

A LEVEL

Examiners' report

FURTHER MATHEMATICS B (MEI)

H645

For first teaching in 2017

Y422/01 Summer 2019 series

Version 1

Contents

Introduction3

Paper Y422 series overview4

Section A overview6

 Question 1 (c)6

 Question 2 (a)6

 Question 2 (d)6

 Question 3 (a)7

 Question 3 (b)7

 Question 3 (c)7

Section B overview8

 Question 4 (c)8

 Question 4 (e)9

 Question 5 (b)9

 Question 5 (c)10

 Question 6 (a) (i)11

 Question 6 (a) (ii)12

 Question 6 (a) (iii)12

 Question 6 (b) (i)13

 Question 6 (b) (ii)13

 Question 6 (b) (iii)14

 Question 7 (b)14

 Question 7 (c)14

 Question 7 (d)15

 Question 8 (a)16

 Question 8 (b)17

 Question 8 (c)17

 Question 9 (a)18

 Question 9 (b)18

 Question 9 (e)19

 Question 9 (f)20

 Question 9 (g)20

 Question 9 (h)20

 Question 10 (b)21

 Question 10 (c) (i)21

 Question 10 (c) (ii)21

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.



Would you prefer a Word version?

Did you know that you can save this pdf as a Word file using Acrobat Professional?

Simply click on **File > Save As Other ...** and select **Microsoft Word**

(If you have opened this PDF in your browser you will need to save it first. Simply right click anywhere on the page and select **Save as ...** to save the PDF. Then open the PDF in Acrobat Professional.)

If you do not have access to Acrobat Professional there are a number of **free** applications available that will also convert PDF to Word (search for *pdf to word converter*).



We value your feedback

We'd like to know your view on the resources we produce. By clicking on the icon above you will help us to ensure that our resources work for you.

Paper Y422 series overview

Y422/01 (Statistics Major) is an optional examined component for the new revised A Level examination 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 Y432 (Statistics Minor) 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 content assessed in Y420. 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.

To do well on this paper, candidates need to be comfortable applying their knowledge and understanding to new and unfamiliar contexts. Most of the questions in Y422/01 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.

Candidate performance overview

This paper was very accessible to the majority of candidates with many very good responses seen. There was no evidence that there was insufficient time for candidates to complete the paper. In general, most candidates gave their answers to an appropriate degree of accuracy, but a few over-specified their answers, giving many to 6 or more significant figures. On this occasion this was not penalised, other than in Question 6a(ii) where an estimate was requested.

Candidates who did well on this paper generally did the following:

- Understood how to use linear combinations of independent Normal random variables in Question 3.
- Could carry out and interpret the results of hypothesis tests in Questions 5, 6b (iii) and 8c.
- Understood how to construct confidence intervals and to use those constructed by software.
- Understood how to use probability density functions and cumulative distribution functions.

Candidates who did less well on this paper generally did the following:

- Demonstrated a lack of understanding of which regression line to use for predictions in Questions 6a (i) and 6a (ii).
- Did not understand how to use Normal probability plots in Question 8 to decide which hypothesis test to use.
- Did not realise how to obtain the distribution of T in Question 9.

Section A overview

Section A questions were generally very well done, apart from Question 2(d) and finding the distribution in Question 3.

Question 1 (c)

- 1 A fair six-sided dice is rolled three times.
The random variable X represents the lowest of the three scores.
The probability distribution of X is given by the formula

$$P(X = r) = k(127 - 39r + 3r^2) \text{ for } r = 1, 2, 3, 4, 5, 6.$$

- (c) Draw a graph to illustrate the distribution.

[2]

This question was generally well answered for the first mark for the heights of the lines. Many candidates however omitted either the scales or the labels or occasionally both and were only given the first mark.

Question 2 (a)

- 2 A special railway coach detects faults in the railway track before they become dangerous.
- (a) Write down the conditions required for the numbers of faults in the track to be modelled by a Poisson distribution.

[2]

There were some very clear answers to this part. Some candidates forgot to put their answers in context and so, in this instance, could only get a maximum of one mark. It is advisable to state the answer in the context of the question given to ensure full credit. Others made incorrect statements such as 'the number of faults must occur at a uniform average rate', rather than 'faults must occur at a uniform average rate'.

