Qualification Accredited

OCR
Oxford Cambridge and RSA

A LEVEL

Examiners' report

MATHEMATICS A

H240
For first teaching in 2017

H240/02 Summer 2019 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. 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 can be downloaded from OCR.

Paper 2 series overview

This is the second series of the reformed linear A Level Maths specification, and the first to be sat by candidates following the standard two year programme. The new specification includes a greater emphasis upon modelling, proof and interpretation (rather than calculation). These emphases are reflected in this paper.

Some questions contain specific defined 'command words', in particular the instruction '**In this question you must show detailed working**'. In these questions, candidates are required to demonstrate their understanding of the relevant concepts by showing their working, rather than by presenting an answer gained simply by pressing a few buttons. Consequently, in these questions, marks will not be given unless correct working is seen. Remember that this does not preclude candidates from checking their working using the calculator.

	OCR support	A poster detailing the different command words and what they mean is available here: https://teach.ocr.org.uk/italladdsup

Conversely, some questions have slightly lower mark tariffs than seen in the legacy assessment, where candidates are expected to make efficient use of the full range of functions on their calculator. These are also signposted by the 'command words' used in the question: 'Find', 'Calculate', 'Write down'.

This paper covered pure mathematics in section A and statistics, including use of the large data set, in Section B. There was a wide spread of marks, including a significant proportion of candidates above 90%.

Poor handwriting was a very significant issue - more so than in the past. There were cases where marks simply could not be awarded because the handwriting was indecipherable.

Sometimes a candidate needed more space than was available for a particular question. There were two blank pages at the end of the answer booklet that were available for this purpose, but some candidates ignored these and wrote on additional sheets. This is, of course, quite acceptable. However, the use the extra pages within the booklet, rather than extra sheets, simplifies matters for candidates, for centre administration, for scanners and for markers.

The written answers required in all parts of Question 4 and Question 11 parts (b)(ii), (c) and (d) involving interpretation of data were variable in quality. Centres might be advised to note the following. The provision of an answer space that includes about six lines is intended to accommodate those candidates who insist on writing short essays. However, the best answers generally can be written on one line, or two at the most. The number of marks, rather than the number of answer space lines is a better rough guide to the amount of detail needed in a response.

Section A overview

Many candidates' algebra was good in this Section, although the omission of brackets was not uncommon. Incorrect "cancelling" was sometimes seen, particularly in Questions 1(a)(i) and 5(b).

The specification contain explanations of command words such as "Show that …" and "Determine . . ." and of the instruction "**In this question you must show detailed reasoning.**" The marking of Questions 3(b), 5(a), 5(b) and 7 reflected this new aspect of the specification. Responses which included less than adequate explanation or working did not score full marks. In particular, correct answers given to these questions without justification scored few marks. Trial and improvement were not regarded as an acceptable method in these questions, (and indeed is sometimes treated as unacceptable even in questions not containing such instructions).

The comments about the interpretation of data in the overview section apply to all parts of Question 4.

Question 1 (a) (i)

1 **(a)** Differentiate the following.

(i) $\frac{x^2}{2x+1}$ **[3]**

Most candidates answered this question well, using either the quotient rule or the product rule. Some candidates simplified their answers incorrectly.

Question 1 (a) (ii)

(ii) $\tan(x^2 - 3x)$ **[2]**

Many correct answers were seen. Some candidates omitted one or both pairs of brackets in their answer. Most knew that the derivative of $\tan x$ is $\sec^2 x$, although a few resorted to the quotient rule applied to $\frac{\sin\left(x^2-3x\right)}{\cos\left(x^2-3x\right)}$. A few candidates gave $\sec^2(2x-3)$. Some candidates used the $\tan(A-B)$ formula. Some then stopped, while others attempted to differentiate the result, usually making errors.

Question 1 (b)

(b) Use the substitution $u = \sqrt{x} - 1$ to integrate $\dfrac{1}{\sqrt{x}-1}$. **[4]**

The majority of candidates found $\dfrac{du}{dx} = 0.5x^{-0.5}$ and attempted to express $x^{-0.5}$ in terms of u. Many of these candidates made mistakes in the algebra. Some found $2\sqrt{x}\int \frac{1}{u}du$. Candidates who found x in terms of u and then found $\dfrac{dx}{du}$ were much less likely to make algebraic errors. Some candidates substituted u for the denominator, but then just replaced dx by du. A few omitted the "$+c$".

