OCR
Oxford Cambridge and RSA

Wednesday 14 October 2020 – Afternoon

A Level Mathematics B (MEI)

H640/02 Pure Mathematics and Statistics

Time allowed: 2 hours

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

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**. You can use extra paper if you need to, but you must clearly show your candidate number, the centre number and the question numbers.
- 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 **20** pages.

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

Formulae A Level Mathematics B (MEI) (H640)

Arithmetic series

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

Geometric series

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

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

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

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

Differentiation

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

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

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

Integration

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

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

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

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

Trigonometric identities

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

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

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

Numerical methods

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

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

Probability

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

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

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

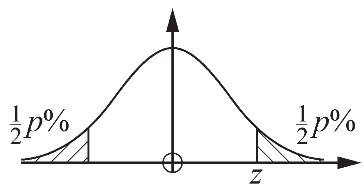
Mean of X is np

Hypothesis testing for the mean of a Normal distribution

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

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

$v = u + at$

$s = ut + \frac{1}{2}at^2$

$s = \frac{1}{2}(u + v)t$

$v^2 = u^2 + 2as$

$s = vt - \frac{1}{2}at^2$

Motion in two dimensions

$\mathbf{v} = \mathbf{u} + \mathbf{a}t$

$\mathbf{s} = \mathbf{u}t + \frac{1}{2}\mathbf{a}t^2$

$\mathbf{s} = \frac{1}{2}(\mathbf{u} + \mathbf{v})t$

$\mathbf{s} = \mathbf{v}t - \frac{1}{2}\mathbf{a}t^2$

Answer **all** the questions.

Section A (23 marks)

1 Fig. 1 shows triangle ABC.

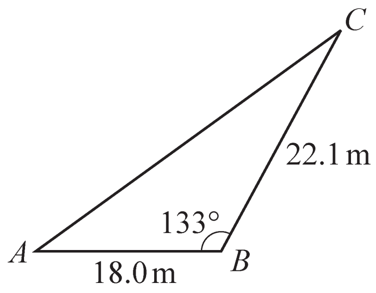


Fig. 1

Calculate the area of triangle ABC, giving your answer correct to 3 significant figures. **[2]**

2 Fig. 2 shows a sector of a circle of radius 8 cm.

The angle of the sector is 2.1 radians.

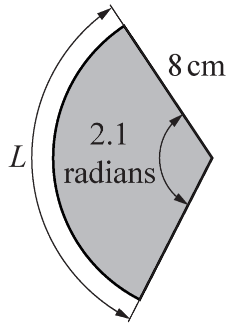


Fig. 2

(a) Calculate the length of the arc L. **[1]**

(b) Calculate the area of the sector. **[2]**

3 You are given that $y = 4x + \sin 8x$.

(a) Find $\frac{dy}{dx}$.

(b) Find the smallest positive value of x for which $\frac{dy}{dx} = 0$, giving your answe

4 Fig. 4 shows a cumulative frequency diagram for the time spent revising mathematics by year 11 students at a certain school during a week in the summer term.

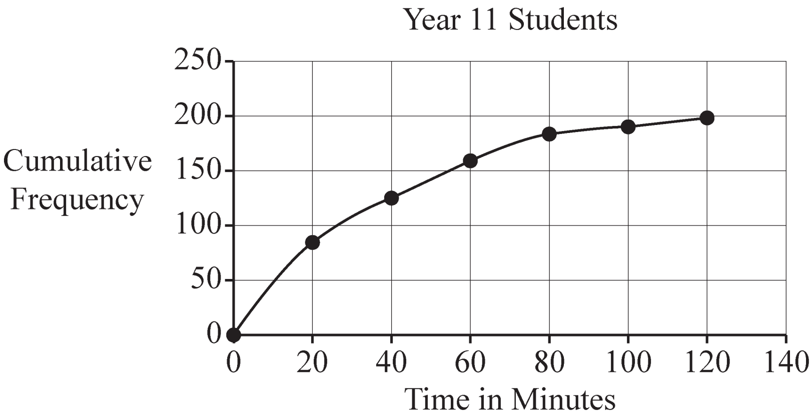


Fig. 4

(a) Use the diagram to estimate the median time spent revising mathematics by these students. **[1]**

A teacher comments that 90% of the students spent less than an hour revising mathematics during this week.

