Maths with Matt

Physics and maths resources

OCR A Level Mathematics B (MEI)

H640/02 - Pure Mathematics and Statistics

Tuesday 13 June 2023

Questions and Fully Worked Model Solutions

This booklet places the **question and the solution together**. The mathematics is set out using proper fractions, powers, roots and exact notation, with extra written explanation so that a student can see *why* each step is being taken.

Solutions have been checked against the OCR Tuesday 13 June 2023 mark scheme. The aim is a clear **full-mark model response**, not merely a list of final answers.

Question wording is reproduced from the user-provided 2023 paper for study/reference. OCR retains copyright in the original examination material.

Contents

Question 1 2

Question 2 3

Question 3 4

Question 4 5

Question 5 6

Question 6 7

Question 7 8

Question 8 9

Question 9 11

Question 10 13

Question 11 14

Question 12 15

Question 13 16

Question 14 17

Question 15 19

Question 16 20

Question 17 21

Question 18 22

www.mathswithmatt.co.uk

Question 1

Question 1

Determine the sum of the infinite geometric series $9 - 3 + 1 - \frac{1}{3} + \frac{1}{9} + \ldots$. **(3)**

Model solution

This is geometric because each term is obtained by multiplying by the same ratio, $r = -\frac{1}{3}$. Since $|r| < 1$, an infinite sum exists.

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

$\frac{27}{4}$ *(3)*

www.mathswithmatt.co.uk

Question 2

Question 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)**

Model solution

(a) Completing the square puts the equation into centre--radius form.

$$(x-6)^2 - 36 + (y+4)^2 - 16 + 3 = 0, \qquad (x-6)^2 + (y+4)^2 = 49.$$

The radius is 7. *(2)*

(b) Comparing with $(x-a)^2 + (y-b)^2 = r^2$*, the centre is read directly from the brackets.*

Centre $(6, -4)$. *(1)*

www.mathswithmatt.co.uk

Question 3

Question 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)**

Model solution

A negative power means take the reciprocal; the power $\frac{3}{2}$ means square-root then cube.

$$\left(\frac{64}{49}\right)^{-\frac{3}{2}} = \left(\frac{49}{64}\right)^{\frac{3}{2}} = \left(\frac{7}{8}\right)^{3} = \frac{343}{512}.$$

$m = 343$ ***and*** $n = 512$. ***(3)***

www.mathswithmatt.co.uk

Question 4

Question 4

A biased octagonal dice has faces numbered from 1 to 8. The discrete random variable X is the score when it is rolled once.

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

(a) Determine p. **(2)**
(b) Find $P(X \geq 4)$. **(1)**
The dice is rolled 30 times. (c) Determine the probability that a score of 8 occurs exactly twice. **(2)**

Model solution

(a) The probabilities in a complete distribution add to 1.

$$7p + 3p = 1 \quad \Rightarrow \quad p = 0.1.$$

$p = 0.1$. *(2)*

(b) Scores $4, 5, 6, 7$ *each have probability* p*, and score* 8 *has probability* $3p$.

$$P(X \geq 4) = 4p + 3p = 0.7.$$

0.7 *(1)*

(c) Each roll is an independent trial with $P(8) = 3p = 0.3$*, so a binomial model is appropriate.*

$$P(X = 2) = \binom{30}{2}(0.3)^2(0.7)^{28} = 0.001817ldots$$

0.0018 *(approximately).* *(2)*

Question 5

Question 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)**

Model solution

Subtract the position vectors to obtain the displacement from A to B, then use Pythagoras on its components.

$$\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA} = \begin{pmatrix} 2 \\ -2 \end{pmatrix}, \qquad AB = \sqrt{2^2 + (-2)^2} = \sqrt{8}.$$

$AB = 2\sqrt{2}$. *(3)*

www.mathswithmatt.co.uk

Question 6

Question 6

The parametric equations of a circle are $x = 2\cos\theta - 3$ and $y = 2\sin\theta + 1$. Determine its cartesian equation in the form $(x-a)^2 + (y-b)^2 = k$, where a, b, k are integers. **(4)**

