Revision

18

Revision of Chapters 13–17

18A Technology-free questions

1 A factory has two machines, Machine I and Machine II. Machine I produces 800 items each day, and on average 20 are faulty. Machine II produces 700 items each day, and on average 14 are faulty.

a What is the probability that an item selected at random from the days production is faulty?

b If the item selected is faulty, what is the probability that it was produced by Machine II?

2 A tomato farmer knows that $\frac{1}{5}$ of the tomatoes seedlings he plants will not survive. He plants four tomato seedlings in each of his planter boxes.

a What is the probability that none of tomato plants in a box will survive?

b What is the probability that at least one of tomato plants in a box will survive?

c A customer buys six boxes, each with four plants. What is the probability that all the plants survive? Express your answers in the form $\left(\frac{a}{b}\right)^c$, where a, b and c are positive integers.

3 The function

$$f(x) = \begin{cases} k\cos(\pi x) & \text{if } \frac{3}{2} < x < \frac{5}{2} \\ 0 & \text{otherwise} \end{cases}$$

is a probability density function for the continuous random variable X.

a Find the value of k.

b Find the median of X.

c Find $\Pr\left(X < \frac{7}{4} \,\middle|\, X < 2\right)$.

d Find $\Pr\left(X > \frac{9}{4} \,\middle|\, X > \frac{7}{4}\right)$.

ISBN 978-1-009-11049-5

Cambridge University Press

4 The random variable X has the following probability distribution.

x	0	1	2	3	4
$\Pr(X = x)$	0.3	0.2	0.1	0.3	0.1

Find:

a $\Pr(X > 3 \mid X > 1)$

b $\Pr(X > 1 \mid X \leq 3)$

c the mean of X

d the variance of X

5 The continuous random variable X has probability density function given by

$$f(x) = \begin{cases} kx(6 - x) & \text{if } 0 < x < 6 \\ 0 & \text{otherwise} \end{cases}$$

Find:

a the value of k

b $\Pr(X < 4)$

c the median of X

d the mean of X

e $\Pr(X < 2 \mid X < 3)$

f $\Pr(X > 2 \mid X < 4)$

6 Eight coloured balls are placed in a box: 3 red balls, 2 black balls, 2 green balls and a yellow ball. A ball is randomly withdrawn from the box and not returned, and then a second ball is randomly withdrawn.

a What is the probability of withdrawing a red ball first and a green ball second?

b What is the probability of obtaining one green and one red ball?

c What is the probability that the second ball withdrawn is not red, given that the first ball withdrawn is red?

d What is the probability that neither of the two balls withdrawn is red?

e What is the probability of obtaining two balls of the same colour?

7 Two events A and B are such that $\Pr(A) = \frac{4}{7}$ and $\Pr(B) = \frac{1}{3}$. Find $\Pr(A' \cap B)$ if:

a $\Pr(A \cup B) = \frac{5}{7}$

b A and B are mutually exclusive.

8 Two events A and B are such that $\Pr(A) = \frac{3}{4}$, $\Pr(B \mid A) = \frac{1}{5}$ and $\Pr(B' \mid A') = p$.

a Find $\Pr(B)$ in terms of p.

b Find $\Pr(A \mid B)$ in terms of p.

9 The random variable X has probability distribution:

x	1	2	3	4	5
$\Pr(X = x)$	a	0.3	0.1	0.2	b

a Given that $E(X) = 2.34$, find the values of a and b.

b Find $\text{Var}(X)$.

10 The probability that a certain football club has all its first-team players fit is 0.7. When the club has a fully fit team, the probability of it winning a home game is 0.9. When the team is not fully fit, the probability of winning a home game is 0.4.

a Find the probability that the team will win its next home game.

b Given that the team did not win its last home game, find the probability that the team was fully fit.

ISBN 978-1-009-11049-5 Cambridge University Press

Revision

11 The random variable X has probability density function:

$$f(x) = \begin{cases} (x-a)(2a-x) & \text{if } a \le x \le 2a \\ 0 & \text{otherwise} \end{cases}$$

a Show that $a^3 = 6$. **b** Find $E(X)$.

12 A machine has a probability of $\frac{1}{8}$ of manufacturing a defective part. The parts are packed in boxes of 16.

a What is the expected number of defective parts in a box?

b What is the probability that the number of defective parts in the box is less that the expected number? Express you answer in the form $\frac{ab^{15}}{c^{16}}$ where a, b and $c \in Z^+$.

13 A continuous random variable X has probability density function

$$f(x) = \begin{cases} 0.004x - 0.04 & \text{if } 10 \le x < 20 \\ -0.001x + 0.06 & \text{if } 20 \le x \le 60 \\ 0 & \text{otherwise} \end{cases}$$

Find $\Pr(15 \le X \le 30)$.

14 An experiment consists of five independent trials. Each trial results in either a success or a failure. The probability of success in a trial is p.

a If the probability of at least one success is 0.99968, what is the value of p?

b Write down an expression for $\Pr(X = 3)$ in terms of p, and find the value of p which maximises the probability of exactly three successes.

15 The random variable X is normally distributed with mean 40 and standard deviation 2. If $\Pr(36 < X < 44) = q$, find $\Pr(X > 44)$ in terms of q.

16 The probability density function of a random variable X is given by

$$f(x) = \begin{cases} 2(1-x) & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the value a of X such that $\Pr(X < a) = \frac{3}{4}$.

17 A biased coin is tossed three times. On each toss, the probability of a head is p.

a Find, in terms of p, the probability that all three tosses show tails.

b If the probability of three tails is equal to 8 times the probability of three heads, find p.