(b) Determine whether the information in the diagram supports this comment. **[1]**

5 The first n terms of an arithmetic series are

$17 + 28 + 39 + \ldots + 281 + 292.$

(a) Find the value of n. **[1]**

(b) Find the sum of these n terms. **[2]**

6 **(a)** Find the first three terms in ascending powers of x of the binomial expansio

(b) State the range of values of x for which this expansion is valid.

7 You are given that $\mathrm{P}(A) = 0.6$, $\mathrm{P}(B) = 0.5$ and $\mathrm{P}(A \cup B)' = 0.2$.

(a) Find $\mathrm{P}(A \cap B)$. **[2]**

(b) Find $\mathrm{P}(A \mid B)$. **[2]**

(c) State, with a reason, whether A and B are independent. **[1]**

BLANK PAGE

Answer **all** the questions.

Section B (77 marks)

8 Rosella is carrying out an investigation into the age at which adults retire fro
where she lives. She collects a sample of size 50, ensuring this comprises of 25
retired men and 25 randomly selected retired women.

(a) State the name of the sampling method she uses. **[1]**

Fig. 8.1 shows the data she obtains in a frequency table and Fig. 8.2 shows these data displayed in a histogram.

Age in years at retirement	45 –	50 –	55 –	60 –	65 –	70 –	75 – 80
Frequency density	0.4	1.8	2.4	2.2	1.8	1.2	0.2

Fig. 8.1

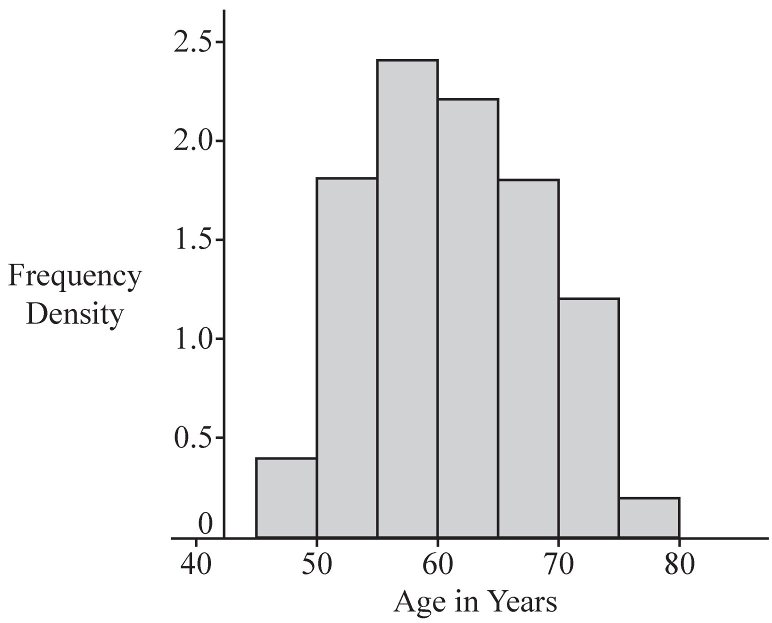


Fig. 8.2

(b) How many people in the sample are aged between 50 and 55? **[1]**

Rosella obtains a list of the names of all 4960 people who have retired in the city during the previous month.

(c) Describe how Rosella could collect a sample of size 200 from her list using

- systematic sampling such that every item on the list could be selected,
- simple random sampling. **[4]**

Rosella collects two simple random samples, one of size 200 and one of size 50
histograms in Fig. 8.3 show the data from the sample of size 200 on the left an
sample of 500 on the right.

