Qualification Accredited

OCR
Oxford Cambridge and RSA

Examiners' report

A LEVEL

FURTHER MATHEMATICS B (MEI)

OCR Level 3 Advanced GCE in Further Mathematics B (MEI)

H645

For first teaching in 2017

Y422/01 Summer 2025 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 responses 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 Teach Cambridge.

Paper Y422 series overview

Statistics Major, Y422/01, is an optional examined component for GCE Further Mathematics B (MEI). The component provides two-thirds of the total mark for the applied part of the course and includes the content of Statistics Minor, Y432/01, together with various other topics. The component focuses on

- Discrete random variables, including the Poisson, geometric and discrete uniform distributions.
- Bivariate data, including Pearson's product moment correlation coefficient, Spearman's rank correlation coefficient and regression analysis.
- Chi-squared tests for contingency tables and for goodness of fit.
- Continuous random variables, including probability density and cumulative distribution functions, the Normal distribution and the continuous uniform distribution.
- Statistical inference including confidence intervals and hypothesis tests, using the Central Limit Theorem if necessary.
- Simulation.

Candidates are expected to know the content of A Level Mathematics and the Core Pure mandatory paper for Further Mathematics (Y420/01). Candidates should have gained experience during their course of spreadsheets or other software to explore data sets and to conduct hypothesis tests and construct confidence intervals. They should also have had experience of using a spreadsheet to simulate a random variable.

Most of the questions in this paper are in context, and many require interpretation in addition to understanding. Questions may also require candidates to comment about the modelling assumptions underlying their answers.

As in previous years, in general, candidates did well in questions involving calculations, but rather less well in questions requiring reasons or explanations.

Candidates who did well on this paper generally:	Candidates who did less well on this paper generally:
• carried out hypothesis tests: giving the hypotheses appropriately in terms of the population, using correct terminology, identifying critical values and giving non-assertive conclusions • chose appropriate levels of accuracy for their answers • used calculators effectively to calculate probabilities, find a correlation coefficient and find expectation and variance • applied their knowledge and understanding to new and unfamiliar contexts.	• correctly found a correlation coefficient, showing full working • correctly found a contribution for a chi-squared test • found it difficult to give clear answers to questions requiring explanations or comments • made errors in carrying out hypothesis tests such as not mentioning the population, using the wrong critical value or giving the wrong conclusion or a conclusion that was too assertive.

Section A overview

This section consists of more straightforward questions. Most questions in this section were very well-answered, apart from Question 1 (c) and Question 4 (b). Question 1 (c) required candidates to find a combined probability based on a discrete uniform distribution. Question 4 (b) involved the geometric distribution where many candidates, having correctly found the mean and variance, did not use a correct method to find the final answer.

Question 1 (a)

1 A fair eight sided dice has its sides labelled 1, 2, …, 8. The random variable X represents the score when the dice is rolled once.

(a) Find $\mathrm{P}(X > 4)$. **[1]**

Almost all candidates correctly answered this question.

Question 1 (b)

(b) Find each of the following.

- $\mathrm{E}(X)$
- $\mathrm{Var}(X)$ **[3]**

Almost all candidates correctly answered this question.

Question 1 (c)

The dice is rolled 3 times.

(c) Determine the exact probability that one of the scores is less than 4 and the other two are greater than 4. **[3]**

Most candidates realised that they had to multiply three probabilities together, but some then forgot to multiply their answer by 3 or multiplied instead by 3!. Some candidates had an error in one or other of the probabilities.

Question 2 (a)

2 At a toy factory, wooden blocks of approximate heights 20 mm, 30 mm and 50 mm are made in red, yellow and green respectively. The heights of the blocks in mm are modelled by independent random variables which are Normally distributed with means and standard deviations as shown in the table.

Colour	Mean	Standard deviation
Red	20	0.8
Yellow	30	0.9
Green	50	1.2

In parts **(a)**, **(b)** and **(c)**, the blocks are selected randomly and independently of one another.

(a) Find the probability that the height of a red block is less than 19 mm. **[1]**

Almost all candidates correctly answered this question.

Question 2 (b)

(b) A tower is made of 15 blocks stacked on top of each other consisting of 5 red blocks, 5 yellow blocks and 5 green blocks.