18 Let X be a continuous random variable with probability density function

$$f(x) = \begin{cases} 2x\cos x^2 & \text{if } 0 < x < \sqrt{\frac{\pi}{2}} \\ 0 & \text{elsewhere} \end{cases}$$

ISBN 978-1-009-11049-5 Cambridge University Press

a Show by differentiation that $\sin x^2$ is an antiderivative of $2x \cos x^2$.

b Calculate $\Pr\left(\sqrt{\frac{\pi}{3}} < X < \sqrt{\frac{\pi}{2}}\right)$.

c Find the value of m such that $\Pr(X \le m) = \frac{1}{2}$.

19 Consider a bag containing three blue and seven red balls.

a What is p, the proportion of blue balls in the bag?

b Samples of size 3 are taken from the bag without replacement. If $\hat{P}$ is a random variable describing the possible values of the sample proportion $\hat{p}$ of blue balls in the sample, list the possible values that $\hat{P}$ can take.

c Find $\Pr(\hat{P} = 0)$.

20 In a large population the proportion of left handed people is $\frac{1}{5}$. Let $\hat{P}$ be the random variable that represents that the sample proportion of left handed people in a sample of size n. Find the smallest integer value of n such that the standard deviation of $\hat{P}$ is less than or equal to $\frac{1}{20}$.

21 There are n identical black balls and n identical white balls. A blue box contains 3 black balls and $n - 3$ white balls. A red box contains $n - 3$ black balls and 3 white balls. A ball is taken at random from the red box and put in the blue box. A ball is then taken at random from the blue box.

a Find the probability, in terms of n, that the ball taken from the blue box is:

i black **ii** white.

b Find the probability, in terms of n, that the first ball is black given that the second is white.

22 Let X be a random variable with mean μ and variance σ^2. Show that, if $Z = \frac{X - \mu}{\sigma}$, then $E(Z) = 0$ and $\text{Var}(Z) = 1$.

18B Multiple-choice questions

1 A box contains 12 red balls and 4 green balls. A ball is selected at random from the box and not replaced, and then a second ball is drawn. The probability that the two balls are both the same colour is equal to

A $\frac{11}{20}$ **B** $\frac{1}{20}$ **C** $\frac{37}{64}$ **D** $\frac{1}{5}$ **E** $\frac{3}{5}$

2 Two events A and B are independent, where $\Pr(A) = \frac{\Pr(B)}{3}$, and $\Pr(A \cup B) = 0.5325$. $\Pr(A)$ is equal to

A 0.0675 **B** 0.15 **C** 0.45 **D** 0.60 **E** 0.4213

ISBN 978-1-009-11049-5 Cambridge University Press

Revision

3 A test consists of six true/false questions. The probability that a student who guesses will obtain six correct answers is

A 0.9844 **B** 0.0278 **C** 0.5 **D** 0.0156 **E** 0.17

4 A random variable X has the following probability distribution.

x	1	2	3	4
$\Pr(X = x)$	$4c^2$	$5c^2$	$4c^2$	$3c^2$

$\Pr(X < \mu)$, where μ is the mean of X, is

A 0.5 **B** 0.5625 **C** 0.375 **D** 0.25 **E** 0.3125

5 Suppose that a spinner numbered 1, 2, 3, 4, 5, 6 is spun until a ‘3’ appears, and the number of spins is noted. The sample space for this random experiment is

A $\{1, 2, 3, 4, 5, 6\}$ **B** $\{0, 1, 2, 3, 4, 5, 6\}$ **C** $\{1, 2, 3, 4, \dots\}$
D $\{3\}$ **E** $\{1, 2, 3\}$

Questions 6–9 refer to the following probability distribution.

x	4	6	7	9
$\Pr(X = x)$	0.3	0.2	0.1	0.4

6 For this probability distribution, the mean, $E(X)$, is equal to

A 6.7 **B** 0.275 **C** 6.5 **D** 2.75 **E** 2.59

7 For this probability distribution, the variance, $\mathrm{Var}(X)$, is equal to

A 19.45 **B** 4.41 **C** 6.7 **D** 2.1 **E** 0.61

8 Let $Y = 2X - 1$, where X has the probability distribution given in the table. The probability distribution of Y is

A

y	4	6	7	9
$\Pr(Y = y)$	0.3	0.2	0.1	0.4

B

y	8	12	14	18
$\Pr(Y = y)$	0.3	0.2	0.1	0.4

C

y	4	6	7	9
$\Pr(Y = y)$	0.6	0.4	0.2	0.8

D

y	6	10	12	16
$\Pr(Y = y)$	0.3	0.2	0.1	0.4

E

y	7	11	13	17
$\Pr(Y = y)$	0.3	0.2	0.1	0.4

ISBN 978-1-009-11049-5 Cambridge University Press

9 Let $Z = 4 - X$, where X has the probability distribution given in the table. The variance of Z is

A 11.59 **B** 2.1 **C** 0.41 **D** 4.41 **E** −0.41

10 Suppose that in Melbourne the probability of the temperature exceeding 30°C on a particular day is 0.6 if the temperature exceeded 30°C on the previous day, and 0.25 if it did not. If the temperature exceeds 30°C on Monday, then the probability that it exceeds 30°C on Wednesday is

A 0.36 **B** 0.10 **C** 0.60 **D** 0.30 **E** 0.46

11 If a random variable X is such that $\mathrm{E}(X) = 11$ and $\mathrm{E}(X^2) = 202$, then the standard deviation of X is equal to

A 191 **B** 13.82 **C** 9 **D** 3.72 **E** 81

12 A set of test scores has a probability distribution with mean $\mu = 50$ and standard deviation $\sigma = 10$. Which of the following intervals contains about 95% of the test scores?