Question 1 (c)

(c) Integrate $\dfrac{x-2}{2x^2-8x-1}$. **[2]**

Many candidates recognised that the numerator is a multiple of the derivative of the denominator. Most of these gave $\ln|2x^2 - 8x - 1|$, but some multiplied by 4, instead of dividing by 4. Some omitted brackets or "$+c$". Those who did not recognise the basic form of the integrand attempted various strategies such as partial fractions or integration by parts, but without success.

Question 2 (a)

2 **(a)** Find the coefficient of x^5 in the expansion of $(3-2x)^8$. **[1]**

Most candidates answered this correctly, although a few omitted the "–" sign.

Question 2 (b) (i)

(b) **(i)** Expand $(1+3x)^{0.5}$ as far as the term in x^3. **[3]**

Many answered correctly. Some made an error with the "3" and gave $1 + \frac{3}{2}x - \frac{3}{8}x^2 + \frac{3}{16}x^3$. Others omitted the "3" altogether.

Question 2 (b) (ii)

(ii) State the range of values of x for which your expansion is valid. **[1]**

Again, many answered this correctly. Some incorrect answers were $|3x| < 1$, $x < \frac{1}{3}$, $x > -\frac{1}{3}$, $|x| < 1$ and $x < \left|\frac{1}{3}\right|$.

Question 2 (b) (iii)

A student suggests the following check to determine whether the expansion obtained in part **(b)(i)** may be correct.

"Use the expansion to find an estimate for $\sqrt{103}$, correct to five decimal places, and compare this with the value of $\sqrt{103}$ given by your calculator."

(iii) Showing your working, carry out this check on your expansion from part **(b)(i)**. **[3]**

This part was challenging for most candidates, with candidates not spotting the link to the previous work. Some used their series successfully, but lost a mark by not showing explicitly their substitution of $x = 0.01$ into the series. The majority attempted $1+3x = 103$, giving $x = 34$ and then substituted $x = 34$ into their expansion, apparently not recognising that this is invalid. A few candidates managed to answer correctly with other substitutions.

Question 3 (a) (i)

3 **(a)** A circle is defined by the parametric equations $x = 3+2\cos\theta$, $y = -4+2\sin\theta$.

(i) Find a cartesian equation of the circle. **[2]**

Many candidates found the standard cartesian form of the equation, although a few made sign errors. Some began with $(x-a)^2+(y-b)^2 = r^2$ and attempted to find the values of *a*, *b* and *r*. Some found *a* and *b* correctly, but not *r*.

Other candidates gave clumsy, but correct, alternative forms such as $\cos^{-1}\left(\frac{x-3}{2}\right) = \sin^{-1}\left(\frac{y+4}{2}\right)$

or $y = -4+2\sin\left(\cos^{-1}\left(\frac{x-3}{2}\right)\right)$ or $y = -4+2\sqrt{1-\left(\frac{x-3}{2}\right)^2}$.

Some candidates gave answers that were correct equations in themselves, but involved a mixture of x, y and θ, for example $y = (x-3)\tan\theta - 4$ or $x-3-2\cos\theta = y+4-2\sin\theta$

Some candidates started by finding x^2+y^2, which unfortunately leads nowhere. Disappointingly, many of these candidates also squared incorrectly, such as $(3+2\cos\theta)^2 = 9+4\cos^2\theta$

A significant number of candidates did not recognise that the key to this problem is the identity $\sin^2\theta+\cos^2\theta = 1$. Some candidates (perhaps a reflex triggered by anything involving parametric equations) found $\frac{dy}{dx}$ by dividing $\frac{dy}{d\theta}$ by $\frac{dx}{d\theta}$.

Question 3 (a) (ii)

(ii) Write down the centre and radius of the circle. **[1]**

Most candidates who answered part (a)(i) correctly also answered this part correctly. Those whose answer to answers to (a)(i) were incorrect, but in the standard form, usually answered this part correctly following through their answer to (a)(i). Candidates whose answers to (a)(i) were in terms of *x y* and *θ* sometimes also answered this part in terms of *x y* and *θ*.

Question 3 (b)

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

The curve S is defined by the parametric equations $x = 4\cos t$, $y = 2\sin t$. The line L is a tangent to S at the point given by $t = \frac{1}{6}\pi$.

Find where the line L cuts the x-axis. **[6]**

Most candidates recognised that the starting point was to find the gradient of the curve at the relevant point. Most of these successfully found $\frac{dy}{dx}$ by dividing $\frac{dy}{dt}$ by $\frac{dx}{dt}$. A few made errors while substituting $t = \frac{\pi}{6}$, such as omitting the "–" sign. Many candidates carried out the next two steps (finding the equation of the line and substituting $y = 0$) correctly. However, arithmetical errors were legion.

