SCOTTISH
QUALIFICATIONS
AUTHORITY

2004 Applied Mathematics

Advanced Higher – Section A

Finalised Marking Instructions

Advanced Higher Applied Mathematics 2004
Solutions for Section A (Statistics 1 and 2)

A1. (a) Stratified — **1**
and Quota [or Quota (convenience)] — **1**
(b) Approach (a) should be best — **1**
since (b) is not random (other forms e.g. Glasgow not typical, biased) — **1**

A2. (a) $F \sim Bin(192, 0.002)$. — **1 for distribution**, **1 for parameters**

(b) $P(F \geqslant 3) = 1 - P(F \leqslant 2)$ — **1**
$= 1 - (0.6809 + 0.2620 + 0.0501)$ — **1**
$= 0.0070$ — **1**

Notes: applying a Poisson distribution loses (at least) one mark; a Normal distribution loses two marks.

(c) Approximate using the $Poi(0.384)$ — **1 for distribution**, **1 for parameters**

A3. Assume that yields are normally distributed . — **1**
[Random or independent will not do.]
$\bar{x} = 404.2;\ s = 10.03$ — **1**
$t = 2.776$ — **1**
A 95% confidence interval for the mean yield, μ, is given by:–
$$\bar{x} \pm t \frac{s}{\sqrt{n}}$$ **1**
$$404.2 \pm 2.776 \frac{10.03}{\sqrt{5}}$$
404.2 ± 12.45 — **1**
or (391.75, 416.65).
The fact that the confidence interval does not include 382 provides *evidence*, at the 5% level of significance, of a change in the mean yield. (Stating it *is* changed loses one mark.) — **1**
Note: the third and fourth marks are lost if a z interval is used.

A4. TNE = 3% of 500 = 15 — **1**
With maximum allowable standard deviation
$P(\text{weight} < 485) = 0.025$ — **1**
$$\Rightarrow \frac{485 - 505}{\sigma} = -1.96$$ **1,1**
$$\Rightarrow \sigma = \frac{20}{1.96} = 10.2$$ **1**

There will be a small probability of obtaining a content weight less than 470g with the normal model. — **1**

A5. Assume that the distributions of times Before and After have the same shape. 1

Notes: a Normal distribution with the same shape is a valid comment.
Independent, random, Normal (without shape) are not valid.

Null hypothesis H_0: Median After = Median Before
Alternative hypothesis H_1: Median After < Median Before 1

Time	19	29	31	35	37	39	39	41	42	43	45	52	59	64
Period	A	A	B	A	A	B	A	A	B	B	A	B	B	B
Rank	1	2	3	4	5	6.5	6.5	8	9	10	11	12	13	14

Rank sum for After times = 37.5 1

$W - \frac{1}{2}n(n+1) = 37.5 - 28 = 9.5$

$P(W - \frac{1}{2}n(n+1) < 10)$

$= \dfrac{125}{3432}$

$= 0.036$ 1

Since this value is *less than* 0.05 the null hypothesis
would be rejected in favour of the alternative, 1
indicating evidence of improved performance. 1

Notes:
As the computed value, 9.5, is not in the tables, a range of values for the probability was acceptable.
A Normal approximation was accepted.

A6.

Cream	A	B	C
Obs. No. of purchasers	66	99	75
Exp. No. of purchasers	80	80	80

1

$$X^2 = \sum \frac{(O-E)^2}{E}$$

$$= \frac{(66-80)^2}{80} + \frac{(99-80)^2}{80} + \frac{(75-80)^2}{80}$$

$= 2.45 + 4.5125 + 0.3125 = 7.275$ 1

with 2 d.f. 1

The critical value of chi-squared at the 5% level is 5.991 1
so the null hypothesis would be rejected.
i.e. there is *evidence* of a preference. 1

The fact that the p-value is less than 0.05 confirms 1
rejection of the null hypothesis at the 5%level of significance.
Note: using a two-tail test loses a mark.

