

Wednesday 14 October 2020 – Afternoon

A Level Mathematics B (MEI)

H640/02 Pure Mathematics and Statistics

Time allowed: 2 hours

8 2 5 7 6 0 8 4

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.

Section A (23 marks)

1 Fig. 1 shows triangle *ABC*.

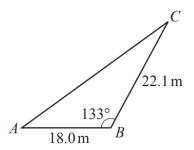


Fig. 1

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

[2]

Area =
$$\frac{1}{2}$$
 × 22.1 × 18.0 × sin 133° = 145.46... ≈ 145 (35f)

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

The angle of the sector is 2.1 radians.

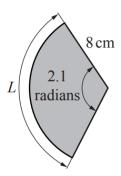


Fig. 2

(a) Calculate the length of the arc L.

[1]

(b) Calculate the area of the sector.

[2]

(a) Arc Length =
$$r\theta = 8 \times 2.1 = 16.8 \text{ cm}$$

(b.) Area of Sector =
$$\frac{\theta^{\circ}}{360} \pi r^2 = \frac{2.1}{2\pi} \times \pi \times 8^2 = 67.2 \text{ cm}^2$$

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

(a) Find
$$\frac{dy}{dx}$$
. [2]

(b) Find the smallest positive value of x for which $\frac{dy}{dx} = 0$, giving your answer in an exact form. [2]

(a.)
$$y = 4z + \sin 8z$$

$$y = \sin \alpha x \rightarrow \frac{dy}{dx} = a\cos \alpha x$$

$$\frac{dy}{dx} = 4 + 8\cos 8x$$

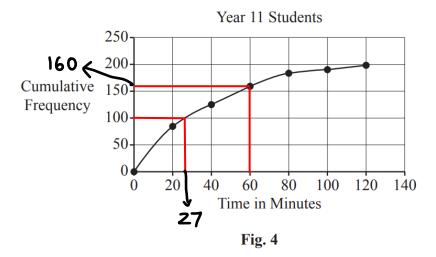
(b)
$$\frac{dy}{dz} = 0 \rightarrow 4 + 8\cos 8x = 0$$

 $\cos 8z = -\frac{4}{8} = -\frac{1}{2}$
 $8z = \cos^{-1}(-\frac{1}{2}) = \frac{2\pi}{3}$

$$x = \frac{2\pi}{3} \div 8 = \frac{\pi}{12}$$

$$\therefore z = \frac{\pi}{12}$$

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.



(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]

(a.) Total Students = 200
$$\frac{200}{2}$$
 = 100

(b.)
$$\frac{160}{200} \times 100 = 80\%$$

80% of the students spent less than an hour revising maths. ... Teacher's comment isn't supported in the diagram.

5 The first n terms of an arithmetic series are

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

- (a) Find the value of n. [1]
- (b) Find the sum of these *n* terms. [2]

(a)
$$17, 28, 39, ..., 281, 292$$

$$\therefore$$
 nth Term = $11n + 6$

$$N = \frac{292 - 6}{11} = 26$$
 : $n = 26$

(b.)
$$S_n = \frac{1}{2}n(a+l)$$

$$S_{26} = \frac{1}{2}(26)(17+292) = 4017$$

- 6 (a) Find the first three terms in ascending powers of x of the binomial expansion of $(1+4x)^{\frac{1}{2}}$. [3]
 - (b) State the range of values of x for which this expansion is valid. [1]

(a.)
$$(1+z)^n = 1 + nx + \frac{n(n-1)}{2!}x^2$$

$$(1+4z)^{\frac{1}{2}} = 1 + \frac{1}{2}(4z) + \frac{\frac{1}{2}(\frac{1}{2}-1)}{2!} (4z)^{2}$$

$$= 1 + 2z - \frac{1}{8}(16z^{2})$$

$$= 1 + 2z - 2z^{2}$$

(b.) Given $(1+2)^n$, expansion is valid for |x| < 1. $|4x| < 1 \Rightarrow |x| < \frac{1}{4}$

: Expansion is valid for $121 < \frac{1}{4}$.

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

(a) Find
$$P(A \cap B)$$
. [2]

(b) Find
$$P(A|B)$$
. [2]

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

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

 $P(A \cup B) = 1 - P(A \cup B)' = 1 - 0.2 = 0.8$
 $P(A \cap B) = 0.6 + 0.5 - 0.8 = 0.3$
 $\therefore P(A \cap B) = 0.3$

(b.)
$$P(A1B) = \frac{P(A \cap B)}{P(B)} = \frac{0.3}{0.5} = 0.6$$

(C) Events A & B are independent if
$$P(A \cap B) = P(A) \times P(B)$$
.

$$P(A) \times P(B) = 0.6 \times 0.5 = 0.3 = P(A \cap B)$$

.. A & B are independent.

