Faculty of Actuaries **Institute of Actuaries**

REPORT OF THE BOARD OF EXAMINERS

September 2003

Subject 103 — Stochastic Modelling

EXAMINERS' REPORT

Introduction

The attached subject report has been written by the Principal Examiner with the aim of helping candidates. The questions and comments are based around Core Reading as the interpretation of the syllabus to which the examiners are working. They have however given credit for any alternative approach or interpretation which they consider to be reasonable.

J Curtis
Chairman of the Board of Examiners

11 November 2003

EXAMINATIONS

September 2003

Subject 103 — Stochastic Modelling

EXAMINERS' REPORT

1 (i) Suppose that $M_n = i$. Then the value of M_{n+1} depends entirely on the outcome of the next spin of the roulette wheel (this being independent of M_n) and the value of M_n. Hence M_n is a Markov chain.

(ii) Let X_n be the winning number from the n^{th} spin of the roulette wheel. Then $P(X_n = i) = 1/37$ for all $i = 0, 1, 2, \ldots, 36$.

If $j > i$ then $P(M_{n+1} = j \mid M_n = i) = P(X_{n+1} = j) = 1/37$;
If $j < i$ then $P(M_{n+1} = j \mid M_n = i) = 0$;
If $j = i$ then $P(M_{n+1} = j \mid M_n = i) = P(X_{n+1} \le i) = (i + 1)/37$.

i.e. $p_{ij} = \left\{ \begin{array}{ll} 1/37 & \text{if } \; j > i \\ (i+1)/37 & \text{if } \; j = i \\ 0 & \text{if } \; j < i \end{array} \right\}$

(iii) M is aperiodic, as it can stay in the same state with positive probability. It is not, however, irreducible, since it is not possible to return to state j from any state $k > j$.

(iv) State 36 is absorbing, so the only stationary distribution, which is also the limiting distribution, is $\pi_{36} = 1$, $\pi_i = 0$ for all other i.

The general reasoning type answers required for parts (i), (iii) and (iv) were quite well done in general. There were some difficulties with part (ii) However candidates fared less well on part (ii), and in many cases candidates struggling on part (ii) then failed to attempt the later parts.

2 (i) The Poisson process and the standard Brownian motion both possess the independent increments property.

$\mathrm{Cov}(X(t), X(t + s)) = \mathrm{Cov}(X(t), X(t)) + \mathrm{Cov}(X(t), X(t + s) - X(t)) = \lambda t + 0$, by the independent increments property.

(ii) $\mathrm{Cov}(B(t), B(t + s)) = \mathrm{Cov}(B(t), B(t)) + \mathrm{Cov}(B(t), B(t + s) - B(t) = t$.

(iii) (a) A Lévy process is a continuous-time process with stationary, independent increments. Alternatively, a Lévy process can be defined as a sum of three (indpendent) components: a constant drift, a multiple of Brownian motion and a purely discontinuous random component such as a compound Poisson process.

(b) The increments of Y are the weighted sum of the increments of B, X_1 and X_2, so are stationary and independent.

(c) $\mathrm{Cov}(Y(t), Y(t+s)) = \sigma^2\mathrm{Cov}(B(t), B(t), B(t+s)) + \kappa^2\mathrm{Cov}(X_1(t), X_1(t+s)) + \kappa^2\mathrm{Cov}(X_2(t), X_2(t+s)) = (\sigma^2 + 2\kappa^2\lambda)t$. All other terms vanish by independence.

The question demanded straightforward manipulation of the independent increments property and the covariance function.

3 (i) (a) The condition for X to be stationary is that the roots of the equation

$$1 - \alpha_1 z - \alpha_2 z^2 = 0$$

should lie outside the unit circle.

(b) The roots are $\dfrac{-\alpha_1 \pm \sqrt{\alpha_1^2 + 4\alpha_2}}{2\alpha_2}$.

In the given instance these are $\alpha_1 \pm \sqrt{\alpha_1^2 - 2}$.

If $\alpha_1^2 > 2$ then we require that

$$|\alpha_1| - \sqrt{\alpha_1^2 - 2} > 1 \text{ and } |\alpha_1| + \sqrt{\alpha_1^2 - 2} > 1$$

which is equivalent to

$$\alpha_1^2 - 2 < \left(|\alpha_1| - 1\right)^2$$

implying that $|\alpha_1| < 1.5$.

