

Answers

1 Sampling

1.1 Terms and notation

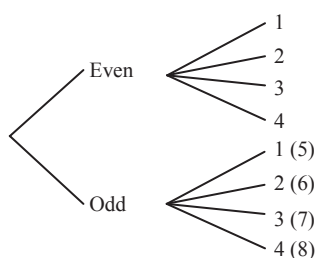
- 1 (i) A *census* is when we collect data from the entire population.
- (ii) A *sample* is when we collect data from a sub-group of the population.
- (iii) To ensure valid results our population should be surveyed *randomly*.
- (iv) *Bias* is when some members of the population are more likely to be selected than others.
- (v) The *sampling frame* is a list of all of the members of the population.
- (vi) The measures of centre and spread we calculate from a sample are called sample *statistics*.
- (vii) The mean and standard deviation of the whole are called population *parameters*.
- (viii) An *unbiased* estimate is one where, although individual estimates will vary, the process gives an accurate result 'on average'.

2

	Population	From a sample
Mean	μ	$\bar{x}$
Standard deviation	σ	s
Variance	σ^2	s^2

1.2 Sampling and sampling techniques

- 1 (i) Because each number 1–8 is not equally likely.
- (ii) Throw one die.
If even, pick from 1–4. If odd, pick from 5–8.
Throw again and choose.



Repeat process for 2nd choice.

Ignore repeats on 2nd choice.

- (iii) Each person is equally likely to be chosen.

- 2 (i) (a) All shoppers at the store.
- (b) Only on one day, bias in choosing people.
- (c) Give the questionnaire to the 1st, 5th, 10th person to enter the shop every hour for a whole week.
- (ii) (a) All people in the country.
- (b) Many! Only TV viewers, only people available to watch at 7 p.m., only a certain demographic watching, only those with strong opinions will call, etc.
- (c) Many ways, including cluster sampling cities first, then suburbs, then randomly calling people (followed up by post if needed).
- (iii) (a) All cars in the city.
- (b) Only one part of the city, on only one morning, is being considered.
- (c) Go to licence register and choose every 10th car on the list until they have enough cars.
- (iv) (a) All students at the school.
- (b) Only morning, only students who catch the bus, not a random selection.
- (c) Randomly choose 2 or 3 students in each class of the school to give the survey to.
- (v) (a) Every voter in the town.
- (b) Only people available then can be involved, only those home at the time (excludes people at work).
- (c) Choose every 100th name on the list, post survey and follow up with calls.

Further practice

- 1 (i) Anyone in the class with surnames starting with less frequent letters (Q, Z, etc.) will have less chance of being chosen.
- (ii) Number all students 1–30.
Use the calculator to produce random 2-digit numbers. Ignoring numbers outside the range and any repeats, choose 4 numbers.

Past exam questions

- 1 (i) Only students who use canteen can be in the survey.
The five may be friends.
- (ii) She looks at the first two digits and continues on through the numbers, ignoring numbers > 82 or any repeats.

2 (i) $0.392 \times 150 = 58.8$ so choose person 59

(ii) $\frac{104}{150} = 0.693$ and $\frac{103}{150} = 0.687$

So any number between 0.687 and 0.693 would result in a choice of 104.

(iii) It may be that you end up with duplicates.

3 573, 43 (or 043), 289

Stretch and challenge

1 Students' own answers.

2 Continuous random variables

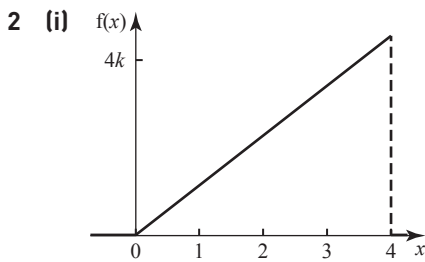
2.1 Probability density function

1 (i) No, area under the line is $\frac{1}{2} \times 2 \times 2 = 2$

(ii) Yes

(iii) No, $f(x) < 0$ for $x > 1$

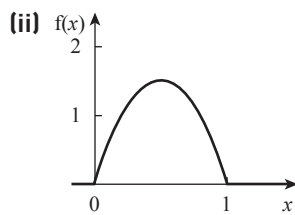
(iv) No, area under lines > 1



$$A = \frac{1}{2} \times 4 \times 4k = 1$$

$$8k = 1$$

$$k = \frac{1}{8}$$



$$\int_0^1 kx(x-1) dx = 1$$

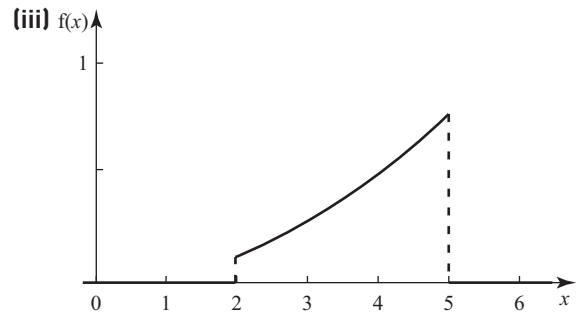
$$k \int_0^1 (x^2 - x) dx = 1$$

$$k \left[\frac{x^3}{3} - \frac{x^2}{2} \right]_0^1 = 1$$

$$k \left[\frac{1}{3} - \frac{1}{2} \right] = 1$$

$$-\frac{1}{6}k = 1$$

$$k = -6$$



$$A = \int_2^5 kx^2 dx = 1$$

$$k \left[\frac{x^3}{3} \right]_2^5 = 1$$

$$k \left[\frac{5^3}{3} - \frac{2^3}{3} \right] = 1$$

$$39k = 1$$

$$k = \frac{1}{39}$$

3 (i) $A = \int_0^2 kx(2-x) dx = 1$

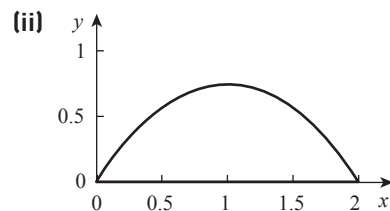
$$k \int_0^2 (2x - x^2) dx = 1$$

$$k \left[x^2 - \frac{x^3}{3} \right]_0^2 = 1$$

$$k \left[2^2 - \frac{2^3}{3} \right] = 1$$

$$\frac{4}{3}k = 1$$

$$k = \frac{3}{4}$$



(iii) $P(1 < x \leq 1.5)$

$$= \int_1^{1.5} \frac{3}{4}x(2-x) dx$$

$$= \frac{3}{4} \left[x^2 - \frac{x^3}{3} \right]_1^{1.5}$$

$$= \frac{3}{4} \left[\left(1.5^2 - \frac{1.5^3}{3} \right) - \left(1^2 - \frac{1^3}{3} \right) \right]$$

$$= \frac{11}{32} \approx 0.344$$

$$4 \quad (i) \quad \int_1^4 \frac{k}{x^2} dx = 1$$

$$k \int_1^4 x^{-2} dx = 1$$

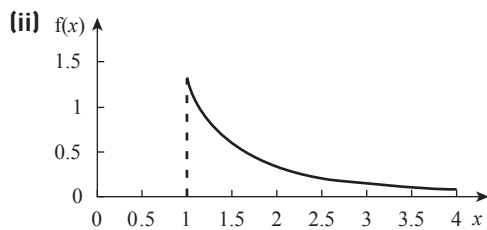
$$k \left[\frac{x^{-1}}{-1} \right]_1^4 = 1$$

$$k \left[-\frac{1}{x} \right]_1^4 = 1$$

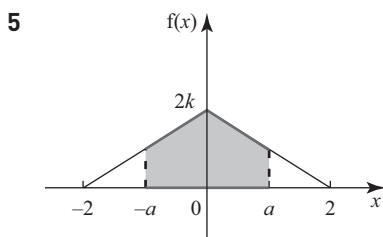
$$k \left[-\frac{1}{4} - \left(-\frac{1}{1} \right) \right] = 1$$

$$k \left(\frac{3}{4} \right) = 1$$

$$k = \frac{4}{3}$$



$$\begin{aligned} (iii) \quad P(X \geq 2) &= \frac{4}{3} \int_2^4 \frac{1}{x^2} \\ &= \frac{4}{3} \left[-\frac{1}{x} \right]_2^4 \\ &= \frac{4}{3} \left[-\frac{1}{4} - \left(-\frac{1}{2} \right) \right] \\ &= \frac{1}{3} \end{aligned}$$



$$(i) \quad A = \frac{1}{2} \times 4 \times 2k = 1 \Rightarrow k = \frac{1}{4}$$

$$(ii) \quad P(|X| \leq a)$$

$$= P(-a \leq X \leq a)$$

$$= 2 \times P(0 \leq X \leq a)$$

$$\text{When } x = a, f(x) = \frac{1}{4}(2 - a) = \frac{1}{2} - \frac{1}{4}a$$

$$\text{So } P(|X| \leq a) = 2 \times \left(\frac{\frac{1}{2} - \frac{1}{4}a + \frac{1}{2}}{2} \right) \times a$$

$$= \left(1 - \frac{1}{4}a \right) \times a$$

$$= a - \frac{1}{4}a^2$$

$$6 \quad (i) \quad A = 4k + \frac{1}{2} \times 2 \times k = 1$$

$$5k = 1$$

$$k = \frac{1}{5} = 0.2$$

$$m(\text{line}) = \frac{-\frac{1}{5}}{2} = -\frac{1}{10} = -0.1 \Rightarrow a = -0.1$$

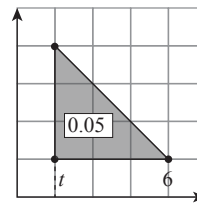
$$y = -0.1x + b$$

$$0 = -0.1 \times 6 + b \Rightarrow b = 0.6$$

$$(ii) \quad P(0 \leq x \leq 4) = 0.8$$

$$P(4 \leq x \leq t) = 0.15 \Rightarrow P(t \leq x \leq 6) = 0.05$$

$$y = -0.1t + 0.6$$



$$A = \frac{1}{2}(6 - t)(-0.1t + 0.6) = 0.05$$

$$\frac{1}{2}(-0.6t + 3.6 + 0.1t^2 - 0.6t) = 0.05$$

$$0.1t^2 - 1.2t + 3.6 = 0.1$$

$$0.1t^2 - 1.2t + 3.5 = 0$$

$$t = 5 \text{ or } 7$$

$$\text{Since } t < 6, t = 5$$

2.2 Mean, median and variance of continuous random variables

$$1 \quad (i) \quad \int_0^k \frac{x^3}{4} dx = 1$$

$$\left[\frac{x^4}{16} \right]_0^k = 1$$

$$\frac{k^4}{16} = 1$$

$$k^4 = 16$$

$$k = 2$$

$$(ii) \quad \text{Median, } m$$

$$\int_0^m \frac{x^3}{4} dx = 0.5$$

$$\left[\frac{x^4}{16} \right]_0^m = 0.5$$

$$\frac{m^4}{16} = 0.5$$

$$m^4 = 8$$

$$m = 1.68 \text{ (3 s.f.)}$$

$$2 \text{ (i)} \int_0^1 k(x - x^5) dx = 1$$

$$k \left[\frac{x^2}{2} - \frac{x^6}{6} \right]_0^1 = 1$$

$$k \left(\frac{1}{2} - \frac{1}{6} \right) = 1$$

$$\frac{1}{3}k = 1$$

$$k = 3$$

$$\text{(ii)} \int_0^m 3(x - x^5) dx = 0.5$$

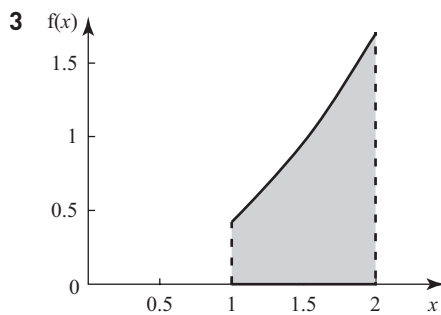
$$3 \left[\frac{x^2}{2} - \frac{x^6}{6} \right]_0^m = 0.5$$

$$3 \left[\frac{m^2}{2} - \frac{m^6}{6} \right] = \frac{1}{2}$$

$$\frac{m^2}{2} - \frac{m^6}{6} = \frac{1}{6}$$

$$3m^2 - m^6 = 1$$

$$m^6 - 3m^2 + 1 = 0$$



$$\text{(i)} \int_1^2 kx^2 dx = 1$$

$$k \left[\frac{x^3}{3} \right]_1^2 = 1$$

$$k \left[\frac{2^3}{3} - \frac{1^3}{3} \right] = 1$$

$$k \left[\frac{7}{3} \right] = 1$$

$$k = \frac{3}{7}$$

$$\text{(ii)} E(X) = \int_1^2 xf(x) dx$$

$$= \int_1^2 x \times \frac{3}{7}x^2 dx$$

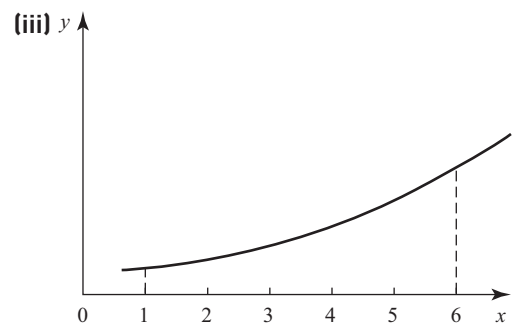
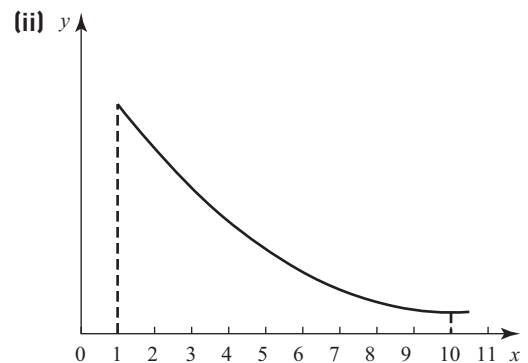
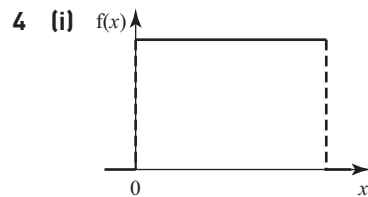
$$= \frac{3}{7} \left[\frac{x^4}{4} \right]_1^2$$

$$= \frac{3}{7} \left[\frac{2^4}{4} - \frac{1^4}{4} \right]$$

$$= \frac{3}{7} \left[\frac{15}{4} \right]$$

$$= \frac{45}{28} = 1.61 \text{ (3 s.f.)}$$

$$\begin{aligned} \text{(iii)} \text{Var}(X) &= \int_1^2 x^2 f(x) dx - \mu^2 \\ &= \int_1^2 x^2 \times \frac{3}{7}x^2 dx - \left(\frac{45}{28} \right)^2 \\ &= \frac{3}{7} \left[\frac{x^5}{5} \right]_1^2 - \left(\frac{45}{28} \right)^2 \\ &= \frac{3}{7} \left[\frac{2^5}{5} - \frac{1^5}{5} \right] - \left(\frac{45}{28} \right)^2 \\ &= \frac{3}{7} \left[\frac{31}{5} \right] - \left(\frac{45}{28} \right)^2 \\ &= 0.0742 \text{ (3 s.f.)} \end{aligned}$$



$$5 \text{ (i)} \int_0^5 kx(5 - x) dx = 1$$

$$\int_0^5 k(5x - x^2) dx = 1$$

$$k \left[\frac{5x^2}{2} - \frac{x^3}{3} \right]_0^5 = 1$$

$$k \left(\frac{5 \times 5^2}{2} - \frac{5^3}{3} \right) = 1$$

$$\frac{125}{6}k = 1$$

$$k = \frac{6}{125} = 0.048$$

$$\begin{aligned}
 \text{(ii)} \quad P(0 \leq X \leq 1) &= \int_0^1 \frac{6}{125} x(5-x) dx \\
 &= \frac{6}{125} \left[\frac{5x^2}{2} - \frac{x^3}{3} \right]_0^1 \\
 &= \frac{6}{125} \left(\frac{5}{2} - \frac{1}{3} \right) \\
 &= \frac{13}{125} = 0.104
 \end{aligned}$$

$$P(\text{two failures}) = 0.104^2 = 0.0108 \text{ (3 s.f.)}$$

$$\begin{aligned}
 \text{(iii)} \quad P(0 \leq X \leq 3) &= \int_0^3 \frac{6}{125} x(5-x) dx \\
 &= \frac{6}{125} \left[\frac{5x^2}{2} - \frac{x^3}{3} \right]_0^3 \\
 &= \frac{6}{125} \left(\frac{5 \times 3^2}{2} - \frac{3^3}{3} \right) \\
 &= \frac{81}{125} = 0.648
 \end{aligned}$$

$$\begin{aligned}
 P(\text{exactly one failure}) \\
 &= {}^2C_1 \times 0.648^1 \times 0.352^1 \\
 &= 0.456 \text{ (3 s.f.)}
 \end{aligned}$$

$$\begin{aligned}
 6 \text{ (i)} \quad \int_{25}^{\infty} Ax^{-\frac{5}{2}} dx &= 1 \\
 A \left[\frac{x^{-\frac{3}{2}}}{-\frac{3}{2}} \right]_{25}^{\infty} &= 1 \\
 A \left[-\frac{2}{3\sqrt{x^3}} \right]_{25}^{\infty} &= 1 \\
 A \left[0 + \frac{2}{375} \right] &= 1 \\
 A &= \frac{375}{2} = 187.5
 \end{aligned}$$

$$\begin{aligned}
 \text{(ii)} \quad \int_{25}^m \frac{375}{2} x^{-\frac{5}{2}} dx &= 0.5 \\
 \frac{375}{2} \left[-\frac{2}{3\sqrt{x^3}} \right]_{25}^m &= 0.5 \\
 \left[-\frac{2}{3\sqrt{m^3}} - \left(-\frac{2}{3\sqrt{25^3}} \right) \right] &= \frac{1}{375} \\
 -\frac{2}{3\sqrt{m^3}} + \frac{2}{375} &= \frac{1}{375} \\
 -\frac{2}{3\sqrt{m^3}} &= -\frac{1}{375} \\
 3\sqrt{m^3} &= 750 \\
 \sqrt{m^3} &= 250
 \end{aligned}$$

$$m = (250)^{\frac{2}{3}}$$

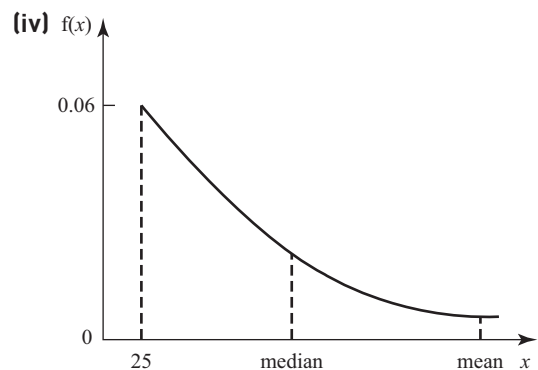
$$m = 39.7 \text{ (3 s.f.)}$$

Median salary is \$39 700 (3 s.f.)

(iii) Mean

$$\begin{aligned}
 &= \int_{25}^{\infty} x \times \left(\frac{375}{2} x^{-\frac{5}{2}} \right) dx \\
 &= \int_{25}^{\infty} \left(\frac{375}{2} x^{-\frac{3}{2}} \right) dx \\
 &= \frac{375}{2} \left[\frac{x^{-\frac{1}{2}}}{-\frac{1}{2}} \right]_{25}^{\infty} \\
 &= \frac{375}{2} \left[-\frac{2}{\sqrt{x}} \right]_{25}^{\infty} \\
 &= \frac{375}{2} \left[0 + \frac{2}{\sqrt{25}} \right] \\
 &= 75
 \end{aligned}$$

Mean salary is \$75 000



Median is the 'middle' value so not influenced by high salaries.

High salaries push the mean up significantly.

