm{corr}(A, C)\sigma_A\sigma_C = 0.2 \times 2.5 \times 1.3 = 0.65$.

b) We have

$$\mathrm{E}(T) = \mathrm{E}(A + B + C) = \mathrm{E}(A) + \mathrm{E}(B) + \mathrm{E}(C) = 5.5 + 3.4 + 4.5 = 13.4,$$

and

$$\begin{aligned}
\mathrm{var}(T) &= \mathrm{var}(A + B + C) \\
&= \mathrm{var}(A + C) + \mathrm{var}(B) \text{ (since } B \text{ is independent of } A \text{ and } C) \\
&= \mathrm{var}(A) + \mathrm{var}(C) + 2\mathrm{cov}(A, C) + \mathrm{var}(B) \\
&= 2.5^2 + 1.3^2 + 2 \times 0.65 + 2.6^2 \\
&= 16.
\end{aligned}$$

c) The calculation time of section B is the sum of 100 independent and identically distributed calculation times, so the central limit theorem implies that the distribution of B is approximately normal, and $B \dot\sim \mathcal{N}(3.4, 6.76)$.

d) Since T is a sum of normal variables, it is approximately normal, with $T \dot\sim \mathcal{N}(13.4, 16)$. Thus

$$\Pr(T \le 10) = \Pr\left(\frac{T - 13.4}{\sqrt{16}} \le \frac{10 - 13.4}{\sqrt{16}}\right) = \Phi(-0.85) \approx 0.20,$$

and

$$\Pr(T \ge 20) = \Pr\left(\frac{T - 13.4}{\sqrt{16}} \ge \frac{20 - 13.4}{\sqrt{16}}\right) = 1 - \Phi(1.65) \approx 0.05.$$

Solution 87 We note that $\overline{X} \dot\sim \mathcal{N}(\mu, \mu/n)$ using the central limit theorem, and apply the delta method with $g(u) = 2\sqrt{u}$, giving $g'(u) = u^{-1/2}$. Therefore

$$Y = g(\overline{X}) \dot\sim \mathcal{N}\left\{g(\mu), g'(\mu)^2 \times \mu/n\right\} = \mathcal{N}(2\sqrt{\mu}, 1/n), \quad n \to \infty.$$

Thus the variance of Y does not depend on μ, at least to this order of approximation, and therefore the square root transformation is variance-stabilizing for the Poisson distribution.

Solution 88 Using the results on linear combinations of normal variables, $\overline{X} \sim \mathcal{N}(\mu, \sigma^2/n)$, and if we apply the delta method with $g(u) = 1/u$, we have $g'(u) = -1/u^2$, provided $u \neq 0$. Therefore

$$Y = g(\overline{X}) \dot\sim \mathcal{N}\left\{g(\mu), g'(\mu)^2 \times \sigma^2/n\right\} = \mathcal{N}\{1/\mu, \sigma^2/(n\mu^4)\}, \quad n \to \infty,$$

provided that $\mu \neq 0$. Note that if X has units of length (say), then its mean and its variance have units of length and length^2, so $1/X$ has units of 1/length and its variance has units of 1/length^2, agreeing with the distribution here.

If $\mu = 0$, then for all n, we have $\Pr(\overline{X} < 0) = \Pr(\overline{X} > 0) = 1/2$, so the distribution of Y will concentrate at $\pm\infty$ with equal probabilities as $n \to \infty$.

Chapter 7

Solution 89 The median is a location measure in the same way as the average, as it finds a typical or central value that best describes the data. The median is called robust because it is very little influenced by a single outlier (or even a few outliers), unlike the average.

Solution 90 Only b): All the observations are equal, because $s^2 = (n-1)^{-1}\sum(x_j - \overline{x})^2 = 0$.

The only way we can observe $s^2 = 0$ is if $(x_j - \overline{x}) = 0$ for all j, and this means that $x_1 = \cdots = x_n = \overline{x}$. There is no implication here that n is small or that $\overline{x} = 0$, and if the data were normally distributed, then it would be impossible to observe two identical values, because the normal density is continuous.

Solution 91 The empirical covariance measures the association of the variables X and Y but is unbounded. We can give as counter-example to c) the observations $(x_1, y_1) = (1, -1)$ and $(x_2, y_2) = (1, 1)$, for which the covariance is nil. Unlike the correlation, the covariance depends on the units of X and Y.

