GCE A LEVEL MARKING SCHEME

SUMMER 2023

A LEVEL
FURTHER MATHEMATICS
UNIT 5 FURTHER STATISTICS B
1305U50-1

INTRODUCTION

This marking scheme was used by WJEC for the 2023 examination. It was finalised after detailed discussion at examiners' conferences by all the examiners involved in the assessment. The conference was held shortly after the paper was taken so that reference could be made to the full range of candidates' responses, with photocopied scripts forming the basis of discussion. The aim of the conference was to ensure that the marking scheme was interpreted and applied in the same way by all examiners.

It is hoped that this information will be of assistance to centres but it is recognised at the same time that, without the benefit of participation in the examiners' conference, teachers may have different views on certain matters of detail or interpretation.

WJEC regrets that it cannot enter into any discussion or correspondence about this marking scheme.

WJEC GCE A LEVEL FURTHER MATHEMATICS

UNIT 5 FURTHER STATISTICS B

SUMMER 2023 MARK SCHEME

Qu.	Solution	Mark	Notes
1 (a)	($\sum x$ = 2759 $\sum x^2$ = 846 081)		
	$\hat{\mu} = 306.555\ldots$	B1	At least 1dp
	$\hat{\sigma}^2 = \frac{1}{8}(846081 - 9 \times 306.555\ldots^2) = \frac{331}{9} = 36.777\ldots$	M1A1	M1 for appropriate use of calculator or Use of $\hat{\sigma}^2 = \frac{1}{n-1}(\sum x^2 - n\overline{x}^2)$ Allow 33.7122 from rounding $\hat{\mu}$ to 306.56 M1A0 for 40.6096… from $\overline{x} = 306.55$ M1A0 for 6.12 from $\overline{x} = 306.6$ FT their $\hat{\mu}$ for M1 only, provided $\hat{\sigma}^2 > 0$
(b)	$H_0: \mu = 305$ $H_1: \mu > 305$	B1	
	DF = 8	B1	si
	CV = 1.860	B1	FT their DF
	$t = \frac{306.5555\ldots - 305}{\sqrt{\frac{36.7777\ldots}{9}}}$	M1	FT their $\hat{\mu}$ and $\hat{\sigma}^2$
	$t = 0.7695\ldots$	A1	cao Accept 0.77 from correct working Allow 0.806 from $\hat{\mu} = 306.56$ and $\hat{\sigma}^2 = 33.71(22)$
	Since $0.7695 < 1.860$ there is insufficient evidence to reject H_0.	m1	FT their t Dep on use of t-distribution.
	There is insufficient evidence to say that this is an old kettle.	A1	cso
(c)	Valid factor. e.g. the initial water temperature. e.g. the initial kettle temperature. e.g. the ambient temperature. e.g. the volume of water. e.g. the voltage going to the kettle. e.g. the mineral content of the water	E1	
		Total [11]	