Model solution

Isolate $\sin\theta$ ***and*** $\cos\theta$***, then use*** $\sin^2\theta + \cos^2\theta = 1$ ***to eliminate the parameter.***

$$\cos\theta = \frac{x+3}{2}, \qquad \sin\theta = \frac{y-1}{2},$$

$$\left(\frac{x+3}{2}\right)^2 + \left(\frac{y-1}{2}\right)^2 = 1.$$

$(x+3)^2 + (y-1)^2 = 4.$ ***(4)***

www.mathswithmatt.co.uk

Question 7

Question 7

The coefficient of x^8 in $(2x + k)^{12}$, where k is a positive integer, is $79\,200\,000$. Determine k. **(4)**

Model solution

For the x^8 term, choose $2x$ eight times and k four times in the binomial expansion.

$$\binom{12}{8}(2x)^8k^4 = 79\,200\,000x^8,$$

$$495 \times 256 \times k^4 = 79\,200\,000 \quad \Rightarrow \quad k^4 = 625.$$

$k = 5$, *since k is positive.* ***(4)***

www.mathswithmatt.co.uk

Question 8

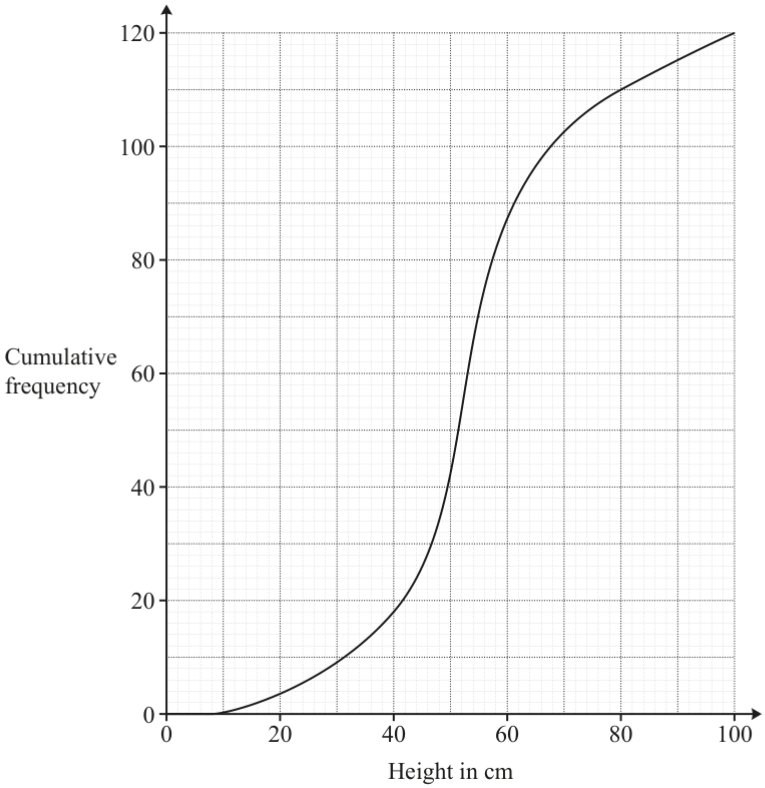
Question 8

A garden centre displays coniferous hedging plants in 10 rows of 120 plants. An employee records the height of every plant in the front row.

(a) Explain whether these data form a simple random sample. **(1)**

The cumulative frequency curve for the 120 recorded heights is shown below.

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.

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

(b) Show that the data support this statement. **(4)**

(c) Explain whether all samples of 120 plants would necessarily support the statement. **(1)**

Model solution

(a) A simple random sample gives every possible sample an equal chance of selection. Restricting selection to the front row does not do this.

No: not every plant, and hence not every possible sample, has an equal probability of selection. ***(1)***

(b) Read cumulative frequencies from the curve. About 18 plants are below 40 cm and about 110 are below 80 cm, so the difference gives the number in the stated interval.

$$110 - 18 = 92, \qquad \frac{92}{120} \times 100\% = 76.7\%.$$

About $76.7\% > 75\%$, so the data support the owner's statement. ***(4)***

(c) A different sample can have different heights, so its proportion in the interval can differ.