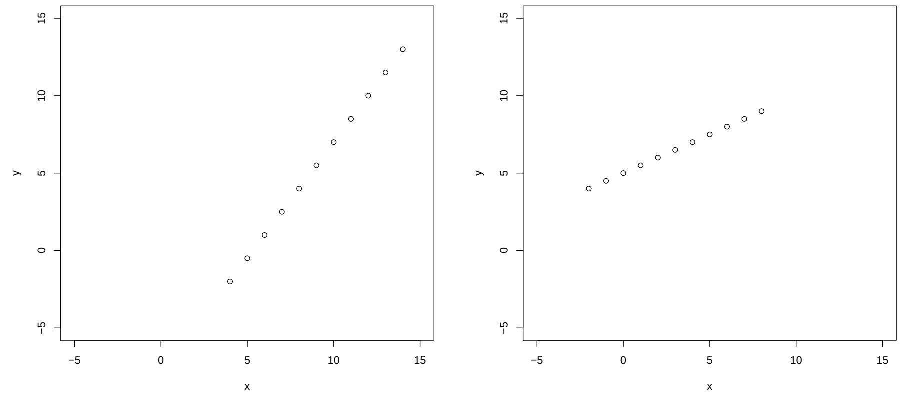
Solution 92 The empirical correlation is dimensionless, measures the (linear) association between the variables X and Y, and lies between -1 and 1, so b) and c) are true. A counter-example to d) is the observations $(x_1, y_1) = (1, -1)$ and $(x_2, y_2) = (1, 1)$, for which the correlation is zero.

Solution 93 The empirical correlation is defined as

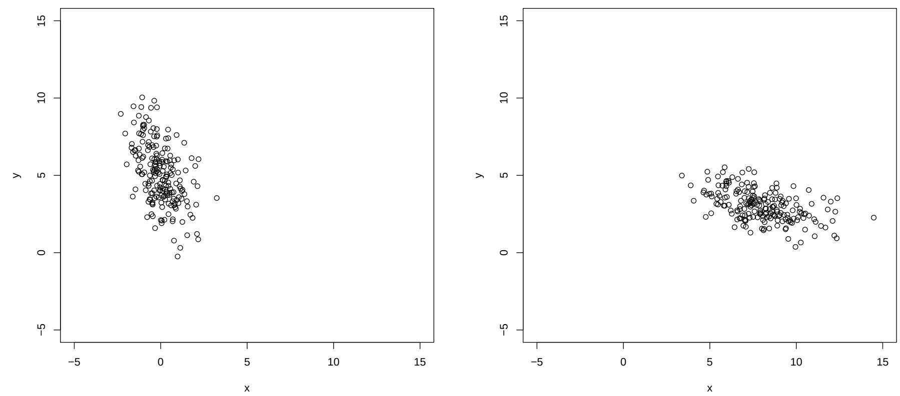
$$r_{XY} = \frac{1}{n-1}\frac{\sum_{i=1}^n (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^n (x_i - \overline{x})^2 \sum_{i=1}^n (y_i - \overline{y})^2}},$$

where $\overline{x}$ and $\overline{y}$ are the averages of $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$. If $r_{XY} = -\pm$, the points line on a straight line of positive/negative slope. If $r_{XY} = 0$ the cloud of points shows no *linear* relation: there may be no relation, or there may be a nonlinear relation.

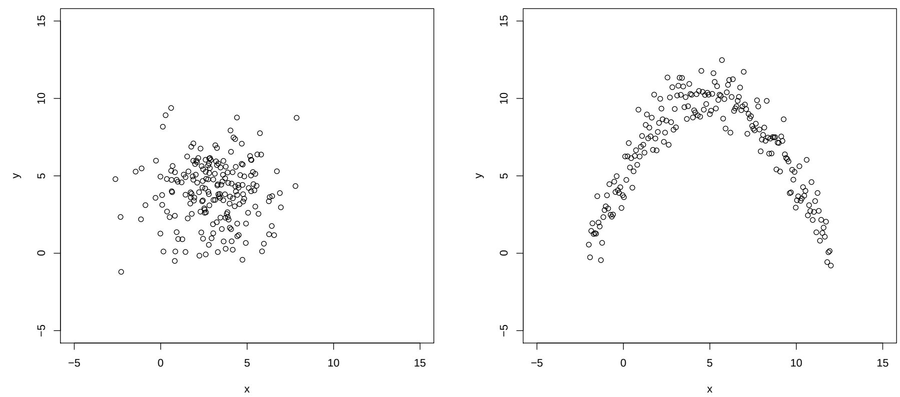
Solution 94 (i) If the empirical correlation is 1, then the points (x_i, y_i) are perfectly aligned and the corresponding line has a strictly positive slope. Possible data configurations are



(ii) If the empirical correlation is -0.5, then the points (x_i, y_i) form a cloud that has a rough negative trend. Possible data configurations are



(iii) If the empirical correlation is 0, then the points (x_i, y_i) form a cloud that doesn't exhibit a linear trend. Possible data configurations are



Chapter 8

Solution 95 The mean and variance are respectively α/λ and α/λ^2, so the estimators are the solutions to the equations

$$\frac{\alpha}{\lambda} = \overline{X} = n^{-1} \sum_{j=1}^{n} X_j, \quad \frac{\alpha}{\lambda^2} = n^{-1} \sum_{j=1}^{n} (X_j - \overline{X})^2,$$

i.e.,

$$\tilde{\lambda} = \overline{X} / \left\{ n^{-1} \sum_{j=1}^{n} (X_j - \overline{X})^2 \right\}, \quad \tilde{\alpha} = (\overline{X})^2 / \left\{ n^{-1} \sum_{j=1}^{n} (X_j - \overline{X})^2 \right\}.$$

Solution 96 (a) Since $X \sim \mathcal{N}(\mu, \sigma^2)$, we have

$$\mathrm{E}(Y^r) = \mathrm{E}(e^{rX}) = M_X(r) = \exp(r\mu + r^2\sigma^2/2), \quad r \in \mathbb{R},$$

which gives the stated equations after a little work with $M_X(1)$ and $M_X(2)$.