It is worth noting that some candidates combined these steps into one, not giving explicitly the equation of the line before substituting $y = 0$. This is not incorrect and was not penalised, but it possibly led to more errors than would otherwise have occurred. It also certainly made the marking more difficult, and sometimes marks could not be awarded because candidates' working was not clear enough.

Some candidates found the gradient of the normal and used this as the gradient of the tangent.

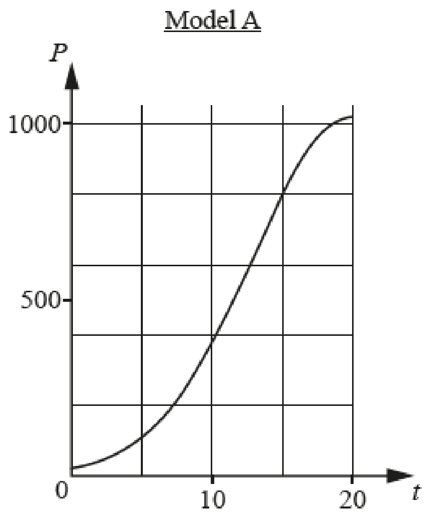
Some candidates retained "*t*" in their working until a late stage, only substituting $t = \frac{\pi}{6}$ at the end. This led to working being more complex and confusing than necessary, and sometime led to errors along the way.

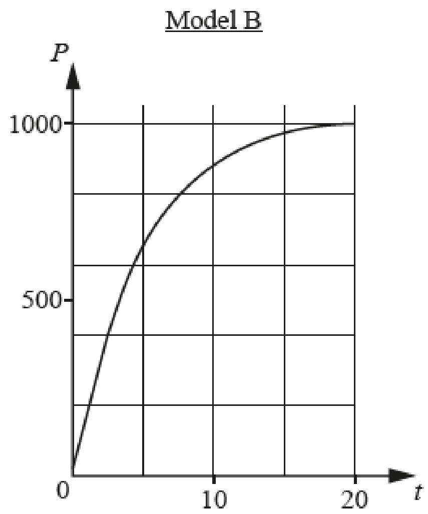
A few candidates found the point where the curve cuts the *x*-axis, instead of where the line cuts the *x*-axis.

Question 4

Centres are referred to the general comment on Question 4 under "Section A Overview" above. For the types of answers that were acceptable in all parts of Question 4, centres are referred to the published mark scheme.

4 A species of animal is to be introduced onto a remote island. Their food will consist only of various plants that grow on the island. A zoologist proposes two possible models for estimating the population P after t years. The diagrams show these models as they apply to the first 20 years.





(a) Without calculation, describe briefly how the rate of growth of P will vary for the first 20 years, according to each of these two models. **[1]**

Question 4 (a)

Some candidates were unable to distinguish between P and the rate of growth of P.

A few gave an answer for Model A, but not Model B. Some gave an incomplete answer for Model A, such as "The rate of growth increases and then becomes constant." Others tried to give a single answer that covered both models, such as "Growth overall decreases to zero" or "The rate of growth is greater under Model B than Model A."

A good succinct answer is as follows:

A: The rate of growth starts at zero, then increases and then decreases to zero.
B: The rate of growth is initially high and decreases to zero.

An acceptable succinct answer is as follows:

A: The rate of growth increases and then decreases.
B: The rate of growth decreases.

Question 4 (b) (i)

The equation of the curve for model A is $P = 20 + 1000\mathrm{e}^{-\frac{(t-20)^2}{100}}$.

The equation of the curve for model B is $P = 20 + 1000\left(1 - \mathrm{e}^{-\frac{t}{5}}\right)$.

(b) Describe the behaviour of P that is predicted for $t > 20$

(i) using model A, **[1]**

This part (and the next) depend on an ability to interpret the relevant equation involving an exponential function. Some candidates were able to interpret the equation correctly without quoting particular values of *P*. Others calculated values but were unable to give a proper interpretation. In some cases, candidates' answers to this part fitted the next part, and vice versa, suggesting a misunderstanding of the two equations. Some examples of inadequate or incorrect answers were as follows.

P will decrease. *P* decreases exponentially after $t = 20$. P tends to zero. *P* will grow exponentially. The population will grow to about 1000 and then drop. *P* will tend to 1020. *P* will grow, but very slowly. The curve will be a reflection of the given curve.

A good succinct answer is as follows:

After $t = 20$ *P* will decrease, tending to 20.