Question 2 (d)

- (d) On a particular day the coach is used to check 10 randomly chosen 1 km lengths of track.
Find the probability that exactly 1 fault, in total, is found.

[3]

There are 2 ways of correctly answering this question. The simpler method is to realise that the distribution is simply Poisson(3.2) and to calculate $P(X = 1)$ using this distribution. The harder method is to find $P(X = 0) = 0.7261$ and $P(X = 1) = 0.2324$ using Poisson(0.32), and then use these as in Exemplar 1. Many candidates mistakenly thought that the distribution was simply B(10, 0.2324) which did not get any credit.

Exemplar 1

$Y = 0.2X$
$Y \sim Po(0.32)$
$P(Y=1) = 0.23236769$
$P(Y=0) = 0.72614903$
$P(\text{there is exactly 1 fault}) = 10 \times 0.23236769 \times 0.72614903^9$
$= 0.1304390402$
$= 0.1304 \text{ to } 4\text{s.f.}$

Question 3 (a)

- 3 The weights of bananas sold by a supermarket are modelled by a Normal distribution with mean 205 g and standard deviation 11 g.

(a) Find the probability that the total weight of 5 randomly selected bananas is at least 1 kg. [2]

This question was generally well answered although some candidates multiplied the original variance by 5^2 rather than by 5.

Question 3 (b)

When a banana is peeled the change in its weight is modelled as being a reduction of 35%.

(b) Find the probability that the weight of a randomly selected peeled banana is at most 150 g. [3]

There were many correct responses to this part. The first mark was gained by almost all candidates, but then the variance was fairly often incorrectly multiplied by 0.65 rather than by 0.65^2 .

Question 3 (c)

Andy makes smoothies. Each smoothie is made using 2 peeled bananas and 20 strawberries from the supermarket, all the items being randomly chosen. The weight of a strawberry is modelled by a Normal distribution with mean 22.5 g and standard deviation 2.7 g.

(c) Find the probability that the total weight of a smoothie is less than 700 g. [4]

Again, this part was often answered correctly, usually by those candidates who had got part (b) correct. Those who did not could still get 2 method marks and most got the first, as they found the new mean correctly. However very few showed the method which they had used to find the new variance, and so those who had the wrong variance in part (b) could not usually be given the second mark.

Section B overview

There were many very good responses to most of the questions in Section B. However, lower ability candidates did struggle with Questions 8, 9 and 10.

Question 4 (c)

- 4 Shellfish in the sea near nuclear power stations are regularly monitored for levels of radioactivity. On a particular occasion, the levels of caesium-137 (a radioactive isotope) in a random sample of 8 cockles, measured in becquerels per kilogram, were as follows.

2.36 2.97 2.69 3.00 2.51 2.45 2.21 2.63

Software is used to produce a 95% confidence interval for the level of caesium-137 in the cockles. The output from the software is shown in Fig. 4. The value for 'SE' has been deliberately omitted.

T Estimate of a Mean ▼

Confidence Level

Sample

Mean

s

N

Result

T Estimate of a Mean

Mean	2.6025
s	0.2793
SE	
N	8
df	7
Interval	2.6025 ± 0.2335

Fig. 4

- (c) In the software output shown in Fig. 4, SE stands for standard error. Find the standard error in this case.

[2]

It seems that many candidates do not know the meaning of standard error, as a number of wrong answers were seen. However, part (d) was usually answered correctly, which suggests that candidates know how to calculate confidence intervals, but do not know all of the terms associated with them.

Question 4 (e)

- (e) State how, using this sample, a wider confidence interval could be produced. [1]

Most candidates gave a correct response to this part, but some discussed significance level instead of confidence level. Others stated that a higher confidence level should be used but then suggested 90% as being suitable or some other value lower than 95%.

Question 5 (b)

- 5 In an investigation into the possible relationship between smoking and weight in adults in a particular country, a researcher selected a random sample of 500 adults. The adults in the sample were classified according to smoking status (non-smoker, light smoker or heavy smoker, where light smoker indicates less than 10 cigarettes per day) and body weight (underweight, normal weight or overweight).

Fig. 5 is a screenshot showing part of the spreadsheet used to calculate the contributions for a chi-squared test. Some values in the spreadsheet have been deliberately omitted.

