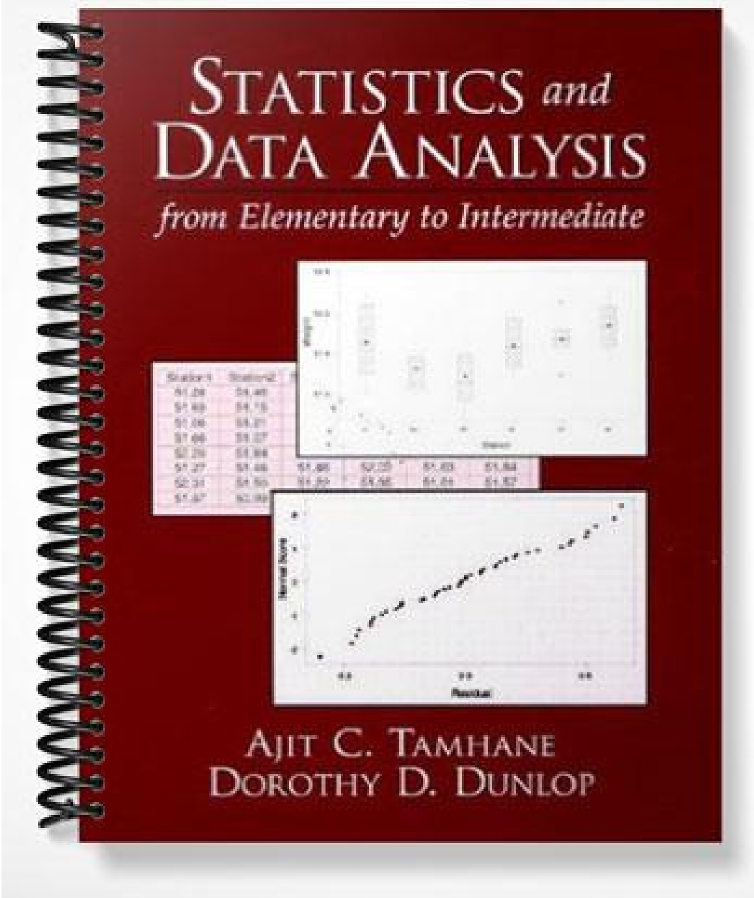
SOLUTIONS MANUAL



Instructor's Solutions Manual

Brent Logan

Statistics and Data Analysis

from Elementary to Intermediate

Ajit C. Tamhane
Dorothy D. Dunlop

PRENTICE HALL, Upper Saddle River, NJ 07458

Acquisitions Editor: Kathy Boothby Sestak
Supplement Editor: Joanne Wendelken
Special Projects Manager: Barbara A. Murray
Production Editor: Wendy Rivers
Supplement Cover Manager: Paul Gourhan
Supplement Cover Designer: PM Workshop Inc.
Manufacturing Buyer: Alan Fischer

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

ISBN 0-13--022578-9

Prentice-Hall International (UK) Limited, London
Prentice-Hall of Australia Pty. Limited, Sydney
Prentice-Hall Canada, Inc., Toronto
Prentice-Hall Hispanoamericana, S.A., Mexico
Prentice-Hall of India Private Limited, New Delhi
Pearson Education Asia Pte. Ltd., Singapore
Prentice-Hall of Japan, Inc., Tokyo
Editora Prentice-Hall do Brazil, Ltda., Rio de Janeiro

CONTENTS

Chapter 2 Solutions

Solutions to Section 2.1

2.1 (a)

Die Result	Coin(s) Result	Number of Outcomes
1	{(H),(T)}	2
2	{(H,H),...,(T,T)}	4
3	{(H,H,H),...,(T,T,T)}	8
4	{ (H,H,H,H),...,(T,T,T,T) }	16
5	{ (H,H,H,H,H),...,(T,T,T,T,T) }	32
6	{ (H,H,H,H,H,H),...,(T,T,T,T,T,T) }	64

There are a total of 126 outcomes.

(b)

Die Result	Number of Heads	Number of Outcomes
1	{(0),(1)}	2
2	{(0),(1),(2)}	3
3	{(0),(1),...,(3)}	4
4	{ (0),(1),...,(4) }	5
5	{ (0),(1),...,(5) }	6
6	{ (0),(1),...,(6) }	7

There are a total of 27 outcomes.

2.2 (a) $T \cap N^c$.

(b) $(T \cup R)^c$.

(c) $(T \cup R \cup N)^c$.

2.3 **Result 1**: Since A and A^c are mutually exclusive,

$$P(A) + P(A^c) = P(A \cup A^c) = P(S) = 1$$

and

$$P(A^c) = 1 - P(A).$$

Result 3:

$$P(A) = P((A \cap B) \cup (A \cap B^c)) = P(A \cap B) + P(A \cap B^c).$$

since $A \cap B$ and $A \cap B^c$ are mutually exclusive.

Result 2: (Uses Result 3)

$$P(A \cup B) = P((A \cap B) \cup (A \cap B^c) \cup (A^c \cap B)).$$

Since all of these are mutually exclusive,

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

Result 4:

$$P(A \cap B^c) = P(A) - P(A \cap B) = P(A) - P(B).$$

Since

$$P(A \cap B^c) = P(A) - P(B) \geq 0,$$
$$P(A) \geq P(B).$$

2.4

$$\begin{aligned} P(A \cap B) &= 1 - P\left((A \cap B)^c\right) \\ &= 1 - P\left(A^c \cup B^c\right) \\ &\geq 1 - [P(A^c) + P(B^c)] \\ &\geq 1 - [1 - P(A) + 1 - P(B)] \\ &\geq P(A) + P(B) - 1. \end{aligned}$$

2.5 (a) $\binom{52}{2} = 1326$.

(b) $\binom{4}{2} = 6$.

(c) $\binom{4}{1}\binom{48}{1} + \binom{4}{2} = 198$.

2.6 Out of the 54 numbers, 6 of them will be chosen for the lottery and 48 will not. So for the grand prize, the probability of winning is

$$\frac{\binom{6}{6}\binom{48}{0}}{\binom{54}{6}} = \frac{1}{25,827,165}.$$

For the second prize, the probability of winning is

$$\frac{\binom{6}{5}\binom{48}{1}}{\binom{54}{6}} = \frac{288}{25,827,165}.$$

For the third prize, the probability of winning is

$$\frac{\binom{6}{4}\binom{48}{2}}{\binom{54}{6}} = \frac{16,920}{25,827,165}.$$

2.7 The total number of moves is $m + n$. On each move, the path can either go right or up. So the number of paths between (0,0) and (m, n) is the same as the number of ways to fill $m + n$ moves with m rights (and the rest ups), or $\binom{m+n}{m}$.

2.8 First split the n objects into two groups, one with n_1 and the other with n_2. Then we can get r total objects by adding up the combinations selecting i from group 1 and $r - i$ from group 2, so that

$$\binom{n}{r} = \sum_i \binom{n_1}{i}\binom{n_2}{r-i}.$$

If $r < n_1$ then we can only select up to n_1 objects from the first group and the upper limit of the sum is n_1. Otherwise, it is r, which yields $\min(n_1, r)$. Similarly, if $r > n_2$ then we can only select up to n_2 objects from the second group, so $r - i \leq n_2$ and the lower limit is $r - n_2$. Otherwise, we can select r objects yielding a lower limit on i of 0. So the final result is

$$\binom{n}{r} = \sum_{i=\max(0, r-n_2)}^{\min(n_1, r)} \binom{n_1}{i}\binom{n_2}{r-i}.$$

2.9

$$\begin{aligned}\binom{n-1}{r-1}+\binom{n-1}{r} &= \frac{(n-1)!}{(r-1)!(n-r)!}+\frac{(n-1)!}{r!(n-r)!}\\ &= \frac{n!}{r!(n-r)!}\left(\frac{r}{n}+\frac{n-r}{n}\right)\\ &= \binom{n}{r}.\end{aligned}$$

n	Coefficients
0	1
1	1 1
2	1 2 1
3	1 3 3 1
4	1 4 6 4 1
5	1 5 10 10 5 1

2.10 (a) $\binom{7}{2}=21$.

(b) $\binom{7}{2}(2)^2(-3)^5=-20,412$.

(c) $\frac{7!}{2!2!3!}=210$.

2.11 (a)

$$\frac{\binom{12}{3}}{\binom{34}{3}}=\frac{220}{5984}.$$

(b)

$$\frac{\binom{2}{2}\binom{32}{1}}{\binom{34}{3}}=\frac{32}{5984}.$$

(c)

$$\frac{\binom{2}{2}\binom{32}{1}}{\binom{34}{3}}+\frac{\binom{12}{2}\binom{22}{1}}{\binom{34}{3}}+\frac{\binom{20}{2}\binom{14}{1}}{\binom{34}{3}}=\frac{4144}{5984}.$$

(d)

$$\frac{\binom{2}{2}\binom{32}{1}}{\binom{34}{3}}+\frac{\binom{12}{2}\binom{22}{1}}{\binom{34}{3}}+\frac{\binom{20}{2}\binom{14}{1}}{\binom{34}{3}}+\frac{\binom{12}{3}\binom{22}{0}}{\binom{34}{3}}+\frac{\binom{20}{3}\binom{14}{0}}{\binom{34}{3}}=\frac{5504}{5984}.$$

2.12 (a)

$$\frac{25}{30}\times\frac{5}{29}=0.144.$$

(b)

$$\frac{25}{30}\times\frac{5}{29}+\frac{5}{30}\times\frac{4}{29}=0.167.$$

(c)

$$\frac{5}{30} \times \frac{4}{29} \times \frac{3}{28} = 0.002.$$

2.13 (a)

$$P(T \cap N^c) = P(T) - P(T \cap N) = 0.77 - 0.45 = 0.32.$$

(b)

$$P((T \cup R)^c) = 1 - P(T \cup R) = 1 - P(T) - P(R) + P(T \cap R) = 1 - .77 - .47 + .29 = 0.05.$$

(c)

$$\begin{aligned} P((T \cup R \cup N)^c) &= 1 - P(N) - P(T \cup R) + P(N \cap (T \cup R)) \\ &= 1 - 0.63 - 0.95 + P((N \cap T) \cup (N \cap R)) \\ &= -0.58 + P(N \cap T) + P(N \cap R) - P(N \cap T \cap R) \\ &= -0.58 + 0.45 + 0.21 - 0.06 = 0.02. \end{aligned}$$

2.14 If we order the 12 kids by team, then there are 12! ways to assign the performance ranks. However, within each team, order is irrelevant, so we need to divide out the 3! ways of ordering the three kids per team. So the final number of ways of ranking the teams is

$$\frac{12!}{3!3!3!3!} = 369,600.$$

2.15 For all four suits to be represented, one suit must have 2 cards and the other suits must have 1 card each. There are $\binom{4}{1}$ ways to choose the suit with 2 cards, and there are $\binom{13}{2}$ ways to select the 2 cards from that suit. For the remaining suits, there are $\binom{13}{1}$ ways to select the 1 card from that suit. So the final probability is

$$\frac{\binom{4}{1}\binom{13}{2}\binom{13}{1}\binom{13}{1}\binom{13}{1}}{\binom{52}{5}} = \frac{685,464}{25,827,165}.$$

Solutions to Section 2.2

2.16

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.3}{0.6} = 0.5.$$

$$P(B|C) = \frac{P(B \cap C)}{P(C)} = \frac{0.3}{0.6} = 0.5.$$

$$P(A|B^c) = \frac{P(A \cap B^c)}{P(B^c)} = \frac{P(A) - P(A \cap B)}{1 - P(B)} = \frac{0.5 - 0.3}{1 - 0.6} = 0.5.$$

$$P(B^c|A \cap C) = \frac{P(A \cap B^c \cap C)}{P(A \cap C)} = \frac{P(A \cap C) - P(A \cap B \cap C)}{P(A \cap C)} = \frac{0.2 - 0.1}{0.2} = 0.5.$$

2.17 (a) For A and B to be mutually exclusive, $P(A \cap B) = 0$ or $P(A \cup B) = P(A) + P(B)$. Then

$$0.8 = 0.4 + p,$$

which means that $p = 0.4$.

