Qualification Accredited

Oxford Cambridge and RSA

AS LEVEL

Examiners' report

MATHEMATICS A

H230

For first teaching in 2017

H230/01 Summer 2023 series

Version 1

Contents

Introduction

Our examiners' reports are produced to offer constructive feedback on candidates' performance in the examinations. They provide useful guidance for future candidates.

The reports will include a general commentary on candidates' performance, identify technical aspects examined in the questions and highlight good performance and where performance could be improved. A selection of candidate answers is also provided. The reports will also explain aspects which caused difficulty and why the difficulties arose, whether through a lack of knowledge, poor examination technique, or any other identifiable and explainable reason.

Where overall performance on a question/question part was considered good, with no particular areas to highlight, these questions have not been included in the report.

A full copy of the question paper and the mark scheme can be downloaded from OCR.

Paper 1 series overview

The overall standard was similar to 2022 (and, in turn, to 2019), with very few candidates scoring extremely high or very low marks.

However, the questions requiring verbal, written answers were notably less well answered than in 2022. Many candidates gave insufficiently specific answers, reciting facts rather than addressing the specific request in the question and so did not gain marks on these questions.

As in 2022, many candidates lost marks for not showing sufficient working. Centres should make sure candidates are aware of the specific meaning of command words, in particular 'Determine', and the significance of '…you must show detailed reasoning'. The 2022 examiners' report contains a helpful explanation of how much working is required in these types of questions.

Candidates who did well on this paper generally:	Candidates who did less well on this paper generally:
• demonstrated a good level of understanding on both the Pure and Statistics sections of the paper • made appropriate use of their calculator (e.g. for statistical calculations), but also showed detailed working when required • read the questions carefully and provided an appropriately specific answer • used their familiarity with the Large data set to provide sensible, plausible explanations in Question 12.	• did not show sufficient working (especially on 'Determine', 'Show that' and 'detailed reasoning' questions) • recited generic facts in response to 'Explain' questions, rather than addressing the specific request in the question • appeared to lack understanding of statistical concepts (e.g. hypothesis testing and interpreting data from a histogram) • were unable to use their familiarity with the Large data set.

OCR support

A student guide and a classroom poster to help reinforce the A Level Maths command words are available to be downloaded.

Section A overview

The Pure section of this paper was generally well answered, but most candidates did better in Section B. Candidates were generally able to perform the required operations, demonstrating understanding across the specification. However, many candidates struggled with questions relating to proof and did not show sufficient working in 'Determine' or 'detailed reasoning' questions. Candidates also appeared to struggle with 'unseen' concepts, especially when relating knowledge across Pure and Statistics.

Question 1 (a)

1 **(a)** Prove that $\cos x + \sin x \tan x \equiv \dfrac{1}{\cos x}$ (where $x \neq \frac{1}{2}n\pi$ for any odd integer n). **[3]**

This question was generally well answered by most candidates. Some candidates showed a lack of rigour by writing down the given identity and manipulating one or both sides. This was permitted as long as it was clear what they were doing (e.g. multiplying both sides by cos(x) needed to be clearly shown). Some candidates wrote the given identity followed by the Pythagorean identity with no intermediate working, which could not be credited.

In 'Show that' and proof questions, students should be encouraged to start with one side and manipulate to reach the other, and indicate clearly when they have reached a conclusion. More candidates used the alternative method in the mark scheme than attempted to combine terms on the LHS over a common denominator. Some candidates' notation was unclear, for example sin x^2 instead of $\sin^2 x$. Another common error was using sin + cos = 1 (which is missing the squares).

Question 1 (b)

(b) Solve the equation $2\sin^2 x = \cos^2 x$ for $0° \leqslant x \leqslant 180°$. **[2]**

The majority of candidates were able to correctly manipulate this equation. However, of those who obtained the correct equation for the first M1, many either did not square root or forgot to consider more than one solution. Some missed the second solution in the given interval (achieving M1A0), while others gave additional solutions outside the interval.

Question 2 (a) (i)

2 **(a)** The points A, B and C have position vectors $\begin{pmatrix}-4\\3\end{pmatrix}$, $\begin{pmatrix}-3\\6\end{pmatrix}$ and $\begin{pmatrix}-1\\12\end{pmatrix}$ respectively.

(i) Show that B lies on AC. **[2]**

Many candidates chose to approach this question in a 'non-vector' method, using standard coordinate geometry, which was permitted (but required appropriate rigour). For example, finding the equation of the line between A and C, then showing that B satisfies the equation of that line. A number of candidates found two of the required vectors (e.g. $\overrightarrow{AB}$ and $\overrightarrow{BC}$) then showed that $\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}$ and claimed this proved that B lies on the line AC (which it doesn't prove). A method stating two gradients could only be credited if candidates showed that the gradients of AB and BC were 3 **and** stated that B was a common point (showing that, for instance, AB and AC have the same gradient and share the point A does not prove that B lies on the line [segment] AC). Many candidates confused x and y when attempting to find a gradient. Centres should encourage students to be familiar with vector manipulation, as this was the most straightforward way to answer this question.

Question 2 (a) (ii)

(ii) Find the ratio $AB : BC$. **[1]**

This was reasonably well answered, but some candidates did not appear to recognise the difference in notation between $\overrightarrow{AB}$ and AB. Many candidates left responses in the form of a ratio of vectors rather than lengths.

Candidates who had not computed the vectors AB and BC needed to do so in this part. Some candidates made an error in comparing vectors, or gave the ratio the wrong way around (which could not be credited unless they specified the ratio as $BC : AB$). Only ratios in the form $p : q$ were acceptable here (a 'ratio' of vectors could not be credited).

