Qualification Accredited

OCR
Oxford Cambridge and RSA

A LEVEL

Examiners' report

FURTHER MATHEMATICS A

H245

For first teaching in 2017

Y542/01 Summer 2024 series

Contents

© OCR 2024

Introduction

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

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

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

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

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 > Export to** 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).

© OCR 2024

Paper 1 series overview

The overall standard of work on this paper was high. Candidates found much of it accessible, and most were well prepared to deal with not only hypothesis tests but also probability questions involving selections and arrangements. However, as in the past, there were weaknesses in answering questions that require explaining and written communication, particularly those on modelling.

In general, candidates seemed to have enough time to complete the paper.

Most candidates were able to state hypotheses and conclusions to significance tests well.

Perhaps the weakest aspect of candidates' performance was that many did not use language with sufficient precision in mathematical contexts. Many did not use or interpret technical terms with sufficient precision, but the issue goes wider than that. This is discussed in detail in several places in this report. The precise use of language is fundamental in Statistics, and candidates need to be aware of the precise meaning of what they write.

Candidates who did well on this paper generally:	Candidates who did less well on this paper generally:
• used technical terms precisely and correctly • read questions carefully • gave modelling assumptions and conclusions in terms of the context of the question • fully understood the fact that 'sample mean' and 'population mean' are different things and usualy take different values • appreciated that the central limit theorem refers to the distribution of the sample mean • understood and used the difference between 'sample mean' and 'expected value'.	• used technical terms imprecisely, as if they were using the words in everyday speech • read questions too hastily and jumped to wrong conclusions about what they were asked to do • gave modelling assumptions and conclusions in terms of 'events' • assumed that 'sample mean' and 'population mean' are interchangeable • confused 'the distribution of the sample mean' with 'the distribution of the sample' • treated 'sample mean' and 'expected value' as if they were the same thing.

Misconception

It seems that many candidates do not truly understand the distinction between sample mean and expected value. For example, in Question 5 (b) the issue was whether a sample mean and a sample variance of 3.55 and 5.6475 indicated that a Poisson distribution was unsuitable. Many candidates said that it was unsuitable because in a Poisson distribution, the mean had to be equal to the variance. But this refers to the expected values, not to the sample mean and variance. Even if the expected mean and variance are equal, you would not expect the sample mean and variance to be exactly equal for any particular sample. The issue is: if the sample mean and variance are close together (not equal), it is plausible that the population mean and variance are equal (and 'equal' here does not mean 'approximately equal') Likewise, in Question 8 (a) many candidates simply equated $E(T)$ to the sample mean as if they were the same thing.

It is possible that many such candidates, if asked to explain the difference, would be able to give the answer, but under pressure they tend to fall back to their comfort zone, in which, it seems, there is no difference between them.

© OCR 2024

Question 1 (a)

1 A discrete random variable X has the following distribution, where a, b and c are constants.

x	0	1	2	3
$P(X=x)$	a	b	c	0.1

It is given that $E(X) = 1.25$ and $Var(X) = 0.8875$.

(a) Determine the values of a, b and c. **[5]**

Responses were almost always right, apart from those candidates who used $Var(X) = E(X^2) - E(X)$, $E(X^2) + [E(X)]^2$ or just $E(X^2)$.

Question 1 (b)

(b) The random variable Y is defined by $Y = 7 - 2X$.

Write down the value of $Var(Y)$. **[1]**

Responses were almost always right.

Question 1 (c)

(c) Twenty independent observations of X are obtained. The number of those observations for which $X = 3$ is denoted by T.

Find the value of $Var(T)$. **[2]**

Many candidates did not read the question carefully enough. Many wrote down the variance of the sum, or the mean, of 20 observations of X, and did not use T at all. Those who identified the distribution of T as binomial were usually successful.