Question 4 (b) (ii)

(ii) using model B. **[1]**

Some of the comments on the previous part apply also to this part. Some examples of inadequate or incorrect answers were as follows.

The growth of the population will slow until it stops. *P* will rise but increasingly slowly. *P* will increase. *P* will increase exponentially. *P* tends to 1000. The population is at a fixed number. *P* will decay exponentially. *P* will decrease and tend to 20.

A good succinct answer is as follows:

After $t = 20$ *P* will continue to increase, tending to 1020.

Question 4 (c) (i)

There is only a limited amount of food available on the island, and the zoologist assumes that the size of the population depends on the amount of food available and on no other external factors.

(c) State what is suggested about the long-term food supply by

(i) model A, **[1]**

Candidates whose answers to (b)(i) were incorrect were unlikely to answer this part correctly. Some examples of inadequate or incorrect answers were as follows.

Food will decrease. Food will not sustain a large population. It is limited. Food will cap growth at 1000, but after that the population will degrade as there is less food. Consistent long term food supply. There's enough in the long term. There will be less food available in the long term.
It will not be enough.

Two good succinct answers are as follows.

The food will almost run out.
There will only be enough food to support 20 animals.

Question 4 (c) (ii)

(ii) model B. **[1]**

Many candidates answered this part well, recognising that the population equation implies that, in the long term, the food supply will not run out, but will be limited. Some examples of inadequate or incorrect answers were as follows.

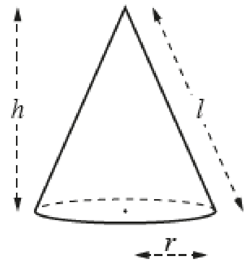
Food will be enough. Food will not run out. Food will be limited. There will be plenty of food. Food will be enough to support 20 animals. It will last longer as the population decreases. There will always be enough food for the population to grow. Food will nearly run out. The food supply will start to decrease. There will be lots of food for 10 years, then it decreases.

Two good succinct answers are as follows.

The food supply will be just sufficient to support about 1020 animals.
There will be enough food to sustain a stable population.

Question 5 (a)

5



For a cone with base radius r, height h and slant height l, the following formulae are given.

Curved surface area, $S = \pi r l$

Volume, $V = \frac{1}{3}\pi r^2 h$

A container is to be designed in the shape of an inverted cone with no lid. The base radius is r m and the volume is V m^3.

The area of the material to be used for the cone is 4π m^2.

(a) Show that $V = \frac{1}{3}\pi\sqrt{16r^2 - r^6}$. **[4]**

Many candidates gave convincing answers, showing sufficient steps to gain the marks. A particularly neat method, that was seen, involved finding h in terms of *l* and *r*, noting that *lr* = 4, and substituting both these into the formula for *V*.

However, a great deal of incorrect algebra was seen, for example $\sqrt{\frac{16}{r^2} - r^2} = \frac{4}{r} - r$ and $r^2\sqrt{\frac{16}{r^2} - r^2} = \sqrt{16 - r^4}$. Some candidates gave incorrect working and then "fudged" their working in order obtain the given answer. Many candidates did not appear to be aware of the fact that the key was to express everything in terms of *r*.

Question 5 (b)

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

It is given that V has a maximum value for a certain value of r.

Find the maximum value of V, giving your answer correct to 3 significant figures. **[5]**

Most attempted to differentiate V, with varying levels of success. Some made arithmetical errors. Others had $\frac{dV}{dr} = \frac{\pi}{6}\left(16r^2 - r^6\right)^{0.5} \times \left(32r - 6r^5\right)$ or $\frac{\pi}{6}\sqrt{32r - 6r^5}$. Some differentiated correctly, but then (incorrectly) set $16r^2 - r^6 = 0$, as well as (correctly) setting $32r - 6r^5 = 0$.

The 'Detailed reasoning' instruction indicates that the solution must lead to a conclusion by showing a detailed and complete analytical method; this meant that candidates were expected to give a reason for rejecting one or both of $r = 0$ and $r = -\frac{2}{\sqrt[4]{3}}$. Similarly, a 'Trial and Improvement' method to identify $r = 1.52$, seen by a few candidates, was not considered an adequate method for full credit.

Question 6

6 Shona makes the following claim.

"n is an **even** positive integer greater than $2 \Rightarrow 2^n - 1$ is **not** prime"

Prove that Shona's claim is true. **[4]**

Not many candidates answered this successfully, although many at least began by writing $2^{2k} - 1$ and thereby gained a mark. Some gained another mark by factorising this expression correctly, but very few went on to show that neither factor was equal to 1.

