Cambridge Assessment International Education

Specimen Paper Answers

Paper 4

Cambridge International AS & A Level Further Mathematics 9231

For examination from 2020

Version 1

Cambridge Pathway

In order to help us develop the highest quality resources, we are undertaking a continuous programme of review; not only to measure the success of our resources but also to highlight areas for improvement and to identify new development needs.

We invite you to complete our survey by visiting the website below. Your comments on the quality and relevance of our resources are very important to us.

www.surveymonkey.co.uk/r/GL6ZNJB

Would you like to become a Cambridge International consultant and help us develop support materials?

Please follow the link below to register your interest.

www.cambridgeinternational.org/cambridge-for/teachers/teacherconsultants/

Contents

Introduction

The main aim of this booklet is to exemplify standards for those teaching Cambridge International AS & A Level Further Mathematics 9231, and to show examples of model answers to the 2020 Specimen Paper 4 Paper 4 assesses the syllabus content for Further Probability & Statistics. We have provided answers for each question in the specimen paper, along with examiner comments explaining where and why marks were awarded. Candidates need to demonstrate the appropriate techniques, as well as applying their knowledge when solving problems.

You will need to use the mark scheme alongside this document. This can be found on the School Support Hub (www.cambridgeinternational.org/support) on the 'Syllabus materials' tab – scroll down to the bottom of the page where the Specimen Paper materials are.

Individual examination questions may involve ideas and methods from more than one section of the syllabus content for that component. The main focus of examination questions will be the AS & A Level Further Mathematics syllabus content. However, candidates may need to make use of prior knowledge and mathematical techniques from previous study, as listed in the introduction to section 3 of the syllabus.

There are five to seven structured questions in Paper 1; candidates must answer **all** questions. Questions are of varied lengths and often contain several parts, labelled (a), (b), (c), which may have sub-parts (i), (ii), (iii), as needed. Some questions might require candidates to sketch graphs or diagrams, or draw accurate graphs.

Candidates are expected to answer directly on the question paper. All working should be shown neatly and clearly in the spaces provided for each question. New questions often start on a fresh page, so more answer space may be provided than is needed. If additional space is required, candidates should use the lined page at the end of the question paper, where the question number or numbers must be clearly shown.

Past exam resources and other teacher support materials are available on the School Support Hub (www.cambridgeinternational.org/support).

Assessment overview

There are three routes for Cambridge International AS & A Level Further Mathematics. Candidates may combine components as shown below.

Route 1 **AS Level only** (Candidates take the AS components in the same series)	**Paper 1** **Further Pure Mathematics 1**	**Paper 2** **Further Pure Mathematics 2**	**Paper 3** **Further Mechanics**	**Paper 4** **Further Probability & Statistics**
Either	✓	Not available for AS Level	✓	
Or	✓			✓

Route 2 **A Level (staged over two years)**	**Paper 1** **Further Pure Mathematics 1**	**Paper 2** **Further Pure Mathematics 2**	**Paper 3** **Further Mechanics**	**Paper 4** **Further Probability & Statistics**
Either Year 1 AS Level	✓		✓	
Year 2 Complete the A Level		✓		✓
Or Year 1 AS Level	✓			✓
Year 2 Complete the A Level		✓	✓	

Route 3 **A Level** (Candidates take the A Level components in the same series)	**Paper 1** **Further Pure Mathematics 1**	**Paper 2** **Further Pure Mathematics 2**	**Paper 3** **Further Mechanics**	**Paper 4** **Further Probability & Statistics**
Year 2 full A Level	✓	✓	✓	✓

Paper 4 – Further Probability & Statistics

- Written examination, 1 hour 30 minutes, 50 marks
- 5 to 7 structured questions based on the Further Probability & Statistics subject content
- Candidates answer all questions
- Externally assessed by Cambridge International
- 40% of the AS Level
- 20% of the A Level

Offered as part of AS Level or A Level.

Assessment objectives

The assessment objectives (AOs) are the same for all papers:

AO1 Knowledge and understanding

- Show understanding of relevant mathematical concepts, terminology and notation
- Recall accurately and use appropriate mathematical manipulative techniques

AO2 Application and communication

- Recognise the appropriate mathematical procedure for a given situation
- Apply appropriate combinations of mathematical skills and techniques in solving problems
- Present relevant mathematical work, and communicate corresponding conclusions, in a clear and logical way

Weightings for assessment objectives

The approximate weightings (± 5%) allocated to each of the AOs are summarised below.