© OCR 2024

Question 2 (a)

2 A newspaper article claimed that "taller dog owners have taller dogs as pets". Alex investigated this claim and obtained data from a random sample of 16 fellow students who owned exactly one dog. The results are summarised as follows, where the height of the student, in cm, is denoted by h and the height, in cm, of their dog is denoted by d.

$$n = 16 \quad \sum h = 2880 \quad \sum d = 660 \quad \sum h^2 = 519276 \quad \sum d^2 = 30000 \quad \sum hd = 119425$$

(a) Calculate the value of Pearson's product moment correlation coefficient for the data. **[2]**

Responses were almost always right, although 0.40086... should not be rounded to 0.4.

Question 2 (b)

(b) State what your answer tells you about a scatter diagram illustrating the data. **[1]**

Answers were often imprecise. The question required a description of the scatter diagram, so 'weak positive correlation' was inadequate. 'Points are widely scattered' is also insufficient as it is only the vertical scattering that matters. Stronger responses were similar to 'points do not lie close to the line'.

Question 2 (c)

(c) Use the data to test, at the 5% significance level, the claim of the newspaper article. **[5]**

Generally this part was well answered. It is best to state the hypotheses in terms of population parameters, for example H_0: $\rho = 0$, H_1: $\rho > 0$ and then to interpret ρ as the population correlation coefficient between students' and dogs' heights. A few wrote H_1: $\rho > 1$. The conclusions were generally well expressed.

When defining ρ, 'population product-moment correlation coefficient' should not be abbreviated to PPMCC, as this could simply mean 'Pearson's PMCC'. 'Population PMCC' is acceptable.

© OCR 2024

Question 2 (d)

(d) Explain whether the answer to part **(a)** would be likely to be different if the dogs' weights had been used instead of their heights. **[1]**

Many responses were imprecise, for example, 'not different as weight is proportional to height'. Many candidates used 'different' or 'proportional' in vague, informal terms (to mean 'not very similar' or 'generally increasing'). Candidates need to realise the word 'proportional' is used in a precise way in mathematics. Some seemed to claim that there was an exact linear relationship between height and weight.

Exemplar 1

2(d)	It likely would not as ~~height~~ weight is likely just a linear coding of the height. ~~...~~ Linear coding doesn't affect the pmcc calculated so the result of the test is the same.

In this context, linear coding means that heights and weights are related by an exact formula of the form $w = a + bh$, so that on a scatter diagram showing heights and weights, all the points lie exactly on a straight line. However, this answer suggests that 'linear coding' means only an approximately increasing relationship. Statements in this question needed to be precise.

The correct answer is that it is almost certain that the answer would be different, as it is extremely unlikely that (for this sample) height and weight are related by an exact linear relationship. Obviously, there is no exact relationship in the population as a whole; thinking in terms of scale factors, it might be reasonable to suggest that there was some correlation between weight and $(\text{height})^3$, although of course this too would be far from perfect.

Assessment for learning

This question would give good opportunities for discussion of the way in which words used in common speech need to be used with far greater precision in mathematics.

© OCR 2024

Question 3 (a)

3 Research suggests that the mean reading age of a child about to start secondary school is 10.75. The reading ages, X years, of a random sample of 80 children who were about to start secondary school in a particular district were measured, and the results are summarised as follows.

$n = 80 \qquad \sum x = 893 \qquad \sum x^2 = 10267$

(a) Test at the 5% significance level whether the mean reading age of children about to start secondary school in this district is **not** 10.75. **[10]**

This intricate although routine question was generally well answered. Common errors included not multiplying the sample variance by 80/79, omitting the factor of 80 in the variance of the sample mean, and – particularly relevant in the context of this paper as a whole – confusing the roles of 10.75 and 11.1625. The null hypotheses requires that the sample mean $\bar{X}$ has the distribution N(10.75, 3.78/80), and then the probability that $\bar{X} > 11.1625$ is found. It is wrong to use N(11.1625, 3.78/80) and P(< 10.75).

