OCR
RECOGNISING ACHIEVEMENT

General Certificate of Education
Advanced Supplementary (AS) and Advanced Level
former Oxford and Cambridge modular syllabus

MEI STRUCTURED MATHEMATICS 5518

Statistics 6

Tuesday **12 JUNE 2001** Afternoon 1 hour 20 minutes

Additional materials:
  Answer paper
  Graph paper
  Students' Handbook

TIME 1 hour 20 minutes

INSTRUCTIONS TO CANDIDATES

Write your name, Centre number and candidate number in the spaces provided on the answer paper/ answer booklet.

Answer any **three** questions.

Write your answers on the separate answer paper provided.

If you use more than one sheet of paper, fasten the sheets together.

INFORMATION FOR CANDIDATES

The approximate allocation of marks is given in brackets [ ] at the end of each question or part question.

You are advised that an answer may receive no marks unless sufficient detail of the working is shown on the answer paper to indicate that a correct method is being used.

This question paper consists of 6 printed pages and 2 blank pages.

[Turn over

Option 1: Estimation

1 [In parts (i) and (ii) of this question, you are *not* required to verify that any turning-point you find is a maximum.]

The safety officer of a factory has undertaken a survey of the numbers of accidents occurring per week over a period of n weeks. The number of accidents in week i is denoted by x_i (for $i = 1, 2, \ldots, n$). Each week is assumed to be independent of all other weeks in respect of numbers of accidents. It is also assumed that the random variable X underlying the x_i has the Poisson distribution with parameter θ, so that

$$\mathrm{P}(X = x) = \frac{\mathrm{e}^{-\theta}\theta^{x}}{x!} \quad \text{for } x = 0, 1, 2, \ldots.$$

(i) Write down the likelihood of the set of the x_i. Hence find the maximum likelihood estimate $\hat{\theta}$ of θ. [5]

(ii) The probability that there are no accidents in a week is $\phi = \mathrm{e}^{-\theta}$. Rewrite the likelihood as a function of ϕ instead of θ. Hence show that the maximum likelihood estimate of ϕ is

$$\hat{\phi} = \mathrm{e}^{-\bar{x}}$$

where $\bar{x}$ is the average number of accidents per week in the survey. [7]

(iii) Comment on the relationship between $\hat{\theta}$ and $\hat{\phi}$. [1]

(iv) It may be shown that, when n is large, $\hat{\phi}$ is unbiased and has underlying variance

$$\frac{-\phi^2 \ln\phi}{n}.$$

An alternative estimate of ϕ is $\tilde{\phi} = r/n$ where r is the number of weeks with no accidents; it may be shown that this also is unbiased and has underlying variance

$$\frac{\phi(1-\phi)}{n}.$$

By considering the difference

$$1 - \phi - (-\phi\ln\phi)$$

and showing that this has its minimum at $\phi = 1$, deduce that, when n is large, $\hat{\phi}$ is to be preferred to $\tilde{\phi}$. [7]

Option 2: Bivariate distributions

2 [Numerical answers in this question should be given as fractions in their lowest terms.]

X and Y are discrete random variables whose joint distribution is given in the table.

		values of Y		
		1	2	3
	1	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{4}$
values of X	2	$\frac{1}{8}$	$\frac{1}{24}$	0
	3	$\frac{1}{12}$	$\frac{1}{24}$	$\frac{1}{24}$

(i) Find the marginal distribution of X and its mean and variance. [5]

(ii) Find the marginal distribution of Y. [1]

(iii) Find the conditional distribution of X given that $Y = 1$; hence find $\mathrm{E}(X \mid Y = 1)$, i.e. the conditional mean of X given that $Y = 1$.

Similarly find $\mathrm{E}(X \mid Y = 2)$ and $\mathrm{E}(X \mid Y = 3)$. [8]

(iv) Consider the quantity $\mathrm{E}(X \mid Y)$ regarded as a function of the random variable Y, say $\mathrm{g}(Y)$. In part **(iii)**, you calculated the value of this function for each of the three possible values of Y. In part **(ii)**, you calculated the probability that Y takes each of its possible values. Use these results to calculate $\mathrm{E}(\mathrm{g}(Y))$. [2]

(v) Now re-write $\mathrm{g}(Y)$ as $\mathrm{E}(X \mid Y)$ and let $\mathrm{E}_Y(\mathrm{E}(X \mid Y))$ denote the expectation you found in part **(iv)**. Verify that $\mathrm{E}_Y(\mathrm{E}(X \mid Y)) = \mathrm{E}(X)$.