Assessment objectives as an approximate percentage of each component

Assessment objective	Weighting in components %			
	Paper 1	Paper 2	Paper 3	Paper 4
AO1 Knowledge and understanding	45	45	45	45
AO2 Application and communication	55	55	55	55

Assessment objectives as an approximate percentage of each qualification

Assessment objective	Weighting in AS Level %	Weighting in A Level %
AO1 Knowledge and understanding	45	45
AO2 Application and communication	55	55

Question 1

1 **(a)** State briefly the circumstances under which a non-parametric test of significance should be used rather than a parametric test. [1]

When the population cannot be assumed to be normally distributed

The level of pollution in a river was measured at 12 randomly chosen locations. The results, in suitable units, are shown below, where higher values represent greater pollution.

5.62 5.73 6.55 6.81 6.10 5.75 5.87 6.47 5.86 6.26 6.99 5.91

Examiner comment

The single B1 mark is awarded for any reasonable statement of the circumstances, centred on lack of knowledge about the population distribution. For example, '*or where nothing can be assumed about the underlying population distribution*' is also acceptable.

Candidates may perhaps choose to expand on this key point by ruling out large sample sizes, so that the Central Limit Theorem and hence the normal distribution cannot be used, but this is not required for the award of the B1.

The availability of only one mark and the provision of only three lines for the candidate's response serve to reinforce the question's request for brevity.

(b) Use a Wilcoxon signed-rank test to test whether the average pollution level in the river is more than 6.00. Use a 5% significance level. [6]

H_0: median of population = 6.00; H_1: median of population > 6.00 (1)

Deviations and signed ranks:

−0.38	−0.27	0.55	0.81	0.10	−0.25	−0.13	0.47	−0.14	0.26	0.99	−0.09
		10	11	2			9		6	12	
−8	−7				−5	−3		−4			−1

$P = 50$, $Q = 28$, $T = 28$ (2)

$28 > 17$ so accept H_0:

there is insufficient evidence that the average pollution level is more than 6.00. (3)

Examiner comment

B1 is awarded for correctly stating both the null and alternative hypotheses. A reasonable attempt to find the deviations and the resulting signed ranks earns M1, with A1 awarded if all these values are correct. A second A1 is awarded if the Wilcoxon test statistic T is also found correctly.

A reasonable attempt to compare the value found for T with the appropriate critical value in order to reach a conclusion earns M1. A full statement of the conclusion that is consistent with the value found for T and the correct critical value earns a follow-through A1.

(1) There is not a symbol assigned to the median of a population in the *Notation List*, so if a candidate chooses to use a symbol then it should be defined.

(2) The meaning of P, Q and T should be sufficiently clear here without explanation since the question specifies the Wilcoxon signed-rank test and these are the corresponding symbols used in the *List of formulae and statistical tables (MF19)*. Similarly, use of the critical value 17 implies that it has been found from the table with $n = 12$ and one-tailed significance level of 5%. This is, however, about the **minimum** level of working that can reasonably be shown.

(3) It is good practice to end by stating the conclusion explicitly in the context of what the question specifies is to be tested.

Question 2

2 Each of 200 identically biased dice is thrown repeatedly until an even number is obtained. The number of throws needed is recorded and the results are summarised in the following table.

Number of throws	1	2	3	4	5	6	⩾7
Frequency	126	43	22	3	5	1	0

Carry out a goodness of fit test, at the 5% significance level, to test whether Geo(0.6) is a satisfactory model for the data. [7]

H_0: Geo(0·6) fits the data; [H_1: Geo(0·6) does not fit the data] (1)

Expected values:

120 48 19·2 7·68 3·072 1·2288 0·8192 (2)

Combine last three cells to give 5·12 (3)

$X^2 = 0·3 + 0·5208 + 0·4083 + 2·8519 + 0·15125 = 4·23$ (4)

Comparing 4·23 with $\chi_{4,\,0.95}^2 = 9·488$, (5)

9·49 > 4·23 so Geo(0·6) is a satisfactory model for the data (6)

Examiner comment

M1 is awarded for a reasonable attempt to find the expected values, with A1 awarded if they are all correct (or implied by a combined value). A second M1 is awarded if the cells are combined appropriately to ensure all the expected values are not less than 5. A reasonable attempt to find the value of X^2 earns M1, with A1 awarded for a value correct to at least 3SF. Comparing the value found for X^2 with an appropriate critical value in order to reach a conclusion earns M1, and a full statement of the conclusion based on correct values earns A1.

