Qualification Accredited

Oxford Cambridge and RSA

AS LEVEL

Examiners' report

FURTHER MATHEMATICS A

H235
For first teaching in 2017

Y532/01 Summer 2018 series

Version 1

Contents

Introduction

Our examiners' reports are produced to offer constructive feedback on candidates' performance in the examinations. They provide useful guidance for future candidates. The reports will include a general commentary on candidates' performance, identify technical aspects examined in the questions and highlight good performance and where performance could be improved. The reports will also explain aspects which caused difficulty and why the difficulties arose, whether through a lack of knowledge, poor examination technique, or any other identifiable and explainable reason.

Where overall performance on a question/question part was considered good, with no particular areas to highlight, these questions have not been included in the report. A full copy of the question paper can be downloaded from OCR.

Paper Y532/01 series overview

This was the first sitting of the AS Further Mathematics optional Statistics component for the new specification. Many candidates demonstrated confidence on the calculation questions and many correct numerical answers were seen, even where the questions were not routine. Questions requiring more descriptive responses were in general not answered so well, particularly when the topic did not appear on the old S2 specification. Topics that were well answered included the calculations involving the Poisson and geometric distributions, and those focused on counting arguments. Modelling assumptions were less well answered; some candidates confused assumptions with conditions, or did not focus their response to the context of the question. Many candidates found the paper challenging, and in particular, questions 7(iii) and 8 were found difficult.

It might be helpful to spell out the issues of modelling conditions and assumptions. Some find that a distinction between the two is helpful.

With this distinction, the conditions for a geometric distribution are that you are counting the number of trials up to and including the first success, and that the outcome of each trial can be classified as "success" or "fail". The condition for a Poisson distribution is that you are counting the number of separate events within a fixed interval of time or space. There is rarely any doubt as to whether these conditions are satisfied and candidates should usually use them as indicators of what distribution is being considered.

Modelling assumptions for a geometric distribution are that the outcome of each event is independent of the outcome of all other events, and that the probability of a success is the same for each event. The assumptions for a Poisson distribution are that events occur randomly, independently and at constant average rate. There is often doubt as to whether these assumptions hold in a given situation.

The questions in this paper focused on the modelling assumptions, rather than the conditions. Candidates are expected to identify what assumptions are not stated or clearly implied by what is given in the question.

Question 1(i)

1 A book reviewer estimates that the probability that he receives a delivery of books to review on any one weekday is 0.1. The first weekday in September on which he receives a delivery of books to review is the Xth weekday of September.

(i) State an assumption needed for X to be well modelled by a geometric distribution. **[1]**

The "constant probability" assumption was implied by the first sentence of the question, so that the missing assumption was "book deliveries need to be received independently". Generic responses, such as "events occur independently", without reference to the context, should be avoided. Conditions such as "there must be only two outcomes" are also clearly implied by the question. Learners are expected to be able to identify which of the conditions or assumptions are not implied by the question.

Question 1(ii)

(ii) Find $\mathrm{P}(X = 11)$. **[2]**

This was generally very well done. A few candidates gave 0.035 as their answer, presumably thinking that this was correct to 3 significant figures.

Question 1(iii)

(iii) Find $\mathrm{P}(X \leqslant 8)$. **[2]**

Again generally very well answered. Those who attempted to add up 8 terms often made accuracy errors. Only a few calculated $1 – 0.9^9$ instead of $1 – 0.9^8$ or omitted the "1 –".

Question 1(iv)

(iv) Find $\mathrm{Var}(X)$. **[2]**

Almost always correctly answered.

Question 1(v)

(v) Give a reason why a geometric distribution might not be an appropriate model for the first weekday in a calendar year on which the reviewer receives a delivery of books to review. **[1]**

This turned out to be quite a subtle question. The issue is: does the constant probability assumption (or independence) apply to the first few weekdays of the year, as opposed to other days in the year (where the assumptions are assumed to hold already). Correct comments included references to holidays at the beginning of the year, a build-up of deliveries following Christmas, or even the idea that more books might be published right at the start of the year. Suggestions such as "deliveries might always occur on the same day of the week" did not gain credit as this would deny the validity of a geometric distribution at other times of the year, including September.

