

Mark Scheme

Summer 2023

Pearson Edexcel GCE

In Mathematics (9MA0)

Paper 31 Statistics

Edexcel and BTEC Qualifications

Edexcel and BTEC qualifications are awarded by Pearson, the UK's largest awarding body. We provide a wide range of qualifications including academic, vocational, occupational and specific programmes for employers. For further information visit our qualifications websites at www.edexcel.com or www.btec.co.uk. Alternatively, you can get in touch with us using the details on our contact us page at www.edexcel.com/contactus.

Pearson: helping people progress, everywhere

Pearson aspires to be the world's leading learning company. Our aim is to help everyone progress in their lives through education. We believe in every kind of learning, for all kinds of people, wherever they are in the world. We've been involved in education for over 150 years, and by working across 70 countries, in 100 languages, we have built an international reputation for our commitment to high standards and raising achievement through innovation in education. Find out more about how we can help you and your students at: www.pearson.com/uk

Summer 2023
Publications Code 9MA0_31_2306_MS
All the material in this publication is copyright
© Pearson Education Ltd 2023

General Marking Guidance

- All candidates must receive the same treatment. Examiners must mark the first candidate in exactly the same way as they mark the last.
- Mark schemes should be applied positively. Candidates must be rewarded for what they have shown they can do rather than penalised for omissions.
- Examiners should mark according to the mark scheme not according to their perception of where the grade boundaries may lie.
- There is no ceiling on achievement. All marks on the mark scheme should be used appropriately.
- All the marks on the mark scheme are designed to be awarded. Examiners should always award full marks if deserved, i.e. if the answer matches the mark scheme. Examiners should also be prepared to award zero marks if the candidate's response is not worthy of credit according to the mark scheme.
- Where some judgement is required, mark schemes will provide the principles by which marks will be awarded and exemplification may be limited.
- When examiners are in doubt regarding the application of the mark scheme to a candidate's response, the team leader must be consulted.
- Crossed out work should be marked UNLESS the candidate has replaced it with an alternative response.

EDEXCEL GCE MATHEMATICS

General Instructions for Marking

1. The total number of marks for the paper is 50.

2. The Edexcel Mathematics mark schemes use the following types of marks:

- **M** marks: method marks are awarded for 'knowing a method and attempting to apply it', unless otherwise indicated.
- **A** marks: Accuracy marks can only be awarded if the relevant method (M) marks have been earned.
- **B** marks are unconditional accuracy marks (independent of M marks)
- Marks should not be subdivided.

3. Abbreviations

These are some of the traditional marking abbreviations that will appear in the mark schemes.

- bod – benefit of doubt
- ft – follow through
- the symbol $\surd$ will be used for correct ft
- cao – correct answer only
- cso - correct solution only. There must be no errors in this part of the question to obtain this mark
- isw – ignore subsequent working
- awrt – answers which round to
- SC: special case
- oe – or equivalent (and appropriate)
- dep – dependent
- indep – independent
- dp decimal places
- sf significant figures
- $*$ The answer is printed on the paper
- $\sqsubset$ The second mark is dependent on gaining the first mark

4. For misreading which does not alter the character of a question or materially simplify it, deduct two from any A or B marks gained, in that part of the question affected.

5. Where a candidate has made multiple responses and indicates which response they wish to submit, examiners should mark this response.

   If there are several attempts at a question which have not been crossed out, examiners should mark the final answer which is the answer that is the most complete.

6. Ignore wrong working or incorrect statements following a correct answer.

7. Mark schemes will firstly show the solution judged to be the most common response expected from candidates. Where appropriate, alternatives answers are provided in the notes. If examiners are not sure if an answer is acceptable, they will check the mark scheme to see if an alternative answer is given for the method used.