	A	B	C	D	E	F
1	Observed frequencies					
2		Underweight	Normal	Overweight	Totals	
3	Non-smoker	8	52	178	238	
4	Light smoker	10	40	68	118	
5	Heavy smoker	5	47	92	144	
6	Totals	23	139	338	500	
7						
8	Expected frequencies					
9	Non-smoker	10.9480	66.1640	160.8880		
10	Light smoker	5.4280		79.7680		
11	Heavy smoker		40.0320	97.3440		
12						
13	Contributions to the test statistic					
14	Non-smoker	0.7938		1.8200		
15	Light smoker	3.8510	1.5785	1.7361		
16	Heavy smoker	0.3982	1.2129	0.2934		
17						

Fig. 5

- (b) Complete the hypothesis test at the 1% level of significance. [6]

There were many fully correct responses to this part. The majority of candidates stated their hypotheses in terms of association, but some used independence instead, which was perfectly valid. A small number used 'relationship' in place of 'association' and this was not given any credit. Some lower ability candidates made a correct comparison ' $14.72 > 13.28$ ' but then came to the wrong conclusion stating that 'hence we can accept H_0 '.

Question 5 (c)

- (c) For each smoking status, give a brief interpretation of the largest of the three contributions to the test statistic. [3]

This was less successfully answered, with many candidates simply stating which was the largest contribution for each smoking status. Others, as in Exemplar 2, again correctly identified which was the largest contribution for each status and then tried to explain why the contributions were as they were, without stating whether there were less or more observed than expected. In questions such as this, candidates should be advised that it is essential to explicitly state whether there are less or more observed than expected.

Exemplar 2

Non-smoker's ~~have~~ have normal weight because they are healthier.

Light smoker's are underweight because ^{caused} smoking takes away their appetite.

Heavy smoker's are normal weight because they are used to the smoking, so it may not affect their diet.

Question 6 (a) (i)

- 6 (a) A researcher is investigating the date of the 'start of spring' at different locations around the country.

A suitable date (measured in days from the start of the year) can be identified by checking, for example, when buds first appear for certain species of trees and plants, but this is time-consuming and expensive. Satellite data, measuring microwave emissions, can alternatively be used to estimate the date that land-based measurements would give.

The researcher chooses a random sample of 12 locations, and obtains land-based measurements for the start of spring date at each location, together with relevant satellite measurements. The scatter diagram in Fig. 6.1 shows the results; the land-based measurements are denoted by x days and the corresponding values derived from satellite measurements by y days.

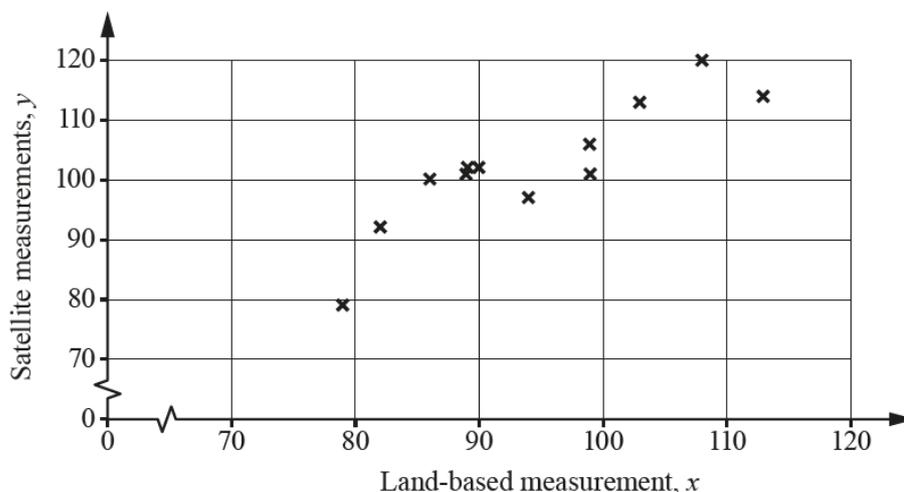


Fig. 6.1

Fig. 6.2 shows part of a spreadsheet used to analyse the data. Some rows of the spreadsheet have been deliberately omitted.

