2007 Applied Mathematics

Advanced Higher – Statistics

Finalised Marking Instructions

General Marking Principles

These principles describe the approach taken when marking Advanced Higher 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.)

In the detailed Marking Instructions which follow, marks are shown alongside the line for which they are awarded. There are two codes used, M and E. M indicates a method mark, so in question B6, M1 means a method mark for using the partial fractions to work out the are. E is shorthand for error. For example, 2E1, means that a correct answer is awarded 2 marks but that 1 mark is deducted for each error.

Advanced Higher Applied Mathematics 2007
Section A – Statistics

A1. (a) These two probabilities do not take into account the gender profile of the population of junior executives. **1**

(b)
$$P(S \mid M) = \frac{P(M \cap S)}{P(M)} \qquad \mathbf{1}$$
$$= \frac{P(S \cap M)}{P(M)} = \frac{P(S)P(M \mid S)}{P(M)} \qquad \mathbf{1}$$
$$= \frac{0.75 \times 0.9}{0.8} = 0.844 \qquad \mathbf{1}$$
$$P(S \mid F) = \frac{P(S)P(F \mid S)}{P(F)} = \frac{0.75 \times 0.1}{0.2} = 0.375 \qquad \mathbf{1}$$

Since $P(S \mid M)$ is much greater than $P(S \mid F)$ there is evidence of possible discrimination against females for promotion. **1**

[Alternative methods such as Venn or tree diagrams are acceptable.]

A2. (a) Method A is a form of systematic sampling. **1**

(b) Method B is a form of stratified sampling. **1**

(c) Method A would be relatively quick to implement **1**
whereas Method B would be more time-consuming in
that mould numbers would have to be checked. **1**
Method B is to be recommended as it has the potential
to reveal problems associated with individual moulds. **1**

A3. (a) The critical region is $z > 2.33$ **1**

$\Rightarrow \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} > 2.33$ **1**

$\Rightarrow \bar{x} > \mu + 2.33\frac{\sigma}{\sqrt{n}}$

$\Rightarrow \bar{x} > 2250 + 103 = 2353$ **1**

(b) The required probability is:

$P(\bar{X} < 2353)$ **1**

$= P\left(Z < \frac{2353 - 2400}{140/\sqrt{10}}\right)$ **1**

$= P(Z < -1.06)$ **1**

$= 0.145$ **1**

A4. H_0 : Median (following refit) = 12
H_1 : Median (following refit) < 12 **1**

In the sample there are 11 scores less than 12, two equal to 12 and two greater than 12.

Thus the p-value is $F(2)$ **1**

for the Bin $(13, 0.5)$ distribution. **1**

p-value = $\left({}^{13}C_0 + {}^{13}C_1 + {}^{13}C_2\right)0.5^{13}$

$= (1 + 13 + 78)\,0.5^{13} = 0.011.$ **1**

Since the p-value is less than 0.05, the null hypothesis would be rejected, **1**

providing evidence that the new stabilisers have led to a reduction in the median score for the passengers' assessment of the motion experience during cruising. **1**

A5. (a) The assumptions are that the data are for a random sample from a normally distributed population. **1**

A 95% confidence interval for the mean is:–

$\bar{x} \pm t_{a/2,\,n-1}\dfrac{s}{\sqrt{n}}$ **1**

$\approx 146.6 \pm 2.101 \times \dfrac{17.9}{\sqrt{19}}$ **1, 1**

$= 146.6 \pm 8.6$ or $[138.0, 155.2]$ **1**

(b) Since the confidence interval does not include 137.4 **1**

the data provide evidence that mean systolic blood pressure in the population of diabetic men aged 50 to 59 differs significantly from the mean in the population of men aged 50 to 59 in good health. **1**

A6. (a) The standard deviation of a Poisson distribution is the square root of the mean i.e. 2 in this case. **1**

$P(X > \mu + 2\sigma) = P(X > 8)$ **1**

$= 1 - 0.9786 = 0.0214$ **1**

(b) The smallest integer k, such that $P(X \leqslant k) > 0.999$, is required. **1**

Hence $k = 11$. **1**

A7. (a) It is assumed that the errors ε_i are normally distributed with mean 0 and constant variance. **1**

$$s^2 = \frac{\text{SSR}}{n-2} = \frac{11581321}{38} = 552.06^2$$ **1**

$$t = \frac{b}{s/\sqrt{S_{xx}}} = \frac{28.19}{552.06/\sqrt{53637.186}} = 11.83$$ **1**

The critical region for a two-tailed t-test at the 0.1% level of significance with 38 df is $|t| > 3.566$. **1**

Since t lies in the critical region there is very strong evidence that the slope parameter is non-zero. **1**

[5% or 1% levels are acceptable.]