For the conclusion, a double negative is needed, for instance 'there is insufficient evidence that the mean reading age is not 10.75'. Most candidates were able to give this.

Question 3 (b)

(b) A student wrote: "Although we do not know that the distribution of X is normal, the central limit theorem allows us to assume that it is, as the sample size is large." This statement is incorrect. Give a corrected version of the student's statement. **[1]**

Many candidates, perhaps helped by previous examiner reports, were able to give the correct answer that it is the sample mean $(\bar{X})$, not *X*, that can be assumed to have a normal distribution. However, there are still many candidates who did not understand this and who gave wrong answers.

Misconception

In most previous reports, attention has been drawn to widespread misunderstanding of the Central Limit Theorem. Many candidates think that they can assume that the distribution of *X* (the 'parent distribution') is normal, but this is false. Others say that it can be assumed that the sample is normally distributed, but the (expected) distribution of a random sample of single items is the same as the distribution of the parent population. The bigger the sample, the closer the distribution of the sample is expected to be to that of the parent population.

In this question there are two random variables: the value (*X*) of a single observation and the mean $(\bar{X})$ of a sample of 80 observations. These two random variables have different distributions, and the one that can be assumed to have a normal distribution is the sample mean $\bar{X}$. No sample, however large, can turn a non-normal distribution into a normal one.

© OCR 2024

Exemplar 2

3(b)	Given a large enough Sample size, all distributions approximate to Normal So CLT allows us to use the Normal distribution as an approximation for the distr. of X

This appears to be a very widespread misunderstanding. Candidates need to distinguish the distribution of a random sample of single items taken from the population from the distribution of the sample mean of a random sample of size 80. A histogram of results from the first distribution will simply show the shape of the parent distribution, and the larger the sample, the nearer to the shape of the parent distribution it is likely to be – certainly not normal, unless the parent distribution is normal.

If by contrast you pick, say, 15 samples each of size 80, calculate the sample means of each of the 15 samples and plot those on a histogram, you should get something recognisably normal. (Note that it is the size of each sample (here 80) that has to be large, and not the number of separate random samples (here 15) obtained.)

Assessment for learning

The last two paragraphs suggest a valuable classroom demonstration on a spreadsheet to show histograms of the two distributions. To get a good histogram it might be better to choose more than 15 random samples but picking samples of size 25 from a continuous distribution on (0, 1) works well. For the uniform distribution you can use interval widths of 0.1, and for the distribution of sample means, interval widths of 0.02. The Excel syntax for selecting a random number from a uniform distribution on (0, 1) is =RAND().

Question 4 (a)

4 (a) Write down the number of ways of choosing 5 objects from 12 distinct objects. **[1]**

Responses were almost always right.

© OCR 2024

Question 4 (b)

(b) Each possible set of 5 different integers selected from the integers 1, 2, …, 12 is obtained, and for each set, the sum of the 5 integers is found. The sum S can take values between 15 and 50 inclusive. Part of the frequency distribution of S is shown in the following table, together with the cumulative frequencies.

S	15	16	17	18	19	20	21	22	23
Frequency	1	1	2	3	5	7	10	13	17
Cumulative Frequency	1	2	4	7	12	19	29	42	59

Use these numbers to determine the critical region for a 1-tail Wilcoxon rank-sum test at the 2% significance level when $m = 5$ and $n = 7$. **[2]**

Candidates seemed to find this challenging. The question tested the specification item 'understand the basis of non-hypothesis tests', and it was necessary to use the result from part (a). Most correct answers used 0.02×792 = 15.84 and compared this with 12 and 19. (The principle is that if the null hypotheses is true, all 792 orders are equally likely.) As there are only 12 different orders that produce a sum of 19 or fewer, the probability of obtaining a total of 19 or fewer purely by chance is less than 2% (12/792 = 0.015…). Those who attempted to use a normal approximation or tables had not followed the instruction in the question to 'use these numbers'.