Many candidates misinterpreted the question to mean:

"Prove that the following statement is **not** true: n is an even number greater than 2 $\Rightarrow 2^n - 1$ is prime."

They then disproved this statement by assuming it to be true and quoting a numerical example that contradicted it and claimed to have answered the question using Proof by Contradiction.

Some candidates calculated $2^2 - 1$, $2^4 - 1$, $2^6 - 1$, $2^8 - 1$ etc., up to about $2^{20} - 1$ and noted that in each case the result is not prime, and incorrectly assumed that the result has been proved by "Proof by exhaustion".

A few candidates (who presumably were studying Further Maths) gave a proof by induction, proving that, for n even and greater than 2, $2^n - 1$ is divisible by 3.

Question 7

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

Use the substitution $u = 6x^2 + x$ to solve the equation $36x^4 + 12x^3 + 7x^2 + x - 2 = 0$. **[5]**

Candidates who started by finding u^2 were generally successful. Many other candidates started by dividing the left hand side of the given equation by $6x^2 + x$. Those who handled the remainder correctly generally went on to be successful. Many, however, after dividing and finding a remainder of 2, said that

$36x^4 + 12x^3 + 7x^2 + x - 2 = \left(6x^2 + x\right)\left(6x^2 + x + 1\right) - \frac{2}{6x^2 + x}$, which is incorrect.

Some candidates arrived correctly at the two equations $6x^2 + x - 1 = 0$ and $6x^2 + x + 2 = 0$. However, because this question contains the instruction "**In this question you must show detailed reasoning**", candidates needed to show (not just state) that the second equation has no real roots. Many candidates lost a mark by failing to do this.

This question did prove to be challenging for many candidates, with a range of attempts involving a great deal of algebraic manipulation, in terms of *x* and *u*. Some used their calculator to find the roots directly, and then used these in a "fudged" attempt to provide some justification. Whilst the use of calculators to check work should be encouraged, this question required the explicit use of the substitution to be seen in order to gain credit.

Section B overview

In Questions 9(b) and 12(c)(ii), where the command word is "Determine", correct answers without sufficient working did not receive full marks.

On the other hand, it is worth noting that a small mark tariff for a question is sometimes a strong pointer to the fact that candidates are free to use their calculator functions, without necessarily showing working. Examples are Questions 8(b), 9(a)(i) and 9(a)(ii). This is a feature of the reformed qualification where efficient use of technology is expected (a change from the legacy, where there was always an expectation that all working should be shown).

The comments about interpretation of data in the overview section apply to Question 11, parts (b)(ii), (c) and (d).

Question 8 (a)

8 The stem-and-leaf diagram shows the heights, in centimetres, of 17 plants, measured correct to the nearest centimetre.

5	5 7 9 9
6	3 4 5 5 5 9 9
7	4 5 7 9 9
8	
9	9

Key: 5 | 6 means 56

(a) Find the median and inter-quartile range of these heights. **[3]**

Most candidates found the median correctly. However, many candidates used a slightly incorrect method to find the quartiles. For example: $\frac{17}{4} = 4.25$, so the Lower Quartile is 60. Textbooks vary in the method for finding the quartiles, but in the case of these 17 heights, any correct method leads to the value 61cm. Perhaps the simplest way to arrive at this is to find the median of the lowest 8 heights.

Question 8 (b)

(b) Calculate the mean and standard deviation of these heights. **[2]**

Most candidates found both values correctly, although some made arithmetical errors. Those who simply entered the values in the calculator and used the relevant functions were generally successful. Those who used the "long-hand" method sometimes made arithmetical errors. It is worth noting that there were only two marks for this question, which is a strong pointer to the fact that candidates are free to use their calculator functions, without showing working, as is the command word 'Calculate'.

Question 8 (c)

(c) State one advantage of using the median rather than the mean as a measure of average for these heights. **[1]**

Questions almost identical to this one have been asked (and commented upon in Examiners' reports) for many years, in legacy paper Statistics 1 (4732). Nevertheless, many candidates made the usual error of giving a correct generic answer, but not related to the context of the question data. Many wrote statements that began with "If there are any outliers..." without apparently having considered whether there are actually any outliers in the data in this question. Some other candidates gave inadequate reasons such as "The median is the middle value, and so is more representative". Some claimed that the median is "more accurate" for one of a number of possible reasons.

Question 9 (a) (i)

9 **(a)** The masses, in grams, of plums of a certain kind have the distribution N(55, 18).

