

Solutions to Exercises 1

1. Data types:

- (a) Categorical
- (b) Continuous
- (c) Ordinal
- (d) Discrete

2. (a) Systematic sampling.

- (b) Quasi-random.

3. Relative frequency table for data on number of calls:

Calls made	Tally	Frequency	Relative Frequency %
0		5	25
1		6	30
2		3	15
3		3	15
4		2	10
5		1	5
Totals		20	100

4. Relative frequency table for length of call data:

Length of call	Tally	Frequency	Relative Frequency %
$250 \leq \text{length} < 270$		1	4
$270 \leq \text{length} < 290$		6	24
$290 \leq \text{length} < 310$		9	36
$310 \leq \text{length} < 330$		7	28
$330 \leq \text{length} < 350$		2	8
Totals		25	100

Solutions to Exercises 2

1. The stem and leaf plot is

[illegible]

Stem	Leaf
1	1
1	2
1	3
1	4
1	5
1	6
1	7
1	8
1	9
2	0
2	1
2	2
2	3
2	4
2	5
2	6
2	7
2	8
2	9
3	0
3	1
3	2
3	3
3	4
3	5
3	6
3	7
3	8
3	9
4	0
4	1
4	2
4	3
4	4
4	5
4	6
4	7
4	8
4	9
5	0
5	1
5	2
5	3
5	4
5	5
5	6
5	7
5	8
5	9
6	0
6	1
6	2
6	3
6	4
6	5
6	6
6	7
6	8
6	9
7	0
7	1
7	2
7	3
7	4
7	5
7	6
7	7
7	8
7	9
8	0
8	1
8	2
8	3
8	4
8	5
8	6
8	7
8	8
8	9
9	0
9	1
9	2
9	3
9	4
9	5
9	6
9	7
9	8
9	9

$$n = 50, \quad \text{stem unit} = 1\text{kg}, \quad \text{leaf unit} = 0.1\text{kg}.$$

Comment: The distribution is roughly symmetric around a value of about 10kg. Most of the potatoes weigh between about 8 and 12kg. There are no obvious outliers.

2. A histogram is more appropriate, because the data are continuous.
3. The frequency polygons show a clear increase in sales due to the advertising campaign (the “After” polygon is shifted to the right), suggesting that the campaign has been a success.

Solutions to Exercises 3

1. (a) Let the weight of sack i be x_i . The sum of the weights is

$$\sum_{i=1}^{50} x_i = 505.8 \text{ kg.}$$

Therefore the sample mean is

$$\bar{x} = \frac{1}{50} \sum_{i=1}^{50} x_i = \frac{505.8}{50} = 10.116 \text{ kg.}$$

- (b) The observations into increasing order.

8.1	8.2	8.5	8.7	8.8	8.9	9.2	9.3	9.3	9.4
9.5	9.5	9.6	9.6	9.6	9.7	9.7	9.9	9.9	10.0
10.0	10.0	10.0	10.0	10.1	10.2	10.2	10.2	10.3	10.3
10.4	10.4	10.4	10.5	10.6	10.6	10.6	10.6	10.6	10.7
10.8	10.9	11.0	11.2	11.3	11.3	11.3	11.5	11.6	12.8

As there are $n = 50$ observations, the median M is the $(50 + 1)/2 = 25\frac{1}{2}$ th smallest observation, that is, half way between the 25th and 26th smallest observations, which are 10.1 and 10.2, respectively:

$$M = 10.1 + \frac{1}{2}(10.2 - 10.1) = \frac{10.1 + 10.2}{2} = 10.15 \text{ kg.}$$

- (c) The range of the data is the difference between the largest and smallest observations. Here the minimum and maximum values are $\min = 8.1$ kg and $\max = 12.8$ kg and so the range is

$$\text{Range} = \max - \min = 12.8 - 8.1 = 4.7 \text{ kg.}$$

- (d) The interquartile range is the difference between the upper and lower quartiles. As there are $n = 50$ observations, the lower quartile is the $(50 + 1)/4 = 12\frac{3}{4}$ th smallest observation, that is, three quarters of the way between the 12th and 13th smallest observations, which are 9.5 and 9.6, respectively:

$$Q_1 = 9.5 + \frac{3}{4} \times (9.6 - 9.5) = 9.575 \text{ kg.}$$

Similarly, the upper quartile is the $3(50 + 1)/4 = 38\frac{1}{4}$ th smallest observation, that is, a quarter of the way between the 38th and 39th smallest observations. As both of these observations are 10.6 kg, we have

$$Q_3 = 10.6 + \frac{1}{4}(10.6 - 10.6) = 10.6 \text{ kg.}$$

Therefore, the interquartile range is

$$\text{IQR} = Q_3 - Q_1 = 10.6 - 9.575 = 1.025 \text{ kg.}$$

- (e) The sample standard deviation is best calculated using your calculator in statistics mode. However we can also do the calculation the long way. We find $\sum_{i=1}^{50} x_i^2 = 5157.54$. Therefore the sample variance is

