PROBABILITY

STAT 220/230
PROBLEM SOLUTIONS
SPRING 2022 Edition

STAT 220/230 PROBLEM SOLUTIONS

By Chris Springer Revised by Jerry Lawless, Don McLeish and Cyntha Struthers

Spring 2022 Edition

Contents

2. SOLUTIONS TO CHAPTER 2 PROBLEMS

2.1 (a) Label the profs A, B, C and D.

$$S = \{AA, AB, AC, AD, BA, BB, BC, BD, CA, CB, CC, CD, DA, DB, DC, DD\}$$

(b) $1/4$

2.2 (a) A sample space is $\{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$. All outcomes are equally probable with probability $\frac{1}{8}$.

(b)

$$P\left(\text{two heads}\right) = P\left(\{HHT, HTH, THH\}\right) = \frac{3}{8}$$

(c)

$$P\left(\text{two consecutive tails}\right) = P\left(\{HTT, TTH\}\right) = \frac{2}{8}$$

2.3 (a) A suitable sample space is

$$\begin{gathered} S = \{(1,2), (1,3), (1,4), (1,5), (2,3), (2,4), (2,5), (3,4), (3,5), (4,5), \\ (2,1), (3,1), (4,1), (5,1), (3,2), (4,2), (5,2), (4,3), (5,3), (5,4)\} \end{gathered}$$

All outcomes are equally probable with probability $\frac{1}{20}$.

(b)

$$P\left(\text{both numbers are odd}\right) = P\left(\{(1,3), (1,5), (3,5), (3,1), (5,1), (5,3)\}\right) = \frac{6}{20}$$

(c)

$$\begin{aligned} &P\left(\text{two numbers are consecutive}\right) \\ &= P\left(\{(1,2), (2,3), (3,4), (4,5), (2,1), (3,2), (4,3), (5,4)\}\right) = \frac{8}{20} \end{aligned}$$

2.4 (a) Let XWYZ represent the outcome that X is in W's envelope, W is in X's envelope, Y is in Y's envelope and Z is in Z's envelope. Similarly let ZXYW represent the outcome that Z is in W's envelope, X is in X's envelope, Y is in Y's envelope and W is in Z's envelope. With this notation the set of all possible outcomes are the $4! = 24$ possible arrangements of the letters WXYZ as listed below:

$$\begin{aligned} S = \{&WXYZ, XWYZ, YWXZ, ZWXY, \\ &WXZY, XWZY, YWZX, ZWYX, \\ &WYXZ, XYWZ, YXWZ, ZXWY, \\ &WYZX, XYZW, YXZW, ZXYW, \\ &WZXY, XZWY, YZWX, ZYXW, \\ &WZYX, XZYW, YZXW, ZYWX\} \end{aligned}$$

(b) $A = \{WXYZ, WXZY, WYXZ, WYZX, WZXY, WZYX\}$
$B = \{XWZY, XYZW, XZWY, YWZX, YZWX, YZXW, ZWXY, ZYXW, ZYWX\}$
$C = \{WXZY, WYXZ, WZYX, ZXYW, YXWZ, XWYZ\}$
$D = \emptyset$
(c)

$$P(A) = \frac{6}{24}, \quad P(B) = \frac{9}{24}, \quad P(C) = \frac{6}{24}, \quad P(D) = P(\emptyset) = 0$$

2.5 (a) Let ijk represent the outcome "ball 1 is in box i, ball 2 is in box j and ball 3 is in box k" where $i, j, k = 1, 2, 3$. Then

$$\begin{aligned} S = \{&111, 222, 333, 112, 121, 211, 113, 131, 311, 221, 212, 122, \\ &223, 232, 322, 331, 313, 133, 332, 323, 233, 123, 132, 213, 231, 312, 321\} \end{aligned}$$

(b) Since $A = \{222, 333, 223, 232, 322, 332, 323, 233\}$, $P(A) = \frac{8}{27}$.
Since $B = \{333\}$, $P(B) = \frac{1}{27}$.
Since $C = \{123, 132, 213, 231, 312, 321\}$, $P(C) = \frac{6}{27} = \frac{2}{9}$
(c)

$$P(A) = \frac{(n-1)^3}{n^3}, \quad P(B) = \frac{(n-2)^3}{n^3}, \quad P(C) = \frac{n(n-1)(n-2)}{n^3}$$

(d)

$$P(A) = \frac{(n-1)^k}{n^k}, \quad P(B) = \frac{(n-2)^k}{n^k}, \quad P(C) = \frac{n(n-1)\cdots(n-k+1)}{n^k}$$

2.6 (a) 0.018 (b) 0.020 (c) $18/78 = 0.231$

2.7 (b) 0.978

3. SOLUTIONS TO CHAPTER 3 PROBLEMS

3.1

$$(a)\ \frac{(4)\,6^{(5)}}{7^{(6)}} \qquad (b)\ \frac{(5)\,5^{(4)}}{7^{(6)}} \qquad (c)\ \frac{(10)\,5^{(4)}}{7^{(6)}}$$

3.2

$$(a)\ \frac{6^{(4)}}{6^4} \qquad (b)\ \frac{\frac{4!}{2!2!}}{6^4} \qquad (c)\ \frac{\binom{6}{2}\frac{4!}{2!2!}}{6^4}$$

3.3

$$(a)\ \frac{7^{(5)}}{7^5} \qquad (b)\ \frac{7}{7^5}=\frac{1}{7^4} \qquad (c)\ \frac{(7)\,(6)\binom{5}{2}}{7^5} \qquad (d)\ 1-\frac{6^5}{7^5} \qquad (e)\ \frac{5^5}{7^5}$$

3.4 (a)

$$(i)\ \frac{(n-1)^k}{n^k} \qquad (ii)\ \frac{n^{(k)}}{n^k}$$

(b) All n^k outcomes are equally likely. That is, all n floors are equally likely to be selected, and each person's selection is unrelated to each other person's selection. Both assumptions are doubtful since people may be travelling together (e.g. same family) and the floors may not have equal traffic (e.g. more likely to use the stairs for going up 1 floor than for 10 floors);

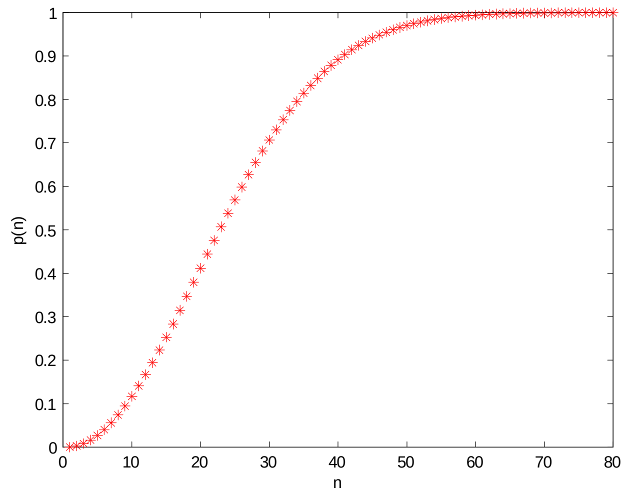
3.5

$$\frac{\binom{4}{2}\binom{12}{4}\binom{36}{7}}{\binom{52}{13}}$$

3.6

$$(a)\ \frac{1}{\frac{10!}{3!3!2!1!1!}}=\frac{1}{\frac{10!}{3!3!2!}} \qquad (b)\ \frac{\frac{8!}{3!3!}}{\frac{10!}{3!3!2!}} \qquad (c)\ 2\left(\frac{\frac{8!}{3!2!}}{\frac{10!}{3!3!2!}}\right)+\frac{\frac{8!}{3!3!}}{\frac{10!}{3!3!2!}}$$

$$(d)\ \frac{\frac{8!}{3!2!}}{\frac{10!}{3!3!2!}} \qquad (e)\ \frac{\frac{7!}{2!2!}}{\frac{10!}{3!3!2!}}$$



3.7

$$(a)\ \frac{\binom{10}{3}}{10^{(3)}} = \frac{1}{3!} \qquad (b)\ \frac{\binom{10}{3}}{10^{3}} = \frac{1}{3!}\frac{10^{(3)}}{10^{3}}$$

3.8 (a) The probability that every person has a different birthday is

$$\frac{365^{(n)}}{365^{n}}$$

(b)

$$p(n) = 1 - \frac{365^{(n)}}{365^{n}} \text{ for } n = 1, 2, \ldots, 365$$

(c) The plot of $p(n)$ is given below: $p(23) = 0.5073$ so there if there are 23 or more people in the room then the probability at least two people have the same birthday is greater than 0.5.

3.9

$$(a)\ \frac{1}{n} \qquad (b)\ \frac{2}{n}$$

3.10 (a) For nine tickets the sets of 3 tickets which form an arithmetic sequence are

$$\begin{aligned} A = \{&\{1,2,3\},\{2,3,4\},\{3,4,5,\},\{4,5,6\},\{5,6,7\},\{6,7,8\},\{7,8,9\}, \\ &\{1,3,5\},\{2,4,6\},\{3,5,7\},\{4,6,8\},\{5,7,9\}, \\ &\{1,4,7\},\{2,5,8\},\{3,6,9\}, \\ &\{1,5,9\}\} \end{aligned}$$

and

$$P(A) = \frac{7+5+3+1}{\binom{9}{3}}$$

(b) For $2n+1$ tickets

$$\begin{aligned} A = \{ & \{1,2,3\}, \{2,3,4\}, \ldots, \{2n-1, 2n, 2n+1\}, \\ & \{1,3,5\}, \{2,4,6\}, \ldots, \{2n-3, 2n-1, 2n+1\}, \\ & \vdots \\ & \{1, n, 2n-1\}, \{2, n+1, 2n\}, \{3, n+2, 2n+1\}, \\ & \{1, n+1, 2n+1\}\} \end{aligned}$$

and

$$\begin{aligned} P(A) &= \frac{(2n-1)+(2n-3)+\cdots+3+1}{\binom{2n+1}{3}} = \frac{1+3+\cdots+(2n-3)+(2n-1)}{\binom{2n+1}{3}} \\ &= \frac{1+2+3+4+\cdots+(2n-3)+(2n-2)+(2n-1)-[2+4+\cdots+(2n-2)]}{\binom{2n+1}{3}} \\ &= \frac{1+2+3+4+\cdots+(2n-3)+(2n-2)+(2n-1)-2\,[1+2+\cdots+(n-1)]}{\binom{2n+1}{3}} \\ &= \frac{\sum\limits_{i=1}^{2n-1} i - 2\sum\limits_{i=1}^{n-1} i}{\binom{2n+1}{3}} = \frac{\frac{(2n-1)(2n)}{2} - 2\left[\frac{(n-1)n}{2}\right]}{\binom{2n+1}{3}} = \frac{n^2}{\binom{2n+1}{3}} \end{aligned}$$

3.11 (a)

$$(i)\ \frac{\frac{4!}{2!2!}}{10^4} \qquad (ii)\ \frac{4!}{10^4} \qquad (b)\ \frac{10^{(4)}}{10^4}$$

3.12 (a)

$$\frac{\binom{6}{2}\binom{19}{3}}{\binom{25}{5}}$$

(b) Let $N =$ the unknown number of deer in the area. We know that the proportion of these deer which have been tagged is $6/N$. The proportion of deer in the sample of 5 deer who have been tagged is $2/5$. It seems reasonable to estimate the population proportion $6/N$ using the sample proportion $2/5$. Solving $6/N = 2/5$ gives $N = 15$ as an estimate of the number of deer in the area.

3.13

$$(a)\ \frac{\binom{6}{3}\binom{43}{3}}{\binom{49}{6}} \qquad (b)\ \frac{\binom{6}{0}\binom{43}{6}}{\binom{49}{6}} \qquad (c)\ \frac{\binom{6}{x}\binom{43}{6-x}}{\binom{49}{6}} \text{ for } x = 0, 1, \ldots, 6$$

3.14 (a)

$$5^{(3)} \cdot 26^3$$

(b)

$$(i)\ \ \frac{1}{26^3} \quad (ii)\ \ \frac{\binom{3}{2}5^2 \times 21}{26^3} \quad (iii)\ \ \frac{21 \times 21 \times 26}{26^3} \quad (iv)\ \ 1 - \frac{25^3}{26^3}$$

(c)

$$(i)\ \ \frac{1}{26^{(3)}} \quad (ii)\ \ \frac{\binom{3}{2}5^{(2)} \times 21}{26^{(3)}} \quad (iii)\ \ \frac{21^{(2)} \times 24}{26^{(3)}} \quad (iv)\ \ 1 - \frac{25^{(3)}}{26^{(3)}}$$

3.15 (a) The probability of at least one collision is one minus the probability of no collisions or

$$\begin{aligned} & 1 - \frac{M^{(n)}}{M^n} \\ &= 1 - \left(\frac{M}{M}\right)\left(\frac{M-1}{M}\right)\left(\frac{M-2}{M}\right)\cdots\left(\frac{M-(n-1)}{M}\right) \\ &= 1 - \left(1 - \frac{1}{M}\right)\left(1 - \frac{2}{M}\right)\cdots\left(1 - \frac{n-1}{M}\right) \end{aligned}$$

(b)

$$\begin{aligned} & 1 - \left(1 - \frac{1}{M}\right)\left(1 - \frac{2}{M}\right)\cdots\left(1 - \frac{n-1}{M}\right) \\ &\approx 1 - \left(e^{-1/M}\right)\left(e^{-2/M}\right)\cdots\left(e^{-(n-1)/M}\right) \\ &= 1 - \exp\left(-\frac{1}{M}\sum_{i=1}^{n-1} i\right) = 1 - e^{-\frac{1}{M}n(n-1)/2} \end{aligned}$$

(c) We want

$$1 - \frac{M^{(n)}}{M^n} \leq 0.5$$

or approximately

$$1 - e^{-\frac{1}{M}n(n-1)/2} \leq 0.5$$

Solving

$$1 - e^{-\frac{1}{M}n(n-1)/2} = 0.5$$

gives

$$n\,(n-1) = 2M \log\left(\frac{1}{1-0.5}\right) = 2M \log 2$$

or $n \approx \sqrt{2M \log 2}$. Therefore n should be less than $\sqrt{2M \log 2}$.

(d) $\sqrt{2M \log 2} = M^{1/2}\sqrt{2 \log 2} \approx 1.18M^{1/2} \approx M^{1/2}$ so if $M = 2^L$, then $\sqrt{2M \log 2} \approx 2^{L/2}$.

3.16

$$(a)\ \ 1 - \frac{\binom{48}{3}}{\binom{50}{3}} \qquad (b)\ \ 1 - \frac{\binom{45}{2}}{\binom{47}{2}} \qquad (c)\ \ \frac{\binom{48}{3}}{\binom{50}{5}}$$

3.17 Let $Q = \{\text{heads on quarter}\}$ and $D = \{\text{heads on dime}\}$. Then

$$\begin{aligned}
&P(\text{Both heads at same time}) \\
&= P(QD \cup \overline{Q}\,\overline{D}QD \cup \overline{Q}\,\overline{D}\,\overline{Q}\,\overline{D}\,QD \cup \cdots) \\
&= (0.6)(0.5) + (0.4)(0.5)(0.6)(0.5) + (0.4)(0.5)(0.4)(0.5)(0.6)(0.5) + \cdots \\
&= \frac{(0.6)(0.5)}{1 - (0.4)(0.5)} = 3/8 \ \text{ by the Geometric series}
\end{aligned}$$

3.18 Solution not provided.

4. SOLUTIONS TO CHAPTER 4 PROBLEMS

4.1

$$0.75,\ 0.6,\ 0.65,\ 0,\ 1,\ 0.35,\ 1$$

4.2

$$P(A) = 0.01 \quad P(B) = 0.72 \quad P(C) = (0.9)^3 \quad P(D) = (0.5)^3 \quad P(E) = (0.5)^2$$
$$P(B \cap E) = 0.12 \quad P(B \cup D) = 0.785 \quad P(B \cup D \cup E) = 0.886$$
$$P((A \cup B) \cap D) = 0.065 \quad P(A \cup (B \cap D)) = 0.07$$

4.3

$$\begin{aligned} P(A|\overline{B}) &= \frac{P(A \cap \overline{B})}{P(\overline{B})} = \frac{P(A) - P(A \cap B)}{1 - P(B)} \\ &= \frac{P(A) - P(B)P(A|B)}{1 - P(B)} = \frac{0.3 - (0.4)(0.5)}{1 - 0.4} = \frac{1}{6} \end{aligned}$$

4.4 (a) $(0.7)^8$ (b) $(0.9)^8$ (c) $(0.6)^8$ (d) $1 - \left[(0.7)^8 + (0.9)^8 - (0.6)^8\right]$

4.5 Since A and B are independent events $P(A \cap B) = (0.3)(0.2) = 0.06$.
Therefore $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.3 + 0.2 - 0.06 = 0.44$.

4.6 Since $E = A \cup B$, $F = A \cap B$ and E and F are independent events we have

$$P(E \cap F) = P((A \cup B) \cap (A \cap B)) = P(A \cap B)$$

and

$$P(E \cap F) = P(E)P(F) = P(A \cup B)P(A \cap B)$$

which implies

$$P(A \cup B)P(A \cap B) = P(A \cap B)$$

or

$$P(A \cap B)[1 - P(A \cup B)] = 0$$

This statement holds only if $P(A \cap B) = 0$ or $P(A \cup B) = 1$.
But $1 = P(A \cup B) = 1 - P\left(\bar{A} \cap \bar{B}\right)$ which implies $P\left(\bar{A} \cap \bar{B}\right) = 0$. Therefore either $P(A \cap B) = 0$ or $P\left(\bar{A} \cap \bar{B}\right) = 0$ as required.

4.7 A necessary and sufficient condition is

$$\frac{f}{F} = \frac{m}{M}$$

4.8 Note that A and B are independent events since you are given that A and $\bar{B}$ are independent events (see solution to 4.3.3).

$$\begin{aligned}
P(A \cup B) &= 0.15 + (1 - 0.3) - (0.15)(1 - 0.3) \\
&= 0.745 \\
P(B \cap D|A) &= \frac{P(B \cap D \cap A)}{P(A)} = \frac{P(B \cap A)}{P(A)} \quad \text{since } A \subset D \\
&= \frac{P(B)\,P(A)}{P(A)} = P(B) \\
&= 0.7 \\
P\left(B \cup \bar{D}\right) &= P\left(\overline{\bar{B} \cap D}\right) = 1 - P\left(\bar{B} \cap D\right) \\
&= 1 - P\left(D|\bar{B}\right) P\left(\bar{B}\right) \\
&= 1 - (0.8)(0.3) \\
&= 0.76 \\
P\left(C|\bar{A} \cup B\right) &= \frac{P\left(C \cap \left(\bar{A} \cup B\right)\right)}{P\left(\bar{A} \cup B\right)} = \frac{P\left(\left(C \cap \bar{A}\right) \cup (C \cap B)\right)}{P\left(\bar{A} \cup B\right)} \\
&= \frac{P(C \cup (C \cap B))}{P\left(\bar{A} \cup B\right)} \\
&= \frac{P(C)}{P\left(\bar{A} \cup B\right)} \\
&= \frac{P(C)}{P\left(\bar{A}\right) + P(B) - P\left(\bar{A}\right) P(B)} \\
&= \frac{0.1}{0.85 + 0.7 - (0.85)(0.7)} \\
&= 0.1047
\end{aligned}$$

4.9 Let C_i be the event that component i is working, $i = 1, 2, 3, 4$. Using Rule 4b and the fact that the components function independently we have

$$
\begin{aligned}
& P\,(\text{system is working properly}) \\
& = P\,((C_1 \cap C_2) \cup (C_1 \cap C_4) \cup (C_3 \cap C_4)) \\
& = P\,(C_1 \cap C_2) + P\,(C_1 \cap C_4) + P\,(C_3 \cap C_4) \\
& - P\,((C_1 \cap C_2) \cap (C_1 \cap C_4)) - P\,((C_1 \cap C_2) \cap (C_3 \cap C_4)) \\
& - P\,((C_1 \cap C_4) \cap (C_3 \cap C_4)) + P\,((C_1 \cap C_2) \cap (C_1 \cap C_4) \cap (C_3 \cap C_4)) \\
& = P\,(C_1 \cap C_2) + P\,(C_1 \cap C_4) + P\,(C_3 \cap C_4) \\
& - P\,(C_1 \cap C_2 \cap C_4) - P\,(C_1 \cap C_2 \cap C_3 \cap C_4) \\
& - P\,(C_1 \cap C_3 \cap C_4) + P\,(C_1 \cap C_2 \cap C_3 \cap C_4) \\
& = P\,(C_1)\,P\,(C_2) + P\,(C_1)\,P\,(C_4) + P\,(C_3)\,P\,(C_4) \\
& - P\,(C_1)\,P\,(C_2)\,P\,(C_4) - P\,(C_1)\,P\,(C_3)\,P\,(C_4) \\
& = (0.9)\,(0.8) + (0.9)\,(0.6) + (0.7)\,(0.6) - (0.9)\,(0.8)\,(0.6) - (0.9)\,(0.7)\,(0.6) \\
& = 0.87
\end{aligned}
$$

4.10

$$(a) \quad \binom{5}{3}(0.7)^3(0.3)^2 \qquad (b) \quad (0.7)^4(0.3)^1 \qquad (c) \quad \binom{4}{2}(0.7)^3(0.3)^2$$

4.11 (a) Since students answer independently

$$
\begin{aligned}
P\,(\text{all 3 student get the correct answer}) &= P\,(A \cap B \cap C) \\
&= P\,(A)\,P\,(B)\,P\,(C) \\
&= (0.9)\,(0.7)\,(0.4) \\
&= 0.252
\end{aligned}
$$

(b)

$$
\begin{aligned}
& P\,(\text{exactly two students get the correct answer}) \\
& = P\left(A \cap B \cap \bar{C}\right) + P\left(A \cap \bar{B} \cap C\right) + P\left(\bar{A} \cap B \cap C\right) \\
& = P\,(A)\,P\,(B)\,P\left(\bar{C}\right) + P\,(A)\,P\left(\bar{B}\right)P\,(C) + P\left(\bar{A}\right)P\,(B)\,P\,(C) \\
& = (0.9)\,(0.7)\,(0.6) + (0.9)\,(0.3)\,(0.4) + (0.1)\,(0.7)\,(0.4) \\
& = 0.514
\end{aligned}
$$

(c)

$$\begin{aligned}
&P\left(C \text{ is wrong} \mid 2 \text{ students correct}\right)\\
&= \frac{P\left(C \text{ is wrong and 2 students correct}\right)}{P\left(2 \text{ students correct}\right)}\\
&= \frac{P\left(A \cap B \cap \bar{C}\right)}{0.514} = \frac{0.378}{0.514} = 0.7354
\end{aligned}$$

4.12 (a) 0.48

(b) $(0.6)(0.5) + (0.48)(0.5) = 0.54$

(c) $\frac{(0.6)(0.5)}{0.54} = \frac{5}{9}$

4.13 (a) $(0.05)(0.3) + (0.04)(0.6) + (0.02)(0.1) = 0.041$

(b) $1 - (1 - 0.041)^{10} = 0.342$

4.14 (a) $0.1225, \;\; 0.175$

(b) 0.395

4.15 The probability of C for the first $n-1$ trials and then A occurs on the n'th trial is $r^{n-1}p$. Add over all $n \geq 1$ using Geometric series.

4.16

4.17 (a) The probability all three positions show a flower is

$$\left(\frac{2}{10}\right)\left(\frac{6}{10}\right)\left(\frac{2}{10}\right) = 0.024$$

(b) Suppose there are $m \geq 1$ flowers on wheel 1, $n \geq 1$ flowers on wheel 2, and $10 - m - n \geq 1$ on wheel 3. The probability all three positions show a flower is

$$\left(\frac{m}{10}\right)\left(\frac{n}{10}\right)\left(\frac{10-m-n}{10}\right) = \frac{mn\left(10-m-n\right)}{10^3}$$

Let $f(n,m) = mn(10 - m - n)$. We want to minimize $f(n,m)$ subject to the restrictions $m \geq 1$, $n \geq 1$, and $10 - m - n \geq 1$. For each value of m, $f(n,m) = mn(10 - m - n)$ is a quadratic function of n which is minimized for $n = 1$ or $n = 9 - m$. Now $f(1,m) = f(9 - m, m) = m(9 - m)$ which is minimized for $m = 1$ or $m = 8$. Now $f(1,1) = f(1,8) = f(8,1) = 8$ and the values of $(m, n, 10 - m - 9)$ which minimize the probability all three positions show a flower are $(1,1,8)$, $(1,8,1)$, and $(8,1,1)$.

4.18 (a) $0.010 + 0.016 + 0.040 = 0.066$

(b) $0.066 + 0.010 = 0.076$

(c) $0.185 + 0.683 + 0.016 + 0.04 = 0.924$ or $1 - 0.076 = 0.924$

(d)

$$\frac{0.185 + 0.683}{0.066 + 0.185 + 0.683} = 0.929$$

(e) The events are not independent since

$$\begin{aligned} 0.010 &= P\left(\text{unemployed} \cap \text{no certificate, diploma or degree}\right) \\ &\neq P\left(\text{unemployed}\right) P\left(\text{no certificate, diploma or degree}\right) = (0.066)\,(0.076) \end{aligned}$$

4.19

$$\begin{aligned} (a)\ P\left(\text{Yes}\right) &= P\left(\text{Yes}\,|B\right) P\left(B\right) + P\left(\text{Yes}\,|A\right) P\left(A\right) \\ &= p\left(\frac{80}{100}\right) + \left(\frac{2}{12}\right)\left(\frac{20}{100}\right) = \frac{4p}{5} + \frac{1}{30} \\ (b) \quad p &= \frac{(30x/n) - 1}{24} \\ (c)\ P\left(B|\text{Yes}\right) &= \frac{P\left(\text{Yes}\,|B\right) P\left(B\right)}{P\left(\text{Yes}\right)} = \frac{\frac{4p}{5}}{\frac{4p}{5} + \frac{1}{30}} = \frac{24p}{1 + 24p} \end{aligned}$$

4.20 $0.9,\ 0.061,\ 0.078$

4.21 (a)

$$\begin{aligned} P\left(A\right) &= P\left(\text{Message contains the word Viagra}\right) \\ &= P\left(A|\text{Spam}\right) P\left(\text{Spam}\right) + P\left(A|\text{Not Spam}\right) P\left(\text{Not Spam}\right) \\ &= (0.2)\,(0.5) + (0.001)\,(0.5) = 0.1005 \end{aligned}$$

(b)

$$\begin{aligned} P\left(\text{Spam}|A\right) &= \frac{P\left(A|\text{Spam}\right) P\left(\text{Spam}\right)}{P\left(A\right)} = \frac{(0.2)\,(0.5)}{0.1005} = 0.995 \\ P\left(\text{Not Spam}|A\right) &= 1 - 0.995 = 0.005 \end{aligned}$$

(c)

$$P\left(\text{declared as Spam}|\text{Spam}\right) = P\left(A|\text{Spam}\right) = 0.2$$

4.22 (a)

$$\begin{aligned}
P(A_1A_2A_3) &= P(A_1A_2A_3|\text{Spam})\,P(\text{Spam}) + P(A_1A_2A_3|\text{Not Spam})\,P(\text{Not Spam})\\
&= P(A_1|\text{Spam})\,P(A_2|\text{Spam})\,P(A_3|\text{Spam})\,P(\text{Spam})\\
&+ P(A_1|\text{Not Spam})\,P(A_2|\text{Not Spam})\,P(A_3|\text{Not Spam})\,P(\text{Not Spam})\\
&= (0.2)\,(0.1)\,(0.1)\,(0.5) + (0.005)\,(0.004)\,(0.005)\,(0.5)\\
&= 0.00100005
\end{aligned}$$

(b)

$$P(\text{Spam}|A_1A_2A_3) = \frac{P(A_1A_2A_3|\text{Spam})\,P(\text{Spam})}{P(A_1A_2A_3)} = \frac{(0.2)\,(0.1)\,(0.1)\,(0.5)}{0.00100005} = 0.99995$$

(c)

$$\begin{aligned}
P\left(A_1A_2\bar{A}_3\right) &= P\left(A_1A_2\bar{A}_3|\text{Spam}\right)P(\text{Spam}) + P\left(A_1A_2\bar{A}_3|\text{Not Spam}\right)P(\text{Not Spam})\\
&= (0.2)\,(0.1)\,(0.9)\,(0.5) + (0.005)\,(0.004)\,(0.995)\,(0.5)\\
&= 0.00900995
\end{aligned}$$

$$P\left(\text{Spam}|A_1A_2\bar{A}_3\right) = \frac{P\left(A_1A_2\bar{A}_3|\text{Spam}\right)P(\text{Spam})}{P\left(A_1A_2\bar{A}_3\right)} = \frac{(0.2)\,(0.1)\,(0.9)\,(0.5)}{0.00900995} = 0.99889$$

(d)

$$\begin{aligned}
P(\text{declared as Spam}|\text{Spam}) &= P(A_1 \cup A_2 \cup A_3|\text{Spam})\\
&= P\left(\overline{\bar{A}_1 \cap \bar{A}_2 \cap \bar{A}_3}|\text{Spam}\right) = 1 - P\left(\bar{A}_1 \cap \bar{A}_2 \cap \bar{A}_3|\text{Spam}\right)\\
&= 1 - P\left(\bar{A}_1|\text{Spam}\right)P\left(\bar{A}_2|\text{Spam}\right)P\left(\bar{A}_3|\text{Spam}\right)\\
&= 1 - (0.8)\,(0.9)\,(0.9) = 0.352
\end{aligned}$$

which is larger than 0.2.

(e)

$$\begin{aligned}
&P(\text{declared as Spam})\\
&= P(\text{declared as Spam}|\text{Spam})\,P(\text{Spam}) + P(\text{declared as Spam}|\text{Not Spam})\,P(\text{Not Spam})\\
&= [1-(0.8)\,(0.9)\,(0.9)]\,(0.5) + [1-(0.995)\,(0.996)\,(0.995)]\,(0.5)\\
&= 0.18296755
\end{aligned}$$

$$\begin{aligned}
P(\text{Spam}|\text{declared as Spam}) &= \frac{P(\text{declared as Spam}|\text{Spam})\,P(\text{Spam})}{P(\text{declared as Spam})}\\
&= \frac{[1-(0.8)\,(0.9)\,(0.9)]\,(0.5)}{0.18296755} = 0.961919
\end{aligned}$$

(f)

$$\begin{aligned}
P\left(A_1|\text{declared as Spam}\right) &= \frac{P\left(\text{declared as Spam}|A_1\right)P\left(A_1\right)}{P\left(\text{declared as Spam}\right)} \\
&= \frac{(1)\,P\left(A_1\right)}{P\left(\text{declared as Spam}\right)} \\
&= \frac{P\left(A_1|\text{Spam}\right)P\left(\text{Spam}\right)+P\left(A_1|\text{Not Spam}\right)P\left(\text{Not Spam}\right)}{P\left(\text{declared as Spam}\right)} \\
&= \frac{(0.2)\,(0.5)+(0.005)\,(0.5)}{0.18296755} \\
&= 0.560209
\end{aligned}$$

4.23 (a) Note that

$$P\left(\text{feature present}|F\right) = rP\left(\text{feature present}|\bar{F}\right) = 0.02r$$

Therefore

$$\begin{aligned}
P\left(F|\text{feature present}\right) &= \frac{P\left(\text{feature present}|F\right)P\left(F\right)}{P\left(\text{feature present}\right)} \\
&= \frac{P\left(\text{feature present}|F\right)P\left(F\right)}{P\left(\text{feature present}|F\right)P\left(F\right)+P\left(\text{feature present}|\bar{F}\right)P\left(\bar{F}\right)} \\
&= \frac{(0.02r)\,(0.0005)}{(0.02r)\,(0.0005)+(0.02)\,(0.9995)} \\
&= \frac{r}{r+1999}
\end{aligned}$$

For $r = 10$, 30 and 50 we obtain $0.005,\ \ 0.0148,\ \ 0.0244$ respectively.

(b)

$$P\left(\text{flagged}\right) = P\left(\text{feature present}\right) = (0.02r)\,(0.0005)+(0.02)\,(0.9995)$$

If $r = 50$ we have

$$P\left(\text{flagged}\right) = (0.02)\,(50)\,(0.0005)+(0.02)\,(0.9995) = 0.02049$$

or 2.049% of transactions will be flagged.

