PROBABILITY AND STATISTICS FOR COMPUTER SCIENTISTS

Instructor's Solution Manual

Michael Baron

Department of Mathematical Sciences
University of Texas at Dallas, Richardson, Texas 75083-0688

Table of Contents

Chapter 2

2.1 An outcome is the chosen pair of chips. The sample space in this problem consists of 15 pairs: AB, AC, AD, AE, AF, BC, BD, BE, BF, CD, CE, CF, DE, DF, EF (or 30 pairs if the order of chips in each pair matters, i.e., AB and BA are different pairs).

All the outcomes are equally likely because two chips are chosen at random.

One outcome is 'favorable', when both chips in a pair are defective (two such pairs if the order matters).

Thus,

$$P(\text{both chips are defective}) = \frac{\text{number of favorable outcomes}}{\text{total number of outcomes}} = \boxed{1/15}$$

2.2 Denote the events:

$$\begin{aligned} M &= \{ \text{ problems with a motherboard } \} \\ H &= \{ \text{ problems with a hard drive } \} \end{aligned}$$

We have:

$$\boldsymbol{P}\{M\} = 0.4, \ \boldsymbol{P}\{H\} = 0.3, \quad \text{and} \quad \boldsymbol{P}\{M \cap H\} = 0.15.$$

Hence,

$$\boldsymbol{P}\{M \cup H\} = \boldsymbol{P}\{M\} + \boldsymbol{P}\{H\} - \boldsymbol{P}\{M \cap H\} = 0.4 + 0.3 - 0.15 = 0.55,$$

and

$$\boldsymbol{P}\{\text{fully functioning MB and HD}\} = 1 - \boldsymbol{P}\{M \cup H\} = \boxed{0.45}$$

2.3 Denote the events,

$$\begin{aligned} I &= \{\text{the virus enters through the internet}\} \\ E &= \{\text{the virus enters through the e-mail}\} \end{aligned}$$

Then

$$\begin{aligned} P\{\bar{E} \cap \bar{I}\} &= 1 - P\{E \cup I\} = 1 - (P\{E\} + P\{I\} - P\{E \cap I\}) \\ &= 1 - (.3 + .4 - .15) = \boxed{\ 0.45\ } \end{aligned}$$

It may help to draw a Venn diagram.

2.4 Denote the events,

$$C = \{ \text{ knows C/C++ } \}, \quad F = \{ \text{ knows Fortran } \}.$$

Then

(a) $\boldsymbol{P}\{\bar{F}\} = 1 - \boldsymbol{P}\{F\} = 1 - 0.6 = \boxed{0.4}$

(b) $$\begin{aligned} \boldsymbol{P}\left\{\bar{F} \cap \bar{C}\right\} &= 1 - \boldsymbol{P}\left\{F \cup C\right\} = 1 - (\boldsymbol{P}\left\{F\right\} + \boldsymbol{P}\left\{C\right\} - \boldsymbol{P}\left\{F \cap C\right\}) \\ &= 1 - (0.7 + 0.6 - 0.5) = 1 - 0.8 = \boxed{0.2} \end{aligned}$$

(c) $\boldsymbol{P}\left\{C \backslash F\right\} = \boldsymbol{P}\left\{C\right\} - \boldsymbol{P}\left\{F \cap C\right\} = 0.7 - 0.5 = \boxed{0.2}$

(d) $\boldsymbol{P}\left\{F \backslash C\right\} = \boldsymbol{P}\left\{F\right\} - \boldsymbol{P}\left\{F \cap C\right\} = 0.6 - 0.5 = \boxed{0.1}$

(e) $\boldsymbol{P}\left\{C \mid F\right\} = \dfrac{\boldsymbol{P}\left\{C \cap F\right\}}{\boldsymbol{P}\left\{F\right\}} = \dfrac{0.5}{0.6} = \boxed{0.8333}$

(f) $\boldsymbol{P}\left\{F \mid C\right\} = \dfrac{\boldsymbol{P}\left\{C \cap F\right\}}{\boldsymbol{P}\left\{C\right\}} = \dfrac{0.5}{0.7} = \boxed{0.7143}$

2.5 Denote the events:

$$\begin{aligned} D_1 &= \{\text{first test discovers the error}\} \\ D_2 &= \{\text{second test discovers the error}\} \\ D_3 &= \{\text{third test discovers the error}\} \end{aligned}$$

Then

$$\begin{aligned} &\boldsymbol{P}\left\{\text{ at least one discovers }\right\} = \boldsymbol{P}\left\{D_1 \cup D_2 \cup D_3\right\} \\ &= 1 - \boldsymbol{P}\left\{\bar{D_1} \cap \bar{D_2} \cap \bar{D_3}\right\} \\ &= 1 - (1 - 0.2)(1 - 0.3)(1 - 0.5) = 1 - 0.28 = \boxed{0.72} \end{aligned}$$

We used the complement rule and independence.

2.6 Let $A = \{\text{arrive on time}\}$, $W = \{\text{good weather}\}$. We have

$$\boldsymbol{P}\left\{A \mid W\right\} = 0.8,\ \boldsymbol{P}\left\{A \mid \bar{W}\right\} = 0.3,\ \boldsymbol{P}\left\{W\right\} = 0.6$$

By the Law of Total Probability,

$$\begin{aligned} \boldsymbol{P}\left\{A\right\} &= \boldsymbol{P}\left\{A \mid W\right\}\boldsymbol{P}\left\{W\right\} + \boldsymbol{P}\left\{A \mid \bar{W}\right\}\boldsymbol{P}\left\{\bar{W}\right\} \\ &= (0.8)(0.6) + (0.3)(0.4) = \boxed{0.60} \end{aligned}$$

2.7 Organize the data. Let $D = \{\text{detected}\}$, $I = \{\text{via internet}\}$, $E = \{\text{via e-mail}\} = \overline{I}$. Notice that the question about detection already assumes that the spyware *has entered* the system. This is the sample space, and this is why $\boldsymbol{P}\left\{I\right\} + \boldsymbol{P}\left\{E\right\} = 1$. We have

$$\boldsymbol{P}\left\{I\right\} = 0.7,\quad \boldsymbol{P}\left\{E\right\} = 0.3,\quad \boldsymbol{P}\left\{D \mid I\right\} = 0.6,\quad \boldsymbol{P}\left\{D \mid E\right\} = 0.8.$$

By the Law of Total Probability,

$$\boldsymbol{P}\left\{D\right\} = (0.6)(0.7) + (0.8)(0.3) = \boxed{0.66}$$

2.8 Let $A_1 = \{$1st device fails$\}$, $A_2 = \{$2nd device fails$\}$, $A_3 = \{$3rd device fails$\}$.

$$\begin{aligned}
\boldsymbol{P}\{\text{ on time }\} &= \boldsymbol{P}\{\text{ all function }\}\\
&= \boldsymbol{P}\left\{\overline{A}_1 \cap \overline{A}_2 \cap \overline{A}_3\right\}\\
&= \boldsymbol{P}\left\{\overline{A}_1\right\}\boldsymbol{P}\left\{\overline{A}_2\right\}\boldsymbol{P}\left\{\overline{A}_3\right\} && \text{(independence)}\\
&= (1-0.01)(1-0.02)(1-0.02) && \text{(complement rule)}\\
&= \boxed{0.9508}
\end{aligned}$$

2.9 $\boldsymbol{P}\{\text{at least one fails}\} = 1 - \boldsymbol{P}\{\text{all work}\} = 1 - (.96)(.95)(.90) = \boxed{0.1792}$.

2.10
$$\begin{aligned}
\boldsymbol{P}\{A \cup B \cup C\} &= 1 - \boldsymbol{P}\left\{\bar{A} \cap \bar{B} \cap \bar{C}\right\} = 1 - \boldsymbol{P}\left\{\bar{A}\right\}\boldsymbol{P}\left\{\bar{B}\right\}\boldsymbol{P}\left\{\bar{C}\right\}\\
&= 1 - (1-0.4)(1-0.5)(1-0.2) = \boxed{0.76}
\end{aligned}$$

2.11 (a) $\boldsymbol{P}\{\text{at least one test finds the error}\}$
$$\begin{aligned}
&= 1 - \boldsymbol{P}\{\text{all tests fail to find the error}\}\\
&= 1 - (1-0.1)(1-0.2)(1-0.3)(1-0.4)(1-0.5)\\
&= 1 - (0.9)(0.8)(0.7)(0.6)(0.5) = \boxed{0.8488}
\end{aligned}$$

(b) The difference between events in (a) and (b) is the probability that *exactly one* test finds an error. This probability equals

$\boldsymbol{P}\{\text{exactly one test finds the error}\}$
$$\begin{aligned}
&= \boldsymbol{P}\{\text{test 1 find the error, the others don't find}\}\\
&\quad + \boldsymbol{P}\{\text{test 2 find the error, the others don't find}\} + \ldots\\
&= (0.1)(1-0.2)(1-0.3)(1-0.4)(1-0.5)\\
&\quad + (1-0.1)(0.2)(1-0.3)(1-0.4)(1-0.5) + \ldots\\
&= (0.1)(0.8)(0.7)(0.6)(0.5) + (0.9)(0.2)(0.7)(0.6)(0.5)\\
&\quad + (0.9)(0.8)(0.3)(0.6)(0.5) + (0.9)(0.8)(0.7)(0.4)(0.5)\\
&\quad + (0.9)(0.8)(0.7)(0.6)(0.5) = 0.3714.
\end{aligned}$$

Then

$\boldsymbol{P}\{\text{at least two tests find the error}\}$
$$\begin{aligned}
&= \boldsymbol{P}\{\text{at least one test finds the error}\}\\
&\quad - \boldsymbol{P}\{\text{exactly one test finds the error}\}\\
&= 0.8488 - 0.3714 = \boxed{0.4774}
\end{aligned}$$

(c) $\boldsymbol{P}\{\text{all tests find the error}\} = (0.1)(0.2)(0.3)(0.4)(0.5) = \boxed{0.0012}$

2.12 Let $A_j = \{$ dog j detects the explosives $\}$.

$\boldsymbol{P}\{\text{at least one dog detects}\} = 1 - \boldsymbol{P}\{\text{all four dogs don't detect}\}$

$= 1 - \boldsymbol{P}\left\{\bar{A}_1\right\}\boldsymbol{P}\left\{\bar{A}_2\right\}\boldsymbol{P}\left\{\bar{A}_3\right\}\boldsymbol{P}\left\{\bar{A}_4\right\}$

$= 1 - (1-0.6)^4 = \boxed{0.9744}$

2.13 Let A_j be the event {Team j detects a problem}. Then
$\boldsymbol{P}\{\text{at least one team detects}\} = 1 - \boldsymbol{P}\{\text{no team detects}\}$
$= 1 - P\{\bar{A}_1 \cap \bar{A}_2 \cap \bar{A}_3\} = 1 - P\{\bar{A}_1\} P\{\bar{A}_2\} P\{\bar{A}_3\}$
$= 1 - (1 - 0.8)(1 - 0.8)(1 - 0.8) = \boxed{0.992}$.

2.14 (a) The total number of possible passwords is

$$P(26,6) = (26)(25)(24)(23)(22)(21) = 165,765,600$$

because there are 26 letters in the alphabet, they should be all different in the password, and the order of characters is important. The password is guessed (favorable outcome) if it is among the 1,000,000 attempted passwords. Then

$$\begin{aligned} \boldsymbol{P}\{\text{guess the password}\} &= \frac{\text{number of favorable passwords}}{\text{total number of passwords}} \\ &= \frac{1,000,000}{165,765,600} = \boxed{0.0060} \end{aligned}$$

(b) Now we can use 52 characters, and the order is still important. Then the total number of passwords is

$$P(52,6) = (52)(51)(50)(49)(48)(47) = 14,658,134,400,$$

and

$$\boldsymbol{P}\{\text{guess the password}\} = \frac{1,000,000}{14,658,134,400} = \boxed{0.000068}$$

(c) Letters can be repeated in passwords, therefore, the total number of passwords is

$$P_r(52,6) = 52^6,$$

and

$$\boldsymbol{P}\{\text{guess the password}\} = \frac{10^6}{52^6} = \boxed{0.000051}$$

(d) Adding the digits brings the number of possible characters to 62. Then the total number of passwords is

$$P_r(62,6) = 62^6,$$

and

$$\boldsymbol{P}\{\text{guess the password}\} = \frac{10^6}{62^6} = \boxed{0.000018}$$

The more characters we use the lower is the probability for a spyware to break into the system.

2.15 Let $A = \{\text{Error in the 1st block}\}$ and $B = \{\text{Error in the 2nd block}\}$. Then $P\{A\} = 0.2$, $P\{B\} = 0.3$, and $P\{A \cap B\} = 0.06$ by independence;
$P\{\text{ error in program }\} = P\{A \cup B\} = 0.2 + 0.3 - 0.06 = 0.44.$

Then, by the definition of conditional probability,

$$P\{A \cap B \mid A \cup B\} = \frac{P\{A \cap B\}}{P\{A \cup B\}} = \frac{0.06}{0.44} = \boxed{0.1364}$$

Or, by the Bayes Rule,

$$\begin{aligned} P\{A \cap B \mid A \cup B\} &= \frac{P\{A \cup B \mid A \cap B\} P\{A \cap B\}}{P\{A \cup B\}} \\ &= \frac{(1)(0.06)}{0.44} = \boxed{0.1364} \end{aligned}$$

2.16 Organize the data. Let $D = \{\text{defective part}\}$. We are given:

$$\begin{array}{r|r} P\{S1\} = 0.5 & P\{D|S1\} = 0.05 \\ P\{S2\} = 0.2 & P\{D|S2\} = 0.03 \\ P\{S3\} = 0.3 & P\{D|S3\} = 0.06 \end{array}$$

We need to find $P\{S1|D\}$.

(a) By the Law of Total Probability:

$$\begin{aligned} \boldsymbol{P}\{D\} &= P\{D|S1\} P\{S1\} + P\{D|S2\} P\{S2\} + P\{D|S3\} P\{S3\} \\ &= (0.5)(0.05) + (0.2)(0.03) + (0.3)(0.06) = \boxed{0.049} \end{aligned}$$

(b) Bayes Rule:

$$P\{S1|D\} = \frac{P\{D|S1\} P\{S1\}}{P\{D\}} = \frac{(0.5)(0.05)}{0.49} = \boxed{25/49 \text{ or } 0.5102}$$

2.17 Let $D = \{\text{defective part}\}$. We are given:

$$\begin{array}{r|r} P\{X\} = 0.24 & P\{D|X\} = 0.05 \\ P\{Y\} = 0.36 & P\{D|Y\} = 0.10 \\ P\{Z\} = 0.40 & P\{D|Z\} = 0.06 \end{array}$$

Combine the Bayes Rule and the Law of Total Probability.

$$\begin{aligned} \boldsymbol{P}\{Z \mid D\} &= \frac{\boldsymbol{P}\{D|Z\} P\{Z\}}{P\{D|X\} P\{X\} + P\{D|Y\} P\{Y\} + P\{D|Z\} P\{Z\}} \\ &= \frac{(0.06)(0.40)}{(0.05)(0.24) + (0.10)(0.36) + (0.06)(0.40)} \\ &= \boxed{1/3 \text{ or } 0.3333} \end{aligned}$$

2.18 Let $C = \{\text{ correct }\}$, $G = \{\text{ guessing }\}$. It is given that:

$$\boldsymbol{P}\{\bar{G}\} = 0.75, \quad \boldsymbol{P}\{C \mid \bar{G}\} = 0.9, \quad \boldsymbol{P}\{C \mid G\} = 1/4 = 0.25.$$

Also, $\boldsymbol{P}\{G\} = 1 - 0.75 = 0.25$.

Then, by the Bayes Rule,

$$\begin{aligned}\boldsymbol{P}\{G \mid C\} &= \frac{\boldsymbol{P}\{C \mid G\} P\{G\}}{\boldsymbol{P}\{C \mid G\} P\{G\} + \boldsymbol{P}\{C \mid \bar{G}\} P\{\bar{G}\}} \\ &= \frac{(0.25)(0.25)}{(0.25)(0.25) + (0.9)(0.75)} = \boxed{0.0847}\end{aligned}$$

2.19 Let $D = \{\text{defective part}\}$ and $I = \{\text{inspected electronically}\}$. By the Bayes Rule,

$$\begin{aligned}\boldsymbol{P}\{I|D\} &= \frac{\boldsymbol{P}\{D|I\} \boldsymbol{P}\{I\}}{\boldsymbol{P}\{D|I\} \boldsymbol{P}\{I\} + \boldsymbol{P}\{D|\bar{I}\} \boldsymbol{P}\{\bar{I}\}} \\ &= \frac{(1-0.95)(0.20)}{(1-0.95)(0.20) + (1-0.7)(1-0.20)} = \boxed{0.0400}\end{aligned}$$

2.20 Let $S = \{\text{steroid user}\}$ and $N = \{\text{test is negative}\}$.

It is given that $\boldsymbol{P}\{S\} = 0.05$, $\boldsymbol{P}\{\bar{N}|S\} = 0.9$, and $\boldsymbol{P}\{\bar{N}|\bar{S}\} = 0.02$.

By the complement rule, $\boldsymbol{P}\{\bar{S}\} = 0.95$, $\boldsymbol{P}\{N|S\} = 0.1$, and $\boldsymbol{P}\{N|\bar{S}\} = 0.98$.

By the Bayes Rule,

$$\begin{aligned}\boldsymbol{P}\{S|N\} &= \frac{\boldsymbol{P}\{N|S\} \boldsymbol{P}\{S\}}{\boldsymbol{P}\{N|S\} \boldsymbol{P}\{S\} + \boldsymbol{P}\{N|\bar{S}\} \boldsymbol{P}\{\bar{S}\}} \\ &= \frac{(0.1)(0.05)}{(0.1)(0.05) + (0.98)(0.95)} = \boxed{5/936 \text{ or } 0.00534}\end{aligned}$$

2.21 At least one of the first three components works with probability

$$1 - \boldsymbol{P}\{\text{all three fail}\} = 1 - (0.3)^3 = 0.973.$$

At least one of the last two components works with probability

$$1 - \boldsymbol{P}\{\text{both fail}\} = 1 - (0.3)^2 = 0.91.$$

Hence, the system operates with probability $(0.973)(0.91) = \boxed{0.8854}$

2.22 (a) The scheme of cities A, B, and C and all five highways is similar to Exercise 2.21. Similarly to this exercise, there exists an open route from city A to city C with probability

$$\left\{1 - (0.2)^3\right\}\left\{1 - (0.2)^2\right\} = \boxed{0.9523}$$

(b-α) If the new highway is built between cities A and B, it will be the 4-th highway connecting A and B. Then the probability of an open route from city A to city C becomes

$$\left\{1 - (0.2)^4\right\}\left\{1 - (0.2)^2\right\} = \boxed{0.9585}$$

(b-β) If the new highway is built between B and C, it will be the 3rd highway connecting these cities. Then the probability of an open route from city A to city C is

$$\left\{1-(0.2)^3\right\}\left\{1-(0.2)^3\right\} = \boxed{0.9841}$$

(b-γ) Finally, if the new highway is built between A and C, then
$\boldsymbol{P}\{$ at least one open route from A to C $\}$

$$= \left\{ \begin{array}{c} \text{a new direct route} \\ \text{from A to C is open} \end{array} \bigcup \begin{array}{c} \text{a route from A to B to C} \\ \text{is open, see question (a)} \end{array} \right\}$$
$$= \quad 1-(1-0.2)(1-0.9523) = \boxed{0.9618}$$

2.23 (a) $(0.9)(0.8) = \boxed{0.72}$

(b) $1 - \{1-(0.9)(0.8)\}\{1-(0.7)(0.6)\} = \boxed{0.8376}$

(c) $1-(1-0.9)(1-0.8)(1-0.7) = \boxed{0.994}$

(d) $\{1-(1-0.9)(1-0.7)\}\{1-(1-0.8)(1-0.6)\} = \boxed{0.8924}$

(e) $\{1-(1-0.9)(1-0.6)\}\{1-(1-0.8)(1-0.7)(1-0.5)\} = \boxed{0.9312}$

2.24 A customer is unaware of defects, so he buys 6 random laptops. The outcomes are equally likely, so each probability can be computed as

$$\frac{\text{number of favorable outcomes}}{\text{total number of outcomes}}$$

(a) $$\boldsymbol{P}\{\text{exactly } 2\} = \frac{\binom{5}{2}\binom{5}{4}}{\binom{10}{6}} = \frac{\frac{5\cdot 4}{2}\cdot 5}{\frac{10\cdot 9\cdot 8\cdot 7}{4\cdot 3\cdot 2\cdot 1}} = \boxed{\frac{5}{21} \text{ or } 0.238}$$

(b) This is a conditional probability because $\{X \geq 2\}$ is given. We need

$$\boldsymbol{P}\{X=2 \mid X \geq 2\} = \frac{\boldsymbol{P}\{X=2 \cap X \geq 2\}}{\boldsymbol{P}\{X \geq 2\}} = \frac{\boldsymbol{P}\{X=2\}}{\boldsymbol{P}\{X \geq 2\}}$$

where $\boldsymbol{P}\{X=2\} = 5/21$ is already computed in (a), and

$$\boldsymbol{P}\{x \geq 2\} = 1 - P(X=1) = 1 - \frac{\binom{5}{1}\binom{5}{5}}{\binom{10}{6}} = 1 - \frac{5\cdot 1}{\frac{10\cdot 9\cdot 8\cdot 7}{4\cdot 3\cdot 2\cdot 1}} = \frac{41}{42}$$

Notice that $\boldsymbol{P}\{X=0\} = 0$ because there are only 5 good computers, so among the purchased 6 computers there has to be at least 1 defective. So,

$$\boldsymbol{P}\{X=2 \mid X \geq 2\} = \frac{\boldsymbol{P}\{X=2\}}{\boldsymbol{P}\{X \geq 2\}} = \frac{5/21}{41/42} = \boxed{10/41 \text{ or } 0.244}.$$

2.25 Our sample space consists of birthdays of all $N = 30$ students. The total number of outcomes in it is

$$\mathcal{N}_T = P_r(365, N) = 365^N.$$

It is easier to count the outcomes where all students are born on *different* days. The number of such outcomes is

$$\mathcal{N}_F = P(365, N) = \frac{365!}{(365-N)!} = (365)(364)\dots(366-N).$$

Then

$$\begin{aligned} P(N) &= \boldsymbol{P}\{\text{at least two share birthdays}\} \\ &= 1 - \boldsymbol{P}\{\text{all born on different days}\} \\ &= 1 - \frac{\mathcal{N}_F}{\mathcal{N}_T} = 1 - \left(\frac{365}{365}\right)\left(\frac{364}{365}\right)\cdots\left(\frac{366-N}{365}\right). \end{aligned}$$

For $N = 30$, we get

$$P(30) = 1 - 0.2937 = \boxed{0.7063}$$

(b) Evaluating $P(N)$ for different N, we see that $P(22) = 0.4757$ and $P(23) = 0.5073$. Hence, we need at least $\boxed{23}$ students in order to find birthday matches with a probability above 0.5.

2.26 The sample space consists of all (unordered) sets of three computers selected from six computers in the lab. Favorable outcomes are sets of three computers with non-defective hard drives. We have

$$\mathcal{N}_T = C(6,3) = \frac{(6)(5)(4)}{(3)(2)(1)} = 20;$$

$$\mathcal{N}_F = C(4,3) = 4;$$

therefore,

$$\boldsymbol{P}\{\text{no hard drive problems}\} = \frac{\mathcal{N}_F}{\mathcal{N}_T} = \frac{4}{20} = \boxed{0.2}$$

2.27 The sample space consists of all unordered sets of five computers selected from 18 computers in the store. Favorable outcomes are sets of five non-defective computers (that come from a subset of $18 - 6 = 12$. Then

$$\mathcal{N}_T = C(18,5) = \frac{(18)(17)(16)(15)(14)}{(5)(4)(3)(2)(1)};$$

$$\mathcal{N}_F = C(12,5) = \frac{(12)(11)(10)(9)(8)}{(5)(4)(3)(2)(1)};$$

therefore,

$$\boldsymbol{P}\left\{\begin{array}{c}\text{five computers}\\ \text{without defects}\end{array}\right\} = \frac{\mathcal{N}_F}{\mathcal{N}_T} = \frac{(12)(11)(10)(9)(8)}{(18)(17)(16)(15)(14)} = \boxed{\frac{11}{119} \quad \text{or} \quad 0.0924}$$

2.28 The sample space consists of sequences of 6 answers where each answer is one of 4 possible answers, say, A, B, C, or D. Then a sequence of 6 answers is a 6-letter word written with letters A, B, C, and D with replacement. The student guesses, therefore, all outcomes are equally likely.

The total number of outcomes is

$$\mathcal{N}_T = P_r(4,6) = 4^6 = 4096.$$

Favorable outcomes occur when the student guesses at least 3 answers correctly. This includes 3, 4, 5, and 6 correct answers. The correctly answered questions are chosen at random from 6 questions. Then, a correct answer is given to each of the chosen questions. Also, an incorrect answer to each remaining question is chosen out of 3 possible incorrect answers. Altogether, the number of favorable outcomes is

$$\begin{aligned}\mathcal{N}_F &= C(6,3)(3^3) + C(6,4)(3^2) + C(6,5)(3^1) + C(6,6)(3^0) \\ &= \frac{(6)(5)(4)}{(3)(2)(1)}(27) + \frac{(6)(5)}{(2)(1)}(9) + (6)(3) + 1 = 694.\end{aligned}$$

$$\boldsymbol{P}\{\text{he will pass}\} = \frac{\mathcal{N}_F}{\mathcal{N}_T} = \frac{694}{4096} = \boxed{0.1694}$$

One can also use the complement rule for a little shorter solution.

2.29 Outcomes are sets of four databases selected from nine. Favorable outcomes are such sets where at least 2 databases have a keyword, out of 5 such databases (and the remaining ones don't have a keyword, so they come from the remaining 4 databases). Then

$$\mathcal{N}_T = C(9,4) = \frac{(9)(8)(7)(6)}{(4)(3)(2)(1)} = 126,$$

$$\begin{aligned}\mathcal{N}_F &= C(5,2)C(4,2) + C(5,3)C(4,1) + C(5,4)C(4,0) \\ &= (10)(6) + (10)(4) + (5)(1) = 105,\end{aligned}$$

and

$$\boldsymbol{P}\{\text{at least two have the keyword}\} = \frac{\mathcal{N}_F}{\mathcal{N}_T} = \frac{105}{126} = \boxed{\frac{5}{6} \text{ or } 0.8333}$$

2.30 (a) All outcomes are listed in the table below. According to the problem, they are equally likely.

Outcome	The older child	The younger child	Who is met
1	girl	girl	the older girl
2	girl	girl	the younger girl
3	girl	boy	the girl
4	girl	boy	the boy
5	boy	girl	the girl
6	boy	girl	the boy
7	boy	boy	the older boy
8	boy	boy	the younger boy

(b) $\boldsymbol{P}\{BB\} = \boldsymbol{P}\{\text{outcomes } 7,\, 8\} = 1/4,$
$\boldsymbol{P}\{BG\} = \boldsymbol{P}\{\text{outcomes } 5,\, 6\} = 1/4,$
$\boldsymbol{P}\{GB\} = \boldsymbol{P}\{\text{outcomes } 3,\, 4\} = 1/4.$

(c) Meeting Jimmy automatically eliminates outcomes 1, 2, 3, and 5. The remaining outcomes are

Outcome	The older child	The younger child	Who is met
4	girl	boy	the boy
6	boy	girl	the boy
7	boy	boy	the older boy
8	boy	boy	the younger boy

Two remaining outcomes form the event BB whereas BG and GB have only one outcome each. Therefore, given that you met a boy,
$\boldsymbol{P}\{BB \mid \text{met Jimmy}\} = \boldsymbol{P}\{\text{outcomes } 7,\, 8 \mid \text{met Jimmy}\} = 1/2,$
$\boldsymbol{P}\{BG \mid \text{met Jimmy}\} = \boldsymbol{P}\{\text{outcome } 6 \mid \text{met Jimmy}\} = 1/4,$
$\boldsymbol{P}\{GB \mid \text{met Jimmy}\} = \boldsymbol{P}\{\text{outcome } 4 \mid \text{met Jimmy}\} = 1/4.$

(d) $\boldsymbol{P}\{\text{Jimmy has a sister} \mid \text{met Jimmy}\}$
$= \boldsymbol{P}\{\text{outcomes } 4,\, 6 \mid \text{met Jimmy}\} = 1/2.$

2.31 According to (2.2),

$$\overline{A \cap B \cap C \cap \ldots} = \overline{A} \cup \overline{B} \cup \overline{C} \cup \ldots.$$

Then, events $A, B, C, \ldots$ are disjoint (i.e., $A \cap B \cap C \cap \ldots = \varnothing$) if and only if

$$\overline{A} \cup \overline{B} \cup \overline{C} \cup \ldots = \overline{A \cap B \cap C \cap \ldots} = \overline{\varnothing} = \Omega.$$

We see that the union of $\overline{A}, \overline{B}, \overline{C}, \ldots$ equals the entire sample space in this case. By Definition 2.9 $\overline{A}, \overline{B}, \overline{C}, \ldots$ are exhaustive.

2.32 *Intuitive solutions:*

(a) Independent events A and B occur independently of each other. Hence, they *don't occur* independently of each other. Every time when A (or B) does not occur, its complement occurs. Hence, the complements of A and B are also independent of each other.

(b) Being disjoint is a very strong dependence because disjoint events completely eliminate each other. The only way for such events to be independent is when one of these events is *always* eliminated. Such an event must have probability 0.

(c) Being exhaustive is also a strong type of dependence because one event absolutely has to cover all the parts of Ω that are not covered by the other event. The only way for such events to be independent is when one of the events covers all the parts of Ω regardless of the other event. Such event should be the entire sample space, Ω.

Mathematical solutions:

(a) Using (2.2),

$$\begin{aligned}
\boldsymbol{P}\left\{\overline{A}\cap\overline{B}\right\} &= \boldsymbol{P}\left\{\overline{A\cup B}\right\} = 1-\boldsymbol{P}\{A\cup B\} \\
&= 1-(\boldsymbol{P}\{A\}+\boldsymbol{P}\{B\}-\boldsymbol{P}\{A\cap B\}) \\
&= 1-\boldsymbol{P}\{A\}-\boldsymbol{P}\{B\}+\boldsymbol{P}\{A\}\boldsymbol{P}\{B\} \\
&\qquad (\textit{because } A \textit{ and } B \textit{ are independent}) \\
&= (1-\boldsymbol{P}\{A\})(1-\boldsymbol{P}\{B\}) \\
&= \boldsymbol{P}\left\{\overline{A}\right\}\boldsymbol{P}\left\{\overline{B}\right\}.
\end{aligned}$$

Hence, $\overline{A}$ and $\overline{B}$ are independent.

(b) If A and B are independent and disjoint, then

$$0=\boldsymbol{P}\{A\cap B\}=\boldsymbol{P}\{A\}\boldsymbol{P}\{B\},$$

which can only happen when $\boldsymbol{P}\{A\}=0$ or $\boldsymbol{P}\{B\}=0$.

(c) If A and B are independent and exhaustive, then

$$\begin{aligned}
1 &= \boldsymbol{P}\{A\cup B\}=\boldsymbol{P}\{A\}+\boldsymbol{P}\{B\}-\boldsymbol{P}\{A\cap B\} \\
&= \boldsymbol{P}\{A\}+\boldsymbol{P}\{B\}-\boldsymbol{P}\{A\}\boldsymbol{P}\{B\}.
\end{aligned}$$

Then

$$0=1-(\boldsymbol{P}\{A\}+\boldsymbol{P}\{B\}-\boldsymbol{P}\{A\}\boldsymbol{P}\{B\})=(1-\boldsymbol{P}\{A\})(1-\boldsymbol{P}\{B\}),$$

which can only happen when $\boldsymbol{P}\{A\}=1$ or $\boldsymbol{P}\{B\}=1$.

2.33 Generalizing (2.4), we prove that for any events $E_1,\ldots,E_n$,

$$\begin{aligned}
&\boldsymbol{P}\{E_1\cup\ldots\cup E_n\} \\
&\quad=\sum_{i\le n}\boldsymbol{P}\{E_i\}-\sum_{1\le i<j\le n}\boldsymbol{P}\{E_i\cap E_j\}+\sum_{1\le i<j<k\le n}\boldsymbol{P}\{E_i\cap E_j\cap E_k\}-\ldots \\
&\qquad-(-1)^n\boldsymbol{P}\{E_1\cap\ldots\cap E_n\}.
\end{aligned}$$

This can be proved by induction.

For $n=2$ events, this formula is given by (2.4).

Suppose the formula is true for n events. Let A denote their overall union,

$A = E_1 \cup \ldots \cup E_n$. Then for any event E_{n+1},

$$\begin{aligned}
\boldsymbol{P}\{E_1 \cup \ldots \cup E_{n+1}\} &= \boldsymbol{P}\{A \cup E_{n+1}\} \\
&= \boldsymbol{P}\{A\} + \boldsymbol{P}\{E_{n+1}\} - \boldsymbol{P}\{A \cap E_{n+1}\} \\
&= \sum_{i \le n+1} \boldsymbol{P}\{E_i\} - \sum_{1 \le i < j \le n} \boldsymbol{P}\{E_i \cap E_j\} + \sum_{1 \le i < j < k \le n} \boldsymbol{P}\{E_i \cap E_j \cap E_k\} \\
&\quad - \ldots - (-1)^n \boldsymbol{P}\{E_1 \cap \ldots \cap E_n\} - \boldsymbol{P}\{A \cap E_{n+1}\}.
\end{aligned}$$

Also, since the formula is assumed true for n events,

$$\begin{aligned}
\boldsymbol{P}\{A \cap E_{n+1}\} &= \boldsymbol{P}\{(E_1 \cap E_{n+1}) \cup \ldots \cup (E_n \cap E_{n+1})\} \\
&= \sum_{i \le n} \boldsymbol{P}\{E_i \cap E_{n+1}\} - \sum_{1 \le i < j \le n} \boldsymbol{P}\{E_i \cap E_j \cap E_{n+1}\} + \ldots \\
&\quad - (-1)^n \boldsymbol{P}\{E_1 \cap \ldots \cap E_{n+1}\}.
\end{aligned}$$

Altogether,

$$\begin{aligned}
&\boldsymbol{P}\{E_1 \cup \ldots \cup E_{n+1}\} \\
&= \sum_{i \le n+1} \boldsymbol{P}\{E_i\} - \sum_{1 \le i < j \le n+1} \boldsymbol{P}\{E_i \cap E_j\} + \sum_{1 \le i < j < k \le n+1} \boldsymbol{P}\{E_i \cap E_j \cap E_k\} \\
&\quad - \ldots - (-1)^{n+1} \boldsymbol{P}\{E_1 \cap \ldots \cap E_{n+1}\}.
\end{aligned}$$

This proves the formula for $(n+1)$ events. By induction, the formula is proved for any $n \ge 2$.