Qu.	Solution	Mark	Notes
2 (a)	$\mathrm{E}(T_1) = \dfrac{3\mathrm{E}(\overline{X}) + 7\mathrm{E}(\overline{Y})}{10}$	M1	
	$\mathrm{E}(T_1) = \dfrac{3\mu + 7\mu}{10}$		
	$\mathrm{E}(T_1) = \mu$, therefore T_1 is an unbiased estimator for μ.	A1	Convincing
(b)	$\mathrm{E}(T_2) = \dfrac{\mathrm{E}(\overline{X}) + a^2\mathrm{E}(\overline{Y})}{1+a}$		
	To be an unbiased estimator for μ		
	$\dfrac{\mu + a^2\mu}{1+a} = \mu$	M1	Forming an equation in μ. si
	$1 + a^2 = 1 + a$	A1	oe
	$a = 0$ or $a = 1$. a is positive ∴ $a = 1$ (so $T_2 = \frac{\overline{X}+\overline{Y}}{2}$)	A1	Must reject $a = 0$. If M0, then SC1 for verification only
(c)	$\mathrm{Var}(T_1) = \dfrac{3^2 \times \mathrm{Var}(\overline{X}) + 7^2 \times \mathrm{Var}(\overline{Y})}{10^2}$	M1	Use of $\mathrm{Var}(cW) = c^2\mathrm{Var}(W)$
	$\mathrm{Var}(T_1) = \dfrac{9 \times \frac{\sigma^2}{20} + 49 \times \frac{k\sigma^2}{25}}{100}$	M1	Use of $\mathrm{Var}(\overline{W}) = \mathrm{Var}(W)/n$
	$\mathrm{Var}(T_1) = \dfrac{45\sigma^2 + 196k\sigma^2}{10000} = \dfrac{\sigma^2}{10000}(45 + 196k)$	A1	oe, cao $\mathrm{Var}(T_1) = \frac{9\sigma^2}{2000} + \frac{49k\sigma^2}{2500}$
	$\mathrm{Var}(T_2) = \dfrac{1}{4}(\mathrm{Var}(\overline{X}) + \mathrm{Var}(\overline{Y}))$ $\mathrm{Var}(T_2) = \dfrac{1}{4}\left(\dfrac{\sigma^2}{20} + \dfrac{k\sigma^2}{25}\right)$	M1	
	$\mathrm{Var}(T_2) = \dfrac{\sigma^2}{400}(5 + 4k)$	A1	oe $\mathrm{Var}(T_2) = \frac{\sigma^2}{80} + \frac{k\sigma^2}{100}$ If left in terms of a $\mathrm{Var}(T_2) = \dfrac{\sigma^2(5 + 4a^4k)}{100(1+a)^2}$

Qu.	Solution	Mark	Notes
2 (d)	$\frac{\sigma^2}{400}(5+4k) = \frac{45\sigma^2 + 196k\sigma^2}{10000}$	M1	M1 for setting their $\text{Var}(T_1) = \text{Var}(T_2)$
	$\frac{10000}{400}(5+4k) = 45 + 196k$ or $25(5+4k) = 45 + 196k$	m1	Forming an equation in k
	$125 + 100k = 45 + 196k$		
	$k = \frac{5}{6}$ ***ag**	A1	Convincing.
(e)	$V = \text{Var}(T_3) = (1-\lambda)^2 \times \text{Var}(\overline{X}) + \lambda^2 \times \text{Var}(\overline{Y})$ $V = \text{Var}(T_3) = (1-\lambda)^2 \times \frac{\sigma^2}{20} + \lambda^2 \times \frac{k\sigma^2}{25}$	B1	cao
	$\frac{\mathrm{d}V}{\mathrm{d}\lambda} = \frac{-2(1-\lambda)\sigma^2}{20} + \frac{2\lambda k\sigma^2}{25}$	M1	M1 for expression for $\frac{\mathrm{d}V}{\mathrm{d}\lambda}$ At least 1 term correct
	Smallest variance is when $\frac{\mathrm{d}V}{\mathrm{d}\lambda} = 0$	M1	M1 for setting $\frac{\mathrm{d}V}{\mathrm{d}\lambda} = 0$ and attempt to solve.
	$\frac{2\lambda k\sigma^2}{25} = \frac{2(1-\lambda)\sigma^2}{20}$		
	$\lambda k = \frac{5}{4}(1-\lambda)$		
	$\lambda k + \frac{5\lambda}{4} = \frac{5}{4}$	m1	m1 for λ on one side of equation. Provided M1M1 awarded
	$\lambda\left(\frac{4k}{4} + \frac{5}{4}\right) = \frac{5}{4}$		
	$\lambda = \frac{5}{4k+5}$	A1	cao
	$\frac{\mathrm{d}^2V}{\mathrm{d}\lambda^2} = \frac{\sigma^2}{10} + \frac{2k\sigma^2}{25} > 0$		
	Therefore, it is a minimum.	E1	E1 for verifying minimum, oe method
		Total [19]	