5. SOLUTIONS TO CHAPTER 5 PROBLEMS

5.1 (a)

$$1 = \sum_{x \in A} f(x) = 0.1c + 0.2c + 0.5c + c + 0.2c = 2c \text{ so } c = 1/2 = 0.5$$

$$P(X > 2) = P(X = 3) + P(X = 4) = c + 0.2c = 1.2c = 0.6$$

(b)

x	0	1	2	3	4
$F(x) = P(X \leq x)$	0.05	0.15	0.4	0.9	1

5.2 (a) $4k^2 = 1$ so $k = \frac{1}{2} = 0.5$ $P(2 < X \leq 4) = P(X \leq 4) - P(X \leq 2) = 0.5 - 0.2 = 0.3$

(b)

x	1	2	3	4	5	Total
$f(x)$	0.05	0.15	0.05	0.25	0.5	1

5.3 (a)

$$P(X = 5) = P(X \leq 5) - P(X \leq 4) = \left(1 - 2^{-5}\right) - \left(1 - 2^{-4}\right) = \frac{1}{32}$$

$$P(X \geq 5) = 1 - P(X \leq 4) = 1 - \left(1 - 2^{-4}\right) = \frac{1}{16}$$

(b) For $x = 1, 2, \ldots$

$$\begin{aligned} f(x) &= P(X = x) = P(X \leq x) - P(X \leq x - 1) \\ &= F(x) - F(x-1) = \left(1 - 2^{-x}\right) - \left(1 - 2^{-x+1}\right) \\ &= 2^{-x}(2 - 1) = 2^{-x} \end{aligned}$$

5.4 (a) (i)

x	1	2	3	4	5	6	7	8	9	Total
$f_X(x)$	$\frac{2}{10^{(2)}}$	$\frac{4}{10^{(2)}}$	$\frac{6}{10^{(2)}}$	$\frac{8}{10^{(2)}}$	$\frac{10}{10^{(2)}}$	$\frac{12}{10^{(2)}}$	$\frac{14}{10^{(2)}}$	$\frac{16}{10^{(2)}}$	$\frac{18}{10^{(2)}}$	1

or

$$f_X(x) = \frac{2x}{10^{(2)}}, \quad x = 1, 2, \ldots, 9$$

(ii)

y	1	2	3	4	5	6	7	8	9
$f_Y(y)$	$\frac{2}{10^{(2)}}$	$\frac{2}{10^{(2)}}$	$\frac{4}{10^{(2)}}$	$\frac{4}{10^{(2)}}$	$\frac{6}{10^{(2)}}$	$\frac{6}{10^{(2)}}$	$\frac{8}{10^{(2)}}$	$\frac{8}{10^{(2)}}$	$\frac{10}{10^{(2)}}$

y	10	11	12	13	14	15	16	17	Total
$f_Y(y)$	$\frac{8}{10^{(2)}}$	$\frac{8}{10^{(2)}}$	$\frac{6}{10^{(2)}}$	$\frac{6}{10^{(2)}}$	$\frac{4}{10^{(2)}}$	$\frac{4}{10^{(2)}}$	$\frac{2}{10^{(2)}}$	$\frac{2}{10^{(2)}}$	1

or

$$f_Y(y) = \begin{cases} \frac{10-|y-9|}{10^{(2)}}, & y = 1, 3, 5 \ldots, 17 \\ \frac{9-|y-9|}{10^{(2)}}, & y = 2, 4, 6 \ldots, 18 \end{cases}$$

(b) (i)

x	0	1	2	3	4	5	6	7	8	9	Total
$f_X(x)$	$\frac{1}{10^2}$	$\frac{3}{10^2}$	$\frac{5}{10^2}$	$\frac{7}{10^2}$	$\frac{9}{10^2}$	$\frac{11}{10^2}$	$\frac{13}{10^2}$	$\frac{15}{10^2}$	$\frac{17}{10^2}$	$\frac{19}{10^2}$	1

or

$$f_X(x) = \frac{2x+1}{10^2}, \quad x = 0, 1, \ldots, 9$$

(ii)

y	0	1	2	3	4	5	6	7	8
$f_Y(y)$	$\frac{1}{10^2}$	$\frac{2}{10^2}$	$\frac{3}{10^2}$	$\frac{4}{10^2}$	$\frac{5}{10^2}$	$\frac{6}{10^2}$	$\frac{7}{10^2}$	$\frac{8}{10^2}$	$\frac{9}{10^2}$

y	9	10	11	12	13	14	15	16	17	18	Total
$f_Y(y)$	$\frac{10}{10^2}$	$\frac{9}{10^2}$	$\frac{8}{10^2}$	$\frac{7}{10^2}$	$\frac{6}{10^2}$	$\frac{5}{10^2}$	$\frac{4}{10^2}$	$\frac{3}{10^2}$	$\frac{2}{10^2}$	$\frac{1}{10^2}$	1

or

$$f_Y(y) = \frac{10 - |y-9|}{10^2}, \quad y = 0, 1, \ldots, 18$$

5.5 (a)

$$\sum_{x=0}^{\infty} p\,(1-p)^x = \frac{p}{1-(1-p)} = 1 \text{ by the Geometric series}$$

(b) For $x = 0, 1, \ldots$

$$P(X < x) = 1 - P(X \ge x) = 1 - \sum_{t=x}^{\infty} p(1-p)^t = 1 - \frac{p(1-p)^x}{1-(1-p)}$$
$$= 1 - (1-p)^x \quad \text{by the Geometric series}$$

(c)

$$\begin{aligned} P(X \text{ is odd}) &= P(X=1) + P(X=3) + P(X=5) + \cdots \\ &= p(1-p) + p(1-p)^3 + p(1-p)^5 + \cdots \\ &= \frac{p(1-p)}{1-(1-p)^2} \quad \text{by the Geometric series} \end{aligned}$$

(d)

$$\begin{aligned} P(X \text{ is divisible by } 3) &= P(X=0) + P(X=3) + P(X=6) + \cdots \\ &= p + p(1-p)^3 + p(1-p)^6 + \cdots \\ &= \frac{p}{1-(1-p)^3} \quad \text{by the Geometric series} \end{aligned}$$

(e)

$$\begin{aligned} P(R=0) &= P(X=0) + P(X=4) + P(X=8) + \cdots \\ &= p + p(1-p)^4 + p(1-p)^8 + \cdots \\ &= \frac{p}{1-(1-p)^4} \quad \text{by the Geometric series} \end{aligned}$$

The cases $r = 1, 2, 3$ can be done similarly to obtain

$$P(R=r) = \frac{p(1-p)^r}{1-(1-p)^4}, \quad r = 0, 1, 2, 3$$

5.6 (a) The probability function of $X =$ number of defective chips in sample of twenty is

$$P(X=x) = \frac{\binom{50}{x}\binom{950}{20-x}}{\binom{1000}{20}}, \quad x = 0, 1, \ldots, 20$$

(b)

$$P(X \ge 2) = \sum_{x=2}^{20} \frac{\binom{50}{x}\binom{950}{20-x}}{\binom{1000}{20}} = 1 - \sum_{x=0}^{1} \frac{\binom{50}{x}\binom{950}{20-x}}{\binom{1000}{20}}$$

(c) Since $n = 20$ draws is small compared to $N = 1000 =$ total number of items this probability can be approximated using the Binomial approximation

$$\begin{aligned} P(X \geq 2) &\approx 1 - \sum_{x=0}^{1} \binom{20}{x} (0.05)^x (0.95)^{20-x} \\ &= 1 - (0.95)^{20} - 20 (0.05) (0.95)^{19} = 0.264 \end{aligned}$$

5.7

$$(a)\ \frac{4}{15} \qquad (b)\ \frac{\binom{74}{y}\binom{76}{12-y}}{\binom{150}{12}} \qquad (c)\ 0.0176$$

5.8 Let $X =$ the number of corrupted bits. Then $X \sim Binomial\left(10^4, 10^{-5}\right)$.

(a)

$$P(X = 0) = \binom{10^4}{0} \left(10^{-5}\right)^0 \left(1 - 10^{-5}\right)^{10^4} = \left(1 - 10^{-5}\right)^{10^4} = 0.904837$$

(b)

$$\begin{aligned} P(X \leq 1) &= \left(1 - 10^{-5}\right)^{10^4} + \binom{10^4}{1} \left(10^{-5}\right)^1 \left(1 - 10^{-5}\right)^{10^4 - 1} \\ &= \left(1 - 10^{-5}\right)^{10^4} + 10^4 \left(10^{-5}\right) \left(1 - 10^{-5}\right)^{10^4 - 1} = 0.9953216 \end{aligned}$$

(c) Since $n = 10^4$ is large and $p = 10^{-5}$ is small, the Poisson approximation to the Binomial may be used with $\mu = \left(10^4\right)\left(10^{-5}\right) = 0.1$.

$$\begin{aligned} P(X = 0) &\approx \frac{(0.1)^0 e^{-0.1}}{0!} = e^{-0.1} = 0.9048374 \\ P(X \leq 1) &\approx e^{-0.1} + \frac{(0.1)^1 e^{-0.1}}{1!} = 1.1 e^{-0.1} = 0.9953212 \end{aligned}$$

5.9 (a)

$$1 - \frac{\binom{500}{0}\binom{499500}{10}}{\binom{500000}{10}} = 0.009955209$$

$$\begin{aligned} &\approx 1 - \binom{10}{0} (0.001)^0 (0.999)^{10} \\ &= 1 - (0.999)^{10} \\ &\approx 0.009955120 \end{aligned}$$

The Binomial approximation to the Hypergeometric is valid since the number of draws, $n = 10$, is small relative to the the total number of items, $N = 500000$.

(b)

$$1 - \frac{\binom{500}{0}\binom{499500}{2000}}{\binom{500000}{2000}} = 0.865342$$

$$\approx 1 - \binom{2000}{0}(0.001)^0(0.999)^{2000} = 1 - (0.999)^{2000} = 0.8648001$$

$$\approx 1 - \frac{e^{-2}2^0}{0!} = 1 - e^{-2} = 0.864665$$

The Binomial approximation to the Hypergeometric is valid since the number of draws, $n = 2000$, is still small relative to the the total number of items, $N = 500000$. The Poisson approximation to the Binomial is valid since $n = 2000$ is large, and $p = 0.001$ is small.

5.10 (a)

$$\frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} = \frac{\frac{r^{(x)}}{x!}\frac{(N-r)^{(n-x)}}{(n-x)!}}{\frac{N^{(n)}}{n!}} = \frac{n!}{x!\,(n-x)!}\frac{r^{(x)}\,(N-r)^{(n-x)}}{N^{(n)}} = \binom{n}{x}\frac{r^{(x)}\,(N-r)^{(n-x)}}{N^{(n)}}$$

Substituting $r = pN$ we obtain

$$\begin{aligned}\frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} &= \binom{n}{x}\frac{(pN)^{(x)}\,(N-pN)^{(n-x)}}{N^{(n)}}\\ &= \binom{n}{x}\frac{(pN)^{(x)}\,[(1-p)\,N]^{(n-x)}}{N^{(x)}\,(N-x)^{(n-x)}}\end{aligned}$$

(b) Since

$$\begin{aligned}\frac{(pN)^{(x)}}{N^{(x)}} &= \frac{(pN)\,(pN-1)\,(pN-2)\cdots(pN-x+1)}{N\,(N-1)\,(N-2)\cdots(N-x+1)}\\ &= p^x\left(\frac{1-\frac{1}{pN}}{1-\frac{1}{N}}\right)\left(\frac{1-\frac{2}{pN}}{1-\frac{2}{N}}\right)\cdots\left(\frac{1-\frac{x-1}{pN}}{1-\frac{x-1}{N}}\right)\end{aligned}$$

then

$$\lim_{N\to\infty}\frac{(pN)^{(x)}}{N^{(x)}} = p^x$$

Similarly

$$\frac{[(1-p)\,N]^{(n-x)}}{(N-x)^{(n-x)}} = (1-p)^{n-x}$$

and thus

$$\lim_{N\to\infty}\binom{n}{x}\frac{(pN)^{(x)}\,[(1-p)\,N]^{(n-x)}}{N^{(x)}\,(N-x)^{(n-x)}} = \binom{n}{x}p^x\,(1-p)^{n-x}$$

(c) This result justifies the Binomial approximation to the Hypergeometric when N gets large but $p = r/N$, the proportion of type 1 items, and n, the number of items drawn, are held fixed.

5.11

$$P(X=x)=\frac{\binom{35}{x}\binom{70}{7}}{\binom{105}{x+7}}\frac{63}{105-(x+7)}\ \text{ for }\ x=0,1,\ldots,35$$

5.12 (a) $(0.99)^{20}=0.8179;\ 1-(0.99)^{20}=0.1821$

(b) $(0.99)^{30}=0.7397$

(c)

$$\sum_{y=96}^{\infty}\binom{y+3}{3}(0.01)^4(0.99)^y=1-\sum_{y=0}^{95}\binom{y+3}{3}(0.01)^4(0.99)^y$$

5.13 (a)

$$\left(\frac{8}{9}\right)^{15}=0.1709$$

(b)

$$\left(\frac{8}{9}\right)^{24}\left(\frac{1}{9}\right)=0.006578$$

(c)

$$\left(\frac{8}{9}\right)^{60}+60\left(\frac{1}{9}\right)\left(\frac{8}{9}\right)^{59}=0.007249$$

5.14 (a)

$$P(X\geq x)=p(1-p)^x+p(1-p)^{x+1}+p(1-p)^{x+2}+\cdots=\frac{p(1-p)^x}{1-(1-p)}=(1-p)^x\ \text{ for } x=0,1,\ldots$$

(b) $X=0$

5.15 (b) Let $X=$ number of requests in a one second interval. Then $X\sim Poisson(2)$.

$$P(X\geq 3)=1-P(X\leq 2)=1-\sum_{x=0}^{2}\frac{2^x e^{-2}}{x!}$$
$$=1-5e^{-2}=0.3233$$

(c) Let $Y=$ number of requests in a one minute $=60$ second interval. Then $Y\sim Poisson(2\times 60)$.

$$P(Y>125)=\sum_{y=126}^{\infty}\frac{(120)^y e^{-120}}{y!}=1-\sum_{y=0}^{125}\frac{(120)^y e^{-120}}{y!}$$
$$=0.3038\quad\text{(calculated using R)}$$

5.16 0.9989

5.17

$$(a)\ \ 0.0758 \quad (b)\ \ 0.0488 \quad (c)\ \ \binom{10}{y}(e^{-10\lambda})^y(1-e^{-10\lambda})^{10-y} \quad (d)\ \ \lambda = 0.12$$

5.18 (a) 0.0769 (b) $0.2019;$ 0.4751

5.19 (a)

$$\binom{7}{5}(0.8)^5(0.2)^2 = 0.2753$$

(b)

$$\binom{6}{2}(0.8)^5(0.2)^2 = 0.1966$$

(c) Let $X =$ number of power failures in one month. Since power failures occur independently of each other at a uniform rate through the months of the year, with little chance of 2 or more occurring simultaneously then $X \sim Poisson\,(\mu)$. To determine μ we note that

$$0.8 = P\,(X=0) = \frac{\mu^0 e^{-\mu}}{0!} = e^{-\mu} \text{ or } \mu = -\ln(0.8)$$

Therefore

$$\begin{aligned} P\,(> 1 \text{ power failure in one month}) &= P\,(X>1) = 1 - P\,(X \le 1) \\ &= 1 - P\,(X=0) - P\,(X=1) \\ &= 1 - 0.8 - \frac{[-\ln(0.8)]^1\, e^{\ln(0.8)}}{1!} = 0.02149 \end{aligned}$$

5.20 (a) Let $X =$ number of spruce budworms in a one hectare plot. Then $X \sim Poisson\,(\lambda)$ since spruce budworms are distributed through a forest according to a Poisson process so that the average is λ per hectare.

$$\begin{aligned} &P\,(\text{a one hectare plots contains at least } k \text{ spruce budworms}) \\ &= P\,(X \ge k) = 1 - P\,(X \le k-1) \\ &= 1 - \sum_{x=0}^{k-1} \frac{\lambda^x e^{-\lambda}}{x!} \end{aligned}$$

Now

$$\begin{aligned} &P\,(\text{at least 1 of } n \text{ one hectare plots contains at least } k \text{ spruce budworms}) \\ &= 1 - P\,(\text{none of } n \text{ one hectare plots contain at least } k \text{ spruce budworms}) \\ &= 1 - \binom{n}{0}\left[1 - \sum_{x=0}^{k-1}\frac{\lambda^x e^{-\lambda}}{x!}\right]^0 \left[\sum_{x=0}^{k-1}\frac{\lambda^x e^{-\lambda}}{x!}\right]^n \\ &= 1 - \left[\sum_{x=0}^{k-1}\frac{\lambda^x e^{-\lambda}}{x!}\right]^n \end{aligned}$$

(b) (Could probably argue for other answers also). Budworms may not be distributed at a uniform rate over the forest and they may not occur singly.

5.21 (a) 0.2048 (b) 0.0734 (c) 0.428 (d) 0.1404

5.22

$$(b)\ \ p = \frac{(4.5)^{11}\, e^{-4.5}}{11!} = 0.004264$$

$$(c)\ \ 1 - \sum_{x=0}^{10} \frac{(4.5)^{x}\, e^{-4.5}}{x!} = 0.006669$$

$$(d)\ \ \binom{20}{2} p^2 \,(1-p)^{18} = 0.003199$$

$$(e)\ \ (i) \binom{999}{11} p^{12} (1-p)^{988} = \binom{999}{11} p^{11} (1-p)^{988} p = 0.00001244$$

$$(ii)\ \ \binom{999}{11} p^{11} (1-p)^{988} p \approx \frac{\mu^{11} e^{-\mu}}{11!} p = 0.00001265$$

On the first 999 attempts we essentially have a Binomial distribution with $n = 999$ (large), $p = 0.004264$ (near 0) so the Poisson approximation can be used with $\mu = 999\,(0.004264) = 4.2601$.

5.23

$$(a)\ \ P\,(\text{no bubbles}) = \frac{(0.96)^0\, e^{-0.96}}{0!} = e^{-0.96}$$

$$(b)\ \ P\,(\text{more than one bubble}) = 1 - \left[\frac{(0.96)^0\, e^{-0.96}}{0!} + \frac{(0.96)^1\, e^{-0.96}}{1!}\right]$$
$$= 1 - 1.96 e^{-0.96} = 0.2495$$

$$(c)\ \ P\,(X = x) = \binom{n}{x} \left(e^{-0.96}\right)^x \left(1 - e^{-0.96}\right)^{n-x};\ \ x = 0, 1, \ldots, n$$

$$(d)\ \ \sum_{x=11}^{100} \binom{100}{x} \left(1 - 1.96 e^{-0.96}\right)^x \left(1.96 e^{-0.96}\right)^{100-x}$$
$$= 1 - \sum_{x=0}^{10} \binom{100}{x} \left(1 - 1.96 e^{-0.96}\right)^x \left(1.96 e^{-0.96}\right)^{100-x}$$
$$= 0.9999$$

$$(e)\ \ \frac{(0.8\lambda)^0\, e^{-0.8\lambda}}{0!} = e^{-0.8\lambda} \geq 0.5 \ \text{ or } \ \lambda \leq -1.25 \ln(0.5) \ \text{ bubbles per } m^2$$

(f) $e^{-0.8\lambda} + (0.8\lambda)\, e^{-0.8\lambda} \geq 0.95$ which must be solved numerically to obtain $\lambda \approx 0.4442$.

5.24 (a) 0.555 (b) 0.828; 0.965 (c) 0.789; 0.946 (d) $n = 1067$

5.25 (a) $\binom{x-1}{999}(0.3192)^{1000}(0.6808)^{x-1000}$

(b) 0.002, 0.051, 0.350, 0.797

(c) $\binom{3200}{y}(0.3192)^{y}(0.6808)^{3200-y}$; 0.797

5.26 (a) A key will be assigned to a given list (Success) with probability $1/M$ and not assigned (Failure) with probability $1 - 1/M$. Since the keys are assigned independently, we have a sequence of n Bernoulli trials. The probability that exactly x of the n keys are assigned to a given list is given by the Binomial distribution

$$\binom{n}{x}\left(\frac{1}{M}\right)^{x}\left(1-\frac{1}{M}\right)^{n-x} \quad \text{for } x = 0, 1, \ldots, n$$

(b) If n the number of keys is large and M the size of the hash table is large so that $1/M$ is small then we have

$$\binom{n}{x}\left(\frac{1}{M}\right)^{x}\left(1-\frac{1}{M}\right)^{n-x} \approx \frac{\alpha^{x}e^{-\alpha}}{x!} \quad \text{for } x = 0, 1, \ldots$$

by the Poisson approximation to the Binomial, where $\alpha = n\,(1/M) = n/M$.

(c) If $\alpha = 10$

$$P\,(X \geq x) \approx e^{-10}\left(\frac{10e}{x}\right)^{x}$$

Thus

$$P\,(X \geq 15) \approx e^{-10}\left(\frac{10e}{15}\right)^{15} = 0.3389$$

and

$$P\,(X \geq 20) \approx e^{-10}\left(\frac{10e}{20}\right)^{20} = 0.0210$$

5.27 (b)

$$1 - \sum_{x=0}^{2}\frac{(0.75)^{x}\,e^{-0.75}}{x!} = 0.04051$$

(c) $e^{-0.75(2)} = 0.2231$

(d) $e^{-0.75(3)} = 0.1054$

(e) $0.75e^{-0.75} = 0.3543$

$$\binom{10}{7}\left(0.75e^{-0.75}\right)^{7}\left(1 - 0.75e^{-0.75}\right)^{3} = 0.02263$$

5.28 (a) Let X_3 = number of bits which are flipped in transmission in a group of three bits. Then $X_3 \sim Binomial\,(3, p)$. A group of three repeated bits will be decoded correctly if none or one bit is flipped in transmission. Therefore

$$\begin{aligned}
&P\,(\text{a group of three repeated bits is decoded correctly}) \\
&= P\,(X_3 \le 1) = P\,(X_3 = 0) + P\,(X_3 = 1) \\
&= \binom{3}{0} p^0 (1-p)^3 + \binom{3}{1} p^1 (1-p)^2 \\
&= (1-p)^3 + 3p\,(1-p)^2
\end{aligned}$$

(b) If no ECC is used then a message of length four is decoded correctly if no bits are flipped in transmission which occurs with probability $(1-p)^4$.

(c) If TRC is used then the original message of length four is sent as a string of length twelve. The message is decoded correctly if each group of three is decoded correctly. Using the result from (a)

$$\begin{aligned}
&P\,(\text{original message of length four is decoded correctly}) \\
&= \left[(1-p)^3 + 3p\,(1-p)^2\right]^4
\end{aligned}$$

(d) See table given in Problem 5.29 (c). For $p = 0.2$ the probability that the message is not decoded correctly exceeds 50% if no ECC is used whereas if TRC is used the probability the message is decoded correctly is approximately 64%. As p decreases in value the probability the message is decoded correctly increases for both TRC and no ECC as one would expect. For TRC the probability the message is decoded correctly is nearly 1 for $p = 0.01$.

(e) Let X_5 = number of bits which are flipped in transmission in a group of five bits. Then $X_5 \sim Binomial\,(5, p)$. A group of five repeated bits will be decoded correctly if none, one or two bits are flipped in transmission. Therefore

$$\begin{aligned}
&P\,(\text{a group of five repeated bits is decoded correctly}) \\
&= P\,(X_5 \le 2) = P\,(X_5 = 0) + P\,(X_5 = 1) + P\,(X_5 = 2) \\
&= \binom{5}{0} p^0 (1-p)^5 + \binom{5}{1} p^1 (1-p)^4 + \binom{5}{2} p^2 (1-p)^3 \\
&= (1-p)^5 + 5p\,(1-p)^2 + 10p^2 (1-p)^3
\end{aligned}$$

(f) Let X_k = number of bits which are flipped in transmission in a group of k bits. Then $X_k \sim Binomial\,(k, p)$. A group of k repeated bits will be decoded correctly if $0, 1, \ldots,$ or $(k-1)/2$ bits are flipped in transmission. Therefore

$$P(k,p) = P\,(\text{a group of } k \text{ repeated bits is decoded correctly})$$
$$= P\left(X_k \leq \frac{k-1}{2}\right)$$
$$= \sum_{x=0}^{(k-1)/2} \binom{k}{x} p^x (1-p)^{k-x}$$

	$P(k,p)$	k = 3	5	7	9
p	0.2	0.896	0.9421	0.9667	0.9969
	0.1	0.972	0.9914	0.9973	0.9999
	0.05	0.9928	0.9988	0.9998	1.0000

As the number of repeated bits, k, is increased the probability that a group of k repeated bits is decoded correctly approaches 1 for each value of p.

5.29 (a) A correctable message is received using Hamming(7,4) if none or one of the seven bits sent are flipped which occurs with probability

$$\binom{7}{0} p^0 (1-p)^7 + \binom{7}{1} p^1 (1-p)^6$$
$$= (1-p)^7 + 7p(1-p)^6$$

(b) See values given in the table in (c).

(c)

No. bits sent	ECC	p = 0.2	0.1	0.05	0.01
4	No ECC	0.4906	0.6561	0.8145	0.9696
12	TRC	0.6445	0.8926	0.9713	0.9988
7	Hamming(7,4)	0.5767	0.8503	0.9556	0.9980

As before, the probabilities increase as the value of p decreases. Of most interest is to compare TRC which requires 12 bits to Hamming(7,4) which requires only 7 bits. Hamming(7,4) performs almost as well as TRC for fewer bits sent.

7. SOLUTIONS TO CHAPTER 7 PROBLEMS

7.1 $E(X) = 4,\ E(X^2) = 17.6,\ Var(X) = E(X^2) - [E(X)]^2 = 17.6 - 16 = 1.6$
$E(X) = 2.5,\ E(X^2) = 7.2,\ Var(X) = E(X^2) - [E(X)]^2 = 7.2 - 6.25 = 0.95$

7.2 $E(X) = 2.775,\ E(X^2) = 10.275,$
$Var(X) = E(X^2) - [E(X)]^2 = 10.275 - (2.775)^2 = 2.574375$

7.3 (a) The person wins $\$2^x$ if x tosses are needed for $x = 1, 2, 3, 4, 5$ but loses $\$256$ if $x > 5$. Note that

$$P(X > 5) = \frac{1}{2^6} + \frac{1}{2^7} + \frac{1}{2^8} + \cdots = \frac{1/2^6}{1 - \frac{1}{2}} = \frac{1}{2^5}$$

Let $W =$ winnings. Then the probability function for W is

w	2	2^2	2^3	2^4	2^5	-256	Total
$P(W = w)$	$\frac{1}{2}$	$\left(\frac{1}{2}\right)^2 = \frac{1}{2^2}$	$\frac{1}{2^3}$	$\frac{1}{2^4}$	$\frac{1}{2^5}$	$\frac{1}{2^5}$	1

Therefore

$$E(W) = 2\left(\frac{1}{2}\right) + 2^2\left(\frac{1}{2^2}\right) + 2^3\left(\frac{1}{23}\right) + 2^4\left(\frac{1}{2^4}\right) + 2^5\left(\frac{1}{2^5}\right) + (-256)\left(\frac{1}{2^5}\right)$$
$$= 5 - 8 = -3 \text{ dollars}$$

(b)

$$E(W^2) = (2)^2\left(\frac{1}{2}\right) + (2^2)^2\left(\frac{1}{2^2}\right) + (2^3)^2\left(\frac{1}{2^3}\right) + (2^4)^2\left(\frac{1}{2^4}\right) + (2^5)^2\left(\frac{1}{2^5}\right) + (-256)^2\left(\frac{1}{2^5}\right)$$
$$= 2110$$

and

$$Var(W) = E(W^2) - [E(W)]^2 = 2110 - (-3)^2 = 2101 \text{ (dollars)}^2$$

7.4 (a) Let $Y =$ number of cups drunk by Yasmin and let $Z =$ number of cups drunk by Zack. Then $Y = 2X^2$ and $Z = |2X - 1|$.

$$\begin{aligned} E(Y) = E\left(2X^2\right) &= \sum_{x=0}^{5}(2x^2)P(X=x) = 2\sum_{x=0}^{5}x^2P(X=x) \\ &= 2[(0)^2(0.09) + (1)^2(0.10) + (2)^2(0.25) \\ &\qquad + (3)^2(0.40) + (4)^2(0.15) + (5)^2(0.01)] = 14.7 \end{aligned}$$

On average Yasmin drinks 14.7 cups of coffee per week.

$$\begin{aligned} E(Z) = E(|2X-1|) &= \sum_{x=0}^{5}|2x-1|P(X=x) \\ &= |2(0)-1|(0.09) + |2(1)-1|(0.10) + |2(2)-1|(0.25) \\ &\qquad + |2(3)-1|(0.40) + |2(4)-1|(0.15) + |2(5)-1|(0.01) \\ &= 4.08 \end{aligned}$$

On average Zack drinks 4.08 cups of coffee per week.

(b) Since

$$\begin{aligned} E\left(Y^2\right) = E\left[\left(2X^2\right)^2\right] = E\left(4X^4\right) &= 4\sum_{x=0}^{5}x^4P(X=x) \\ &= 4[(0)^4(0.09) + (1)^4(0.10) + (2)^4(0.25) \\ &\qquad + (3)^4(0.40) + (4)^4(0.15) + (5)^4(0.01)] = 324.6 \end{aligned}$$

therefore

$$Var(Y) = E\left(Y^2\right) - [E(Y)]^2 = 324.6 - (14.7)^2 = 108.51$$

Since

$$\begin{aligned} E\left(Z^2\right) = E\left(|2X-1|^2\right) &= \sum_{x=0}^{5}|2x-1|^2P(X=x) \\ &= |2(0)-1|^2(0.09) + |2(1)-1|^2(0.10) + |2(2)-1|^2(0.25) \\ &\qquad + |2(3)-1|^2(0.40) + |2(4)-1|^2(0.15) + |2(5)-1|^2(0.01) \\ &= 20.6 \end{aligned}$$

therefore

$$Var(Z) = E\left(Z^2\right) - [E(Z)]^2 = 20.6 - (4.08)^2 = 3.9536$$

7.5 (a) Using the results of Problem 3.6.2 we have

$$\begin{aligned} E(X) &= \sum_{x=1}^{\infty} xp(1-p)^x = p(1-p)\sum_{x=1}^{\infty} x(1-p)^{x-1} \\ &= p(1-p)\frac{1}{[1-(1-p)]^2} = \frac{1-p}{p} \end{aligned}$$

$$\begin{aligned} E[X(X-1)] &= \sum_{x=2}^{\infty} x(x-1)p(1-p)^x \\ &= p(1-p)^2\sum_{x=1}^{\infty} x(x-1)(1-p)^{x-2} \\ &= p(1-p)^2\frac{2}{[1-(1-p)]^3} = \frac{2(1-p)^2}{p^2} \end{aligned}$$

and

$$\begin{aligned} Var(X) &= E[X(X-1)] + E(X) - [E(X)]^2 \\ &= \frac{2(1-p)^2}{p^2} + \frac{1-p}{p} - \left(\frac{1-p}{p}\right)^2 \\ &= \left(\frac{1-p}{p}\right)^2 + \frac{p(1-p)}{p^2} \\ &= \frac{(1-p)}{p^2}(1-p+p) \\ &= \frac{1-p}{p^2} \end{aligned}$$

(b) Let $Y =$ the number of trials to obtain the first success. Then $Y = X+1$ and therefore

$$E(Y) = E(X+1) = E(X) + 1 = \frac{1-p}{p} + 1 = \frac{1}{p}$$

7.6 Let X be the number of tests applied to an individual.

$$\begin{aligned} P(X=1) &= P(\text{individual tests negative}) \\ &= P(\text{tests negative}|D)\,P(D) + P\left(\text{tests negative}|\bar{D}\right)P\left(\bar{D}\right) \\ &= 0 + (0.95)(0.98) = 0.931 \end{aligned}$$

$$\begin{aligned} P(X=2) &= P(\text{individual tests positive}) \\ &= P(\text{tests positive}|D)\,P(D) + P\left(\text{tests positive}|\bar{D}\right)P\left(\bar{D}\right) \\ &= 1 - 0.931 = 0.069 \end{aligned}$$

Then the expected cost per person is $\$10(0.931) + \$110(0.069) = \$16.90$.

7.7 (a) For test B, $P\,(\text{test positive}) = 0.02$ so the expected number of cases detected among 150 person is $150\,(0.02) = 3$ cases.

(b) For test A, $P\,(\text{test positive}) = (0.8)\,(0.02) + (0.05)\,(0.98) = 0.065$ so the expected cost is $2000 + 100\,(2000)\,(0.065) = 15000$ and the expected number of cases detected is $2000\,(0.02)\,(0.8) = 32$ cases.

7.8 (a) The possible number of tests done is 1 if all k people are negative or $k+1$ tests if at least one person is positive.

x	1	$k+1$	Total
$P\,(X = x)$	$(1-p)^k$	$1-(1-p)^k$	1

so

$$\begin{aligned} E\,(X) &= (1)\,(1-p)^k + (k+1)\left[1-(1-p)^k\right] \\ &= k+1-k\,(1-p)^k \end{aligned}$$

(b)

$$\begin{aligned} \text{Expected number of tests for } \frac{n}{k} \text{ groups} &= \frac{n}{k}\left[k+1-k\,(1-p)^k\right] \\ &= n + \frac{n}{k} - n(1-p)^k \end{aligned}$$

which gives $1.01n$, $0.249n$, $0.196n$ for $k = 1, 5, 10$.

7.9 (a) If you bet $\$1$ on 10 consecutive plays then your expected net winnings equals

$$10\left[(1)\left(\frac{18}{37}\right) + (-1)\left(\frac{19}{37}\right)\right] = 10\left(\frac{-1}{37}\right) = -\frac{10}{37} \text{ dollars}$$

If you bet $\$10$ on a single play then your expected net winnings equals

$$(10)\left(\frac{18}{37}\right) + (-10)\left(\frac{19}{37}\right) = -\frac{10}{37} \text{ dollars}$$

(b) If you bet $\$1$ on 10 consecutive plays then the probability you make a profit is the probability you win 6 or more times which equals

$$\sum_{x=6}^{10} \binom{10}{x}\left(\frac{18}{37}\right)^x\left(\frac{19}{37}\right)^{10-x} = 0.3442$$

If you bet $\$10$ on a single play then the probability you make a profit is $\frac{18}{37} = 0.4865$

7.10 Expected winnings per dollar spent is equal to

$$(20)\left(\frac{2}{10}\right)\left(\frac{6}{10}\right)\left(\frac{2}{10}\right)+(10)\left(\frac{4}{10}\right)\left(\frac{3}{10}\right)\left(\frac{3}{10}\right)+(5)\left(\frac{4}{10}\right)\left(\frac{1}{10}\right)\left(\frac{5}{10}\right)$$
$$= 0.94 \text{ dollars}$$

7.11 (a)

$$\text{Expected net profit} = \sum_{i=1}^{n}(a_i-1)\,p_i+(-1)\,p_{n+1}=\sum_{i=1}^{n}a_ip_i-\sum_{i=1}^{n+1}p_i=\sum_{i=1}^{n}a_ip_i-1$$

(b)

$$\text{Expected net profit} = 3\,(0.1)+5\,(0.04+0.04+0.04)-1=-0.10 \text{ dollars}$$

(c) The expected profit is

$$-0.05=\sum_{i=1}^{n}db_ip_i+(-1)\,p_{n+1}=\sum_{i=1}^{n}d\left(\frac{1}{p_i}\right)p_i+(-1)\,p_{n+1}=dn-p_{n+1}$$

so

$$d=(p_{n+1}-0.05)\,/n$$

If $n=10$ and $p_{n+1}=0.7$ then $d=0.065$.

7.12 (a) Let X_i be the winnings when attempting Question i first and let i be the event Question i is answered correctly, $i=A,B$.

$$\begin{aligned}
E\,(X_A) &= \sum_{x\in S} xP(X_A=x)\\
&=(100+200)P(A\cap B)+(100)P(A\cap\overline{B})+(0)P(\overline{A})\\
&=(300)P(A)P(B)+(100)P(A)P(\overline{B})\\
&=(300)(0.8)(0.6)+(100)(0.8)(0.4)\\
&=176\\
E\,(X_B) &= \sum_{x\in S} xP(X_B=x)\\
&=(100+200)P(B\cap A)+(200)P(B\cap\overline{A})+(0)P(\overline{B})\\
&=(300)P(B)P(A)+(200)P(B)P(\overline{A})\\
&=(300)(0.8)(0.6)+(200)(0.6)(0.2)\\
&=168
\end{aligned}$$

Therefore the expected winnings are greatest when attempting Question A first.

(b)

$$
\begin{aligned}
E\left(X_A\right) &= (300)P(A\cap B) + (100)P(A\cap \overline{B}) + (-50)P(\overline{A}) \\
&= (300)(0.8)(0.6) + (100)(0.8)(0.4) - (50)(0.2) \\
&= 166 \\
E\left(X_B\right) &= (300)P(B\cap A) + (200)P(B\overline{\cap A}) + (-50)P(\overline{B}) \\
&= (300)(0.6)(0.8) + (200)(0.6)(0.2) - (50)(0.4) \\
&= 148
\end{aligned}
$$

Question A should still be attempted first to maximize expected winnings.

7.13 Let $N =$ net profit. Then $N = 59.5n - 25 - 200X^2$ where $X \sim Binomial\,(n, 0.05)$ which implies $E\,(X) = n\,(0.05)$ and $Var\,(X) = n\,(0.05)\,(0.95)$. Also since $Var\,(X) = E\left(X^2\right) - [E\,(X)]^2$ we know $E\left(X^2\right) = Var\,(X) + [E\,(X)]^2$. The expected net profit is

$$
\begin{aligned}
E\,(N) &= E\left(59.5n - 25 - 200X^2\right) = 59.5n - 25 - 200E\left(X^2\right) \\
&= 59.5n - 25 - 200\left\{Var\,(X) + [E\,(X)]^2\right\} \\
&= 50n - 0.5n^2 - 25
\end{aligned}
$$

which is maximized for $n = 50$.