If, on the other hand, $\alpha_1^2 < 2$, then the roots are imaginary and satisfy

$$|z|^2 = \alpha_1^2 + (2 - \alpha_1^2) = 2$$

so that the condition is automatically satisfied.

(ii) The spectral density satisfies

$$H_1(\omega) f_X\left(\omega\right) = H_2(\omega) f_e(\omega),$$

where H_1 is the transfer function associated with $(1 - \alpha_1 B - \alpha_2 B^2)$, H_2 the transfer function associated with $(1 + \beta B)$, $f_X(\omega)$ is the spectral density of X and $f_e(\omega)$ is the spectral density of e.

We have

$$H_1(\omega) = |1-\alpha_1 e^{i\omega} - \alpha_2 e^{2i\omega}|^2,\ H_2(\omega) = |1 + \beta e^{i\omega}|^2 \text{ and } f_e(\omega) = \sigma^2/(2\pi).$$

Therefore

$$f_X(\omega) = \frac{\sigma^2\left(1+\beta^2+2\beta\cos(\omega)\right)}{2\pi\ \left(1+\alpha_1^2+\alpha_2^2-2\left(\alpha_1-\alpha_1\alpha_2\right)\cos\left(\omega\right)+2\alpha_2\cos\left(2\omega\right)\right)}$$

There was good understanding of the ARIMA process, which meant that candidates successfully derived the quadratic equation in part (i), though some were let down by their knowledge of complex numbers. Part (ii) was a straightforward application of the transfer function: the fact that marks were on average slightly lower seems to indicate that it had not been learned especially well.

4 (i) $P'_{u,u}(t) = -\lambda_u P_{u,u}(t) + \lambda_d P_{u,d}(t)$.

(ii) Note that $P_{u,d}(t) = 1 - P_{u,u}(t)$. Therefore we have

$$P'_{u,u}(t) = \lambda_d - (\lambda_u + \lambda_d) P_{u,u}(t),$$

implying that

$$\frac{d}{dt}\left[e^{(\lambda_u+\lambda_d)t} P_{u,u}(t)\right] = \lambda_d e^{(\lambda_u+\lambda_d)t}.$$

Together with the boundary condition $P_{u,u}(0) = 1$, this gives the required solution.

(iii)

$$E\left[U_t \mid Y_0 = y_u\right] = \int_0^t E\left[I_s \mid Y_0 = y_u\right] ds = \int_0^t P\left[Y_s = y_u \mid Y_0 = y_u\right] ds = \int_0^t P_{u,u}(s)\, ds.$$

Applying the previous part, this is equal to

$$\frac{\lambda_d}{\lambda_u+\lambda_d} t + \frac{\lambda_u}{(\lambda_u+\lambda_d)^2}\left(1-e^{-(\lambda_u+\lambda_d)t}\right).$$

(iv) $E\left[X_t \mid Y_0 = y_u\right] = y_d t + (y_u - y_d) E\left[U_t \mid Y_0 = y_u\right]$

$$= \left\{ y_d + (y_u - y_d)\frac{\lambda_d}{\lambda_d+\lambda_u}\right\} t + \frac{\lambda_u (y_u - y_d)}{(\lambda_u+\lambda_d)^2}(1-e^{-(\lambda_u+\lambda_d)t}).$$

In general candidates were able to score well on the first two parts, although a common mistake here was to have the exponential parameters λ_d and λ_u transposed in the Kolmogorov Forward Equations. In such cases appropriate credit was given for valid attempts at part (ii).

5 (i) $M_n = e^{\theta S_n - cn}$ will be a martingale if $E(e^{\theta S_{n+1} - c(n+1)} | F_n) = e^{\theta S_n - cn}$.

(ii) This will happen if $e^c = Ee^{\theta X_{n+1}} = pe^{\theta} + qe^{-\theta}$, i.e. if $pe^{2\theta} - e^{\theta+c} + q = 0$.

Therefore

$$\theta = \ln\left(\frac{e^c \pm \sqrt{e^{2c} - 4pq}}{2p}\right).$$

(iii) For the OST to hold we require that M is bounded or T_1 is bounded or $M_{n\wedge T_1}$ is bounded.