A $(40, 60)$ **B** $(30, 70)$ **C** $(20, 80)$ **D** $(46.84, 53.16)$ **E** $(43.68, 56.32)$

13 If three fair coins are tossed, what is the probability that there are at least two heads?

A $\frac{1}{3}$ **B** $\frac{6}{7}$ **C** $\frac{1}{4}$ **D** $\frac{1}{2}$ **E** $\frac{1}{8}$

14 Let X be a binomial random variable with parameters $n = 400$ and $p = 0.1$. Then $\mathrm{E}(X)$, the mean of X, is equal to

A 36 **B** 6 **C** 40 **D** 6.32 **E** 360

15 Which of the following does *not* define a binomial random variable?

A A die is rolled 10 times, and the number of sixes observed.
B A die is rolled until a six is obtained, and the number of rolls counted.
C A die is rolled five times, and the number of even numbers showing observed.
D A sample of 20 people is chosen from a large population, and the number of females counted.
E A student guesses the answer to every question on a multiple-choice test, and the number of correct answers is noted.

16 Let X be a binomial random variable with parameters $n = 900$ and $p = 0.2$. The standard deviation of X is equal to

A 18 **B** 144 **C** 180 **D** 13.42 **E** 12

17 Let X be a binomial random variable with a variance of 9.4248. If $n = 42$, then the probability of success p is equal to

A 0.45 **B** 0.22 **C** 0.34 **D** 0.68 **E** 0.34 or 0.66

ISBN 978-1-009-11049-5 Cambridge University Press

18 If p is the probability of success in one trial, then $\binom{7}{5} p^5(1-p)^2$ is the probability of

A exactly two failures **B** exactly two successes **C** at least two failures
D exactly five failures **E** more failures than successes

19 The proportion of female students at a particular university is 0.2. A sample of 10 students is chosen at random from the entire student population. What is the probability that the sample contains exactly four female students?

A 0.0881 **B** 0.5000 **C** 0.0328 **D** 0.0016 **E** 0.9672

20 Mai decides to call five friends to invite each of them to a party. The probability of a friend not being at home when Mai calls is p. An expression for the probability that Mai finds at least one of her friends at home is given by

A p^5 **B** $1-(1-p)^5$ **C** $5p(1-p)^4$ **D** $1-p^5$ **E** $5p^4(1-p)$

21 If a random variable X has probability density function given by

$$f(x) = \begin{cases} kx^3 + \frac{3}{4}x & \text{if } 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

then k is equal to

A $-\frac{3}{16}$ **B** $\frac{6}{25}$ **C** $-\frac{9}{16}$ **D** $-\frac{1}{8}$ **E** $-\frac{3}{8}$

22 If a random variable X has probability density function given by

$$f(x) = \begin{cases} \frac{1}{9}(4x - x^2) & \text{if } 0 < x < 3 \\ 0 & \text{otherwise} \end{cases}$$

then $\Pr(X \leq 2)$ is closest to

A 0.6667 **B** 0.4074 **C** 0.5926 **D** 0.4444 **E** 0.5556

23 A random variable X has probability density function:

$$f(x) = \begin{cases} \frac{8}{3}(1 - x) & \text{if } 0 < x < \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$$

The median, that is the value of m such that $\Pr(X < m) = 0.5$, of X is closest to

A 0.222 **B** 0.667 **C** 0.250 **D** 1.791 **E** 0.209

24 A random variable X has probability density function:

$$f(x) = \begin{cases} 2\left(1 - \frac{1}{x^2}\right) & \text{if } 1 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

The mean of X is closest to

A 1 **B** 1.614 **C** 2 **D** 1.5 **E** 0.609

ISBN 978-1-009-11049-5 Cambridge University Press

25 The probability of obtaining a z-value which falls between $z = -1.0$ and $z = 0$ for a standard normal distribution is approximately

A 0.05 **B** 0.20 **C** 0.34 **D** 0.68 **E** 0.16

26 For a normal probability distribution, which of the following is true?

A The mean is always positive.
B No value can be more than four standard deviations away from the mean.
C The area under the normal curve is approximately equal to 1.
D The standard deviation is always positive.
E The standard deviation is less than the mean.

27 If X is a normally distributed random variable with mean $\mu = 2$ and standard deviation $\sigma = 0.5$, then the probability that X is greater than 2.6 is closest to

A 0.8849 **B** 0.9918 **C** 0.1151 **D** 0.0082 **E** 0.0302

28 If X is a normally distributed random variable with mean $\mu = 2$ and standard deviation $\sigma = 2$, then the probability that X is less than -2 is

A 0.1587 **B** 0.8413 **C** 0.9772 **D** 0.1228 **E** 0.0228

29 If X is a normally distributed random variable with mean $\mu = 3$ and variance $\sigma^2 = 0.4$, then the probability that X is greater than -2.73 is

A 1 **B** 0 **C** 0.9115 **D** 0.0885 **E** 0.5537

30 If X is a normally distributed random variable with mean $\mu = 2$ and variance $\sigma^2 = 4$, then $\Pr(1 < X < 2.5)$ is

A 0.5987 **B** 0.2902 **C** 0.6915 **D** 0.4013 **E** 0.3085

31 An automatic dispensing machine fills cups with cordial. If the amount of cordial in the cup is a normally distributed random variable with a mean of 50 mL and a standard deviation of 2 mL, then 90% of the cups contain more than

A 44.87 mL **B** 53.29 mL **C** 46.71 mL **D** 52.56 mL **E** 47.44 mL

32 Lengths of blocks of cheese are found to be normally distributed with a mean of 10 cm and a variance of 0.5. Then 95% of the blocks of cheese are shorter than