2.34 Let $A_i = \bar{E}_i$ for $i = 1, \ldots, n$. According to (2.2),

$$\overline{A}_1 \cap \ldots \cap \overline{A}_n = \overline{A_1 \cup \ldots \cup A_n}$$

Therefore,

$$\overline{E_1 \cap \ldots \cap E_n} = \overline{\overline{A}_1 \cap \ldots \cap \overline{A}_n} = A_1 \cup \ldots \cup A_n = \overline{E}_1 \cup \ldots \cup \overline{E}_n$$

Chapter 3

3.1 Possible values of X are: 0, 1, and 2.

(a) The pmf is:

$$\begin{aligned}
P(0) &= \boldsymbol{P}\{\text{both files are not corrupted}\} \\
&= (1-0.4)(1-0.3) = 0.42, \\
P(1) &= \boldsymbol{P}\left\{\begin{array}{c}\text{1st is corrupted,}\\ \text{2nd is not}\end{array}\right\} + \boldsymbol{P}\left\{\begin{array}{c}\text{2nd is corrupted,}\\ \text{1st is not}\end{array}\right\} \\
&= (0.4)(1-0.3) + (0.3)(1-0.4) = 0.46, \\
P(2) &= \boldsymbol{P}\{\text{both are corrupted}\} = (0.4)(0.3) = 0.12.
\end{aligned}$$

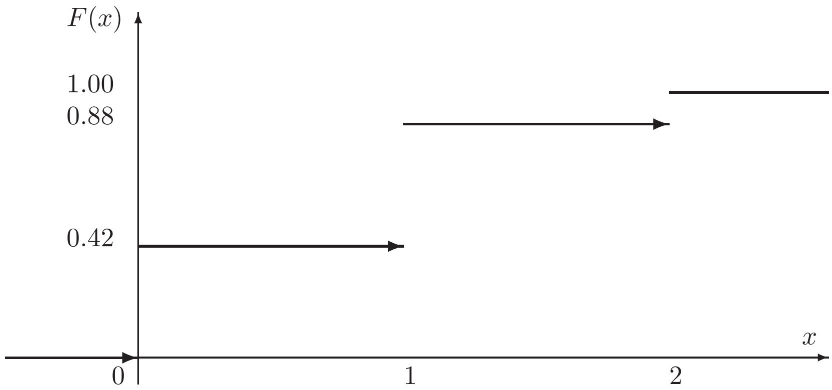


Figure 11.3 *The cdf of X for Exercise 3.1*

(check: $P(0) + P(1) + P(2) = 1$.)

(b) The cdf is given in Figure 11.3.

3.2 Let X be the number of network blackouts, and Y be the loss. Then $Y = 500X$. Compute

$$\begin{aligned}\mathbf{E}(X) &= \sum_x xP(x) = (0)(0.7) + (1)(0.2) + (2)(0.1) = 0.4;\\ \mathrm{Var}(X) &= \sum_x (x-0.4)^2 P(x)\\ &= (0-0.4)^2(0.7) + (1-0.4)^2(0.2) + (2-0.4)^2(0.1) = 0.44.\end{aligned}$$

Hence,

$$\mathbf{E}(Y) = 500\,\mathbf{E}(X) = (500)(0.4) = \boxed{\text{200 dollars}}$$

and

$$\mathrm{Var}(Y) = 500^2\,\mathrm{Var}(X) = (250,000)(0.44) = \boxed{\text{110,000 squared dollars}}$$

3.3 Find the distribution of X, then compute $\mathbf{E}(X)$ and $\mathrm{Var}(X)$. Since there is only 1 error in the entire program, the number of errors in 3 selected blocks may be either 0 or 1, where

$$\begin{aligned}P(1) &= \boldsymbol{P}\{\text{error in 3 selected blocks}\} = 3/5,\\ P(0) &= \boldsymbol{P}\{\text{error in 2 other blocks}\} = 2/5.\end{aligned}$$

Then

$$\begin{aligned}
\mathbf{E}(X) &= \sum_x xP(x) = (0)(2/5) + (1)(3/5) = \boxed{3/5} \\
\mathbf{E}(X^2) &= \sum_x x^2P(x) = (0)^2(2/5) + (1)^2(3/5) = 3/5; \\
\mathrm{Var}(X) &= \mathbf{E}(X^2) - \mathbf{E}^2(X) = 3/5 - (3/5)^2 = \boxed{6/25 \text{ or } 0.24}
\end{aligned}$$

Shortcut: Actually, $X^2 = X$ for $X = 0$ or 1, therefore, $\mathbf{E}(X^2) = \mathbf{E}(X) = 3/5$.

3.4 We can complete a table,

x	$P(x)$	$xP(x)$	x^2	$x^2P(x)$
1	1/6	1/6	1	1/6
2	1/6	2/6	4	4/6
3	1/6	3/6	9	9/6
4	1/6	4/6	16	16/6
5	1/6	5/6	25	25/6
6	1/6	6/6	36	36/6
		21/6 or 3.5		91/6

$$\mathbf{E}(X) = \sum_x xP(x) = \boxed{3.5}$$

$$\mathrm{Var}(X) = \mathbf{E}(X^2) - \mathbf{E}^2(X) = 91/6 - (7/2)^2 = \boxed{35/12 \text{ or } 2.9167}$$

We can also compute $\mathrm{Var}(X)$ as

$$\mathrm{Var}(X) = \sum_x (x - 3.5)^2 P(x)$$

$$= (1/6)(.5^2 + 1.5^2 + 2.5^2 + 2.5^2 + 1.5^2 + .5^2) = \boxed{2.9167}$$

3.5 Let X be the number of randomly chosen programs that need to be updated. There are 5 programs overall that need an update and 7 that don't need an update. Outcomes are sets of 4 chosen programs, and they are equally likely. The total number of outcomes is

$$\mathcal{N}_T = \binom{12}{4} = \frac{12 \cdot 11 \cdot 10 \cdot 9}{4 \cdot 3 \cdot 2 \cdot 1} = 495.$$

Thus, X has a pmf

$$\begin{aligned}
P(0) &= \frac{\binom{5}{0}\binom{7}{4}}{495} = \frac{(1)(7 \cdot 6 \cdot 5 \cdot 4)/(4 \cdot 3 \cdot 2 \cdot 1)}{495} = \frac{35}{495} = \frac{7}{99}, \\
P(1) &= \frac{\binom{5}{1}\binom{7}{3}}{495} = \frac{(5)(7 \cdot 6 \cdot 5)/(3 \cdot 2 \cdot 1)}{495} = \frac{175}{495} = \frac{35}{99}, \\
\text{etc.} & \quad \ldots\ldots\ldots
\end{aligned}$$

(a) $\boldsymbol{P}\{X \geq 2\} = 1 - P(0) - P(1) = \dfrac{57}{99} = \boxed{\dfrac{19}{33} \text{ or } 0.5758}$

(b) Two solutions will follow: one is direct but long, the other is short and elegant.

Long solution. Compute the entire pmf of X,

$$\begin{aligned} P(2) &= \frac{\binom{5}{2}\binom{7}{2}}{495} = \frac{\frac{5\cdot 4}{2}\frac{7\cdot 2}{2}}{495} = \frac{210}{495} = \frac{42}{99}, \\ P(3) &= \frac{\binom{5}{3}\binom{7}{1}}{495} = \frac{(5\cdot 4/2)(7)}{495} = \frac{70}{495} = \frac{14}{99}, \\ P(4) &= \frac{\binom{5}{4}\binom{7}{0}}{495} = \frac{(5)(1)}{495} = \frac{5}{495} = \frac{1}{99} \end{aligned}$$

Then

$$\begin{aligned} \mathbf{E}(X) &= \sum_x xP(x) \\ &= (0)\left(\frac{7}{99}\right) + (1)\left(\frac{35}{99}\right) + (2)\left(\frac{42}{99}\right) + (3)\left(\frac{14}{99}\right) + (4)\left(\frac{1}{99}\right) \\ &= \boxed{\frac{5}{3} \text{ or } 1.6667} \end{aligned}$$

Short solution. For each of the four chosen programs $i = 1, 2, 3, 4$, define a random variable

$$X_i = \begin{cases} 1 & \text{if program } i \text{ needs an upgrade} \\ 0 & \text{if program } i \text{ does not need an upgrade} \end{cases}$$

Each X_i has Bernoulli distribution with parameter

$$p = \boldsymbol{P}\{\text{if program } i \text{ needs an upgrade}\} = \frac{5}{12},$$

because 5 programs out of 12 must be upgraded. Then

$$\mathbf{E}(X) = \mathbf{E}(X_1+X_2+X_3+X_4) = p+p+p+p = (4)\left(\frac{5}{12}\right) = \boxed{\frac{5}{3} \text{ or } 1.6667}$$

3.6 Compute the pmf of X. It takes two possible values, 0 and 1, with probabilities

$$P(0) = \frac{\binom{2}{0}\binom{4}{2}}{\binom{6}{2}} = 0.4 \text{ and } P(1) = 1 - P(0) = 0.6.$$

The expectation of X equals

$$\mathbf{E}(X) = (0)P(0) + (1)P(1) = \boxed{0.6}$$

3.7 Let X_i be the number of home runs in game i for $i = 1, 2$. Compute

$$\begin{aligned}
\mathbf{E}(X) &= \sum_x xP(x) = (0)(0.4) + (1)(0.4) + (2)(0.2) = 0.8; \\
\mathbf{E}(X^2) &= \sum_x x^2 P(x) = (0)(0.4) + (1)(0.4) + (4)(0.2) = 1.2; \\
\mathrm{Var}(X) &= \mathbf{E}(X^2) - \mathbf{E}^2(X) = 1.2 - 0.8^2 = 0.56.
\end{aligned}$$

Then use the fact that $Y = X_1 + X_2$, where X_1 and X_2 are independent.

$$\begin{aligned}
\mathbf{E}(Y) &= \mathbf{E}(X_1) + \mathbf{E}(X_2) = 0.8 + 0.8 = \boxed{1.6} \\
\mathrm{Var}(Y) &= \mathrm{Var}(X_1) + \mathrm{Var}(X_2) = 0.56 + 0.56 = \boxed{1.12}
\end{aligned}$$

3.8 Searching for the right password, out of possible four, the user may try between 0 and 3 wrong passwords. The correct password is equally likely to be the first, second, third, or fourth one that she tries. Therefore, X has the pmf

$$P(0) = P(1) = P(2) = P(3) = \frac{1}{4}$$

Then compute

$$\mathbf{E}(X) = \sum_{x=0}^{3} xP(x) = (0 + 1 + 2 + 3)\left(\frac{1}{4}\right) = \boxed{1.5}$$

and

$$\begin{aligned}
\mathrm{Var}(X) &= \sum_{x=0}^{3} (x - 1.5)^2 P(x) \\
&= \left((-1.5)^2 + (-0.5)^2 + (0.5)^2 + (1.5)^2\right)\left(\frac{1}{4}\right) = \boxed{1.25}
\end{aligned}$$

3.9 We are given $\mu = 40$ seconds and $\sigma = 5$ seconds. Using Chebyshev's inequality (3.8) on p. 58 with $\varepsilon = 20$ seconds, we get

$$\begin{aligned}
\boldsymbol{P}\{X > 60\} &\le \boldsymbol{P}\{|X - 40| > 20\} \\
&\le \boldsymbol{P}\{|X - \mu| > \varepsilon\} \le \left(\frac{\sigma}{\varepsilon}\right)^2 = \left(\frac{5}{20}\right)^2 = 1/16.
\end{aligned}$$

Thus, the probability of spending more than 1 minute (60 seconds) does not exceed 1/16.

3.10 *Solution 1.*

Let X be the number of accidents on Thursday, and Y be the number of accidents on Friday. Using independence, compute the joint distribution

$P(x,y) = P(x)P(y)$,

$P(x,y)$		x = 0	1	2
y	0	.36	.12	.12
	1	.12	.04	.04
	2	.12	.04	.04

Then

$$\boldsymbol{P}\{X < Y\} = P(0,1) + P(0,2) + P(1,2) = .12 + .12 + .04 = \boxed{0.28}$$

Solution 2.

Since X and Y have the same distribution, events $\{X < Y\}$ and $\{X > Y\}$ occur with the same probability. The only other case is $X = Y$ and the probability of this is

$$\boldsymbol{P}\{X = Y\} = P(0,0) + P(1,1) + P(2,2) = (0.6)^2 + (0.2)^2 + (0.2)^2 = 0.44.$$

Then

$$\boldsymbol{P}\{X < Y\} = \frac{1 - 0.44}{2} = \boxed{0.28}$$

3.11 The sample space for this experiment consists of all 36 pairs of integer numbers between 1 and 6. All outcomes are equally likely. Hence, the probability of each outcome is $1/36$.

(a) Both X and Y take values 1, 2, 3, 4, 5, and 6. X is the smaller, and Y is the larger, hence $X \le Y$. Therefore,

- if $x = y$, then $P_{(X,Y)}(x,y) = 1/36$;
- if $x < y$, then $P_{(X,Y)}(x,y) = 2/36 = 1/18$ because it includes 2 outcomes;
- if $x > y$, then $P_{(X,Y)}(x,y) = 0$.

The joint pmf is

$P_{(X,Y)}(x,y)$		y = 1	2	3	4	5	6	$P_X(x)$
x	1	1/36	1/18	1/18	1/18	1/18	1/18	11/36
	2	0	1/36	1/18	1/18	1/18	1/18	9/36
	3	0	0	1/36	1/18	1/18	1/18	7/36
	4	0	0	0	1/36	1/18	1/18	5/36
	5	0	0	0	0	1/36	1/18	3/36
	6	0	0	0	0	0	1/36	1/36
	$P_Y(y)$	1/36	3/36	5/36	7/36	9/36	11/36	1

(Check that $\sum P_X(x) = 1$, $\sum P_Y(y) = 1$, and $\sum P_{(X,Y)}(x,y) = 1$.)

(b) They are dependent. For example,

$$0 = P_{(X,Y)}(6,1) \neq P_X(6)P_Y(1) = (1/36)(1/36).$$

(c) To find the marginal pmf of X, add the joint probabilities $P_X(1) = 11/36$, $P_X(2) = 9/36 = 1/4$, $P_X(3) = 7/36$, $P_X(4) = 5/36$, $P_X(5) = 3/36 = 1/12$, $P_X(6) = 1/36$.

(d) $\boldsymbol{P}\{Y = 5 \mid X = 2\} = P_{(X,Y)}(2,5)/P_X(2) = (1/18)/(1/4) = \boxed{2/9}$

3.12 (a) Compute the marginal distributions by the Addition Rule, as in the table.

$P(x,y)$		$x=0$	$x=1$	$P_Y(y)$
y	0	0.5	0.2	0.7
	1	0.2	0.1	0.3
$P_X(x)$		0.7	0.3	

Then check:

$$P_X(0)P_Y(0) = (0.7)(0.7) = 0.49 \neq 0.5 = P(0,0).$$

Hence, X and Y are dependent.

(b) Variables $(X+Y)$ and $(X-Y)$ are also dependent.

Explanation 1. If $(X+Y)$ is odd then $(X-Y)$ is odd. If $(X+Y)$ is even then $(X-Y)$ is even. Hence, they are dependent.

Explanation 2. For each pair of (X,Y), we can compute $V = (X+Y)$ and $W = (X-Y)$:

$P_{(V,W)}(v,w)$		$v=x+y$: 0	1	2	$P_W(w)$
$w = x - y$	-1	0	0.2	0	0.2
	0	0.5	0	0.1	0.6
	1	0	0.2	0	0.2
	$P_V(v)$	0.5	0.4	0.1	

We see, for example, that

$$P_V(0)P_W(1) = (0.5)(0.2) \neq 0 = P_{(V,W)}(0,1).$$

Hence, V and W are dependent.

3.13 From the given joint distribution, we find that X and Y can take values 0, 1, and 2. To find the distributions (pmf) of Z, U, and V,

(1) identify possible values of each variable;

(2) compute the probability of each possible value;

(3) check that the sum of probabilities equals 1.

(a) Possible values of $Z = X + Y$ are: 0, 2, and 4. The pmf is

$$\begin{aligned} P_Z(0) &= P(0,0) = 0.2 \\ P_Z(2) &= P(0,2) + P(2,0) + P(1,1) = 0.7 \\ P_Z(4) &= P(2,2) = 0.1 \end{aligned}$$

(b) Possible values of $U = X - Y$ are: -2, 0, and 2. The pmf is

$$\begin{aligned} P_U(-2) &= P(0,2) = 0.3 \\ P_U(0) &= P(0,0) + P(1,1) + P(2,2) = 0.4 \\ P_U(2) &= P(2,0) = 0.3 \end{aligned}$$

(c) Possible values of $V = XY$ are: 0, 1, and 4. The pmf is

$$\begin{aligned} P_V(0) &= P(0,0) + P(0,2) + P(2,0) = 0.8 \\ P_V(1) &= P(1,1) = 0.1 \\ P_V(4) &= P(2,2) = 0.1 \end{aligned}$$

3.14 We can write a table of possible values of X, Y, and Z.

x	y	$z = xy$	$P(x,y)$
1	2	2	0.10
1	3	3	0.40
2	1	2	0.06
2	2	4	0.10
2	3	6	0.10
3	1	3	0.06
3	2	6	0.04
4	1	4	0.10
4	2	8	0.04

From this table, $Z = XY$ takes values 2, 3, 4, 6, and 8 cents. Its pmf is

$$\begin{aligned} P_Z(2) &= 0.10 + 0.06 = 0.16 \\ P_Z(3) &= 0.40 + 0.06 = 0.46 \\ P_Z(4) &= 0.10 + 0.10 = 0.20 \\ P_Z(6) &= 0.10 + 0.04 = 0.14 \\ P_Z(8) &= 0.04 \end{aligned}$$

3.15 (a) $1 - P_{(X,Y)}(0,0) = 1 - 0.52 = 0.48$.

(b) Compute the marginal distributions of X and Y,

$$P_X(x) = \sum_{y=0}^{2} P(x,y) \text{ and } P_Y(y) = \sum_{x=0}^{2} P(x,y)$$

$P(x,y)$		x = 0	1	2	$P_Y(y)$
y	0	0.52	0.20	0.04	0.76
	1	0.14	0.02	0.01	0.17
	2	0.06	0.01	0	0.07
$P_X(x)$		0.72	0.23	0.05	

Variables X and Y are dependent. For example, $P_{(X,Y)}(2,2) \neq P_X(2)P_Y(2)$ because $(0.05)(0.07) \neq 0$.

3.16 Find the marginal pmf of X and Y,

$$\begin{aligned} P_X(0) &= P(0,0)+P(0,1) = 0.6+0.1 = 0.7, \\ P_X(1) &= P(1,0)+P(1,1) = 0.1+0.2 = 0.3; \\ P_Y(0) &= P(0,0)+P(1,0) = 0.6+0.1 = 0.7, \\ P_Y(1) &= P(0,1)+P(1,1) = 0.1+0.2 = 0.3. \end{aligned}$$

(a) Variables X and Y are dependent. For example, $P_{(X,Y)}(0,0) \neq P_X(0)P_Y(0)$ because $(0.7)(0.7) \neq 0.6$.

(b) Compute expectations $\mathbf{E}(X)$ and $\mathbf{E}(Y)$. We know from (a) that these two variables have the same distribution, and thus, they have the same expected value

$$\mathbf{E}(X) = \mathbf{E}(Y) = (0)(0.7) + (1)(0.3) = 0.3.$$

Then

$$\mathbf{E}(X+Y) = \mathbf{E}(X) + \mathbf{E}(Y) = \boxed{0.6}$$

Alternatively, we can find the pmf of $Z = X+Y$ and use it to compute $\mathbf{E}(Z)$.

3.17 Let X be the profit made on 1 share of A, and Y be the profit made on 1 share of B. Each of these variables has the distribution

x	$P(x)$
-1	0.5
1	0.5

Then

$$E(X) = E(Y) = (-1)(0.5) + (1)(0.5) = 0$$

and

$$Var(X) = Var(Y) = E(X-0)^2 = (-1)^2(0.5) + (1)^2(0.5) = 1$$

The risk of each portfolio is measured by the variance of its return Z:

(a) $Z = 100X;\ \mathrm{Var}(Z) = 100^2\,\mathrm{Var}(X) = 10,000.$

(b) $Z = 50X + 10Y;\ \mathrm{Var}(Z) = 50^2\,\mathrm{Var}(X) + 10^2\,\mathrm{Var}(Y) = 2,600.$

(c) $Z = 40X + 12Y;\ \mathrm{Var}(Z) = 40^2\,\mathrm{Var}(X) + 12^2\,\mathrm{Var}(Y) = 1,744.$

Thus, the *third portfolio* has the lowest risk.

3.18 The dollar profit Z equals

$$Z = (100 \text{ shares})(10 \text{ dollars per share}) \left(\frac{X}{100} \text{ profit} \right) = 10X$$

for portfolio (a), and similarly, $5X + 5Y$ for portfolio (b), and $10Y$ for portfolio (c).

Compute expectations and variances of X and Y.

$$\begin{aligned}
\mathbf{E}(X) &= (-3)(0.3) + (0)(0.2) + (3)(0.5) = 0.6, \\
\mathrm{Var}(X) &= (-3)^2(0.3) + (0)^2(0.3) + (3)^2(0.5) - 0.6^2 = 7.2 - 0.36 = 6.84; \\
\mathbf{E}(Y) &= (-3)(0.4) + (3)(0.6) = 0.6, \\
\mathrm{Var}(Y) &= (-3)^2(0.4) + (3)^2(0.6) - 0.6^2 = 9 - 0.36 = 8.64.
\end{aligned}$$

Now,

(a) (a) $\mathbf{E}(Z) = \mathbf{E}(10X) = 10\,\mathbf{E}(X) = 6$,
$\mathrm{Var}(Z) = \mathrm{Var}(10X) = 100\,\mathrm{Var}(X) = 684$.

(b) $\mathbf{E}(Z) = \mathbf{E}(5X + 5Y) = 5\,\mathbf{E}(X) + 5\,\mathbf{E}(Y) = 6$,
$\mathrm{Var}(Z) = \mathrm{Var}(5X + 5Y) = 25\,\mathrm{Var}(X) + 25\,\mathrm{Var}(Y) = 387$.

(c) $\mathbf{E}(Z) = \mathbf{E}(10Y) = 10\,\mathbf{E}(Y) = 6$,
$\mathrm{Var}(Z) = \mathrm{Var}(10Y) = 100\,\mathrm{Var}(Y) = 864$.

The least risky portfolio is (b); the most risky portfolio is (c).

3.19 Find the expectations and variances of X and Y:

$$\begin{aligned}
\mathbf{E}(X) &= (-2)(0.5) + (2)(0.5) = 0, \\
\mathrm{Var}(X) &= (-2)^2(0.5) + (2)^2(0.5) - 0^2 = 4; \\
\mathbf{E}(Y) &= (-1)(0.8) + (4)(0.2) = 0, \\
\mathrm{Var}(Y) &= (-1)^2(0.8) + (4)^2(0.2) - 0^2 = 4.
\end{aligned}$$

Expectations and variances of the total profit for the three listed strategies are:

(a) $\mathbf{E}(100X) = 100\,\mathbf{E}(X) = 0$,
$\mathrm{Var}(100X) = 100^2\,\mathrm{Var}(X) = 40,000$

(b) $\mathbf{E}(100Y) = 100\,\mathbf{E}(Y) = 0$,
$\mathrm{Var}(100Y) = 100^2\,\mathrm{Var}(Y) = 40,000$

(c) $\mathbf{E}(50X + 50Y) = 50\,\mathbf{E}(X) + 50\,\mathbf{E}(Y) = 0$,
$\mathrm{Var}(50X + 50Y) = 50^2\,\mathrm{Var}(X) + 50^2\,\mathrm{Var}(Y) = 10,000$

Each strategy has expected profit 0. Splitting the investment between A and B (diversifying the portfolio) is the least risky strategy.

3.20 (a) We need to compute $\boldsymbol{P}\{X = 3\}$, where X is the number of defective computers ("successes") in a shipment of 20 ("trials"). It has Binomial distribution with parameters $n = 20$ and $p = 0.05$. From Table A2,

$$P(X = 3) = P(X \leq 3) - P(X \leq 2) = .9841 - .9245 = \boxed{0.0596}$$

(b) Let Y be the number of defective computers among the first four. It has Binomial distribution with $n = 4$ and $p = 0.05$. From Table A2,

$$\begin{aligned} & \boldsymbol{P}\{\text{at least 5 computers are tested until 2 defective ones are found}\} \\ = \; & \boldsymbol{P}\{\text{among the first 4 computers, at most 1 is defective}\} \\ = \; & \boldsymbol{P}\{Y \leq 1\} = \boxed{0.9860} \end{aligned}$$

The problem can also be solved directly, but computing $\boldsymbol{P}\{W \geq 5\}$, where W is the Negative Binomial ($k = 2$, $p = 0.05$) number of computers the engineer has to test in order to find 2 defective computers:

$$\begin{aligned} \boldsymbol{P}\{W \geq 5\} &= 1 - P(2) - P(3) - P(4) \\ &= 1 - 0.05^2 - (2)(0.05)^2(0.95) - (3)(0.05)^2(0.95)^2 \\ &= \boxed{0.9860} \end{aligned}$$

3.21 Let X be the number of computers entered by the virus. Each of the 20 computers is either entered or not, thus X is the number of "successes" in $n = 20$ Bernoulli trials. Hence, X has Binomial distribution with $n = 20$ and $p = 0.4$. From Table A2,

$$\boldsymbol{P}\{X \geq 10\} = 1 - \boldsymbol{P}\{X \leq 9\} = 1 - 0.7553 = \boxed{0.2447}$$

3.22 We need $\boldsymbol{P}\{X > 3\}$, where X is the number of defective computers in this sample. It is the number of successes in $n = 16$ Bernoulli trials, where each computer is a "trial" (defective or not) and a defective computer is a "success". Therefore, X has Binomial distribution with $n = 16$, $p = 0.05$. From Table A2,

$$\boldsymbol{P}\{X > 3\} = 1 - F(3) = 1 - 0.9930 = \boxed{0.0070}$$

3.23 (a) We need to find $\boldsymbol{P}\{X \geq 4\}$, where X is the number of canceled classes out of 15 remaining. This is the number of "successes" (cancelations) in 15 trials (scheduled classes), hence it has Binomial distribution with $n = 15$ and $p = 0.05$. From Table A2,

$$\boldsymbol{P}\{X \geq 4\} = 1 - \boldsymbol{P}\{X \leq 3\} = 1 - 0.9945 = \boxed{0.0055}$$

(b) Two solutions are possible.

First solution. We need $\boldsymbol{P}\{Y = 10\}$, where Y is the number of classes until three classes are canceled. This is the number of trials until there are 3 successes. Hence, Y is Negative Binomial with $k = 3$ and $p = 0.05$.

$$\boldsymbol{P}\{Y = 10\} = \binom{9}{2}(0.05)^3(0.95)^7 = \frac{(9)(8)}{2}(0.05)^3(0.95)^7 = \boxed{0.00314}$$

Second solution. Let Z be the number of cancelations out of the first 9 classes. Then Z is Binomial with $n = 9$ and $p = 0.05$;

$$\boldsymbol{P}\{Z = 2\} = 0.9916 - 0.9288 = 0.0628,$$

from Table A2, and

$$\boldsymbol{P}\{Y = 10\} = \boldsymbol{P}\{Z = 2\}(0.05) = (0.0628)(0.05) = \boxed{0.00314}$$

because in order to have the third success on the tenth trial, there must be 2 successes in the first 9 trials, and the tenth trial must be a success.

3.24 (a) Let X be the number of sites with a keyword, which is Binomial($n = 10$, $p = 0.2$). From Table A2,

$$\boldsymbol{P}\{X \geq 5\} = 1 - \boldsymbol{P}\{X \leq 4\} = 1 - 0.9672 = \boxed{0.0328}$$

(b) Let Y be the number of sites visited until a site with a keyword is found, which is Geometric($p = 0.2$).

$$\boldsymbol{P}\{Y \geq 5\} = (1-p)^4 = \boxed{0.4096}$$

3.25 (a) Let X be the number of users who don't close Windows properly before someone does. It is the number of trials before the first success, and therefore, $(X + 1)$ has Geometric distribution with $p = 1 - 0.1 = 0.9$. Compute the expectation,

$$\mathbf{E}(X) = \mathbf{E}(X+1) - 1 = \frac{1}{0.9} - 1 = \boxed{0.1111}$$

(b) Let Y be the number of users (out of the next ten) who close Windows properly. It has Binomial distribution with $n = 10$ and $p = 0.9$. From Table A2,

$$\boldsymbol{P}\{Y = 8\} = F(8) - F(7) = 0.2639 - 0.0702 = \boxed{0.1937}$$

3.26 (a) Let X be the number of damaged files. This is the number of "successes" (damaged files) out of 20 "trials" (files), thus, it has Binomial distribution with $n = 20$, $p = 0.2$. From Table A2,

$$\boldsymbol{P}\{X \geq 5\} = 1 - F(4) = 1 - 0.6296 = \boxed{0.3704}$$

(b) The number of files to be checked has Negative Binomial distribution with parameters $k = 3$ (successes, undamaged files) and $p = 0.8$ (probability of an undamaged file). However, we can use Table A2 as follows:

$$\begin{aligned}
&\boldsymbol{P}\{\text{need to check at least 6 files}\} \\
&= \boldsymbol{P}\{\text{ 5 files are not enough }\} \\
&= \boldsymbol{P}\{\text{ there at most 2 undamaged files among the first 5}\} \\
&= F_X(2) = \boxed{0.0579},
\end{aligned}$$

where X, the number of undamaged files among the first five, has Binomial distribution with $n = 5$ and $p = 0.8$.

3.27 We need to find $\boldsymbol{P}\{X \geq 5\}$ and $\boldsymbol{P}\{X = 5\}$, where X is the number of arrived messages during the next hour. This is the number of "rare" events, which can be any nonnegative integer number. Also, their distribution has only one parameter, the arrival rate $\lambda = 9$. Therefore, X has Poisson distribution with parameter $\lambda = 9$. From Table A3,

(a) $\boldsymbol{P}\{X \geq 5\} = 1 - \boldsymbol{P}\{X \leq 4\} = 1 - .055 = \boxed{.945}$

(b) $\boldsymbol{P}\{X = 5\} = \boldsymbol{P}\{X \leq 5\} - \boldsymbol{P}\{X \leq 4\} = .116 - .055 = \boxed{.061}$

3.28 Let X be the Poisson(λ) number of received messages. Then

$$\mu = \mathbf{E}(X) = \lambda, \quad \sigma = \mathrm{Std}(X) = \sqrt{\lambda},$$

and by Chebyshev inequality,

$$\begin{aligned} \boldsymbol{P}\{X > 4\lambda\} &= \boldsymbol{P}\{X - \lambda > 3\lambda\} \leq \boldsymbol{P}\{|X - \lambda| > 3\lambda\} \\ &= \boldsymbol{P}\left\{\frac{|X - \lambda|}{\sqrt{\lambda}} > 3\sqrt{\lambda}\right\} \leq \frac{1}{(3\sqrt{\lambda})^2} = \frac{1}{9\lambda} \end{aligned}$$

3.29 Denote the events: $H = \{\text{high risk}\}$, $L = \{\text{low risk}\}$, $N = \{\text{no accidents}\}$. The number of accidents is the number of "rare events", discrete, ranging from 0 to infinity, thus it has a Poisson distribution. We have:

$$\boldsymbol{P}\{H\} = 0.2, \ \boldsymbol{P}\{L\} = 0.8, \ \boldsymbol{P}\{N|H\} = 0.368, \ \boldsymbol{P}\{N|L\} = 0.905$$

(from Table A3, with $\lambda = 1$ and $\lambda = 0.1$).