Misconception

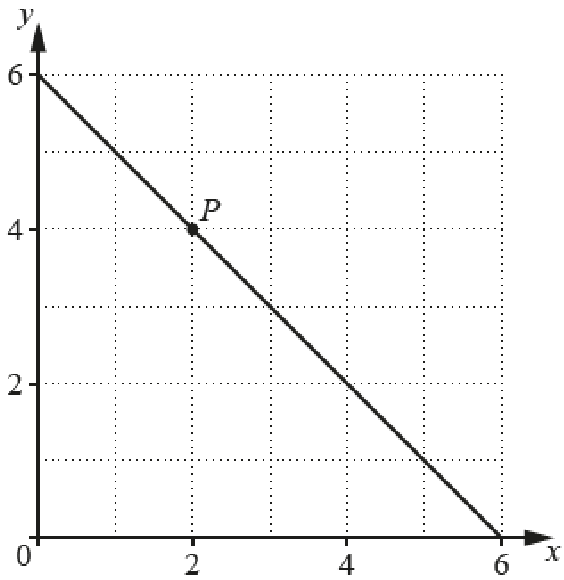
Although in this case it was possible to answer the first part of the question using only coordinate geometry, the vector method was by far the most straightforward.

In Question 2 (a) (ii) many candidates did not appear to understand that the question required a ratio of lengths.

Question 2 (b)

(b) The diagram shows the line $x+y=6$ and the point $P(2, 4)$ that lies on the line.

A copy of the diagram is given in the Printed Answer Booklet.



The distinct point Q also lies on the line $x+y=6$ and is such that $|\overrightarrow{OQ}|=|\overrightarrow{OP}|$, where O is the origin.

Find the magnitude and direction of the vector $\overrightarrow{PQ}$. **[3]**

Many candidates were able to correctly identify the point Q as (4, 2). Some candidates misread the question and gave the magnitude of $\overrightarrow{OP}$ or $\overrightarrow{OQ}$ instead of $\overrightarrow{PQ}$ as required.

Relatively few candidates correctly computed the direction of $\overrightarrow{PQ}$ (as -45° or 315°); in general candidates should know that the term 'direction' in questions relating to vectors means 'compute the anti-clockwise angle to the positive *x*-axis'. Note that 135° could only be credited if specified as a bearing. Candidates who made use of the diagram in the answer booklet (while not required) were more likely to achieve 2 or 3 marks on this question.

Question 2 (c)

(c) The point R is such that $\overrightarrow{PR}$ is perpendicular to $\overrightarrow{OP}$ and $|\overrightarrow{PR}|=\frac{1}{2}|\overrightarrow{OP}|$.

Write down the position vectors of the **two** possible positions of the point R. **[2]**

Many candidates omitted this part, and relatively few candidates achieved one or both marks (including those who had obtained full or nearly full marks across parts (a) and (b)). Many appeared to not have understood the request in the question and computed some related vectors or magnitudes without reaching the required vectors for R. Again use of the diagram appeared to correlate with success.

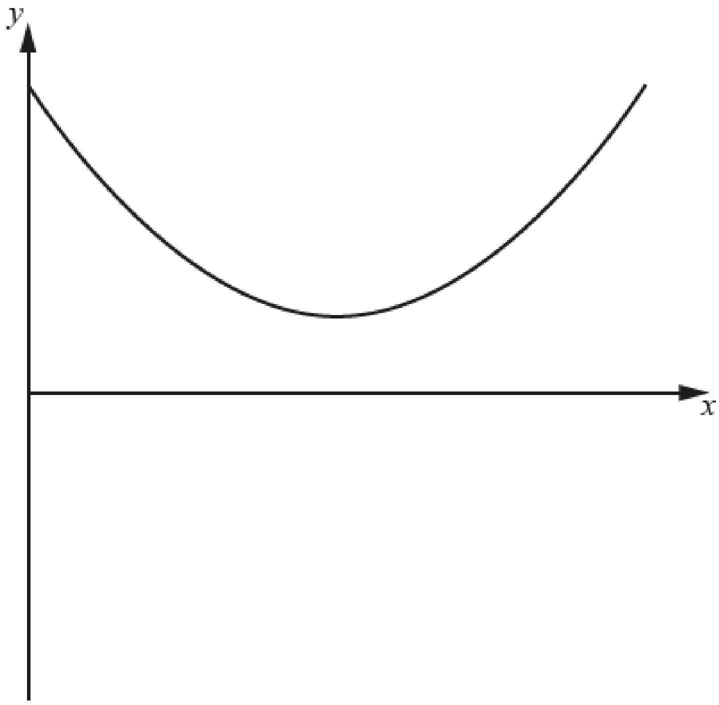
The command words 'Write down' indicates that very little algebraic working was required. Those candidates who found the gradient of *OP* (i.e. 2) and then the corresponding perpendicular gradient (i.e. -0.5) and drew a line of this gradient through *P* were most likely to produce the correct answers. Those who tried to rely on algebra and use the distance formula would often get stuck by not realising they needed to substitute the equation of the line through *P* ($y = -0.5x + 5$) to get their final answers.

Some candidates did not give their responses in the required form (position vectors) and gave the co-ordinates of *R* instead.

Question 3 (a)

3 The diagram shows the graph of $y = \mathrm{f}(x)$, where $\mathrm{f}(x)$ is a **quadratic** function of x.

A copy of the diagram is given in the Printed Answer Booklet.



(a) On the copy of the diagram in the Printed Answer Booklet, draw a possible graph of the gradient function $y = \mathrm{f}'(x)$. **[3]**

This question was reasonably well answered, with many candidates achieving at least 1 or 2 marks. Some candidates gave a non-straight line or a 'modulus-like' graph that showed some (but not fully correct) qualitative understanding of differentiation (or perhaps that they had not noted the instruction in the question that the given function was quadratic). A common incorrect response was a straight line with positive gradient, but starting from the positive *y*-axis (showing that candidates did not appreciate the need for the gradient to be negative in the first half of the graph). Centres should encourage candidates to carefully check their responses to questions like this. Straight lines that did not extend to the *y*-axis could not achieve full marks.