(i) Find the probability that a plum chosen at random has a mass between 50.0 and 60.0 grams. **[1]**

This question was well answered.

Question 9 (a) (ii)

(ii) The heaviest 5% of plums are classified as extra large.

Find the minimum mass of extra large plums. **[1]**

This question was also well answered.

Question 9 (a) (iii)

(iii) The plums are packed in bags, each containing 10 randomly selected plums.

Find the probability that a bag chosen at random has a total mass of less than 530 g. **[4]**

A disappointingly small number of candidates recognised that this question could be answered using the distribution of $\overline{X}$, the sample mean. Those who did recognise this were generally successful. A few candidates used the formulae (not in the specification, but acceptable) $\mathrm{E}(X_1 + X_2 + X_3 + \ldots X_n) = n\mathrm{E}(X)$ and $\mathrm{Var}(X_1 + X_2 + X_3 + \ldots X_n) = n\mathrm{Var}(X)$, and answered correctly. However, some used $n^2\mathrm{Var}(X)$, and so obtained a standard deviation of $\sqrt{1800}$ instead of $\sqrt{180}$.

Many assumed that each plum had to have a mass of less than 53 g. These found $(\mathrm{P}(X = 53))^{10}$.

Question 9 (b)

(b) The masses, in grams, of apples of a certain kind have the distribution $N(67, \sigma^2)$. It is given that half of the apples have masses between 62 g and 72 g.

Determine σ. **[5]**

Many candidates recognised that they needed to use either $\phi^{-1}(0.25)$ or $\phi^{-1}(0.75)$ and went on to obtain the correct answer. Some others used $\phi^{-1}(0.5)$, obtaining no marks.

No candidate attempting a Trial and Improvement method showed sufficient detail

10 The level, in grams per millilitre, of a pollutant at different locations in a certain river is denoted by the random variable X, where X has the distribution $N(\mu, 0.000\,040\,9)$.

In the past the value of μ has been 0.0340.

This year the mean level of the pollutant at 50 randomly chosen locations was found to be 0.0325 grams per millilitre.

Test, at the 5% significance level, whether the mean level of pollutant has changed. **[7]**

to be awarded full marks and generally scored 3 out of 5. See published mark scheme for example of level of detail required. Question 10

Many candidates appeared to have been well-drilled in hypothesis test questions and gave very good answers. Some of these candidates lost one mark because they did not define the "μ" in their hypotheses. Others made arithmetical errors in their calculation of a probability or a critical value. Some other common errors were as follows.

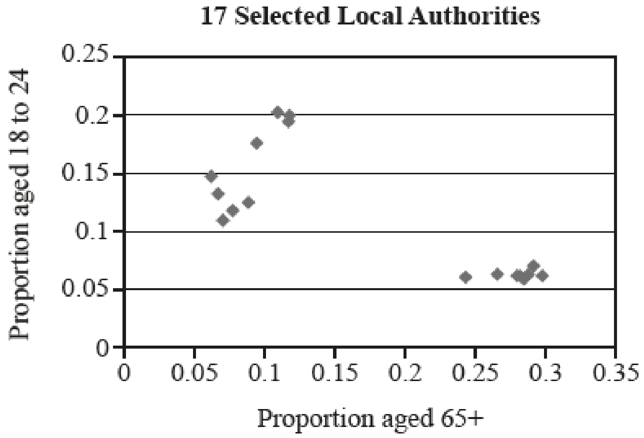
Misunderstanding the N(..., ...) notation, e. g. using a variance of $0.000\,040\,9^2$.
Failure to divide the variance by 50.
H_1: $\mu = 0.0325$ or H_1: $\mu < 0.0340$.
Use of 1.645 instead of 1.96 in finding the critical value or critical region, despite a correct H_1.
Claiming that because $0.0486 > 0.025$, H_0 should be rejected.
Giving the final conclusion as a definite statement, such as "The mean pollutant level has not changed".

Centres are referred to the exemplars at the end of the published mark scheme for further guidance.

Question 11

Centres are referred to the general comment on Question 11 under "Section A Overview" above. For the types of answers that were acceptable in Question 11 parts (b)(ii), (c) and (d), centres are referred to the published mark scheme.

11 A trainer was asked to give a lecture on population profiles in different Local Authorities (LAs) in the UK. Using data from the 2011 census, he created the following scatter diagram for 17 selected LAs.



