



GCE

Further Mathematics B (MEI)

Y422/01: Statistics major

Advanced GCE

Mark Scheme for November 2020

OCR (Oxford Cambridge and RSA) is a leading UK awarding body, providing a wide range of qualifications to meet the needs of candidates of all ages and abilities. OCR qualifications include AS/A Levels, Diplomas, GCSEs, Cambridge Nationals, Cambridge Technicals, Functional Skills, Key Skills, Entry Level qualifications, NVQs and vocational qualifications in areas such as IT, business, languages, teaching/training, administration and secretarial skills.

It is also responsible for developing new specifications to meet national requirements and the needs of students and teachers. OCR is a not-for-profit organisation; any surplus made is invested back into the establishment to help towards the development of qualifications and support, which keep pace with the changing needs of today's society.

This mark scheme is published as an aid to teachers and students, to indicate the requirements of the examination. It shows the basis on which marks were awarded by examiners. It does not indicate the details of the discussions which took place at an examiners' meeting before marking commenced.

All examiners are instructed that alternative correct answers and unexpected approaches in candidates' scripts must be given marks that fairly reflect the relevant knowledge and skills demonstrated.

Mark schemes should be read in conjunction with the published question papers and the report on the examination.

© OCR 2020

Y422/01

Mark Scheme

November 2020

Text Instructions

Annotations and abbreviations

Annotation in scoris	Meaning
✓ and ✖	
BOD	Benefit of doubt
FT	Follow through
ISW	Ignore subsequent working
M0, M1	Method mark awarded 0, 1
A0, A1	Accuracy mark awarded 0, 1
B0, B1	Independent mark awarded 0, 1
E	Explanation mark 1
SC	Special case
^	Omission sign
MR	Misread
BP	Blank page
Highlighting	
Other abbreviations in mark scheme	Meaning
E1	Mark for explaining a result or establishing a given result
dep*	Mark dependent on a previous mark, indicated by *. The * may be omitted if only previous M mark.
cao	Correct answer only
oe	Or equivalent
rot	Rounded or truncated
soi	Seen or implied
www	Without wrong working
AG	Answer given
a wrt	Anything which rounds to
BC	By Calculator
DR	This indicates that the instruction In this question you must show detailed reasoning appears in the question.

Subject-specific Marking Instructions for AS Level Mathematics B (MEI)

- a Annotations must be used during your marking. For a response awarded zero (or full) marks a single appropriate annotation (cross, tick, M0 or ^) is sufficient, but not required.

For responses that are not awarded either 0 or full marks, you must make it clear how you have arrived at the mark you have awarded and all responses must have enough annotation for a reviewer to decide if the mark awarded is correct without having to mark it independently.

It is vital that you annotate standardisation scripts fully to show how the marks have been awarded.

Award NR (No Response)

- if there is nothing written at all in the answer space and no attempt elsewhere in the script
- OR if there is a comment which does not in any way relate to the question (e.g. 'can't do', 'don't know')
- OR if there is a mark (e.g. a dash, a question mark, a picture) which isn't an attempt at the question.

Note: Award 0 marks only for an attempt that earns no credit (including copying out the question).

If a candidate uses the answer space for one question to answer another, for example using the space for 8(b) to answer 8(a), then give benefit of doubt unless it is ambiguous for which part it is intended.

- b An element of professional judgement is required in the marking of any written paper. Remember that the mark scheme is designed to assist in marking incorrect solutions. Correct solutions leading to correct answers are awarded full marks but work must not always be judged on the answer alone, and answers that are given in the question, especially, must be validly obtained; key steps in the working must always be looked at and anything unfamiliar must be investigated thoroughly. Correct but unfamiliar or unexpected methods are often signalled by a correct result following an apparently incorrect method. Such work must be carefully assessed. When a candidate adopts a method which does not correspond to the mark scheme, escalate the question to your Team Leader who will decide on a course of action with the Principal Examiner. If you are in any doubt whatsoever you should contact your Team Leader.

- c The following types of marks are available.