Some candidates drew a line that was a tangent to the given curve. Many recognised that the gradient at the minimum would be zero and thus had their graph cutting, or touching, the *x*-axis vertically below the minimum.

Question 3 (b)

(b) State the gradient of the graph of $y = f''(x)$. **[1]**

Many candidates interpreted this correctly, although '2' was a common incorrect response (from f''(x) = 2, but not recognising that the gradient of this horizontal straight line is 0). This question did not require part (a), although correct responses often followed a 'reasonably correct' interpretation in part (a).

Question 4 (a)

4 A curve has equation $y = e^{3x}$.

(a) Determine the value of x when $y = 10$. **[2]**

This was well answered by the majority of candidates, with most being able to correctly take logs of both sides and reach the required answer. Note that as a 'Determine' question some working was required and most candidates followed this instruction adequately. Quite a few candidates rounded incorrectly (giving 0.767 or 0.77 as their final response), but most of these had reached an exact form in their working first so few lost marks for this error.

Question 4 (b)

(b) Determine the gradient of the tangent to the curve at the point where $x = 2$. **[2]**

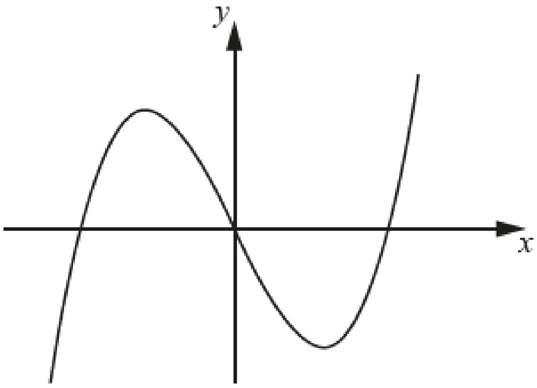
Most candidates recognised the need to differentiate and the majority of these got the correct derivative. Common errors included $-3e^{2x}$, $3xe^{3x}$ or getting $y = e^6$ and then gradient = $6e^5$.

Some candidates seemed unsure about the degree of accuracy required (with many giving their response to 2dp or similar); as per the rubric on the front cover, 3sf is the default unless otherwise stated.

Question 5

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

The diagram shows part of the graph of $y = x^3 - 4x$.



Determine the total area enclosed by the curve and the x-axis. **[6]**

This was a 'detailed reasoning' question, so appropriate working throughout was required to gain full marks.

Quite a number of candidates did not show any working when finding the roots, which meant that they could not be given the first B mark (in a 'detailed reasoning' question, some evidence of the method used is required). Some candidates showed a valid factorisation, but did not state the roots explicitly (because of the nature of this question, the second B mark was given if the roots were seen to be 'used' later, but candidates should be reminded of the importance of showing detailed reasoning).

Candidates were required to show clear evidence of recognising and correctly handling the negative area for full marks. Some candidates switched the limits to 'handle' the negative area, which was permissible as long as their responses were consistent with their limits. Similarly for candidates who used modulus signs to 'handle' negative area (as long as it was clear what they had done). In general candidates should be encouraged to show all their working and clearly indicate negative areas, then explicitly handle (rather than going back and retrospectively adjust their working to remove or introduce - signs).

The majority of candidates were good at showing the integrated expression but the substitution of limits was not always clearly shown. Centres should advise candidates to be careful not to write mathematically incorrect statements. In this case notation such as -4 = 4 or -8 = 8 was condoned, and the final M mark could be given for candidates who recognised the need to 'match' the signs of the area, (e.g. -4 + (-4) = -8), but the final A mark required a correct, positive answer from fully correct working. Relatively few candidates used the symmetry of the curve to find the other half area (which was an acceptable method as long as it was done clearly).

Care should be taken with the negative integral, e.g. a clear modulus sign used or comment that $A_1 = \int_0^2 (x^3 - 4x)\mathrm{d}x = -4$ gives an area of 4 sq units below the x-axis. Simply changing a negative sign to a positive sign ($A_1 = \int_0^2 (x^3 - 4x)\mathrm{d}x = +4$) was penalised.

The main errors were dealing with signs in the integrals. Most candidates recognised the need to consider the two areas separately and it was rare to see the single integral $\int_{-2}^2 (x^3 - 4x)\mathrm{d}x = 0$ as an answer.

Exemplar 1

$\cancel{\int (x^3}$ $\quad x^3 - 4x = 0 \quad x = 2$ or $x = -2$ or $x = 0$

$\int_0^2 -(x^3 - 4x)\,dx + \int_{-2}^0 (x^3 - 4x)\,dx$

$= -\left[\frac{1}{4}x^4 - 2x^2\right]_0^2 + \left[\frac{1}{4}x^4 - 2x^2\right]_{-2}^0$

$= \cancel{\frac{56}{15}} \quad \cancel{4 + 4} \quad -\left(\left(\frac{1}{4} \times 2^4 - 2 \times 4\right) - (0 - 0)\right) + \left((0 - 0) - \left(\frac{1}{4} \times (-2)^4 - 2 \times 4\right)\right)$

$= \cancel{8} \quad 4 + 4$

$= 8$

This candidate has split the area correctly, but did not provide any working to support their roots (so could not be given the first B mark). They have correctly integrated the given expression and then handled the negative area correctly. Ideally we would have seen a more explicit recognition of the negative area (for instance showing –(-4) + 4 before the penultimate line), but this was sufficient for the remaining marks. This was awarded 5/6.