General Principles for Mechanics Marking

(But note that specific mark schemes may sometimes override these general principles)

- Rules for M marks: correct no. of terms; dimensionally correct; all terms that need resolving (i.e. multiplied by cos or sin) are resolved.
- Omission or extra g in a resolution is an accuracy error not method error.
- Omission of mass from a resolution is a method error.
- Omission of a length from a moments equation is a method error.
- Omission of units or incorrect units is not (usually) counted as an accuracy error.
- DM indicates a dependent method mark i.e. one that can only be awarded if a previous specified method mark has been awarded.
- Any numerical answer which comes from use of g = 9.8 should be given to 2 or 3 SF.
- Use of g = 9.81 should be penalised once per (complete) question.
  N.B. Over-accuracy or under-accuracy of correct answers should only be penalised *once* per complete question. However, premature approximation should be penalised every time it occurs.
- Marks must be entered in the same order as they appear on the mark scheme.
- In all cases, if the candidate clearly labels their working under a particular part of a question i.e. (a) or (b) or (c),……then that working can only score marks for that part of the question.
- Accept column vectors in all cases.
- Misreads – if a misread does not alter the character of a question or materially simplify it, deduct two from any A or B marks gained, bearing in mind that after a misread, the subsequent A marks affected are treated as A ft
- Mechanics Abbreviations

  M(A) Taking moments about A.

  N2L Newton's Second Law (Equation of Motion)

  NEL Newton's Experimental Law (Newton's Law of Impact)

  HL Hooke's Law

  SHM Simple harmonic motion

  PCLM Principle of conservation of linear momentum

  RHS, LHS Right hand side, left hand side

Qu 1	Scheme	Marks	AO
(a)	[0.13 + 0.25 =] **0.38**	B1 **(1)**	1.1b
(b)	Independence implies: e.g. $\left[\mathrm{P}(B\cap C)=\mathrm{P}(B)\times\mathrm{P}(C) \Rightarrow\right]$ $0.3=(0.3+0.05+0.25)\times(0.3+p)$	M1	1.1b
	So p = **0.2**	A1	1.1b
	[Sum of probabilities = 1 gives] q = **0.07**	B1ft **(3)**	1.1b
(c)	$[\mathrm{P}(A\mid B')=]\dfrac{\mathrm{P}(A\cap B')}{\mathrm{P}(B')}$ or $\dfrac{0.13}{(1-0.6)}$ or $\dfrac{0.13}{(0.13+"0.2"+"0.07")}$	M1	1.1b
	$=\dfrac{13}{40}$ or **0.325**	A1 **(2)**	1.1b
		(6 marks)	

Notes

(a) B1 for 0.38 (or exact equivalent)

If answers are given on Venn Diagram and in the script then the script takes precedence.

(b) M1 for a correct equation in p or P(C) only.
May be implied by an answer of p = 0.2 **provided** this does not come from incorrect working.
Condone missing brackets if they get 0.2
Other rules for independence will give simple rearrangements of this equation.

Beware If p = 0.2 comes from incorrect working, we've seen $p=\frac{0.6}{0.3}=0.2$, score M0A0

A1 for p = 0.2 (or exact equivalent)
B1ft for q = 0.07 (or exact equivalent) ft their p i.e. q = 0.27 – "0.2" where 0 „ p „ 0.27

(c) M1 for a correct ratio of probability expressions or a correct ratio of probabilities ft their values of p and q (provided both probabilities) or letters p and q
A1 for 0.325 or exact equivalent. Correct answer only will score 2/2
NB on epen this is labelled M1 but treat it as A1

Qu 2	Scheme	Marks	AO
(a)	Comment in context about either **independence** or **random** packing e.g. "prizes must be placed in packets at random/independently of each other" **or** about **constant probability** e.g. "the probability of a packet containing a prize is constant/ the same/fixed"	B1 **(1)**	3.5b
(b)(i)	[P($T = 6$) =] 0.17273… awrt **0.173**	B1	1.1b
(ii)	[P($T < 3$) = P($T \leq 2$) =] 0.061587… awrt **0.0616**	B1 **(2)**	1.1b
(c)	[K= no. of boxes with fewer than 3 packets containing a prize] $K \sim$ B(5, "0.0616") P($K = 2$) = 0.031344… in the range **[0.0313~0.0314]**	M1 A1 **(2)**	1.1b 1.1b
(d)	$H_0 : p = \frac{1}{7}$ $H_1 : p < \frac{1}{7}$ [X = no of packets containing a prize] X~B(110, $\frac{1}{7}$) [P($X \leq 9$)] = 0.038292… [Significant result or reject H_0] E.g. there is evidence to support Kamil's claim	B1 M1 A1 A1 **(4)**	2.5 3.3 3.4 2.2b
		(9 marks)	