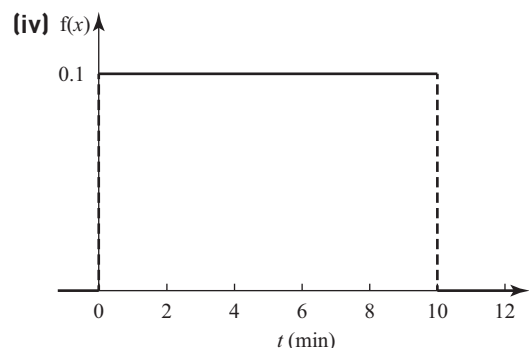
The company would use the mean to attract potential employees.

2.3 The uniform (rectangular) distribution

1 (i) X can take on any value in the interval $0 \leq X \leq 10$

(ii) $0 \leq X \leq 10$

(iii) $P(0 \leq X \leq 5) = 0.5$



$$\text{(v)} \quad f(x) = \begin{cases} 0.1 & 0 \leq x \leq 10 \\ 0 & \text{otherwise} \end{cases}$$

2 (i) $4k = 1 \Rightarrow k = \frac{1}{4} = 0.25$

(ii) Median = 3

(iii) Mean = 3

$$\begin{aligned} E(X) &= \int_1^5 x \times 0.25 \, dx \\ &= 0.25 \left[\frac{x^2}{2} \right]_1^5 \\ &= 0.25 \left[\frac{5^2}{2} - \frac{1^2}{2} \right] \\ &= 0.25[12] \\ &= 3 \end{aligned}$$

(iv) $\text{Var}(X) = \int x^2 f(x) - \mu^2$

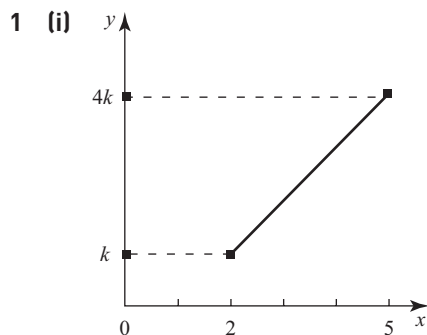
$$\begin{aligned} &= \int_1^5 (x^2 \times 0.25) \, dx - 3^2 \\ &= 0.25 \left[\frac{x^3}{3} \right]_1^5 - 3^2 \\ &= 0.25 \left[\frac{5^3}{3} - \frac{1^3}{3} \right] - 3^2 \\ &= 0.25 \left[\frac{124}{3} \right] - 3^2 \\ &= \frac{4}{3} \end{aligned}$$

3 (i) Area = $(p^2 - p) \times k = 1$

$$k = \frac{1}{p^2 - p}$$

(ii) Median = $\frac{p^2 - p}{2}$

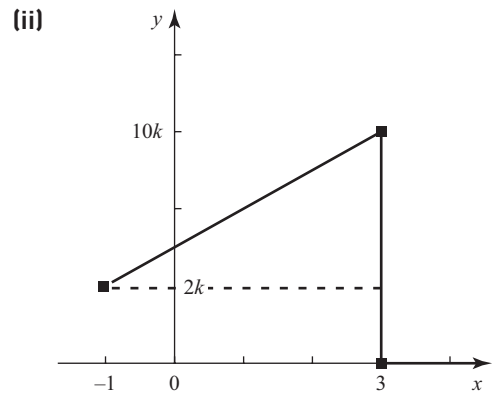
Further practice



$$\frac{1}{2} \times 3 \times 3k + 3k = 1$$

$$\frac{15}{2}k = 1$$

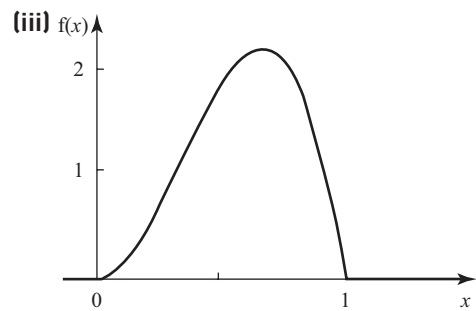
$$k = \frac{2}{15}$$



$$\frac{1}{2} \times 4 \times 8k + 4 \times 2k = 1$$

$$24k = 1$$

$$k = \frac{1}{24}$$



$$k \left[\frac{x^3}{3} - \frac{x^4}{4} \right]_0^1 = 1$$

$$k \left[\frac{1^3}{3} - \frac{1^4}{4} \right] = 1$$

$$\frac{1}{12}k = 1$$

$$k = 12$$

2 (i) $\int_0^1 k(1-x) \, dx = 1$

$$k \left[x - \frac{x^2}{2} \right]_0^1 = 1$$

$$k \left[1 - \frac{1}{2} \right] = 1$$

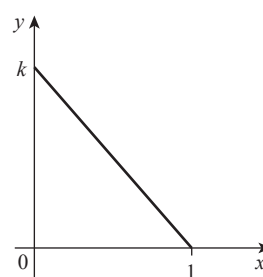
$$\frac{1}{2}k = 1$$

$$k = 2$$

or

$$A = \frac{1}{2} \times 1 \times k = 1$$

$$k = 2$$



$$\begin{aligned}
 \text{(ii)} \quad E(X) &= \int_0^1 x \times 2(1-x) dx \\
 &= 2 \left[\frac{x^2}{2} - \frac{x^3}{3} \right]_0^1 \\
 &= 2 \left[\frac{1}{2} - \frac{1}{3} \right] \\
 &= \frac{1}{3}
 \end{aligned}$$

$$\begin{aligned}
 \text{Var}(X) &= \int_0^1 x^2 \times 2(1-x) dx - \left(\frac{1}{3}\right)^2 \\
 &= 2 \left[\frac{x^3}{3} - \frac{x^4}{4} \right]_0^1 - \left(\frac{1}{3}\right)^2 \\
 &= 2 \left[\frac{1}{3} - \frac{1}{4} \right] - \frac{1}{9} \\
 &= 2 \times \frac{1}{12} - \frac{1}{9} \\
 &= \frac{1}{18}
 \end{aligned}$$

$$\begin{aligned}
 \text{(iii)} \quad \int_0^{1-\frac{1}{\sqrt{2}}} 2(1-x) dx \\
 &= 2 \left[x - \frac{x^2}{2} \right]_0^{1-\frac{1}{\sqrt{2}}} \\
 &= 2 \left[1 - \frac{1}{\sqrt{2}} - \frac{\left(1 - \frac{1}{\sqrt{2}}\right)^2}{2} \right] \\
 &= 2 \left[1 - \frac{1}{\sqrt{2}} - \left(\frac{1 - \frac{2}{\sqrt{2}} + \frac{1}{2}}{2} \right) \right] \\
 &= 2 \left[1 - \frac{1}{\sqrt{2}} - \frac{1}{2} + \frac{1}{\sqrt{2}} - \frac{1}{4} \right] \\
 &= 2 \left[\frac{1}{4} \right] \\
 &= \frac{1}{2}
 \end{aligned}$$

$$\begin{aligned}
 3 \quad \text{(i)} \quad \int_{25}^a \frac{1}{4\sqrt{t}} dt &= 0.8 \\
 \frac{1}{4} \int_{25}^a t^{-\frac{1}{2}} dt &= 0.8 \\
 \frac{1}{4} [2\sqrt{t}]_{25}^a &= 0.8 \\
 \frac{1}{4} [2\sqrt{a} - 2\sqrt{25}] &= 0.8 \\
 2\sqrt{a} - 10 &= 3.2 \\
 2\sqrt{a} &= 13.2 \\
 \sqrt{a} &= 6.6 \\
 a &= 43.56 \text{ min}
 \end{aligned}$$

$$\begin{aligned}
 \text{(ii)} \quad E(T) &= \int_{25}^{49} t \times \frac{1}{4\sqrt{t}} dt \\
 &= \frac{1}{4} \int_{25}^{49} t^{\frac{1}{2}} dt \\
 &= \frac{1}{4} \left[\frac{2\sqrt{t^3}}{3} \right]_{25}^{49} \\
 &= \frac{1}{4} \left[\frac{2\sqrt{49^3}}{3} - \frac{2\sqrt{25^3}}{3} \right] \\
 &= \frac{109}{3} = 36.3 \text{ min (or 36 min 20 s)}
 \end{aligned}$$

$$\begin{aligned}
 \text{Var}(T) &= \int_{25}^{49} t^2 \times \frac{1}{4\sqrt{t}} dt - \left(\frac{109}{3}\right)^2 \\
 &= \frac{1}{4} \int_{25}^{49} t^{\frac{3}{2}} dt - \left(\frac{109}{3}\right)^2 \\
 &= \frac{1}{4} \left[\frac{2\sqrt{t^5}}{5} \right]_{25}^{49} - \left(\frac{109}{3}\right)^2 \\
 &= \frac{1}{4} \left[\frac{2\sqrt{49^5}}{5} - \frac{2\sqrt{25^5}}{5} \right] - \left(\frac{109}{3}\right)^2 \\
 &= 48.1 \text{ min}^2 \text{ (3 s.f.)}
 \end{aligned}$$

$$\begin{aligned}
 4 \quad \text{(i)} \quad \int_2^k \frac{2}{5} x dx &= 1 \\
 \left[\frac{x^2}{5} \right]_2^k &= 1 \\
 \frac{k^2}{5} - \frac{2^2}{5} &= 1 \\
 \frac{k^2}{5} &= \frac{9}{5} \\
 k^2 &= 9 \\
 k &= 3 \quad (\text{ignore } k = -3)
 \end{aligned}$$

$$\begin{aligned}
 \text{(ii)} \quad E(X) &= \int_2^3 x \left(\frac{2}{5} x \right) dx \\
 &= \frac{2}{5} \left[\frac{x^3}{3} \right]_2^3 \\
 &= \frac{2}{5} \left[\frac{3^3}{3} - \frac{2^3}{3} \right] \\
 &= \frac{38}{15} = 2.53 \text{ (3 s.f.)}
 \end{aligned}$$

$$\begin{aligned}
 \text{(iii)} \quad P(X < E(X)) &= P(X < 2.53) \\
 &= \int_2^{2.53} \frac{2}{5} x dx \\
 &= \frac{2}{5} \left[\frac{x^2}{2} \right]_2^{2.53} \\
 &= \frac{2}{5} \left[\frac{(2.53)^2}{2} - \frac{2^2}{2} \right] \\
 &= 0.484 \text{ (3 s.f.)}
 \end{aligned}$$

(iv) Since $P(X < E(X)) < 0.5$, the median is greater than the mean.

$$5 \text{ (i)} \quad \int_2^{\infty} \frac{k}{x^4} dx = 1$$

$$k \int_2^{\infty} x^{-4} dx = 1$$

$$k \left[\frac{x^{-3}}{-3} \right]_2^{\infty} = 1$$

$$k \left[-\frac{1}{3x^3} \right]_2^{\infty} = 1$$

$$k \left[0 + \frac{1}{24} \right] = 1$$

$$k = 24$$

$$\begin{aligned} \text{(ii)} \quad P(2 \leq X \leq 3) &= \int_2^3 \frac{24}{x^4} dx \\ &= 24 \left[-\frac{1}{3x^3} \right]_2^3 \\ &= 24 \left[-\frac{1}{3 \times 3^3} - \left(-\frac{1}{3 \times 2^3} \right) \right] \\ &= 24 \left[\frac{19}{648} \right] \\ &= \frac{19}{27} = 0.704 \text{ (3 s.f.)} \end{aligned}$$

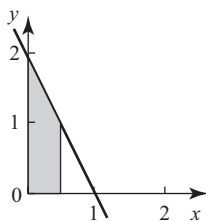
$$\begin{aligned} \text{(iii)} \quad E(X) &= \int_2^{\infty} x \times \frac{24}{x^4} dx \\ &= 24 \int_2^{\infty} x^{-3} dx \\ &= 24 \left[\frac{x^{-2}}{-2} \right]_2^{\infty} \\ &= 24 \left[-\frac{1}{2x^2} \right]_2^{\infty} \\ &= 24 \left[0 + \frac{1}{8} \right] \\ &= 3 \end{aligned}$$

Past exam questions

1 (i) (a) X

(b) V – high and low values are likely

$$\begin{aligned} \text{(ii)} \quad A &= \frac{2+1}{2} \times 0.5 \\ &= 0.75 \end{aligned}$$



or

$$\begin{aligned} A &= \int_0^{0.5} (2 - 2x) dx \\ &= \left[2x - x^2 \right]_0^{0.5} \\ &= 2 \times 0.5 - 0.5^2 \\ &= 0.75 \end{aligned}$$

$$\text{(iii)(a)} \quad \int_0^1 ax^n dx = 1$$

$$a \left[\frac{x^{n+1}}{n+1} \right]_0^1 = 1$$

$$a \left[\frac{1}{n+1} \right] = 1$$

$$a = n + 1$$

$$\text{(b)} \quad E(X) = \int_0^1 x \times ax^n dx$$

$$= \int_0^1 ax^{n+1} dx$$

$$= a \left[\frac{x^{n+2}}{n+2} \right]_0^1$$

$$= \frac{a}{n+2}$$

$$\frac{a}{n+2} = \frac{5}{6}$$

$$6a = 5(n+2)$$

$$6a = 5n + 10 \text{ and } a = n + 1$$

$$\Rightarrow 6(n+1) = 5n + 10$$

$$6n + 6 = 5n + 10$$

$$n = 4$$

$$a = 5$$

$$2 \text{ (i)} \quad \int_0^{\frac{2}{3}\pi} k \sin x dx = 1$$

$$k \left[-\cos x \right]_0^{\frac{2}{3}\pi} = 1$$

$$k \left[-\cos \frac{2}{3}\pi - (-\cos 0) \right] = 1$$

$$k \left[-(-0.5) + 1 \right] = 1$$

$$1.5k = 1$$

$$k = \frac{2}{3}$$

$$\text{(ii)} \quad \int_0^m \frac{2}{3} \sin x dx = 0.5$$

$$\left[-\frac{2}{3} \cos x \right]_0^m = 0.5$$

$$-\frac{2}{3} \cos m - \left(-\frac{2}{3} \cos 0 \right) = 0.5$$

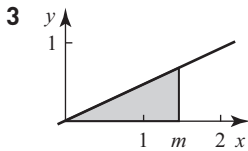
$$-\frac{2}{3} \cos m + \frac{2}{3} = 0.5$$

$$-\frac{2}{3} \cos m = -\frac{1}{6}$$

$$\cos m = \frac{1}{4}$$

$$m = 1.32 \text{ (3 s.f.)}$$

$$\begin{aligned}
 \text{(iii)} \quad E(X) &= \int_0^{\frac{2}{3}\pi} x \times \frac{2}{3} \sin x \, dx \\
 &= \frac{2}{3} \int_0^{\frac{2}{3}\pi} x \sin x \, dx \\
 &= \frac{2}{3} \left[-x \cos x \right]_0^{\frac{2}{3}\pi} - \int_0^{\frac{2}{3}\pi} (-\cos x) \, dx \\
 &= \frac{2}{3} \left[-x \cos x + \sin x \right]_0^{\frac{2}{3}\pi} \\
 &= \frac{2}{3} \left[\left(-\frac{2}{3}\pi \times \cos \frac{2}{3}\pi + \sin \frac{2}{3}\pi \right) - (0 + \sin 0) \right] \\
 &= \frac{2}{3} \left[-\frac{2}{3}\pi \times -\frac{1}{2} + \frac{\sqrt{3}}{2} \right] \\
 &= \frac{2}{3} \left[\frac{\pi}{3} + \frac{\sqrt{3}}{2} \right] \\
 &= \frac{2\pi}{9} + \frac{\sqrt{3}}{3} \\
 &= \frac{2\pi + 3\sqrt{3}}{9} \\
 &= 1.28 \text{ (3 s.f.)}
 \end{aligned}$$



$$\begin{aligned}
 \frac{1}{2} \times m \times \frac{1}{2}m &= \frac{1}{2} \\
 \frac{1}{4}m^2 &= \frac{1}{2} \\
 m^2 &= 2 \\
 m &= \sqrt{2} = 1.41 \text{ (3 s.f.)}
 \end{aligned}$$

or

$$\begin{aligned}
 \int_0^m \left(\frac{1}{2}x \right) dx &= \frac{1}{2} \\
 \left[\frac{x^2}{4} \right]_0^m &= \frac{1}{2} \\
 \frac{m^2}{4} &= \frac{1}{2} \\
 m^2 &= 2 \\
 m &= \sqrt{2} = 1.41 \text{ (3 s.f.)}
 \end{aligned}$$

4 (i) $\sigma_X, \sigma_Z, \sigma_Y, \sigma_W$

(ii) (a) $\mu_X = 0$

$$\begin{aligned}
 \text{Var}(X) &= \int_{-3}^3 \left(x^2 \times \frac{1}{18}x^2 \right) dx - 0^2 \\
 &= \frac{1}{18} \int_{-3}^3 x^4 \, dx \\
 &= \frac{1}{18} \left[\frac{x^5}{5} \right]_{-3}^3 \\
 &= \frac{1}{18} \left[\frac{3^5}{5} - \frac{(-3)^5}{5} \right] \\
 &= \frac{27}{5} \\
 \sigma_X &= \sqrt{\frac{27}{5}} = 2.32 \text{ (3 s.f.)}
 \end{aligned}$$

$$\begin{aligned}
 \text{(b)} \quad P(X > 2.324\dots) &= \int_{2.324}^3 \left(\frac{1}{18}x^2 \right) dx \\
 &= \left[\frac{x^3}{54} \right]_{2.324}^3 \\
 &= 0.268 \text{ (3 s.f.)}
 \end{aligned}$$

$$\begin{aligned}
 \text{(c)} \quad P(X > 2 \times 2.324\dots) &= P(X > 4.648\dots) \\
 &= 0
 \end{aligned}$$

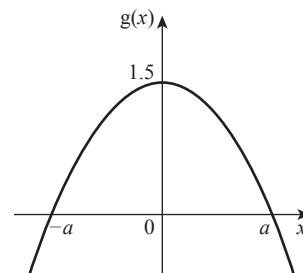
5 (i) (a) $\frac{1}{2} \times 2 \times k = 1 \Rightarrow k = 1$

(b) f_2 – the area under the curve is clearly greater than 1.

f_3 – the curve is below the x -axis; probability cannot be less than 0.

$$\begin{aligned}
 \text{(ii) (a)} \quad \int_{-a}^a 6(a^2 - x^2) dx &= 1 \\
 6 \left[a^2x - \frac{x^3}{3} \right]_{-a}^a &= 1 \\
 6 \left[\left(a^3 - \frac{a^3}{3} \right) - \left(-a^3 - \frac{(-a)^3}{3} \right) \right] &= 1 \\
 6 \left[\left(a^3 - \frac{a^3}{3} \right) - \left(-a^3 + \frac{a^3}{3} \right) \right] &= 1 \\
 6 \left[2a^3 - \frac{2a^3}{3} \right] &= 1 \\
 6 \left[\frac{4a^3}{3} \right] &= 1 \\
 8a^3 &= 1 \\
 a^3 &= \frac{1}{8} \\
 a &= \frac{1}{2}
 \end{aligned}$$

(b) By symmetry $E(X) = 0$



$$\begin{aligned}
 \text{(c)} \quad \text{Var}(X) &= \int_{-0.5}^{0.5} \left(x^2 \times 6 \left(\frac{1}{4} - x^2 \right) \right) dx - 0^2 \\
 &= \int_{-0.5}^{0.5} \left(\frac{3}{2}x^2 - 6x^4 \right) dx \\
 &= \left[\frac{x^3}{2} - \frac{6x^5}{5} \right]_{-0.5}^{0.5} \\
 &= \left[\frac{0.5^3}{2} - \frac{6(0.5)^5}{5} \right] - \left[\frac{(-0.5)^3}{2} - \frac{6(-0.5)^5}{5} \right] \\
 &= \frac{1}{20} = 0.05
 \end{aligned}$$

Stretch and challenge

1 (i) $f(x) = \frac{k}{e^{kx}} > 0$ for $x \geq 0$

$$\begin{aligned}\int_0^{\infty} ke^{-kx} &= [-e^{-kx}]_0^{\infty} \\ &= -e^{-kr} - (-e^0) \\ \text{As } r \rightarrow \infty, e^{-kr} &\rightarrow 0 \\ &= 0 - (-1) \\ &= 1\end{aligned}$$