No; another sample of 120 need not give the same result and may not support the statement. *(1)*

www.mathswithmatt.co.uk

Question 9

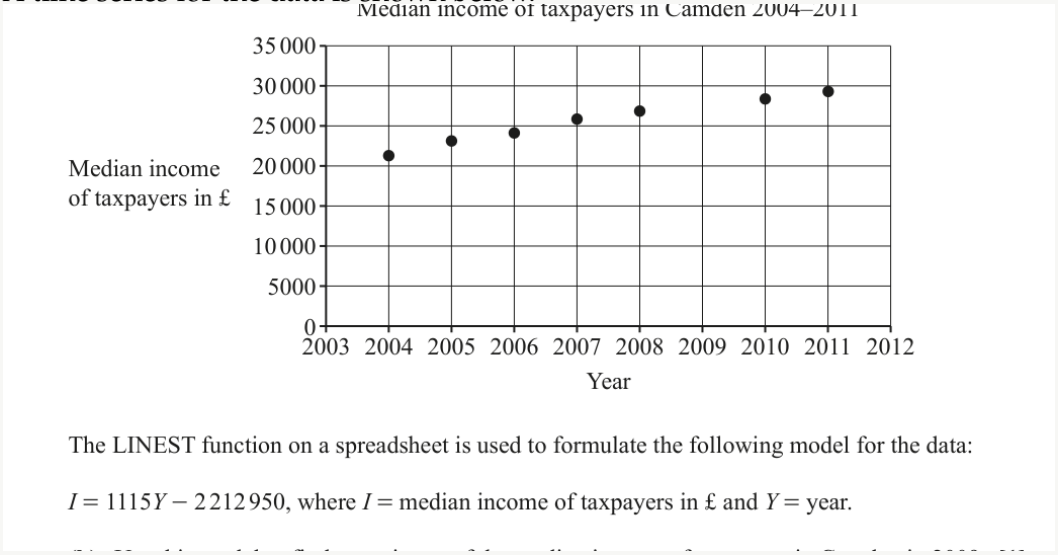
Question 9

The pre-release material concerns median income of taxpayers in different areas of London. For Camden, the following data are given; the year 2004 refers to financial year 2003/04.

Year	2004	2005	2006	2007	2008	2009	2010	2011
Median income in £	21300	23200	24200	25900	26900	#N/A	28400	29400

(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.

The LINEST function on a spreadsheet is used to formulate the model $I = 1115Y - 2\,212\,950$, where I is median income in pounds and Y is year.

(b) Estimate the median income in Camden in 2009. **(1)**

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

The Croydon value for 2009 is also unavailable. (d) Explain whether the model in part (b) would reasonably estimate it. **(1)**

Model solution

(a) The table contains data for only selected years, rather than all Camden taxpayers or all relevant data.

It is a sample. *(1)*

(b) Substitute $Y = 2009$ into the fitted line.

$$I = 1115(2009) - 2\,212\,950 = 27\,085.$$

Estimated median income: $\$27\,085$ (about $\$27\,100$). *(1)*

(c) The estimate lies between known years, and the plotted values follow an approximately straight increasing trend.

It is interpolation between 2008 and 2010, and a linear model appears reasonable for Camden's time series. *(2)*

(d) A model fitted to one borough's income level and trend should not be assumed to apply to another borough.

No: median incomes, and their changes over time, can differ substantially between Croydon and Camden. *(1)*

www.mathswithmatt.co.uk

Question 10

Question 10

Determine the exact value of $\int_0^{\frac{\pi}{4}} 4x\cos(2x)\,dx$. **(5)**

Model solution

The integrand is a product of an algebraic factor and a trigonometric factor, so integration by parts reduces the power of x. Take $u = 4x$ and $dv = \cos(2x)\,dx$.

$$\int 4x\cos(2x)\,dx = 2x\sin(2x) - \int 2\sin(2x)\,dx = 2x\sin(2x) + \cos(2x).$$

$$[2x\sin(2x) + \cos(2x)]_0^{\frac{\pi}{4}} = \frac{\pi}{2} - 1.$$

$\frac{\pi - 2}{2}$. *(5)*

www.mathswithmatt.co.uk

Question 11

Question 11

In this question you must show detailed reasoning. The variables x and y satisfy $\frac{dy}{dx}$ directly proportional to $\sqrt{x}$. 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)**

Model solution

(a) "Directly proportional" gives a constant multiple of $\sqrt{x}$. Use the given gradient to determine the constant.

$$\frac{dy}{dx} = k\sqrt{x}, \qquad 3 = k\sqrt{4} = 2k, \qquad k = \frac{3}{2}.$$

$\frac{dy}{dx} = \frac{3}{2}\sqrt{x}$. *(3)*

(b) Integrate the gradient, then use the point on the curve to find the integration constant.