Given that

$$\mathrm{E}\left(X^2 \mid Y = 1\right) = \tfrac{34}{9}, \qquad \mathrm{E}\left(X^2 \mid Y = 2\right) = \tfrac{19}{8}, \qquad \mathrm{E}\left(X^2 \mid Y = 3\right) = \tfrac{15}{7},$$

show similarly that $\mathrm{E}_Y(\mathrm{E}(X^2 \mid Y)) = \mathrm{E}(X^2)$. [4]

Option 3: Markov chains

3 A supermarket chain stocks three brands (T, U and V) of a particular canned food and uses a simple Markov chain model of customer behaviour. Brand T is the best known, and at each purchase a customer who bought brand T last time has probability $\frac{3}{4}$ of buying it again; otherwise, the customer is equally likely to buy brand U or brand V. For each of brands U and V, a customer who bought the brand last time has probability $\frac{1}{2}$ of doing so again, probability $\frac{1}{3}$ of switching to brand T and probability $\frac{1}{6}$ of switching to the remaining brand.

Write down the transition matrix for the Markov chain model. [4]

Find the long-run proportion of purchases (the "market share") for each of the brands. [6]

A customer bought brand T last time. Find the expected number of consecutive further occasions on which this customer will purchase brand T, *showing your working clearly.* [5]

For the steady state, find the expected number of consecutive further occasions on which a randomly chosen customer will purchase the brand bought last time. [5]

Option 4: Analysis of variance

4 **(i)** State the usual one-way analysis of variance model, including the assumptions about the experimental error, for a situation having k treatments with n_i observations on the ith treatment, with x_{ij} denoting the jth observation on the ith treatment. Interpret the parameters in the model. [5]

(ii) A firm making agricultural products is carrying out a trial of four fertilizers for growing wheat. Several plots of soil are prepared, and the trial is carried out under carefully controlled conditions. The yields, in kg per plot, at the end of the trial are as follows.

Fertilizer A	33.9	33.6	35.4	32.8
Fertilizer B	37.2	36.4	35.5	33.8
Fertilizer C	33.1	31.7	33.7	
Fertilizer D	39.8	38.4	38.6	37.1

[For information, the sum of these data items is 531.0 and the sum of their squares is 18 879.82.]

Carry out the customary one-way analysis of variance to examine whether there appear to be differences among the fertilizers. Use a 1% level of significance. Display your working clearly. [10]

(iii) Let W denote the within-samples sum of squares,

$$W = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(x_{ij} - \bar{x}_i\right)^2,$$

and let s_i^2 denote the sample variance for the ith sample,

$$s_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} \left(x_{ij} - \bar{x}_i\right)^2,$$

where $\bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}$. Also let σ^2 denote the common population variance underlying all the observations. Use the result that the underlying distribution of s_i^2 is $\frac{\sigma^2}{n_i - 1} \chi^2_{n_i - 1}$ to deduce that $\sum_{j=1}^{n_i} \left(x_{ij} - \bar{x}_i\right)^2$ has underlying distribution $\sigma^2 \chi^2_{n_i - 1}$.

Hence deduce the underlying distribution of W, using the following result: the random variable obtained as the sum of independent random variables each having χ^2 distributions has itself a χ^2 distribution, and its number of degrees of freedom is the sum of the numbers of degrees of freedom of the random variables in the sum. [5]

[Turn over

Option 5: Regression

5 A weighing machine is thought to be subject to a constant (but unknown) additive bias δ. There are also the customary experimental errors when using the machine so, when an object of (unknown) true weight θ is weighed on the machine, the observed weight Y is a random variable given by

$$Y = \theta + \delta + e$$

where e represents the experimental error, assumed to have zero mean and variance σ^2.

Three parcels have to be weighed on the machine. Their unknown true weights are $\theta_1, \theta_2, \theta_3$. As it is necessary to estimate these weights and the bias δ, four independent observations Y_1, Y_2, Y_3, Y_4 will be made. Two weighing schemes are proposed. In both cases, the first three observations consist of weighing each parcel by itself.