(ii) Median = 10 $\Rightarrow \int_0^{10} ke^{-kx} dx = \frac{1}{2}$

$$\begin{aligned}[-e^{-kx}]_0^{10} &= \frac{1}{2} \\ -e^{-10k} - (-e^0) &= \frac{1}{2} \\ -e^{-10k} + 1 &= \frac{1}{2} \\ -e^{-10k} &= -\frac{1}{2} \\ e^{-10k} &= \frac{1}{2} \\ -10k &= \ln\left(\frac{1}{2}\right) \\ k &= \frac{\ln\left(\frac{1}{2}\right)}{-10} \\ k &= 0.0693 \text{ (3 s.f.)}\end{aligned}$$

(iii) $P(X > 20) = \int_{20}^{\infty} 0.0693e^{-0.0693x} dx$

$$\begin{aligned}&= [-e^{-0.0693x}]_{20}^{\infty} \\ &= -e^{-0.0693r} - (-e^{-0.0693 \times 20}) \\ \text{As } r \rightarrow \infty, e^{-0.0693r} &\rightarrow 0 \\ &= 0 + e^{-0.0693 \times 20} \\ &= 0.250 \text{ (3 s.f.)}\end{aligned}$$

(iv) $P(0 < X < 1) = \int_0^1 ke^{-kx} dx$

$$\begin{aligned}&= [-e^{-kx}]_0^1 \\ &= -e^{-k} - (-e^0) \\ &= 1 - e^{-k} \\ &= 1 - \frac{1}{e^k} = \frac{e^k - 1}{e^k}\end{aligned}$$

(v) $P(X < x) = \int_0^x ke^{-kx} dx$

$$\begin{aligned}&= [-e^{-kx}]_0^x \\ &= [-e^{-kx} - (-e^0)] \\ &= -e^{-kx} + 1 \\ &= 1 - e^{-kx} \\ &= 1 - \frac{1}{e^{kx}} \\ &= \frac{e^{kx} - 1}{e^{kx}}\end{aligned}$$

2 (i) $\int_0^3 \frac{x^3}{18} dx = \left[\frac{x^4}{54} \right]_0^3$

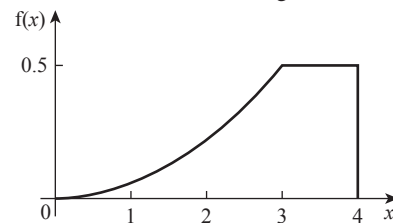
$$= \frac{3^4}{54} = 0.5$$

$P(3 < x \leq 4) = 0.5$

$\Rightarrow k = 0.5$

(ii) $E(X) = \int_0^3 x \times \frac{x^2}{18} dx + \int_3^4 x \times 0.5 dx$

$$\begin{aligned}&= \left[\frac{x^4}{72} \right]_0^3 + \left[\frac{x^2}{4} \right]_3^4 \\ &= \frac{3^4}{72} + \left[\frac{4^2}{4} - \frac{3^2}{4} \right] \\ &= 1.125 + 1.75 \\ &= 2.875 = 287.5 \text{ kg}\end{aligned}$$



(iii) Profit = $50 \times 1.125 + 40 \times 1.75$

$= \$126.25$

3 (i) $P(X \leq x) = \int_{-\frac{1}{2}}^x \left(\frac{1}{2} \pi \cos \pi x \right) dx$

$$\begin{aligned}&= \frac{1}{2} \pi \left[\frac{1}{\pi} \sin \pi x \right]_{-\frac{1}{2}}^x \\ &= \frac{1}{2} \left[\sin \pi x - \sin \left(\pi \times -\frac{1}{2} \right) \right] \\ &= \frac{1}{2} [\sin \pi x + 1]\end{aligned}$$

(iii)

$$\begin{aligned}
 \text{Var}(X) &= \int_{-\frac{1}{2}}^{\frac{1}{2}} x^2 \left(\frac{1}{2} \pi \cos \pi x \right) dx - 0^2 \\
 &= \frac{1}{2} \pi \int_{-\frac{1}{2}}^{\frac{1}{2}} x^2 \cos \pi x dx \\
 &= \frac{1}{2} \pi \left[x^2 \times \frac{1}{\pi} \sin \pi x - \int_{-\frac{1}{2}}^{\frac{1}{2}} 2x \times \frac{1}{\pi} \sin \pi x dx \right] \\
 &= \frac{1}{2} \pi \left[\frac{x^2}{\pi} \sin \pi x - \frac{2}{\pi} \left(x \times -\frac{1}{\pi} \cos \pi x - \int_{-\frac{1}{2}}^{\frac{1}{2}} -\frac{1}{\pi} \cos \pi x dx \right) \right] \\
 &= \frac{1}{2} \pi \left[\frac{x^2}{\pi} \sin \pi x - \frac{2}{\pi} \left(-\frac{x}{\pi} \cos \pi x + \frac{1}{\pi} \times \frac{1}{\pi} \sin \pi x \right) \right]_{-\frac{1}{2}}^{\frac{1}{2}} \\
 &= \frac{1}{2} \pi \left[\left(\frac{1}{2} \right)^2 \frac{\sin \frac{\pi}{2}}{\pi} - \frac{2}{\pi} \left(-\frac{1}{2\pi} \cos \frac{\pi}{2} + \frac{1}{\pi^2} \sin \frac{\pi}{2} \right) \right] \\
 &\quad - \left[\left(-\frac{1}{2} \right)^2 \frac{\sin -\frac{\pi}{2}}{\pi} - \frac{2}{\pi} \left(\frac{1}{2\pi} \cos -\frac{\pi}{2} + \frac{1}{\pi^2} \sin -\frac{\pi}{2} \right) \right] \\
 &= \frac{1}{2} \pi \left[\frac{1}{4\pi} - \frac{2}{\pi} \left(\frac{1}{\pi^2} \right) - \left(-\frac{1}{4\pi} - \frac{2}{\pi} \left(-\frac{1}{\pi^2} \right) \right) \right] \\
 &= \frac{1}{2} \pi \left[\frac{1}{4\pi} - \frac{2}{\pi^3} + \frac{1}{4\pi} - \frac{2}{\pi^3} \right] \\
 &= \frac{1}{2} \pi \left[\frac{1}{2\pi} - \frac{4}{\pi^3} \right] \\
 &= \frac{1}{4} - \frac{2}{\pi^2} \\
 SD(X) &= \sqrt{\frac{1}{4} - \frac{2}{\pi^2}}
 \end{aligned}$$

(iii) $P(-0 \leq X \leq 0) = 2 \times P(-0 \leq X \leq 0)$

$$\begin{aligned}
 &= 2 \times \int_0^{\sqrt{\frac{1}{4} - \frac{2}{\pi^2}}} \frac{1}{2} \pi \cos \pi x dx \\
 &= \pi \left[\frac{1}{\pi} \sin \pi x \right]_0^{\sqrt{\frac{1}{4} - \frac{2}{\pi^2}}} \\
 &= \sin \pi \left(\sqrt{\frac{1}{4} - \frac{2}{\pi^2}} \right) \\
 &= 0.632 \text{ (3 s.f.)}
 \end{aligned}$$

(iv) X : number of people who make a purchase
(out of all the people visiting the store).

$$X \sim B(100, 0.92)$$

$$H_0: p = 0.92, H_1: p > 0.92$$

$$2 \text{ (i) } E(X) = 0.5 \times 10 = 5$$

$$\text{(ii) } H \sim B(10, 0.5)$$

$$P(H \geq 8) = 0.0547 \text{ (3 s.f.)}$$

$$\text{(iii) } 0.0547 = 5.47\% > 5\%$$

Do not reject H_0

Not enough evidence to say that the coin is biased.

$$3 \text{ (i) } E(X) = 0.35 \times 25 = 8.75$$

$$\text{(ii) } P(X \leq 3) = 0.0097 = 0.97\% (< 10\%)$$

(iii) H_0 is not true at the 10% significance level.

There is evidence to say that the support is less than 35%.

$$4 \text{ (i) } E(X) = 0.02 \times 400 = 8$$

$$\text{(ii) } P(X \leq 3)$$

$$= 0.0410 \text{ (3 s.f.)} = 4.10\%$$

(iii) $4.10\% > 4\%$ so do not reject H_0

$$5 \text{ (i) } E(X) = 100 \times 0.92 = 92$$

$$\text{(ii) } P(X \geq 98)$$

$$= 0.0113 \text{ (3 s.f.)}$$

(iii) Since $0.0113 < 0.05$, there is sufficient evidence at the 5% significance level to conclude that the manager is correct.

$$6 \text{ (i) (a) } P(X \leq 2) = 0.0192$$

$$P(X \geq 10) = 0.0192$$

$$P(X \leq 2 \text{ or } X \geq 10) = 0.0384$$

4% or 5% significance level

$$\text{(b) } P(X \leq 2 \text{ or } X \geq 10)$$

$$= P(X \leq 2) + P(X \geq 10)$$

$$= 0.0192 + 0.0192 = 0.0384 = 3.84\%$$

$$\text{(ii) } P(X \geq 9)$$

$$= 0.0729 = 7.29\%$$

 $7.29\% > 5\%$ so accept H_0

$$7 \text{ (i) } H_0: p = 0.25, H_1: p < 0.25$$

$$\text{(ii) } X \sim B(20, 0.25)$$

$$P(X \leq x) \approx 0.05$$

$$P(X \leq 1) = 0.0243 = 2.43\%$$

$$P(X \leq 2) = 0.0913 = 9.13\%$$

The critical region is $X \leq 1$ (iii) $X = 3$ is outside the critical region so do not reject the null hypothesis.

There is insufficient evidence to conclude that the programme has been effective.

3 Hypothesis testing using the binomial distribution

3.1 Binomial hypothesis tests

1 (ii) X : number of voters that support Party A.

$$X \sim B(25, 0.35)$$

$$H_0: p = 0.35, H_1: p < 0.35$$

(iii) X : number of apples that have developed bruising after 1 week.

$$X \sim B(400, 0.02)$$

$$H_0: p = 0.02, H_1: p < 0.02$$

- 8 (i)
- $X \sim B(20, 0.1)$

$$H_0: p = 0.1, H_1: p > 0.1$$

$$P(X \geq 5) = 1 - P(X < 5) \\ = 0.0432 = 4.32\%$$

$$4.32\% < 10\%, \text{ reject } H_0$$

There is evidence to conclude that Samantha has special ability.

- (ii)
- $P(X \geq x) < 0.1$

$$P(X \geq 4) = 0.133 = 13.3\%$$

$$X \geq 5 \text{ gives the value closest to } 10\%$$

So the rejection region is $X \geq 5$

- 9
- $X \sim B\left(30, \frac{1}{6}\right)$

$$H_0: p = \frac{1}{6}, H_1: p < \frac{1}{6}$$

$$P(X \leq 2) = 0.103 \text{ (3 s.f.)} = 10.3\%$$

$$10.3\% > 5\%, \text{ do not reject } H_0$$

There is no evidence that the die is biased.

- 10
- $X \sim B(9, 0.65)$

$$H_0: p = 0.65, H_1: p \neq 0.65$$

$$P(X \leq 2) = 0.0112$$

$$P(X \leq 3) = 0.0536$$

$$\text{So } X \leq 2$$

$$P(X \geq 8) = 0.121$$

$$P(X = 9) = 0.0207$$

$$X = 9 \text{ gives closest value to } 2.5\%$$

Claim rejected for $X = 9$ and $X \leq 2$

- 11 (i)
- $H_0: p = 0.15, H_1: p < 0.15$

$$(ii) P(X \leq 1) = 0.176 = 17.6\%$$

$$17.6\% > 5\%, \text{ do not reject } H_0$$

- (iii) If
- $n \leq 18, P(X \leq 0) > 5\%$

Not sufficient evidence to reject H_0 , even if no patients fail to turn up.

3.2 Type I and Type II errors

		Decision	
		H_0 accepted	H_0 rejected
Reality	H_0 is true	OK	Type I error
	H_0 is false	Type II error	OK

- 2 (i)
- $P(X \geq 8) = 5.47\%$

$$P(X \geq 9) = 1.07\%$$

Critical region is $X \geq 9$

$$P(\text{Type I error}) = 0.0107$$

- (ii)
- $X \sim B(10, 0.6)$

$$P(\text{Type II error}) = P(X < 9 | p = 0.6)$$

$$= 1 - [P(X \geq 9) | p = 0.6]$$

$$= 1 - 0.0464$$

$$= 0.954 \text{ (3 s.f.)}$$

- (iii) The probability that Mo concludes that the coin is not biased when it actually is biased.

- 3 (i)
- $X \sim B(14, 0.35)$

$$H_0: p = 0.35, H_1: p < 0.35$$

$$P(X \leq 3) = 0.221 \text{ (3 s.f.)} = 22.1\%$$

$$22.1\% > 5\%, \text{ do not reject } H_0$$

There is insufficient evidence to conclude that the support has decreased.

- (ii)
- $X \sim B(14, 0.35)$

$$P(X \leq 1) = 0.0205$$

$$P(X \leq 2) = 0.0839$$

$$\text{Critical region is } X \leq 1$$

$$P(\text{type I error}) = 0.0205$$

- (iii)
- $X \sim B(14, 0.2)$

$$P(\text{type II error}) = P(X > 1 | p = 0.2)$$

$$= 1 - [P(X \leq 1 | p = 0.2)]$$

$$= 1 - 0.198 = 0.802 \text{ (3 s.f.)}$$

- 4
- $X \sim B\left(7, \frac{4}{7}\right)$

- (i)
- $H_0: p = \frac{4}{7}, H_1: p > \frac{4}{7}$

$$P(X \geq 6) = 0.124$$

$$P(X \geq 7) = 0.0199$$

$$\text{Critical region is } X \geq 7$$

$$P(\text{Type I error}) = 0.0199$$

This is the probability the council will conclude that the level of contamination is rising, when really it is not rising.

- (ii) Type II error: probability that the council will conclude that the level of contamination is not rising, when really it is rising. Type II error is more important to avoid for public safety.

- 5 (i)
- $H_1: p > \frac{1}{6}$

- (ii)
- $N \sim B\left(240, \frac{1}{6}\right)$

$$N \approx N\left(40, 33\frac{1}{3}\right)$$

$$P(N > n) \leq 0.05$$

$$\frac{n - 40}{\sqrt{33\frac{1}{3}}} = 1.645 \Rightarrow n = 49.5 \text{ (3 s.f.)}$$

$$N \geq 50$$

- (iii)
- $N \sim B\left(240, \frac{1}{4}\right) \Rightarrow N \approx N(60, 45)$

$$P(\text{Type II error}) = P\left(N \leq 49 \mid p = \frac{1}{4}\right)$$

$$= P\left(Z \leq \frac{49.5 - 60}{\sqrt{45}}\right)$$

$$= P(Z \leq -1.565)$$

$$= 0.0588$$

Further practice

- 1 (i) $H_0: p = 0.1, H_1: p > 0.1$
 (ii) X : number of faulty tiles, $X \sim B(12, 0.1)$
 $P(X = 0) = 0.9^{12} = 0.2824$
 $P(X = 1) = \binom{12}{1} 0.1^1 0.9^{11} = 0.3766$
 $P(X = 2) = \binom{12}{2} 0.1^2 0.9^{10} = 0.2301$
 $P(X = 3) = \binom{12}{3} 0.1^3 0.9^9 = 0.0852$
 $P(X \geq 2) = 0.341 (> 0.05)$
 $P(X \geq 3) = 0.111 (> 0.05)$
 $P(X \geq 4) = 0.0256 (< 0.05)$
 Critical region: $X \geq 4$
 (iii) 3 does not lie in critical region, do not reject H_0
 Not sufficient evidence to show that cheaper way increases the proportion of faulty tiles.
- 2 (i) H_0 is rejected when H_0 is true.
 The die is not biased, but Ishani concludes that it is biased.
 X : number of fours in six rolls, $X \sim B(6, 0.25)$
 $P(\text{Type I error}) = P(X = 5 \text{ or } 6)$

$$= \binom{6}{5} (0.25)^5 (0.75)^1 + \binom{6}{6} (0.25)^6 (0.75)^0$$

$$= 0.00464 \text{ (3 s.f.)}$$

 (ii) $P(\text{Type II error})$
 $= P(X \leq 4 | p = 0.5) = 0.891$
- 3 X : number of goals scored from 18 shots.
 $H_0: p = 0.74, H_1: p > 0.74$
 Under H_0 , $X \sim B(18, 0.74)$
 $P(X \geq 16) = 0.116$
 $0.116 > 0.10$, do not reject H_0
 Not sufficient evidence to show that he has improved.

Past exam questions

- 1 $H_0: p = 0.5, H_1: p > 0.5$
 $H \sim B(12, 0.5)$
 $P(X > 9) = P(X = 10, 11, 12)$

$$= \binom{12}{10} (0.5)^{10} (0.5)^2 + \binom{12}{11} (0.5)^{11} (0.5)^1 + \binom{12}{12} (0.5)^{12}$$

$$= 0.0193$$

 $\therefore$ Significance level is 1.93%

- 2 (i) X : number of packets that contain a gift.
 $X \sim B(20, 0.25)$
 $H_0: p = 0.25, H_1: p < 0.25$

$$P(X \leq 1) = \binom{20}{0} (0.25)^0 (0.75)^{20} + \binom{20}{1} (0.25)^1 (0.75)^{19}$$

$$= 0.0243 \text{ (3 s.f.)}$$

$$P(X \leq 2) = 0.0243 + \binom{20}{2} (0.25)^2 (0.75)^{18}$$

$$= 0.0913 \text{ (3 s.f.)} > 5\%$$

 $\therefore$ Critical region is $X \leq 1$.

(ii) $P(\text{Type I error}) = 0.0243$

(iii) 2 is outside the rejection region, so accept H_0 .
 There is no evidence that the proportion of packets with free gifts is less than 25%.

- 3 (i) X : number of 5s in the sample of 25 digits.
 $H_0: p = 0.2, H_1: p < 0.2$
 Under H_0 , $X \sim B(25, 0.2)$

$$P(X \leq 1) = P(X = 0) + P(X = 1)$$

$$= (0.8)^{25} + \binom{25}{1} (0.2)^1 (0.8)^{24}$$

$$= 0.0274$$

Since $0.0274 > 0.025$, accept H_0 . There is not enough evidence to suggest Max's claim is correct that the digit 5 is appearing less often than it should.

(ii) Normal distribution, with $\mu = 1000 \times 0.2 = 200$
 $\sigma^2 = 1000 \times 0.2 \times 0.8 = 160$
 $Y \sim N(200, 160)$

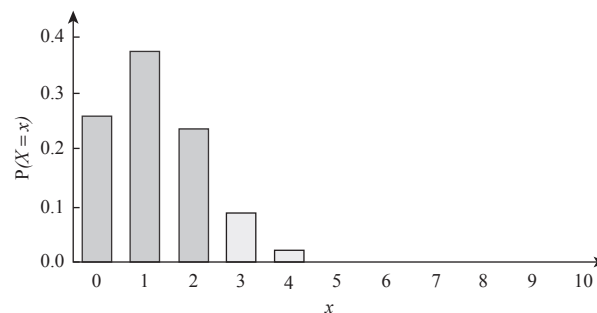
(iii) Accept H_0 when it is false,
 i.e. conclude that the machine is producing the right proportion of 5s when it is actually producing less than it should.

- 4 (i) X : number of correctly predicted race winners.
 $H_0: p = 0.125, H_1: p > 0.125$
 (ii) Under H_0 , $X \sim B(10, 0.125)$
 $P(\text{Type I error}) = P(X \geq 3)$

$$= 1 - P(X = 0, 1, 2)$$

$$= 1 - \left[(0.875)^{10} + \binom{10}{1} (0.125)^1 (0.875)^9 + \binom{10}{2} (0.125)^2 (0.875)^8 \right]$$

$$= 0.120 \text{ (3 s.f.)}$$



(iii) Significance level is 12%

- 5 (i) X : number of packets that contain a gift.

$$H_0: p = 0.3, H_1: p < 0.3$$

- (ii) Under H_0 , $X \sim B(20, 0.3)$

$$P(\text{Type I error}) = P(\text{reject } H_0)$$

$$P(X \leq 2)$$

$$= \left[(0.7)^{20} + \binom{20}{1}(0.3)^1 (0.7)^{19} + \binom{20}{2}(0.3)^2 (0.7)^{18} \right]$$

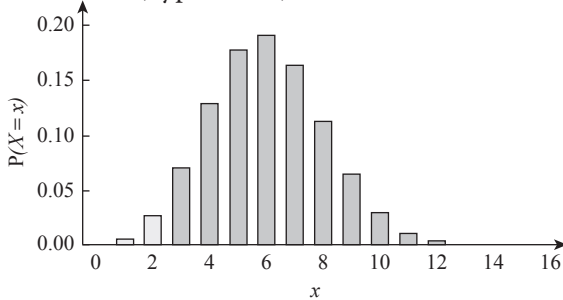
$$= 0.0355 < 0.05$$

$$P(X \leq 3)$$

$$= 0.0355 + \binom{20}{3}(0.3)^3 (0.7)^{17}$$

$$= 0.107 > 0.05$$

$$\text{So } P(\text{Type I error}) = 0.0355$$



- (iii) Since the rejection region is $X \leq 2$

$X = 3$ is outside this so accept H_0 . There is not enough evidence to indicate that the proportion of packets with a free gift is less than 30%.