Determine the probability that the tower is at least 495 mm high. **[4]**

Most candidates answered this correctly, although some multiplied the three variances by 5^2 rather than by 5.

Question 2 (c)

(c) Determine the probability that a tower made of 3 red blocks will be at least 1 mm higher than a tower made of 2 yellow blocks. **[3]**

Again, most candidates answered this correctly, although some again multiplied by 3^2 and 2^2 rather than by 3 and 2 or made other errors in finding the variance.

Question 3 (a)

3 Potatoes are sold in sacks of nominal weight 25 kg. As part of a trading standards check, a random sample of 50 sacks from a large batch is selected and software is used to produce a 95% confidence interval for the mean weight of potatoes per sack. An extract from the software is shown below.

Sample Mean	24.878
Standard Deviation	0.5664
Standard Error	0.0801
Sample Size	50
Confidence Level	0.95
Interval	24.878 ± 0.157

(a) Show how the standard error was calculated. **[1]**

Almost all candidates correctly answered this question.

Question 3 (b)

(b) State the confidence interval in the form $a < \mu < b$. **[1]**

Almost all candidates correctly answered this question.

Question 3 (c)

(c) Explain whether the confidence interval suggests that the mean weight of potatoes per sack is different from 25 kg. **[1]**

Although there were many correct responses, many candidates gave answers which either suggested that due to 25 being at the upper end of the interval, it was unlikely that the mean weight was 25kg, or that the mean was definitely 25kg.

Question 3 (d)

(d) Determine whether your conclusion to part **(c)** would have been different if the confidence level had been 90%. **[4]**

Most candidates found the new confidence interval correctly, gaining 3 marks. The final mark was for stating that the conclusion would be the same. Some did not get this final mark due to similar reasons as in part (c).

Question 4 (a) (i)

4 Gabi is practising basketball. On each attempt to score a basket, the probability that Gabi scores is modelled as 0.4. It is assumed that each attempt is independent of previous attempts. The number of attempts which Gabi takes in order to score a basket for the first time is denoted by the random variable X.

(a) (i) Calculate $P(X = 5)$. **[1]**

Almost all candidates answered this question correctly.

Question 4 (a) (ii)

(ii) Calculate $P(X > 5)$. **[1]**

Most candidates answered this correctly, although many gave the complement of the answer, i.e. $P(X \leqslant 5)$.

Question 4 (b)

(b) Determine the probability that X is within one standard deviation of its mean. **[5]**

Only a few candidates did not know how to find the mean and standard deviation. Once these were correctly found, many candidates forgot that they were dealing with a geometric distribution and tried to use a Normal distribution. Others did use a geometric distribution but had the wrong range of values or the right range but worked it out incorrectly.

Question 4 (c)

(c) Determine the probability that Gabi scores at least 5 baskets in 20 attempts. **[2]**

Most candidates realised that they had to use a binomial distribution and answered this correctly.

Question 4 (d)

(d) Determine the probability that Gabi scores the 5th basket on their 20th attempt. **[2]**

Many correct responses were seen. Some candidates realised that they had to multiply a B (19, 0.4) probability by 0.4 but did not find the binomial probability correctly.

Question 4 (e)

(e) Comment on the validity of the assumption that on each attempt to score a basket, the probability that Gabi scores is 0.4, independent of previous attempts. **[1]**

This question was very well-answered, with some candidates suggesting that, with more practice. the probability would increase, while others suggested that Gabi would get tired so it would reduce.

Section B overview

This section consists of more of a mixture of question styles with less straightforward questions. The only question that proved to be very challenging was Question 12 involving a proof relating to the mean and variance of a binomial distribution. Question 7 (c) and Question 7 (d) involving comments on a table of correlation coefficients were also fairly challenging, as was Question 11 (f). The remainder of the questions were generally well-answered, although some candidates found considerable difficulties with Question 8 (c) involving a Normal distribution hypothesis test and Question 10 (c) involving the Central Limit Theorem.

Question 5 (a)

5 The random variable X has a continuous uniform distribution over $[-4, 2]$.

(a) Find the variance of X. **[2]**

Almost all candidates answered this correctly.