(b) The moment estimators solve the equations

$$\overline{Y} = e^{\psi}, \quad n^{-1} \sum_{j=1}^{n} (Y_j - \overline{Y})^2 = e^{2\psi}(e^{\sigma^2} - 1)$$

simultaneously, so they are (after a little algebra)

$$\tilde{\psi} = \log(\overline{Y}), \quad \tilde{\sigma}^2 = \log \left[1 + \log \left\{ n^{-1} \sum_{j=1}^{n} (Y_j - \overline{Y})^2 / \overline{Y}^2 \right\} \right].$$

Solution 97 The likelihood is

$$L(\theta) \;=\; \prod_{i=1}^{n} f_\theta(y_i) = \begin{cases} 2^{-n}\theta^n\theta^{2n}e^{-\theta\sum_{i=1}^n y_i}\prod_{i=1}^n y_i^2, & \theta > 0, \\ 0, \quad \theta \le 0, \end{cases}$$

and this equals

$$L(\theta) = C \exp\left\{ -\theta \sum_{i=1}^{n} y_i + 3n \log\theta \right\}, \quad \theta > 0,$$

where C does not depend on the parameter θ, The function $\theta \mapsto \ell(\theta) = 3n\log\theta - \theta\sum_{i=1}^n y_i$ is concave and the equation $d\ell(\theta)/d\theta = 0$ has exactly one root, so $\ell(\theta)$ reaches its maximum at

$$\widehat{\theta} = \frac{3n}{\sum_{i=1}^n y_i},$$

which is the maximum likelihood estimator of θ.

Solution 98 The likelihood is

$$L(\theta) \;=\; \prod_{i=1}^{n} e^{-\theta}\frac{\theta^{z_i}}{z_i!} = e^{-2n\theta}\frac{(2\theta)^{z_1+\cdots+z_n}}{z_1!\ldots z_n!}, \qquad \theta > 0.$$

so

$$L(\theta) \propto \exp\left\{ (z_1 + \cdots + z_n)\log 2\theta - 2n\theta \right\}, \quad \theta > 0,$$

where the constant of proportionality does not depend on the parameter θ. The function $L(\cdot)$ reaches its maximum at

$$\widehat{\theta} = \frac{\sum_{i=1}^n z_i}{2n} = \overline{z}/2,$$

which is the maximum likelihood estimator of θ.

Solution 99 The density of the $\mathcal{U}(0, b)$ distribution is

$$f(x) = b^{-1}, \quad 0 \le x \le b.$$

Since the observations are independent, the likelihood based on a random sample that has taken values $x_1, \ldots, x_n$ is

$$L(b) = f(x_1) \times \cdots \times f(x_n) = b^{-n}, \quad 0 \le x_1, \ldots, x_n \le b,$$

which reaches its maximum when b takes the smallest possible value for which $L(b) > 0$, i.e., when $b = \max x_i$. The maximum likelihood estimate of b is therefore $\widehat{b} = \max_{i=1}^n x_i$, and the corresponding estimator is $\max_{i=1}^n X_i$.

Solution 100 (a) With U_a denoting the event '$X_1 \sim U(0,a)$', and likewise for U_b, we have

$$
\begin{aligned}
F_{X_1}(x) &= \begin{cases} 0, & x < 0, \\ \Pr(X_1 \le x \mid U_a)\Pr(U_a) + \Pr(X_1 \le x \mid U_b)\Pr(U_b), & 0 \le x \le b, \\ 1, & x > b, \end{cases} \\
&= \begin{cases} 0, & x < 0, \\ p\frac{x}{a} + (1-p)\frac{x}{b}, & 0 \le x \le a, \\ p + (1-p)\frac{x}{b}, & a \le x \le b, \\ 1, & x > b, \end{cases}
\end{aligned}
$$

and the corresponding density function is

$$
f_{X_1}(x) = \begin{cases} 0, & x < 0, \\ \frac{p}{a} + \frac{(1-p)}{b}, & 0 \le x \le a, \\ \frac{(1-p)}{b}, & a \le x \le b, \\ 0, & x > b. \end{cases}
$$

(b) The variables being supposed independent, each of the X_i will belong to the interval $[0,a]$ with probability $F_{X_1}(a)$. Thus $N_a \sim B\{n, \tilde{p} = p + (1-p)a/b\}$, with expectation $\mathrm{E}(N_a) = n\tilde{p}$ and variance $\mathrm{var}(N_a) = n\tilde{p}(1-\tilde{p})$.