Section B (77 marks)

- Rosella is carrying out an investigation into the age at which adults retire from work in the city where she lives. She collects a sample of size 50, ensuring this comprises of 25 randomly selected retired men and 25 randomly selected retired women.
 - (a) State the name of the sampling method she uses.

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

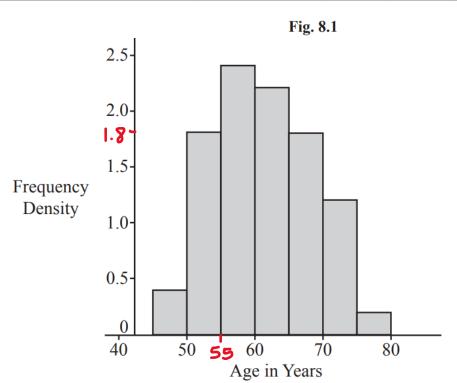


Fig. 8.2

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

[1]

[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 500, from her list. The histograms in Fig. 8.3 show the data from the sample of size 200 on the left and the data from the sample of 500 on the right.

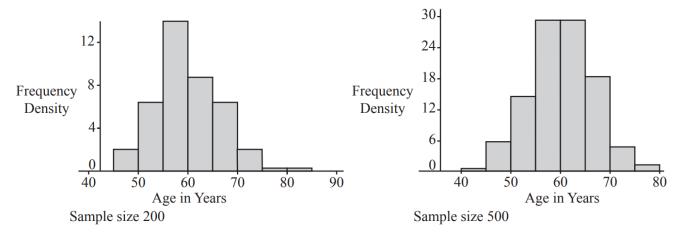


Fig. 8.3

(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	
Σχ	30025.7601	
Σx ²	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]

- (a.) Quota Sampling
- (b.) Frequency = class width x Frequency Density $f = (55-50) \times 1.8 = 9$

:. 9 people are aged 50-55.

(c.) Systematic Sampling:

4960 - 200 = 24.8

- ·Select every 24th number on the list
- Start randomly between n=1 and $n \ge 184$ and stop when 200 have been selected.

Simple Random Sampling:

- Assign each person in list by a unique identifier from 1 to 4960.
- · Generate <u>random numbers</u> between 1 and 4960 until a sample of 200 has been selected.
- (d.) Normal Distribution is reasonable to model the age of retirement because as the size of the sample increases, the shape of the distribution appears more "normal".
- (e.) $X \sim N(Mean, Standard Deviation^2) \rightarrow X \sim N(60.0515, 6.5783^2)$ P(X > 65) = 0.2259... $0.2259... \times 4960 = 1120.71... \approx 1121$

:1121 people are aged over 65.

(f.) Limitation: There may be seasonal fluctuations, such as teachers retiring in August

- A company supplies computers to businesses. In the past the company has found that computers are kept by businesses for a mean time of 5 years before being replaced. Claud, the manager of the company, thinks that the mean time before replacing computers is now different.
 - (a) Describe how Claud could obtain a cluster sample of 120 computers used by businesses the company supplies. [1]

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	
Σχ	586.2566	
Σx ²	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]
- (a.) To obtain a cluster sample of 120 computers, randomly select N different businesses and then randomly select P different computers from each of the businesses.

(b) Ho: M = 5 2 Tailed Test. .. Significance = 2.5% = 0.025

 $H_1: \mu \neq 5 \rightarrow H_1$ takes this form as it is testing whether the mean length of time is different to 5.

m is the population mean time for which computers are kept before being replaced.

Use Normal Distribution, where $\mu = 5$, $\sigma = \frac{2.7054^2}{120}$. $X \sim N(5, \frac{2.7054^2}{120})$ used to find $P(\bar{x} < 4.8855)$.

 $P(\bar{x} < 4.8855) = 0.32 > 0.025$

- .. Reject H. Accept Ho.
- ... There is insufficient evidence to suggest at 5% significance level that the mean length of time is not 5 years.

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

$$\int_{0.01}^{0.05} y \, \mathrm{d}x \approx \ln 5.$$

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

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

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.

y
-7.79967E-07
-2.90821E-06
y
4.54066E-06
-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]

(a.)
$$\sin \theta \gtrsim \Theta \Rightarrow \therefore \sin 2z = 2z \otimes \sin z = z$$
.

$$\therefore y = \frac{2x - x}{x \cdot x} = \frac{z}{x^2} = \frac{1}{z}$$

$$\int_{0.01}^{0.05} \int_{0.01}^{1} dx = \left[\ln x \right]_{0.01}^{0.05} = \ln 0.05 - \ln 0.01 = \ln \left(\frac{0.05}{0.01} \right) = \ln 5$$