	Notes
(a)	B1 **May use idea of independent events:** a suitable reason, **in context**, covering idea of random packing or packets filled independently. Should mention key words/ideas of: prizes in packets **or** packets in boxes **May use idea of constant probability**. Must see key words underlined in scheme. Idea of probability with "independence" or "not affected by other packets" is B0 B0 for: **Idea of only 2 cases**. E.g. Packet contains a prize or not or **Idea of a fixed number of trials**. E.g. Need a fixed number of packets in each box
(b)(i)	B1 for awrt 0.173
(ii)	B1 for awrt 0.0616
(c)	M1 for sight of B(5, "0.0616") or ${}^5C_2 ("0.0616")^2 (1-"0.0616")^3$ ft their answer to (b)(ii). A1 for an answer in the range [0.0313 to 0.0314] Use of 0.0616 gives 0.031356..ans only 2/2
(d)	B1 for both hypotheses correct in terms of p or π M1 for selecting an appropriate model, may be implied by 1st A1 or P($X = 9$) = 0.0199(2…) 1st A1 for 0.038 or better or allow 0.04 with sight of P($X \leq 9$)
ALT	**Critical Region.** Allow CR of $X \leq 9$ (or $X < 10$) provided a supporting probability is seen e.g. A1 for correct CR plus P($X \leq 10$) = 0.0718… (accept 2sf or 1sf if prob statement seen) 2nd A1 (dep on 1st A1 but indep of hyp's) for a suitable conclusion in context that suggests support for (Kamil's) claim **or** states that there is evidence that proportion /probability/chance of packets containing a prize is less than $\frac{1}{7}$ Do not award 2nd A1 for contradictory statements e.g. "not significant" so "supports claim"
Normal	Sight of $N\left(\frac{110}{7}, \frac{660}{49}\right)$ or awrt 13.5 or probability of 0.045(20..) or 0.033(66..) scores M1

Qu 3	Scheme	Marks	AO
(a)	Need to replace tr with a numerical value	M1	1.2
	Value of tr is between 0 and 0.05 suggest using e.g 0.025 , 0 <u>or</u> value „ 0.05	A1	1.1b
		(2)	
(b)(i)	$\left[\bar{x} = \frac{389.3 \sim 390.8}{184}\right] = 2.119\ldots$ awrt **<u>2.12</u>** allow $\frac{195}{92}$ or $2\frac{11}{92}$	B1	1.1b
(ii)	$\left[\sigma =\right]\sqrt{\frac{(\text{awrt})4336}{184} - "\bar{x}^2"}$ <u>or</u> allow $\left[\sigma^2 =\right]\frac{(\text{awrt})4336}{184} - "\bar{x}^2"$ <u>or</u> awrt 19.1	M1	1.1b
	$= 4.367\ldots$ awrt **<u>4.37</u>**	A1	1.1b
		(3)	
(c)(i)	Only covers May~Oct (so not a suitable sample)	B1	1.1b
(ii)	e.g. Winter months are <u>missing</u> when we'd expect <u>more rain</u> so expect estimate in (b)(i) to be an <u>underestimate</u> (oe)	B1	2.4
		(2)	
		(7 marks)	

Notes

(a) M1 for recognising that tr must be replaced (oe) with a numerical value
The following examples would score M0: The tr values are worth 0 so ignore (not replacing) <u>or</u> must remove outliers <u>or</u> fill gaps in table <u>or</u> make widths the same <u>or</u> need to find mid-points
A1 for using a suitable value: e.g. 0.025 (or allow 0) i.e. any value in [0, 0.05]
(these give $\sum x = 390$ (3sf), use of 0.05 gives 390.8, use of 0 gives 389.3 allow in (b)(i))

(b)(i) B1 for awrt 2.12 <u>or</u> allow simplified fraction or mixed number. B0 for $\frac{390}{184}$

(ii) M1 for a correct expression for standard deviation **or** variance. Allow $\sum x^2$ = awrt 4336
Ignore their label σ or σ^2 Can ft their mean

A1 for awrt 4.37 [Use of *s* gives 4.3791… so for correct use seen allow awrt 4.38]

SC **Using *n* = 155** Allow M1 for expression $[\sigma =]\sqrt{\frac{(\text{awrt})4336}{155} - "\bar{x}^2"} = \sqrt{21.64...}$ or 4.65…

Part (c) can effectively be marked together.

