GCE A LEVEL MARKING SCHEME

SUMMER 2022

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

INTRODUCTION

This marking scheme was used by WJEC for the 2022 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 2022 MARK SCHEME

Qu. No.	Solution	Mark	Notes
1	$\bar{x} = 15.37$	B1	
	Standard error = $\sqrt{\frac{0.9}{10}}$	B1	SE^2 = $\frac{0.9}{10}$
	Use of $\bar{x} \pm z \times \text{SE}$	M1	FT their $\bar{x}$ and SE $\neq \sqrt{0.9}$
	$=15.37 \pm 1.6449 \times \sqrt{\frac{0.9}{10}}$	A1	1.645 or better
	$[14.88, 15.86]$	A1	cao
		Total [5]	

Qu. No.	Solution	Mark	Notes
2 (a)(i)	$\mathrm{P}(X > 14) = 0.5793$	M1A1	M1 for correct method (calculator or standardizing)
(ii)	$\mathrm{P}(X > 14 \text{ for two out of three}) = 0.5793^2 \times 0.4207 \times 3$	M1	Ft for M1A1 "their (i)" and "1-(i)" awrt 0.423 or 0.424
	$= 0.4235$	A1	
		(4)	
(b)	Let $T = X_1 + X_2 + X_3 + \cdots + X_8$ $\mathrm{E}(T) = 120$ $\mathrm{Var}(T) = 8\mathrm{Var}(X)$ $\mathrm{Var}(T) = 200$ $\mathrm{P}(T > 160) = 0.00234 \text{ (3sf)}$	B1 M1 A1 A1	cao 0.00233 from tables
		(4)	
(c)	Let $A = X_1 + X_2 + X_3$ Let $B = X_1 + X_2 + X_3 + X_4 + X_5$ $A \sim \mathrm{N}(45, 75)$ and $B \sim \mathrm{N}(75, 125)$	B1	si
	Consider $U = B - 2A$ $\mathrm{E}(U) = -15$	M1 A1	M1A0 for $\mathrm{E}(U) = 105$ from $U = 2B - A$
	$\mathrm{Var}(U) = \mathrm{Var}(B) + 2^2\mathrm{Var}(A)$ $= 425$	M1 A1	M1A1 for $\mathrm{Var}(U) =$ 575 from $U = 2B - A$
	$\mathrm{P}(U > 0)$	m1	Dependent on 1st M1 and $U = B - 2A$ m0A0 if $U = 2B - A$
	$= 0.2334$	A1	cao 0.23270 from tables.
		(7)	
		Total [15]	

Qu. No.	Solution	Mark	Notes
3 (a)	Valid reason e.g. No knowledge of underlying distribution Data are ordinal. Interval scale assumption may not be valid.	E1 (1)	
(b)	H_0: Students from the north and the south of the county are similarly stressed. H_1: Students from the north and the south of the county are NOT similarly stressed.	B1	$H_0: \eta_N = \eta_S$ $H_1: \eta_N \neq \eta_S$
	Upper critical value = 22 Lower critical value = $5 \times 5 - 22 = 3$	B1	For either CV
	Use of the formula $U = \sum\sum z_{ij}$	M1	Attempt to use
	U = 4 + 0 + 4 + 3 + 4 = 15 OR U = 1 + 5 + 1 + 2 + 1 = 10	A1	
	Since 15 < 22 OR 10 > 3 and there is insufficient evidence to reject H_0. There is not enough evidence to say that students from the North and from the South have different stress levels.	B1 E1 (6)	FT their CV and U cso
(c)	Valid improvement. e.g. Bigger sample size. e.g. Use a control test to see if students from the North and South are generally more stressed.	E1 (1)	
		Total [8]	

