OCR
Oxford Cambridge and RSA

Tuesday 13 June 2023 – Afternoon

A Level Mathematics B (MEI)

H640/02 Pure Mathematics and Statistics

Time allowed: 2 hours

9917853939

You must have:
- the Printed Answer Booklet
- a scientific or graphical calculator

QP

INSTRUCTIONS
- Use black ink. You can use an HB pencil, but only for graphs and diagrams.
- Write your answer to each question in the space provided in the **Printed Answer Booklet**. If you need extra space use the lined pages at the end of the Printed Answer Booklet. The question numbers must be clearly shown.
- Fill in the boxes on the front of the Printed Answer Booklet.
- Answer **all** the questions.
- Where appropriate, your answer should be supported with working. Marks might be given for using a correct method, even if your answer is wrong.
- Give your final answers to a degree of accuracy that is appropriate to the context.
- Do **not** send this Question Paper for marking. Keep it in the centre or recycle it.

INFORMATION
- The total mark for this paper is **100**.
- The marks for each question are shown in brackets **[ ]**.
- This document has **16** pages.

ADVICE
- Read each question carefully before you start your answer.

© OCR 2023 [603/1002/9]
DC (ST/SG) 321177/4

OCR is an exempt Charity

stemup.co.uk

Formulae A Level Mathematics B (MEI) (H640)

Arithmetic series

$S_n = \frac{1}{2}n(a+l) = \frac{1}{2}n\{2a+(n-1)d\}$

Geometric series

$S_n = \dfrac{a(1-r^n)}{1-r}$

$S_\infty = \dfrac{a}{1-r}$ for $|r| < 1$

Binomial series

$(a+b)^n = a^n + {}^n\mathrm{C}_1 a^{n-1}b + {}^n\mathrm{C}_2 a^{n-2}b^2 + \ldots + {}^n\mathrm{C}_r a^{n-r}b^r + \ldots + b^n \qquad (n \in \mathbb{N}),$

where ${}^n\mathrm{C}_r = {}_n\mathrm{C}_r = \binom{n}{r} = \dfrac{n!}{r!(n-r)!}$

$(1+x)^n = 1 + nx + \dfrac{n(n-1)}{2!}x^2 + \ldots + \dfrac{n(n-1)\ldots(n-r+1)}{r!}x^r + \ldots \qquad (|x| < 1,\ n \in \mathbb{R})$

Differentiation

$\mathrm{f}(x)$	$\mathrm{f}'(x)$
$\tan kx$	$k\sec^2 kx$
$\sec x$	$\sec x \tan x$
$\cot x$	$-\mathrm{cosec}^2 x$
$\mathrm{cosec}\, x$	$-\mathrm{cosec}\, x \cot x$

Quotient Rule $y = \dfrac{u}{v}, \dfrac{\mathrm{d}y}{\mathrm{d}x} = \dfrac{v\dfrac{\mathrm{d}u}{\mathrm{d}x} - u\dfrac{\mathrm{d}v}{\mathrm{d}x}}{v^2}$

Differentiation from first principles

$\mathrm{f}'(x) = \lim_{h \to 0} \dfrac{\mathrm{f}(x+h) - \mathrm{f}(x)}{h}$

Integration

$\displaystyle\int \frac{\mathrm{f}'(x)}{\mathrm{f}(x)}\,\mathrm{d}x = \ln|\mathrm{f}(x)| + c$

$\displaystyle\int \mathrm{f}'(x)\big(\mathrm{f}(x)\big)^n\,\mathrm{d}x = \frac{1}{n+1}\big(\mathrm{f}(x)\big)^{n+1} + c$

Integration by parts $\displaystyle\int u\frac{\mathrm{d}v}{\mathrm{d}x}\,\mathrm{d}x = uv - \int v\frac{\mathrm{d}u}{\mathrm{d}x}\,\mathrm{d}x$

Small angle approximations

$\sin\theta \approx \theta$, $\cos\theta \approx 1 - \frac{1}{2}\theta^2$, $\tan\theta \approx \theta$ where θ is measured in radians

© OCR 2023

stemup.co.uk

Trigonometric identities

$\sin(A \pm B) = \sin A \cos B \pm \cos A \sin B$

$\cos(A \pm B) = \cos A \cos B \mp \sin A \sin B$

$\tan(A \pm B) = \dfrac{\tan A \pm \tan B}{1 \mp \tan A \tan B} \qquad \left(A \pm B \neq (k + \frac{1}{2})\pi\right)$

Numerical methods

Trapezium rule: $\displaystyle\int_a^b y\,\mathrm{d}x \approx \tfrac{1}{2}h\{(y_0 + y_n) + 2(y_1 + y_2 + \ldots + y_{n-1})\}$, where $h = \dfrac{b-a}{n}$

The Newton-Raphson iteration for solving $\mathrm{f}(x) = 0$: $x_{n+1} = x_n - \dfrac{\mathrm{f}(x_n)}{\mathrm{f}'(x_n)}$

Probability

$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$

$\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B \mid A) = \mathrm{P}(B)\mathrm{P}(A \mid B)$ **or** $\mathrm{P}(A \mid B) = \dfrac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}$