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

$$10 = 4^{\frac{3}{2}} + c = 8 + c \quad \Rightarrow \quad c = 2.$$

$y = x^{\frac{3}{2}} + 2$. *(3)*

www.mathswithmatt.co.uk

Question 12

Question 12

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

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

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

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

Model solution

(a) The $\pm$ means that every permitted input has two outputs.

It is a one-to-many mapping, so it is not a function. *(1)*

(b) Composition $gh(x)$ means substitute $h(x)$ into g.

$$gh(x) = g(x^2 + 2) = \frac{x^2 + 2}{x^2 + 2 - 3} = \frac{x^2 + 2}{x^2 - 1}.$$

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

(c) The input to g must be greater than 3.

$$x^2 + 2 > 3 \quad \Rightarrow \quad x^2 > 1.$$

$x < -1$ ***or*** $x > 1$. *(1)*

www.maths

Question 13

Question 13

A supermarket advertises mean apple mass 0.14 kg. After a poor growing season, quality control believes the mean is less than 0.14 kg and tests this at the 5% level. A random sample gives $n = 80$, mean 0.1316, $\sigma = 0.0198$, and $s = 0.0199$ (kg).

(a) State the null and alternative hypotheses, defining the parameter. **(2)**

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

(c) Determine the critical region. **(2)**

(d) Carry out the test, giving a conclusion in context. **(3)**

Model solution

(a) Let μ be the population mean mass, in kg, of this variety of apple sold by the chain. The stated suspicion makes this a lower-tailed test.

$H_0 : \mu = 0.14$, $H_1 : \mu < 0.14$. *(2)*

(b) Under H_0, use the sample standard deviation as the estimate of population standard deviation.

$$\overline{X} \sim N\left(0.14, \frac{0.0199^2}{80}\right).$$

$\overline{X} \sim N\left(0.14, \dfrac{0.0199^2}{80}\right)$. *(2)*

(c) The lower 5% critical value is 1.645 standard errors below the null mean.

$$0.14 - 1.645\left(\frac{0.0199}{\sqrt{80}}\right) = 0.1363\ldots$$

Critical region: $\overline{X} < 0.136$ kg (approximately). *(2)*

(d) Compare the observed sample mean with this lower-tail critical region.

$$0.1316 < 0.136.$$

Reject H_0. There is sufficient evidence at the 5% level to suggest that the mean apple mass is less than 0.14 kg. *(3)*