Question 6 (a)

6 (a) Determine the **two** real roots of the equation $8x^6 + 7x^3 - 1 = 0$. **[3]**

Some candidates did not identify this as a quadratic in x^3 and so either struggled to solve it or provided solutions only by calculator (SC marks for this were in the mark scheme). The command word 'Determine' requires working to be shown.

Of those who did realise, some obtained values for x^3, and then went on to state that these were values for x (often because they had stated 'let $x = x^3$' as part of their working) and hence did not cube root in order to get values of x. Others recognised their values were for x^3, but thought that negative values could not be cube rooted. Nevertheless, most candidates obtained all 3 marks.

Question 6 (b)

(b) Determine the coordinates of the stationary points on the curve $y = 8x^7 + \frac{49}{4}x^4 - 7x$. **[4]**

Most candidates successfully differentiated the given expression. However, many did not show the process of setting this = 0 and solving (e.g. dividing through by a factor of 7 to obtain the equation from (a)). Again the command word 'Determine' means that working is required; some candidates appeared to simply assume that the answers would be the x-values from part (a). While most candidates were able to correctly calculate the corresponding y-values, some candidates missed this instruction and so were unable to access the final A mark.

Question 6 (c)

(c) For each of the stationary points, use the value of $\frac{d^2y}{dx^2}$ to determine whether it is a maximum or a minimum. **[4]**

Most candidates obtained a reasonable second derivative (FT was allowed here from their first derivative in part (b)). Note that the question asked candidates to 'use the value' of the second derivative, so we needed to see an explicit statement that this was < or > 0 at each point before concluding whether each point was a local minimum or maximum. If values were given, they needed to be correct. FT was allowed from candidates' second derivative and/or their values of *x*. Some candidates misinterpreted the question and differentiated the equation from part (a) ($8x^6$ + …), which if done correctly was awarded M1A0M1A0.

Question 7 (a) (i)

7 **(a)** Two real numbers are denoted by a and b.

(i) Write down expressions for the following.

- The mean of the squares of a and b
- The square of the mean of a and b **[1]**

The majority of candidates answered this part correctly, but there were a number of errors, primarily from not interpreting the question correctly. A common incorrect response was $\frac{(a+b)^2}{2}$ for the square of the mean.

Question 7 (a) (ii)

(ii) Prove that the mean of the squares of a and b is greater than or equal to the square of their mean. **[3]**

Very few candidates provided a fully correct solution to this part. The alternative method in the scheme was much more common than the 'difference' method. However, many candidates made errors when multiplying out the ('squared') brackets. A good number of candidates (who multiplied out correctly) reached a valid inequality, but very few spotted the need to factorise to $(a - b)^2$ to prove that this was true by showing it > 0. Candidates should be encouraged to approach 'Prove' or 'Show that' questions (even those for relatively few marks) with appropriate rigour, for instance noting that they are writing an inequality that they seek to prove, rather than assuming that the given result is true.

Many candidates used specific values to demonstrate the result, which gained no marks.

Question 7 (b)

(b) You are given that the result in part **(a)(ii)** is true for any two or more real numbers.

Explain what this result shows about the variance of a set of data. **[1]**

Very few candidates answered this part correctly, mainly because of failing to complete part 7 (a) (ii) (even though it could be answered independently). Some candidates identified a connection between the variance and the given statement, but did not connect this to the conclusion that the variance is always $\geqslant 0$.

Question 8

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

A circle has equation $x^2 + y^2 - 6x - 4y + 12 = 0$. Two tangents to this circle pass through the point (0, 1).

You are given that the scales on the x-axis and the y-axis are the same.

Find the angle between these two tangents. **[7]**

The scheme's first method (finding the centre of the circle and applying coordinate geometry) was most common. Many candidates made a good attempt at completing the square in *x* and *y* to find the centre and radius of the circle (although there were quite a few errors, e.g. concluding that the centre was (-3, -2) or (2, 3)). Many of these candidates then did not draw a correct diagram, often assuming that (0, 1) was on the circle and looking for a tangent through this point. Around half of candidates were able to show sufficient geometric understanding to achieve the middle M mark, but very few then went on to provide a complete solution using trigonometry to calculate the required angle.

Fewer candidates attempted the alternative algebraic method, of using $y = mx + 1$ as the equation of the line through (0, 1) and attempting to find the values of *m* that resulted in this being a tangent to the circle. Those who did however were generally more successful, substituting in correctly and recognising the need to use the determinant.

Exemplar 2

$x^2 + y^2 - 6x - 4y + 12 = 0$

centre $= (3, 2)$

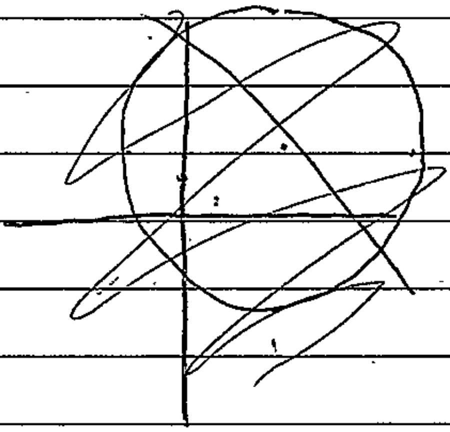
Radius $= \cancel{3}\ 1$

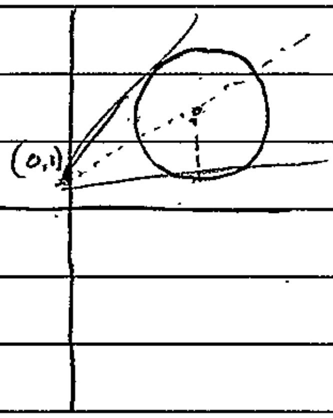
$(x - 3)^2 - 9 + (y - 2)^2 - 4 + 12 = 0$