By the Bayes' Rule,

$$\begin{aligned} \boldsymbol{P}\{H|N\} &= \frac{\boldsymbol{P}\{N|H\}\boldsymbol{P}\{H\}}{\boldsymbol{P}\{N|H\}\boldsymbol{P}\{H\} + \boldsymbol{P}\{N|L\}\boldsymbol{P}\{L\}} \\ &= \frac{(0.368)(0.2)}{(0.368)(0.2) + (0.905)(0.8)} = \boxed{0.0923} \end{aligned}$$

3.30 (a) Let X be the number of components that pass the inspection. It is the number of successes in 20 Bernoulli trials, thus it has Binomial distribution with $n = 20$ and $p = 0.8$. From Table A2,

$$\boldsymbol{P}\{X \geq 18\} = 1 - F_X(17) = 1 - 0.7939 = \boxed{0.2061}$$

(b) Let Y be the number of components should be inspected until a component that passes inspection is found. It is the number of trials needed to see the first success, thus it has Geometric distribution with $p = 0.8$.

$$\mathbf{E}(Y) = \frac{1}{p} = \boxed{1.25 \text{ components}}$$

3.31 Let X be the number of crashed computers. This is the number of "successes" (crashed computers) out of 4,000 "trials" (computers), with the probability of success 1/800. Thus, it has Binomial distribution with parameters $n = 4,000$ (large) and $p = 1/800$ (small), that is approximately Poisson with

$$\lambda = np = 5.$$

Using Table A3 with parameter 5,

(a) $\boldsymbol{P}\{X < 10\} = F(9) = \boxed{0.968}$

(b) $\boldsymbol{P}\{X = 10\} = F(10) - F(9) = 0.986 - 0.968 = \boxed{0.018}$

3.32 (a) The number X of computer shutdowns during one year (12 months) averages

$$\lambda = (0.25)(12) = 3 \text{ shutdowns per year.}$$

From Table A3 with $\lambda = 3$,

$$\boldsymbol{P}\{X \geq 3\} = 1 - F_X(2) = 1 - 0.423 = \boxed{0.577}$$

(b) Let Y be the number of months with exactly 1 computer shutdown. Each month is a Bernoulli trial because it either has exactly 1 shutdown or not. Hence, Y has Binomial distribution with $n = 12$ and

$$\begin{aligned} p &= \boldsymbol{P}\{\text{ 1 shutdown in one month }\} = e^{-0.25}\frac{0.25^1}{1!} \\ &= 0.25\,e^{-0.25} = 0.1947. \end{aligned}$$

Then,

$$\begin{aligned} \boldsymbol{P}\{Y \geq 3\} &= 1 - P_Y(0) - P_Y(1) - P_Y(2) \\ &= 1 - (1 - 0.1947)^{12} - (12)(0.1947)(1 - 0.1947)^{11} \\ &\quad - \frac{(12)(11)}{2}(0.1947)^2(1 - 0.1947)^{10} = \boxed{0.4228} \end{aligned}$$

3.33 The number of files X affected by the virus has Binomial distribution with $n = 250$ (large) and $p = 0.032$ (small), which is approximately Poisson with

$$\lambda = np = 8.$$

From Table A3,

$$\boldsymbol{P}\{X > 7\} = 1 - F(7) = \boxed{0.547}$$

3.34 Let X be the number of accidents, and T be the event {thunderstorm}.

During the thunderstorm,

$$\boldsymbol{P}\{X = 7 \mid T\} = F(7) - F(6) = 0.220 - 0.130 = 0.09.$$

When there is no thunderstorm,

$$\boldsymbol{P}\left\{X = 7 \mid \bar{T}\right\} = F(7) - F(6) = 0.949 - 0.889 = 0.06$$

(from Table A3 with parameters $\lambda = 10$ and $\lambda = 4$).

By the Bayes Rule,

$$\begin{aligned}
\boldsymbol{P}\{T \mid X=7\} &= \frac{\boldsymbol{P}\{X=7 \mid T\}\boldsymbol{P}\{T\}}{\boldsymbol{P}\{X=7 \mid T\}\boldsymbol{P}\{T\} + \boldsymbol{P}\left\{X=7 \mid \bar{T}\right\}\boldsymbol{P}\left\{\bar{T}\right\}} \\
&= \frac{(0.09)(0.6)}{(0.09)(0.6)+(0.06)(1-0.6)} = \boxed{9/13 \text{ or } 0.6923}
\end{aligned}$$

Given that there were 7 accidents yesterday, the (conditional) probability that there was a thunderstorm is 0.6923.

3.35 At any time, the number of terminals X that are ready to transmit has Binomial distribution with $n = 10$ and $p = 0.7$. From Table A2,

$$\boldsymbol{P}\{X=6\} = F(6) - F(5) = 0.3504 - 0.1503 = \boxed{0.2001}$$

3.36 We need $\boldsymbol{P}\{X > 4\}$, where X is the number of breakdowns during 21 weeks. This is the number of rare events, averaging 1 per 3 weeks, or 7 per 21 weeks. Thus, X is Poisson with $\lambda = 7$, and from Table A3,

$$\boldsymbol{P}\{X > 4\} = 1 - F(4) = 1 - 0.173 = \boxed{0.827}$$

3.37 (a)

$$\begin{aligned}
\mathrm{Var}(X) &= \mathbf{E}\{X - \mathbf{E}(X)\}^2 = \mathbf{E}\left\{X - 2X\,\mathbf{E}(X) + \mathbf{E}^2(X)\right\} \\
&= \mathbf{E}(X^2) - 2\,\mathbf{E}(X)\,\mathbf{E}(X) + \mathbf{E}\,\mathbf{E}^2(X) \\
&= \mathbf{E}(X^2) - 2\,\mathbf{E}^2(X) + \mathbf{E}^2(X) = \mathbf{E}(X^2) - \mathbf{E}^2(X)
\end{aligned}$$

(b)

$$\begin{aligned}
\mathrm{Cov}(X,Y) &= \mathbf{E}\{(X - \mathbf{E}(X))(Y - \mathbf{E}(Y))\} \\
&= \mathbf{E}\{XY - X\,\mathbf{E}(Y) - Y\,\mathbf{E}(X) + \mathbf{E}(X)\,\mathbf{E}(Y)\} \\
&= \mathbf{E}(XY) - \mathbf{E}(X)\,\mathbf{E}(Y) - \mathbf{E}(Y)\,\mathbf{E}(X) + \mathbf{E}\,\mathbf{E}(X)\,\mathbf{E}(Y) \\
&= \mathbf{E}(XY) - 2\,\mathbf{E}(X)\,\mathbf{E}(Y) + \mathbf{E}(X)\,\mathbf{E}(Y) \\
&= \mathbf{E}(XY) - \mathbf{E}(X)\,\mathbf{E}(Y)
\end{aligned}$$

3.38

$$\begin{aligned}
&\mathrm{Cov}(aX + bY + c, dZ + eW + f) \\
&\quad= \mathbf{E}\{((aX + bY + c) - \mathbf{E}(aX + bY + c)) \\
&\qquad\qquad \cdot ((dZ + eW + f) - \mathbf{E}(dZ + eW + f))\} \\
&\quad= \mathbf{E}\{(aX + bY + c - a\,\mathbf{E}(X) - b\,\mathbf{E}(Y) - c) \\
&\qquad\qquad \cdot (dZ + eW + f - d\,\mathbf{E}(Z) - e\,\mathbf{E}(W) - f)\} \\
&\quad= \mathbf{E}\{ad(X - \mathbf{E}(X))(Z - \mathbf{E}(Z)) + ae(X - \mathbf{E}(X))(W - \mathbf{E}(W)) \\
&\qquad\quad + bd(Y - \mathbf{E}(Y))(Z - \mathbf{E}(Z)) + be(Y - \mathbf{E}(Y))(W - \mathbf{E}(W))\} \\
&\quad= ad\,\mathrm{Cov}(X,Z) + ae\,\mathrm{Cov}(X,W) + bd\,\mathrm{Cov}(Y,Z) + be\,\mathrm{Cov}(Y,W)
\end{aligned}$$

Chapter 4

4.1 Find k from the condition $\int f(x)dx = 1$:

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{1}^{+\infty} \frac{k}{x^4}dx = -\left.\frac{k}{3x^3}\right|_{x=1}^{+\infty} = \frac{k}{3} = 1.$$

Hence, $\boxed{\text{k=3}}$.

Next, compute the cumulative distribution function:

$$F(x) = \int_{1}^{x} f(y)dy = \int_{1}^{x} \frac{3}{y^4}dy = -\left.\frac{1}{y^3}\right|_{y=1}^{x} = \boxed{1 - \frac{1}{x^3} \text{ for } x \geq 1}$$

($F(x) = 0$ for $x < 1$).

Then

$$\boldsymbol{P}\{X > 2\} = 1 - F(2) = 1 - \left(1 - \frac{1}{8}\right) = \boxed{1/8 \text{ or } 0.125}$$

4.2 (a) Find C from the condition $\int f(x)dx = 1$:

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{0}^{10} C(10-x)^2 dx = -\left.\frac{C(10-x)^3}{3}\right|_{x=0}^{10} = \frac{1000\,C}{3} = 1.$$

Hence, $\boxed{C = 0.003}$.

(b)
$$\begin{aligned} \boldsymbol{P}\{1 < X < 2\} &= \int_{1}^{2} 0.003(10-y)^2 dy = -\left.\frac{0.003(10-y)^3}{3}\right|_{y=1}^{2} \\ &= -0.001(8^3 - 9^3) = \boxed{0.217} \end{aligned}$$

4.3 For any density, $\int f(x)dx = 1$. Therefore,

$$\int f(x)dx = \int_{0}^{1} k(1-x^3)dx = k\left(x - \frac{x^4}{4}\right)\Big|_{0}^{1} = k\left(1 - \frac{1}{4}\right) = \frac{3}{4}k = 1,$$

hence, $\boxed{\text{k} = 4/3}$. Then,

$$\begin{aligned} \boldsymbol{P}\{X < 1/2\} &= \frac{4}{3}\int_{0}^{1/2}(1-x^3)dx = \frac{4}{3}\left(x - \frac{x^4}{4}\right)\Big|_{0}^{1/2} = \frac{4}{3}\left(\frac{1}{2} - \frac{1}{64}\right) \\ &= \boxed{\frac{31}{48} \text{ or } 0.6458} \end{aligned}$$

4.4 (a) Find C from the condition $\iint f(x,y)dx\,dy = 1$:

$$\iint f(x,y)dx\,dy = \int_{y=0}^{1}\int_{x=-1}^{1} C(x^2+y)dx\,dy = C\int_{y=0}^{1}\left(\frac{x^3}{3} + xy\right)\Big|_{x=-1}^{1} dy$$

$$= C \int_{y=0}^{1} \left(\frac{2}{3} + 2y \right) dy = C \left(\frac{2}{3}y + y^2 \right) \Big|_{y=0}^{1} = \frac{5C}{3} = 1.$$

Hence, $\boxed{C = 3/5 \text{ or } 0.6}$.

(b) $$\boldsymbol{P}\{Y < 0.6\} = \int_{y=0}^{0.6} \int_{x=-1}^{1} 0.6(x^2 + y) dx\, dy \qquad \text{[similarly to (a)]}$$

$$= 0.6 \int_{y=0}^{0.6} \left(\frac{2}{3} + 2y \right) dy = 0.6 \left(\frac{2}{3}y + y^2 \right) \Big|_{y=0}^{0.6} = 0.6(0.4 + 0.6^2) = \boxed{0.456}$$

Computing the conditional probability $P\{Y < 0.6 \mid X = 0.5\}$ directly by formula (2.7) yields an indeterminate expression

$$P\{Y < 0.6 \mid X = 0.5\} = \frac{P\{Y < 0.6 \ \cap \ X = 0.5\}}{\boldsymbol{P}\{X = 0.5\}} = \frac{0}{0} = \ldots ?$$

because X is a continuous variable.

We can resolve this by using the L'Hospital Rule,

$$P\{Y < 0.6 \mid X = 0.5\} = \lim_{\varepsilon \to 0} \boldsymbol{P}\{Y < 0.6 \mid 0.5 \leq X \leq 0.5 + \varepsilon\}$$

$$= \lim_{\varepsilon \to 0} \frac{\boldsymbol{P}\{Y < 0.6 \ \cap \ 0.5 \leq X \leq 0.5 + \varepsilon\}}{\boldsymbol{P}\{0.5 \leq X \leq 0.5 + \varepsilon\}}$$

$$= \lim_{\varepsilon \to 0} \frac{\int_{y=0}^{0.6} \int_{x=0.5}^{0.5+\varepsilon} f(x,y) dx\, dy}{\int_{y=0}^{1} \int_{x=0.5}^{0.5+\varepsilon} f(x,y) dx\, dy} = \lim_{\varepsilon \to 0} \frac{\frac{d}{d\varepsilon} \int_{y=0}^{0.6} \int_{x=0.5}^{0.5+\varepsilon} f(x,y) dx\, dy}{\frac{d}{d\varepsilon} \int_{y=0}^{1} \int_{x=0.5}^{0.5+\varepsilon} f(x,y) dx\, dy}$$

$$= \lim_{\varepsilon \to 0} \frac{\int_{y=0}^{0.6} f(x,y) \Big|_{x=0.5+\varepsilon} dy}{\int_{y=0}^{1} f(x,y) \Big|_{x=0.5+\varepsilon} dy} = \left. \frac{\int_{y=0}^{0.6} f(x,y) dy}{\int_{y=0}^{1} f(x,y) dy} \right|_{x=0.5}$$

$$= \frac{\int_{y=0}^{0.6} C(0.5^2 + y) dy}{\int_{y=0}^{1} C(0.5^2 + y) dy} = \frac{0.5^2 \cdot 0.6 + 0.6^2/2}{0.5^2 \cdot 1 + 1^2/2} = \frac{0.33}{0.75} = \boxed{0.44}$$

By the way, $\frac{f(x,y)}{\int_0^1 f(x,y) dy} = \frac{f(x,y)}{f_X(x)} = f(y|x)$ is the *conditional density* of Y given $X = x$. Thus, to find the conditional probability of $Y < 0.6$, we integrated the conditional density.

4.5 (a) Find K from the condition $\int f(x) dx = 1$:

$$\int f(x) dx = \int_0^{10} (K - x/50) dx = Kx - \frac{x^2}{2 \cdot 50} \Big|_0^{10} = 10K - 1 = 1.$$

Solving for K, we get $K = \boxed{0.2}$.

(b) $\boldsymbol{P}\{X<5\} = \int_0^5 (0.2 - x/50)dx = 0.2x - \left.\frac{x^2}{2\cdot 50}\right|_0^5 = 1 - 0.25 = \boxed{0.75}$

$$\begin{aligned}\text{(c)}\quad \mathbf{E}(X) &= \int x f(x)dx = \int_0^{10} x(0.2 - x/50)dx = \left.\frac{0.2x^2}{2} - \frac{x^3}{3\cdot 50}\right|_0^{10} \\ &= 10 - \frac{20}{3} = \boxed{3\frac{1}{3} \text{ or } 3.3333 \text{ years}}\end{aligned}$$

4.6 Denote Exponential(λ) times for the 3 blocks by X_1, X_2, and X_3, and let

$$X = \max_i X_i$$

be the time it takes to compile the whole program. Find the cdf, then the pdf of X and use the latter to compute the expectation $\mathbf{E}(X)$.

For an Exponential (λ) time X_i,

$$\mathbf{E}(X_i) = 1/\lambda = 5 \text{ min},$$

therefore, $\lambda = 0.2 \text{ min}^{-1}$.

Compute the cdf of X,

$$\begin{aligned} F_X(x) &= \boldsymbol{P}\left\{\max_i X_i \le x\right\} = \boldsymbol{P}\left\{\bigcap_{i=1}^{3}(X_i \le x)\right\} \\ &= \prod_{i=1}^{3} \boldsymbol{P}\{X_i \le x\} = \left(1 - e^{-0.2x}\right)^3, \quad x > 0. \end{aligned}$$

Differentiate it to find the pdf,

$$f_X(x) = F_X'(x) = 0.6\left(1 - e^{-0.2x}\right)^2 e^{-0.2x}, \quad x > 0.$$

Now we can compute $\mathbf{E}(X)$ as

$$\begin{aligned} \mathbf{E}(X) &= \int x\, f_X(x)dx = 0.6\int_0^\infty x\left(1 - e^{-0.2x}\right)^2 e^{-0.2x}dx \\ &= 0.6\int_0^\infty \left(xe^{-0.2x} - 2xe^{-0.4x} + xe^{-0.6x}\right)dx \\ &= 0.6\left(\frac{\Gamma(2)}{0.2^2} - \frac{2\Gamma(2)}{0.4^2} + \frac{\Gamma(2)}{0.6^2}\right) \\ &= 15 - 7.5 + 1\frac{2}{3} = \boxed{\frac{55}{6} \text{ or } 9.17 \text{ minutes}} \end{aligned}$$

We used the gamma-function ($\Gamma(2) = 1! = 1$), but one can also take the three integrals by parts.

4.7 The time it takes to process print 1 job is Exponential(λ). It has expectation $\mathbf{E}(X) = 1/\lambda = 12$ seconds. Hence, $\lambda = 1/12 \text{ sec}^{-1}$.

The time T until the job is finished is the sum of 3 Exponential times (the remaining time of the currently active job is also Exponential because of the memoryless property). Thus, T has Gamma distribution with parameters $\alpha = 3$ and $\lambda = 1/12$.

By the Gamma-Poisson formula,

$$\boldsymbol{P}\{T < 60 \text{ seconds}\} = \boldsymbol{P}\{X \geq 3\},$$

where X is a Poisson variable with parameter $\lambda t = (1/12)(60) = 5$. Then

$$\boldsymbol{P}\{T < 60\} = \boldsymbol{P}\{X \geq 3\} = 1 - F(2) = 1 - 0.125 = \boxed{0.875}$$

from Table A3 with parameter 5.

4.8 Notice that 6 months = 1/2 years. Using the Gamma-Poisson formula with a Gamma(2,2) variable X and a Poisson($2 \cdot \frac{1}{2} = 1$) variable Y,

$$\boldsymbol{P}\{X < 1/2\} = \boldsymbol{P}\{Y \geq 2\} = 1 - F(1) = 1 - e^{-1} - 1e^{-1} = 1 - 0.736 = \boxed{0.264},$$

by the formula of Poisson distribution or by Table A3.

Alternatively, one can take an integral,

$$\boldsymbol{P}\{X < 1/2\} = \int_0^{1/2} f(x)dx = \int_0^{1/2} 4xe^{-2x}dx = 1 - e^{-1} - e^{-1}$$

by parts.

4.9 The time T until the third breakdown has Gamma distribution with parameters $\alpha = 3$ and $\lambda = 1/5$ months^{-1}.

(a) By the Gamma-Poisson formula with a Poisson($\lambda t = 1/5 \cdot 9 = 1.8$) variable X and Table A3,

$$\boldsymbol{P}\{T \leq 9\} = \boldsymbol{P}\{X \geq 3\} = 1 - F_X(2) = 1 - 0.731 = \boxed{0.269}$$

(b) $$\boldsymbol{P}\{T > 16 \mid T > 12\} = \frac{\boldsymbol{P}\{T > 16 \cap T > 12\}}{\boldsymbol{P}\{T > 12\}} = \frac{\boldsymbol{P}\{T > 16\}}{\boldsymbol{P}\{T > 12\}}$$

$$= \frac{\boldsymbol{P}\{X_1 < 3\}}{\boldsymbol{P}\{X_2 < 3\}} = \frac{e^{-3.2}(1 + 3.2 + 3.2^2/2)}{0.570} = \boxed{0.666}$$

by the Gamma-Poisson formula, the formula of Poisson pmf, and Table A3, where X_1 has Poisson distribution with parameter $(1/5)(16) = 3.2$ and X_2 has Poisson distribution with parameter $(1/5)(12) = 2.4$.

4.10 Denote the events:

$$\begin{aligned} A &= \{\text{ first specialist processes the order }\} \\ B &= \{\text{ second specialist processes the order }\} \\ C &= \{\text{ the order takes more than 30 minutes (1/2 hr)}\} \end{aligned}$$

Using Exponential distributions with parameters $\lambda_1 = 3$ hrs^{-1} and $\lambda_2 = 2$

hrs^{-1},

$$\boldsymbol{P}\{C \mid A\} = e^{-(3)(1/2)} = e^{-1.5} \text{ and } \boldsymbol{P}\{C \mid B\} = e^{-(2)(1/2)} = e^{-1}$$

Also, $\boldsymbol{P}\{A\} = 0.6$ and $\boldsymbol{P}\{B\} = 0.4$. By the Bayes Rule,

$$\boldsymbol{P}\{A \mid C\} = \frac{0.6\boldsymbol{P}\{C \mid A\}}{0.6\boldsymbol{P}\{C \mid A\} + 0.4\boldsymbol{P}\{C \mid B\}} = \frac{0.6e^{-1.5}}{0.6e^{-1.5} + 0.4e^{-1}} = \boxed{0.4764}$$

4.11 Both lifetimes T_A and T_B have Exponential distribution with parameter $\lambda = 1/\mathbf{E}(T) = 0.1$ years^{-1}.

(a) $\boldsymbol{P}\{T_A > 10 \cup T_B > 10\} = 1 - \boldsymbol{P}\{T_A \leq 10 \cap T_B \leq 10\}$

$$= 1 - \left(1 - e^{-(10)(0.1)}\right)\left(1 - e^{-(10)(0.1)}\right) = \boxed{0.600}$$

(b) The lifetime of the satellite $T = \max\{T_A, T_B\}$ has a cdf

$$\begin{aligned} F_T(t) &= \boldsymbol{P}\{\max\{T_A, T_B\} \leq t\} = F_{T_A}(t)F_{T_B}(t) \\ &= \left(1 - e^{-0.1t}\right)^2 = 1 - 2e^{-0.1t} + e^{-0.2t} \end{aligned}$$

Differentiating, we find the pdf

$$f_T(t) = F_T'(t) = 0.2e^{-0.1t} - 0.2e^{-0.2t}$$

The expected lifetime is

$$\begin{aligned} \mathbf{E}(T) &= \int t f_T(t)dt = \int_0^\infty \left(0.2te^{-0.1t} - 0.2te^{-0.2t}\right) dt \\ &= 0.2\left(\frac{\Gamma(2)}{0.1^2} - \frac{\Gamma(2)}{0.2^2}\right) = 0.2(100 - 25) = \boxed{15 \text{ years}} \end{aligned}$$

4.12 The time T it takes to process a package is a sum of 5 independent Exponential times. Therefore, it has Gamma distribution with parameters $\alpha = 5$ and $\lambda = 1/2$ min^{-1}.

By the Gamma-Poisson formula,

$$\boldsymbol{P}\{T < 8\} = \boldsymbol{P}\{X \geq 5\} = 1 - F_X(4) = 1 - 0.629 = \boxed{0.371}$$

from Table A3, where X is Poisson with parameter $\lambda t = (1/2)(8) = 4$.

4.13 The time T it takes to download 3 files is a sum of 3 independent Exponential variables. Therefore, it has Gamma distribution with parameters $\alpha = 3$ and $\lambda = 1/25$ sec^{-1}.

By the Gamma-Poisson formula,

$$\boldsymbol{P}\{T > 70\} = \boldsymbol{P}\{X < 3\} = F_X(2) = \boxed{0.469}$$

from Table A3, where X is Poisson with parameter $\lambda t = (1/25)(70) = 2.8$.

4.14 (a) For this Gamma distribution, $\mathbf{E}(X) = \alpha/\lambda = 20$ and $\mathrm{Var}(X) = \alpha/\lambda^2 = 10^2 = 100$. Solving these equations for α and λ, we get

$$\alpha = \mathbf{E}^2(X)/\mathrm{Var}(X) = 20^2/100 = \boxed{4}$$

and

$$\lambda = \mathbf{E}(X)/\mathrm{Var}(X) = 20/100 = \boxed{0.2}$$

(b) Use the Poisson-Gamma formula with a Poisson variable Y that has a parameter $\lambda(15) = 3$. From Table A3,

$$\boldsymbol{P}\{X < 15\} = \boldsymbol{P}\{Y \geq 4\} = 1 - \boldsymbol{P}\{Y \leq 3\} = 1 - 0.647 = \boxed{0.353}$$

4.15 (a) Component A works at least 2 years with probability

$$\boldsymbol{P}\{A\} = \boldsymbol{P}\{T_A \geq 2\} = \boldsymbol{P}\{X < 3\} = F_X(2) = 0.677$$

where T_A has Gamma(3,1) distribution, and X is Poisson with parameter $\lambda(2) = (1)(2) = 2$.
Component B works at least 2 years with probability

$$\boldsymbol{P}\{B\} = \boldsymbol{P}\{X_B \geq 2\} = \boldsymbol{P}\{Y < 2\} = F_Y(1) = 0.092,$$

where T_B has Gamma(2,2) distribution, and Y is Poisson with parameter $\lambda(2) = (2)(2) = 4$.
Then, by independence of A and B,

$$\begin{aligned} \boldsymbol{P}\left\{\begin{array}{c}\text{system works}\\ \text{at least 2 years}\end{array}\right\} &= \boldsymbol{P}\{A \cap B\} = \boldsymbol{P}\{A\}\,\boldsymbol{P}\{B\} \\ &= (.677)(.092) = \boxed{0.062} \end{aligned}$$

(b) $\boldsymbol{P}\{\text{ B failed } \cap \text{ A worked } \mid \text{ system failed }\}$

$$\begin{aligned} &= \frac{\boldsymbol{P}\{\text{ B failed } \cap \text{ A worked } \cap \text{ system failed }\}}{\boldsymbol{P}\{\text{ system failed }\}} \\ &= \frac{\boldsymbol{P}\{\text{ B failed } \cap \text{ A worked }\}}{1 - \boldsymbol{P}\{\text{ system worked }\}} \\ &= \frac{(1 - 0.092)(0.677)}{1 - 0.062} = \boxed{0.655} \end{aligned}$$

4.16 Apply the Central Limit Theorem. A continuity correction is not needed because the lifetime is already a continuous random variable.

$$\begin{aligned} &\boldsymbol{P}\left\{\frac{S_{400}}{400} < 5012\right\} = \boldsymbol{P}\{S_{400} < (400)(5012)\} \\ &= \boldsymbol{P}\left\{Z < \frac{(400)(5012) - 400\mu}{\sigma\sqrt{400}}\right\} \approx \Phi\left(\frac{5012 - 5000}{100/\sqrt{400}}\right) \\ &= \Phi(2.4) = \boxed{0.9918} \end{aligned}$$

from Table A4.

4.17 Here $n = 82$, $\mu = 15$ sec, $\sigma = \sqrt{16} = 4$ sec. Convert the time into seconds, 20 min = 1200 sec.

By the Central Limit Theorem,

$$\boldsymbol{P}\{X_1 + \ldots + X_{82} < 1200\} = \boldsymbol{P}\left\{\frac{X_1 + \ldots + X_{82} - n\mu}{\sigma\sqrt{n}} < \frac{1200 - n\mu}{\sigma\sqrt{n}}\right\}$$
$$= \boldsymbol{P}\left\{Z < \frac{1200 - (82)(15)}{4\sqrt{82}}\right\} = \boldsymbol{P}\{Z < -0.83\} \approx \Phi(-0.83) = \boxed{0.2033}$$

Continuity correction is not needed, because X_i, the times, are already continuous variables.

4.18 (a) The number X of defective chips in the sample has Binomial distribution with parameters $n = 400$ and $p = 0.06$, expectation $\mu = np = 24$, and standard deviation $\sigma = \sqrt{np(1-p)} = 4.75$.

By the Central Limit Theorem (with a suitable continuity correction),

$$\boldsymbol{P}\{20 \le X \le 25\} = \boldsymbol{P}\{19.5 \le X \le 25.5\}$$
$$= \boldsymbol{P}\left\{\frac{19.5 - 24}{4.75} \le Z \le \frac{25.5 - 24}{4.75}\right\} = \boldsymbol{P}\{-0.95 \le Z \le 0.32\}$$
$$\approx \Phi(0.32) - \Phi(-0.95) = 0.6255 - 0.1711 = \boxed{0.4544}$$

from Table A4.

(b) The number Y of inspectors that find between 20 and 25 defective chips in their samples has Binomial distribution with parameters $n = 40$ and $p = 0.4544$, computed in (a), expectation $\mu = np = 18.176$, and standard deviation $\sigma = \sqrt{np(1-p)} = 3.149$. Then

$$\boldsymbol{P}\{X \ge 8\} = \boldsymbol{P}\{X \ge 7.5\} = \boldsymbol{P}\left\{Z \ge \frac{7.5 - 18.176}{3.149}\right\}$$
$$= \boldsymbol{P}\{Z \ge -3.39\} \approx \Phi(3.39) = \boxed{0.9997}$$

from Table A4.

4.19 Here $n = 80$, $\mu = 0.6$, and $\sigma = 0.4$.

$$\boldsymbol{P}\{47 \le S_{80} \le 50\} = \boldsymbol{P}\left\{\frac{47 - n\mu}{\sigma\sqrt{n}} \le Z \le \frac{50 - n\mu}{\sigma\sqrt{n}}\right\}$$
$$= \boldsymbol{P}\left\{\frac{47 - 48}{0.4\sqrt{80}} \le Z \le \frac{50 - 48}{0.4\sqrt{80}}\right\} \approx \Phi(0.56) - \Phi(-0.28)$$
$$= 0.7123 - 0.3897 = \boxed{0.3226}$$

from Table A4.

4.20 The number of damaged files has Binomial distribution with $n = 2400$,

$p = 0.35$, $\mu = np = 840$, and $\sigma = \sqrt{np(1-p)} = 23.37$. Then

$$\begin{aligned} \boldsymbol{P}\{800 \le X \le 850\} &= \boldsymbol{P}\{799.5 < X < 850.5\} \\ &= \boldsymbol{P}\left\{\frac{799.5-840}{23.37} < Z < \frac{850.5-840}{23.37}\right\} = \boldsymbol{P}\{-1.73 < Z < 0.45\} \\ &\approx \Phi(0.45) - \Phi(-1.73) = 0.6736 - 0.0418 = \boxed{0.6318} \end{aligned}$$

from Table A4.

4.21 Exponential variables $X_1, ..., X_{70}$ have $\mu = 1/\lambda = 0.2$ min and $\sigma = 1/\lambda = 0.2$ min. Using the Central Limit Theorem,

$$\begin{aligned} \boldsymbol{P}\{X_1 + \ldots + X_{70} < 12\} &= \boldsymbol{P}\left\{\frac{X_1 + \ldots + X_{70} - n\mu}{\sigma\sqrt{n}} < \frac{12 - n\mu}{\sigma\sqrt{n}}\right\} \\ &= \boldsymbol{P}\left\{Z < \frac{12 - 70 \cdot 0.2}{0.2\sqrt{70}}\right\} = \boldsymbol{P}\{Z < -1.20\} \approx \boxed{0.1151} \end{aligned}$$

from Table A4.

4.22 Denote the events $A = \{\text{Printer I}\}$ and $B = \{T < 1\}$, where T is the printing time. We know that $\boldsymbol{P}\{A\} = 0.4$, $\boldsymbol{P}\{\bar{A}\} = 0.6$.

For Printer I with Exponential time, $E(T) = 2 = 1/\lambda$, hence $\lambda = 0.5$ and

$$\boldsymbol{P}\{B \mid A\} = \boldsymbol{P}\{T < 1 \mid A\} = 1 - e^{-(0.5)(1)} = 0.393.$$

For Printer II with Uniform time, the density of T is $f(t) = 1/5$ for t between 0 and 5, and

$$\boldsymbol{P}\{B \mid \bar{A}\} = \boldsymbol{P}\{T < 1 \mid \bar{A}\} = \int_0^1 \frac{1}{5}\, dt = 0.2$$

(or, instead of integration, draw the graph of this density and find the area under it from 0 to 1; it is the area of a rectangle).

By the Bayes Rule,

$$\begin{aligned} \boldsymbol{P}\{A \mid B\} &= \frac{\boldsymbol{P}\{B \mid A\}\boldsymbol{P}\{A\}}{\boldsymbol{P}\{B \mid A\}\boldsymbol{P}\{A\} + \boldsymbol{P}\{B \mid \bar{A}\}\boldsymbol{P}\{\bar{A}\}} \\ &= \frac{(0.393)(0.4)}{(0.393)(0.4) + (0.2)(0.6)} = \boxed{0.567} \end{aligned}$$

4.23 Let T be the connection time, $A = \{\text{Line I}\}$, and $B = \{\text{Line II}\}$, where $\boldsymbol{P}\{A\} = 0.8$ and $\boldsymbol{P}\{B\} = 0.2$.

For Line I,

$$\boldsymbol{P}\{T > 0.5 \text{ min} \mid A\} = \boldsymbol{P}\{X < 3\} = F_X(2) = 0.92,$$

using Table A3 and the Gamma-Poisson formula with a Gamma(3,2) variable T and a Poisson$(2 \cdot 0.5 = 1)$ variable X.

For Line II,

$$\boldsymbol{P}\{T > 30\,\text{sec} \mid B\} = \int_{20}^{30} \frac{1}{50-20}\,dt = 1/3 \text{ or } 0.3333.$$

Then, by the Law of Total Probability,

$$\boldsymbol{P}\{T > 30\,\text{sec}\} = (0.8)(0.92) + (0.2)(0.3333) = \boxed{0.8027}$$

4.24 (a) We have $n = 68$, $\mu = 15$ sec, and $\sigma = \sqrt{11}$ sec. By the Central Limit Theorem,

$$\boldsymbol{P}\{S_{68} < 720\,\text{sec}\} = \boldsymbol{P}\left\{Z < \frac{720 - n\mu}{\sigma\sqrt{n}}\right\} = \boldsymbol{P}\left\{Z < \frac{720-(68)(15)}{\sqrt{11}\sqrt{68}}\right\}$$

$$\approx \Phi(-10.97) = \boxed{0.00 \text{ (practically 0)}}$$

(see the last line of Table A4)

(b) We are given that

$$\boldsymbol{P}\{S_N < 600\,\text{sec}\} = 0.95.$$

That is,

$$0.95 = \boldsymbol{P}\left\{Z < \frac{600 - N\mu}{\sigma\sqrt{N}}\right\} \approx \Phi\left(\frac{600-N\mu}{\sigma\sqrt{N}}\right).$$

On the other hand,

$$0.95 = \Phi(1.645)$$

from Table A4, because $\Phi(1.64) = 0.9495$ and $\Phi(1.65) = 0.9505$. Therefore,

$$\frac{600 - N\mu}{\sigma\sqrt{N}} = \frac{600 - 15N}{\sqrt{11N}} = 1.645$$

It remains to solve this equation for N:

$$(600 - 15N)^2 = 1.645^2(11N), \quad 360,000 - 18,000N + 225N^2 = 30N,$$

$$225N^2 - 18,030N + 360,000 = 0,$$

$$N = \frac{18,030 \pm \sqrt{18,030^2 - (4)(225)(360,000)}}{(2)(225)} = 40 \pm 2 = 38 \text{ or } 42.$$

Notice that $600 - 15N$ is positive, therefore, $N < 40$. Thus, the answer is $\boxed{38}$. The new version of the package requires 38 new files.