(i) Suppose that the fourth observation consists of a weighing taken with no parcels. Thus

$$Y_i = \theta_i + \delta + e_i \qquad \text{for } i = 1, 2, 3,$$
$$Y_4 = \delta + e_4.$$

Show that one of the normal equations for least-squares estimators of $\theta_1, \theta_2, \theta_3$ and δ is

$$Y_1 - \theta_1 - \delta = 0.$$

Obtain the other normal equations. Solve these equations. (You are not required to prove that your solutions give a minimum.) Obtain the variances of the estimators. [8]

(ii) Suppose now that the fourth observation consists of all the parcels weighed together. Obtain and solve the normal equations for least-squares estimators of $\theta_1, \theta_2, \theta_3$ and δ in this case (again you are not required to prove that your solutions give a minimum). [*Hint.* You might find it helpful in solving the equations to note that the solutions satisfy $\hat{\delta} + \hat{\theta}_1 + \hat{\theta}_2 + \hat{\theta}_3 = Y_4$.]

Obtain the variances of the estimators in this case. [8]

(iii) Discuss which of the two weighing schemes is better. [4]

Mark Scheme

Q.1	(i) $L = \frac{e^{-\theta}\theta^{x_1}}{x_1!} \times \frac{e^{-\theta}\theta^{x_2}}{x_2!} \times \dots \times \frac{e^{-\theta}\theta^{x_n}}{x_n!}$	M1 1	General product form Fully correct	
	$= \frac{e^{-n\theta}\theta^{\Sigma x_i}}{x_1!x_2!\dots x_n!}$			
	$\ell n L = -n\theta + \Sigma x_i \ell n\theta$ (+ constant)	M1		
	$\frac{\partial \ell n L}{\partial \theta} = -n + \frac{\Sigma x_i}{\theta} = 0$	M1		
	$\underline{\underline{\rightarrow \hat{\theta} = \bar{x}}}$	1		5
	(ii) $\phi = P(X = 0) = e^{-\theta}$			
	$L = \frac{\varphi^n(-\ell n\varphi)^{\Sigma x_i}}{x_1!x_2!\dots x_n!}$	1		
	$\ell n L = n\,\ell n\phi + \Sigma x_i \ell\, n(-(-\ell n\phi)$ (+ constant)	2		
	$\frac{\partial \ell n L}{\partial \varphi} = \frac{n}{\varphi} + \Sigma x_i\, \frac{1}{-\ell n\varphi} \cdot \frac{-1}{\varphi}$	M1, 1		
	$= 0 \rightarrow n\,\ell n\phi + \Sigma x_i = 0$	1		
	$\rightarrow \ell n\hat{\varphi} = -\bar{x},\ \underline{\underline{\hat{\varphi} = e^{-\bar{x}}}}$	1		7
	(iii) We have $\widehat{e^{-\theta}} = e^{-\hat{\theta}}$	1	Award for any equivalent statement/explanation	1
	(iv) $\mathrm{Var}(\tilde{\varphi}) - \mathrm{Var}(\hat{\varphi}) = \frac{\varphi(1-\varphi)}{n} - \frac{-\varphi^2 \ell n\varphi}{n}$			
	$\therefore$ Consider $y = 1 - \phi + \phi\,\ell n\phi$. We have $\frac{dy}{d\varphi} =$ $-1 + \varphi . \frac{1}{\varphi} + \ell n\varphi = 0 \rightarrow \ell n\varphi = 0$ i.e. $\varphi = 1$	1		
	Also $\frac{d^2y}{d\varphi^2} = \frac{1}{\varphi} > 0$	1		
	$\therefore\ \phi = 1$ is a minimum	1		
	and this gives $y = 0$	1		
	$\therefore\ y > 0$ for all other ϕ (in particular, for $0 < \phi < 1$)	1		
	$\therefore\ \mathrm{Var}(\tilde{\varphi}) > \mathrm{Var}(\hat{\varphi})$	1		
	so $\hat{\varphi}$ is preferred	1		7