Question 5 (b)

(b) Determine the probability that if two independent values of X are taken, one is less than zero and the other is greater than zero. **[2]**

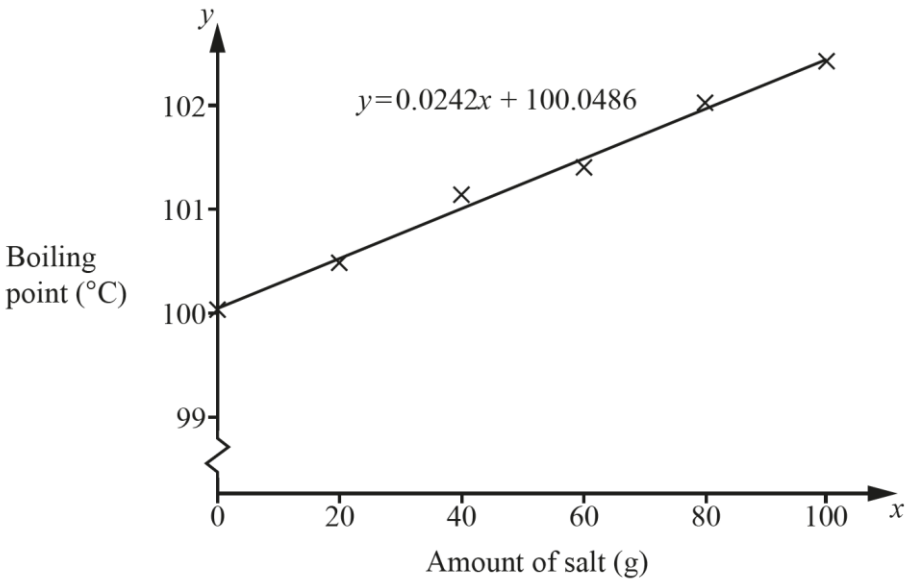
Although candidates had used the correct formula for a continuous uniform distribution in part (a), many tried to use a discrete uniform distribution here, working out $\frac{4}{7} \times \frac{2}{7}$ and sometimes then multiplying by 2. Some used the correct distribution but forgot to multiply by 2.

Question 6 (a)

6 A student is investigating the link between the amount of salt dissolved in water and the boiling point of the water. The student dissolves fixed amounts of salt, x g, into water and then measures the boiling point, y °C, of the water.

The data are shown in the table together with a scatter diagram to illustrate them. The equation of the regression line of y on x is also shown.

Amount of salt, x g	0	20	40	60	80	100
Boiling point, y °C	100.03	100.49	101.15	101.41	102.04	102.44



(a) State what type of variable 'Boiling point' is. **[1]**

Most candidates answered this correctly, although a few said 'independent' rather than dependent.

Question 6 (b)

(b) Explain why 'Amount of salt' has been plotted on the horizontal axis. **[1]**

Again, most candidates answered this correctly, although most candidates who made the error in part (a) made a similar error in part (b).

Question 6 (c)

(c) Calculate an estimate of the boiling point with the following amounts of salt.

- 50 g
- 200 g

[2]

Although most candidates knew what to do, many gave their answers to too many decimal places. Since the answers were estimates, they should have been given to 1 or at most 2 decimal places. Answers of 101.2586 and 104.8886 were very common.

Question 6 (d)

(d) Comment on the reliability of each of the estimates. **[2]**

Almost all candidates scored at least 1 mark for stating that the second estimate was not reliable due to extrapolation. Fewer gained the other mark since they only mentioned 'interpolation' and not the goodness of fit. Some thought that as the value for 50g was close to other points then that was a further reason for it being reliable but did not mention the general closeness of all the points to the line of best fit.

Question 6 (e)

(e) Determine the residual for the data point with 60 g of salt. **[2]**

Again, almost all candidates got at least 1 mark, but many subtracted the value of 101.41 from the estimated value, rather than the other way around and so did not get the second mark.

Question 7 (a)

7 A student is investigating the relationship between different electricity generation methods and cost of electricity in a particular country. The student first checks whether there is any correlation between the cost per unit of electricity, x euros, and the amount of electricity being generated by wind, y GW. The data from 30 observations are summarised as follows.