Stretch and challenge

- 1 $X \sim B(n, p)$

$$H_0: p = 0.4, H_1: p < 0.4$$

$$P(X = 0) = 0.6^n \leq 0.05$$

$$n \log 0.6 \leq \log 0.05$$

$$\Rightarrow n \geq 5.86$$

$$\text{So } n \geq 6$$

Also need n such that

$$P(X = 0) + P(X = 1) \leq 0.05$$

$$0.6^n + \binom{n}{1}(0.4)^1 (0.6)^{n-1} \leq 0.05$$

$$n \leq 9$$

$$6 \leq n \leq 9$$

- 2 $X \sim B(10, p)$

$$\text{Var}(X) = np(1-p) = 10p(1-p)$$

$$10p(1-p) \geq 1.6$$

$$p^2 - p + 0.16 \leq 0, \Rightarrow p = \frac{1}{5} \text{ or } \frac{4}{5}$$

$$p = \frac{1}{5}, \text{ make 5 alternatives for each question.}$$

4 Hypothesis testing and confidence intervals using the normal distribution

4.1 Unbiased estimates

- 1 (i)

Sample	Mean	Variance	Standard deviation	Median	Minimum	Maximum
1, 3	$\frac{1+3}{2} = 2$	1	1	2	1	3
1, 4	2.5	2.25	1.5	2.5	1	4
1, 8	4.5	12.25	3.5	4.5	1	8
3, 4	3.5	0.25	0.5	3.5	3	4
3, 8	5.5	6.25	2.5	5.5	3	8
4, 8	6	4	2	6	4	8
Mean of samples	4	4.33	1.83	4	2.17	5.83
Population mean	4	6.5	2.55	3.5	1	8

- (ii) The mean.

- (iii) When the population is symmetric (like the normal distribution) the median will be an unbiased estimator.

(iv) $\frac{2}{3} \times 6.5 = 4.33$ (3 s.f.) = $E(V)$

$$\begin{aligned} \text{(v)} \quad s^2 &= \frac{n}{n-1} \left(\frac{\sum x^2}{n} - \bar{x}^2 \right) \\ &= \frac{n}{n-1} \left(\frac{\sum x^2}{n} \right) - \frac{n}{n-1} \left(\frac{\sum x}{n} \right)^2 \\ &= \frac{\sum x^2}{n-1} - \frac{(\sum x)^2}{n(n-1)} \\ &= \frac{1}{n-1} \left(\sum x^2 - \frac{(\sum x)^2}{n} \right) \end{aligned}$$

$$2 \quad \bar{x} = \frac{\Sigma x}{n} = \frac{996}{8} = 124.5 \text{ min}$$

$$s^2 = \frac{1}{8-1} \left(128\,440 - \frac{(996)^2}{8} \right) = 634 \text{ min}^2$$

or

$$s^2 = \frac{8}{8-1} \left(\frac{128\,440}{8} - 124.5^2 \right) = 634 \text{ min}^2$$

$$3 \quad \bar{x} = \frac{3492}{24} = 145.5 \text{ s}$$

$$s^2 = \frac{1}{24-1} \left(508\,397 - \frac{(3492)^2}{24} \right) = 13.5 \text{ s}^2 \text{ (3 s.f.)}$$

$$4 \quad \text{Let } y = x - 200$$

$$\Sigma y = 102, \Sigma y^2 = 1143.5$$

$$\bar{y} = \frac{102}{12} = 8.5 \text{ m}$$

$$\bar{x} = \bar{y} + 200 = 208.5 \text{ m}$$

$$s_y^2 = \frac{1}{12-1} \left(1143.5 - \frac{(102)^2}{12} \right) = 25.1 \text{ m}^2 \text{ (3 s.f.)}$$

$$s_y = \sqrt{25.1}$$

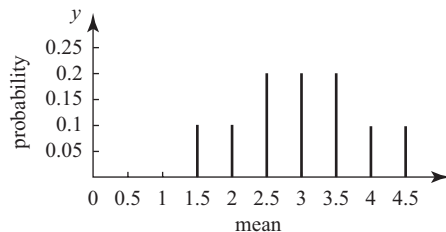
$$= 5.01 \text{ m (3 s.f.)}$$

$$s_y = s_x \Rightarrow s_x = 5.01 \text{ m (3 s.f.)}$$

4.2 The Central Limit Theorem

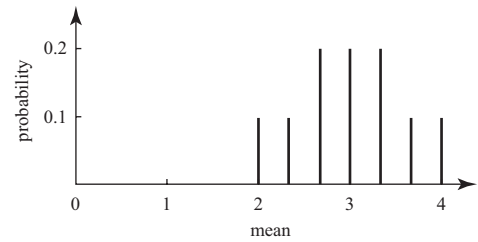
1 (i)

Numbers drawn	Mean	Probability
1, 2	1.5	0.1
1, 3	2	0.1
1, 4	2.5	0.1
1, 5	3	0.1
2, 3	2.5	0.1
2, 4	3	0.1
2, 5	3.5	0.1
3, 4	3.5	0.1
3, 5	4	0.1
4, 5	4.5	0.1



(ii)

Numbers drawn	Mean	Probability
1, 2, 3	2	0.1
1, 2, 4	2.33	0.1
1, 2, 5	2.67	0.1
1, 3, 4	2.67	0.1
1, 3, 5	3	0.1
1, 4, 5	3.33	0.1
2, 3, 4	3	0.1
2, 3, 5	3.33	0.1
2, 4, 5	3.67	0.1
3, 4, 5	4	0.1



$$(iii) \quad \mu = \frac{1+2+3+4+5}{5} = 3$$

$$\sigma^2 = \frac{1^2+2^2+3^2+4^2+5^2}{5} - 3^2 = 2$$

$$(iv) \quad 1.5, 2, 2.5, 2.5, 3, 3, 3.5, 3.5, 4, 4.5$$

$$\bar{x} = \frac{30}{10} = 3$$

$$s^2 = \frac{97.5}{10} - 3^2 = 0.75$$

The mean is the same as the population mean.

The variance is less than the population variance.

$$(v) \quad 2, 2.33, 2.67, 2.67, 3, 3, 3.33, 3.33, 3.67, 4$$

$$\bar{x} = \frac{30}{10} = 3$$

$$s^2 = \frac{93.3}{10} - 3^2 = \frac{1}{3}$$

The mean is the same as the population mean.

The variance is less than the population variance.

$$2 \quad X \sim N(50, 10^2)$$

$$\bar{X} \sim N\left(50, \frac{10^2}{125}\right)$$

$$\begin{aligned} P(\bar{X} > 52) &= P\left(Z > \frac{52-50}{\frac{10}{\sqrt{125}}}\right) \\ &= P(Z > 2.236) \\ &= 1 - 0.9873 \\ &= 0.0127 \text{ (3 s.f.)} \end{aligned}$$

$$3 \quad \text{Var}(\bar{X}) = \frac{\sigma^2}{45} = 32$$

$$\sigma^2 = 1440$$

$$\sigma = 37.9 \text{ (3 s.f.)}$$

- 4 (i) X : length of fish in the lake.

$$\bar{X} \sim N\left(22, \frac{2.6^2}{30}\right)$$

$$\begin{aligned} P(\bar{X} < 21) &= P\left(Z < \frac{21-22}{\frac{2.6}{\sqrt{30}}}\right) \\ &= P(Z < -2.107) \\ &= 1 - 0.9824 \\ &= 0.0176 \text{ (3 s.f.)} \end{aligned}$$

- (ii) No, since the sample size was large enough (>30) by the CLT the sample means are approximately normally distributed.

- 5 $X \sim N(4, 16)$, $\bar{X} \sim N\left(4, \frac{16}{n}\right)$

(i) $E(\bar{X}) = \mu = 4$

(ii) $\text{Var}(\bar{X}) = \frac{16}{n} = 0.16 \Rightarrow n = 100$

- (iii) (i) Would be the same.

- (ii) Would also be the same.

- 6 X : weight of a packet of chips.

(i) $\bar{X} \sim N\left(150, \frac{1.5^2}{15}\right)$

$$\begin{aligned} P(\bar{X} < 149) &= P\left(Z < \frac{149-150}{\frac{1.5}{\sqrt{15}}}\right) \\ &= P(Z < -2.582) \\ &= 1 - 0.9951 \\ &= 0.0049 \text{ (2 s.f.)} \end{aligned}$$

- (ii) No, the sample size (15) was not large enough to guarantee that $\bar{X}$ is approximately normal.

- 7 $X \sim B(20, p)$

$$\Rightarrow E(X) = 20p \text{ and}$$

$$\text{Var}(X) = 20p(1-p)$$

$$E(\bar{X}) = \mu = 20p = 4.8 \Rightarrow p = 0.24$$

$$\text{Var}(\bar{X}) = \frac{\sigma^2}{n} = \frac{20 \times 0.24 \times (1-0.24)}{12} = 0.304$$

- 8 (i) $\int_0^{10} kx \, dx = 1$

$$k \left[\frac{x^2}{2} \right]_0^{10} = 1$$

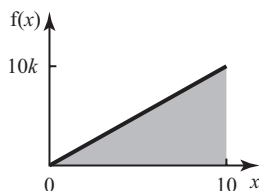
$$k \left[\frac{10^2}{2} - 0 \right] = 1$$

$$50k = 1$$

$$k = \frac{1}{50}$$

OR

$$\frac{1}{2} \times 10 \times 10k = 1 \Rightarrow k = \frac{1}{50}$$



$$\begin{aligned} \text{(ii)} \quad E(X) &= \int_0^{10} x \times \left(\frac{1}{50}x\right) dx \\ &= \frac{1}{50} \left[\frac{x^3}{3} \right]_0^{10} \\ &= \frac{1}{50} \left[\frac{10^3}{3} - 0 \right] \\ &= \frac{20}{3} \end{aligned}$$

$$\begin{aligned} \text{Var}(X) &= \int_0^{10} x^2 \times \left(\frac{1}{50}x\right) dx - \left(\frac{20}{3}\right)^2 \\ &= \frac{1}{50} \left[\frac{x^4}{4} \right]_0^{10} - \left(\frac{20}{3}\right)^2 \\ &= \frac{1}{50} \left[\frac{10^4}{4} - 0 \right] - \left(\frac{20}{3}\right)^2 \\ &= \frac{50}{9} \end{aligned}$$

$$\text{(iii)} \quad \bar{X} \sim N\left(\frac{20}{3}, \frac{50}{9}\right)$$

$$\begin{aligned} P(\bar{X} > 7.2) &= P\left(Z > \frac{7.2 - \frac{20}{3}}{\sqrt{\frac{50}{9}}}\right) \\ &= P(Z > 2.024) \\ &= 0.0215 \text{ (3 s.f.)} \end{aligned}$$

4.3 Hypothesis tests – normal distribution

- 1 (i) $H_0: \mu = 58$

$$H_1: \mu > 58$$

- (ii) Under H_0 , $\bar{X} \sim N\left(58, \frac{8^2}{80}\right)$

$$\frac{a-58}{\frac{8}{\sqrt{80}}} = 1.645$$

$$a = 1.645 \times \frac{8}{\sqrt{80}} + 58 = 59.5$$

- (iii) 64% is in the rejection region, so reject H_0 and conclude that there is evidence that the mean has increased.

- 2 $H_0: \mu = 50$, $H_1: \mu < 50$

(i) 0.02

- (ii) Type II error.

3 X: light bulb intensity.

(i) $H_0: \mu = 1720, H_1: \mu < 1720$

(ii) $\bar{X} \sim N\left(1720, \frac{90^2}{30}\right)$

$$\begin{aligned} P(\bar{X} < 1703) &= P\left(Z < \frac{1703 - 1720}{\frac{90}{\sqrt{30}}}\right) \\ &= P(Z < -1.035) \\ &= 0.150 \text{ (3 s.f.)} \end{aligned}$$

Since $0.150 > 0.05$, accept H_0

There is insufficient evidence to conclude that the intensities are lower.

4 $X \sim N(27\,500, 4000^2)$

(i) $H_0: \mu = 27\,500, H_1: \mu > 27\,500$

(ii) $\bar{X} \sim N\left(27\,500, \frac{4000^2}{15}\right)$

$$\begin{aligned} P(\bar{X} > 30\,450) &= P\left(Z > \frac{30\,450 - 27\,500}{\frac{4000}{\sqrt{15}}}\right) \\ &= P(Z > 2.856) \\ &= 0.0022 \end{aligned}$$

Since $0.00214 < 0.05$, reject H_0 .

There is evidence at the 5% level that the mean lifetime has increased.

5 X: level of protein.

(i) $\bar{X} \sim N\left(12.5, \frac{2^2}{25}\right)$

$$\begin{aligned} P(\bar{X} > 13.2) &= P\left(Z > \frac{13.2 - 12.5}{\frac{2}{\sqrt{25}}}\right) \\ &= P(Z > 1.75) \\ &= 0.0401 \text{ (3 s.f.)} \end{aligned}$$

$\alpha = 4.01$

(ii) $P(\text{Type II error}) = P(\bar{X} < 13.2 \mid \mu = 13.5)$

$$\begin{aligned} &= P\left(Z < \frac{13.2 - 13.5}{\frac{2}{\sqrt{25}}}\right) \\ &= P(Z < -0.75) \\ &= 0.227 \text{ (3 s.f.)} \end{aligned}$$

6 (i) X: level of hormone.

$H_0: \mu = 0.436, H_1: \mu \neq 0.436$

$$\begin{aligned} s^2 &= \frac{1}{120-1} \left(22.8 - \frac{(47.4)^2}{120} \right) \\ &= 0.0343 \text{ (3 s.f.)} \end{aligned}$$

Under $H_0, \bar{X} \sim N\left(0.436, \frac{0.0343}{120}\right)$

$\bar{x} = \frac{47.4}{120} = 0.395$

$$\begin{aligned} P(\bar{X} < 0.395) &= P\left(Z < \frac{0.395 - 0.436}{\sqrt{\frac{0.0343}{120}}}\right) \\ &= P(Z < -2.426) \\ &= 0.0077 \end{aligned}$$

Since $0.00762 > 0.005$, accept H_0 .

(ii) No, because the sample size (120) is large so by the CLT the sample mean distribution is approximately normal.

7 X: weight of cereal.

$X \sim N(\mu, 15^2)$

$\bar{X} \sim N\left(\mu, \frac{15^2}{16}\right)$

(i) $H_0: \mu = 375, H_1: \mu > 375$

$z(0.975) = 1.960$

$$1.96 = \frac{\bar{x} - 375}{\frac{15}{\sqrt{16}}}$$

$\bar{x} = 382.35$

(ii) $P(\text{Type II error}) = P(\bar{X} < 382.35 \mid \mu = 385)$

$$\begin{aligned} &= P\left(Z < \frac{382.35 - 385}{\frac{15}{\sqrt{16}}}\right) \\ &= P(Z < -0.707) \\ &= 0.240 \text{ (3 s.f.)} \end{aligned}$$

(iii) $z(0.025) = -1.960$

$$-1.960 = \frac{382.35 - \mu}{\frac{15}{\sqrt{16}}}$$

$\mu = 389.7$

$\mu > 390 \text{ (3 s.f.)}$

(iv) $P(\text{Type I error}) = 0.025$

$$\begin{aligned} P(\text{at least 1 Type I error}) &= 1 - P(\text{no errors}) \\ &= 1 - 0.975^2 \\ &= 0.0494 \text{ (3 s.f.)} \end{aligned}$$

$$\begin{aligned} 8 \text{ (i)} \quad X &\sim B\left(270, \frac{1}{2}\right) \\ X &\approx N(135, 67.5) \\ \frac{k-135}{\sqrt{67.5}} &= 2.326 \\ k &= 154.11... \\ X &\geq 155 \end{aligned}$$

$$\begin{aligned} \text{(ii)} \quad P(\text{Type I error}) &= P(X \geq 155) \\ &= P\left(Z \geq \frac{154.5 - 135}{\sqrt{67.5}}\right) \\ &= P(Z \geq 2.373) \\ &= 0.0088 \end{aligned}$$

$$\begin{aligned} \text{(iii)} \quad X &\sim B(270, 0.58) \\ X &\approx N(156.6, 65.772) \\ P(\text{Type II error}) &= P(X \leq 154 | p = 0.58) \\ &= P\left(Z < \frac{154.5 - 156.6}{\sqrt{65.772}}\right) \\ &= P(Z < -0.259) \\ &= 0.398 \end{aligned}$$

4.4 Confidence intervals

$$\begin{aligned} 1 \text{ (i)} \quad \bar{x} &= 63.8 \\ 99\% \text{ CI:} \\ 63.8 \pm 2.576 \times \frac{\sqrt{34}}{\sqrt{10}} \\ 63.8 \pm 4.75 \\ [59.1, 68.5] \end{aligned}$$

$$\begin{aligned} \text{(ii)} \quad 2.576 \times \frac{\sqrt{34}}{\sqrt{n}} &\leq 0.4 \\ 2.576 \times \sqrt{34} &\leq 0.4\sqrt{n} \\ 37.551... &\leq \sqrt{n} \\ 1410.1... &\leq n \\ n &\geq 1411 \end{aligned}$$

$$\begin{aligned} 2 \text{ (i)} \quad \bar{x} &= \frac{930.4}{80} = 11.63 \\ s^2 &= \frac{1}{80-1} \left(11024.88 - \frac{(930.4)^2}{80} \right) = 2.59 \text{ (3 s.f.)} \end{aligned}$$

$$\begin{aligned} \text{(ii)} \quad 90\% \text{ CI:} \\ 11.63 \pm 1.645 \times \frac{\sqrt{2.59...}}{80} \\ 11.63 \pm 0.296 \\ [11.3, 11.9] \end{aligned}$$

$$\begin{aligned} \text{(iii)} \quad 2 \times 1.645 \times \frac{\sqrt{2.59...}}{\sqrt{n}} &\leq 0.4 \\ 13.227... &\leq \sqrt{n} \\ 174.9... &\leq n \\ n &\geq 175 \end{aligned}$$

$$\begin{aligned} 3 \quad s &= 2.6, n = 90, z(98\%) = 2.326 \\ \text{Width} &= 2 \times 2.326 \times \frac{2.6}{\sqrt{90}} \\ &= 1.27 \text{ (3 s.f.)} \end{aligned}$$

$$\begin{aligned} 4 \text{ (i)} \quad s^2 &= \frac{1}{50-1} \left(33\,544 - \frac{(1295)^2}{50} \right) \\ &= 0.07143 \text{ kg} \end{aligned}$$

$$\begin{aligned} \text{(ii)} \quad 95\% \text{ CI:} \\ \frac{1295}{50} \pm 1.96 \times \frac{\sqrt{0.07143}}{\sqrt{50}} \\ 25.9 \pm 0.0741 \\ [25.83, 25.97] \end{aligned}$$

(iii) Since 25 is *not* in the CI, we can conclude that there is no evidence that the machine is overfilling the bags.

$$\begin{aligned} 5 \text{ (i)} \quad \bar{x} &= 55 \\ 94\% \text{ CI:} \\ 55 \pm 1.881 \times \frac{15}{\sqrt{15}} \\ 55 \pm 7.275 \\ [47.7, 62.3] \end{aligned}$$

(ii) There is a probability of 0.94 that this interval encloses the true population mean, μ .
or

If we repeated the survey 100 times, 94 of the CIs we constructed would contain μ .

