2012 Applied Mathematics

Advanced Higher – Statistics

Finalised Marking Instructions

Advanced Higher Applied Mathematics 2012
Statistics Solutions

A1.

$$\mathrm{P}\left(R \mid Rf\right)$$

$$= \frac{\mathrm{P}(Rf \cap R)}{\mathrm{P}\left(Rf\right)} = \frac{\mathrm{P}(R \cap Rf)}{\mathrm{P}\left(Rf\right)}$$

$$= \frac{\mathrm{P}(R)\,\mathrm{P}\left(Rf \mid R\right)}{\mathrm{P}\left(R\right)\,\mathrm{P}\left(Rf \mid R\right) + \mathrm{P}\left(\overline{R}\right)\,\mathrm{P}\left(Rf \mid \overline{R}\right)}$$ **1**

$$= \frac{0{\cdot}2 \times 0{\cdot}9}{0{\cdot}2 \times 0{\cdot}9 + 0{\cdot}8 \times 0{\cdot}05}$$ **1,1**

$$= \frac{0{\cdot}18}{0{\cdot}18 + 0{\cdot}04} = \frac{9}{11}$$ **1**

{other methods acceptable}

A2. (a) Quota or convenience sampling. **1**

Telephone contact rules out certain members of the general public from inclusion in the sample. **1**

(b) Assuming that the sample may be regarded as a random one from the population, an approximate 95% confidence interval is **1**

$$p \pm 1{\cdot}96\sqrt{\frac{p(1 - p)}{n}}$$ **1**

where $p = \frac{539}{1013}$ and $n = 1013$ **1**

giving a 95% CI of $(0{\cdot}5014,\ 0{\cdot}5628)$. **1**

The lower confidence limit exceeds 50% so that the claim is supported. **1**

A3. Assuming that the weights of climbers and packs are independent we have: **1**

$$T = C_1 + C_2 + \ldots + C_8 + P_1 + P_2 + \ldots + P_8$$ **1**

$$\mu_T = 80 + 80 + \ldots + 30 + 30 + \ldots = 8 \times 80 + 8 \times 30 = 880$$ **1**

$$\sigma_T^2 = 16 + 16 + \ldots + 4 + 4 + \ldots = 8 \times 16 + 8 \times 4 \approx 12{\cdot}65^2$$ **1**

$$\mathrm{P}(T > 900)$$

$$= \mathrm{P}\left(Z > \frac{900 - 880}{12{\cdot}65}\right)$$ **1**

$$= \mathrm{P}\left(Z > 1{\cdot}58\right) = 0{\cdot}0571$$ **1**

A4. (a)

$$\bar{x} = 500{\cdot}265 \quad \textbf{1}$$

$$H_0 : \mu = 500{\cdot}30 \qquad H_1 : \mu \neq 500{\cdot}30 \quad \textbf{1}$$

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{500{\cdot}265 - 500{\cdot}30}{0{\cdot}1 / \sqrt{10}} \approx -1{\cdot}11 \quad \textbf{1}$$

Since $z > -1{\cdot}96$ we accept the null hypothesis **1**
and there is no evidence that the mean differs from 500.30ml **1**

(b) A t-test would be required if the fill volume standard deviation is unknown. **1**
This could change the conclusion since both the standard deviation and the critical value will be different. **1**

A5.

$$t_c = 2{\cdot}069 \quad \textbf{1}$$

$$|t| > 2{\cdot}069$$

$$\Rightarrow \frac{|r|}{\sqrt{\dfrac{1 - r^2}{23}}} > 2{\cdot}069 \quad \textbf{1}$$

$$\Rightarrow r^2 > 4{\cdot}281 \times \frac{1 - r^2}{23} \quad \textbf{1}$$

$$\Rightarrow 6{\cdot}373r^2 > 1$$

$$\Rightarrow r^2 > 0{\cdot}157 \quad \textbf{1}$$

$$\Rightarrow |r| > 0{\cdot}40$$

Display the data in a scatter plot as there could be **1**
a non-linear relationship between Y and X. **1**

A6. (a) H_0 : extinction counts follow a Poisson distribution.
H_1: they do not follow a Poisson distribution. **1**

(b)

$$P(X \leqslant 1) = f(0) + f(1)$$

$$= e^{-4{\cdot}21} + \frac{4{\cdot}21e^{-4{\cdot}21}}{1!} \quad \textbf{1}$$

$$= 0{\cdot}01485 + 0{\cdot}06250 = 0{\cdot}0773$$

Expected frequency = $76 \times 0{\cdot}0773 = 5{\cdot}88$ **1**

The amalgamation of two sets of frequencies is to comply with the guideline that around 80% of the expected frequencies should be more than 5. **1**

(c) The critical value of chi-squared with $9 - 1 - 1 = 7$ df
for a test at the 0·1% significance level is 24·321. **1**