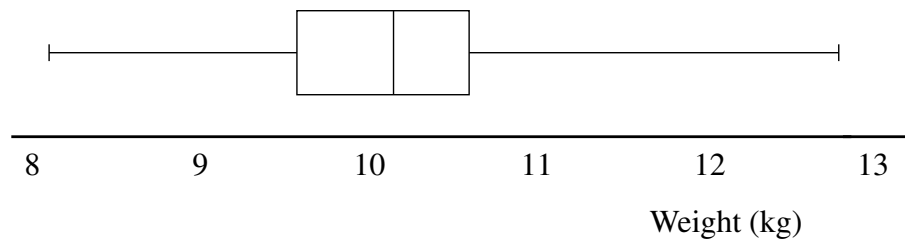
$$\begin{aligned} s^2 &= \frac{1}{n-1} \left\{ \sum_{i=1}^n x_i^2 - n\bar{x}^2 \right\} \\ &= \frac{1}{49} (5157.54 - 50 \times 10.116^2) \\ &= \frac{40.8672}{49} \\ &= 0.8340245 \text{ kg}^2. \end{aligned}$$

The sample standard deviation is

$$s = \sqrt{s^2} = \sqrt{0.8340245} \simeq 0.913 \text{ kg}.$$

Note that this is the value obtained on calculators using the s or $x\sigma_{n-1}$ button and not the σ or $x\sigma_n$ button.

- (f) The box plot is constructed from the minimum (8.1), lower quartile (9.575), median (10.15), upper quartile (10.6) and maximum (12.8), which were found earlier and looks like this:



Comment: The distribution of weights is roughly symmetric about a median of around 10 kg.

- (g) Grouping these data into a frequency table gives

Class	Class Interval	Frequency
1	$8.0 \leq x < 8.5$	2
2	$8.5 \leq x < 9.0$	4
3	$9.0 \leq x < 9.5$	4
4	$9.5 \leq x < 10.0$	9
5	$10.0 \leq x < 10.5$	14
6	$10.5 \leq x < 11.0$	9
7	$11.0 \leq x < 11.5$	5
8	$11.5 \leq x < 12.0$	2
9	$12.0 \leq x < 12.5$	0
10	$12.5 \leq x < 13.0$	1
Total		$n = 50$

(The class intervals and frequencies can be determined from the stem and leaf plot in Exercise 1 of Chapter 2.)

- (h) The modal class in the grouped frequency table is the class with the largest frequency, that is class 5, with $10.0 \leq x < 10.5$.
2. (a) Mark is correct. There is one very large cookie which weighs 51.3g; this might distort the sample mean and sample standard deviation (as their calculation involves all the data).
- (b) The median and inter-quartile range are more appropriate measures of location and spread for these data as they are more robust to outliers.

Ordered data: 22.4, 23.4, 24.9, 25.4, 25.6, 26.0, 26.3, 26.5, 27.1, 51.3

$n = 10$ so the median is the $(n + 1)/2 = 11/2 = 5.5$ th smallest observation, i.e. $25.6 + 0.5 \times (26.0 - 25.6) = 25.8g$.

Q_1 is the $(n + 1)/4 = 11/4 = 2.75$ th smallest observation, i.e. $23.4 + 0.75 \times (24.9 - 23.4) = 24.525g$.

Q_3 is the $3 \times (n + 1)/4 = 3 \times 11/4 = 8.25$ th smallest observation, i.e. $26.5 + 0.25 \times (27.1 - 26.5) = 26.65g$.

Therefore, the inter-quartile range is

$$IQR = Q_3 - Q_1 = 26.65 - 24.525 = 2.125g$$

3. Females have the more variable visit lengths as the standard deviation for females is greater than that for males.

Solutions to Exercises 4

- (a) E and F are **dependent**. Rumour has it that you have to be intelligent to get to University! :-) Someone with a high IQ is more likely to be accepted for a University place than someone with a low IQ.
 - (b) E_1 and E_2 are **dependent**. An individual's financial position is likely to affect a bank's decision to extend his overdraft.
- In this question we simply count up the number of students who fit the description and divide this number by the total number of students in the sample.

- (a) There are 4 female students out of 18. So the probability that the student is female is

$$\frac{4}{18} = 0.2222.$$

- (b) There are 6 students with weights greater than 70kg. So the probability that the student's weight is greater than 70kg. is

$$\frac{6}{18} = 0.3333.$$

- (c) There are 4 students with weights greater than 70kg and shoe-sizes greater than 8. So the probability of choosing such a student is

$$\frac{4}{18} = 0.2222.$$

- (d) There are 8 students with weights greater than 70kg or shoe-sizes greater than 8. So the probability of choosing such a student is

$$\frac{8}{18} = 0.4444.$$

- We are told $P(\text{Branded}) = 0.85$, $P(\text{Own-brand}) = 0.6$, $P(\text{Branded and Own-brand}) = 0.5$ and we require $P(\text{neither branded nor own-brand})$, where $P(\text{neither branded nor own-brand}) = 1 - P(\text{Branded or own-brand})$.