(b)
$$y = \frac{1}{x} = x^{-1}$$

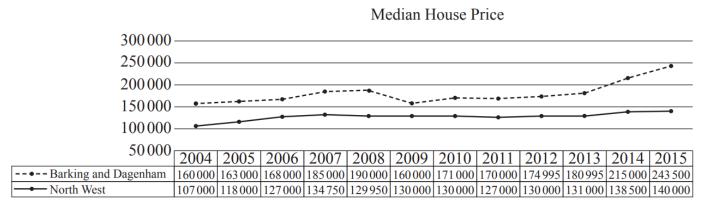
$$\frac{dy}{dx} = -1x^{-2} = -\frac{1}{x^2}$$

$$x = 0.01 \Rightarrow \frac{1}{2} = \frac{1}{0.01^2} = -10000$$

(c)
$$4.54066 = -06 = 4.54066 \times 10^{-6}$$

Sign of y changes between x = 0.947745 and x = 0.9477475. Sign change for 5 dp.

The pre-release material contains information concerning median house prices over the period 2004 – 2015. A spreadsheet has been used to generate a time series graph for two areas: the London borough of "Barking and Dagenham" and "North West". This is shown together with the raw data in Fig. 11.1.



--•-- 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 - 9657847, 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

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 information about median house prices in Barking and Dagenham, which is shown in Fig. 11.2.

2016	2017		
£290000	£300000		

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.
- (a.) House prices are generally higher in London boroughs than elsewhere in the country, so Dr Procter's suggestion is probably wrong.

(b)
$$P = 4897 \, \text{Y} - 9657847$$

 $Y = 2016 \rightarrow P = 4897(2016) - 9657847 = 214505$
 $Y = 2017 \rightarrow P = 4897(2017) - 9657847 = 219402$

(c.) Gradient,
$$m = \frac{243500 - 215000}{2015 - 2014} = 28500$$

(2014, 215000) & m = 28500

$$y - 215000 = 28500(x - 2014)$$

(d.)
$$Y = 2016 \Rightarrow P = 28500(2016) - 57184000 = 272000$$

 $Y = 2017 \Rightarrow P = 28500(2017) - 57184000 = 300 500$

(e.) Dr Procter's model is a very poor fit.

Professor Jackson's model is a good fit. It works well for 2017, but not so well for 2016.

(f.) Neither of the medels are more reliable because they use extrapolation.

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
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]

(a.)
$$2p + q + 0.2 + 0.3 = 1 \Rightarrow q = 0.5 - 2p$$

 $2pq = 0.06 \Rightarrow 2p(0.5 - 2p) = 0.06$
 $p - 4p^2 = 0.06$
 $4p^2 - p + 0.06 = 0$
 $(3p - 20)(p - 10) = 0$
 $p = 0.15 \quad p = 0.1$
 $p = 0.15 \Rightarrow q = 0.5 - 2(0.15) = 0.2$
 $p = 0.1 \Rightarrow q = 0.5 - 2(0.1) = 0.3$
Since $q < 2p$, $q = 0.2$ and $p = 0.15$.

(b.)
$$X \sim B(10,0.2)$$

 $P(X=1) = {}^{10}C_{1} \cdot q^{1} \cdot (1-q_{1})^{10-1}$
 $= {}^{10 \cdot 0.2 \cdot 0.8}$
 $= 0.2684...$
 $\sim 0.268 (3sf)$

(c) Ho:
$$P = 0.2$$
 Significance = 5% = 0.05
 H_1 : $P > 0.2$
 $X \sim B(100, 0.2)$
 $P(X \le 27) = 0.9658...$
 $P(X \ge 28) = 1 - P(X \le 27) = 0.03415... < 0.05$

- . Reject Ho. Accept Hi.
- ... There is enough evidence to suggest at 5% significance level that the probability of a score of 1 is greater than 0.2.

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

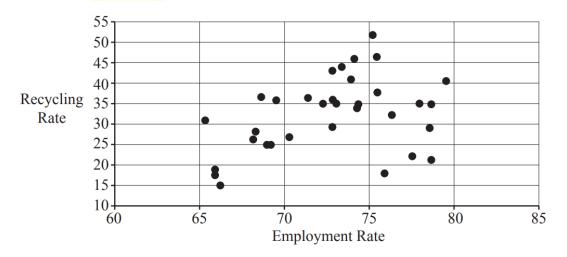


Fig. 13.1

The product moment correlation coefficient for this sample is $\frac{0.37154}{0.033}$ and the associated p-value is $\frac{0.033}{0.033}$.

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.

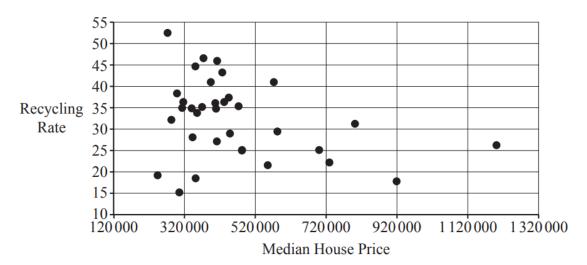


Fig. 13.2

The product moment correlation coefficient for this sample is $\frac{-0.33278}{0.058}$ and the associated p-value is $\frac{0.058}{0.058}$.

