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

Y432/01 Summer 2025 series

Version 1

ocr.org.uk/maths

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 Y432 series overview

This is the Statistics Minor option component of the MEI Further Mathematics course.

This component was accessible to most candidates, who typically attempted to answer all questions.

Questions involving calculations (Questions 1, 2 (a), 3 (c), 3 (d), 4 (a), 5 (c) and 7 (a)) were usually done well. These accounted for about one-third of the marks in the component.

For the two questions involving hypothesis tests (Questions 5 (d) and 6 (b)) candidates demonstrated a good understanding of the required structure. Hypotheses were stated, critical values were found, and comparisons made with test statistics. Conclusions were then usually suitably non-assertive. To improve further, candidates should make sure that the null hypothesis is chosen correctly and includes just the right amount of detail. They should also take care in identifying the correct critical value.

Candidates did not perform as well on 'explain' and 'comment' questions (Questions 2 (b), 2 (c), 4 (b) and 6 (a)). Centres should continue to make sure that sufficient attention is given to these questions when preparing candidates.

The conditions required for a probability distribution to be an appropriate model and the conditions for a test to be valid need to be clearly understood and expressed in the context of the question (Questions 3 (a), 3 (b), 5 (a) (i), 5 (a) (ii), 5 (a) (iii) and 5 (b)).

Candidates needed to be able to find the variance from a probability distribution table (Question 7) where the probability was given algebraically, as well as summing independent variables to create a new probability distribution table. Confidence with algebraic manipulation will help candidates with these types of questions.

Useful information regarding sampling, bivariate data and modelling assumptions can be found in the specification for this component.

Candidates should be encouraged to write neatly. Some responses to the 'explain', 'comment' and 'state' questions were challenging to read.

Candidates who did well on this paper generally:	Candidates who did less well on this paper generally:
• made appropriate and accurate use of statistical and probability functions on their calculator • consistently carried out calculations accurately, giving answers to a sufficient degree of accuracy • carried out tests accurately stating correct hypotheses, identifying correct critical values, and drawing non-assertive conclusions • used subject-specific terminology correctly • were able to state modelling assumptions correctly within the context of the question • were confident in using algebra when problem solving with probability distributions.	• carried out standard calculations correctly but did not always maintain sufficient accuracy in their working • carried out basic test procedures reliably, but expressed their hypotheses incompletely or inaccurately • found it difficult to identify the key points in 'explain' and 'comment' questions • were less confident in using algebra when problem solving with probability distributions.

Question 1 (a)

1 A discrete random variable, X, has a Poisson distribution with mean 4.

(a) Write down the value of $\mathrm{Var}(X)$. **[1]**

This question was answered correctly by most candidates.

Question 1 (b)

(b) Find $\mathrm{P}(X = 3)$. **[1]**

This question was answered correctly by most candidates.

Question 1 (c)

(c) Find $\mathrm{P}(X > 3)$. **[2]**

This question was answered correctly by most candidates. Candidates that did make mistakes generally had included the $X = 3$ term.

Question 1 (d)

Another discrete random variable, Y, which is independent of X, has a Poisson distribution with mean 5.

(d) Find $\mathrm{P}(X + Y \leqslant 9)$. **[2]**

This question was correctly answered by most candidates; candidates clearly understood that $\mathrm{Po}(9)$ was needed to find this probability.

Question 2 (a)

2 In an experiment, the speed of an object, $v\,\text{m}\,\text{s}^{-1}$, at time t seconds, is recorded at intervals of 10 seconds from $t = 0$ to $t = 80$.

The summary statistics are as follows.

$n = 8$, $\Sigma t = 280$, $\Sigma t^2 = 14000$, $\Sigma v = 876.28$, $\Sigma v^2 = 133310.2$, $\Sigma tv = 43190$

A linear function of t is used as a model for v.

(a) By first determining the equation of an appropriate regression line, estimate the value of v when $t = 45$. **[3]**

The approach that most candidates took in this question was generally successful, but in most cases the final mark was not given due to candidates stating the final answer to three decimal places. The most successful candidates also considered the degree of accuracy for their final answer, giving it to two or three significant figures.

Assessment for learning

In this question, the accuracy of the final answer matters.