Question 4 (c)

(c) A student says that, for a Wilcoxon rank-sum test on samples of size m and n, where m and n are large, the mean and variance of the test statistic R_m are 200 and $616\frac{2}{3}$ respectively.

Show that at least one of these values must be incorrect. **[3]**

This was generally well answered. Most could obtain the two relevant equations for *m* and *n* and eliminate the common factor ($m + n + 1$) to get $n = 18.5$. All that was then needed was the comment that *n* had to be an integer, so at least one of the values was impossible. There were some numerical errors in working with $616\frac{2}{3}$., some taking it to be $616 \times \frac{2}{3}$. Some wasted time by finding *m* as well as *n*; once *n* has been found not to be an integer, no more work is needed.

© OCR 2024

Question 5 (a)

5 Some bird-watchers study the song of chaffinches in a particular wood. They investigate whether the number, N, of separate bursts of song in a 5 minute period can be modelled by a Poisson distribution. They assume that a burst of song can be considered as a single event, and that bursts of song occur randomly.

(a) State **two** further assumptions needed for N to be well modelled by a Poisson distribution. **[2]**

As often in the past, questions about modelling assumptions were challenging to candidates. Candidates struggle to distinguish between the conditions under which a Poisson distribution may be considered (the random variable is the number of separate events that occur in a given time or space interval, and there is no upper limit), and assumptions. Assumptions are usually in doubt; conditions are not in doubt and should not be mentioned.

In the case of a Poisson distribution, the 'independence' assumption is usually quite well expressed, but the other assumption seems widely misunderstood. Many candidates write down a learnt phrase in such a way that they show it has no meaning for them. For instance, 'bursts of song must occur at an average constant rate', or 'there is an average frequency per time period' are meaningless. Every year, the answer 'they must occur at a constant rate' suggests that candidates may not appreciate the randomness involved; 'constant rate' would mean exactly evenly spaced, with the same number in each time interval, like a ticking clock. To show that this is not being over-critical, each year there are answers such as 'there must be the same number of bursts of song in each 5-minute interval'.

Finally, as frequently noted before, 'singly' is not an assumption for the Poisson distribution. There is no reason why a Poisson distribution should be invalid just because some bursts of song overlap.

Assessment for learning

The following commonly-seen statements are inappropriate:

- 'There must be the same number of bursts of song in each 5-minute interval'
- 'Bursts of birdsong occur at a constant rate'
- 'The probability of a burst occurring is constant'.

In a 5-minute period there can easily be more than one burst of song. A stronger response would be 'the probability in a sufficiently small time interval is constant', but this begs the question of what 'sufficiently small' means. Even for small intervals, the probability is proportional to the length of the interval. The probability of an event in a single (small) time interval is what is meant by the rate. Trying to give a precise statement in terms of probability is complicated, which is why 'the probability of an event occurring is constant' is not an acceptable answer. 'Constant average rate' is a stronger response.

Some wrote, 'Bursts of song occur at a constant, average rate'. The comma suggests that 'average rate' means 'a rate that does not fluctuate much', whereas these two words form a technical term represented by the parameter λ.

© OCR 2024

Question 5 (b)

The bird-watchers record the value of N in each of 60 periods of 5 minutes. The mean and variance of the results are 3.55 and 5.6475 respectively.

(b) Explain what this suggests about the validity of a Poisson distribution as a model in this context. **[2]**

As stated above, this was another question that showed much confusion between sample and population mean and variance. A very common answer was that for a Poisson distribution the mean and variance had to be equal, but here they were not. A precise explanation might start, 'As the expected mean and variance of a Poisson distribution are equal, the sample mean and variance should be similar'. It was sufficient to say that the mean and the variance were not close together.

Candidates need to give nuanced conclusions, and not 'the Poisson distribution is not valid', which is too definite.