Qu.	Solution	Mark	Notes
3 (a)	$\bar{x}\left(=\frac{4014}{90}\right)=44.6$	B1	si
	Standard error = $\sqrt{\frac{4.7^2}{90}}$	B1	si
	Use of $\bar{x}\pm z\times \mathrm{SE}$	M1	FT their $\bar{x}$ and SE
	= 44.6$\pm$2.5758 $\times\sqrt{\frac{4.7^2}{90}}$	A1	provided $\neq 4.7$ for M1A1 Must show working. From tables 2.576
	$[43.3, 45.9]$	A1	cao
(b)	Because the confidence level has decreased, the width is narrower.	E1	Condone width will be smaller.
(c)	$\bar{x}=\frac{49.9+52.6}{2}=51$		
	Use of $\bar{x}\pm 1.96\times\frac{\sigma}{\sqrt{n}}$	M1	
	Upper limit of 95% CI is given by $51+1.96\times\frac{\sigma}{\sqrt{100}}=52.6$ oe	A1	
	OR		
	$2\times\frac{\sigma}{\sqrt{100}}\times 1.96=3.2$	(A1)	
	$\sigma=8.163\ldots..$	A1	cao
(d)	Valid comment. Eg. The confidence intervals suggest that athletes who compete in the 400m have lower RHR on average.	E1	FT their CI from (a)
	Valid reason. Eg. The confidence interval for the athletes who compete in the 400m lies entirely below the confidence interval for the athletes who compete in the discus.	E1	Condone 'Non-overlapping confidence intervals'.
	Valid comment and reason. e.g. The RHR of athletes who compete in the discus event are possibly more varied, as the width of the CI is wider and the confidence level is lower.	(E2)	
		Total [11]	

<table>
<tr><th>Qu.</th><th>Solution</th><th>Mark</th><th>Notes</th></tr>
<tr><td>4(a)(i)</td><td>H_0: The population median difference between the number of social media followers before and after appearing on the television show is 0.

H_1: The population median difference, when subtracting the number of social media followers before appearing on the show from the number of social media followers after appearing on the show, is positive.

OR $H_0: \eta_d = 0$ $H_1: \eta_d > 0$ where η_d is the median difference in numbers of followers before and after appearing on the show, $\eta_d = \eta_{\text{after}} - \eta_{\text{before}}$</td><td>B1</td><td>Both
oe</td></tr>
<tr><td></td><td><table><tr><td>Contestant</td><td>A</td><td>B</td><td>C</td><td>D</td><td>E</td><td>F</td><td>G</td><td>H</td><td>I</td></tr><tr><td>Difference</td><td>361</td><td>-10</td><td>751</td><td>0</td><td>603</td><td>-239</td><td>-56</td><td>270</td><td>187</td></tr></table></td><td>B1</td><td>Accept differences with opposite signs.</td></tr>
<tr><td></td><td>Ranks<table><tr><td>Contestant</td><td>A</td><td>B</td><td>C</td><td>D</td><td>E</td><td>F</td><td>G</td><td>H</td><td>I</td></tr><tr><td>Ranks</td><td>6</td><td>1</td><td>8</td><td>-</td><td>7</td><td>4</td><td>2</td><td>5</td><td>3</td></tr></table></td><td>M1
A1</td><td>M1 either attempt at ranks. FT one slip in difference for A1</td></tr>
<tr><td></td><td>W^+ = Sum of positive ranks (W^- = Sum of negative ranks)
= 6 + 8 + 7 + 5 + 3 = 29 (= 1 + 4 + 2 = 7)</td><td>M1
A1</td><td></td></tr>
<tr><td></td><td>n = 8
Upper CV = 28 (Lower CV = 8)</td><td>B1</td><td>Must match direction of H_1</td></tr>
<tr><td></td><td>Since 29 > 28 (OR 7 < 8) there is sufficient evidence to reject H_0.</td><td>m1</td><td></td></tr>
<tr><td></td><td>There is evidence to suggest that appearing on the show may increase the number of social media followers Llŷr has.</td><td>A1</td><td>cso</td></tr>
<tr><td>(ii)</td><td>Valid comment
e.g. He should apply to appear on the show if he wants more social media followers.
e.g. He should not apply to appear on the show if he doesn’t want more social media followers.</td><td>E1</td><td></td></tr>
<tr><td>(b)</td><td>The underlying distribution of the differences may not be normally distributed.</td><td>E1</td><td>Or equivalent statement</td></tr>
<tr><td></td><td>Data are paired.</td><td>E1</td><td></td></tr>
<tr><td></td><td></td><td>**Total**
[12]</td><td></td></tr>
</table>