A7. (a) The fitted value is 13.791 **1**
with residual 10.209. **1**

(b) The wedge-shaped plot casts doubt on the assumption of constant variance of Y_i. (i.e. variance not constant) **1**

(c) Satisfactory now since *variance* seems to be *more constant*. **1**
Note: A phrase such as 'more randomly scattered' is acceptable.

(d) The residuals are normally distributed. **1**

A8. (a)

Pre	36	45	30	63	48	52	44	44	45	51	39	44
Post	39	42	33	70	53	51	48	51	51	51	42	50
Post – Pre	3	−3	3	7	5	−1	4	7	6	0	3	6
Sign	1	−1	1	1	1	−1	1	1	1	0	1	1

Assume that differences are independent. **1**
H_0: Median (Post – Pre) = 0 [or $\eta_d = 0$]
H_1: Median (Post – Pre) > 0 [or $\eta_d > 0$] **1**
Under H_0 the differences Bin (11,0.5) with $b = 2$. **1**
$P(B \leqslant 2) = (C_0^{11} + C_1^{11} + C_2^{11})0.5^{11}$
$= (1 + 11 + 55)0.5^{11} = 0.0327$. **1**

Since 0.0327 < 0.05 the null hypothesis is rejected and there **1**
is *evidence* that the median PCS-12 score has gone up. **1**

Note: applying a two-tailed test loses a mark.

(b) $H_0 : \mu_{\text{Post}} = 50$
$H_1 : \mu_{\text{Post}} \neq 50$ **1**
$\bar{x} = 48.42$

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{48.42 - 50}{10/\sqrt{12}} = -0.55.$$ **1,1**

The critical region is $|z| > 1.96$ at the 5% level of significance. **1**
Since −0.55 is not in the critical region, the null hypothesis is accepted indicating that the Post-operation scores are consistent with a population mean of 50. **1**

Note: a correct use of probability comparisons gets full marks.

A9. (a) P (Alaskan fish classified as Canadian)

$= P(X > 120 \mid X \sim N(100, 20^2))$ **1**

$= P\left(Z > \dfrac{120 - 100}{20}\right)$

$= P(Z > 1)$ **1**

$= 0.1587$ **1**

(b) The probability is the same as in (a) because of symmetry. **1**

(c) P (Canadian origin | Alaskan predicted)

$= \dfrac{\text{P (Alaskan predicted and Canadian origin)}}{\text{P (Alaskan predicted)}}$ **1**

$= \dfrac{\text{P (Alaskan predicted but Canadian origin)}}{\text{P (Ala pred and Alaskan)} + \text{P (Ala pred but Canadian)}}$ **1**

$= \dfrac{0.4 \times 0.1587}{0.6 \times 0.8413 + 0.4 \times 0.1587}$ **1,1**

$= \dfrac{0.06348}{0.50478 + 0.06348}$

$= 0.112.$ **1**

Note: Alternative methods acceptable e.g. Venn or Tree Diagrams

A10. The number, X, of inaccurate invoices in samples of n will have the Bin (n, p) distribution so

$V(X) = npq$ **1**

$\quad = np(1 - p)$

$\Rightarrow V(\text{Proportion}) = V\left(\frac{1}{n}X\right) = \dfrac{1}{n^2} V(X)$ **1**

$\quad = \dfrac{p(1 - p)}{n}$ **1**

$\Rightarrow$ Standard deviation of Proportion $= \sqrt{\dfrac{p(1 - p)}{n}}$.

(a) $\text{UCL} = p + 3\sqrt{\dfrac{p(1 - p)}{n}}$

$= 0.12 + 3\sqrt{\dfrac{0.12 \times 0.88}{150}}$

$= 0.12 + 0.08 = 0.20.$ **1**

$\text{LCL} = 0.12 - 0.08 = 0.04$ **1**

(b) The fact that the point for Week 30 falls below the lower chart limit provides evidence of a drop in the proportion of inaccurate invoices. **1**
or: 8 consecutive points fell below the centre line.

(c) A new chart should be constructed (or set new limits) **1**
using an estimate of p for calculation of limits which is based on data collected since the process change. **1**