Sample variance

$s^2 = \dfrac{1}{n-1}S_{xx}$ where $S_{xx} = \Sigma(x_i - \bar{x})^2 = \Sigma x_i^2 - \dfrac{(\Sigma x_i)^2}{n} = \Sigma x_i^2 - n\bar{x}^2$

Standard deviation, $s = \sqrt{\text{variance}}$

The binomial distribution

If $X \sim \mathrm{B}(n, p)$ then $\mathrm{P}(X = r) = {}^n\mathrm{C}_r p^r q^{n-r}$ where $q = 1 - p$

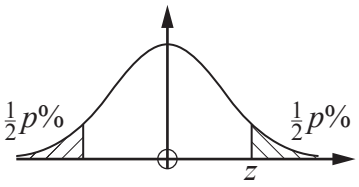
Mean of X is np

Hypothesis testing for the mean of a Normal distribution

If $X \sim \mathrm{N}(\mu, \sigma^2)$ then $\overline{X} \sim \mathrm{N}\left(\mu, \dfrac{\sigma^2}{n}\right)$ and $\dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim \mathrm{N}(0, 1)$

Percentage points of the Normal distribution

p	10	5	2	1
z	1.645	1.960	2.326	2.576



Kinematics

Motion in a straight line	Motion in two dimensions
$v = u + at$	$\mathbf{v} = \mathbf{u} + \mathbf{a}t$
$s = ut + \frac{1}{2}at^2$	$\mathbf{s} = \mathbf{u}t + \frac{1}{2}\mathbf{a}t^2$
$s = \frac{1}{2}(u + v)t$	$\mathbf{s} = \frac{1}{2}(\mathbf{u} + \mathbf{v})t$
$v^2 = u^2 + 2as$	
$s = vt - \frac{1}{2}at^2$	$\mathbf{s} = \mathbf{v}t - \frac{1}{2}\mathbf{a}t^2$

Section A (25 marks)

1 Determine the sum of the infinite geometric series $9-3+1-\frac{1}{3}+\frac{1}{9}+\ldots$ **[3]**

2 The equation of a circle is

$x^2-12x+y^2+8y+3=0$.

(a) Find the radius of the circle. **[2]**

(b) State the coordinates of the centre of the circle. **[1]**

3 **In this question you must show detailed reasoning.**

Find the smallest possible positive integers m and n such that $\left(\frac{64}{49}\right)^{-\frac{3}{2}}=\frac{m}{n}$. **[3]**

4 A biased octagonal dice has faces numbered from 1 to 8. The discrete random variable X is the score obtained when the dice is rolled once. The probability distribution of X is shown in the table below.

x	1	2	3	4	5	6	7	8
$P(X=x)$	p	p	p	p	p	p	p	$3p$

(a) Determine the value of p. **[2]**

(b) Find the probability that a score of at least 4 is obtained when the dice is rolled once. **[1]**

The dice is rolled 30 times.

(c) Determine the probability that a score of 8 occurs exactly twice. **[2]**

5 You are given that $\overrightarrow{OA}=\begin{pmatrix}3\\-1\end{pmatrix}$ and $\overrightarrow{OB}=\begin{pmatrix}5\\-3\end{pmatrix}$.

Determine the exact length of AB. **[3]**

© OCR 2023

stemup.co.uk

6 The parametric equations of a circle are

$x = 2\cos\theta - 3$ and $y = 2\sin\theta + 1$.

Determine the cartesian equation of the circle in the form $(x-a)^2 + (y-b)^2 = k$, where a, b and k are integers. **[4]**