$n = 30 \quad \sum x = 2.219 \quad \sum y = 357.7 \quad \sum x^2 = 0.2368 \quad \sum y^2 = 4648 \quad \sum xy = 25.01$

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

Determine the product moment correlation coefficient. **[4]**

Most candidates scored all 4 marks, but some did not follow the 'detailed reasoning' instruction. These candidates usually showed the working for each of S_{xx}, S_{yy} and S_{xy} but then just quoted the formula for the correlation coefficient and wrote down -0.2744, without showing the calculation. As such, they could only be given the first 2 marks.

Question 7 (b)

(b) Carry out a hypothesis test at the 5% level to investigate whether there is any correlation between the cost per unit of electricity and the amount of electricity generated by wind. **[5]**

Although most candidates knew how to carry out the test, there were many errors in the hypotheses and/or the final answer in context. Many candidates missed out 'population' in their definition of rho, or, if they gave the hypothesis in words, they again missed out 'population'. In the final answer, some candidates incorrectly stated that there was sufficient evidence to suggest that there was correlation. Others said that there was 'no evidence' rather than 'insufficient evidence'. This is not entirely correct but was allowed.

Assessment for learning

In questions involving hypothesis testing, the hypotheses should always mention the population (other than in goodness of fit tests). If a parametric test is being carried out, the parameter (in this case ρ) should be defined as the population parameter. Thus, the definition of ρ should be 'ρ is the population product moment correlation coefficient between the cost of electricity and the amount of electricity being generated by wind'. In the case of non-parametric tests, the hypotheses are usually given in words, but they must again include the word 'population'.

The conclusions in such tests should always be given in context. In this case 'There is *insufficient* evidence at the 5 percent level to suggest that there *is correlation* between the cost of electricity and the amount of electricity being generated by wind'. A conclusion stated in reverse, such as 'there is *sufficient* evidence to suggest that there is *no correlation* between the cost of electricity and the amount of electricity being generated by wind' is incorrect.

The conclusion should have an element of doubt. A statement such as 'There is insufficient evidence to suggest that there is correlation, and so the cost of electricity and the amount of electricity being generated by wind are *not correlated*' is too assertive. This would not be given the mark for the conclusion due to the final overly assertive part of the sentence.

Question 7 (c)

Table 7.1 shows the values of the product moment correlation coefficients between 5 variables. These variables are the cost per unit of electricity in euros and the amount of electricity generated in GW by four different types of generation, for a very large sample of observations.

For example, the correlation coefficient between the amount of electricity generated by wind and the amount of electricity generated by gas is −0.59. The 'storage' category consists of battery and other types of electricity taken from stored reserves. All of the correlations are significant at the 10% level.

Table 7.1

	Cost	Gas	Wind	Solar	Storage
Cost	1.00				
Gas	0.42	1.00			
Wind	−0.26	−0.59	1.00		
Solar	−0.18	−0.21	−0.07	1.00	
Storage	0.12	0.35	−0.38	−0.30	1.00

Table 7.2 shows standard guidelines for effect sizes.

Table 7.2

Product moment correlation coefficient	Effect size
$0.1 \leqslant r < 0.3$	Small
$0.3 \leqslant r < 0.5$	Medium
$r \geqslant 0.5$	Large

The student analyses these data for effect size.

(c) Explain how the very large sample size relates to the interpretation of the correlation coefficient of −0.07 between the amount of electricity generated by wind and the amount of electricity generated by solar, as shown in **Table 7.1**. **[2]**

Most candidates gave responses such as 'for large samples the critical value is very low' or 'a large sample size means that the value given by the data is very close to the actual correlation coefficient in the population' which got the first mark. Rather fewer got the second mark for stating that the correlation was very close to zero and of little practical use. Although the question stated that all the correlations were significant at the 10 percent level, some candidates stated that the correlation 'might' be significant.