Question 14

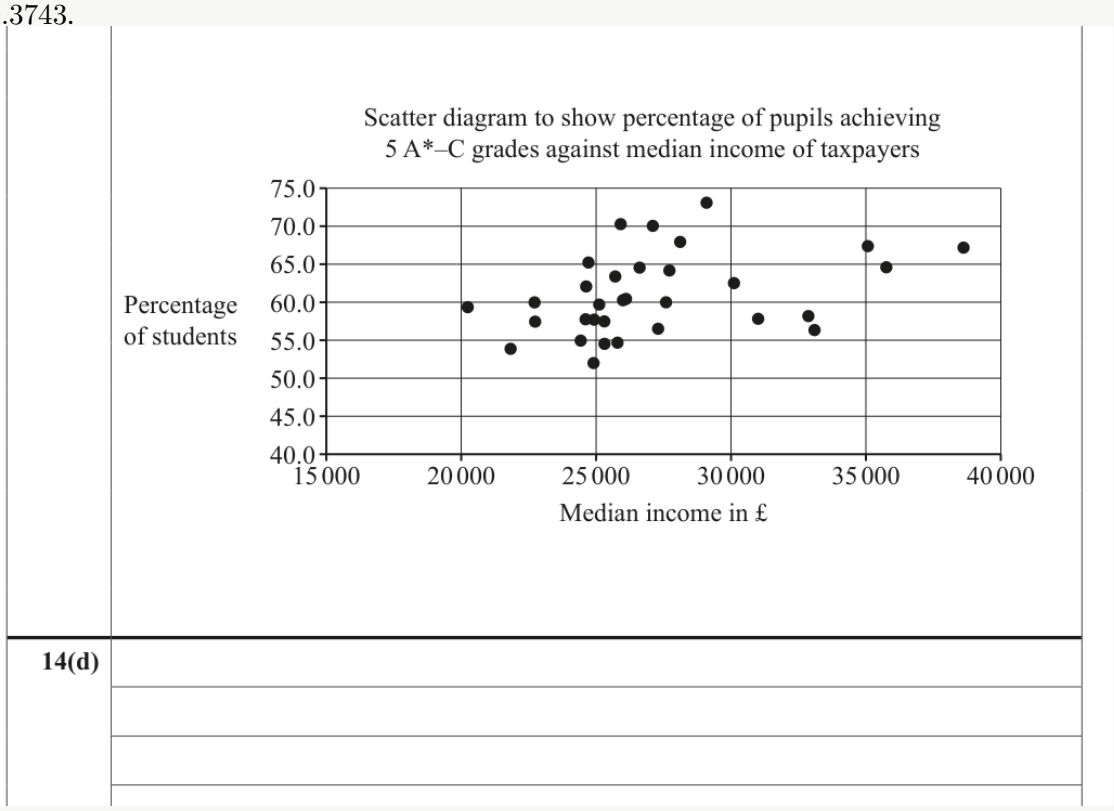
Question 14

The pre-release material gives median income of taxpayers and the percentage of pupils at the end of KS4 achieving 5 or more GCSEs at grade A*–C, including English and Maths, for London areas. For 2014/15, the data include:

	Median income in £	Percentage achieving 5+ A*–C
City of London	61100	#N/A
Barking and Dagenham	21800	54.0
Barnet	27100	70.1
Bexley	24400	55.0
Brent	22700	60.0
Bromley	28100	68.0

(a) With reference to these data, explain how they should be cleaned before analysis. **(1)**

After cleaning, the scatter diagram is shown below. The product moment correlation coefficient is 0.3743.



14(d)

(b) Give two reasons why a linear model may not be appropriate. **(2)**

Further analysis gives mean median income 27216, standard deviation 4177.5, mean percentage 61.0, and standard deviation 5.32. Three outliers were identified.

(c) Determine the outlier ranges for both variables and identify the three outliers on the scatter diagram. **(4)**

(d) The outliers are removed and the correlation coefficient recalculated. Explain whether the new value is between 0.3743 and 1, or between 0 and 0.3743. **(1)**

Model solution

(a) A row with a missing value cannot supply a paired observation for the scatter diagram or correlation calculation.

Discard City of London, because its percentage value is unavailable. **(1)**

(b) A linear model requires a roughly linear pattern and a correlation close to 1 or -1.

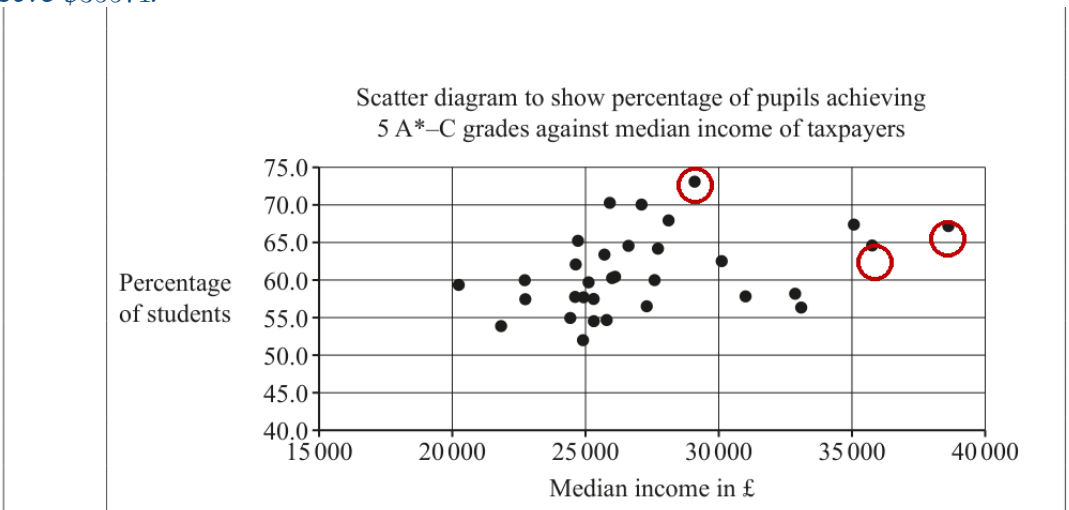
The scatter does not look linear, and 0.3743 is not close to 1. **(2)**