7 The coefficient of x^8 in the expansion of $(2x+k)^{12}$, where k is a positive integer, is 79 200 000.

Determine the value of k. **[4]**

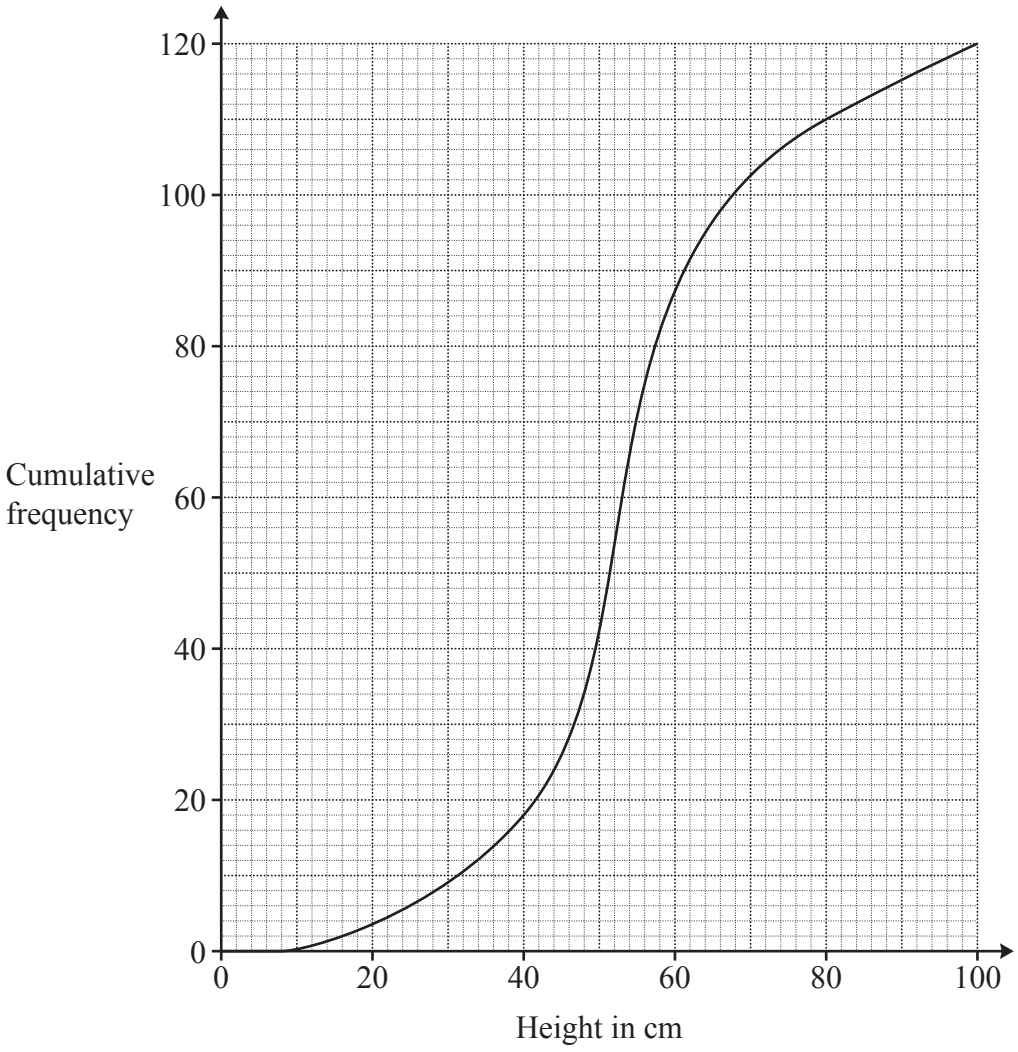
stemup.co.uk

Section B (75 marks)

8 A garden centre stocks coniferous hedging plants. These are displayed in 10 rows, each of 120 plants. An employee collects a sample of the heights of these plants by recording the height of each plant on the front row of the display.

(a) Explain whether the data collected by the employee is a simple random sample. **[1]**

The data are shown in the cumulative frequency curve below.



The owner states that at least 75% of the plants are between 40 cm and 80 cm tall.

(b) Show that the data collected by the employee supports this statement. **[4]**

(c) Explain whether all samples of 120 plants would necessarily support the owner's statement. **[1]**

© OCR 2023

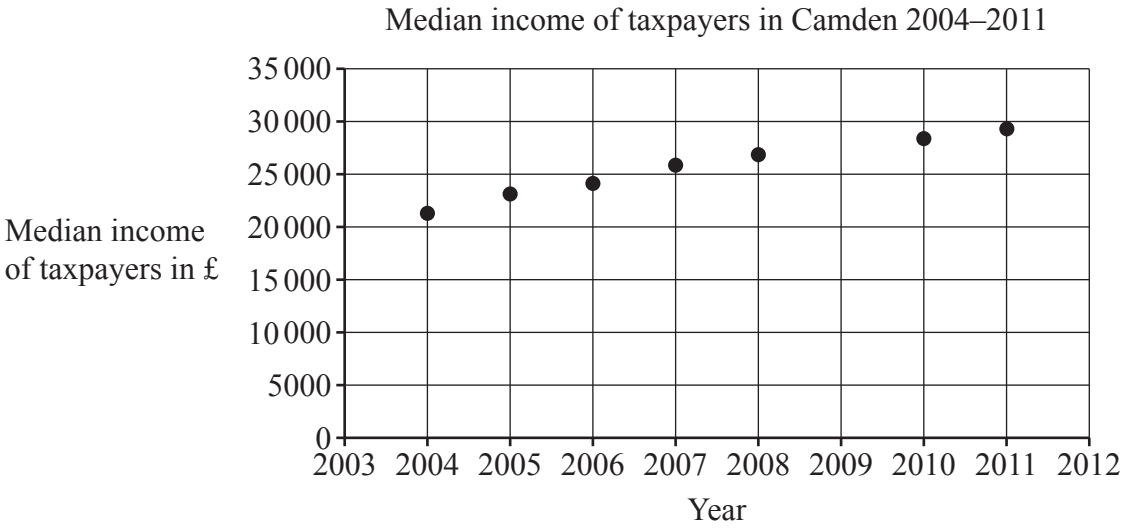
stemup.co.uk

9 The pre-release material contains information concerning the median income of taxpayers in different areas of London. Some of the data for Camden is shown in the table below. The years quoted in this question refer to the end of the financial years used in the pre-release material. For example, the year 2004 in the table refers to the year 2003/04 in the pre-release material.

Year	2004	2005	2006	2007	2008	2009	2010	2011
Median Income in £	21 300	23 200	24 200	25 900	26 900	#N/A	28 400	29 400

(a) Explain whether these data are a sample or a population of Camden taxpayers. **[1]**

A time series for the data is shown below.