(b) For A and B to be independent, $P(A \cap B) = P(A)P(B)$. Then

$$P(A \cap B) = P(A) + P(B) - P(A \cup B) = 0.4 + p - 0.8 = P(A)P(B) = 0.4p,$$

which means that $p = 2/3$.

2.18 (a)

$$P(N|R) = \frac{P(N \cap R)}{P(R)} = \frac{0.21}{0.47} = 0.447.$$

(b)

$$P(R^c|T) = \frac{P(R^c \cap T)}{P(T)} = \frac{P(T) - P(R \cap T)}{P(T)} = \frac{0.77 - 0.29}{0.77} = 0.623.$$

(c)

$$P(T^c|N \cap R) = \frac{P(T^c \cap N \cap R)}{P(N \cap R)} = \frac{P(N \cap R) - P(N \cap R \cap T)}{P(N \cap R)} = \frac{0.21 - 0.06}{0.21} = 0.714.$$

2.19 (a) Let E_1 and F_1 denote whether the first return has an error or is flagged, respectively. Then

$$P(E_1 \cap F_1) = P(F_1|E_1)P(E_1) = 0.9 \times \frac{5}{30} = 0.15.$$

(b)

$$P(F_1) = P(F_1|E_1)P(E_1) + P(F_1|E_1^c)P(E_1^c) = 0.15 + 0.02 \times \frac{25}{30} = 0.167.$$

(c) Let E_2 and F_2 denote whether the second return has an error or is flagged, respectively. Then

$$P(E_2) = P(E_2|E_1)P(E_1) + P(E_2|E_1^c)P(E_1^c) = \frac{4}{29} \times \frac{5}{30} + \frac{5}{29} \times \frac{25}{30} = 0.167$$

and

$$P(F_2 \cap E_2) = P(F_2|E_2)P(E_2) = 0.9 \times 0.167 = 0.15.$$

2.20 Let A_i and P_i denote the events that an ace or a face card is drawn on the ith draw, respectively. Then

$$\begin{aligned} P(\text{ ace before face card}) &= P(A_1) + P((A_1 \cup P_1)^c)P(A_2) + \ldots \\ &= \frac{4}{52} + \left(\frac{36}{52}\right)\frac{4}{52} + \left(\frac{36}{52}\right)^2 \frac{4}{52} + \ldots \\ &= \frac{4}{52}\sum_{i=0}^{\infty}\left(\frac{36}{52}\right)^i \\ &= \frac{4}{52}\left(\frac{1}{1 - \frac{36}{52}}\right) = 0.25. \end{aligned}$$

2.21

(a) Let F and M denote the genetic contributions from the father and mother respectively. Then the probabilities that a father contributes A or a are, respectively,

$$P(F = A) = P(F = A|AA)P(AA) + P(F = A|Aa)P(Aa) = p_0 + q_0/2.$$

and

$$P(F = a) = P(F = a|aa)P(aa) + P(F = a|Aa)P(Aa) = r_0 + q_0/2.$$

Then the probability that the first generation is AA is

$$p_1 = P(F = A, M = A) = P(F = A)P(M = A) = [P(F = A)]^2 = (p_0 + q_0/2)^2.$$

Similarly, the probability that the first generation is Aa is

$$\begin{aligned} q_1 &= P(F = A, M = a) + P(F = a, M = A) \\ &= P(F = A)P(M = a) + P(F = a)P(M = A) \\ &= 2P(F = A)P(F = a) = 2(p_0 + q_0/2)(r_0 + q_0/2). \end{aligned}$$

Finally, the probability that the first generation is aa is

$$r_1 = P(F = a, M = a) = P(F = a)P(M = a) = [P(F = a)]^2 = (r_0 + q_0/2)^2.$$

(b) Similar to (a),

$$p_2 = (p_1 + q_1/2)^2, \ q_2 = 2(p_1 + q_1/2)(r_1 + q_1/2), \text{ and } r_2 = (r_1 + q_1/2)^2.$$

Then

$$\begin{aligned} p_2 &= \left[(p_0 + q_0/2)^2 + 2(p_0 + q_0/2)(r_0 + q_0/2)/2\right]^2 \\ &= [(p_0 + q_0/2)(p_0 + q_0/2 + r_0 + q_0/2)]^2 \\ &= (p_0 + q_0/2)^2 \, (1)^2, \\ q_2 &= 2\left[(p_0 + q_0/2)^2 + 2(p_0 + q_0/2)(r_0 + q_0/2)/2\right] \\ &\quad \times \left[(r_0 + q_0/2)^2 + 2(p_0 + q_0/2)(r_0 + q_0/2)/2\right] \\ &= 2\,(p_0 + q_0/2)\,(r_0 + q_0/2)\,, \\ r_2 &= \left[(r_0 + q_0/2)^2 + 2(p_0 + q_0/2)(r_0 + q_0/2)/2\right]^2 \\ &= [(r_0 + q_0/2)(r_0 + q_0/2 + p_0 + q_0/2)]^2 \\ &= (r_0 + q_0/2)^2 \, (1)^2. \end{aligned}$$

For the recursive proof, assume that this set of equations is true for n. Then

$$\begin{aligned} p_{n+1} &= (p_n + q_n/2)^2 \\ &= \left[(p_0 + q_0/2)^2 + 2(p_0 + q_0/2)(r_0 + q_0/2)/2\right]^2 \\ &= [(p_0 + q_0/2)(p_0 + q_0/2 + r_0 + q_0/2)]^2 \\ &= (p_0 + q_0/2)^2\,, \\ q_{n+1} &= 2(p_n + q_n/2)(r_n + q_n/2) \\ &= 2\left[(p_0 + q_0/2)^2 + 2(p_0 + q_0/2)(r_0 + q_0/2)/2\right] \end{aligned}$$

$$
\begin{aligned}
&\quad \times \left[(r_0 + q_0/2)^2 + 2(p_0 + q_0/2)(r_0 + q_0/2)/2\right] \\
&= 2\,(p_0 + q_0/2)\,(r_0 + q_0/2)\,, \\
r_{n+1} &= (r_n + q_n/2)^2 \\
&= \left[(r_0 + q_0/2)^2 + 2(p_0 + q_0/2)(r_0 + q_0/2)/2\right]^2 \\
&= [(r_0 + q_0/2)(r_0 + q_0/2 + p_0 + q_0/2)]^2 \\
&= (r_0 + q_0/2)^2\,.
\end{aligned}
$$

2.22 Place of residence and opinion on a tax increase are not independent, since

$$P(\text{ Yes and City}) = \frac{100}{1000} = 0.1 \neq P(\text{Yes})P(\text{City}) = \frac{400}{1000} \times \frac{400}{1000} = 0.16.$$

2.23

$$
\begin{aligned}
P(A_1 \cup A_2 \cup \ldots \cup A_n) &= 1 - P((A_1 \cup A_2 \cup \ldots \cup A_n)^c) \\
&= 1 - P(A_1^c \cap A_2^c \cap \ldots \cap A_n^c) \\
&= 1 - P(A_1^c)P(A_2^c) \cdots P(A_n^c) \\
&= 1 - (1-p)^n.
\end{aligned}
$$

Using $p = 0.9$, for $n = 2$ the reliability is 0.99, for $n = 3$, the reliability is 0.999, and for $n = 4$, the reliability is 0.9999. As n gets larger, the reliability approaches 1.

2.24 (a) The event that there is current from A to C is

$$(R_1 \cap R_2) \cup (R_3) \cup (R_4 \cap R_5).$$

(b) The probability that there is current from A to C is

$$
\begin{aligned}
P(\text{Current}) &= P((R_1 \cap R_2) \cup (R_3) \cup (R_4 \cap R_5)) \\
&= P(R_1 \cap R_2) + P(R_3) + P(R_4 \cap R_5) \\
&\quad - P(R_1 \cap R_2 \cap R_3) - P(R_1 \cap R_2 \cap R_4 \cap R_5) - P(R_3 \cap R_4 \cap R_5) \\
&\quad + P(R_1 \cap R_2 \cap R_3 \cap R_4 \cap R_5) \\
&= (0.9)^2 + 0.9 + (0.9)^2 - (0.9)^3 - (0.9)^4 - (0.9)^3 + (0.9)^5 \\
&= 0.996.
\end{aligned}
$$

2.25 (a) Let D be the event that an appliance is defective. Then

$$P(B \cap D) = P(D|B)P(B) = 0.08 \times 0.37 = 0.0296.$$

(b)

$$P(D) = P(A \cap D) + P(B \cap D) = P(D|A)P(A) + 0.0296 = 0.04 \times 0.63 + 0.0296 = 0.0548.$$

(c)

$$P(B|D) = \frac{P(B \cap D)}{P(D)} = \frac{0.0296}{0.0548} = 0.5401.$$

2.26

(a) Let E be the event that a tax return contains an error, and let F be the event that a tax return is flagged. Then

$$P(E \cap F) = P(F|E)P(E) = 0.85 \times 0.15 = 0.1275.$$

(b)

$$P(F) = P(E \cap F) + P(E^c \cap F) = 0.1275 + P(F|E^c)P(E^c) = 0.1275 + 0.05 \times 0.85 = 0.17.$$

(c)

$$P(E^c|F^c) = \frac{P(E^c \cap F^c)}{P(F^c)} = \frac{P(F^c|E^c)P(E^c)}{1 - P(F)} = \frac{0.95 \times 0.85}{1 - 0.17} = 0.973.$$

2.27 (a) Let D and ND refer to the events where a person has or doesn't have the disease, respectively. Then

$$P(D|+) = \frac{P(+|D)P(D)}{P(+|D)P(D) + P(+|ND)P(ND)} = \frac{0.99 \times 0.1}{0.99 \times 0.1 + 0.02 \times 0.9} = 0.846.$$

(b)

$$P(ND|-) = \frac{P(-|ND)P(ND)}{P(-|ND)P(ND) + P(-|D)P(D)} = \frac{0.98 \times 0.9}{0.98 \times 0.9 + 0.01 \times 0.1} = 0.999.$$

The diagnostic test appears pretty reliable, although it is less reliable in identifying true positives than true negatives.