Qu.	Solution	Mark	Notes
5 (a)	$\overline{X} \sim \mathrm{N}\left(75, \frac{10^2}{5}\right)$	B1	si oe
	$\mathrm{P}(\overline{X} < 70) = 0.13177$	M1A1	
	ALTERNATIVE METHOD $T = X_1 + X_2 + \cdots + X_5$		
	$T \sim N(375{,}500)$	(B1)	
	$P(T < 350) = 0.13177$ ...	(M1A1)	
(b)	$\overline{X} \sim \mathrm{N}\left(75, \frac{10^2}{n}\right)$	B1	si
	$\mathrm{P}(\overline{X} > 80) \approx 0.007$		
	$\mathrm{P}\left(Z > \frac{80 - 75}{\sqrt{\frac{100}{n}}}\right) \approx 0.007$	M1	Standardising accept (75 – 80) for numerator
	$\frac{80 - 75}{\sqrt{\frac{100}{n}}} \approx 2.4572$	M1B1	M1 for correct standardisation set equal to $2 \le k \le 3$ B1 for 2.457 or better
	$n = 24$	A1	cao
(c)	Let $T = X_1 + X_2 + X_3 + Y_1 + Y_2 + Y_3 + Y_4$		
	$\mathrm{E}(T) = 497$ $\mathrm{Var}(T) = 3 \times 100 + 4 \times 36$ $\mathrm{Var}(T) = 444$ $\mathrm{P}(T > 500) = 0.44339\ldots$	B1 M1 A1 B1	 From tables 0.44433
(d)	Valid assumption. e.g. the workers do not carry any extra baggage. e.g. mass of workers' clothes may be ignored.	E1	
		Total **[13]**	

Qu.	Solution	Mark	Notes
6 (a)	H_0: The median numbers of races entered by competitors who are club members and those who are not club members are the same. H_1: The median number of races entered by competitors who are club members is **more** than the median number of races entered by those who are not club members.	B1	Accept $H_0: \eta_1 = \eta_2 \qquad H_1: \eta_1 > \eta_2$
	Use of the formula $U = \sum\sum z_{ij}$ U = 1 + 6 + 6 + 3 + 2 + 4 OR U = 5 + 0 + 0 + 3 + 4 + 2	M1	oe
	U = 22 OR U = 14	A1	
	Upper critical value is 29 OR Lower CV is 7	B1	
	22 < 29 OR 14 >7, there is insufficient evidence to reject H_0.	m1	
	There is insufficient evidence to suggest that athletes race more frequently if they are members of a triathlon club.	A1	cso
(b)	The samples are independent rather than paired.	E1 **Total** **[7]**	E0 for data is ordinal Ignore spurious additional comments
7 (a)	(SE of difference of means) $=\sqrt{\frac{0.75^2}{6}+\frac{0.6^2}{5}}$	M1	Award M1 for Var = $\frac{0.75^2}{6}+\frac{0.6^2}{5}$
	= 0.407…	A1	si
	$2.2 - k\sqrt{\frac{0.75^2}{6}+\frac{0.6^2}{5}} = 1.25$	M1	Condone > FT their SE provided ≠ $\sqrt{0.75^2+0.6^2}$
	k = 2.333…	A1	cao
	Probability from calculator = 0.99018…… Or 0.99010 from tables	M1	FT their k for M1A1
	Largest value of p is 98.	A1	Accept 98.04
(b)	Valid assumption e.g.The time trials are all done on the same terrain. She suffers no mechanical problems. She doesn't get quicker because she's fitter. She isn't slower because she's tired. Weather conditions are similar. Wears the same clothing.	E1 **Total** **[7]**	