The LINEST function on a spreadsheet is used to formulate the following model for the data:

$I = 1115Y - 2\,212\,950$, where I = median income of taxpayers in £ and Y = year.

(b) Use this model to find an estimate of the median income of taxpayers in Camden in 2009. **[1]**

(c) Give **two** reasons why this estimate is likely to be close to the true value. **[2]**

The median income of taxpayers in Croydon in 2009 is also not available.

(d) Use your knowledge of the pre-release material to explain whether the model used in part **(b)** would give a reasonable estimate of the missing value for Croydon. **[1]**

10 Determine the exact value of $\int_0^{\frac{\pi}{4}} 4x\cos 2x \, \mathrm{d}x$. **[5]**

stemup.co.uk

11 **In this question you must show detailed reasoning.**

The variables x and y are such that $\frac{dy}{dx}$ is directly proportional to the square root of x.

When $x = 4$, $\frac{dy}{dx} = 3$.

(a) Find $\frac{dy}{dx}$ in terms of x. **[3]**

When $x = 4$, $y = 10$.

(b) Find y in terms of x. **[3]**

12 It is given that

- $f(x) = \pm\frac{1}{\sqrt{x}}, x > 0$
- $g(x) = \frac{x}{x-3}, x > 3$
- $h(x) = x^2 + 2, x \in \mathbb{R}$.

(a) Explain why $f(x)$ is **not** a function. **[1]**

(b) Find $gh(x)$. **[2]**

(c) State the domain of $gh(x)$. **[1]**

stemup.co.uk

13 A large supermarket chain advertises that the mean mass of apples of a certain variety on sale in their stores is 0.14 kg.

Following a poor growing season, the head of quality control believes that the mean mass of these apples is less than 0.14 kg and she decides to carry out a hypothesis test at the 5% level of significance.

She collects a random sample of this variety of apple from the supermarket chain and records the mass, in kg, of each apple. She uses software to analyse the data. The results are summarised in the output below.

n	80
Mean	0.1316
σ	0.0198
s	0.0199
Σx	10.525
Σx^2	1.4161
Min	0.1
Q1	0.12
Median	0.132
Q3	0.1435
Max	0.19

(a) State the null hypothesis and the alternative hypothesis for the test, defining the parameter used. **[2]**

(b) Write down the distribution of the sample mean for this hypothesis test. **[2]**

(c) Determine the critical region for the test. **[2]**

(d) Carry out the test, giving your conclusion in context. **[3]**

stemup.co.uk

14 The pre-release material contains information concerning the median income of taxpayers in £ and the percentage of all pupils at the end of KS4 achieving 5 or more GCSEs at grade A*–C, including English and Maths, for different areas of London.

Some of the data for 2014/15 is shown in **Fig. 14.1**.

Fig. 14.1

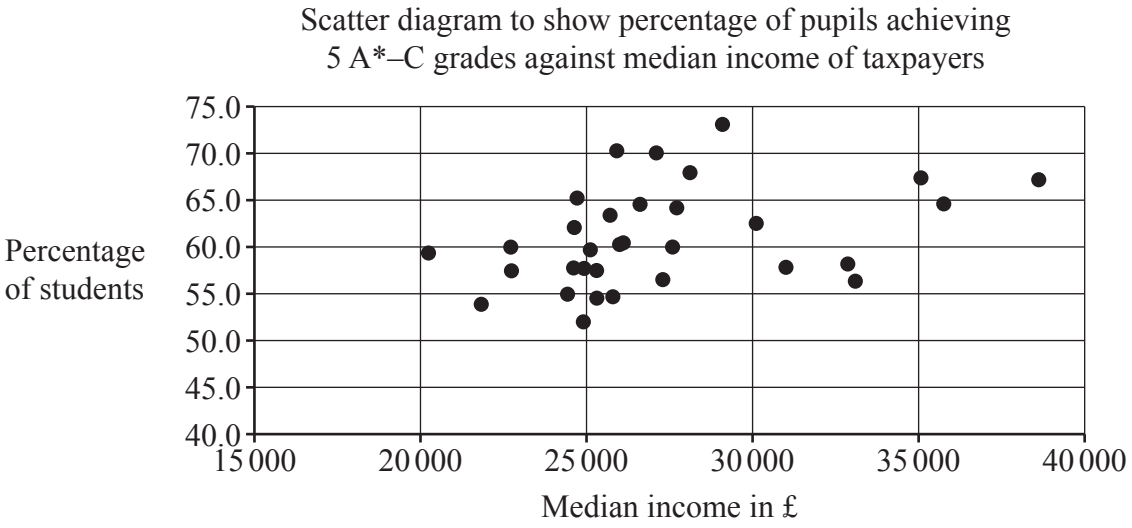
	Median Income of Taxpayers in £	Percentage of Pupils Achieving 5 or more A*–C, including English and Maths
City of London	61 100	#N/A
Barking and Dagenham	21 800	54.0
Barnet	27 100	70.1
Bexley	24 400	55.0
Brent	22 700	60.0
Bromley	28 100	68.0

A student investigated whether there is any relationship between median income of taxpayers and percentage of pupils achieving 5 or more GCSEs at grade A*–C, including English and Maths.