(c)

$$P(D|+) = \frac{P(+|D)P(D)}{P(+|D)P(D) + P(+|ND)P(ND)} = \frac{0.99 \times 0.001}{0.99 \times 0.001 + 0.02 \times 0.999} = 0.047.$$

(d) For rare diseases, too many false positives would appear in the screening program, and it would not be very effective in identifying people with the disease.

Solutions to Section 2.3

2.28 (a) For this to be a p.m.f.,

$$\sum_x f(x) = c(1/2) + c(1/4) + c(1/8) + c(1/16) = 1,$$

or $c = 16/15 = 1.067$.

(b) The c.d.f. is

$$F(x) = \begin{cases} 0 & \text{if } x < 1 \\ 1.067/2 = 0.533 & \text{if } 1 \leq x < 2 \\ 0.533 + 1.067/4 = 0.8 & \text{if } 2 \leq x < 3 \\ 0.8 + 1.067/8 = 0.933 & \text{if } 3 \leq x < 4 \\ 1 & \text{if } x \geq 4. \end{cases}$$

2.29 (a)

x	$f(x)$	$F(y)$, for $x \le y < x+1$
0	6/36	6/36
1	10/36	16/36
2	8/36	24/36
3	6/36	30/36
4	4/36	34/36
5	2/36	1

(b)

$$P(0 < x \le 3) = F(3) - F(0) = 24/36.$$

$$P(1 \le x < 3) = F(2) - F(0) = 18/36.$$

2.30 (a) For this to be a p.m.f.,

$$\sum_x f(x) = \sum_{x=1}^{\infty} \frac{1}{n(n+1)} = 1.$$

Since

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

then

$$\sum_{x=1}^{\infty} = \lim_{n\to\infty} \frac{n}{n+1} = 1,$$

and $f(x)$ is a p.m.f.

(b) From (a) the c.d.f. is

$$F(x) = \begin{cases} 0 & \text{if } x < 1 \\ \frac{i}{i+1} & \text{if } i \le x < i+1 \end{cases}$$

2.31 For x between n and N, the random variable $X = x$ when the largest chip is x and the remaining $n-1$ chips are smaller than x. Out of the $x-1$ chips smaller than x, we want to choose $n-1$, so there are

$$\binom{x-1}{n-1}$$

ways to do this. The total number of ways to choose n chips is

$$\binom{N}{n},$$

so the p.m.f. is

$$P(X = x) = \frac{\binom{x-1}{n-1}}{\binom{N}{n}}$$

2.32 (a) For this to be a p.d.f.,

$$\int_x f(x) = \int_0^1 0.5\, dx + \int_1^3 (0.5 + c(x-1))\, dx = 1.$$

Since

$$\int_x f(x) = .5x|_0^1 + .5x|_1^3 + \frac{c}{2}(x-1)^2|_1^3$$
$$= 1.5 + \frac{c}{2}(4-0) = 1.5 + 2c = 1,$$

then c must be -0.25.

(b) Using the value of c found above, the new p.d.f. for $1 \le X < 3$ is

$$f(x) = 0.5 - 0.25(x-1) = 0.75 - 0.25x.$$

The c.d.f. for x between 0 and 1 is

$$F(x) = \int_0^x 0.5\,dx = 0.5x.$$

At $x = 1$, $F(1) = 0.5$, so the c.d.f. for x between 1 and 3 is

$$F(x) = F(1) + \int_1^x (0.75 - 0.25x)\,dx = 0.5 + (0.75x - 0.125x^2)_1^x = -.125x^2 + 0.75x - .125.$$

Then the final c.d.f. is given by

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ 0.5x & \text{if } 0 \le x < 1 \\ -.125x^2 + .75x - .125 & \text{if } 1 \le x < 3 \\ 1 & \text{if } x \ge 3. \end{cases}$$

2.33 (a) Continuous.

(b)

$$P(1 \le X \le 3) = F(3) - F(1) = 0.8 - 0.4 = 0.4.$$

(c)

$$P(X \ge 1) = 1 - F(1) = 1 - 0.4 = 0.6.$$

2.34 (a) Discrete.

(b)

$$P(1 \le X < 2) = P(X = 1) = 0.8 - 0.4 = 0.4.$$

(c)

$$P(X \ge 1) = 1 - F(0) = 1 - 0.4 = 0.6.$$

Solutions to Section 2.4

2.35 (a) The p.d.f. is

$$f(x) = \begin{cases} \frac{1}{N} & \text{if } x = 1, 2, \ldots, N \\ 0 & \text{otherwise}. \end{cases}$$

(b) The mean is

$$E(X) = \sum_x xf(x) = \sum_{x=1}^{N} x \times \frac{1}{N} = \frac{1}{N} \times \frac{N(N+1)}{2} = \frac{N+1}{2}.$$

The variance is

$$\begin{aligned}
\mathrm{Var}(X) &= \sum_x x^2 f(x) - \left(\frac{N+1}{2}\right)^2 \\
&= \frac{1}{N}\sum_{x=1}^{N} x^2 - \left(\frac{N+1}{2}\right)^2 \\
&= \frac{N(N+1)(2N+1)}{6N} - \left(\frac{N+1}{2}\right)^2 \\
&= \frac{N+1}{2}\left[\frac{2N+1}{3} - \frac{N+1}{2}\right] \\
&= \frac{(N+1)(N-1)}{12}.
\end{aligned}$$

(c) For a single die, $N = 6$, so that $E(X) = 7/2 = 3.5$ and $\mathrm{Var}(X) = (7)(5)/12 = 2.917$.

2.36 (a) The mean is

$$E(X) = \sum_x xf(x) = 0 \times 0.1 + 1 \times 0.2 + \ldots + 8 \times 0.02 = 2.57.$$

The variance is

$$\mathrm{Var}(X) = \sum_x x^2 f(x) - E(X)^2 = 0^2 \times 0.1 + 1^2 \times 0.2 + \ldots + 8^2 \times 0.02 - 2.57^2 = 3.545.$$

(b) The skewness is

$$\beta_3 = \frac{E[(X-\mu)^3]}{\sigma^3} = \frac{(0-2.57)^3 \times 0.1 + \ldots + (8-2.57)^3 \times 0.02}{(3.545)^{3/2}} = 0.948.$$

The distribution is positively skewed.

2.37 (a)

$$\begin{aligned}
E(X) &= \sum_x xf(x) = \sum_{x=n}^{N} x \frac{\binom{x-1}{n-1}}{\binom{N}{n}} \\
&= \frac{1}{\binom{N}{n}} \sum_{x=n}^{N} \frac{x(x-1)!}{(n-1)!(x-n)!} \\
&= \frac{n}{\binom{N}{n}} \sum_{x=n}^{N} \binom{x}{n} \\
&= \frac{n}{\binom{N}{n}} \binom{N+1}{n+1} \\
&= \frac{n(n!(N-n)!)}{N!} \frac{(N+1)!}{(n+1)!(N-n)!} \\
&= \frac{n(N+1)}{n+1}.
\end{aligned}$$

(b) We could estimate N by solving

$$X = \frac{n(N+1)}{n+1}$$

for N, which yields

$$\hat{N} = \frac{x(n+1)}{n} - 1.$$

2.38

$$E(X) = \sum_x x f(x) = \sum_{x=1}^{\infty} x \frac{1}{x(x+1)} = \sum_{x=1}^{\infty} \frac{1}{x+1} = \infty.$$

2.39 (a) The expected demand is

$$E(X) = \sum_x x f(x) = 1 \times 0.01 + 2 \times 0.04 + \ldots + 10 \times 0.05 = 6.19.$$

(b) If $X \leq n$ then

$$\text{Profit} = 0.5X - 1(n - X) = 1.5X - n.$$

If $X > n$ then

$$\text{Profit} = 0.5n - 0.75(X - n) = 1.25n - 0.75X.$$

(c) The table of profits is below

Demand (x)	1	2	3	4	5	6	7	8	9	10
Probability ($f(x)$)	0.01	0.04	0.05	0.10	0.15	0.20	0.20	0.10	0.10	0.05
Profit ($n = 5$)	-3.5	-2	-0.5	1	2.5	1.75	1	0.25	-0.5	-1.25
Profit($n = 6$)	-4.5	-3	-1.5	0	1.5	3	2.25	1.5	0.75	0
Profit($n = 7$)	-5.5	-4	-2.5	-1	0.5	2	3.5	2.75	2	1.25

The expected profits are obtained by computing the weighted average of the profits, weighed by the probability of each occurring.

$$E(\text{Profit}) = \begin{cases} 0.798 & \text{if } n = 5 \\ 1.26 & \text{if } n = 6 \\ 1.205 & \text{if } n = 7. \end{cases}$$

$n = 6$ maximizes the expected profit.