Centres can help candidates by drawing attention to the instructions on the front of the exam paper:

- Where appropriate, your answer should be supported by working. Marks might be given for using a correct method, even if your answer is wrong.
- Give your final answers to a degree of accuracy that is appropriate to the context.

Question 2 (b)

(b) Comment on the reliability of your estimate in part **(a)**. **[1]**

Overall, this was a very well-answered question, with nearly all candidates responding with terms 'reliable' and 'interpolation'. This was a marked improvement from previous years. Occasionally, candidates calculated the product moment correlation coefficient, which allowed for a different response about reliability.

Question 2 (c)

(c) Explain why in this scenario it would be inappropriate to use a value of v to estimate a value of t using **any** regression line. **[1]**

Many candidates correctly stated that t was non-random. Many also incorrectly focused on the dependence or independence of the variables (the relationship between the variables) rather than addressing the method of data selection (recorded at ten-minute intervals).

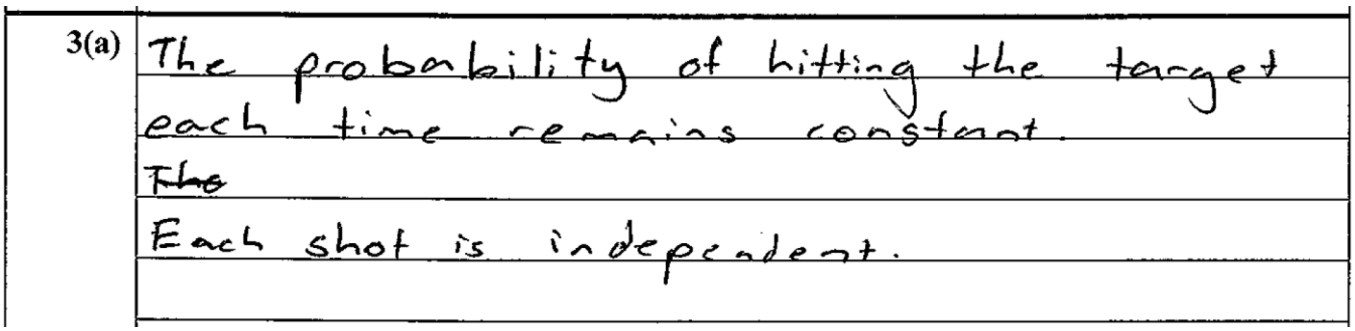
Question 3 (a)

3 An archer fires arrows one by one at a target until an arrow hits the target. The number of arrows fired up to and including the arrow that hits the target is denoted by X. The archer considers using a geometric model for the distribution of X.

(a) State **two** assumptions required for such a model to be appropriate. **[2]**

Many candidates were able to identify one of the assumptions correctly, with correct use of the context. Comments about probability were generally well-made, but many candidates referred to probability when talking about independence. Successful candidates were able to get the second mark by stating 'the independence of the arrows hitting the target', rather than the arrows being fired.

Exemplar 1



This exemplar shows how the candidate was able to identify the probability assumption within the context of the question to get 1 mark. For the second statement, the outcome of the shot has not been defined fully. The outcome of each shot at hitting the target, or not, is independent of the outcome of other attempts.

Assessment for learning

Centres can help candidates by giving them practice with identifying outcomes of trials within a real-world context.

Question 3 (b)

For the remainder of this question you may assume that a geometric model is appropriate for the distribution of X.

The archer keeps records of the values of X. He collects a sample of these values and uses them to estimate the variance of X.

(b) Give **two** features that the sample should have. **[2]**

Most candidates were able to answer that the sample should be large or random for one of the 2 marks. Many stated that the sample should be 'representative of the population', which was also appropriate.

Question 3 (c)

The estimate of the variance of X calculated from the sample is 12.

(c) Determine an estimate of the probability that when the archer fires a single arrow at the target, the arrow hits the target. **[3]**

This was a successfully attempted question by nearly all candidates. Many also found the negative probability, $\mathrm{p} = -\frac{1}{3}$, and correctly identified that they needed to rule it out as a solution.

Question 3 (d)

(d) Use your answer to part **(c)** to write down an estimate of the expected number of arrows that the archer must fire at the target up to and including the first arrow that hits the target. **[1]**

This was correctly answered by nearly all candidates.