M

A suitable method has been selected and *applied* in a manner which shows that the method is essentially understood. Method marks are not usually lost for numerical errors, algebraic slips or errors in units. However, it is not usually sufficient for a candidate just to indicate an intention of using some method or just to quote a formula; the formula or idea must be applied to the specific problem in hand, e.g. by substituting the relevant quantities into the formula. In some cases the nature of the errors allowed for the award of an M mark may be specified.

A method mark may usually be implied by a correct answer unless the question includes the DR statement, the command words “Determine” or “Show that”, or some other indication that the method must be given explicitly.

A

Accuracy mark, awarded for a correct answer or intermediate step correctly obtained. Accuracy marks cannot be given unless the associated Method mark is earned (or implied). Therefore M0 A1 cannot ever be awarded.

B

Mark for a correct result or statement independent of Method marks.

E

A given result is to be established or a result has to be explained. This usually requires more working or explanation than the establishment of an unknown result.

Unless otherwise indicated, marks once gained cannot subsequently be lost, e.g. wrong working following a correct form of answer is ignored. Sometimes this is reinforced in the mark scheme by the abbreviation *isw*. However, this would not apply to a case where a candidate passes through the correct answer as part of a wrong argument.

- d When a part of a question has two or more ‘method’ steps, the M marks are in principle independent unless the scheme specifically says otherwise; and similarly where there are several B marks allocated. (The notation ‘dep*’ is used to indicate that a particular mark is dependent on an earlier, asterisked, mark in the scheme.) Of course, in practice it may happen that when a candidate has once gone wrong in a part of a question, the work from there on is worthless so that no more marks can sensibly be given. On the other hand, when two or more steps are successfully run together by the candidate, the earlier marks are implied and full credit must be given.
- e The abbreviation FT implies that the A or B mark indicated is allowed for work correctly following on from previously incorrect results. Otherwise, A and B marks are given for correct work only – differences in notation are of course permitted. A (accuracy) marks are not given for answers obtained from incorrect working. When A or B marks are awarded for work at an intermediate stage of a solution, there may be various alternatives that are equally acceptable. In such cases, what is acceptable will be detailed in the mark scheme. If this is not the case, please escalate the question to your Team Leader who will decide on a course of action with the Principal Examiner.
- Sometimes the answer to one part of a question is used in a later part of the same question. In this case, A marks will often be ‘follow through’. In such cases you must ensure that you refer back to the answer of the previous part question even if this is not shown within the image zone. You may find it easier to mark follow through questions candidate-by-candidate rather than question-by-question.

- f Unless units are specifically requested, there is no penalty for wrong or missing units as long as the answer is numerically correct and expressed either in SI or in the units of the question. (e.g. lengths will be assumed to be in metres unless in a particular question all the lengths are in km, when this would be assumed to be the unspecified unit.)
We are usually quite flexible about the accuracy to which the final answer is expressed; over-specification is usually only penalised where the scheme explicitly says so.
- When a value is **given** in the paper only accept an answer correct to at least as many significant figures as the given value.
 - When a value is **not given** in the paper accept any answer that agrees with the correct value to **2 s.f.** unless a different level of accuracy has been asked for in the question, or the mark scheme specifies an acceptable range.
NB for Specification A the rubric specifies 3 s.f. as standard, so this statement reads “3 s.f.”
- Follow through should be used so that only one mark in any question is lost for each distinct accuracy error.
Candidates using a value of 9.80, 9.81 or 10 for g should usually be penalised for any final accuracy marks which do not agree to the value found with 9.8 which is given in the rubric.
- g Rules for replaced work and multiple attempts:
- If one attempt is clearly indicated as the one to mark, or only one is left uncrossed out, then mark that attempt and ignore the others.
 - If more than one attempt is left not crossed out, then mark the last attempt unless it only repeats part of the first attempt or is substantially less complete.
 - if a candidate crosses out all of their attempts, the assessor should attempt to mark the crossed out answer(s) as above and award marks appropriately.
- h For a genuine misreading (of numbers or symbols) which is such that the object and the difficulty of the question remain unaltered, mark according to the scheme but following through from the candidate's data. A penalty is then applied; 1 mark is generally appropriate, though this may differ for some units. This is achieved by withholding one A or B mark in the question. Marks designated as cao may be awarded as long as there are no other errors. If a candidate corrects the misread in a later part, do not continue to follow through. E marks are lost unless, by chance, the given results are established by equivalent working. Note that a miscopy of the candidate's own working is not a misread but an accuracy error.
- i If a calculator is used, some answers may be obtained with little or no working visible. Allow full marks for correct answers provided that there is nothing in the wording of the question specifying that analytical methods are required such as the bold “In this question you must show detailed reasoning”, or the command words “Show” and “Determine. Where an answer is wrong but there is some evidence of method, allow appropriate method marks. Wrong answers with no supporting method score zero. If in doubt, consult your Team Leader.
- j If in any case the scheme operates with considerable unfairness consult your Team Leader.