7.14 (a) Let $N =$ the number of trick-or-treaters that arrive in the first half hour. Since arrivals follow a Poisson process then $N \sim Poisson\,(6)$.

$$
\begin{aligned}
P(5 \le N \le 7) &= P(N=5) + P(N=6) + P(N=7) \\
&= \frac{e^{-6}(6)^5}{5!} + \frac{e^{-6}(6)^6}{6!} + \frac{e^{-6}(6)^7}{7!} \\
&= e^{-6}\frac{1296}{7} \\
&= 0.45892
\end{aligned}
$$

(b) Let $M =$ the number of trick-or-treaters that arrive per 3.5 hours. Therefore $M \sim Poisson(\mu)$ where $\mu = 3.5 \times 12 = 42$. Since $E\,(M) = 42$, the expected number of trick-or-treaters over the whole evening is 42.

(c) Using M as (b), we note that

$$
\frac{P(M=x)}{P(M=x+1)} = \frac{e^{-42}\frac{42^x}{x!}}{e^{-42}\frac{42^{x+1}}{(x+1)!}} = \frac{42^x}{42^{x+1}}\frac{(x+1)!}{x!} = \frac{x+1}{42}
$$

This is < 1 if $x < 41$ and > 1 if $x > 41$. It follows that outcomes become more likely as x increases to a maximum of 41 and then less likely thereafter. The two outcomes, $X = 41$ and $X = 42$ are both equally likely (and occur with probability $e^{-42}\frac{42^{41}}{41!} = e^{-42}\frac{42^{42}}{42!} \approx 0.0614$).

(d) Since $Var(M) = \mu$, the variance of the number of trick-or-treaters arriving over the whole evening is 42.

7.15 Let $X =$ the number of weeks the stock price increases in value by $\$1$, then $X \sim Binomial\left(13, \frac{1}{2}\right)$. The price of the stock in 13 weeks is $S = 50 + X - (13 - X) = 37 + 2X$. The return from the option is $R = \max(37 + 2X - 55, 0) = \max(2X - 18, 0)$. Therefore

$$
\begin{aligned}
E(R) &= 0 + \sum_{x=10}^{13} (2x - 18)\binom{13}{x}\left(\frac{1}{2}\right)^{x}\left(1 - \frac{1}{2}\right)^{13-x} \\
&= \left[2\binom{13}{10} + 4\binom{13}{11} + 6\binom{13}{12} + 8\binom{13}{13}\right]\left(\frac{1}{2}\right)^{13} \\
&= \frac{485}{4096}
\end{aligned}
$$

7.16 Let $X =$ time to display the information.

(a) Without cache, $X = 50 + 70 + 50 = 170$ only so the expected time to display is 170.

(b) With a cache, $X = 10$ or $10 + 50 + 70 + 50 = 180$ (since the cache is always searched first) with probabilities 0.2 and 0.8 respectively, so the expected time to display is $10(0.2) + 180(0.8) = 146$ ms.

(c) Let p be the probability of a cache hit. Solving $170 = 10p + 180(1 - p)$ gives $p = 0.059$. Therefore even with only around a 6% probability of a cache hit it is still worthwhile to use a cache!

8. SOLUTIONS TO CHAPTER 8 PROBLEMS

8.1 (a) Solving

$$1=\int_{-\infty}^{\infty} f(x)dx=0+k\int_{-1}^{1}\left(1-x^2\right)dx+0=k\int_{-1}^{1}\left(1-x^2\right)dx$$

gives $k=3/4$.

$$F(x)=\begin{cases}0 & x\leq -1\\ \frac{3}{4}\int_{-1}^{x}\left(1-u^2\right)du=\frac{1}{4}\left(2+3x-x^3\right) & -1<x<1\\ 1 & x\geq 1\end{cases}$$

(b) We need to find c such that

$$\begin{aligned}0.95&=P\left(-c\leq X\leq c\right)\\&=\frac{1}{4}\left[\left(2+3c-c^3\right)-\left(2-3c+c^3\right)\right]\\&=\frac{1}{4}\left(6c-2c^3\right)=\frac{1}{2}\left(3c-c^3\right)\end{aligned}$$

or

$$c^3-3c+1.9=0$$

This cubic equation must be solved numerically which gives $c=0.811$.

(c) $\mu=E\left(X\right)=0$ since the p.d.f. is symmetric about $x=0$.

$$\sigma^2=Var\left(X\right)=E\left(X^2\right)-0^2=E\left(X^2\right)=0.75\int_{-1}^{1}x^2\left(1-x^2\right)dx=0.2$$

$$\sigma=sd\left(X\right)=\sqrt{0.2}$$

(d) For $0 < y < 1$

$$G(y) = P(Y \le y) = P\left(X^2 \le y\right) = P(-\sqrt{y} \le X \le \sqrt{y}) = F(\sqrt{y}) - F(-\sqrt{y})$$

$$\begin{aligned}
g(y) &= \frac{d}{dy} G(y) = f(\sqrt{y}) \frac{d}{dy} \sqrt{y} - f(-\sqrt{y}) \frac{d}{dy}(-\sqrt{y}) = f(\sqrt{y}) \frac{1}{2\sqrt{y}} + f(-\sqrt{y}) \frac{1}{2\sqrt{y}} \\
&= 2f(\sqrt{y}) \frac{1}{2\sqrt{y}} \quad \text{since } f(-\sqrt{y}) = f(\sqrt{y}) \text{ by symmetry of } f(x) \\
&= \frac{3}{4}(1-y) \frac{1}{\sqrt{y}} = \frac{3}{4}\left(y^{-1/2} - y^{1/2}\right)
\end{aligned}$$

Therefore

$$g(y) = \begin{cases} \frac{3}{4}\left(y^{-1/2} - y^{1/2}\right) & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

8.2 (a) The area under the p.d.f. and above the x-axis consists of two triangles with area $= \frac{1}{2}\left(\frac{1}{2}\right)(2) + \frac{1}{2}\left(\frac{1}{2}\right)(2) = 1$ and therefore $\int\limits_{-\infty}^{\infty} f(x)dx = 1$.

(b)

$$P(0.25 \le X \le 0.8) = \int_{0.25}^{0.8} f(x)dx = \int_{0.25}^{0.5} 4x dx + \int_{0.5}^{0.8} 4(1-x)dx = 0.375 + 0.42 = 0.795$$

(c) Since the p.d.f. is symmetric about the line $x = 1/2$, the median is equal to $1/2$.
To find the 10th percentile we need to find the value c such that $0.1 = F(c)$. Since $0.5 = F(1/2)$ we know that c must lie between 0 and $1/2$. Therefore c is the solution to

$$0.1 = \int_0^c 4x dx = 2x^2|_0^c = 2c^2$$

which gives $c = \sqrt{0.05} \approx 0.2236$.

(d) Since the p.d.f. is symmetric about the line $x = 1/2$, the mean is $\mu = E(X) = 1/2$.

$$E\left(X^2\right) = \int_{-\infty}^{\infty} x^2 f(x)dx = 0 + \int_0^{1/2} 4x^3 dx + \int_{1/2}^{1} 4x^2(1-x)dx + 0 = \frac{7}{24}$$

$$Var(X) = E\left(X^2\right) - [E(X)]^2 = \frac{7}{24} - \left(\frac{1}{2}\right)^2 = \frac{1}{24}$$

(e) For $-1 < y < 1$

$$G(y) = P(Y \le y) = P(2(X - 1/2) \le y) = P\left(X \le \frac{y+1}{2}\right) = F\left(\frac{y+1}{2}\right)$$

$$g(y) = \frac{d}{dy} G(y) = f\left(\frac{y+1}{2}\right) \frac{d}{dy}\left(\frac{y+1}{2}\right) = f\left(\frac{y+1}{2}\right)\left(\frac{1}{2}\right)$$

Therefore

$$g(y)=\begin{cases} y+1 & -1<y\le 0\\ 1-y & 0<y<1\\ 0 & \text{otherwise}\end{cases}$$

(f) For $0<z<1$

$$H(z)=P(Z\le z)=P\left(X^3\le z\right)=P\left(X\le z^{1/3}\right)=F\left(z^{1/3}\right)$$

since $P(X<0)=0$. Therefore

$$\begin{aligned} h(z)&=\frac{d}{dz}H(z)=f\left(z^{1/3}\right)\frac{d}{dz}z^{1/3}=f\left(z^{1/3}\right)\frac{1}{3z^{2/3}}\\ &=\begin{cases} 4z^{-1/3}/3 & 0<z\le(0.5)^3\\ 4\left(z^{-2/3}-z^{-1/3}\right)/3 & (0.5)^3<z<1\\ 0 & \text{otherwise}\end{cases}\end{aligned}$$

8.3 For $0<y<1$

$$\begin{aligned} G(y)&=P(Y\le y)=P\left(\frac{X+10}{20}\le y\right)=P(X\le 20y-10)\\ &=F(20y-10)\\ g(y)&=\frac{d}{dy}G(y)=f(20y-10)\frac{d}{dy}(20y-10)=\left(\frac{1}{20}\right)(20)=1\end{aligned}$$

which is the probability density function of a $U(0,1)$ random variable.

8.4

$$\begin{aligned} P(|X-\mu|\le 2\sigma)&=P(\mu-2\sigma\le X\le\mu+2\sigma)=2P(\mu\le X\le\mu+2\sigma)\quad\text{by symmetry}\\ &=2\left[P(X\le\mu+2\sigma)-P(X\le\mu)\right]\\ &=2\left[P(X\le\mu+2\sigma)-0.5\right]\quad\text{since } P(X\le\mu)=P(X>\mu)=0.5\\ &=2P(X\le\mu+2\sigma)-1\end{aligned}$$

8.5 (a) For females

$$P(X>80|X>30)=\frac{P(X>80,X>30)}{P(X>30)}=\frac{P(X>80)}{P(X>30)}=\frac{0.704}{0.996}=0.707$$

$$P(X>90|X>30)=\frac{P(X>90)}{P(X>30)}=\frac{0.396}{0.996}=0.398$$

For males

$$P(X>80|X>30)=\frac{P(X>80)}{P(X>30)}=\frac{0.603}{0.989}=0.610$$
$$P(X>90|X>30)=\frac{P(X>90)}{P(X>30)}=\frac{0.273}{0.989}=0.276$$

(b)

$$\begin{aligned}P(X>90)&=P(X>90|\text{Female})\,P(\text{Female})+P(X>90|\text{Male})\,P(\text{Male})\\&=(0.396)(0.49)+(0.273)(0.51)=0.333\end{aligned}$$

8.6 (a) $f(x)$ is a probability density function for $\theta>-1$, since for $\theta>-1$, $f(x)\geq 0$ for all $x\in\Re$ and

$$\int_{-\infty}^{\infty}f(x)dx=\int_0^1(\theta+1)\,x^{\theta}dx=x^{\theta+1}|_0^1=1$$

(b)

$$P(X\leq 0.5)=\int_0^{0.5}(\theta+1)\,x^{\theta}dx=x^{\theta+1}|_0^{0.5}=(0.5)^{\theta+1}$$

(c) For $k=1,2,\ldots$

$$\begin{aligned}E\left(X^k\right)&=\int_{-\infty}^{\infty}x^kf(x)dx=\int_0^1x^k(\theta+1)\,x^{\theta}dx=(\theta+1)\int_0^1x^{k+\theta}dx\\&=\frac{\theta+1}{\theta+k+1}x^{\theta+k+1}|_0^1=\frac{\theta+1}{\theta+k+1}\end{aligned}$$

Therefore

$$E(X)=\frac{\theta+1}{\theta+2},\quad E\left(X^2\right)=\frac{\theta+1}{\theta+3}$$

and

$$Var(X)=E\left(X^2\right)-[E(X)]^2=\frac{\theta+1}{\theta+3}-\left(\frac{\theta+1}{\theta+2}\right)^2=\frac{\theta+1}{(\theta+2)^2(\theta+3)}$$

(d) For $y>1$

$$\begin{aligned}G(y)&=P(Y\leq y)=P\left(\frac{1}{X}\leq y\right)=P\left(X>\frac{1}{y}\right)=1-P\left(X\leq\frac{1}{y}\right)\\&=1-F\left(\frac{1}{y}\right)\ \text{ since }\ F(x)=P(X\leq x)\\g(y)&=\frac{d}{dy}G(y)=\frac{d}{dy}\left[1-F\left(\frac{1}{y}\right)\right]\\&=-f\left(\frac{1}{y}\right)\frac{d}{dy}\left(\frac{1}{y}\right)=f\left(\frac{1}{y}\right)\left(\frac{1}{y^2}\right)=\frac{\theta+1}{y^{\theta+2}}\ \text{ for } y>1\end{aligned}$$

Therefore

$$g(y) = \begin{cases} \frac{\theta+1}{y^{\theta+2}} & y > 1 \\ 0 & \text{otherwise} \end{cases}$$

8.7 (a)

$$1 = \int_{-\infty}^{\infty} f(x)dx = 0 + k\int_{0}^{\infty} xe^{-x^2/\theta}dx = k\lim_{b\to\infty}\left[-\frac{\theta}{2}e^{-x^2/\theta}|_0^b\right]$$
$$= k\left(\frac{\theta}{2}\right)\left[1 - \lim_{b\to\infty} e^{-b^2/\theta}\right] = k\left(\frac{\theta}{2}\right) \text{ and therefore } k = \frac{2}{\theta}$$

$$F(x) = \begin{cases} 0 & x \le 0 \\ \int_0^x \frac{2u}{\theta}e^{-u^2/\theta}du = 1 - e^{-x^2/\theta} & x > 0 \end{cases}$$

(b)

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx = 0 + \int_0^{\infty} x\frac{2x}{\theta}e^{-x^2/\theta}dx = \int_0^{\infty}\frac{2x^2}{\theta}e^{-x^2/\theta}dx \quad \text{let } y = \frac{x^2}{\theta}, dy = \frac{2x}{\theta}dx$$
$$= \int_0^{\infty}\theta^{1/2}y^{1/2}e^{-y}dy = \theta^{1/2}\int_0^{\infty} y^{3/2-1}e^{-y}dy = \theta^{1/2}\Gamma\left(\frac{3}{2}\right) = \theta^{1/2}\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right) = \frac{\sqrt{\theta\pi}}{2}$$

$$E\left(X^2\right) = \int_{-\infty}^{\infty} x^2f(x)dx = 0 + \int_0^{\infty} x^2\frac{2x}{\theta}e^{-x^2/\theta}dx = \int_0^{\infty}\frac{2x^3}{\theta}e^{-x^2/\theta}dx \quad \text{let } y = \frac{x^2}{\theta}, dy = \frac{2x}{\theta}dx$$
$$= \int_0^{\infty}\theta ye^{-y}dy = \theta\int_0^{\infty} y^{2-1}e^{-y}dy = \theta\Gamma(2) = \theta(1) = \theta$$
$$Var(X) = E\left(X^2\right) - [E(X)]^2 = \theta - \left(\frac{\sqrt{\theta\pi}}{2}\right)^2 = \theta\left(\frac{4-\pi}{4}\right)$$

(c) For $y > 0$

$$G(y) = P(Y \le y) = P\left(\frac{X^2}{\theta} \le y\right) = P\left(X \le \sqrt{\theta y}\right) \quad \text{since } P(X \le 0) = 0$$
$$= F\left(\sqrt{\theta y}\right)$$
$$g(y) = \frac{d}{dy}G(y) = f\left(\sqrt{\theta y}\right)\frac{d}{dy}\sqrt{\theta y} = f\left(\sqrt{\theta y}\right)\frac{\sqrt{\theta}}{2\sqrt{y}} = \frac{2\sqrt{\theta y}}{\theta}e^{-y}\frac{\sqrt{\theta}}{2\sqrt{y}}$$
$$= e^{-y}$$

which is the probability density funciton of an $Exponential\,(1)$ random variable. Therefore $Y = X^2/\theta \sim Exponential\,(1)$.

8.8 Since X is the diameter of the sphere

$$Y = \left(\frac{4\pi}{3}\right)\left(\frac{X}{2}\right)^3 = \frac{\pi}{6}X^3 \text{ and } X = \left(\frac{6}{\pi}Y\right)^{1/3}$$

Since $X \sim U\,(0.6, 1)$

$$f\,(x) = \frac{1}{1-0.6} = \frac{5}{2} \text{ for } 0.6 < x < 1$$

The range of Y is

$$\frac{\pi}{6}\,(0.6)^3 < y < \frac{\pi}{6}\,(1)^3 \text{ or } 0.036\pi \le y \le \frac{\pi}{6}$$

Therefore for $0.036\pi \le y \le \frac{\pi}{6}$

$$\begin{aligned} G\,(y) &= P\,(Y \le y) = P\left(\frac{\pi}{6}X^3 \le y\right) = P\left(X \le \left(\frac{6}{\pi}y\right)^{1/3}\right) \\ &= F\left(\left(\frac{6}{\pi}y\right)^{1/3}\right) \end{aligned}$$

and

$$\begin{aligned} g(y) &= \frac{d}{dy}G\,(y) = \frac{d}{dy}\left[F\left(\left(\frac{6}{\pi}y\right)^{1/3}\right)\right] \\ &= f\left(\left(\frac{6}{\pi}y\right)^{1/3}\right)\frac{d}{dy}\left(\frac{6}{\pi}y\right)^{1/3} \\ &= \frac{5}{2}\left(\frac{6}{\pi}\right)^{\frac{1}{3}}\left(\frac{1}{3}\right)y^{-2/3} \\ &= \left(\frac{5}{6}\right)\left(\frac{6}{\pi}\right)^{\frac{1}{3}}y^{-\frac{2}{3}} \end{aligned}$$

Therefore

$$g\,(y) = \begin{cases} \left(\frac{5}{6}\right)\left(\frac{6}{\pi}\right)^{\frac{1}{3}}y^{-\frac{2}{3}} & 0.036\pi \le y \le \frac{\pi}{6} \\ 0 & \text{otherwise} \end{cases}$$

8.9 (a) Let X = magnitude of earthquake. Then $X \sim Exponential\,(2.5)$. The probability an earthquake has a magnitude greater than 5 is $P\,(X > 5) = e^{-5/2.5} = e^{-2}$.

(b) The probability that among 3 earthquakes there are none with a magnitude greater than 5 is $\binom{3}{0}\left(e^{-2}\right)^0(1-e^{-2})^3 = (1-e^{-2})^3$.

(c) By the memoryless property of the Exponential distribution

$$P\,(X > 5|X > 4) = P\,(X > 1) = e^{-1/2.5} = e^{-0.4}$$

8.10 Let $X =$ lifetime of this type of light bulb in hours. Then $X \sim Exponential\,(1000)$.

(a) $E\,(X) = 1000$ hours and $sd\,(X) = \sqrt{(1000)^2} = 1000$ hours.

(b) Let $Y =$ lifetime of this type of lightbulb in days. Then $Y = X/24$ and

$$\begin{aligned} E\,(Y) &= E\left(\frac{X}{24}\right) \\ &= \frac{1}{24}E\,(X) = \frac{1}{24}\,(1000) \\ &= \frac{125}{3} \text{ days} \\ Var\,(Y) &= Var\left(\frac{X}{24}\right) \\ &= \left(\frac{1}{24}\right)^2 Var\,(X) \\ &= \left(\frac{1}{24}\right)^2 (1000)^2 = \left(\frac{1000}{24}\right)^2 \\ &= \left(\frac{125}{3}\right)^2 \\ sd\,(Y) &= \sqrt{\left(\frac{125}{3}\right)^2} = \frac{125}{3} \text{ days} \end{aligned}$$

(c) Solving

$$0.5 = P\,(X > m) = e^{-m/1000}$$

gives

$$m = 1000 \ln 2 \approx 693.14 \text{ hours}$$

8.11 (a) Let $X =$ waiting time until the next accident. Since accidents occur according to a Poisson process with rate $\lambda = 0.5$ accidents per day, then X has an Exponential distribution with mean $\theta = 1/\lambda = 1/0.5 = 2$ days. Since $E\,(Y) = \theta$ if $Y \sim Exponential\,(\theta)$ then $E\,(X) = 2$ days is the expected waiting time until the next accident.

(b)

$$\begin{aligned} &P\,(\text{waiting time is less than 12 hours}) \\ &= P\,(\text{waiting time is less than } 0.5 \text{ days}) \\ &= P\,(X < 0.5) = 1 - e^{-0.5/2} = 1 - e^{-0.25} \end{aligned}$$

(c) By the memoryless property of the Exponential distribution

$$P\,(X > 1 | X > 0.5) = P\,(X > 0.5) = e^{-0.25}$$

8.12 (a) Let X be the lifetime kilometer-age of the car. Then X has an Exponential distribution with mean 20 thousand kilometers. By the memoryless property of the Exponential distribution,

$$P(X > 20 + 10|X > 10) = P(X > 20) = e^{-20/20} = e^{-1} = 0.3679$$

(b) If $X =$ the lifetime kilometer-age of the car (in thousands of kilometers) is a $Uniform(0, 40)$ random variable then

$$P(X > 20 + 10|X > 10) = \frac{P(X > 30, X > 10)}{P(X > 10)} = \frac{P(X > 30)}{P(X > 10)}$$
$$= \frac{\int_{30}^{40} \frac{1}{40}dx}{\int_{10}^{40} \frac{1}{40}dx} = \frac{1}{3}$$

8.13 Let $X =$ waiting time in days between server crashes. Since on average there are three server crashes per day, $X \sim Exponential\,(1/3)$.

(a) Since 8 hours = $1/3$ day, the probability that the waiting time between two consecutive crashes is greater than 8 hours is

$$P\left(X > \frac{1}{3}\right) = \int_{1/3}^{\infty} 3e^{-3x}dx = e^{-3(1/3)} = e^{-1} = 0.368$$

(b) By the memoryless property of the Exponential distribution

$$P\left(X > \frac{1}{24} + \frac{1}{3}|X > \frac{1}{3}\right) = P\left(X > \frac{1}{24}\right) = e^{-3(1/24)} = 0.882$$

since one hour = $1/24$ days.

(c) Let $N =$ number of crashes in a day. Then $N \sim Poisson\,(3)$ and the probability that there are fewer than three crashes in a day is

$$P\,(N < 3) = P\,(N \leq 2) = \sum_{n=0}^{2} \frac{3^n e^{-3}}{n!} = 0.423$$

8.14 (a) Clearly, $f(x) \geq 0$ for any $x \in \mathbb{R}$. Since

$$\int\limits_{-\infty}^{\infty} f(x)dx = 0 + \int\limits_{0}^{\infty} \frac{x^{\alpha-1}}{\Gamma(\alpha)\beta^{\alpha}} e^{-\frac{x}{\beta}} dx$$
$$= \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \int_0^{\infty} x^{\alpha-1} e^{-\frac{x}{\beta}} dx \text{ let } \frac{x}{\beta} = y$$
$$= \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \int_0^{\infty} (\beta y)^{\alpha-1} e^{-y} d(\beta y)$$
$$= \frac{1}{\Gamma(\alpha)} \int_0^{\infty} y^{\alpha-1} e^{-y} dy = 1$$

therefore f is a legitimate probability density function.

(b)

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_0^{\infty} x \frac{x^{\alpha-1}}{\Gamma(\alpha)\beta^{\alpha}} e^{-\frac{x}{\beta}} dx = \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} \beta = \alpha\beta$$

Similarly,

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx = \int_0^{\infty} x^2 \frac{x^{\alpha-1}}{\Gamma(\alpha)\beta^{\alpha}} e^{-\frac{x}{\beta}} dx = \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} \beta^2 = \alpha(\alpha+1)\beta^2$$

Hence,

$$Var(X) = E\left(X^2\right) - \left[E\left(X\right)\right]^2 = \alpha\beta^2$$

(c) For $\alpha = 1$, we have $\Gamma(1) = 1$. Then,

$$f(x) = \begin{cases} \frac{1}{\beta} e^{-\frac{x}{\beta}} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

This is the probability density function of an $Exponential\left(\beta\right)$ random variable.

8.15 $0.06681,\ 0.24173,\ 0.38292,\ 0.2417,\ 0.06681$

8.16 Since $X \sim N\left(10, 16\right)$, then to find the 20th percentile we need to find c such that $P\left(X \leq c\right) = 0.2$ or $P\left(Z \leq \frac{c-10}{4}\right) = 0.2$ where $Z \sim N\left(0, 1\right)$.

From the inverse Normal table we have $P\left(Z \leq 0.8416\right) = 0.8$ which implies $P\left(Z \leq -0.8416\right) = 0.2$ which gives $\frac{c-10}{4} = -0.8416$ or $c = 6.6336$.

To find the 40th percentile we need to find c such that $P\left(X \leq c\right) = 0.4$ or $P\left(Z \leq \frac{c-10}{4}\right) = 0.4$.

From the inverse Normal table we have $P\left(Z \leq 0.2533\right) = 0.6$ which implies $P\left(Z \leq -0.2533\right) = 0.4$ which gives $\frac{c-10}{4} = -0.2533$ or $c = 8.9868$.

To find the 60th percentile we need to find c such that $P(X \leq c) = 0.6$ or $P\left(Z \leq \frac{c-10}{4}\right) = 0.6$.
From the inverse Normal table we have $P(Z \leq 0.2533) = 0.6$ which gives
$\frac{c-10}{4} = 0.2533$ or $c = 11.0132$.

To find the 80th percentile we need to find c such that $P(X \leq c) = 0.8$ or $P\left(Z \leq \frac{c-10}{4}\right) = 0.8$.
From the inverse Normal table we have $P(Z \leq 0.8416) = 0.8$ which gives
$\frac{c-10}{4} = 0.8416$ or $c = 13.3664$.

8.17 (a) Since $X \sim N\left(2, (0.01)^2\right)$

$$P(X < 2) = P\left(\frac{X-2}{0.01} < \frac{2-2}{0.01}\right) = P(Z < 0) = 0.5 \text{ where } Z \sim N(0,1)$$

(b) Find c such that

$$0.9 = P(|X-\mu| \leq c) = P\left(\frac{|X-\mu|}{0.01} \leq \frac{c}{0.01}\right)$$

Since $P(|Z| \leq 1.6449) = 0.9$ then $c = (0.01)(1.6449) = 0.016449$.
(c) $X \sim N\left(\mu, (0.01)^2\right)$. We want μ such that $P(X < 2) < 0.01$. Since

$$P(X < 2) = P\left(Z < \frac{2-\mu}{0.01}\right) \text{ where } Z \sim N(0,1)$$

and $P(Z < -2.3263) = 0.01$, therefore

$$\frac{2-\mu}{0.01} < -2.3263 \text{ or } \mu > 2.023263$$

8.18 Let X be the bolt's diameter. Then $X \sim N(1.2, (0.005)^2)$.

$$\begin{aligned} P(X > 1.21 \text{ or } X < 1.19) &= P(X > 1.21) + P(X < 1.19) \\ &= P\left(\frac{X-1.2}{0.005} > \frac{1.21-1.2}{0.005}\right) + P\left(\frac{X-1.2}{0.005} < \frac{1.19-1.2}{0.005}\right) \\ &= P(Z > 2) + P(Z < -2) \text{ where } Z \sim N(0,1) \\ &= 1 - P(Z < 2) + [1 - P(Z < 2)] \\ &= 2\left[1 - P(Z < 2)\right] = 2\left[1 - F(2)\right] = 2(1 - 0.97725) = 0.0455 \end{aligned}$$

8.19 The average wholesale price per egg is

$$\begin{aligned} &5P\left(Z < \frac{37-40}{2}\right) + 6P\left(\frac{37-40}{2} < Z < \frac{42-40}{2}\right) + 7P\left(Z > \frac{42-40}{2}\right) \\ &= 5P(Z < -1.5) + 6P(-1.5 < Z < 1) + 7P(Z > 1) \text{ where } Z \sim N(0,1) \\ &= 5(1 - 0.93319) + 6(0.84134 + 0.93319 - 1) + 7(1 - 0.84134) \\ &= 6.092 \text{ cents} \end{aligned}$$

8.20 Let $X =$ lifetime of computer chip. Then $X \sim N\left(5 \times 10^6, \left(5 \times 10^5\right)^2\right)$.

(a) Since

$$P\left(X \leq 6 \times 10^6\right) = P\left(Z \leq \frac{6 \times 10^6 - 5 \times 10^6}{5 \times 10^5}\right) \text{ where } Z \sim N(0,1)$$
$$= P\left(Z \leq 2\right) = 0.97725$$

the proportion of computer chips that last less than 6×10^6 hours is 0.97725.

(b) Since

$$P\left(X > 4 \times 10^6\right) = P\left(Z > \frac{4 \times 10^6 - 5 \times 10^6}{5 \times 10^5}\right) \text{ where } Z \sim N(0,1)$$
$$= P\left(Z > -2\right) = P\left(Z < 2\right) = 0.97725$$

the proportion of computer chips that last longer than 4×10^6 hours is 0.97725.

(c) Let $Y =$ the lifetime of improved computer chip. Then $Y \sim N\left(\mu_{new}, \left(5 \times 10^5\right)^2\right)$. The manufacturer wants

$$0.95 \leq P\left(Y > 4.5 \times 10^6\right) = P\left(Z > \frac{4.5 \times 10^6 - \mu_{new}}{5 \times 10^5}\right) \text{ where } Z \sim N(0,1)$$

Since $P\left(Z \leq 1.6449\right) = 0.95$, $P\left(Z > -1.6449\right) = 0.95$ the manufacturer should choose μ_{new} such that

$$\frac{4.5 \times 10^6 - \mu_{new}}{5 \times 10^5} \leq -1.6449$$

or

$$\mu_{new} \geq 4.5 \times 10^6 + 1.6449 \times 5 \times 10^5 = 4.5 \times 10^6 + 0.82249 \times 10^6 = 5.32245 \times 10^6$$

Therefore the new mean should be at least 5.32245×10^6.

8.21 (a) Let $X =$ temperature of CPU. Since $X \sim N\left(60, 5^2\right)$

$$P\left(X > 75\right) = P\left(Z > \frac{75 - 60}{5}\right) = P\left(Z > 3\right) = 1 - 0.99865 = 0.00135 \text{ where } Z \sim N(0,1)$$

(b) Since $P\left(Z < 1.2816\right) = 0.9$ where $Z \sim N(0,1)$, therefore

$$\frac{c - 60}{5} = 1.2816 \text{ or } c = 66.408$$

(c) Since $P\,(Z > 2.3263) = 0.01$ where $Z \sim N(0,1)$, therefore

$$\frac{95 - \mu_{new}}{5} = 2.3263 \text{ or } \mu_{new} = 83.3685$$

8.22 (a)

$$\begin{aligned} P\,(\text{false negative}) &= P\,(X < d) \text{ if } X \sim N\left(\mu_1, \sigma_1^2\right) \\ &= P\,(X < 5) \text{ if } X \sim N\left(10, (6)^2\right) \\ &= P\left(Z < \frac{5-10}{6}\right) \text{ where } Z \sim N\,(0,1) \\ &\approx P\,(Z < -0.83) = 1 - P\,(Z < 0.83) = 1 - 0.79673 = 0.20327 \end{aligned}$$

$$\begin{aligned} P\,(\text{false positive}) &= P\,(X \geq d) \text{ if } X \sim N\left(\mu_0, \sigma_0^2\right) \\ &= P\,(X \geq 5) \text{ if } X \sim N\left(0, (4)^2\right) \\ &= P\left(Z \geq \frac{5-0}{4}\right) \text{ where } Z \sim N\,(0,1) \\ &= 1 - P\,(Z < 1.25) = 1 - 0.89435 = 0.10565 \end{aligned}$$

(b)

$$\begin{aligned} P\,(\text{false negative}) &= P\,(X < 5) \text{ if } X \sim N\left(10, (3)^2\right) \\ &= P\left(Z < \frac{5-10}{3}\right) \text{ where } Z \sim N\,(0,1) \\ &\approx P\,(Z < -1.67) = 1 - P\,(Z < 1.67) = 1 - 0.95254 = 0.04746 \end{aligned}$$

$$\begin{aligned} P\,(\text{false positive}) &= P\,(X \geq 5) \text{ if } X \sim N\left(0, (3)^2\right) \\ &= P\left(Z \geq \frac{5-0}{3}\right) \text{ where } Z \sim N\,(0,1) \\ &= 1 - P\,(Z < 1.67) = 1 - 0.95254 = 0.04746 \end{aligned}$$

8.23 (a) False positive probabilities are $P(Z > d/3) = 0.0475$, 0.092, 0.023 for $Z \sim N\,(0,1)$ and $d = 5$, 4, 6 in (i), (ii), (iii).

False negative probabilities are $P(Z < (d-10)\,/3) = 0.0475$, 0.023, 0.092 for $Z \sim N\,(0,1)$ and $d = 5, 4, 6$ in (i), (ii), (iii).

(b) The factors are the security (proportion of spam in email) and proportion of legitimate messages that are filtered out.

8.24 (a) S_n is the time we need to wait until the n'th hit occurs.

(b) If $S_n \leq t$, then the n'th hit has happened sometime in $(0,t]$. Therefore, $X_t \geq n$, because X_t counts the number of hits in $(0,t]$. Conversely, if $X_t \geq n$, it means that the number of hits occurred up to time t is at least n. Therefore, the n'th hit has happened sometime in $(0,t]$, that is, $S_n \leq t$.

(c) We know that $X_t \sim Poisson(\lambda t)$. Therefore,

$$\begin{aligned} P(S_n \leq t) = P(X_t \geq n) &= 1 - P(X_t < n) \\ &= 1 - \sum_{j=0}^{n-1} P(X_t = j) \\ &= 1 - \sum_{j=0}^{n-1} \frac{e^{-\lambda t}(\lambda t)^j}{j!} \end{aligned}$$

(d) The probability density function of S_n is given by

$$\begin{aligned} f_{S_n}(t) &= \frac{d}{dt} P(S_n \leq t) \\ &= -\sum_{j=0}^{n-1} \left[-\lambda e^{-\lambda t} \frac{(\lambda t)^j}{j!} + j\lambda(\lambda t)^{j-1} \frac{e^{-\lambda t}}{j!} \right] \\ &= \lambda e^{-\lambda t} \left\{ 1 + \sum_{j=1}^{n-1} \left[\frac{(\lambda t)^j}{j!} - \frac{(\lambda t)^{j-1}}{(j-1)!} \right] \right\} \\ &= \lambda e^{-\lambda t} \left\{ 1 + \left[\frac{(\lambda t)^{n-1}}{(n-1)!} - \frac{(\lambda t)^0}{0!} \right] \right\} \\ &= \frac{\lambda^n t^{n-1} e^{-\lambda t}}{(n-1)!} \end{aligned}$$

This is for $t > 0$. For $t < 0$, we simply have $f_{S_n}(t) = 0$. Therefore,

$$f_{S_n}(t) = \begin{cases} \frac{\lambda^n t^{n-1} e^{-\lambda t}}{(n-1)!} & \text{if } t > 0 \\ 0 & \text{if } t < 0 \end{cases}$$

Noting that $\Gamma(n) = (n-1)!$, we have

$$f_{S_n}(t) = \begin{cases} \frac{\lambda^n t^{n-1} e^{-\lambda t}}{\Gamma(n)} & \text{if } t > 0 \\ 0 & \text{if } t < 0 \end{cases}$$

which is a Gamma random variable with parameters $\alpha = n$ and $\beta = 1/\lambda$.