(c) The likelihood for p is

$$
L(p) = \left(\frac{p}{a} + \frac{1-p}{b}\right)^{N_a} \left(\frac{1-p}{b}\right)^{n-N_a}, \quad 0 \le p \le 1,
$$

and differentiation with respect to p yields

$$
\frac{\partial \log L(p)}{\partial p} = N_a \frac{b-a}{pb + (1-p)a} - \frac{n - N_a}{1-p},
$$

thus giving

$$
\widehat{p} = \frac{N_a b - na}{n(b-a)}.
$$

Solution 101 Let $Z_j = X_j - \theta$, so that $\mathrm{E}(Z_j) = 0$, $\mathrm{var}(Z_j) = (b-a)^2/12$, and the Z_j are independent and identically distributed. Note that $\mathrm{E}(\overline{X}) = \theta$, and that

$$
\left(\overline{X} - \theta\right)^2 = \left\{\sum_{i=1}^n \left(\frac{X_i}{n} - \frac{\theta}{n}\right)\right\}^2 = \left(\frac{1}{n}\sum_{j=1}^n Z_j\right)^2 = \frac{1}{n^2}\sum_{j=1}^n Z_j^2 + \frac{1}{n^2}\sum_{i \ne j} Z_i Z_j.
$$

Since $\mathrm{E}(Z_j) = 0$, we have that $\mathrm{E}(Z_j^2) = \mathrm{var}(Z_j)$ and $\mathrm{E}(Z_i Z_j) = 0$ if $i \ne j$, by independence, so the mean square error of $\overline{X}$ as an estimator of θ is

$$
\mathrm{E}\left\{\left(\overline{X} - \theta\right)^2\right\} = \frac{1}{n^2}\sum_{j=1}^n \mathrm{var}(Z_j) = \frac{1}{n}\mathrm{var}(Z_1) = \frac{(b-a)^2}{12n}.
$$

Solution 102 (a) We have $\mathrm{E}(\overline{T}) = \theta/2$ and $\mathrm{var}(\overline{T}) = \theta^2/(12n)$, so $\mathrm{E}(\widehat{\theta}_1) = \theta$ and $\mathrm{var}(\widehat{\theta}_1) = \theta^2/(3n) \to 0$ as $n \to \infty$.

(b) Now $\Pr(M_n < m) = \Pr(T_1 < m, \ldots, T_n < m)$, so

$$
\Pr(M_n < x) = \begin{cases} 0, & x \le 0, \\ (x/\theta)^n, & x \in [0, \theta], \\ 1, & x > \theta, \end{cases}
$$

from which the density nx^{n-1}/θ^n for $0 < x < \theta$ is easily obtained. Integration then gives $\mathrm{E}(M_n) = n\theta/(n+1)$ and $\mathrm{E}(M_n^2) = n\theta^2/(n+2)$, so $\mathrm{var}(M_n) = n\theta^2/\{(n+2)(n+1)^2\}$.

(c) We take $\widehat{\theta}_2 = (n+1)M_n/n$, which gives

$$\mathrm{E}(\widehat{\theta}_2) = \theta, \quad \mathrm{var}(\widehat{\theta}_2) = (n+1)^2 \mathrm{var}(M_n)/n^2 = \theta^2/\{n(n+2)\} \to 0, \quad n \to \infty.$$

(d) Since $\mathrm{var}(\widehat{\theta}_2) < \mathrm{var}(\widehat{\theta}_1)$ for $n > 10$ and both estimators are unbiased, we should use $\widehat{\theta}_2$ whenever $n > 10$ and otherwise use $\widehat{\theta}_1$.

(e) M_n and $\widehat{\theta}_2$ converge towards θ in probability because, for any $\epsilon > 0$ and as $n \to \infty$,

$$\Pr(|M_n - \theta| > \epsilon) = \Pr(\theta - M_n > \epsilon) = \left(\frac{\theta - \epsilon}{\theta}\right)^n \to 0,$$

and

$$\Pr(|\widehat{\theta}_2 - \theta| > \epsilon) = \Pr\left(\left|M_n - \frac{n}{n+1}\theta\right| > \frac{n}{n+1}\epsilon\right) = \Pr\left\{M_n < \frac{n}{n+1}(\theta - \epsilon)\right\} + \Pr\left\{M_n > \frac{n}{n+1}(\theta + \epsilon)\right\} \to 0,$$

as can easily be verified using the distribution function of M_n.

Solution 103 The proportion of the population who gave a false answer is estimated as $12/120 = 0.1$. Let X be the random variable representing the number of wrong answers, that is $X \sim \mathrm{Bin}(120, 0.1)$ (assuming that the answers were independent). We can apply the central limit theorem to deduce a confidence interval with $1 - \alpha = 0.95$, and $\sigma^2 = np(1-p)$. We want the quantile $z_{1-\alpha/2}$ satisfying $\Pr(Z \le z) = 1 - \alpha/2 = 0.975$, i.e., $z_{0.975} = 1.96$. Thus the limits are $0.1 \pm 1.96 \times \sqrt{\frac{0.1 \times (1-0.1)}{120}} = [0.046, 0.154]$.