Question 2(i)

2 The probability distribution for the discrete random variable W is given in the table.

w	1	2	3	4
$\mathrm{P}(W = w)$	0.25	0.36	x	x^2

(i) Show that $\mathrm{Var}(W) = 0.8571$. **[7]**

Many very good responses to this question were seen. Candidates were expected to solve a quadratic equation and to exclude the solution $x = -1.3$ explicitly, with a reason such as "the probability must be between 0 and 1". Many did this well. As the result was a given answer, sufficient detail of the calculation of $\mathrm{E}(W)$ and $\mathrm{E}(W^2)$ was needed. Those who calculated $\mathrm{E}(W - \mathrm{E}(W))^2$ needed to give enough detail to convince the Examiners that they had carried out the calculation correctly.

Question 2(ii)

(ii) Find $\mathrm{Var}(3W+6)$. **[1]**

Generally very well answered, with almost everybody avoiding the obvious traps. A few candidates insisted on using their wrong answer from part (i); the point of giving 0.8571 in part (i) was so that candidates did not do this.

Question 3(i)

3 In the manufacture of fibre optical cable (FOC), flaws occur randomly. Whether any point on a cable is flawed is independent of whether any other point is flawed. The number of flaws in 100 m of FOC of standard diameter is denoted by X.

(i) State a further assumption needed for X to be well modelled by a Poisson distribution. **[1]**

The modelling assumption "flaws must occur at constant average rate" (or at "uniform rate") needed to be given in context, so that "events occur at constant average rate" did not gain credit. The specimen assessment material and practice papers have made it clear that "flaws occur singly" should not be given as a modelling assumption (this issue is discussed in the article "Misunderstandings in A-level Statistics", *The Mathematical Gazette*, March 2014). "Constant rate" is also not correct as it implies that flaws occur at fixed, unvarying intervals. Randomness and independence are stated in the question. As on the old S2 papers, a number of candidates gave "flaws must occur with constant probability", which is a confusion with the assumptions needed for the binomial distribution.

Question 3(ii)

Assume now that X can be well modelled by the distribution $\mathrm{Po}(0.7)$.

(ii) Find the probability that in 300 m of FOC of standard diameter there are exactly 3 flaws. **[2]**

Almost always correct.

Question 3(iii)

The number of flaws in 100 m of FOC of a larger diameter has the distribution Po(1.6).

(iii) Find the probability that in 200 m of FOC of standard diameter and 100 m of FOC of the larger diameter the total number of flaws is at least 4. **[3]**

Very many correct responses were seen. Some candidates seemed to be unaware of the additive property of the Poisson distribution and attempted to find all possible combinations of the two random variables; as there were 15 possibilities, such candidates almost never succeeded. Some attempted to use a combined mean but did not calculate it correctly.

Question 4(i)

4 Judith believes that mathematical ability and chess-playing ability are related. She asks 20 randomly chosen chess players, with known British Chess Federation (BCF) ratings X, to take a mathematics aptitude test, with scores Y. The results are summarised as follows.

$$n = 20,\ \Sigma x = 3600,\ \Sigma x^2 = 660\,500,\ \Sigma y = 1440,\ \Sigma y^2 = 105\,280,\ \Sigma xy = 260\,990$$

(i) Calculate the value of Pearson's product-moment correlation coefficient r. **[2]**

Almost always correct. Most candidates showed some working, but they sometimes made numerical errors that could have been avoided by checking their answer against an answer obtained directly from calculator software.

Question 4(ii)

(ii) State an assumption needed to be able to carry out a significance test on the value of r. **[1]**

This request seemed unfamiliar to a majority of candidates. Either of the answers "the data must have a bivariate normal distribution" or "the points should form an approximate ellipse on a scatter diagram" was acceptable. If this assumption does not hold, the hypothesis test is not valid.