Question 7 (d)

(d) Comment briefly on what the student might infer from these tables regarding different electricity generation methods and cost of electricity. **[2]**

This question was not very well-answered. Many candidates did not read the question properly and discussed the correlation between electricity generated by wind and that generated by gas. The question specifically stated different generating methods and the cost of electricity. Of those candidates that did answer the question that was asked, only a few got both marks. Many mentioned positive correlations between electricity generated by gas and the cost of electricity, but rather fewer mentioned effect sizes. Since the question says tables rather than table there should be some discussion of effect sizes to score the second mark.

Question 8 (a)

8 A teacher is investigating whether students run more quickly round an athletics track if they run in a clockwise direction rather than in an anti-clockwise direction. She thinks that students will run more quickly on average in a clockwise direction.

The teacher selects 40 students at random and records how long it takes for each of them to run one lap of the track in a clockwise direction. The next day she records how long it takes for each of these students to run one lap in an anti-clockwise direction.

(a) Explain why the teacher does **not** get the students to run the second lap immediately after they have completed their first lap. **[1]**

This was very well-answered, with almost all candidates mentioning that the students would be fatigued after the first lap.

Question 8 (b)

(b) Explain why the teacher notes the names of each student, so that she knows which student took which time on each day, rather than just noting the 40 times on each day. **[1]**

Only a few candidates mentioned 'pairing', which was the ideal answer, but the mark was also given for discussing the individual differences in the lap times.

Question 8 (c)

The differences in times, d seconds (anti-clockwise time – clockwise time), are summarised as follows:

$\sum d = 72.2, \quad \sum d^2 = 1510$

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

Carry out a hypothesis test at the 5% significance level to investigate whether there is any evidence to suggest that on average students run more quickly round the track in a clockwise direction. **[10]**

Some candidates thought that a test based on Spearman's rank correlation coefficient was required. Such candidates occasionally scored 1 mark out of 10 if they happened to find the population mean, but most got 0 marks. Candidates who realised that a test based on the Normal distribution or the *t* distribution was required usually got at least 7 marks and some were given full marks. The main error which many candidates made was to use the wrong critical value, usually 1.96 rather than 1.645. Many more candidates than in previous series used an approach where they worked out the *p* value and compared this to 0.05, usually successfully. A few missed out 'population' in the definition of μ or had the wrong alternative hypothesis.

Question 8 (d)

(d) Comment briefly on any possible problem with the teacher's method. **[1]**

The expected answer was that weather conditions on the two days might be different, but most candidates gave some other unsuitable reason and so did not get the mark.

Question 9 (a)

9 A consultant is modelling the number of trades per minute of a particular share in a stock exchange. The number of trades per minute in a random sample of 200 minutes is shown in the table.

Number of trades per minute	0	1	2	3	4	5	6	7	$\geqslant 8$
Frequency	23	42	46	45	20	15	7	2	0

The consultant thinks that the number of trades per minute may be modelled by a Poisson distribution.

(a) Use the data in the table to estimate the value of the parameter λ for an appropriate Poisson model. **[2]**

Almost all candidates answered this question correctly.

Question 9 (b)

The consultant decides to carry out a chi-squared goodness of fit test to investigate further. The screenshot below shows part of a spreadsheet to assess the goodness of fit of the distribution Po(λ), using the value of λ estimated from the data.

	A	B	C	D	E
1	Number of trades per minute	Frequency	Poisson probability	Expected frequency	Chi-squared contribution
2	0	23	0.0907	18.1436	1.2999
3	1	42			
4	2	46			0.7484
5	3	45	0.2090	41.8028	0.2445
6	4	20	0.1254	25.0817	1.0296
7	5	15	0.0602	12.0392	0.7281
8	$\geqslant 6$	9	0.0357	7.1345	0.4878
9					

(b) Calculate the missing values in each of the following cells.

- C3
- D3
- E3

[4]

Again, almost all candidates answered this question correctly.

Question 9 (c)

(c) Explain why the frequencies for 6, 7 and $\geqslant 8$ trades have been combined into the single category of $\geqslant 6$ trades, as shown in the spreadsheet. **[1]**

Once again, almost all candidates answered this question correctly.

Question 9 (d)

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

Carry out the test at the 5% significance level. **[6]**

Candidates usually got either 5 or 6 marks or at most 2 marks for this question. The main error in the responses that scored 5 marks was to have the wrong conclusion despite correctly accepting H_0 and having correct hypotheses. Of the responses that got 2 marks, these were usually for having correct hypotheses and for finding the correct test statistic (by follow through if they had the wrong value in part (b)).