Using the addition rule,

$$\begin{aligned} P(\text{Branded or own-brand}) &= P(\text{Branded}) + P(\text{Own-brand}) - P(\text{Branded and own-brand}) \\ &= 0.85 + 0.6 - 0.5 \\ &= 0.95 \end{aligned}$$

and so

$$\begin{aligned} P(\text{neither branded nor own-brand}) &= 1 - P(\text{Branded or own-brand}) \\ &= 1 - 0.95 \\ &= 0.05 \end{aligned}$$

Therefore, 5% of customers buy neither branded nor own-brand breakfast cereal.

[You could also use a Venn diagram or a table to compute this probability.]

Solutions to Exercises 5

1. (a) $P(\text{Male and aged 30 to 45}) = \frac{20}{200} = 0.1.$

(b) $P(\text{Male}|\text{aged 30 to 45}) = \frac{20}{50} = 0.4.$

Alternatively,

$$P(\text{aged 30 to 45}) = \frac{20}{200} + \frac{30}{200} = \frac{50}{200} = 0.25,$$

therefore

$$P(\text{Male}|\text{aged 30 to 45}) = \frac{P(\text{Male and aged 30 to 45})}{P(\text{aged 30 to 45})} = \frac{0.1}{0.25} = 0.4.$$

(c) $P(45 \text{ or under}|\text{Female}) = \frac{40+30}{80} = \frac{70}{80} = 0.875.$

Alternatively,

$$\begin{aligned} P(45 \text{ or under and Female}) &= P(\text{Under 30 and Female}) + P(30 \text{ to 45 and Female}) \\ &= \frac{40}{200} + \frac{30}{200} = \frac{70}{200} = 0.35, \end{aligned}$$

and

$$P(\text{Female}) = \frac{40}{200} + \frac{30}{200} + \frac{10}{200} = \frac{80}{200} = 0.4,$$

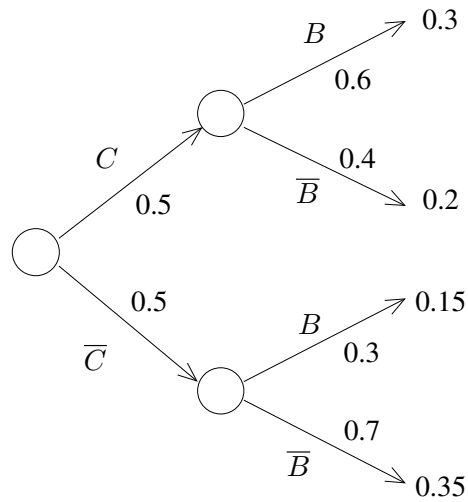
therefore

$$P(45 \text{ or under}|\text{Female}) = \frac{P(45 \text{ or under and Female})}{P(\text{Female})} = \frac{0.35}{0.4} = 0.875.$$

(d) We assume that the selection is done without replacement. We use the multiplication rule:

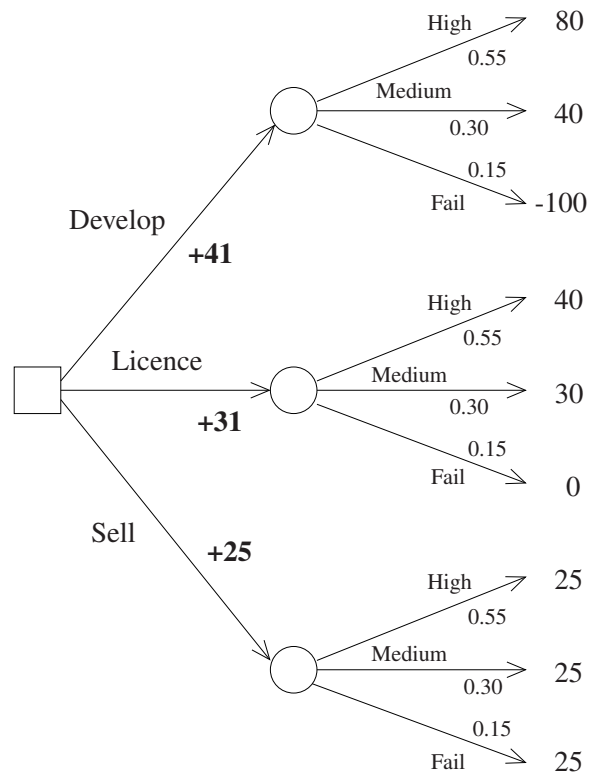
$$\begin{aligned} P(\text{both Male}) &= P(1\text{st Male and 2nd Male}) \\ &= P(1\text{st Male}) \times P(2\text{nd Male}|1\text{st Male}) \\ &= \frac{120}{200} \times \frac{119}{199} \\ &= \frac{357}{995} \simeq 0.359. \end{aligned}$$