Question 4 (a) (i)

4 The discrete random variable, X, has a uniform distribution over the values $\{5, 6, ..., n\}$.

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

In part **(a)** you should assume that $n = 24$.

(i) Using the formula $\mathrm{E}(X) = \sum x_i p_i$, determine the value of $\mathrm{E}(X)$ by direct calculation. **[3]**

Candidates answered this question very well. Most candidates recognised that they needed to show sufficient workings to demonstrate the use of the expectation formula, rather than find the midpoint of 5 and 24. Errors were more likely when unnecessarily evaluating $\mathrm{E}(x)$ in terms of n before substituting a value in or using the expectation formula to calculate $\mathrm{E}(X - 4)$.

Question 4 (a) (ii)

(ii) Use the substitution $Y = X - 4$ and the given formula in the formulae booklet to determine $\mathrm{Var}(X)$. **[2]**

Most candidates were able to use the formula correctly. Some candidates did not mention Y at all, so could only get a maximum of 1 mark for $\mathrm{Var}(X)$.

Question 4 (a) (iii)

(iii) Show that $\mathrm{P}(X > \mathrm{E}(X)) = \frac{1}{2}$. **[1]**

This was a generally well-answered question, but it should be noted that a full justification is needed for a 'show' question. Candidates must show that the given result is true by providing an explanation to cover the argument.

Question 4 (b)

(b) A student states that $\mathrm{P}(X > \mathrm{E}(X)) = \frac{1}{2}$ for **all** $n \geqslant 24$.

Explain whether the student is correct. **[1]**

Many candidates recognised that the student was incorrect when n is odd, which was enough for the mark. There were generally a mix of responses from the candidates, with few deciding that a counterexample such as when $n = 25$ would be a quick and easy way to explain their response, instead most candidates wrote lengthy explanations to justify their response.

Question 5 (a) (i)

5 Long-term data collected at a certain farm indicates that 30% of eggs that it produces are large and the rest are categorised as medium.

Eggs are collected in batches of 4 and the number of large eggs in each batch, X, is recorded.

It is suggested that a binomial distribution might be appropriate to model X.

(a) (i) State **two** reasons, within the context of this question, why a binomial model might be appropriate to model X. **[2]**

There was a mixed response from candidates to this question. Some did not respond in the context of the question; that is, not referring to the eggs correctly. References to the modelling assumptions were often seen, as opposed to the reasons for believing that a binomial model might be appropriate, despite Question 5 (a) (ii) then going on to ask for the assumptions.

Question 5 (a) (ii)

(ii) State **two** assumptions, within the context of this question, that must be made for a binomial model to be appropriate to model X. You should not refer to the reasons given in part **(a)(i)**. **[2]**

Successful candidates were able to give the assumptions in context, such as by stating that 'the size of the eggs are independent'. Many candidates were able to correctly comment on a fixed probability in context; however, many candidates incorrectly wrote about the probability of independence, like in Question 3 (a).

Assessment for learning

Centres can help candidates by giving them practice with examples that help them identify the trial within a real-world context.

Question 5 (a) (iii)

(iii) Write down the binomial distribution which might be a suitable model for the distribution of X. **[1]**

This was correctly answered by nearly all candidates.

Question 5 (b)

The farm manager investigates whether the distribution in part **(a)(iii)** is a good model for X.

(b) State **one** reason why taking data from a sample of 6 batches would **not** be appropriate to carry out the investigation. **[1]**

Most candidates referred to the small sample size, but further explanation was needed of the implication of such a small sample. Successful candidates correctly stated that a small sample size would not be likely to represent the population or '6 batches is too small to draw conclusions about the population'.

Question 5 (c)

The farm manager takes a random sample of 100 batches. The results are summarised in **Table 5.1**.

Table 5.1

x	0	1	2	$\geqslant 3$
Frequency	31	36	30	3

Table 5.2, shows expected frequencies for the values of X in 100 batches, assuming that X has the distribution from part **(a)(iii)**. The table is incomplete.

Table 5.2

x	0	1	2	$\geqslant 3$
Expected frequency	24.01			

(c) Complete the copy of **Table 5.2** in the **Printed Answer Booklet**. Give the expected frequencies to **2** decimal places. **[2]**

Most candidates answered this question correctly.