Q.2	(i) and (ii) $\begin{array}{cc\|ccc\|c} & Y\to & 1 & 2 & 3 & \\ \hline & 1 & \frac{1}{6} & \frac{1}{4} & \frac{1}{4} & \frac{2}{3} \\ X & 2 & \frac{1}{8} & \frac{1}{24} & 0 & \frac{1}{6} \\ & 3 & \frac{1}{12} & \frac{1}{24} & \frac{1}{24} & \frac{1}{6} \\ \hline & & \frac{3}{8} & \frac{1}{3} & \frac{7}{24} & 1 \end{array}$	M1 } A1 } For marg (*X*) A1 For marg (*Y*)	FT THROUGHOUT, but A0 for Var(*x*) if negative. Accept fractions not in lowest terms, but **DEDUCT 1 FROM TOTAL** if this has been done.	
	$E(X)=1\times\frac{2}{3}+2\times\frac{1}{6}+3\times\frac{1}{6}=\frac{3}{2}$	A1		
	$E(X^2)=1^2\times\frac{2}{3}+2^2\times\frac{1}{6}+3^2\times\frac{1}{6}=\frac{17}{6}$	A1		
	$\therefore \text{Var}(X)=\frac{17}{6}-\left(\frac{3}{2}\right)^2=\frac{2}{12}$	A1		6
	(iii) $\begin{array}{cccc} & P(X=x\mid Y=1) & P(X=x\mid Y=2) & P(X=x\mid Y=3) \\ x=1 & \frac{4}{9} & \frac{3}{4} & \frac{6}{7} \\ 2 & \frac{1}{3} & \frac{1}{8} & 0 \\ 3 & \frac{2}{9} & \frac{1}{8} & \frac{1}{7} \end{array}$	M1 A1 A1 A1	Award once	
	$E(X\mid Y=1)=\frac{16}{9}$ $E(X\mid Y=2)=\frac{11}{8}$ $E(X\mid Y=3)=\frac{9}{7}$	M1 A1 A1 A1	Award once	8
	(iv) $E[g(Y)]\ (\equiv E[E[X\mid Y])$ $=\frac{16}{9}\times\frac{3}{8}+\frac{11}{8}\times\frac{1}{3}+\frac{9}{7}\times\frac{7}{24}=\frac{18}{12}=\frac{3}{2}$	M1 A1		2
	(v) answer to (iv) = $E(X)$ as found earlier	1		
	$E\left[E\left(X^2\mid Y\right)\right]=\frac{34}{9}\times\frac{3}{8}+\frac{19}{8}\times\frac{1}{3}+\frac{15}{7}\times\frac{7}{24}=\frac{68}{24}=\frac{17}{6}$	M1 A1		
	= $E(X^2)$ as earlier	1		4

Q.3	Transition matrix $\mathbf{P} = \begin{array}{cc} & \begin{array}{ccc} \text{T} & \text{U} & \text{V} \end{array} \\ \begin{array}{c} \text{T} \\ \text{U} \\ \text{V} \end{array} & \begin{bmatrix} \frac{3}{4} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \end{bmatrix} \end{array}$	A4		4
	$\pi = \pi\mathbf{P}$ with $\Sigma\pi_i = 1$ $\left.\begin{array}{l} \pi_1 = \frac{3}{4}\pi_1 + \frac{1}{3}\pi_2 + \frac{1}{3}\pi_3 \\ \pi_2 = \frac{1}{8}\pi_1 + \frac{1}{2}\pi_2 + \frac{1}{6}\pi_3 \\ \pi_3 = \frac{1}{8}\pi_1 + \frac{1}{6}\pi_2 + \frac{1}{2}\pi_3 \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{array}\right\}$ Solutions are $\pi_1 = \frac{4}{7}(0.5714)$ $\pi_2 = \frac{3}{14}(0.2143)$ $= \pi_3$	M2 M1 A1 A1 A1 A1 A1	If solutions wrong, allow A1 (out of 3) if they add to 1	6
	Want $1 \times \frac{3}{4} \times \frac{1}{4} + 2 \times \left(\frac{3}{4}\right)^2 \times \left(\frac{1}{4}\right) + 3 \times \left(\frac{3}{4}\right)^3 \times \left(\frac{1}{4}\right) + \ldots$ $= \frac{1}{4}\left\{\frac{3}{4} + 2 \times \left(\frac{3}{4}\right)^2 + 3 \times \left(\frac{3}{4}\right)^3 + \ldots\right\} = 3$	M2 M2 A1	METHOD for summing series must be CLEAR (e.g. use of "GP of GPs"). Do not accept write-down of answer, or of statements such as "$\alpha/(1-\alpha)$".	5
	Previous answer is E[consecutive further purchases of T] Similarly, E[consecutive further purchases of U] $\left(= \frac{\frac{1}{2}}{1-\frac{1}{2}}\right) = 1$	1	These marks depend on an attempt having been made on combining the results (i.e. almost, on the *following* M2	
	E[consecutive further purchases of V] $\left(= \frac{\frac{1}{2}}{1-\frac{1}{2}}\right) = 1$	1		
	∴ In steady state, E[consecutive further purchases of same brand] = $3 \times \frac{4}{7} + 1 \times \frac{3}{14} + 1 \times \frac{3}{14} = \frac{15}{7}(2.1429)$	M2 A1		5