Exemplar 3

5(b)	Mean and variance of poisson distribution should be equal. since the measured mean and variance is different a poisson distribution is likely not valid

This is the common incorrect answer from candidates who have not distinguished the expected values of the mean and the variance from the mean and variance of a sample.

© OCR 2024

Question 5 (c)

The complete results are shown in the table.

n	0	1	2	3	4	5	6	7	8	$\geqslant 9$
Frequency	10	3	7	8	13	6	6	2	5	0

The bird-watchers carry out a χ^2 goodness of fit test at the 5% significance level.

(c) State suitable hypotheses for the test. **[1]**

Most candidates gave almost correct hypotheses, but as explained in previous reports, the only fully correct statement of the null hypothesis is that *N* (the number of bursts of birdsong) has a Poisson distribution. On this occasion candidates could be given the mark for looser statements such as 'the Poisson distribution is a good model', 'data can be well modelled by…', 'data is consistent with Poisson', or 'data follows/fits Poisson' – but *not* 'there is evidence that…' (as statements about evidence belong only in the conclusion). Nor is the answer 'the data can be modelled by a Poisson distribution' adequate as anything can be modelled, by almost anything.

Further, it was important for candidates to appreciate that they are not testing for a distribution with a specific parameter, such as Po(3.55). The question concerns a general Poisson distribution, with mean estimated from the data. This of course affects the number of degrees of freedom, and hence the critical value, in the test.

Question 5 (d)

(d) Determine the contribution to the test statistic for $n = 3$. **[3]**

Often well done, but some candidates tried to use the mean of 3.55 and 5.6475 as their value of λ, which is not the correct procedure. As the command word is 'Determine', working was needed, and not just a value obtained from a calculator.

Question 5 (e)

(e) The total value of the test statistic, obtained by combining the cells for $n \leqslant 1$ and also for $n \geqslant 6$, is 9.202, correct to 4 significant figures.

Complete the goodness of fit test. **[3]**

A surprising number of candidates tried to work out the critical value for themselves, usually wrongly, and sometimes adding in 9.202 as one of the terms. Such candidates lost time, but not marks. Many used the wrong critical value, typically 11.07 or 7.815. For these critical values, however, they could be given 2 of the 3 marks. A double negative is needed in the conclusion: 'there is insufficient evidence that the distribution is not Poisson', or equivalent.

© OCR 2024

Question 5 (f)

(f) It is known that chaffinches are more likely to sing in the presence of other chaffinches.

Explain whether this fact affects the validity of a Poisson model for N. **[1]**

Many candidates understood the right idea but expressed themselves poorly. Most knew that the information in the question meant that bursts of song were not independent, and so the validity of the Poisson model was reduced.

Candidates should contextualise their answers, and not refer merely to 'events'.

Question 6 (a)

6 A bag contains 6 identical blue counters and 5 identical yellow counters.

(a) Three counters are selected at random, without replacement.

Find the probability that at least two of the counters are blue. **[2]**

This question was often answered very well. Over the years there has been a substantial improvement in the standard of answers in a topic that is often considered hard. Candidates should in general be encouraged to use ${}^nC_r$ and ${}^nP_r$ methods, and not multiplication of probabilities where it is all too easy to omit crucial constants.

Responses were quite often right, especially from those who used $({}^6C_2 \times 5 + {}^6C_3) / {}^{11}C_3$.

Question 6 (b)

All 11 counters are now arranged in a row in a random order.

(b) Find the probability that all the yellow counters are next to each other. **[2]**

Responses were often right. A common wrong answer was $7 / {}^{11}P_5$ rather than $7 / {}^{11}C_5$.

Question 6 (c)

(c) Find the probability that no yellow counter is next to another yellow counter. **[3]**

Most candidates could find the denominator of ${}^{11}C_5$ but quite a few used a numerator of 1.