4.25 Let X and Y be the Uniform(2,8) arrival times of the two customers.

(a) The time T of the earlier arrival is

$$T = \min\{X, Y\}.$$

It has the cdf

$$F_T(t) = \boldsymbol{P}\{\min\{X,Y\} \le t\} = \boldsymbol{P}\{X \le t \cup Y \le t\}$$

$$= 1 - \boldsymbol{P}\{X > t\}\,\boldsymbol{P}\{Y > t\} = 1 - \left(\frac{8-t}{8-2}\right)^2 = -\frac{7}{9} + \frac{4t}{9} - \frac{t^2}{36}$$

for $2 \leq t \leq 8$.

The density of T is

$$f_T(t) = F_T'(t) = \frac{4}{9} - \frac{t}{18}$$

Then, the expected time of the first arrival is

$$\mathbf{E}(T) = \int t f_T(t)dt = \int_2^8 \left(\frac{4t}{9} - \frac{t^2}{18}\right) dt = \left(\frac{2t^2}{9} - \frac{t^3}{54}\right)\Bigg|_{t=2}^{t=8} = \boxed{4}$$

(b) Long solution. The time W of the later arrival is

$$W = \max\{X, Y\}.$$

It has the cdf

$$\begin{aligned} F_W(w) &= \boldsymbol{P}\{\max\{X,Y\} \leq w\} = \boldsymbol{P}\{X \leq w \cap Y \leq w\} \\ &= \boldsymbol{P}\{X \leq w\}\boldsymbol{P}\{Y \leq w\} = \left(\frac{w-2}{8-2}\right)^2 \end{aligned}$$

for $2 \leq w \leq 8$.

The density of W is

$$f_W(w) = F_W'(w) = \frac{w-2}{18}$$

Then, the expected time of the second arrival is

$$\mathbf{E}(W) = \int w f_W(w)dw = \int_2^8 \frac{w^2 - 2w}{18}\, dw = \left(\frac{w^3}{54} - \frac{w^2}{18}\right)\Bigg|_{w=2}^{w=8} = \boxed{6}$$

Short solution. We can use our result in (a) by noticing that

$$T + W = X + Y,$$

and for Uniform(2,6) variables X and Y,

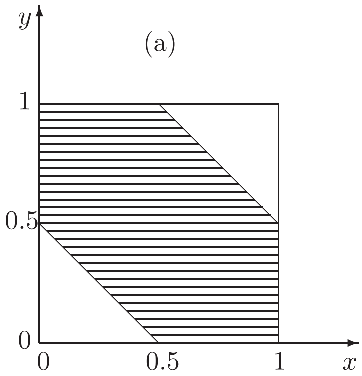
$$\mathbf{E}(X) = \mathbf{E}(Y) = 5.$$

Then

$$\mathbf{E}(W) = \mathbf{E}(X) + \mathbf{E}(Y) - \mathbf{E}(T) = 5 + 5 - 4 = \boxed{6}$$

Expected times of the earlier and the later arrivals are 4 pm and 6 pm. (By the way, the expected arrival time of each customer is 5 pm!)

4.26 Analytic solutions. For the Standard Uniform distribution of X and Y, $f_X(x) = 1$, $f_Y(y) = 1$, and $f_{(X,Y)}(x,y) = f_X(x)f_Y(y) = 1$ for $0 \leq x \leq 1$ and $0 \leq y \leq 1$.



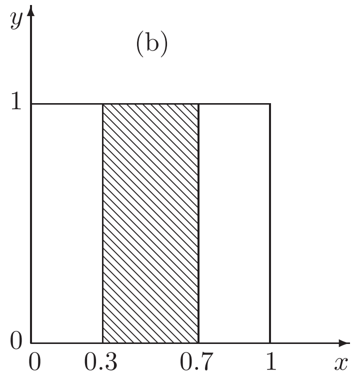


Figure 11.4 *Geometric solutions of Exercise 4.26*

(a) $$\boldsymbol{P}\{0.5 < (X+Y) < 1.5\} = \iint_{\{0.5<(X+Y)<1.5\}} f_{(X,Y)} dydx$$

$$= \int_{x=0}^{0.5}\int_{y=0.5-x}^{1} 1\,dydx + \int_{x=0.5}^{1}\int_{y=0}^{1.5-x} 1\,dydx$$

$$= \int_{x=0}^{0.5}(0.5+x)dx + \int_{x=0.5}^{1}(1.5-x)dx$$

$$= (0.5)(0.5) + \frac{0.5^2}{2} + (1.5)(0.5) - \frac{1^2-0.5^2}{2} = \boxed{0.75}$$

(b) Since X and Y are independent,

$$\boldsymbol{P}\{0.3 < X < 0.7 \mid Y > 0.5\} = \boldsymbol{P}\{0.3 < X < 0.7\} = \int_{0.3}^{0.7} 1\,dx = \boxed{0.4}$$

<u>Geometric solutions</u>. The joint density of X and Y equals 1 in the square where $0 \le x \le 1$ and $0 \le y \le 1$. Therefore, probabilities in (a) and (b) can be found as areas of the corresponding regions, as in Figure 11.4.

(a) $$\boldsymbol{P}\{0.5 < (X+Y) < 1.5\} = 1 - \begin{array}{c}\text{area of two}\\ \text{triangles}\end{array} = 1 - 2\left(\frac{0.5^2}{2}\right) = \boxed{0.75}$$

(b) $$\boldsymbol{P}\{0.3 < X < 0.7 \mid Y > 0.5\} = \boldsymbol{P}\{0.3 < X < 0.7\}$$

$$= \begin{array}{c}\text{area of the}\\ \text{rectangle}\end{array} = \boxed{0.4}$$

4.27 For any integer $x, y \ge 0$,

$$\boldsymbol{P}\{X > x+y \mid X > y\} = \frac{\boldsymbol{P}\{X > x+y \cap X > y\}}{\boldsymbol{P}\{X > y\}} = \frac{\boldsymbol{P}\{X > x+y\}}{\boldsymbol{P}\{X > y\}}$$

$$= \frac{(1-p)^{x+y}}{(1-p)^y} = (1-p)^x = \boldsymbol{P}\{X > x\}$$

Chapter 5

5.1 First, generate a Uniform(0,1) random variable U.

Inverse Transform method. Find the cdf

$$F(x) = \int_0^x f(t)dt = 1.5 \int_0^x t^{1/2} dt = 1.5 \left(\frac{t^{3/2}}{3/2} \right) \Big|_{t=0}^{t=x} = x^{3/2}$$

Compute $X = F^{-1}(U)$ by solving the equation $F(X) = U$:

$$X^{3/2} = U, \quad \boxed{X = U^{2/3}}$$

For $U = 0.001$, we generate $X = (0.001)^{2/3} = \boxed{0.01}$

Method of rejection cannot be applied because it requires at least two generated Uniform variables.

5.2 Use Algorithm 5.2 to generate continuous variables is (a)-(g) and Algorithm 5.3 to generate discrete variables. Notice that solutions and answers to (a)-(g) are *not unique*.

(a) $X = -\frac{1}{\lambda}\ln(U) = \boxed{-\frac{1}{2.5}\ln(U) = 0.3693}$

(b) Let

$$X = \begin{cases} 1 & \text{if} \quad U < 0.77 \\ 0 & \text{if} \quad U \geq 0.77 \end{cases}$$

For $U = 0.3972$, we generate $X = \boxed{1}$.

(c) The cdf F of the Binomial(15,0.4) distribution of X is given in Table A2. Let

$$X = \min\{x \geq 0 \text{ such that } F(x) > U\}.$$

For $U = 0.3972$, we have $F(4) = 0.2173 < U < F(5) = 0.4032$. Therefore, we let $X = \boxed{5}$.

(d) Compute $F(0) = 0.2$, $F(2) = 0.6$, $F(7) = 0.9$, and $F(11) = 1.0$, and let

$$X = \min\{x \in \{0, 2, 7, 11\} \text{ such that } F(x) > U\}.$$

For $U = 0.3972$, we let $X = \boxed{2}$.

(e) Compute the cdf

$$F(x) = \int_0^x 3t^2 dt = x^3.$$

Then let

$$X = F^{-1}(U) = \boxed{\sqrt[3]{U} = 0.7351}$$

(f) Compute the cdf

$$F(x) = \int_{-1}^x 1.5t^2 dt = 0.5t^3 \Big|_{t=-1}^{t=x} = 0.5\left(x^3 + 1\right).$$

Then solve the equation $F(X) = U$ and let

$$X = F^{-1}(U) = \boxed{\sqrt[3]{2U-1} = -0.5902}$$

(g) Compute the cdf

$$F(x) = \frac{1}{12}\int_0^x t^{1/3}dt = \frac{1}{16}x^{4/3}$$

Then let

$$X = F^{-1}(U) = (16U)^{3/4} = \boxed{8\,U^{3/4} = 4.0026}$$

It is good to verify that each generated variable belongs to the range of possible values of X.

5.3 *Inverse Transform method*. Compute the cdf

$$F(x) = \frac{1}{2}\int_{-1}^x (1+t)dt = \frac{1}{4}(1+t)^2\Big|_{t=-1}^{t=x} = \frac{1}{4}(1+x)^2.$$

Then let

$$X = F^{-1}(U) = \boxed{\sqrt{4U} - 1}$$

Start reading Table A1 from "randomly selected" 11th and 5th column. For $U = 0.0841$, we get $X = -0.4200$.

Rejection method. Select a bounding box $[-1,1] \times [0,1]$. Continue reading Table A1 until we find Uniform variables $-1 < X < 1$ and $0 < Y < 1$ such that $Y < f(X)$.

The first pair of Standard Uniform numbers is $(U_x, U_y) = (0.9342, 0.3759)$. Let $Y = U_y = 0.3759$. At the same time, convert U_x to a Uniform(-1,1) variable

$$X = 2U - 1.$$

That is, let $X = 2(0.9342) - 1 = 0.8684$. We have

$$f(X) = \frac{1}{2}(1 + 0.8684) = 0.9342 > Y.$$

The point is accepted, so the generated value of X is 0.8684.

5.4 Compute the cdf

$$F(x) = \frac{2}{9}\int_{-1}^x (1+t)dt = \frac{1}{9}(1+t)^2\Big|_{t=-1}^{t=x} = \frac{1}{9}(1+x)^2.$$

Then let

$$X = F^{-1}(U) = \sqrt{9U} - 1 = \boxed{3\sqrt{U} - 1 = 0.4685}$$

5.5 Compute the cdf

$$F(x) = \frac{1}{3}\int_0^x t^2 dt = \frac{1}{9}t^3\Big|_{t=-1}^{t=x} = \frac{1}{9}\left(x^3+1\right).$$

Then let

$$X = F^{-1}(U) = \boxed{\sqrt[3]{9U-1} = 1.8371}$$

5.6 We need 2 Standard Uniform variables U and V. The first variable will be used to determine the mechanic, then the second variable will be used to generate the service time.

Generating a random mechanic according to the given probabilities, we let the first mechanic change your oil filter if $U < 1/5$, otherwise the service will be performed by the second mechanic.

Then generate the Exponential service time as

$$X = \begin{cases} -\frac{1}{5}\ln(V) & \text{if} \quad U < 1/5 \\ -\frac{1}{20}\ln(V) & \text{if} \quad U \geq 1/5 \end{cases}$$

5.7 (a) Generate two sequences of Standard Uniform variables,

$$U_1, U_2, \ldots, U_N \text{ and } V_1, V_2, \ldots, U_N.$$

For each $i = 1, \ldots, N$, let

$$\begin{aligned} X_i &= \min\{x \geq 0 \text{ such that } F_{\lambda=3}(x) > U_i\} \\ &= \min\left\{x \geq 0 \text{ such that } e^{-3}\sum_{k=0}^{x}\frac{3^k}{k!} > U_i\right\} \end{aligned}$$

and

$$\begin{aligned} Y_i &= \min\{y \geq 0 \text{ such that } F_{\lambda=5}(y) > V_i\} \\ &= \min\left\{y \geq 0 \text{ such that } e^{-5}\sum_{k=0}^{y}\frac{5^k}{k!} > V_i\right\}. \end{aligned}$$

Then estimate $\boldsymbol{P}\{X > Y\}$ by the proportion of i such that $X_i > Y_i$.

(b) Partition the interval $[0,1]$ into 52 subintervals,

$$A_1 = [0,\ 1/52], A_2 = [1/52,\ 2/52], \ldots, A_{52} = [51/52,\ 1].$$

Assign each card to an interval. Then generate a long sequence of Standard Uniform variables $U_1, \ldots, U_N$.

In each poker hand, the first card is chosen according to the first Uniform variable U_1. If $U_1 \in A_j$, then card j is generated. The second card is generated according to U_2. However, if U_2 also belongs to A_j, then use U_3, U_4, etc., until a different card is generated.

Having generated a large number of poker hands, estimate the probability of a royal-flush by a proportion of generated hands with a royal-flush.

(c) Generate two sequences of Standard Uniform variables U_i and V_i for $i = 1, \ldots, N$ and let

$$X_i = \begin{cases} -\dfrac{1}{5}\ln(V_i) & \text{if} \quad U_i < 1/5 \\ -\dfrac{1}{20}\ln(V_i) & \text{if} \quad U_i \geq 1/5 \end{cases}$$

as in Exercise 5.6. Then estimate the probability by the proportion of generated service times X_i that exceed 35.

(d) For $\alpha = 0.05$ and $\varepsilon = 0.005$, the size

$$N \geq 0.25\left(\frac{z_{\alpha/2}}{\varepsilon}\right)^2 = 0.25\left(\frac{1.96}{0.005}\right)^2 = \boxed{38{,}416}$$

is sufficient. One can reduce this size by using a preliminary estimate, or an upper bound for the probability of interest. For example, in (b), the probability of a royal flush is less than 0.05, therefore, the size

$$N \geq (0.05)(1 - 0.05)\left(\frac{z_{\alpha/2}}{\varepsilon}\right)^2 = (0.05)(0.95)\left(\frac{1.96}{0.005}\right)^2 = 7,299.04$$

is sufficient. That is, in (b), we can use 7,300 Monte Carlo runs.

5.11 Let us derive the cdf of $X = F^{-1}(U)$,

$$\begin{aligned} F_X(x) &= \boldsymbol{P}\{X \leq x\} = \boldsymbol{P}\left\{F^{-1}(U) \leq x\right\} = \boldsymbol{P}\left\{F(F^{-1}(U)) \leq F(x)\right\} \\ &= \boldsymbol{P}\{U \leq F(x)\} = F(x), \end{aligned}$$

where we used monotonicity of the cdf F and the Standard Normal distribution of U. Thus, $F_X(x) = F(x)$.

5.12 If $x_0 < x_1 < x_2 < \ldots$, then we have

$$\begin{aligned} p_0 &= P(x_0) &&= F(x_0), \\ p_0 + p_1 &= P(x_0) + P(x_1) &&= F(x_1), \\ p_0 + p_1 + p_2 &= P(x_0) + P(x_1) + P(x_2) &&= F(x_2), \text{ etc.} \end{aligned}$$

Algorithm 5.3 takes the smallest value x with $F(x) > U$. It is such x_i that $F(x_i) > U$ whereas $F(x_{i-1}) \leq U$, which means

$$F(x_{i-1}) \leq U < F(x_i).$$

This places U into the subinterval A_i defined in Algorithm 5.1. According to this algorithm, $X = x_i$ is generated. Hence, both algorithms result in the same value x_i.

5.13 First, we notice that $X = -2\ln(U_1)$ has Exponential(0.5) distribution (Example 5.10). Then, we start with the Exponential variable X with pdf

$f(x) = \frac{1}{2}e^{-x/2}$ and the Uniform variable U_2 with pdf 1 and make the following substitutions of variables:

(A) For every u, let

$$x = \frac{z_1^2}{\cos^2(2\pi u)}; \quad \frac{dx}{dz_1} = \frac{2z_1}{\cos^2(2\pi u)}$$

(B) For every z_1, let

$$u = \frac{1}{2\pi}\tan^{-1}\left(\frac{z_2}{z_1}\right); \quad \frac{1}{\cos^2(2\pi u)} = 1 + \tan^2(2\pi u) = 1 + \left(\frac{z_2}{z_1}\right)^2;$$

$$\frac{du}{dz_2} = \frac{1}{2\pi z_1}\frac{1}{1+(z_2/z_1)^2}$$

The given probability can be computed as follows.

$$\begin{aligned}
&\boldsymbol{P}\left\{Z_1 \le a \,\cap\, Z_2 \le b\right\} = \boldsymbol{P}\left\{\sqrt{X}\cos(2\pi U_2) \le a \,\cap\, \sqrt{X}\sin(2\pi U_2) \le b\right\} \\
&= \iint_{\left\{\begin{array}{l}\sqrt{x}\cos(2\pi u) \le a \\ \sqrt{x}\sin(2\pi u) \le b\end{array}\right\}} \frac{1}{2}e^{-\frac{1}{2}x} \cdot 1 \cdot dx\, du \\
&\overset{(A)}{=} \iint_{\left\{\begin{array}{l}z_1 \le a \\ z_1\tan(2\pi u) \le b\end{array}\right\}} \frac{1}{2}e^{-\frac{1}{2}\frac{z_1^2}{\cos^2(2\pi u)}}\frac{2z_1}{\cos^2(2\pi u)}dz_1 du \\
&\overset{(B)}{=} \iint_{\left\{\begin{array}{l}z_1 \le a \\ z_2 \le b\end{array}\right\}} e^{-\frac{1}{2}z_1^2\left\{1+\left(\frac{z_2}{z_1}\right)\right\}^2} z_1\left\{1+\left(\frac{z_2}{z_1}\right)\right\}^2 \frac{1}{2\pi z_1}\frac{dz_1 dz_2}{1+(z_2/z_1)^2} \\
&= \iint_{\left\{\begin{array}{l}z_1 \le a \\ z_2 \le b\end{array}\right\}} \frac{1}{2\pi e^{-\frac{1}{2}\left(z_1^2+z_2^2\right)}}dz_1 dz_2 \\
&= \int_{-\infty}^{a}\frac{1}{\sqrt{2\pi}}\,e^{-\frac{1}{2}z_1^2}dz_1\int_{-\infty}^{b}\frac{1}{\sqrt{2\pi}}\,e^{-\frac{1}{2}z_2^2}dz_2 = \Phi(a)\Phi(b).
\end{aligned}$$

The joint cdf of Z_1 and Z_2 factors into the product of marginal Standard Normal cdfs. Hence, Z_1 and Z_2 are independent Standard Normal random variables.

5.14 Recall that $\mathbf{E}(\hat{p}) = p$ and $\mathrm{Std}(\hat{p}) = \sqrt{p(1-p)/N}$. According to (3.8),

$$\boldsymbol{P}\left\{|\hat{p}-p| > \varepsilon\right\} = \boldsymbol{P}\left\{\frac{|\hat{p}-p|}{\sqrt{p(1-p)/N}} > \frac{\varepsilon}{\sqrt{p(1-p)/N}}\right\} \le \frac{p(1-p)}{N\varepsilon^2}$$

Any $N \ge \dfrac{p(1-p)}{\alpha\varepsilon^2}$ satisfies the inequality

$$\frac{p(1-p)}{N\varepsilon^2} \le \alpha,$$

and therefore, guarantees

$$\boldsymbol{P}\left\{|\hat{p}-p| > \varepsilon\right\} \le \alpha.$$

If a good guess p^* of p is available, one can choose

$$N \geq \frac{p^*(1-p^*)}{\alpha\varepsilon^2}$$

Also, since $p(1-p) \leq 1/4$ for all $0 \leq p \leq 1$, any number

$$N = \frac{1}{4\alpha\varepsilon^2}$$

is sufficient.

5.15 According to Algorithm 5.4, we generate pairs of random variables ("trials") until we have a pair (X, Y) such that $Y \leq f(X)$ ("success"). The number of trials needed for the first success has Geometric(p) distribution, where

$$p = \boldsymbol{P}\{Y \leq f(X)\} = \frac{1}{(b-a)c}$$

Using the Uniform distribution of the point (X, Y) in the bounding rectangle $[a, b] \times [0, c]$, we computed the probability p as the ratio of areas of the region $\{(x, y) : y \leq f(x)\}$ (which equals 1 because $f(x)$ is the density) and the bounding rectangle.

5.16 Algorithm 5.5 with a bounding box $[a, b] \times [0, c]$ produces an estimator

$$\hat{\mathcal{I}} = c(b-a)\hat{p}$$

with $\hat{p}$ being the proportion of points (U_i, V_i) such that $V_i \leq g(U_i)$. Variance of $\hat{\mathcal{I}}$ is

$$\begin{aligned}
\operatorname{Var}(\hat{\mathcal{I}}) &= c^2(b-a)^2 \operatorname{Var}(\hat{p}) = \frac{c^2(b-a)}{N} p(1-p) \\
&= \frac{c^2(b-a)}{N} \frac{\mathcal{I}}{c(b-a)} \left(1 - \frac{\mathcal{I}}{c(b-a)}\right) = \frac{\mathcal{I}}{N} \{c(b-a) - \mathcal{I}\},
\end{aligned}$$

where $p = \boldsymbol{P}\{V_i \leq g(U_i)\} = \mathcal{I}/\{c(b-a)\}$.

The improved Monte Carlo integration method estimates $\mathcal{I}$ by the average value of $(b-a)g(U_i)$. This estimator has variance

$$\begin{aligned}
\operatorname{Var}(\hat{\mathcal{I}}) &= \frac{(b-a)^2}{N} \operatorname{Var}(g(U_i)) = \frac{(b-a)^2}{N} \left(\mathbf{E}g^2(U_i) - \mathbf{E}^2 g(U_i)\right) \\
&\leq \frac{(b-a)^2}{N} \left(c\,\mathbf{E}g(U_i) - \mathbf{E}^2 g(U_i)\right) \\
&= \frac{(b-a)^2}{N} \left(c\frac{\mathcal{I}}{b-a} - \frac{\mathcal{I}^2}{(b-a)^2}\right) = \frac{\mathcal{I}}{N} \{c(b-a) - \mathcal{I}\}.
\end{aligned}$$

This derivation relied on $\mathbf{E}g(U_i) = \mathcal{I}/(b-a)$ and the inequality $g(U_i) \leq c$ that was assumed when we selected the bounding box.

Thus, the second method results in a lower (or equal) variance of $\hat{\mathcal{I}}$. Equality can hold only in trivial cases when $g(x) \equiv 0$ for all x or $g(x) \equiv c$ for all x. In all the other cases, the second method has a lower variance of $\hat{\mathcal{I}}$.

Chapter 6

6.1 (a) This Markov chain has 3 states: 0, 1, and 2 students in the lab. Transition probabilities are given as

$$p_{00} = p_{01} = p_{11} = p_{12} = p_{21} = p_{22} = 0.5.$$

The other transitions have probability 0:

$$p_{02} = p_{10} = p_{20} = 0.$$

Then, the transition probability matrix is

$$P = \begin{pmatrix} 0.5 & 0.5 & 0 \\ 0 & 0.5 & 0.5 \\ 0 & 0.5 & 0.5 \end{pmatrix}$$

(b) This Markov chain is not regular because transitions from states 1 and 2 to state 0 are not possible, and therefore, $p_{10}^{(h)} = p_{20}^{(h)} = 0$ for any number of steps h.

(c) Since the number of students cannot change from 1 or 2 to 0, there may be 0 students in the lab *only* if there were no students there each hour between 7 am and 10 am. Therefore,

$$p_{00}^{(3)} = p_{00}p_{00}p_{00} = 0.5^3 = \boxed{0.125}$$

6.2 (a) $P^{(2)} = P \cdot P = \begin{pmatrix} 0.4 & 0.6 \\ 0.6 & 0.4 \end{pmatrix} \begin{pmatrix} 0.4 & 0.6 \\ 0.6 & 0.4 \end{pmatrix} = \begin{pmatrix} 0.52 & 0.48 \\ 0.48 & 0.52 \end{pmatrix}$

(b) There are 3 transitions between 5:30 and 8:30, thus we need to compute $p_{11}^{(3)}$. The 3-step transition probability matrix is

$$P^{(3)} = P^{(2)}P = \begin{pmatrix} 0.52 & 0.48 \\ 0.48 & 0.52 \end{pmatrix} \begin{pmatrix} 0.4 & 0.6 \\ 0.6 & 0.4 \end{pmatrix} = \begin{pmatrix} 0.496 & ... \\ ... & ... \end{pmatrix}$$

(there is no need to compute the entire matrix). Hence, $p_{11}^{(3)} = \boxed{0.496}$.

6.3 (a) Let "black" be state 1 and "brown" be state 2. Arranging the given probabilities in a matrix, get

$$P = \begin{pmatrix} 0.6 & 0.4 \\ 0.2 & 0.8 \end{pmatrix}.$$

(b) There are 2 transitions between Rex and his grandchild.

$$P_{21}^{(2)} = p_{22}p_{21} + p_{21}p_{11} = (.8)(.2) + (.2)(.6) = \boxed{0.28}$$

6.4 (a) Let $X(n) = 1$ if the n-th light is green; $X(n) = 2$ if it is red.

$$P = \begin{pmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{pmatrix}$$

(b) $P_{12}^{(2)} = P_{12}^2 = (0.6)(0.4) + (0.4)(0.7) = \boxed{0.52}$

(c) We need to find $\pi_2 = \lim_{n\to\infty}\{X(n) = 2\}$. To find the steady-state distribution, solve the system $\pi = \pi P$ along with $\pi_1 + \pi_2 = 1$. We have

$$\pi_1 = 0.6\pi_1 + 0.3\pi_2, \qquad \pi_2 = 0.4\pi_1 + 0.7\pi_2,$$

from where $\pi_1 = (3/4)\pi_2$. Thus, $(3/4)\pi_2 + \pi_2 = (7/4)\pi_2 = 1$, and

$$\pi_2 = \boxed{4/7 \text{ or } 0.5714}$$

6.5 Let "sunny" be state 1 and "rainy" be state 2. Write the transition probability matrix,

$$P = \begin{pmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{pmatrix}$$

(each row total is 1).

April 1 next year is so many transitions away that we should better use the steady-state distribution. To find it, solve the system

$$\left\{ \begin{array}{l} \pi P = \pi \\ \sum \pi_i = 1 \end{array} \right. ; \quad \left\{ \begin{array}{rcl} 0.8\pi_1 + 0.4\pi_2 & = & \pi_1 \\ 0.2\pi_1 + 0.6\pi_2 & = & \pi_2 \\ \pi_1 + \pi_2 & = & 1 \end{array} \right. ;$$

$$\left\{ \begin{array}{rcl} 0.4\pi_2 & = & 0.2\pi_1 \\ \pi_1 + \pi_2 & = & 1 \end{array} \right. ; \quad \left\{ \begin{array}{rcl} \pi_1 & = & 2\pi_2 \\ 3\pi_2 & = & 1 \end{array} \right. ; \quad \left\{ \begin{array}{rcl} \pi_1 & = & 2/3 \\ \pi_2 & = & 1/3 \end{array} \right. .$$

The probability that April 1 next year is rainy is $\pi_2 = \boxed{1/3 \text{ or } 0.3333}$

6.6 The transition probability matrix is given as

$$P = \begin{pmatrix} 0.8 & 0.2 \\ 0.1 & 0.9 \end{pmatrix}$$

Also, we have $X(0) = 2$ (idle mode) with probability 1, i.e., $P_0 = (0,\ 1)$.

(a) $P_2 = P_0 P^2 = (0, 1) \begin{pmatrix} 0.8 & 0.2 \\ 0.1 & 0.9 \end{pmatrix}^2 = (0.17,\ 0.83)$. That is, $X(2)$, the state after 2 transitions, is busy with probability 0.17 and idle with probability 0.83.

(b) Solve the system of equations

$$\left\{ \begin{array}{l} \pi P = \pi \\ \pi_1 + \pi_2 = 1 \end{array} \right. ; \quad \left\{ \begin{array}{l} 0.8\pi_1 + 0.1\pi_2 \ = \ \pi_1 \\ 0.2\pi_1 + 0.9\pi_2 \ = \ \pi_2 \\ \pi_1 + \pi_2 = 1 \end{array} \right. ;$$

$$\left\{ \begin{array}{l} 0.2\pi_1 = 0.1\pi_2 \\ 0.2\pi_1 = 0.1\pi_2 \\ \pi_1 + \pi_2 = 1 \end{array} \right. ; \quad \left\{ \begin{array}{l} \pi_2 = 2\pi_1 \\ \pi_1 + 2\pi_1 = 1 \end{array} \right. ; \quad \boxed{\begin{array}{c} \pi_1 = 1/3 \\ \pi_2 = 2/3 \end{array}}$$

6.7 (a) Since probabilities in each row add to 1, the missing values should be filled as follows,

$$P = \begin{pmatrix} 0.3 & 0.7 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

(b) According to this transition probability matrix, transitions from state 1 to 1 and 2, from 2 to 3, and from 3 to 1 have positive probabilities. In 4 steps, all transitions are possible; we can reach any state from any other state. This can be checked by a state diagram or by computing the 4-step transition probability matrix:

$$\begin{pmatrix} 0.4281 & 0.5089 & 0.063 \\ 0.09 & 0.21 & 0.7 \\ 0.727 & 0.063 & 0.21 \end{pmatrix}$$

(c) Solve the system

$$\left\{ \begin{array}{l} \pi P = \pi \\ \sum \pi_i = 1 \end{array} \right. ; \quad \left\{ \begin{array}{l} 0.3\pi_1 + \pi_3 = \pi_1 \\ 0.7\pi_1 = \pi_2 \\ \pi_2 = \pi_3 \\ \sum \pi_i = 1 \end{array} \right. ; \quad \left\{ \begin{array}{l} \pi_3 = 0.7\pi_1 \\ \pi_2 = 0.7\pi_1 \\ \pi_2 = \pi_3 \\ \sum \pi_i = 1 \end{array} \right. ;$$

As we expect, one equation follows from the others. Next,

$$\left\{ \begin{array}{l} \pi_3 = 0.7\pi_1 \\ \pi_2 = 0.7\pi_1 \\ \pi_1 + 0.7\pi_1 + 0.7\pi_1 = 1 \end{array} \right. ; \quad \left\{ \begin{array}{l} \pi_1 = 1/2.4 = 5/12 \\ \pi_2 = 0.7\pi_1 = 7/24 \\ \pi_3 = 0.7\pi_1 = 7/24 \end{array} \right.$$

Answer: $\boxed{(5/12,\ 7/24,\ 7/24) \text{ or } (0.4167,\ 0.2917,\ 0.2917)}$

6.8 The problem describes the following transition probability matrix,

$$P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 1 & 0 & 0 \end{pmatrix}.$$

Find the steady state distribution by solving

$$\left\{ \begin{array}{l} \pi P = \pi \\ \sum \pi_i = 1 \end{array} \right. ; \quad \left\{ \begin{array}{l} \frac{1}{2}\pi_2 + \pi_3 = \pi_1 \\ \frac{1}{2}\pi_1 = \pi_2 \\ \frac{1}{2}\pi_1 + \frac{1}{2}\pi_2 = \pi_3 \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{array} \right. ; \quad \left\{ \begin{array}{l} \pi_1 = \frac{1}{2}\pi_2 + \pi_3 \\ \pi_2 = \frac{1}{2}\pi_1 = \frac{1}{4}\pi_2 + \frac{1}{2}\pi_3 \\ \frac{1}{2}\pi_1 + \frac{1}{2}\pi_2 = \pi_3 \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{array} \right. ;$$

$$\left\{ \begin{array}{l} \pi_1 = \frac{1}{2}\pi_2 + \pi_3 = \frac{4}{3}\pi_3 \\ \pi_2 = \frac{2}{3}\pi_3 \\ \frac{2}{3}\pi_3 + \frac{1}{3}\pi_3 = \pi_3 \\ \frac{4}{3}\pi_3 + \frac{2}{3}\pi_3 + \pi_3 = 1 \end{array} \right. ; \left\{ \begin{array}{l} \pi_1 = \frac{4}{3}\pi_3 = \frac{4}{9} \\ \pi_2 = \frac{2}{3}\pi_3 = \frac{2}{9} \\ \pi_3 = \frac{1}{3} \end{array} \right.$$

Answer: $\boxed{(4/9,\ 2/9,\ 1/3) \text{ or } (0.4444,\ 0.2222,\ 0.3333)}$

6.9 We have $\lambda = 6$ min^{-1} $= 0.1$ sec^{-1} and $\Delta = 2$ sec. Compute $p = \lambda\Delta = (0.1)(2) = 0.2$.

(a) Ten seconds represent $10/\Delta = 10/2 = 5$ frames. The number of tasks during 5 frames has Binomial($n = 5, p = 0.2$) distribution. From Table A2,

$$\boldsymbol{P}\{X > 2\} = 1 - F(2) = 1 - 0.9421 = \boxed{0.0579}$$

(b) The number of tasks during 100 seconds (50 frames) has Binomial($n = 50, p = 0.2$) distribution. For this large n, we can use Normal approximation with an appropriate continuity correction, $\mu = np = 10$, $\sigma = \sqrt{np(1-p)} = 2.83$,

$$\boldsymbol{P}\{X > 20\} = \boldsymbol{P}\{X > 20.5\} = \boldsymbol{P}\left\{\frac{X-\mu}{\sigma} > \frac{20.5-\mu}{\sigma}\right\}$$

$$= \boldsymbol{P}\{Z > 3.71\} = 1 - 0.9999 = \boxed{0.0001}$$

Notice that probabilities in (a) and (b) are not the same.