A 11.39 cm **B** 8.84 cm **C** 11.16 cm **D** 8.61 cm **E** 9.18 cm

33 Assume that X is a normally distributed random variable with mean $\mu = 1$ and variance $\sigma^2 = 2.25$. If $\Pr(\mu - k < X < \mu + k) = 0.7$, then $k =$

A 1.555 **B** 1.037 **C** 0.787 **D** 0.524 **E** 2.332

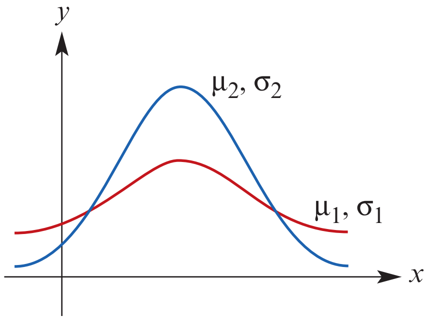
34 The weight of a packet of biscuits is known to be normally distributed with a mean of 1 kg. If a packet is more than 0.05 kg underweight, it is unacceptable. If it is found that 3% of packets are unacceptable, then the standard deviation of the weight is

A 1.881 **B** 0.027 **C** 10.488 **D** 0.030 **E** 37.616

ISBN 978-1-009-11049-5 Cambridge University Press

Revision

35 The diagram shows the probability density functions of two normally distributed random variables, one with mean μ_1 and standard deviation σ_1, and the other with mean μ_2 and standard deviation σ_2.
Which of the following statements is true?



A $\mu_1 = \mu_2,\ \sigma_1 < \sigma_2$
B $\mu_1 = \mu_2,\ \sigma_1 > \sigma_2$
C $\mu_1 > \mu_2,\ \sigma_1 = \sigma_2$
D $\mu_1 < \mu_2,\ \sigma_1 = \sigma_2$
E $\mu_1 = \mu_2,\ \sigma_1 = \sigma_2$

36 If the heights of a certain population of men are normally distributed with a mean of 173 cm and a variance of 25, then about 68% of men in the population have heights in the interval (in cm)

A (148, 198) **B** (168, 178) **C** (163, 183) **D** (158, 188) **E** (123, 223)

37 In a random sample of 200 people, 38% said they would rather watch tennis on television than attend the match. An approximate 95% confidence interval for the proportion of people in the population who prefer to watch tennis on television is

A (0.136, 0.244) **B** (0.313, 0.447) **C** (0.285, 0.475)
D (0.255, 0.505) **E** (0.292, 0.468)

38 A 99% confidence interval for the proportion of people in the a certain country who live in an apartment is (0.284, 0.336). The sample proportion from which this interval was constructed is

A 0.052 **B** 0.289 **C** 0.293
D 0.310 **E** 0.620

39 For a fixed sample, an increase in the level of confidence will lead to a confidence interval which is

A narrower **B** wider **C** unchanged **D** asymmetric
E cannot be determined from the information given

40 Which of the following statements are true?

I The lower the level of confidence, the smaller the confidence interval.
II The larger the sample size, the smaller the confidence interval.
III The smaller the sample size, the smaller the confidence interval.
IV The higher the level of confidence, the smaller the confidence interval.

A I and II **B** I and III **C** II only **D** II and IV **E** none of these

ISBN 978-1-009-11049-5 Cambridge University Press

41 If a researcher decreases her sample size by a factor of 2, then the width of a 95% confidence interval would

A increase by a factor of 2
B increase by a factor of $\sqrt{2}$
C decrease by a factor of $\sqrt{2}$
D decrease by a factor of 4
E none of these

42 In a certain country it is known that 25% of people own cats. A random sample of 30 people is to be selected. If $\hat{P}$ is the proportion of people in the sample who own cats, then (do not use a normal approximation)

$$\Pr\left(\hat{P} \geq \frac{3}{10}\right)$$

is closest to

A 0.8030
B 0.6736
C 0.3264
D 0.9675
E 0.1966

18C Extended-response questions

1 A fish shop catches their trout from two different lakes, lake A and lake B. The weight of the trout caught from lake A is normally distributed with a mean of $\mu = 3.6$ kg and a standard deviation of $\sigma = 0.5$ kg.

a Find the probability that a trout caught from lake A weighs more than 4.25 kg, correct to four decimal places.

b The probability that a trout caught from lake A weighs more than k kg is 0.9. Find the value of k, correct to two decimal places.

The weight of trout from lake B is modelled by the following probability density function

$$b(x) = \begin{cases} \frac{\pi}{6}\cos\left(\frac{\pi(2x-7)}{6}\right) & 2 \leq x \leq 5 \\ 0 & \textit{elsewhere} \end{cases}$$

c Determine the mean weight of the fish caught from lake B, correct to two decimal places.

d Find the probability that a randomly selected trout caught from lake B will have a weight greater than 4.25kg.

e The probability that a trout caught from lake B weighs more than c kg is 0.9. Find the value of c, correct to two decimal places.

The fish shop catches 60% of their trout from lake A, and 40% from lake B.

f Find the probability that a randomly selected trout weighs more than 4.0 kg, giving your answer correct to three decimal places.

g Find the probability that a randomly selected trout which weighs more than 4.0 kg was caught in lake A, giving your answer correct to three decimal places.

ISBN 978-1-009-11049-5 Cambridge University Press

h If the fish shop packs the combined catch of trout into boxes of 6 for distribution to restaurants, find the probability that at least one of the trout in a randomly chosen box weighs more than 4 kg, giving your answer correct to two decimal places.

i Suppose that the fish are packed in boxes of size n for distribution to restaurants, and that $\hat{P}_n$ is the random variable which represents the proportion of fish in the box that weigh more than 4.0 kg.