(b) Fitted value = 4103 + 28.19 × 144 = 8162 **1**

Residual = Observed value – Fitted value

= 8640-8162 = 478. **1**

(c) The non-linear appearance. **1**

(d) $y = 2305 + 64.8x - 0.1644x^2$

$\Rightarrow \frac{dy}{dx} = 64.8 - 0.3288x$ **1**

$\frac{dy}{dx} = 0 \Rightarrow x = \frac{64.8}{0.3288} \approx 197$ **1**

The plot indicates that an intake of around 200 appears to be optimal. **1**

A8. (a) Null hypothesis: - Consumption of sea-food and contraction of food-poisoning are independent.

Alternative hypothesis: - Consumption of sea-food and contraction of food-poisoning are not independent. **1**

(b)

	Well	Ill	
Ate	21	9	**30**
Did not	38	4	**42**
	59	**13**	**72**

Estimates are P(Ate) = 30/72 and P(Ill) = 13/72. **1,1**

(c) If the null hypothesis is true then the events "Ate seafood" and "Contracted food-poisoning" are independent so: –

Estimate of P(Both Ate and Ill) = P(Ate) × P(Ill) = 30/72 × 13/72 **1**

Hence the expected frequency of both Ate and Ill = 72 × 30/72 × 13/72. **1**

(d)

	Well	Ill	
Ate	21 24·58	9 5·42	**30**
Did not	38 34·42	4 7·58	**42**
	59	**13**	**72**

1

$$x^2 = \frac{(21 - 24.58)^2}{24.58} + \frac{(9 - 5.42)^2}{5.42} + \frac{(38 - 34.42)^2}{34.42} + \frac{(4 - 7.58)^2}{7.58}$$

$= 0.52 + 2.36 + 0.37 + 1.69$

$= 4.94$ with 1 degree of freedom **1,1**

The critical value of chi-squared is 3.84 with level of significance 5%. Since 4.94 exceeds 3.84 the test **1** provides evidence of an association between consumption of seafood and contraction of food-poisoning. **1**

A9. (a) $E(Y) = (b - a)E(X) + a$

$= \frac{1}{2}(b - a) + a = \frac{1}{2}(a + b)$ **1**

$V(Y) = (b - a)^2 V(X) + 0$ **1, 1**

$= \frac{1}{12}(b - a)^2$

(b) Total time = Set-up time + Build-time + Test-time **1**

E[Total time] = 12 + 90 + 18 = 120 **1**

Range = $(b - a)$ for the $U(a, b)$ distribution

V[Total time] $= 6^2/12 + 12^2/12 + 6^2/12$ **1**

$= 3 + 12 + 3 = 18$ **1**

The required standard deviation is $\sqrt{18} \approx 4.24$ **1**

(c) $\Phi(u) = 0.9 \Rightarrow u = 1.28$ **1**

$\Rightarrow \frac{T - 120}{4.24} = 1.28$ **1**

$\Rightarrow T = 120 + 1.28 \times 4.24 = 125.4$ **1**

Section B
Section B – Mathematics for Applied Mathematics

B1.

$$\int_0^{\pi/6} x \sin 3x \, dx = \left[x \int \sin 3x \, dx - \int 1.\left(-\frac{1}{3}\cos 3x\right)\right]_0^{\pi/6} \quad \textbf{2E1}$$

$$= \left[x\left(-\frac{1}{3}\cos 3x\right) + \frac{1}{9}\sin 3x \right]_0^{\pi/6} \quad \textbf{1}$$

$$= -\frac{\pi}{18}\cos\frac{\pi}{2} + \frac{1}{9}\sin\frac{\pi}{2} - (0 + 0) \quad \textbf{1}$$

$$= \frac{1}{9} \quad \textbf{1}$$

B2.