(c) Use the two-standard-deviation rule: values more than two standard deviations from their mean are outliers.

$$27216 \pm 2(4177.5) = 18861 \text{ and } 35571,$$

$$61.0 \pm 2(5.32) = 50.36 \text{ and } 71.64.$$

Thus the three points to circle are the point above 71.64%, and the two points with median income above $\$35571$.



14(d)

Income outliers: $m < \$18861$ or $m > \$35571$. Percentage outliers: $p < 50.36$ or $p > 71.64$. **(4)**

(d) Those extreme points make the cloud appear more linear than it is; removing them leaves a more curved/scattered pattern.

The new coefficient is between 0 and 0.3743. **(1)**

Question 15

Question 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 where $y = 1$, in the form $ax + by + c = 0$. **(7)**

Model solution

First find the point. Then implicit differentiation is needed because y is defined only through an equation involving both x and y.

$$\ln 1 + x^3(1) = 8 \quad \Rightarrow \quad x = 2.$$

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

At $(2, 1)$, this gives the tangent gradient.

$$(1+8)\frac{dy}{dx} + 12 = 0 \quad \Rightarrow \quad \frac{dy}{dx} = -\frac{4}{3}.$$

The normal gradient is the negative reciprocal, $\frac{3}{4}$.

$$y - 1 = \frac{3}{4}(x - 2) \quad \Rightarrow \quad 3x - 4y - 2 = 0.$$

$3x - 4y - 2 = 0.$ ***(7)***

www.mathswithmatt.co.uk

Question 16

Question 16

Research shows that 16% of young adults smoke cigarettes. Two young adults are selected at random.
(a) Determine the probability that one smokes and the other does not. **(2)**
Research also shows that 75% drink alcohol and 66% drink alcohol but do not smoke.
(b) Determine the probability that a randomly selected young adult smokes but does not drink alcohol. **(2)**
(c) A young adult who drinks alcohol is selected. Determine the probability that this adult smokes. **(2)**
(d) Using part (c), explain whether smoking and drinking alcohol are independent events. **(2)**

Model solution

(a) There are two mutually exclusive orders: smoker then non-smoker, or non-smoker then smoker.

$$2(0.16)(0.84) = 0.2688.$$

0.2688. *(2)*

(b) Of the 75% who drink, 66% do not smoke, leaving the intersection of drinking and smoking. Subtract this from the total smoking probability.

$$P(S \cap D) = 0.75 - 0.66 = 0.09, \qquad P(S \cap D') = 0.16 - 0.09 = 0.07.$$

0.07. *(2)*

(c) Use conditional probability, restricting attention to drinkers.

$$P(S \mid D) = \frac{P(S \cap D)}{P(D)} = \frac{0.09}{0.75} = 0.12.$$

0.12. *(2)*

(d) Independence would require the conditional smoking probability among drinkers to equal the overall smoking probability.

They are not independent, because $P(S \mid D) = 0.12 \neq 0.16 = P(S)$. *(2)*

Question 17

Question 17

In this question you must show detailed reasoning. Solve $2\sin x + \sec x = 4\cos x$, where $-\pi < x < \pi$. **(6)**

Model solution

Because** $\sec x = 1/\cos x$, **divide by** $\cos x$ **to express the equation in** $\tan x$. **Values with** $\cos x = 0$ **are not in the original equation.

$$2\tan x + \sec^2 x = 4.$$

Use** $\sec^2 x = 1 + \tan^2 x$ **to form a quadratic.

$$\tan^2 x + 2\tan x - 3 = 0 \quad \Rightarrow \quad (\tan x - 1)(\tan x + 3) = 0.$$

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

Give every solution in the stated interval.

$$x = \frac{\pi}{4},\ -\frac{3\pi}{4},\ -1.249\ldots,\ 1.893\ldots$$

$$x = -\frac{3\pi}{4},\ -1.249\ldots,\ \frac{\pi}{4},\ 1.893\ldots. \qquad \textit{(6)}$$

www.mathswithmatt.co.uk

Question 18

Question 18

Riley records daily household water consumption x, in litres, for a random sample of 120 days from the previous twelve months. Summary statistics are $\sum x = 31164.7$, $\sum x^2 = 8101050.91$, and $n = 120$.