Y422/01

Mark Scheme

November 2020

Question		Answer	Marks	AOs	Guidance
1	(a)	$P(\text{all 4}) = \frac{4}{10} \times \frac{3}{9} \times \frac{2}{8} \times \frac{1}{7} = \frac{1}{210}$	M1 A1 [2]	1.1 1.1	AG
1	(b)	$E(X) = 1.6$ $\text{Var}(X) = 0.64$	B1 B1 [2]	1.1a 1.1	BC BC
1	(c)	Loss $L = 1.00 - 0.4X$ $E(L) = 0.36$ (36 pence) $\text{SD}(L) = 0.4 \times 0.8 = 0.32$ (32 pence)	B1 M1 A1 [3]	3.3 1.1 1.1	
1	(d)	Expected amount = $1.00 - (0.25E(X) + 100 \times \frac{1}{210})$ 12.4 pence loss	M1 A1 [2]	1.1 1.1	Allow M1 if use original distribution but miss out $1 \times 1/210$ term (12.3809...) Allow answer rounded to 12 (pence) provided there is working
2	(a) (i)	People with the antigen occur randomly and occurrences are independent with constant probability 0.00025. The number of people with the antigen out of 1200 is being counted, so a binomial distribution is appropriate. Because $n (= 1200)$ is large and $p (= 0.00025)$ is small a Poisson distribution is also appropriate	E1 E1 E1 [3]	2.4 2.4 2.4	For partial explanation of binomial For full explanation For explanation of Poisson.
2	(a) (ii)	Binomial (1200, 0.00025) or Poisson (0.3) $P(X = 3) = 0.0033$ $P(X > 3) = 0.0003$	M1 A1 A1 [3]	3.3 1.1 1.1	NB both distributions give same answer to 4 dp
2	(b)	$P(\leq 5000 \text{ needed}) = 1 - P(0 \text{ or } 1 \text{ in } 5000 \text{ have antigen})$ Use B(5000, 0.00025) or Po(1.25) $= 1 - 0.6446 = 0.3554$	M1 B1 A1 [3]	3.3 3.4 1.1	

Y422/01

Mark Scheme

November 2020

Question		Answer	Marks	AOs	Guidance
3	(a)	Mean of 2 large bags $\sim N(201.3, 1.7^2/2)$	M1	3.3	For Normal and mean Allow M1 for total $N(402.6, 5.78)$ even if not used For correct variance
		i.e. $N(201.3, 1.445)$ $P(\text{Mean} > 200) = 0.8603$	A1 B1 [3]	1.1 3.4	
3	(b)	Need distribution of 8 small – 2 medium – 1 large mean = $8 \times 51.5 - 2 \times 100.7 - 1 \times 201.3$ variance = $8 \times 1.1^2 + 2 \times 1.6^2 + 1 \times 1.7^2$ so distribution is $N(9.3, 17.69)$ $P(> 0) = 0.9865$	M1	3.1b	For distribution required. so eg correct mean or variance Method for mean Method for variance Correct distribution
			M1	3.3	
			M1	1.1	
			A1	1.1	
			A1	3.4	
[5]					
4	(a)	Because the sample size is large so by the central limit theorem the sample mean is approximately Normally distributed	E1	2.4	Must mention sample mean for second mark
			E1 [2]	2.4	
4	(b)	$-0.0286 < \mu < 0.2632$	B1 [1]	1.1	
4	(c)	This is $\sigma/\sqrt{60}$	B1 [1]	3.4	
4	(d)	The confidence interval contains zero Which suggests that the two measurements broadly agree.	E1	3.4	Allow valid alternative answers such as 'Although the interval contains zero almost all of the interval is above zero, which suggests that on average the digital gauge may be reading lower than the traditional gauge.'
			E1 [2]	2.4	