2.40 (a) The c.d.f. is

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ \int_0^x \frac{1}{(1+z)^2}\, dz = \frac{-1}{1+z}\Big|_0^x = 1 - \frac{1}{1+x)} & \text{if } x \geq 0 \end{cases}$$

(b) To find the pth quantile, set

$$p = F(x) = 1 - \frac{1}{1+x}$$

and solve for x. In this case,

$$x = \frac{1}{1-p} - 1$$

For $p = 0.5$,

$$x_{(0.5)} = \frac{1}{1-0.5} - 1 = 1.$$

(c)

$$E(\sqrt{X}) = \int_0^\infty \frac{\sqrt{x}}{(1+x)^2}\,dx$$

Let $x = \tan^2\theta$, so that $dx = 2\tan\theta\sec^2\theta\,d\theta$. Then

$$\begin{aligned} E(\sqrt{X}) &= \int_0^{\pi/2} \frac{2\tan^2\theta\sec^2\theta}{(sec^2\theta)^2}\,d\theta \\ &= 2\int_0^{\pi/2}\sin^2\theta\,d\theta \\ &= 2\frac{\pi}{4} = \frac{\pi}{2}. \end{aligned}$$

2.41 (a) The c.d.f. is

$$F(x) = \begin{cases} 0 & \text{if } x < 1 \\ \int_1^x 2z^{-3}\,dz = -z^{-2}|_1^x = 1 - x^{-2} & \text{if } x \geq 1 \end{cases}$$

(b) To find the pth quantile, set

$$p = F(x) = 1 - x^{-2}$$

and solve for x. In this case,

$$x = \sqrt{\frac{1}{1-p}}$$

For $p = 0.5$,

$$x_{(0.5)} = \sqrt{\frac{1}{1-0.5}} = \sqrt{2} = 1.414.$$

(c) The mean is

$$E(X) = \int_1^\infty x \times 2x^{-3}\,dx = \int_1^\infty 2x^{-2}\,dx = -2x^{-1}|_1^\infty = 2.$$

Since

$$E(X^2) = \int_1^\infty x^2 \times 2x^{-3}\,dx = \int_1^\infty 2/x\,dx = 2\ln x|_1^\infty = \infty.$$

the variance is also ∞.

2.42 (a) For this to be a c.d.f.,

$$\int_0^1 f(x) = \int_0^1 cx(1-x) = [cx^2/2 - cx^3/3]_0^1 = c/6 = 1,$$

or $c = 6$.

(b) The c.d.f. is

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ \int_0^x 6z(1-z)\,dz = (3z^2 - 2z^3)|_0^x = (3x^2 - 2x^3) & \text{if } 0 \leq x < 1 \\ 1 & \text{if } x \geq 1 \end{cases}$$

(c) The mean is

$$E(X) = \int_0^1 x \times 6x(1-x)\,dx = \int_0^1 (6x^2 - 6x^3)\,dx = (2x^3 - 1.5x^4)_0^1 = 0.5.$$

Since

$$E(X^2) = \int_0^1 x^2 \times 6x(1-x)\,dx = \int_0^1 (6x^3 - 6x^4)\,dx = (1.5x^4 - 1.2x^5)_0^1 = 0.3.$$

the variance is

$$\text{Var}(X) = E(X^2) - E(X)^2 = 0.3 - 0.5^2 = 0.05.$$

2.43 Let $Y = aX + b$. Then

$$E(Y) = aE(X) + b$$

and

$$\begin{aligned}
\text{Var}(Y) &= E\left[(Y - E(Y))^2\right] \\
&= E\left[(aX + b - (aE(X) + b))^2\right] \\
&= a^2 E\left[(X - E(X))^2\right] \\
&= a^2 \text{Var}(X)
\end{aligned}$$

2.44 (a) For X with a finite range $0, 1, \ldots, N$,

$$\begin{aligned}
E(X) &= \sum_x x f(x) = \sum_{x=1}^{N} x(F(x) - F(x-1)) \\
&= 1(F(1) - F(0) + 2(F(2) - F(1)) + \ldots + N(F(N) - F(N-1)) \\
&= NF(N) - \sum_{x=0}^{N-1} F(x) \\
&= \sum_{x=0}^{N-1} [F(N) - F(x)] \\
&= \sum_{x=0}^{N-1} [1 - F(x)]
\end{aligned}$$

As $N \to \infty$,

$$E(X) = \sum_{x=0}^{\infty} [1 - F(x)].$$

(b) For $F(x) = 1 - (1-p)^x$,

$$E(X) = \sum_{x=0}^{\infty} [1 - (1 - (1-p)^x)] = \sum_{x=0}^{\infty} (1-p)^x = \frac{1}{1 - (1-p)} = 1/p.$$

2.45

(a) For X with a finite range $[0, A]$,

$$E(X) = \int_0^A x f(x)\, dx.$$

Let $u = x, dv = f(x)dx$. Then $du = dx$ and $v = F(x)$, and integration by parts yields

$$E(X) = xF(x)|_0^A - \int_0^A F(x)\, dx = A - \int_0^A F(x)\, dx = \int_0^A (1 - F(x))\, dx.$$

As $A \to \infty$,

$$E(X) = \int_0^\infty (1 - F(x))\, dx.$$

(b) For $F(x) = 1 - e^{-\lambda x}$,

$$E(X) = \int_0^\infty [1 - (1 - e^{-\lambda x})]\, dx = \int_0^\infty e^{-\lambda x}\, dx = \frac{-1}{\lambda} e^{-\lambda x}|_0^\infty = 1/\lambda.$$

2.46 (a) Let Y_i be Bernoulli(p). Then

$$\psi_{Y_i}(t) = E\left(e^{tY_i}\right) = \sum_{y=0}^{1} e^{ty} f(y) = pe^t + q.$$

Then

$$\psi_X(t) = \psi_{Y_1}(t) \cdots \psi_{Y_n}(t) = [\psi_{Y_i}(t)]^n = (pe^t + q)^n.$$

(b) The first derivative of the moment generating function is

$$\psi'_X(t) = n(pe^t + q)^{n-1} pe^t.$$

Then

$$E(X) = \psi'_X(0) = n(p + q)^{n-1} p = np.$$

The second derivative is

$$\psi''_X(t) = n(n-1)(pe^t + q)^{n-2} p^2 e^{2t} + n(pe^t + q)^{n-1} pe^t.$$

Then

$$E(X^2) = \psi''_X(0) = n(n-1)(p + q)^{n-2} p + n(p + q)^{n-1} p = np[(n-1)p + 1],$$

and

$$\text{Var}(X) = E(X^2) - E(X)^2 = np[(n-1)p + 1] - (np)^2 = np(1-p).$$

2.47 (a)

$$\begin{aligned}
\psi_X(t) &= E\left(e^{tX}\right) \\
&= \sum_{x=0}^{\infty} e^{tx} \frac{\lambda^x e^{-\lambda}}{x!} \\
&= \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x e^{-\lambda}}{x!} \\
&= \frac{e^{-\lambda}}{e^{-\lambda e^t}} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x e^{-(\lambda e^t)}}{x!}
\end{aligned}$$

Since this sum is the sum of a Poisson(λe^t) p.m.f. from 0 to ∞, it reduces to 1 and

$$\psi_X(t) = e^{\lambda(e^t-1)}.$$

The first derivative of the moment generating function is

$$\psi'_X(t) = (\lambda e^t)e^{\lambda(e^t-1)}.$$

Then

$$E(X) = \psi'_X(0) = \lambda.$$

The second derivative is

$$\psi''_X(t) = (\lambda^2 e^{2t})e^{\lambda(e^t-1)} + (\lambda e^t)e^{\lambda(e^t-1)}.$$

Then

$$E(X^2) = \psi''_X(0) = \lambda^2 + \lambda,$$

and

$$\text{Var}(X) = E(X^2) - E(X)^2 = \lambda^2 + \lambda - (\lambda)^2 = \lambda.$$

(b) Let $Y = \sum_i X_i$. Then

$$\psi_Y(t) = \psi_{X_1}(t)\cdots\psi_{X_n}(t) = e^{\lambda_1(e^t-1)}\cdots e^{\lambda_n(e^t-1)} = e^{(e^t-1)\sum_i \lambda_i}.$$

This is the moment generating function of a Poisson ($\sum_i \lambda_i$) random variable.

Solutions to Section 2.5

2.48 (a)

$$\begin{aligned}\text{Cov}(aX+b, cY+d) &= E\left[(aX+b-(aE(x)+b))\,(cY+d-(cE(Y)+d))\right]\\ &= acE\left[(X-E(X))(Y-E(Y))\right]\\ &= ac\text{Cov}(X,Y).\end{aligned}$$

(b)

$$\begin{aligned}\text{Var}(X\pm Y) &= E\left\{[(X\pm Y)-(E(X)\pm E(Y))]^2\right\}\\ &= E\left\{[(X-E(X))\pm(Y-E(Y))]^2\right\}\\ &= E\left[(X-E(X))^2\right] \pm 2E\left[(X-E(X))(Y-E(Y))\right] + E\left[(Y-E(Y))^2\right]\\ &= \text{Var}(X)+\text{Var}(Y)\pm 2\text{Cov}(X,Y).\end{aligned}$$

(c) If X and Y are independent, then

$$\begin{aligned}E(XY) &= \int_{x,y} xyf(x,y)\,dxdy\\ &= \int_x\int_y xyf(x)f(y)\,dx\,dy\\ &= \int_y yf(y)\,dy\int_x xf(x)\,dx = E(X)E(Y)\end{aligned}$$

and

$$\text{Cov}(X,Y) = E(XY) - E(X)E(Y) = 0.$$

2.49

$$
\begin{aligned}
\operatorname{Cov}(X,Y) &= \operatorname{Cov}(X_1+X_2, X_1-X_2)\\
&= E\left[(X_1+X_2-(E(X_1)+E(X_2))(X_1-X_2-(E(X_1)-E(X_2))\right]\\
&= E\left[((X_1-E(X_1))+(X_2-E(X_2)))\,((X_1-E(X_1))-(X_2-E(X_2)))\right]\\
&= E\left[(X_1-E(X_1))^2-(X_2-E(X_2))^2\right]\\
&= V(X_1)-V(X_2)=0.
\end{aligned}
$$

So the correlation ρ is also 0.

2.50

$$
\begin{aligned}
\operatorname{Cov}(Y_1,Y_2) &= \operatorname{Cov}(X_0-X_1, X_0-X_2)\\
&= \operatorname{Var}(X_0)-\operatorname{Cov}(X_0,X_1)-\operatorname{Cov}(X_0,X_2)+\operatorname{Cov}(X_1,X_2)\\
&= \sigma_0^2.
\end{aligned}
$$

So the correlation is

$$
\rho = \frac{\sigma_0^2}{\sqrt{\sigma_0^2+\sigma_1^2}\sqrt{\sigma_0^2+\sigma_2^2}}.
$$

2.51

$$
\begin{aligned}
\operatorname{Cov}(\bar{X}, X_i-\bar{X}) &= \operatorname{Cov}\left(\frac{1}{n}\sum_j X_j, X_i-\frac{1}{n}\sum_j X_j\right)\\
&= \operatorname{Cov}\left(\frac{1}{n}\sum_j X_j, \frac{n-1}{n}X_i-\frac{1}{n}\sum_{j\neq i} X_j\right)\\
&= \frac{1}{n^2}\left[\operatorname{Cov}\left(\sum_j X_j, (n-1)X_i\right)-\operatorname{Cov}\left(\sum_j X_j, \sum_{j\neq i} X_j\right)\right]\\
&= \frac{1}{n^2}\left[(n-1)\operatorname{Var}(X_i)-\sum_{j\neq i}\operatorname{Var}(X_j)\right]\\
&= \frac{1}{n^2}\left[(n-1)\sigma^2-(n-1)\sigma^2\right]\\
&= 0.
\end{aligned}
$$

2.52 (a)

(x,y)	t	$f(x,y)$
(0,0)	0	0.3
(0,1)	1	0.18
(0,2)	2	0.12
(1,0)	1	0.15
(1,1)	2	0.09
(1,2)	3	0.06
(2,0)	2	0.05
(2,1)	3	0.03
(2,2)	4	0.02

(b)

t	$f(t)$
0	0.3
1	0.33
2	0.26
3	0.09
4	0.02

(c)

$$E(T) = \sum_t tf(t) = 0 \times 0.3 + 1 \times 0.33 + \ldots + 4 \times 0.02 = 1.2.$$

$$\text{Var}(T) = \sum_t t^2 f(t) - E(T)^2 = 0^2 \times 0.3 + 1^2 \times 0.33 + \ldots + 4^2 \times 0.02 - (1.2)^2 = 1.06.$$

(d) If heart and lung problems were positively correlated, $E(T)$ would be slightly decreased. Since 0, 2, and 4 (when $X = Y$) would be more frequent and most of the probability is located at $X = 0$, a positive correlation would increase the likelihood that Y was also low, and decrease the mean. The variance would be increased because, with the correlation between X and Y, extreme values like 0 or 4 would be more likely.