In this instance, if $c \geq 0$ and $\theta \geq 0$ then $M_n < e^{\theta}$ for all $n \leq T_1$. But if $c < 0$ or $\theta \leq 0$ then there is no such upper bound and it is not safe to assume that the OST can be applied.

(iv) When $c > 0$, one root for θ is positive, the other negative, since $pe^{2\theta} - e^{\theta+c} + q < 0$ when $\theta = 0$. We need the positive root.

Applying the OST, $E(M_T) = M_0 = 1$ as long as $c \geq 0$ and $\theta \geq 0$. This implies that

$$1 = E(e^{\theta S_{T_1} - cT_1}) = e^{\theta} E(e^{-cT_1}).$$

Thus

$$E(e^{-cT_1}) = e^{-\theta} = \frac{2p}{e^c + \sqrt{e^{2c} - 4pq}} = \frac{e^c - \sqrt{e^{2c} - 4pq}}{2q}.$$

In a number of cases, candidates covered some of part (ii) under part (i) — credit was given in these cases. In general candidates were able to score well on the bookwork required for part (iii) although candidates were less successful in tackling the final part.

6 (i) We have

$$\gamma_k = \mathrm{Cov}(X_t, X_{t-k}) = \frac{1}{(m+1)^2}\mathrm{Cov}\left(\sum_{r=0}^{m} e_{t-r}, \sum_{r=0}^{m} e_{t-k-r}\right)$$

$$= \frac{1}{(m+1)^2}\sum_{r=0}^{m}\sum_{r=0}^{m}\mathrm{Cov}(e_{t-r}, e_{t-k-r}).$$

Clearly if $k > m$, all terms are zero, so that $\gamma_k = 0$.

For $0 \le k \le m$, there are exactly $(m - k + 1)$ non-zero terms, and each of these covariance terms equals σ_e^2. Thus

$$\gamma_k = \begin{cases} 0 & k > m \\ \frac{m-k+1}{(m+1)^2}\sigma_e^2 & k = 0, 1, 2, \ldots, m. \end{cases}$$

The autocorrelation function is

$$\rho_k = \begin{cases} 1 & k = 0 \\ \frac{m-k+1}{m+1} & k = 1, 2, \ldots, m \\ 0 & k > m \end{cases}$$

(ii) For the process to be invertible, we require that the roots of the characteristic equation should be greater than 1 in absolute value.

We can rewrite the MA model with the aid of the backward shift operator B as follows:

$$X_t - \mu = \frac{1}{3}(1 + B + B^2)\, e_t.$$

The roots of the characteristic equation

$$1 + B + B^2 = 0$$

are $B = -½ + ½ i\sqrt{3}$, $B = -½ - ½ i\sqrt{3}$.

In both cases $|B| = 1$. Thus the process is not invertible.

In general candidates made reasonable attempts along the right lines, although this did not always result in the correct autocorrelation function being calculated. Where possible, some credit was given for attempts at the second part based on an incorrect part (i).

7 (i) The original data are clearly subject to seasonal variation, and the size of the seasonal fluctuations is increasing in line with the value of the underlying quantity. This suggests that the seasonal variation is multiplicative rather than additive, in which case taking the logarithm is the sensible thing to do. In addition to this, a look at the plot of y_t against time confirms that the variation is much more regular.

(ii) Seasonal variation is a predictable pattern of deterministic variation in the mean of the process which is cyclic, i.e it repeats after a fixed number of time periods, usually corresponding to a year of elapsed time.

A linear trend is a deterministic pattern of variation in the mean of the process which is linearly dependent on the time variable, i.e. is of the form $a + bt$.

There are various possible answers. Possible methods include:

(a) Estimate the trend by linear regression and remove it, then, for each month, calculate the sample mean value for that particular month over all five years. From every detrended observation subtract off the appropriate seasonal mean to obtain seasonally adjusted data.

(b) Remove the linear trend by differencing the data once, then remove the seasonal variation by seasonal differencing. In other words, $z_t = (1 - B^{12})(1 - B)y_t$.

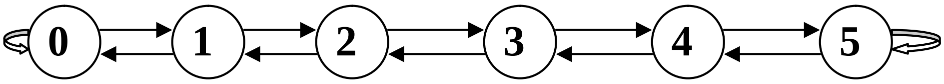
(iii) (a) The fact that the sample ACF is not near 1 for small lags is the most obvious pointer to the stationarity of the adjusted data set.