$\cancel{(x-3)^2 + (y-2)^2 - 25}$

$(x - 3)^2 + (y - 2)^2 - 1 = 0$

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



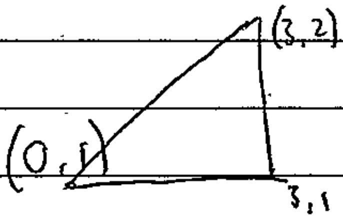


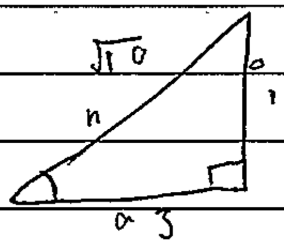
Centre $(3, 2)$ point $(0, 1)$ tangent at $(3, 1)$

Gradient $= \frac{1}{5}$

$y - 1 = \frac{1}{5}(x - 0)$

$y = \frac{1}{5}x + 1$





SOH
CAH
TOA

$\cos^{-1}\left(\frac{3}{\sqrt{10}}\right) = 18.43$

$18.43 \times 2 = 36.87°$

This is an example of an excellent response to this question. The erroneous gradient of $\frac{1}{5}$ was ignored because the candidate then restarts their attempt in the final 3 lines. This candidate has drawn an adequate diagram (albeit not fully labelled) and then gone on to show correct geometrical understanding with their final triangle, which has allowed them to compute the required angle.

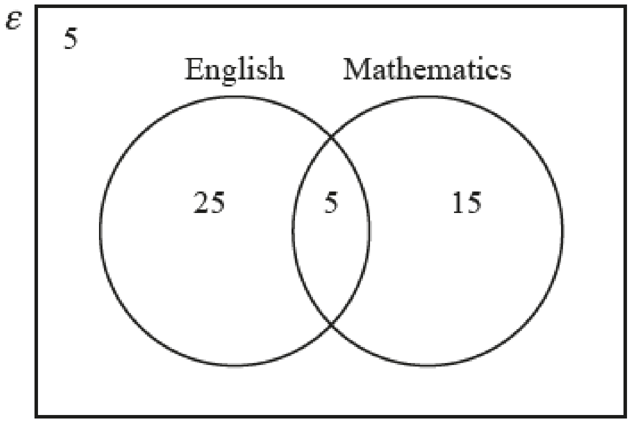
Section B overview

The majority of questions in the Statistics section of this paper were well answered, with candidates demonstrating a good understanding of the various techniques tested. However, many candidates struggled to provide sufficiently specific explanations in the 'Explain' questions and few candidates achieved full marks on the Large data set question (Question 12).

Question 9 (a)

9 In a survey, 50 people were asked whether they had passed A-level English and whether they had passed A-level Mathematics.

The numbers of people in various categories are shown in the Venn diagram.



(a) A person is chosen at random from the 50 people.

Find the probability that this person has passed A-level Mathematics. **[1]**

Almost all candidates answered this correctly.

Question 9 (b)

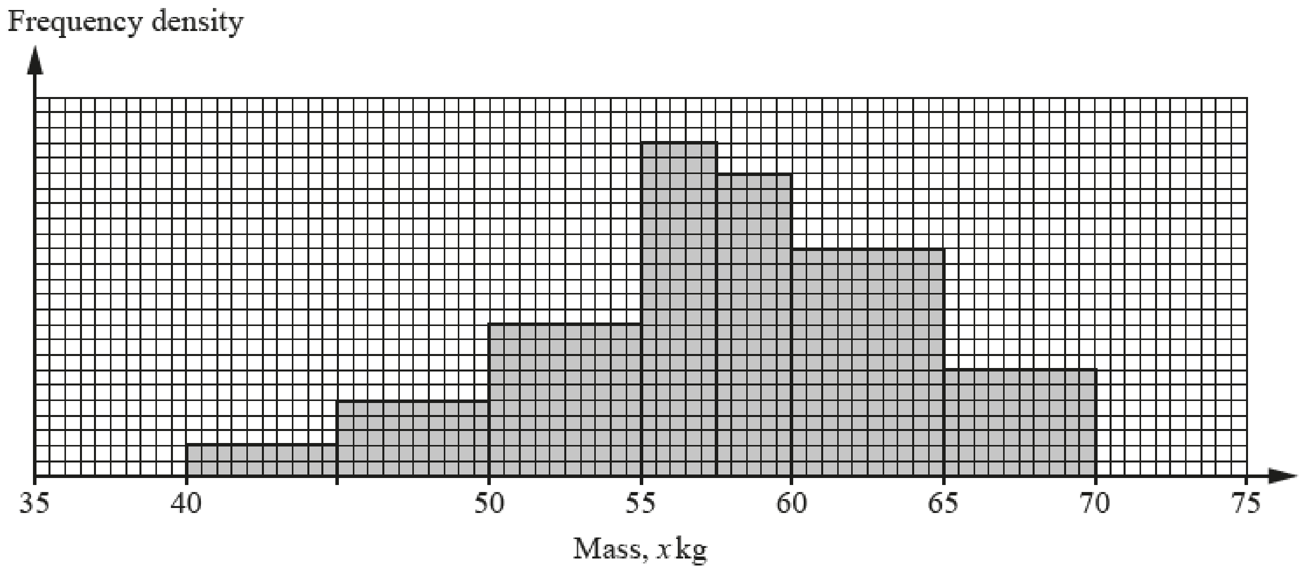
(b) Two people are chosen at random, **without replacement**, from those who have passed A-level in at least one of the two subjects.