	A	B	C	D	E	F
1		x	y	x^2	y^2	xy
2		90	102	8100	10404	9180
3						
10						
11						
12		94	97	8836	9409	9118
13		99	101	9801	10201	9999
14	Sum	1131	1227	107783	126725	116724
15						

Fig. 6.2

- (i) Calculate the equation of a regression line suitable for estimating the land-based date of the start of spring from satellite measurements. [5]

This was usually well answered, but some candidates found the equation of the regression line of y on x rather than x on y . This is of course unsuitable for estimating the land-based date from satellite measurements. These candidates were however usually given 2 method marks, one for finding the gradient and one for finding the constant term.

Question 6 (a) (ii)

- (ii) Using this equation, estimate the land-based date of the start of spring for the following dates from satellite measurements.
- 95 days
 - 60 days

[2]

This part required an estimate and so an over-accurate answer was penalised. Candidates who gave their answers to more than one decimal place were only given 1 mark out of the 2 available. Follow through was not available for those who had the wrong regression line (the regression line of y on x) in part a(i).

Question 6 (a) (iii)

- (iii) Comment on the reliability of each of your estimates.

[2]

Almost all candidates were given the second mark but very few gained the first since they did not discuss the goodness of fit of the regression line. They instead simply stated that because the prediction was within the existing data, it was reliable. Exemplar 3 shows a response which was given both marks. The candidate states that interpolation has been used and then correctly comments that the data do not appear to follow a straight line. Candidates should be advised that in commenting on the reliability of an estimate from a regression line, they should always discuss the fit of the line in addition to whether interpolation or extrapolation has been used.

Exemplar 3

The 95 day estimate is ~~a~~ fairly reliable as it is interpolation and looks like it fits the trend reasonably well although the data does not look like it follows a straight line

The 60 day estimate is less reliable ~~as~~ as it is extrapolating and the trend may not continue outside of the graph.

Question 6 (b) (i)

- (b) The researcher is also investigating whether there is any correlation between the average temperature during a month in spring and the total rainfall during that month at a particular location. The average temperatures in degrees Celsius and total rainfall in mm for a random selection, over several years, of 10 spring months at this location are as follows.

Temperature	4.2	7.1	5.6	3.5	8.6	6.5	2.7	5.9	6.7	4.1
Rainfall	18	26	42	76	15	43	84	53	66	36

The researcher plots the scatter diagram shown in Fig. 6.3 to check which type of test to carry out.

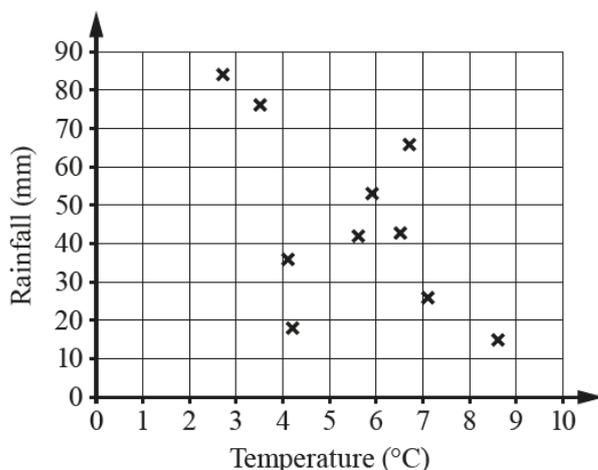


Fig. 6.3

- (i) Explain why the researcher might come to the conclusion that a test based on Pearson's product moment correlation coefficient may be valid. [2]

Although this is a standard question, over one third of candidates did not get full marks. Most mentioned the fact the data on the scatter diagram is very approximately elliptical. Some slightly questionable comments about bivariate Normality were condoned but unfortunately quite a few went on to say that the data (rather than the distribution of the parent population) was perhaps therefore bivariate Normal, thus getting only 1 mark out of 2. It is advisable to state explicitly that the parent population is bivariate Normal, rather than simply a vague comment about bivariate Normality.

Question 6 (b) (ii)

- (ii) Find the value of Pearson's product moment correlation coefficient. [2]

Many candidates calculated S_{xx} , S_{yy} and S_{xy} in order to find the value of the correlation coefficient. However, there were no method marks available for using this method, since, as it states in the specification, 'The use of a calculator is expected for calculation from raw data.'. Candidates should be aware that unless there is a specific instruction 'In this question you must show detailed reasoning' then they should use the most efficient method, which in this case is the use of a calculator.