(1) Even though no mark has been allocated for them here, it is good practice to state the hypotheses on which a test is based (or at least the null hypothesis where the alternative hypothesis is obvious, as is the case here). (2) It would be acceptable to round the expected values to 4SF as it is not obvious in advance what effect a rounding error in a particular expected value will have on the accuracy of the value found for X^2. (3) While not necessary, it would be good practice to state why cells are combined, namely to ensure all the expected values are not less than 5. (4) Candidates will differ in the amount of working they choose to write down, but there should be sufficient working to indicate the method being used in case an arithmetical error adversely affects the final value. (5) Similarly it is advisable to indicate that the critical value is found here from the χ^2-distribution table with 4 degrees of freedom and $p = 0.95$. (6) As in the previous question, it is good practice to end by stating the conclusion explicitly in the context of what the question specifies should be tested.

Question 3

3 Employees at a particular company have been working seven hours each day, from 9 am to 4 pm. To try to reduce absence, the company decides to introduce ‘flexi-time’ and allow employees to work their seven hours each day at any time between 7 am and 9 pm. For a random sample of 10 employees, the numbers of hours of absence in the year before and the year after the introduction of flexi-time are given in the following table.

Employee	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*
Before	42	35	96	74	20	5	78	45	146	0
After	34	32	100	72	31	2	61	35	140	0

Test, at the 10% significance level, whether the population mean number of hours of absence has decreased following the introduction of flexi-time, stating any assumption that you make. [8]

Assume the population of differences is normally distributed (1)

$H_0: \mu_{before} - \mu_{after} = 0;\ H_1: \mu_{before} - \mu_{after} > 0$ (2)

Differences d_i: 8 3 −4 2 −11 3 17 10 6 0 (3)

$\bar{d} = 3.4$, $s^2 = \left(\frac{1}{9}\right)\left(648 - \frac{34^2}{10}\right) = 59.16$ (4)

$$t = \frac{\bar{d}}{\left(\frac{s}{\sqrt{10}}\right)} = 1.40$$

Comparing 1.40 with $t_{9,\ 0.9} = 1.383$, (5)

the population mean number of hours has decreased.

Examiner comment

B1 is awarded for any reasonable statement of an assumption that incorporates the key point of normal population distribution. A second B1 is awarded for correctly stating both the null and alternative hypotheses. Attempting to find the differences between the given paired values earns M1, as does using them to find both their mean and a suitable estimate of the population variance. A further M1 is awarded for finding the test statistic *t*, with A1 for a value correct to at least 3SF. Finally B1 is awarded for comparing the value found for *t* with the correct critical value. A follow-through B1 is awarded for a full statement of the conclusion based on the correct critical value and consistent with the value found for *t*.

Although not explicitly stated in the question, the appropriate test here is clearly a paired sample *t*-test. Use instead of another test, for example, a 2-sample *t*-test, will considerably reduce the number of marks which may be earned. (1) An assumption that both the populations are normally distributed would be acceptable. (2) The hypotheses can be framed in terms of the means of the two populations instead of the mean of the population of differences. Any potentially ambiguous notation should be defined, e.g. subscripts *A* and *B* could refer to the values in the first (*A*) and second (*B*) years respectively, or conversely to **A**fter and **B**efore.

(3) The values found for the differences could of course all have opposite signs.

(4) It is good practice to find an **unbiased** estimate of the population variance, but a biased estimate is acceptable provided $\sqrt{9}$ replaces $\sqrt{10}$ in the formula for *t*.

(5) As in the previous question, it is good practice to indicate the origin of the critical value and to state the conclusion explicitly in the context of what the question specifies should be tested.

