EXAMINATIONS

September 2004

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.

M Flaherty
Chairman of the Board of Examiners

23 November 2004

© Faculty of Actuaries
© Institute of Actuaries

he examiners were pleased to note that the overall quality of answers on this final sitting of subject 103 was high and that many of the candidates demonstrated a good knowledge of the principles and practice of stochastic modelling. As always, credit was awarded for comments which showed that candidates had an understanding of the topic covered in a question, even if the calculations gave the wrong answer due to some mathematical error. Question 7 was particularly well answered, with Questions 2 and 4 not far behind. Questions 6 and 10 had the lowest proportion of good answers; it is possible that time pressure played a role in the case of Question 10.

1 (i) Let n_{ij} denote the number of direct transitions from state i to state j, with n_{i+} the total number of transitions out of state i. Then $\hat{p}_{ij} = n_{ij} / n_{i+}$.

(ii) Model fitting aims to find the best-fitting model in a given class. But it is conceivable that even the best-fitting model in the class does not fit very well. Model validation is a set of procedures to test the adequacy of the fit.

(iii) Sensitivity analysis is part of model validation. The purpose is to determine whether the behaviour of the fitted model would be substantially different if the parameter values were slightly different from the estimates already obtained.

The technique involves simulating the fitted process a large number of times, using several simulations for each of a number of slightly different parameter values, then examining the output of the simulation to attempt to identify systematic differences.

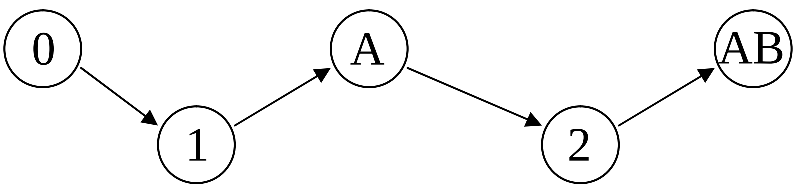
It is important that the same sequence of random numbers be used in each of the sets of simulations to ensure comparability.

Many candidates failed to mention the importance of using the same sequence of pseudo-random numbers. Apart from that, most answers showed good knowledge of the principles of modelling.

2 (i) It is inadequate because someone who has never suffered from disease A or B is not in the same position as someone who has suffered from one or both in the past but is currently healthy.

The state space should be extended by splitting state 0 into 3: "0: Has never suffered from A or B", "A: Has suffered from A but is now healthy", "AB: Has suffered both A and B but is now healthy."

(ii)



(iii) Only a time-inhomogeneous model can properly reflect the dependence of both the recovery rate for A and the infection rate for B on the age of the person.

(iv) If, in a population taken as a whole, the number of people in each age group is roughly constant over time, then the age-dependent transition rates of the individuals who make up the population can be "averaged out" to give a time-homogeneous model which works perfectly adequately given that national medical services are generally only concerned with total numbers falling ill.
This question was answered well in general. Where candidates lost marks it was often due to mis-specifying the additional states in part (i). Splitting state A into "Has recovered from Disease A and is aged below 18" and "Has recovered from Disease A and is aged 18 or more" is reasonable when modelling an entire population, but does not lead to a time-homogeneous Markov model when applied to a single individual, since one's 18th birthday does not occur at a random time. However, answers along these lines with good explanations were given full marks.

3 (i) Set X_i as follows:

$$X_i = \begin{cases} A & \text{if } 0 \le U_i \le 1/5 \\ B & \text{if } 1/5 < U_i \le 9/20 \\ C & \text{if } 9/20 < U_i \end{cases}$$

(ii) Use the inverse transformation method.

The distribution function is $F(x) = \int_4^x \frac{10-t}{18}\,dt = 1 - \frac{(10-x)^2}{36}$ for $4 \le x \le 10$.

Solving the equation gives $F^{-1}(u) = 10 - 6\sqrt{1-u}$

So we can set $X_i = 10 - 6\sqrt{1-U_i}$

or alternatively we could use $X_i = 10 - 6\sqrt{U_i}$

(iii) Use acceptance-rejection method:

Let $V_1 = \pi\, U_1$, so that V_1 is uniformly distributed on $[0, \pi]$ and has density function $g(x) = 1/\pi$ over that range.

We define

$$C = \sup_{0<x<\pi} f(x)/g(x) = \sup_{0<x<\pi} 2\sin^2 x = 2.$$

If $U_2 < \sin^2 V_1$ let $X_1 = V_1$; otherwise reject this value and select a new pair U_1, U_2. Repeat for other X_i

Answers to parts (i) and (ii) were generally good. For part (iii) many candidates only described the general theory without specifying g(x) or calculating the constant C.

4 (i) A standard Brownian motion $\{B_t\}$ is defined by the following properties:

- $B(0) = 0$ and B_t has independent increments; $B_t - B_s$ is independent of B_u for $0 \le u \le s$ and $s \le t$.
- B_t has stationary and Gaussian increments; $B_t - B_s \sim N(0, t-s)$.
- B_t has continuous sample paths, i.e. $t \to B_t$ is continuous.