Find the least value of n such that $\Pr\left(\hat{P}_n < \frac{1}{n}\right) < 0.1$

2 Electronic sensors of a certain type fail when they become too hot. The temperature at which a randomly chosen sensor fails is T°C, where T is modelled as a normal random variable with mean $\mu = 95$ and standard deviation $\sigma = 5$.

a If $\Pr(T < a) = 0.8$, find a to one decimal place.

b Find the probability, correct to three decimal places, that the temperature at which the sensor fails is less than 112°C, given that it was still working at 108°C.

c The manufacturer of the sensors has determined that 81.86% of the sensors fail at temperatures between 90°C and 105°C. A new improved version of the sensor is being developed, which will have mean fail temperature of $\mu = k$ and a standard deviation $\sigma = 5$. Find the values of k, so that 81.86% of the new sensors fail at temperatures between 95°C and 110°C.

d In a laboratory test another type of sensor, 98% of a random sample of sensors continued working at a temperature of 80°C, and 4% continued working at 110°C. Determine, correct to one decimal place, estimates for the values of μ and σ.

Another manufacturer of sensors claims that 10% of their sensors will continue working at 110°C. Assuming that this claim is true:

e Find, correct to three decimal places, the probability that in a box of 20 sensors, at least two of them will continue working at 110°C.

f Find the minimum number of sensors that should be packed in a box to ensure that that probability that at least one sensor will continue working at 110°C is at least 90%.

To test this companies claim, a random sample of m sensors was tested, and an approximate 95% confidence interval for the proportion of sensors continues working at 110°C was found to be (0.0186, 0.1414).

g Find the value of $\hat{p}$ that was used to obtain this approximate 95% confidence interval.

h Find the size of the sample from which this approximate 95% confidence interval was obtained. Give your answer to the nearest whole number.

3 A train is declared to be 'on time' if it arrives at its destination within 5 minutes either side of the advertised arrival time; otherwise, it is declared early or late. The time difference between the actual arrival time and the scheduled arrival time, T, is normally distributed with a mean of 3 minutes and a standard deviation of 3 minutes. The time of arrival of a particular train on any one day is independent of the time of arrival of the flight on any other day.

a Calculate, correct to four decimal places, the probability that:

ISBN 978-1-009-11049-5 Cambridge University Press

i On any given day, the train arrives on time.

ii On any given day, the train arrives late.

The train company needs to pay a fine if they are consistently late. If the train is late less than twice in a 7 day week they are not penalised, but if they are late twice in a week they are fined \$3000, and if they are late more than twice they are fined \$10 000.

b Construct a probability distribution table for F, the random variable which represents the amount the train company might be fined in a week, in dollars.

c Use the table to calculate the expected value and the standard deviation of F, correct to two decimal places.

The operators of the station continuously monitor the proportion of trains which arrive at the station which are classified as late. Of the 268 trains which arrives on particular day, 33 were late.

d Find the 95% confidence interval for the proportion of trains which are late, giving the values correct to three decimal places.

The difference between the scheduled arival time and the actual arrival time, D, for different station is given by the probability density function :

$$f(d) = \frac{1}{2\pi}\left(1 + \sin\left(\frac{d}{2}\right)\right), \quad \pi \le d \le 3\pi$$

e Find correct to one decimal place the expected value of D.

f Find correct to four decimal places the probability that one of these trains will arrive more than 8 minutes after the scheduled arrival time.

4 A factory has two machines that produce widgets. Machine A and Machine B. The time taken, X seconds, to produce a widget using machine A has probability density function given by

$$f(x) = \begin{cases} k(x-8)(12-x)^2 & \text{if } 8 \le x \le 12 \\ 0 & \text{otherwise} \end{cases}$$

a Find the value of k.

b Find the mean time it takes to produce a widget from Machine A, giving you answer correct to one decimal place.

c In a sample of 80 widgets selected from the production of Machine A, how many would be expected to have taken more than 10 seconds to produce?

d What is the probability that a widget takes more than 11 seconds to produce, given that it took at least 10 seconds to produce? Give you answer correct to four decimal places.

The time taken, Y seconds to produce a widget using machine B is normally distributed, with a mean of 10 seconds and a standard deviation of 1.2 seconds.

e Find the probability, correct to four decimal places, that a randomly selected widget from the production of Machine B has taken between 8.5 and 10.5 seconds to produce.

ISBN 978-1-009-11049-5 Cambridge University Press

f If $\Pr(Y > a) = 0.05$, find a to one decimal place.

g A widget which takes more that 12.2 seconds to produce is considered to be unprofitable.

i Suppose that a random sample of 64 widgets is selected from Machine B. Find, correct to four decimal places, the probability that at least one of them would be unprofitable?

ii If the probability that m or more widgets in a sample of size 64 from Machine B are unprofitable is less that 0.05%, find the smallest value of m, where m is an integer.

iii For a random sample of 64 widgets from Machine B, $\hat{P}$ is the random variable that represents the proportion of widgets in the sample that are unprofitable. Find the expected value and standard deviation of $\hat{P}$ correct to three decimal places.

h Suppose that the widgets manufactured by machine A, and those produced machine B are combined at the end of each day. If the probability that widget selected at random from the combined production was produced in less than 10 seconds is 0.6125, what proportion of the widgets were produced by Machine A that day?

5 In each of a sequence of trials, the probability of the occurrence of a certain event is $\frac{1}{2}$, except that this event cannot occur in two consecutive trials.

a Show that the probability of this event occurring:

i exactly twice in three trials is $\frac{1}{4}$ **ii** exactly twice in four trials is $\frac{1}{2}$.

b What is the probability of this event occurring exactly twice in five trials?