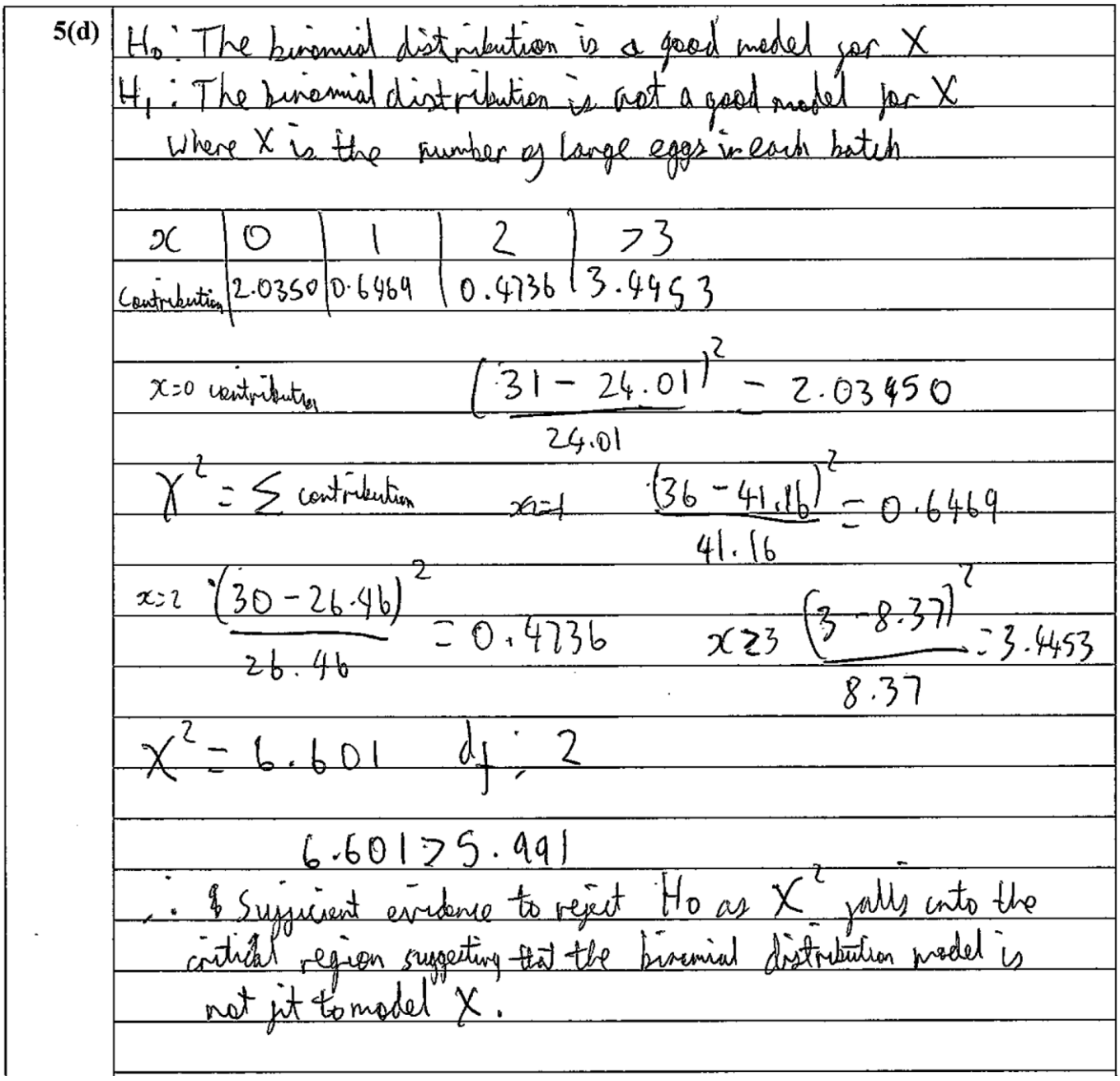
Question 5 (d)

(d) Carry out a χ^2 test, at the 5% significance level, for goodness of fit of the model in part **(a)(iii)**. **[5]**

Candidates often did not refer to the specific model, $\mathrm{B}(4, 0.3)$, in the hypothesis despite that being what the hypothesis test is going on to test. Many candidates chose the incorrect critical value to compare their test statistic to. There has been a clear improvement in candidates conclusions to the hypothesis test compared to previous years, using 'sufficient evidence to suggest that' or 'insufficient evidence to suggest that'.