8.25 (a) To show that $E\left(X\right)$ does not exist we need to show the improper integral

$$\int\limits_{-\infty}^{\infty} \frac{\alpha x}{\pi(\alpha^2+x^2)}dx$$

does not converge. By a change of variable

$$\int\limits_{-\infty}^{\infty} \frac{x}{\alpha^2+x^2}dx = \int\limits_{-\infty}^{\infty} \frac{x}{1+x^2}dx$$

For this integral to converge the integral

$$\int\limits_{1}^{\infty} \frac{x}{1+x^2}dx$$

must converge. Since for $x \geq 1$

$$\frac{x}{1+x^2} \geq \frac{x}{x^2+x^2} = \frac{1}{2}\left(\frac{1}{x}\right) \text{ and } \frac{1}{2}\int\limits_{1}^{\infty} \frac{1}{x}dx \text{ diverges}$$

therefore by the Comparison Test for Improper Integrals the integral $\int\limits_{1}^{\infty} \frac{x}{1+x^2}dx$ does not converge and thus $E\left(X\right)$ does not exist. It follows that the variance $Var\left(X\right) = E\left[\left(X-\mu\right)^2\right]$ does not exist since $E\left(X\right) = \mu$ does not exit.

(b) Let $Y = X^{-1}$. Show that Y has a Cauchy distribution with parameter α^{-1}. Since $P(Y \leq y) = P(X^{-1} \leq y)$ and X can be positive or negative, we have two cases:

Case 1: Let $y > 0$. Then,

$$P\left(X^{-1} \leq y\right) = P\left(X^{-1} \leq y, X > 0\right) + P\left(X^{-1} \leq y, X < 0\right) = P\left(X > \frac{1}{y}\right) + P\left(X < 0\right)$$

This follows since for $y > 0$, $\left\{X^{-1} \leq y, X > 0\right\} = \left\{X \geq \frac{1}{y}\right\}$ and $\left\{X^{-1} \leq y, X < 0\right\} = \left\{X < 0\right\}$. Also, since the probability density function of X is symmetric around the origin, $P(X < 0) = 0.5$. Therefore,

$$P\left(X^{-1} \leq y\right) = \frac{3}{2} - P\left(X \leq \frac{1}{y}\right)$$

Case 2: Let $y < 0$. In this case, $\left\{X^{-1} \leq y\right\} = \left\{y^{-1} \leq X < 0\right\}$. Then,

$$P\left(X^{-1} \leq y\right) = P\left(y^{-1} \leq X < 0\right) = P(X < 0) - P\left(X \leq y^{-1}\right) = 0.5 - P\left(X \leq y^{-1}\right)$$

Therefore

$$P(Y \leq y) = \begin{cases} 1.5 - P\left(X \leq y^{-1}\right) & y > 0 \\ 0.5 - P\left(X \leq y^{-1}\right) & y < 0 \end{cases}$$

For any $y \neq 0$,

$$\begin{aligned} f_Y(y) &= \frac{d}{dy}P(Y \leq y) = -\frac{d}{dy}P\left(X \leq y^{-1}\right) \\ &= y^{-2} f_X\left(y^{-1}\right) \\ &= \frac{\alpha}{\pi(\alpha^2 y^2 + 1)} = \frac{\alpha^{-1}}{\pi\left[y^2 + (\alpha^{-1})^2\right]} \end{aligned}$$

Therefore $Y = X^{-1}$ is a Cauchy random variable with parameter α^{-1}.

(c)

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(u)du = \int_{-\infty}^{x} \frac{\alpha}{\pi(\alpha^2 + u^2)} du \\ &= \lim_{b - -\infty} \left[\frac{1}{\pi} \arctan\left(\frac{u}{\alpha}\right)\right]_b^x \\ &= \frac{1}{\pi}\left[\arctan\left(\frac{x}{\alpha}\right) - \lim_{b--\infty} \arctan\left(\frac{b}{\alpha}\right)\right] \\ &= \frac{1}{\pi}\left[\arctan\left(\frac{x}{\alpha}\right) + \frac{\pi}{2}\right] \\ &= \frac{1}{2} + \frac{1}{\pi}\arctan\left(\frac{x}{\alpha}\right) \quad \text{for } x \in \Re \end{aligned}$$

Since F is increasing, the inverse cumulative distribution function is in fact the inverse function of F, so

$$F^{-1}(s) = \alpha \tan\left[\pi\left(s - \frac{1}{2}\right)\right], \quad s \in [0, 1]$$

(d) Suppose $U \sim U(-1, 0)$. We know that if $V \sim U(0, 1)$, then $F^{-1}(V)$ is a Cauchy random variable with parameter α. However if $U \sim U(-1, 0)$ then $1 + U \sim U(0, 1)$. Therefore, $g(U) = F^{-1}(1 + U) = \alpha \tan\left[\pi\left(U + 0.5\right)\right]$ is the desired function.

9. SOLUTIONS TO CHAPTER 9 PROBLEMS

9.1 (a) The marginal probability functions $f_1(x)$ and $f_2(y)$ are given in the table

	$f(x,y)$	x 0	1	2	$f_2(y) = P(Y=y)$
y	0	0.15	0.1	0.05	0.3
	1	0.35	0.2	0.15	0.7
	$f_1(x) = P(X=x)$	0.5	0.3	0.2	1

(b) X and Y are not independent random variables since

$$P(X=1, Y=0) = 0.1 \neq P(X=1)P(Y=0) = (0.3)(0.3) = 0.09$$

(c) $P(X > Y) = f(1,0) + f(2,0) + f(2,1) = 0.3$
(d) Conditional probability function of X given $Y = 0$:

x	0	1	2	Total
$P(X=x \mid Y=0)$	$\frac{0.15}{0.3} = \frac{3}{6}$	$\frac{0.1}{0.3} = \frac{2}{6}$	$\frac{0.05}{0.3} = \frac{1}{6}$	1

(e) Probability function of $T = X + Y$:

t	0	1	2	3	Total
$P(T=t)$	0.15	$0.1 + 0.35 = 0.45$	$0.05 + 0.2 = 0.25$	0.15	1

9.2 (a)

						y						
	$p(x,y)$	0	1	2	3	4	5	6	7	8	9	$p_1(x)$
	0	0.096	0	0	0	0	0	0.004	0	0	0	0.1
	1	0	0.1	0	0	0	0	0	0	0	0	0.1
	2	0	0	0.1	0	0	0	0	0	0	0	0.1
	3	0	0	0	0.1	0	0	0	0	0	0	0.1
x	4	0	0	0	0	0.098	0	0	0.002	0	0	0.1
	5	0	0	0	0	0	0.095	0	0	0	0.005	0.1
	6	0.004	0	0	0	0	0	0.096	0	0	0	0.1
	7	0	0	0	0	0.002	0	0	0.098	0	0	0.1
	8	0	0	0	0	0	0	0	0	0.1	0	0.1
	9	0	0	0	0	0	0.005	0	0	0	0.095	0.1
	$p_2(y)$	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1

Since $P(X=0,Y=0)=0.096 \neq P(X=0)\,P(Y=0)=(0.1)(0.1)=0.01$, therefore X and Y are not independent random variables.

(b)

$$\begin{aligned} P(X=Y) &= \sum_{\substack{(x,y):\\ x=y}} p(x,y) \\ &= 2(0.096)+4(0.1)+2(0.098)+2(0.095) \\ &= 0.978 \end{aligned}$$

(c) P(number 5 is identified incorrectly|number is a five) $= p(5,9)/0.1 = 0.005/0.1 = 0.05$.

9.3 (a) The joint probability function of X and Y is

$$\frac{\binom{2}{x}\binom{1}{y}\binom{7}{3-x-y}}{\binom{10}{3}} \quad x=0,1,2; \quad y=0,1; \quad x+y\leq 3$$

(b) The marginal probability function of X is

$$f_1(x)=\frac{\binom{2}{x}\binom{8}{3-x}}{\binom{10}{3}} \quad x=0,1,2$$

The marginal probability function of Y is

$$f_2(y)=\frac{\binom{1}{y}\binom{9}{3-y}}{\binom{10}{3}} \quad y=0,1$$

(c)

$$\begin{aligned}
P(X=Y) &= \sum_{\substack{(x,y):\\ x=y}} f(x,y) = \sum_{\substack{(x,y):\\ x=y}} P(X=x,Y=y) \\
&= P(X=0,Y=0) + P(X=1,Y=1) \\
&= \frac{\binom{2}{0}\binom{1}{0}\binom{7}{3}}{\binom{10}{3}} + \frac{\binom{2}{1}\binom{1}{1}\binom{7}{1}}{\binom{10}{3}} = \frac{49}{120}
\end{aligned}$$

(d)

$$P(X=1|Y=0) = \frac{P(X=1,Y=0)}{P(Y=0)} = \frac{\frac{\binom{2}{1}\binom{1}{0}\binom{7}{2}}{\binom{10}{3}}}{\frac{\binom{1}{0}\binom{9}{3}}{\binom{10}{3}}} = \frac{1}{2}$$

9.4 (a) The event "x yellow balls on 1st 2 draws and y yellow balls on 4 draws" only happens if x yellow balls are drawn on the first 2 draws and the remaining $y-x$ yellow balls are drawn on the last 2 draws. The joint probability function of X and Y is

$$\begin{aligned}
&P(X=x,Y=y) \\
&= P(x \text{ yellow balls on 1st 2 draws and } y-x \text{ yellow balls on last 2 draws}) \\
&= P(x \text{ yellow balls on 1st 2 draws}) \\
&\times P(y-x \text{ yellow balls on last 2 draws } \textbf{given } x \text{ yellow balls on 1st 2 draws}) \\
&= \frac{\binom{5}{x}\binom{3}{2-x}}{\binom{8}{2}} \frac{\binom{5-x}{y-x}\binom{3-(2-x)}{2-(y-x)}}{\binom{6}{2}} \\
&= \frac{\binom{5}{x}\binom{3}{2-x}\binom{5-x}{y-x}\binom{x+1}{2+x-y}}{\binom{8}{2}\binom{6}{2}} \quad \text{for } x=0,1,2;\;\; y=\max(1,x),x+1,x+2
\end{aligned}$$

(b) Since $X =$ number of yellow balls in first 2 draws without replacement, the marginal distribution of X is Hypergeometric with marginal probability function

$$P(X=x) = \frac{\binom{5}{x}\binom{3}{2-x}}{\binom{8}{2}} \quad x=0,1,2$$

Similarly since $Y =$ number of yellow balls in all 4 draws without replacement, the marginal probability function of Y is

$$P(Y=y) = \frac{\binom{5}{y}\binom{3}{4-y}}{\binom{8}{4}} \quad y=1,2,3,4$$

Since

$$P(X=0,Y=3) = 0 \neq P(X=0)\,P(Y=3) = \frac{\binom{3}{2}}{\binom{8}{2}} \frac{\binom{5}{3}\binom{3}{1}}{\binom{8}{4}}$$

therefore X and Y are not independent random variables.

9.5 (a)

$$P(X+Y>1)=\sum_{\substack{(x,y):\\ x+y>1}}P(X=x,Y=y)=1-\sum_{\substack{(x,y):\\ x+y\leq 1}}P(X=x,Y=y)$$
$$=1-[P(X=0,Y=0)+P(X=0,Y=1)+P(X=1,Y=0)]$$
$$=1-P(X=0)P(Y=0)-P(X=0)P(Y=1)$$
$$-P(X=1)P(Y=0)$$

since X and Y are independent random variables. Since $X \sim Poisson\,(0.1)$ and $Y \sim Poisson\,(0.05)$

$$P(X+Y>1)$$
$$=1-\left(e^{-0.1}\right)\left(e^{-0.05}\right)-\left(e^{-0.1}\right)\frac{(0.05)^1\,e^{-0.05}}{1!}-\frac{(0.1)^1\,e^{-0.1}}{1!}\left(e^{-0.05}\right)$$
$$=1-e^{-0.15}\,(1+0.05+0.1)=1-1.15e^{-0.15}$$

(b) $E(X+Y)=E(X)+E(Y)=0.1+0.05=0.15$ and $Var(X+Y)=Var(X)+Var(Y)=0.1+0.05=0.15$.

9.6 (a) Note that

$$f(x,y)=P(X=x,Y=y)=\frac{2^{x+y}e^{-4}}{x!y!}=\frac{2^xe^{-2}}{x!}\times\frac{2^ye^{-2}}{y!}$$

for $x=0,1,\ldots$
and $y=0,1,\ldots$

We recognize $\frac{2^xe^{-2}}{x!}$ as the probability function of a $Poisson\,(2)$ random variable. Therefore $X \sim Poisson\,(2)$ and $Y \sim Poisson\,(2)$ independently.

(b) Since $X \sim Poisson\,(2)$ and $Y \sim Poisson\,(2)$ independently, therefore $X+Y \sim Poisson\,(4)$ by Theorem 29.

9.7 First note that

$$P(Y=y)=\sum_{\text{all } x}P(X=x,Y=y)$$
$$=\sum_{\text{all } x}P(Y=y|X=x)P(X=x) \quad \text{by the Product Rule}$$
$$=\sum_{\text{all } x}f_2(y|x)\,f_1(x)$$

Since $X \sim Poisson\,(\mu)$
$$f_1(x) = \frac{\mu^x e^{-\mu}}{x!} \text{ for } x = 0, 1, \ldots.$$
Also, for a given number x of defective items produced, the number, Y, detected has a Binomial distribution with $n = y$ and $p = 0.9$, assuming each inspection takes place independently so
$$f\,(y|x) = \binom{x}{y} (0.9)^y \, (0.1)^{x-y} \text{ for } y = 0, 1, \ldots, x$$

Therefore
$$\begin{aligned} f(x,y) &= f_1(x) f_2\,(y|x) \\ &= \frac{\mu^x e^{-\mu}}{x!} \frac{x!}{y!(x-y)!} (0.9)^y (0.1)^{x-y} \\ &\text{for } \begin{matrix} y = 0, 1, \ldots, x \\ x = 0, 1, \ldots \end{matrix} \quad \text{or} \quad \begin{matrix} x = y, y+1, \ldots \\ y = 0, 1, \ldots \end{matrix} \end{aligned}$$

To get $f_1\,(x|y)$ we need $f_2(y)$. We have
$$f_2(y) = \sum_{\text{all } x} f(x,y) = \sum_{x=y}^{\infty} \frac{\mu^x e^{-\mu}}{y!(x-y)!} (0.9)^y (0.1)^{x-y}$$
($x \geq y$ since the number of defective items produced can't be less than the number detected)
$$= \frac{(0.9)^y e^{-\mu}}{y!} \sum_{x=y}^{\infty} \frac{\mu^x (0.1)^{x-y}}{(x-y)!}$$

Then
$$\begin{aligned} f_2(y) &= \frac{(0.9\mu)^y e^{-\mu}}{y!} \sum_{x=y}^{\infty} \frac{(0.1\mu)^{x-y}}{(x-y)!} \\ &= \frac{(0.9\mu)^y e^{-\mu}}{y!} \left[\frac{(0.1\mu)^0}{0!} + \frac{(0.1\mu)^1}{1!} + \frac{(0.1\mu)^2}{2!} + \cdots \right] \\ &= \frac{(0.9\mu)^y e^{-\mu}}{y!} e^{0.1\mu} \quad \text{by the Exponential series} \\ &= \frac{(0.9\mu)^y e^{-0.9\mu}}{y!} \quad \text{for } y = 0, 1, \ldots \end{aligned}$$

Therefore
$$\begin{aligned} f_1\,(x|y) &= \frac{f(x,y)}{f_2(y)} = \frac{\frac{\mu^x e^{-\mu} (0.9)^y (0.1)^{x-y}}{y!((x-y)!}}{\frac{(0.9)^y \mu^y e^{-.9\mu}}{y!}} \\ &= \frac{(0.1\mu)^{x-y} e^{-0.1\mu}}{(x-y)!} \quad \text{for } x = y, y+1, y+2, \ldots \end{aligned}$$

9.8 (a) The probability that string 1 breaks exactly x times is

$$P(X = x) = \binom{5}{x}\left(\frac{1}{6}\right)^x\left(\frac{5}{6}\right)^{5-x}$$

that is, $X \sim Binomial(5, \frac{1}{6})$.

The probability string 1 breaks three or more times is

$$P\left(X \geq 3\right) = \textstyle\sum_{x=3}^{5} P(X = x) = \displaystyle\binom{5}{x}\left(\frac{1}{6}\right)^x\left(\frac{5}{6}\right)^{5-x} = \frac{23}{648}$$

The probability is the same for each of strings $2, 3, \ldots, 6$, so the probability that any one string breaks three or more times is

$$6\sum_{x=3}^{5}\binom{5}{x}\left(\frac{1}{6}\right)^x\left(\frac{5}{6}\right)^{5-x} = \frac{23}{108} = 0.21296$$

(b) Let $Y =$ number of strings that break during a concert. Then $Y \sim Poisson\,(1)$. Suppose the guitarist has n packages of strings and the total number of breaks is y.

If $n \geq y$, then the probability that any one string breaks more than n times is 0.

If $n < y < 2n$, then the probability that any one string breaks more than n times is

$$6\sum_{x=n+1}^{y}\binom{y}{x}\left(\frac{1}{6}\right)^x\left(\frac{5}{6}\right)^{y-x}$$

If $2n \leq y$, probability that any one string breaks more than n times is less than or equal to

$$6\sum_{x=n+1}^{y}\binom{y}{x}\left(\frac{1}{6}\right)^x\left(\frac{5}{6}\right)^{y-x}$$

Let $X =$ maximum number of breaks of any one string. Then

$$\begin{aligned}
P(X > n) &= \sum_{y=0}^{\infty} P(X > n|Y = y)P(Y = y) \\
&\leq 6e^{-1}\sum_{y=n+1}^{\infty}\sum_{x=n+1}^{y}\binom{y}{x}\left(\frac{1}{6}\right)^x\left(\frac{5}{6}\right)^{y-x}\frac{1}{y!} \\
&= 6e^{-1}\sum_{x=n+1}^{\infty}\left(\frac{1}{6}\right)^x\frac{1}{x!}\sum_{y=x+1}^{\infty}\left(\frac{5}{6}\right)^{y-x}\frac{1}{(y-x)!}
\end{aligned}$$

If we replace $y - x$ by j this becomes

$$6e^{-1}\sum_{x=n+1}^{\infty}\left(\frac{1}{6}\right)^x\frac{1}{x!}\sum_{j=1}^{\infty}\left(\frac{5}{6}\right)^j\frac{1}{j!}$$

and the last summation is the exponential series expansion of $e^{5/6}$ with the first term 1 missing. This becomes

$$6e^{-1}\sum_{x=n+1}^{\infty}\left(\frac{1}{6}\right)^{x}\frac{1}{x!}(e^{5/6}-1) = 6(1-e^{-5/6})\sum_{x=n+1}^{\infty}\left(\frac{1}{6}\right)^{x}\frac{1}{x!}e^{-1/6}$$
$$= 6(1-e^{-5/6})\left(1-\sum_{x=0}^{n}\left(\frac{1}{6}\right)^{x}\frac{1}{x!}e^{-1/6}\right)$$

This is approximately 3.3924 times the probability that a Poisson$\left(\frac{1}{6}\right)$ random variable is greater than n. This needs to be less than 0.01. When $n=1$ we obtain 0.042 and when $n=2$ we obtain

$$6(1-e^{-5/6})\left(1-\sum_{x=0}^{1}\left(\frac{1}{6}\right)^{x}\frac{1}{x!}e^{-1/6}\right) = 2.311\,2\times10^{-3}$$

so that $n=2$ packages is sufficient.

9.9 (a) The conditional probability function of X given $N=n$ is

$$P(X=x|N=n) = \binom{n}{x}s^{x}(1-s)^{n-x}$$

Therefore

$$\begin{aligned}
P(X=x) &= \sum_{n=0}^{\infty}P(X=x|N=n)P(N=n)\\
&= \sum_{n=x}^{\infty}\binom{n}{x}s^{x}(1-s)^{n-x}\frac{R_o^n}{n!}e^{-R_0}\\
&= \frac{s^x R_o^x}{x!}e^{-R_0}\sum_{n=x}^{\infty}\frac{[(1-s)R_o]^{n-x}}{(n-x)!}\\
&= \frac{s^x R_o^x}{x!}e^{-R_0}e^{(1-s)R_0}\\
&= \frac{(sR_0)^x}{x!}e^{-sR_0} \quad \text{for } x=0,1,2,...
\end{aligned}$$

which is the probability function of a $Poisson\,(sR_0)$ random variable.

(b) In order that $E(X)$ be less than one we require

$$\begin{aligned}
E(X) = sR_0 &= (1-h)R_0 < 1\\
\text{or }\ 1-h &< \frac{1}{R_0}\\
\text{or }\ h &> 1-\frac{1}{R_0}
\end{aligned}$$

This critical value $1 - \frac{1}{R_0}$ of h is the fraction of the population that is required to have immunity (through vaccinations for example) in order that the number of infected in the population will not increase. In the special case that $R_0 = 2.5$ we would require $h > 0.60$. Since X is Poisson(SR_0), $Var(X) = sR_0$.

(c) The expected number of new cases after one period is

$$x_1 = x_0 R_o s_0 = 25$$

At the end of the first period the number of susceptibles is reduced by the number of these new infections, that is the new value of s is

$$s_1 = s_0\left(1 - \frac{x_1}{m}\right) = 1\left(1 - \frac{25}{10000}\right) = 0.9975$$

Therefore the expected number of new infections in the second period is

$$x_2 = x_1 R_0 s_1 = 25(2.5)\,(0.9975) = 62.344$$

Again we reduce the number of susceptibles by these new infections,

$$s_2 = s_1\left(1 - \frac{x_2}{m}\right) = (0.9975)\left(1 - \frac{62.344}{10000}\right) = 0.99118$$

The expected number of new infections in the next period is

$$x_3 = x_2 R_0 s_2 = x_1 R_0^2 s_1 s_2 = 62.344(2.5)\,(0.99118) = 154.49$$

$$s_3 = s_2\left(1 - \frac{x_3}{m}\right) = 0.99118\left(1 - \frac{154.49}{10000}\right) = 0.97587$$

and finally

$$x_4 = x_3 R_0 s_3 = 154.49(2.5)\,(0.97587) = 376.91$$

in general

$$x_k = x_0 R_0^k s_0 s_1 s_2 \cdots s_{k-1}$$

Notice that this is an approximately exponential increase until the fraction of susceptibles are substantially less than one. Indeed when $R_0 = 2.5$ and $s_0 = x_0 = 1$ and $m = 10000$, we have a decrease in the number of infections $x_k < x_{k-1}$ if and only if $R_0 s_{k-1} < 1$ or

$$s_{k-1} < \frac{1}{R_0}$$

that is, the fraction of susceptibles after $k-1$ periods is less than or equal $\frac{1}{R_0}$.

9.10 Let $X_i =$ the number of offspring of type i in a sample of size 40, $i = 1, 2, 3, 4$. Then

$$\begin{aligned}
&P\left(X_1 = x_1, X_2 = x_2, X_3 = x_3, X_4 = x_4\right) \\
&= \frac{40!}{x_1!x_2!x_3!x_4!}\left(\frac{3}{16}\right)^{x_1}\left(\frac{5}{16}\right)^{x_2}\left(\frac{5}{16}\right)^{x_3}\left(\frac{3}{16}\right)^{x_4} \\
&x_i = 0, 1, \ldots;\ i = 1, 2, 3, 4 \text{ and } x_1 + x_2 + x_3 + x_4 = 40
\end{aligned}$$

(a)

$$\begin{aligned}
&P\left(X_1 = 10, X_2 = 10, X_3 = 10, X_4 = 10\right) \\
&= \frac{40!}{10!10!10!10!}\left(\frac{3}{16}\right)^{10}\left(\frac{5}{16}\right)^{10}\left(\frac{5}{16}\right)^{10}\left(\frac{3}{16}\right)^{10} \\
&= \frac{40!}{(10!)^4}\left(\frac{3}{16}\right)^{20}\left(\frac{5}{16}\right)^{20}
\end{aligned}$$

(b) The probability of a type 1 or type 2 offspring is $\frac{3}{16} + \frac{5}{16} = \frac{1}{2}$. Therefore $X_1 + X_2 \sim Binomial\left(40, \frac{1}{2}\right)$ and

$$P\left(X_1 + X_2 = 16\right) = \binom{40}{16}\left(\frac{1}{2}\right)^{40}$$

(c)

$$\begin{aligned}
P\left(X_1 = 10 | X_1 + X_2 = 16\right) &= \frac{P\left(X_1 = 10, X_1 + X_2 = 16\right)}{P\left(X_1 + X_2 = 16\right)} \\
&= \frac{P\left(X_1 = 10, X_2 = 6\right)}{P\left(X_1 + X_2 = 16\right)} \\
&= \frac{\frac{40!}{10!6!24!}\left(\frac{3}{16}\right)^{10}\left(\frac{5}{16}\right)^{6}\left(\frac{8}{16}\right)^{24}}{\frac{40!}{16!24!}\left(\frac{8}{16}\right)^{40}} \\
&= \binom{16}{10}\left(\frac{3}{8}\right)^{10}\left(\frac{5}{8}\right)^{6}
\end{aligned}$$

9.11 Let X = number of bacteria in 50 cubic centimeters of water. Then X has a $Poisson(2.5)$ distribution.

$$P(X = x) = \frac{(2.5)^x e^{-2.5}}{x!} \quad x = 0, 1, \ldots$$

Then if (X_0, X_1, X_{2+}) represent the number of samples with $0, 1,$and 2 or more bacteria in the five samples, having respectively probabilities $e^{-2.5}$, $2.5e^{-2.5}$ and $1 - 3.5e^{-2.5}$, we have

$$P(X_0 = 1, X_1 = 2, X_{2+} = 2) = \frac{5!}{1!2!2!}\left(e^{-2.5}\right)^1\left(2.5e^{-2.5}\right)^2\left(1 - 3.5e^{-2.5}\right)^2$$

9.12 Let $X =$ the lifetime of a light bulb. Then $X \sim Exponential\,(1000)$.

(a)

$$\begin{aligned} P\left(X < 500\right) &= 1 - e^{-500/1000} = 1 - e^{-0.5} \\ P\left(500 < X < 1000\right) &= e^{-0.5} - e^{-1} \\ P\left(1000 < X < 1500\right) &= e^{-1} - e^{-1.5} \\ P\left(X > 1500\right) &= e^{-1.5} \end{aligned}$$

(b) Let A be the event: 15 light bulbs last less than 500 hours, 15 light bulbs last between 500 and 1000 hours, and 10 light bulbs last between 1000 and 1500 hours.

$$P\left(A\right) = \frac{50!}{15!15!10!10!}\left(1 - e^{-0.5}\right)^{15}\left(e^{-0.5} - e^{-1}\right)^{15}\left(e^{-1} - e^{-1.5}\right)^{10}\left(e^{-1.5}\right)^{10}$$

(c)

$$\begin{aligned} &P\left(10 \text{ or more light bulbs last longer than } 1500 \text{ hours}\right) \\ &= 1 - \sum_{y=0}^{9}\binom{50}{y}\left(e^{-1.5}\right)^{y}\left(1 - e^{-1.5}\right)^{50-y} \end{aligned}$$

9.13 From Chapter 8, Problem 15 we have $P\left(A\right) = 0.06681$, $P\left(B\right) = 0.24173$, $P\left(C\right) = 0.38292$, $P\left(D\right) = 0.2417$, $P\left(F\right) = 0.06681$.

(a)

$$\begin{aligned} &P\left(5\ A\text{'s, } 15\ B\text{'s, } 10\ C\text{'s and } 15\ D\text{'s}\right) \\ &= \frac{50!}{5!15!10!15!5!}\left(0.06681\right)^{5}\left(0.24173\right)^{15}\left(0.38292\right)^{10}\left(0.24173\right)^{15}\left(0.06681\right)^{5} \end{aligned}$$

(b)

$$\begin{aligned} &P\left(\text{ at least } 45 \text{ students have marks above an } F\right) \\ &= \sum_{y=45}^{50}\binom{50}{y}\left(0.93319\right)^{y}\left(0.06681\right)^{50-y} \end{aligned}$$

(c) Let $X =$ number of students who receive A's and let $Y =$ the number of students that receive B's in a class of 50 students. Then the joint probability function of X and Y is

$$\begin{aligned} P\left(X = x, Y = y\right) &= \frac{50!}{x!y!\left(50 - x - y\right)!}\left(0.06681\right)^{x}\left(0.24173\right)^{y}\left(0.69146\right)^{50-x-y} \\ & x, y = 0, 1, \ldots;\ \ x + y \leq 50 \end{aligned}$$

9.14 (a)

$$P(X_1 = x_1, \ldots, X_6 = x_6) = \frac{10!}{x_1! x_2! \cdots x_6!} (0.1)^{x_1} (0.05)^{x_2} (0.05)^{x_3} (0.15)^{x_4} (0.15)^{x_5} (0.5)^{x_6}$$
$$x_i = 0, 1, \ldots; \;\; x_1 + x_2 + \cdots + x_6 = 10$$

(b)

$$\begin{aligned}
&P(\text{at least one apartment fire given } 4 \text{ fire-related calls}) \\
&= P(X_3 \geq 1 | X_1 + X_2 + X_3 + X_4 = 4) = 1 - P(X_3 = 0 | X_1 + X_2 + X_3 + X_4 = 4) \\
&= 1 - \frac{P(X_3 = 0, X_1 + X_2 + X_3 + X_4 = 4)}{P(X_1 + X_2 + X_3 + X_4 = 4)} = 1 - \frac{P(X_3 = 0, X_1 + X_2 + X_4 = 4)}{P(X_1 + X_2 + X_3 + X_4 = 4)} \\
&= 1 - \frac{\frac{10!}{0!4!6!}(0.05)^0 (0.1 + 0.05 + 0.15)^4 (0.65)^6}{\frac{10!}{4!6!}(0.35)^4 (0.65)^6} = 1 - \frac{(0.3)^4}{(0.35)^4} = 1 - \left(\frac{6}{7}\right)^4
\end{aligned}$$

(c) Since $X_i \sim Binomial(10, p_i)$ then $E(X_i) = 10p_i$. The total cost T is given by

$$T = 100(5X_1 + 5X_2 + 7X_3 + 20X_4 + 4X_5 + 2X_6).$$

The expected cost is

$$\begin{aligned}
E(T) &= 100[5E(X_1) + 5E(X_2) + 7E(X_3) + 20E(X_4) + 4E(X_5) + 2E(X_6)] \\
&= 100(10)[5(0.1) + 5(0.05) + 7(0.05) + 20(0.15) + 4(0.15) + 2(0.5)] \\
&= 5700 \text{ dollars}
\end{aligned}$$

9.15 (a) The joint probability function of X and Y is

$$P(X = x, Y = y) = \frac{(9 + x + y)!}{x! y! 9!} p^x q^y (1 - p - q)^{10} \quad x, y = 0, 1, 2, \ldots$$

(b)

$$\begin{aligned}
P(X = x) &= \sum_{y=0}^{\infty} \frac{(9 + x + y)!}{x! y! 9!} p^x q^y (1 - p - q)^{10} \\
&= \frac{(9 + x)!}{x! 9!} p^x (1 - p - q)^{10} \sum_{y=0}^{\infty} \frac{(9 + x + y)!}{y! (9 + x)!} q^y \\
&= \frac{\binom{9+x}{x} p^x (1 - p - q)^{10}}{(1 - q)^{10+x}} \sum_{y=0}^{\infty} \binom{(10 + x) + y - 1}{y} q^y (1 - q)^{10+x} \\
&= \binom{9+x}{x} p^x (1 - p - q)^{10} (1 - q)^{-10-x} \quad x = 0, 1, \ldots
\end{aligned}$$

where the sum is equal to one since

$$\sum_{y=0}^{\infty} \binom{k+y-1}{y} q^y (1-q)^k = 1$$

because it is a sum over all values for a Negative Binomial probability function.

(c) The conditional probability function of Y given $X = x$ is

$$P(Y = y|X = x) = \frac{P(X = x, Y = y)}{P(X = x)} = \frac{\frac{(9+x+y)!}{x!y!9!} p^x q^y (1-p-q)^{10}}{\frac{(9+x)!}{x!9!} p^x (1-p-q)^{10} (1-q)^{-10-x}}$$
$$= \binom{(10+x)+y-1}{y} q^y (1-q)^{x+10} \quad y = 0, 1, 2, \ldots$$

which we recognize as a Negative Binomial probability function with $k = 10 + x$ and $p = 1 - q$.

9.14 (a)

$$P(X_1 = 0, X_2 = 2, X_3 = 0, X_4 = 1, X_5 = 3, X_6 = 1)$$
$$= P(X_1 = 0) P(X_2 = 2) P(X_3 = 0) P(X_4 = 1) P(X_5 = 3) P(X_6 = 1)$$
$$= \left(\frac{1^0 e^{-1}}{0!}\right)\left(\frac{1^2 e^{-1}}{2!}\right)\left(\frac{1^0 e^{-1}}{0!}\right)\left(\frac{1^1 e^{-1}}{1!}\right)\left(\frac{1^3 e^{-1}}{3!}\right)\left(\frac{1^1 e^{-1}}{1!}\right)$$
$$= \frac{e^{-6}}{12} \text{ for } \theta > 0$$

(b)

$$P(X_1 = 0, X_2 = 2, X_3 = 0, X_4 = 1, X_5 = 3, X_6 = 1)$$
$$= P(X_1 = 0) P(X_2 = 2) P(X_3 = 0) P(X_4 = 1) P(X_5 = 3) P(X_6 = 1)$$
$$= \left(\frac{\theta^0 e^{-\theta}}{0!}\right)\left(\frac{\theta^2 e^{-\theta}}{2!}\right)\left(\frac{\theta^0 e^{-\theta}}{0!}\right)\left(\frac{\theta^1 e^{-\theta}}{1!}\right)\left(\frac{\theta^3 e^{-\theta}}{3!}\right)\left(\frac{\theta^1 e^{-\theta}}{1!}\right)$$
$$= \frac{\theta^7 e^{-6\theta}}{12} \quad \text{for } \theta > 0$$

9.17

	$f(x,y)$	x: 0	1	2	$f_2(y)$
y	0	0.15	0.1	0.05	0.3
	1	0.35	0.2	0.15	0.7
	$f_1(x)$	0.5	0.3	0.2	1

$$\begin{aligned}
E(X) &= 0 + 0.3 + 2(0.2) = 0.7, \;\; E\left(X^2\right) = 0 + 0.3 + (2)^2(0.2) = 1.1, \\
Var(X) &= 1.1 - (0.7)^2 = 0.61 \\
E(Y) &= 0.7, \;\; E\left(Y^2\right) = 0.7, \;\; Var(Y) = 0.7 - (0.7)^2 = 0.21 \\
E(XY) &= 0.2 + 2(0.15) = 0.5 \\
Cov(X,Y) &= 0.5 - (0.7)(0.7) = 0.01 \\
\rho &= \frac{0.01}{\sqrt{(0.61)(0.21)}} = 0.02794
\end{aligned}$$

9.18 Note that

$$\begin{aligned}
Cov(X,Y) &= \rho\sqrt{Var(X)\,Var(Y)} \\
&= (-0.7)\sqrt{(13)(34)} \\
&= (-0.7)\sqrt{442}
\end{aligned}$$

$$\begin{aligned}
Var(X-2Y) &= Var(X) + (-2)^2\,Var(Y) + 2(1)(-2)\,Cov(X,Y) \\
&= 13 + 4(34) - 4(-0.7)\sqrt{442} = 149 + 2.8\sqrt{442} \\
&= 207.867
\end{aligned}$$

9.19

$$\begin{aligned}
Cov(X+Y, X-Y) &= Var(X) - Cov(X,Y) + Cov(X,Y) - Var(Y) \\
&= Var(X) - Var(Y) = 1 - 2 \\
&= -1
\end{aligned}$$

9.20 (a)

$$\begin{aligned}
E(U) &= E(X+Y) = E(X) + E(Y) = 2(0.5) + 2(0.5) = 2 \\
E(V) &= E(X-Y) = E(X) - E(Y) = 2(0.5) - 2(0.5) = 0
\end{aligned}$$

(b)

$$\begin{aligned}
Var(U) &= Var(X+Y) = Var(X) + Var(Y) = 2(0.5)^2 + 2(0.5)^2 = 1 \\
Var(V) &= Var(X-Y) = Var(X) + Var(Y) = 2(0.5)^2 + 2(0.5)^2 = 1
\end{aligned}$$

(c)

$$Cov(U,V) = Cov(X+Y, X-Y) = Var(X) - Var(Y) = 0$$

Although $Cov(U,V) = 0$, U and V are not independent since $P(U=0) \neq 0; P(V=1) \neq 0$ but $P(U=0, V=1) = 0$.