(c)(i) B1 for a comment mentioning that data is just from May~Oct (so not representative of the whole year).
Just saying "only 184 days so not representative" is B0, **must mention May ~ Oct**

(ii) B1 for comment that <u>missing</u>/<u>winter</u> months expected to have more rain (oe) **and** "underestimate"(oe)

We are looking for all 3 of these ideas here:
1. A statement or implication that missing data is from winter or different months.
2. A suggestion about the rainfall in these months (probably more rain).
3. A statement about the impact on the estimate in (b)(i) <u>equivalent</u> to saying it would be an underestimate or the (actual) mean will be higher.

SC If you see "Leeming or N or NE has <u>less</u> rain in winter months" – please send to review

Qu 4	Scheme	Marks	AO
(a)	[Let N = height from region A; P(N > 180) =] 0.24937… awrt **0.249**	B1 **(1)**	1.1b
(b)	$H_0 : \mu = 175.4$ $H_1 : \mu \neq 175.4$	B1	2.5
	[S = height from region B] $\bar{S} \sim N\left(175.4, \frac{6.8^2}{52}\right)$ Allow σ^2 = awrt 0.889	M1	3.3
	[P($\bar{S}$ > 177.2)] = 0.02814…	A1	3.4
	[0.028… > 0.025, Not sig, do not reject H_0] Insufficient evidence to support student's claim	A1 **(4)**	2.2b
(c)	[p-value = $2 \times 0.02814...$ =] 0.05628… in range **0.056~0.06** or **5.6(%)~6(%)**	B1ft **(1)**	1.2
		(6 marks)	

	Notes
(a)	B1 for awrt 0.249
(b)	B1 for both hypotheses correct in terms of μ (**See below for one-tail test**)
	M1 for selecting the correct model, may be implied by standardisation using correct values or may be implied by a correct value in 1st A1 e.g.(Prob =) 0.028 or awrt 0.972, (Z =) 1.9(08..) (CV=) 177.25 Condone use of S (or any other letter) instead of $\bar{S}$ Condone use of $\bar{S} \sim N\left(177.2, \frac{6.8^2}{52}\right)$ but **this will lose 2nd A mark**
	1st A1 for probability of awrt 0.028 (allow 0.03 if P($\bar{S}$ > 177.2) is seen) Condone 1 – 0.02814 … = 0.9718…(awrt 0..972) **only if clearly compared with 0.975**
ALT	Allow Z = 1.9(088…) and comparison with 1.96 (or better: calc gives 1.95996…) or CR of $\left[\bar{S}\right]$… 177.248…(awrt 177.25) Allow $\left[\bar{S}\right]$ >/≥177.248…(awrt 177.25) Implied by diagram or correct interpretation of inequality with their CV (Ignore any attempt at a lower CR for $\bar{S}$)
	2nd A1 (dep on 1st A1 and use of correct model. Use of N(177.2, …) scores A0) for a conclusion using context: e.g. does not support student's claim or e.g. insufficient evidence of a difference in heights Do not allow 2nd A mark for contradictory statements e.g. "significant" so "no support for claim"
(c)	B1ft for answer in range 0.056~0.06 or 5.6%~6% (Ranges are inclusive, condone missing %) (can ft their probability, provided < 0.5, from part (b) but not 0.025 leading to 5%)
NB	**One-tail test [Max of 3/5 for (b) and (c)]** In (b) B0 (hypotheses) M1(model as above) 1st A1[for probability or Z compared with 1.6449 or CR $\left[\bar{S}\right]$… or > 176.95… (awrt 177)] 2nd A1 for conclusion in context that supports claim **or** "heights of men from B is different from/greater than from A" In (c) B0