(b) The clue here is the highest lag for which the ACF or PACF is significantly different from 0. Looking at the sample ACF we might suggest that a MA(3) might fit, as the sample ACF is roughly zero for $k > 3$. Similarly, a look at the sample PACF seems to indicate an AR(3). But it might well be possible to find an ARMA(1,1) which would fit adequately. In other words, $d = 0$ and either $p = 0$, $q = 3$ or $p = 3$, $q = 0$ or $p = 1$, $q = 1$.

(iv) We have $y_t = m_t + z_t$, where m_t represents deterministic variation and z_t is a purely random component. The process of seasonal adjustment and detrending has produced an estimate $\hat{m}_t$ for m_t which can be extrapolated into the future. Thus we have $\hat{y}_{60}(1) = \hat{m}_{61} + \hat{z}_{60}(1)$, which in turn leads to $\hat{x}_{60}(1) = \exp(\hat{y}_{60}(1))$.

Neither the discussion of how to deal with seasonal variation nor the practical part to do with model identification was especially well tackled.

8 (i) Transition graph:



(Transition probabilities are α to the right and $1 - \alpha$ to the left.)

The transition matrix is

$$P = \begin{pmatrix} 1-\alpha & \alpha & 0 & 0 & 0 & 0 \\ 1-\alpha & 0 & \alpha & 0 & 0 & 0 \\ 0 & 1-\alpha & 0 & \alpha & 0 & 0 \\ 0 & 0 & 1-\alpha & 0 & \alpha & 0 \\ 0 & 0 & 0 & 1-\alpha & 0 & \alpha \\ 0 & 0 & 0 & 0 & 1-\alpha & \alpha \end{pmatrix}$$

(ii) (a) The equations $\pi P = \pi$ read

$$\begin{aligned} \pi_0(1-\alpha) + \pi_1(1-\alpha) &= \pi_0 \\ \pi_0\alpha + \pi_2(1-\alpha) &= \pi_1 \\ \pi_1\alpha + \pi_3(1-\alpha) &= \pi_2 \\ \pi_2\alpha + \pi_4(1-\alpha) &= \pi_3 \\ \pi_3\alpha + \pi_5(1-\alpha) &= \pi_4 \\ \pi_4\alpha + \pi_5\alpha &= \pi_5 \end{aligned}$$

(b) Discard last equation and solve first one in terms of π_0:

$$\pi_1(1-\alpha) = \pi_0(1-(1-\alpha))$$

$$\therefore \pi_1 = \frac{\alpha}{1-\alpha}\pi_0$$

$$\pi_2(1-\alpha) = \pi_0\left(\frac{\alpha}{1-\alpha} - \alpha\right) = \pi_0 \frac{\alpha^2}{1-\alpha}$$

$$\therefore \pi_2 = \left(\frac{\alpha}{1-\alpha}\right)^2 \pi_0$$

In general $\pi_j = \left(\dfrac{\alpha}{1-\alpha}\right)^j \pi_0$.

(c) Find π_0 by normalisation:

$$1 = \sum_{j=0}^{n} \pi_j = \pi_0 \sum_{j=0}^{5} \left(\frac{\alpha}{1-\alpha}\right)^j = \pi_0 \frac{1-\left(\frac{\alpha}{1-\alpha}\right)^6}{1-\frac{\alpha}{1-\alpha}}$$

$$\therefore \pi_0 = \frac{1-\frac{\alpha}{1-\alpha}}{1-\left(\frac{\alpha}{1-\alpha}\right)^6}.$$

(iii) For a consistent profit the company requires that $\min_{0<\alpha\le 0.5} P(\alpha) > 0$, where $P(\alpha)$ is the expected profit when annual claim rate is α.

Expected long-term annual income from one customer is $\sum_{j=0}^{5} \pi_j P_j$, where P_j is the premium payable in discount level j, and expected annual claims $= C(1-\alpha)$, so expected profit is

$$P(\alpha) = \sum_{j=0}^{5} \rho^j P_j \frac{1-\rho}{1-\rho^6} - C(1-\alpha), \text{ where } \rho = \frac{\alpha}{1-\alpha}.$$

Candidates showed a good understanding of Markov chains here, with many candidates achieving a high score. For the final part of the question, a large number of candidates only gave a cursory explanation and were unable to score the full marks.

9 (i) Use the inverse distribution function method.