He selected the 17 LAs using the following method. The proportions of people aged 18 to 24 and aged 65+ in any Local Authority are denoted by P_{young} and P_{senior} respectively. The trainer used a spreadsheet to calculate the value of $k = \frac{P_{\text{young}}}{P_{\text{senior}}}$ for each of the 348 LAs in the UK. He then used specific ranges of values of k to select the 17 LAs.

(a) Estimate the ranges of values of k that he used to select these 17 LAs. **[2]**

Question 11 (a)

Few candidates answered this correctly. Most candidates did not understand which pairs of proportions to divide in order to find the limiting values of *k*. A helpful approach would be to draw a line from the origin, just meeting the point in a cluster nearest to the line *y* = *x*. This point provides the correct pair of proportions for one of the limiting values of *k*. Then repeat for the other cluster.

Question 11 (b) (i)

(b) Using the 17 LAs the trainer carried out a hypothesis test with the following hypotheses.

H_0: There is no linear correlation in the population between P_{young} and P_{senior}.
H_1: There is negative linear correlation in the population between P_{young} and P_{senior}.

He found that the value of Pearson's product-moment correlation coefficient for the 17 LAs is -0.797, correct to 3 significant figures.

(i) Use the table on page 9 to show that this value is significant at the 1% level. **[2]**

Most candidates answered this correctly. Two common errors were "–0.797 < 0.5577, hence significant" and –0.797 < –0.6055 (i.e. a two-tail test).

Question 11 (b) (ii)

The trainer concluded that there is evidence of negative linear correlation between P_{young} and P_{senior} in the population.

(ii) Use the diagram to comment on the reliability of this conclusion. **[2]**

Many candidates thought the statement was reliable, because the diagram does show negative correlation. Many other candidates said the conclusion was not reliable because the diagram omits data from a large number of LAs. A few noted the key point that the points on the diagram are in two clusters. But not many explained that this results in an apparent (but false) appearance of good correlation. A few candidates noted, correctly, that the points cannot be contained roughly within an ellipse and therefore the distribution was probably not bivariate normal, so the use of Pearson's coefficient is invalid.

Question 11 (c)

(c) Describe one outstanding feature of the population in the areas represented by the points in the bottom right hand corner of the diagram. **[1]**

Many candidates gave good answers. Some, incorrectly, stated that the diagram showed that relevant LAs had a large <u>number</u> of older residents.

Question 11 (d)

(d) The trainer's audience included representatives from several universities.

Suggest a reason why the diagram might be of particular interest to these people. **[1]**

There was a large variety of answers. Perhaps the best answers referred to the usefulness of the diagram in choosing areas where a university might concentrate its advertising. Any answer was accepted that implied that the diagram enabled members of the audience to identify areas where there was a high proportion of young people.

Question 12 (a)

12 A random variable X has probability distribution defined as follows.

$$P(X = x) = \begin{cases} kx & x = 1, 2, 3, 4, 5, \\ 0 & \text{otherwise,} \end{cases}$$

where k is a constant.

(a) Show that $P(X = 3) = 0.2$. **[3]**

Most candidates answered this question correctly. Many, however, (perhaps misunderstanding the phrase "where *k* is a constant") assumed that all five probabilities were equal, and therefore each probability was 0.2

(b) Show in a table the values of X and their probabilities. **[2]**

Question 12 (b)

Almost all candidates answered this correctly, or incorrectly but consistently with their method for part (a).

Question 12 (c) (i) and (c)(ii)

In both these parts, some candidates either assumed that the probabilities for all the values of *X* were equal, or (which is equivalent) drew up a two-way table of all the possible outcomes and counted up the relevant ones. These were able to gain only a few marks. Some candidates used one of these incorrect methods, despite having answered parts (a) and (b) correctly.

(c) Two independent values of X are chosen, and their total T is found.

(i) Find $P(T = 7)$. **[3]**

Question 12 (c) (i)

Many candidates recognised that for $T = 7$ could be found by considering 2+5 and 3+4, but forgot to consider 5+2 and 4+3.

Exemplar 1

$2 + 5$ or $3 + 4$

$P(2+5) = \frac{2}{15} \times \frac{1}{3} = \frac{2}{45}$

$P(3+4) = \frac{1}{5} \times \frac{4}{15} = \frac{4}{75}$

$\left(\frac{2}{45} + \frac{4}{75}\right) = \frac{22}{225}$

This candidate has shown clear working, but forgotten to multiple by two in order to consider all the ways to score $T = 7$.