6 Katia and Mikki play a game in which a fair six-sided die is thrown five times:

- Katia will receive \$1 from Mikki if there is an odd number of sixes
- Mikki will receive \$$x$ from Katia if there is an even number of sixes.

Find the value of x so that the game is fair. (Note that the number 0 is even.)

7 A newspaper seller buys papers for 50 cents and sells them for 75 cents, and cannot return unsold papers. Daily demand has the following distribution, and each day's demand is independent of the previous day's demand.

Number of customers	24	25	26	27	28	29	30
Probability	0.05	0.10	0.10	0.25	0.25	0.15	0.10

If the newspaper seller stocks too many papers, a loss is incurred. If too few papers are stocked, potential profit is lost because of the excess demand. Let s represent the number of newspapers stocked, and X the daily demand.

a If P is the newspaper seller's profit for a particular stock level s, find an expression for P in terms of s and X.

b Find the expected value of the profit, $\mathrm{E}(P)$, when $s = 26$.

ISBN 978-1-009-11049-5 Cambridge University Press

c Hence find an expression for the expected profit when s is unknown.

d By evaluating the expression for expected profit for different values of s, determine how many papers the newspaper seller should stock.

8 Anne and Jane play a game against each other, which starts with Anne aiming to throw a bean bag into a circle marked on the ground.

a The probability that the bean bag lands entirely inside the circle is $\frac{1}{2}$, and the probability that it lands on the rim of the circle is $\frac{1}{3}$.

i Show that the probability that the bean bag lands entirely outside the circle is $\frac{1}{6}$.

ii What is the probability that two successive throws land outside the circle?

iii What is the probability that, for two successive throws, the first lands on the rim of the circle and the second inside the circle?

b Jane then shoots at a target on which she can score 10, 5 or 0. With any one shot, Jane scores 10 with probability $\frac{2}{5}$, scores 5 with probability $\frac{1}{10}$, and scores 0 with probability $\frac{1}{2}$. With exactly two shots, what is the probability that her total score is:

i 20 **ii** 10?

c If the bean bag thrown by Anne lands outside the circle, then Jane is allowed two shots at her target; if the bean bag lands on the rim of the circle, then Jane is allowed one shot; if it lands inside the circle, then Jane is not allowed any shots. Find the probability that Jane scores a total of 10 as a result of any one throw from Anne.

9 A large taxi company determined that the distance travelled annually by each taxi is normally distributed with a mean of 80 000 km and a standard deviation of 20 000 km.

a What is the probability that a randomly selected taxi will travel between 56 000 km and 60 000 km in a year?

b What percentage of taxis can be expected to travel either less than 48 000 km or more than 96 000 km in a year?

c How many of the 250 taxis in the fleet are expected to travel between 48 000 km and 96 000 km in a year?

d At least how many kilometres would be travelled by 85% of the taxis?

10 The weight of cereal in boxes, packed by a particular machine, is normally distributed with a mean of μ g and a standard deviation of $\sigma = 5$ g.

a A box is considered underweight if it weighs less than 500 g.

i Find the proportion of boxes that will be underweight if $\mu = 505$ g.

ii Find the value of μ required to ensure that only 1% of boxes are underweight.

b As a check on the setting of the machine, a random sample of five boxes is chosen and the setting changed if more than one of them is underweight. Find the probability that the setting of the machine is changed if $\mu = 505$ g.

ISBN 978-1-009-11049-5

11 The queuing time, X minutes, at the box office of a movie theatre has probability density function:

$$f(x) = \begin{cases} kx(100 - x^2) & \text{if } 0 \le x \le 10 \\ 0 & \text{otherwise} \end{cases}$$

a Find:

i the value of k **ii** the mean of X

iii the probability that a moviegoer will have to queue for more than 3 minutes

iv the probability that a moviegoer will have to queue for more than 3 minutes, given that she queues for less than 7 minutes.

b If 10 moviegoers go independently to the theatre, find the probability that at least five of them will be required to queue for more than 3 minutes.

12 Jam is packed in tins of nominal net weight 1 kg. The actual weight of jam delivered to a tin by the filling machine is normally distributed about the mean weight set on the machine, with a standard deviation of 12 g.

a If the machine is set to 1 kg, find the probability that a tin chosen at random contains less than 985 g.

b It is a legal requirement that no more than 1% of tins contain less than the nominal weight. Find the minimum setting of the filling machine which will meet this requirement.

13 In a factory, machines A, B and C are all producing springs of the same length. Of the total production of springs in the factory, machine A produces 35% and machines B and C produce 25% and 40% respectively. Of their production, machines A, B and C produce 3%, 6% and 5% defective springs respectively.

a Find the probability that:

i a randomly selected spring is produced by machine A and is defective

ii a randomly selected spring is defective.

b Given that a randomly selected spring is defective, find the probability that it was produced by machine C.

c Given that a randomly selected spring is not defective, find the probability that it was produced by either machine A or machine B.

ISBN 978-1-009-11049-5 Cambridge University Press

14 An electronic game comes with five batteries. The game only needs four batteries to work. But because the batteries are sometimes faulty, the manufacturer includes five of them with the game. Suppose that X is the number of good batteries included with the game. The probability distribution of X is given in the following table.

x	0	1	2	3	4	5
$\Pr(X = x)$	0.01	0.02	0.03	0.04	0.45	0.45

a Use the information in the table to:

i find μ, the expected value of X

ii find σ, the standard deviation of X, correct to one decimal place

iii find, exactly, the proportion of the distribution that lies within two standard deviations of the mean

iv find the probability that a randomly selected game works, i.e. find $\Pr(X \geq 4)$.

b The electronic games are packed in boxes of 20. Whether or not an electronic game in a box will work is independent of any other game in the box working. Let Y be the number of working games in a box.

i Name the distribution of Y.

ii Find the expected number of working games in a box.

iii Find the standard deviation of the number of working games in a box.

iv Find the probability that a randomly chosen box will contain at least 19 working games.

