Mark Scheme (Results)

Summer 2022

Pearson Edexcel GCE
AL Further Mathematics (9FM0)
Paper 4B Further Statistics 2

Summer 2022
Question Paper Log Number P72096A*
Publications Code 9FM0_4B_2206_MS*

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 75.

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 $\sqrt{}$ 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
- ⌈ 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 <u>and indicates which response they wish to submit</u>, examiners should mark this response.
If there are several attempts at a question <u>which have not been crossed out</u>, examiners should mark the final answer which is the answer that is the <u>most complete</u>.

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.

Question	Scheme	Marks	AOs
1(a)	[20 × 45.5 + 2080] = 2990 [kg/ha]	B1	3.4
		(1)	
(b)	(A residual is the) difference between the observed value (oe) and the predicted value (oe) (of the dependent variable)	B1	1.2
		(1)	
(c)	$1666567 = 1774155(1-r^2)$	M1	1.1b
	$r = 0.246...$ awrt 0.246	A1	1.1b
		(2)	
(d)(i)	Since *r* is close to 0/weak correlation	B1	2.4
(ii)	e.g. (For $t > 20$…) the residuals do not appear randomly scattered about 0.	B1	3.5a
		(2)	
(e)	Kwame's conclusion cannot be supported using RSS since the two values of RSS do not have the same units.	B1	2.3
		(1)	
			(7 marks)

Notes

(a)	**B1:** cao
(b)	**B1:** Correct definition. Allow equivalent wording. Distance from regression line on its own is B0, but allow if vertical distance or *y* is referenced
(c)	**M1:** Use of correct expression for *r* or r^2 Allow use of $S_{ty} = 45.5 \times 52.0$ [=2366] or RSS: $1774155 - \frac{(S_{ty})^2}{52} = 1666567 \rightarrow [S_{ty} = 2365.2...]$ and then $r = \frac{\text{awrt 2365 or awrt 2366}}{\sqrt{52.0 \times 1774155}}$ **A1:** awrt 0.246 (–0.246 or ± 0.246 scores M1A0)
(d)(i) **(ii)**	**B1:** Correct explanation **B1:** Correct evaluation of the fit of the model's residuals (e.g. variance either side of $t = 20$ does not appear to be the same) 'residuals not randomly scattered' on its own is B0.
(e)	**B1:** Correct assessment of the conclusion involving the units/size of the variables used to calculate the RSS

Question	Scheme	Marks	AOs
2(a)	1.96	B1	3.3
	$57.2 \pm '1.96' \times \frac{1.2}{\sqrt{120}}$	M1	2.1
	(56.985…, 57.414…) awrt (57.0, 57.4)	A1	1.1b
		(3)	
(b)(i)	$\bar{Y} - \bar{W} \sim$ N.......	M1	3.3
(ii)	$\left(0, \frac{1.2^2}{120} + \frac{0.9^2}{140}\right)$	A1A1 **(3)**	1.1b1.1b
	Central limit theorem applies so we do not need to know the distributions of Y and W/ Allows us to assume that sample means ($\bar{Y}$ and $\bar{W}$) are normally distributed.	B1	2.4
		(1)	
(c)	$H_0: \mu_{y(ellow)} = \mu_{w(hite)}$ $H_1: \mu_{y(ellow)} > \mu_{w(hite)}$	B1	2.5
	$z = \frac{57.2 - 56.9}{\sqrt{\frac{1.2^2}{120} + \frac{0.9^2}{140}}} = 2.24949...$	M1 A1	3.1b 1.1b
	CV = 1.6449 [or p-value = 0.01224…] p = awrt 0.01	B1	1.1b
	[Reject H_0] Significant evidence to **support Jamie's claim/mean weight** of the population of **yellow** tennis **balls** is **greater** than **mean weight** of the population of **white** tennis **balls**.	A1	2.2b
		(5)	
			(12 marks)

Notes	
(a)	**B1:** Understanding that sample mean can be modelled using a Normal distribution with z = 1.96 Allow 1.959…or better from calculator **M1:** Setting up confidence interval $57.2 \pm z \times \frac{1.2}{\sqrt{120}}$ **A1:** awrt (57.0, 57.4) Must come from correct z value = 1.96 (or better).
(b)(i)(ii)	**M1:** Translating context into a Normal distribution model **A1:** Correct mean **A1:** Correct variance (allow awrt 0.0178 or exact fraction $\frac{249}{14000}$) **B1:** Correct explanation about the distributions of Y and W or $\bar{Y}$ and $\bar{W}$
(c)	**B1:** Both hypotheses (oe) correct with correct notation (if using μ_x and μ_y these must be defined). **M1:** Standardising using normal distribution test statistic for difference of two means with known variance **A1:** awrt 2.25 **B1:** Correct critical value 1.6449 or better, [or p = 0.01 or better from correct working] **A1:** Drawing a correct inference in context. Do not allow contradictory statements, e.g. 'Do not reject H_0, so Jamie's claim is supported'