9.21 (a) For a trinomial distribution there are three possible outcomes. Call these possible outcomes A, B and C where $P(A) = p$, $P(B) = q$ and $P(C) = 1 - p - q$. Since the joint distribution of X and Y is given as

$$f(x,y) = \frac{n!}{x!y!(n-x-y)!} p^x q^y (1-p-q)^{n-x-y} \quad \text{for} \quad \begin{array}{l} x = 0, 1, \ldots, n \\ y = 0, 1, \ldots, n \\ \text{and } x + y \leq n \end{array}$$

then the random variable X counts the number of times outcome A occurs and Y counts the number of times outcome B occurs. Therefore the random variable $T = X + Y$ counts the number of times outcome A or B occurs. Since $P(A \cup B) = p + q$ then
$T = X + Y \sim Binomial(n, p + q)$.

(b) Since $T = X + Y \sim Binomial(n, p + q)$ then $E(T) = n(p + q)$ and
$Var(T) = n(p + q)(1 - p - q)$.

(c) The variance of T can also be written as

$$\begin{aligned} Var(T) &= Var(X + Y) \\ &= Var(X) + Var(Y) + 2Cov(X, Y) \end{aligned}$$

Solving for $Cov(X, Y)$, we obtain

$$Cov(X, Y) = \frac{1}{2}\left[Var(T) - Var(X) - Var(Y)\right]$$

Since the marginal distributions of a Multinomial distribution are Binomial we know
$X \sim Binomial(n, p)$ and $Y \sim Binomial(n, q)$ and thus $Var(X) = np(1 - p)$ and
$Var(Y) = nq(1 - q)$. Therefore

$$\begin{aligned} Cov(X, Y) &= \frac{1}{2}\left[Var(T) - Var(X) - Var(Y)\right] \\ &= \frac{1}{2}\left[n(p + q)(1 - p - q) - np(1 - p) - nq(1 - q)\right] \\ &= \frac{1}{2}(-npq - npq) = -npq \end{aligned}$$

We would expect the covariance to be negative since we know that if X is large (number of A outcomes is large) then Y must be small (number of B outcomes is small) since the total number of trials n is fixed.

9.22 (a)

$$\begin{aligned} &E(X) \\ &= 0(0.2) + 1(0.25) + 2(0.35) + 3(0.1) + 4(0.05) + 5(0.02) + 6(0.02) + 7(0.01) + 8(0.01) \\ &= 1.76 \end{aligned}$$

(b)

$$
\begin{aligned}
P(Y=y) &= \sum_{x=y}^{8} \binom{x}{y} \left(\frac{1}{2}\right)^{x} f(x) \\
E(Y) &= \sum_{y=0}^{8} \sum_{x=y}^{8} y \binom{x}{y} \left(\frac{1}{2}\right)^{x} f(x) \quad \text{change the order of summation} \\
&= \sum_{x=0}^{8} f(x) \left[\sum_{y=0}^{x} y \binom{x}{y} \left(\frac{1}{2}\right)^{x} \right] \quad \text{term in } [\ ] \text{ is mean of } Binomial\left(x, \frac{1}{2}\right) \text{ r.v.} \\
&= \sum_{x=0}^{8} f(x) \left[x \left(\frac{1}{2}\right) \right] \\
&= \frac{1}{2} \sum_{x=0}^{8} x f(x) \\
&= \frac{1}{2} E(X) = \frac{1}{2}(1.76) = 0.88
\end{aligned}
$$

9.23

$$
E[g(X,Y)] = \sum_{\text{all }(x,y)} g(x,y) f(x,y) \le \sum_{\text{all }(x,y)} b f(x,y) = b
$$

Similarly $E[g(X,Y)] \ge a$.

9.24 The optimal weights are

$$
\begin{aligned}
&w_1 = \frac{1}{c\sigma_1^2},\;\; w_2 = \frac{1}{c\sigma_2^2},\;\; w_3 = \frac{1}{c\sigma_3^2} \text{ where } c = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} + \frac{1}{\sigma_3^2} \\
&\text{and } \sigma_1 = 0.2, \quad \sigma_2 = 0.3, \quad \sigma_3 = 0.4
\end{aligned}
$$

9.25

$$
E(\bar{X}) = \frac{1}{n} \sum_{i=1}^{n} E(X_i) = \frac{1}{n} \sum_{i=1}^{n} \theta = \frac{1}{n}(n\theta) = \theta
$$

and since the X_i's are independent random variables

$$
\begin{aligned}
Var(\bar{X}) &= \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n} Var(X_i) \\
&= \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n} \theta = \left(\frac{1}{n}\right)^2 (n\theta) = \frac{\theta}{n} \to 0 \quad \text{as } n \to \infty
\end{aligned}
$$

9.26

$$
E(\bar{X}) = \frac{1}{n} \sum_{i=1}^{n} E(X_i) = \frac{1}{n} \sum_{i=1}^{n} \frac{1-\theta}{\theta} = \frac{1}{n}\left[n \left(\frac{1-\theta}{\theta} \right) \right] = \frac{1-\theta}{\theta}
$$

and since the X_i's are independent random variables

$$\text{and } Var\left(\bar{X}\right) = \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n} Var\left(X_i\right)$$
$$= \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n} \frac{1-\theta}{\theta^2} = \left(\frac{1}{n}\right)^2 \left[n\left(\frac{1-\theta}{\theta^2}\right)\right] = \frac{1}{n}\left(\frac{1-\theta}{\theta^2}\right) \to 0 \quad \text{as } n \to \infty$$

9.27
$$E\left(\bar{X}\right) = \frac{1}{n}\sum_{i=1}^{n} E\left(X_i\right) = \frac{1}{n}\sum_{i=1}^{n}\theta = \frac{1}{n}\left(n\theta\right) = \theta$$

and since the X_i's are independent random variables

$$\text{and } Var\left(\bar{X}\right) = \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n} Var\left(X_i\right)$$
$$= \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n}\theta^2 = \left(\frac{1}{n}\right)^2 \left(n\theta^2\right) = \frac{\theta^2}{n} \to 0 \text{ as } n \to \infty$$

9.28 (a) $E\left(X_i^2\right) = Var\left(X_i\right) + \left[E\left(X_i^2\right)\right] = \sigma^2 + \mu^2.$

(b)
$$E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E\left(X_i\right) = \frac{1}{n}\sum_{i=1}^{n}\mu = \frac{1}{n}\left(n\mu\right) = \mu$$

Since the X_i's are independent random variables

$$Var(\bar{X}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n} Var\left(X_i\right) = \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n}\sigma^2$$
$$= \left(\frac{1}{n}\right)^2 \left(n\sigma^2\right) = \frac{\sigma^2}{n}$$
$$E\left[(\bar{X})^2\right] = \left[E\left(\bar{X}\right)\right]^2 + Var\left(\bar{X}\right) = \mu^2 + \frac{\sigma^2}{n}$$

(c)

$$E\left(S^2\right) = \frac{1}{n-1}\left\{\sum_{i=1}^{n} E\left(X_i^2\right) - nE\left[\left(\bar{X}\right)^2\right]\right\}$$
$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}\left(\mu^2+\sigma^2\right) - n\left(\mu^2 + \frac{\sigma^2}{n}\right)\right]$$
$$= \frac{1}{n-1}\left[n\left(\mu^2+\sigma^2\right) - n\mu^2 - \sigma^2\right]$$
$$= \frac{1}{n-1}\left[\left(n-1\right)\sigma^2\right]$$
$$= \sigma^2$$

9.29 (a) Since $X \sim G(-1.4, 1.5)$ and $Y \sim N(-2.1, 4)$ independently, then

$$E(X+Y) = E(X) + E(Y) = -1.4 + (-2.1) = -3.5$$

and

$$Var(X+Y) = Var(X) + Var(Y) = (1.5)^2 + 4 = 2.25 + 4 = 6.25 = (2.5)^2$$

so $X + Y \sim G(-3.5, 2.5)$ and

$$\begin{aligned} P(X+Y > -6) &= P\left(Z > \frac{-6-(-3.5)}{2.5}\right) = P\left(Z > \frac{-2.5}{2.5}\right) \text{ where } Z \sim N(0,1) \\ &\approx P(Z > -1) = P(Z \leq 1) = 0.84134 \end{aligned}$$

(b) Since $X \sim G(-1.4, 1.5)$ and $Y \sim N(-2.1, 4)$ independently, $-2X + Y \sim N(0.7, 13)$ and

$$P(-2X + Y < 3) = P\left(Z < \frac{3-0.7}{\sqrt{13}}\right) \approx P(Z < 0.64) = 0.73891$$

(c) Since $X \sim G(-1.4, 1.5)$ and $Y \sim N(-2.1, 4)$ independently, $Y - X \sim N(-0.7, 6.25)$ and

$$\begin{aligned} P(Y < X) = P(Y - X < 0) &= P\left(Z < \frac{0-(-0.7)}{2.5}\right) \text{ where } Z \sim N(0,1) \\ &= P\left(Z < \frac{0.7}{2.5}\right) = P(Z < 0.28) \\ &= 0.61026 \end{aligned}$$

9.30 (a) 1.22

(b) 17.67%

9.31 (a) Let X = amount of wine in a bottle. Then $X \sim N(1.05, 0.0004)$.

$$\begin{aligned} &P(\text{bottle contains less than 1 liter}) \\ &= P(X < 1) = P\left(Z < \frac{1-1.05}{0.02}\right) \text{ where } Z \sim N(0,1) \\ &= P(Z < -2.5) = 1 - P(Z \leq 2.5) = 1 - 0.99379 = 0.00621 \end{aligned}$$

A bottle is labelled as containing 1 liter. What is the probability the bottle contains less than 1 liter?

(b) Let V = volume of a cask. Then $V \sim N(22, 0.16)$. Let X_i = amount of wine in the ith bottle, $i = 1, 2, \ldots, 20$. Then $X_i \sim N(1.05, 0.0004)$, $i = 1, 2, \ldots, 20$ independently. Therefore $T = \sum\limits_{i=1}^{20} X_i \sim N(20(1.05), 20(0.0004))$ or $T \sim N(21, 0.008)$.

$$P(\text{contents of 20 bottle fit inside}) = P(V \geq T) = P(V - T \geq 0)$$

Since $V \sim N(22, 0.16)$ independently of $T \sim N(21, 0.008)$, $V - T \sim N(22 - 21, 0.16 + 0.008)$ or $V - T \sim N(1, 0.168)$.
Therefore

$$\begin{aligned} P(V - T \geq 0) &= P\left(Z \geq \frac{0-1}{\sqrt{0.168}}\right) \text{ where } Z \sim N(0,1) \\ &\approx P(Z \geq -2.44) = P(Z \leq 2.44) = 0.99266 \end{aligned}$$

9.32 0.4134

9.33 (a) 0.0062 (b) 0.0771

9.34 (a) Since $\bar{X}$ is a linear combination of independent Normal random variables it has a Normal distribution.

(b)

$$E(\bar{X}) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n}\mu = \frac{1}{n}(n\mu) = \mu$$

and since the X_i's are independent random variables

$$Var(\bar{X}) = \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n} Var(X_i) = \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n}\sigma^2 = \left(\frac{1}{n}\right)^2 (n\sigma^2) = \frac{\sigma^2}{n}$$

(c)

$$\begin{aligned} P\left(\left|\bar{X} - \mu\right| \leq 1.96\sigma/\sqrt{n}\right) &= P(|Z| \leq 1.96) \quad \text{where } Z \sim N(0,1) \\ &= 2P(Z \leq 1.96) - 1 = 2(0.975) - 1 = 0.95 \end{aligned}$$

(d) We want $P\left(\left|\bar{X} - \mu\right| \leq 1.0\right) \geq 0.95$ where $\bar{X} \sim G(\mu, 12/\sqrt{n})$ or

$$\begin{aligned} P\left(\left|\bar{X} - \mu\right| \leq 1.0\right) &= P\left(\frac{\left|\bar{X} - \mu\right|}{12/\sqrt{n}} \leq \frac{1.0}{12/\sqrt{n}}\right) \\ &= P\left(|Z| \leq \frac{\sqrt{n}}{12}\right) \geq 0.95 \text{ where } Z \sim N(0,1) \end{aligned}$$

Since $P(|Z| \leq 1.96) = 0.95$ we want $\sqrt{n}/12 \geq 1.96$ or $n \geq (1.96)^2(144) = 553.2$. Therefore n should be at least 554.

9.35 Let $T = X_1 + X_2 + \cdots + E(X_5) =$ number of adjacent pairs of unlike beads in a necklace. Since $E(X_1) = E(X_2) = \cdots = E(X_5)$ and

$$\begin{aligned} E(X_1) &= P(X_1 = 1) \\ &= P(\text{Bead 1 is Pink and Bead 2 is Blue}) + P(\text{Bead 1 is Blue and Bead 2 is Pink}) \\ &= \left(\frac{2}{3}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{3}\right)\left(\frac{2}{3}\right) = \frac{4}{9} \end{aligned}$$

therefore

$$E(T) = 5\left(\frac{4}{9}\right) = \frac{20}{9}$$

Now $Var(X_1) = Var(X_2) = \cdots = Var(X_5)$ and

$$Var(X_1) = P(X_1 = 1)[1 - P(X_1 = 1)] = \left(\frac{4}{9}\right)\left(\frac{5}{9}\right) = \frac{20}{81}$$

To find $Cov(X_1, X_2)$ we note that

$$\begin{aligned} E(X_1 X_2) &= P(X_1 = 1, X_2 = 1) \\ &= P(\text{Bead 1 is Pink, Bead 2 is Blue, Bead 3 is Pink}) \\ &+ P(\text{Bead 1 is Blue, Bead 2 is Pink, Bead 3 is Blue}) \\ &= \left(\frac{2}{3}\right)\left(\frac{1}{3}\right)\left(\frac{2}{3}\right) + \left(\frac{1}{3}\right)\left(\frac{2}{3}\right)\left(\frac{1}{3}\right) = \frac{6}{27} = \frac{18}{81} \end{aligned}$$

and therefore

$$Cov(X_1, X_2) = E(X_1 X_2) - E(X_1)E(X_2) = \frac{18}{81} - \left(\frac{4}{9}\right)\left(\frac{4}{9}\right) = \frac{2}{81}$$

Now $Cov(X_1, X_2) = Cov(X_2, X_3) = Cov(X_3, X_4) = Cov(X_4, X_5) = Cov(X_5, X_1) = \frac{2}{81}$ and all other covariances are zero. Therefore

$$Var(T) = 5\left(\frac{20}{81}\right) + 2(5)\left(\frac{2}{81}\right) = \frac{100 + 20}{81} = \frac{40}{27}$$

9.36 $p^3(4 + p)$; $4p^3(1 - p^3) + p^4(1 - p^4) + 8p^5(1 - p^2)$

9.37 Let

$$X_i = \begin{cases} 1 & \text{student answers question } i \text{ correctly} \\ 0 & \text{otherwise} \end{cases} \qquad i = 1, 2, \ldots, 100$$

(a) If the student guesses randomly then

$$\begin{aligned} P(X_i = 1) &= P(\text{student answers question } i \text{ correctly}) \\ &= P(\text{student knows the answer}) \\ &+ P(\text{student does not know the answer but guesses correctly}) \\ &= p_i + (1 - p_i)\left(\frac{1}{5}\right) = \frac{4}{5}p_i + \frac{1}{5} = \frac{1}{5}(1 + 4p_i) = q_i, \ \ i = 1, 2, \ldots, 100 \end{aligned}$$

Then $X_i \sim Binomial(1, q_i)$ with

$$\begin{aligned} E(X_i) &= q_i = \frac{1}{5}(1 + 4p_i) \\ Var(X_i) &= q_i(1 - q_i) = \frac{1}{5}(1 + 4p_i) \cdot \frac{4}{5}(1 - p_i) = \frac{4}{25}(1 + 4p_i)(1 - p_i) \\ &= \frac{4}{25}[(1 - p_i) + 4p_i(1 - p_i)] \ \ \text{for } i = 1, 2, \ldots, 100 \end{aligned}$$

Let $S = \sum\limits_{i=1}^{100} X_i$. Then

$$E(S) = E\left(\sum_{i=1}^{100} X_i\right) = \sum_{i=1}^{100} E(X_i) = \sum_{i=1}^{100} \frac{1}{5}(1 + 4p_i) = \frac{100}{5} + \frac{4}{5}\sum_{i=1}^{100} p_i$$

and

$$Var(S) = Var\left(\sum_{i=1}^{100} X_i\right) = \sum_{i=1}^{100} Var(X_i) = \frac{4}{25}\sum_{i=1}^{100}[(1 - p_i) + 4p_i(1 - p_i)].$$

The student's total mark is given by

$$T = \sum_{i=1}^{100} X_i - \left(\frac{1}{4}\right)\left(100 - \sum_{i=1}^{100} X_i\right) = \frac{5}{4}\sum_{i=1}^{100} X_i - 25 = \frac{5}{4}S - 25.$$

Therefore

$$E(T) = E\left(\frac{5}{4}S - 25\right) = \frac{5}{4}E(S) - 25 = \frac{5}{4}\left[\frac{100}{5} + \frac{4}{5}\sum_{i=1}^{100} p_i\right] - 25 = \sum_{i=1}^{100} p_i$$

and

$$\begin{aligned} Var(T) &= Var\left(\frac{5}{4}S - 25\right) = \frac{25}{16}Var(S) = \frac{25}{16} \cdot \frac{4}{25}\sum_{i=1}^{100}[(1 - p_i) + 4p_i(1 - p_i)] \\ &= \sum_{i=1}^{100} p_i(1 - p_i) + \frac{1}{4}\left(100 - \sum_{i=1}^{100} p_i\right) \end{aligned}$$

as required.

(b) If the student does not guess then $X_i \sim Binomial\,(1, p_i)$ with $E\,(X_i) = p_i$ and $Var\,(X_i) = p_i\,(1 - p_i)$. The student's total mark is $S = \sum\limits_{i=1}^{100} X_i$ with

$$E\,(S) = E\left(\sum_{i=1}^{100} X_i\right) = \sum_{i=1}^{100} E\,(X_i) = \sum_{i=1}^{100} p_i$$

and

$$Var\,(S) = Var\left(\sum_{i=1}^{100} X_i\right) = \sum_{i=1}^{100} Var\,(X_i) = \sum_{i=1}^{100} p_i\,(1 - p_i)$$

as required.

(c) (i) If $p_i = 0.9$ then $Var(T) = 11.5$ and $Var(S) = 9$.

(ii) If $p_i = 0.5$ then $Var(T) = 37.5$ and $Var(S) = 25$.

9.38 (a) $X =$ the number of keys assigned to a given list has a $Binomial\left(n, \frac{1}{M}\right)$ distribution. The expected number of keys assigned to a given list is

$$E\,(X) = n\left(\frac{1}{M}\right) = \frac{n}{M}$$

(b) Consider slot i in the hash table and let $S_i = 1$ if the slot is empty and $S_i = 0$ otherwise, $i = 1, 2, \ldots, M$. Then

$$P\,(S_i = 1) = \left(1 - \frac{1}{M}\right)^n$$

and

$$E\,(S_i) = (1)\left(1 - \frac{1}{M}\right)^n + (0)\left[1 - \left(1 - \frac{1}{M}\right)^n\right] = \left(1 - \frac{1}{M}\right)^n$$

Now $S = S_1 + S_2 + \cdots + S_M =$ the number of empty slots and the expected number of empty slots is

$$\begin{aligned} E\,(S) &= E\,(S_1 + S_2 + \cdots + S_M) \\ &= E\,(S_1) + E\,(S_2) + \cdots + E\,(S_m) \\ &= M\left(1 - \frac{1}{M}\right)^n \end{aligned}$$

(c) We first note that

$$\text{number of collisions} = n - \text{ number of occupied slots}$$

and

$$\text{number of occupied slots} = M - \text{ number of empty slots}$$

so

$$\begin{aligned}\text{number of collisions} &= n - \ (M - \text{ number of empty slots}) \\ &= n - M + \text{ number of empty slots}\end{aligned}$$

Using the result from (b) we have

$$\begin{aligned}E\,(\text{number of collisions}) &= n - M + E\,(\text{number of empty slots}) \\ &= n - M + M\left(1 - \frac{1}{M}\right)^n\end{aligned}$$

(d) Let $X_i =$ number of keys in the table when a total of i slots are assigned for the first time, $i = 1, 2, \ldots, M$. Then $T = \sum\limits_{i=1}^{M} X_i =$ number of keys in the table when every slot has at least one key for the first time. Now $X_1 = 1$ with probability one so $E\,(X_1) = 1$. $X_2 =$ the number of keys assigned when a second slot is assigned for the first time in a sequence of Bernoulli trials where a success is "a second slot is chosen for the first time" and $P\,(\text{Success}) = \frac{M-1}{M}$. Recall if $X \sim Geometric\,(p)$ then $E\,(X) = (1-p)\,/p$. Therefore

$$E\,(X_2) = 1 + \frac{1 - \frac{M-1}{M}}{\frac{M-1}{M}} = \frac{\frac{M-1}{M} + 1 - \frac{M-1}{M}}{\frac{M-1}{M}} = \frac{M}{M-1}$$

Similarly, $X_3 =$ the number of keys assigned when a third slot is assigned for the first time in a sequence of Bernoulli trials where a success is "a third slot is chosen for the first time" and $P\,(\text{Success}) = \frac{M-2}{M}$. Therefore $E\,(X_3) = \frac{M}{M-2}$. Continuing in this manner we find

$$E\,(T) = \sum_{i=1}^{M} E\,(X_i) = \sum_{i=1}^{M} \frac{M}{M-i+1} = M \sum_{j=1}^{M} \frac{1}{j}$$

which is the sum of the first M terms in a harmonic series which does not have a closed form. Using the approximation

$$\sum_{j=1}^{M} \frac{1}{j} \approx \ln M$$

we have

$$E\,(T) = M \sum_{j=1}^{M} \frac{1}{j} \approx M \ln M$$

9.39 Suppose P is $N \times N$ and let $\mathbf{1}$ be a column vector of ones of length N. Consider the probability vector corresponding to the discrete Uniform distribution $\pi = \frac{1}{N}\mathbf{1}$. Then

$$\pi^T P = \frac{1}{N}\mathbf{1}^T P = \frac{1}{N}\left(\sum_{i=1}^{N} P_{i1}, \sum_{i=1}^{N} P_{i2}, \ldots, \sum_{i=1}^{N} P_{iN}\right) = \frac{1}{N}\mathbf{1}^T = \pi^T$$

since P is doubly stochastic. Therefore π is a stationary distribution of the Markov chain.

9.40 The transition matrix is

$$P=\begin{bmatrix} 0 & 1 & 0 \\ \frac{2}{3} & 0 & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & 0 \end{bmatrix}$$

from which, solving $\pi^T P = \pi^T$ and rescaling so that the sum of the probabilities is one, we obtain $\pi^T = (0.4, 0.45, 0.15)$, the long run fraction of time spent in cities A,B,C respectively.

9.41 By arguments similar to those in Section 9.3, the limiting matrix has rows all identically π^T where the vector π^T are the stationary probabilities satisfying $\pi^T P = \pi^T$ and

$$P=\begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

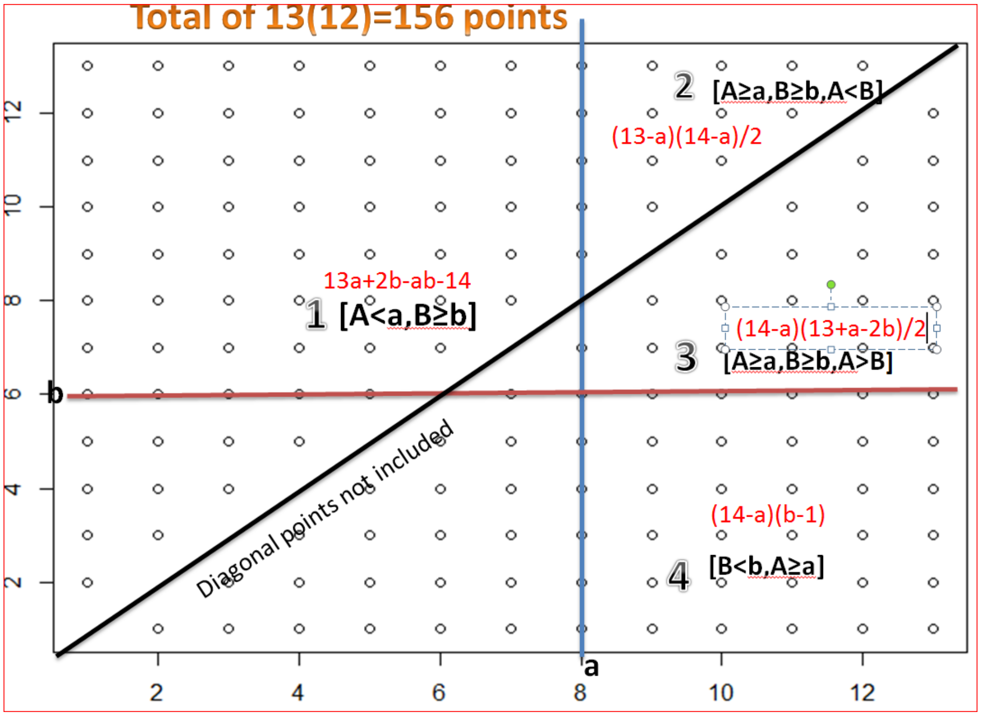
The solution is $\pi^T = (0.1, 0.6, 0.3)$ and the limit is

$$\begin{bmatrix} 0.1 & 0.6 & 0.3 \\ 0.1 & 0.6 & 0.3 \\ 0.1 & 0.6 & 0.3 \end{bmatrix}$$

9.42 If today is raining, the probability of Rain, Nice, Snow three days from now is obtainable from the first row of the matrix P^3, that is, $(0.406\ 0.203\ 0.391)$. The probabilities of the three states in five days, given (1) today is raining (2) today is nice (3) today is snowing are the three rows of the matrix P^5. In this case call rows are identical to three decimals; they are all equal the equilibrium distribution $\pi^T = (0.400, 0.200, 0.400)$.

9.43 We begin this solution by describing a suitable sample space. Players A and B are dealt one card each without replacement from a deck of 13 cards numbered $1, 2, \ldots, 13$. There are $(13)(12) = 156$ possible outcomes. Let A represent the number on player A's card and let B represent the number on player B's card. The possible outcomes are the points shown in the following diagram. The diagonal points are not included in the sample space.

Recall that before the game begins, Player A choses an integer number a between 1 and 13 and similarly Player B choses an integer number b between 1 and 13. Suppose $b \le a$, that is, player B has chosen a number b less than or equal to the number chosen by player A. The region $[A < a, B \ge b]$, which is labelled 1 in the diagram, corresponds to the event that player A has drawn a card less than their chosen number a and player B has chosen a card greater than or equal to their chosen number b. There are a total of $(14-b)(a-1)$ points in this region less the points $\{(b,b), \ldots, (a-1, a-1)\}$ on the diagonal, so a total of $(14-b)(a-1) - (a-b) = 13a + 2b - ab - 14$ points.



The number of points in the rectangular region $[A \geq a, B < b]$, which is labelled 4, is $(14-a)(b-1)$.

Consider the triangular region $[A \geq a, B \geq b, A < B]$ which is labelled 2. The number of points in the region excluding the points on the diagonal is

$$1 + 2 + \cdots + (13-a) = \sum_{i=1}^{13-a} i = \frac{(13-a)\,(14-a)}{2}$$

To obtain the number of points in the triangular region labelled 3, we note that the number of points in the rectangular region comprised of regions 2 and 3 is $(14-a)(13-b)$. Therefore the number of points in region 3 is $(14-a)(13-b)$ minus the number of points in region 2 which is

$$(14-a)(13-b) - \frac{(13-a)\,(14-a)}{2} = \frac{(14-a)(13+a-2b)}{2}$$

Note that the points in the region $[A < a, B < b]$ correspond to a draw and no player wins any money.

In summary

Region	Number of points	Player A winnings	Player B winnings
2. $A \geq a, B \geq b, A < B$	$\frac{(13-a)(14-a)}{2}$	-6	$+6$
3. $A \geq a, B \geq b, A > B$	$\frac{(14-a)(13+a-2b)}{2}$	$+6$	-6
4. $A \geq a, B < b$	$(14-a)(b-1)$	$+1$	-1
1. $A < a, B \geq b$	$13a + 2b - ab - 14$	-1	$+1$

(a) If $b \leq a$ then the probability player B wins given that both players raise is determined by the ratio of the number of points in region 2 to the total number of points in regions in 2 and 3, that is,

$$P(\text{player } B \text{ wins}|R) = \frac{\frac{(13-a)(14-a)}{2}}{\frac{(13-a)(14-a)}{2} + \frac{(14-a)(13+a-2b)}{2}} = \frac{13-a}{2(13-b)} \text{ if } b \leq a$$

Note that $P(\text{player } B \text{ wins}|R) \leq \frac{1}{2}$ if $b \leq a$.
If $b \leq a$ then the probability player A wins given that both players raise is

$$P(\text{player } A \text{ wins}|R) = 1 - \frac{13-a}{2\,(13-b)} = \frac{a-2b+13}{2\,(13-b)} \text{ if } b \leq a$$

To determine the probabilities if $b > a$, we reverse the roles of a and b in the above discussion to obtain

$$P(\text{player } B \text{ wins}|R) = \frac{b-2a+13}{2\,(13-a)} \text{ if } b > a$$
$$P(\text{player } A \text{ wins}|R) = \frac{13-b}{2\,(13-a)} \text{ if } b > a$$

Therefore

$$P(\text{player } A \text{ wins}|R) - P(\text{player } B \text{ wins}|R) = \begin{cases} \frac{a-b}{13-b} & \text{if } b \leq a \\ \frac{a-b}{13-a} & \text{if } b > a \end{cases}$$

and the expected value of player A's winnings, given that both players raise is

$$6[P(\text{player } A \text{ wins}|R) - P(\text{player } B \text{ wins}|R)] = \frac{6\,(a-b)}{13-\min(a,b)}$$

(b) To determine the answer for $b = 1$ we first solve the more general case where b is arbitrary and the amount by which each player raises is $r-1$ (in the given question $r-1=5$) so that the total stake is r.

By counting the total number of points in the regions labeled 2 and 3, we obtain

$$\begin{aligned} P(R) &= \frac{1}{156}\left[\frac{(13-a)(14-a)}{2} + \frac{(14-a)(13+a-2b)}{2}\right] \\ &= \frac{(14-a)\,(13-b)}{156} \end{aligned}$$

Assuming any raise is by an amount $r-1$ and the case $a \geq b$ we have

$$\begin{aligned} &E\,(\text{player } A \text{ winnings}) \\ &= rP(R)[P(\text{player } A \text{ wins}|R) - P(\text{player } B \text{ wins}|R)] + P(A \geq a, B < b) - P(A < a, B \geq b) \\ &= r\left(\frac{a-b}{13-b}\right)\frac{(14-a)\,(13-b)}{156} + \frac{(14-a)(b-1)}{156} - \frac{(13a+2b-ab-14)}{156} \\ &= \frac{(a-b)\,(14r-ar-12)}{156} \end{aligned}$$

For $r = 6$ we obtain

$$E\,(A \text{ winnings}) = \frac{(a-b)\,(12-a)}{26} \tag{1}$$

To determine the maximum of (1) over a we solve

$$\frac{d}{da}\,(a-b)\,(12-a) = \frac{d}{da}\left(12a - a^2 - 12b + ab\right) = 12 - 2a + b = 0$$

and obtain

$$a = \frac{b}{2} + 6$$

Note that if $\frac{b}{2} \geq 6$ or $b \geq 12$, then the optimal choice of a is b. Otherwise, the optimal choice of a is the integer closest to $\frac{b}{2} + 6$, that is, the optimal choice is $\frac{1}{2}b + 6$ if b is even and either of the two integers closest to $\frac{1}{2}b + 6$ if b is odd. The maximum expected winnings for player A obtains by substituting for a in (1) which gives

$$\max_a E(\text{player } A \text{ winnings}) = \begin{cases} \frac{1}{104}\,(b-12)^2 & \text{if } b \text{ is even} \\ \frac{1}{104}\,(11-b)\,(13-b) & \text{if } b \text{ is odd} \end{cases}$$

Finally, if $b = 1$ so that B always raises then

$$E\,(\text{player } A \text{ winnings}) = \frac{(a-1)\,(12-a)}{26}$$

and

$$\max_a E(\text{player } A \text{ winnings}) = \frac{1}{104}\,(11-1)\,(13-1) = \frac{15}{13}$$

which occurs for either $a = 6$ or $a = 7$.

(c) Based on the solution given in (b) we note that for $b = 11$, the optimal choice of a is the closest to $\frac{11}{2} + 6 = 11.5$ which $a = 11$ or 12 and

$$\max_a E(\text{player } A \text{ winnings}) = \frac{1}{104}\,(11-11)\,(13-11) = 0$$

Thus, in this case, player A has no strategy which provides a positive expected profit.

(d) Note that in order to maximize $E\,(\text{player } A \text{ winnings}) = \frac{1}{26}\,(a-b)\,(12-a)$, player A wishes to choose a to be an integer close to $\frac{1}{2}b + 6$. Provided that $b \leq 10$ this always returns a positive value for $E\,(\text{player } A \text{ winnings}) = \frac{1}{26}\,(a-b)\,(12-a)$ when $a > b$. However if $b = 11$, the expected return to player A is always 0 when $a \geq b$ and < 0 otherwise. So the (minimax) strategy is $a = 11$ and $b = 11$. This is the minimax strategy since we maximized over a and then minimixed that result over b. For the minimax solution the expected winnings to both players is 0.