Qu 5	Scheme	Marks	AO
(a)	$P(S \cap \{X = 50\}) = P(S \cap \{X = 80\})$[= a constant, V] $\Rightarrow b \times \frac{k}{50} = c \times \frac{k}{80}$	M1	3.1a
	May see: $\frac{k}{50} = \frac{V}{b}$ and $\frac{k}{80} = \frac{V}{c}$ (condone any letter for V even S)		
	So $\underline{c = \frac{8}{5}b}$ *	A1cso*	1.1b
		(2)	
(b)	$d = 2b$ **or** $a = \frac{2}{5}b$ **or** $c = 4a$ **or** $d = 5a$ **or** $d = \frac{5}{4}c$	M1 A1	2.1 3.3
	$\frac{2}{5}b + b + \frac{8}{5}b + 2b = 1$	M1	2.1
	$\Rightarrow 5b = 1$ so $b = \frac{1}{5}$ (o.e.)	A1	1.1b
	$\underline{\underline{a = \frac{2}{25} \quad b = \frac{1}{5} \quad c = \frac{8}{25} \quad d = \frac{2}{5}}}$	A1	3.2a
		(5)	
(c)	[Experiment suggests for Nav] $P(S \mid \{X = 100\}) = 0.3 \Rightarrow k = 30$		
	or $0.3 = \frac{V}{0.4} \Rightarrow V = 0.12$	B1	2.4
	So model won't work since		
	$P(S \mid X = 20) = \frac{30}{20}$ or $\frac{0.12}{0.08}$ and so would be greater than 1		
		(1)	
		(8 marks)	

	Notes
(a)	M1 for use of $P(S \mid X = x) \times P(X = x)$ for $x = 50$ and $x = 80$ (Must see k or their V) Any expression or equation MUST be based on the probability statements in qu.
*	A1cso for rearranging to required result, no incorrect work seen, condone poor notation
NB	Use of values e.g. $b = \frac{50}{20+50+80+100}$ to prove (a) is M0A0 but scores 2nd M1A1 in (b)
	Marks for (b) may be awarded for work seen in (a)
(b)	1st M1 for at least one other relationship (either probability the subject) from the list. 1st A1 for a second different relationship (either probability the subject) from the list. or Allow for: $\frac{ak}{20} = \frac{bk}{50} = \frac{ck}{80} = \frac{dk}{100}$ for 1st M1 1st A1 2nd M1 for using or stating sum of prob's = 1 May be implied by one correct probability. 2nd A1 for one correct probability e.g. $b = \frac{1}{5}$ or exact equivalent such as 0.2 3rd A1 for all correct probabilities. Allow exact equivalents e.g. $c = 0.32$ **Sight of correct distribution or list of probs with no obvious incorrect working is 5/5**
(c)	B1 for deducing $k = 30$ **and** giving a suitable example to show model breaks down

Qu 6	Scheme	Marks	AO
(a)	$2\times4.2\,,\,4\times4\,,\,4\times3.5\,,10\times1$ $(=8.4+16+14+10=48.4)$	M1	1.1b
	[So P($10 < T < 30$) =] $\left[\dfrac{48.4}{90}\right]=\dfrac{121}{225}=0.53777...$ **<u>0.53~0.54</u>** (2sf OK)	A1	1.1b
		(2)	
(b)	(Not suitable as) data is not symmetric <u>or</u> is skew (normal is symmetric) ("Even" distribution or a diagram <u>on its own</u> is not enough so B0)	B1 **(1)**	2.4
(c)	$\int xe^{-x}\ (dx)=\int x\mathrm{d}(-e^{-x})$	M1	2.1
	$=\left[-xe^{-x}\right]-\int\left(-e^{-x}\right)\ (dx)$ $(+c)$	A1	1.1b
	$\int_0^n xe^{-x}\ (dx)=\left[-xe^{-x}-e^{-x}\right]_0^n=\left(-ne^{-n}-e^{-n}\right)-\left[-(0)-1\right]$	dM1	1.1b
	$=\underline{1-(n+1)e^{-n}}$ (*)	A1cso*	1.1b
		(4)	
(d)	Require area = 90 i.e. $k\int_{(0)}^{(n)} xe^{-x}\,dx=90$ (ignore limits)	M1	3.1a
	Using the result in part (c) with $n = 4$ gives $k\left[1-5e^{-4}\right]=90$	M1	2.1
	($k =$) **<u>99</u>**(.0729…) (*****)	A1cso*	1.1b
		(3)	
(e)(i)	[P($10 < T < 30$) =] 0.64863… awrt **<u>0.649</u>**	B1 **(1)**	1.1b
(ii)	[No. of patients =] $(99)\left[\left(1-4e^{-3}\right)-\left(1-2e^{-1}\right)\right]$ $(=53.1..)$	M1	3.4
	Prob $=\dfrac{0.5366...\times99}{90}=0.59027...$[or 0.5907...] = awrt **<u>0.590</u>** or **<u>0.591</u>**	A1 **(2)**	3.2a
(f)	eg Patients might stay longer than 40 hours (Can ignore other comments unless clearly contradictory.)	B1 **(1)**	3.5b
		(14 marks)	