Question	Scheme	Marks	AOs
3	$5a+8b=169$	B1	1.1b
	$0.16a+0.01b^2=4$	B1	1.1b
	$a=33.8-1.6b \rightarrow 0.16(33.8-1.6b)+0.01b^2=4$	M1	2.1
	$0.01b^2-0.256b+1.408=0 \rightarrow b=\dfrac{0.256\pm\sqrt{0.256^2-4(0.01)(1.408)}}{2(0.01)}$	M1	1.1b
	$b=8,\ a=21$ (reject $b=17.6,\ a=5.64$ since $a\in\mathbb{Z}^+$)	A1A1	1.1b 2.2a
			(6 marks)

Notes

B1: Correct equation for the means
B1: Correct equation for the variances (allow $0.4^2a+0.1^2b^2=4$)
M1: Attempt to solve two simultaneous equations in *a* and *b* by eliminating one variable
M1: Attempt to solve their quadratic (must be seen if answers are incorrect)
A1: *b* = 8 or *a* = 21 or both sets of values of *a* and *b* without rejecting
A1: Choosing correct pair of solutions $b=8,\ a=21$ only

Question	Scheme	Marks	AOs
4 (a)	d: 2 3 4 3 0 4 12	M1	3.1b
	$\overline{d} = \pm 4$ $s_d = = \sqrt{\frac{1}{6}('198' - 7('4')^2)} = \sqrt{14.333...} = 3.7859\ldots$	M1	1.1b
	$H_0 : \mu_d = 0$ $H_1 : \mu_d > 0$	B1	2.5
	$t = \pm \dfrac{"\pm 4"}{\frac{"3.78..."}{\sqrt{7}}}$	M1	1.1b
	$= \pm\ 2.795\ldots\ ....$ awrt ± 2.80	A1	1.1b
	Critical value $t_6 = \pm 1.943$	B1	1.1b
	[2.80 > 1.943,] therefore there is sufficient evidence to support the doctor's belief/evidence to suggest resting heart rate is reduced.	A1	2.2b
		(7)	
(b)	**Differences** in resting heart rates must be normally distributed for the test to be valid.	B1	2.4
		(1)	
			(8 marks)

Notes

(a)	**M1:** For understanding paired t-test is required and attempting to find differences at least 5 correct, allow ± (may be implied by correct $\overline{d}$ and s_d) **M1:** Complete method for $\overline{d}$ and s_d or $(s_d)^2$ **B1:** Correct model for differences with both hypotheses correct in terms of μ / μ_d (sign of H_1 must be compatible with their $\overline{d}$) **M1:** Method for finding test statistic with their values **A1:** awrt ± 2.80 (allow awrt ± 2.8 from correct working) **B1:** Correct critical ± 1.943 (or better) with compatible sign **A1:** Correct comparison to deduce that the doctor's belief is supported. Must be consistent with their CV and their test statistic (dependent upon all M marks). SC: Difference of means test apply scheme but also allow 2nd B1 for $t_{12} = 1.782$
(b)	**B1:** Correct modelling assumption

Question	Scheme	Marks	AOs
5(a)	The concentration of <u>air pollutant</u> (for each site) follows a normal distribution	B1	2.4
		(1)	
(b)	$F_{12,8}(0.01) = 5.67$ or $F_{8,12}(0.01) = 4.5(0)$	B1	3.3
	$\frac{6.39}{s_B^2} > '5.67'$ $\quad \frac{s_B^2}{6.39} > '4.50'$	M1M1	1.1b 2.1
	$(0<)s_B^2 < 1.12698...$, $s_B^2 > 28.755$	A1	1.1b
		(4)	
(c)	$\chi^2_{12,0.005} = 28.3$ or $\chi^2_{12,0.995} = 3.074$	M1	3.3
	$\frac{12 \times 6.39}{"28.3"} < \sigma_A^2 < \frac{12 \times 6.39}{"3.074"}$	M1	1.1b
	$2.7095... < \sigma_A^2 < 24.94469...$	A1	1.1b
		(3)	
			(8 marks)

Notes

(a)	**B1:** Correct modelling assumption. Allow <u>air pollutant</u> samples are independent. Samples of air pollutant are normally distributed is B0
(b)	**B1:** Setting up either F statistic **M1:** Either correct inequality (condone equation) **M1:** Understanding both correct inequalities are needed (condone equation) **A1:** Correct range of values with awrt 1.13 and awrt 28.8 $s_B^2 < 1.12698...$ and $s_B^2 > 28.755$ is A0
(c)	**M1:** For realising a χ^2 distribution must be used as a model with at least one value correct . Allow $\chi^2_{12,0.01} = 26.217$ or $\chi^2_{12,0.99} = 3.571$ for this mark. **M1:** Setting up 99% CI **A1:** Correct CI with awrt 2.71 and awrt 24.9