<table>
<tr>
<td>Q.4</td>
<td>(i) $x_{ij} = \mu + a_i + e_{ij}$
$e_{ij} \sim$ ind $N(0, \sigma^2)$
μ is population grand mean for whole experiment
α_i is population mean amount by which i'th treatment differs from μ</td>
<td>2
1
1
1</td>
<td>Allow 'uncorrelated' for 'ind N'; condone absence of mean = 0</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>(ii) Totals
A 135.7
B 142.9
C 98.5
D 153.9
531.0

'Correction factor' = $\frac{531.0^2}{15}$ = 18797.4
Total SS = 18879.82 – CF
= 82.42
Between fertilisers SS
$= \frac{135.7^2}{4} + \frac{142.9^2}{4} + \frac{98.5^2}{3} + \frac{153.9^2}{4} - \text{CF} = 66.711$
Residual SS (by subtraction) = 82.42 – 66.711 = 15.709
<table>
<tr><th>Source of variation</th><th>SS</th><th>df</th><th>MS</th><th>MS ratio</th></tr>
<tr><td>Between fertilisers</td><td>66.711</td><td>3</td><td>22.237</td><td>15.572</td></tr>
<tr><td>Residual</td><td>15.709</td><td>11</td><td>1.428</td><td></td></tr>
<tr><td>Total</td><td>82.42</td><td>14</td><td></td><td></td></tr>
<tr><td></td><td>M1</td><td>2</td><td>M1</td><td>M1 A1</td></tr>
</table>
Refer 15.572 to $F_{3,11}$
Upper 1% pt is 6.22
Significant
Seems fertilisers not all the same</td>
<td>M1
2
M1
M1 A1
1
1
1
1</td>
<td></td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>(iii) $\sum_{j=1}^{n_i}\left(x_{ij}-\bar{x}_i\right)^2=\left(n_i-1\right)s_i^2\sim\sigma^2\chi^2_{n_i-1}$
We have $W=\sum_{i=1}^{k}\left\{\sum_{j}(x_{ij}-\bar{x}_i)^2\right\}$
$\sim\sum_{i=1}^{k}\left\{\sigma^2\chi^2_{n_i-1}\right\}$
These χ^2s are independent
$=\sigma^2\chi^2_{\Sigma(n_i-1)}\left(=\sigma^2\chi^2_{N-k}\right)$</td>
<td>1
1
1
1
1</td>
<td></td>
<td>5</td>
</tr>
</table>