6.10 We need $\boldsymbol{P}\{X > 3\}$ where X is the number of messages received during a 2-minute interval. This X has Binomial distribution with parameters

$$n = 2 \text{ min}/\Delta = 2 \text{ min}/0.25 \text{ min} = 8$$

and

$$p = \lambda\Delta = (1\,\text{min}^{-1})(0.25 \text{ min}) = 0.25.$$

From Table A2, $\boldsymbol{P}\{X > 3\} = 1 - \boldsymbol{P}\{X \le 3\} = 1 - 0.8862 = \boxed{0.1138}$.

6.11 (a) $\Delta = p/\lambda = 0.1/(2\,\text{min}^{-1}) = \boxed{0.05 \text{ min or } 3 \text{ sec}}$

(b) The number of frames during 1 hour is

$$n = \frac{3600\,\text{sec}}{3\,\text{sec}} = 1200.$$

The number of jobs during 1200 frames is Binomial with $n = 1200$ and $p = 0.1$.

$$\mathbf{E}(X) = np = \boxed{120 \text{ jobs}}$$

$$\text{Std}(X) = \sqrt{np(1-p)} = \boxed{\sqrt{108} \text{ or } 10.39 \text{ jobs}}$$

6.12 (a) When $p \le 0.1$ and $\lambda = 2\,\text{min}^{-1}$,

$$\Delta = p/\lambda \le 0.1/(2\,\text{min}^{-1}) = \boxed{0.05 \text{ min or } 3 \text{ sec}}$$

Thus, any frame size up to 3 seconds is sufficient.

(b) Using 3-second frames, the number of frames during 5 minutes is

$$n = \frac{300\,\text{sec}}{3\,\text{sec}} = 100.$$

The number of landings $X(100)$ during 100 frames has Binomial distribution with $n = 100$ and $p = 0.1$.

$$\boldsymbol{P}\{X(100) = 0\} = (1-p)^n = 0.9^{100} = \boxed{0.000027}$$

(c) The number of frames during the next hour is

$$n = \frac{3600\,\text{sec}}{3\,\text{sec}} = 1200.$$

Thus, the number of landings $X(1200)$ during the next hour has Binomial distribution with $n = 1200$ and $p = 0.1$. With such a large n and a moderate p, this distribution is approximately Normal with parameters $\mu = np = 120$ and $\sigma = \sqrt{np(1-p)} = 10.39$. Using the continuity correction and Table A4,

$$\boldsymbol{P}\{X(1200) > 100\} = \boldsymbol{P}\{X(1200) > 100.5\} = \boldsymbol{P}\left\{Z > \frac{100.5 - 120}{10.39}\right\}$$

$$= 1 - \Phi(-1.88) = \Phi(1.88) = \boxed{0.9699}$$

6.13 We have $\lambda = 1/12\,\text{sec}^{-1}$ and $\Delta = 2$. Therefore, $p = \lambda\Delta = 1/6$,

$$\mathbf{E}(T) = 1/\lambda = \boxed{12 \text{ sec}}$$

and

$$\text{Var}(T) = \Delta^2 \frac{1-p}{p^2} = \boxed{120 \text{ sec}^2}$$

6.14 We have $\lambda = 3 \text{ min}^{-1}$ and $\Delta = 5$ sec $= 1/12$ min. Then $p = \lambda\Delta = 0.25$ is the probability of a new connection during a given frame. During 3 minutes, there are $3/(1/12) = 36$ frames. The number of new connections X during 36 frames has Binomial($n = 36$, $p = 0.25$) distribution that is approximately Normal with

$$\mu = np = 9 \quad \text{and} \quad \sigma = \sqrt{np(1-p)} = 2.60$$

Then

$$\boldsymbol{P}\{X > 10\} = \boldsymbol{P}\{X > 10.5\} = \boldsymbol{P}\left\{\frac{X-\mu}{\sigma} > \frac{10.5-9}{2.60}\right\}$$

$$= \boldsymbol{P}\{Z > 0.58\} = 1 - 0.7190 = \boxed{0.2810}$$

from Table A4.

For the time T between connections,

$$\mathbf{E}(T) = 1/\lambda = (1/3)\,\text{min} = \boxed{20 \text{ sec}}$$

and

$$\text{Std}(T) = \Delta\frac{\sqrt{1-p}}{p} = (5\,\text{sec})\frac{\sqrt{0.75}}{0.25} = \boxed{17.32 \text{ sec}}$$

6.15 (a) We are given $\lambda = 12\,\text{min}^{-1} = 0.2\,\text{sec}^{-1}$ and $p = 0.15$. Then

$$\Delta = p/\lambda = 0.15/0.2 = \boxed{0.75 \text{ seconds}}$$

(b) For the time T between two consecutive connections,

$$\mathbf{E}(T) = 1/\lambda = 1/0.2 = \boxed{5 \text{ sec}}$$

and

$$\text{Std}(T) = \frac{\Delta}{p}\sqrt{1-p} = \frac{0.75}{0.15}\sqrt{0.85} = \boxed{4.61 \text{ sec}}$$

6.16 The arrival rate is $\lambda = 40\,\text{hr}^{-1} = (2/3)\ \text{min}^{-1}$, and the frame length is $\Delta = (1/30)$ min. Then, the probability of a new message during any given frame is

$$p = \lambda\Delta = (2/3)(1/30) = 1/45.$$

There are $n = (30 \text{ min})/\Delta = 900$ frames between 10 am and 10:30 am, hence, the number of new messages X during this time has Binomial distribution with $n = 900$ and $p = 1/45$,

$$\mathbf{E}(X) = np = \boxed{20 \text{ messages}}$$

and

$$\text{Std}(X) = \sqrt{np(1-p)} = \boxed{4.42 \text{ messages}}$$

6.17 There is no restriction on the number of arrived messages arriving at random times with the given arrival rate. Hence, Poisson process is the appropriate model for the number of arrived messages.

The number of messages X during the next hour has Poisson distribution with parameter $\lambda = 9\ \text{hr}^{-1}$. From Table A3,

(a) $\boldsymbol{P}\{X \geq 5\} = 1 - F(4) = 1 - 0.055 = \boxed{0.945}$

(b) $\boldsymbol{P}\{X = 5\} = F(5) - F(4) = 0.116 - 0.055 = \boxed{0.061}$

6.18 First, compute $p = \lambda\Delta = (1/15\,\text{sec})(5\,\text{sec}) = 1/3$, the probability of a new arrival during any given frame.

Then X, the number of arrivals during 200 min, or 12000 sec, has Binomial distribution with parameters $n = 12000/\Delta = 12000/5 = 2400$ and $p = 1/3$. Using the Central Limit Theorem with a suitable continuity correction,

$$\begin{aligned} \boldsymbol{P}\{X \leq 750\} &= \boldsymbol{P}\{X \leq 750.5\} = \boldsymbol{P}\left\{Z \leq \frac{750.5 - np}{\sqrt{np(1-p)}}\right\} \\ &= \boldsymbol{P}\{Z \leq -2.14\} = \Phi(-2.14) = \boxed{0.0162} \end{aligned}$$

6.19 We need $\boldsymbol{P}\{X > 5\}$, where X is the number of power outages during 3 months. It has Poisson distribution with parameter $\lambda t = (3)(3) = 9$ months^{-1}. From Table A3,

$$\boldsymbol{P}\{X > 5\} = 1 - \boldsymbol{P}\{X \le 5\} = 1 - 0.116 = \boxed{0.884}$$

6.20 We need $\boldsymbol{P}\{X > 5\}$, where X is the number of telephone calls during 12 months. It has Poisson distribution with parameter $\lambda t = (1/3)(12) = 4$ min^{-1}. From Table A3,

$$\boldsymbol{P}\{X > 5\} = 1 - \boldsymbol{P}\{X \le 5\} = 1 - 0.785 = \boxed{0.215}$$

6.21 The number of blackouts X during any month is the number of "rare" events, with possible values 0, 1, 2, ... Thus, X has Poisson distribution with parameter $\lambda = 5$.

(a) From Table A3,

$$\boldsymbol{P}\{X > 3\} = 1 - F(3) = 1 - 0.265 = \boxed{0.735}$$

(b) Poisson distribution has $\mathbf{E}(X) = \mathrm{Var}(X) = \lambda = 5$. The monthly cost equals $1500\,X$ because an amount of \$1500 is paid for each blackout. Thus,

$$\mathbf{E}(\text{monthly cost}) = \mathbf{E}(1500X) = 1500\,\mathbf{E}(X) = \boxed{7500 \text{ dollars}}$$

and

$$\mathrm{Std}(\text{monthly cost}) = 1500\,\mathrm{Std}(X) = 1500\sqrt{5} = \boxed{3354 \text{ dollars } 10 \text{ cents}}$$

6.22 (a) Let X be the number of connections during 2 minutes. It has Poisson distribution with parameter $\lambda t = (5)(2) = 10$. No offer is made if there are fewer than three connections. By Table A3,

$$\boldsymbol{P}\{\text{ no offer }\} = \boldsymbol{P}\{X \le 2\} = \boxed{0.003}$$

(b) The time T_3 of the third connection (and therefore, the first offer) has Gamma distribution with parameters $\alpha = 3$ and $\lambda = 5$. Then

$$E(T_3) = \alpha/\lambda = \boxed{3/5 \text{ or } 0.6 \text{ min}}$$

and

$$\mathrm{Var}(T_3) = \alpha/\lambda^2 = \boxed{3/25 \text{ or } 0.12 \text{ min}^2}$$

6.23 Let X be the number of times Mr. Z is caught drinking and driving during 10 years. It has Poisson distribution with parameter $\lambda t = (1/4)(10) = 2.5$.

Keeping the driver's license is equivalent to being caught no more than twice. Then,

$$\boldsymbol{P}\left\{\begin{array}{c}\text{Mr. Z keeps his}\\ \text{driver's license}\end{array}\right\} = \boldsymbol{P}\left\{X \le 2\right\} = \boxed{0.544}$$

6.24 For each i, let us find the steady-state distribution of Y_i. First, write the transition probability matrix for Y_i,

$$P = \begin{pmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{pmatrix}.$$

Then solve the system of equations,

$$\left\{\begin{array}{l} \pi P = \pi \\ \sum \pi_i = 1 \end{array}\right. ; \quad \left\{\begin{array}{lcl} 0.8\pi_0 + 0.5\pi_1 & = & \pi_0 \\ 0.2\pi_0 + 0.5\pi_1 & = & \pi_1 \\ \pi_0 + \pi_1 & = & 1 \end{array}\right. ;$$

$$\left\{\begin{array}{l} 0.5\pi_1 = 0.2\pi_0 \\ \pi_0 + 0.4\pi_0 = 1 \end{array}\right. ; \quad \left\{\begin{array}{lcl} \pi_0 & = & 1/1.4 = 5/7 \\ \pi_1 & = & 0.4\pi_0 = 2/7 \end{array}\right.$$

It means that at each moment, each user is connected with probability $2/7$ and disconnected with probability $5/7$ independently of the other user. Thus, the number of connected users $X = Y_1 + Y_2$ has Binomial distribution with parameters $n = 2$ and $p = 2/7$:

$$\begin{array}{lclcl} \boldsymbol{P}\{X=0\} & = & (5/7)^2 & = & 25/49 \text{ or } 0.5102 \\ \boldsymbol{P}\{X=1\} & = & 2(2/7)(5/7) & = & 20/49 \text{ or } 0.4082 \\ \boldsymbol{P}\{X=2\} & = & (2/7)^2 & = & 4/49 \text{ or } 0.0816 \end{array}$$

6.27 Arrival times should be distributed rather unevenly. Indeed, they are separated by Exponential interarrival times. Looking at the graph of Exponential density, we see that an Exponential variable takes small values with rather high probabilities. This is why we see many short periods between arrivals. However, once in a while an Exponential variable takes a large value that corresponds to a long period of time with no arrivals. There is no limit on the longest interarrival time.

Chapter 7

7.1 We have the frame length $\Delta = 10\,\text{sec} = (1/6)$ min, the arrival rate $\lambda_A = 10\,\text{hrs}^{-1} = (1/6)\ \text{min}^{-1}$, the mean service time $\mu_S = 2$ min, thus the service rate is $\lambda_S = (1/2)\ \text{min}^{-1}$. Then,

$$p_A = \lambda_A \Delta = 1/36 \quad \text{and} \quad p_S = \lambda_S \Delta = 1/12,$$

and

$$P = \begin{pmatrix} \frac{35}{36} & \frac{1}{36} & 0 & 0 & \dots \\ \frac{35}{36}\frac{1}{12} & \frac{1}{36}\frac{1}{12}+\frac{35}{36}\frac{11}{12} & \frac{1}{36}\frac{11}{12} & 0 & \dots \\ 0 & \frac{35}{36}\frac{1}{12} & \frac{1}{36}\frac{1}{12}+\frac{35}{36}\frac{11}{12} & \frac{1}{36}\frac{11}{12} & \dots \\ 0 & 0 & \frac{35}{36}\frac{1}{12} & \frac{1}{36}\frac{1}{12}+\frac{35}{36}\frac{11}{12} & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

$$= \begin{pmatrix} \frac{35}{36} & \frac{1}{36} & 0 & 0 & \dots \\ \frac{35}{432} & \frac{386}{432} & \frac{11}{432} & 0 & \dots \\ 0 & \frac{35}{432} & \frac{386}{432} & \frac{11}{432} & \dots \\ 0 & 0 & \frac{35}{432} & \frac{386}{432} & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} \text{ or } \begin{pmatrix} .972 & .028 & 0 & 0 & \dots \\ .081 & .894 & .025 & 0 & \dots \\ 0 & .081 & .894 & .025 & \dots \\ 0 & 0 & .081 & .894 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

7.2 We are given $\lambda_A = 2\text{ min}^{-1}$, $\lambda_S = 4\text{ min}^{-1}$, $\Delta = 0.2$ min, and $C = 2$. Compute probabilities

$$p_A = \lambda_A \Delta = (2)(0.2) = 0.4, \quad p_S = \lambda_S \Delta = (4)(0.2) = 0.8,$$

and the transition probability matrix (with $X \in \{0,1,2\}$ due to the limited capacity)

$$\begin{aligned} P &= \begin{pmatrix} 1-p_A & p_A & 0 \\ p_S(1-p_A) & (1-p_A)(1-p_S)+p_Ap_S & p_A(1-p_S) \\ 0 & p_S(1-p_A) & 1-p_S(1-p_A) \end{pmatrix} \\ &= \begin{pmatrix} 0.6 & 0.4 & 0 \\ 0.48 & 0.44 & 0.08 \\ 0 & 0.48 & 0.52 \end{pmatrix}. \end{aligned}$$

Then solve the steady-state equations

$$\left\{ \begin{array}{l} \pi P = \pi \\ \sum \pi_i = 1 \end{array} \right. \Rightarrow \left\{ \begin{array}{lcl} 0.6\pi_0 + 0.48\pi_1 & = & \pi_0 \\ 0.4\pi_0 + 0.44\pi_1 + 0.48\pi_2 & = & \pi_1 \\ 0.08\pi_1 + 0.52\pi_2 & = & \pi_2 \\ \pi_0 + \pi_1 + \pi_2 & = & 1 \end{array} \right.$$

$$\Rightarrow \left\{ \begin{array}{lcl} \pi_0 = 1.2\pi_1 & & \\ (0.4)(1.2\pi_1) + 0.44\pi_1 + 0.48\pi_2 & = & \pi_1 \\ 0.08\pi_1 + 0.52\pi_2 & = & \pi_2 \\ 1.2\pi_1 + \pi_1 + \pi_2 & = & 1 \end{array} \right. \Rightarrow \left\{ \begin{array}{l} \pi_0 = 1.2\pi_1 \\ 0.48\pi_2 = 0.08\pi_1 \\ 0.48\pi_2 = 0.08\pi_1 \\ 2.2\pi_1 + \pi_2 = 1 \end{array} \right.$$

$$\Rightarrow \left\{\begin{array}{l} \pi_0 = 1.2\pi_1 \\ \pi_1 = 6\pi_2 \\ 2.2(6\pi_2) + \pi_2 = 1 \end{array}\right. \qquad \Rightarrow \boxed{\left\{\begin{array}{lcccl} \pi_0 & = & 1.2\pi_1 & = & 0.5070 \\ \pi_1 & = & 6\pi_2 & = & 0.4225 \\ \pi_2 & = & 1/14.2 & = & 0.0704 \end{array}\right.}$$

7.3 We have $\Delta = 2$ min, $\lambda_A = 1/\mu_A = 1/10$ min^{-1}, $\lambda_S = 1/\mu_S = 1/6$ min^{-1}. Therefore, $p_A = \lambda_A\Delta = 1/5$ and $p_S = \lambda_S\Delta = 1/3$.

There are 2 frames between 10:00 am and 10:04 am. We need the *conditional probability* $\boldsymbol{P}\{X_2 = 2 \mid X_0 = 0\}$. Since the number of cars at the wash center can change by at most 1 during each frame, this probability equals

$$p_{02}^{(2)} = p_{01}p_{12} = p_A(p_A(1-p_S)) = (1/5)(1/5)(2/3) = \boxed{2/75 \text{ or } 0.0267}$$

7.4 Mary's telephone queuing system has the arrival rate $\lambda_A = 10$ hr^{-1}= 1/6 min^{-1}, the service rate $\lambda_S = 1/\mu_S = 1/2$ min^{-1}, the frame length $\Delta = 1$ min, and a limited capacity $C = 2$.

Compute probabilities

$$p_A = \lambda_A\Delta = 1/6, \quad p_S = \lambda_S\Delta = 1/2,$$

and the transition probability matrix

$$P = \begin{pmatrix} 1-p_A & p_A & 0 \\ p_S(1-p_A) & (1-p_A)(1-p_S)+p_Ap_S & p_A(1-p_S) \\ 0 & p_S(1-p_A) & 1-p_S(1-p_A) \end{pmatrix}$$

$$= \begin{pmatrix} 5/6 & 1/6 & 0 \\ 5/12 & 1/2 & 1/12 \\ 0 & 5/12 & 7/12 \end{pmatrix}.$$

Then solve the steady-state equations

$$\left\{\begin{array}{l} \pi P = \pi \\ \sum \pi_i = 1 \end{array}\right. \quad \Rightarrow \left\{\begin{array}{lcl} (5/6)\pi_0 + (5/12)\pi_1 & = & \pi_0 \\ (1/6)\pi_0 + (1/2)\pi_1 + (5/12)\pi_2 & = & \pi_1 \\ (1/12)\pi_1 + (7/12)\pi_2 & = & \pi_2 \\ \pi_0 + \pi_1 + \pi_2 & = & 1 \end{array}\right.$$

$$\Rightarrow \left\{\begin{array}{l} \pi_0 = (5/2)\pi_1 \\ (1/6)(5/2)\pi_1 + (1/2)\pi_1 + (5/12)\pi_2 = \pi_1 \\ \pi_2 = (1/5)\pi_1 \\ (5/2)\pi_1 + \pi_1 + (1/5)\pi_1 = 1 \end{array}\right.$$

$$\Rightarrow \left\{\begin{array}{l} \pi_0 = (5/2)\pi_1 \\ (5/12)\pi_1 + (6/12)\pi_1 + (1/12)\pi_1 = \pi_1 \quad \text{(always true)} \\ \pi_2 = (1/5)\pi_1 \\ 3.7\pi_1 = 1 \end{array}\right.$$

$$\Rightarrow \left\{\begin{array}{lcl} \pi_0 & = & 25/37 \\ \pi_1 & = & 10/37 \\ \pi_2 & = & 2/37 \end{array}\right.$$

Then, the fraction of time Mary spends using her telephone is

$$\boldsymbol{P}\{X \geq 1\} = 1 - \pi_0 = \boxed{12/37 \text{ or } 0.3243}$$

7.5 This is a Bernoulli queuing process with limited capacity $C = 2$ (one customer getting service and one customer waiting). Its parameters are:

$$\lambda_A = 1/10\,\text{min}^{-1}, \quad \lambda_S = 1/15\,\text{min}^{-1}, \quad \Delta = 3 \text{ min}.$$

Compute the probabilities $p_A = \lambda_A \Delta = 0.3$ and $p_S = \lambda_S \Delta = 0.2$.

Then, $p_S(1-p_A) = 0.14$ and $p_A(1-p_S) = 0.24$, and we have the transition probability matrix

$$P = \begin{pmatrix} 0.7 & 0.3 & 0 \\ 0.14 & 0.62 & 0.24 \\ 0 & 0.14 & 0.86 \end{pmatrix}$$

Solve $\pi P = \pi$ along with $\pi_0 + \pi_1 + \pi_2 = 1$.

The first equation is

$$0.7\pi_0 + 0.14\pi_1 = \pi_0 \quad \Rightarrow \quad 0.3\pi_0 = 0.14\pi_1 \quad \Rightarrow \quad \pi_0 = 7/15\pi_1$$

The second equation is

$$0.3\pi_0 + 0.62\pi_1 + 0.14\pi_2 = \pi_1 \quad \Rightarrow \quad 0.14\pi_1 + 0.62\pi_1 + 0.14\pi_2 = \pi_1$$

$$\Rightarrow \quad 0.24\pi_1 = 0.14\pi_2 \quad \Rightarrow \quad \pi_2 = 12/7\pi_1$$

Check that the third equation,

$$0.24\pi_1 + 0.86\pi_2 = (0.24)(7/12)\pi_2 + 0.86\pi_2 = \pi_2,$$

follows from the first two, as we expect.

From the normalizing condition $\pi_0 + \pi_1 + \pi_2 = 1$,

$$\frac{7}{15}\pi_1 + \pi_1 + \frac{12}{7}\pi_1 = \frac{334}{105}\pi_1 = 1 \quad \Rightarrow \quad \boxed{\left\{\begin{array}{l} \pi_0 = 49/334 \text{ or } 0.1467 \\ \pi_1 = 105/334 \text{ or } 0.3144 \\ \pi_2 = 180/334 \text{ or } 0.5389 \end{array}\right.}$$

7.6 This queuing system has capacity $C = 2$, arrival rate $\lambda_A = 5$ hrs^{-1} or $(1/12)$ min^{-1}, service rate $\lambda_S = (1/20)$ min^{-1}, and frames $\Delta = 4$ min. Then

$$p_A = \lambda_A \Delta = (1/12)(4) = 1/3 \quad \text{and} \quad p_S = \lambda_S \Delta = (1/20)(4) = 1/5.$$

The transition probability matrix is

$$P = \begin{pmatrix} 1-p_A & p_A & 0 \\ p_S(1-p_A) & (1-p_A)(1-p_S)+p_Ap_S & p_A(1-p_S) \\ 0 & p_S(1-p_A) & 1-p_S(1-p_A) \end{pmatrix}$$

$$= \begin{pmatrix} 2/3 & 1/3 & 0 \\ 2/15 & 3/5 & 4/15 \\ 0 & 2/15 & 13/15 \end{pmatrix}.$$

Solve the steady-state equations

$$\left\{ \begin{array}{l} \pi P = \pi \\ \sum \pi_i = 1 \end{array} \right. \Rightarrow \left\{ \begin{array}{lcl} (2/3)\pi_0 + (2/15)\pi_1 & = & \pi_0 \\ (1/3)\pi_0 + (3/5)\pi_1 + (2/15)\pi_2 & = & \pi_1 \\ (4/15)\pi_1 + (13/15)\pi_2 & = & \pi_2 \\ \pi_0 + \pi_1 + \pi_2 & = & 1 \end{array} \right.$$

$$\Rightarrow \left\{ \begin{array}{l} \pi_0 = (2/5)\pi_1 \\ (1/3)(2/5)\pi_1 + (3/5)\pi_1 + (2/15)\pi_2 = \pi_1 \\ \pi_2 = 2\pi_1 \\ (2/5)\pi_1 + \pi_1 + 2\pi_1 = 1 \end{array} \right.$$

$$\Rightarrow \left\{ \begin{array}{l} \pi_0 = (2/5)\pi_1 \\ (11/15)\pi_1 + (2/15)(2\pi_1) = \pi_1 \quad \text{(always true)} \\ \pi_2 = 2\pi_1 \\ (17/5)\pi_1 = 1 \end{array} \right.$$

$$\Rightarrow \boxed{\left\{ \begin{array}{lcccl} \pi_0 & = & 2/17 & = & 0.1176 \\ \pi_1 & = & 5/17 & = & 0.2941 \\ \pi_2 & = & 10/17 & = & 0.5882 \end{array} \right.}$$

7.7 (a) Here $\lambda_A = 4/5 \text{ min}^{-1}$ and $p = 0.05$. Then

$$\Delta = p_A/\lambda_A = 0.05/(4/5) = \boxed{1/16 \text{ min or } 3.75 \text{ sec}}$$

(b) Let us see if 50 messages during 1 hour (960 frames) is "too many" for the arrival rate $\lambda_A = 4/5 \text{ min}^{-1}$. The number of messages arrived during 1 hour has Binomial distribution with parameters $n = 960$ and $p = 0.05$. Compute the probability

$$\boldsymbol{P}\{X \geq 50\} = \boldsymbol{P}\left\{ Z > \frac{49.5 - np}{\sqrt{np(1-p)}} \right\} = \boldsymbol{P}\{Z > 0.22\} = 0.4129.$$

This is not a low probability; observing 50 or more arrivals is not unlikely, so there is no indication that the arrival rate is on the increase.

7.8 This system has capacity $C = 3$, arrival rate $\lambda_A = 8 \text{ hrs}^{-1}$or $(2/15) \text{ min}^{-1}$, service rate $\lambda_S = (1/3) \text{ min}^{-1}$, and frames $\Delta = 5$ sec or $(1/12)$ min. Then

$$p_A = \lambda_A\Delta = (2/15)(1/12) = 1/90 \quad \text{and} \quad p_S = \lambda_S\Delta = (1/3)(1/12) = 1/36.$$

Then, the transition probability matrix is

$$P = \begin{pmatrix} 1-p_A & p_A & 0 & 0 \\ p_S(1-p_A) & \begin{array}{c}(1-p_A)(1-p_S)\\+p_Ap_S\end{array} & p_A(1-p_S) & 0 \\ 0 & p_S(1-p_A) & \begin{array}{c}(1-p_A)(1-p_S)\\+p_Ap_S\end{array} & p_A(1-p_S) \\ 0 & 0 & p_S(1-p_A) & 1-p_S(1-p_A) \end{pmatrix}$$

$$= \begin{pmatrix} 89/90 & 1/90 & 0 & 0 \\ 89/3240 & 3116/3240 & 35/3240 & 0 \\ 0 & 89/3240 & 3116/3240 & 35/3240 \\ 0 & 0 & 89/3240 & 3151/3240 \end{pmatrix}$$

$$\text{or} \begin{pmatrix} 0.9889 & 0.0111 & 0 & 0 \\ 0.0275 & 0.9617 & 0.0108 & 0 \\ 0 & 0.0275 & 0.9617 & 0.0108 \\ 0 & 0 & 0.0275 & 0.9725 \end{pmatrix}$$

7.9 This M/M/1 queuing system has the arrival rate $\lambda_A = 5\ \text{min}^{-1}$ and the mean service time $\mu_S = 4$ sec or (1/15) min. Then

$$r = \lambda_A/\lambda_S = \lambda_A\mu_S = 5/15 = 1/3.$$

(a) $\boldsymbol{P}\{X=2\} = \pi_2 = (1-r)r^2 = (2/3)(1/3)^2 = \boxed{2/27 \text{ or } 0.0741}$

(b) $\mathbf{E}(R) = \dfrac{\mu_S}{1-r} = \dfrac{4\text{ sec}}{2/3} = \boxed{6 \text{ seconds}}$

7.10 Description of this queuing system matches M/M/1 queue with the arrival rate $\lambda_A = 12\ \text{hrs}^{-1} = 0.2\ \text{min}^{-1}$ and the mean service time $\mu_S = 2$ min. Its utilization is $r = \lambda_A\mu_S = 0.4$.

(a) $\mathbf{E}(R) = \dfrac{\mu_S}{1-r} = \dfrac{2}{1-0.4} = \boxed{3\tfrac{1}{3} \text{ min or 3 min 20 sec}}$
The job is expected to be printed at 12:03:20.

(b) $\boldsymbol{P}\{X>2\} = \sum_{k=3}^{\infty}(1-r)r^k = r^3 = 0.4^3 = \boxed{0.064}$

7.11 This M/M/1 system has the arrival rate $\lambda_A = 20\ \text{hr}^{-1}$ or $1/3\ \text{min}^{-1}$, the mean service time $\mu_S = 2$ min, and the service rate $\lambda_S = 1/2\ \text{min}^{-1}$. The arrival-to-service ratio is $r = \lambda_A/\lambda_S = (1/3)/(1/2) = 2/3$.

(a) Regardless of the arrival time,

$$\mathbf{E}(W) = \frac{r\mu_S}{1-r} = \boxed{4 \text{ minutes}}$$

(b) $\boldsymbol{P}\{X=0\} = 1 - r = \boxed{1/3}$

7.12 The problem describes an M/M/1 queuing process with $\lambda_A = 10/25 = 0.4$ min^{-1}, $\mu_S = 2$ min, and $r = \lambda_A \mu_S = 0.8$.

(a) $\mathbf{E}(X) = \dfrac{r}{1-r} = 0.8/0.2 = \boxed{4 \text{ customers}}$

$\mathbf{E}(X_w) = \dfrac{r^2}{1-r} = 0.8^2/0.2 = \boxed{3.2 \text{ customers}}$

(b) $\boldsymbol{P}\{\text{ busy }\} = r = \boxed{0.8}$

(c) $\boldsymbol{P}\{X \geq 6\} = \sum_{6}^{\infty} \pi_x = \sum_{6}^{\infty} (1-r)r^x = r^6 = (0.8)^6 = \boxed{0.2621}$

7.13 This M/M/1 queuing system has $\mu_A = 5$ min and $\mu_S = 3$ min. Then, its utilization equals $r = \mu_S/\mu_A = 0.6$.

(a) $\mathbf{E}(R) = \mu_S/(1-r) = 3/0.4 = \boxed{7.5 \text{ min}}$

(b) $\boldsymbol{P}\{X < 2\} = P(0) + P(1) = (1-r) + (1-r)r = 1 - r^2 = \boxed{0.64}$

(c) Customers have to wait only if there are customers ahead of them still in the system. Therefore,

$$\boldsymbol{P}\{W > 0\} = \boldsymbol{P}\{X > 0\} = r = \boxed{0.6}$$

7.14 We are given the arrival rate $\lambda_A = 1/4.5 = 2/9$ min^{-1}and the mean response time $\mathbf{E}(R) = 12/5$ min. This is *not* an M/M/1 queuing system because the service times are not Exponential.

(a) According to the Little's Law,

$$\mathbf{E}(X) = \lambda_A \, \mathbf{E}(R) = (12/5)(2/9) = \boxed{24/45 \text{ or } 0.5333 \text{ jobs}}$$

(b) The number of arrivals N during a 3-hour (180-minute) period has Poisson distribution with parameter

$$\lambda_A t = (2/9)(180) = 40.$$

That is, we expect to see 40 arrivals during this time. How unlikely is it to observe only 20 (or even fewer) arrivals, if the expected interarrival time is still 4.5 minutes? Compute

$$\boldsymbol{P}\{N \leq 20\} = \sum_{x=0}^{20} \frac{e^{-40} 40^x}{x!} = 0.00037.$$

This probability is extremely low; it is very unlikely to see only 20 arrivals under the current conditions. Thus, there is a strong evidence that the arrival rate has fallen, and the expected interarrival time has increased.

Remark: (Computation) Is there any shortcut for the computation of $\boldsymbol{P}\{N \leq 20\}$? A Normal approximation is not accurate for this Poisson distribution. We computed the probability directly by the formula of Poisson pmf. Instead, we can provide two simple bounds for this probability.

One is obtained from Chebyshev's inequality,

$$\boldsymbol{P}\{N \leq 20\} \leq \boldsymbol{P}\{|N-40| \geq 20\} \leq \frac{\mathrm{Var}(N)}{20^2} = \frac{40}{400} = 0.1.$$

The other bound can be obtained from Table A3. The largest parameter λ in this table is 30, and the probability of $N \leq 20$ is 0.035 under this λ. This probability is even lower for $\lambda = 40$. Therefore,

$$\boldsymbol{P}\{N \geq 20\} \leq 0.035.$$

7.15 The described system is M/M/1 with $\mu_A = 10$ min and $\mu_S = 3$ min. Hence, $\lambda_A = 1/10\ \mathrm{min}^{-1}$, $\lambda_S = 1/3\ \mathrm{min}^{-1}$, and

$$r = \lambda_A/\lambda_S = 0.3.$$

(a) $\mathbf{E}(X) = \dfrac{r}{1-r} = \boxed{3/7 \text{ or } 0.4286}$

(b) $\boldsymbol{P}\{X=0\} = 1 - r = \boxed{0.7 \text{ or } 70\% \text{ of time}}$

(c) $\mathbf{E}(R) = \dfrac{\mu_S}{1-r} = \dfrac{3}{0.7} = \boxed{30/7 \text{ or } 4.2857 \text{ min}}$

7.16 The arrival rate is $\lambda_A = 5\ \mathrm{min}^{-1}$ or $(1/12)\ \mathrm{sec}^{-1}$. The expected service time is 2 sec for all messages plus 7 sec for 40% of messages that contain attachments, so that

$$\mu_S = 2 + (0.4)(7) = 4.8 \text{ sec}$$

By the Little's Law,

$$\mathbf{E}(X) = \lambda_A\, \mathbf{E}(R) = \frac{\mathbf{E}(R)}{12}$$

Let us find another equation relating $\mathbf{E}(X)$ and $\mathbf{E}(R)$.

During a long period of time, T seconds, we expect $(T/12)$ arrived jobs. Among them, 40% of jobs take an average of 7 sec to process attachments. Overall, attachment processing takes $(T/12)(0.4)(7) = 7T/30$ sec, or $(7/30)$ of the time. Thus, when a new job arrives,

$$\boldsymbol{P}\{\text{ server processes attachments }\} = 7/30$$

We evaluated this probabilities as a long-run proportion.

