2014 Applied Mathematics - Statistics

Advanced Higher

Finalised Marking Instructions

Part One: General Marking Principles for Applied Mathematics – Statistics – Advanced Higher
This information is provided to help you understand the general principles you must apply when marking candidate responses to questions in this Paper. These principles must be read in conjunction with the specific Marking Instructions for each question.

(a) Marks for each candidate response must always be assigned in line with these general marking principles and the specific Marking Instructions for the relevant question.

(b) Marking should always be positive ie, marks should be awarded for what is correct and not deducted for errors or omissions.

GENERAL MARKING ADVICE: Applied Mathematics – Statistics – Advanced Higher

The marking schemes are written to assist in determining the "minimal acceptable answer" rather than listing every possible correct and incorrect answer. The following notes are offered to support Markers in making judgements on candidates' evidence, and apply to marking both end of unit assessments and course assessments.

These principles describe the approach taken when marking Advanced Higher Applied Mathematics papers. For more detailed guidance please refer to the detailed Marking Instructions.

1 The main principle is to give credit for the skills demonstrated and the criteria met. Failure to have a correct method may not preclude a candidate gaining credit for their solution.

2 The answer to one part of a question, even if incorrect, can be accepted as a basis for subsequent dependent parts of the question.

3 The following are not penalised:

- working subsequent to a correct answer (unless it provides firm evidence that the requirements of the question have not been met)
- legitimate variation in numerical values/algebraic expressions.

4 Full credit will only be given where the solution contains appropriate working. Where the correct answer might be obtained by inspection or mentally, credit may be given.

5 Sometimes the method to be used in a particular question is explicitly stated; no credit will be given where a candidate obtains the correct answer by an alternative method.

6 Where the method to be used in a particular question is not explicitly stated in the question paper, full credit is available for an alternative valid method. (Some likely alternatives are included but these should not be assumed to be the only acceptable ones.)

Part Two: Marking Instructions for each Question

Section A

Question			Expected Answer(s)	Max Mark	Additional Guidance
A	**1**	**(a)**	$P(S)$ $= P(\text{Pos}).P(S\|\text{Pos}) + P(\text{Neu}).P(S\|\text{Neu})$ $+P(\text{Neg}).P(S\|\text{Neg})$ $= 0\cdot 6\times 0\cdot 55 + 0\cdot 3\times 0\cdot 25 + 0\cdot 1\times 0\cdot 05$ $= 0\cdot 330 + 0\cdot 075 + 0\cdot 005$ $= 0\cdot 41$	**2**	
A	**1**	**(b)**	$P(\text{Pos}\|\bar{S}) = \dfrac{P(\bar{S}\cap \text{Pos})}{P(\bar{S})}$ $= \dfrac{P(\text{Pos}\cap \bar{S})}{1-P(S)}$ $= \dfrac{P(\text{Pos}).P(\bar{S}\|\text{Pos})}{1-P(S)}$ $= \dfrac{0\cdot 6\times 0\cdot 45}{0\cdot 59}$ $= \dfrac{27}{59}\ (0\cdot 46)$ **(other methods are acceptable)**	**4**	

Question			Expected Answer(s)	Max Mark	Additional Guidance
A	**2**	**(a)**	In order to take a simple random sample, a list of all sixth year pupils in Scottish schools would be required and such a list is not readily available.	**1**	
A	**2**	**(b)**	A simple random sample of the 376 schools would be taken. You could then visit each of the chosen schools to obtain the required information from sixth year pupils.	**2**	
A	**2**	**(c)**	A random sample of schools would be taken in each region. You could then visit each of the selected schools to obtain the required information from sixth year pupils.	**1**	
A	**2**	**(d)**	An advantage of cluster sampling is that it can reduce the cost of obtaining data through reducing travel costs for interviewers. An advantage of stratified sampling is that it yields more precise estimates.	**2**	

Question			Expected Answer(s)	Max Mark	Additional Guidance
A	**3**		Ho: there is no difference in population blood group distribution H1: there is a difference <table><tr><td>**Group**</td><td>O</td><td>A</td><td>B</td><td>AB</td></tr><tr><td>O_i</td><td>190</td><td>249</td><td>18</td><td>10</td></tr><tr><td>E_i</td><td>210·2</td><td>186·8</td><td>51·4</td><td>18·7</td></tr></table> $x^2 = \sum \frac{(O-E)^2}{E}$ $x^2 = \frac{(190-210\cdot 2)^2}{210\cdot 2} + \frac{(249-186\cdot 8)^2}{186\cdot 8} + \frac{(18-51\cdot 4)^2}{51\cdot 4} + \frac{(10-18\cdot 7)^2}{18\cdot 7}$ $= 1\cdot 93 + 20\cdot 71 + 21\cdot 68 + 4\cdot 03 = 48\cdot 35$ with 3 d.f. The critical value of chi-squared at the 0·1% level is 16·266. Since 48·35 exceeds this value Ho is rejected and the data provide very strong evidence that the blood group proportions amongst people of Hawaiian ethnic origin differ from those amongst the Caucasian segment of the USA population eg the Hawaiians have significantly more people in group A and less in the other three groups.	**6**	