2. (a) $P(C) = 0.5$.
 (b) $P(\overline{C}) = 1 - P(C) = 0.5$.
 (c) $P(\overline{B}|C) = 1 - P(B|C) = 1 - 0.6 = 0.4$.
 (d) $P(\overline{B}|\overline{C}) = 1 - P(B|\overline{C}) = 1 - 0.3 = 0.7$.
 (e) $P(C \text{ and } B) = P(C) \times P(B|C) = 0.5 \times 0.6 = 0.3$.
 (f) $P(B) = P(C \text{ and } B) + P(\overline{C} \text{ and } B) = 0.5 \times 0.6 + 0.5 \times 0.3 = 0.3 + 0.15 = 0.45$.
 [A tree diagram may be of some use here.]



Solutions to Exercises 6

1. (a) The decision tree is as follows.



(b) The *EMV* (in £ millions) of each possible decision can be calculated as follows:

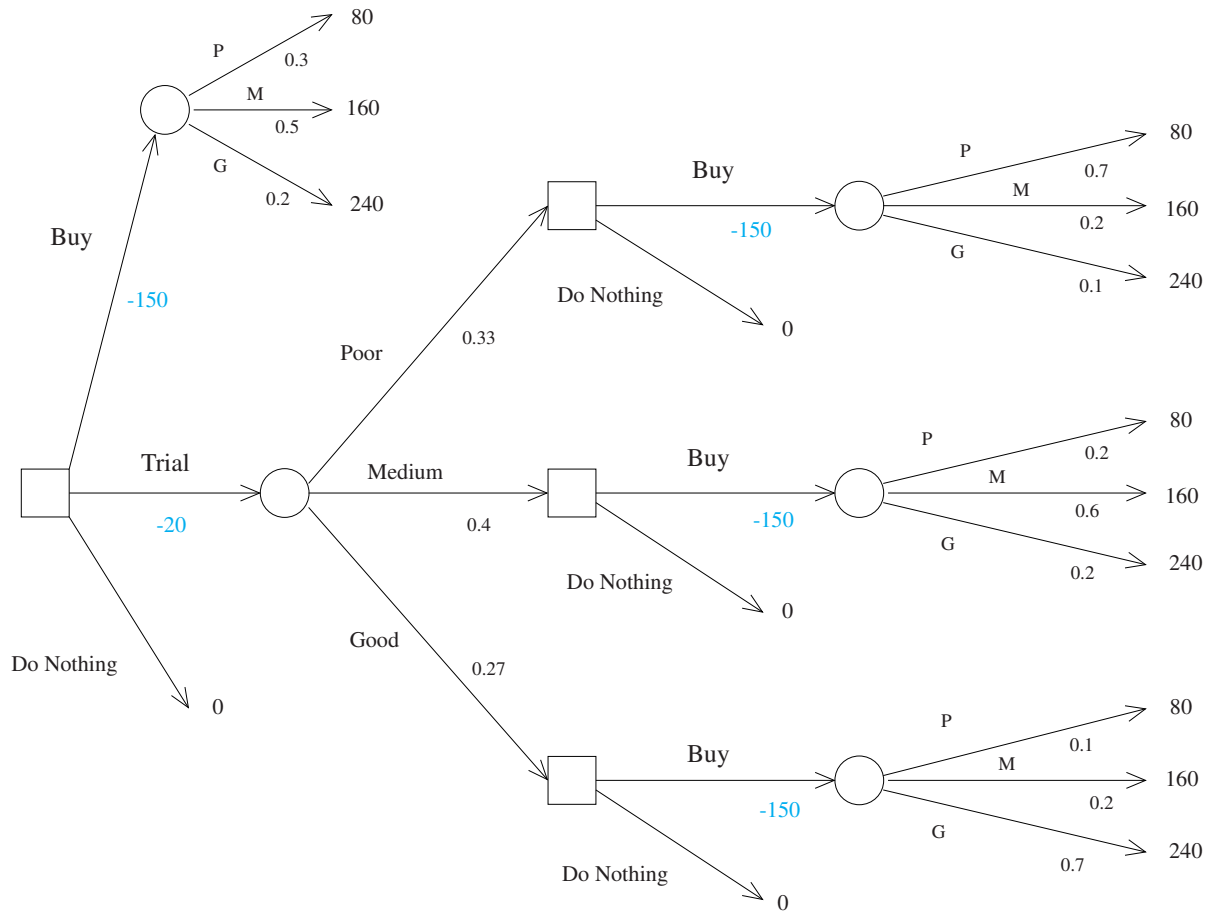
$$EMV(\text{Develop}) = 0.55 \times 80 + 0.30 \times 40 + 0.15 \times (-100) = 41$$

$$EMV(\text{Licence}) = 0.55 \times 40 + 0.30 \times 30 + 0.15 \times 0 = 31$$

$$EMV(\text{Sell}) = 0.55 \times 25 + 0.30 \times 25 + 0.15 \times 25 = 25.$$

On the basis of expected monetary value, the optimal decision for the company is to develop the technology themselves; this gives an *EMV* of £41 million.