	Notes
(a)	M1 for an attempt to find the number between 10 and 30 (2 correct products or 48 or 48.4 seen) A1 for 2sf answer in [0.53 ~ 0.54] NB use of 48 gives 0.5333… [Correct ans implies 2/2]
(b)	B1 for a comment suggesting not suitable based on (lack of) symmetry <u>or</u> "not bell shaped"
(c)	1st M1 for attempting integration by parts in right direction. Must have $u = x$ and $v=\pm e^{-x}$ 1st A1 for a correct first step, correct first integration and expression for second integral 2nd dM1 (dep on 1st M1) for all integration attempted and some use of at least one limit * 2nd A1 for cso with no incorrect working seen. Minimum is correct int and use of limits seen.
(d)	1st M1 for realising need area under the curve (implied by the integral) = 90 2nd M1 for use of (c) with $n = 4$ and set = 90 May be implied by sight of 99.07… or better * A1cso for $k = 99$ or awrt 99.1 **NB** Allow use of $k = 99$ and show area = awrt 89.9 with a conclusion to score 3/3
(e)(i)	B1 for awrt 0.649
(ii)	M1 for use of (c) with $n = 1$ and $n = 3$ Don't need the 99. Implied by sight of awrt 0.54 A1 for awrt 0.590 or awrt 0.591 Allow 0.59 from correct working seen.
(f)	B1 eg for comment, in context, about the upper limit for <u>time</u> (<u>t or x</u>)(time/hour may be implied)

Notes on Question 5

The question essentially uses the definition of P($A \mid B$) given in the formula booklet.

In particular $\mathrm{P}\left(S \mid \{X = x\}\right) = \dfrac{\mathrm{P}\left(S \cap \{X = x\}\right)}{\mathrm{P}\left(X = x\right)}$ [1]

The first "blob" tells us that $\mathrm{P}\left(S \mid \{X = x\}\right) = \dfrac{k}{x}$ where k is a constant.

The second "blob" tells us that $\mathrm{P}\left(S \cap \{X = x\}\right)$ is the same for all x so $\mathrm{P}\left(S \cap \{X = x\}\right) = V$ where V is a constant.

Using these results in [1] gives $\dfrac{k}{x} = \dfrac{V}{\mathrm{P}(X = x)}$ [2]

Line 1 of MS for part (a) uses $V = \mathrm{P}\left(X = x\right) \times \dfrac{k}{x}$ for $x = 50$ and $x = 80$

Line 2 of MS for part (a) uses [2] with $x = 50$ and $x = 80$

Other implications

Equation [1] can be rearranged to give $\mathrm{P}\left(X = x\right) = x \times \dfrac{V}{k}$ [3]

So when $a + b + c + d = 1$ is used this gives $1 = \dfrac{V}{k}\left(20 + 50 + 80 + 100\right)$ or $\dfrac{V}{k} = \dfrac{1}{250}$ [4]

In particular if we use this relationship in [3] the probabilities a, b, c and d can simply be written down for example $b = \dfrac{50}{250}$ as given in the **NB** in the notes on the MS.

The point is that k and V will vary according to equation [4] but as part (c) shows there are some restrictions on the values k, and therefore V, can take.

Since $\dfrac{k}{x}$ is a probability then, ignoring the trivial cases*, $0 < \dfrac{k}{x} < 1$ and the "restricting" value of x is clearly $x = 20$ so $0 < k < 20$ and from [4] we get $0 < V < \dfrac{20}{250} = \dfrac{2}{25} = a$

So the restrictions on k and on V are given by the shortest distance and its associated probability.

* $k = 0$ would say Tisam can never get the ball in the cup no matter what the distance.

$k = 20$ says she always gets the ball in the cup for any distance.

Pearson Education Limited. Registered company number 872828
with its registered office at 80 Strand, London, WC2R 0RL, United Kingdom