Y422/01

Mark Scheme

November 2020

Question		Answer	Marks	AOs	Guidance
5	(a)	If all did the same test first, the experience gained in the first test might affect their performance in the second test	E1 E1 [2]	3.4 2.4	Allow 1 mark for 'To avoid bias' or 'because one test might affect the other test' oe
5	(b)	$\bar{x} = 12.375$ $\bar{y} = 11.79375$ $b = \frac{S_{xy}}{S_{xx}} = \frac{2554.87 - (198.0 \times 188.7 / 16)}{2936.92 - 198.0^2 / 16} = \frac{219.7075}{486.67}$ $= 0.4515$ For correct line (y on x) hence regression line equation is: $y - \bar{y} = b(x - \bar{x})$ $\Rightarrow y - 11.79375 = 0.4515(x - 12.375)$ $\Rightarrow y = 0.4515x + 6.207$	M1 A1 B1 M1 A1 [5]	1.1a 1.1 1.1 3.3 1.1	For attempt at gradient (b) For 0.4515 cao For equation of line FT their <i>b</i>
5	(c)	Prediction for 12 is 11.6 Prediction for 25 is 17.5	B1 B1 [2]	3.4 1.1	FT Allow B1B0 if answers given to more than 2 dp FT
5	(d)	Because the points do not lie very close to the line, the first prediction is only moderately reliable. The second prediction is rather less reliable because in addition it is extrapolation.	E1 E1 [2]	2.2a 2.4	Allow 1 mark for either not very close to line and so not very reliable or for second value is extrapolation so unreliable.
5	(e)	Coordinates (18, 1.9) The expert should check whether this data item is genuine and if not then remove it from the analysis	B1 E1 [2]	1.1 3.5c	Allow y-coordinate between 1.7 and 2

Y422/01

Mark Scheme

November 2020

Question		Answer	Marks	AOs	Guidance
6	(a)	$H_0: \rho = 0$, $H_1: \rho \neq 0$ (two-tailed test)	B1	3.3	For both hypotheses
		where ρ is the population correlation coefficient (between wind-speed and nitrogen dioxide level)	B1	2.5	For defining ρ
		For $n = 10$, 5% critical value = 0.2144	B1	3.4	For critical value
		Since $0.3231 > 0.2144$ the result is significant. There is sufficient evidence to reject H_0	M1	1.1	For comparison leading to a conclusion
		There is sufficient evidence at the 10% level to suggest that there is correlation between wind-speed and nitrogen dioxide level	A1 [5]	2.2b	FT for conclusion in words
6	(b)	The population from which the data are drawn must be bivariate Normally distributed.	B1 [1]	2.5	Do not allow 'Data must be bivariate Normally distributed' but condone 'Data must be from a bivariate Normal distribution'
6	(c)	For very large samples, the sample correlation coefficient is very often significant. However as in this case those significant correlations which are very close to zero are of very little real use.	E1 E1 [2]	3.5a 3.5a	oe eg 'large sample size means lower values of r can be significant'
6	(d)	Although four out of five of the correlation coefficients are statistically significant, very little use can be made of most of them as the effect sizes are small. However the effect size of 0.25 between wind-speed and NO_2 suggests that the windier it is, the higher the level of NO_2 pollution.	E1 E1 [2]	2.2b 2.2b	