2. (a)



(b)

$$\begin{aligned}
 EMV(\text{Buy with no trial}) &= (0.3 \times 80 + 0.5 \times 160 + 0.2 \times 240) - 150 \\
 &= 152 - 150 \\
 &= 2 \quad (\text{that is, } \pounds 2,000)
 \end{aligned}$$

$$\begin{aligned}
 EMV(\text{Buy} | \text{Trial is poor}) &= (0.7 \times 80 + 0.2 \times 160 + 0.1 \times 240) - 150 \\
 &= 112 - 150 \\
 &= -38 \quad (\text{that is, a loss of } \pounds 38,000)
 \end{aligned}$$

$$\begin{aligned}
 EMV(\text{Buy} | \text{Trial is medium}) &= (0.2 \times 80 + 0.6 \times 160 + 0.2 \times 240) - 150 \\
 &= 160 - 150 \\
 &= 10 \quad (\text{that is, } \pounds 10,000)
 \end{aligned}$$

$$\begin{aligned}
 EMV(\text{Buy}|\text{Trial is good}) &= (0.1 \times 80 + 0.2 \times 160 + 0.7 \times 240) - 150 \\
 &= 208 - 150 \\
 &= 58 \quad (\text{that is, } \pounds 10,000)
 \end{aligned}$$

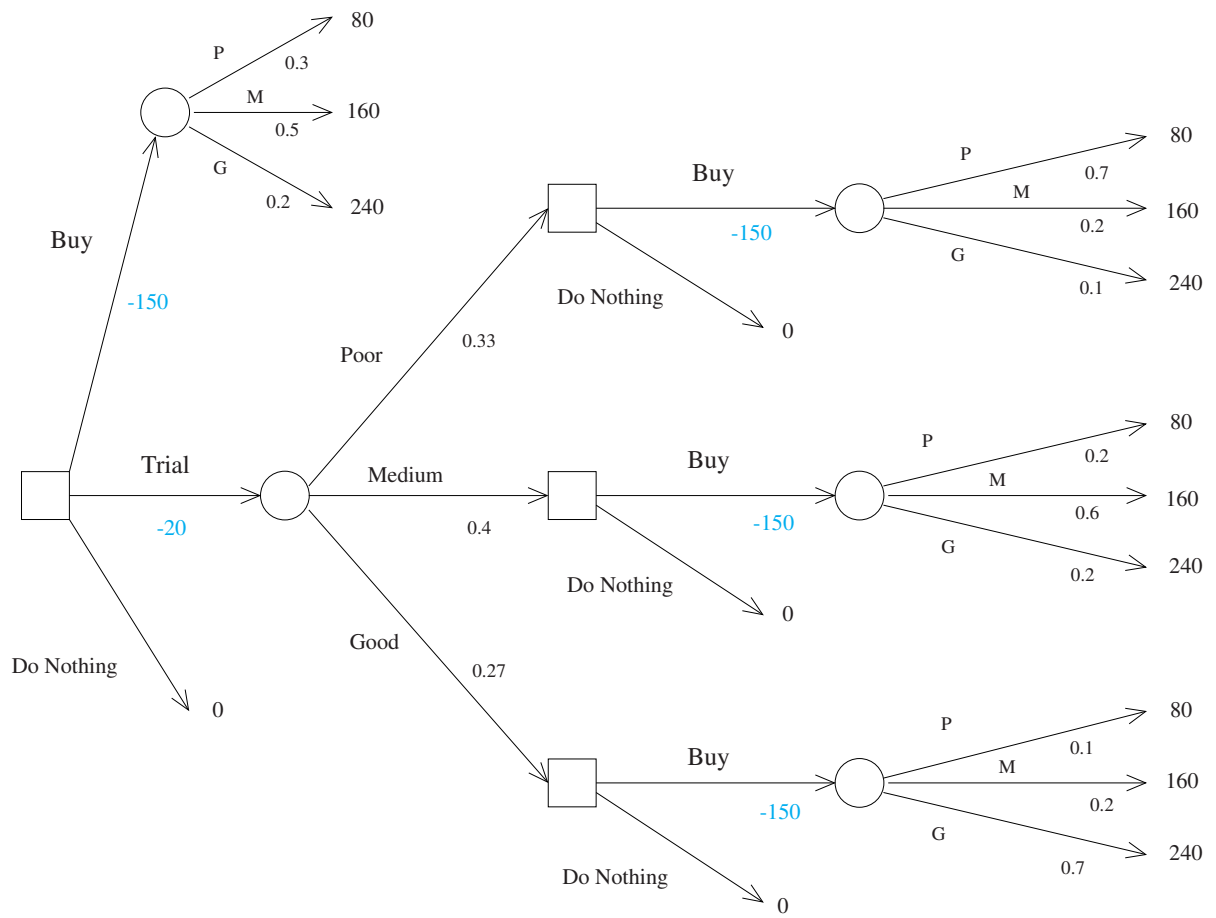
$EMV(\text{Trial is poor}) = \max(-38, 0) = 0$ (optimal choice is to do nothing if trial is poor).