Question 4(iii)

(iii) Assume now that the assumption in part **(ii)** is valid. Test at the 5% significance level whether there is evidence that chess players with higher BCF ratings are better at mathematics. **[4]**

Most candidates found the correct critical value. They also gave their conclusions in context and with acknowledgement of the uncertainty involved. A good conclusion might be, "Reject H_0. There is significant evidence that chess players with higher BCF ratings are better at mathematics." Problems arose with stating the hypotheses. Good statements were of one of two forms:

H_0: $\rho = 0$, H_1: $\rho > 0$, where ρ is the population product moment correlation coefficient

OR

H_0: there is no association between BCF rating and mathematical ability. H_1: there is positive association between BCF rating and mathematical ability.

A definition of ρ is crucial, including mention of the word "population" It was also important to include the word "positive" in H_1, since its absence would make their test 2-tailed instead of 1-tailed.

A number of candidates stated their hypotheses the wrong way round: "H_0: there is positive association …". Some wrote: H_0: there is no evidence of association …", but "evidence" refers only to the conclusion. The hypotheses refer to the actual underlying distribution, where either there is positive association or there is not.

Question 4(iv)

(iv) There are two different grading systems for chess players, the BCF system and the international ELO system. The two sets of ratings are related by

$$\text{ELO rating} = 8 \times \text{BCF rating} + 650.$$

Magnus says that the experiment should have used ELO ratings instead of BCF ratings. Comment on Magnus's suggestion. **[1]**

A large number of candidates correctly said that there would be no effect because linear coding does not affect the value of the pmcc. Some candidates said that ELO ratings have a wider geographical validity than BCF ratings, but this is irrelevant to the context of the experiment.

Question 5(i)

5 (i) A team of 9 is chosen at random from a class consisting of 8 boys and 12 girls.
Find the probability that the team contains no more than 3 girls. **[4]**

It was pleasing to see a large number of correct answers to this challenging question, candidates were obviously well prepared for this sort of question. Inevitably there were some who attempted to use the less efficient method of multiplying probabilities together, but in view of the large number of different possibilities it is not surprising that almost no one accurately completed the question by this method.

Question 5(ii)

(ii) A group of n people, including Mr and Mrs Laplace, are arranged at random in a line. The probability that Mr and Mrs Laplace are placed next to each other is less than 0.1. Find the smallest possible value of n. **[4]**

Again this question was largely well answered. Many candidates were able to obtain $2/n$ and compare it with 0.1. Others used $1/n$, or gave their final answers as $n = 20$, or $n \geq 21$. Those who used trial and improvement rarely found an answer close to the correct one.

Question 6

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

The random variable T has a binomial distribution. It is known that $\mathrm{E}(T) = 5.625$ and the standard deviation of T is 1.875. Find the values of the parameters of the distribution. **[5]**

Many fully correct answers were seen to this challenging question. Most realised that 1.875 had to be squared. The most efficient method of solution is to divide 1.875^2 by 5.625 to obtain q. Candidates were able to gain credit by stating that np was 5.625 even if they were unable to progress further.

Question 7(i)

7 An environmentalist measures the mean concentration, c milligrams per litre, of a particular chemical in a group of rivers, and the mean mass, m pounds, of fish of a certain species found in those rivers. The results are given in the table.

c	1.94	1.78	1.62	1.51	1.52	1.4
m	6.5	7.2	7.4	7.6	8.3	9.7

(i) State which, if either, of m and c is an independent variable. **[1]**

In this question neither variable was independent. Many thought that c was independent so that m depended on it, but c is not controlled; this is a genuine bivariate situation.

Question 7(ii)

(ii) Calculate the equation of the least squares regression line of c on m. **[3]**

Some candidates found the equation of the least squares regression line of *m* on *c*, rather than *c* on *m*. Consequently the answer $m = 15.7 - 4.83c$ was seen almost as often as the correct answer $c = 2.85 - 0.157m$. Many worked with *x* and *y*, so that $y = 15.7 - 4.83x$ was also commonly seen.