(a) With reference to **Fig. 14.1**, explain how the data should be cleaned before any analysis can take place. **[1]**

After the data was cleaned, the student used software to draw the scatter diagram shown in **Fig. 14.2**.

Fig. 14.2



The student calculated that the product moment correlation coefficient for these data is 0.3743.

stemup.co.uk

(b) Give **two** reasons why it may not be appropriate to use a linear model for the relationship between median income of taxpayers in £ and the percentage of all pupils at the end of KS4 achieving 5 or more GCSEs at grade A*–C. **[2]**

The student carried out some further analysis. The results are shown in **Fig. 14.3**.

Fig. 14.3

	median income of taxpayers in £	percentage of pupils achieving 5+ A*–C
mean	27216	61.0
standard deviation	4177.5	5.32

The student identified **three** outliers in total.

(c)

- Use the information in **Fig. 14.3** to determine the range of values of the median income of taxpayers in £ which are outliers.
- Use the information in **Fig. 14.3** to determine the range of values of the percentage of all pupils at the end of KS4 achieving 5 or more GCSEs at grade A*–C which are outliers.
- On the copy of **Fig. 14.2** in the **Printed Answer Booklet**, circle the **three** outliers identified by the student. **[4]**

The student decided to remove these outliers and recalculate the product moment correlation coefficient.

(d) Explain whether the new value of the product moment correlation coefficient would be between 0.3743 and 1 or between 0 and 0.3743. **[1]**

stemup.co.uk

15 In this question you must show detailed reasoning.

The equation of a curve is

$\ln y + x^3 y = 8$.

Find the equation of the normal to the curve at the point where $y = 1$, giving your answer in the form $ax + by + c = 0$, where a, b and c are constants to be found. **[7]**

16 Research conducted by social scientists has shown that 16% of young adults smoke cigarettes.

Two young adults are selected at random.

(a) Determine the probability that one smokes cigarettes and the other doesn't. **[2]**

The same research has also shown that

- 75% of young adults drink alcohol.
- 66% of young adults drink alcohol, but do **not** smoke cigarettes.

(b) Determine the probability that a young adult selected at random **does** smoke cigarettes, but **does not** drink alcohol. **[2]**

(c) A young adult who drinks alcohol is selected at random. Determine the probability that this young adult smokes cigarettes. **[2]**

(d) Using your answer to part **(c)**, explain whether the event that a young adult selected at random smokes cigarettes is independent of the event that a young adult selected at random drinks alcohol. **[2]**

17 In this question you must show detailed reasoning.

Solve the equation $2\sin x + \sec x = 4\cos x$, where $-\pi < x < \pi$. **[6]**

© OCR 2023

stemup.co.uk

18 Riley is investigating the daily water consumption, in litres, of his household.
He records the amount used for a random sample of 120 days from the previous twelve-month period.

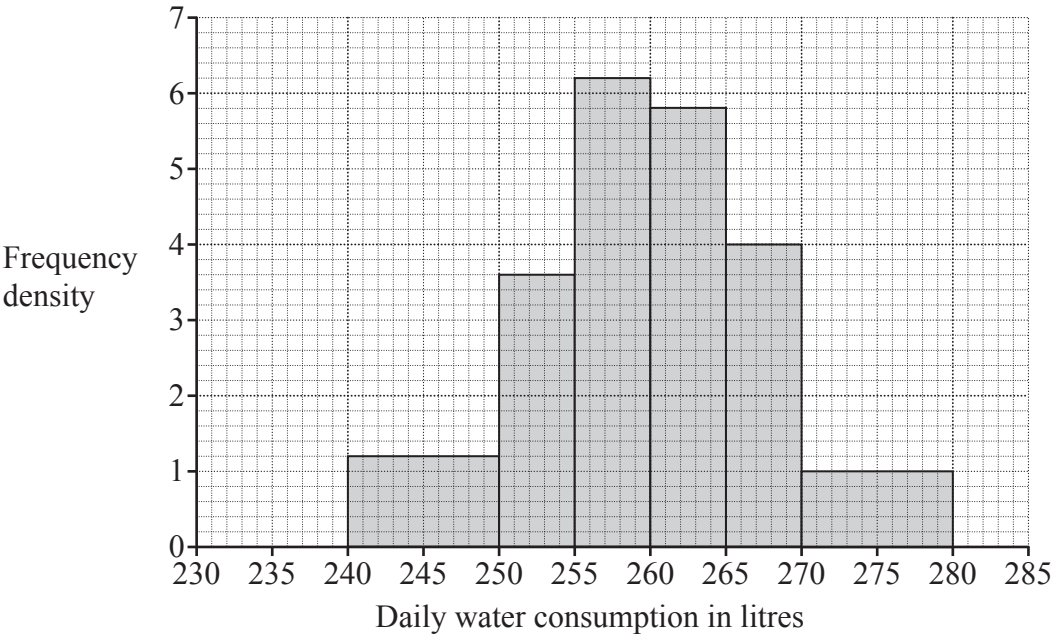
The daily water consumption, in litres, is denoted by x.