Question	Scheme	Marks	AOs
6(a)	$\text{E}(K)=\frac{2\text{E}(X)}{n(n+1)}+\frac{4\text{E}(X)}{n(n+1)}+\frac{6\text{E}(X)}{n(n+1)}+...+\frac{2n\text{E}(X)}{n(n+1)}$ oe	M1	3.1a
	$=\frac{1}{n(n+1)}\left(\sum_{r=1}^{n}2r\right)\text{E}(X)=\frac{1}{n(n+1)}\left(\frac{(2+2n)n}{2}\right)\text{E}(X)(=\text{E}(X))$	M1	2.1
	$=\mu$ (therefore K is an unbiased estimator of μ)	A1cso	1.1b
	$\text{E}(L)=\frac{2\text{E}(X)}{3}+\frac{(n-2)\text{E}(X)}{3(n-2)}=\frac{2\text{E}(X)}{3}+\frac{1\text{E}(X)}{3}(=\text{E}(X))$	M1	1.1b
	$=\mu$ (therefore L is an unbiased estimator of μ)	A1cso	1.1b
		(5)	
(b)(i)	$\text{Var}(K)=\text{Var}\left(\frac{2X}{n(n+1)}\right)+\text{Var}\left(\frac{4X}{n(n+1)}\right)+...+\text{Var}\left(\frac{2n(X)}{n(n+1)}\right)$	M1	3.1a
	$=\frac{2^2}{n^2(n+1)^2}\text{Var}(X)+\frac{4^2}{n^2(n+1)^2}\text{Var}(X)+...+\frac{(2n)^2}{n^2(n+1)^2}\text{Var}(X)$	M1	1.1b
	$=\frac{\sum_{r=1}^{n}(2r)^2}{n^2(n+1)^2}\text{Var}(X)=\frac{4\sum_{r=1}^{n}r^2}{n^2(n+1)^2}\text{Var}(X)=\frac{\frac{4}{6}(n)(n+1)(2n+1)}{n^2(n+1)^2}\sigma^2$	M1	2.1
	$=\frac{2(2n+1)}{3n(n+1)}\sigma^2$	A1	1.1b
(ii)	$\text{Var}(L)=\text{Var}\left(\frac{X_1+X_2}{3}\right)+\text{Var}\left(\frac{X_3+X_4+...+X_n}{3(n-2)}\right)$	M1	1.1b
	$=\frac{2}{9}\text{Var}(X)+\frac{(n-2)}{9(n-2)^2}\text{Var}(X)$	M1	2.1
	$=\frac{2n-3}{9(n-2)}\sigma^2$	A1	1.1b
		(7)	
(c)	For large values of n $\text{Var}(K)\to 0$ $\text{Var}(L)\to\frac{2}{9}(\sigma^2)$	M1	2.1
	(Since both are unbiased,) the better estimator is the one with the smaller variance or $0<\frac{2}{9}(\sigma^2)$	M1	2.4
	Therefore K is the better estimator and Korhan wins the challenge.	A1	2.2a
		(3)	
			(15 marks)

Notes

(a)	**M1:** Using independence to set up expression for $\text{E}(K)$ **M1:** Use of $\sum_{r=1}^{n}r\left(=\frac{n(n+1)}{2}\right)$ **A1cso:** Correct conclusion from correct working **M1:** Using independence to set up expression for $\text{E}(L)$ **A1cso:** Correct conclusion from correct working
(b)(i)	**M1:** Using independence to set up expression for $\text{Var}(K)$ **M1:** Use of $\text{Var}(aX)=a^2\text{Var}(X)$ Implied by $\frac{2^2}{n^2(n+1)^2}$

(ii)	**M1:** Use of $\sum_{r=1}^{n} r^2$ **A1:** Correct equivalent expression for Var(K) oe **M1:** Using independence to set up expression for Var(L) **M1:** Use of $\text{Var}(aX) = a^2\text{Var}(X)$ and understanding $\text{Var}(X_1 + X_2) = 2\text{Var}(X)$ Implied by either correct term **A1:** Correct equivalent expression for Var(L) oe
(c)	**M1**: Finding correct limits as n gets larger for each expression allow ft Note: Solving $\dfrac{2n-3}{9(n-2)} = \dfrac{2(2n+1)}{3n(n+1)} \to n = -0.53,\ 2.46,\ 4.57$ so allow comments relating to $n \geq 5\,(4.57)$ **M1:** Correct explanation **A1:** Deducing that Korhan is the winner (dependent upon both M marks)