15 In a study of the prevalence of red hair in a certain country, researchers collected data from a random sample of 1800 adults.

a Of the 1000 females in the sample, they found that 10% had red hair. Calculate an approximate 95% confidence interval for the proportion with red hair in the female population.

b Of the 800 males in the sample, they found that 10% had red hair. Calculate an approximate 95% confidence interval for the proportion with red hair in the male population.

c Why is the width of the confidence interval for males different from the width of the confidence interval for females?

d How should the sample of 1800 adults be chosen to ensure that the widths of the two confidence intervals are the same when the sample proportions are the same?

e Assume that there are 1000 females and 800 males in the sample, and that the proportion of females in the sample with red hair is 10%. What sample proportion of red-headed males would result in the 95% confidence interval for the proportion with red hair in the female population and the 95% confidence interval for the proportion with red hair in the male population being of the same width?

ISBN 978-1-009-11049-5 Cambridge University Press

Revision

18D Algorithms and pseudocode

Skill-sheet

An introduction to pseudocode is given in Appendix A of this book and the reader is referred to that appendix for explanations of the terms used in this section. You may like to use a device to implement the algorithms in this section; see the coding appendices in the Interactive Textbook for instructions.

Algorithms for simulation to estimate probabilities

For this section, we introduce two useful pseudocode functions:

- The function *random*( ) generates a random real number in the interval (0, 1).
- The function *randint*(n, m) generates a random integer between n and m inclusive.

1 We start with an algorithm to estimate by simulation the probability of getting a six when a dice is thrown.

```
input N
count ← 0
for i from 1 to N
    outcome ← randint(1, 6)
    if outcome = 6 then
        count ← count + 1
    end if
end for
estimate ← count / N
print estimate
```

We use the variables:

- N for the number of times to roll the die
- *outcome* for the result of the current roll
- *count* to keep a running tally of the number of sixes obtained.

We use a `for` loop to simulate rolling a die N times.

In each pass of the `for` loop, we generate a random integer between 1 and 6 to simulate rolling the die. If the outcome is a six, then we add one to the count.

The estimate of the probability is the number of sixes divided by the number of rolls.

a Change the algorithm to estimate

i the probability of an even number.

ii the probability of a number less than 5.

b Two dice are thrown and the numbers on their uppermost faces added. Use nested loops and simulation to find the probability of obtaing a sum between 6 and 10 inclusive.

2 Let $A = \{0, 1, 2, \dots, 9\}$. One digit is randomly selected from A. Describe an algorithm to estimate by simulation that the probability that:

a a 9 is obtained.

b an even number is obtained.

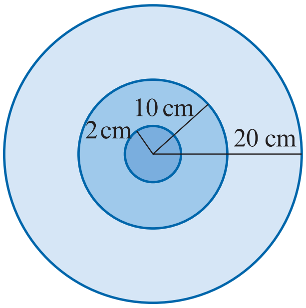
c a number greater than or equal to 3 and less than or equal to 7 is obtained.

d a second number is randomly selected and it is the same as the first. (We are of course assuming replacement.)

ISBN 978-1-009-11049-5 Cambridge University Press

We saw the following question in Chapter 13. Simulation can be used to obtain estimates of the probabilities which we can consider from the game of darts.

3 A dartboard consists of three circular sections, with radii of 2 cm, 10 cm and 20 cm respectively, as shown in the diagram.
When a dart lands in the centre circle the score is 100 points, in the middle circular section the score is 20 points and in the outer circular section the score is 10 points.



A throw is valid if it hits within the circular target.We are only concerned with valid throws. Each dart is equally likely to land at any point on the circular dartboard, and none lands on the lines. Here is an algorithm which records the number of valid throws and the number of each score obtained out of 1000 attempts.

```
counta ← 0
countb ← 0
countc ← 0
validthrow ← 0
for i from 1 to 1000
    x ← random(−20, 20)
    y ← random(−20, 20)
    if x^2 + y^2 ≤ 400 then
        validthrow ← validthrow + 1
        if x^2 + y^2 ≥ 100 then
            counta ← counta + 1
        else if x^2 + y^2 ≥ 4 then
            countb ← countb + 1
        else
            countc ← countc + 1
        end if
    end if
end for
print validthrow, counta, countb, countc
```

We use a for loop to repeat the throw 1000 times.

We use an if statement to restrict our throws to valid thows.

We use an if statements determine the number of ech of the score types.

We use the variables:

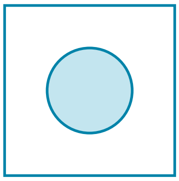
- x for the x-coordinate of a point in the plane, $x \in [-20, 20]$.
- y for the y-coordinate of a point in the plane, $y \in [-20, 20]$.
- *counta* to give the number of hits in the outer ring.
- *countb* to give the number of hits in the middle ring.
- *countc* to give the number of hits in the 'bullseye'.

a Adapt the code to estimate the probability of a
  - **i** valid throw.
  - **ii** score of 100 given that the throw is valid.
  - **iii** score of 20 given that the throw is valid.
  - **iv** score of 10 given that the throw is valid.

b Adapt the code to estimating the probabilities associated with total score when two throws are taken and the scores summed. Only valid throws are considered.The scores that can be attained are 20, 30, 40, 110, 120 and 200.

ISBN 978-1-009-11049-5 Cambridge University Press

4 A dart is thrown at the board shown. The board is a square of side length 50 cm, and the circle has radius R cm. The dart is equally likely to hit any point on the board. If it hits the shaded region inside the circle, the score is 50 points; otherwise, the score is 10 points.



