A2 Further Mathematics Unit 5: Further Statistics B

Solutions and Mark Scheme

Qu. No.	Solution	Mark	AO	Notes
1(a)(i)	Upper quartile = $\mu + 0.6745\sigma$ = 32 + 0.6745 × 4 = 34.7 This is the time that is exceeded on 25% of the days.	M1 A1 E1	AO3 AO1 AO2	
(ii)	Let $T = X_1 + X_2 + X_3 + X_4 + X_5$ Then $E(T) = 160$ Var(T) = 5Var(X) Var(T) = 80 $P(T > 170) = 0.132$	 B1 M1 A1 B1	 AO3 AO3 AO1 AO1	
(b)	Consider $U = X - 2Y$ $E(U) = -4$ Var(U) = Var(X) + 4Var(Y) = 32 We require P(U > 0) = 0.240	M1 A1 M1 A1 M1 A1 **[13]**	AO3 AO1 AO3 AO1 AO3 AO1	
2(a)	$\Sigma x = 691,\ \Sigma x^2 = 47762.32$ $\hat{\mu} = 69.1$ $s^2 = \sum \frac{x^2}{n-1} - \frac{\left(\sum x\right)^2}{n(n-1)}$ = 1.58 DF = 9 t value = 2.262 Standard error = $\frac{s}{\sqrt{n}} = \frac{\sqrt{1.58}}{\sqrt{10}}$ Confidence limits = $\bar{x} \pm t \times \frac{s}{\sqrt{n}}$ $= 69.1 \pm 2.262 \times \frac{\sqrt{1.58}}{\sqrt{10}}$ leading to [68.2,70.0]	 B1 M1 A1 B1 B1 B1 M1 A1 A1	 AO1 AO3 AO1 AO1 AO1 AO1 AO3 AO1 AO1	
(b)	The value of μ either lies in the interval or it does not, there is no question of a probability being involved. EITHER The confidence interval is an observed value of a random interval which contains μ with probability 0.95. OR If the process is carried out a large number of times, we would expect 95% of the confidence intervals obtained to contain μ.	E1 E1 **[11]**	AO2 AO2	

Qu. No.	Solution	Mark	AO	Notes
3(a)	H_0 : The petrol consumptions of models A and B are the same	B1	AO3	B0 for saying that the **mean** petrol consumption is the same
	H_1 : The petrol consumptions of models A and B are not the same	B1	AO3	For correctly identifying the alternative hypothesis as two-sided
(b)	From tables upper crit value = 31 Therefore lower crit value = 36 – 31 = 5 The critical region is $(U \geq 31) \cup (U \leq 5)$	B1 B1 B1	AO1 AO2 AO2	
(c)	Use of the formula $U = \sum \sum z_{ij}$ $U = 1 + 6 + 2 + 6 + 6 + 3$ $= 24$	M1 A1	AO3 AO1	
	The conclusion is that there is no difference in petrol consumption of the two models because 24 is not in the critical region.	B1 B1 **[9]**	AO3 AO2	
4(a)	$\hat{p} = \frac{1242}{1800} = 0.69$	B1	AO3	
	ESE = $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ $= \sqrt{\frac{0.69 \times 0.31}{1800}}$ $= 0.0109(0107\ldots)$	M1 A1	AO1 AO1	
	95% confidence limits are $\hat{p} \pm z \times \text{ESE}$ $0.69 \pm 1.96 \times 0.0109..$ giving [0.669,0.711]	M1 A1 A1	AO3 AO2 AO1	
(b)(i)	$\hat{p} = \frac{0.672 + 0.732}{2} = 0.702$ Number of people = 0.702 × 1000 = 702	B1 B1	AO3 AO1	
(ii)	$0.732 - 0.672 = 2z\sqrt{\frac{0.702 \times 0.298}{1000}}$ $z = 2.07417..$ Prob from tables = 0.98077 or 0.98097 from calc Confidence level = 96.2%	M1 A1 A1 A1 **[12]**	AO3 AO1 AO1 AO2	

Qu. No.	Solution	Mark	AO	Notes
5(a)	$H_0 : \mu_M = \mu_F ; H_1 : \mu_M \neq \mu_F$	B1	AO3	
(b)	Let X= male weight, Y=female weight ($\sum x = 39.2; \sum y = 46.6$) $\bar{x} = 4.9;$ $\bar{y} = 4.66$	 B1 B1	 AO1 AO1	
	SE of diff of means= $\sqrt{\frac{0.5^2}{8} + \frac{0.5^2}{10}}$ $= 0.237\ldots$	M1 A1	AO2 AO1	 Award m0 if no working seen
	Test statistic = $\frac{4.9 - 4.66}{0.237...}$ $= 1.01$ Prob from tables = 0.1562	m1 A1 A1	AO1 AO1 AO1	 From calculator, prob = 0.1558 FT 'their' test statistic From calculator, p-value = 0.3116
	p-value = 0.3124	B1	AO2	FT 'their' p-value
	Insufficient evidence to conclude that there is a difference in mean weight between males and females.	B1 **[10]**	AO3	
6(a)	The differences are 5 –2 8 10 –6 12 –4 7 9 1	B1	AO3	
	The signs may be omitted at this stage. The ranks are 4 2 7 9 5 10 3 6 8 1	M1 A1	AO3 AO1	Attempting to rank absolute values All correct
	W = Sum of positive ranks = 4 + 7 + 9 +10 + 6 + 8 + 1 = 45 The critical value is 44.	M1 A1 B1	AO3 AO1 AO1	
(b)	The conclusion at this significance level is that Method B gives on average a higher reading than Method A because 45 > 44	B1 E1 **[8]**	AO3 AO2	

Qu. No.	Solution	Mark	AO	Notes
7(a)	$E(X) = \theta + 3(1-3\theta) + 5\times 2\theta$ $= 2\theta + 3$ Var(X) = $\theta + 9(1-3\theta) + 25\times 2\theta - (2\theta+3)^2$ $= \theta + 9 - 27\theta + 50\theta - 4\theta^2 - 12\theta - 9$ $= 4\theta(3-\theta)$	M1 A1 M1 A1	AO1 AO1 AO2 AO2	
(b)(i)	Consider $E(V) = \dfrac{E(\bar{X})-3}{2}$ $= \dfrac{2\theta+3-3}{2} = \theta$ (Therefore V is unbiased)	M1 A1	AO2 AO2	
(ii)	Var(V) = $\dfrac{\text{Var}(\bar{X})}{4}$ $= \dfrac{\theta(3-\theta)}{n}$	M1 A1	AO3 AO1	
(c)	Y is B(n,θ) So $E(Y) = n\theta$ $E(W) = E\left(\dfrac{Y}{n}\right) = \theta$ (Therefore W is unbiased) Var(W) = $\dfrac{\text{Var}(Y)}{n^2}$ $= \dfrac{\theta(1-\theta)}{n}$	M1 A1 A1 M1 A1	AO3 AO2 AO2 AO2 AO1	
(d)	$\dfrac{\text{Var}(V)}{\text{Var}(W)} = \dfrac{\theta(3-\theta)}{n} \div \dfrac{\theta(1-\theta)}{n}$ $= \dfrac{(3-\theta)}{(1-\theta)}$ It follows that W is the better estimator since it has the smaller variance	M1 A1 B1 B1 **[17]**	AO3 AO1 AO2 AO2	