Question	Scheme	Marks	AOs
7	$P = 2(L + \frac{40}{L})$	M1	3.1a
	$\mathrm{f}(l) = \frac{1}{6}$	B1	1.1b
	$\mathrm{E}(P) = \mathrm{E}(2(L + \frac{40}{L})) = \int_4^{10} \frac{1}{6}(\text{"}2l + \frac{80}{l}\text{"})\,\mathrm{d}l$	M1	2.1
	$= \left[\frac{1}{6}l^2 + \frac{40}{3}\ln\lvert l\rvert\right]_4^{10}$	M1A1	1.1b 1.1b
	$= (\frac{1}{6}(100) + \frac{40}{3}\ln 10) - (\frac{1}{6}(16) + \frac{40}{3}\ln 4))$	M1	1.1b
	$= 26.217...$ awrt 26.2	A1	1.1b
			(7 marks)
	Notes		
	M1: Finding an expression for the perimeter in terms of L **B1:** Correct distribution for L (may be implied by $\mathrm{E}(L) = 7$) **M1:** Setting up integral for expectation of perimeter (allow 2 separate integrals e.g. $2(7) + \int_4^{10} \frac{1}{6}(\frac{80}{l})\,\mathrm{d}l$) **M1:** Attempt to integrate an expression for expectation of perimeter (allow two separate expressions) **A1:** Correct integration **depM1:** (dep on previous M1) Use of correct limits 10 and 4 **A1:** awrt 26.2 Note: exact value is $14 + \dfrac{40\ln(2.5)}{3}$		

Question	Scheme	Marks	AOs
8(a)	$1.5m - 0.25m^2 - 1.25 = 0.5 \quad (\rightarrow 0.25m^2 - 1.5m + 1.75 = 0)$	M1	1.1b
	$m = 3 - \sqrt{2} \qquad$ (reject $m = 3 + \sqrt{2}$)	A1	2.2a
		(2)	
(b)	P(X < 1.6 | X > 1.2) = $\dfrac{\mathrm{P}(1.2 < X < 1.6)}{\mathrm{P}(X > 1.2)}$	M1	3.1a
	$\dfrac{\mathrm{F}(1.6) - \mathrm{F}(1.2)}{1 - \mathrm{F}(1.2)}$	M1	1.1b
	$= \dfrac{32}{81}$	A1	1.1b
		(3)	
(c)	$\mathrm{P}(Y \leqslant y) = \mathrm{P}\left(\dfrac{1}{X} \leqslant y\right)$		
	$= \mathrm{P}\left(X \geqslant \frac{1}{y}\right) = 1 - \mathrm{F}(\frac{1}{y})$	M1	3.1a
	$= 1 - (\frac{1.5}{y} - 0.25(\frac{1}{y})^2 - 1.25)$	M1	1.1b
	$\mathrm{F}(y) = \begin{cases} 0 & y < \frac{1}{3} \\ 2.25 - \frac{1.5}{y} + 0.25(\frac{1}{y})^2 & \frac{1}{3} \le y \le 1 \\ 1 & y > 1 \end{cases}$	A1 A1	2.1 1.1b
		(4)	
(d)	$\mathrm{f}(y) = \dfrac{\mathrm{d}}{\mathrm{d}y}\big(\mathrm{F}(y)\big) = 1.5y^{-2} - 0.5y^{-3} \rightarrow \mathrm{f}'(y) = -3y^{-3} + 1.5y^{-4}$	M1	3.1a
	$\mathrm{f}'(y) = -3y^{-3} + 1.5y^{-4} = 0$	depM1	1.1b
	$1.5y^{-4} = 3y^{-3} \rightarrow \frac{1.5}{y^4} = \frac{3}{y^3}$		
	[Mode of Y =] 0.5 (since $\mathrm{f}''(0.5) = -48 < 0$)	A1	1.1b
		(3)	
			(12 marks)

Notes

(a)	**M1:** Correct equation for m **A1:** $m = 3 - \sqrt{2}$ only (isw if exact answer is given then rounded)
(b)	**M1:** Determining the two probabilities required to find the probability **M1:** Correct ratio of probabilities **A1:** allow awrt 0.395
(c)	**M1:** Realising that $\mathrm{P}\left(X \geqslant \frac{1}{y}\right)$ is necessary **M1:** Correct use of F(x) **A1:** Determining correct limits (allow $\underline{<}$ for < etc.) **A1:** All lines of cdf correct (ignoring limits) must be in terms of y SC: If 0 scored, then [F(y) =] $\frac{1.5}{y} - 0.25(\frac{1}{y})^2 - 1.25 \quad \frac{1}{3} \le y \le 1$ scores M0M0A1A0
(d)	**M1:** Realising that the cdf must be differentiated twice **depM1:** (dep on previous M1) Equating their $\mathrm{f}'(y) = 0$ with attempt to solve **A1:** 0.5 cao