Solution 104 (a) With n measurements with average $\overline{x}$, the confidence interval at level $(1-\alpha) = 0.9$ for μ has limits $\overline{x} \pm \sigma z_{1-\alpha/2}/\sqrt{n}$, so its length is $2\sigma z_{1-\alpha/2}/\sqrt{n}$. To halve its length, we must quadruple n to equal 100, that is, take 75 further measurements.

(b) Let $\alpha = 0.1$ and $\alpha' = 0.05$. To obtain a 95% confidence interval of the same length as the initial one, we need n' measurements such that $\sqrt{n'} \simeq \sqrt{n} \times z_{1-\alpha'/2}/z_{1-\alpha/2} \simeq 5.958$, i.e., $n' = 35.5$, which amounts to 11 extra measurements.

Solution 105 a) For a normal sample of size n with unknown mean and variance, a confidence interval for the mean is based on the Student t distribution with $\nu = n-1$ degrees of freedom. Thus $\nu = 5$ for the xs and $\nu = 11$ for the ys, and a standard computation gives $[47.44, 50.96]$ as the 95% confidence interval for μ_1, and $[47.29, 49.51]$ as that for μ_2.

b) This time, σ is known, and we use the normal distribution. Again this is a standard computation, this time using the fact that $\sqrt{6}(\frac{\overline{X}-\mu_1}{\sigma_1}) \sim N(0,1)$ to arrive at the interval $[47.94, 50.47]$. Similar computations for the second sample lead to $[47.42, 49.38]$.

c) $\overline{X} - \overline{Y}$ is a linear combination of normal variables, so it is normal, of expectation $\mu_1 - \mu_2$ and variance $\sigma_1^2/6 + \sigma_2^2/12 = 2/3$. The 90% confidence interval for $\mu_1 - \mu_2$ equals $[(49.2 - 48.4) - 1.6445 \times \sqrt{2/3}, (49.2 - 48.4) + 1.6445 \times \sqrt{2/3}] = [-0.54, 2.14]$.

Solution 106 Let X denote the number of coffees drunk annually by any given employee. From the wording, $X \sim N(\mu, \sigma^2)$, with both parameters unknown. The data give $\overline{x} = 500$ and $s^2 = 100^2$. A 95% confidence interval for μ is

$$[\overline{x} - t_{n-1}(1-\alpha/2)s/\sqrt{n}, \overline{x} + t_{n-1}(1-\alpha/2)s/\sqrt{n}],$$

where $n = 300$ and $t_{n-1}(1-\alpha/2)$ is the $(1-\alpha/2)$-quantile of the t_{n-1} distribution, here equal to $z_{1-\alpha/2}$ since $n = 300$ is big, and $\alpha = 0.05$. This gives $[488.68, 511.32]$ coffees as the required confidence interval for μ.

A 95% confidence interval for the variance σ^2 of X is given by

$$\left[\frac{(n-1)s^2}{\chi^2_{n-1}(1-\alpha/2)}, \frac{(n-1)s^2}{\chi^2_{n-1}(\alpha/2)}\right],$$

where $\chi^2_{n-1}(\alpha/2)$ is the $\alpha/2$-quantile of the χ^2_{n-1} distribution at $(n-1)$ degrees of freedom. This gives $[8572.49, 11818.65]$ coffees^2.

Solution 107 a) The average is $\overline{x} = 150.3$ grams and the sample standard deviation is $s = 33.2$ grams. Supposing that the weight of the apples is normally distributed, the bounds of a 90% confidence interval can be calculated as follows: $150.3 \pm t_9(0.95) \times 33.2/\sqrt{10}$, where $t_9(0.95) = 1.833$, giving $[131.0, 169.5]$ grams.

b) A 90% confidence interval covers the true average weight of an apple 9 times out of 10, when computed for many independent samples of apples.

Solution 108 a) We have:

$$\begin{aligned}\overline{x} &= \tfrac{1}{1000}(9 \times 2001 + 21 \times 2003 + \cdots + 3 \times 2021) \simeq 2010.73, \\ s^2 &= \tfrac{1}{999}\{9 \times (2001 - 2010.73)^2 + 21 \times (2003 - 2010.73)^2 + \cdots + 3 \times (2023 - 2010.73)^2\} \simeq 12.81.\end{aligned}$$

Clearly we can take $\widehat{\mu} = \overline{x}$ and $\widehat{\sigma} = \sqrt{s^2}$.

b) Let $Z \sim \mathcal{N}(0, 1)$, so that $\Pr\{Z > 1.96\} \simeq 0.025, \Pr\{Z > 2.58\} \simeq 0.005$. Since the number of pots is very large and because the standard normal distribution is the limiting case of the Student for large degrees of freedom, we use a normal confidence interval, rather than a Student t interval. Thus a 95% confidence interval for μ is $[\widehat{\mu} - \frac{1.96\times\widehat{\sigma}}{\sqrt{1000}}, \widehat{\mu} + \frac{1.96\times\widehat{\sigma}}{\sqrt{1000}}] \simeq [2010.51, 2010.95]$ grams.