Summary statistics for Riley's sample are given below.

$\sum x = 31164.7 \quad \sum x^2 = 8101050.91 \quad n = 120$

(a) Calculate the sample mean giving your answer correct to **3** significant figures. **[1]**

Riley displays the data in a histogram.



(b) Find the number of days on which between 255 and 260 litres were used. **[1]**

(c) Give **two** reasons why a Normal distribution may be an appropriate model for the daily consumption of water. **[2]**

Riley uses the sample mean and the sample variance, both correct to **3** significant figures, as parameters of a Normal distribution to model the daily consumption of water.

(d) Use Riley's model to calculate the probability that on a randomly chosen day the household uses less than 255 litres of water. **[2]**

(e) Calculate the probability that the household uses less than 255 litres of water on **at least** 5 days out of a random sample of 28 days. **[2]**

Turn over for question 18(f)

stemup.co.uk

The company which supplies the water makes charges relating to water consumption which are shown in the table below.

Standing charge per day in pence	7.8
Charge per litre in pence	0.18

(f) Adapt Riley's model for daily water consumption to model the daily **charges** for water consumption. **[3]**

END OF QUESTION PAPER

© OCR 2023

stemup.co.uk

BLANK PAGE

© OCR 2023

stemup.co.uk

OCR
Oxford Cambridge and RSA

Copyright Information

OCR is committed to seeking permission to reproduce all third-party content that it uses in its assessment materials. OCR has attempted to identify and contact all copyright holders whose work is used in this paper. To avoid the issue of disclosure of answer-related information to candidates, all copyright acknowledgements are reproduced in the OCR Copyright Acknowledgements Booklet. This is produced for each series of examinations and is freely available to download from our public website (www.ocr.org.uk) after the live examination series.

If OCR has unwittingly failed to correctly acknowledge or clear any third-party content in this assessment material, OCR will be happy to correct its mistake at the earliest possible opportunity.

For queries or further information please contact The OCR Copyright Team, The Triangle Building, Shaftesbury Road, Cambridge CB2 8EA.

OCR is part of Cambridge University Press & Assessment, which is itself a department of the University of Cambridge.

© OCR 2023

stemup.co.uk

Surname	StemUp Model Solution
Answer booklet	Write the question number in the left-hand margin at the start of each answer.

Question number

1 $a = 9, \quad r = \frac{-3}{9} = \frac{-1}{3}$

$S_\infty = \frac{a}{1-r} = \frac{9}{1-(-1/3)} = \frac{9}{4/3} = \frac{27}{4}$

Sum to infinity $= \frac{27}{4} \quad (= 6.75)$

2 (a) $x^2 - 12x + y^2 + 8y + 3 = 0$

$(x-6)^2 - 36 + (y+4)^2 - 16 + 3 = 0$

$(x-6)^2 + (y+4)^2 = 49 = 7^2$

Radius $= 7$

2 (b) Centre $(6, -4)$

3 $\left(\frac{64}{49}\right)^{-3/2} = \left(\frac{49}{64}\right)^{3/2}$ (reciprocal)

$= \left(\frac{\sqrt{49}}{\sqrt{64}}\right)^3 = \left(\frac{7}{8}\right)^3$ (square root)

$= \frac{7^3}{8^3} = \frac{343}{512}$ (cube)

$343 = 7^3$ and $512 = 2^9$ have no common factor, so

$m = 343, \quad n = 512$

4 (a) $\Sigma P(X = x) = 1$

$7p + 3p = 1 \Rightarrow 10p = 1$

Question number	
	$p = 0.1$
4 (b)	$P(X \geqslant 4) = P(4) + P(5) + P(6) + P(7) + P(8)$
	$= 4p + 3p = 7p = 0.7$
4 (c)	$P(\text{score of 8}) = 3p = 0.3$
	Y = number of 8s in 30 rolls, $Y \sim B(30, 0.3)$
	$P(Y = 2) = {}^{30}C_2 \times 0.3^2 \times 0.7^{28}$
	$= 0.0018008... = 0.00180$ (3 s.f.)
5	$\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA} = \begin{pmatrix} 5 \\ -3 \end{pmatrix} - \begin{pmatrix} 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \end{pmatrix}$
	$\lvert AB \rvert = \sqrt{(2^2 + (-2)^2)} = \sqrt{8}$
	$AB = 2\sqrt{2}$
6	$\cos\theta = \frac{x+3}{2}, \quad \sin\theta = \frac{y-1}{2}$
	Using $\cos^2\theta + \sin^2\theta = 1$:
	$\left(\frac{x+3}{2}\right)^2 + \left(\frac{y-1}{2}\right)^2 = 1$
	$(x+3)^2 + (y-1)^2 = 4$
7	General term: ${}^{12}C_r\,(2x)^r\,k^{12-r}$
	x^8 term ($r = 8$): ${}^{12}C_8\,(2x)^8\,k^4$
	${}^{12}C_8 = 495, \quad 2^8 = 256$
	$495 \times 256 \times k^4 = 79\,200\,000$
	$126\,720\,k^4 = 79\,200\,000$
	$k^4 = 625$
	$k = 5$ (k is a positive integer)