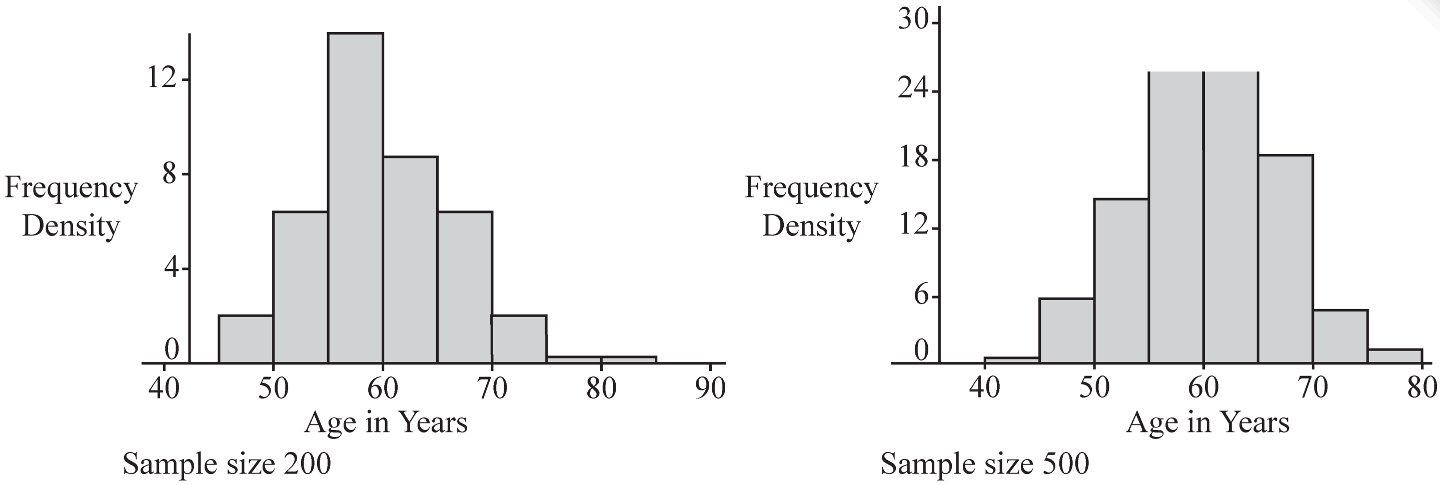


Fig. 8.3

(d) With reference to the histograms shown in Fig. 8.2 and Fig. 8.3, explain why it appears reasonable to model the age of retirement in this city using the Normal distribution. **[1]**

Summary statistics for the sample of 500 are shown in Fig. 8.4.

Statistics	
n	500
Mean	60.0515
σ	6.5717
s	6.5783
Σx	30025.7601
Σx²	1824686.322
Min	36.0793
Q1	55.2573
Median	59.9202
Q3	64.4239
Max	81.742

Fig. 8.4

(e) Use an appropriate Normal model based on the information in Fig. 8.4 to estimate the number of people aged over 65 who retired in the city in the previous month. **[4]**

(f) Identify a limitation in using this model to predict the number of people aged over 65 retiring in the **following** month. **[1]**

9 A company supplies computers to businesses. In the past the company has fou are kept by businesses for a mean time of 5 years before being replaced. Claud, company, thinks that the mean time before replacing computers is now different

(a) Describe how Claud could obtain a cluster sample of 120 computers usec company supplies.

Claud decides to conduct a hypothesis test at the 5% level to test whether there is evidence to suggest that the mean time that businesses keep computers is not 5 years. He takes a random sample of 120 computers. Summary statistics for the length of time computers in this sample are kept are shown in Fig. 9.

Statistics	
n	120
Mean	4.8855
σ	2.6941
s	2.7054
Σx	586.2566
Σx²	3735.1475
Min	0.1213
Q1	2.5472
Median	4.8692
Q3	7.0349
Max	9.9856

Fig. 9

(b) **In this question you must show detailed reasoning.**

- State the hypotheses for this test, explaining why the alternative hypothesis takes the form it does.
- Use a suitable distribution to carry out the test. **[8]**

10 In this question you must show detailed reasoning.

The equation of a curve is

$$y = \frac{\sin 2x - x}{x \sin x}.$$

(a) Use the small angle approximation given in the list of formulae on pages 2–3 of this question paper to show that

$$\int_{0.01}^{0.05} y \, dx \approx \ln 5.$$ **[4]**

(b) Use the same small angle approximation to show that

$$\frac{dy}{dx} \approx -10000$$ at the point where $x = 0.01$. **[2]**

The equation $y = 0$ has a root near $x = 1$. Joan uses the Newton-Raphson method to find this root. The output from the spreadsheet she uses is shown in Fig. 10.1.

n	0	1	2	3	4	5	6	7
x_n	1	0.958509	0.950084	0.948261	0.94786	0.947772	0.947753	0.947748

Fig. 10.1

Joan carries out some analysis of this output. The results are shown in Fig. 10.2.