Question 6 (b) (iii)

- (iii) Carry out a test at the 5% significance level to investigate whether there is any correlation between temperature and rainfall. [5]

In this type of question, candidates should give their hypotheses in terms of ρ and then define ρ , in this case as the population correlation coefficient between temperature and rainfall. Many of those candidates who tried to do this did not get full marks since they did not fully define ρ . Others gave their hypotheses in words, which could also gain full marks if they mentioned population, but most did not do so. The remaining parts of the test were very well done, with almost all candidates correctly comparing the test statistic with the critical value. This could either be done by comparing $|-0.5638|$ with 0.6319 or by comparing -0.5638 with -0.6319 (but not -0.5638 with 0.6319).

Question 7 (b)

- 7 A swimming coach believes that times recorded by people using stopwatches are on average 0.2 seconds faster than those recorded by an electronic timing system. In order to test this, the coach takes a random sample of 40 competitors' times recorded by both methods, and finds the differences between the times recorded by the two methods. The mean difference in the times (electronic time minus stopwatch time) is 0.1442 s and the standard deviation of the differences is 0.2580 s.

- (b) Explain whether there is evidence to suggest that the coach's belief is correct. [2]

This part was well answered although some candidates simply stated that since the whole of the confidence interval was above zero then the coach was probably correct, without mentioning that the interval contained 0.2. Such candidates were given 1 mark out of the 2 available.

Question 7 (c)

- (c) Explain how you can calculate the confidence interval in part (a) even though you do not know the distribution of the parent population of differences. [2]

There was a generous first mark for mentioning the Central Limit Theorem. The second mark required candidates to mention 'the distribution of the sample mean (of differences)' and under one third of them actually did so.

Question 7 (d)

- (d) If the coach wanted to produce a 95% confidence interval of width no more than 0.12 s, what is the minimum sample size that would be needed, assuming that the standard deviation remains the same? [3]

Although there were many correct responses, some candidates rounded their answer to 71 rather than realising that the value of 71 would be too small. Exemplar 4 shows a candidate who forgot to halve 0.12 and so only got one mark out of three. There is a further error since the value of 17.76 should have been rounded up to 18 rather than down to 17.

Exemplar 4

$$0.12 \geq \frac{1.96 \times 0.258}{\sqrt{n}}$$

$$\therefore \frac{1.96 \times 0.258}{0.12} \geq \sqrt{n}$$

$$\therefore 4.1214 \geq \sqrt{n}$$

$$\therefore n \leq 17.76$$

$$\therefore n = 17$$

Question 8 (a)

- 8 A student doing a school project wants to test a claim which she read in a newspaper that drinking a cup of tea will improve a person's arithmetic skills. She chooses 13 students from her school and gets each of them to drink a cup of tea. She then gives each of them an arithmetic test. She knows that the average score for this test in students of the same age group as those she has chosen is 33.5. The scores of the students she tests, arranged in ascending order, are as follows.

26 28 29 30 31 32 34 42 49 54 55 56 61

The student decides to use software to draw a Normal probability plot for these data, and to carry out a Normality test as shown in Fig. 8.

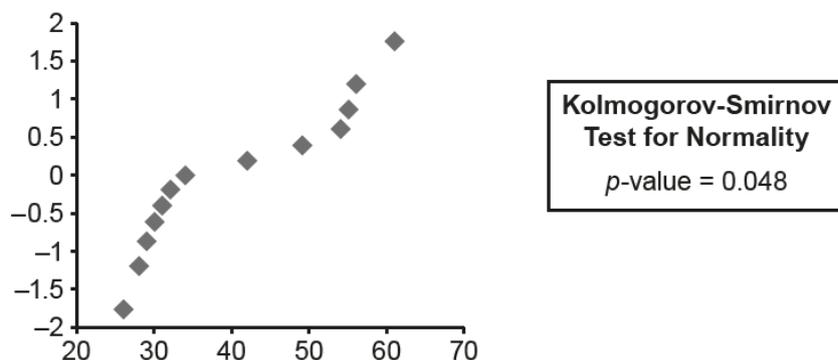


Fig. 8

- (a) The student uses the output from the software to help in deciding on a suitable hypothesis test to use for investigating the claim about drinking tea. Explain what the student should conclude. [3]