$EMV(\text{Trial is medium}) = \max(10, 0) = 10$ (optimal choice is Buy if trial is medium).

$EMV(\text{Trial is good}) = \max(58, 0) = 58$ (optimal choice is Buy if trial is good).

$$\begin{aligned}
 EMV(\text{Trial}) &= (0.33 \times 0 + 0.4 \times 10 + 0.27 \times 58) - 20 \\
 &= 19.66 - 20 \\
 &= -0.34 \quad (\text{that is, a loss of } \pounds 340)
 \end{aligned}$$

and $EMV(\text{Do nothing}) = 0$. Therefore the optimal decision is to buy without a trial, as this maximises the EMV.



Solutions to Exercises 7

1. $P(X = 0) = 0.2$ as the probabilities must sum up to 1. That is, $P(X = 0) = 1 - (0.1 + 0.2 + 0.3 + 0.2) = 1 - 0.8 = 0.2$.

2. (a) Use the formula

$$P(X = r) = {}^nC_r \times p^r \times (1 - p)^{n-r}$$

with $n = 4$, $r = 0, 1, 2, 3, 4$, $p = \frac{1}{6}$ and $1 - p = \frac{5}{6}$ to get:

$$P(X = 0) = {}^4C_0 \times \left(\frac{1}{6}\right)^0 \times \left(\frac{5}{6}\right)^4 = \left(\frac{5}{6}\right)^4 = 0.4823$$

$$P(X = 1) = {}^4C_1 \times \frac{1}{6} \times \left(\frac{5}{6}\right)^3 = 0.3858$$

$$P(X = 2) = {}^4C_2 \times \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^2 = 0.1157$$

$$P(X = 3) = {}^4C_3 \times \left(\frac{1}{6}\right)^3 \times \frac{5}{6} = 0.0154$$

$$P(X = 4) = {}^4C_4 \times \left(\frac{1}{6}\right)^4 \times \left(\frac{5}{6}\right)^0 = \left(\frac{1}{6}\right)^4 = 0.0008,$$

and so the full probability distribution for X is:

r	0	1	2	3	4
$P(X = r)$	0.4823	0.3858	0.1157	0.0154	0.0008

- (b) $P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = 0.4823 + 0.3858 + 0.1157 = 0.9838$.

- (c) $P(X > 2) = 1 - P(X \leq 2) = 1 - 0.9838 = 0.0162$.

- (d) Mean: $E[X] = n \times p = 4 \times \frac{1}{6} = \frac{2}{3} \simeq 0.6667$.

Variance: $\text{Var}(X) = n \times p \times (1 - p) = 4 \times \frac{1}{6} \times \frac{5}{6} = \frac{5}{9} \simeq 0.5556$.

- (e) Rolling no sixes is most likely, $P(X = 0) = 0.4823$.

Solutions to Exercises 8

1. (a) Binomial, $n = 10$, $p = 0.3$; so $X \sim \text{Bin}(10, 0.3)$.
(b) Poisson: $\lambda = 7$; so $Y \sim \text{Po}(7)$.
(c) Binomial: $n = 12$, $p = 0.1$; so $Z \sim \text{Bin}(12, 0.1)$.
2. (a) Assuming that calls are answered independently, $X \sim \text{Bin}(8, 0.4)$.
(b) The mean and variance are

$$\begin{aligned}E[X] &= np = 8 \times 0.4 = 3.2 \\ \text{Var}(X) &= np(1 - p) = 8 \times 0.4 \times 0.6 = 1.92\end{aligned}$$

and so

$$SD(X) = \sqrt{\text{Var}(X)} = \sqrt{1.92} \simeq 1.386.$$

(c)

$$\begin{aligned}P(X = 7) &= {}^nC_r p^r (1 - p)^{n-r} \\ &= {}^8C_7 \times 0.4^7 \times (1 - 0.4)^1 \\ &\simeq 0.00786.\end{aligned}$$