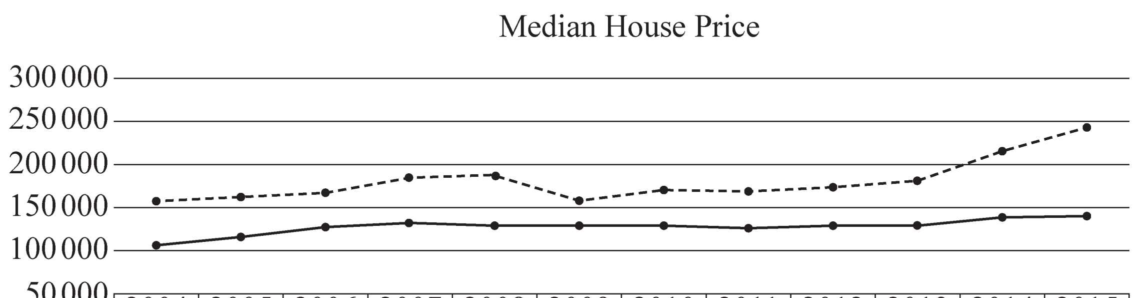
x	y
0.9477475	–7.79967E–07
0.9477485	–2.90821E–06
x	y
0.947745	4.54066E–06
0.947755	–1.67417E–05

Fig. 10.2

(c) Consider the information in Fig. 10.1 and Fig. 10.2.

- Write 4.54066E–06 in standard mathematical notation.
- State the value of the root as accurately as you can, justifying your answer. **[3]**

11 The pre-release material contains information concerning median house pric
2004 – 2015. A spreadsheet has been used to generate a time series graph for two
borough of “Barking and Dagenham” and “North West”. This is shown togethe
in Fig. 11.1.



	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Barking and Dagenham	160 000	163 000	168 000	185 000	190 000	160 000	171 000	170 000	174 995	180 995	215 000	243 500
North West	107 000	118 000	127 000	134 750	129 950	130 000	130 000	127 000	130 000	131 000	138 500	140 000

Barking and Dagenham — North West

Fig. 11.1

Dr Procter suggests that it is unusual for median house prices in a London borough to be consistently higher than those in other parts of the country.

(a) Use your knowledge of the large data set to comment on Dr Procter’s suggestion. **[1]**

Dr Procter wishes to predict the median house price in Barking and Dagenham in 2016. She uses the spreadsheet function LINEST to find the equation of the line of best fit for the given data. She obtains the equation

$P = 4897Y - 9\,657\,847$, where P is the median house price in pounds and Y is the calendar year, for example 2015.

(b) Use Dr Procter’s equation to predict the median house price in Barking and Dagenham in

- 2016
- 2017. **[2]**

Professor Jackson uses a simpler model by using the data from 2014 and 2015 only to form a straight-line model.

(c) Find the equation Professor Jackson uses in her model. **[2]**

(d) Use Professor Jackson’s equation to predict the median house price in Barking and Dagenham in

- 2016
- 2017. **[2]**

Professor Jackson carries out some research online. She finds some informa house prices in Barking and Dagenham, which is shown in Fig. 11.2.

2016	2017
£290 000	£300 000

Fig. 11.2

(e) Comment on how well

- Dr Procter's model fits the data,
- Professor Jackson's model fits the data. **[2]**

(f) Explain which, if any, of the models is likely to be more reliable for predicting median house prices in Barking and Dagenham in 2020. **[1]**

12 In this question you must show detailed reasoning.

A 5-sided spinner can give scores of 1, 2, 3, 4 or 5. After observing a large number of spins, Elaine models the probability distribution of X, the score on the spinner, as shown in Fig. 12.

x	1	2	3	4	5
$\mathrm{P}(X = x)$	0.2	0.3	p	p	q

Fig. 12

When the spinner is spun twice, the probability of obtaining a total score of 9 is 0.06.

(a) Given that $q < 2p$, determine the values of p and q. **[6]**

(b) The spinner is spun 10 times. Calculate the probability that exactly one 5 is obtained. **[2]**

Elaine's teacher believes that the probability that the spinner shows a 1 is greater than 0.2. The spinner is spun 100 times and gives a score of 1 on 28 occasions.

(c) Conduct a hypothesis test at the 5% level to determine whether there is any evidence to suggest that the probability of obtaining a score of 1 is greater than 0.2. **[7]**

13 The pre-release material contains information concerning median house prices, employment rates. Fig. 13.1 shows a scatter diagram of recycling rate against employment rate for a random sample of 33 regions.