(iii) Every member of the population has an equal chance of being included in the sample.

4.5 Confidence intervals for a proportion

$$\begin{aligned} 1 \text{ (i)} \quad \hat{p} &= \frac{18}{45} = 0.4 \\ 95\% \text{ CI for } p: \\ 0.4 \pm 1.96 \sqrt{\frac{0.4(1-0.4)}{45}} \\ 0.4 \pm 0.143... \\ [0.257, 0.543] \end{aligned}$$

$$\begin{aligned} \text{(ii)} \quad 1.96 \times \sqrt{\frac{0.4(1-0.4)}{n}} &\leq 0.1 \\ \sqrt{\frac{0.24}{n}} &\leq \frac{0.1}{1.96} \\ \frac{0.24}{n} &\leq \left(\frac{0.1}{1.96}\right)^2 \\ 0.24 \div \left(\frac{0.1}{1.96}\right)^2 &\leq n \\ 92.198... &\leq n \end{aligned}$$

$$n \geq 93$$

(iii) Increase the sample size.

Reduce the level of confidence.

2 $\hat{p} = \frac{27}{48} = 0.5625$

98% CI:

$$0.5625 \pm 2.326 \times \sqrt{\frac{0.5625(1-0.5625)}{48}}$$

$$0.5625 \pm 0.1665$$

$$[0.396, 0.729]$$

3 $\hat{p} = 0.307$

95% CI:

$$0.307 \pm 1.96 \sqrt{\frac{0.307(1-0.307)}{500}}$$

$$0.307 \pm 0.0404$$

$$[0.267, 0.347]$$

Since 0.288 (or 28.8%) is in the CI, and the probability is 95% that this interval encloses the population mean, there is no evidence of a significant change in market share.

Further practice

- 1 X : time rats take to complete a maze.

$$X \sim N(9, 1.6^2)$$

$$H_0: \mu = 9, H_1: \mu \neq 9$$

From the sample

$$\bar{x} = 10.46 \text{ min}$$

$$\bar{X} \sim N\left(9, \frac{1.6^2}{5}\right)$$

$$P(\bar{X} > 10.46) = P\left(Z > \frac{10.46 - 9}{\frac{1.6}{\sqrt{5}}}\right)$$

$$= P(Z > 2.040)$$

$$= 0.021 \text{ (3 s.f.)}$$

Since $0.021 < 0.025$, reject H_0 . There is evidence that the mean time has changed.

2 (i) $\bar{x} = \frac{52\,800}{150} = 352 \text{ ml}$

$$s^2 = \frac{1}{150-1} \left(18\,586\,776 - \frac{(52\,800)^2}{150} \right)$$

$$= 7.89 \text{ ml}^{-2}$$

(ii) $\bar{X} \approx N\left(352, \frac{7.89}{40}\right)$

$$P(\bar{X} < 351) = P\left(Z < \frac{351 - 352}{\sqrt{\frac{7.89}{40}}}\right)$$

$$= P(Z < -2.251)$$

$$= 0.0122 \text{ (3 s.f.)}$$

(iii) Yes, as the original distribution of volumes was not given.

- 3 (i) X : time to complete exam.

$$X \sim N(73.4, 3.2^2)$$

$$H_0: \mu = 73.4, H_1: \mu > 73.4$$

$$\bar{X} \sim N\left(73.4, \frac{3.2^2}{n}\right)$$

$$\frac{74.2 - 73.4}{\frac{3.2}{\sqrt{n}}} = 1.871$$

$$0.8 = \frac{1.871 \times 3.2}{\sqrt{n}}$$

$$n = 56$$

We are assuming the standard deviation for this group is the same as the population value of 3.2 min.

(ii) $Z_{\text{crit}} = 1.751$

Since $1.871 > 1.751$, reject H_0 and conclude there is evidence that the teaching method is causing the students to take longer.

- 4 X : length of butterfly.

$$X \sim N(\mu, \sigma^2)$$

(i) $\bar{x} = \frac{774}{120} = 6.45 \text{ cm}$

$$s^2 = \frac{1}{120-1} \left(5087 - \frac{774^2}{120} \right) = 0.796 \text{ cm}^2$$

- (ii) 98% CI is:

$$6.45 \pm 2.326 \times \frac{\sqrt{0.796}}{\sqrt{120}}$$

$$6.45 \pm 0.189$$

$$[6.26, 6.64]$$

(iii) $0.02^3 = 0.000\,008$

- 5 (i) $H_0: \mu = 8500, H_1: \mu > 8500$

$$\text{Under } H_0, \bar{X} \approx N\left(8500, \frac{2100^2}{80}\right)$$

$$P(\bar{X} > 9050)$$

$$= P\left(Z > \frac{9050 - 8500}{\frac{2100}{\sqrt{80}}}\right)$$

$$= P(Z > 2.343) = 0.0095$$

Since $0.0095 < 0.02$, reject H_0 . There is evidence that the mean fees have increased.

Assume that the standard deviation from the population can be used for sample.

- (ii) Yes. The original population distribution was unknown.

(iii) $P(\text{Type I error}) = 0.02$

(iv) Let k be the boundary for the acceptance region.

$$\frac{k - 8500}{\frac{2100}{\sqrt{80}}} = 2.054$$

$$k = 8982.3$$

$P(\text{Type II error})$

$$= P(X \leq 8982.3 | \mu = 9500)$$

$$= P\left(Z \leq \frac{8982.3 - 9500}{\frac{2100}{\sqrt{80}}}\right)$$

$$= P(Z \leq -2.205) = 0.0137 \text{ (3 s.f.)}$$

6 (i) $\hat{p} = \frac{70}{400} = 0.175$

99% CI is:

$$0.175 \pm 2.576 \times \sqrt{\frac{0.175(1 - 0.175)}{400}}$$

$$0.175 \pm 0.0489$$

$$[0.126, 0.224]$$

(ii) $z \times \sqrt{\frac{0.175(1 - 0.175)}{400}} = 0.04$

$$z = 2.105$$

$$P(Z < 2.105) = 0.9823$$

$$1 - 0.9823 = 0.0177$$

$$\text{Total probability} = 1 - 2 \times 0.0177 = 0.9646$$

96.5% confidence interval

Past exam questions

1 (i) 95% CI:

$$41.2 \pm 1.96 \times \frac{\sqrt{32.6}}{\sqrt{50}}$$

$$41.2 \pm 1.583$$

$$[39.6, 42.8]$$

(ii) $2 \times \frac{1}{16} = \frac{1}{8} = 0.125 = 12.5$

$$\alpha = 100 - 12.5 = 87.5\%$$

2 (i) $\bar{x} = \frac{332 + 334 + \dots + 333}{8} = 331.125 \text{ ml}$

$$= 331 \text{ ml (3 s.f.)}$$

$$s^2 = \frac{1}{8-1} \left(877179 - \frac{(2649)^2}{8} \right)$$

$$= 4.125 \text{ or } 4.13 \text{ (3 s.f.)}$$

(ii) 98% CI

$$331 \pm 2.326 \times \frac{\sqrt{4.20}}{\sqrt{50}}$$

$$331 \pm 0.674$$

$$[330, 332] \text{ or } [330.3, 331.7]$$

(iii) 333 is not contained in the CI from (ii).

So claim is *not* supported.

3 X : number of questions correct.

$$X \sim B(100, 0.2)$$

$$H_0: p = 0.2, H_1: p > 0.2$$

$$\text{Under } H_0, X \sim B(100, 0.2)$$

$$\approx N(20, 16)$$

$$P(X \geq 27) = P\left(Z \geq \frac{26.5 - 20}{\sqrt{16}}\right)$$

$$= P(Z \geq 1.625) = 0.0521 \text{ (3 s.f.)}$$

Since $0.0521 > 0.05$, accept H_0 .

Claim is not justified at the 5% level.

4 N : number of words contained in the leading article each day.

$$N \sim N(352, 29)$$

(i) $\bar{N} \sim N\left(352, \frac{29}{10}\right)$

(ii) $P(\bar{N} > 354) = P\left(Z > \frac{354 - 352}{\sqrt{\frac{29}{10}}}\right)$

$$= P(Z > 1.174)$$

$$= 1 - 0.8798$$

$$= 0.120 \text{ (3 s.f.)}$$

5 (i) T : time taken to get to work (min).

$$T \sim N(35.2, 5.6^2)$$

$$H_0: \mu = 35.2, H_1: \mu < 35.2$$

$$\text{Under } H_0, \bar{T} \sim N\left(35.2, \frac{5.6^2}{25}\right)$$

$$P(\bar{T} < 34.7) = P\left(Z < \frac{34.7 - 35.2}{\sqrt{\frac{5.6^2}{25}}}\right)$$

$$= P(Z < -0.446)$$

$$= 1 - 0.6722$$

$$= 0.328 \text{ (3 s.f.)}$$

$$P(\text{Type I error}) = 0.328 \text{ (3 s.f.)}$$

(ii) A Type II error occurs when we accept the null hypothesis when it is false. In this case we will reject the null hypothesis so no chance of a Type II error.

6 (i) $\hat{p} = \frac{0.5672 + 0.6528}{2} = 0.61$

(ii) $z \times \sqrt{\frac{0.61(1 - 0.61)}{700}} = 0.0428$

$$z = 2.322$$

$$P(Z < 2.322) = 0.9899$$

$$1 - 0.9899 = 0.0101$$

$$\text{Total probability} = 1 - 2 \times 0.0101 = 0.9798$$

98% confidence interval

7 X: computer battery life

$$X \sim N(\mu, 1.6^2)$$

$$H_0: \mu = 11, H_1: \mu > 11$$

$$\text{Under } H_0, \bar{X} \approx N\left(11, \frac{1.6^2}{n}\right)$$

$$P(Z < z_1) = 0.95 \Rightarrow z_1 = 1.645$$

$$P(Z < z_2) = 0.975 \Rightarrow z_2 = 1.96$$

$$\frac{11.8 - 11}{\frac{1.6}{\sqrt{n}}} = 1.645 \Rightarrow n = 10.8 \text{ (3 s.f.)}$$

$$\frac{11.8 - 11}{\frac{1.6}{\sqrt{n}}} = 1.96 \Rightarrow n = 15.4 \text{ (3 s.f.)}$$

$$11 \leq n \leq 15$$

$$8 \text{ (i)} \quad \bar{x} = \frac{6.7}{200} = 0.0335$$

$$s^2 = \frac{1}{200-1} \left(0.2312 - \frac{(6.7)^2}{200} \right) = 0.000\,033\,9$$

(ii) X: level of pollutant (g/mL).

$$H_0: \mu = 0.034, H_1: \mu \neq 0.034$$

$$\text{Under } H_0, \bar{X} \approx N\left(0.034, \frac{0.000\,033\,9}{200}\right)$$

$$\begin{aligned} P(\bar{X} < 0.0335) &= P\left(Z < \frac{0.0335 - 0.034}{\sqrt{\frac{0.000\,033\,9}{200}}}\right) \\ &= P(Z < -1.214) \\ &= 1 - 0.8876 \\ &= 0.112 \text{ (3 s.f.)} \end{aligned}$$

Since $0.112 > 0.05$, accept H_0 . There is insufficient evidence to say that the mean pollutant level has changed.

Stretch and challenge

 1 (i) X: number of intervals that contain μ

$$X \sim B(10, 0.95)$$

$$E(X) = 10 \times 0.95 = 9.5$$

$$\text{Var}(X) = 10 \times 0.95 \times 0.05 = 0.475$$

$$\text{SD}(X) = \sqrt{0.475} = 0.689 \text{ (3 s.f.)}$$

$$(ii) \quad P(X = 10) = 0.95^{10} = 0.599 \text{ (3 s.f.)}$$

$$(iii) \quad c^{10} \geq 0.95$$

$$c \geq \sqrt[10]{0.95}$$

$$c \geq 0.995 >$$

$$99.5\%$$

$$2 \text{ (i)} \quad E(\bar{X}_1 - \bar{X}_2) = \mu_1 - \mu_2$$

$$\text{Var}(\bar{X}_1 - \bar{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

$$\bar{X}_1 - \bar{X}_2 \sim N\left(\mu_1 - \mu_2, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right)$$

$$(ii) \quad \bar{x}_N - \bar{x}_T = 72.5 - 64.3 = 8.2$$

$$H_0: \bar{x}_N - \bar{x}_T = 0, H_1: \bar{x}_N - \bar{x}_T > 0$$

$$\text{Var}(\bar{x}_N - \bar{x}_T) = \frac{12.8^2}{32} + \frac{12.8^2}{25} = 11.6736$$

Under H_0 ,

$$\bar{x}_N - \bar{x}_T \sim N(0, 11.6736)$$

$$\begin{aligned} P(\bar{x}_N - \bar{x}_T \geq 8.2) &= P\left(Z \geq \frac{8.2 - 0}{\sqrt{11.6736}}\right) \\ &= P(Z \geq 2.40) \\ &= 0.0082 \end{aligned}$$

Since $0.0082 < 0.01$, reject H_0 . Conclude that there is evidence that there has been improvement in the test scores.

$$\begin{aligned} (iii) \quad 2 \times 1.96 \times \sqrt{\frac{12.8^2}{n} + \frac{12.8^2}{n}} &\leq 3 \\ \sqrt{\frac{327.68}{n}} &\leq \frac{3}{2 \times 1.96} \\ \frac{\sqrt{327.68} \times 2 \times 1.96}{3} &\leq \sqrt{n} \\ 23.65... &\leq \sqrt{n} \\ 559.47 &\leq n \\ n &\geq 560 \end{aligned}$$

$$(iv) \quad E(\bar{P}_1 - \bar{P}_2) = \hat{p}_1 - \hat{p}_2$$

$$\text{Var}(\bar{P}_1 - \bar{P}_2) = \frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}$$

$$\bar{P}_1 - \bar{P}_2 \sim N\left(\hat{p}_1 - \hat{p}_2, \frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}\right)$$

5 The Poisson distribution

5.1 The Poisson distribution and modelling with a Poisson distribution

$$1 \text{ (i)} \quad P(X = 2) = \frac{e^{-1.8} 1.8^2}{2!} = 0.268 \text{ (3 s.f.)}$$

$$(ii) \quad P(X = 5) = \frac{e^{-1.8} 1.8^5}{5!} = 0.0260 \text{ (3 s.f.)}$$

$$(iii) \quad P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-1.8} = 0.835 \text{ (3 s.f.)}$$

$$(iv) \quad P(X = 2) + P(X = 3) + P(X = 4) = 0.501 \text{ (3 s.f.)}$$

$$2 \text{ (i)} \quad X \sim \text{Po}(3)$$

$$P(X = 0) = e^{-3} = 0.0498 \text{ (3 s.f.)}$$

$$(ii) \quad Y \sim \text{Po}(1.5)$$

$$P(Y = 2) = 0.251 \text{ (3 s.f.)}$$

$$(iii) \quad Z \sim \text{Po}(6)$$

$$P(Z \geq 2) = 1 - P(Z = 0, 1) = 0.983 \text{ (3 s.f.)}$$

$$3 \text{ (i)} \quad \sum f = 120$$

$$\sum fx = 312$$

$$\frac{312}{120} = 2.6$$

(ii) The variance, s^2 , is 3.51. Since the mean and variance are not equal, the Poisson model is inappropriate.

- 4 (i) $Y \sim \text{Po}(8)$
 $P(Y = 10) = 0.0993$ (3 s.f.)
- (ii) $Z \sim \text{Po}(1.33)$
 $P(Z \geq 2) = 1 - P(Z = 0.1) = 0.385$ (3 s.f.)
- 5 $P(X \geq 1) = 0.32$
 $1 - P(X = 0) = 0.32$
 $P(X = 0) = 0.68$
 $\frac{e^{-\lambda} \lambda^0}{0!} = 0.68$
 $e^{-\lambda} = 0.68$
 $\lambda = -\ln 0.68 = 0.386$ (3 s.f.)
- 6 (i) mean = variance = 0.15
- (ii) $X \sim \text{Po}(0.15)$
 $P(X > 1) = 1 - P(X = 0, 1) = 0.0102$ (3 s.f.)
- (iii) $Y \sim \text{B}(n, 0.0102)$
 $P(Y \geq 1) = 1 - P(Y = 0) = 0.95$
 $P(Y = 0) = 0.9898^n = 0.05$
 $n = 292$ or 293
- 7 (i) W : number of accidents on a weekday.
 $W \sim \text{Po}(0.8)$, $P(W = 3) = 0.0383$ (3 s.f.)
 S : number of accidents on Saturday or Sunday.
 $S \sim \text{Po}(1.4)$, $P(S = 3) = 0.113$ (3 s.f.)
 $0.0383 \times \frac{5}{7} + 0.113 \times \frac{2}{7} = 0.0596$ (3 s.f.)
- (ii) $P(W = 0) = 0.449$ (3 s.f.)
 $P(S = 0) = 0.247$ (3 s.f.)
 X : number of accidents on any given day.
 $P(X = 0) = 0.449 \times \frac{5}{7} + 0.247 \times \frac{2}{7} = 0.391$ (3 s.f.)
 $P(\text{weekday} | X = 0) = \frac{P(\text{weekday} \cap (X = 0))}{P(X = 0)}$
 $= \frac{P(W = 0)}{P(X = 0)} = \frac{\frac{5}{7} \times 0.449}{0.391} = 0.820$ (3 s.f.)
- 8 (i) $\frac{50}{18} = 2.78$ (3 s.f.)
- (ii) $X \sim \text{Po}(2.8)$
 $P(X < 4) = 0.692$ (3 s.f.)
- (iii) $Y \sim \text{Po}(5.6)$
 $P(Y = 3) = 0.108$ (3 s.f.)
- (iv) $\sigma^2 = 2.8, \sigma = 1.67$ (3 s.f.)
 $a = 2.8 - 3 \times 1.67$
 $= -2.22$ (3 s.f.)
 Since $a \geq 0$, $a = 0$
 $b = 2.8 + 3 \times 1.67$
 $= 7.82$ (3 s.f.)

- (v) $P(X = 0) = 0.0608$ (3 s.f.)
 $Z \sim \text{B}(12, 0.0608)$
 $P(Z = 3) = \binom{12}{3} 0.0608^3 (1 - 0.0608)^9$
 $= 0.0281$ (3 s.f.)

5.2 The sum of two or more Poisson distributions

- 1 (i) $X + Y \sim \text{Po}(8)$
- (ii) $X - Y$ not Poisson
- (iii) $2X + 3Y$ not Poisson
- (iv) $4X + 4Y$ not Poisson
- 2 (i) $T \sim \text{Po}(5)$
- (ii) $P(T = 1) = P(X + Y = 1) = \frac{e^{-5} 5^1}{1!} = 0.0337$ (3 s.f.)
- (iii) $P(T \geq 2) = 1 - P(T = 0 \text{ or } 1)$
 $= 1 - e^{-5}(1 + 5)$
 $= 0.960$ (3 s.f.)
- 3 (i) X : number of accidents per day in N direction
 Y : number of accidents per day in S direction
 Z : number of accidents in 7 days in both directions
 $X \sim \text{Po}(0.25)$
 $Y \sim \text{Po}(0.4)$
 Number of accidents per day in both directions is Poisson with mean $0.25 + 0.4 = 0.65$
 $Z \sim \text{Po}(7 \times 0.65)$
 $Z \sim \text{Po}(4.55)$
 $P(Z = 0) = e^{-4.55} = 0.0106$
- (ii) K : number of accidents in 2 weeks.
 $K \sim \text{Po}(9.1)$
 $P(K \geq 3) = 1 - e^{-9.1} \left(1 + 9.1 + \frac{9.1^2}{2!} \right) = 0.00994$ (3 s.f.)
- (iii) L : number of accidents in 10 days.
 M : number of accidents in N direction in 10 days.
 N : number of accidents in S direction in 10 days.
 $M \sim \text{Po}(2.5)$, $N \sim \text{Po}(4)$, $L \sim \text{Po}(6.5)$
 $P(M = 2 | L = 6) = \frac{P(M = 2 \cap L = 6)}{P(L = 6)}$
 $= \frac{P(M = 2 \cap N = 4)}{P(L = 6)}$
 $= \frac{\frac{e^{-2.5} 2.5^2}{2!} \times \frac{e^{-4} 4^4}{4!}}{\frac{e^{-6.5} 6.5^6}{6!}}$
 $= 0.318$
- 4 (i) Cars arrive at random.
 Cars arrive at an average rate.
 The arrival of one car does not affect the arrival of another car.