2.53 (a)

$$E(X) = -200 \times 0.2 + 400 \times 0.8 = 280.$$

$$\sigma_X = \sqrt{(-200 - 280)^2 \times 0.2 + (400 - 280)^2 \times 0.8} = 240.$$

(b)

$$E(Y) = -100 \times 0.1 + 300 \times 0.9 = 260.$$

$$\sigma_Y = \sqrt{(-100 - 260)^2 \times 0.1 + (300 - 260)^2 \times 0.9} = 120.$$

(c)

u	$f(u)$	v	$f(v)$
−300	0.02	−500	0.18
100	0.18	−100	0.02
300	0.08	100	0.72
700	0.72	500	0.08

(d)

$$E(U) = -300 \times 0.02 + \ldots + 700 \times 0.72 = 540.$$

$$\sigma_U = \sqrt{(-300)^2 \times 0.02 + \ldots + (700)^2 - (540)^2} = 268.328.$$

$$E(V) = -500 \times 0.18 + \ldots + 300 \times 0.08 = 20.$$

$$\sigma_V = \sqrt{(-500)^2 \times 0.18 + \ldots (500)^2 \times 0.08 - (20)^2} = 268.328.$$

(e)

$$E(U) = 280 + 260 = 540.$$

$$E(V) = 280 - 260 = 20.$$

$$\sigma_U = \sigma_V = \sqrt{(240)^2 + (120)^2} = 268.328.$$

2.54 (a)

x	$f(x)$	y	$f(y)$
100	0.5	0	0.25
250	0.5	100	0.25
		200	0.5

X and Y are not independent because

$$P(X = 100, Y = 0) = 0.2 \neq P(X = 100)P(Y = 0) = 0.5 \times 0.25 = 0.125.$$

(b)

y	$f(y\|x = 100)$	$f(y\|x = 250)$
0	0.4	0.1
100	0.2	0.3
200	0.4	0.6

(c)

$$E(XY) = \sum_{x,y} xyf(x,y) = 100 \times 100 \times \frac{1500}{15000} + \ldots + 250 \times 200 \times \frac{4500}{15000} = 23750,$$

$$E(X) = 175,$$

and

$$E(Y) = 0 \times 0.25 + 100 \times 0.25 + 200 \times 0.5 = 125.$$

Then

$$\text{Cov}(X,Y) = E(XY) - E(X)E(Y) = 23750 - (175)(125) = 1875.$$

2.55 (a)

$$f_X(x) = \int_x^1 8xy\,dy = \left.\frac{8xy^2}{2}\right|_x^1 = 4x - 4x^3.$$

$$f_Y(y) = \int_0^y 8xy\,dx = \left.\frac{8yx^2}{2}\right|_0^y = 4y^3.$$

Since

$$f(x,y) \neq f_X(x)f_Y(y)$$

X and Y are not independent.

(b)

$$f(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{8xy}{4x(1-x^2)} = \frac{2y}{1-x^2}.$$

(c)

$$\begin{aligned} E(XY) &= \int_0^1 \int_x^1 xy8xy\,dy\,dx \\ &= \int_0^1 8x^2 \int_x^1 y^2\,dy\,dx \\ &= \int_0^1 8x^2y^3/3|_x^1\,dx \\ &= \frac{8}{3}\int_0^1 (x^2 - x^5)\,dx \\ &= \frac{8}{3}\left(x^3/3 - x^6/6\right)_0^1 \\ &= 4/9 = 0.444. \end{aligned}$$

$$E(X) = \int_0^1 x(4x - 4x^3)\,dx = \left(4x^3/3 - 4x^5/5\right)_0^1 = 8/15 = 0.533.$$

$$E(Y) = \int_0^1 y(4y^3)\,dy = 4y^5/5\Big|_0^1 = 4/5 = 0.8.$$

Then

$$\text{Cov}(X,Y) = E(XY) - E(X)E(Y) = 0.444 - (0.533)(0.8) = 0.017.$$

Solutions to Section 2.6

2.56 (a) Since

$$E(X_i) = 0$$

and

$$\text{Var}(X_i) = (-1)^2 \times 0.5 + (1)^2 \times 0.5 = 1,$$

then

$$E(Y) = 10E(X_i) = 0.$$

and

$$\text{Var}(Y) = 10\text{Var}(X_i) = 10.$$

(b) An upper bound is

$$P(|Y| \geq 4) \leq \frac{\sigma_Y^2}{4^2} = \frac{10}{16} = 0.625.$$

(c) An upper bound is

$$P(|\bar{X}| \geq 4) \leq \frac{\sigma_{\bar{X}}^2}{4^2} = \frac{1/10}{16} = 0.006.$$

2.57 (a)

$$P(|\bar{X} - \mu| \leq c) = 1 - P(|\bar{X} - \mu| \geq c) \geq 1 - \frac{\sigma_{\bar{X}}^2}{c^2} = 1 - \frac{1}{nc^2}.$$

(b) To assure a probability of at least 0.95, i.e.

$$1 - \frac{1}{nc^2} = 0.95$$

then

$$n = \frac{1}{0.05c^2} = \frac{20}{c^2}.$$

For $c = 0.1$, n must be 2000. For $c = 1.0$, n must be 20. For $c = 2.0$, n must be 5.

2.58

$$P(15 \le X \le 25) = P(|X-20| \le 5) = 1 - P(|X-20| \ge 5) \ge 1 - \frac{4.4^2}{5^2} = 0.2256.$$

Solutions to Section 2.7

2.59 (a) Let X be the number of occupied lanes, then X is binomial with $n = 10$ and $p = 0.75$. Then

$$P(X \le 9) = 1 - P(X = 10) = 1 - \binom{10}{10}(0.75)^{10}(1-0.75)^0 = 0.944.$$

(b)

$$E(X) = np = 10 \times 0.75 = 7.5$$

and

$$\sigma_X = \sqrt{np(1-p)} = \sqrt{10 \times 0.75 \times (1-0.75)} = 1.369.$$

2.60 (a) Let X be the number of times a participant will be selected, then X is binomial with $n = 12$ and $p = 0.1$. Then

$$E(X) = np = 12 \times 0.1 = 1.2.$$

(b)

$$P(X = 2) = \binom{12}{2}(0.1)^2(1-0.1)^{10} = 0.230.$$

(c) Now $p = 0.2$, so that

$$P(X = 2) = \binom{12}{2}(0.2)^2(1-0.2)^{10} = 0.283.$$

2.61 (a)

$$P(X \le 2) = \frac{\binom{10}{0}\binom{30}{4-0}}{\binom{40}{4}} + \frac{\binom{10}{1}\binom{30}{4-1}}{\binom{40}{4}} + \frac{\binom{10}{2}\binom{30}{4-2}}{\binom{40}{4}} = 0.300 + 0.444 + 0.214 = 0.958.$$

(b)

$$\begin{aligned} P(X \le 2) &\approx \binom{4}{0}(0.25)^0(1-0.25)^4 + \binom{4}{1}(0.25)^1(1-0.25)^3 + \binom{4}{2}(0.25)^2(1-0.25)^2 \\ &= 0.316 + 0.422 + 0.211 = 0.949. \end{aligned}$$

The approximation is fairly accurate.

2.62 (a) X has a hypergeometric distribution with $N = 50$, $M = 2$, and $n = 3$. Then

$$f(x) = \frac{\binom{2}{x}\binom{48}{3-x}}{\binom{50}{3}} \text{ for } x = 0, 1, 2.$$

(b)

$$P(X=0)=\frac{\binom{2}{0}\binom{48}{3-0}}{\binom{50}{3}}=0.882.$$

(c) Using $p=M/N=2/50=0.04$,

$$P(X=0)\approx\binom{3}{0}(0.04)^0(1-0.04)^3=0.885.$$

The approximation is very accurate.

2.63 (a) Let X be the number of emissions in a week. X has a Poisson distribution with $\lambda=0.25$. Then

$$P(X\geq 1)=1-P(X=0)=1-\frac{e^{-0.25}(0.25)^0}{0!}=0.221.$$

(b) Let Y be the number of emissions in a year. Y has a Poisson distribution with $\lambda=0.25\times 52=13$. Then

$$P(Y\geq 1)=1-P(Y=0)=1-\frac{e^{-13}(13)^0}{0!}=1.000.$$

2.64 (a) X has a Binomial distribution with $n=200$ and $p=1/20=0.05$. Then

$$P(X\geq 5)=1-P(X\leq 4)=1-\sum_{i=0}^{4}\binom{200}{i}(0.05)^i(1-0.05)^{200-i}$$

(b) The Poisson distribution fits because an error on a page is a rare event, and the number of pages in the manuscript is large. Using $\lambda=0.05\times 200=10$, and the table of cumulative Poisson probabilities,

$$P(X\geq 5)=1-P(X\leq 4)=1-\sum_{i=0}^{4}\frac{e^{-10}10^i}{i!}=1-0.029=0.971.$$

2.65 (a) Using the table of cumulative Poisson probabilities,

$$P(X>10)=1-\sum_{i=0}^{10}\frac{e^{-5}5^i}{i!}=1-0.986=0.014.$$

(b) The distribution needs to be renormalized by the probability that $X\leq 10$, so the p.m.f. is

$$f(x)=\frac{e^{-5}5^x}{x!}\frac{1}{0.986}.$$

Now the total probability sums to 1. Then

$$E(X)=\sum_{x=0}^{10}xf(x)=4.910.$$

2.66 (a) $E(X_1)=np_1=2$, $E(X_2)=np_2=4$, $E(X_3)=np_3=6$, and $E(X_4)=np_4=8$.