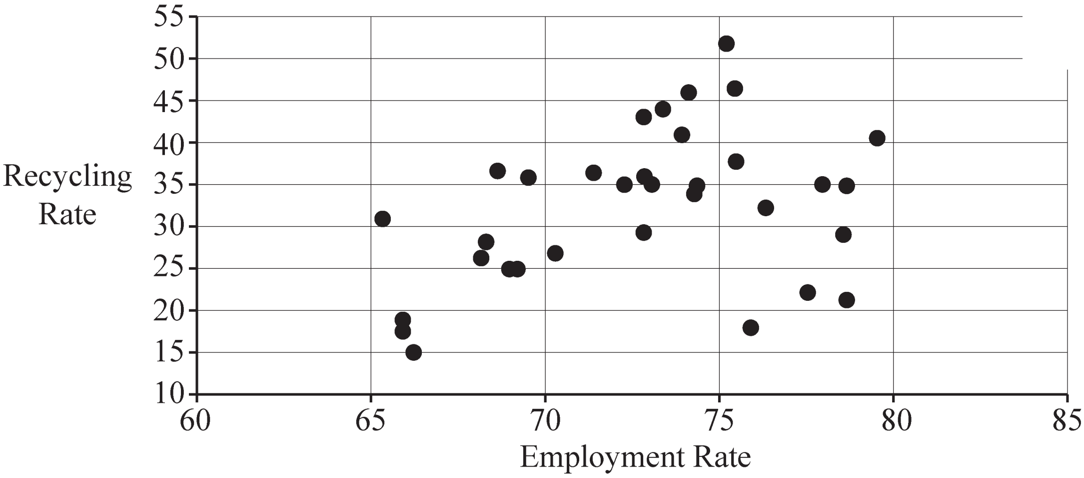


Fig. 13.1

The product moment correlation coefficient for this sample is 0.37154 and the associated p-value is 0.033.

Lee conducts a hypothesis test at the 5% level to test whether there is any evidence to suggest there is positive correlation between recycling rate and employment rate. He concludes that there is no evidence to suggest positive correlation because $0.033 \approx 0$ and $0.37154 > 0.05$.

(a) Explain whether Lee's reasoning is correct. **[2]**

Fig. 13.2 shows a scatter diagram of recycling rate against median house price for a random sample of 33 regions.

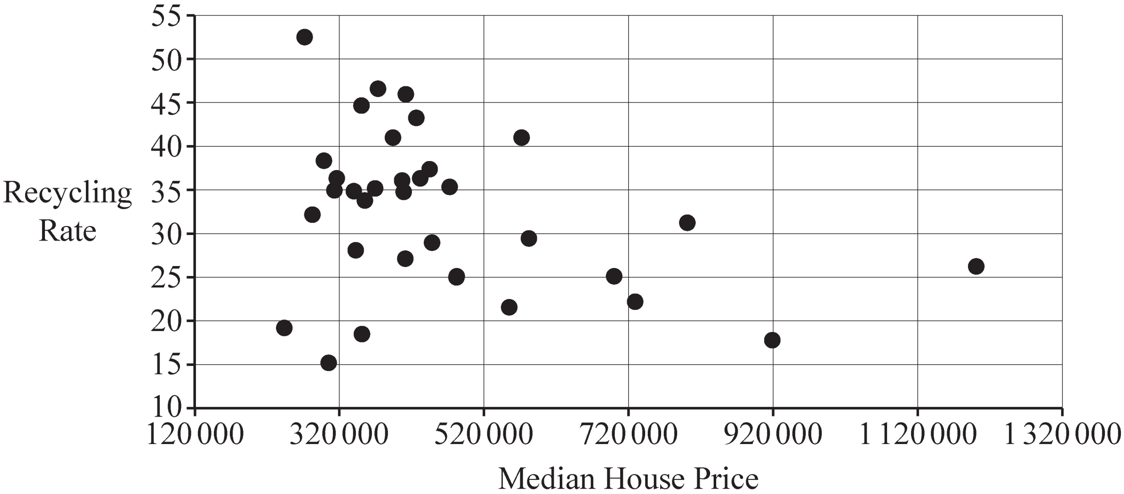


Fig. 13.2

The product moment correlation coefficient for this sample is −0.33278 and the
is 0.058.