- (iii) X : number of cars arriving at the gas station in a 2-minute interval.

$$X \sim \text{Po}(7.6)$$

$$P(X \geq 4) = 1 - e^{-7.6} \left[1 + 7.6 + \frac{7.6^2}{2!} + \frac{7.6^3}{3!} \right] \\ = 0.945 \text{ (3 s.f.)}$$

- (iii) Z : total number of arrivals at the gas station in a 3-minute interval.

$$Z \sim \text{Po}(13.8)$$

$$P(Z < 2) = 0.0000150 \text{ (3 s.f.)}$$

- (iv) Y : number of trucks arriving at the gas station in a 1-minute interval.

$$Y \sim \text{Po}(0.8)$$

$$P(X \geq 1) = 0.978 \text{ (3 s.f.)}$$

$$P(Y = 1) = 0.359 \text{ (3 s.f.)}$$

$$P(X \geq 1 \text{ and } Y = 1) = 0.978 \times 0.359 = 0.351 \text{ (3 s.f.)}$$

5.3 The Poisson approximation to the binomial distribution

- (i) Yes, $\lambda = 4$

(ii) Yes, $\lambda = 2.72$

(iii) Can't use approximation.
- (i) Faults in capacitors are independent; the probability that a capacitor is faulty is the same for each capacitor.
Capacitor is either faulty or not.
There are a fixed number of capacitors.

(ii) n large, p small, $np < 5$
 $\lambda = 0.4$

(iii) $P(X \geq 3) = 1 - P(X = 0, 1, 2)$

$$= 1 - \left[e^{-0.4} \left(1 + 0.4 + \frac{0.4^2}{2!} \right) \right] \\ = 0.00793 \text{ (3 s.f.)}$$
- (i) X : number of 'out of shape' cans in a crate.
 $X \sim \text{B}(24, 0.001)$
Since $np < 5$ use Poisson approximation.
 $X \approx \text{Po}(0.024)$
 $P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-0.024} = 0.0237 \text{ (3 s.f.)}$

(ii) Y : number of boxes that have at least 1 'out of shape' can.
 $Y \sim \text{B}(30, 0.0237)$
 $P(Y > 2)$
 $= 1 - P(Y = 0, 1, 2)$

$$= 1 - \left[\binom{30}{0} 0.0237^0 0.976^{30} + \binom{30}{1} 0.0237^1 0.976^{29} \right. \\ \left. + \binom{30}{2} 0.0237^2 0.976^{28} \right] \\ = 1 - 0.966... \\ = 0.0336 \text{ (3 s.f.)}$$

- (iii) C : number of 'out of shape' cans in 30 boxes.

$$C \sim \text{B}(720, 0.001)$$

Since $np = 0.72 (< 5)$ use Poisson approximation.

$$C \approx \text{Po}(0.72)$$

$$P(C > 2) = 1 - P(C = 0, 1, 2)$$

$$= 1 - \left[e^{-0.72} \left(1 + 0.72 + \frac{0.72^2}{2!} \right) \right] \\ = 1 - 0.96338... \\ = 0.0366 \text{ (3 s.f.)}$$

- 4 (i) $X \sim \text{B}(94, 0.05)$

- (ii) n large, p small, $np < 5$

- (iii) Y : number of no-shows per day in August.

$$Y \approx \text{Po}(4.7)$$

(a) $P(Y = 4) = \frac{e^{-4.7} 4.7^4}{4!} = 0.185 \text{ (3 s.f.)}$

(b) $P(Y \geq 4) = 1 - P(Y = 0, 1, 2, 3)$

$$= 1 - \left[e^{-4.7} \left(1 + 4.7 + \frac{4.7^2}{2!} + \frac{4.7^3}{3!} \right) \right]$$

$$1 - 0.3097 = 0.690$$

- (iv) Z : number of days that there are enough rooms.

$$Z \sim \text{B}(31, 0.690)$$

$$P(Z = 31) = 0.690^{31} = 0.0000102 \text{ (3 s.f.)}$$

5.4 Using the normal distribution as an approximation for the Poisson distribution

- A : the number of accidents in one year
 $A \sim \text{Po}(36)$
Since $\lambda > 15$, use normal approximation.
 $A \approx \text{N}(36, 36)$

$$P(A \geq 50) = P\left(Z \geq \frac{49.5 - 36}{\sqrt{36}}\right) \\ = P(Z \geq 2.25) \\ = 1 - 0.9878 \\ = 0.0122$$
- (i) X : number of calls in 5 minutes.
 $X \sim \text{Po}(3.2)$
 $P(X = 1) = 0.130 \text{ (3 s.f.)}$

(ii) Y : number of calls in 1 minute.
 $Y \sim \text{Po}(0.64)$
 $P(Y = 1) = 0.337 \text{ (3 s.f.)}$

(iii) $0.337^5 = 0.00438 \text{ (3 s.f.)}$

(iv) I : number of calls per hour.
 $I \sim \text{Po}(38.4)$
Since $\lambda > 15$
 $I \approx \text{N}(38.4, 38.4)$
 $P(I \leq 45) = P\left(Z \leq \frac{45.5 - 38.4}{\sqrt{38.4}}\right) \\ = P(Z \leq 1.146) \\ = 0.874 \text{ (3 s.f.)}$

(v) Bank: assumptions likely to be valid.
Emergency room: calls not likely to be independent.

- 3 (i)** Meteors are seen randomly.
Meteors occur independently.
The average rate at which meteors are seen is constant.

- (ii)** X : number of meteors seen in 1 minute.
 $X \sim \text{Po}(1.3)$

(a) $P(X = 1) = \frac{e^{-1.3} 1.3^1}{1!} = 0.354$ (3 s.f.)

(b) $P(X \geq 4) = 1 - P(X \leq 3)$

$$= 1 - \left[e^{-1.3} \left(1 + 1.3 + \frac{1.3^2}{2!} + \frac{1.3^3}{3!} \right) \right]$$

$$= 0.0431$$
 (3 s.f.)

- (iii)** Y : number of meteors in 10 min, $Y \sim \text{Po}(13)$

$$P(Y = 10) = \frac{e^{-13} 13^{10}}{10!} = 0.0859$$

- (iv)** W : number of meteors in one hour, $W \sim \text{Po}(78)$
Since $\lambda > 15$, use normal approximation.
 $W \approx N(78, 78)$

$$P(W \geq 100) = P\left(Z \geq \frac{99.5 - 78}{\sqrt{78}}\right)$$

$$= P(Z \geq 2.434)$$

$$= 1 - 0.9926 = 0.0074$$

- (v)** M : no meteors in t minutes.

$$M \sim \text{Po}(1.3t)$$

$$P(M \geq 1) \geq 0.99$$

$$1 - e^{-1.3t} \geq 0.99$$

$$0.01 \geq e^{-1.3t}$$

$$\ln 0.01 \geq -1.3t$$

$$t \geq 3.54$$
 (3 s.f.)

Smallest integer value of t is 4.

5.5 Hypothesis test for the mean of a Poisson distribution

- 1** $H_0: \lambda = 8, H_1: \lambda < 8$
 $P(X \leq 2) = 0.0138$ (3 s.f.)
 $0.0138 < 0.05$, reject H_0
 There is enough evidence to suggest that the changes have decreased the number of accidents.

- 2 (i)** $H_0: \lambda = 2.4, H_1: \lambda > 2.4$

(ii) $P(X \geq 4) = 0.221$ (3 s.f.)

$$P(X \geq 5) = 0.0959$$
 (3 s.f.)

5 or more hurricanes.

- (iii)** $X \sim \text{Po}(3.6)$

$$P(\text{Type II error}) = P(X < 5 \mid \lambda = 3.6)$$

$$= e^{-3.6} \left(1 + 3.6 + \frac{3.6^2}{2!} + \frac{3.6^3}{3!} + \frac{3.6^4}{4!} \right)$$

$$= 0.706$$
 (3 s.f.)

- 3** X : number of taxis in 5 hours.

$$X \sim \text{Po}(21), X \approx N(21, 21)$$

$$H_0: \lambda = 21, H_1: \lambda < 21$$

$$P(X \leq 14) = P\left(Z \leq \frac{14.5 - 21}{\sqrt{21}}\right)$$

$$= P(Z \leq -1.418)$$

$$= 0.078$$

$0.078 > 0.05$, do not reject H_0

Not enough evidence to prove that the suspicions are correct.

- 4 (i)** $X \sim \text{Po}(32), X \approx N(32, 32)$

$$H_0: \lambda = 32, H_1: \lambda \neq 32$$

$$P(X \geq 36) = P\left(Z \geq \frac{35.5 - 32}{\sqrt{32}}\right)$$

$$= P(Z \geq 0.619)$$

$$= 0.268$$

$26.8\% > 0.5\%$, do not reject H_0

No evidence to suggest that the number of faults has changed.

- (ii)** $P(X \geq a) \leq 0.005 \Rightarrow P\left(Z \geq \frac{a - 32}{\sqrt{32}}\right) \leq 0.005$

$$\frac{a - 32}{\sqrt{32}} = 2.576 \Rightarrow a = 46.6$$
 (3 s.f.)

H_0 rejected for $X \geq 47$ and $X \leq 17$ (by symmetry).

- (iii)** $P(X \geq 47) = P\left(Z \geq \frac{46.5 - 32}{\sqrt{32}}\right)$

$$= P(Z \geq 2.563)$$

$$= 0.0052$$

By symmetry, $P(X \leq 17) = 0.0052$

Actual significance level is 2×0.0052

$$= 0.0104 = 1.04\%$$
 (3 s.f.)

- (iv)** 0.0104 (3 s.f.)

- (v)** $X \sim \text{Po}(48), X \approx N(48, 48)$

$$P(\text{Type II error}) = P(18 \leq X \leq 46 \mid \lambda = 48)$$

$$= P\left(\frac{17.5 - 48}{\sqrt{48}} \leq Z \leq \frac{46.5 - 48}{\sqrt{48}}\right)$$

$$= P(-4.402 \leq Z \leq -0.217)$$

$$\approx 1 - P(Z \leq 0.217)$$

$$= 0.414$$

Further practice

- 1 (i)** Errors occur independently and randomly.
Errors occur at an average rate.
Errors occur one at a time.
- (ii)** X : number of errors per page
 $X \sim \text{Po}(0.85)$
 $P(X \geq 2) = 0.209$ (3 s.f.)

- (iii) Y : number of errors in 10 pages

$$Y \sim \text{Po}(8.5)$$

$$P(Y = 10) = 0.110 \text{ (3 s.f.)}$$

- (iv) $P(Y \leq k) < 0.1$

$$P(Y \leq 4) = 0.0744, P(Y \leq 5) = 0.1496$$

$$k = 4$$

- (v) I : number of errors in 30 pages

$$I \sim \text{Po}(25.5), I \approx N(25.5, 25.5)$$

$$P(I \leq 30) = P\left(Z \leq \frac{30.5 - 25.5}{\sqrt{25.5}}\right)$$

$$= P(Z \leq 0.9901\dots)$$

$$= 0.839 \text{ (3 s.f.)}$$

- 2 (i) (a) X : number of butterflies in 1 minute.

$$X \sim \text{Po}(2.1)$$

$$P(X \geq 2) = 0.620 \text{ (3 s.f.)}$$

- (b) Y : number of butterflies in 5 minutes.

$$Y \sim \text{Po}(10.5)$$

$$P(8 \leq Y \leq 10) = e^{-10.5} \left[\frac{10.5^8}{8!} + \frac{10.5^9}{9!} + \frac{10.5^{10}}{10!} \right]$$

$$= 0.342 \text{ (3 s.f.)}$$

- (ii) I : number of butterflies in 1 hour.

$$I \sim \text{Po}(126), I \approx N(126, 126)$$

$$P(I \geq 130) = P\left(Z \leq \frac{129.5 - 126}{\sqrt{126}}\right)$$

$$= P(Z \geq 0.3118\dots)$$

$$= 0.378 \text{ (3 s.f.)}$$

- (iii) Events are not independent of each other.

- (iv) S : number of single butterflies, $S \sim \text{Po}(1.7)$

$$T$$
: number of butterflies in pairs, $T \sim \text{Po}(0.2)$

$$P(\text{No more than 3 total})$$

$$= P(S \leq 3 \text{ and } T = 0) \text{ OR } P((S = 0 \text{ or } 1) \text{ and } T = 1)$$

$$= 0.9068 \times 0.8187 + 0.4932 \times 0.1637$$

$$= 0.823 \text{ (3 s.f.)}$$

- 3 (i) X : number of faulty laptops.

$$X \sim B(850, 0.005)$$

$$X \approx \text{Po}(4.25)$$

$$p \text{ small, } n \text{ large, } np < 5$$

- (ii) (a) $P(X > 2) = 1 - P(X = 0, 1, 2)$

$$= 1 - \left[e^{-4.25} \left(1 + 4.25 + \frac{4.25^2}{2!} \right) \right]$$

$$= 0.796 \text{ (3 s.f.)}$$

$$(b) \frac{e^{-4.25} 4.25^n}{n!} > \frac{e^{-4.25} 4.25^{n+1}}{(n+1)!}$$

$$\frac{(n+1)!}{n!} > \frac{4.25^{n+1}}{4.25^n}$$

$$n+1 > 4.25$$

$$n > 3.25$$

$$\text{Smallest value is } n = 4.$$

- 4 (i) Y : number of calls in 5 minutes.

$$Y \sim \text{Po}(9)$$

$$P(Y > 3) = 0.979 \text{ (3 s.f.)}$$

- (ii) Z : number of calls in 48 hours.

$$Z \sim \text{Po}(5184), Z \approx N(5184, 5184)$$

$$P(Z < 5000) = P\left(Z < \frac{4999.5 - 5184}{\sqrt{5184}}\right)$$

$$= P(Z < -2.5625)$$

$$= 0.0052$$

- 5 (i) $Z = X + Y$

$$Z \sim \text{Po}(11)$$

$$P(X = 2 \text{ and } Y = 8 | X + Y = 10)$$

$$= \frac{P(X = 2 \text{ and } Y = 8)}{P(X + Y = 10)}$$

$$= \frac{0.147 \times 0.130}{0.119} = 0.160 \text{ (3 s.f.)}$$

- (ii) $P(X = r) = \frac{3}{4} P(X = r + 1)$

$$\frac{e^{-4} 4^r}{r!} = \frac{3}{4} \frac{e^{-4} 4^{r+1}}{(r+1)!}$$

$$\frac{(r+1)!}{r!} = \frac{3}{4} \times \frac{4^{r+1}}{4^r}$$

$$4(r+1) = 3 \times 4$$

$$4r + 4 = 12$$

$$r = 2$$

Past exam questions

- 1 $X \sim B\left(6000, \frac{2}{10\,000}\right)$

$$E(X) = np = 6000 \times \frac{2}{10\,000} = 1.2$$

$$\text{Since } np < 5 \text{ use Poisson.}$$

$$X \sim \text{Po}(1.2)$$

$$P(X > 2) = 1 - P(X = 0, 1, 2)$$

$$= 1 - e^{-1.2} \left(1 + 1.2 + \frac{1.2^2}{2} \right)$$

$$= 0.121 \text{ (3 s.f.)}$$

- 2 $\lambda = 2.1$ / month, $\lambda = 6.3$ (3 months)

$$(i) P(X \leq 1) = e^{-6.3} (1 + 6.3) = 0.0134$$

$$P(X \leq 2) = e^{-6.3} \left(1 + 6.3 + \frac{6.3^2}{2} \right) = 0.0498 > 0.02$$

$$\text{Critical region is } X \leq 1$$

$$(ii) P(\text{Type I error}) = P(X \leq 1) = 0.0134$$

$$(iii) H_0: \lambda = 6.3, H_1: \lambda < 6.3$$

$$3 \text{ is outside the critical region, so accept } H_0.$$

$$\text{There is no evidence that the mean has decreased.}$$

$$(iv) \lambda = 25.2 \text{ (12 months)}$$

$$\text{Since } \lambda > 15 \text{ use normal approximation.}$$

$$P(X < 20) = P\left(Z < \frac{19.5 - 25.2}{\sqrt{25.2}}\right)$$

$$= P(Z < -1.135)$$

$$= 0.128 \text{ (3 s.f.)}$$

- 3 X : number of lions seen per day standard $X \sim \text{Po}(0.8)$
 Y : number of lions seen per day off-road $Y \sim \text{Po}(2.7)$

$$\begin{aligned} \text{(i)} \quad P(X \geq 2) &= 1 - P(X = 0, 1) \\ &= 1 - e^{-0.8}(1 + 0.8) \\ &= 0.191 \text{ (3 s.f.)} \end{aligned}$$

$$\begin{aligned} \text{(ii)} \quad \lambda &= 3 \times 0.8 + 2 \times 2.7 = 7.8 \\ P(\text{total} < 5) &= P(0, 1, 2, 3, 4) \\ &= e^{-7.8} \left(1 + 7.8 + \frac{7.8^2}{2!} + \frac{7.8^3}{3!} + \frac{7.8^4}{4!} \right) \\ &= 0.112 \text{ (3 s.f.)} \end{aligned}$$

$$\begin{aligned} \text{(iii)} \quad \lambda &= 0.8n \\ L: \text{ number of lions seen in } n \text{ days.} \\ P(L = 0) &< 0.1 \\ e^{-0.8n} &< 0.1 \\ -0.8n &< \ln 0.1 \\ n &> \frac{\ln 0.1}{-0.8} \\ n &> 2.878... \\ \text{Smallest value of } n &\text{ is } 3. \end{aligned}$$

$$\begin{aligned} \text{4 (i)} \quad R &= X_1 + X_2 + X_3 \\ E(R) &= E(X_1 + X_2 + X_3) = 3 \times 1.6 = 4.8 \\ P(R < 4) &= P(R = 0, 1, 2, 3) \\ &= e^{-4.8} \left(1 + 4.8 + \frac{4.8^2}{2!} + \frac{4.8^3}{3!} \right) \\ &= 0.294 \text{ (3 s.f.)} \end{aligned}$$

$$\begin{aligned} \text{(ii)} \quad S &= X_1 + X_2 + \dots + X_n \\ E(S) &= 1.6n \\ \frac{e^{-1.6n}(1.6)^4}{4!} &= \frac{16e^{-1.6n}}{3} \times \frac{(1.6n)^2}{2!} \\ (1.6n)^2 &= \frac{16 \times 4!}{3 \times 2!} \\ 2.56n^2 &= 64 \\ n^2 &= 25 \\ n &= 5 \end{aligned}$$

$$\begin{aligned} \text{(iii)} \quad T &= X_1 + X_2 + \dots + X_{40} \\ E(T) &= 40 \times 1.6 = 64 \\ T &\sim \text{Po}(64) \\ \text{Since } \lambda > 15 &\text{ use normal approximation.} \\ T &\approx N(64, 64) \\ P(T > 75) &= P\left(Z > \frac{75.5 - 64}{\sqrt{64}}\right) \\ &= P(Z > 1.438) \\ &= 0.0753 \end{aligned}$$

- 5 X : number of calls in a 5-minute period.
 $X \sim \text{Po}(\lambda)$

Since $\lambda > 30$ use normal approximation.

$$\begin{aligned} X &\approx N(\lambda, \lambda) \\ P(X > 55.5) &= 0.01 \\ P\left(Z > \frac{55.5 - \lambda}{\sqrt{\lambda}}\right) &= 0.01 \\ \frac{55.5 - \lambda}{\sqrt{\lambda}} &= 2.326 \\ 55.5 - \lambda &= 2.326\sqrt{\lambda} \\ \lambda + 2.326\sqrt{\lambda} - 55.5 &= 0 \\ \text{Let } t &= \sqrt{\lambda} \\ t^2 + 2.326t - 55.5 &= 0 \\ t &= 6.344 \text{ or } -8.670 \\ \sqrt{\lambda} &= 6.377 \text{ or } -8.703 \\ \lambda &= 40.7 \text{ (3 s.f.)} \end{aligned}$$

- 6 (i) (a) S : number of defective spoons in sample of 80.

$$\begin{aligned} S &\sim B(80, 0.01), S \approx \text{Po}(0.8) \\ F: \text{ number of defective forks in sample of 60.} \\ F &\sim B(60, 0.015), F \approx \text{Po}(0.9) \\ P(S \geq 1) \text{ and } P(F \geq 1) \\ &= (1 - e^{-0.8}) \times (1 - e^{-0.9}) \\ &= 0.551 \times 0.593 \\ &= 0.327 \text{ (3 s.f.)} \end{aligned}$$

- (b) T : total number of defective items.

$$\begin{aligned} \lambda &= 80 \times 0.01 + 60 \times 0.015 = 1.7 \\ P(T < 3) &= P(T = 0, 1, 2) \\ &= e^{-1.7} \left(1 + 1.7 + \frac{1.7^2}{2} \right) \\ &= 0.757 \text{ (3 s.f.)} \end{aligned}$$

$$\begin{aligned} \text{(ii)} \quad P(X = 1) &= p \Rightarrow e^{-\lambda} \lambda = p \\ P(X = 2) &= 1.5p \Rightarrow \frac{e^{-\lambda} \lambda^2}{2!} = 1.5p \end{aligned}$$

Substituting the value of p into 2nd equation:

$$\begin{aligned} \frac{e^{-\lambda} \lambda^2}{2!} &= 1.5e^{-\lambda} \lambda \\ e^{-\lambda} \lambda^2 &= 3e^{-\lambda} \lambda \\ \lambda^2 &= 3\lambda \end{aligned}$$

$$\lambda^2 - 3\lambda = 0$$

$$\lambda(\lambda - 3) = 0$$

$$\lambda = 0 \text{ or } 3$$

When $\lambda = 3$

$$p = e^{-3} \times 3 = 0.149 \text{ (3 s.f.)}$$

7 (i) $X \sim \text{Po}(42)$

(a) Since λ is large, $X \approx N(42, 42)$

$$\begin{aligned} P(X \geq 40) &= P\left(Z \geq \frac{39.5 - 42}{\sqrt{42}}\right) \\ &= P(Z \geq -0.386) \\ &= 0.650 \text{ (3 s.f.)} \end{aligned}$$

(b) Since λ is large (> 15) normal distribution is appropriate.