© OCR 2024

Question 6 (d)

(d) Find the probability that the counters are arranged in such a way that **both** of the following conditions hold.

- Exactly three of the yellow counters are next to one another.
- Neither of the other two yellow counters is next to a yellow counter. **[3]**

This was quite challenging, but nevertheless many right answers were seen. The best plan was to draw a diagram with six blue counters and seven spaces, one space between each counter and one at each end. Then the set of three counters can be placed in seven ways, and the remaining two single counters is ${}^6C_2$ ways.

The most common wrong answer was ${}^7C_2$ / ${}^{11}C_5$ = 5/66. Presumably this came from simply selecting three spaces where the yellow counters could go, without distinguishing which of the spaces was occupied by three counters.

Question 6 (e)

(e) Explain whether the answer to part **(d)** would be different if the yellow counters were numbered 1, 2, 3, 4 and 5, so that they are not identical. **[1]**

Some candidates saw that this would make no difference, either because numbering the counters makes no difference to the conditions, or because both numerator and denominator are multiplied by the same factor (5).

Question 7 (a) (i)

7 The coordinates of a set of 10 points are denoted by (x_i, y_i) for $i = 1, 2, \ldots, 10$.
For a particular set of values of (x_i, y_i) and any constants a and b it can be shown that

$$\Sigma(y_i - a - bx_i)^2 = 10(11 - a - 6b)^2 + 126\left(b - \frac{83}{42}\right)^2 + \frac{139}{14}.$$

(a) **(i)** Explain why $\Sigma(y_i - a - bx_i)^2$ is minimised by taking $b = \frac{83}{42}$ and $a = 11 - 6b$. **[1]**

This was not the standard question asking what is meant by 'least squares', although many candidates answered it as if it were. A strong response would be that the squared terms on the right-hand side are automatically minimised by setting them equal to 0.

© OCR 2024

Question 7 (a) (ii)

(ii) Hence explain why the equation of the regression line of y on x for these points is given by the corresponding values of a and b (so that the equation is $y = \frac{83}{42}x - \frac{6}{7}$). **[1]**

Answering 'because this minimises the expression' merely repeats part (a) (i). It was necessary to explain that it minimises the sum of the squared vertical distances from the points to the line. Many candidates omitted to say 'squared'.

Question 7 (b)

(b) State which of the following terms **cannot** apply to the variable X if the regression line of y on x can be used for estimating values of Y.

Dependent Independent Controlled Response **[1]**

Perhaps because of the negative in the question, this part was answered less well. *Y* might be the dependent or the response variable, so these are the ones that *X* cannot be.

Question 7 (c)

(c) Use the regression line to estimate the value of y corresponding to $x = 8$. **[1]**

Reponses were almost always right.

Question 7 (d)

(d) State what must be true of the value $x = 8$ if the estimate in part **(c)** is to be reliable. **[1]**

Most knew that extrapolation had to be avoided, but there were many answers that were poorly expressed. It is correct to say, '8 must be within the range of the xi values' (and some candidates used the even more accurate term 'domain'), but not that '8 must be one of the values in the data set', or even 'the point must be on the line', both of which were seen several times.

© OCR 2024

Question 7 (e)

(e) Variables u and v are related to x and y by the following relationships.

$u = 2 + 4x \quad v = 8 - 2y$

Show that the gradient of the regression line of v on u is very close to -1. **[3]**

This was often done well. It contrasts interestingly with the common question 'do these substitutions change the equation of the regression line?' which is often wrongly answered by saying 'no, as this is linear coding'. There were several successful methods: substituting the rearranged formulae into the equation of *y* on *x*; using what was effectively a chain rule $\frac{dv}{du} = \frac{dy}{dx} \times \frac{dv}{dy} \div \frac{du}{dx}$ or equivalents such as

$b' = S_{uv} / S_{uu} = (S_{xy} \times (-2\times4)) / (S_{xx} \times 4^2)$; or even finding two points on the *y*-on-*x* line, converting to *u* and *v* and finding the equation of the new line.