Fig. 13.3 shows summary statistics for the median house prices for the data in tl

Statistics	
n	33
Mean	465467.9697
σ	201236.1345
s	204356.2606
Σx	15360443
Σx^2	8486161617387
Min	243500
Q1	342500
Median	410000
Q3	521000
Max	1200000

Fig. 13.3

(b) Use the information in Fig. 13.3 and Fig. 13.2 to show that there are at least two outliers. **[2]**

(c) Describe the effect of removing the outliers on

- the product moment correlation coefficient between recycling rate and median house price,
- the *p*-value associated with this correlation coefficient,

in each case explaining your answer. **[2]**

All 33 items in the sample are areas in London. A student suggests that it is very unlikely that only areas in London would be selected in a random sample.

(d) Use your knowledge of the pre-release material to explain whether you think the student's suggestion is reasonable. **[1]**

14 In this question you must show detailed reasoning.

Fig. 14 shows the graphs of $y = \sin x \cos 2x$ and $y = \frac{1}{2} - \sin 2x \cos x$.

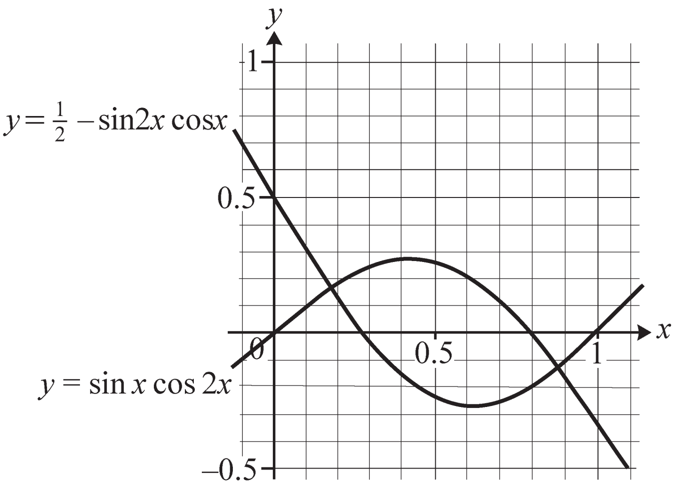


Fig. 14

Use integration to find the area between the two curves, giving your answer in an exact form. **[8]**

15 Functions $\mathrm{f}(x)$ and $\mathrm{g}(x)$ are defined as follows.

$\mathrm{f}(x) = \sqrt{x}$ for $x > 0$ and $\mathrm{g}(x) = x^3 - x - 6$ for $x > 2$.

The function $\mathrm{h}(x)$ is defined as

$\mathrm{h}(x) = \mathrm{fg}(x)$.

(a) Find $\mathrm{h}(x)$ in terms of x and state its domain. **[2]**

(b) Find $\mathrm{h}(3)$. **[1]**

Fig. 15 shows $\mathrm{h}(x)$ and $\mathrm{h}^{-1}(x)$, together with the straight line $y = x$.

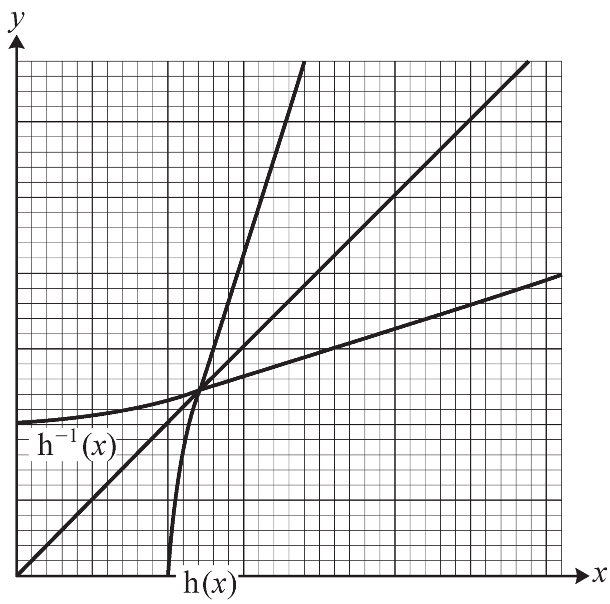


Fig. 15

(c) Determine the gradient of $y = \mathrm{h}^{-1}(x)$ at the point where $y = 3$. **[4]**

END OF QUESTION PAPER

BLANK PAGE

BLANK PAGE