Question 4

4 The number, x, of a certain type of sea shell was counted at 60 randomly chosen sites, each one metre square, along the coastline in country A. The number, y, of the same type of sea shell was counted at 50 randomly chosen sites, each one metre square, along the coastline in country B. The results are summarised as follows, where $\bar{x}$ and $\bar{y}$ denote the sample means of x and y respectively.

$$\bar{x} = 29.2 \qquad \Sigma(x - \bar{x})^2 = 4341.6 \qquad \bar{y} = 24.4 \qquad \Sigma(y - \bar{y})^2 = 3732.0$$

Find a 95% confidence interval for the difference between the mean number of sea shells, per square metre, on the coastlines in country A and in country B. [7]

$s_x{}^2 = \frac{4341.6}{59} = 73.59$, $s_y{}^2 = \frac{3732.0}{49} = 76.16$ (1)

$s^2 = \frac{s_x{}^2}{60} + \frac{s_y{}^2}{50} = 2.750$ (2) (or s = 1.658)

Confidence interval:

$(29.2 - 24.4) \pm 1.96\, s = 4.8 \pm 3.25$ (3)

Examiner comment

M1 is awarded for a reasonable attempt to estimate either of the population variances (which may be implied). An A1 is awarded for finding both of them correct to at least 4SF. A second M1 is awarded for a reasonable attempt to estimate from them the combined variance, with an A1 awarded for a value correct to at least 4SF. Finally, a third M1 is awarded for a reasonable attempt to find the required confidence interval, using the value found for the combined variance and some tabular z-value. A1 is awarded for the correct two-sided z-value 1.96 and another A1 for the confidence interval (in any appropriate form) correct to at least 3SF.

(1) The meaning of the notation s_x, s_y and s is sufficiently clear here that no explanation is necessary. Other relevant subscripts are equally acceptable, such as X, Y or A, B. Note that $s_x{}^2$, $s_y{}^2$ need not appear explicitly, since they may be incorporated into the expression used to find s^2.

(2) While it is expected that the combined variance will be used, a pooled estimate of the common variance might be accepted provided the population variances are explicitly assumed by the candidate to be equal. This will result in a slight difference in the confidence interval, possibly obscured by rounding.

(3) An equivalent form of the confidence interval is equally acceptable, such as -4·8 ± 3.25, (1.55, 8.05) or (-8.05, -1.55).

Question 5

5 The continuous random variable X has probability density function f given by

$$f(x) = \begin{cases} 0 & x < 0, \\ \frac{6}{5}x & 0 \leqslant x \leqslant 1, \\ \frac{6}{5}x^{-4} & x > 1. \end{cases}$$

(a) Find $P(X > 1)$. [1]

For $0 \leqslant x \leqslant 1$, $F(x) = \frac{3x^2}{5}$

so $P(X > 1) = 1 - F(1) = \frac{2}{5}$

An alternative might be: for $x > 1$

$P(X > 1) = \int_1^\infty \frac{6}{5}x^{-4}\,dx = \left[-\frac{2x^{-3}}{5}\right]_1^\infty = \frac{2}{5}$

Examiner comment

B1 is awarded for the correct result, found by any valid method. Noting, for example, that f(x) is linear over $0 \leqslant x \leqslant 1$, F(1) may be found from the area $\frac{3}{5}$ of the corresponding triangle. Both an exact decimal form 0.4 and a fractional form $\frac{2}{5}$ of the answer are of equally acceptable here.

(b) Find the median value of X. [2]

Median m satisfies $F(m) = \frac{1}{2}$ so $\frac{3m^2}{5} = \frac{1}{2}$, $m = \sqrt{\left(\frac{5}{6}\right)}$

Examiner comment

M1 is awarded for any valid method which would lead to the median value of X, with A1 for the correct value. This can be the exact value $\sqrt{\left(\frac{5}{6}\right)}$ or an approximate value correct to at least 3SF, here 0.913.

(c) Given that E(X) = 1, find the variance of X. [3]

$$E(X^2) = \int_0^1 \left(\frac{6}{5}\right) x^3 dx + \int_1^\infty \left(\frac{6}{5}\right) x^{-2} dx$$

$$= \left[\left(\frac{3}{10}\right)x^4\right]_0^1 + \left[-\left(\frac{6}{5}\right)x^{-1}\right]_1^\infty$$

$$= 0.3 + 1.2 = 1.5$$

$$Var(X) = 1.5 - 1^2 = 0.5$$

Examiner comment

M1 is awarded for a reasonable attempt to find E(X^2) by integration, with A1 for its correct value. A follow-through accuracy mark A1 is awarded for finding a value of the required variance which is consistent with the (possibly incorrect) value found for E(X^2), though the mark is withheld if incorrect working should yield a negative variance.