For the 99% confidence interval we replace 1.96 by 2.58 and obtain $[2010.44, 2011.02]$ grams.

Solution 109 a) Since $n = 1000$ is quite large, we can suppose that the average salary follows a normal distribution. A 90% confidence interval is therefore (with $\alpha = 0.1$)

$$[\overline{x} - z_{1-\alpha/2}s/\sqrt{n}, \overline{x} + z_{1-\alpha/2}s/\sqrt{n}] = [47375, 48624] \text{ CHF}.$$

b) This comes down to testing the hypothesis H: "the average salary is 50000 CHF". From the previous question, we can reject H at level 90% (because 50000 is not in the 90% confidence interval), so the statement is not reasonable. The true average seems to be lower than 50000.

Solution 110 a) We have $\overline{x} = 9$ minutes and $s^2 = 6.25$ minutes^2.

b) If X is the random variable for the conversation time, we seek $\Pr(X \geq 10)$. If we suppose that $X \sim \mathcal{N}(\overline{x}, s^2)$, which seems reasonable since $\overline{x}$ and s^2 are estimates of $\mathrm{E}(X_i)$ and $\mathrm{var}(X_i)$ and we have assumed that the data are normal, and write $\tilde{X} = (X - \overline{x})/s$, the required probability is $\Pr(\tilde{X} \geq 0.4) = 1 - \Phi(0.4) = 0.345$.

c) The null hypothesis H_0 specifies the mean but not the variance, and under it $T = (\overline{X} - 8)/\sqrt{S^2/9} \sim t_8$. Under the alternative hypothesis $\mathrm{E}(\overline{X}) < 8$, so negative values of T would be evidence against H_0; thus we seek to compute $p_{\text{obs}} = \Pr_0(T \leq t_{\text{obs}})$, small values of which will suggest that H_0 is false. The observed value of T is $t = (9 - 8)/\sqrt{6.25/9} = 1.2$, and $p = \Pr(T \leq 1.2) = 0.868$. Since this exceeds 0.05, we accept H_0. If anything, the data suggest that the average length of a call exceeds 8 minutes.

Solution 111 An optimal rejection region of level α is calculated using the Neyman–Pearson lemma. Write $r = \sum_{j=1}^n y_j$, and let $f_0(y) = \prod_{j=1}^n e^{-y_j}$ and $f_1(y) = \prod_{j=1}^n \lambda e^{-\lambda y_j}$ denote the densities of the sample under the null and alternative hypotheses. Then

$$\frac{f_1(y)}{f_0(y)} = \lambda^n e^{-r(1-\lambda)},$$

which is increasing in r. Therefore we set

$$\mathcal{Y}_\alpha = \left\{ y_1, \ldots, y_n \in \mathbb{R}_+ : \sum_{j=1}^n y_j \leq r_\alpha \right\},$$

and we want to find r_α such that $\alpha = \Pr_0(Y \in \mathcal{Y}_\alpha) = \Pr_0(R \leq r_\alpha)$. Thus r_α is the α quantile of the distribution of $R = \sum_{j=1}^n Y_j$, computed under the null hypothesis that $Y_1, \ldots, Y_n \overset{\text{iid}}{\sim} \exp(1)$. The sum of n standard independent exponential variables follows a $\Gamma(n, 1)$ distribution (to be checked via the moment generating function, if you are unsure of this), giving $r_\alpha = q_{\Gamma,\alpha}$, the α^{th} quantile of a $\Gamma(n, 1)$ distribution. We reject H_0 in favour of H_1 at level α if we observe the event $R > r_\alpha$.

Chapter 9

Solution 112 a) The cumulative distribution function of R is

$$\Pr(R \le r) = \begin{cases} 0, & r < 0, \\ \frac{\pi r^2}{\pi \tau^2} = (\frac{r}{\tau})^2, & 0 \le r \le \tau, \\ 1, & r > \tau, \end{cases}$$

so the density function equals zero outside the interval $[0, \tau]$ and $f(r) = (2r)/\tau^2$ inside this interval.