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

Statistic	s
n	33
Mean	465467.9697
σ	201236.1345
S	204356.2606
Σχ	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]
- (a.) Lee is wrong because he should make the comparison of 0.033 with 0.05 and he should make the comparison of 0.37154 with 0.
- (b) $\mu + 2sd = 465467 + 2(204356) = 874179$
 - ... Regions with median house prices greater than 874179 are outliers.
 - .. Outliers: 920000 & 1200000

- (c.) The proce would probably be closer to 0 because the Scatter is less well modelled by a straight line.

 The p-value would increase because a value which is obser to 0 is more likely assuming no correlation.
- (d.) The Student's suggestion is reasonable, since there are other regions defined in the LDS.

14 In this question you must show detailed reasoning.

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

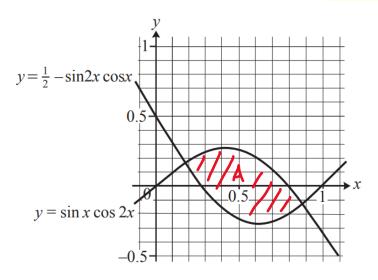


Fig. 14

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

$$\frac{1}{2} - \sin 2\pi \cos x = \sin x \cos 2x$$

$$\sin x \cos 2x + \sin 2x \cos x = \frac{1}{2}$$

$$Sin(A \pm B) = SinAcosB \pm SinBcosA$$

$$\therefore \sin x \cos 2x + \sin 2x \cos x = \sin (x + 2x) = \sin 3x$$

$$3x = \sin^{-1}(\frac{1}{2}) = \frac{\pi}{6}, \frac{5\pi}{6} \rightarrow \therefore x = \frac{\pi}{18}, \frac{5\pi}{18}$$

$$\frac{38}{18} \sin x \cos 2x - \left(\frac{1}{2} - \sin 2x \cos x\right) dx$$

$$= \int \sin 3x - \frac{1}{2} dx$$

$$= \left[-\frac{\cos 3x}{3} - \frac{x}{2}\right] \frac{51}{18}$$

$$= \left[-\frac{\cos (3 \times \frac{51}{18})}{3} - \frac{(\frac{511}{18})}{2}\right] - \left(-\frac{\cos (3 \times \frac{11}{18})}{3} - \frac{(\frac{11}{18})}{2}\right]$$

$$= \frac{13}{6} - \frac{511}{36} + \frac{13}{6} + \frac{13}{36}$$

$$= \frac{\sqrt{3}}{3} - \frac{11}{9}$$

$$\therefore A = \sqrt{\frac{3}{3}} - \sqrt{\frac{1}{9}} \quad \text{units}^2$$

15 Functions f(x) and g(x) are defined as follows.

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

The function h(x) is defined as

h(x) = fg(x).

(a) Find h(x) in terms of x and state its domain.

[1]

[2]

(b) Find h(3).

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

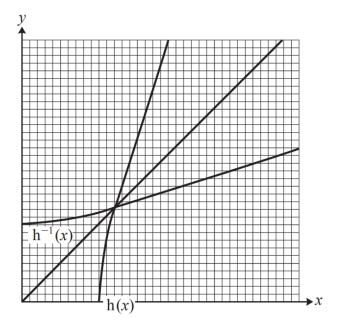


Fig. 15

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

(a.)
$$h(x) = fg(x) = \sqrt{x^3 - x - 6}$$

:
$$h(x) = \sqrt{x^3 - x - 6}$$
 for $x > 2$.

(b)
$$h(3) = \sqrt{3^3 - 3 - 6} = \sqrt{18} = 3\sqrt{2}$$

$$-1.h(3) = 3\sqrt{2}$$

(c)
$$h(x) = \int x^3 - x - 6$$

 $y = \int x^3 - x - 6$
Let $x = y : x = \int y^3 - y - 6$
 $x^2 = y^3 - y - 6$

Differentiating
$$x^2 = y^3 - y - 6$$
:
$$2x = 3y^2 \frac{dy}{dx} - \frac{dy}{dx}$$

$$2x = \frac{dy}{dx} (3y^2 - 1)$$

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

$$y = 3 \Rightarrow h(3) = 3\sqrt{2}$$

$$\frac{dy}{dx} = \frac{2(3\sqrt{2})}{3(3^2) - 1} = \frac{3\sqrt{2}}{13}$$

$$\therefore \text{ Gradient} = \frac{3\sqrt{2}}{13}$$