(b) Using $\text{Var}(X_i) = np_i(1-p_i)$ and $\text{Cov}(X_i, X_j) = -np_ip_j$, the variance-covariance matrix is

$$\Sigma = \begin{bmatrix} 1.8 & -0.4 & -0.6 & -0.8 \\ -0.4 & 3.2 & -1.2 & -1.6 \\ -0.6 & -1.2 & 4.2 & -2.4 \\ -0.8 & -1.6 & -2.4 & 4.8 \end{bmatrix}.$$

(c) Using

$$\rho_{ij} = \frac{\text{Cov}(X_i, X_j)}{\sigma_i \sigma_j}$$

the correlation matrix is

$$\rho = \begin{bmatrix} 1 & -0.167 & -0.218 & -0.272 \\ -0.167 & 1 & -0.327 & -0.408 \\ -0.218 & -0.327 & 1 & -0.535 \\ -0.272 & -0.408 & -0.535 & 1 \end{bmatrix}.$$

(d)

$$P(X_1 = 2, X_2 = 4, X_3 = 6, X_4 = 8) = \frac{20!}{2!4!6!8!}(0.1)^2(0.2)^4(0.3)^6(0.4)^8 = 0.013.$$

2.67 (a) Let X_1, X_2, and X_3 be the numbers of low, middle, and upper income people surveyed out of 4, respectively. Then X_1, X_2, and X_3 have a multinomial distribution with $n = 4$, $p_1 = 0.3$, $p_2 = 0.45$, and $p_3 = 0.25$. Then

$$P(X_1 = 1, X_2 = 1, X_3 = 2) = \frac{4!}{1!1!2!}(0.3)^1(0.45)^1(0.25)^2 = 0.101.$$

(b)

$$P(X_1 = 0, X_2 = 4, X_3 = 0) = \frac{4!}{0!4!0!}(0.3)^0(0.45)^4(0.25)^0 = 0.041.$$

(c) The number of people with high income, X_3, is marginally binomial with $p = 0.25$. Then

$$P(X_3 = 0) = \binom{4}{0}(0.25)^0(1-0.25)^4.$$

2.68

$$\begin{aligned} P(N_1 = n_1, N_2 = n_2, \ldots, N_c = n_c | N = n) &= \frac{f(n_1, n_2, \ldots, n_c)}{f_N(n)} \\ &= \frac{\frac{e^{-\lambda_1}(\lambda_1)^{n_1}}{n_1!} \cdots \frac{e^{-\lambda_c}(\lambda_c)^{n_c}}{n_c!}}{\frac{e^{-\lambda}(\lambda)^n}{n!}} \\ &= \frac{n!}{n_1!n_2!\cdots n_c!}\left(\frac{\lambda_1}{\lambda}\right)^{n_1}\left(\frac{\lambda_2}{\lambda}\right)^{n_2}\cdots\left(\frac{\lambda_2}{\lambda}\right)^{n_c} \\ &= \frac{n!}{n_1!\cdots n_c!}p_1^{n_1}\cdots p_2^{n_2}. \end{aligned}$$

2.69

(a) Out of the first $x-1$ trials, we need $r-1$ successes, and then the final trial must be a success. The number of ways to do this is

$$\binom{x-1}{r-1}.$$

Then since there are r successes and $x-r$ failures,

$$f(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \text{ for } x = r, r+1, \ldots.$$

(b) Let X_i be i.i.d. geometric random variables with success probability p. Then we obtain r successes by letting $X = X_1 + \ldots + X_r$. Then

$$E(X) = \sum_{i=1}^{r} E(X_i) = r \times \frac{1}{p}$$

and

$$\text{Var}(X) = \sum_{i=1}^{r} V(X_i) = r \times \frac{1-p}{p^2}.$$

2.70 (a) Let E denote the event that the Eastern Conference team wins the series. Then

$$P(E) = P(E,\ 4 \text{ games}) + P(E,\ 5 \text{ games}) + P(E,\ 6 \text{ games}) + P(E,\ 7 \text{ games}).$$

For the Eastern Conference team to win in j games, it must have won 3 out of the first $j-1$ games, as well as the last one. So

$$P(E,\ j \text{ games}) = \binom{j-1}{3} p^4 q^{j-4}$$

and

$$P(E) = \binom{3}{3} p^4 q^0 + \binom{4}{3} p^4 q^1 + \binom{5}{3} p^4 q^2 + \binom{6}{3} p^4 q^3.$$

(b) Let W denote the event that the Western Conference team wins the series. Then the formula for the probability that the Western Conference team wins in j games is similar to that of the Eastern Conference, but with p and q switched. Then

$$\begin{aligned} P(\text{Series ends in } j \text{ games}) &= P(E,\ j \text{ games}) + P(W,\ j \text{ games}) \\ &= \binom{j-1}{3} p^4 q^{j-4} + \binom{j-1}{3} q^4 p^{j-4} \\ &= \binom{j-1}{3} \left[p^4 q^{j-4} + q^4 p^{j-4}\right]. \end{aligned}$$

Solutions to Section 2.8

2.71 (a)

$$P(90 \le X \le 100) = \frac{100-90}{100-70} = 0.333.$$

(b) To find the 90th percentile, set

$$P(b \le X \le 100) = \frac{100 - b}{100 - 70} = 0.1.$$

Solving for b gives $b = 97$.

2.72 (a) If city A is at mile 0 and city C is at mile 75, then city B is at mile 25. Then the probability that the car is towed more than 10 miles is

$$P(10 \le X \le 15) + P(35 \le X \le 65) = \frac{5}{75} + \frac{30}{75} = 0.467.$$

(b) The probability that the car is towed more than 20 miles is

$$P(45 \le X \le 55) = \frac{10}{75} = 0.133.$$

(c)

$$\begin{aligned} P(\text{Towed} > 10 \text{ miles} | X > 20) &= \frac{P(\text{Towed} > 10 \text{ miles} \cap X > 20)}{P(X > 20)} \\ &= \frac{P(35 \le X \le 65)}{P(20 \le X \le 75)} \\ &= \frac{30/75}{55/75} = 0.545. \end{aligned}$$

2.73 (a) To find the pth quantile, set

$$F(x) = 1 - e^{-0.1x} = p$$

and solve for x, which yields

$$x = -10 \ln(1 - p).$$

For $p = 0.5$, the median is 6.931. For $p = 0.75$, the 75th percentile is 13.863.

(b) There are no jobs in 15 minutes if the arrival time of the first job is above 15 minutes, so

$$P(\text{No jobs in 15 minutes}) = P(X > 15) = 1 - F(15) = e^{-0.1(15)} = 0.223.$$

2.74 (a) If the mean time to failure is 10,000 hours, then $\lambda = 1/10000$. To find the median, set

$$F(x) = 1 - e^{-x/10000} = 0.5$$

and solve for x, which yields

$$x = -10000 \ln(1 - 0.5) = 6931.472.$$

(b)

$$P(X \ge 1000) = 1 - F(1000) = e^{-1000/10000} = 0.905.$$

(c) Because of the memoryless property of the exponential distribution,

$$P(X \ge 2000 | X \ge 1000) = P(X \ge 1000) = 0.905.$$

2.75

Let X_i be the failure time of the ith bulb. Then X_i is exponential with failure rate $\lambda = 1/10000$. Since $T = X_1 + X_2 + \ldots + X_5$ is the sum of 5 exponential distributions with failure rate $\lambda = 1/10000$, T has a gamma distribution with the same failure rate $\lambda = 1/10000$ and $r = 5$. Then

$$E(T) = \frac{r}{\lambda} = \frac{5}{1/10000} = 50,000$$

and

$$\text{Var}(T) = \frac{r}{\lambda^2} = \frac{5}{(1/10000)^2} = 500,000,000.$$

2.76 Let X be the beta random variable under study. For $E(X)$ to be $3/4$,

$$\frac{a}{a+b} = \frac{3}{4}$$

or $a = 3b$. For $\text{Var}(X)$ to be $3/32$,

$$\frac{ab}{(a+b)^2(a+b+1)} = \frac{3b^2}{(4b)^2(4b+1)} = \frac{3}{32}$$

or

$$3 \times 32 = 3 \times 16 \times (4b+1)$$

or $b = 1/4$. Then solving for a yields $a = 3/4$.

2.77

$$E(X) = \frac{a}{a+b} = \frac{3}{3+1} = 0.75.$$

$$\text{Var}(X) = \frac{ab}{(a+b)^2(a+b+1)} = \frac{3 \times 1}{(3+1)^2(3+1+1)} = 0.0375.$$

Then solving

$$E(X) = \frac{E(T) - 70}{30}$$

for $E(T)$ yields

$$E(T) = 30E(X) + 70 = 92.5.$$

Similarly, solving

$$\text{Var}(X) = \frac{\text{Var}(T)}{30^2}$$

for $\text{Var}(T)$ yields

$$\text{Var}(T) = 30^2\text{Var}(X) = 33.75.$$

Solutions to Section 2.9

2.78 (a)

$$P(Z \le 1.68) = 0.9535,$$
$$P(Z > 0.75) = 1 - 0.7734 = 0.2266,$$
$$P(Z \le -2.42) = 0.0078,$$
$$P(Z > -1) = 1 - 0.1587 = 0.8413.$$

(b)

$$P(1 \le Z \le 2) = 0.9772 - 0.8413 = 0.1359,$$
$$P(-2 \le Z \le -1) = 0.1359,$$
$$P(-1.5 \le Z \le 1.5) = 0.9332 - 0.0668 = 0.8664,$$
$$P(-1 \le Z \le 2) = 0.9772 - 0.1587 = 0.8185.$$

(c)

$$P(Z \le z_{.1}) = 1 - 0.1 = 0.9,$$
$$P(Z > -z_{.05}) = 1 - 0.05 = 0.95,$$
$$P(z_{.25} \le Z \le z_{.01}) = 0.25 - 0.01 = 0.24,$$
$$P(-z_{.25} \le Z \le z_{.01}) = 0.75 - 0.01 = 0.74.$$

2.79 (a)

$$z_{.3} = 0.525, \ z_{.15} = 1.04, \text{ and } z_{.075} = 1.44.$$

(b) Since $x_p = \mu + z_p\sigma$,

$$x_{.3} = 4 + (0.525) \times 3 = 5.575,$$
$$x_{.15} = 4 + (1.04) \times 3 = 7.12,$$
$$x_{.075} = 4 + (1.44) \times 3 = 8.32.$$