(ii) $Y \sim B(60, 0.02)$

(a) $\mu = 60 \times 0.02 = 1.2$ so $Y \approx \text{Po}(1.2)$

$$\begin{aligned} P(Y > 2) &= 1 - P(Y = 0, 1, 2) \\ &= 1 - e^{-1.2} \left(1 + 1.2 + \frac{1.2^2}{2!}\right) \\ &= 0.121 \text{ (3 s.f.)} \end{aligned}$$

(b) $\mu = 60 \times 0.02 = 1.2 < 5$ so Poisson is appropriate.

Stretch and challenge

1 (i) X : number of earthquakes per year. $X \sim \text{Po}(\lambda)$

$$\begin{aligned} P(X \geq 1) &= \frac{1}{3} \Rightarrow P(X = 0) = \frac{2}{3} \\ &\Rightarrow e^{-\lambda} = \frac{2}{3} \\ &\Rightarrow \lambda = -\ln \frac{2}{3} = 0.405 \end{aligned}$$

$$\text{Mean} = \lambda = 0.405 \text{ (3 s.f.)}$$

$$\text{Standard deviation} = \sqrt{\lambda} = 0.637 \text{ (3 s.f.)}$$

$$(ii) P(X = x) = \frac{e^{-\lambda t} (\lambda t)^x}{x!}$$

$$\lambda = 0.405, t = 4$$

$$\begin{aligned} \Rightarrow P(X = x) &= \frac{e^{-0.405 \times 4} (0.405 \times 4)^x}{x!} \\ &= \frac{e^{-1.622} (1.622)^x}{x!} \end{aligned}$$

$$\begin{aligned} P(X \leq 2) &= P(X = 0, 1, 2) \\ &= e^{-1.622} \left(1 + 1.622 + \frac{1.622^2}{2!}\right) \\ &= 0.778 \text{ (3 s.f.)} \end{aligned}$$

$$(iii) P(X = 1) = \frac{e^{-0.405} (0.405)^1}{1!} = 0.270$$

$$\begin{aligned} &\text{P(Exactly one earthquake in each of the next 3 years)} \\ &= 0.270^3 = 0.0198 \text{ (3 s.f.)} \end{aligned}$$

$$(iv) P(X = 0) = \frac{e^{-0.405t} (0.405t)^0}{0!} = e^{-0.405t}$$

(v) $P(X = 0)$ for $t \geq 16$

$$\begin{aligned} &= e^{-16\lambda} + e^{-17\lambda} + e^{-18\lambda} + e^{-19\lambda} + \dots \\ &= (e^{-\lambda})^{16} + (e^{-\lambda})^{17} + (e^{-\lambda})^{18} + (e^{-\lambda})^{19} + \dots \\ &= \left(\frac{2}{3}\right)^{16} + \left(\frac{2}{3}\right)^{17} + \left(\frac{2}{3}\right)^{18} + \left(\frac{2}{3}\right)^{19} + \dots \\ &\text{GP with } a = \left(\frac{2}{3}\right)^{16}, r = \frac{2}{3} \end{aligned}$$

$$S_{\infty} = \frac{\left(\frac{2}{3}\right)^{16}}{1 - \frac{2}{3}} = 0.00457 \text{ (3 s.f.)}$$

(vi) Since $P(T > t) = P(X = 0)$

$$\begin{aligned} P(T > t) &= e^{-\lambda t} \\ F(t) &= P(T \leq t) = 1 - e^{-\lambda t} \\ f(t) &= F'(t) = -e^{-\lambda t} \times -\lambda = \lambda e^{-\lambda t} \end{aligned}$$

2 $X \sim B(n, p)$

$$\text{let } \lambda = np, \text{ so } p = \frac{\lambda}{n}$$

$$\begin{aligned} P(X = r) &= \frac{n!}{r!(n-r)!} p^r (1-p)^{n-r} \\ &= \frac{1}{r!} \frac{n!}{(n-r)!} \left(\frac{\lambda}{n}\right)^r \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-r} \\ &\approx \frac{n^r}{r!} \left(\frac{\lambda}{n}\right)^r \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-r} \\ &\quad \left(\text{because } \frac{n!}{(n-r)!} \rightarrow n^r \text{ as } n \rightarrow \infty\right) \\ &= \frac{1}{r!} \lambda^r \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-r} \\ &= \frac{1}{r!} \lambda^r \left(1 - \frac{\lambda}{n}\right)^n \\ &\quad \left(\text{because } (1-p) \approx 1\right) \\ &\approx \frac{1}{r!} \lambda^r e^{-\lambda} \end{aligned}$$

$$\text{Since } e = \lim_{x \rightarrow \infty} \left(1 + \frac{1}{x}\right)^x, \text{ define a number } x = -\frac{n}{\lambda}$$

$$\lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n = \lim_{x \rightarrow \infty} \left(1 + \frac{1}{x}\right)^{x(-\lambda)} = e^{-\lambda}$$

$$\begin{aligned} 3 (i) \text{ Number of pairs of people} &= \binom{n}{2} = \frac{n!}{2!(n-2)!} \\ &= \frac{n(n-1)}{2} \end{aligned}$$

$$N \sim B\left(\frac{n(n-1)}{2}, \frac{3}{365}\right)$$

$$N \approx \text{Po}\left(\frac{3n(n-1)}{730}\right)$$

The approximation is valid when $np < 5$

$$\frac{3n(n-1)}{730} < 5 \Rightarrow 3n(n-1) < 3650$$

$$3n^2 - 3n - 3650 < 0$$

$$-34 \leq n \leq 35$$

$$\text{Since } n > 0, n \leq 35$$

$$(ii) \quad P(N \geq 1) = 1 - P(N = 0)$$

$$= 1 - e^{-\frac{3n(n-1)}{730}}$$

$$1 - e^{-\frac{3n(n-1)}{730}} = 0.5$$

$$e^{-\frac{3n(n-1)}{730}} = 0.5$$

$$-\frac{3n(n-1)}{730} = \ln 0.5$$

$$-\frac{3n(n-1)}{730} = -\ln 2$$

$$3n(n-1) = 730 \ln 2$$

$$3n^2 - 3n - 730 \ln 2 = 0$$

$$n = 13.5 \text{ or } -12.5$$

$$n = 14$$

$$(iii) \quad \text{Number of triples} = \frac{n!}{3!(n-3)!}$$

$$= \frac{n(n-1)(n-2)}{6}$$

$$P(\text{All 3 have the same birthdate}) = \frac{1}{365^2}$$

$$B \sim B\left(\frac{n(n-1)(n-2)}{6}, \frac{1}{365^2}\right)$$

$$E(B) = \frac{n(n-1)(n-2)}{799\,350}$$

$$(iv) \quad B \approx \text{Po}\left(\frac{n(n-1)(n-2)}{799\,350}\right)$$

(approximation valid if $n \leq 159$)

$$n = 100, \text{ so } B \approx \text{Po}(1.21)$$

$$P(B \geq 1) = 1 - P(B = 0) = 1 - e^{-1.21} = 0.703 \text{ (3 s.f.)}$$

4 X : number of doctors who call in sick per weekend.

$$X \sim \text{Po}(1.6)$$

$$P(X = 2) = 0.258 \text{ (3 s.f.)}$$

$$P(X \geq 1) = 0.798 \text{ (3 s.f.)}$$

$$P(X = 2 | X \geq 1) = \frac{P(X = 2)}{P(X \geq 1)}$$

$$= \frac{0.258}{0.798} = 0.324$$

$$(iv) \quad \text{Var}(2X) = 2^2 \times \text{Var}(X)$$

$$= 4 \times 2.21$$

$$= 8.84$$

$$2 \quad (i) \quad E(X - 1) = E(X) - 1 = 5 - 1 = 4$$

$$(ii) \quad \text{Var}(X + 2) = \text{Var}(X) = 2^2 = 4$$

$$(iii) \quad E(2 - X) = 2 - E(X) = 2 - 5 = -3$$

$$(iv) \quad \text{sd}(2 - X) = \text{sd}(X) = 2$$

$$(v) \quad E(2X + 3) = 2 \times E(X) + 3 = 2 \times 5 + 3 = 13$$

$$(vi) \quad \text{Var}(3X - 2) = 3^2 \times \text{Var}(X) = 9 \times 2^2 = 36$$

$$(vii) \quad \text{Var}(X) = E(X^2) - \mu^2 \Rightarrow 4 = E(X^2) - 5$$

$$\Rightarrow E(X^2) = 9$$

$$(viii) \quad E[(X - 2)^2] = E(X^2 - 4X + 4)$$

$$= E(X^2) - 4E(X) + 4$$

$$= 9 - 4 \times 5 + 4$$

$$= -7$$

$$(ix) \quad \text{Var}(4 - 5X) = (-5)^2 \text{Var}(X)$$

$$= 25 \times 2^2$$

$$= 100$$

3 X : price of TV before GST.

$$E(X) = 650 \quad \text{Var}(X) = 40$$

$$E(1.15X + 20) = 1.15 \times 650 + 20$$

$$= \$767.50$$

$$\text{Var}(1.15X + 20) = 1.15^2 \times 40$$

$$= \$52.90$$

$$4 \quad P(X = 0) = \left(\frac{3}{4}\right)^3 = \frac{27}{64}$$

$$P(X = 3) = \left(\frac{1}{4}\right)^3 = \frac{1}{64}$$

$$P(X = 1) = \binom{3}{1} \left(\frac{1}{4}\right)^1 \left(\frac{3}{4}\right)^2 = \frac{27}{64}$$

$$P(X = 2) = \binom{3}{2} \left(\frac{1}{4}\right)^2 \left(\frac{3}{4}\right) = \frac{9}{64}$$

(i)	x	0	1	2	3
	$P(X = x)$	$\frac{27}{64}$	$\frac{27}{64}$	$\frac{9}{64}$	$\frac{1}{64}$

$$(ii) \quad E(X) = 0 \times \frac{27}{64} + 1 \times \frac{27}{64} + 2 \times \frac{9}{64} + 3 \times \frac{1}{64}$$

$$= 0.75$$

(iii)

$$\text{Var}(X) = 0^2 \times \frac{27}{64} + 1^2 \times \frac{27}{64} + 2^2 \times \frac{9}{64} + 3^2 \times \frac{1}{64} - 0.75^2$$

$$= \frac{9}{16} = 0.5625$$

$$(iv) \quad E(2X + 1) = 2 \times 0.75 + 1 = 2.5$$

$$(v) \quad \text{Var}(4X - 3) = 4^2 \times \frac{9}{16} = 9$$

6 Linear combinations of random variables

6.1 Expected value and variance, and algebraic results

$$1 \quad (i) \quad E(X) = 0 \times 0.1 + 1 \times 0.3 + 2 \times 0.2 + 4 \times 0.4$$

$$= 2.3$$

$$(ii) \quad \text{Var}(X) = 0^2 \times 0.1 + 1^2 \times 0.3 + 2^2 \times 0.2$$

$$+ 4^2 \times 0.4 - 2.3^2$$

$$= 7.5 - 2.3^2$$

$$= 2.21$$

$$(iii) \quad E(2X) = 2 \times E(X) = 2 \times 2.3 = 4.6$$

6.2 The sums and differences of independent random variables

1 (i)

$$E(X) = 0 \times 0.2 + 1 \times 0.3 + 2 \times 0.5 = 1.3$$

$$E(Y) = 0 \times 0.4 + 1 \times 0.5 + 3 \times 0.1 = 0.8$$

$$\text{Var}(X) = 0^2 \times 0.2 + 1^2 \times 0.3 + 2^2 \times 0.5 - 1.3^2 = 0.61$$

$$\text{Var}(Y) = 0^2 \times 0.4 + 1^2 \times 0.5 + 3^2 \times 0.1 - 0.8^2 = 0.76$$

(ii)

$x + y$	$^{0.4}0$	$^{0.5}2$	$^{0.1}3$
$^{0.2}0$	$0^{(0.08)}$	$1^{(0.1)}$	$3^{(0.02)}$
$^{0.3}1$	$1^{(0.12)}$	$2^{(0.15)}$	$4^{(0.03)}$
$^{0.5}2$	$2^{(0.2)}$	$3^{(0.25)}$	$5^{(0.05)}$

$x - y$	0	1	3
0	$0^{(0.08)}$	$-1^{(0.1)}$	$-3^{(0.02)}$
1	$1^{(0.12)}$	$0^{(0.15)}$	$-2^{(0.03)}$
2	$2^{(0.2)}$	$1^{(0.25)}$	$-1^{(0.05)}$

$x + y$	0	1	2	3	4	5
Prob	0.08	0.22	0.35	0.27	0.03	0.05

$x - y$	-3	-2	-1	0	1	2
Prob	0.02	0.03	0.15	0.23	0.37	0.2

$$\begin{aligned} \text{(iii)} \quad E(X + Y) &= 0 \times 0.08 + 1 \times 0.22 + 2 \times 0.35 \\ &\quad + 3 \times 0.27 + 4 \times 0.03 + 5 \times 0.05 \\ &= 2.1 \end{aligned}$$

$$\begin{aligned} E(X - Y) &= -3 \times 0.02 + (-2) \times 0.03 + (-1) \times 0.15 \\ &\quad + 0 \times 0.23 + 1 \times 0.37 + 2 \times 0.2 \\ &= 0.5 \end{aligned}$$

$$\begin{aligned} \text{Var}(X + Y) &= 0^2 \times 0.08 + 1^2 \times 0.22 + 2^2 \times 0.35 \\ &\quad + 3^2 \times 0.27 + 4^2 \times 0.03 + 5^2 \times 0.05 - 2.1^2 \\ &= 1.37 \end{aligned}$$

$$\begin{aligned} \text{Var}(X - Y) &= (-3)^2 \times 0.02 + (-2)^2 \times 0.03 + (-1)^2 \times 0.15 \\ &\quad + 0^2 \times 0.23 + 1^2 \times 0.37 + 2^2 \times 0.2 - 0.5^2 \\ &= 1.37 \end{aligned}$$

$$\begin{aligned} \text{(iv)} \quad E(X + Y) &= E(X) + E(Y) \\ E(X - Y) &= E(X) - E(Y) \\ \text{Var}(X + Y) &= \text{Var}(X) + \text{Var}(Y) \\ \text{Var}(X - Y) &= \text{Var}(X) + \text{Var}(Y) \end{aligned}$$

$$\begin{aligned} \text{2 (i)} \quad E(P + Q) &= E(P) + E(Q) \\ &= 8 + 3 \\ &= 11 \end{aligned}$$

$$\begin{aligned} \text{(ii)} \quad \text{Var}(P + Q) &= \text{Var}(P) + \text{Var}(Q) \\ &= 5 + 10 \\ &= 15 \end{aligned}$$

$$\begin{aligned} \text{(iii)} \quad E(2P + 5Q) &= 2 \times 8 + 5 \times 3 \\ &= 31 \end{aligned}$$

$$\text{(iv)} \quad \text{Var}(P - Q) = 5 + 10 = 15$$

$$\begin{aligned} \text{(v)} \quad \text{Var}(3P + 2Q) &= 3^2 \times 5 + 2^2 \times 10 \\ &= 85 \end{aligned}$$

$$\begin{aligned} \text{(vi)} \quad \text{Var}(P - 3Q) &= \text{Var}(P) + (-3)^2 \times \text{Var}(Q) \\ &= 5 + 9 \times 10 \\ &= 95 \end{aligned}$$

$$\begin{aligned} \text{3 (i)} \quad E(X_1 + X_2) &= 250 + 250 = 500\text{g} \\ \text{Var}(X_1 + X_2) &= 5^2 + 5^2 = 50\text{g}^2 \\ \Rightarrow \text{sd} &= \sqrt{50} = 7.07\text{g} \end{aligned}$$

$$\begin{aligned} \text{(ii)} \quad E(Y_1 + Y_2 + Y_3 + Y_4) &= 125 \times 4 = 500\text{g} \\ \text{Var}(Y_1 + Y_2 + Y_3 + Y_4) &= 4 \times 5^2 = 100\text{g}^2 \\ \text{sd} &= \sqrt{100} = 10\text{g} \end{aligned}$$

4 (i)

x	1	2	3	4
$P(Y=x)$	0.25	0.25	0.25	0.25

$$\begin{aligned} E(X) &= 1 \times 0.25 + 2 \times 0.25 + 3 \times 0.25 + 4 \times 0.25 \\ &= 2.5 \end{aligned}$$

$$\begin{aligned} \text{Var}(X) &= 1^2 \times 0.25 + 2^2 \times 0.25 + 3^2 \times 0.25 \\ &\quad + 4^2 \times 0.25 - 2.5^2 \\ &= 7.5 - 2.5^2 \\ &= 1.25 \end{aligned}$$

(ii)

+	1	2	3	4
1	2	3	4	5
2	3	4	5	6
3	4	5	6	7
4	5	6	7	8

$x_1 + x_2$	2	3	4	5	6	7	8
Prob.	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	$\frac{1}{16}$

$$\begin{aligned} E(X_1 + X_2) &= 2 \times \frac{1}{16} + 3 \times \frac{2}{16} + 4 \times \frac{3}{16} + 5 \times \frac{4}{16} \\ &\quad + 6 \times \frac{3}{16} + 7 \times \frac{2}{16} + 8 \times \frac{1}{16} \\ &= 5 \end{aligned}$$

$$\begin{aligned} \text{Var}(X_1 + X_2) &= 2^2 \times \frac{1}{16} + 3^2 \times \frac{2}{16} + 4^2 \times \frac{3}{16} + 5^2 \times \frac{4}{16} \\ &\quad + 6^2 \times \frac{3}{16} + 7^2 \times \frac{2}{16} + 8^2 \times \frac{1}{16} - 5^2 \\ &= 2.5 \end{aligned}$$

(iii)	$ x_1 - x_2 $	1	2	3	4
	1	0	1	2	3
	2	1	0	1	2
	3	2	1	0	1
	4	3	2	1	0
	$ x_1 - x_2 $	0	1	2	3
	Prob.	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{2}{16}$

$$E(|X_1 - X_2|) = 0 \times \frac{4}{16} + 1 \times \frac{6}{16} + 2 \times \frac{4}{16} + 3 \times \frac{2}{16}$$

$$= \frac{20}{16} = 1.25$$

$$\text{Var}(|X_1 - X_2|) = 0^2 \times \frac{4}{16} + 1^2 \times \frac{6}{16}$$

$$+ 2^2 \times \frac{4}{16} + 3^2 \times \frac{2}{16} - 1.25^2$$

$$= \frac{15}{16} = 0.9375 = 0.938 \text{ (3 s.f.)}$$

(iv)	$2x_1$	2	4	6	8
	Prob.	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

$$E(2X_1) = 2 \times \frac{1}{4} + 4 \times \frac{1}{4} + 6 \times \frac{1}{4} + 8 \times \frac{1}{4}$$

$$= 5$$

$$\text{Var}(2X_1) = 2^2 \times \frac{1}{4} + 4^2 \times \frac{1}{4}$$

$$+ 6^2 \times \frac{1}{4} + 8^2 \times \frac{1}{4} - (5^2)$$

$$= 5$$

- (v) (a) T (b) F (c) T
 (d) T (e) F

- 5 (i) $T = W + B$

$$E(T) = 12 + 25 = 37 \text{ mins}$$

$$\text{Var}(T) = 1.2^2 + 1.6^2 = 4$$

$$T \sim N(37, 4)$$

- (ii) $P(T > 40) = P\left(Z > \frac{40 - 37}{2}\right)$
 $= P(Z > 1.5)$
 $= 1 - 0.9332$
 $= 0.0668 \text{ (3 s.f.)}$

- (iii) $Z(0.98) = 2.054$

$$2.054 = \frac{x - 37}{2}$$

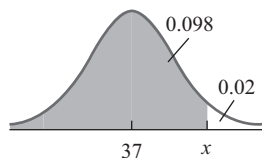
$$x = 41.108 \text{ min}$$

98% of all trips take 41.108 min or less.