Describe an algorithm using simulation to estimate the probabilities associated with throwing three darts at the dart board for different values of R.

5 **A simple random walk** This problem is an example of a one-dimensional random walk. Random walks have applications, in Finance, Physics, Genetics, Ecology and many other fields.

Consider moving along a number line which is numbered from −5 to 5. Units are metres. Start at the origin and toss a coin to decide whether to go one metre to the left or go one metre to the right, heads to the right and tails to the left. Repeat this every metre. For how many moves will you stay in this interval of the number line? You are finished when the throw of the coin takes you out of the interval$[-5, 5]$. Once out, you cannot get back in.

We give an algorithm to perform such a walk a thousand times and find the average number of moves it takes before you move out of this interval of the number line.

```
sum ← 0
for j from 1 to 1000
    count ← 0
    x ← 0
    while x ≤ 5 and x ≥ −5
        A ← random( )
            if A < 0.5 then
                x ← x + 1
            else
                x ← x − 1
            end if
            count ← count + 1
    end while
    sum ← sum + count
    average ← sum/j
end for

print average
```

We use a for loop to repeat our walk 1000 times.

We use a while loop to keep our walk going while we are on the interval $[-5, 5]$

Inside the while loop we use the variables:

- x for the current position on a number line
- *count* to give the number of moves before we move beyond $[-5, 5]$
- A to record if you get a head or a tail. If a head add one. If a tail subtract 1.

Inside the for loop we use the variables:

- *sum* to to total the counts
- *average* derived by dividing sum by$j = 1000$.

The result is the average number of moves that are taken before moving out of $[-5, 5]$.

a Rewrite the algorithm with the probability that you move to the right being 0.6 and to the left 0.4.

b Consider the number line, $(-\infty, \infty)$, marked off at the integers $\dots, -3, -2, -1, 0, 1, 2, 3, \dots$. Describe an algorithm to simulate moving N times and giving the final position. Use: probability of moving to the right = probability of moving to the left = 0.5.

ISBN 978-1-009-11049-5 Cambridge University Press

6 Random walks in two dimensions This is a more difficult question. You start at the origin and you can go North, South, East or West for one unit, each with probability 0.25. Describe an algorithm in pseudocode which gives a position after N moves.

Algorithms using simulation for other purposes

7 A question that is difficult answer and impossible with school mathematics is to find the average distance between two points in a geometric shape on the Cartesian plane.

We start with the square with vertices at the points $(0, 0), (1, 0), (1, 1), (0, 1)$.We can do this with simulation. (N has to be large, for example take $N = 100\,000$)

```
x ← 0; y ← 0; u ← 0; z ← 0
average ← 0
for j from 1 to N
    x ← random( )
    y ← random( )
    u ← random( )
    z ← random( )
    d ← √((x − u)² + (y − z)²)
    average ← (average × (j − 1) + d)/j
end for

print average
```

We use a for loop to repeat the distance calculation N times.

- The coordinates of the two points for each pass are (x, y) and (u, z)
- d is the calculated distance betwen the two points
- *average* is the ongoing average of the distances.

Write the code to determine the average distance between two points in the unit circle.

Note: Have the centre at the origin and so the only points to consider are those such that $x^2 + y^2 \leq 1$.

Algorithms for calculating exact probabilities

8 In the previous questions, we have seen algorithms used to obtain estimates. We can also use algorithms to calculate exact probabilities.
For example, we already know how to find the probability of achieving a certain sum when two dice are rolled and the two numbers added. Here there are 36 equally likely outcomes to consider. However, if we roll three dice, then there are 216 outcomes, and if we roll four dice, then there are 1296 outcomes. When there is a very large number of equally likely outcomes, we can use a computer to calculate an exact probability.
Three dice are rolled and the sum of the three numbers on the uppermost faces recorded. We describe an algorithm to find the probability that this sum is between 4 and 11 inclusive.

ISBN 978-1-009-11049-5 Cambridge University Press

- Let i, j and k represent the numbers obtained on the three dice. We use three loops to run through all the possible values of i, j and k.
- The variable *count* keeps a running tally of the number of outcomes (i, j, k) such that $4 \le i + j + k \le 11$.
- The variable *total* keeps a running tally of the total number of outcomes (i, j, k). We know that the final value of *total* will be $6 \times 6 \times 6 = 216$.

```
total ← 0
count ← 0
for i from 1 to 6
    for j from 1 to 6
        for k from 1 to 6
            total ← total + 1
            if 4 ≤ i + j + k ≤ 11 then
                count ← count + 1
            end if
        end for
    end for
end for
print count/total
```

a Three sets A, B and C are all equal to $\{0, 1, 2, \ldots, 9\}$. One digit is randomly selected from each of the sets. Describe an algorithm to calculate the exact probability that:

i the sum of the three digits is greater than 5 and less than 20.

ii the sum is greater than 20 given that it is greater than 15.

iii the digit from set A plus twice the digit from set B and 3 times the digit from set C is greater than 40.

b Let $A = \{0, 1, 2, \ldots, 9\}$ and $B = \{12, 13, 14, 15, 16\}$. One digit is randomly selected from each of the sets. Describe an algorithm to calculate the exact probability that:

i the sum of the two digits is greater than 16 and less than 20.

ii the sum is greater than 20 given that it is greater than 16.

9 **Binomial $\binom{n}{r}$**

a Prove that

$$\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}$$

b Use this result to construct an algorithm described by pseudocode to evaluate $\binom{n}{r}$ as a function of n and r.

ISBN 978-1-009-11049-5 Cambridge University Press