Qu. No.	Solution	Mark	Notes
4 (a)	$\hat{p} = \frac{940}{2000} = 0.47$	B1	
	$\text{ESE} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$		
	$= \sqrt{\frac{0.47 \times 0.53}{2000}}$	M1	FT their $\hat{p}$ for M1A1
	$= 0.01116 \ldots$	A1	
	95% confidence limits are $\hat{p} \pm z \times \text{ESE}$	M1	FT their $\hat{p}$ and ESE for M1A1
	$0.47 \pm 1.96 \times 0.01116 \ldots$	A1	
	Giving $[0.448, 0.492]$	A1	cao
		(6)	
(b)	Two valid reasons. Eg. We have used an approximate value for p (in calculating the standard error). The binomial distribution has been approximated by the normal distribution. No continuity correction has been used.	E2 (2)	E1 for one reason
(c)			Full FT their $\hat{p}$
	$2.5758 \ldots \times \sqrt{\frac{0.53 \times 0.47}{n}} \le 0.02$	M1	Attempt at equation or inequality with $2.5758, n$ and 0.02 oe
	$\frac{2.5758 \ldots \times \sqrt{0.53 \times 0.47}}{0.02} \le \sqrt{n}$	A1	Correct equation or inequality. 4132.4295… from $Z_{0.995} = 2.576$
	$n \ge 4131.88 \ldots ..$	A1	2133 from $Z_{0.995} = 2.576$
	Therefore, an additional (4132 – 2000 =) 2132 people.	A1	
		(4)	
		Total [12]	

Qu. No.	Solution	Mark	Notes
5 (a)	$\bar{x} = \frac{2163}{50}$		
	$= 43.26$	B1	
	$s^2 = \frac{1}{49} \times \left(98508 - \frac{2163^2}{50}\right)$	M1	
	$s^2 = 100.7473(469\ldots)$ or $s = 10.0(3729779\ldots)$	A1	si
	H_0: $\mu = 38$ H_1: $\mu > 38$	B1	
	Under H_0, $\bar{X} \sim N\left(38, \frac{100.747\ldots}{50}\right)$		*Alternative method*
	p-value = P($\bar{X} > 43.26$ | H_0 is true)	M1	M1 for Test statistic = $\frac{43.26-38}{10.03729779\ldots/\sqrt{50}}$ if standardising.
	p-value = 0.000105	A1	A1 p-value from tables = 0.00010
	Since $p << 0.05$ there is strong evidence to reject H_0.	m1	
	There is strong evidence to reject the laboratory's claim that the average time taken for test results to be returned is 38 hours.	A1	FT from their p-value and corresponding conclusion above
	Valid Headline implying failure on the part of the laboratory. e.g. Lab lets down patients. e.g. Laboratory takes longer than claimed to process test results.	B1 (9)	
(b)	Valid explanation. e.g. Because n is large, the central limit theorem allows us to use the normal distribution. e.g. Because n is large, the CLT allows us to assume that the distribution of the sample mean is normal.	E1 (1)	
(c)	Valid explanation. e.g. Random sampling eliminates the bias that may occur from taking a batch from, say, the same day. e.g. 50 consecutive results might all come from a time when the process is having a good, or bad, run. Randomisation avoids this.	E1 (1)	
(d)(i)	A t-test because the sample size is small.	E1	
(ii)	The assumption would be that the time taken for results to be returned is normally distributed.	E1 (2)	
		Total [13]	

Qu. No.	Solution	Mark	Notes
6	H_0: median = 4.2 H_1: median ≠ 4.2	B1	Both $H_0: \eta = 4.2$ $H_1: \eta \neq 4.2$ Condone mean
	(see table below)	M1 A1	Differences
	$W^+ = 5 + 4 + 9 + 1 + 3 + 8$ OR $W^- = 7 + 2 + 6$	M1	Attempt at summing ranks
	= 30 = 15	A1	For A1, FT their differences provided M1M1 awarded. M1A0M1A0 if rank of 1 assigned to diff = 0 FT n = 10 if M1A0M1A0 awarded
	Upper CV = 43 OR Lower CV = 2	B1	FT their TS and CV
	Because 30 < 43 OR 15 > 2, there is insufficient evidence to reject H_0.	B1	cso
	The test suggests that the zoologist should abandon his studies on this population.	E1	
		Total [8]	

Diff $x - 4.2$	+ 0.8	- 1.0	+ 0.7	- 0.2	- 0.9	0	+ 1.9	+ 0.1	+ 0.6	+ 1.7
Rank	5	7	4	2	6	.	9	1	3	8