(d)

$$\begin{aligned}P(X < 2) &= P(X = 0) + P(X = 1) \\ &= {}^8C_0 \times 0.4^0 \times (1 - 0.4)^{8-0} + {}^8C_1 \times 0.4^1 \times (1 - 0.4)^{8-1} \\ &= 0.01679616 + 0.08957952 \\ &= 0.10637568.\end{aligned}$$

3. (a) $Y \sim \text{Po}(4)$
(b) The mean and variance are

$$\begin{aligned}E[Y] &= \lambda = 4 \\ \text{Var}(Y) &= \lambda = 4\end{aligned}$$

and so

$$SD(Y) = \sqrt{\text{Var}(Y)} = \sqrt{4} = 2.$$

(c)

$$\begin{aligned}P(Y = 6) &= \frac{\lambda^r e^{-\lambda}}{r!} \\ &= \frac{4^6 \times e^{-4}}{6!} \\ &= 0.1041956346.\end{aligned}$$

(d)

$$\begin{aligned}P(Y \leq 2) &= P(Y = 0) + P(Y = 1) + P(Y = 2) \\&= \frac{4^0 \times e^{-4}}{0!} + \frac{4^1 \times e^{-4}}{1!} + \frac{4^2 \times e^{-4}}{2!} \\&= e^{-4} + 4e^{-4} + 8e^{-4} \\&= 0.01831563889 + 0.07326255555 + 0.1465251111 \\&\simeq 0.2381.\end{aligned}$$

(e)

$$\begin{aligned}P(Y > 2) &= 1 - P(Y \leq 2) \\&= 1 - 0.2381 \\&= 0.7619.\end{aligned}$$

Solutions to Exercises 9

1. Let X denote the time taken to deliver the goods. Then X has a normal distribution with mean $\mu = 16$ days and standard deviation $\sigma = 2.5$ days, that is $X \sim N(16, 2.5^2)$.

(a) The probability of a delivery being late is

$$\begin{aligned} P(X > 20) &= 1 - P(X \leq 20) \\ &= 1 - P\left(Z \leq \frac{20 - 16}{2.5}\right) \\ &= 1 - P(Z \leq 1.6) \\ &= 1 - 0.9452 \quad (\text{from tables}) \\ &= 0.0548, \end{aligned}$$

therefore, 5.48% of customers will receive their order late.

(b) The probability that customers receive their orders between 10 and 15 days is

$$\begin{aligned} P(10 < X < 15) &= P(X < 15) - P(X \leq 10) \\ &= P\left(Z < \frac{15 - 16}{2.5}\right) - P\left(Z \leq \frac{10 - 16}{2.5}\right) \\ &= P(Z < -0.4) - P(Z \leq -2.4) \\ &= 0.3446 - 0.0082 \quad (\text{from tables}) \\ &= 0.3364. \end{aligned}$$

Therefore, around 33.6% of customers receive their goods between 10 and 15 days after placing their orders.

(c) In the new processing system, Y has a normal distribution with mean $\mu = 16$ days and standard deviation $\sigma = 1.5$ days, so $Y \sim N(16, 1.5^2)$. The probability of a delivery on time is

$$\begin{aligned} P(Y < 20) &= P\left(Z < \frac{20 - 16}{1.5}\right) \\ &= P(Z < 2.67) \\ &= 0.9962 \quad (\text{from tables}). \end{aligned}$$

So around 99.6% of customers would receive their orders on time. This is an increase on the previous system of over 5%.

2. The amount of liquid dispensed, X , has a normal distribution with mean $\mu = 200\text{ml}$ and standard deviation $\sigma = 15\text{ml}$, that is $X \sim N(200, 15^2)$.

(a) The probability that a cup contains less than 170ml is

$$\begin{aligned}P(X < 170) &= P\left(Z < \frac{170 - 200}{15}\right) \\&= P(Z < -2.00) \\&= 0.0228 \quad (\text{from tables}),\end{aligned}$$

therefore, 2.28% of cups will contain less than the minimum permissible volume.

(b) The probability that a cup contains over 225ml is

$$\begin{aligned}P(X > 225) &= 1 - P(X \leq 225) \\&= 1 - P\left(Z \leq \frac{225 - 200}{15}\right) \\&= 1 - P(Z \leq 1.67) \\&= 1 - 0.9525 \quad (\text{from tables}) \\&= 0.0475,\end{aligned}$$

therefore, 4.75% of cups contain over 225ml.