9.44 (a) The permutation X_{j+1} after $j+1$ requests depends only on the permutation X_j before and the record requested at time $j+1$. Thus the new state depends only on the old state X_j (without knowing the previous states) and the record currently requested.

(b) For example the long-run probability of the state (i,j,k) is $q_i p_j$, where $q_i = \frac{p_i}{1-p_i}$.

(c) The probability that record j is in position k is

p_j	for $k = 1$
$p_j(Q - q_j)$	for $k = 2$
$1 - p_j(1 + Q - q_j)$	for $k = 3$

where $Q = \sum_{i=1}^{3} q_i$. The expected cost of accessing a record in the long run is

$$\sum_{j=1}^{3} \{p_j^2 + 2p_j^2(Q - q_j) + 3p_j[1 - p_j(1 + Q - q_j)]\}$$

Substituting $p_1 = 0.1$, $p_2 = 0.3$, $p_3 = 0.6$ gives $q_1 = \frac{1}{9}$, $q_2 = \frac{3}{7}$, $q_3 = \frac{6}{4}$ and $Q = \frac{1}{9} + \frac{3}{7} + \frac{6}{4} = 2.0397$ and the expected cost is 1.7214.

(d) If they are in random order, the expected cost is $1(\frac{1}{3}) + 2(\frac{1}{3}) + 3(\frac{1}{3}) = 2$. If they are ordered in terms of decreasing p_j, the expected cost is $p_3^2 + 2p_2^2 + 3p_1^2 = 0.57$.

9.46 Let $J =$ index of maximum. $P(J = j) = \frac{1}{N}$, for $j = 1, 2, \ldots, N$. Let $A =$ "your strategy chooses the maximum".

A occurs only if $J > k$ and if $\max\{X_i; k < i < J\} < \max\{X_i; 1 \leq i \leq k\}$. Given $J = j > k$, the probability of this is the probability that $\max\{X_i; 1 \le i < j\}$ occurs among the first k values, which occurs with probability is $\frac{k}{j-1}$. Therefore,

$$P(A) = \textstyle\sum_j P(A|J=j)P(J=j) = \sum_{j=k+1}^{N} P(A|J=j)\dfrac{1}{N}$$
$$= \sum_{j=k+1}^{N} \frac{k}{j-1}\frac{1}{N} = \frac{k}{N}\left(\frac{1}{k} + \frac{1}{k+1} + \ldots + \frac{1}{N-1}\right) \approx \frac{k}{N}\ln\left(\frac{N}{k}\right)$$

Note that the value of x maximizing $x\ln(1/x)$ is $x = e^{-1} \approx 0.37$ so roughly, the best k is Ne^{-1}. The probability that you select the maximum is approximately $e^{-1} \approx 0.37$.

9.47 (a) By definition

$$f_1(x|y) = \frac{d}{dx}P(X \le x|Y=y)$$

and

$$f_2(y|x) = P(Y=y|X=x) = \frac{f_1(x|y)\,P(Y=y)}{f_1(x)}$$

Note that

$$\int_{-\infty}^{\infty} f_1(x|y)\,dx = \int_{-\infty}^{\infty} \frac{d}{dx}P(X \le x|Y=y)\,dx$$
$$= \lim_{x\to\infty} P(X \le x|Y=y) - \lim_{x\to\infty} P(X \le x|Y=y) = 1 - 0 = 1$$

Since

$$f_2(y|x) = P(Y=y|X=x) = \frac{f_1(x|y)\,P(Y=y)}{f_1(x)}$$

we have

$$f_1(x|y)\,P(Y=y) = f_2(y|x)\,f_1(x)$$

and

$$\int_{-\infty}^{\infty} f_1(x|y)\,P(Y=y)\,dx = \int_{-\infty}^{\infty} f_2(y|x)\,f_1(x)\,dx$$

but

$$\int_{-\infty}^{\infty} f_1(x|y)\,P(Y=y)\,dx = P(Y=y)\int_{-\infty}^{\infty} f_1(x|y)\,dx = P(Y=y)$$

and therefore

$$\int_{-\infty}^{\infty} f_2(y|x)\,f_1(x)\,dx = P(Y=y) = f_2(y)$$

as required.

(b) Since $X \sim U(0,1)$ and $P(Y=y|X=x) = f_2(y|x) = \binom{n}{y} x^y (1-x)^{n-y}$ therefore

$$f_2(y) = P(Y=y) = \int_{-\infty}^{\infty} f_2(y|x) f_1(x)\, dx = \int_0^1 \binom{n}{y} x^y (1-x)^{n-y} (1)\, dx$$
$$= \binom{n}{y} \int_0^1 x^y (1-x)^{n-y}\, dx = \frac{n!}{y!\,(n-y)!} \frac{y!\,(n-y)!}{(n+1)!} = \frac{1}{n+1}$$

(c) To find the conditional probability density function of X given $Y=0$ we note that

$$P(X \le x|Y=0) = P(X \le x|\,|X| \le 1) = \frac{P(X \le x, |X| \le 1)}{P(|X| \le 1)}$$

and

$$P(X \le x, |X| \le 1) = \begin{cases} 0 & \text{if } x < -1 \\ P(-1 < X \le x) & \text{if } |x| \le 1 \\ P(|X| \le 1) & \text{for } x > 1 \end{cases}$$
$$= \begin{cases} 0 & \text{if } x < -1 \\ F_1(x) - F_1(-1) & \text{if } |x| \le 1 \\ P(|X| \le 1) & \text{for } x > 1 \end{cases}$$

where $F_1(x) = P(X \le x)$. Note also that

$$\frac{d}{dx} P(X \le x, |X| \le 1) = \begin{cases} 0 & \text{if } x < -1 \\ f_1(x) & \text{if } |x| \le 1 \\ 0 & \text{for } x > 1 \end{cases}$$

and therefore the conditional probability density function of X given $Y=0$ is

$$f_1(x|0) = \frac{d}{dx} P(X \le x|Y=0) = \begin{cases} 0 & \text{if } x < -1 \\ \frac{f_1(x)}{P(|X| \le 1)} & \text{if } |x| \le 1 \\ 0 & \text{for } x > 1 \end{cases}$$

10. SOLUTIONS TO CHAPTER 10 PROBLEMS

10.1 (a) Since $E(X_i) = 1/2$ and $Var(X_i) = 1/24$, $E(S) = 100(1/2) = 50$ and $Var(S) = 100(1/24) = 25/6$. Since S is the sum of independent and identically distributed random variables then by the Central Limit Theorem $S = \sum_{i=1}^{100} X_i$ will have approximately a $N(50, 25/6)$ distribution.

Therefore

$$
\begin{aligned}
P(49.0 \leq S \leq 50.5) &\approx P\left(\frac{49-50}{\sqrt{25/6}} \leq Z \leq \frac{50.5-50}{\sqrt{25/6}}\right) \quad \text{where } Z \sim N(0,1) \\
&= P(-0.4899 \leq Z \leq 0.2449) \\
&\approx P(Z \leq 0.24) + P(Z \leq 0.49) - 1 \\
&= 0.59484 + 0.68793 - 1 \\
&= 0.28277
\end{aligned}
$$

(b) If $X_i \sim U(0,1)$ then $E(X_i) = 1/2$ and $Var(X_i) = 1/12$. Then $E(S) = 100(1/2) = 50$ and $Var(S) = 100(1/12) = 25/3$. Since S is the sum of independent and identically distributed random variables then by the Central Limit Theorem $S = \sum_{i=1}^{100} X_i$ will have approximately a $N(50, 25/3)$ distribution.

Therefore

$$
\begin{aligned}
P(49.0 \leq S \leq 50.5) &\approx P\left(\frac{49.0-50}{\sqrt{25/3}} \leq Z \leq \frac{50.5-50}{\sqrt{25/3}}\right) \quad \text{where } Z \sim N(0,1) \\
&= P(-0.3464 \leq Z \leq 0.1732) \\
&\approx P(Z \leq 0.17) + P(Z \leq 0.35) - 1 \\
&= 0.5675 + 0.63683 - 1 \\
&= 0.20433
\end{aligned}
$$

10.2 Recall that if the student does not guess

$$E(S) = \sum_{i=1}^{100} p_i \text{ and } Var(S) = \sum_{i=1}^{100} p_i (1 - p_i)$$

where

$$S = \sum_{i=1}^{100} X_i$$

is their total mark.

If the student guesses then their total mark T is

$$T = \sum_{i=1}^{100} X_i - \left(\frac{1}{4}\right)\left(100 - \sum_{i=1}^{100} X_i\right) = \frac{5}{4}\sum_{i=1}^{100} X_i - 25 = \frac{5}{4}S - 25$$

where

$$E(S) = \frac{100}{5} + \frac{4}{5}\sum_{i=1}^{100} p_i \text{ and } Var(S) = \frac{4}{25}\sum_{i=1}^{100}\left[(1 - p_i) + 4p_i(1 - p_i)\right]$$

(a) If $p_i = 0.45$ and student does not guess then $E(S) = 100(0.45) = 45$ and $Var(S) = 100(0.45)(0.55) = 24.75$. Since S is the sum of independent and identically distributed random variables then by the Central Limit Theorem $S = \sum_{i=1}^{100} X_i$ will have approximately a $N(45, 24.75)$ distribution. Therefore

$$\begin{aligned} P(S \geq 50) &\approx P\left(Z \geq \frac{49.5 - 45}{\sqrt{24.75}}\right) \quad \text{where } Z \sim N(0,1) \\ &= 1 - P(Z \leq 0.9045) \approx 1 - P(Z \leq 0.90) = 1 - 0.81954 \\ &= 0.18045 \end{aligned}$$

If $p_i = 0.45$ and the student guesses then $E(S) = \frac{100}{5} + \frac{4}{5}(100)(0.45) = 56$ and $Var(S) = \frac{4}{25}[100(0.55) + 400(0.45)(0.55)] = 24.64$.

$$\begin{aligned} P(T \geq 50) &= P\left(\frac{5}{4}S - 25 \geq 50\right) = P(S \geq 60) \approx P\left(Z \geq \frac{59.5 - 56}{\sqrt{24.64}}\right) \quad \text{where } Z \sim N(0,1) \\ &= 1 - P(Z \leq 0.7051) \approx 1 - P(Z \leq 0.71) = 1 - 0.76115 \\ &= 0.23885 \end{aligned}$$

(b) If $p_i = 0.55$ and the student does not guess then $E(S) = 100(0.55) = 55$ and $Var(S) = 100(0.55)(0.45) = 24.75$. Since S is the sum of independent and identically distributed random variables then by the Central Limit Theorem $S = \sum_{i=1}^{100} X_i$ will have approximately a

$N(55, 24.75)$ distribution. Therefore

$$\begin{aligned} P(S \geq 50) &\approx P\left(Z \geq \frac{49.5 - 55}{\sqrt{24.75}}\right) \quad \text{where } Z \sim N(0,1) \\ &= 1 - P(Z \leq -1.1055) \approx P(Z \leq 1.11) \\ &= 0.86650 \end{aligned}$$

If $p_i = 0.55$ and the student guesses then $E(S) = \frac{100}{5} + \frac{4}{5}(100)(0.55) = 64$ and
$Var(S) = \frac{4}{25}[100(0.45) + 400(0.55)(0.45)] = 23.04$.

$$\begin{aligned} P(T \geq 50) &= P\left(\frac{5}{4}S - 25 \geq 50\right) = P(S \geq 60) \approx P\left(Z \geq \frac{59.5 - 64}{\sqrt{23.04}}\right) \quad \text{where } Z \sim N(0,1) \\ &= 1 - P(Z \leq -0.9375) \approx P(Z \leq 0.94) \\ &= 0.82639 \end{aligned}$$

If $p_i = 0.45$ then the best strategy for passing is to guess and if $p_i = 0.55$ then the best strategy for passing is to not guess.

10.3 We to find n such that

$$P\left(\left|\frac{X}{n} - 0.16\right| \leq 0.03\right) \geq 0.95$$

where $X \sim Binomial(n, 0.16)$. By the Normal approximation to the Binomial

$$\begin{aligned} P\left(\left|\frac{X}{n} - 0.16\right| \leq 0.03\right) &= P\left(\left|\frac{X - 0.16n}{\sqrt{n(0.16)(0.84)}}\right| \leq \frac{0.03n}{\sqrt{n(0.16)(0.84)}}\right) \\ &\approx P\left(|Z| \leq 0.08183\sqrt{n}\right) \quad \text{where } Z \sim N(0,1) \end{aligned}$$

Since $P(|Z| \leq 1.96) = 0.95$ then we want $0.08183\sqrt{n} \geq 1.96$ or
$n \geq (1.96/0.08183)^2 = (23.95)^2 = 573.6$. Therefore n should be at least 574.

10.4 (a) Expected number of tests $= (1)(0.98)^{20} + (21)\left[1 - (0.98)^{20}\right] = 7.6478$

Variance of number of tests $= (1)^2(0.98)^{20} + (21)^2\left[1 - (0.98)^{20}\right] - (7.6478)^2 = 88.7630$

(b) For 2000 people the expected number of tests is $(100)(7.6478) = 764.78$, and the variance of the number of tests is $(100)(88.7630) = 8876.30$, since people within pooled samples are independent and each pooled sample is independent of each other pooled sample.

(c) Let N = number of tests for 2000 people. Now $N = \sum_{i=1}^{100} N_i$ where N_i = number of tests required in the ith group of 20 people. Since N is the sum of 100 independent and identically distributed random variables then by the Central Limit Theorem N has approximately a

$N(764.78, 8876.30)$ distribution. Since the possible values for N are $n = 100, 120, 140, \ldots, 2100$ the continuity correction is $20/2 = 10$. Therefore

$$\begin{aligned} P(N > 800) &\approx P\left(Z \geq \frac{(800+10) - 764.78}{\sqrt{8876.30}}\right) \quad \text{where } Z \sim N(0,1) \\ &= 1 - P(Z \leq 0.48) = 1 - 0.68439 \\ &= 0.31561 \end{aligned}$$

10.5 Since $X \sim Binomial(60, 0.8)$, then

$$E(X) = 60(0.8) = 48 \text{ and } Var(X) = 60(0.8)(0.2) = 9.6$$

Since $Y \sim Binomial(62, 0.8)$, then

$$E(Y) = 62(0.8) = 49.6 \text{ and } Var(Y) = 62(0.8)(0.2) = 9.92$$

Now $E(X - Y) = 48 - 49.6 = -1.6$ and since X and Y are independent random variables

$$Var(X - Y) = Var(X) + Var(Y) = 9.6 + 9.92 = 19.52$$

By the Normal approximation to the Binomial, X has approximately a $N(48, 9.6)$ distribution and Y has approximately a $N(49.6, 9.92)$ distribution. Since X and Y are independent random variables then $X - Y$ has approximately a $N(-1.6, 19.52)$ distribution. Therefore

$$\begin{aligned} P(|X - Y| \geq 3) &= 1 - P(|X - Y| < 3) = 1 - P(-3 < X - Y < 3) \\ &\approx 1 - P\left(\frac{-2.5 - (-1.6)}{\sqrt{19.52}} \leq Z \leq \frac{2.5 - (-1.6)}{\sqrt{19.52}}\right) \quad \text{where } Z \sim N(0,1) \\ &= 1 - P(-0.20 \leq Z \leq 0.93) \\ &= 1 - [P(Z \leq 0.93) - 1 + P(Z \leq 0.20)] \\ &= 2 - 0.82381 - 0.57926 \\ &= 0.59693 \end{aligned}$$

10.6 (a) Let $X =$ number of unemployed people in a sample of 10000 persons. Then $X \sim Binomial(10000, 0.07)$. By the Normal approximation to the Binomial X has approximately a $N(700, 651)$ distribution. Therefore

$$\begin{aligned} P(675 \leq X \leq 725) &\approx P\left(\frac{675 - 700}{\sqrt{651}} \leq Z \leq \frac{725 - 700}{\sqrt{651}}\right) \quad \text{where } Z \sim N(0,1) \\ &= P(|Z| \leq 0.98) = 2P(Z \leq 0.98) - 1 \\ &= 2(0.83646) - 1 \\ &= 0.67292 \end{aligned}$$

Note than since $n = 10000$ is very large a continuity correction has not been used.

(b) We need to find n such that

$$P\left(0.069 \leq \frac{X}{n} \leq 0.071\right) = P\left(\left|\frac{X}{n} - 0.07\right| \leq 0.001\right) \geq 0.95$$

where $X \sim Binomial\,(n, 0.07)$. By the Normal approximation to the Binomial

$$\begin{aligned} P\left(\left|\frac{X}{n} - 0.07\right| \leq 0.001\right) &= P\left(\left|\frac{X - 0.07n}{\sqrt{n\,(0.07)\,(0.93)}}\right| \leq \frac{0.001n}{\sqrt{n\,(0.07)\,(0.93)}}\right) \\ &\approx P\left(|Z| \leq 0.003919\sqrt{n}\right) \quad \text{where } Z \sim N\,(0,1) \end{aligned}$$

Since $P\,(|Z| \leq 1.96) = 0.95$ then we want $0.003919\sqrt{n} \geq 1.96$ or
$n \geq (1.96/0.003919)^2 = (500.1276)^2 = 250127.6$. Therefore n should be at least $250,128$.

10.7 Let $X =$ number of requests in a one minute $= 60$ second interval. Then $X \sim Poisson\,(2 \times 60)$. Since $\mu = 120$ is large we can use the Normal approximation to the Poisson.

(a)

$$\begin{aligned} P\,(110 \leq X \leq 135) &= \sum_{x=110}^{135} \frac{(120)^x\, e^{-120}}{x!} \\ &= 0.7502 \quad \text{(calculated using R)} \end{aligned}$$

$$\begin{aligned} P\,(110 \leq X \leq 135) &\approx P\left(\frac{109.5 - 120}{\sqrt{120}} \leq Z \leq \frac{135.5 - 120}{\sqrt{120}}\right) \quad \text{where } Z \sim N\,(0,1) \\ &= P\,(-0.96 \leq Z \leq 1.41) \\ &= P\,(Z \leq 1.41) - P\,(Z \leq -0.96) \\ &= P\,(Z \leq 1.41) - [1 - P\,(Z \leq 0.96)] \\ &= P\,(Z \leq 1.41) + P\,(Z \leq 0.96) - 1 \\ &= 0.92073 + 0.83147 - 1 \\ &= 0.7522 \end{aligned}$$

(b)

$$\begin{aligned} P\,(X > 150) &= \sum_{x=151}^{\infty} \frac{(120)^x\, e^{-120}}{x!} = 1 - \sum_{x=0}^{150} \frac{(120)^x\, e^{-120}}{x!} \\ &= 0.003552 \quad \text{(calculated using R)} \end{aligned}$$

$$
\begin{aligned}
P(X > 150) &\approx P\left(Z \geq \frac{150.5 - 120}{\sqrt{120}}\right) \quad \text{where } Z \sim N(0,1) \\
&= P(Z \geq 2.78) \\
&= 1 - 0.99728 \\
&= 0.00272
\end{aligned}
$$

(c) Let X_i = waiting time between requests $(i-1)$ and i, $i = 1, 2, \ldots, 600$. Then X_i has an Exponential distribution with mean $1/2 = 0.5$ seconds and variance $(1/2)^2 = 0.25$ (seconds)^2. The waiting time until the 600'th request is $S = X_1 + X_2 + \cdots + X_{600}$. Since S is the sum of independent and identically distributed random variables then by the Central Limit Theorem S will have approximately a $N(600(0.5), 600(0.25)) = N(300, 150)$ distribution.

$$
\begin{aligned}
P(S < (4.5)(60)) &= P(S < 270) \\
&\approx P\left(Z < \frac{270 - 300}{\sqrt{150}}\right) \quad \text{where } Z \sim N(0,1) \\
&= P(Z < -2.45) = 1 - P(Z < 2.45) = 1 - 0.99286 \\
&= 0.00714
\end{aligned}
$$

Note that a continuity correction is not used since S is a continuous random variable.

10.8 (a) By the Central Limit Theorem we have

$$
\frac{X - np}{\sqrt{np(1-p)}} \sim N(0,1) \quad \text{approximately}
$$

which implies

$$
\frac{\frac{X}{n} - p}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0,1) \quad \text{approximately}
$$

Therefore

$$
\begin{aligned}
&P\left(\frac{X}{n} - 1.645\sqrt{\frac{p(1-p)}{n}} \leq p \leq \frac{X}{n} + 1.645\sqrt{\frac{p(1-p)}{n}}\right) \\
&= P\left(-1.645 \leq \frac{\frac{X}{n} - p}{\sqrt{\frac{p(1-p)}{n}}} \leq 1.645\right) \\
&\approx P(-1.645 \leq Z \leq 1.645) \quad \text{where } Z \sim N(0,1) \\
&= 2P(Z \leq 1.645) - 1 \\
&= 2(0.95) - 1 = 0.9
\end{aligned}
$$

(b) Since $X_i \sim Poisson\,(\mu)$, $i = 1, 2, \ldots, n$ where n is large then by the Central Limit Theorem

$$\frac{\bar{X} - \mu}{\sqrt{\frac{\mu}{n}}} \sim N\,(0, 1) \quad \text{approximately}$$

Therefore

$$\begin{aligned}
&P\left(\bar{X} - 1.96\sqrt{\frac{\mu}{n}} \leq \mu \leq \bar{X} + 1.96\sqrt{\frac{\mu}{n}}\right) \\
&= P\left(-1.96 \leq \frac{\bar{X} - \mu}{\sqrt{\frac{\mu}{n}}} \leq 1.96\right) \\
&\approx P\,(-1.96 \leq Z \leq 1.96) \quad \text{where } Z \sim N\,(0, 1) \\
&= 2P\,(Z \leq 1.96) - 1 \\
&= 2\,(0.975) - 1 = 0.95
\end{aligned}$$

(c) Since $X_i \sim Exponential\,(\theta)$, $i = 1, 2, \ldots, n$ where n then by the Central Limit Theorem

$$\frac{\bar{X} - \theta}{\sqrt{\frac{\theta^2}{n}}} \sim N\,(0, 1) \quad \text{approximately}$$

Therefore

$$\begin{aligned}
&P\left(\bar{X} - 2.576\sqrt{\frac{\theta^2}{n}} \leq \theta \leq \bar{X} + 2.576\sqrt{\frac{\theta^2}{n}}\right) \\
&= P\left(-2.576 \leq \frac{\bar{X} - \theta}{\sqrt{\frac{\theta^2}{n}}} \leq 2.576\right) \\
&\approx P\,(-2.576 \leq Z \leq 2.576) \quad \text{where } Z \sim N\,(0, 1) \\
&= 2P\,(Z \leq 2.576) - 1 \\
&= 2\,(0.995) - 1 = 0.99
\end{aligned}$$

10.9 (a) If you play n times then your expected profit is

$$E\,(S) = n\,[(1)\,(0.49) + (-1)\,(0.51)] = -0.02n$$

and the variance of your profit is

$$Var\,(S) = n\left[(1)^2\,(0.49) + (-1)^2\,(0.51) - (-0.02)^2\right] = 0.9996n$$

Since S is the sum of independent and identically distributed random variables then, by the Central Limit Theorem, S has approximately a $N(-0.02n, 0.9996n)$ distribution.

(b) If $n=20$, the possible values of S are $x=-20,-18,\ldots,-2,0,2,18,20$ and the continuity correction is $2/2=1$.

$$\begin{aligned} P(S\geq 0) &\approx P\left(Z\geq \frac{(0-1)-(-0.02)(20)}{\sqrt{0.9996\,(20)}}\right) \quad \text{where } Z\sim N(0,1) \\ &= P(Z\geq -0.13) = P(Z\leq 0.13) = 0.55172 \end{aligned}$$

If $n=50$, the possible values of S are $x=-50,-48,\ldots,-2,0,2,48,50$.

$$\begin{aligned} P(S\geq 0) &\approx P\left(Z\geq \frac{(0-1)-(-0.02)(50)}{\sqrt{0.9996\,(50)}}\right) \quad \text{where } Z\sim N(0,1) \\ &= P(Z\geq 0) = 0.5 \end{aligned}$$

If $n=100$, the possible values of S are $x=-100,-98,\ldots,-2,0,2,98,100$.

$$\begin{aligned} P(S\geq 0) &\approx P\left(Z\geq \frac{(0-1)-(-0.02)(100)}{\sqrt{0.9996\,(100)}}\right) \quad \text{where } Z\sim N(0,1) \\ &= P(Z\geq 0.10) = 1-P(Z\leq 0.10) = 1-0.53983 = 0.46017 \end{aligned}$$

The more you play, the smaller your chance of winning.

(c) For the casino owner, Y has approximately a $N(0.02n, 0.9996n)$ distribution. For $n=100,000$, Y has approximately a $N(2000, 99960)$ distribution. We want to find c such that $P(Y>c)=0.99$. Since

$$P(Y>c)\approx P\left(Z>\frac{(c+1)-2000}{\sqrt{99960}}\right) = P\left(Z>\frac{c-1999}{\sqrt{99960}}\right) \quad \text{where } Z\sim N(0,1)$$

and $P(Z>-2.3263)=0.99$ then

$$\frac{c-1999}{\sqrt{99960}} = -2.3263 \;\text{ or }\; c = 1999-(2.3263)\left(\sqrt{99960}\right) = 1263.506$$

With probability 0.99 the casino owner's profit is at least \$1263.51.

10.10 (a) Let T be the number of hearts which turn up. Then $T\sim Binomial\,(3,1/6)$ with $E(T)=3\,(1/6)=1/2$ and $Var\,(T)=3\,(1/6)\,(5/6)=5/12$. The profit for one play is $T-1$ with $E(T-1)=1/2-1=-1/2$ and $Var\,(T-1)=Var\,(T)=5/12$. If you play the game n times then your expected profit is

$$E(X)=n\,[E(T-1)]=-\frac{n}{2}$$

and the variance of your profit is

$$Var\,(X)=n\,[Var\,(T-1)]=\frac{5n}{12}$$

Since S is the sum of independent and identically distributed random variables then, by the Central Limit Theorem, S has approximately a $N(-n/2, 5n/12)$ distribution.

(b) (i) If $n = 10$

$$\begin{aligned} P(S > 0) &= P\left(Z > \frac{0 + 0.5 - (-5)}{\sqrt{50/12}}\right) = P(Z > 2.69) = 1 - P\left(Z < 2.69\right) \\ &= 1 - 0.99643 = 0.00357 \end{aligned}$$

(ii) If $n = 50$

$$P(S > 0) = P\left(Z > \frac{0 + 0.5 - (-25)}{\sqrt{250/12}}\right) = P(Z > 5.58677) \approx 0$$

10.11 (a)

$$\begin{aligned} M\left(t\right) &= E\left(e^{tX}\right) = \sum_{x=0}^{\infty} e^{tx} p\left(1-p\right)^{x} = p\sum_{x=0}^{\infty}\left[\left(1-p\right)e^{t}\right]^{x} \\ &= \frac{p}{1-(1-p)e^{t}} \quad \text{by the Geometric series for } t < -\ln(1-p) \\ &= \frac{p}{1-qe^{t}} \quad \text{where } q = 1-p \end{aligned}$$

(b)

$$\begin{aligned} M'\left(t\right) &= \frac{d}{dt}\left[p\left(1-qe^{t}\right)^{-1}\right] = p\left(-1\right)\left(1-qe^{t}\right)^{-2}\left(-qe^{t}\right) \\ &= pqe^{t}\left(1-qe^{t}\right)^{-2} = \frac{pqe^{t}}{\left(1-qe^{t}\right)^{2}} \\ E(X) &= M'(0) = \frac{pq}{p^{2}} = \frac{q}{p} \end{aligned}$$

$$\begin{aligned} M''\left(t\right) &= \frac{d}{dt}\left[pqe^{t}\left(1-qe^{t}\right)^{-2}\right] = pq\left[e^{t}\left(-2\right)\left(1-qe^{t}\right)^{-3}\left(-qe^{t}\right) + e^{t}\left(1-qe^{t}\right)^{-2}\right] \\ &= \frac{pqe^{t}\left[2qe^{t} + \left(1-qe^{t}\right)\right]}{\left(1-qe^{t}\right)^{3}} \\ E(X^{2}) &= M''(0) = \frac{pq\left(2q+1-q\right)}{\left(1-q\right)^{3}} = \frac{pq\left(1+q\right)}{p^{3}} = \frac{q\left(1+q\right)}{p^{2}} \end{aligned}$$

$$Var\left(X\right) = E(X^{2}) - \left[E(X)\right]^{2} = \frac{q\left(1+q\right)}{p^{2}} - \left(\frac{q}{p}\right)^{2} = \frac{q}{p^{2}}$$

10.12

$$M(t) = E\left(e^{tX}\right) = \frac{1}{b-a+1}\sum_{x=a}^{b} e^{xt} = \frac{e^{at}-e^{(b+1)t}}{(1-e^t)(b-a+1)} \text{ for } t \neq 0$$

$$M'(t) = \frac{1}{b-a+1}\sum_{x=a}^{b}\frac{d}{dt}e^{xt} = \frac{1}{b-a+1}\sum_{x=a}^{b} xe^{xt}$$
$$E(X) = M'(0) = \frac{1}{b-a+1}\sum_{x=a}^{b} x = \frac{1}{2(b-a+1)}\left[b(b+1)-(a-1)a\right]$$

$$M''(t) = \frac{1}{b-a+1}\sum_{x=a}^{b}\frac{d^2}{dt^2}e^{xt} = \frac{1}{b-a+1}\sum_{x=a}^{b} x^2e^{xt}$$
$$E(X^2) = M''(0) = \frac{1}{b-a+1}\sum_{x=a}^{b} x^2 = \frac{1}{6(b-a+1)}\left[b(b+1)(2b+1)-(a-1)a(2a-1)\right]$$

10.13 (a) Since X only takes on values $0, 1, 2$ the moment generating function of X is

$$\begin{aligned} M(t) &= e^{t(0)}P(X=0) + e^{t(1)}P(X=1) + e^{t(2)}P(X=2) \\ &= P(X=0) + e^{t}P(X=1) + e^{2t}P(X=2) \end{aligned}$$

Taking two derivatives with respect to t we have

$$\begin{aligned} M'(t) &= e^{t}P(X=1) + 2e^{2t}P(X=2) \\ M''(t) &= e^{t}P(X=1) + 4e^{2t}P(X=2) \end{aligned}$$

Since $M'(0) = E(X) = 1$ and $M''(0) = E\left(X^2\right) = 1.5$ we have

$$1 = E(X) = M'(0) = P(X=1) + 2P(X=2)$$

and

$$1.5 = E\left(X^2\right) = M''(0) = P(X=1) + 4P(X=2)$$

Solving these two equations in two unknowns gives $P(X=2) = 0.25$ and $P(X=1) = 0.5$ and thus $P(X=0) = 0.25$. Therefore

$$M(t) = 0.25 + 0.5e^{t} + 0.25e^{2t} \quad \text{for } t \in \Re$$

(b)

$$M^{(3)}(t) = e^{t}P(X=1) + 8e^{2t}P(X=2) = e^{t}(0.5) + 8e^{2t}(0.25)$$

and $E\left(X^3\right) = M^{(3)}(0) = 0.5 + 2 = 2.5$

$$M^{(4)}(t) = e^t P(X = 1) + 16e^{2t} P(X = 2) = e^t (0.5) + 16e^{2t} (0.25)$$
and $E\left(X^4\right) = M^{(4)}(0) = 0.5 + 4 = 4.5$

(c) Given the first two moments $E(X) = m_1$ and $E(X^2) = m_2$, there is a unique solution to the equations $p_0 + p_1 + p_2 = 1, p_1 + 2p_2 = m_1, p_1 + 4p_2 = m_2$ where $p_i = P(X = x), x = 1, 2, 3.$

10.14 (a) Expand $M(t)$ in a power series in powers of e^t, that is

$$\begin{aligned} M(t) &= \frac{1}{3e^{-t} - 2} = \frac{\frac{1}{3}e^t}{1 - \frac{2}{3}e^t} \\ &= \frac{1}{3}e^t \sum_{i=0}^{\infty} \left(\frac{2}{3}e^t\right)^i \quad \text{by the Geometric series if } \left|\frac{2}{3}e^t\right| < 1 \\ &= \sum_{i=0}^{\infty} \frac{1}{3}\left(\frac{2}{3}\right)^i e^{t(i+1)} \\ &= \sum_{x=1}^{\infty} \frac{1}{3}\left(\frac{2}{3}\right)^{x-1} e^{tx} \end{aligned}$$

Therefore

$$P(X = x) = \text{coefficient of } e^{xt} = \frac{1}{3}\left(\frac{2}{3}\right)^{x-1} \quad \text{for } x = 1, 2, \ldots$$

which we recognize as being the probability function of X = the total number of trials until the first success in a sequence of Bernoulli trials with $P(S) = \frac{1}{3}$.

(b)

$$\begin{aligned} M(t) &= e^{2\left(e^t - 1\right)} = e^{-2} e^{2e^t} \\ &= e^{-2} \sum_{x=0}^{\infty} \frac{\left(2e^t\right)^x}{x!} \quad \text{by the Exponential series for } t \in \Re \\ &= \sum_{x=0}^{\infty} \frac{e^{-2} 2^x}{x!} e^{tx} \end{aligned}$$

Therefore

$$P(X = x) = \text{coefficient of } e^{xt} = \frac{2^x e^{-2}}{x!} \quad \text{for } x = 0, 1, \ldots$$

which we recognize as being the probability function of a $Poisson\,(2)$ random variable.

10.15 (a)

$$\begin{aligned} M(t) &= E\left(e^{tX}\right) = \int_0^{\infty} e^{xt} \frac{1}{\theta} e^{-x/\theta} dx = \frac{1}{\theta} \int_0^{\infty} e^{-x\left(\frac{1}{\theta} - t\right)} dx \\ &= \frac{1}{\theta} \frac{1}{\left(\frac{1}{\theta} - t\right)} = \frac{1}{1 - \theta t} \quad \text{if } t < \frac{1}{\theta} \end{aligned}$$

If $t \geq \frac{1}{\theta}$, the integral $\int\limits_0^\infty e^{-x\left(\frac{1}{\theta}-t\right)}dx$ does not converge and the moment generating function does not exist for $t \geq \frac{1}{\theta}$.

(b)

$$M'(t) = \frac{d}{dt}(1-\theta t)^{-1} = (-1)(1-\theta t)^{-2}(-\theta) = \theta(1-\theta t)^{-2}$$
$$E(X) = M'(0) = \theta$$

$$M''(t) = \frac{d^2}{dt^2}\left[\theta(1-\theta t)^{-2}\right] = \theta\left[(-2)(1-\theta t)^{-3}(-\theta)\right] = 2\theta^2(1-\theta t)^{-3}$$
$$E\left(X^2\right) = M''(0) = 2\theta^2$$

$$Var(X) = E(X^2) - [E(X)]^2 = 2\theta^2 - (\theta)^2 = \theta^2$$

10.16 Recall that if $X \sim N\left(\mu, \sigma^2\right)$ then the moment generating function of X is $M_X(t) = e^{\mu t + \sigma^2 t^2/2}$ for $t \in \Re$. If $X_i \sim N(1,2)$, $i = 1, 2, \ldots, n$ then the moment generating function of X_i is $M_i(t) = E\left(e^{tX_i}\right) = e^{t+t^2}$ for $t \in \Re$, $i = 1, 2, \ldots, n$.