Time she should leave is 8:30 a.m. - 41.108 min.

= 8:30 a.m. - 42 min

= 7:48 a.m.



- 6 X: Wage from houses sold per month.

Y: Wage from units sold per month.

S: Monthly salary.

$$E(X) = 0.4\% \text{ of } \$524\,000 = \$2096$$

$$E(Y) = 0.2\% \text{ of } \$285\,000 = \$570$$

$$\text{Var}(X) = (0.004)^2 \times 9200^2 = 1354.24$$

$$\text{Var}(Y) = (0.002)^2 \times 4800^2 = 92.16$$

$$S = 0.8X + 0.6Y + 500$$

$$E(S) = 0.8 \times 2096 + 0.6 \times 570 + 500$$

$$= \$2518.80$$

$$\text{Var}(S) = 0.8^2 \times 1354.24 + 0.6^2 \times 92.16$$

$$= 899.89$$

$$\text{sd}(S) = \sqrt{899.89}$$

$$= \$30$$

- 7 (i) $F - S \sim N(7, 19.6^2)$

$$P(|F - S| > 6)$$

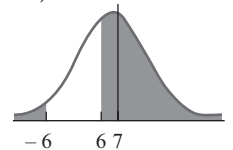
$$= P(F - S > 6) + P(F - S < -6)$$

$$= P\left(Z > \frac{6 - 7}{19.6}\right) + P\left(Z < \frac{-6 - 7}{19.6}\right)$$

$$= P(Z > -0.051) + P(Z < -0.663)$$

$$= 0.5203 + (1 - 0.7464)$$

$$= 0.774 \text{ (3 s.f.)}$$



- (ii) The standard deviation would be less.

Pairing fathers and sons would mean less variation.

- 8 $W \sim N(750, 2.6^2)$

- (i) $W_1 + W_2 \sim N(1500, 2 \times 2.6^2)$

$$P(W_1 + W_2 < 1495 \text{ g})$$

$$= P\left(Z < \frac{1495 - 1500}{\sqrt{13.52}}\right)$$

$$= P(Z < -1.360)$$

$$= 1 - 0.9131$$

$$= 0.0869 \text{ (3 s.f.)}$$

- (ii) $W_1 - W_2 \sim N(0, 13.52)$

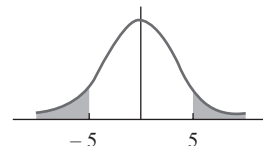
$$= P(W_1 - W_2 > 5) + (W_2 - W_1 > 5)$$

$$= 2 \times P(W_1 - W_2 > 5)$$

$$= 2 \times P\left(Z > \frac{5 - 0}{\sqrt{13.52}}\right)$$

$$= 2 \times (1 - 0.9131)$$

$$= 0.174 \text{ (3 s.f.)}$$



$$9 \quad H \sim N(12.5, 0.04^2), B \sim N(12.4, 0.01^2) \\ B - H \sim N(-0.1, 0.0017)$$

$$(i) \quad P(B - H < 0)$$

$$= P\left(Z < \frac{0 - (-0.1)}{\sqrt{0.0017}}\right) \\ = P(Z < 2.425) \\ = 0.992 \text{ (3 s.f.)}$$

$$(ii) \quad P(B - H < 0) \geq 0.999$$

$$B - H \sim N(12.4 - h, 0.0017)$$

$$3.090 = \frac{0 - (12.4 - h)}{\sqrt{0.0017}} \\ h = 12.53 \text{ cm}$$

$$10 \quad N \sim N(45, 8^2), L \sim N(55, 10^2)$$

$$E(C) = 4 \times 45 + 3 \times 55 \\ = 345$$

$$\text{Var}(C) = 4^2 \times 8^2 + 3^2 \times 10^2 \\ = 1924$$

$$C \sim N(345, 1924)$$

$$P(C \geq 390) = P\left(Z \geq \frac{390 - 345}{\sqrt{1924}}\right) \\ = P(Z \geq 1.026) \\ = 1 - 0.8475 \\ = 0.1524 = 0.153 \text{ (3 s.f.)}$$

i.e. approx. 15.3% of applicants will get an interview.

6.3 More than two independent random variables

$$1 \quad (i) \quad X + Y + Z \sim N(35, 14)$$

$$(ii) \quad X - Y - Z \sim N(-11, 14)$$

$$(iii) \quad E(2X + 3Y + Z) = 2 \times 12 + 3 \times 8 + 15 = 63$$

$$\text{Var}(2X + 3Y + Z) = 2^2 \times 4 + 3^2 \times 1 + 9 = 34$$

$$2X + 3Y + Z \sim N(63, 34)$$

$$(iv) \quad E(X - 2Y + 5Z) = 12 - 2 \times 8 + 5 \times 15 = 71$$

$$\text{Var}(X - 2Y + 5Z) = 4 + 2^2 \times 1 + 5^2 \times 9 = 233$$

$$X - 2Y + 5Z \sim N(71, 233)$$

$$2 \quad T = A + B + C + D$$

$$E(T) = 62 + 70.4 + 57.6 + 52.8 \\ = 242.8 \text{ s} = 4 \text{ min } 2.8 \text{ s}$$

$$\text{Var}(T) = 0.21^2 + 0.12^2 + 0.18^2 + 0.35^2 \\ = 0.2134$$

$$T \sim N(242.8, 0.2134)$$

$$P(T < 242) = P\left(Z < \frac{242 - 242.8}{\sqrt{0.2134}}\right) \\ = P(Z < -1.732) \\ = 1 - 0.9584 \\ = 0.0416 \text{ (3 s.f.)}$$

$$3 \quad (i) \quad W: \text{weight of truck} + 10 \text{ containers.}$$

$$E(W) = 2450 + 10 \times 42 \\ = 2870 \text{ kg}$$

$$\text{Var}(W) = \text{Var}(T + C_1 + C_2 + \dots + C_{10}) \\ = 85^2 + 10 \times 3^2 \\ = 7315 \text{ kg}^2$$

$$W \sim N(2870, 7315)$$

$$(ii) \quad P(W > 3000) = P\left(Z > \frac{3000 - 2870}{\sqrt{7315}}\right) \\ = P(Z > 1.520) \\ = 1 - 0.9357 \\ = 0.0643 \\ = 6.43\% \text{ (3 s.f.)}$$

$$(iii) \quad W \sim N(2450 + 10x, 7315)$$

x : average weight of one container.

$$Z(0.995) = 2.576$$

$$2.576 = \frac{3000 - (2450 + 10x)}{\sqrt{7315}}$$

$$220.32 = 550 - 10x$$

$$10x = 329.68$$

$$x = 33.0 \text{ kg (3 s.f.)}$$

$$4 \quad (i) \quad E(T) = 2 \times 0.1 + 3 \times 0.05 + 0.15 = 0.5 \text{ mm}$$

$$\text{Var}(T) = \text{Var}(A_1 + A_2 + B_1 + B_2 + B_3 + C_1) \\ = 2 \times 0.02^2 + 3 \times 0.01^2 + 0.03^2 \\ = 0.002 \text{ mm}^2$$

$$T \sim N(0.5, 0.002)$$

$$(ii) \quad P(T > 0.6) = P\left(Z > \frac{0.6 - 0.5}{\sqrt{0.002}}\right) \\ = P(Z > 2.236) \\ = 1 - 0.9873 \\ = 0.0127 \\ = 1.27\% \text{ (3 s.f.)}$$

$$5 \quad C: \text{amount of coffee dispensed in one serve.}$$

$$C \sim N(210, 3^2)$$

$$C_1 + C_2 + C_3 + C_4 \sim N(840, 6^2)$$

$$P(C_1 + C_2 + C_3 + C_4 > 850) = P\left(Z > \frac{850 - 840}{6}\right) \\ = P(Z > 1.667) \\ = 1 - 0.9522 \\ = 0.0478 \text{ (3 s.f.)}$$

$$6 \quad T = W_1 + W_2 + W_3 + W_4 + W_5 + E_1 + E_2$$

$$(i) \quad E(T) = 5 \times 450 + 2 \times 1080 = 4410 \text{ MB}$$

$$\text{Var}(T) = 5 \times 70^2 + 2 \times 45^2 = 28\,550 \text{ MB}^2$$

$$T \sim N(4410, 28\,550)$$

$$(ii) \quad P(E_1 + E_2 > W_1 + W_2 + W_3 + W_4 + W_5)$$

$$= P(E_1 + E_2 - (W_1 + \dots + W_5) > 0)$$

$$E(E_1 + E_2 - (W_1 + \dots + W_5)) = 2 \times 1080 - 5 \times 450 \\ = -90 \text{ MB}$$

$$\text{Var}(E_1 + E_2 - (W_1 + \dots + W_5)) = 2 \times 45^2 + 5 \times 70^2 \\ = 28\,550 \text{ MB}^2$$

$$P(E_1 + E_2 - (W_1 + \dots + W_5) > 0)$$

$$= P\left(Z > \frac{0 - (-90)}{\sqrt{28\,550}}\right)$$

$$= P(Z > 0.533)$$

$$= 1 - 0.7029$$

$$= 0.297 \text{ (3 s.f.)}$$

Further practice

$$1 \quad P \sim N(14, 1.8^2), W \sim N(7.8, 0.6^2)$$

$$F = 0.8P + 0.2W$$

$$E(F) = 0.8 \times 14 + 0.2 \times 7.8$$

$$= 12.8 \text{ (3 s.f.)}$$

$$\text{Var}(F) = 0.8^2 \times 1.8^2 + 0.2^2 \times 0.6^2$$

$$= 2.088$$

$$\text{SD}(F) = \sqrt{2.088}$$

$$= 1.44 \text{ (3 s.f.)}$$

$$2 \quad CD \sim N(15, 0.2^2)$$

$$\text{Case} \sim N(80, 1.4^2)$$

$$T = CD_1 + CD_2 + \dots + CD_{12} + C_1 + \dots + C_{12} + B$$

$$E(T) = 12 \times 15 + 12 \times 80 + 320$$

$$= 1460 \text{ g}$$

$$\text{Var}(T) = 12 \times 0.2^2 + 12 \times 1.4^2$$

$$= 24 \text{ g}^2$$

$$P(T > 1470) = P\left(Z > \frac{1470 - 1460}{\sqrt{24}}\right)$$

$$= P(Z > 2.041)$$

$$= 1 - 0.9793$$

$$= 0.0207 \text{ (3 s.f.)}$$

$$3 \quad (i) \quad P: \text{weight of passenger } P \sim N(76, 9^2)$$

$$A: \text{weight of plane empty } A \sim N(36\,400, 320^2)$$

$$T: \text{total weight of plane + 80 passengers}$$

$$T = A + P_1 + P_2 + \dots + P_{80}$$

$$E(T) = 36\,400 + 80 \times 76$$

$$= 42\,480 \text{ kg}$$

$$\text{Var}(T) = 320^2 + 80 \times 9^2$$

$$= 108\,880 \text{ kg}^2$$

$$P(42\,000 < T < 42\,600)$$

$$= P\left(\frac{42\,000 - 42\,480}{\sqrt{108\,880}} < Z < \frac{42\,600 - 42\,480}{\sqrt{108\,880}}\right)$$

$$= P(-1.455 < Z < 1.576)$$

$$= (0.9272 - 0.5) + (0.9425 - 0.5)$$

$$= 0.870 \text{ (3 s.f.)}$$

$$(ii) \quad P(500P \geq A) = P(500P - A \geq 0)$$

$$E(500P - A) = 500 \times 76 - 36\,400$$

$$= 1600 \text{ kg}$$

$$\text{Var}(500P - A) = 500^2 \times 9^2 + 320^2$$

$$= 20\,352\,400 \text{ kg}^2$$

$$P(500P - A \geq 0) = P\left(Z \geq \frac{0 - 1600}{\sqrt{20\,352\,400}}\right)$$

$$= P(Z \geq -0.355)$$

$$= 0.639 \text{ (3 s.f.)}$$

$$4 \quad (i) \quad \text{Var}(S) = \text{Var}(X_1 + X_2 + X_3 + X_4)$$

$$= 4 \times 1.4$$

$$= 5.6$$

$$(ii) \quad \text{Var}(T) = \text{Var}(4X + 3)$$

$$= 4^2 \times \text{Var}(X)$$

$$= 4^2 \times 1.4$$

$$= 22.4$$

$$(iii) \quad E(S - T) = E(X_1 + X_2 + X_3 + X_4 - (4X + 3))$$

$$= 3.5 + 3.5 + 3.5 + 3.5 - (4 \times 3.5 + 3)$$

$$= -3$$

$$\text{Var}(S - T) = \text{Var}(X_1 + X_2 + X_3 + X_4) + \text{Var}(4X + 3)$$

$$= 4 \times 1.4 + 4^2 \times 1.4$$

$$= 28$$

Past exam questions

$$1 \quad E(T) = 0.13 + 0.14 + 0.10 = 0.37 \text{ mm}$$

$$\text{Var}(T) = 0.02^2 + 0.03^2 + 0.01^2 = 0.0014 \text{ mm}^2$$

$$T \sim N(0.37, 0.0014)$$

$$P(T < 0.30) = P\left(Z < \frac{0.3 - 0.37}{\sqrt{0.0014}}\right)$$

$$= P(Z < -1.871)$$

$$= 0.0306 \text{ (3 s.f.)}$$

2 $E(3X - Y) = 3 \times 6.5 - 7.4 = 12.1$

$$\text{Var}(3X - Y) = 3^2 \times 14 + 15 = 141$$

$$3X - Y \sim N(12.1, 141)$$

$$\begin{aligned} P(3X - Y < 20) &= P\left(Z < \frac{20 - 12.1}{\sqrt{141}}\right) \\ &= P(Z < 0.665) \\ &= 0.747 \text{ (3 s.f.)} \end{aligned}$$

- 3 (i) M : number of minutes bike is hired for.

C : cost of hire.

$$C = 500 + 3M$$

$$E(C) = 500 + 3 \times 142 = 926 \text{ c or } \$9.26$$

$$\text{Var}(C) = 3^2 \times 35^2 = 11\,025$$

$$\text{sd}(C) = \sqrt{11\,025} = 105 \text{ c or } \$1.05$$

- (ii) T : total amount paid by 6 people.

$$E(T) = 6 \times 926 = 5556 \text{ c or } \$55.56$$

$$\text{Var}(T) = 6 \times 11\,025 = 66\,150 \text{ c}^2$$

$$\text{sd}(T) = \sqrt{66\,150} = 257 \text{ c or } \$2.57$$

- 4 S : weight of a carton of sugar.

$$S \sim N(78.8, 12.6^2)$$

F : weight of a carton of flour.

$$F \sim N(62.0, 10.0^2)$$

- (i) $S_1 + S_2 + \dots + S_8 + F_1 + F_2 + F_3 \sim N(816.4, 1570.08)$

$$P(S_1 + S_2 + \dots + S_8 + F_1 + F_2 + F_3 > 900)$$

$$= P\left(Z > \frac{900 - 816.4}{\sqrt{1570.08}}\right)$$

$$= P(Z > 2.110)$$

$$= 1 - 0.9826$$

$$= 0.0174 \text{ (3 s.f.)}$$

- (ii) $P(S < F) = P(S - F < 0)$

$$S - F \sim N(16.8, 258.76)$$

$$P(S - F < 0) = P\left(Z < \frac{0 - 16.8}{\sqrt{258.76}}\right)$$

$$= P(Z < -1.044)$$

$$= 1 - 0.8517$$

$$= 0.148 \text{ (3 s.f.)}$$

- 5 I : mass of iron ore produced per day.

$$I \sim N(7.0, 0.46^2)$$

$$I_1 + I_2 + \dots + I_{10} \sim N(70, 2.116)$$

$$P(I_1 + I_2 + \dots + I_{10} > 71) = P\left(Z > \frac{71 - 70}{\sqrt{2.116}}\right)$$

$$= P(Z > 0.687)$$

$$= 1 - 0.754$$

$$= 0.246 \text{ (3 s.f.)}$$

Stretch and challenge

1 $M \sim N(\mu, \sigma^{-2})$

$$F \sim N(\mu - 15, \sigma^{-2})$$

$$P(M_1 + M_2 > F_1 + F_2 + F_3) = 0.25$$

$$P(M_1 + M_2 - F_1 + F_2 + F_3 > 0) = 0.25$$

$$E(M_1 + M_2 - F_1 - F_2 - F_3)$$

$$= \mu + \mu - (\mu - 15) - (\mu - 15) - (\mu - 15)$$

$$= 45 - \mu$$

$$\text{Var}(M_1 + M_2 - F_1 - F_2 - F_3) = 5\sigma^2$$

$$\therefore P\left(Z > \frac{0 - (45 - \mu)}{\sqrt{5}\sigma}\right) = 0.25$$

$$\frac{\mu - 45}{\sqrt{5}\sigma} = 0.674$$

$$\mu - 45 = 1.5071\sigma$$

$$8\sigma^2 - 45 = 1.5071\sigma$$

$$8\sigma^2 - 1.5071\sigma - 45 = 0$$

$$\sigma = 2.47 \text{ (3 s.f.)}$$

$$\mu = 48.7 \text{ (3 s.f.)}$$

- 2 (i) T : total weight

$$T = M_1 + M_2 + C_1 + C_2 + C_3$$

$$T \sim N(750, 500)$$

$$P(T > 800) = P\left(Z > \frac{800 - 750}{\sqrt{500}}\right)$$

$$= P(Z > 2.236)$$

$$= 1 - 0.9873$$

$$= 0.0127 \text{ (3 s.f.)}$$

- (ii) M : amount (in kg) of macadamias in 2 scoops.

$$M \sim N(0.3, 0.0002)$$

C : amount (in kg) of cashews in 3 scoops.

$$C \sim N(0.45, 0.0003)$$

$$\text{Cost of macadamias} = \$12M$$

$$\text{Cost of cashews} = \$9C$$

$$P(12M < 9C) = P(12M - 9C < 0)$$

$$E(12M - 9C) = 12 \times 0.3 - 9 \times 0.45$$

$$= -0.45$$

$$\text{Var}(12M - 9C) = 12^2 \times 0.0002 + 9^2 \times 0.0003$$

$$= 0.0531$$

$$P(12M - 9C < 0) = P\left(Z < \frac{0 - (-0.45)}{\sqrt{0.0531}}\right)$$

$$= P(Z < 1.953)$$

$$= 0.975 \text{ (3 s.f.)}$$