In Section 7.4.1, we computed the expected response time $\mathbf{E}(R)$ as the sum of service times of $(X+1)$ jobs, that is, $(X-1)$ waiting jobs, one new arrived job, and the the remaining time of the currently served job. We could do it for M/M/1 systems because all service times were memoryless, so the remaining service time was still Exponential with the same parameter.

However, the given system is not M/M/1 because messages containing attachments are not processed in an Exponential amount of time. If the server is processing attachments (with probability 7/30) at the time when a new job arrives, this remaining service time has the expectation of 7 sec instead of $\mu_S = 4.8$ sec. With this correction, we have

$$\begin{aligned}\mathbf{E}(R) &= (\mathbf{E}(X)+1)\mu_S + \left(\frac{7}{30}\right)(7-4.8) \\ &= \left(\frac{\mathbf{E}(R)}{12}+1\right)(4.8) + 0.5133 = 0.4\,\mathbf{E}(R) + 5.3133.\end{aligned}$$

Solving this equation, we get $\mathbf{E}(R) = 5.3133/0.6 = \boxed{8.8555 \text{ sec}}$

7.17 This system has $\lambda_A = 2$ min^{-1}, $\lambda_S = 1$ min^{-1}, and $\Delta = 1/60$ min.

(a) Compute

$$p_A = \lambda_A \Delta = 1/30; \quad p_S = \lambda_S \Delta = 1/60.$$

The transition probability matrix is

$$P = \begin{pmatrix} 1-p_A & p_A \\ (1-p_A)p_S & 1-(1-p_A)p_S \end{pmatrix} = \begin{pmatrix} 29/30 & 1/30 \\ 29/1800 & 1771/1800 \end{pmatrix}$$

Solve $\pi P = \pi$, $\pi_0 + \pi_1 = 1$ for π:

$$\left\{\begin{array}{lcl} \frac{29}{30}\pi_0 + \frac{29}{1800}\pi_1 &=& \pi_0 \\ \frac{1}{30}\pi_0 + \frac{1771}{1800}\pi_1 &=& \pi_1 \\ \pi_0 + \pi_1 &=& 1 \end{array}\right. \Rightarrow \left\{\begin{array}{l} \pi_0 = \frac{29}{60}\pi_1 \\ \left(\frac{29}{60}+1\right)\pi_1 = 1 \end{array}\right. \Rightarrow \boxed{\left\{\begin{array}{l} \pi_0 = \frac{29}{89} \\ \pi_1 = \frac{60}{89} \end{array}\right.}$$

(b) Let N be the number of calls during 1 hour between 2 pm and 3 pm. It has Binomial distribution with $n = 3600$ (the number of frames) and $p = p_A = 1/30$.

$$\boldsymbol{P}\{X > 150\} = \boldsymbol{P}\left\{Z > \frac{150.5 - np}{\sqrt{np(1-p)}}\right\} = \boldsymbol{P}\{Z > 2.83\} = \boxed{0.0023}$$

7.18 We are given the arrival rate $\lambda_A = 1/6$ min^{-1}and the mean response time $\mathbf{E}(R) = 20$ min. By the Little's Law.

$$\mathbf{E}(X) = \lambda_A\,\mathbf{E}(R) = \boxed{3.3333 \text{ jobs}}$$

7.19 For the whole M/M/1 system,

$$\mathbf{E}(X) = \frac{r}{1-r} \quad \text{and} \quad \mathbf{E}(R) = \frac{\mu_S}{1-r},$$

so that

$$\lambda_A\,\mathbf{E}(R) = \frac{\lambda_A \mu_S}{1-r} = \frac{r}{1-r} = \mathbf{E}(X).$$

For the waiting line only,

$$\mathbf{E}(X_w) = \frac{r^2}{1-r} \quad \text{and} \quad \mathbf{E}(W) = \frac{\mu_S r}{1-r},$$

so that

$$\lambda_A \, \mathbf{E}(W) = \frac{\lambda_A \mu_S r}{1-r} = \frac{r^2}{1-r} = \mathbf{E}(X_w).$$

For the server only,

$$\mathbf{E}(X_s) = r \quad \text{and} \quad \mathbf{E}(S) = \mu_S,$$

so that

$$\lambda_A \, \mathbf{E}(S) = \lambda_A \mu_S = r = \mathbf{E}(X_s).$$

Chapter 8

8.1 (a) Naturally, the first digit of each number goes to the stem, and the second digit goes to the leaf. The leaf on the left represents the blocked intrusion attempts during the first 14 days, and the leaf on the right represents 20 days after the change.

	2	1
7 8	3	2 2 3 5 6 6 7 8 9 9
3 3 7 9	4	3 4 5 5 6 8 9
0 0 4 6 6 8 9	5	3 3
0	6	

(b) Before the change of firewall settings,

$$\min(X) = 37, \; \hat{Q}_1 = 43, \; \hat{M} = 50, \; \hat{Q}_3 = 56, \; \max(X) = 60.$$

The interquartile range is $56 - 43 = 13$, all the data are within 1.5(IQR) from $\hat{Q}_1$ or $\hat{Q}_3$, and thus, we do not suspect any outliers.

After the change,

$$\min(X) = 21, \; \hat{Q}_1 = 35, \; \hat{M} = 39, \; \hat{Q}_3 = 46, \; \max(X) = 53.$$

In fact, according to Definition 8.7, any number between 35 and 36 is the first quartile, and any number between 45 and 46 is the third quartile. Next, $IQR = 10$, and again, we see no outliers in this sample.

For the boxplots, compute sample means: $\bar{X} = 50$ before the change, and $\bar{X} = 40.2$ after the change. Parallel boxplots are in Figure 11.5.

(c) The stem-and-leaf plots and boxplots show that the daily number of intrusion attempts has decreased in general. Every element of the five-number summary has reduced following the change in firewall settings,

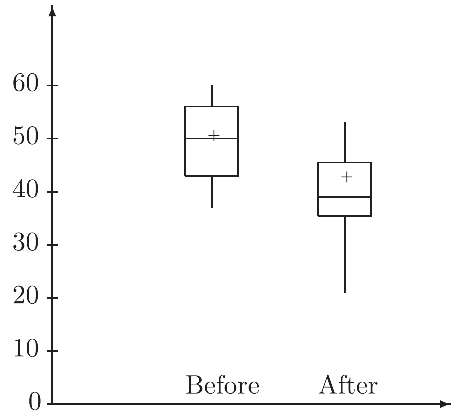


Figure 11.5 *Parallel boxplots for Exercise 8.1*

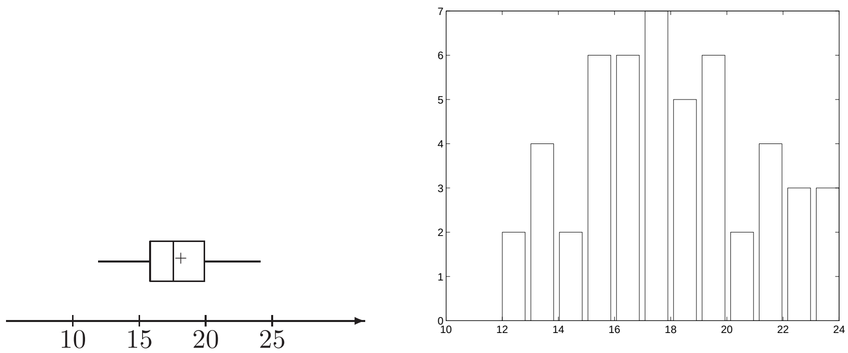


Figure 11.6 *A boxplot and a histogram for Exercise 8.2*

suggesting that any number of intrusion attempts is now exceeded with a lower probability.

8.2 (a) By Definitions 8.3 and 8.8,

$$\bar{X} = 17.9540, \quad \mathrm{Var}(X) = 9.9682, \quad \mathrm{Std}(X) = 3.1573.$$

(b) Order the data from smallest to largest. The minimum is 11.9, the maximum is 24.1. For the sample of size 50, the median $\hat{M}$ is any number between the 25th and the 26th smallest, i.e., between 17.5 and 17.6. The first quartile $\hat{Q}_1$ is the 13th smallest, i.e., 15.8 (exceeds $22\% \leq 25\%$ of the sample; is exceeded by $74\% \leq 75\%$ of the sample), and the third quartile $\hat{Q}_3$ is the 37th smallest, i.e., 19.9. The five-point summary is

$$(11.9,\ 15.8,\ 17.55,\ 19.9,\ 24.1).$$

The boxplot is on Figure 11.6.

(c) $IQR = \hat{Q}_3 - \hat{Q}_1 = 4.1$. Compute

$$\hat{Q}_1 - 1.5(IQR) = 9.65 \text{ and } \hat{Q}_3 + 1.5(IQR) = 26.05.$$

Both $\min(X) = 11.9$ and $\max(X) = 24.1$ are within the range $[9.65, 26.05]$, hence the sample has no obvious outliers.

(d) The histogram in Figure 11.6 does not have a bell shape and does not support the assumption of a Normal distribution.

8.3 According to Chebyshev's inequality,

$$\begin{aligned}\boldsymbol{P}\{X \geq 127.78\} &= \boldsymbol{P}\{X - 48.2333 \geq 79.5467\} \\ &\leq \boldsymbol{P}\left\{\frac{|X - 48.2333|}{26.517} \geq \frac{79.5467}{26.517}\right\} \\ &= \boldsymbol{P}\left\{\frac{|X-\mu|}{\sigma} \geq 3\right\} \leq \frac{1}{9}\end{aligned}$$

Thus, $\boldsymbol{P}\{X < 127.78\} \geq 8/9$.

8.4 Standard Normal quartiles equal about ± 0.675 (Table A4), and the Standard Normal IQR is $2(0.675) = 1.35$.

For a Normal(μ, σ) random variable, the quartiles equal $\mu \pm 0.675\sigma$, and the interquartile range equals 1.35σ. The probability to take a value within 1.5 interquartile ranges from its quartiles equals

$$\begin{aligned}&\boldsymbol{P}\{Q_1 - 1.5(IQR) \leq X \leq Q_3 + 1.5(IQR)\} \\ &= \boldsymbol{P}\left\{-0.675 - 1.5(1.35) \leq \frac{X-\mu}{\sigma} \leq 0.675 + 1.5(1.35)\right\} \\ &= \boldsymbol{P}\{-2.70 \leq Z \leq 2.70\} \\ &= \Phi(2.70) - \Phi(-2.70) = 2(0.0035) = \boxed{0.0070}\end{aligned}$$

8.5 The time plot is on Figure 11.7, left. There is a steady nonlinear growth of the population although during the last century it may have become linear. The U.S. population increases every decade.

8.6 (a) The sample mean is 13.2143 (mln), the sample median is 12.8 (mln), and the variance is 82.5143 (mln^2).

(b) The time plot is on Figure 11.7, right. It shows an increasing trend until year 1960. After 1960, the increments in the total population are decreasing. The population grows every decade, but after 1960 its growth is slowing. Due to an obvious time trend, the data do not follow the same distribution, therefore, the sample mean and sample variance in (a) are not unbiased estimates of the population means and variances. On the average, the U.S. population has increased by 13.2143 mln. people every 10 years.

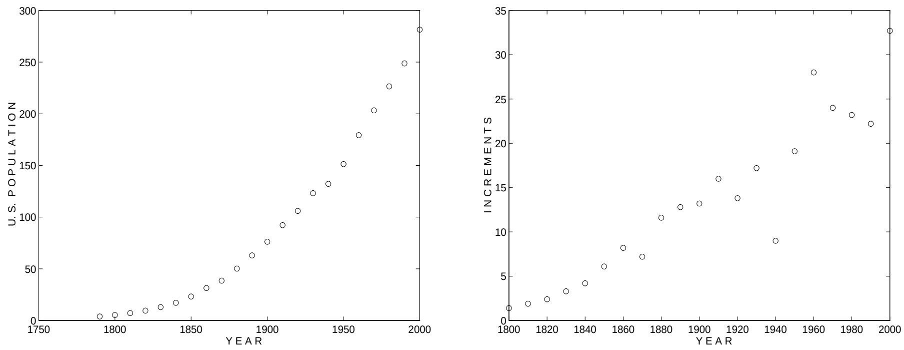


Figure 11.7 *Time plots for Exercises 8.5 and 8.6*

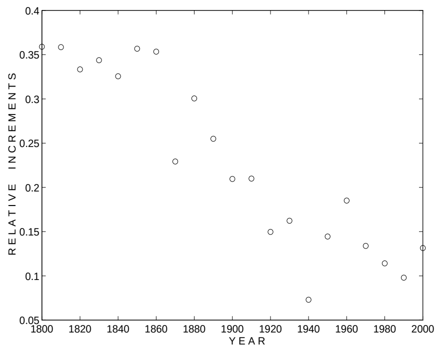


Figure 11.8 *Time plot for Exercise 8.7*

8.7 (a) The sample mean is 0.2298, the sample median is 0.21, and the variance is 0.0099. That is, the U.S. population has increased by 22.98%, on the average, every 10 years.

(b) The time plot is on Figure 11.8. We see that in general, the proportional (percentage) population change decreases, and the trend is almost linear. In 1800–1860, the population increased by about 35% each decade whereas it never exceeded 20% since 1920.

(c) An increasing trend of the absolute increments in Exercise 8.6 corresponds to a decreasing trend of the relative increments in Exercise 8.7. We should expect a rather strong *negative correlation*: large absolute increments correspond to small relative increments, and vice versa.

Indeed, the correlation coefficient equals -0.7683.

A steady reduction of relative increments is not surprising. Had the population increased by a constant percent every decade, it would have grown exponentially fast. However, the human population does not grow at an exponential rate.

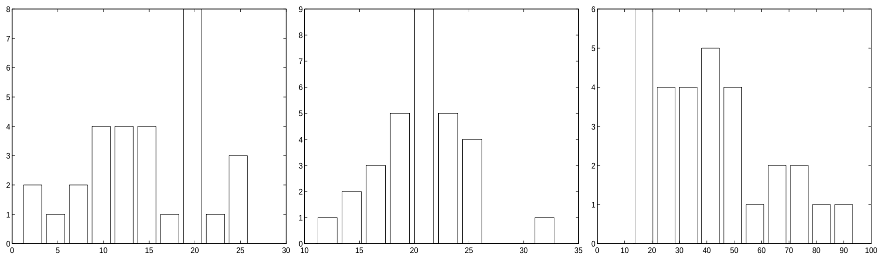


Figure 11.9 *Histograms of three data sets in Exercise 8.8*

On the other hand, had the population increased by a constant absolute number every decade, it would have grown linearly. The increments increase, therefore the population grows faster than a linear function.

Thus, increasing absolute increments and decreasing relative increments show that the U.S. population growth is faster than linear but slower than exponential. As we'll see in Section 10.3, a *quadratic* model is rather accurate for the U.S. population.

8.8 (a) Histograms of the given data sets are shown in Figure 11.9. The first distribution is slightly left-skewed, the second is symmetric, and the third is right-skewed.

(b) Set 1: $\bar{X} = 14.9667$, $\hat{M} = 15.5$. As expected for a left-skewed distribution, $\bar{X} < \hat{M}$.

Set 2: $\bar{X} = 20.8333$, $\hat{M} = 21.0$. As expected for a symmetric distribution, $\bar{X} \approx \hat{M}$.

Set 3: $\bar{X} = 41.3$, $\hat{M} = 39.5$. As expected for a right-skewed distribution, $\bar{X} > \hat{M}$.

8.9 (a) $\bar{X} = 49.6$, $\hat{M} = 47.5$, $\hat{Q}_1 = 43$, $\hat{Q}_3 = 52$, and $s = 23.4767$. For a sample of size 10, any number between the 5th and the 6th smallest observations is a median, the 3rd smallest observation is $\hat{Q}_1$, and the 3rd largest observation is $\hat{Q}_3$.

(b) Compute $IQR = \hat{Q}_3 - \hat{Q}_1 = 9$, $\hat{Q}_1 - 1.5(IQR) = 29.5$, and $\hat{Q}_3 + 1.5(IQR) = 65.5$. One observation, 105, is far outside of the 1.5(IQR) range. So many registered new accounts during the sixth day is likely an outlier.

(c) Without 105, we have $\bar{X} = 43.4444$, $\hat{M} = 45$, $\hat{Q}_1 = 43$, $\hat{Q}_3 = 51$, and $s = 13.9204$. Now, for a sample of size 9, the median is the 5th smallest observation, the 3rd smallest is $\hat{Q}_1$, and the 3rd largest observation is $\hat{Q}_3$.

(d) When we deleted the outlier, the mean and the standard deviation decreased significantly. Especially the standard deviation because variation of the data is much smaller without $x_6 = 105$. Quartiles did not change that much. They are *robust* measures, not very sensitive to outliers.

Chapter 9

9.1 (a) Compute

$$\mu_1 = \mathbf{E}(X) = \sum xP(x) = 3\theta + 7(1-\theta) = 7 - 4\theta$$

and

$$m_1 = \bar{X} = \frac{3+3+3+3+3+7+7+7}{8} = 4.5.$$

Solve the equation

$$\mu_1 = m_1 \quad \Rightarrow \quad 7 - 4\theta = 4.5 \quad \Rightarrow \quad \theta = \frac{7-4.5}{4} \quad \Rightarrow \quad \boxed{\hat{\theta} = 0.625}$$

(b) Write the joint probability mass function

$$P(3,3,3,3,3,7,7,7) = \theta \cdot \theta \cdot \theta \cdot \theta \cdot \theta \cdot (1-\theta)(1-\theta)(1-\theta) = \theta^5(1-\theta)^3$$

and maximize it in θ. Take logarithm,

$$\ln P(\boldsymbol{x}) = 5\ln\theta + 3\ln(1-\theta).$$

Differentiate, equate the derivative to 0 and solve for θ,

$$\frac{\partial}{\partial\theta}\ln P(\boldsymbol{x}) = \frac{5}{\theta} - \frac{3}{1-\theta} = 0 \quad \Rightarrow \quad 5 - 5\theta = 3\theta \quad \Rightarrow \quad \boxed{\hat{\theta} = 5/8 \text{ or } 0.625}$$

It turns out in this problem that the method of moments and the maximum likelihood estimators coincide.

A shortcut can be used here. Define

$$Y = \begin{cases} 1 & \text{if} \quad X = 3 \\ 0 & \text{if} \quad X = 7 \end{cases}$$

Then estimate parameter θ from the Y-sample $(1,1,1,1,1,0,0,0)$ taken from Bernoulli(θ) distribution.

9.2 (a) For this Geometric sample, $\mu_1 = 1/p$ and $\bar{X} = 4$. Solving the equation

$$\frac{1}{p} = 4,$$

we obtain $\boxed{\hat{p} = 0.25}$.

(b) The joint pmf of $\boldsymbol{x}$ is

$$P(\boldsymbol{x}) = \prod_{i=1}^{5} p(1-p)^{x_i - 1} = p^5(1-p)^{\sum x_i - 5} = p^5(1-p)^{15}$$

Then

$$\ln P(\boldsymbol{x}) = 5\ln p + 15\ln(1-p),$$

and the maximum likelihood estimator is obtained by solving the equation

$$\frac{\partial}{\partial p}\ln P(\boldsymbol{x}) = \frac{5}{p} - \frac{15}{1-p} = 0 \;\Rightarrow\; 5 - 5p = 15p,$$

so that $\hat{p} = 5/20 = \boxed{0.25}$.

9.3 (a) Method of moments. Use two moments, $\mu_1 = (a+b)/2$ and, for example,

$$\mu_2' = \mathrm{Var}(X) = (b-a)^2/12.$$

Find $\hat{a}$ and $\hat{b}$ by solving the system

$$\begin{cases} \dfrac{a+b}{2} & = & \bar{X} \\ \dfrac{(b-a)^2}{12} & = & S^2 \end{cases} \Rightarrow \begin{cases} a+b & = & 2\bar{X} \\ b-a & = & S\sqrt{12} \end{cases}$$

$$\Rightarrow \boxed{\hat{a} = \frac{2\bar{X} - S\sqrt{12}}{2}, \quad \hat{b} = \frac{2\bar{X} + S\sqrt{12}}{2}}$$

Maximum likelihood. The joint density

$$f(\boldsymbol{x}) = \begin{cases} \left(\dfrac{1}{b-a}\right)^n & \text{if } a \le x_1, \ldots, x_n \le b \\ 0 & \text{otherwise} \end{cases}$$

is monotonically increasing in a and decreasing in b. is maximized at the largest value of a and the smallest value of b where this density is not 0. These are

$$\boxed{\hat{a} = \min(x_i) \text{ and } \hat{b} = \max(x_i)}$$

(b) Method of moments. Equate $\mu_1 = 1/\lambda$ to $\bar{X}$ and solve: $\boxed{\hat{\lambda} = 1/\bar{X}}$.

Maximum likelihood. Start with the joint density

$$f(\boldsymbol{x}) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i},$$

then

$$\ln f(\boldsymbol{x}) = \sum_{i=1}^{n} (\ln \lambda - \lambda x_i),$$

$$\frac{\partial}{\partial \lambda} \ln f(\boldsymbol{x}) = \sum_{i=1}^{n} \left(\frac{1}{\lambda} - x_i\right) = \frac{n}{\lambda} - \sum x_i = 0,$$

(find roots of the derivative in order to maximize the density) so that $\hat{\lambda} = n/\sum x_i = \boxed{1/\bar{X}}$.

(c) Method of moments. Equate μ to $\bar{X}$ and trivially obtain $\boxed{\hat{\mu} = \bar{X}}$.

Maximum likelihood. The joint density is

$$f(\boldsymbol{x}) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x_i - \mu)^2}{2\sigma^2}\right\}.$$

Then

$$\begin{aligned}\ln f(\boldsymbol{x}) &= -n\ln(\sigma\sqrt{2\pi}) - \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\sigma^2} \qquad (11.1)\\ &= -n\ln(\sigma\sqrt{2\pi}) - \frac{n\mu^2 - 2\sum x_i\mu + \sum x_i^2}{2\sigma^2}\end{aligned}$$

is a parabola in terms of μ, and it is maximized at $\hat{\mu} = \sum x_i/n = \boxed{\bar{X}}$.

(d) Method of moments. The first population moment is not a function of σ, thus we equate the second (central, for simplicity) moment σ^2 to the second moment S^2 and obtain $\boxed{\hat{\sigma} = S}$.
Since μ is known, we can also use the estimator

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$$

instead of S^2.
Maximum likelihood. Differentiate $\ln f(\boldsymbol{x})$ in (11.1) with respect to σ,

$$\frac{\partial}{\partial\sigma}\ln f(\boldsymbol{x}) = -\frac{n}{\sigma} + \frac{\sum(x_i-\mu)^2}{\sigma^3} = 0,$$

(find roots of the derivative in order to maximize the density) so that $\boxed{\hat{\sigma} = \sqrt{\frac{\sum(x_i-\mu)^2}{n}}}$.

(e) Method of moments. When both μ and σ are unknown, we equate the first and second (central, for simplicity) moments and trivially obtain $\boxed{\hat{\mu} = \bar{X}}$ and $\hat{\sigma}^2 = S^2$ so that $\boxed{\hat{\sigma} = S}$.
Maximum likelihood. First, maximizing (11.1) in μ, we again obtain $\boxed{\hat{\mu} = \bar{X}}$, regardless of the value of σ. Then, substitute the obtained maximizer into $\ln f(\boldsymbol{x})$ for the unknown value of μ and maximize the resulting function in terms of σ. We get the same answer as in (d) with the unknown parameter μ replaced by $\bar{X}$, $\boxed{\hat{\sigma} = \sqrt{\frac{\sum(x_i-\bar{X})^2}{n}} = S}$.

9.4 Method of moments. Compute

$$\mu_1 = \mathbf{E}(X) = \int_0^1 xf(x)dx = \int_0^1 \theta x^\theta dx = \left.\frac{\theta x^{\theta+1}}{\theta+1}\right|_{x=0}^{x=1} = \frac{\theta}{\theta+1},$$

equate it to

$$m_1 = \bar{X} = \frac{0.4+0.7+0.9}{3} = \frac{2}{3}$$

and solve for θ,

$$\frac{\theta}{\theta+1} = \frac{2}{3} \Longrightarrow \boxed{\hat{\theta} = 2}$$

Method of maximum likelihood. The joint density is

$$f(\boldsymbol{x}) = \prod_{i=1}^{3} \theta x_i^{\theta-1}$$

Take logarithm,

$$\ln f(\boldsymbol{x}) = \sum_{i=1}^{3} \{\ln\theta + (\theta-1)\ln x_i\} = 3\ln\theta + (\theta-1)\sum_{i=1}^{3}\ln x_i$$

Take derivative, equate to 0, and solve for θ,

$$\frac{\partial \ln f}{\partial\theta} = \frac{3}{\theta} + \sum_{i=1}^{3}\ln x_i = 0,$$

$$\hat\theta = -3/\sum_{i=1}^{3}\ln x_i = -3/(\ln 0.4 + \ln 0.7 + \ln 0.9) = \boxed{2.1766}$$

9.5 The first moment can be computed by integration,

$$\begin{aligned}\mu &= \int x f(x)dx = \frac{1}{2}\int_0^\infty \left(\frac{1}{\theta}xe^{-x/\theta} + \frac{1}{10}xe^{-x/10}\right) \\ &= \frac{1}{2}\left(\frac{1}{\theta}(\theta^2\Gamma(2)) + \frac{1}{10}(10^2\Gamma(2))\right) = \frac{1}{2}(\theta+10).\end{aligned}$$

Alternatively, we can notice that the given density is a *mixture* (Section **??**) of two Exponential distributions. A random variable that has such a density is Exponential($1/\theta$) with probability 0.5 and Exponential(1/10) with probability 0.5. Therefore, its expectation equals $0.5\theta + 0.5(10)$.

Then we solve the equation

$$\mu = \bar X \;\Rightarrow\; \frac{1}{2}(\theta+10) = \frac{150}{10} \;\Rightarrow\; \boxed{\hat\theta = 20}$$

9.6 Unbiasedness of s_p^2 follows from unbiasedness of sample variances s_x^2 and s_y^2 (proved in Section **??**):

$$\mathbf{E}(s^2) = \mathbf{E}\left(\frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}\right) = \frac{(n-1)\sigma^2 + (m-1)\sigma^2}{n+m-2} = \sigma^2.$$

9.7 Under the null hypotheses stated in Table 9.1,

$$\mathbf{E}(\bar X) = \mu = \mu_0, \quad \mathbf{E}(\hat p) = p = p_0,$$

$$\mathbf{E}(\bar X - \bar Y) = \mu_X - \mu_Y = D, \quad \text{and} \quad \mathbf{E}(\hat p_1 - \hat p_2) = p_1 - p_2 = D.$$

This verifies column 3. Next, compute variances,

$$\mathrm{Var}(\bar X) = \frac{\sigma^2}{n}, \quad \mathrm{Var}(\hat p) = \frac{p(1-p)}{n} \quad \text{(Section 9.3.2)},$$

$$\mathrm{Var}(\bar{X}-\bar{Y}) = \mathrm{Var}(\bar{X}) + \mathrm{Var}(\bar{Y}) = \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m},$$

and

$$\mathrm{Var}(\hat{p}_1-\hat{p}_2) = \mathrm{Var}(\hat{p}_1) + \mathrm{Var}(\hat{p}_2) = \frac{p_1(1-p_1)}{n} + \frac{p_2(1-p_2)}{m}$$

Finally, Z-statistics in the last column of Table 9.1 are obtained by standardizing parameter estimators under H_0. From each parameter estimator $\hat{\theta}$, its expectation (under H_0) is subtracted, then the result is divided by $\mathrm{Std}(\hat{\theta})$. For tests about proportions, the standard deviation is unknown, and we estimate it by replacing p with $\hat{p}$. Then the resulting Z-statistics have an approximately Normal distribution if sample sizes are sufficiently large.

9.8 (a) $\bar{X} \pm z_{0.05}\frac{\sigma}{\sqrt{n}} = 37.7 \pm (1.645)\frac{9.2}{\sqrt{100}} = \boxed{37.7 \pm 1.5 \text{ or } [36.2, 39.2]}$

(b) Test $H_0 : \mu = 35$ vs $H_A : \mu > 35$. Reject H_0 if the test statistic $Z > z_{0.01} = 2.326$. The observed test statistic is

$$Z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} = \frac{37.7-35}{9.2/\sqrt{100}} = 2.9348,$$

belongs to the rejection region. Therefore, reject H_0 in favor of H_A. Yes, these data provide significant evidence that the mean number of concurrent users is greater than 35.

9.9 (a) $\bar{X} \pm z_{0.025}\frac{\sigma}{\sqrt{n}} = 42 \pm (1.96)\frac{5}{\sqrt{64}} = \boxed{42 \pm 1.225 \text{ or } [40.775, 43.225]}$

(b) $\boldsymbol{P}\{40.775 \le X \le 43.225\} = \boldsymbol{P}\left\{\frac{40.775-\mu}{\sigma} \le Z \le \frac{43.225-\mu}{\sigma}\right\}$

$= \boldsymbol{P}\left\{\frac{40.775-40}{5} \le Z \le \frac{43.225-40}{5}\right\}$

$= \Phi(0.645) - \Phi(0.155) = 0.7406 - 0.5616 = \boxed{0.1790}$,

using Table A4.

The individual value, the time of one installation, is not very likely to belong to the computed 95% confidence interval. The interval is computed for the population mean μ. The probability of 0.95 refers to the proportion of confidence intervals, in a long run, that contain μ.

9.10 (a) The standard deviation is unknown. Therefore, the interval is

$$\bar{X} \pm t_{\alpha/2}s/\sqrt{n},$$

where $\alpha = 1 - 0.90 = 0.10$, $n = 3$, $t_{\alpha/2} = t_{0.05} = 2.920$ (with 2 d.f.), $\bar{X} = (30+50+70)/3 = 50$, and

$$s = \sqrt{\frac{(30-50)^2+(50-50)^2+(70-50)^2}{n-1}} = \sqrt{\frac{800}{2}} = 20,$$

Then, the interval is

$$50 \pm 2.920 \frac{20}{\sqrt{3}} = \boxed{50 \pm 33.7 \text{ or } [16.3; 83.7]}$$

(b) Hypothesis $H_0 : \mu = 80$ is *not rejected* against alternative $H_A : \mu \neq 80$ at the 10% level because the 90% confidence interval for μ contains 80.

This is a sufficient explanation, but you may also perform a test,

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{50 - 80}{20/\sqrt{3}} = -2.598$$

It belongs to the acceptance region $[-2.920; 2.920]$, therefore, H_0 is not rejected. The data does *not* provide a significant evidence against H_0.

9.11 (a) Find $\hat{p} = 24/200 = 0.12$. Then for $\alpha = 1 - 0.96 = 0.04$, find $z_{\alpha/2} = z_{0.02} = 2.054$ (the easiest way is to use Table A6 with ∞ degrees of freedom)

$$\begin{aligned} \hat{p} \pm z_{0.02} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} &= 0.12 \pm (2.054) \sqrt{\frac{0.12(1-0.12)}{200}} \\ &= \boxed{0.12 \pm 0.047 \text{ or } [0.073, 0.167]} \end{aligned}$$

(b) Test $H_0 : p \leq 0.1$ (or $H_0 : p = 0.1$) vs $H_A : p > 0.1$. Disproving the manufacturer's claim means rejecting H_0 in favor of this H_A.

The observed test statistic is

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} = \frac{0.12 - 0.1}{\sqrt{\frac{0.12(1-0.12)}{200}}} = 0.8704.$$

In order to consider different significance levels, let us compute the P-value,

$$P = \boldsymbol{P}\{Z > 0.8704\} = 1 - \Phi(0.8704) = 1 - 0.8078 = 0.1922,$$

from Table A4.

The P-value exceeds 0.03 and 0.04. Therefore, we *do not* have a significance evidence, at the mentioned levels, to disprove the manufacturer's claim.

9.12 Test $H_0 : p_1 = p_2$ vs $H_A : p_1 > p_2$. Higher quality means lower proportion of defective items.

Given $\hat{p}_1 = 0.12$ from a sample of size $n = 200$ and $\hat{p}_2 = 13/150 = 0.0867$ from a sample of size $m = 150$, we compute the pooled proportion

$$\hat{p}(\text{pooled}) = \frac{n\hat{p}_1 + m\hat{p}_2}{n + m} = \frac{24 + 13}{200 + 150} = 0.1057.$$

Then, the test statistic is

$$Z = \frac{0.12 - 0.0867}{\sqrt{(0.1057)(1 - 0.1057)\left(\frac{1}{200} + \frac{1}{150}\right)}} = 1.0027$$

Finally, we compute the P-value

$$P = \boldsymbol{P}\{Z > 1.0027\} = 1 - 0.8413 = 0.1587$$

(Table A4), it is rather large, and we conclude that there is *no significance evidence* that the quality of items produced by the new supplier is higher than the quality of items in Exercise 9.11.

9.13 (a) The standard deviation is known, therefore we construct a Z-interval.

$$\bar{X} \pm z_{0.025}\frac{\sigma}{\sqrt{n}} = 0.62 \pm (1.96)\frac{0.2}{\sqrt{52}} = \boxed{0.62 \pm 0.054 \text{ or } [0.566, 0.674]}$$

(b) If $\mu = 0.6$, then

$$\boldsymbol{P}\left\{\bar{X} \geq 0.62\right\} = \boldsymbol{P}\left\{\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \geq \frac{0.62 - 0.6}{0.2/\sqrt{52}}\right\} = \boldsymbol{P}\{Z \geq 0.721\}$$

$$= 1 - \Phi(0.721) = 1 - 0.7642 = \boxed{0.2358},$$

from Table A4.