Y422/01

Mark Scheme

November 2020

Question		Answer	Marks	AOs	Guidance
7	(a)	The population from which the sample is drawn must be Normally distributed	E1 [1]	1.2	Do not allow 'The sample is Normally distributed'
7	(b)	Because the Normal probability plot appears to be linear.	E1 [1]	3.5a	Do not allow 'linear correlation'
7	(c)	<p>Sample mean = 285.2 Sample SD = 15.30</p> <p>Confidence interval is given by</p> $285.2 \pm 2.571 \times \frac{15.30}{\sqrt{6}}$ <p>269.1 < μ < 301.3 or 285.2 \pm 16.1</p>	<p>B1</p> <p>B1</p> <p>M1</p> <p>M1</p> <p>A1</p> <p>M1</p> <p>A1</p> <p>[7]</p>	<p>1.1</p> <p>1.1</p> <p>3.4</p> <p>1.1a</p> <p>1.1</p> <p>1.1</p> <p>1.1</p> <p>1.1</p>	<p>DR</p> <p>BC Allow 285</p> <p>BC</p> <p>For general form</p> <p>For 5 degrees of freedom</p> <p>For 2.571</p> <p>For $\frac{15.30}{\sqrt{6}}$</p> <p>OE</p> <p>Condone 269 < μ < 301 provided working seen, but must be from 285.2, not from 285</p>

Y422/01

Mark Scheme

November 2020

Question	Answer	Marks	AOs	Guidance	
8	<p>Est. population mean $= \frac{51.92}{40} = 1.298$</p> <p>Est. population variance $= \frac{1}{39} \left(70.57 - \frac{51.92^2}{40} \right)$ $= 0.0815$</p> <p>Test is based on a Normal distribution $H_0: \mu = 1.25$, $H_1: \mu \neq 1.25$,</p> <p>where μ is the population mean (electricity consumption)</p> <p>Test statistic is $\frac{1.298 - 1.25}{\sqrt{0.0815/40}}$ $= 1.064$</p> <p>Critical value (2-tailed) at 5% level is 1.96 $1.064 < 1.96$ so not significant (do not reject H_0)</p> <p>Insufficient evidence to suggest that the website claim is not true</p>	<p>B1</p> <p>M1</p> <p>A1</p> <p>B1</p> <p>B1</p> <p>B1</p> <p>M1</p> <p>B1</p> <p>M1</p> <p>A1</p> <p>[10]</p>	<p>1.1</p> <p>1.1</p> <p>1.1</p> <p>3.4</p> <p>1.1a</p> <p>1.2</p> <p>3.3</p> <p>1.1</p> <p>2.2b</p> <p>3.5a</p>	<p>DR</p> <p>$= \frac{1}{39} \times 3.1778$</p> <p>or est. pop sd = 0.2854 soi</p> <p>Both correct; if stated in words only, must include 'population'</p> <p>For definition of μ in context</p> <p>Allow wrong mean or sd here</p> <p>FT wrong (sensible) c.v. and test statistic if calculation is of right form for M1 but not for A1</p>	<p>Alternative: calculate $P(X > 1.298)$ using $N(1.25, 0.0815/40)$ $= 0.856$ Use of 0.975 (not 0.95)</p>

Y422/01

Mark Scheme

November 2020

Question		Answer	Marks	AOs	Guidance
9	(a)	$p = 1.72/10$ $= 0.172$	M1 A1 [2]	3.3 1.1	
9	(b)	Cell C2 = 0.1515 Cell D2 = 22.7193 Cell E2 = $\frac{(22.7193 - 39)^2}{22.7193}$ $= 11.6668$	B1 B1 M1 A1 [4]	3.4 1.1 1.1a 1.1	Allow 22.725 from 0.1515 For correct form Allow 11.6557 from 22.725
9	(c)	Because otherwise at least one expected frequency would be less than 5 so too small for the test to be valid.	E1 [1]	3.5b	For 'less than 5 so invalid'
9	(d)	H ₀ : the binomial model is appropriate H ₁ : the binomial model is not appropriate $X^2 = 23.32$ Refer to χ_3^2 Critical value at 1% level = 11.34 $23.32 > 11.34$ Result is significant There is insufficient evidence to suggest that the binomial model is a good fit.	B1 B1 B1 B1 M1 A1 [6]	2.4 1.1 3.4 1.1 1.1 2.2b	For both. Ignore any reference to value of binomial parameter. For degrees of freedom = 3 soi For comparison with critical value Conclusion in context
9	(e)	The contribution of 11.67 suggests that far more trays have no rotten peaches than would be expected if a binomial model were appropriate. The contribution of 6.22 suggests that far more trays have four or more rotten peaches than would be expected if a binomial model were appropriate. The other three smaller contributions suggest that the numbers of trays with 1, 2 or 3 rotten peaches are very roughly as expected if a binomial model were appropriate.	E1 E1 E1 [3]	3.5a 3.5a 3.5a	Allow answers suggesting less than expected for 2 trays (or similar valid answers).