(a) The moment generating function of $Y = -3X_1 + 4$ is

$$\begin{aligned}
M_Y(t) &= E\left(e^{tY}\right) = E\left(e^{t(-3X_1+4)}\right) \\
&= e^{4t}E\left(e^{(-3t)X_1}\right) = e^{4t}M_1(-3t) = e^{4t}e^{-3t+(-3t)^2} \\
&= e^{t+9t^2} \quad \text{for } t \in \Re
\end{aligned}$$

which is the moment generating function of a $N(1,18)$ random variable. By the Uniqueness Theorem $Y \sim N(1,18)$.

(b) The moment generating function of $T = X_1 + X_2$ is

$$\begin{aligned}
M_T(t) &= E\left(e^{tT}\right) = E\left(e^{t(X_1+X_2)}\right) \\
&= E\left(e^{tX_1}\right)E\left(e^{tX_2}\right) \quad \text{since } X_1 \text{ and } X_2 \text{ are independent random variables} \\
&= \left(e^{t+t^2}\right)\left(e^{t+t^2}\right) \\
&= e^{2t+2t^2} \quad \text{for } t \in \Re
\end{aligned}$$

which is the moment generating function of a $N(2,4)$ random variable. By the Uniqueness Theorem, $T \sim N(2,4)$.

(c) The moment generating function of $S_n = X_1 + X_2 + \ldots + X_n$ is

$$\begin{aligned} M_{S_n}(t) &= E\left(e^{tS_n}\right) = E\left(e^{t(X_1+X_2+\ldots+X_n)}\right) \\ &= E\left(\prod_{i=1}^{n} e^{tX_i}\right) = \prod_{i=1}^{n} E\left(e^{tX_i}\right) \quad \text{since the } X_i\text{'s are independent random variables} \\ &= \prod_{i=1}^{n} M_i(t) = \prod_{i=1}^{n} e^{t+t^2} \\ &= e^{nt+nt^2} \quad \text{for } t \in \Re \end{aligned}$$

which is the moment generating function of a $N(n, 2n)$ random variable. By the Uniqueness Theorem, $S_n \sim N(n, 2n)$.

(d) The moment generating function of $Z = (2n)^{-1/2}(S_n - n)$ is

$$\begin{aligned} M_Z(t) &= E\left(e^{tZ}\right) = E\left(e^{t(2n)^{-1/2}(S_n-n)}\right) \\ &= e^{t\left(2^{-1/2}\right)n^{1/2}} E\left(e^{t(2n)^{-1/2}S_n}\right) = e^{t\left(2^{-1/2}\right)n^{1/2}} M_{S_n}\left(t\,(2n)^{-1/2}\right) \\ &= e^{t\left(2^{-1/2}\right)n^{1/2}} e^{n\left(t(2n)^{-1/2}\right)+n\left(t(2n)^{-1/2}\right)^2} \\ &= e^{t^2/2} \quad \text{for } t \in \Re \end{aligned}$$

which is the moment generating function of a $N(0, 1)$ random variable. By the Uniqueness Theorem, $Z \sim N(0, 1)$.

10.17 Since $X \sim Poisson(\lambda_1)$, the moment generating function of X is

$$M_X(t) = e^{-\lambda_1+\lambda_1 e^t}$$

Since $Y \sim Poisson(\lambda_2)$, the moment generating function of Y is

$$M_Y(t) = e^{-\lambda_2+\lambda_2 e^t}$$

Since X and Y are independent random variables, the moment generating function of the sum $X + Y$ is the product of the moment generating functions, that is,

$$M_X(t)M_Y(t) = e^{-\lambda_1+\lambda_1 e^t} e^{-\lambda_2+\lambda_2 e^t} = e^{-(\lambda_1+\lambda_2)+(\lambda_1+\lambda_2)e^t}$$

Note that this is the moment generating function of a Poisson distribution with parameter $\lambda_1+\lambda_2$. Therefore by the Uniqueness Theorem $X + Y \sim Poisson(\lambda_1 + \lambda_2)$.

10.18 (a) The moment generating function of X is

$$M(t) = E\left(e^{tX}\right) = \int_0^\infty e^{xt} \frac{1}{\theta^2} x e^{-x/\theta} dx = \frac{1}{\theta^2} \int_0^\infty x e^{-\left(\frac{1}{\theta}-t\right)x} dx = \frac{1}{(1-\theta t)^2} \quad \text{if } t < \frac{1}{\theta}$$

(b) From the solution to Chapter 10, Problem 15 we have

$$M_X(t) = M_Y(t) = \frac{1}{1-\theta t} \quad \text{for } t < \frac{1}{\theta}$$

Therefore the moment generating function of $S = Z + Y$ is

$$M_S(t) = E\left[e^{t(X+Y)}\right] = E(e^{tX})E(e^{tY}) = M_X(t)M_Y(t) = \frac{1}{(1-\theta t)^2} \quad \text{for } t < \frac{1}{\theta}$$

and since this is the moment generating function of the distribution obtained in (a), S must have the probability density function $f(s) = \frac{1}{\theta^2} s e^{-s/\theta}$ for $s > 0$.

10.19

10.20 Let $Y =$ total change over day. Given $N = n$, Y has a $N(0, n\sigma^2)$ distribution and therefore

$$\begin{aligned}
E\left(e^{tY}|N=n\right) &= \exp\left(\frac{n\sigma^2 t^2}{2}\right) \\
M_Y(t) = E\left(e^{tY}\right) &= \sum_{n=0}^{\infty} E[e^{tY}|N=n]P(N=n) \\
&= e^{-\lambda}\sum_{n=0}^{\infty} \exp\left(\frac{n\sigma^2 t^2}{2}\right)\frac{\lambda^n}{n!} \\
&= e^{-\lambda}\sum_{n=0}^{\infty} \frac{(e^{\sigma^2 t^2/2}\lambda)^n}{n!} \\
&= \exp(-\lambda + e^{\sigma^2 t^2/2}\lambda) \quad \text{by the Exponential series}
\end{aligned}$$

This is not a moment generating function we have seen in this course. The mean is $M_Y'(0) = 0$ and the variance is $M_Y''(0) = \lambda\sigma^2$.

11. SAMPLE TESTS

Sample Midterm 1

1. Four students were late for an exam. Their excuse was that the car they shared had a flat tire on the way. The instructor, suspecting that they were not telling the truth, asked them each separately which tire went flat. Assume each student will randomly pick one of the tires 1, 2, 3, or 4, independently of each other. Find the probability of each of the following events:

(a) $A =$ they all pick the same tire.

(b) $B =$ they all pick a different tire.

(c) $C =$ at least two of them pick the same tire.

(d) $D =$ exactly one student picks tire 1 and exactly one student picks tire 3.

2. The letters of the word PROBABILITY are arranged at random to form a “word”. Find the probability of each of the following events:

(a) $A =$ the word ends with the letter Y

(b) $B =$ the two B’s occur side by side in the word

(c) $C =$ the word ends with the letter Y and the two B’s occur side by side

(d) $D =$ the word does not end with the letter Y and the B’s do not occur side by side.

3. In a class of 60 students, 40% are international students. Five students are chosen at random.

(a) If the students are chosen **without** replacement, what is the probability none of them are international students?

(b) If the students are chosen **without** replacement, what is the probability at least two of them are international students?

(c) If the students are chosen **with** replacement, what is the probability none of them are international students?

(d) If the students are chosen **with** replacement, what is the probability exactly one of them is an international student?

4. (a) A and B are mutually exclusive events with $P(A) = 0.6$ and $P(B) = 0.3$. Find $P(A \cap B)$ and $P(A \cup B)$.

(b) A and B are independent events with $P(A) = 0.5$ and $P(B) = 0.1$. Find $P(A \cap B)$ and $P(A \cup B)$.

(c) $P(A) = 0.4$, $P(B) = 0.6$, and $P(A|B) = 0.5$. Find $P(B|A)$ and $P\left(\bar{B}|A\right)$.

(d) $P(B) = 0.3$, $P(A|B) = 0.6$, and $P\left(A|\bar{B}\right) = 0.2$. Find $P(A)$ and $P\left(\bar{A}\right)$.

5. Students Aziz, Bo and Chun each independently write a tutorial test. The probability of passing the test is 0.8 for Aziz, 0.6 for Bo, and 0.7 for Chun.

(a) Find the probability that at least one of them passes the test.

(b) Find the probability that exactly two of them pass the test.

(c) If exactly two of them pass the test, what is the probability it was Bo who did not pass the test?

6. In 2013, 10% of all immigrants to Canada were refugees. Forty-five percent of the refugees were under 25 years old, and 30% of the non-refugee immigrants were under 25 years old. A person is chosen at random from those who immigrated to Canada in 2013.

(a) What is the probability the randomly chosen person is a refugee and under 25 years old?

(b) What is the probability the randomly chosen person is a non-refugee immigrant and under 25 years old?

(c) What is the probability the randomly chosen person is under 25 years old?

(d) If the randomly chosen person is under 25 years old, then what is the probability the person is a refugee?

Sample Midterm 2

1. Traffic accidents at the intersection of University Avenue and Westmount Road occur according to a Poisson process with an average rate of 0.5 accidents per day. If no accidents occur during a week of seven days (Sunday to Saturday), the week is declared a "Safe-Week".

(a) Find the probability of a Safe-Week.

(b) Find the probability that in a period of 10 non-overlapping weeks there is at most 1 Safe-Week.

(c) Find the probability that there are 6 accidents during the two-week period November 1-14.

(d) Given that 6 accidents occurred during the two-week period November 1-14, find the probability that the first week (November 1-7) was a Safe-Week.

(e) Suppose an accident has just occurred. What is the expected waiting time until the next accident?

2. Suppose the random variable X has a $Geometric\,(p)$ distribution.

(a) Prove that $P(X \geq x) = (1-p)^x$ for $x = 0, 1, 2, \ldots$.
(b) Prove that $P(X \geq x + y \big| X \geq x) = P(X \geq y)$ for all non-negative integers x and y.

(c) Prove that $E\,(X) = (1-p)\,/p$. Be sure to show all your work.

(d) If p is the probability of success in a sequence of Bernoulli trials then find the expected total number of trials to obtain the first success.

3. X is a continuous random variable with cumulative distribution function

$$F(x) = \begin{cases} 0 & \text{if } x \leq 0, \\ 2x^2 & \text{if } 0 < x \leq \frac{1}{2} \\ 4x - 2x^2 - 1 & \text{if } \frac{1}{2} < x < 1 \\ 1 & \text{if } x \geq 1 \end{cases}$$

(a) Find $f\,(x)$, the probability density function of X, for all $x \in \Re$.

(b) Find $P(X \geq 0.2)$.

(c) Show $E\,(2X+1) = 2$.

(d) If $Var\,(X) = 1/24$, then find $E\left(X^2\right)$ WITHOUT using integration.

4. In a large population the probability a randomly chosen person has a rare disease is 0.02. An inexpensive diagnostic test gives a false positive result (person does not have the disease but the test says they do) with probability 0.05 and a false negative result (person has the disease but the test says they don't) with probability 0.01. The inexpensive test costs $\$10$. If a person tests positive they are given a more expensive diagnostic test that costs $\$100$ which correctly identifies all persons with the disease.

(a) What is the expected cost per person for this testing protocol?

(b) To reduce the number of cases being missed due to false negative results, a second test is added to the testing protocol above as follows: If a person tests negative on the first test using the inexpensive test then the person is tested again using the inexpensive test. If the second test is negative then no more testing is done. If the second test is positive then the person is tested with the more expensive test. What is the expected cost per person for this testing protocol?

5. A continuous random variable X has probability density function

$$f(x) = \theta x^{\theta - 1} \quad \text{for } 0 < x < 1$$

and zero otherwise where $\theta > 0$ is a constant.

(a) Find $P\left(X \leq 0.25\right)$.

(b) Find $E\left(X^k\right)$ for $k = 1, 2, \ldots$.

(c) Let $Y = -\theta \ln X$. Show that $Y \sim Exponential(1)$. Be sure to show all your work.

6. For each of the functions in the table indicate with a ✓ which of the statements A-M is true. For example, statement A is true for all these functions so there is a ✓ in each box in the column labelled A.

	A	B	C	D	E	F	G	H	I	J	K	L	M
the p.f. $f(x)$ of a discrete r.v.	✓			$*$									
the c.d.f. $F(x)$ of a discrete r.v.	✓												
the p.d.f. $f(x)$ of a continuous r.v.	✓			$*$									
the c.d.f. $F(x)$ of a continuous r.v.	✓												

$*$ = do not use this box

A: The value of the function is always non-negative.

B: Every value of the function lies in the interval $[0, 1]$.

C: The limit of the function as $x \to \infty$ equals 1.

D: The limit of the function as $x \to -\infty$ equals 0.

E: The domain of the function is countable.

F: The domain of the function is $\Re$.

G: The function is non-decreasing for all $x \in \Re$.

H: The function is increasing for all $x \in \Re$.

I: The function is right-continuous for all $x \in \Re$.

J: The function is continuous for all $x \in \Re$.

K: The sum of the function over all values of x equals 1.

L: The area bounded by the graph of the curve of the function and the $x-$axis equals 1.

M: The derivative of the function is equal to $P(X = x)$.

Sample Final Exam

Part A: Circle the letter corresponding to the correct answer.

1. Three numbers are drawn at random WITH replacement from the digits $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. The probability that there is a repeated number among the three numbers drawn is:

   A: $\frac{3\times 10^2+10}{10^3}$

   B: $1 - \frac{3\times 10^2}{10^3}$

   C: $1 - \frac{10\times 9\times 8}{10^3}$

   D: $\frac{10\times 9+10}{10^3}$

   E: $\frac{3\times 10\times 8}{10\times 9\times 8}$

2. If two events A and B are independent and mutually exclusive, then:

   A: this is impossible

   B: A must have a probability 1

   C: both A and B must have probability 1

   D: both A and B must have probability 0

   E: either A or B (or both) have probability 0

3. In a specific population 50% of all people are males. Five percent of the males are colour-blind, and 0.25% of the females are colour-blind. If a randomly chosen person is colour-blind, then the probability, to 3 decimal places, that the person is a male is:

   A: 0.050

   B: 0.025

   C: 0.952

   D: 0.026

   E: None of the above

4. Sharks normally attack swimmers at Myhammy Beach on average about one day in 200. There have been no shark attacks in the last 400 days. The probability of this happening is approximately:

   A: $\frac{1}{2}$

   B: e^{-1}

   C: $2e^{-2}$

   D: e^{-2}

   E: none of the above

5. Suppose new posts on a forum occur independently at a constant rate of 3 posts per half hour. The probability that exactly 20 non-overlapping minutes in a half-hour period contain no new posts is:

   A: $\binom{30}{20}\left(e^{-0.1}\right)^{10}\left(1-e^{-0.1}\right)^{20}$

   B: $\binom{30}{20}\left(e^{-0.1}\right)^{20}\left(1-e^{-0.1}\right)^{10}$

   C: $\binom{29}{19}\left(e^{-0.1}\right)^{20}\left(1-e^{-0.1}\right)^{10}$

   D: $\binom{30}{20}\left(e^{-0.1}\right)^{30}$

   E: none of the above

6. Suppose X is a non-negative random variable with $E\left(X^2\right) = 6$ and $Var\left(X\right) = 2$ then

   A: $E\left(X\right) = 2$

   B: $E\left(X\right) = 4$

   C: $E\left(X\right) = -2$

   D: $E\left(X\right) = 6$

   E: there is not enough information to determine $E\left(X\right)$.

7. A certain river floods every year. Suppose the low-water mark is set at one meter and the high-water mark is modeled by the random variable X with cumulative distribution function:

$$F(x) = \begin{cases} 1 - \frac{1}{x^2} & x \geq 1 \\ 0 & x < 1 \end{cases}$$

   The probability that the high-water mark is greater than 3m but less than 4m is:

   A: $\frac{137}{144}$

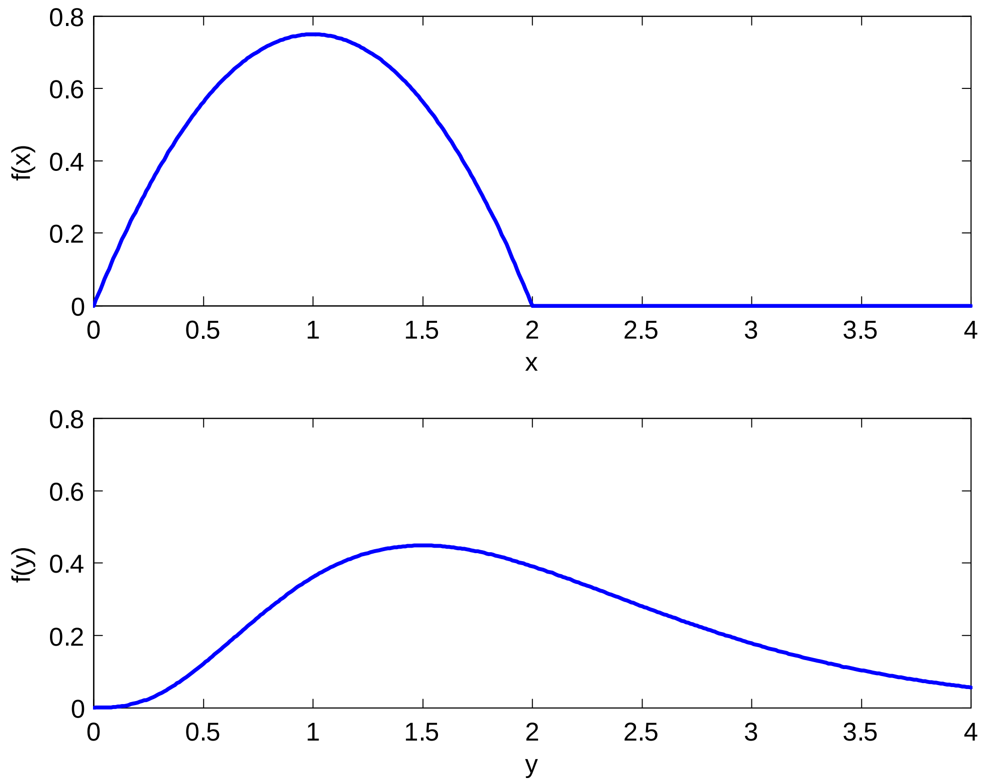
   B: $\frac{7}{144}$

   C: $\frac{16}{144}$

   D: $\frac{9}{144}$

   E: none of the above

8. In the top half of the graph below is the probability density function of the random variable X and in the bottom half is the probability density function of the random variable Y. Assume the probability density function equals 0 outside the visible area of the graphs.



Which one of the following statements is false?

A: $E(Y) > E(X)$

B: $E(X) = 1$

C: $P(X = 1) = P(Y = 1)$

D: $sd(X) > sd(Y)$

E: $Var(X) \leq 1$

9. Suppose that $X \sim Uniform(1, 6)$ and $Y \sim Uniform(1, 20)$. Which one of the following statements is true?

   A: $P(X > 3) < P(Y > 10)$

   B: $P(X > 3) = P(Y > 10)$

   C: $P(X > 3) > P(Y > 10)$

   D: Not enough information to determine.

10. Suppose $X \sim Exponential(2)$. Then $P\left(X < 3|X > 1\right)$ is equal to:

    A: $\frac{1}{2}e^{-1}$

    B: e^{-1}

    C: $1 - e^{-1}$

    D: $1 - e^{-2}$

    E: e^{-2}

11. Suppose X is a random variable with $Var\left(X\right) > 0$. Which one of the following statements is true?

    A: $E(X^2) > [E(X)]^2$

    B: $E(X^2) = [E(X)]^2$

    C: $E(X^2) < [E(X)]^2$

    D: Not enough information to determine.

12. Average daily caffeine consumption is 165 mg. Ninety-nine percent of people consume less than 380 mg. Assuming daily caffeine consumption follows a Normal distribution, the standard deviation σ is:

    A: 130.7

    B: 107.5

    C: 167.8

    D: 92.4

    E: none of the above

13. Suppose $X \sim Poisson(2)$, $Y \sim Poisson(3)$, and that X and Y are independent. The joint probability function of X and Y is:

A: $f(x,y) = e^{-6}\frac{2^x 3^y}{x!y!}, \;\; x = 0,1,2,...; y = 0,1,2,...$

B: $f(x,y) = e^{-5}\frac{5^{x+y}}{x!y!}, \;\; x = 0,1,2,...; y = 0,1,2,...$

C: $f(x,y) = e^{-5}\frac{2^x 3^y}{x!y!}, \;\; x = 0,1,2,...; y = 0,1,2,...$

D: $f(x,y) = e^{-6}\frac{6^{x+y}}{x!y!}, \;\; x = 0,1,2,...; y = 0,1,2,...$

E: none of the above

14. Suppose $X \sim N(-2,1)$, $Y \sim N(2,4)$ and $Z \sim N(0,1)$ independently. Let $W = -3X + Y + 2Z$. Which one of the following statements is true?

A: $W \sim N(3,9)$

B: $W \sim N(8,9)$

C: $W \sim U(-4,17)$

D: $W \sim N(8,17)$

E: $W \sim N(3,17)$

15. The random variable which would be the LEAST accurately approximated using the Central Limit Theorem is:

A: the sum on 40 fair 6-sided dice.

B: the average grade of 913 students in STAT 230.

C: the total waiting time for 5 events in a Poisson process with rate $\lambda = 10$ events per hour.

D: the number of Heads in 50 flips of a fair coin.

E: the number of events in 5 hours in a Poisson process with rate $\lambda = 10$ events per hour.

16. If $X \sim Binomial(100, 0.4)$ then $P(X \geq 45)$ is best approximated by:

A: $P\left(Z \geq \frac{45.5-40}{\sqrt{24}}\right)$ where $Z \sim N(0,1)$

B: $P\left(Z \geq \frac{44.5-40}{\sqrt{24}}\right)$ where $Z \sim N(0,1)$

C: $P\left(Z \geq \frac{45-40}{\sqrt{24}}\right)$ where $Z \sim N(0,1)$

D: $P\left(Z \geq \frac{46.5-40}{\sqrt{24}}\right)$ where $Z \sim N(0,1)$

E: none of the above

Part B: Fill in the blank

1. For each of (a) to (j) choose the appropriate name of the distribution for the random variable X from the following list: **Discrete Uniform, Hypergeometric, Binomial, Negative Binomial, Geometric, Poisson, Continuous Uniform, Exponential, Normal, and Multinomial**:

   (a) A researcher is interested in studying a rare disease among beavers in Algonquin National Park. The researcher decides to capture and test beavers until the first beaver with the disease is found. $X =$ number of disease-free beavers tested by the researcher.

   ____

   (b) The pointer on a circular spinner is spun. $X =$ point on the circumference of the circle at which the pointer stops (assume almost no friction).

   ____

   (c) An instructor has n identical looking keys in the bottom of her knapsack and only one of the keys opens the door to her office. She draws a key from her knapsack and tries to open her office door. If the key does not work she draws another key. She continues this process until she obtains the correct key. $X =$ the draw on which she obtains the correct key where the draws are numbered 1 (1st draw), 2 (2nd draw), etc.

   ____

   (d) The probability of winning any prize in a weekly lottery is p. Jamie decides to purchase one lottery ticket each week until s/he wins 3 prizes. $X =$ number of weeks in which s/he wins no prizes.

   ____

   (e) Electrical power failures in a large Canadian city occur independently of each other throughout the year at a uniform rate with little chance of more than one failure on a given day.
   $X =$ number of power failures in a month.

   ____

   (f) In a shipment of N smartphones there are D defective smartphones. A sample of n smartphones are chosen at random and tested. $X =$ number of defective smartphones in the sample.

   ____

   (g) Aziz, Bo and Chow play a game together in which Aziz wins with probability p, Bo wins with probability q, and Chow wins with probability r $(p + q + r = 1)$. They play the game n times. $X =$ number of times Aziz or Bo wins.

   ____

(h) Since men tend to have larger feet on average than women a very long footprint at a crime scene might indicate the criminal is male. A criminal investigator randomly selects 100 males and measures their right foot in centimeters. $X =$ length of right foot in centimeters of a randomly chosen male from the 100 measured.

____

(i) Hits on a particular website occur independently of each other at a uniform rate throughout the day with little chance of more than one hit in a one minute interval. $X =$ waiting time between consecutive hits on the website.

____

(j) In a very large city, the probability that a randomly chosen person supports a new bylaw banning Christmas decorations until after Remembrance Day is equal to p. A sample of 100 people are selected at random. $X =$ number of people in sample who support the bylaw.

____

2. Here are five concepts covered in STAT 230:

A: Bernoulli trials

B: Poisson process

C: Binomial approximation to the Hypergeometric

D: Poisson approximation to the Binomial

E: Central Limit Theorem

For each of the following statements indicate with a letter **A**, **B**, **C**, **D**, or **E** which of the above concepts is **best** associated with that statement.

(a) n random variables are independent and identically distributed with mean μ and variance σ^2.

____

(b) Events occur at a uniform rate over time.

____

(c) Trials are independent.

____

(d) The probability p of one of only two possible outcomes is constant on each trial.

____

(e) The number of random draws n made without replacement from a population of two types of items is small relative to the size of the population.

____

(f) The probability of 2 or more events in a sufficiently short period of time is approximately zero.

____

(g) The probability p of one of only two possible outcomes is constant and small on each trial.

____

(h) The number of events occurring in non-overlapping time intervals are independent.

____

(i) The number of random variables n in the sum or average approaches ∞.

____

(j) The number of independent trials n is large.

____

Part C: Long Answer

1. X and Y are discrete random variables with joint probability function

$f(x,y)$ $P(X=x,Y=y)$		x 0	1	2	
	1	0.15	0.05	0.15	
y	0	0.15	0.05	0.20	
	-1	0.10	0.00	0.15	

   (a) Are X and Y independent random variables? Justify your answer.

   (b) Find the covariance of X and Y.

   (c) Find the correlation coefficient of X and Y.

   (d) Find $Var(2X - Y + 1)$.

   (e) Tabulate the conditional probability function of X given $Y = 0$.

   (f) Tabulate the probability function of $T = X + Y$.

2. The weights of full-term babies born in Ontario are Normally distributed with mean $\mu = 3.5$ kg and standard deviation $\sigma = 0.5$ kg.

   (a) What proportion of full-term babies born in Ontario weigh more than 4.25 kg?

   (b) What proportion of full-term babies born in Ontario weigh between 3.1 and 4.25 kg?

   (c) What proportion of full-term babies born in Ontario have weights within one standard deviation of the mean?

   (d) A sample of 9 babies is drawn from all full-term babies born in Ontario in 2014. Give an expression for the probability that exactly 1 baby weighs more than 4.25 kg, exactly 5 babies weigh between 3.1 kg and 4.25 kg, and exactly 3 babies weigh less than 3.1 kg. You do not need to evaluate the expression.

   (e) A sample of 9 babies is drawn from all full-term babies born in Ontario in 2014. What is the probability that their average weight exceeds 3.4 kg?

   (f) Let $\bar{X} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$ be the average weight of n babies chosen at random. Find the smallest value of n such that $P\left(\left|\bar{X} - 3.5\right| \leq 0.05\right) \geq 0.9$.

3. Suppose X is a continuous random variable with probability density function:

$$f(x) = \begin{cases} \frac{1}{\theta^2} x e^{-x/\theta} & x \geq 0 \\ & \\ 0 & \text{otherwise} \end{cases}$$

   (a) Using the Gamma function or integration by parts show that $E(X^k) = \theta^k (k+1)!$ for $k = 1, 2, \ldots$.

   (b) Use the result given in (a) to find $E(X)$ and $Var(X)$.

   (c) Find the probability density function of $Y = \sqrt{X}$.

   (d) Suppose that $X_1, X_2, \ldots, X_{98}$ are independent random variables, each having the probability density function $f(x)$. Let $\bar{X} = \sum\limits_{i=1}^{98} X_i/98$ denote the sample mean. Use a suitable approximation to calculate the probability

$$P\left(\frac{|\bar{X} - 2\theta|}{\theta/7} < 1.15\right)$$

4. Ten friends go to an all-you-can-eat sushi restaurant and sit at one large round table. Each person likes spicy food with probability 0.6, independently of each other. We say a "match" occurs when two people sitting next to each other BOTH like spicy food or BOTH do not like spicy food. Let

$$X_i = \begin{cases} 1 & \text{if there is a "match" between person } i \text{ and person } i+1 \\ 0 & \text{otherwise} \end{cases}$$

for $i = 1, 2, ..., 10$ where person 11 is defined to be person 1 since they are at a circular table.

   (a) Find the expected value of X_i.

   (b) Find the expected total number of "matches" at the table.

   (c) Find the variance of X_i.

   (d) Show that the covariance between X_1 and X_2 is exactly 0.0096.

   (e) Find the variance of the total number of "matches" at the table.

12. SOLUTIONS TO SAMPLE TESTS

Sample Midterm 1 Solutions

1. (a) $A =$ they all pick the same tire.

Sample space $= \{1111, 1122, 1234, \ldots, 4444\}$ = the set of all 4^4 permutations of the numbers $1, 2, 3, 4$ with repeats. All outcomes are equally probable.

Since $A = \{1111, 2222, 3333, 4444\}$ then

$$P(A) = \frac{4}{4^4} = \frac{1}{4^3} = \frac{1}{64} = 0.016$$

(b) $B =$ they all pick a different tire.

$B =$ the set of all $4!$ permutations of the numbers $1, 2, 3, 4$ without repeats. Therefore

$$P(B) = \frac{4!}{4^4} = \frac{3}{32} = 0.094$$

(c) $C =$ at least two of them pick the same tire.

Since the complement of the event 'at least 2 of them pick the same tire' is the event 'they all pick a different tire' therefore

$$P(C) = 1 - P(B) = 1 - \frac{4!}{4^4} = 1 - \frac{3}{32} = \frac{29}{32} = 0.906$$

(d) $D =$ exactly one student picks tire 1 and exactly one student picks tire 3.

$D =$ the set of all $4!$ permutations of the numbers 1234, all $\frac{4!}{2!1!1!}$ permutations of the numbers 1322, and all $\frac{4!}{2!1!1!}$ permutations of the numbers 1344

$$P(D) = \frac{4! + (2)\,\frac{4!}{2!1!1!}}{4^4} = \frac{24 + 24}{4^4} = \frac{3}{16} = 0.188$$

2. (a) $A =$ the word ends with the letter Y

Sample space = the set of all $\frac{11!}{2!2!}$ permutations of the letters BBIIPROALTY. All outcomes are equally probable.

There is only 1 way to place the Y. The remaining 10 letters can be arranged in $\frac{10!}{2!2!}$ ways. Therefore

$$P(A) = \frac{(1)\frac{10!}{2!2!}}{\frac{11!}{2!2!}} = \frac{1}{11} = 0.091$$

(b) $B =$ the two B's occur side by side in the word

Consider BB as one letter. The number of arrangements of the letters $\boxed{\text{BB}}$IIPROALTY is $\frac{10!}{2!}$. Therefore

$$P(B) = \frac{\frac{10!}{2!}}{\frac{11!}{2!2!}} = \frac{2}{11} = 0.182$$

(c) $C =$ the word ends with the letter Y and the two B's occur side by side

There is only 1 way to place the Y and we consider BB as one letter. The number of arrangements of $\boxed{\text{BB}}$IIPROALT is $\frac{9!}{2!}$.