Question 12 (c) (ii)

(ii) Given that $T = 7$, determine the probability that one of the values of X is 2. **[4]**

Most candidates attempted to use the conditional probability formula and, correctly, placed their answer to part (c)(i) in the denominator of a fraction. But many were unable to identify the pairs of values that gave rise to "*T* = 7 AND one of the values is 2".

Question 13

13 It is known that 26% of adults in the UK use a certain app. A researcher selects a random sample of 5000 adults in the UK. The random variable X is defined as the number of adults in the sample who use the app.

Given that $P(X < n) < 0.025$, calculate the largest possible value of n. **[5]**

Many candidates, sensibly, used the Normal approximation to the Binomial distribution, which gave rise to n = 1239.2. However, most then just wrote down their answer (either 1239 or 1240) without checking the actual values of P($n \leq a$) (or P($n < a$)) for values of a close to 1239.

Other candidates used the "Binomial cdf" function, trying various values of a until they arrived at the largest value that gave P($n < a$) < 0.025.

A few used the "Inverse binomial" function, which gave rise to n = 1239, but most of these candidates did not proceed to check the actual values of P($n \leq a$) (or P($n < a$)) for values of a close to 1239.

Many candidates, whichever method they used, showed a misunderstanding of the "Binomial cdf" function on their calculator. Their working showed that they thought that the "Binomial cdf" function gives the value of P($n < a$) rather than P($n \leq a$). So for example, some wrote P(n < 1239) = 0.0251, which is incorrect. Others "fudged" the issue by writing, for example "1239 $\rightarrow$ 0.0251". Best practice would be to show explicitly P($n \leq$ 1239) = 0.0251 and P($n \leq$ 1238) = 0.0233, and hence state that the largest possible value of n is 1239.

Copyright

Supporting you

For further details of this qualification please visit the subject webpage.

Review of results

If any of your students' results are not as expected, you may wish to consider one of our review of results services. For full information about the options available visit the OCR website. If university places are at stake you may wish to consider priority service 2 reviews of marking which have an earlier deadline to ensure your reviews are processed in time for university applications.

activeresults

Review students' exam performance with our free online results analysis tool. Available for GCSE, A Level and Cambridge Nationals.

It allows you to:

- review and run analysis reports on exam performance
- analyse results at question and/or topic level*
- compare your centre with OCR national averages
- identify trends across the centre
- facilitate effective planning and delivery of courses
- identify areas of the curriculum where students excel or struggle
- help pinpoint strengths and weaknesses of students and teaching departments.

*To find out which reports are available for a specific subject, please visit ocr.org.uk/administration/support-and-tools/active-results/

Find out more at ocr.org.uk/activeresults

CPD Training

Attend one of our popular CPD courses to hear exam feedback directly from a senior assessor or drop in to an online Q&A session.

Please find details for all our courses on the relevant subject page on our website.

www.ocr.org.uk

The small print

OCR Resources: *the small print*

OCR's resources are provided to support the delivery of OCR qualifications, but in no way constitute an endorsed teaching method that is required by OCR. Whilst every effort is made to ensure the accuracy of the content, OCR cannot be held responsible for any errors or omissions within these resources. We update our resources on a regular basis, so please check the OCR website to ensure you have the most up to date version.

This resource may be freely copied and distributed, as long as the OCR logo and this small print remain intact and OCR is acknowledged as the originator of this work.

Our documents are updated over time. Whilst every effort is made to check all documents, there may be contradictions between published support and the specification, therefore please use the information on the latest specification at all times. Where changes are made to specifications these will be indicated within the document, there will be a new version number indicated, and a summary of the changes. If you do notice a discrepancy between the specification and a resource please contact us at: resources.feedback@ocr.org.uk.

Whether you already offer OCR qualifications, are new to OCR, or are considering switching from your current provider/awarding organisation, you can request more information by completing the Expression of Interest form which can be found here: www.ocr.org.uk/expression-of-interest

Please get in touch if you want to discuss the accessibility of resources we offer to support delivery of our qualifications: resources.feedback@ocr.org.uk

Looking for a resource?

There is now a quick and easy search tool to help find **free** resources for your qualification:

www.ocr.org.uk/i-want-to/find-resources/

www.ocr.org.uk

OCR Customer Support Centre

General qualifications

Telephone 01223 553998

Facsimile 01223 552627

Email general.qualifications@ocr.org.uk

OCR is part of Cambridge Assessment, a department of the University of Cambridge. *For staff training purposes and as part of our quality assurance programme your call may be recorded or monitored.*

Cambridge Assessment

CERTIFIED ISO 9001

UKAS MANAGEMENT SYSTEMS 001