9.14 (a) To see if there is *any* significant difference between servers A and B, we test $H_0 : \mu_A = \mu_B$ (or $\mu_A - \mu_B = 0$) vs $H_A : \mu_A \neq \mu_B$. The 95% confidence interval in Example 9.19 on p. 281 is $[-1.4, -0.2]$. It does not contain the value 0 that we are testing, hence the difference between the two servers *is significant* at the 5% level.

(b) The test statistic is already computed in Example 9.28 on p. 295, and it equals -2.7603. The P-value for this two-sided test is

$$P = 2\boldsymbol{P}\{t > |-2.7603|\} \boxed{\text{is between 0.01 and 0.02}}$$

(Table A6 with 25 degrees of freedom, already computed by Satterthwaite approximation in Example 9.19).

We conclude that there is a significant difference between servers A and B at a level of 2% or higher, and the difference is not significant at a level of 1% or lower.

(c) A faster server should have a shorter execution time. Thus we test $H_0 : \mu_A = \mu_B$ vs $H_A : \mu_A < \mu_B$. For this one-sided test, the P-value equals

$$P = \boldsymbol{P}\{t < -2.7603\} \boxed{\text{is between 0.005 and 0.01}}$$

This is rather significant. At a 1% level of significance and any level above that, we have a significant evidence that server A is faster than server B.

9.15 To see if there is significant difference between the two towns, test H_0 : $p_1 = p_2$ vs $H_A : p_1 \neq p_2$.

As in Example 9.15, $n_1 = 70$, $n_2 = 100$, $\hat{p}_1 = 0.6$, and $\hat{p}_2 = 0.59$. Compute the pooled proportion,

$$\hat{p}(\text{pooled}) = \frac{(70)(0.6) + (100)(0.59)}{70 + 100} = 0.5941.$$

Then the test statistic is

$$Z = \frac{0.6 - 0.59}{\sqrt{(0.5941)(1 - 0.5941)\left(\frac{1}{70} + \frac{1}{100}\right)}} = 0.1307$$

The P-value equals

$$P = 2\boldsymbol{P}\{Z > |0.1307|\} = 2(1 - 0.5517) = \boxed{0.8966}$$

(Table A4). This is a very high P-value, thus there is no significant difference between the support of the candidate in the two towns.

9.16 Here $n_1 = 250$, $n_2 = 300$, $\hat{p}_1 = 10/250 = 0.04$, and $\hat{p}_2 = 18/300 = 0.06$.

(a) A 98% confidence interval for $p_1 - p_2$ is

$$\begin{aligned} & \hat{p}_1 - \hat{p}_2 \pm z_{0.02/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \\ = \; & 0.04 - 0.06 \pm 2.326\sqrt{\frac{(0.04)(0.96)}{250} + \frac{(0.06)(0.94)}{300}} \\ = \; & \boxed{-0.02 \pm 0.043 \text{ or } [-0.063, 0.023]} \end{aligned}$$

(b) The null hypothesis $H_0 : p_1 = p_2$ is *not rejected* against the two-sided alternative $H_A : p_1 \neq p_2$ ($p_1 - p_2 = 0$) at the 2% level because the 98% confidence interval for $p_1 - p_2$ contains 0. No, there is no significant difference between the quality of the two lots.

9.17 For the test of $H_0 : \mu = 5000$ against $H_A : \mu > 5000$, the test statistic $Z_{\text{obs}} = 2.5$ is computed in Example 9.23. Based on it, compute the P-value

$$P = \boldsymbol{P}\{Z > 2.5\} = 1 - \Phi(2.5) = 1 - 0.9938 = \boxed{0.0062}$$

The P-value is rather low, hence it does indicate a significant increase in the number of concurrent users.

9.18 From the data in Exercise 8.1, compute $\bar{X} = 50$, $\bar{Y} = 40.2$, $s_X^2 = 58$, $s_Y^2 = 63.33$, and the pooled variance estimator

$$s_p^2 = \frac{(n-1)s_X^2 + (m-1)s_Y^2}{n + m - 2} = 61.1625.$$

Sample sizes are $n = 14$ and $m = 20$. The pooled variance estimator is used because of the assumption of equal variances.

(a) Samples are not large, therefore we use the T-distribution with $n+m-2 = 32$ degrees of freedom. The 95% confidence interval for $\mu_1 - \mu_2$ is

$$\begin{aligned}
&\bar{X} - \bar{Y} \pm t_{0.05/2}\sqrt{s_p^2\left(\frac{1}{n} + \frac{1}{m}\right)} \\
&= \quad 50 - 40.2 \pm (2.037)\sqrt{(61.1625)\left(\frac{1}{14} + \frac{1}{20}\right)} \\
&= \quad \boxed{9.8 \pm 5.55 \text{ or } [4.25, 15.35]}
\end{aligned}$$

(b) Test the null hypothesis $H_0 : \mu_1 = \mu_2$ against the alternative hypothesis $H_A : \mu_1 > \mu_2$ that reflects reduction in the rate of intrusion attempts.
Assuming equal variances: the test statistic is

$$t = \frac{\bar{X} - \bar{Y}}{\sqrt{s_p^2\left(\frac{1}{n} + \frac{1}{m}\right)}} = \frac{50 - 40.2}{\sqrt{(61.1625)\left(\frac{1}{14} + \frac{1}{20}\right)}} = 3.5960.$$

Using Table A6 with $n - m + 2 = 32$ degrees of freedom, obtain

$$P = \boldsymbol{P}\{t > 3.5960\} = \boxed{\text{between } 0.0005 \text{ and } 0.001}$$

Not assuming equal variances: the test statistic is

$$t = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_X^2}{n} + \frac{s_Y^2}{m}}} = \frac{50 - 40.2}{\sqrt{\frac{58}{14} + \frac{63.33}{20}}} = 3.6248.$$

In this case, degrees of freedom are computed by Satterthwaite approximation,

$$\nu = \frac{\left(\frac{s_X^2}{n} + \frac{s_Y^2}{m}\right)^2}{\frac{s_X^4}{n^2(n-1)} + \frac{s_Y^4}{m^2(m-1)}} \approx 29.$$

Using Table A6 with 29 degrees of freedom, obtain

$$P = \boldsymbol{P}\{t > 3.6248\} = \boxed{\text{between } 0.0005 \text{ and } 0.001}$$

In both cases, there is a very significant evidence that the average number of intrusion attempts per day has decreased after the change of firewall settings.

9.19 After dropping the constant coefficients, the Inverse Gamma prior density of θ is expressed as

$$\pi(\theta) \sim \frac{e^{-1/\theta\beta}}{\theta^{\alpha+1}}$$

(a) A sample of size n from Exponential(θ^{-1}) distribution has density

$$f(\boldsymbol{x}|\theta^{-1}) = \left(\theta^{-1}\right)^n e^{-\theta^{-1}\sum x_i}$$

Then the posterior distribution is

$$\begin{aligned}\pi(\theta|x) &\sim \pi(\theta)f(\boldsymbol{x}|\theta^{-1}) \sim \frac{e^{-1/\theta\beta}}{\theta^{\alpha+1}} \cdot \frac{e^{-\sum x_i/\theta}}{\theta^n} \\ &= \frac{e^{-(\sum x_i+1/\beta)/\theta}}{\theta^{n+\alpha+1}},\end{aligned}$$

which is the Inverse Gamma density with parameters

$$\boxed{\alpha_x = n + \alpha \text{ and } \beta_x = \frac{1}{\sum x_i + 1/\beta} = \frac{\beta}{\beta\sum x_i + 1}}$$

(b) Generalizing to Gamma(r, θ^{-1}) distribution (and dropping constant coefficients again), we write the marginal density

$$f(x|\theta^{-1}) = \left(\theta^{-1}\right)^r e^{-\theta^{-1}x_i},$$

and for a sample of size n,

$$f(\boldsymbol{x}|\theta^{-1}) = \left(\theta^{-r}\right)^n e^{-\sum x_i/\theta}$$

Then the posterior distribution is

$$\begin{aligned}\pi(\theta|x) &\sim \pi(\theta)f(\boldsymbol{x}|\theta^{-1}) \sim \frac{e^{-1/\theta\beta}}{\theta^{\alpha+1}} \cdot \frac{e^{-\sum x_i/\theta}}{\theta^{rn}} \\ &= \frac{e^{-(\sum x_i+1/\beta)/\theta}}{\theta^{rn+\alpha+1}},\end{aligned}$$

where we recognize the Inverse Gamma density with parameters

$$\boxed{\alpha_x = rn + \alpha \text{ and } \beta_x = \frac{1}{\sum x_i + 1/\beta} = \frac{\beta}{\beta\sum x_i + 1}}$$

Posterior distribution belongs to the Inverse Gamma family showing that it is conjugate to the Gamma(r, θ^{-1}) model.

(c) Using our results in (a), compute posterior parameters

$$\alpha_x = n + \alpha = 5 + 3 = 8 \text{ and } \beta_x = \frac{\beta}{\beta\sum x_i + 1} = \frac{3}{(3)(31)+1} = \frac{3}{94}$$

The Bayes estimator of θ is

$$\hat{\theta}_{\mathrm{B}} = \mathbf{E}\{\theta|\boldsymbol{x}\} = \frac{1}{\beta_x(\alpha_x - 1)} = \frac{94/3}{7} = \boxed{94/21 \text{ or } 4.476}$$

(d) Posterior risk of the Bayes estimator equals posterior variance

$$\mathrm{Var}(\theta|\boldsymbol{x}) = \frac{1}{\beta_x^2(\alpha_x - 1)^2(\alpha_x - 2)} = \frac{(94/3)^2}{(7)^2(6)} = \boxed{3.34}$$

9.20 In the solution to Exercise 8.2, we computed $\bar{X} = 17.9540$ (thousands of people). Also, we are given $n = 50$, $\sigma = 4$ and the Normal($\mu = 14, \tau = 2$) prior distribution of θ.

(a) The Bayes estimator of θ is

$$\hat{\theta}_B = \frac{n\bar{X}/\sigma^2 + \mu/\tau^2}{n/\sigma^2 + 1/\tau^2} = \frac{(50)(17.9540)/4^2 + 14/2^2}{50/4^2 + 1/2^2} = \boxed{17.6611}$$

(b) The 90% HPD credible set for θ is

$$\begin{aligned} \mu_x \pm z_{0.05}\tau_x &= 17.6611 \pm (1.645)(0.5443) \\ &= \boxed{17.6611 \pm 0.8954 \text{ or } [16.7657, 18.5565]} \end{aligned}$$

where $\mu_x = \hat{\theta}_B = 17.6611$ and τ_x is the posterior standard deviation

$$\tau_x = \frac{1}{\sqrt{n/\sigma^2 + 1/\tau^2}} = \frac{1}{\sqrt{50/4^2 + 1/2^2}} = 0.5443.$$

Interpretation: given the observed data and the assumed prior distribution, there is a 90% posterior probability that the parameter θ belongs to the interval $[16.7657, 18.5565]$.

(c) Test $H_0 : \theta \le 16$ vs $H_A : \theta > 16$. Given the Normal($\mu_x = 17.6611, \tau_x = 0.5443$) posterior distribution of θ, compute posterior probabilities

$$\begin{aligned} \boldsymbol{P}\{H_0|\boldsymbol{x}\} &= \boldsymbol{P}\{\theta \le 16|\boldsymbol{x}\} = \Phi\left(\frac{16 - 17.6611}{0.5443}\right) \\ &= \Phi(-3.0518) = 0.0011 \end{aligned}$$

and $\boldsymbol{P}\{H_A|\boldsymbol{x}\} = 1 - \boldsymbol{P}\{H_0|\boldsymbol{x}\} = 0.9989$. Posterior probability of the alternative is extremely high, and we conclude that there is a significant evidence that the mean number of concurrent users exceeds 16,000.

9.21 (a) The maximum likelihood and the method of moments estimator is

$$\bar{X} = \boxed{17.9540}$$

(b) $\bar{X} \pm \dfrac{z_{0.05}\sigma}{\sqrt{n}} = 17.9540 \pm \dfrac{(1.645)(4)}{\sqrt{50}}$

$= \boxed{17.9540 \pm 0.9306 \text{ or } [17.0234, 18.8846]}$

Interpretation: in a long run of independent random samples and confidence intervals computed from them by the above procedure, 90% of the obtained confidence intervals will contain the parameter θ.

(c) Test $H_0 : \theta \le 16$ vs $H_A : \theta > 16$. Compute the Z-statistic

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{17.9540 - 16}{4/\sqrt{50}} = 3.4542.$$

This yields the P-value

$$P = \boldsymbol{P}\{Z > 3.4542\} = 1 - 0.9997 = 0.0003,$$

which gives a strong support of H_A and a strong evidence that the mean number of concurrent users exceeds 16,000.

(d) Comparing results of Exercises 9.20 and 9.21, we see that the Bayes estimator is slightly shifted towards the prior mean.

The 90% credible set is also shifted towards the prior mean comparing with the 90% confidence interval. It has a lower margin because in addition to the observed sample, it uses the prior information that makes $\tau_x < \mathrm{Std}(\bar{X}) = \sigma/\sqrt{n}$. The obtained intervals have totally different meanings. In the Bayesian setting, the unknown parameter has a distribution, therefore, it is meaningful to state that it belongs to the obtained interval $[16.7657, 18.5565]$ with probability 90%. In a non-Bayesian setting, the parameter is non-random, and the 90% probability refers only to a long run of random samples.

Both approaches give a strong support to the alternative.

9.22 (a) Notice that such a prior distribution is not unique because we can choose two parameters μ and τ of the Normal prior distribution (Normal family is conjugate to the given Normal model), and we have to satisfy only one condition,

$$\boldsymbol{P}\{5.0 < \theta < 6.0\} = 0.95.$$

One solution is to write $[5.0, 6.0] = \mu \pm z_{0.025}\tau$, letting $\mu = 5.5$ be the middle of the interval, and choosing τ such that

$$6.0 - 5.0 = 1.0 = 2z_{0.025}\tau = 2(1.96)\tau \;\Rightarrow\; \tau = \frac{1.0}{2(1.96)} = 0.255.$$

So, let $\boxed{\mu = 5.5 \text{ and } \tau = 0.255}$.

(b) In Example 9.11, we computed $\bar{X} = 6.5$. Also, we were given $n = 6$ and $\sigma = 2.2$. Then the posterior distribution is Normal(μ_x, τ_x) with parameters

$$\begin{aligned}
\mu_x &= \frac{n\bar{X}/\sigma^2 + \mu/\tau^2}{n/\sigma^2 + 1/\tau^2} = \frac{(6)(6.5)/2.2^2 + 5.5/0.255^2}{6/2.2^2 + 1/0.255^2} = 5.575 \\
\tau_x^2 &= \frac{1}{n/\sigma^2 + 1/\tau^2} = \frac{1}{6/2.2^2 + 1/0.255^2} = 0.060.
\end{aligned}$$

The posterior distribution is $\boxed{\text{Normal}(5.575,\ 0.060)}$.

The Bayes estimator is

$$\hat{\theta}_{\mathrm{B}} = \mu_x = \boxed{5.575}$$

Its Bayes risk of μ_x equals $\mathbf{E}\,\mathbf{E}(\mu_x - \theta)^2$, where the expectation $\mathbf{E}\,\mathbf{E}$

is taken with respect to the joint distribution of $\boldsymbol{X}$ and θ. Recall from (9.23) that the joint density of θ and $\boldsymbol{X}$ equals

$$f(\boldsymbol{x}|\theta)\pi(\theta) = \pi(\theta|\boldsymbol{x})m(\boldsymbol{x}).$$

Then the Bayes risk can also be computed as $\mathbf{E}^X\,\mathbf{E}^\theta_X(\mu_x - \theta)^2$, where the expectation $\mathbf{E}^X$ is taken with respect to the *marginal distribution* of X, and the expectation $\mathbf{E}^\theta_X$ is taken with respect to the *posterior distribution* of θ given $\boldsymbol{X}$.

That is,

$$\text{Bayes risk} = \mathbf{E}^X\,\mathbf{E}^\theta_X(\mu_x - \theta)^2 = \mathbf{E}^X\tau_x^2 = \tau_x^2 = \boxed{\frac{1}{n/\sigma^2 + 1/\tau^2}}$$

because τ_x^2 is a non-random constant.

(c) The 95% HPD credible set for μ is

$$\mu_x \pm z_{0.025}\tau_x = 5.575 \pm (1.96)(0.060) = \boxed{5.575 \pm 0.117 \text{ or } [5.458, 5.692]}$$

The 95% confidence interval in Example 9.11 is much wider. It is based on data with a rather high standard deviation $\sigma = 2.2$. The credible set also uses a rather informative prior distribution whose low standard deviation $\tau = 0.255$ does not allow high variability of θ.

9.23 Let θ be the probability of heads. The probability of ten heads in a row is θ^{10}. A fair coin should have $\theta = 0.5$.

Choose a prior distribution $\pi(\theta)$ with $\pi(0.5) = 0.99$. For example, let

$$\theta = \begin{cases} 0.5 & \text{with probability } 0.99 \\ \text{a Uniform(0,1) random variable} & \text{with probability } 0.01 \end{cases}$$

Then the posterior probability of a fair coin is

$$\begin{aligned}
\pi(\theta = 0.5|\boldsymbol{x}) &= \frac{f(\boldsymbol{x}|\theta = 0.5)\pi(0.5)}{m(\boldsymbol{x})} = \frac{0.5^{10}(0.99)}{0.5^{10}(0.99) + \int_0^1 \theta^{10}d\theta(0.01)} \\
&= \frac{0.5^{10}(0.99)}{0.5^{10}(0.99) + (1/11)(0.01)} = \boxed{0.5154}
\end{aligned}$$

We aren't so sure now that the coin is fair, are we?

9.24 (a) A Uniform$(0,\theta)$ sample has density

$$f(\boldsymbol{x}|\theta) = \left(\frac{1}{\theta}\right)^n, \qquad 0 < x_1, \ldots, x_n < \theta.$$

A conjugate family of densities $\pi(\theta)$ should also contain a negative power of θ. Such a family is $\boxed{\text{Pareto}(\alpha, \sigma)}$.

Indeed, Pareto density is

$$\pi(\theta) \sim \theta^{-\alpha-1}, \qquad \theta > 0.$$

With such a prior, the posterior distribution is

$$\pi(\theta|\boldsymbol{x}) \sim f(\boldsymbol{x}|\theta)\pi(\theta) \sim \left(\frac{1}{\theta}\right)^n \theta^{-\alpha-1} = \theta^{-n-\alpha-1},$$

which is Pareto$(n+\alpha, \sigma)$. It belongs to the same Pareto family, therefore, this family is conjugate to the Uniform model.

(b) See description of Pareto distribution in the Appendix. The Bayes estimator is the posterior expectation,

$$\hat{\theta}_{\mathrm{B}} = \mathbf{E}(\theta|\boldsymbol{x}) = \frac{(n+\alpha)\sigma}{n+\alpha-1}$$

Its posterior risk equals the posterior variance,

$$\mathrm{Var}(\theta|\boldsymbol{x}) = \frac{(n+\alpha)\sigma^2}{(n+\alpha-1)^2(n+\alpha-2)}$$

This posterior risk does not depend on $\boldsymbol{x}$. Therefore, it equals its expectation with respect to the marginal distribution of $\boldsymbol{x}$, which is the Bayes risk

$$\mathbf{E}\,\mathbf{E}(\hat{\theta}_{\mathrm{B}} - \theta)^2 = \mathbf{E}^X\,\mathbf{E}^\theta_X(\hat{\theta}_{\mathrm{B}} - \theta)^2 = \frac{(n+\alpha)\sigma^2}{(n+\alpha-1)^2(n+\alpha-2)}$$

Chapter 10

10.1 We are given: $n = 30$, $\bar{x} = 126$, $s_x = 35$, $\bar{y} = 0.04$, $s_y = 0.01$, and $r = 0.86$. Compute the least squares estimates,

$$\begin{aligned} b_1 &= r\left(\frac{s_y}{s_x}\right) = (0.86)\left(\frac{0.01}{35}\right) = 0.000246 \\ b_0 &= \bar{y} - b_1\bar{x} = 0.04 - (0.000246)(126) = 0.009. \end{aligned}$$

The fitted regression line has an equation

$$\boxed{y = 0.009 + 0.000246x}$$

The time it takes to transmit a 400 Kbyte file is predicted as

$$\hat{y}_* = 0.009 + 0.000246x_* = 0.009 + (0.000246)(400) = \boxed{0.107 \text{ seconds}}$$

10.2 Here $n = 75$, $\bar{x} = 32.2$, $s_x^2 = 6.4$, $\bar{y} = 8.4$, $s_y^2 = 2.8$, and $s_{xy} = 3.6$.

(a) Compute the least squares estimates

$$\begin{aligned} b_1 &= \frac{s_{xy}}{s_x^2} = \frac{3.6}{6.4} = 0.5625 \\ b_0 &= \bar{y} - b_1\bar{x} = 8.4 - (0.5625)(32.2) = -9.7125. \end{aligned}$$

The sample regression line is

$$\boxed{y = -9.7125 + 0.5625x}$$

(b) Compute the sums of squares,

$$\begin{aligned} SS_{\text{TOT}} &= (n-1)s_y^2 = (75-1)(2.8) = 207.2 \\ SS_{\text{REG}} &= b_1^2 S_{xx} = b_1^2 s_x^2(n-1) = (0.5625)^2(6.4)(75-1) = 149.85 \\ SS_{\text{ERR}} &= SS_{\text{TOT}} - SS_{\text{REG}} = 57.35 \end{aligned}$$

Also, compute degrees of freedom

$$\text{df}_{\text{TOT}} = n-1 = 74, \quad \text{df}_{\text{REG}} = 1, \quad \text{and} \quad \text{df}_{\text{ERR}} = n-2 = 73$$

and the mean squares

$$MS_{\text{REG}} = \frac{SS_{\text{REG}}}{\text{df}_{\text{REG}}} = 149.85, \quad MS_{\text{ERR}} = \frac{SS_{\text{ERR}}}{\text{df}_{\text{ERR}}} = \frac{57.35}{73} = 0.7856.$$

Finally, we compute the F-ratio

$$F = \frac{MS_{\text{REG}}}{MS_{\text{ERR}}} = 190.75$$

and find from Table A5 (1 and 73 d.f.) that it is significant at the 1% level.

Complete the ANOVA table:

Source	Sum of squares	Degrees of freedom	Mean squares	F
Model	149.85	1	149.5	190.75
Error	57.35	73	0.7856	
Total	207.2	74		

Predictor X can explain

$$R^2 = SS_{\text{REG}}/SS_{\text{TOT}} = \boxed{0.7232 \text{ or } 72.32\%}$$

of the total variation.

(c) The 99% confidence interval for β_1 is

$$\begin{aligned} b_1 \pm t_{\alpha/2}\frac{s}{\sqrt{S_{xx}}} &= b_1 \pm t_{0.005}\frac{\sqrt{MS_{\text{ERR}}}}{\sqrt{(n-1)s_x^2}} \\ &= 0.5625 \pm (2.648)\frac{\sqrt{0.7856}}{\sqrt{(74)(6.4)}} \\ &= \boxed{0.5625 \pm 0.1078 \text{ or } [0.4547,\ 0.6703]} \end{aligned}$$

where $t_{0.005}$ is obtained from Table A6 with 73 d.f.

This interval does not contain 0, therefore, the slope is significant (significantly different from 0) at a 1% level of significance.

10.3 In this problem, $n = 180$, $\bar{x} = 24,598$, $s_x = 14,634$, $\bar{y} = 23.8$, $s_y = 3.4$, and $r = -0.17$.

(a) Parameters of the regression line predicting the number of miles per gallon based on the mileage are estimated as

$$\begin{aligned} b_1 &= r\left(\frac{s_y}{s_x}\right) = (-0.17)\left(\frac{3.4}{14,634}\right) = -0.0000395 \\ b_0 &= \bar{y} - b_1\bar{x} = 23.8 - (-0.0000395)(24,598) = 24.77. \end{aligned}$$

The fitted regression line has an equation

$$\boxed{y = 24.77 - 0.0000395x}$$

Negative slope means that cars with higher mileage (older cars) make a smaller number of miles per gallon. According to this model, brand new cars (with 0 mileage) make 24.77 miles per gallon.

(b) $R^2 = r^2 = (-0.17)^2 = 0.0289$. This model can explain only 2.89% of the total variation of the number of miles per gallon. This is a rather poor fit. The number of miles per gallon should probably depend on many factors other than mileage.

(c) The number of miles per gallon is predicted by

$$\hat{y}_* = 24.77 - (0.0000395)(35,000) = \boxed{23.39 \text{ miles per gallon}}$$

For the confidence and prediction intervals, compute the standard deviation

$$\begin{aligned} s &= \sqrt{MS_{\text{ERR}}} = \sqrt{\frac{(1-R^2)SS_{\text{TOT}}}{\text{df}_{\text{ERR}}}} = \sqrt{\frac{(1-R^2)s_y^2(n-1)}{n-2}} \\ &= \sqrt{\frac{(1-0.0289)(3.4)^2(179)}{178}} = 3.36, \end{aligned}$$

find the quantile $t_{0.025} \approx 1.972$ from Table A6 with $n-2 = 178$ d.f., and recall that $S_{xx} = (n-1)s_x^2$.

The 95% prediction interval is

$$y_* \pm t_{0.025}\, s\sqrt{1 + \frac{1}{n} + \frac{(x_* - \bar{x})^2}{S_{xx}}}$$

$$= 23.39 \pm (1.972)(3.36)\sqrt{1 + \frac{1}{180} + \frac{(35,000 - 24,598)^2}{(179)(14,634)^2}}$$

$$\boxed{23.39 \pm 6.65 \text{ or } [16.74,\ 30.04]}$$

The 95% confidence interval for the average number of miles per gallon for all cars with such a mileage is

$$y_* \pm t_{0.025}\, s\sqrt{\frac{1}{n} + \frac{(x_* - \bar{x})^2}{S_{xx}}}$$

$$= 23.39 \pm (1.972)(3.36)\sqrt{\frac{1}{180} + \frac{(35,000 - 24,598)^2}{(179)(14,634)^2}}$$

$$\boxed{23.39 \pm 0.61 \text{ or } [22.78,\ 24.00]}$$

10.4 To simplify calculations, we can define our independent variable X as

$$X = \text{year} - 2000.$$

Then $n = 11$, $\sum X_i = 0$, $S_{xx} = \sum X_i^2 = 110$, $\sum Y_i = 383$, $\bar{Y} = 383/11 = 34.82$, $S_{yy} = \sum(Y_i - \bar{Y})^2 = 841.64$, and

$$S_{xy} = \sum(X_i - \bar{X})(Y_i - \bar{Y}) = \sum(X_i - 0)Y_i = 292.$$

(a) One way is to compute

$$R^2 = \frac{S_{xy}^2}{S_{xx}S_{yy}} = \frac{292^2}{(110)(841.64)} = 0.92.$$

Then

$$s^2 = MS_{\text{ERR}} = \frac{(1 - R^2)S_{yy}}{n-2} = \frac{(1 - 0.92)(841.64)}{9} = \boxed{7.48}$$

(b) Test $H_0 : \beta_1 = 1.8$ (\$1000s) vs $H_A : \beta_1 > 1.8$.
Start with the estimator

$$b_1 = \frac{S_{xy}}{S_{xx}} = \frac{292}{110} = 2.65.$$

Then compute the test statistic

$$t = \frac{b_1}{s/\sqrt{S_{xx}}} = \frac{2.65}{\sqrt{7.48}/\sqrt{110}} = 10.16.$$

The P-value is

$$P = \boldsymbol{P}\{t > 10.16\} < 0.0001$$

(Table A6 with 9 d.f.), indicating a very significant evidence that the investment increases by more than \$ 1,800 every year, on the average.

(c) For year 2009, $x_* = 9$, the predicted investment is

$$y_* = b_0 + b_1 x_* = 34.82 + (2.65)(9) = 58.67$$

thousand dollars, where $b_1 = 2.65$ is computed above, and $b_0 = \bar{y} - b_1\bar{x} = 34.82 - 0 = 34.82$.

The 95% prediction interval for actual investment in 2009 is

$$\begin{aligned} & y_* \pm t_{0.025}\, s\sqrt{1 + \frac{1}{n} + \frac{(x_* - \bar{x})^2}{S_{xx}}} \\ = \ & 58.67 \pm (2.262)\sqrt{7.48}\sqrt{1 + \frac{1}{11} + \frac{9^2}{110}} \\ = \ & \boxed{58.67 \pm 8.36 \text{ or } [50.31;\ 67.03]} \end{aligned}$$

(d) It is wrong to say that the actual investment will be between \$50,310 and \$67,030 with probability 0.95. Rather, 95% refers to the long run proportion of prediction intervals constructed by the same procedure as we used in (c). That is, about 95% of prediction intervals computed from a large number of independent samples will contain the actual 2009 investment.

This procedure relied on a linear relation $\mathbf{E}(Y_i) = \beta_0 + \beta_1 X_i$, constant variance $\text{Var}Y_i = \sigma^2$, Normal distribution and independence of responses Y_i.

10.5 (a) In this multivariate regression analysis,

$$\boldsymbol{X} = \begin{pmatrix} 1 & -5 & 0 \\ 1 & -4 & 0 \\ 1 & -3 & 1 \\ 1 & -2 & 0 \\ 1 & -1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 3 & 0 \\ 1 & 4 & 1 \\ 1 & 5 & 1 \end{pmatrix} \quad \text{and} \quad \boldsymbol{y} = \begin{pmatrix} 17 \\ 23 \\ 31 \\ 29 \\ 33 \\ 39 \\ 39 \\ 40 \\ 41 \\ 44 \\ 47 \end{pmatrix}$$

Again, to simplify calculations, we define

$$X = \text{year} - 2000.$$

We then compute

$$\boldsymbol{X}^T\boldsymbol{X} = \begin{pmatrix} 11 & 0 & 7 \\ 0 & 110 & 8 \\ 7 & 8 & 7 \end{pmatrix},$$

$$(\boldsymbol{X}^T\boldsymbol{X})^{-1} = \begin{pmatrix} 0.2971 & 0.0236 & -0.3241 \\ 0.0236 & 0.0118 & -0.0370 \\ -0.3241 & -0.0370 & 0.5093 \end{pmatrix},$$

and

$$\boldsymbol{b} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y} = \begin{pmatrix} 32.2138 \\ 2.3569 \\ 4.0926 \end{pmatrix}$$

That is,

$$\boxed{b_0 = 32.2138, \quad b_1 = 2.3569, \quad b_2 = 4.0926}$$

(b) The estimated regression equation is

$$\hat{y} = 32.2138 + 2.3569x + 4.0926z$$

For a company reporting profit ($z_* = 1$) in 2007 (that is, $x_* = 7$), we predict the investment amount as

$$\hat{y}_* = 32.2138 + 2.3569(7) + 4.0926(1) = \boxed{52.8047 \text{ thousand dollars}}$$

(c) The slope β_2 is the change in the response variable when the dummy variable z changes from 0 to 1. Thus, if the company reports a loss during year 2007 instead of a gain, its expected investment amount reduces by β_2. Our prediction will decrease by $b_2 = 4.0926$ thousand dollars.

(d) The total sum of squares

$$SS_{\text{TOT}} = S_{yy} = \sum (Y_i - \bar{Y})^2 = 841.64$$

is computed in the previous exercise.

The error sum of squares can be computed, say, by filling the table,

Y_i	$\hat{Y}_i$	$Y_i - \hat{Y}_i$	$(Y_i - \hat{Y}_i)^2$
17	20.4293	−3.4293	11.7600
23	22.7862	0.2138	0.0457
31	29.2357	1.7643	3.1128
29	27.5000	1.5000	2.2500
33	33.9495	−0.9495	0.9015
39	36.3064	2.6936	7.2555
39	38.6633	0.3367	0.1134
40	41.0202	−1.0202	1.0408
41	39.2845	1.7155	2.9429
44	45.7340	−1.7340	3.0068
47	48.0909	−1.0909	1.1901

Alternatively, one can multiply matrices,

$$SS_{\text{ERR}} = (\boldsymbol{y} - \hat{\boldsymbol{y}})^T (\boldsymbol{y} - \hat{\boldsymbol{y}}) = (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b})^T (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b})$$

Then

$$SS_{\text{ERR}} = \sum_i (Y_i - \hat{Y}_i)^2 = 33.62$$

and

$$SS_{\text{REG}} = SS_{\text{TOT}} - SS_{\text{ERR}} = 841.64 - 33.62 = 808.02$$

Complete the ANOVA table:

Source	Sum of squares	Degrees of freedom	Mean squares	F
Model	808.02	2	404.01	96.14
Error	33.62	8	4.20	
Total	841.64	10		

Comparing $F_{obs} = 96.14$ against Table A5 with 2 and 8 d.f., we find that the model is significant at the 1% level (in fact, its P-value is less than 0.0001).

(e) From Exercise 10.4,

$$SS_{\text{ERR}}(\textit{Reduced}) = (MS_{\text{ERR}})(\text{df}_{\text{ERR}}) = (9)(7.48) = 67.32 \quad (9 \text{ d.f.})$$

This error sum of squares is obtained from the reduced model where investment is predicted based on the time trend only.

For the full model of Exercise 10.5(a-d),

$$SS_{\text{ERR}}(Full) = 33.62 \quad (8 \text{ d.f.})$$

Significance of the new dummy variable (reporting profit) in addition to the time trend is tested by the partial F-statistic

$$\frac{SS_{\text{EX}}/\text{df}_{\text{EX}}}{MS_{\text{ERR}}(Full)} = \frac{(SS_{\text{ERR}}(\textit{Reduced}) - SS_{\text{ERR}}(Full))/(9-8)}{MS_{\text{ERR}}(Full)}$$

$$= \frac{67.32 - 33.62}{4.20} = 8.024.$$

From Table A5 with 1 and 8 d.f., addition of the new variable is significant at the 5% but not at the 1% level. The P-value of this test is between 0.01 and 0.05.