Question number	
8 (a)	No. Only plants on the front row could be chosen, so not every possible sample has an equal probability of being selected.
8 (b)	From the graph: 18 plants are shorter than 40 cm and 110 plants are shorter than 80 cm. Number between 40 cm and 80 cm = 110 − 18 = 92 Percentage = $\frac{92}{120} \times 100\% = 76.7\%$ 76.7% > 75%, so the data supports the owner's statement.
8 (c)	No. Different samples give different results: another sample (e.g. a different row) might contain more of the tallest or shortest plants, so it would not necessarily support the statement.
9 (a)	A sample, since not all of the data is used (the data for 2009 is missing).
9 (b)	$I = 1115 \times 2009 - 2\,212\,950 = 27\,085$ Estimate: £27 085 (about £27 100)
9 (c)	1. It is interpolation: 2009 lies within the range of the data (between 2008 and 2010). 2. The points lie close to a straight line, so a linear model appears to be justified.
9 (d)	No, not reasonable. Median income varies considerably across different areas of London (Croydon's median income is lower than Camden's), so the Camden model would not suit Croydon.

Question number

10

$u = 4x, \quad dv/dx = \cos 2x \quad \Rightarrow \quad du/dx = 4, \quad v = \frac{1}{2} \sin 2x$

$\int 4x \cos 2x \, dx = 4x \times \frac{1}{2} \sin 2x - \int \frac{1}{2} \sin 2x \times 4 \, dx$

$= 2x \sin 2x - \int 2 \sin 2x \, dx$

$= 2x \sin 2x + \cos 2x$

$$\int_0^{\pi/4} 4x \cos 2x \, dx = \left[2x \sin 2x + \cos 2x \right]_0^{\pi/4}$$

$$= \left(2 \times \frac{\pi}{4} \times \sin \frac{\pi}{2} + \cos \frac{\pi}{2} \right) - (0 + \cos 0)$$

$$= \frac{\pi}{2} + 0 - 1 = \frac{\pi}{2} - 1$$

11 (a)

$$\frac{dy}{dx} \propto \sqrt{x} \quad \Rightarrow \quad \frac{dy}{dx} = k\sqrt{x}$$

$$x = 4, \frac{dy}{dx} = 3: \quad 3 = k\sqrt{4} = 2k \quad \Rightarrow \quad k = \frac{3}{2}$$

$$\frac{dy}{dx} = \frac{3}{2}\sqrt{x}$$

11 (b)

$$y = \int \frac{3}{2} x^{1/2} \, dx = \frac{3}{2} \times \frac{x^{3/2}}{3/2} + c = x^{3/2} + c$$

$x = 4, y = 10: \quad 10 = 4^{3/2} + c = 8 + c \quad \Rightarrow \quad c = 2$

$y = x^{3/2} + 2$

12 (a)

Each value of x maps to two values ($\pm 1/\sqrt{x}$), so it is a one-to-many mapping, and a function cannot be one-to-many.

12 (b)

$$gh(x) = g(x^2 + 2) = \frac{x^2 + 2}{x^2 + 2 - 3}$$

Question number

$gh(x) = \dfrac{x^2 + 2}{x^2 - 1}$

12 (c) Need $h(x) > 3$: $x^2 + 2 > 3 \Rightarrow x^2 > 1$

Domain: $|x| > 1$ $(x < -1$ or $x > 1)$

13 (a) H_0: $\mu = 0.14$ H_1: $\mu < 0.14$

where μ is the population mean mass (in kg) of this variety of apple.

13 (b) $\bar{X} \sim N\left(0.14, \dfrac{0.0199^2}{80}\right)$

i.e. $\bar{X} \sim N(0.14, 4.95 \times 10^{-6})$

13 (c) 1-tailed test at 5%: $z = -1.645$

$c = 0.14 - 1.645 \times \dfrac{0.0199}{\sqrt{80}} = 0.13634...$

Critical region: $\bar{X} < 0.136$

13 (d) Sample mean $\bar{x} = 0.1316 < 0.136$, so it lies in the critical region.

Reject H_0.

There is sufficient evidence at the 5% level to suggest that the mean mass of these apples is less than 0.14 kg.

14 (a) Remove City of London (and any other area with #N/A), since its percentage value is missing.

14 (b) 1. The scatter diagram does not look linear (the points do not lie close to a straight line).