$F(t) = 1 - P(T > t)$, so that

$$F(t) = 1 - \exp\left(-\int_{65}^{t} (a+bx)dx\right) = 1 - \exp\left(-0.5b(t^2 - 65^2) - a(t-65)\right).$$

Rearranging,

$$\frac{b}{2}t^2 + at - \frac{b}{2}65^2 - 65a + \log(1 - F(t)) = 0,$$

or, in other words,

$$F^{-1}(u) = \frac{-a \pm \sqrt{a^2 + (65b)^2 + 2b\left[65a - \log(1-u)\right]}}{b}.$$

Since the simulated variable must be positive, the positive root is required. The method, then, is to generate a pseudo-uniform random variable U in the range [0,1] and to set

$$T = \frac{-a + \sqrt{a^2 + (65b)^2 + 2b\left[65a - \log(1-U)\right]}}{b}.$$

(ii) (a) Y is a Binomial random variable, with parameters $n = 10$ and

$$p = \exp\left(-\int_{65}^{75}(a+bx)dx\right) = \exp(-700b - 10a).$$

Let $G(y)$ be the distribution function of Y and let U be a single uniform pseudo-random variable on [0,1]. Then set $Y = \min\left\{y : G(y) \geq U\right\}$.

(b) Although this method requires a certain amount of computing time to evaluate the distribution function G, this only has to be done once; thereafter, only one value of U is needed to generate each value of Y, as opposed to the ten values of U which are required in the other method.

(iii) It is important to use the same sequence of pseudo-random numbers in each case, otherwise we are not comparing like with like.

Many candidates correctly identified that the inverse transformation method was required, although in a surprisingly high proportion of cases marks were lost because of errors in the algebra / integration. The later parts of this question, dealing with the comparison of two methods of generating discrete random variables, were in general less well done.

10 (i) A geometric Brownian motion can be defined as a process $S_t = S_0\exp(\mu t + \sigma B_t)$, where B_t is a standard Brownian motion. It satisfies the SDE $dS_t = \mu S_t dt + \sigma S_t dB_t$.

Alternatively, use the definition $S_t = S_0\exp(\mu t + \sigma B_t)$ *, which satisfies the SDE* $dS_t = \left[\mu + \tfrac{1}{2}\sigma^2\right]S_t dt + \sigma S_t dB_t$.

(ii) We use the Itô formula (we must have the form where F is a function of t as well as x):

$$dF\left(t, X_t\right) = \left(F_t'\left(t, X_t\right) + \tfrac{1}{2}\sigma^2 F_x''\left(t, X_t\right)\right)dt + \sigma F_x'\left(t, X_t\right)dX_t$$

Let $Y_t = e^{\alpha t}(X_t - c) = F\left(t, X_t\right)$. Then $Y(0) = x_0$ and

$$dY_t = d\left(e^{\alpha t}(X_t - c)\right)$$

$$= \left(\alpha e^{\alpha t}[X_t - c] - \alpha e^{\alpha t}[X_t - c]\right)dt + \sigma e^{\alpha t} dB_t$$

$$= \sigma e^{\alpha t} dB_t$$

Hence $Y(t) = x_0 - c + \sigma \int_{s=0}^{t} e^{\alpha s} dB_s$

and so $X_t = c + e^{-\alpha t}(x_0 - c) + \sigma e^{-\alpha t} \int_{s=0}^{t} e^{\alpha s} dB_s$.

(iii) The required condition for the stationary density π of the diffusion Y solving $dY_t = \mu(Y_t)dt + \sigma(Y_t)dB_t$, from the Core Reading, is

$$\frac{d}{dy}\ [\mu(y)\ \pi(y)] = ½\ \frac{d^2}{dy^2}\ [\sigma^2(y)\ \pi(y)]$$

(iv) Two ways to do this. From (i),

$$X_t \sim N(c + e^{-\alpha t}(x_0 - c), \sigma^2 \int_0^t e^{-2\alpha(t-s)} ds) = N\left(c + e^{-\alpha t}(x_0 - c), \frac{\sigma^2}{2\alpha}\left[1 - e^{-2\alpha t}\right]\right)$$

The limiting distribution, as $t \to \infty$, is $N(c, \sigma^2/(2\alpha))$. A limiting distribution is always stationary.