(a) Calculate the sample mean, correct to 3 significant figures. **(1)**

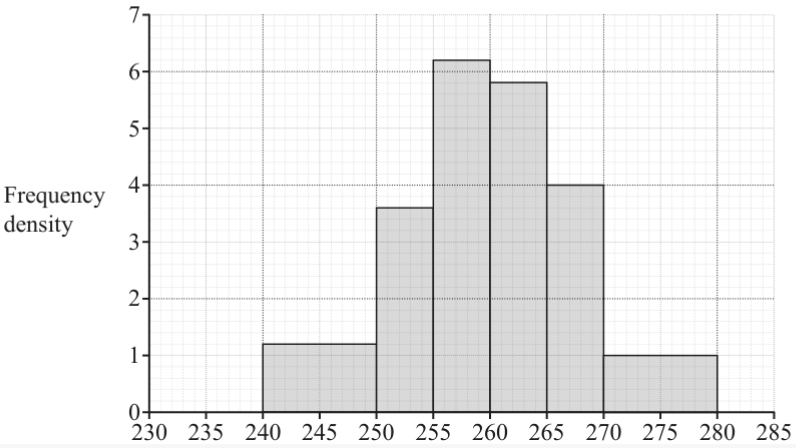
The histogram is shown below.

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 for which consumption was between 255 and 260 litres. **(1)**

(c) Give two reasons why a Normal distribution may be appropriate for daily consumption. **(2)**

Riley uses the sample mean and sample variance, each correct to 3 significant figures, as parameters of a Normal model.

(d) Use the model to calculate $P(x < 255)$. **(2)**

(e) Calculate the probability that consumption is below 255 litres on at least 5 days in a random sample of 28 days. **(2)**

Water charges are a standing charge of 7.8 pence per day plus 0.18 pence per litre.

(f) Adapt Riley's model to model daily water charges. **(3)**

Model solution

(a) Divide the total by the number of days.

$$\bar{x} = \frac{31164.7}{120} = 259.7058\ldots$$

260 *litres (3 significant figures).* *(1)*

(b) Frequency is the area of the histogram bar. The bar from 255 to 260 has frequency density 6.2.

$$5 \times 6.2 = 31.$$

31 *days.* *(1)*

(c) The shape and type of the variable should be checked before using a Normal model.

The data are continuous and the histogram is approximately symmetric and bell-shaped. **(2)**

(d) First calculate the unbiased sample variance, then standardise using Riley's Normal model.

$$s^2 = \frac{8101050.91 - \frac{31164.7^2}{120}}{119} = 62.155\ldots \approx 62.2.$$

$$X \sim N(260, 62.2), \qquad P(X < 255) = \Phi\left(\frac{255 - 260}{\sqrt{62.2}}\right) = 0.263\ldots$$

$P(X < 255) = 0.263$ *(approximately).* **(2)**

(e) For each day, "below 255 litres" has probability $p = 0.263$. Assuming the sampled days are independent, the count is binomial.

$$Y \sim B(28, 0.263), \qquad P(Y \geq 5) = 1 - \sum_{r=0}^{4} \binom{28}{r} (0.263)^r (0.737)^{28-r}.$$

$P(Y \geq 5) = 0.892$ *(approximately).* **(2)**

(f) If C is the daily charge in pence, then $C = 7.8 + 0.18X$. A linear transformation of a Normal variable is Normal: transform its mean and variance.

$$E(C) = 7.8 + 0.18(260) = 54.6, \qquad \mathrm{Var}(C) = 0.18^2(62.2) = 2.01\ldots$$

$C \sim N(54.6, 2.01$ *approximately*$)$, *with charges measured in pence.* **(3)**