Question			Expected Answer(s)	Max Mark	Additional Guidance
A	**4**		Rank sum for yields obtained using compost = 1 + 2 + 3 + 4 + 5+ 7·5 + 9·5 + 12 = 44 $W-\frac{1}{2}n(n+1)=44-36=8$ $\mathrm{P}(W-\frac{1}{2}n(n+1)\le 8)$ $=\frac{67}{12870}$ $=0\cdot 0052$ p-value = 2 × 0·0052 = 0·0104 0·0104 < 0·05 so we would reject Ho at the 5% level and conclude that there is evidence that the median yields differ.	**6**	
A	**5**	**(a)**	The assumption required is that the sample is random. $\hat{p}\pm 1\cdot 96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ $=0\cdot 33\pm 1\cdot 96\sqrt{\frac{0\cdot 33\times 0\cdot 67}{1004}}=(0\cdot 30, 0\cdot 36)$	**3**	
A	**5**	**(b)**	$f(x)=x(1-x)\Rightarrow f'(x)=1-2x=0$ when $x=0\cdot 5$ $M_{\max}=1\cdot 96\sqrt{\frac{0\cdot 25}{n}}$ $=\frac{1\cdot 96\times 0\cdot 5}{\sqrt{n}}\approx 1/\sqrt{n}$	**3**	
A	**5**	**(c)**	$M\le 0\cdot 01\Rightarrow \frac{1}{\sqrt{n}}\le 0\cdot 01$ $\Rightarrow n\ge \frac{1}{0\cdot 01^2}=10000$	**2**	

Question			Expected Answer(s)	Max Mark	Additional Guidance
A	**6**	**(a)**	$\bar{x} = 535\cdot5 \quad s = 5\cdot778$ $t = \dfrac{\bar{x}-\mu}{s/\sqrt{n}} = \dfrac{535\cdot5-540}{5\cdot778/\sqrt{10}} = -2\cdot46$ The critical region is $\lvert t\rvert > 2\cdot262$ Since the observed value of t lies in the critical region, the null hypothesis that μ = 540 would be rejected. The assumption required is that tensile strength is normally distributed.	**5**	
A	**6**	**(b)**	Since 540 does not lie in the interval the null hypothesis H_0:μ = 540 would be rejected at the 5% level of significance.	**2**	
A	**7**	**(a)**	$\mu_T = 30 \quad \sigma_X^2 = \dfrac{400}{3} \quad \mu_Y = 10 \quad \sigma_Y^2 = \dfrac{25}{3}$ $\mu_T = \mu_X + \mu_Y = 40$ $\sigma_T^2 = \sigma_X^2 + \sigma_Y^2 = \dfrac{425}{3} \Rightarrow \sigma_r \approx 11\cdot9$	**4**	
A	**7**	**(b)**	$60 = \mu_r + k\sigma_T \Rightarrow k = \dfrac{60-40}{11\cdot9} \approx 1\cdot68$ $P(\mathrm{T} \geq 60) \leq \dfrac{1}{1+1\cdot68^2} \approx 0\cdot26$	**2**	
A	**7**	**(c)**	If one takes longer to shop one might buy more goods and take longer to checkout. Plot a scattergraph of Y on X and check for a correlation.	**2**	

Question			Expected Answer(s)	Max Mark	Additional Guidance
A	**8**	**(a)**	P(Coin deemed biased|coin is fair) $= \mathrm{P}(X \geq 110)$ where $X \sim \mathrm{B}(200, 0\cdot5)$ $\approx Y \sim \mathrm{N}(100, 50)$ $\mathrm{P}(X \geq 110) \approx \mathrm{P}(Y \geq 109\cdot5)$ $= \mathrm{P}(Z \geq 1\cdot34)$ $= 0\cdot0901$	**6**	
A	**8**	**(b)**	P(Coin deemed fair|coin is biased) $= \mathrm{P}(X < 110 \mid X \sim B(200,\ 0\cdot55)$ $\approx \mathrm{P}(Y \leq 109\cdot5 \mid Y \sim N(110,\ 49\cdot5)$ $= \mathrm{P}(Z \leq -0\cdot07) = 0\cdot4721$	**3**	
A	**8**	**(c)**	The risks could be reduced by increasing the number of tosses from 200.	**1**	
A	**9**	**(a)**	For each additional crow fly mile one can expect to travel, on average, an additional 1·54 miles by road on the optimum route.	**1**	
A	**9**	**(b)**	An urban area has almost certainly got more roads.	**1**	