Candidates would usually be expected to refer to the original context in the conclusion. For example, in this case referring to the eggs. In this question, marks were given for a conclusion that was not written in context.

Exemplar 2



5(d)

H_0: The binomial distribution is a good model for X

H_1: The binomial distribution is not a good model for X

Where X is the number of large eggs in each batch

x	0	1	2	>3
Contribution	2.0350	0.6469	0.4736	3.4453

x=0 contribution $\frac{(31-24.01)^2}{24.01} = 2.03450$

$\chi^2 = \sum$ contribution $x=1 \quad \frac{(36-41.16)^2}{41.16} = 0.6469$

$x=2 \quad \frac{(30-26.46)^2}{26.46} = 0.4736 \qquad x\geq 3 \quad \frac{(3-8.37)^2}{8.37} = 3.4453$

$\chi^2 = 6.601 \quad df: 2$

$6.601 > 5.991$

∴ ∃ Sufficient evidence to reject H_0 as X^2 falls into the critical region suggesting that the binomial distribution model is not fit to model X.

For this exemplar, the candidate was given 3 out of 5 marks. The hypothesis does not refer to the specific model, so they were not given the first mark. They calculated the correct test statistic, so they were given the first M1 mark. They chose the wrong critical value to compare to. Follow-through marks could then be given for the conclusion, which is non-assertive (and the lack of reference to the eggs being allowed in this case), so they were given the final 2 marks.

Question 6 (a)

6 A factory discharges very low concentrations of an inert chemical through an outfall into a nearby body of water.

The concentration of the chemical is measured by sensors at seven known locations at fixed distances from the outfall. Each measurement is known to be affected by various environmental factors. None of these factors are measured or noted.

The manager of the factory needs to assess the relationship between the concentration of the chemical at each sensor and the distance of the sensor from the outfall.

The manager can carry out a hypothesis test **either** for any correlation **or** for any association between the concentration of the chemical at the sensor and the distance of the sensor from the outfall.

(a) Explain why the manager should carry out a test for association rather than one for correlation. **[1]**

Candidates attempted a variety of approaches to answer this question. Successful candidates stated it was not random-on-random or that there might not be an underlying 'bivariate normal population'. Some candidates wrote about factors at sensors but often didn't mention that it could be different effects on each sensor which would affect any correlation.

Question 6 (b)

The concentration of the chemical at each sensor and the distance of each sensor from the outfall, both in arbitrary units, for a particular day after the chemical has been discharged are given in the table below.

Sensor	A	B	C	D	E	F	G
Distance from outfall (arbitrary units)	1.2	6.6	8.9	11.0	15.6	23.2	30.5
Concentration of chemical (arb. units)	621	459	498	610	222	453	420

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

Carry out the appropriate test for association at the 5% significance level. A copy of the table, with spare, empty rows, is given in the **Printed Answer Booklet** to help with any calculations. **[8]**

Most candidates were very good at stating their hypothesis, although many did not state 'in the population' at least once, as was required. When calculating Spearman's rank correlation coefficient, some candidates had incorrectly calculated other test statistics such as the chi-squared value or the PMCC.

It was common to see well-formed, non-assertive conclusions for this hypothesis test.

Question 7 (a)

7 The probability function of the discrete random variable, X, is shown in the table. You are given that α, β and γ are constants.

x	-1	0	1
$\mathrm{P}(X = x)$	α	β	γ

(a) Find $\mathrm{E}(X)$ in terms of α and γ. **[1]**

Most candidates answered this question correctly.

Question 7 (b)

(b) Find $\mathrm{Var}(X)$ in terms of α and γ. **[2]**

This question was generally well-answered. Most candidates expanded the brackets, even though this wasn't necessary.

Question 7 (c)

The random variable W is defined by $W = X_1 + X_2$ where X_1 and X_2 are independent random variables each representing an observation of X.