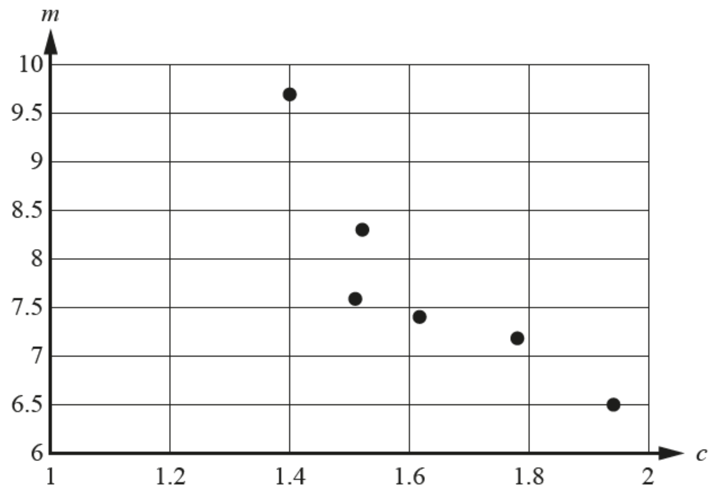
Question 7(iii)

(iii) State what effect, if any, there would be on your answer to part **(ii)** if the masses of the fish had been recorded in kilograms rather than pounds. (1 kg ≈ 2.2 pounds.) **[1]**

Few correct responses were seen here, even though responses from part (ii) were followed through. Many thought that the coefficients were unchanged by linear coding. Others could not work out whether the coefficients had to be divided or multiplied by 2.2.

Question 7(iv)

(iv) The data is illustrated in the scatter diagram. Explain what is meant by 'least squares', illustrating your answer using the copy of this diagram in the Printed Answer Booklet. **[3]**



Candidates needed to draw vertical (or horizontal) displacements from an approximate line of best fit and then say that the sum of the squares of these displacements was minimised. Full credit could not be given for vague statements such as "least squares gives the line nearest to the points". A common incorrect response was to draw diagonal lines from the points to their line.

Question 8

8 The table shows the results of a random sample drawn from a population which is thought to have the distribution U(20).

Range	$1 \leqslant x \leqslant 8$	$9 \leqslant x \leqslant 12$	$13 \leqslant x \leqslant 20$
Observed frequency	12	y	$28-y$

Find the range of values of y for which the data are not consistent with the distribution at the 5% significance level. **[9]**

Many candidates found this question challenging, and in particular many struggled to find the correct expected frequencies. The correct expected frequencies are 16, 8, 16, using the fact that they had to add up to 40. Some thought that the expected frequencies were 8, 4, 8 and hence they combined cells and used the critical value 3.841. Some candidates using one degree of freedom thought that they should use Yates's Correction, but this is not correct; Yates's Correction is not used for goodness-of-fit tests. Other sets of expected frequencies seen quite often were 10.5, 10.5, 10.5 or 12, 6, 12.

Many candidates correctly used $\sum \frac{(O-E)^2}{E} > 5.991$ to obtain a quadratic equation, although some used $\sum \frac{(O-E)^2}{E} < 5.991$. Candidates often got as far as

$3y^2 - 56y + 192.144 > 0 \Rightarrow y < 4.53$ and $y > 14.14$ but the values of y have to be integers less than or equal to 28. The final answer $0 \le y \le 4$ and $15 \le y \le 28$ was rarely seen.

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

Active Results offers a unique perspective on results data and greater opportunities to understand students' performance.

It allows you to:

- Review reports on the **performance of individual candidates**, cohorts of students and whole centres
- **Analyse results** at question and/or topic level
- **Compare your centre** with OCR national averages or similar OCR centres.
- Identify areas of the curriculum where students excel or struggle and help **pinpoint strengths and weaknesses** of students and teaching departments.

http://www.ocr.org.uk/administration/support-and-tools/active-results/

CPD Hub
Your route to OCR's teacher 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.

https://www.cpdhub.ocr.org.uk

The small print

We'd like to know your view on the resources we produce. By clicking on the 'Like' or 'Dislike' button you can help us to ensure that our resources work for you. When the email template pops up please add additional comments if you wish and then just click 'Send'. Thank you.

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

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 Contact Centre

General qualifications
Telephone 01223 553998
Facsimile 01223 552627
Email general.qualifications@ocr.org.uk