Question			Expected Answer(s)	Max Mark	Additional Guidance
A	9	(c)	$\hat{Y}_i = a + bx_i = -0\cdot59 + 1\cdot54 \times 40 = 61\cdot01$ The required prediction interval is given by $\hat{Y}_i \pm t_{\alpha/2,n-2} s\sqrt{1+\frac{1}{n}+\frac{(x_i-\bar{x})^2}{S_{xx}}}$ $= 61\cdot01 \pm 2\cdot069 \times \sqrt{4\cdot669} \times \sqrt{1+\frac{1}{25}+\frac{(40-10\cdot55)^2}{4251\cdot662}}$ $= 61\cdot01 \pm 4\cdot99 = (56\cdot0, 66\cdot0)$ The required confidence level is given by $\hat{Y}_i \pm t_{\alpha/2,n-2} s\sqrt{\frac{1}{n}+\frac{(x_i-\bar{x})^2}{S_{xx}}}$ $= 61\cdot01 \pm 2\cdot069 \times \sqrt{4\cdot669} \times \sqrt{\frac{1}{25}+\frac{(40-10\cdot55)^2}{4251\cdot662}}$ $= 61\cdot01 \pm 2\cdot21 = (58\cdot8, 63\cdot2)$ For a school with crow fly distance 40 miles, the PI indicates that one can be 95% confident that the optimum road distance will lie between 56·0 and 66·0 miles. The CI indicates that one can be 95% confident that mean optimum road distance for schools with crow fly distance 40 miles will lie between 58·8 and 63·2 miles.	7	
A	9	(d)	Since 0•297 exceeds 0•05 the implication is that the null hypothesis that α = 0 cannot be rejected at the 5% level of significance. The linear model could therefore be simplified to $Y_i = \beta x_i + \varepsilon_i \Rightarrow E(Y_i) = \beta x_i$ Thus a crow fly distance of 0 corresponds to an expected optimum road distance of 0 as would be anticipated.	2	

[*END OF SECTION A*]

Section B (Mathematics for Applied Mathematics)

Question			Expected Answer(s)	Max Mark	Additional Guidance
B	**1**		$y = 2x\sqrt{x-1}$ $\frac{dy}{dx} = 2x.\frac{d}{dx}\left(\sqrt{x-1}\right) + \sqrt{x-1} \times \frac{d}{dx}(2x)$ $= 2x.\frac{1}{2}(x-1)^{-\frac{1}{2}} + \sqrt{x-1} \times 2$ Gradient given by $\frac{dy}{dx}$ when $x = 10$, Gradient = $10.(9)^{-\frac{1}{2}} + \sqrt{9} \times 2$ $= \frac{28}{3}$	**4**	**1** product rule **1** first correct term **1** second correct term **1** evaluation (accept decimal equivalent to minimum of 3 sf)
B	**2**	**(a)**	$A + B = \begin{pmatrix} 4 & -7 & 6 \\ k-3 & 9 & -1 \\ 5 & 1 & 1 \end{pmatrix}$	**1**	**1** evaluation
B	**2**	**(b)**	$\det A = 1\begin{vmatrix} 0 & -1 \\ 3 & 0 \end{vmatrix} - 3\begin{vmatrix} k & -1 \\ 5 & 0 \end{vmatrix} + 4\begin{vmatrix} k & 0 \\ 5 & 3 \end{vmatrix}$ $= 1(0 + 3) - 3(0 + 5) + 4(3k - 0)$ $= 12k - 12$	**2**	**1** form of determinant 1 evaluation
B	**2**	**(c)**	$BC = \begin{pmatrix} 3 & -10 & 2 \\ -3 & 9 & 0 \\ 0 & -2 & 1 \end{pmatrix}\begin{pmatrix} 3 & 2 & -6 \\ 1 & 1 & -2 \\ 2 & 2 & -1 \end{pmatrix}$ $= \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$	**1**	**1** evaluation
B	**2**	**(d)**	$BC = 3I$. $B = 3C^{-1}$ or $C = 3B^{-1}$	**2**	**1** identity matrix connection or mention of inverse **1** relationship correct