Y422/01

Mark Scheme

November 2020

Question		Answer	Marks	AOs	Guidance
10	(a)	Estimate of $P(X - 2Y > 0)$ is 0.5 Estimate of $P(X - 2Y > 1)$ is 0.45	B1 B1 [2]	1.1 1.1	
10	(b)	By using more rows in the spreadsheet	E1 [1]	2.4	Condone 'Run more simulations', 'take more samples' or 'increasing the number of values sampled' o.e.
10	(c)	$E(W) = 0$ $\text{Var}(X) = 20 \times 0.3 \times 0.7$ $= 4.2$ $\text{Var}(Y) = 3$ $\text{Var}(X - 2Y) = 4.2 + 2^2 \times 3$ $= 16.2$ Distribution is $N\left(0, \frac{16.2}{50}\right)$ $P(W > 1) = P(\text{Normal} > 1.01) = 0.0380$	B1 M1 A1 B1 M1 A1 M1 M1 A1 [9]	1.1 3.3 1.1 1.1 3.4 1.1 3.3 1.1 1.1	For Normal For parameters (Omitting cc gives 0.0395)
11	(a)	$F(4) = 1$ so $k(8 \times 4^2 - 4^3 - 24) = 1$ So $40k = 1$ so $k = \frac{1}{40}$	B1 [1]	2.1	AG
11	(b)	$P(2.5 \leq T \leq 3.5) = F(3.5) - F(2.5)$ $= \frac{1}{40}(8 \times 3.5^2 - 3.5^3 - 24) - \frac{1}{40}(8 \times 2.5^2 - 2.5^3 - 24)$ $= 0.51875$	M1 A1 [2]	1.1a 1.1	
11	(c)	$0.025(8m^2 - m^3 - 24) = 0.5$ $(8m^2 - m^3 - 24) = 20$ $m^3 - 8m^2 + 44 = 0$	M1 A1 [2]	2.1 1.1	Condoned additional substitution of 2 in equation since this gives zero AG
11	(d)	$F(2.945) = 0.496$ $F(2.955) = 0.501$ So median is 2.95 to 2 dp	B1 E1 [2]	3.4 1.1	For either OR Calculation of $m^3 - 8m^2 + 44$ For 2.945 (= 0.158) For 2.955 (= -0.053) So change of sign

Y422/01

Mark Scheme

November 2020

Question		Answer	Marks	AOs	Guidance
11	(e)	$f(t) = \frac{1}{40}(16t - 3t^2)$ $E(T) = \int_2^4 0.025t(16t - 3t^2)dt$ $= 2.967$ $E(T^2) = \int_2^4 0.025t^2(16t - 3t^2)dt$ $= 9.12$ $\text{Var}(T) = 9.12 - 2.967^2 = 0.3189$ $\text{SD} = 0.5647$ $P(\mu - \sigma < T < \mu + \sigma) = P(2.402 < T < 3.531).$ $= F(3.531) - F(2.402) = 0.586$	B1 M1 A1 M1 A1 M1 A1 [7]	3.1a 1.1 1.1 1.1 1.1 3.3 1.1	BC BC
11	f		B1 B1 [2]	1.1 1.1	For main part of graph For axes and for part where $f(t) = 0$ Numbers not required on $F(t)$ axis but 2 and 4 required on t -axis
11	g	The probability will be less than the Normal probability because the graph of the Normal distribution is more peaked and only the central section of the Normal curve is involved	E1 E1 [2]	2.2a 2.2a	For partial explanation. For full explanation Must clearly state <u>which</u> is greater to get any marks at all.

OCR (Oxford Cambridge and RSA Examinations)
The Triangle Building
Shaftesbury Road
Cambridge
CB2 8EA

OCR Customer Contact Centre

Education and Learning

Telephone: 01223 553998

Facsimile: 01223 552627

Email: general.qualifications@ocr.org.uk

www.ocr.org.uk

For staff training purposes and as part of our quality assurance programme your call may be recorded or monitored