(c) Determine the probability distribution of W in terms of α, β and γ. Your answer should be entered in the empty table in the **Printed Answer Booklet**. You should not need all of the columns. **[4]**

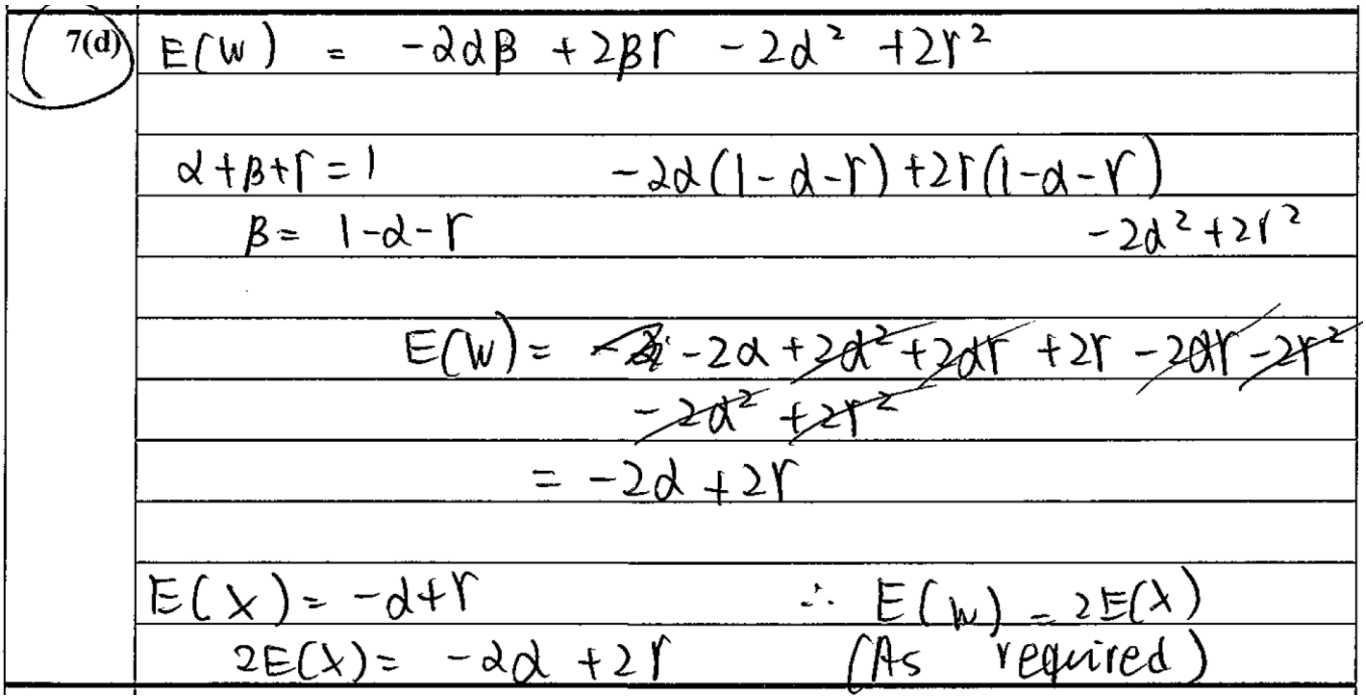
Most candidates found the five possible options of W. The header column was a common place where candidates did not gain marks. When stating the headers, candidates replaced the W with an X which had already been defined within the question so was not correct.

Question 7 (d)

(d) By forming an expression for $\mathrm{E}(W)$ from the table found in **(c)** and then eliminating β, show that $\mathrm{E}(W) = 2\mathrm{E}(X)$. **[4]**

This question was quite well-answered, with most candidates able to take their previous answer from part (c) to find their $\mathrm{E}(W)$. Successful candidates realised that to eliminate β they needed to use the sum of the probabilities, i.e. that $\alpha + \beta + \gamma = 1$.

Exemplar 3



This exemplar shows a clear method for the required result. The cancellation of terms is clear. There are no errors within the working, and there is sufficient working to convincingly demonstrate the required result.

Question 7 (e)

(e) In the case where $\alpha = 0.1$ and $\beta = 0.7$, find $\mathrm{P}(W = 0 | X_1 = 1)$. **[1]**

Successful candidates realised that $W = 0$ and $X_1 = 1$ are not independent events and were able to obtain the answer with ease. Others were able to obtain the answer through longer methods, although mistakes were more common with these methods.

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