Qu. No.	Solution	Mark	Notes
7(a)	$(X+Y) \sim \mathrm{N}(180, 2\sigma^2)$	B1	si
	$\mathrm{P}(180-\sigma < X+Y < 180+\sigma)$ $= \mathrm{P}\left(\frac{180-\sigma-180}{\sqrt{2\sigma^2}} < Z < \frac{180+\sigma-180}{\sqrt{2\sigma^2}}\right)$	M1	
	$= \mathrm{P}\left(\frac{-1}{\sqrt{2}} < Z < \frac{1}{\sqrt{2}}\right)$	M1	SC2 for only doing one side leading to 0.7602 or 0.76115 from tables
	$= 0.52 \ldots$	A1	
		(4)	
(b)	$\mathrm{E}(T_1) = \mathrm{E}(45 + \frac{1}{4}(3X - Y))$		
	$\mathrm{E}(T_1) = 45 + \frac{3}{4}\mathrm{E}(X) - \frac{1}{4}\mathrm{E}(180 - X)$	M1	M1 for either first line.
	$\mathrm{E}(T_1) = 45 + \frac{3}{4}\alpha - 45 + \frac{1}{4}\alpha$	M1	Convincing
	$\mathrm{E}(T_1) = \alpha$	A1	
	T_1 is an unbiased estimator for α		
	$\mathrm{Var}(T_1) = \mathrm{Var}(45 + \frac{1}{4}(3X - Y))$		
	$\mathrm{Var}(T_1) = \frac{9}{16}\mathrm{Var}(X) + \frac{1}{16}\mathrm{Var}(Y)$	M1	
	$\mathrm{Var}(T_1) = \frac{5}{8}\sigma^2$	A1	
	$\mathrm{Var}(T_1) < \sigma^2$	E1	FT their $\mathrm{Var}(T_1) = k\sigma^2$ where $k < 1$
	$\therefore T_1$ is a better estimator than X	(6)	
(c)(i)	$\mathrm{E}(T_2) = \mathrm{E}(\lambda X + (1-\lambda)(180° - Y))$		
	$\mathrm{E}(T_2) = \lambda\alpha + (1-\lambda)(180° - \beta)$	M1	
	$\mathrm{E}(T_2) = \lambda\alpha + (1-\lambda)(\alpha)$		
	$\mathrm{E}(T_2) = \lambda\alpha + \alpha - \lambda\alpha = \alpha$	A1	Convincing
(ii)	$\mathrm{Var}(T_2) = \mathrm{Var}(\lambda X + (1-\lambda)(180° - Y))$		
	$\mathrm{Var}(T_2) = \lambda^2\mathrm{Var}(X) + (1-\lambda)^2\mathrm{Var}(180° - Y)$	M1	
	$\mathrm{Var}(T_2) = \lambda^2\sigma^2 + (1-\lambda)^2\sigma^2$	A1	oe ISW

Qu. No.	Solution	Mark	Notes
(iii)	$\frac{\mathrm{d}}{\mathrm{d}\lambda}\mathrm{Var}(T_2) = 2\lambda\sigma^2 - 2(1-\lambda)\sigma^2$	M1A1	M1 for attempt to differentiate with at least one decrease in power. For A1, FT $\mathrm{Var}(T_2)$ for equivalent difficulty only.
			Use of $\frac{\mathrm{d}}{\mathrm{d}\lambda}\mathrm{Var}(T_2) = 0$
	$\frac{\mathrm{d}}{\mathrm{d}\lambda}\mathrm{Var}(T_2) = 0$ gives the best estimator	M1	
	$2\lambda\sigma^2 = 2(1-\lambda)\sigma^2$		
	$2\lambda = 2 - 2\lambda$		cao
	$\lambda = \frac{1}{2}$	A1	Accept alternative justification
	$\frac{\mathrm{d}^2}{\mathrm{d}\lambda^2}\mathrm{Var}(T_2) = 4\sigma^2 > 0$, therefore minimum	E1	
		(9)	
		Total [19]	