Question			Expected Answer(s)	Max Mark	Additional Guidance
B	**3**		$I = \int x\sin 3x\,dx$ $u = x \qquad dv = \sin 3x$ $du = 1 \qquad v = \int \sin 3x$ $= \frac{-1}{3}\cos 3x$ $I = x.\frac{-1}{3}\cos 3x - \int 1.\frac{-1}{3}\cos 3x\,dx$ $= \frac{-x}{3}\cos 3x + \frac{1}{3}\int \cos 3x\,dx$ $= \frac{-x}{3}\cos 3x + \frac{1}{9}\sin 3x$ $I_0^{2\pi} = \left[\frac{-x}{3}\cos 3x + \frac{1}{9}\sin 3x\right]_0^{2\pi}$ $= \left[\frac{-2\pi}{3}\cos 6\pi + \frac{1}{9}\sin 6\pi\right] - \left[0 + \frac{1}{9}\sin 0\right]$ $= \frac{-2\pi}{3}$	**5**	**1** evidence of integration by parts **1** correct choice of u, dv **1** correct substitution **1** final integration correct 1 evaluation

Question			Expected Answer(s)	Max Mark	Additional Guidance
B	**4**		$\sum_{r=1}^{80} 3r^2 = 3\sum_{r=1}^{80} r^2$ using $\sum_{r=1}^{n} r^2 = \frac{n(n+1)(2n+1)}{6}$ * $3\sum_{r=1}^{80} r^2 = 3\left(\frac{80(81)(2\cdot 80+1)}{6}\right)$ $= 521\,640$	**2**	**1** correct substitution into * **1** evaluation (using incorrect formula – this mark available if of equivalent difficulty eg $\sum_{r=1}^{n} r^2 = \left(\frac{n(n+1)}{2}\right)^2$
B	**5**	**(a)**	$(e^x + 2)^4$ $= 1.(e^x)^4(2)^0 + 4(e^x)^3(2)^1 + 6(e^x)^2(2)^2$ $+ 4.(e^x)^1(2)^3 + 1.(e^x)^0(2)^4$ $= e^{4x} + 8e^{3x} + 24e^{2x} + 32e^x + 16$	**3**	Accept Binomial expansion ***or*** Pascal's Triangle **1** correct coefficients **1** correct powers of e^x and 2 **1** simplification
B	**5**	**(b)**	$\int (e^x + 2)^4 dx$ $= \int (e^{4x} + 8e^{3x} + 24e^{2x} + 32e^x + 16) dx$ $= \frac{e^{4x}}{4} = \frac{8e^{3x}}{3} + \frac{24e^{2x}}{2} + 32e^x + 16x + c$	**2**	**1** correct integration of composite fractions (at least one correct term involving exponential) **1** completion of integral (+ *c* not essential)

Question			Expected Answer(s)	Max Mark	Additional Guidance
B	**6**	**(a)**	10 000 people.	**1**	
B	**6**	**(b)**	$\frac{10000}{N(20000-N)}=\frac{A}{N}+\frac{B}{20000-N}$ 10 000 = $A(20\,000 - N) + BN$ $A=\frac{1}{2},\quad B=\frac{1}{2}$ Using $\frac{10000}{N(20000-N)}dN=dt$ gives $\frac{1}{2}\left(\frac{1}{N}+\frac{1}{20000-N}\right)dN=dt$ Integrating, $\int\left(\frac{1}{N}+\frac{1}{20000-N}\right)dN=\int 2dt$ $\ell nN-\ell n(20000-N)=2t+c$ $\ell n\frac{N}{20000-N}=2t+c$	**5**	**1** appropriate form of partial fractions **1** correct values of A and B **1** separate variables **1** starts integration eg $\int\frac{1}{N}dN$ correct **1** completes integration (moduli signs not required)

Question			Expected Answer(s)	Max Mark	Additional Guidance
B	**6**	**(c)**	Using $\ell n \dfrac{N}{20\,000-N} = 2t + c$ gives $\dfrac{N}{20\,000-N} = e^{2t+c}$ $\dfrac{N}{20\,000-N} = Ke^{2t}\left(\text{where } K = e^{c}\right)$ When $t = 0$, $N = 100$ $\dfrac{100}{19900} = K$ $K = \dfrac{1}{199}$ Hence $N = \left(20\,000 - N\right)\dfrac{e^{2t}}{199}$ $199N = (20\,000 - N)\,e^{2t}$ $N(199 + e^{2t}) = 20\,000e^{2t}$ $N = \dfrac{20\,000e^{2t}}{199 + e^{2t}}$	**4**	**1** accurately converts to exponential form (stating explicitly $K = e^{c}$ not required) **1** interprets initial condition **1** K valve **1** correctly gathers N terms

[*END OF SECTION B***]**

[*END OF QUESTION PAPER***]**