Q.5	(i) $Y_1 = \theta_1 + \delta + e_1 \qquad Y_3 = \theta_3 + \delta + e_3$ $Y_2 = \theta_2 + \delta + e_2 \qquad Y_4 = \delta + e_4$ $\Omega = \Sigma e_i^2 = (Y_1 - \theta_1 - \delta)^2 + (Y_2 - \theta_2 - \delta)^2 + (Y_3 - \theta_3 - \delta)^2 + (Y_4 - \delta)^2$	1		
	$\left.\begin{aligned} \frac{\partial\Omega}{\partial\theta_1} &= -2(Y_1 - \theta_1 - \delta) = 0 \\ \frac{\partial\Omega}{\partial\theta_2} &= -2(Y_2 - \theta_2 - \delta) = 0 \\ \frac{\partial\Omega}{\partial\theta_3} &= -2(Y_3 - \theta_3 - \delta) = 0 \\ \frac{\partial\Omega}{\partial\delta} &= -2(Y_1 - \theta_1 - \delta) - 2(Y_2 - \theta_2 - \delta) - 2(Y_3 - \theta_3 - \delta) - 2(Y_4 - \delta) = 0 \end{aligned}\right\}$	1 M1 A1	Beware printed answer	
	Substitute 1, 2 and 3 in 4 gives $\hat{\delta} = Y_4$	1		
	And $\therefore \left.\begin{aligned} \theta_1 &= Y_1 - Y_4 \\ \theta_2 &= Y_2 - Y_4 \\ \theta_3 &= Y_3 - Y_4 \end{aligned}\right\}$	1		
	Candidates who state these results as obvious <u>before</u> setting up/solving the equations must be <u>absolutely</u> convincing!!			
	Observations are independent and all of variance σ^2 $\therefore \text{Var}(\hat{\delta}) = \sigma^2$	1		
	$\text{Var}(\hat{\theta}_1) = \text{Var}(\hat{\theta}_2) = \text{Var}(\hat{\theta}_3) = 2\sigma^2$	1		8
	(ii) Y_1, Y_2, Y_3 as before, but now $Y_4 = \theta_1 + \theta_2 + \theta_3 + \delta + e_4$ $\Omega = \Sigma e_i^2 = (Y_1 - \theta_1 - \delta)^2 + (Y_2 - \theta_2 - \delta)^2 + (Y_3 - \theta_3 - \delta)^2 + (Y_4 - \theta_1 - \theta_2 - \theta_3 - \delta)^2$ $\left.\begin{aligned} \frac{\partial\Omega}{\partial\theta_1} &= -2(Y_1 - \theta_1 - \delta) - 2(Y_4 - \theta_1 - \theta_2 - \theta_3 - \delta) = 0 \\ \frac{\partial\Omega}{\partial\theta_2} &= -2(Y_2 - \theta_2 - \delta) - 2(Y_4 - \theta_1 - \theta_2 - \theta_3 - \delta) = 0 \\ \frac{\partial\Omega}{\partial\theta_3} &= -2(Y_3 - \theta_3 - \delta) - 2(Y_4 - \theta_1 - \theta_2 - \theta_3 - \delta) = 0 \end{aligned}\right\}$ $\frac{\partial\Omega}{\partial\delta} = -2(Y_1 - \theta_1 - \delta) - 2(Y_2 - \theta_2 - \delta) - 2(Y_3 - \theta_3 - \delta) - 2(Y_4 - \theta_1 - \theta_2 - \theta_3 - \delta) = 0$	A1		
	HINT IN QUESTION THAT: $Y_4 - \hat{\theta}_1 - \hat{\theta}_2 - \hat{\theta}_3 - \hat{\delta} = 0$. Using this, $\hat{\theta}_1 = Y_1 - \hat{\delta}$, $\hat{\theta}_2 = Y_2 - \hat{\delta}, \hat{\theta}_3 = Y_3 - \hat{\delta}$	1		
	and substituting these into the hint gives: $Y_4 - Y_1 + \hat{\delta} - Y_2 + \hat{\delta} - Y_3 + \hat{\delta} - \hat{\delta} = 0$	1		
	$\therefore \hat{\delta} = \frac{1}{2}(Y_1 + Y_2 + Y_3 - Y_4)$	1		
	And $\therefore \hat{\theta}_1 = \frac{1}{2}(Y_1 + Y_4 - Y_2 - Y_3)$ $\hat{\theta}_2 = \frac{1}{2}(Y_2 + Y_4 - Y_1 - Y_3)$ $\hat{\theta}_3 = \frac{1}{2}(Y_3 + Y_4 - Y_1 - Y_2)$	1		
	Variance = $\frac{1}{4} \times 4\sigma^2 = \sigma^2$ for all estimators	2		8
	(iii) Comparing using the variances, scheme (ii) is to be preferred.	E4		4

Examiner’s Report

Statistics 6 (5518)

General Comments

There were 28 candidates from 9 centres. Small numbers, but it is good to know that there are still some candidates entering for this highest level statistics module. Much of the work was of very good quality.

Questions 2 and 3 were answered by very nearly everybody. Questions 1 and 4 were then approximately equally popular as the third question. There were only a very few attempts at question 5.