(d) Find $E\left(\sqrt{X}\right)$. [2]

$$E\left(\sqrt{X}\right) = \int_0^1 \left(\frac{6}{5}\right) x^{\frac{3}{2}} dx + \int_1^\infty \left(\frac{6}{5}\right) x^{\frac{-7}{2}} dx$$

$$= \left[\left(\frac{12}{25}\right)x^{\frac{5}{2}}\right]_0^1 + \left[-\left(\frac{12}{25}\right)x^{\frac{-5}{2}}\right]_1^\infty$$

$$= \frac{24}{25}$$

Examiner comment

M1 is awarded for a reasonable attempt to find $E\left(\sqrt{X}\right)$ by integration, with A1 for its correct value.

Question 6

6 Aisha has a bag containing 3 red balls and 3 white balls. She selects a ball at random, notes its colour and returns it to the bag; the same process is repeated twice more. The number of red balls selected by Aisha is denoted by X.

(a) Find the probability generating function $G_X(t)$ of X. [2]

Probabilities of 0, 1, 2, 3 reds are: $\frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8}$

$$G_X(t) = \frac{1}{8} + \left(\frac{3}{8}\right)t + \left(\frac{3}{8}\right)t^2 + \left(\frac{1}{8}\right)t^3$$

Basant also has a bag containing 3 red balls and 3 white balls. He selects three balls at random, without replacement, from his bag. The number of red balls selected by Basant is denoted by Y.

(b) Find the probability generating function $G_Y(t)$ of Y. [3]

Probabilities of 0, 1, 2, 3 reds are: $\frac{1}{20}, \frac{9}{20}, \frac{9}{20}, \frac{1}{20}$

$$G_Y(t) = \frac{1}{20} + \left(\frac{9}{20}\right)t + \left(\frac{9}{20}\right)t^2 + \left(\frac{1}{20}\right)t^3$$

Examiner comment

B1 is awarded for finding the probabilities correctly, with a follow-through B1 for using the values found for these probabilities (provided they sum to unity) to find $G_X(t)$. In this and the next part of the question, the probabilities can simply be used as the coefficients in $G_X(t)$ without first being written down separately.

Examiner comment

B1 is awarded for finding the probability of either 1 or 2 reds correctly, and a second B1 for finding all four probabilities correctly. A follow-through B1 is then awarded for using the values found for these probabilities (provided they sum to unity) to find $G_Y(t)$.

The random variable Z is the total number of red balls selected by Aisha and Basant.

(c) Find the probability generating function of Z, expressing your answer as a polynomial. [3]

$G_Z(t) = G_X(t) \times G_Y(t)$

$= \left(\frac{1}{160}\right)(1 + 12t + 39t^2 + 56t^3 + 39t^4 + 12t^5 + t^6)$

Examiner comment

B1 is awarded for stating or implying that $G_Z(t) = G_X(t) \times G_Y(t)$, with M1 for a reasonable attempt to expand the product of the two corresponding cubic polynomials and A1 for producing a correct form of the result. While exact values of the coefficients are preferable in the probability generating function, it is acceptable to give inexact decimal forms provided they are correct to at least 3SF.

(d) Use the probability generating function of Z to find E(Z) and Var(Z). [5]

$G_Z(t) = \left(\frac{1}{160}\right)(12 + 78t + 168t^2 + 156t^3 + 60t^4 + 6t^5)$

$E(Z) = \frac{480}{160} = 3$

$G''_Z(t) = \left(\frac{1}{160}\right)(78 + 336t + 468t^2 + 240t^3 + 30t^4)$

$G''_Z(1) = \frac{1152}{160} = 7.2$

$Var(Z) = 7.2 + 3 - 3^2 = 1.2$

Examiner comment

No marks are awarded for simply writing down the relevant formula for E(Z) or Var(Z) since these are given in the *List of formulae and statistical tables (MF19)*. Instead, M1 is awarded for a reasonable attempt to differentiate the polynomial form of $G_Z(t)$ and hence find E(Z) from $G'_Z(1)$, with A1 for a correct value. M1 is then awarded for a reasonable attempt to differentiate the polynomial form already found for $G'_Z(t)$, with another M1 for using $G''_Z(1) + G'_Z(1) - \{G'_Z(1)\}^2$ to find Var(Z) and A1 if this value is correct.

Cambridge Assessment International Education
The Triangle Building, Shaftesbury Road, Cambridge, CB2 8EA, United Kingdom
t: +44 1223 553554
e: info@cambridgeinternational.org   www.cambridgeinternational.org