Question 10 (a)

10 Two independent random variables, X and Y, have distributions B(6, 0.4) and B(3, 0.4) respectively. The random variable $T = X - (Y_1 + Y_2)$ where Y_1 and Y_2 are two independent values of Y.

(a) Determine $P(T = 5)$. **[4]**

This question was not answered well, with around half of candidates getting 0 marks. The main error was to think that a Normal distribution was required, when in fact the binomial probabilities were needed, and this gained no credit unless by chance candidates had first worked out a binomial probability. Most candidates who did use binomial probabilities got full marks.

Question 10 (b)

A spreadsheet has been constructed to simulate the situation, as shown below. There are 100 simulated values of each of X, Y_1, Y_2 and T.

Cells G2 and G3 in the spreadsheet show the number of values of T which are greater than 0, and less than or equal to 0, respectively. Cell G5 shows the mean value of the 100 simulated values of T.

	A	B	C	D	E	F	G
1	X	Y_1	Y_2	$T=X-(Y_1+Y_2)$			
2	3	2	1	0		Number > 0	42
3	2	0	0	2		Number ⩽ 0	58
4	3	1	0	2			
5	1	3	2	−4		Mean value of T	0.25
6	4	2	1	1			
7							
8							
99	4	1	1	2			
100	2	2	2	−2			
101	3	0	3	0			
102							

(b) Use the information in the spreadsheet to write down an estimate of $P(T > 0)$. **[1]**

Almost all candidates answered this question correctly.

Question 10 (c)

As shown in the spreadsheet, the mean value of the 100 simulated values of T is 0.25.

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

Use the Central Limit Theorem to determine an estimate of the probability that the mean of 100 independent values of T is greater than 0.25. **[8]**

In this question, the most common scores were 8, 7 or 0 marks. Most of those who got 7 marks did not use a continuity correction but otherwise did everything correctly. Some got 6 marks, usually using the wrong continuity correction of 0.245 rather than 0.255. A few candidates did not divide the variance of T by 100 and so could only achieve the first 4 marks.

Question 11 (a)

11 The function $F(x)$ is defined as follows.

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{ax+b}{9x+15} - \frac{4}{3} & 0 \leqslant x \leqslant 5 \\ 1 & x > 5 \end{cases}$$

where a and b are constants.

You are given that $F(1) = 0.5$ and $F(2.5) = 0.8$.

(a) Show that $a = 24$ and $b = 20$. **[3]**

Although this question said 'show', some candidates found the two original equations, and then either just wrote $a = 24$, $b = 20$ or stated 'by simultaneous equations $a = 24$, $b = 20$'. Candidates must show their workings to gain all the marks, and such candidates could only gain the first mark.

Question 11 (b) (i)

(b) (i) Show that $F(0) = 0$. **[1]**

Almost all candidates answered this question correctly.

Question 11 (b) (ii)

(ii) Show that $F(x)$ is a non-decreasing function. **[3]**

Many candidates realised that they had to differentiate $F(x)$ and usually did this correctly. Some candidates did not produce a satisfactory conclusion but were still able to gain 2 marks out of 3. Many tried to use reasoning to show that $F(x)$ is non-decreasing, but this was never successful. Some simply substituted a few values of x into $F(x)$ to show that it is non-decreasing, but this approach did not gain any marks.

Question 11 (c) (i)

(c) (i) State another property that is required for $F(x)$ to be a cumulative distribution function. **[1]**

Roughly half of the candidates gained this mark with the other half giving a variety of reasons which were often already implied by the previous statements.

Question 11 (c) (ii)

(ii) Show that the property in part **(c)(i)** is satisfied. **[1]**

This question was answered slightly better than the previous part. Some candidates who said that the area under the graph was equal to 1 in part (b), without stating that it was the graph of the differential of F(x), did show that F(5) = 1.

Question 11 (d)

You are given that F(x) is the cumulative distribution function of the continuous random variable X.