Alternatively, in this instance we have $\mu(x) = -\alpha(x-c)$, $\sigma(x) = \sigma$, so the condition given in (ii) is that

$$-\alpha\pi(x) - \alpha(x - c)\pi'(x) = \tfrac{1}{2}\sigma^2\pi''(x) .$$

The solution to this DE is

$$\pi(x) = const.\exp\left(-\frac{\alpha(x - c)^2}{\sigma^2}\right),$$

which is the density function of $N(c, \sigma^2/(2\alpha))$.

(v) A function of the form Ax^b for all $x > 0$ cannot integrate to 1, no matter what the values of A and b. This means that there is no stationary density function π for the geometric Brownian motion.

It would be surprising if there had been: it is well known that Brownian motion is non-stationary, and therefore that geometric Brownian motion is also a non-stationary process, so cannot possess a stationary density.

Marks were poor for Question 10. Candidates should note that the reference in part (iv) of the question to "the equation in (i)" was incorrect and should have read "the equation in (ii)". Whilst at least some candidates were still able to complete this part of the question, allowance was made in the marking of scripts for this error.

It seemed that candidates had not committed to memory the formula for the equilibrium density of a diffusion process, presumably because it had not been asked before.

11 (i) Model fitting: this occurs after the family of model has been decided and concerns the estimation of the values of parameters. The set of parameters to be estimated is determined by the choice of model family.

Model verification: once the model has been fitted we need to check that the fitted process resembles what has been observed. Generally we produce simulations of the process, using the estimated parameter values, and compare them with the observations.

(ii) The parameters are the rate of leaving state i, λ_i, for each i, and also the jump-chain transition probabilities, r_{ij} for $j \neq i$, where r_{ij} is the conditional probability that the next transition takes the chain to state j given that it is now in state i. Alternatively, one may regard the parameters as being σ_{ij}, where $\sigma_{ii} = -\lambda_i$ and, for $j \neq i$, $\sigma_{ij} = \lambda_i r_{ij}$.

Assumptions of the Markov model are that the duration of holding time in state i has exponential distribution with parameter determined only by i and is independent of anything that happened before the current arrival in state i, and that the destination of the next jump after leaving state i is independent of the holding time in state i and of anything that happened before the chain arrived in state i.

(iii) $\hat{\lambda}_i^{-1}$ is the average duration of each stay in state i. Thus $\hat{\lambda}_1 = \frac{1}{6}$ per minute, or 10 per hour, $\hat{\lambda}_2 = \frac{1}{40}$ per minute or 1.5 per hour, $\hat{\lambda}_3 = \frac{1}{30}$ per minute, or 2 per hour.

$$\hat{r}_{12} = \frac{3}{8},\ \hat{r}_{13} = \frac{5}{8},\ \hat{r}_{21} = \frac{1}{4},\ \hat{r}_{23} = \frac{3}{4},\ \hat{r}_{31} = \frac{7}{8} \text{ and } \hat{r}_{32} = \frac{1}{8}.$$

Thus the generator matrix, in units of hr^{-1}, is

$$\begin{pmatrix} -10 & 3.75 & 6.25 \\ 0.375 & -1.5 & 1.125 \\ 1.75 & 0.25 & -2 \end{pmatrix}$$

(iv) One test should test whether the holding times in each state are exponentially distributed. If $T_{i,k}$ denotes the kth holding time in state i, then the hypothesis is that $T_{i,1}, T_{i,2}, \ldots, T_{i,ni}$ is a sample from an exponential distribution with parameter $\hat{\lambda}_i$: sort the observations into categories, calculate expected number in each category and hence find the X^2 statistic by summing $(O - E)^2/E$. This should be compared with the critical value of the χ^2 distribution with $m - 2$ d.f., where m is the number of categories.

(v) Estimate of expected duration of a visit to state 1 is 6 mins, so this is the estimated time until the 21st transition.

Estimate of expected time between 21st and 22nd transitions is

E(Time | transition is to 2) P(transition is to 2) + E(Time | to 3) P(to 3)
$= 40 \times (3/8) + 30 \times (5/8) = 33.75$ mins.

Attempts to fit the model to the observed data were generally sensible and encouraging, although in a minority of cases the calculation of the parameter estimates betrayed evidence of some confusion. There was a tendency to be less successful as the question continued, with the result that attempts at the final part were of a noticeably lower standard.