Determine the probability that both of these people have passed A-level Mathematics. **[3]**

Fewer than half of candidates achieved full marks here. Many candidates over-complicated it by only using part of the information given in the question, for example drawing a tree diagram and attempting to sum the various combinations (which often resulted in errors, but did obtain the M mark). Those candidates who interpreted the given information carefully were able to write the 'one line' solution as given in the scheme. A common error was $\frac{20}{50} \times \frac{19}{49}$ (i.e. not accounting for the '5' students who did not pass English or Mathematics). Some candidates simplified $\frac{20}{45}$ to $\frac{4}{9}$ before multiplying and although this is valid, it occasionally led to errors in the subtraction from numerator and denominator required for 'without replacement'.

Question 10 (a)

10 The masses of a random sample of 120 boulders in a certain area were recorded. The results are summarized in the histogram.



(a) Calculate the number of boulders with masses between 60 and 65 kg. **[2]**

Most candidates obtained the correct answer of 30 boulders here (although some did so through much more laborious working than required). Those who didn't often forgot to apply the principle of frequency density, i.e. finding the proportion of squares without using the given total of 120.

The majority of candidates were able to work out 120 × $\frac{\text{Area of 60-65 block}}{\text{Total area}}$ = 30 boulders lay in the required range, although the working was not always clear.

The lack of units on the frequency density axis caused problems to those candidates who tried to involve frequency density somewhere in their working and they often produced an answer of 75 (from 5 × 15) as they did not consider that the total area of the histogram must equal the total frequency of 120.

Question 10 (b) (i)

(b) (i) Use midpoints to find estimates of the mean and standard deviation of the masses of the boulders in the sample. **[3]**

Quite a few candidates did not attempt this question (despite several of them having done much of the work in part (a)). Of those who did, a good number obtained the correct answers, but there were numerous small calculation errors (generally being awarded M1A0A0). The M mark was awarded for either one of the values in a reasonable range (even if no working was shown), but the A marks were given only for correct responses accurate to 3sf.

Candidates were expected to use their calculators to find their estimates of the mean and standard deviation, but they should take time to write a table of *x* and f values in their answer booklet to avoid these small calculation errors.

Some candidates did not demonstrate an understanding of how to find the mean for grouped data. $\frac{40 + 70}{2}$ was seen, as was taking an average of the 7 midpoints with no frequency involved.

Question 10 (b) (ii)

(ii) Explain why your answers are only estimates. **[1]**

Candidates should be encouraged to read 'Explain' questions carefully rather than reciting a 'default fact'. In this case, the question was specifically asking for a reason why these estimates of the mean and standard deviation of the **sample** are only estimates. That meant that answers comparing sample statistics to population statistics were not relevant and that the 'standard fact' of 'midpoints were used' was not adequate either. The clearest explanations made a specific reference to the distribution of masses within each class (and noted that this may not be uniform). Note that in this case the scheme did not require answers to be in context, but it is always good practice to do so.

To gain credit for this question, candidates needed to explain why using the midpoints meant the standard deviation was only an estimate and so they needed to show they understood that we did not know how the masses were distributed in each class. No credit was given for 'Because midpoints were used', 'Actual masses not known' or 'Midpoints used instead of actual values'.

'Explain' questions require an explanation, not a generic fact

Candidates should be reminded that Explain questions require an explanation that is specific to the question being asked (often, in context). The request is for candidates to demonstrate their understanding in the context of the question, not to state a fact.

Question 10 (c)

(c) Use your answers to part **(b)(i)** to determine an estimate of the number of outliers, if any, in the distribution. **[2]**

Few candidates applied the correct method here (noting that the question specified 'Use your answers to part (b) (i)' so the only acceptable method was $\mu \pm 2\sigma$). Answers attempting to find and use the IQR could not be credited. The scheme allowed follow-through for candidates who had made an error in calculating either value in (b)(i), but we did require that the estimate was consistent with their values. Centres should make sure that candidates are familiar with this method for estimating the number of outliers.

Many of the candidates who correctly computed $\mu \pm 2\sigma$ often did not go on and provide an estimate for the number of outliers present in the distribution. Candidates should be reminded to read each question carefully and check they have provided a full response.

Question 10 (d)

(d) Give **one** advantage of using a histogram rather than a pie chart in this context. **[1]**

Similarly to part (b) (ii) this question required a specific explanation to the question asked, not a generic fact. Common unacceptable responses included 'histogram shows the spread of data' or 'histograms are easier to read'. Candidates should aim to consider the comparison given (for instance, noting that pie charts can provide relative frequencies, but not actual frequencies).

Question 11 (a)

11 Casey and Riley attend a large school. They are discussing the music preferences of the students at their school. Casey believes that the favourite band of 75% of the students is Blue Rocking. Riley believes that the true figure is greater than 75%.

They plan to carry out a hypothesis test at the 5% significance level, using the hypotheses $\mathrm{H}_0\colon\ p = 0.75$ and $\mathrm{H}_1\colon\ p > 0.75$.

They choose a random sample of 60 students from the school, and note the number, X, who say that their favourite band is Blue Rocking.

They find that $X = 50$.

(a) Assuming the null hypothesis to be true, Riley correctly calculates that $\mathrm{P}(X = 50) = 0.0407$, correct to 3 significant figures.

Riley says that, because this value is less than 0.05, the null hypothesis should be rejected.

Explain why this statement is incorrect. **[1]**

Although intended as a useful 'hint' for part (b), many candidates did not spot the '=' instead of '$\geqslant$' error here and made a general comment on the validity of the hypothesis test instead. Some candidates gave the response 'cumulative', which was deemed too vague. Many candidates recognised the need for $\mathrm{P}(X \geqslant 50)$, although some just had $\mathrm{P}(X > 50)$.

Question 11 (b)

(b) Carry out the test. **[5]**

This hypothesis testing question was generally well answered. As in previous years, the mark scheme is worth reviewing to understand the structured answer that is required. The majority of students appeared well-trained in this structure and followed it closely (noting that, unusually, this paper did not require students to first state their hypotheses). Some candidates made an error in selecting the right probability value (e.g. using P(X = 50)), but could obtain some FT marks.