b) Since the variables $R_1, \ldots, R_n$ are independent, the likelihood based on data $r_1, \ldots, r_n$ is

$$L(\tau) = \prod_{i=1}^{n} f(r_i) = \begin{cases} 0, & 0 \le \tau < m, \\ 2^n (\prod_{i=1}^n r_i)/\tau^{2n}, & \tau \ge m, \end{cases}$$

where $m = \max(r_1, \ldots, r_n)$, and L reaches its maximum at $\widehat{\tau} = m$. Let $M = \max(R_1, \ldots, R_n)$. Since $\Pr(\widehat{\tau} \le r) = \Pr(M \le r) = \Pr(R_1 \le r)^n = (r/\tau)^{2n}$ $(0 < r < \tau)$, $\widehat{\tau}$ has density

$$f_{\widehat{\tau}}(r) = \frac{2n}{\tau^{2n}} r^{2n-1}, \quad 0 < r < \tau,$$

and the bias of $\widehat{\tau}$ is

$$\mathrm{E}(\widehat{\tau}) - \tau = \tfrac{2n}{\tau^{2n}} \int_0^\tau r \times r^{2n-1} dr - \tau = \tfrac{2n}{2n+1}\tau - \tau = -\tau/(2n+1);$$

we see that $\widehat{\tau}$ is biased downwards. An unbiased estimator of τ is $\tilde{\tau} = (2n+1)M/(2n)$.

Solution 113 (a) Let $x_1, \ldots, x_{400} \ge 0$ be the sample. The likelihood is

$$L(k) \;=\; \prod_{i=1}^{400} f_k(x_i) = k^{800} \left(\prod_{i=1}^{400} x_i \right) e^{-k \sum_{i=1}^{400} x_i} \propto \exp\left(800 \log k - k \sum_{i=1}^{400} x_i \right), \quad k > 0,$$

where the constant of proportionality does not depend on k, and the function $k \mapsto 800 \log k - k \sum_{i=1}^{400} x_i$ reaches its maximum at $\widehat{k} = 800/\sum_{i=1}^{400} x_i = 2/\overline{x} = 1$.

(b) The log-likelihood is

$$\ell(k) = 800 \log k - k \sum_{i=1}^{400} x_i, \quad k > 0,$$

plus a constant, and

$$\ell'(k) \;=\; 800/k - \sum_{i=1}^{400} x_i, \quad \ell''(k) = -800/k^2,$$

so the observed information at $\widehat{k}$ equals $J(\widehat{k}) = 800/\widehat{k}^2 = 800$. An approximate confidence interval for k at level $(1-\alpha) = 95\%$ is $[\widehat{k} - J(\widehat{k})^{-1/2} z_{1-\alpha/2}, \widehat{k} + J(\widehat{k})^{-1/2} z_{1-\alpha/2}] = [0.93, 1.07]$ (1000 maravedis)$^{-1}$.

(c) Since the estimated density is xe^{-x}, for $x > 0$, the proportion of families saving less than 1000 maravedis per month can be estimated by

$$\int_0^1 xe^{-x} dx = 1 - 2/e \simeq 0.26.$$

Chapter 10

Solution 114 (a) From the wording, the prior distribution of p is

$$\Pr(p = 0.05) = \Pr(p = 0.1) = 0.5.$$

For fixed p, X follows a binomial distribution of parameters $n = 20$ and p:

$$\Pr(X = k \mid p) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, \ldots, n.$$

(b) From the total probability formula,

$$\begin{aligned}\Pr(X=3) &= \Pr(X=3 \mid p=0.05)\Pr(p=0.05)+\Pr(X=3 \mid p=0.1)\Pr(p=0.1)\\ &= 0.0596\times 0.5+0.1901\times 0.5=0.1249.\end{aligned}$$

(c) Among the 20 components, 3 components have been detected as defective. The posterior probabilities of $p=0.05$ and $p=0.1$ are, from Bayes' theorem,

$$\begin{aligned}\Pr(p=0.05 \mid X=3) &= \frac{\Pr(X=3 \mid p=0.05)\Pr(p=0.05)}{\Pr(X=3)}=\frac{0.0596\times 0.5}{0.1249}=0.2386,\\ \Pr(p=0.1 \mid X=3) &= \frac{\Pr(X=3 \mid p=0.1)\Pr(p=0.1)}{\Pr(X=3)}=\frac{0.1901\times 0.5}{0.1249}=0.7614.\end{aligned}$$

A posteriori it is three times more likely that 10% of the components are defective, rather than 5%.

(d) The posterior expectation of p is

$$\mathrm{E}(p \mid X=3)=0.05\times\Pr(p=0.05 \mid X=3)+0.1\times\Pr(p=0.1 \mid X=3)=0.0881.$$

The posterior variance of p is

$$\begin{aligned}\mathrm{var}(p \mid X=3) &= \mathrm{E}(p^2 \mid X=3)-\mathrm{E}(p \mid X=3)^2\\ &= 0.05^2\times\Pr(p=0.05 \mid X=3)+0.1^2\times\Pr(p=0.1 \mid X=3)-\{\mathrm{E}(p \mid X=3)\}^2\\ &= 0.0004.\end{aligned}$$

(e) The posterior mean of p agrees exactly with neither the employee (who states that $p=0.05$), nor with the inspector (who estimates that $p=0.1$), but the value given by the inspector is nearer the posterior mean.