2. The pmcc of 0.3743 is not close to 1 (only weak correlation).

Question number

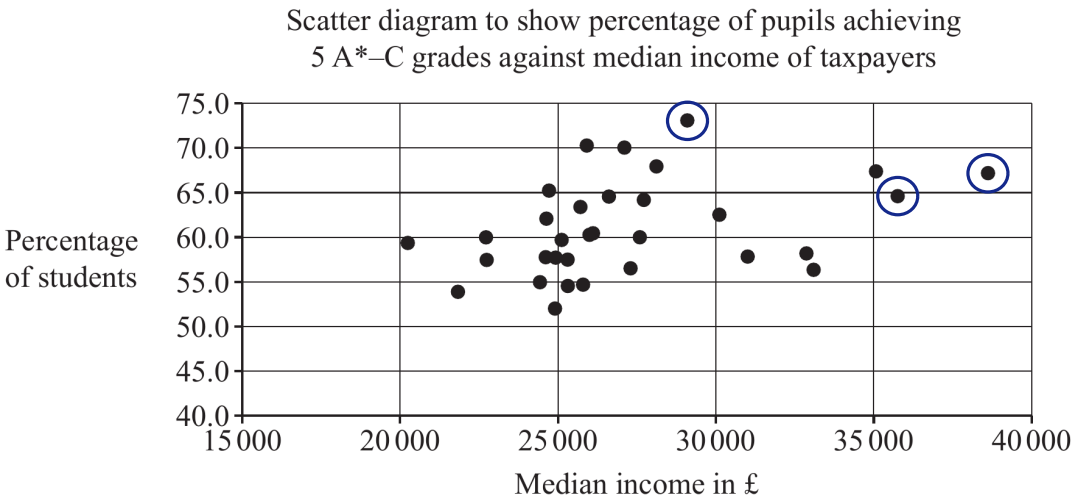
14 (c) Outliers: more than 2 standard deviations from the mean.

Income: $27\,216 \pm 2 \times 4177.5 = 18\,861$ and $35\,571$

Outliers: median income < £18 861 or median income > £35 571

Percentage: $61.0 \pm 2 \times 5.32 = 50.36$ and 71.64

Outliers: percentage < 50.36 or percentage > 71.64



14 (d) Between 0 and 0.3743: the outliers gave a false impression of linearity; without them the scatter is more like a circle with no clear trend, so the pmcc will be smaller.

15 $y = 1$: $\ln 1 + x^3 \times 1 = 8 \Rightarrow x^3 = 8 \Rightarrow x = 2$

Differentiate implicitly with respect to x:

$$\frac{1}{y}\frac{dy}{dx} + x^3\frac{dy}{dx} + 3x^2y = 0$$

At (2, 1): $\frac{dy}{dx} + 8\frac{dy}{dx} + 12 = 0 \Rightarrow 9\frac{dy}{dx} = -12 \Rightarrow \frac{dy}{dx} = -\frac{4}{3}$

Gradient of normal $= \frac{3}{4}$

$y - 1 = \frac{3}{4}(x - 2)$

Question number

$4y - 4 = 3x - 6$

$3x - 4y - 2 = 0$

16 (a) $P(\text{one smokes, other doesn't}) = 2 \times 0.16 \times 0.84$

$= 0.2688$

16 (b) S = smokes cigarettes, D = drinks alcohol

$P(D \cap S) = P(D) - P(D \cap S') = 0.75 - 0.66 = 0.09$

$P(S \cap D') = P(S) - P(S \cap D) = 0.16 - 0.09 = 0.07$

16 (c) $P(S \mid D) = \dfrac{P(S \cap D)}{P(D)} = \dfrac{0.09}{0.75} = 0.12$

16 (d) $P(S \mid D) = 0.12 \neq P(S) = 0.16$

so the events are not independent.

17 $2 \sin x + \sec x = 4 \cos x \quad (\cos x \neq 0)$

Divide by $\cos x$: $2 \tan x + \sec^2 x = 4$

$2 \tan x + 1 + \tan^2 x = 4$

$\tan^2 x + 2 \tan x - 3 = 0$

$(\tan x + 3)(\tan x - 1) = 0$

$\tan x = 1 \quad \text{or} \quad \tan x = -3$

$\tan x = 1$: $x = \dfrac{\pi}{4}$ or $x = \dfrac{\pi}{4} - \pi = -\dfrac{3\pi}{4}$

$\tan x = -3$: $x = -1.2490\ldots$ or $x = -1.2490\ldots + \pi = 1.8925\ldots$

$x = -\dfrac{3\pi}{4}, \ -1.25, \ \dfrac{\pi}{4}, \ 1.89$ (3 s.f.)

18 (a) $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{31164.7}{120} = 259.705\ldots = 260$ (3 s.f.)

Question number

18 (b)

Frequency = frequency density × class width

$= 6.2 \times 5 = 31$ days

18 (c)

1. The histogram is roughly symmetrical about the mean.

2. It is unimodal and roughly bell-shaped.

18 (d)

$S_{xx} = \Sigma x^2 - \frac{(\Sigma x)^2}{n} = 8\,101\,050.91 - \frac{31164.7^2}{120} = 7396.53$

$s^2 = \frac{7396.53}{119} = 62.155\ldots = 62.2$ (3 s.f.)

$X \sim N(260, 62.2)$

$P(X < 255) = 0.26304\ldots = 0.263$ (3 s.f.)

18 (e)

Y = number of days using less than 255 litres, $Y \sim B(28, 0.263)$

$P(Y \geq 5) = 1 - P(Y \leq 4) = 1 - 0.10498\ldots$

$= 0.895$ (3 s.f.)

18 (f)

Daily charge C pence $= 7.8 + 0.18X$

Mean $= 7.8 + 0.18 \times 260 = 54.6$

Variance $= 0.18^2 \times 62.2 = 2.01528 = 2.02$ (3 s.f.)

$C \sim N(54.6, 2.02)$ (C in pence)