Comments on Individual Questions

Question 1 (maximum likelihood)

As has been noted in previous reports, it is pleasing that there were quite a lot of attempts at this question and that many of them were very good. Candidates do not seem to be put off by the technical standard of this work; rather, they get stuck in and get on with it, usually successfully. Perhaps inevitably, the work does prove beyond some candidates, but they are the minority.

Thus, the opening likelihood for the Poisson distribution was usually correctly formed and then correctly manipulated to obtain the sample mean as the estimator of θ. The re-formulation in terms of the new parameter ϕ was likewise usually correct, and the subsequent somewhat harder manipulations to obtain the estimator of ϕ were also usually done carefully and correctly. In the following discussion, however, not all candidates appreciated the point that was being illustrated: ϕ is a function of θ, and the maximum likelihood estimator of ϕ is *the same function* of the maximum likelihood estimator of θ. This is the so-called "invariance property" of maximum likelihood estimators (not the first time it has been explored), and is one of their many useful properties.

Part (iv) was an approach to comparing the efficiencies of the maximum likelihood estimator and another estimator. The question guided candidates through a fairly straightforward method that happens to work for the particular case being explored. Most candidates successfully differentiated the given function and found that its minimum is at $\phi = 1$. Not quite everybody could go on and make the conclusion from that point.

Question 2 (discrete bivariate distribution)

The opening parts of this question proved to be easy pickings for most candidates, but the good news is that parts (iv) and (v), which must presumably have been unfamiliar work, were also usually done very well. These explored what is called the "repeated expectations" method of obtaining expected values in bivariate (or, more generally, multivariate) situations, arguing via the expected value of a conditional expectation. Guidance was given in the question, but it was very pleasing to see most candidates handling the work in an assured and confident way. The candidates might perhaps have thought they were wasting their time in using a second, and in this case longer, method to obtain results that had already been obtained by direct analysis – so it would be as well to make the point here that the "repeated expectations" method is often *very* much easier for finding expected values in harder situations.

There are many numerical answers in the question, too many to quote here. Please see the published mark scheme for details.

Question 3 (Markov chain)

Nearly everyone wrote down the transition matrix without difficulty, and nearly everyone correctly obtained the limiting distribution [4/7, 3/14, 3/14]. The next part, to find the expected number of consecutive further occasions on which brand T will be bought, was well done and in a very important respect showed a welcome considerable improvement on work of the past few years. This refers to giving appropriate indications of method. These reports have repeatedly stated that unsupported writing-down of answers is regarded as in contravention of the global rubric that sufficient details of the working must be shown to indicate that a correct method is being used. This year, most candidates had grasped this point, and perhaps paid heed even more to the explicit instruction "*showing your working clearly*" (in italics!) in the question (but please note that the absence of such an explicit instruction is NOT an invitation to leave out all semblance of working). Thus methods to sum the series in this part were often very carefully set out. A generating function approach was particularly favoured, but other approaches were used too ("GP of GPs", differencing of series, binomial expansions). There were still some candidates who showed no working at all, or perhaps merely quoted "$\alpha/(1 - \alpha)$", often without even saying what α is, and these candidates lost marks.

Proceeding to the last part of the question, some candidates readily saw that this required combination of the answer from the previous part and corresponding answers for the other two brands. Others, however, seemed not to know what to do here.

Answer for penultimate part is 3; for last part is 15/7.

Question 4 (analysis of variance)

This question opened with a requirement to state the model. This is asked in many years, but there are still many candidates who do not know it at all, and many more who do not know it thoroughly. This is disappointing.

Part (ii) required an analysis of variance to be undertaken. All candidates knew what to do, though there were some errors in doing it. A downbeat feature was an increase in the proportion of candidates using cumbersome and inefficient methods for the calculation. The "squared totals" method, which has consistently been exhibited in the published mark schemes, is efficient for hand calculation and is to be commended. For some years, candidates have been moving to this method from others based on calculating variances, but this year there has been a swing back again. In a sense it doesn't matter, as any correct method is of course acceptable. But candidates might as well use an easier one! [Value of test statistic is 15.57, refer to F with 3 and 11 degrees of freedom, critical point is 6.22.] The final part was a derivation of the distribution of the within-sample (or residual) sum of squares. This was often done quite well, making good use of the intermediate results given in the question.

Question 5 (regression)

As already mentioned, there were very few attempts on this question, and unfortunately they were uniformly highly unsuccessful. Please see the published mark scheme for details of the solution.