Solution 115 Let the total number of bugs be N, let the number of bugs found thus far be X, and let $M=N-X$. Since it is possible that there might be no bugs, we represent N using the geometric distribution

$$\Pr(N=n \mid \theta)=\theta(1-\theta)^n,\quad n=0,1,\ldots,\quad 0<\theta<1,$$

and we take $X \mid \{N=n,\theta\}\sim B(n,\theta)$, and $\pi(\theta)=1$ for $0<\theta<1$. We want to find

$$\Pr(M=m \mid X=x)=\Pr(N=m+x \mid X=x),\quad m=0,1,\ldots,$$

which can be written as

$$\frac{\Pr(N=m+x,X=x)}{\Pr(X=x)}=\frac{\int_0^1 \Pr(X=x \mid N=m+x,\theta)\Pr(N=m+x \mid \theta)\pi(\theta)\,d\theta}{\sum_{m=0}^{\infty}\int_0^1 \Pr(X=x \mid N=m+x,\theta)\Pr(N=m+x \mid \theta)\pi(\theta)\,d\theta}.$$

The numerator integral here is

$$\int_0^1 \binom{m+x}{x}\theta^x(1-\theta)^{m+x-x}\times\theta(1-\theta)^{m+x}\times 1\,d\theta=\binom{m+x}{x}\int_0^1\theta^{x+1}(1-\theta)^{2m+x}\,d\theta=\binom{m+x}{x}B(x+2,2m+x+1),$$

where $B(\cdot,\cdot)$ denotes the beta function, so the integral equals

$$\frac{(m+x)!}{x!m!}\frac{(x+1)!(2m+x)!}{(2m+2x+2)!}=(x+1)\frac{(m+x)!(2m+x)!}{m!(2m+2x+2)!},\quad x,m=0,1,\ldots$$

Thus

$$\Pr(M=m \mid X=x)=\frac{(m+x)!(2m+x)!/\{m!(2m+2x+2)!\}}{\sum_{r=0}^{\infty}(r+x)!(2r+x)!/\{r!(2r+2x+2)!\}},\quad m=0,1,\ldots.$$

If $x=0$, this simplifies to

$$\Pr(M=m \mid X=0)=\frac{1/\{(2m+1)(2m+2)\}}{\sum_{r=0}^{\infty}1/\{(2r+1)(2r+2)\}},\quad m=0,1,\ldots;$$

note that the lower sum is finite, because

$$0<\sum_{r=0}^{\infty}\frac{1}{(2r+1)(2r+2)}\le\sum_{r=0}^{\infty}\frac{1}{(r+1)(r+2)}=\sum_{r=0}^{\infty}\left(\frac{1}{r+1}-\frac{1}{r+2}\right)=1,$$

so $\Pr(M = m \mid X = 0)$ gives a well-defined distribution (the probabilities are all positive and their sum equals unity). Now

$$\mathrm{E}(M \mid X = 0) = \frac{\sum_{m=0}^{\infty} m/\{(2m+1)(2m+2)\}}{\sum_{r=0}^{\infty} 1/\{(2r+1)(2r+2)\}}.$$

We have just seen that the lower sum here is finite, but

$$\sum_{m=0}^{\infty} \frac{m}{(2m+1)(2m+2)} = \frac{1}{2}\left\{\sum_{m=0}^{\infty} \frac{2m+1}{(2m+1)(2m+2)} - \sum_{m=0}^{\infty} \frac{1}{(2m+1)(2m+2)}\right\},$$

and first sum here equals $\frac{1}{2}\sum_{m=0}^{\infty}(m+1)^{-1} = +\infty$, while the second is finite. So, as we might expect, $\mathrm{E}(M \mid X = 0) = +\infty$: even if we have found none so far, we can expect there to be an infinite number of bugs in the operating system.

Solution 116 (a) The posterior density is

$$f(\theta \mid t) = \frac{f(t \mid \theta)g(\theta)}{f(t)} \propto f(t \mid \theta)g(\theta) = \lambda\theta e^{-\theta(\lambda+t)} \propto \theta e^{-\theta(\lambda+t)}, \quad \theta > 0,$$

which we recognise as being proportional to the gamma density with shape parameter 2 and scale parameter $\lambda + t$, i.e.,

$$f(\theta \mid t) = (\lambda + t)^2 \theta e^{-\theta(\lambda+t)}, \quad \theta > 0.$$

(b) The posterior density function of θ conditional on $t_1, \ldots, t_n$ is

$$f(\theta \mid t_1, \ldots, t_n) \propto f(t_1, \ldots, t_n \mid \theta)g(\theta) = g(\theta)\prod_{i=1}^{n} f(t_i \mid \theta) = \lambda\theta^n e^{-\theta(\lambda+\sum_{i=1}^{n} t_i)}, \quad \theta > 0,$$

which is maximised when

$$\frac{d}{d\theta}\left\{n \log\theta - \theta\left(\lambda + \sum_{i=1}^{n} t_i\right)\right\} = 0,$$

so the MAP estimate is

$$\widehat{\theta}_{MAP} = \frac{n}{\lambda + \sum_{i=1}^{n} t_i}.$$