2.80 (a) Let X be the weight of coffee in a can. Then

$$P(X < 16) = P\left(Z = \frac{X-\mu}{\sigma} < \frac{16 - 16.1}{0.5}\right) = P(Z < -0.2) = 0.4207.$$

(b)

$$\begin{aligned} P(16 < X < 16.5) &= P\left(\frac{16-16.1}{0.5} < Z = \frac{X-\mu}{\sigma} < \frac{16.5-16.1}{0.5}\right) \\ &= P(-0.2 < Z < 0.8) = 0.7881 - 0.4207 = 0.3674. \end{aligned}$$

(c) To find the 10th percentile, set

$$P\left(Z \le \frac{x_{.9} - \mu}{\sigma}\right) = 0.1$$

or

$$x_{.9} = \mu + z_{.9}\sigma = 16.1 + (-1.28) \times 0.5 = 15.46.$$

2.81 (a) W has a normal distribution.
(b)

$$E(W) = E(U) - E(V) = 160 - 120 = 40$$

and

$$\mathrm{Var}(W) = \mathrm{Var}(X) + \mathrm{Var}(Y) = 30^2 + 25^2 = 1525.$$

(c)

$$P(U-V>50) = P(W>50) = P\left(Z = \frac{W-\mu}{\sigma} > \frac{50-40}{\sqrt{1525}}\right)$$
$$= P(Z>0.256) = 1-0.6020 = 0.398.$$

2.82 (a) $X-Y$ has a normal distribution with

$$E(X-Y) = 0.526-0.525 = 0.001 \text{ and}$$

$$\sigma_{X-Y} = \sqrt{(3)^2+(4)^2} = 5\times 10^{-4}.$$

(b)

$$P(X-Y>0) = P\left(Z=\frac{X-Y-\mu}{\sigma} > \frac{0-0.001}{5\times 10^{-4}}\right) = P(Z>-2) = 1-0.0227 = 0.9773.$$

(c) Since the number of pairs that fit together has a binomial distribution with $n=10$ and $p=0.9772$, this probability is

$$\binom{10}{9}(0.9772)^9(0.0228)^1 + \binom{10}{10}(0.9772)^{10}(0.0228)^0 = 0.185+0.794 = 0.979.$$

2.83 (a) $\bar{X}$ has a normal distribution with mean $\mu = 90$ and SD $= \sigma/\sqrt{n} = 20/\sqrt{25} = 4$.
(b)

$$P(\bar{X}>100) = P\left(Z=\frac{\bar{X}-\mu_{\bar{X}}}{\sigma_{\bar{X}}} > \frac{100-90}{4}\right) = P(Z>2.5) = 0.0062.$$

(c) To find the 90th percentile, set

$$P\left(Z \le \frac{\bar{x}_{0.10}-\mu_{\bar{X}}}{\sigma_{\bar{X}}}\right) = 0.90$$

or

$$\bar{x}_{0.10} = \mu + z_{0.10}\sigma = 90 + (1.28)\times 4 = 95.12.$$

Solutions to Section 2.10

2.84

$$F_Y(y) = P(Y\le y) = P\left(\ln\frac{X}{1-X}\le y\right)$$
$$= P\left(X \le \frac{e^y}{1+e^y}\right)$$
$$= F_X\left(\frac{e^y}{1+e^y}\right)$$
$$= \frac{e^y}{1+e^y}.$$

Then

$$f_Y(y) = F_Y'(y) = \frac{e^y}{1+e^y} - \frac{e^{2y}}{(1+e^y)^2} = \frac{e^y+e^{2y}-e^{2y}}{(1+e^y)^2} = \frac{e^y}{(1+e^y)^2}.$$

2.85

$$
\begin{aligned}
F_Y(y) &= P(Y \le y) = P(\tan(X) \le y) \\
&= P(X \le \arctan(y)) \\
&= \frac{\arctan(y) - (-\pi/2)}{\pi/2 - (-\pi/2)} \\
&= \frac{\arctan(y) + \pi/2}{\pi}.
\end{aligned}
$$

Then

$$
f_Y(y) = F_Y'(y) = \frac{1}{\pi(1+y^2)}.
$$

2.86 The c.d.f of X is

$$
\int_1^x \lambda x^{-(\lambda+1)}\, dx = -x^{-\lambda}|_1^x = 1 - x^{-\lambda}.
$$

Then the c.d.f. of Y is

$$
\begin{aligned}
F_Y(y) &= P(Y \le y) = P(\ln(X) \le y) \\
&= P(X \le e^y) \\
&= 1 - (e^y)^{-\lambda} = 1 - e^{-\lambda y}.
\end{aligned}
$$

So Y has an exponential distribution with rate parameter λ.

2.87 (a) Solving

$$
Y_1 = X_1/X_2 \text{ and } Y_2 = X_2
$$

for X_1 and X_2 yields

$$
X_1 = Y_1Y_2 \text{ and } X_2 = Y_2.
$$

Then the Jacobian of the transformation is

$$
J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} y_2 & y_1 \\ 0 & 1 \end{vmatrix} = y_2.
$$

Finally, the p.d.f. is

$$
\begin{aligned}
g(y_1, y_2) &= f(\psi_1(y_1,y_2), \psi_2(y_1,y_2))|J| \\
&= \frac{1}{2\pi} \exp\left\{ -\frac{(y_1y_2)^2}{2} - \frac{y_2^2}{2} \right\} y_2 \\
&= \frac{y_2}{2\pi} \exp\left\{ -\frac{y_2^2}{2}(1+y_1^2) \right\}.
\end{aligned}
$$

(b)

$$
\begin{aligned}
f(y_1) &= \int_{-\infty}^{\infty} \frac{y_2}{2\pi} \exp\left\{ -\frac{y_2^2}{2}(1+y_1^2) \right\} dy_2 \\
&= 2\int_0^{\infty} \frac{y_2}{2\pi} \exp\left\{ -\frac{y_2^2}{2}(1+y_1^2) \right\} dy_2 \\
&= \frac{1}{\pi}\left(\frac{-1}{1+y_1^2} \right) \exp\left\{ -\frac{y_2^2}{2}(1+y_1^2) \right\}\Bigg|_0^{\infty} \\
&= \frac{1}{\pi(1+y_1^2)} \quad \text{for } -\infty < y_1 < \infty.
\end{aligned}
$$

2.88

Solving

$$Y_1 = \frac{1}{2}(X_1 - X_2) \text{ and } Y_2 = \frac{1}{2}(X_1 + X_2)$$

for X_1 and X_2 yields

$$X_1 = Y_1 + Y_2 \text{ and } X_2 = -Y_1 + Y_2.$$

Then the Jacobian of the transformation is

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix} = 2.$$

Finally, the p.d.f. is

$$\begin{aligned} g(y_1, y_2) &= f(\psi_1(y_1, y_2), \psi_2(y_1, y_2))|J| \\ &= \frac{1}{2}e^{-(y_1+y_2)/2}\frac{1}{2}e^{-(-y_1+y_2)/2} \times 2 \\ &= \frac{1}{2}e^{-y_2/2}. \end{aligned}$$

This p.d.f. is valid only when $x_1 > 0$ (or $y_1 > -y_2$) and $x_2 > 0$ (or $y_1 < y_2$). So the final domain is $-y_2 < y_1 < y_2$. Then the marginal distribution of Y_2 is

$$f(y_2) = \int_{-y_2}^{y_2} \frac{1}{2}e^{-y_2}\, dy_1 = \frac{1}{2}y_1 e^{-y_2}\Big|_{-y_2}^{y_2} = y_2 e^{-y_2}, \; y_2 > 0.$$

Similarly, the marginal distribution of Y_1 is

$$\begin{aligned} f(y_1) &= \int_{|y_1|}^{\infty} \frac{1}{2}e^{-y_2}\, dy_2 \\ &= \frac{-1}{2}e^{-y_2}\Big|_{|y_1|}^{\infty} \\ &= \frac{1}{2}e^{-|y_1|} \text{ for } -\infty < y_1 < \infty. \end{aligned}$$

Since $f(y_1, y_2) \neq f(y_1)f(y_2)$ they are not independent.

Solutions to Chapter 2 Advanced Exercises

2.89 (a) Since there are $n-1$ husband/wife pairs left after the ith man is paired with his wife, there are $(n-1)!$ ways to permute the wives (keeping a fixed ordering of the husbands). Then

$$P(A_i) = \frac{(n-1)!}{n!}.$$

Similarly, after the ith and jth men are paired with their wives, there are $n-2$ husband/wife pairs left, with $(n-2)!$ ways to permute them, yielding

$$P(A_i \cap A_j) = \frac{(n-2)!}{n!}.$$

This reasoning applies for any number of men paired with their wives.

(b) From the inclusion-exclusion formula,

$$\begin{aligned} P(A_1 \cap \ldots \cap A_n) &= \sum_i P(A_i) - \sum_{i \neq j} P(A_i \cap A_j) + \ldots + (-1)^{n-1} P(A_1 \cap \ldots \cap A_n) \\ &= n\frac{(n-1)!}{n!} - \binom{n}{2}\frac{(n-2)!}{n!} + \ldots + (-1)^{n-1}\frac{1}{n!} \\ &= 1 - \frac{1}{2!} + \ldots + (-1)^{n-1}\frac{1}{n!}. \end{aligned}$$

(c)

$$P(A_1 \cap \ldots \cap A_n) = \sum_{i=1}^{n} \frac{(-1)^{i-1}}{i!}.$$

The Taylor Series expansion for e^x is

$$e^x = 1 + x + x^2/2! + \ldots + \frac{x^n}{n!} + \ldots$$

Letting $x = -1$ gives

$$e^{-1} = 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \ldots + \frac{(-1)^n}{n!} + \ldots$$

and

$$1 - e^{-1} = 1 - \frac{1}{2!} + \frac{1}{3!} + \ldots + \frac{(-1)^{n-1}}{n!} + \ldots = \lim_{n \to \infty} P(A_1 \cap \ldots \cap A_n).$$

2.90 (a) For the left matchbox to be emptied on the $(2n - m + 1)$st trial, out of the first $2n - m$ trials, the left pocket must have been chosen exactly n times. There are

$$\binom{2n-m}{n}$$

ways this could happen, each with probability

$$(1/2)^n (1/2)^{n-m}.$$

Then on the final trial, the left matchbox must be chosen(with probability 1/2), so the final probability is

$$\binom{2n-m}{n}(1/2)^n(1/2)^{n-m} \times (1/2) = \binom{2n-m}{n}(1/2)^{n+1}(1/2)^{n-m}.$$