Most candidates gained the first M1 for using the correct binomial and 50. The comparison of their probability with 0.05 was generally clearly shown correctly, but the conclusion of whether to reject or not reject the null hypothesis did not always follow correctly.

Most candidates made an appropriate statement about H_0, but quite a few candidates lost the final mark for not giving their explanation in context or for making an inaccurate 'definite' statement.

Some candidates stated a conclusion along the lines of 'there is insufficient evidence to suggest that the probability of Blue Rocking being the favourite is increased', which shows a misunderstanding of the context (the probability has not been increased, the hypothesis is testing whether there is evidence that it is above a certain value). Some students, having made the correct initial statement ('do not reject H_0') went on to make a definitive statement that the null hypothesis is 'true', e.g. 'the proportion of students whose favourite band is BR is 75%'; this is not correct and is not what the test shows (insignificance does not 'prove' the null hypothesis).

Exemplar 3

let p be the probability of the favorite band of 75% students is Blue Rocking

$H_0 : p = 0.75$

$H_1 : p > 0.75$

$X \sim B(60, 0.75)$

$1 - P(X < 49)$

$1 - 0.914$

$= 0.086$ (3sf)

$\because 0.086 > 0.05$

$\therefore$ reject H_0.

There is sufficient evidence that the probability of the favourite band is of 75% students is Blue rocking, hence reject H_0.

This candidate has performed the first steps of the hypothesis test correctly (although we would ideally see the values to at least 3sf throughout). However, they have drawn the incorrect conclusion from their comparison, to 'reject H_0'. Their conclusion does not correctly follow from this statement, because it asserts that this shows the probability is 75%.

Question 11 (c) (i)

(c) **(i)** State which mathematical model is used in the calculation in part **(b)**, including the value(s) of any parameter(s). **[1]**

This question was asking for the straightforward statement X~B(60,0.75), which many candidates had already obtained and used in part (b). However, quite a few candidates misidentified the 'binomial theorem' and/or did not state the parameters. We condoned 'binomial expansion' if accompanied by the correct values for *n* and *p*. Some candidates stated 'binomial PD' or 'C–', which was allowed, but centres should make sure that students understand the underlying characteristics of these distributions beyond the formulae in their calculator.

Question 11 (c) (ii)

(ii) The random sample was chosen without replacement.

Explain whether this invalidates the model used in part **(b)**. **[1]**

Relatively few candidates obtained the mark here. This question required an answer in context explaining that the non-replacement does invalidate the test. Some candidates were able to recite the standard 'fact' that the binomial distribution does require replacement, but were unable to explain **in context** why this is the case. We needed to see evidence that the lack of replacement means that the probability of success (i.e. choosing a student whose favourite is BR) is not constant with each trial. Some candidates made reference to repeats or 'choosing the same person each time', which was not enough to gain the mark. Some candidates confused the number of trials and the overall number of students in the population from which the sample is drawn.

Question 12 (a)

12 This question deals with information about the populations of Local Authorities (LAs) in the North of England, taken from the 2011 census.

Fig. 1

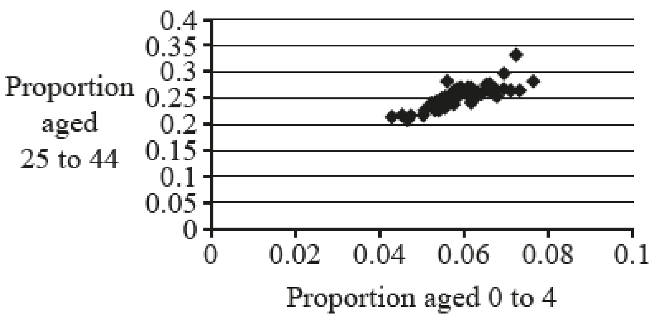


Fig. 2

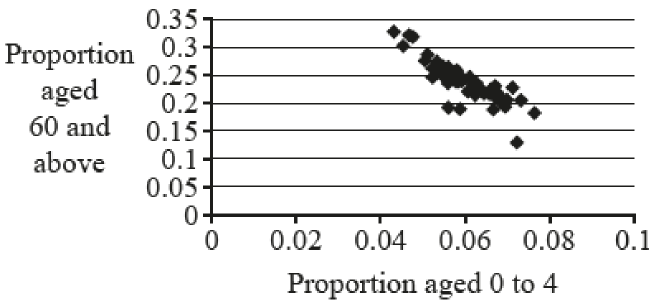


Fig. 1 and **Fig. 2** both show strong correlation, but of two different kinds.

(a) For each diagram, use a single word to describe the kind of correlation shown. **[1]**

Almost all candidates identified these correctly.

Question 12 (b)

(b) For each diagram, suggest a reason, in context, why the correlation is of the particular kind described in part **(a)**. **[2]**

Some candidates gave reasonable, plausible explanations here (e.g. '25-44s are more likely to have/have had young children so there is a correlation with the proportion of 0-4s'). The scheme included a SC for candidates who made an (incorrect) reference to number rather than proportion. Some candidates simply described the correlation, which could not be credited. A number of candidates, despite this being on the LDS, made incorrect statements about how populations change over time, e.g. 'people are living longer' for Fig. 2, which was not sufficient.

Many candidates ignored the labelling on Fig. 1 and Fig. 2 that referred to proportions and gave responses that referred to 'number' of adults and/or children. A good number of others ignored the instruction to 'suggest a reason, in context' and merely stated that in Fig. 1 that as one increased the other also increased, or equivalent and similar wording for Fig. 2. Many recognised that those aged 25 to 44 were more likely to be parents of young children while those aged 60 and above were more likely to have children who were older and many of these could have left their parents' home to set up their own homes.