$$\left(x^3 - \frac{2}{x}\right)^4 = (x^3)^4 + 4(x^3)^3\left(-\frac{2}{x}\right) + 6(x^3)^2\left(-\frac{2}{x}\right)^2 + 4x^3\left(-\frac{2}{x}\right)^3 + \left(-\frac{2}{x}\right)^4 \quad \textbf{2E1}$$

$$= x^{12} - 8x^8 + 24x^4 - 32 + \frac{16}{x^4} \quad \textbf{2E1}$$

B3.

$$x = \frac{t}{t^2 + 1} \Rightarrow \frac{dx}{dt} = \frac{1(t^2 + 1) - t(2t)}{(t^2 + 1)^2} = \frac{1 - t^2}{(t^2 + 1)^2} \quad \textbf{2E1}$$

$$y = \frac{t - 1}{t^2 + 1} \Rightarrow \frac{dy}{dt} = \frac{1(t^2 + 1) - (t - 1)(2t)}{(t^2 + 1)^2} = \frac{1 + 2t - t^2}{(t^2 + 1)^2} \quad \textbf{2E1}$$

$$\Rightarrow \frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{\frac{1 + 2t - t^2}{(t^2 + 1)^2}}{\frac{1 - t^2}{(t^2 + 1)^2}} \quad \textbf{1}$$

$$= \frac{1 + 2t - t^2}{1 - t^2}$$

B4.

$$A = \begin{pmatrix} \lambda & 2 \\ 2 & \lambda - 3 \end{pmatrix}$$

$$\det A = \lambda(\lambda - 3) - 4 \quad \textbf{1}$$

A matrix is singular when its determinant is 0.

$$\lambda^2 - 3\lambda - 4 = 0 \quad \textbf{1}$$

$$(\lambda + 1)(\lambda - 4) = 0$$

$$\lambda = -1 \text{ or } \lambda = 4 \quad \textbf{1}$$

When $\lambda = 3$, $A = \begin{pmatrix} 3 & 2 \\ 2 & 0 \end{pmatrix}$ so $A^{-1} = \frac{1}{-4}\begin{pmatrix} 0 & -2 \\ -2 & 3 \end{pmatrix}$. **1**

B5.

$$x\frac{dy}{dx} - y = x^2e^x$$

$$\frac{dy}{dx} - \frac{1}{x}y = xe^x \qquad \mathbf{1}$$

Integrating factor:

$$\exp\left(\int \frac{-1}{x}\,dx\right) \qquad \mathbf{1}$$

$$= \exp(-\ln x) = x^{-1} \qquad \mathbf{1}$$

$$\frac{d}{dx}\left(\frac{y}{x}\right) = e^x$$

$$\frac{y}{x} = \int e^x dx = e^x + c \qquad \mathbf{1}$$

$$\therefore \quad y = x\left(e^x + c\right)$$

$y = 2$ when $x = 1 \Rightarrow 2 = e + c$

$$\Rightarrow y = x\left(e^x - e + 2\right) \qquad \mathbf{1}$$

B6.

$$\frac{8}{x(x+2)(x+4)} = \frac{A}{x} + \frac{B}{x+2} + \frac{C}{x+4} \qquad \mathbf{1}$$

$$8 = A(x+2)(x+4) + Bx(x+4) + Cx(x+2)$$

$$x = 0 \Rightarrow 8A = 8 \Rightarrow A = 1 \qquad \mathbf{1}$$

$$x = -2 \Rightarrow -4B = 8 \Rightarrow B = -2 \qquad \mathbf{1}$$

$$x = -4 \Rightarrow 8C = 8 \Rightarrow C = 1 \qquad \mathbf{1}$$

$$\frac{8}{x(x+2)(x+4)} = \frac{1}{x} + \frac{-2}{x+2} + \frac{1}{x+4}$$

$$\text{Area} = \int_1^2 \left(\frac{1}{x} + \frac{-2}{x+2} + \frac{1}{x+4}\right)dx \qquad \mathbf{1M}$$

$$= [\ln x - 2\ln(x+2) + \ln(x+4)]_1^2 \qquad \mathbf{1}$$

$$= \left[\ln\frac{x(x+4)}{(x+2)^2}\right]_1^2$$

$$= \ln\frac{12}{16} - \ln\frac{5}{9} = \ln\frac{12}{16} \times \frac{9}{5} = \ln\frac{27}{20} \qquad \mathbf{3E1}$$

[END OF MARKING INSTRUCTIONS]