Since 37·57 exceeds 24·321 the null hypothesis is rejected at the 0.1% level **1**
of significance so it may be concluded that there is strong evidence that extinctions may not be considered as random events in time. **1**

A7 (a) Slope $= \dfrac{S_{xy}}{S_{xx}} = \dfrac{995{\cdot}04}{1592{\cdot}89} \approx 0{\cdot}6247$ **1**

$\text{SSR} = S_{yy} - \dfrac{(S_{xy})^2}{S_{xx}} = 817{\cdot}67 - \dfrac{995{\cdot}04^2}{1592{\cdot}89} \approx 196{\cdot}09$ **1**

$\Rightarrow s^2 = \dfrac{196{\cdot}09}{7} = 5{\cdot}293^2$ **1**

$t = \dfrac{b}{\dfrac{s}{\sqrt{S_{xx}}}} = \dfrac{0{\cdot}6247}{\dfrac{5{\cdot}293}{\sqrt{1592{\cdot}89}}} \approx 4{\cdot}71$ **1**

$4{\cdot}71 > t_{7,\,0{\cdot}995} = 3{\cdot}499$ **1**

and so we would reject the null hypothesis that the slope is zero, at the 1% level. **1**

(b) $a = \bar{y} - b\bar{x} = 26{\cdot}656 - 0{\cdot}6247 \times 64{\cdot}11 \approx -13{\cdot}41$ **1**

$x = 71 \Rightarrow y = -13{\cdot}41 + 0{\cdot}6247 \times 71 \approx 30{\cdot}96$ **1**

(c) Construct a residual plot and check it out for random distribution etc. **1**

Find a prediction interval for the weight of the dog, **1**

A8. (a) The 3-sigma limits are $\mu \pm 3\dfrac{\sigma}{\sqrt{n}}$ **1**

$= 10 \pm 3 \times \dfrac{0{\cdot}2}{\sqrt{4}} = 10 \pm 0{\cdot}3$ i.e. $(9{\cdot}7$ and $10{\cdot}3)$ **1**

The probability that a point plots outwith a 3-sigma limit is

$\text{P}(Z < -3) + \text{P}(Z > 3) = 2\,\text{P}(Z > 3)$ **1**

$= 2(1 - 0{\cdot}9987) \approx 0{\cdot}0026$ **1**

(b) The probability that a point falls above a 2-sigma limit is

$\text{P}(Z > 2) \approx 0{\cdot}0228$ **1**

Since consecutive samples may be regarded as independent the binomial distribution gives the probability of two from three consecutive points above the upper limit to be

$\binom{3}{2}p^2q = \binom{3}{2}0{\cdot}0228^2 \times 0{\cdot}9772 \approx 0{\cdot}0015$ **1**

Doubling takes into account the identical probability of two from three consecutive points below the lower limit. **1**

$2 \times \binom{3}{2}0{\cdot}0228^2 \times 0{\cdot}9772 \approx 0{\cdot}0030$

which is of the same order of magnitue as $0{\cdot}0026$. **1**

(c) $\text{P}(Z > 1) = 0{\cdot}1587$ **1**

$2 \times \binom{5}{4}0{\cdot}1587^4 \times 0{\cdot}8413$ **1**

$= 0{\cdot}0053$ **1**

A9. $H_0 : \eta_W = \eta_{NW}$

$H_1 : \eta_W > \eta_{NW}$ **1**

Rank sum W

$= 1 + 3{\cdot}5 + 5 + 6 + 7 + 8 + 9 + 10 + 11 + 12 + 13 + 18$

$= 103{\cdot}5$ **1**

$\mathrm{E}(W) = \frac{1}{2}n(n + m + 1) = \frac{1}{2} \times 12 \times 25 = 150$

$\mathrm{V}(W) = \frac{1}{12}nm(n + m + 1) = \frac{1}{12} \times 144 \times 25 = 300$ **1**

$\mathrm{P}\left(W \leqslant 103{\cdot}5\right) = \mathrm{P}\left(Z \leqslant \dfrac{104 - 150}{\sqrt{300}}\right)$ **1,1**

$= \mathrm{P}\left(Z \leqslant -2{\cdot}66\right) = 0{\cdot}0039$ **1**

Since $0{\cdot}0039 < 0{\cdot}01$ we reject the null hypothesis at the 1% level. **1**

The null hypothesis $H_0 : \eta_W = \eta_{NW}$ is equivalent to $H_0 : \eta_W - \eta_{NW} = 0$. **1**

The fact that the 95% confidence interval does not include 0 confirms rejection of the null hypothesis at the 5% level of significance. **1**

The trial provides evidence that drinking water before food aids weight loss. **1**