All that was required in this part was to state that the p -value was low and the plot was not approximately straight so a Normal distribution was not appropriate and a Wilcoxon test would be appropriate. This is a standard question and there has been at least one similar question in a practice paper. It is surprising therefore that under half of the candidature scored full marks. Many either mentioned the p -value or the straightness but not both. A number of candidates thought that a t -distribution should be used, sometimes thinking the distribution was Normal and sometimes not. There was no credit available in part (c) for the use of a t -distribution.

Question 8 (b)

- (b) The student's teacher agrees with the student's choice of hypothesis test, but says that even this test may not be valid as there may be some unsatisfactory features in the student's project. Give three features that the teacher might identify as unsatisfactory. [3]

Most candidates gained at least one mark but relatively few got all 3 marks. The majority did try to answer the question but a number gave responses that related to part (a). Exemplar 5 shows a typical response which was given one mark for stating that the sample was small (although the $n < 30$ was not relevant). The comment about unknown population variance was irrelevant. The suggestion that it is unknown whether or not the population follows a Normal distribution was also irrelevant since a Wilcoxon test can be carried out in either case.

Exemplar 5

- Sample is small ($n < 30$)
 - Does not know population variance.
 - not sure if the underlying distribution follows a normal or not.

Question 8 (c)

- (c) Assuming that the student's procedures can be justified, carry out an appropriate test at the 5% significance level to investigate the claim about drinking tea. [7]

As mentioned above, credit was only available if a Wilcoxon test was attempted. Those who did do so usually correctly found the test statistic and the critical value. Often however they only scored 1 mark out of 2 for the hypotheses as although they did state ' H_0 : median = 33.5', they did not mention 'population'. A few did mention population but had a two-tailed alternative hypothesis. Such candidates could get a maximum of 5 marks out of the 7 available.

Question 9 (a)

- 9 Every weekday Jonathan takes an underground train to work. On any weekday the time in minutes that he has to wait at the station for a train is modelled by the continuous uniform distribution over $[0, 5]$.

(a) Find the probability that Jonathan has to wait at least 3 minutes for a train. [1]

A surprisingly large number of candidates did not score the mark here. There were many answers of 0.6, probably due to misreading the question and giving the wrong tail. Rather more surprisingly a considerable number of candidates gave an answer of 0.5. Perhaps they thought that they were dealing with a discrete uniform distribution over the interval $\{0, 1, 2, 3, 4, 5\}$. A few candidates used integration to answer the question which was acceptable but unnecessary.

Question 9 (b)

The total time that Jonathan has to wait on two days is modelled by the continuous random variable X with probability density function given by

$$f(x) = \begin{cases} \frac{1}{25}x & 0 \leq x \leq 5, \\ \frac{1}{25}(10-x) & 5 < x \leq 10, \\ 0 & \text{otherwise.} \end{cases}$$

(b) Find the probability that Jonathan has to wait a total of at most 6 minutes on two days. [3]

This was very well answered with rather strangely over three quarters of candidates getting this part correct as compared to only half getting part (a) correct.

Question 9 (e)

Jonathan's friend suggests that the total waiting time for 5 days, T minutes, will almost certainly be less than 18 minutes. In order to investigate this suggestion, Jonathan constructs the simulation shown in Fig. 9. All of the numbers in the simulation have been rounded to 2 decimal places.

	A	B	C	D	E	F
1	Mon	Tue	Wed	Thu	Fri	Total T
2	1.78	4.36	2.74	3.88	4.64	17.41
3	0.95	1.30	4.83	4.29	1.81	13.18
4	4.27	4.90	4.57	1.41	3.66	18.81
5	0.80	0.06	3.20	1.76	0.35	6.17
6	0.03	4.82	1.26	3.53	0.13	9.77
7	3.88	4.73	1.19	3.75	1.29	14.84
8	4.11	3.54	4.33	0.77	4.50	17.25
9	3.54	0.11	3.85	2.86	1.58	11.94
10	1.87	1.82	3.00	3.53	1.83	12.05
11	4.00	2.98	4.59	1.73	1.76	15.06
12	1.91	3.85	2.08	1.72	2.82	12.38
13	0.10	4.86	2.51	0.52	2.17	10.15
14	1.24	4.26	0.95	1.33	1.78	9.57
15	2.99	0.69	3.85	3.41	2.42	13.36
16	4.67	1.76	2.13	3.48	3.10	15.14
17	1.94	1.07	0.91	0.63	3.34	7.89
18	0.11	2.29	0.71	4.21	0.86	8.18
19	0.43	4.58	4.89	1.86	2.84	14.60
20	4.23	0.88	2.71	4.88	4.20	16.91
21	3.72	4.58	3.11	4.89	3.18	19.49