Question 8 (a)

8 A random sample of 100 students were given a task and the time taken by each student to complete the task was recorded. The maximum time allowed to complete the task was one minute and all students completed the task within the maximum time. The times, T minutes, for the random sample of students are summarised as follows.

$n = 100 \qquad \sum t = 61.88$

A researcher proposes that T can be modelled by the continuous random variable with probability density function

$$f(t) = \begin{cases} \alpha t^{\alpha - 1} & 0 \leqslant t \leqslant 1, \\ 0 & \text{otherwise,} \end{cases}$$

where α is a positive constant.

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

By finding $E(T)$ according to the researcher's model, determine an approximation for the value of α. Give your answer correct to **3** significant figures. **[6]**

This was a question where detailed reasoning was required. As explained above, therefore, it was necessary for candidates to justify why they could equate the formula for E(*T*) to the sample mean. Only a few made any comments at all, and so most were given only 5 marks out of 6. Valid comments included 'the sample mean is an unbiased estimate', 'take E(*T*) to be 0.6811' or even use of the ≈ sign.

Nevertheless, many candidates were able to obtain the correct value of α. A few tried to find the median, or did not use definite limits for the integral, and a common error was $\alpha = 0.6188 \times (\alpha + 1) \Rightarrow$
$\alpha = 0.6188\alpha + 1$.

© OCR 2024

Question 8 (b)

Further information about the times taken for the sample of 100 students to complete the task is given in the table.

Time t	$0 \leqslant t < \frac{1}{3}$	$\frac{1}{3} \leqslant t < \frac{2}{3}$	$\frac{2}{3} \leqslant t \leqslant 1$
Frequency	18	37	45

(b) Using the value of α found in part **(a)**, determine the extent to which the proposed model is a good model. (Do not carry out a goodness of fit test.) **[4]**

This was generally answered well. The best strategy was to use the PDF to find the probabilities of being in the three ranges, multiply them by 100 and compare with the data. Most candidates were able to say that it is a good model because observed and expected frequencies were similar. Those who used the cumulative distribution function were also often successful.

© OCR 2024

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 a free test-building platform, providing unlimited users exclusively for staff at OCR centres with an Interchange account.

Choose from a large bank of questions to build personalised tests and custom mark schemes, with the option to add custom cover pages to simulate real examinations. You can also edit and download complete past papers.

Find out more.

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.

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

Need to get in touch?

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

Call us on
01223 553998

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

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

We really value your feedback

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

I like this **I dislike this**

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

CAMBRIDGE
UNIVERSITY PRESS & ASSESSMENT

OCR is part of Cambridge University Press & 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 2024 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.

OCR operates academic and vocational qualifications regulated by Ofqual, Qualifications Wales and CCEA as listed in their qualifications registers including A Levels, GCSEs, Cambridge Technicals and Cambridge Nationals.

OCR provides resources to help you deliver our qualifications. These resources do not represent any particular teaching method we expect you to use. We update our resources regularly and aim to make sure content is accurate but please check the OCR website so that you have the most up to date version. OCR cannot be held responsible for any errors or omissions in these resources.

Though we make every effort to check our resources, there may be contradictions between published support and the specification, so it is important that you always use information in the latest specification. We indicate any specification changes within the document itself, change the version number and provide a summary of the changes. If you do notice a discrepancy between the specification and a resource, please contact us.

You can copy and distribute this resource in your centre, in line with any specific restrictions detailed in the resource. Resources intended for teacher use should not be shared with students. Resources should not be published on social media platforms or other websites.

OCR acknowledges the use of the following content: N/A

Whether you already offer OCR qualifications, are new to OCR or are thinking about switching, you can request more information using our Expression of Interest form.

Please get in touch if you want to discuss the accessibility of resources we offer to support you in delivering our qualifications.