Misconception

Where a question requires an answer 'in context' this means that answers not given in context will not be credited. In this case, candidates should use their knowledge of the Large data set to suggest a reason.

Question 12 (c)

Fig. 3 is the same as **Fig. 2** but with the point A marked.
Fig. 4 shows information about the same LAs as **Fig. 2** and **Fig. 3**.

Fig. 3

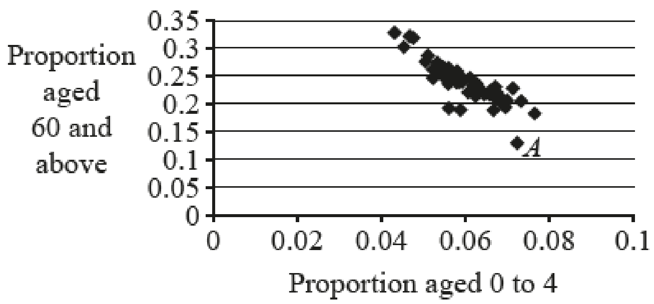
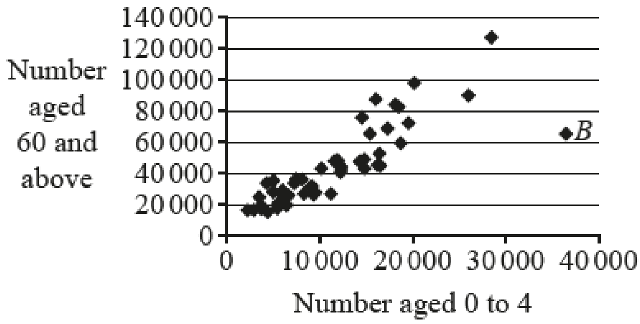


Fig. 4



(c) Point A in **Fig. 3** and point B in **Fig. 4** represent the same LA.

Explain how you can tell that this LA has a large population. **[1]**

Very few candidates gave a creditable response here. We needed to see an explicit link between number and proportion (i.e. between the two figures) to demonstrate that the given LA has a 'large' population (where large is meant in the statistical sense of 'large compared to other LAs'). Many candidates made a general statement or estimated the population (c. 530 000) and simply stated this was a large number, without any comparison (note that 530 000 is not an unusually large population for an LA).

Some candidates made inaccurate statements such as 'large proportion and large number means large population'. The question was specifically looking for candidates to identify the large number and 'medium' proportion of those aged 0-4, or the 'medium' number and low proportion of those aged 60 or over, to demonstrate that its overall population is unusually large compared to other LAs.

Some candidates looked to obtain an estimate of the population size, concluding that this was large with no justification. A number of others made reference to only Fig. 4 ignoring Fig. 3. Only the most successful responses related the two figures and saw candidates comment on the given LA's relationship to the other LAs and hence give an appropriate explanation.

Supporting you

Teach Cambridge

Make sure you visit our secure website **Teach Cambridge** to find the full range of resources and support for the subjects you teach. This includes secure materials such as set assignments and exemplars, online and on-demand training.

Don't have access? If your school or college teaches any OCR qualifications, please contact your exams officer. You can forward them this link to help get you started.

Reviews of marking

If any of your students' results are not as expected, you may wish to consider one of our post-results services. For full information about the options available visit the OCR website.

Access to Scripts

For the June 2023 series, Exams Officers will be able to download copies of your candidates' completed papers or 'scripts' for all of our General Qualifications including Entry Level, GCSE and AS/A Level. Your centre can use these scripts to decide whether to request a review of marking and to support teaching and learning.

Our free, on-demand service, Access to Scripts is available via our single sign-on service, My Cambridge. Step-by-step instructions are on our website.

Keep up-to-date

We send a monthly bulletin to tell you about important updates. You can also sign up for your subject specific updates. If you haven't already, sign up here.

OCR Professional Development

Attend one of our popular CPD courses to hear directly from a senior assessor or drop in to a Q&A session. Most of our courses are delivered live via an online platform, so you can attend from any location.

Please find details for all our courses for your subject on **Teach Cambridge**. You'll also find links to our online courses on NEA marking and support.

Signed up for ExamBuilder?

ExamBuilder is the question builder platform for a range of our GCSE, A Level, Cambridge Nationals and Cambridge Technicals qualifications. Find out more.

ExamBuilder is **free for all OCR centres** with an Interchange account and gives you unlimited users per centre. We need an Interchange username to validate the identity of your centre's first user account for ExamBuilder.

If you do not have an Interchange account please contact your centre administrator (usually the Exams Officer) to request a username, or nominate an existing Interchange user in your department.

Active Results

Review students' exam performance with our free online results analysis tool. It is available for all GCSEs, AS and A Levels and Cambridge Nationals.

Find out more.

Need to get in touch?

If you ever have any questions about OCR qualifications or services (including administration, logistics and teaching) please feel free to get in touch with our customer support centre.

Call us on
01223 553998

Alternatively, you can email us on
support@ocr.org.uk

For more information visit
- **ocr.org.uk/qualifications/resource-finder**
- **ocr.org.uk**
- **facebook.com/ocrexams**
- **twitter.com/ocrexams**
- **instagram.com/ocrexaminations**
- **linkedin.com/company/ocr**
- **youtube.com/ocrexams**

We really value your feedback

Click to send us an autogenerated email about this resource. Add comments if you want to. Let us know how we can improve this resource or what else you need. Your email address will not be used or shared for any marketing purposes.

I like this **I dislike this**

Please note – web links are correct at date of publication but other websites may change over time. If you have any problems with a link you may want to navigate to that organisation's website for a direct search.

CAMBRIDGE
UNIVERSITY PRESS & ASSESSMENT