Therefore

$$P(C) = P(A \cap B) = \frac{(1)\frac{9!}{2!}}{\frac{11!}{2!2!}} = \frac{2}{110} = \frac{1}{55} = 0.018$$

(d) $D =$ the word does not end with the letter Y and the B's do not occur side by side.

$$\begin{aligned} P(D) &= P\left(\overline{A} \cap \overline{B}\right) = P\left(\overline{A \cup B}\right) \quad \text{by De Morgan's Laws} \\ &= 1 - P(A \cup B) \\ &= 1 - [P(A) + P(B) - P(A \cap B)] \quad \text{by the Sum Rule} \\ &= 1 - \left[\frac{1}{11} + \frac{2}{11} - \frac{1}{55}\right] = \frac{41}{55} \\ &= 0.745 \end{aligned}$$

3.(a) Let A_i be the event exactly i international students are chosen, $i = 0, 1, \ldots, 5$. Then

$$P(A_i) = \frac{\binom{24}{i}\binom{36}{5-i}}{\binom{60}{5}} \quad i = 0, 1, \ldots, 5$$

Therefore

$$P(\text{none are international students}) = P(A_0) = \frac{\binom{24}{0}\binom{36}{5}}{\binom{60}{5}} = \frac{\binom{36}{5}}{\binom{60}{5}} = 0.069$$

(b) If the students are chosen **without** replacement

$$\begin{aligned} P(\text{at least 2 are international students}) &= 1 - P(A_0) - P(A_1) \\ &= 1 - \frac{\binom{36}{5}}{\binom{60}{5}} - \frac{\binom{24}{1}\binom{36}{4}}{\binom{60}{5}} \\ &= 1 - 0.069 - 0.259 \\ &= 0.672 \end{aligned}$$

Alternatively

$$\begin{aligned} P(\text{at least 2 are international students}) &= P(A_2) + P(A_3) + P(A_4) + P(A_5) \\ &= \sum_{i=2}^{5} \frac{\binom{24}{i}\binom{36}{5-i}}{\binom{60}{5}} \\ &= 0.3608 + 0.2335 + 0.0700 + 0.0078 \\ &= 0.672 \end{aligned}$$

(c) If the students are chosen **with** replacement

$$P(\text{none are international students}) = \frac{36^5}{60^5} = (0.6)^5 = 0.078$$

(d) If the students are chosen **with** replacement, then the probability we draw the international student first followed by 4 non-international students is $\frac{24 \times 36^4}{60^5} = (0.4)(0.6)^4$. However the international students could also be drawn on the 2nd, 3rd, 4th, and 5th draws. Therefore

$$P(\text{exactly 1 international student}) = 5(0.4)(0.6)^4 = 0.259$$

4. (a) Since A and B are mutually exclusive events $P(A \cap B) = 0$ and

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) = 0.6 + 0.3 \\ &= 0.9 \end{aligned}$$

(b) Since A and B are independent events $P(A \cap B) = P(A)P(B) = (0.5)(0.1) = 0.05$ and

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \quad \text{by the Sum Rule} \\ &= 0.5 + 0.1 - 0.05 \\ &= 0.55 \end{aligned}$$

(c)

$$\begin{aligned} P(A \cap B) &= P(A|B)P(B) \quad \text{by the Product Rule} \\ &= (0.5)(0.6) = 0.3 \end{aligned}$$

$$\begin{aligned} P(B|A) &= \frac{P(A \cap B)}{P(A)} = \frac{0.3}{0.4} = \frac{3}{4} = 0.75 \\ P\left(\bar{B}|A\right) &= 1 - P(B|A) = 1 - 0.75 = 0.25 \end{aligned}$$

(d)

$$\begin{aligned} P(A) &= P(A \cap B) + P\left(A \cap \bar{B}\right) \\ &= P(A|B)P(B) + P\left(A|\bar{B}\right)P\left(\bar{B}\right) \quad \text{by the Product Rule} \\ &= (0.6)(0.3) + (0.2)(0.7) = 0.18 + 0.14 \\ &= 0.32 \end{aligned}$$

$$\begin{aligned} P\left(\bar{A}\right) &= 1 - P(A) = 1 - 0.32 \\ &= 0.68 \end{aligned}$$

5. (a) Let A be the event Aziz passes, B be the event Bo passes, and C be the event Chun passes. These events are independent events with $P(A) = 0.8$, $P(B) = 0.6$, and $P(C) = 0.7$.

$$\begin{aligned}
&P(\text{at least 1 passes}) \\
&= 1 - P(\text{none of them pass}) = 1 - P\left(\overline{A} \cap \overline{B} \cap \overline{C}\right) \\
&= 1 - P\left(\overline{A}\right) P\left(\overline{B}\right) P\left(\overline{C}\right) \quad \text{since the events are independent} \\
&= 1 - (0.2)(0.4)(0.3) = 1 - 0.024 \\
&= 0.976
\end{aligned}$$

(b)

$$\begin{aligned}
&P(\text{exactly 2 pass}) \\
&= P\left(A \cap B \cap \overline{C}\right) + P\left(A \cap \overline{B} \cap C\right) + P\left(\overline{A} \cap B \cap C\right) \\
&= P(A) P(B) P\left(\overline{C}\right) + P(A) P\left(\overline{B}\right) P(C) + P(A) P(B) P\left(\bar{C}\right) \\
&\qquad \text{since the events are independent} \\
&= (0.8)(0.6)(0.3) + (0.8)(0.4)(0.7) + (0.2)(0.6)(0.7) \\
&= 0.144 + 0.224 + 0.084 \\
&= 0.452
\end{aligned}$$

(c)

$$\begin{aligned}
&P(\text{Bo did not pass the test} \mid \text{exactly 2 pass}) \\
&= P\left(A \cap \overline{B} \cap C \mid \text{exactly 2 pass}\right) \\
&= \frac{P\left(A \cap \overline{B} \cap C \ \cap \ \text{exactly 2 pass}\right)}{P(\text{exactly 2 pass})} \\
&= \frac{P\left(A \cap \overline{B} \cap C\right)}{P(\text{exactly 2 pass})} \\
&= \frac{0.224}{0.452} \\
&= 0.496
\end{aligned}$$

6. (a) Let R be the event the person is a refugee and let A be the even the person is under 25 years old.

$$\begin{aligned}
&P\left(\text{person is a refugee and under 25 years old}\right)\\
&= P\left(R\cap A\right) = P\left(A|R\right)P\left(R\right) \quad \text{by the Product Rule}\\
&= \left(0.45\right)\left(0.1\right)\\
&= 0.045
\end{aligned}$$

(b)

$$\begin{aligned}
&P\left(\text{person is a non-refugee immigrant and under 25 years old}\right)\\
&= P\left(\overline{R}\cap A\right)\\
&= P\left(A|\overline{R}\right)P\left(\overline{R}\right) \quad \text{by the Product Rule}\\
&= \left(0.3\right)\left(1-0.1\right)\\
&= 0.27
\end{aligned}$$

(c)

$$\begin{aligned}
&P\left(\text{person is under 25 years old}\right)\\
&= P\left(A\right)\\
&= P\left(R\cap A\right) + P\left(\overline{R}\cap A\right)\\
&= 0.045 + 0.27\\
&= 0.315
\end{aligned}$$

(d)

$$\begin{aligned}
&P\left(\text{person is a refugee} \mid \text{person is under 25 years old}\right)\\
&= P\left(R|A\right)\\
&= \frac{P\left(A\cap R\right)}{P\left(A\right)}\\
&= \frac{0.045}{0.315}\\
&= 0.143
\end{aligned}$$

Sample Midterm 2 Solutions

1. (a) Accidents occur at the average rate of 0.5 accidents per day or $(7)(0.5) = 3.5$ accidents per week (7 days).

$$P(\text{Safe-Week}) = P(0 \text{ accidents in } 1 \text{ week}) = \frac{(3.5)^0 e^{-3.5}}{0!} = e^{-3.5} = 0.030$$

(b) Let $Y =$ number of Safe-Weeks in a 10 week period. Then $Y \sim Binomial\left(10, e^{-3.5}\right)$.

$$\begin{aligned}
&P(\text{there is at most } 1 \text{ Safe-Week in a } 10 \text{ week period})\\
&= P(Y \leq 1)\\
&= \binom{10}{0}\left(e^{-3.5}\right)^0\left(1-e^{-3.5}\right)^{10} + \binom{10}{1}\left(e^{-3.5}\right)^1\left(1-e^{-3.5}\right)^9\\
&= \left(1-e^{-3.5}\right)^{10} + 10e^{-3.5}\left(1-e^{-3.5}\right)^9\\
&= 0.965
\end{aligned}$$

(c) Accidents occur at the average rate of 0.5 accidents per day or $(14)(0.5) = 7$ accidents per two-week period (14 days).

$$P(6 \text{ accidents in } 2 \text{ week period}) = \frac{(7)^6 e^{-7}}{6!} = 0.149$$

(d)

$$\begin{aligned}
&P(0 \text{ accidents the 1st week} \mid 6 \text{ accidents in 2-week period})\\
&= \frac{P(0 \text{ accidents the 1st week and } 6 \text{ accidents in 2-week period})}{P(6 \text{ accidents in 2-week period})}\\
&= \frac{P(0 \text{ accidents the 1st week and } 6 \text{ accidents the 2nd week})}{P(6 \text{ accidents in 2-week period})}\\
&= \frac{\frac{(3.5)^0 e^{-3.5}}{0!}\frac{(3.5)^6 e^{-3.5}}{6!}}{\frac{(7)^6 e^{-7}}{6!}}\\
&= \frac{(3.5)^6}{(7)^6} = (0.5)^6\\
&= 0.016
\end{aligned}$$

(e) Let $Y =$ waiting time until the next accident. Since accidents occur at the average rate of 0.5 accidents per day, then $Y \sim Exponential\left(\frac{1}{0.5}\right)$ and $E(Y) = \frac{1}{0.5} = 2$ days.

2. (a)

$$\begin{aligned} P(X \geq x) &= P(X = x) + P(X = x+1) + P(X = x+2) + \cdots \\ &= p(1-p)^x + p(1-p)^{x+1} + p(1-p)^{x+2} + \cdots \quad \text{which is a Geometric series} \\ &= \frac{p(1-p)^x}{1-(1-p)} \\ &= (1-p)^x \ \text{ for } x = 0, 1, \ldots \end{aligned}$$

(b)

$$\begin{aligned} P(X \geq x+y \,|\, X \geq x) &= \frac{P(X \geq x+y \,\text{and}\, X \geq x)}{P(X \geq x)} = \frac{P(X \geq x+y)}{P(X \geq x)} \\ &= \frac{(1-p)^{x+y}}{(1-p)^x} \\ &= (1-p)^y = P(X \geq y) \end{aligned}$$

which holds for all non-negative integers x and y.

(c) By the Geometric series we have

$$a\sum_{i=0}^{\infty} r^i = \frac{a}{1-r}, \quad |r| < 1$$

By differentiating with respect to r we obtain

$$a\sum_{i=1}^{\infty} i r^{i-1} = \frac{a}{(1-r)^2}, \quad |r| < 1$$

Therefore

$$E(X) = \sum_{x=1}^{\infty} xp(1-p)^x = p(1-p)\sum_{x=1}^{\infty} x(1-p)^{x-1} = \frac{p(1-p)}{[1-(1-p)]^2} = \frac{1-p}{p}$$

(d) Let $N =$ total number of trials to obtain the first success. Then $N = X + 1$ and

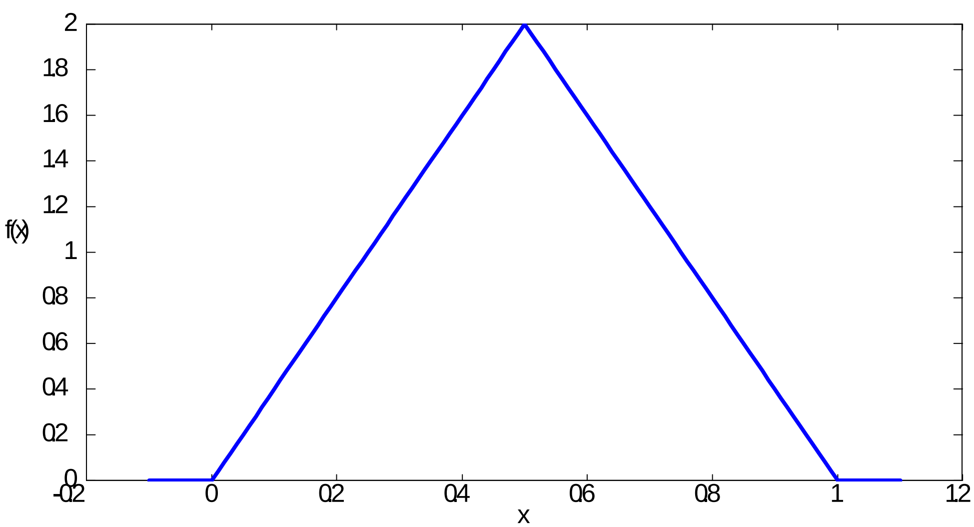
$$\begin{aligned} E(N) &= E(X+1) = E(X) + 1 \\ &= \frac{1-p}{p} + 1 = \frac{1-p+p}{p} \\ &= \frac{1}{p} \end{aligned}$$

3. (a) Since

$$F(x) = \begin{cases} 0 & \text{if } x \le 0, \\ 2x^2 & \text{if } 0 < x \le \frac{1}{2} \\ 4x - 2x^2 - 1 & \text{if } \frac{1}{2} < x < 1 \\ 1 & \text{if } x \ge 1 \end{cases}$$

$$f(x) = \frac{d}{dx}F(x) = \begin{cases} 0 & \text{if } x \le 0, \text{ or if } x \ge 1 \\ 4x & \text{if } 0 < x \le \frac{1}{2} \\ 4 - 4x & \text{if } \frac{1}{2} < x < 1 \end{cases}$$

Note: $F'(x)$ does not exist at $x = 0, 0.5, 1$ so we arbitrarily define $f(0) = 0 = f(1)$, and $f(0.5) = 2$.



(b)

$$P(X \ge 0.2) = 1 - P(X \le 0.2) = 1 - F(0.2) = 1 - 2(0.2)^2 = 1 - 0.08 = 0.92$$

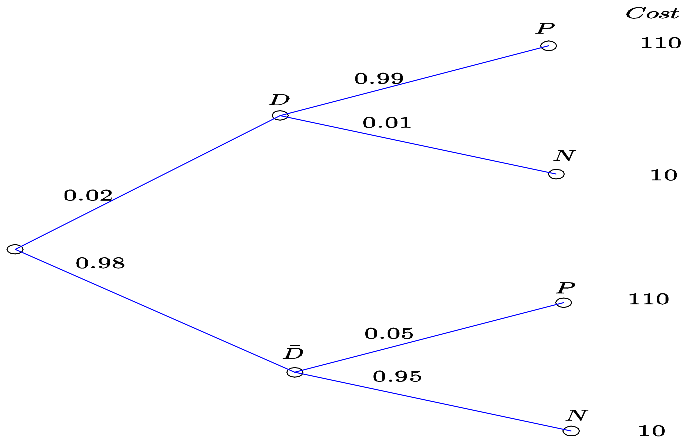
(c) By symmetry of the probability density function $E(X) = \frac{1}{2}$ and therefore

$$E(2X+1) = 2E(X) + 1 = 2\left(\frac{1}{2}\right) + 1 = 2$$

(d)

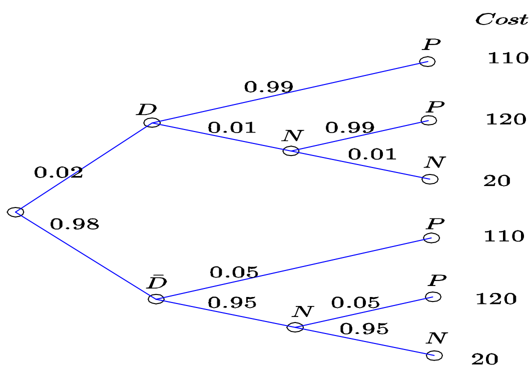
$$E(X^2) = Var(X) + [E(X)]^2 = \frac{1}{24} + \left(\frac{1}{2}\right)^2 = \frac{7}{24} = 0.292$$

4. (a)



$$\text{Expected Cost} = 110\,[(0.02)\,(0.99) + (0.98)\,(0.05)] + 10\,[(0.02)\,(0.01) + (0.98)\,(0.95)] = 16.880$$

(b)



$$\begin{aligned}\text{Expected Cost} &= 110\,[(0.02)\,(0.99) + (0.98)\,(0.05)] \\ &\quad + 20\left[(0.02)\,(0.01)^2 + (0.98)\,(0.95)^2\right] \\ &\quad + 120\,[(0.02)\,(0.01)\,(0.99) + (0.98)\,(0.95)\,(0.05)] \\ &= 30.867\end{aligned}$$

5. (a)

$$
\begin{aligned}
P(X \le 0.25) &= \int_0^{0.25} \theta x^{\theta-1} dx \\
&= x^{\theta}|_0^{0.25} \\
&= (0.25)^{\theta}
\end{aligned}
$$

(b)

$$
\begin{aligned}
E\left(X^k\right) &= \int_0^1 x^k \cdot \theta x^{\theta-1} dx \\
&= \theta \int_0^1 x^{\theta+k-1} dx \\
&= \theta \left(\frac{1}{\theta+k}\right) \left(x^{\theta+k}|_0^1\right) \\
&= \frac{\theta}{\theta+k} \quad \text{for } k = 1, 2, \ldots
\end{aligned}
$$

(c) Let $F(x) = P(X \le x)$ be the cumulative distribution function for X and $G(y) = P(Y \le y)$ be the cumulative distribution function for $Y = -\theta \ln X$. For $y > 0$

$$
\begin{aligned}
G(y) &= P(Y \le y) \\
&= P(-\theta \ln X \le y) \\
&= P\left(X \ge e^{-y/\theta}\right) \\
&= 1 - F\left(e^{-y/\theta}\right)
\end{aligned}
$$

For $y > 0$ the probability density function for Y is

$$
\begin{aligned}
g(y) &= \frac{d}{dy}\left[1 - F\left(e^{-y/\theta}\right)\right] \\
&= -f\left(e^{-y/\theta}\right) \frac{d}{dy}\left(e^{-y/\theta}\right) \quad \text{by the Chain Rule} \\
&= -\theta \left(e^{-y/\theta}\right)^{\theta-1} \left(-\frac{1}{\theta} e^{-y/\theta}\right) \\
&= e^{-y}
\end{aligned}
$$

For $y \le 0$, $g(y) = 0$. Since $g(y)$ is the probability density function for an $Exponential(1)$ random variable we have shown that $Y = -\theta \ln X \sim Exponential(1)$.

6.

	A	B	C	D	E	F	G	H	I	J	K	L	M
the p.f. $f(x)$ of a discrete r.v.	✓	✓		$*$	✓						✓		
the c.d.f. $F(x)$ of a discrete r.v.	✓	✓	✓	✓		✓	✓		✓				
the p.d.f. $f(x)$ of a continuous r.v.	✓			$*$		✓						✓	
the c.d.f. $F(x)$ of a continuous r.v.	✓	✓	✓	✓		✓	✓		✓	✓			

$*=$ do not use this box

A: The value of the function is always non-negative.

B: Every value of the function lies in the interval $[0,1]$.

C: The limit of the function as $x \to \infty$ equals 1.

D: The limit of the function as $x \to -\infty$ equals 0.

E: The domain of the function is countable.

F: The domain of the function is $\Re$.

G: The function is non-decreasing for all $x \in \Re$.

H: The function is increasing for all $x \in \Re$.

I: The function is right-continuous for all $x \in \Re$.

J: The function is continuous for all $x \in \Re$.

K: The sum of the function over all values of x equals 1.

L: The area bounded by the graph of the curve of the function and the $x-$axis equals 1.

M: The derivative of the function is equal to $P(X=x)$.

Sample Exam Solutions

Part A:

1. C
2. E
3. C
4. D
5. B
6. A
7. B
8. D
9. C
10. C
11. A
12. D
13. C
14. D
15. C
16. B

Part B:

1. (a) **Geometric**
   (b) **Continuous Uniform**
   (c) **Discrete Uniform**
   (d) **Negative Binomial**
   (e) **Poisson**
   (f) **Hypergeometric**
   (g) **Binomial**
   (h) **Normal**
   (i) **Exponential**
   (j) **Binomial**

   (a) **E**
   (b) **B**
   (c) **A** (also **D** and **E)**
   (d) **A**
   (e) **C**
   (f) **B**
   (g) **D**
   (h) **B**
   (i) **E**
   (j) **D** (also **E)**

Part C: Long Answer

1. X and Y are discrete random variables with joint probability function

$f(x,y)$ $P(X=x,Y=y)$		x 0	 1	 2	$P(Y=y)$
	1	0.15	0.05	0.15	0.35
y	0	0.15	0.05	0.20	0.40
	-1	0.10	0.00	0.15	0.25
	$P(X=x)$	0.40	0.10	0.50	1.00

(a) Since

$$P(X=1,Y=-1)=0\neq P(X=1)\,P(Y=-1)=(0.1)\,(0.25)$$

therefore X and Y are not independent random variables.

(b)

$$\begin{aligned}
E(X) &= (1)(0.1)+(2)(0.5)=1.1\\
E(Y) &= (1)(0.35)+(-1)(0.25)=0.1\\
E(XY) &= (1)(1)(0.05)+(2)(1)(0.15)+(2)(-1)(0.15)=0.05\\
Cov(X,Y) &= 0.05-(1.1)(0.1)\\
&= -0.06
\end{aligned}$$

(c)

$$\begin{aligned}
E\left(X^2\right) &= (1)^2(0.1)+(2)^2(0.5)=2.1\\
Var(X) &= 2.1-(1.1)^2=0.89\\
E\left(Y^2\right) &= (1)^2(0.35)+(-1)^2(0.25)=0.6\\
Var(Y) &= 0.6-(0.1)^2=0.59\\
\rho(X,Y) &= \frac{-0.06}{\sqrt{(0.89)(0.59)}}\\
&= -0.083
\end{aligned}$$

(d)

$$
\begin{aligned}
Var\left(2X-Y+1\right) &= Var\left(2X-Y\right)\\
&= (2)^2\,Var\left(X\right)+(-1)^2\,Var\left(Y\right)+2\,(2)\,(-1)\,Cov\left(X,Y\right)\\
&= (4)\,(0.89)+0.59+(-4)\,(-0.06)\\
&= 3.56+0.59+0.24\\
&= 4.39
\end{aligned}
$$

(e)

x	0	1	2	$Total$
$P\left(X=x\|Y=0\right)$	$\frac{0.15}{0.40}=0.375$	$\frac{0.05}{0.40}=0.125$	$\frac{0.20}{0.40}=0.5$	1.0

(f)

t	-1	0	1	2	3	$Total$
$P\left(T=t\right)$	0.10	0.15	0.35	0.25	0.15	1

2. The weights of full-term babies born in Ontario are Normally distributed with mean $\mu = 3.5$ kg and standard deviation $\sigma = 0.5$ kg.

   (a) Let $X =$ weight of randomly chosen full-term baby. Then $X \sim N\left(3.5, (0.5)^2\right)$.

$$\begin{aligned}
P(X > 4.25) &= P\left(\frac{X-3.5}{0.5} > \frac{4.25-3.5}{0.5}\right) \\
&= P(Z > 1.5) \quad \text{where } Z \sim N(0,1) \\
&= 1 - P(Z \le 1.5) \\
&= 1 - 0.93319 \\
&= 0.06681 \\
&= 0.067
\end{aligned}$$

   Therefore the proportion of full-term babies born in Ontario that weigh more than 4.25 kg is 0.067.

   (b)

$$\begin{aligned}
P(3.1 \le X \le 4.25) &= P\left(\frac{3.1-3.5}{0.5} \le \frac{X-3.5}{0.5} \le \frac{4.25-3.5}{0.5}\right) \\
&= P(-0.8 \le Z \le 1.5) \quad \text{where } Z \sim N(0,1) \\
&= P(Z \le 1.5) - P(Z \le -0.8) \\
&= P(Z \le 1.5) - [1 - P(Z \le 0.8)] \\
&= P(Z \le 1.5) + P(Z \le 0.8) - 1 = 0.93319 + 0.78814 - 1 \\
&= 0.72133 = 0.721
\end{aligned}$$

   Therefore the proportion of full-term babies born in Ontario that weigh between 3.1 and 4.25 kg is 0.721.

   (c)

$$\begin{aligned}
P(|X-3.5| \le 0.5) &= P\left(\left|\frac{X-3.5}{0.5}\right| \le \frac{0.5}{0.5}\right) \\
&= P(|Z| \le 1) \quad \text{where } Z \sim N(0,1) \\
&= 2P(Z \le 1) - 1 = 2(0.84134) - 1 \\
&= 0.68268 = 0.683
\end{aligned}$$

   Therefore the proportion of full-term babies born in Ontario that have weights within one standard deviation of the mean is 0.683.

(d)

$$\begin{aligned}
&P(\text{exactly } 1 \text{ baby weighs more than } 4.25\text{kg},\\
&\text{exactly } 5 \text{ babies weigh between } 3.1\text{kg and } 4.25\text{kg},\\
&\text{and exactly } 3 \text{ babies weigh less than } 3.1\text{kg})\\
&= \frac{9!}{1!5!3!}(0.067)^1(0.721)^5(0.212)^3
\end{aligned}$$

(e) Let $X_i =$ weight of i'th baby, $i = 1, 2, \ldots, 9$. Then $X_i \sim N\left(3.5, (0.5)^2\right)$, $i = 1, 2, \ldots, 9$ independently and $\bar{X} \sim N\left(3.5, \frac{(0.5)^2}{9}\right)$ or $\bar{X} \sim N\left(3.5, \left(\frac{0.5}{3}\right)^2\right)$.

$$\begin{aligned}
P\left(\bar{X} > 3.4\right) &= P\left(\frac{\bar{X} - 3.5}{0.5/3}\right)\\
&= P\left(Z > -0.6\right) \quad \text{where } Z \sim N(0,1)\\
&= P\left(Z < 0.6\right) = 0.72575 = 0.726
\end{aligned}$$

(f) Since $X_i \sim N\left(3.5, (0.5)^2\right)$, $i = 1, 2, \ldots, n$ independently, then $\bar{X} \sim N\left(3.5, \frac{(0.5)^2}{n}\right)$. We want

$$\begin{aligned}
0.9 \le P\left(\left|\bar{X} - 3.5\right| \le 0.05\right) &= P\left(\frac{\left|\bar{X} - 3.5\right|}{0.5/\sqrt{n}} \le \frac{0.05}{0.5/\sqrt{n}}\right)\\
&= P\left(|Z| \le 0.1\sqrt{n}\right) \quad \text{where } Z \sim N(0,1)\\
&= 2P\left(Z \le 0.1\sqrt{n}\right) - 1
\end{aligned}$$

or

$$0.95 \le P\left(Z \le 0.1\sqrt{n}\right)$$

Since

$$P\left(Z \le 1.6449\right) = 0.95$$

we need

$$0.1\sqrt{n} \ge 1.6449 \text{ or } n \ge (16.449)^2 = 270.570$$

so the smallest value of n is 271.

3. Suppose X is a continuous random variable with probability density function:

$$f(x) = \begin{cases} \frac{1}{\theta^2} x e^{-x/\theta} & x \geq 0 \\ \\ 0 & \text{otherwise} \end{cases}$$

(a)

$$\begin{aligned} E\left(X^k\right) &= \int_0^\infty x^k \frac{1}{\theta^2} x e^{-x/\theta} dx = \frac{1}{\theta^2} \int_0^\infty x^{k+1} e^{-x/\theta} dx \ \text{ let } y = \frac{x}{\theta} \\ &= \frac{1}{\theta^2} \int_0^\infty (\theta y)^{k+1} e^{-y} \theta dy = \theta^k \int_0^\infty y^{k+2-1} e^{-y} dy = \theta^k \Gamma(k+2) \\ &= \theta^k (k+1)! \end{aligned}$$

(b) Let $k = 1$ to obtain

$$E(X) = \theta^1 (1+1)! = 2\theta$$

Let $k = 2$ to obtain

$$E\left(X^2\right) = \theta^2 (2+1)! = 6\theta^2$$

Then

$$Var(X) = E\left(X^2\right) - [E(X)]^2 = 6\theta^2 - (2\theta)^2 = 2\theta^2$$

(c) For $y > 0$ the c.d.f. of Y is

$$\begin{aligned} G(y) = P(Y \leq y) &= P\left(\sqrt{X} \leq y\right) \\ &= P\left(X \leq y^2\right) \\ &= F\left(y^2\right) \end{aligned}$$

where $F(x) = P(X \leq x)$ is the cumulative distribution function of X.

For $y > 0$ the p.d.f. of Y is

$$\begin{aligned} g(y) = \frac{d}{dy} G(y) &= f\left(y^2\right) \frac{d}{dy}\left(y^2\right) \\ &= \frac{1}{\theta^2} y^2 e^{-y^2/\theta} (2y) \\ &= \frac{2}{\theta^2} y^3 e^{-y^2/\theta} \end{aligned}$$

and $g(y) = 0$ for $y \leq 0$.

(d) Since $E\left(X_i\right) = 2\theta$ and $Var\left(X_i\right) = 2\theta^2$ $i = 1, 2, \ldots, 98$ then by the Central Limit Theorem

$$\bar{X} \sim N\left(2\theta, \frac{2\theta^2}{98}\right) \quad \text{approximately}$$

$$\text{or } \bar{X} \sim N\left(2\theta, \left(\frac{\theta}{7}\right)^2\right) \quad \text{approximately}$$

$$\text{or } \frac{\bar{X} - 2\theta}{\theta/7} \sim N\left(0, 1\right) \quad \text{approximately}$$

Therefore

$$\begin{aligned} & P\left(\frac{\left|\bar{X} - 2\theta\right|}{\theta/7} < 1.15\right) \\ & \approx P\left(\left|Z\right| < 1.15\right) \quad \text{where } Z \sim N\left(0, 1\right) \\ & = 2P\left(Z < 1.15\right) - 1 \\ & = 2\left(0.87493\right) - 1 \\ & = 0.74986 = 0.750 \end{aligned}$$

4. Ten friends go to an all-you-can-eat sushi restaurant and sit at one large round table. Each person likes spicy food with probability 0.6, independently of each other. We say a "match" occurs when two people sitting next to each other BOTH like spicy food or BOTH do not like spicy food. Let

$$X_i = \begin{cases} 1 & \text{if there is a "match" between person } i \text{ and person } i+1 \\ 0 & \text{otherwise} \end{cases}$$

for $i = 1, 2, ..., 10$ where person 11 is defined to be person 1 since they are at a circular table.

(a) Let F_i be the event a person likes spicy food and $\bar{F}_i$ be the event a person does not like spicy food.

$$\begin{aligned} E(X_i) &= P(X_i = 1) = P(F_i \cap F_{i+1}) + P\left(\bar{F}_i \cap \bar{F}_{i+1}\right) \\ &= (0.6)(0.6) + (0.4)(0.4) = 0.52 \end{aligned}$$

(b) Let $T = X_1 + X_2 + \cdots + X_{10}$ = total number of matches. Then

$$\begin{aligned} E(T) &= E(X_1) + E(X_2) + \cdots + E(X_{10}) \\ &= 10(0.52) = 5.2 \end{aligned}$$

(c)

$$\begin{aligned} Var(X_i) &= P(X_i = 1)[1 - P(X_i = 1)] \\ &= (0.52)(0.48) = 0.2496 = 0.250 \end{aligned}$$

(d)

$$\begin{aligned} E(X_1 X_2) &= P(X_1 = 1, X_2 = 1) \\ &= P(F_1 \cap F_2 \cap F_3) + P\left(\bar{F}_1 \cap \bar{F}_2 \cap \bar{F}_3\right) \\ &= (0.6)^3 + (0.4)^3 = 0.28 \end{aligned}$$

Therefore

$$\begin{aligned} Cov(X_1, X_2) &= E(X_1 X_2) - E(X_1) E(X_2) \\ &= 0.28 - (0.52)^2 = 0.0096 \end{aligned}$$

(e) Since $Cov(X_1, X_2) = Cov(X_2, X_3) = \cdots = Cov(X_9, X_{10}) = Cov(X_{10}, X_1) = 0.0096$ and all other covariances are equal to zero therefore

$$Var(T) = 10(0.2496) + 2(10)(0.0096) = 2.688$$

13. DISTRIBUTIONS AND $N(0,1)$ TABLES

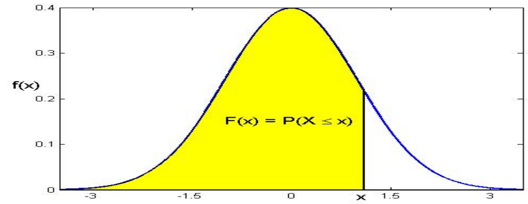
Summary of Discrete Distributions

Notation and Parameters	Probability Function $f(x)$	Mean $E(X)$	Variance $Var(X)$	Moment Generating Function $M(t)$
Discrete Uniform(a,b) $b \geq a$ a,b integers	$\frac{1}{b-a+1}$ $x = a, a+1, \ldots, b$	$\frac{a+b}{2}$	$\frac{(b-a+1)^2-1}{12}$	$\frac{1}{b-a+1}\sum_{x=a}^{b} e^{tx}$ $t \in \Re$
Hypergeometric(N,r,n) $N = 1,2,\ldots$ $n = 0,1,\ldots,N$ $r = 0,1,\ldots,N$	$\frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$ $x = \max(0, n-N+r),$ $\ldots, \min(r,n)$	$\frac{nr}{N}$	$\frac{nr}{N}(1-\frac{r}{N})\frac{N-n}{N-1}$	Not tractable
Binomial(n,p) $0 \leq p \leq 1,\ q = 1-p$ $n = 1,2,\ldots$	$\binom{n}{x}p^x q^{n-x}$ $x = 0,1,\ldots,n$	np	npq	$(pe^t+q)^n$ $t \in \Re$
Bernoulli(p) $0 \leq p \leq 1,\ q = 1-p$	$p^x q^{1-x}$ $x = 0,1$	p	pq	pe^t+q $t \in \Re$
Negative Binomial(k,p) $0 < p \leq 1,\ q = 1-p$ $k = 1,2,\ldots$	$\binom{x+k-1}{x}p^k q^x$ $= \binom{-k}{x}p^k(-q)^x$ $x = 0,1,\ldots$	$\frac{kq}{p}$	$\frac{kq}{p^2}$	$\left(\frac{p}{1-qe^t}\right)^k$ $t < -\ln q$
Geometric(p) $0 < p \leq 1,\ q = 1-p$	pq^x $x = 0,1,\ldots$	$\frac{q}{p}$	$\frac{q}{p^2}$	$\frac{p}{1-qe^t}$ $t < -\ln q$
Poisson(λ) $\lambda \geq 0$	$\frac{e^{-\lambda}\lambda^x}{x!}$ $x = 0,1,\ldots$	λ	λ	$e^{\lambda(e^t-1)}$ $t \in \Re$
Multinomial$(n;p_1,p_2,\ldots,p_k)$ $0 \leq p_i \leq 1$ $i = 1,2,\ldots,k$ and $\sum_{i=1}^{k} p_i = 1$	$f(x_1,x_2,\ldots,x_k) =$ $\frac{n!}{x_1!x_2!\cdots x_k!}p_1^{x_1}p_2^{x_2}\cdots p_k^{x_k}$ $x_i = 0,1,\ldots,n$ $i = 1,2,\ldots,k$ and $\sum_{i=1}^{k} x_i = n$	$E(X_i) = np_i$ $i = 1,2,\ldots,k$	$Var(X_i)$ $= np_i(1-p_i)$ $i = 1,2,\ldots,k$	$M(t_1,t_2,\ldots,t_k)$ $=(p_1e^{t_1}+p_2e^{t_2}+\cdots$ $+p_{k-1}e^{t_{k-1}}+p_k)^n$ $t_i \in \Re$ $i = 1,2,\ldots,k-1$

Summary of Continuous Distributions

Notation and Parameters	**Probability Density Function** $f(x)$	**Mean** $E(X)$	**Variance** $Var(X)$	**Moment Generating Function** $M(t)$
Uniform(a,b) $b > a$	$\frac{1}{b-a}$ $a \le x \le b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{bt}-e^{at}}{(b-a)t}$ $\quad t \neq 0$ $1 \quad t = 0$
Exponential(θ) $\theta > 0$	$\frac{1}{\theta} e^{-x/\theta}$ $x \ge 0$	θ	θ^2	$\frac{1}{1-\theta t}$ $t < \frac{1}{\theta}$
$N(\mu, \sigma^2) = G(\mu, \sigma)$ $\mu \in \Re,\ \sigma^2 > 0$	$\frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/(2\sigma^2)}$ $x \in \Re$	μ	σ^2	$e^{\mu t + \sigma^2 t^2/2}$ $t \in \Re$

N(0,1) Cumulative Distribution Function



This table gives values of $F(x) = P(X \le x)$ for $X \sim N(0,1)$ and $x \ge 0$

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.98300	0.98341	0.98382	0.98422	0.98461	0.98500	0.98537	0.98574
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.98840	0.98870	0.98899
2.3	0.98928	0.98956	0.98983	0.99010	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	0.99461	0.99477	0.99492	0.99506	0.99520
2.6	0.99534	0.99547	0.99560	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.99720	0.99728	0.99736
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.99900
3.1	0.99903	0.99906	0.99910	0.99913	0.99916	0.99918	0.99921	0.99924	0.99926	0.99929
3.2	0.99931	0.99934	0.99936	0.99938	0.99940	0.99942	0.99944	0.99946	0.99948	0.99950
3.3	0.99952	0.99953	0.99955	0.99957	0.99958	0.99960	0.99961	0.99962	0.99964	0.99965
3.4	0.99966	0.99968	0.99969	0.99970	0.99971	0.99972	0.99973	0.99974	0.99975	0.99976
3.5	0.99977	0.99978	0.99978	0.99979	0.99980	0.99981	0.99981	0.99982	0.99983	0.99983

N(0,1) Quantiles: This table gives values of $F^{-1}(p)$ for $p \ge 0.5$

p	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.075	0.08	0.09	0.095
0.5	0.0000	0.0251	0.0502	0.0753	0.1004	0.1257	0.1510	0.1764	0.1891	0.2019	0.2275	0.2404
0.6	0.2533	0.2793	0.3055	0.3319	0.3585	0.3853	0.4125	0.4399	0.4538	0.4677	0.4959	0.5101
0.7	0.5244	0.5534	0.5828	0.6128	0.6433	0.6745	0.7063	0.7388	0.7554	0.7722	0.8064	0.8239
0.8	0.8416	0.8779	0.9154	0.9542	0.9945	1.0364	1.0803	1.1264	1.1503	1.1750	1.2265	1.2536
0.9	1.2816	1.3408	1.4051	1.4758	1.5548	1.6449	1.7507	1.8808	1.9600	2.0537	2.3263	2.5758