(c) The probability that the cup contains between 175ml and 225ml is

$$\begin{aligned}P(175 < X < 225) &= P(X < 225) - P(X \leq 175) \\&= P\left(Z < \frac{225 - 200}{15}\right) - P\left(Z \leq \frac{175 - 200}{15}\right) \\&= P(Z < 1.67) - P(Z \leq -1.67) \\&= 0.9525 - 0.0475 \quad (\text{from tables}) \\&= 0.905.\end{aligned}$$

(d) If a 240ml cup is used then it will overflow with probability

$$\begin{aligned}P(X > 240) &= 1 - P(X \leq 240) \\&= 1 - P\left(Z \leq \frac{240 - 200}{15}\right) \\&= 1 - P(Z \leq 2.67) \\&= 1 - 0.9962 \quad (\text{from tables}) \\&= 0.0038.\end{aligned}$$

Hence, with 10000 drinks we would expect $10000 \times 0.0038 = 38$ to overflow.

Solutions to Exercises 10

1. The amount of time (in minutes) that the coach is delayed, X , has a uniform distribution on $a = -15$ to $b = 45$, that is $X \sim U(-15, 45)$.

(a) The mean of this distribution is

$$E[X] = \frac{a+b}{2} = \frac{(-15)+45}{2} = 15 \text{ minutes,}$$

so that, on average, the coach is 15 minutes late.

Probabilities for this distribution are calculated using

$$\begin{aligned} P(X \leq x) = P(X < x) &= \begin{cases} 0 & \text{for } x < a \\ \frac{x-a}{b-a} & \text{for } a \leq x \leq b \\ 1 & \text{for } x > b \end{cases} \\ &= \begin{cases} 0 & \text{for } x < -15 \\ \frac{x+15}{60} & \text{for } -15 \leq x \leq 45 \\ 1 & \text{for } x > 45. \end{cases} \end{aligned}$$

(b) The probability that the coach is less than 5 minutes late is

$$P(X < 5) = \frac{5+15}{60} = \frac{20}{60} = \frac{1}{3} = 0.3333.$$

(c) The probability that the coach is more than 20 minutes late is

$$P(X > 20) = 1 - P(X \leq 20) = 1 - \frac{20+15}{60} = 1 - \frac{35}{60} = \frac{5}{12} = 0.4167.$$

(d) The probability that the coach arrives between 10.55 and 11.20pm is

$$\begin{aligned} P(-5 < X < 20) &= P(X < 20) - P(X \leq -5) \\ &= \frac{20+15}{60} - \frac{-5+15}{60} \\ &= \frac{7}{12} - \frac{1}{6} \\ &= \frac{5}{12} \\ &= 0.4167. \end{aligned}$$

(e) The probability that the coach arrives before 11pm is

$$P(X < 0) = \frac{0+15}{60} = \frac{15}{60} = \frac{1}{4} = 0.25.$$

2. The time between successive requests, X , has an exponential distribution with parameter $\lambda = 2.5$ minutes, that is $X \sim \text{Exp}(2.5)$.

- (a) The expected time between arrivals of requests is

$$E[X] = \frac{1}{\lambda} = \frac{1}{2.5} = 0.4 \text{ minutes (or 24 seconds)}.$$

Probabilities for this distribution are calculated using

$$\begin{aligned} P(X \leq x) = P(X < x) &= \begin{cases} 1 - e^{-\lambda x} & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases} \\ &= \begin{cases} 1 - e^{-2.5 \times x} & \text{for } x \geq 0 \\ 0 & \text{for } x < 0. \end{cases} \end{aligned}$$

- (b) The probability that the time between requests is less than 1 minute and 30 seconds is

$$P(X < 1.5) = 1 - e^{-2.5 \times 1.5} = 1 - e^{-3.75} = 1 - 0.0235 = 0.9765.$$

- (c) The probability that the time between requests is greater than 1 minute is

$$P(X > 1) = 1 - P(X < 1) = 1 - (1 - e^{-2.5 \times 1}) = e^{-2.5} = 0.0821.$$

- (d) The probability that the time between requests is between 1 minute and 1 minute and 30 seconds is

$$\begin{aligned} P(1 < X < 1.5) &= P(X < 1.5) - P(X \leq 1) \\ &= P(X < 1.5) - P(X \leq 1) \\ &= 1 - e^{-2.5 \times 1.5} - (1 - e^{-2.5 \times 1}) \\ &= e^{-2.5} - e^{-3.75} \\ &= 0.0821 - 0.0235 \\ &= 0.0586. \end{aligned}$$

- (e) The probability that the time between requests is between 30 seconds and 50 seconds is

$$\begin{aligned} P\left(\frac{30}{60} < X < \frac{50}{60}\right) &= P\left(X < \frac{50}{60}\right) - P\left(X \leq \frac{30}{60}\right) \\ &= P\left(X < \frac{5}{6}\right) - P\left(X \leq \frac{1}{2}\right) \\ &= 1 - e^{-2.5 \times 5/6} - (1 - e^{-2.5 \times 1/2}) \\ &= e^{-1.25} - e^{-12.5/6} \\ &= 0.2865 - 0.1245 \\ &= 0.1620. \end{aligned}$$