Fig. 9

Jonathan thinks that he can use the Central Limit Theorem to provide a very good approximation to the distribution of T .

(e) Find each of the following.

- $E(T)$
- $\text{Var}(T)$

[3]

This question caused problems for the vast majority of candidates. Unfortunately, they thought that they had to calculate the expectation and variance of the values in the simulation. What they should have done is to find the expectation and variance of the continuous uniform distribution over $[0, 5]$ and then multiply both by 5 because there are 5 days involved.

Question 9 (f)

- (f) Use the Central Limit Theorem to estimate $P(T > 18)$. [2]

Most candidates got 1 mark out of the 2 available for stating that the distribution was Normal with their mean and variance from part (e). A few mistakenly divided the variance by 5 or by 20. Those who had the correct values in part (e) almost always got both marks here.

Question 9 (g)

- (g) Comment briefly on the use of the Central Limit Theorem in this case. [1]

A surprisingly small proportion of candidates scored the mark here. All that was required was to state that the sample was small, so the Central Limit Theorem might not be very accurate. Some thought that the sample size was 20, which of course is wrong. There was a wide variety of other wrong answers.

Question 9 (h)

Jonathan travels to work on 200 days in a year.

- (h) Find the probability that the total waiting time for Jonathan in a year is more than 510 minutes. [3]

This part was again poorly answered. Many candidates tried to use their wrong answers from part (e). No follow through was allowed in this case. Those candidates who used the correct values were almost always successful in finding the required probability.

Question 10 (b)

10 The probability density function of the continuous random variable X is given by

$$f(x) = \begin{cases} kx^m & 0 \leq x \leq a, \\ 0 & \text{otherwise,} \end{cases}$$

where a , k and m are positive constants.

(b) Find the cumulative distribution function of X in terms of x , a and m .

[4]

The majority of candidates gained at least 3 of the 4 marks available in this part. The two main errors were not to include any limits in their initial integral and to leave the final answer as a partial definition

$F(x) = \frac{x^{m+1}}{a^{m+1}}$, omitting ' $= 0$ for $x < 0$ ' and ' $= 1$ for $x > a$ '. Some had the correct function for $0 < x < a$ but then wrongly stated ' $= 0$ otherwise'.

Question 10 (c) (i)

(c) Given that $P\left(\frac{1}{4}a < X < \frac{1}{2}a\right) = \frac{1}{10}$,

(i) show that $2p^2 - 10p + 5 = 0$, where $p = 2^m$,

[4]

This question caused problems for many candidates with only about one third getting more than 2 marks. Most were able to get the first method mark, even if they had the wrong expression for $F(x)$. However many then did not cancel a^{m+1} from the numerator and denominator. Those who did gained the second method mark. However, many candidates did not then realise that $2^{m+1} = 2p$ and $4^{m+1} = 4p^2$ and so could not get the remaining marks.

Question 10 (c) (ii)

(ii) find the value of m .

[3]

This final part was much more successfully tackled, with over half of the candidature getting all 3 marks. The main error was to give the correct answer, but then also to give the negative value of -0.8275 , and not reject it, which meant that they did not get the final mark. Some candidates found the two values of p correctly but then did not know how to solve the equation $2^m = 4.4365$, despite some of these then stating that $m = \log_2 4.4365$.

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

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.*

© **OCR 2019** Oxford Cambridge and RSA Examinations is a Company Limited by Guarantee. Registered in England. Registered office The Triangle Building, Shaftesbury Road, Cambridge, CB2 8EA. Registered company number 3484466. OCR is an exempt charity.



Cambridge
Assessment