(d) Determine the upper quartile of X. Give your answer as an exact fraction. **[3]**

This question was well-answered, with most candidates getting full credit. Many of the others got 0 marks due to a wide variety of errors.

Question 11 (e)

(e) Find E(X). **[2]**

Those candidates who had found the probability density function (pdf) in part (b) (ii) usually correctly found E(X). Those that did not, occasionally found the pdf in this part, but nearly half of candidates got no marks.

Question 11 (f)

(f) Determine the mode of X by considering the gradient of the probability density function of X. **[2]**

Most candidates found this question challenging, with many unable to differentiate the pdf correctly. Some did differentiate successfully but then thought that the mode was five rather than zero. Only a few candidates gave a fully correct response.

Question 12

12 The random variable X takes the value 1 with probability p and the value 0 with probability $1 - p$. You are given that $\mathrm{E}(X) = p$ and $\mathrm{Var}(X) = p - p^2$.

The random variable $Y \sim B(64, 0.1)$ has mean μ and standard deviation σ.

Use the results given above to **prove** the following in either order.

- $\mu = 6.4$
- $\sigma = 2.4$ **[6]**

This question was the most difficult on the paper, with only very few candidates getting full credit and most getting no marks at all. Of those candidates that had partial marks, many correctly used the given value of $p = 0.1$ to find $\mathrm{E}(X)$ and $\mathrm{Var}(X)$, although these were often not stated explicitly but simply used later on. Some of these then clearly stated that $\mathrm{E}(Y) = 64 \times \mathrm{E}(X)$ and $\mathrm{Var}(Y) = 64 \times \mathrm{Var}(X)$ and so gained SC marks. Only a few got the method marks and only a few stated that $Y = X_1 + X_2 + \cdots + X_{64}$, as per the mark scheme. Those candidates who simply used the formulae $\mathrm{E}(Y) = np$ and $\mathrm{Var}(Y) = npq$ did not get any marks.

Exemplar 1

12

$Y = X_1 + X_2 + \ldots + X_{64}$

$E(Y) = E(X_1 + X_2 \ldots + X_{64})$

$= E(X_1) + E(X_2) + \ldots + E(X_{64})$

$= 64\,E(X)$

$= 64p$

$Var(Y) = Var(X_1 + X_2 + X_3 \ldots + X_{64})$

$= Var(X_1) + Var(X_2) \ldots + Var(X_{64})$

$= 64\,Var(X)$

$= 64(p - p^2)$

Since $p = 0.1$,

$E(Y) = 64p = 6.4$

$Var(Y) = 64(0.1 - 0.1^2) = 64(0.09)$

$= 5.76$

$\therefore \sigma_Y = 2.4$

This exemplar shows an almost perfect response to the question. The only thing that is missing is an explicit statement that E(X) = 0.1 and Var(X) = 0.09. All of the other elements in the mark scheme are present in this response. Candidates could adopt this layout for future questions on this topic.

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

We've made it easier for Exams Officers to download copies of your candidates' completed papers or 'scripts'. 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 professional development 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 exam paper builder platform for a range of our qualifications. Find out more.

Free for all OCR centres with an Interchange account, it allows unlimited users per centre. We require an Interchange username to verify your centre's users for ExamBuilder.

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

Once you have an Interchange username, use the form on this page to sign up for ExamBuilder, or ask an existing user at your centre to create an ExamBuilder account for you.

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 (examined units only).

Find out more.

Tell us what you think

Your feedback plays an important role in how we develop, market, support and resource qualifications now and into the future. We want you and your students to enjoy and get the best out of our qualifications and resources, but to do that we need your honest opinions to tell us whether we're on the right track or not.

You can email your thoughts to resources.feedback@ocr.org.uk or visit our feedback page to learn more about how you can help us improve our resources.

Designing and testing in collaboration with you and your students

Helping young people develop an ethical view of the world

Equality, diversity, inclusion and belonging (EDIB) are part of everything we do

---

Contact the team at:

01223 553998

ocr.org.uk

To stay up to date with all the relevant news about our qualifications, register for email updates at ocr.org.uk/updates

Visit our Online Support Centre at support.ocr.org.uk

CAMBRIDGE
UNIVERSITY PRESS & ASSESSMENT