END OF SECTION A

Section B

B1. The general term is given by

$$\binom{8}{r} x^{2(8-r)}(3x)^r \qquad \mathbf{1}$$

$$= \binom{8}{r} 3^r x^{16-r} \qquad \mathbf{1,1}$$

For x^{13},

$$16 - r = 13 \Rightarrow r = 3 \qquad \mathbf{1}$$

The corresponding coefficient is

$$\frac{8!}{3!\,5!} \times 3^3 = 1512 \qquad \mathbf{1}$$

{Note: some candidates may start from: $\binom{8}{r} x^{2r}(3x)^{8-r}$ leading to $r = 5$.}

B2. (a)

$$y = \frac{x}{x^2 + 4} \Rightarrow \frac{dy}{dx} = \frac{(x^2 + 4) - x.(2x)}{(x^2 + 4)^2} \qquad \mathbf{1M, 1}$$

$$x = 2 \Rightarrow \frac{dy}{dx} = \frac{8 - 8}{8^2} = 0. \qquad \mathbf{1}$$

(b)

$$\int e^{-2t}\,dt = \left(-\frac{1}{2}\right)e^{-2t} + c \qquad \begin{cases} \mathbf{1} \text{ for } \left(-\frac{1}{2}\right) \\ \mathbf{1} \text{ for } e^{-2t} \end{cases}$$

B3. (a)

$$M^2 = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & \lambda \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & \lambda \end{pmatrix} \qquad \mathbf{1M}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 6 & 1 & 0 \\ 0 & 0 & \lambda^2 \end{pmatrix} \qquad \mathbf{1}$$

(b)

$$M^3 = \begin{pmatrix} 1 & 0 & 0 \\ 6 & 1 & 0 \\ 0 & 0 & \lambda^2 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & \lambda \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 9 & 1 & 0 \\ 0 & 0 & \lambda^3 \end{pmatrix} \qquad \mathbf{1}$$

$$M + M^2 + M^3 = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & \lambda \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 6 & 1 & 0 \\ 0 & 0 & \lambda^2 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 9 & 1 & 0 \\ 0 & 0 & \lambda^3 \end{pmatrix}$$

$$= \begin{pmatrix} 3 & 0 & 0 \\ 18 & 3 & 0 \\ 0 & 0 & \lambda + \lambda^2 + \lambda^3 \end{pmatrix} \qquad \mathbf{1}$$

(c)

$$\det M = 1 \times (1 \times \lambda) + 0 + 0 = \lambda \qquad \mathbf{1}$$

Hence the matrix M has an inverse when $\lambda \neq 0$. **1**

B4.

$$\frac{1}{x(x+1)} = \frac{A}{x} + \frac{B}{x+1}$$ **1M**

$$1 = A(x+1) + Bx$$

$$x = 0 \quad \Rightarrow \quad A = 1$$ **1**

$$x = -1 \quad \Rightarrow \quad B = -1$$ **1**

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

$$V = \int \pi y^2 dx \Rightarrow V = \pi \int_1^3 \left(\frac{1}{\sqrt{x^2+x}}\right)^2 dx$$ **1M**

$$= \pi \int_1^3 \left(\frac{1}{x} - \frac{1}{x+1}\right) dx$$ **1**

$$= \pi [\ln x - \ln(x+1)]_1^3$$ **1**

$$= \pi \{[\ln 3 - \ln 4] - [\ln 1 - \ln 2]\}$$

$$= \pi \ln \frac{3}{2} (\approx 1.274 \text{ to 3 s.f.})$$ **1**

B5. (a)

$$\frac{dT}{dx} = k(180 - T)$$

$$\int \frac{dT}{180 - T} = \int k\, dx$$ **1M**

$$-\int \frac{(-1)}{180 - T} dT = \int k\, dx$$

$$-\ln(180 - T) = kx + c$$ **1**

Since $T = 4$ when $x = 0$

$$-\ln 176 = c$$ **1**

$$\Rightarrow \ln(180 - T) - \ln 176 = -kx$$

$$\ln \frac{180 - T}{176} = -kx$$

$$\frac{180 - T}{176} = e^{-kx}$$

$$180 - T = 176e^{-kx}$$ **1**

$$\text{i.e. } T = 180 - 176e^{-kx}.$$

(b) When $x = 1, T = 30$

$$e^{-k} = \frac{150}{176}$$ **1**

$$\Rightarrow k \approx 0{\cdot}16$$ **1**

(c) Using $k = 0{\cdot}16$ and $T = 80$ in $T = 180 - 176e^{-kx}$ gives

$$80 = 180 - 176e^{-0{\cdot}16x}$$ **1**

Hence

$$e^{-0{\cdot}16x} = \frac{100}{176}$$

$$\Rightarrow -0{\cdot}16x = \ln \frac{100}{176}$$ **1**

$$\Rightarrow x \approx 3.533 \text{ hours} \approx 212 \text{ minutes}$$ **1**

So the turkey should be cooked after 3 hours 32 minutes (or 212 minutes).

END OF MARKING INSTRUCTIONS