10.6 Using the hint, we write

$$\begin{aligned} SS_{\text{TOT}} &= \sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}((y_i - \widehat{y}_i) + (\widehat{y}_i - \bar{y}))^2 \\ &= \sum_{i=1}^{n}(y_i - \widehat{y}_i)^2 + \sum_{i=1}^{n}(\widehat{y}_i - \bar{y})^2 + 2\sum_{i=1}^{n}(y_i - \widehat{y}_i)(\widehat{y}_i - \bar{y}) \\ &= SS_{\text{ERR}} + SS_{\text{REG}} + 2\sum_{i=1}^{n}(y_i - \widehat{y}_i)(\widehat{y}_i - \bar{y}). \end{aligned}$$

It remains to demonstrate that

$$\sum_{i=1}^{n}(y_i - \widehat{y}_i)(\widehat{y}_i - \bar{y}) = 0$$

This is true because

$$\begin{aligned}
&\sum_{i=1}^{n}(y_i - \widehat{y}_i)(\widehat{y}_i - \bar{y}) \\
&= \sum_{i=1}^{n}(y_i - b_0 - b_1 x_i)(b_0 + b_1 x_i - \bar{y}) \\
&= \sum_{i=1}^{n}\{y_i - (\bar{y} - b_1\bar{x}) - b_1 x_i\}\{(\bar{y} - b_1\bar{x}) + b_1 x_i - \bar{y}\} \\
&= \sum_{i=1}^{n}\{(y_i - \bar{y}) - b_1(x_i - \bar{x})\} b_1(x_i - \bar{x}) \\
&= b_1\left\{\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x}) - b_1 \sum_{i=1}^{n}(x_i - \bar{x})\right\} \\
&= b_1\left\{S_{xy} - \frac{S_{xy}}{S_{xx}} S_{xx}\right\} = b_1(S_{xy} - S_{xy}) = 0
\end{aligned}$$

10.7

$$\begin{aligned}
R^2 &= \frac{SS_{\text{REG}}}{SS_{\text{TOT}}} = \frac{\sum(b_0 + b_1\, x_i - \bar{y})^2}{S_{yy}} \\
&= \frac{\sum(\bar{y} - b_1\bar{x} + b_1\, x_i - \bar{y})^2}{S_{yy}} = \frac{b_1^2 \sum(x_i - \bar{x})^2}{S_{yy}} \\
&= \frac{(S_{xy}/S_{xx})^2 S_{xx}}{S_{yy}} = \frac{S_{xy}^2}{S_{xx}S_{yy}} = \left(\frac{s_{xy}}{s_x s_y}\right)^2 = r^2
\end{aligned}$$

10.8 In John's regression model, the number of hours of preparation is the predictor X, and the quiz grade is the response variable Y.

We are given: $n = 10$, $\bar{x} = 3.6$, $s_x = 0.5$, $\bar{y} = 82$, $s_y = 14$, and $r = 0.62$.

(a) Compute

$$\begin{aligned}
b_1 &= \frac{r s_y}{s_x} = \frac{(0.62)(14)}{0.5} = 17.36 \\
b_0 &= \bar{y} - b_1\bar{x} = 82 - (17.36)(3.6) = 19.50.
\end{aligned}$$

Thus, the quiz grade can be predicted by the following equation

$$\boxed{\hat{y} = 19.50 + 17.36x}$$

(b) For $x_* = 4$, John can predict his grade as

$$y_* = 19.50 + 17.36(4) = 88.94 \approx 89.$$

(c) In this model,

$$R^2 = r^2 = (0.62)^2 = \boxed{0.3844}$$

The model (the preparation time) explains 38.44% of the total variation of quiz grades. No, most of the total variation remains unexplained.

Then, is the preparation time significant for predicting the quiz grade? Compute the F-statistic

$$F = \frac{SS_{\mathrm{REG}}/\mathrm{df}_{\mathrm{REG}}}{SS_{\mathrm{ERR}}/\mathrm{df}_{\mathrm{ERR}}} = \frac{R^2 SS_{\mathrm{TOT}}/1}{(1-R^2)SS_{\mathrm{TOT}}/(n-2)}$$

$$= \frac{0.3844}{(1-0.3844)/8} = 4.9955$$

From Table A5 with 1 and 8 d.f., we find that the P-value exceeds 0.05, and therefore, the null hypothesis is not rejected at the 5% level of significance. We conclude that John's model does not fit very well. Either the quiz grade is not a linear function of the preparation time, or he should consider other independent variables.

10.9 From the data in Example 10.10 on p. 356, we can obtain:

$$n = 7,\ \bar{x_2} = 9.57,\ s^2_{x_2} = 63.29,\ \bar{y} = 35,\ s^2_y = 242,\ r_{x_2 y} = 0.758$$

Compute the least squares estimates

$$\begin{aligned} b_1 &= r\sqrt{\frac{s^2_y}{s^2_{x_2}}} = (0.758)\sqrt{\frac{242}{63.29}} = 1.48 \\ b_0 &= \bar{y} - b_1\bar{x_2} = 35 - (1.48)(9.57) = 20.84. \end{aligned}$$

The fitted regression line has the equation

$$\boxed{\hat{y} = 20.84 + 1.48x_2}$$

The coefficient of determination is

$$R^2 = r^2 = \boxed{0.575}$$

showing that 57.5% of the total variation of the number of processed requests is explained by the number of tables only.

Next, compute sums of squares,

$$\begin{aligned} SS_{\mathrm{TOT}} &= (n-1)s^2_y = (6)(242) = 1452 \quad (n-1 = 6 \text{ d.f.}) \\ SS_{\mathrm{REG}} &= R^2 SS_{\mathrm{TOT}} = (0.575)(1452) = 834.9 \quad (1 \text{ d.f.}) \\ SS_{\mathrm{ERR}} &= SS_{\mathrm{TOT}} - SS_{\mathrm{REG}} = 1452 - 834.9 = 617.1 \quad (n-2 = 5 \text{ d.f.}) \end{aligned}$$

Based on this, the adjusted R-square is

$$R^2_{\mathrm{adj}} = 1 - \frac{SS_{\mathrm{ERR}}/\mathrm{df}_{\mathrm{ERR}}}{SS_{\mathrm{TOT}}/\mathrm{df}_{\mathrm{TOT}}} = 1 - \frac{617.1/5}{1452/6} = \boxed{0.51}$$

which is lower than the adjusted R-square of the full model ($R^2_{\mathrm{adj}} = 0.68$)

in Example 10.10. According to the adjusted R-square criterion, the full model is better.

Complete the ANOVA table,

Source	Sum of squares	Degrees of freedom	Mean squares	F
Model	834.9	1	834.9	6.77
Error	617.1	5	123.4	
Total	1452	6		

This F-statistic (6.77) is just a little higher than $F_{0.05} = 6.61$ from Table A5 with 1 and 5 d.f. Therefore, the P-value is just below 0.05, and this reduced model predicting the number of processed requests by the number of tables is significant at the 5% level.

10.10 From the data in Exercise 8.5 on p. 250 (with $x = \text{year} - 1800$),

$$n = 22,\ \bar{x} = 95,\ s_x = 64.9,\ \bar{y} = 94.7,\ s_y = 87.3, \quad \text{and} \quad r_{xy} = 0.959$$

(a) Compute

$$\begin{aligned} b_1 &= \frac{rs_y}{s_x} = \frac{(0.959)(87.3)}{64.9} = 1.29 \\ b_0 &= \bar{y} - b_1\bar{x} = 94.7 - (1.29)(95) = -27.9. \end{aligned}$$

The fitted regression line has the equation

$$\boxed{\hat{y} = -27.9 + 1.29x}$$

On the average, the population has increased by 1.29 mln per year.

(b) The coefficient of determination is

$$R^2 = r^2 = (0.959)^2 = \boxed{0.920}$$

For the ANOVA table, compute

$$\begin{aligned} SS_{\text{TOT}} &= (n-1)s_y^2 = (21)(87.3)^2 = 1.60 \cdot 10^5 \\ SS_{\text{REG}} &= R^2 SS_{\text{TOT}} = (0.92)(1.6 \cdot 10^5) = 1.47 \cdot 10^5 \\ SS_{\text{ERR}} &= SS_{\text{TOT}} - SS_{\text{REG}} = 1.3 \cdot 10^4 \end{aligned}$$

Source	Sum of squares	Degrees of freedom	Mean squares	F
Model	$1.47 \cdot 10^5$	1	$1.47 \cdot 10^5$	226.15
Error	$1.3 \cdot 10^4$	20	650	
Total	$1.6 \cdot 10^5$	21		

Very large values of R^2 and F indicate a rather significant model. However, our ANOVA F-statistic should not be compared with Table A5 because the data set violates standard regression assumptions, see Example 10.5 on p. 342.

(c) For years 2010, 2015, and 2020, we have $x_* = 210$, 215, and 220. Then we can use our estimated regression line to predict the population,

$$\hat{y}_* = -27.9 + 1.29x_* = \boxed{243.0, 249.5, \text{ and } 255.9 \text{ million people}}$$

Although being significant, the linear regression model underestimates the population for the recent years and the future. The population of the United States does not grow linearly. Our model is improved in the next exercise.

10.11 (a) In this multivariate regression, define the predictor matrix $\boldsymbol{X}$ with three columns containing $x^{(0)} = 1$, $x^{(1)} =$ (year-1800), and $x^{(2)} =$ (year-1800)^2 and the response vector $\boldsymbol{y}$ as follows,

$$\boldsymbol{X} = \begin{pmatrix} 1 & -10 & 100 \\ 1 & 0 & 0 \\ 1 & 10 & 100 \\ 1 & 20 & 400 \\ 1 & 30 & 900 \\ 1 & 40 & 1600 \\ 1 & 50 & 2500 \\ 1 & 60 & 3600 \\ 1 & 70 & 4900 \\ 1 & 80 & 6400 \\ 1 & 90 & 8100 \\ 1 & 100 & 10000 \\ 1 & 110 & 12100 \\ 1 & 120 & 14400 \\ 1 & 130 & 16900 \\ 1 & 140 & 19600 \\ 1 & 150 & 22500 \\ 1 & 160 & 25600 \\ 1 & 170 & 28900 \\ 1 & 180 & 32400 \\ 1 & 190 & 36100 \\ 1 & 200 & 40000 \end{pmatrix} \quad \text{and} \quad \boldsymbol{y} = \begin{pmatrix} 3.9 \\ 5.3 \\ 7.2 \\ 9.6 \\ 12.9 \\ 17.1 \\ 23.2 \\ 31.4 \\ 38.6 \\ 50.2 \\ 63.0 \\ 76.2 \\ 92.2 \\ 106.0 \\ 123.2 \\ 132.2 \\ 151.3 \\ 179.3 \\ 203.3 \\ 226.5 \\ 248.7 \\ 281.4 \end{pmatrix}.$$

The vector of regression slopes is then estimated by

$$\boldsymbol{b} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y} = \begin{pmatrix} 5.6226 \\ 0.0199 \\ 0.0067 \end{pmatrix}.$$

Thus,

$$\boxed{b_0 = 5.6226, \ b_1 = 0.0199, \ \text{and} \ b_2 = 0.0067}$$

(b) The estimated regression equation is

$$\begin{aligned}\hat{y} &= 5.6226 + 0.0199x^{(1)} + 0.0067x^{(2)} \\ &= 5.6226 + 0.0199(\text{year-1800}) + 0.0067(\text{year-1800})^2\end{aligned}$$

Substituting years 2010, 2015, and 2020, we obtain the forecasts,

$$\boxed{\hat{y}_{2010} = 305.3,\ \hat{y}_{2015} = 319.6, \quad \text{and} \quad \hat{y}_{2020} = 333.2 \text{ million people}}$$

(c) We have $SS_{\text{TOT}} = \sum(y_i - \bar{y})^2 = 159,893$ (here we need more accuracy than in Exercise 10.10 because the error sum of squares is much smaller). Then we can compute

$$\hat{\boldsymbol{y}} = \boldsymbol{X}\boldsymbol{b} = \begin{pmatrix} 6.1 \\ 5.6 \\ 6.5 \\ 8.7 \\ 12.2 \\ 17.1 \\ 23.3 \\ 30.9 \\ 39.7 \\ 50.0 \\ 61.5 \\ 74.4 \\ 88.6 \\ 104.2 \\ 121.1 \\ 139.3 \\ 158.9 \\ 179.8 \\ 202.0 \\ 225.6 \\ 250.5 \\ 276.7 \end{pmatrix}, \quad \boldsymbol{y} - \hat{\boldsymbol{y}} = \begin{pmatrix} -2.2 \\ -0.3 \\ 0.7 \\ 0.9 \\ 0.7 \\ 0.0 \\ -0.1 \\ 0.5 \\ -1.1 \\ 0.2 \\ 1.5 \\ 1.8 \\ 3.6 \\ 1.8 \\ 2.1 \\ -7.1 \\ -7.6 \\ -0.5 \\ 1.3 \\ 0.9 \\ -1.8 \\ 4.7 \end{pmatrix},$$

$$SS_{\text{ERR}} = (\boldsymbol{y} - \hat{\boldsymbol{y}})^T(\boldsymbol{y} - \hat{\boldsymbol{y}}) = 170.0,$$

and

$$SS_{\text{REG}} = SS_{\text{TOT}} - SS_{\text{ERR}} = 159,893 - 170 = 159,723.$$

The quadratic model yields an amazingly high

$$R^2 = \frac{SS_{\text{REG}}}{SS_{\text{TOT}}} = \frac{159,723}{159,893} = \boxed{0.999}$$

Comparing with Exercise 10.10, we see that the quadratic term explains additional $99.9\% - 92.0\% = 7.9\%$ of the total variation. Explaining 99.9% of the total variation, the quadratic model is a very good fit for the U.S. population data.

The complete ANOVA table is

Source	Sum of squares	Degrees of freedom	Mean squares	F
Model	159,723	2	79,862	8928
Error	170	19	8.94	
Total	159,893	21		

Again, the correlation of y_i and y_{i-1} is rather high (estimated as 0.35), therefore, we cannot compare the F-statistic against Table A5 because of the violation of regression assumptions. However, its huge value of 8928 signals an extremely significant model.

(d) For the linear model in Exercise 10.10,

$$R^2_{\text{adj}} = 1 - \frac{MS_{\text{ERR}}}{SS_{\text{TOT}}/\text{df}_{\text{TOT}}} = 1 - \frac{650}{1.6 \cdot 10^5/21} = \boxed{0.915}$$

For the quadratic model in Exercise 10.11,

$$R^2_{\text{adj}} = 1 - \frac{MS_{\text{ERR}}}{SS_{\text{TOT}}/\text{df}_{\text{TOT}}} = 1 - \frac{8.94}{159,893/21} = \boxed{0.9988}$$

For the model with the quadratic term only,

$$y = \beta_0 + \beta_1(\text{year} - 1800)^2,$$

compute the correlation coefficient between the population y and the squared number of years since 1800, $x^{(2)} = (\text{year} - 1800)^2$,

$$r = 0.99946.$$

This yields the following adjusted R-square,

$$\begin{aligned} R^2_{\text{adj}} &= 1 - \frac{SS_{\text{ERR}}/\text{df}_{\text{ERR}}}{SS_{\text{TOT}}/\text{df}_{\text{TOT}}} = 1 - \frac{(1-R^2)SS_{\text{TOT}}/(n-2)}{SS_{\text{TOT}}/(n-1)} \\ &= 1 - \frac{(1-r^2)/20}{1/21} = \boxed{0.9989} \end{aligned}$$

According to the adjusted R-square criterion, the last model, with the quadratic term only, is the best one although it beats the full quadratic model by a very short margin. We may conclude that in the full quadratic model, the linear term does not explain a sufficient portion of the total variation to compensate for one additional parameter (resulting in an additional degree of freedom).

(e) It is apparent from Figure 10.6 on p. 349 that the quadratic model (the full model) fits much better than the linear model. Being very close to the actual points (x_i, y_i), the fitted parabolic curve results in very small residuals, and consequently, a very small error sum of squares. Apart from relatively small deviations, the quadratic model can be used to predict the population rather accurately.

10.12 We can use methods of either Section 10.1 or Section 10.3. In a matrix form,

$$\boldsymbol{X} = \begin{pmatrix} 1 & 0 \\ 1 & 10 \\ 1 & 20 \\ 1 & 30 \\ 1 & 40 \\ 1 & 50 \\ 1 & 60 \\ 1 & 70 \\ 1 & 80 \\ 1 & 90 \\ 1 & 100 \\ 1 & 110 \\ 1 & 120 \\ 1 & 130 \\ 1 & 140 \\ 1 & 150 \\ 1 & 160 \\ 1 & 170 \\ 1 & 180 \\ 1 & 190 \\ 1 & 200 \end{pmatrix}, \quad \boldsymbol{y} = \begin{pmatrix} 1.4 \\ 1.9 \\ 2.4 \\ 3.3 \\ 4.2 \\ 6.1 \\ 8.2 \\ 7.2 \\ 11.6 \\ 12.8 \\ 13.2 \\ 16.0 \\ 13.8 \\ 17.2 \\ 9.0 \\ 19.1 \\ 28.0 \\ 24.0 \\ 23.2 \\ 22.2 \\ 32.7 \end{pmatrix}$$

and the least squares estimates can be obtained as

$$\boldsymbol{b} = (\boldsymbol{X}^T \boldsymbol{X})^{-1} \boldsymbol{X}^T \boldsymbol{y} = \boxed{\begin{pmatrix} -0.495 \\ 0.137 \end{pmatrix}}$$

The estimated regression line for predicting the absolute annual population growth is

$$\hat{y} = -0.495 + 0.137x$$

The time plot of 10-year increments in the U.S. population and our estimated regression line is in Figure 11.10a. Our findings show that the population of the United States grows at a roughly *linearly increasing rate* (justifying again the quadratic model for the population in Exercise 10.11) (also see Exercise 8.6).

10.13 (a) In this problem, the predictor matrix is the same as in Exercise 10.12 above whereas the response vector consists of relative increments,

$$\begin{aligned} \boldsymbol{y} &= (0.36,\ 0.36,\ 0.33,\ 0.34,\ 0.33,\ 0.36,\ 0.35,\ 0.23,\ 0.30,\ 0.26,\ 0.21, \\ &\quad 0.21,\ 0.15,\ 0.16,\ 0.07,\ 0.14,\ 0.19,\ 0.13,\ 0.11,\ 0.10,\ 0.13) \end{aligned}$$

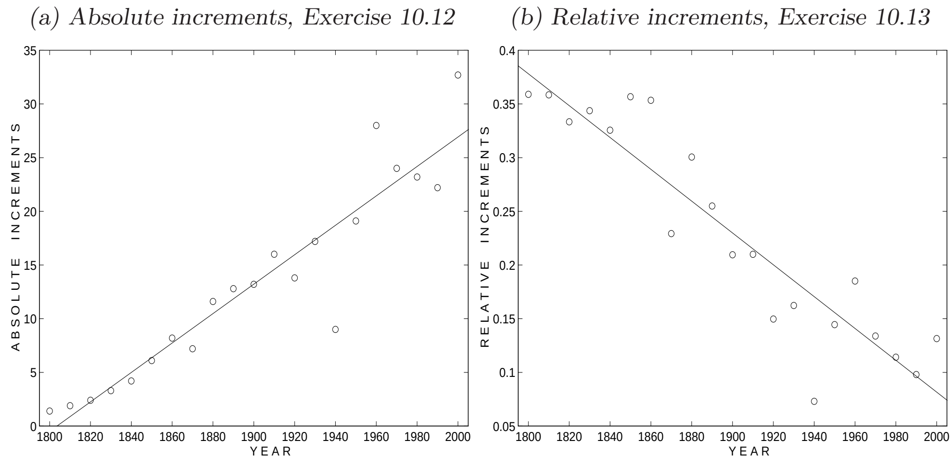


Figure 11.10 *Linear regression models for the absolute and relative growth of the U.S. population (Exercises 10.12 and 10.13)*

Compute the least squares estimates,

$$\boldsymbol{b} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y} = \boxed{\begin{pmatrix} 0.3781 \\ -0.0015 \end{pmatrix}}$$

The estimated regression line for predicting the absolute annual population growth is

$$\hat{y} = 0.3781 - 0.0015x$$

The time plot of 10-year relative increments of the U.S. population and the estimated regression line is in Figure 11.10b. Its *linearly decreasing trend* shows that the proportional 10-year rate of growth of the U.S. population decreases by about 0.15% every year from almost 38% in year 1800 (also see Exercise 8.7; if the proportional growth did not fall, the population would have been growing exponentially fast).

(b) For the ANOVA table, compute

$$\begin{aligned} \bar{y} &= 0.23 \\ SS_{\text{TOT}} &= \sum (y_i - \bar{y})^2 = 0.20 \\ SS_{\text{ERR}} &= \sum (y_i - (0.3781 - 0.0015x_i))^2 = 0.03 \\ SS_{\text{REG}} &= SS_{\text{TOT}} - SS_{\text{ERR}} = 0.17. \end{aligned}$$

Degrees of freedom are: $\text{df}_{\text{TOT}} = n - 1 = 20$, $\text{df}_{\text{ERR}} = n - 2 = 19$, and $\text{df}_{\text{REG}} = 1$. Notice that 22 measurements of the population yield 21 increments.

Complete the ANOVA table,

Source	Sum of squares	Degrees of freedom	Mean squares	F
Model	0.17	1	0.17	107.7
Error	0.03	19	0.0016	
Total	0.20	20		

R-square equals $SS_{\text{REG}}/SS_{\text{TOT}} = \boxed{0.85}$; 85% of the total variation of relative increments is explained by the linear time trend.

(c) The F-statistic of 107.7 is far beyond the critical value of 8.18 from Table A5 with 1 and 19 d.f. Therefore, our linear model is significant at the 0.01 level of significance (based on Table A5, we can only state that the P-value is less than 0.01; in fact, the P-value is even less than 0.0001).

(d) The slope is estimated by $b_1 = -0.0015$. Its standard deviation is estimated by

$$s_{b_1} = \sqrt{\frac{s^2}{S_{xx}}} = \sqrt{\frac{MS_{\text{ERR}}}{\sum(x_i - \bar{x})^2}} = \sqrt{\frac{0.0016}{77,000}} = 0.00014$$

A 95% confidence interval for the regression slope β_1 is

$$\begin{aligned} b_1 \pm t_{0.025} s_{b_1} &= -0.0015 \pm (2.093)(0.00014) \\ &= \boxed{-0.0015 \pm 0.0003 \text{ or } [-0.0018, -0.0012]} \end{aligned}$$

where $t_{0.025} = 2.093$ is obtained from Table A6 with $\text{df}_{\text{ERR}} = 19$ d.f.

(e) According to our results, we predict 10-year the relative growth of

$$y_* = b_0 + b_1 x_* = 0.3781 - (0.0015)(210) = 0.063$$

or 6.3% between years 2000 and 2010. A 95% prediction interval is

$$\begin{aligned} & y_* \pm t_{\alpha/2}\, s \sqrt{1 + \frac{1}{n} + \frac{(x_* - \bar{x})^2}{S_{xx}}} \\ &= 0.063 \pm (2.093)\sqrt{0.0016}\sqrt{1 + \frac{1}{21} + \frac{(210-100)^2}{77,000}} \\ &= \boxed{0.063 \pm 0.107 \text{ or } [-0.044, 0.170]} \end{aligned}$$

For the period between 2010 and 2020, assuming that the linear trend continues, a 95% prediction interval for the relative change of the pop-

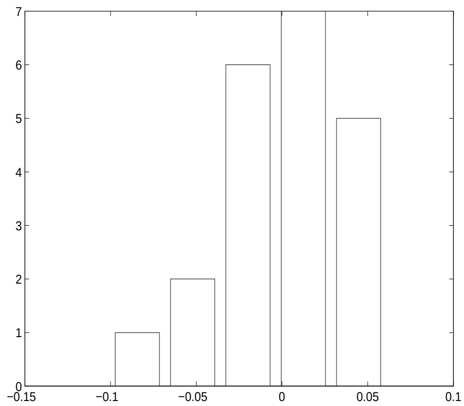


Figure 11.11 *Histogram of regression residuals (Exercises 10.13)*

ulation is

$$
\begin{aligned}
& b_0 + b_1 x_* \pm t_{\alpha/2}\, s\sqrt{1 + \frac{1}{n} + \frac{(x_* - \bar{x})^2}{S_{xx}}} \\
= \quad & 0.3781 - (0.0015)(220) \pm (2.093)\sqrt{0.0016}\sqrt{1 + \frac{1}{21} + \frac{(220 - 100)^2}{77,000}} \\
= \quad & \boxed{0.048 \pm 0.108 \text{ or } [\text{-}0.060,\ 0.156]}
\end{aligned}
$$

(f) Residuals are computed as $e_i = y_i - \widehat{y}_i$; their histogram is given in Figure 11.11. It shows a slightly left-skewed distribution of residuals which does not quite support the assumption of a normal distribution. On the other hand, the sample size is not large enough to show significant deviation from normal distribution (and in fact, rigorous tests for normality performed on these data do not reject the normal distribution of residuals).

10.14 (a) Let $z_i = 1$ for the operational system A and $z_i = 0$ for the operational system B. With the addition of this dummy variable, the predictor matrix and the response vector are

$$
\boldsymbol{X} = \begin{pmatrix} 1 & 6 & 4 & 1 \\ 1 & 7 & 20 & 1 \\ 1 & 7 & 20 & 1 \\ 1 & 8 & 10 & 1 \\ 1 & 10 & 10 & 0 \\ 1 & 10 & 2 & 0 \\ 1 & 15 & 1 & 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{y} = \begin{pmatrix} 40 \\ 55 \\ 50 \\ 41 \\ 17 \\ 26 \\ 16 \end{pmatrix}
$$

Compute the vector of regression slopes

$$\boldsymbol{b} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y} = \begin{pmatrix} 24.20 \\ -0.60 \\ 0.57 \\ 18.78 \end{pmatrix}$$

The estimated regression equation is

$$\boxed{\hat{y} = 24.20 - 0.60x^{(1)} + 0.57x^{(2)} + 18.78z}$$

(b) According to Example 10.10, the model without z (reduced model) has $SS_{\rm ERR} = 308.7$ with 4 d.f. and $SS_{\rm TOT} = 1452$ with 6 d.f.

The model considered here (full model) has the error sum of squares

$$SS_{\rm ERR}(Full) = \sum(y_i - (24.20 - 0.60x_i^{(1)} + 0.57x_i^{(2)} + 18.78z_i))^2 = 126.3$$

with only $n - 4 = 3$ d.f.

The new R-square is

$$R^2 = 1 - \frac{SS_{\rm ERR}}{SS_{\rm TOT}} = 1 - \frac{126.3}{1452} = \boxed{0.913}$$

That is, 91.3% of the total variation of the number of processed requests is explained by the data size, the number of tables, and the operational system. Comparing with $R^2(Reduced) = 0.787$, we conclude that the new variable z explains additional 1.26% of the total variation.

Keeping in mind that R-square is not a fair measure of comparison of two models with different numbers of predictors, let us compute the adjusted R-square,

$$R^2_{\rm adj}(Full) = 1 - \frac{126.3/3}{1452/6} = 0.826$$

Reduced model in Example 10.10 has $R^2_{\rm adj}(Red) = 0.681$. Hence, the new variable improves the goodness of fit, according the adjusted R-square criterion.

(c) Conduct the partial F-test:

$$\begin{aligned} F &= \frac{SS_{\rm ERR}(Red) - SS_{\rm ERR}(Full)}{SS_{\rm ERR}(Full)} \left(\frac{{\rm df}_{\rm ERR}(Full)}{{\rm df}_{\rm ERR}(Red) - {\rm df}_{\rm ERR}(Full)} \right) \\ &= \frac{308.7 - 126.3}{126.3} \left(\frac{3}{1} \right) = 4.33 \end{aligned}$$

From Table A5 with 1 and 3 d.f. we find that the P-value is greater than 0.05, therefore, the new variable is *not significant* at the 5% level (in fact, the exact P-value is 0.1289).

A T-test of $H_0 : \beta_3 = 0$ results in the same P-value and the same conclusion. Its test statistic can be found as $T = \sqrt{F} = 2.08$.

(d) (1) ADJUSTED R-SQUARE. Having three predictor variables, we can fit 8 linear regression models:

$$\begin{array}{ll}
1. & y = \beta_0 + \beta_1 x^{(1)} + \beta_2 x^{(2)} + \beta_3 z + \varepsilon \\
2. & y = \beta_0 + \beta_1 x^{(1)} + \beta_2 x^{(2)} + \varepsilon \\
3. & y = \beta_0 + \beta_1 x^{(1)} + \beta_3 z + \varepsilon \\
4. & y = \beta_0 + \beta_2 x^{(2)} + \beta_3 z + \varepsilon \\
5. & y = \beta_0 + \beta_1 x^{(1)} + \varepsilon \\
6. & y = \beta_0 + \beta_2 x^{(2)} + \varepsilon \\
7. & y = \beta_0 + \beta_3 z + \varepsilon \\
8. & y = \beta_0 + \varepsilon
\end{array}$$

From Examples 10.6, 10.10 and Exercises 10.9 and 10.14(b), we know that $SS_{\text{ERR}}(\text{Model 5}) = 491$, $SS_{\text{ERR}}(\text{Model 2}) = 309$, $SS_{\text{ERR}}(\text{Model 6}) = 617$, and $SS_{\text{ERR}}(\text{Model 1}) = 126$. For other models, compute

$$\boldsymbol{b} = (\boldsymbol{X}^T \boldsymbol{X})^{-1} \boldsymbol{X}^T \boldsymbol{y}, \quad SS_{\text{ERR}} = (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b})^T (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b})$$

choosing the corresponding predictor matrix $\boldsymbol{X}$ every time, and finally,

$$R^2_{\text{adj}} = 1 - \frac{SS_{\text{ERR}}/\text{df}_{\text{ERR}}}{SS_{\text{TOT}}/\text{df}_{\text{TOT}}} = 1 - \frac{SS_{\text{ERR}}/\text{df}_{\text{ERR}}}{1452/6}$$

Model	Predictors	SS_{ERR}	df_{ERR}	R^2_{adj}
1	$x^{(1)},\ x^{(2)},\ z$	126	3	0.83
2	$x^{(1)},\ x^{(2)}$	309	4	0.68
3	$x^{(1)},\ z$	202	4	0.79
4	$x^{(2)},\ z$	133	4	0.86
5	$x^{(1)}$	491	5	0.59
6	$x^{(2)}$	617	5	0.49
7	z	218	5	0.82
8	none	1452	6	0

According to the adjusted R-square criterion, choose Model 4

$$\boxed{y = \beta_0 + \beta_2 x^{(2)} + \beta_3 z + \varepsilon}$$

(2) STEPWISE SELECTION.

Step 1.

Starting with Model 8, which predictor should first enter the regression equation? Looking at the table above, Models 5–7, we see that the dummy variable z alone has the smallest error sum of squares. Therefore, it has the largest SS_{REG} and the largest F-to-enter statistic

$$F_1 = \frac{MS_{\text{REG}}(z)}{MS_{\text{ERR}}(z)} = \frac{SS_{\text{TOT}} - SS_{\text{ERR}}(z)}{SS_{\text{ERR}}(z)/\text{df}_{\text{ERR}}(z)} = \frac{1452 - 218}{218/5} = 28.3$$

It has 1 and 5 d.f. and yields a P-value below 0.01 (Table A5). Then, *variable z enters the equation* and we move from Model 8 to Model 7.

Step 2.

In addition to z, what is the next variable to enter the model? Comparing Models 3 and 4, we see that Model 4 has the smaller SS_{ERR}. Compute the F-to-enter statistic

$$F_2 = \frac{SS_{\text{ERR}}(\text{Model 7}) - SS_{\text{ERR}}(\text{Model 4})}{MS_{\text{ERR}}(\text{Model 4})} = \frac{218 - 133}{133/4} = 2.56$$

with 1 and 4 d.f. With the P-value above 0.05, we conclude that addition of the new variable $x^{(2)}$ did not explain a significant portion of variation, and we stay with Model 7.

The stepwise selection algorithm results in Model 7,

$$\boxed{y = \beta_0 + \beta_3 z + \varepsilon}$$

(3) Backward elimination.

Step 1.

Now we start with the full model (Model 1). Among Models 2–4, the smallest SS_{ERR} and the largest SS_{REG} are obtained from Model 2 predicting y from $x^{(2)}$ and z. Therefore, variable $x^{(1)}$ is the first one to be removed from the model with the F-to-remove statistic

$$F_{-1} = \frac{SS_{\text{ERR}}(\text{Model 4}) - SS_{\text{ERR}}(\text{Model 8})}{MS_{\text{ERR}}(\text{Model 8})} = \frac{133 - 126}{126/3} = 0.17.$$

With 1 and 3 d.f., variable $x^{(1)}$ is not significant in Model 8 (P-value greater than 0.05). We remove it from the equation and move on to Model 4.

Step 2.

Next, without $x^{(1)}$, we compare Models 6 and 7 and find that Model 7 has the smaller SS_{ERR}. Remove variable $x^{(2)}$ and compute the corresponding F-to-remove statistic

$$F_{-2} = \frac{SS_{\text{ERR}}(\text{Model 7}) - SS_{\text{ERR}}(\text{Model 4})}{MS_{\text{ERR}}(\text{Model 4})} = \frac{218 - 133}{133/4} = 2.56$$

(1 and 4 d.f.) This is not significant at the 5% level, and we move further to Model 7 with only one predictor z.

Step 3.

This last step compares Models 7 and 8. We already know (from the F-to-enter statistic F_1 above) that Model 7 is significant at the 1% level. Formally, the F-to-remove statistic is

$$F_{-3} = \frac{MS_{\text{REG}}(z)}{MS_{\text{ERR}}(z)} = F_1 = 28.3$$

Hence, we stay with Model 7

$$\boxed{y = \beta_0 + \beta_3 z + \varepsilon}$$