(b) The probability that either is found empty with m remaining in the other is just the sum of the probabilities for when the left is emptied and when the right is emptied. Since these are identical, this probability is

$$2\binom{2n-m}{n}(1/2)^{n+1}(1/2)^{n-m} = \binom{2n-m}{n}(1/2)^{2n-m}.$$

2.91 (a)

$$\begin{aligned} p_i &= P(\text{ A ruined}|i \text{ dollars }) \\ &= P(\text{ A ruined}|i \text{ dollars, loses next round })P(\text{ loses next round}) \\ &\quad +P(\text{ A ruined}|i \text{ dollars, wins next round })P(\text{wins next round}) \\ &= P(\text{ A ruined}|i-1 \text{ dollars })(1/2) + P(\text{ A ruined}|i+1 \text{ dollars })(1/2) \\ &= \frac{1}{2}p_{i-1} + \frac{1}{2}p_{i+1}. \end{aligned}$$

so

$$p_i/2 - p_{i-1}/2 = p_{i+1}/2 - p_i/2$$

and

$$p_i - p_{i-1} = p_{i+1} - p_i = \delta.$$

(b) The recursive equations are $p_1 = p_0 + \delta = 1 + \delta$, $p_2 = p_1 + \delta = 1 + 2\delta$, and so forth, so that $p_i = 1 + i\delta$. Then at stage $a + b$,

$$0 = p_{a+b} = 1 + (a+b)\delta$$

or $\delta = -1/(a+b)$. Then

$$p_a = 1 + a\delta = 1 - \frac{a}{a+b} = \frac{b}{a+b}.$$

(c)

$$\begin{aligned} p_i &= P(\text{ B ruined}|i \text{ dollars }) \\ &= P(\text{ B ruined}|i \text{ dollars, loses next round })P(\text{ loses next round}) \\ &\quad +P(\text{ B ruined}|i \text{ dollars, wins next round })P(\text{wins next round}) \\ &= P(\text{ B ruined}|i-1 \text{ dollars })(1/2) + P(\text{ B ruined}|i+1 \text{ dollars })(1/2) \\ &= \frac{1}{2}p_{i-1} + \frac{1}{2}p_{i+1}. \end{aligned}$$

so

$$p_i - p_{i-1} = p_{i+1} - p_i = \delta.$$

The recursive equations are $p_1 = p_0 + \delta = 1 + \delta$, $p_2 = p_1 + \delta = 1 + 2\delta$, and so forth, so that $p_i = 1 + i\delta$. Then at stage $a + b$,

$$0 = p_{a+b} = 1 + (a+b)\delta$$

or $\delta = -1/(a+b)$. Then

$$p_b = 1 + b\delta = 1 - \frac{b}{a+b} = \frac{a}{a+b}.$$

Since $p_a + p_b = 1$, there is no probability that the game will go on forever.

2.92 (a)

$$f(x) = f(x|Y=1)P(Y=1) + f(x|Y=2)P(Y=2) = p_1 f_1(x) + p_2 f_2(x).$$

(b)

$$\begin{aligned} E(X) &= \int_x x f(x)\,dx \\ &= \int_x x(p_1 f_1(x) + p_2 f_2(x))\,dx \\ &= p_1 \int_x x f_1(x)\,dx + p_2 \int_x x f_2(x)\,dx \\ &= p_1\mu_1 + p_2\mu_2. \end{aligned}$$

(c)

$$\begin{aligned} E(X^2) &= \int_x x^2 f(x)\,dx \\ &= \int_x x^2(p_1 f_1(x) + p_2 f_2(x))\,dx \\ &= p_1 \int_x x^2 f_1(x)\,dx + p_2 \int_x x^2 f_2(x)\,dx \\ &= p_1(\sigma_1^2 + \mu_1^2) + p_2(\sigma_2^2 + \mu_2^2). \end{aligned}$$

Then

$$\sigma^2 = p_1\sigma_1^2 + p_2\sigma_2^2 + p_1\mu_1^2 + p_2\mu_2^2 - (p_1\mu_1 + p_2\mu_2)^2.$$

2.93 (a) Using

$$E(X) = E[E(X|Y)],$$

and

$$E(X^2) = \text{Var}(X) + [E(X)]^2,$$

then

$$\begin{aligned} \text{Var}(X) &= E(X^2) - [E(X)]^2 \\ &= E\left[E(X^2|Y)\right] - [E(E(X|Y))]^2 \\ &= E\left[\text{Var}(X|Y) + [E(X|Y)]^2\right] - \left\{E\left[E(X|Y)^2\right] - \text{Var}\left[E(X|Y)\right]\right\} \\ &= E\left[\text{Var}(X|Y)\right] + \text{Var}\left[E(X|Y)\right]. \end{aligned}$$

(b) Let $1\{\cdot\}$ denote an indicator variable which takes on the value of 1 if the condition inside the braces is true. Since $E(1\{\cdot\}) = P(\cdot)$, the probability of the condition,

$$\begin{aligned} \text{Var}(X) &= E\left[\text{Var}(X|Y)\right] + \text{Var}\left[E(X|Y)\right] \\ &= E\left[\sigma_1^2 1\{Y=1\} + \sigma_2^2 1\{Y=2\}\right] + \text{Var}\left[\mu_1 1\{Y=1\} + \mu_2 1\{Y=2\}\right] \\ &= \sigma_1^2 p_1 + \sigma_2^2 p_2 + \mu_1^2 p_1(1-p_1) + \mu_2^2 p_2(1-p_2) + 2\mu_1\mu_2\text{Cov}(1\{Y=1\}1\{Y=2\}) \\ &= p_1\sigma_1^2 + p_2\sigma_2^2 + p_1\mu_1^2 + p_2\mu_2^2 - (\mu_1 p_1)^2 - (\mu_2 p_2)^2 + 2\mu_1\mu_2 p_1 p_2 \\ &= p_1\sigma_1^2 + p_2\sigma_2^2 + p_1\mu_1^2 + p_2\mu_2^2 - (p_1\mu_1 + p_2\mu_2)^2. \end{aligned}$$

2.94 (a)

$$\begin{aligned} \{N_t = n\} &\Leftrightarrow \text{ n events have occurred by time t} \\ &\Leftrightarrow \text{ the total time for n events is $\le t$ and the total time for $n+1$ events is $> t$} \\ &\Leftrightarrow \{X \le t, X + T_{n+1} > t\} \end{aligned}$$

(b)

$$\begin{aligned}
P(N_t = n) &= P(X \le t, X + T_{n+1} > t) \\
&= \int_0^t P(X \in [x, x+dx]) P(T_{n+1} > t - x)\, dx \\
&= \int_0^t \frac{\lambda^n x^{n-1} e^{-\lambda x}}{\Gamma(n)} e^{-\lambda(t-x)}\, dx \\
&= \int_0^t \frac{\lambda^n x^{n-1} e^{-\lambda t}}{\Gamma(n)}\, dx \\
&= \frac{\lambda^n e^{-\lambda t}}{\Gamma(n)} \frac{x^n}{n} \Big|_0^t \\
&= \frac{(\lambda t)^n e^{-\lambda t}}{n!},
\end{aligned}$$

since $\Gamma(n) = (n-1)!$. So N_t is Poisson(λt).

2.95 (a)

$$\begin{aligned}
P(X \le t) &= \int_0^t f_X(x)\, dx \\
&= \int_0^t \frac{\lambda^n x^{n-1} e^{-\lambda x}}{\Gamma(n)}\, dx \\
&= \frac{\lambda^n}{\Gamma(n)} \int_0^t x^{n-1} e^{-\lambda x}\, dx.
\end{aligned}$$

Also, $X \le t$ if and only if the number of events at time t is $\ge n$, so

$$P(X \le t) = P(N_t \ge n) = \sum_{i=n}^{\infty} \frac{e^{-\lambda t}(\lambda t)^i}{i!}.$$

(b) Using the table of Poisson cumulative probabilities,

$$P(N_{10} \ge 5) = 1 - \sum_{i=0}^{4} \frac{e^{-5} 5^i}{i!} = 1 - 0.440 = 0.560.$$

2.96 (a)

$$\begin{aligned}
\psi_X(t|N) &= E\left(e^{tX}|N\right) = E\left(e^{t(Y_1+Y_2+\ldots+Y_N)}|N\right) \\
&= E\left(e^{tY_1}\right) \cdots E\left(e^{tY_N}\right) \text{ because independent} \\
&= E\left(e^{tY_i}\right)^N \text{ because identically distributed} \\
&= \left(pe^{t(1)} + qe^{t(0)}\right)^N \\
&= \left(pe^t + q\right)^N.
\end{aligned}$$

(b)

$$\begin{aligned}
E\left(e^t X\right) &= E\left[E\left(e^{tX}\right)|N\right] \\
&= E\left[\left(pe^t+q\right)^N\right] \\
&= \sum_{N=0}^{\infty}\left(pe^t+q\right)^N \frac{e^{-\lambda}\lambda^N}{N!} \\
&= \frac{e^{-\lambda}}{e^{-\lambda(pe^t+q)}}\sum_{N=0}^{\infty}\frac{e^{-\lambda(pe^t+q)}\left[\lambda(pe^t+q)\right]^N}{N!} \\
&= \exp\{-\lambda+\lambda pe^t+\lambda(1-p)\}\times 1 \\
&= \exp\{-\lambda p(e^t-1)\}
\end{aligned}$$

Note that the sum reduces to 1 because it is the sum of the Poisson p.m.f. over the entire range. This unconditional m.g.f. is the m.g.f. of a Poisson(λp) random variable.

2.97 (a) Since the m.g.f. of X is

$$M_X(t) = E\left(e^t X\right) = E\left(e^{t\ln T}\right) = E\left(T^t\right),$$

then

$$E(T) = M_X(1) = e^{\mu(1)+\frac{\sigma^2}{2}(1)^2} = e^{\mu+\sigma^2/2}$$

and

$$E(T^2) = M_X(2) = e^{\mu(2)+\frac{\sigma^2}{2}(2)^2} = e^{2\mu+2\sigma^2}.$$

The variance is then

$$\text{Var}(T) = E(T^2) - E(T)^2 = e^{2\mu+2\sigma^2} - e^{2\mu+\sigma^2} = e^{2\mu+\sigma^2}\left(e^{\sigma^2}-1\right).$$

(b) Since

$$Y = \log_e T = \log_e 10 \times \log_1 0T = \log_e 10 \times X,$$

then

$$Y \sim N\left(log_e 10 \times 4, (\log_e 10)^2 \times (0.5)^2\right).$$

Then by the results of (a),

$$E(T) = \exp\{\mu_Y + \frac{1}{2}\sigma_Y^2\} = \exp\{4\log_e 10 + \frac{1}{2}(0.5\log_e 10)^2\} = 19,400.$$