(ii) Using Itô's Lemma gives

$$d(\log S_t) = \frac{1}{S_t} dS_t - \frac{1}{2} . \frac{1}{S_t^2} (dSt)^2$$

$$= \mu dt + \sigma\, dB_t - \frac{\sigma^2}{2} dt.$$

This implies that

$$\log S_t = \log S_0 + \left(\mu - \frac{\sigma^2}{2} \right) t + \sigma B_t,$$

or, finally,

$$S_t = S_0 e^{\left(\mu - \frac{\sigma^2}{2} \right) t + \sigma B_t}.$$

(iii) We have

$$P\left[S_t > x \mid S_0 = y\right] = P\left[\sigma B_t + \left(\mu - \frac{\sigma^2}{2}\right)t > \log\frac{x}{y}\right]$$

$$= P\left[B_t > \frac{1}{\sigma}\left(\log\frac{x}{y} - \left(\mu - \frac{\sigma^2}{2}\right)t\right)\right]$$

$$= 1 - \Phi\left(\frac{\log\frac{x}{y} - \left(\mu - \frac{\sigma^2}{2}\right)t}{\sigma\sqrt{t}}\right).$$

Substituting the values $x = 120$, $y = 96$, $\mu = 0.2$, $\sigma = 0.15$ and $t = 0.25$ years above we find that the answer is

$$1 - \Phi(2.3461) = 1 - 0.9905 = 0.0095.$$

The calculations in parts (ii) and (iii) were well done. The definition in part (i) caused some problems: it is necessary to mention the stationary, independent increments property; then either of the two remaining properties (continuous sample paths, normally distributed increments) implies the other.

5 (i) The Markov property is clear: the chain jumps either up or down by 1, with probabilities depending only on the current state, not the past history.

$P(X_{n+1} = i+1 \mid X_n = i)$ is the probability that the (n+1)th ball selected is red, which is just 1/N of the number of red balls at time n, which is $N - i$.

(ii) From any state i it is possible to reach any other state j in just $|j - i|$ steps, either all upwards or all downwards. This means that the chain is irreducible.

Every transition takes the chain from an even state to an odd one or vice versa, which implies that the period must be an even number.

On the other hand, starting from state 0 it is possible to return to 0 in two steps. Therefore state 0 has period 2 and, by irreducibility, all states have period 2.

(iii) To find the stationary distribution we can use the relationship suggested by the Detailed Balance Equations:

$$\pi_i \quad \frac{N-i}{N} \quad = \quad \pi_{i+1} \quad \frac{i+1}{N} \text{ for } i = 0, 1, 2, \ldots, N-1.$$

Thus we get the recursive relationship

$$\pi_{i+1} \quad = \quad \frac{N-i}{i+1} \quad \pi_i \text{ for } i = 0, 1, 2, \ldots, N-1$$

Starting with $i = 0$ and working forwards we get

$$\pi_1 = \frac{N}{1}\pi_0$$

$$\pi_2 = \frac{N-1}{2}\pi_1 = \frac{N(N-1)}{(2)(1)}\pi_0$$

$$\pi_3 = \frac{N-2}{3}\pi_2 = \frac{N(N-1)(N-2)}{(3)(2)(1)}\pi_0$$

and in general

$$\pi_i = \frac{N(N-1)(N-2)\cdots(N-i+1)}{i(i-1)(i-2)\cdots(2)(1)}\pi_0 = \frac{N!}{(N-i)!\ \ i!}\pi_0 = \binom{N}{i}\pi_0$$

Alternatively, write down the transition matrix P and use the equation $\pi^T P = \pi^T$ to obtain

$$\frac{1}{N}\pi_1 = \pi_0 \Rightarrow \pi_1 = N\pi_0$$

$$\pi_0 + \frac{2}{N}\pi_2 = \pi_1 \Rightarrow \pi_2 = \frac{N(N-1)}{2}\pi_0$$

$$\frac{N-1}{N}\pi_1 + \frac{3}{N}\pi_3 = \pi_2 \Rightarrow \pi_3 = \frac{N(N-1)(N-2)}{3!}\pi_0, \text{ etc.}$$

To find π_0, we use the fact that $\pi_0 + \pi_1 + \pi_2 + \ldots\ldots + \pi_N = 1$

i.e. $\pi_0 = \left[1 + \sum_{i=1}^{N} \frac{N!}{(N-i)!\ \ i!}\right]^{-1} = \left[\sum_{i=0}^{N} \frac{N!}{i!}\right]^{-1} = \left(2^N\right)^{-1} = \frac{1}{2^N}$

Therefore the stationary distribution $\pi = (\pi_0, \pi_1, \pi_2, \ldots\ldots, \pi_N)$ is given by

$$\pi_i = \binom{N}{i}\ \frac{1}{2^N} \text{ for } i = 0, 1, 2, \ldots, N$$

(iv) $P(X_n = j \mid X_0 = 0)$ does not converge, being alternately zero and non-zero, since X is periodic.

The derivation of the stationary distribution in part (iii) caused difficulties with many candidates, but otherwise candidates showed a good understanding of discrete-time Markov chains.

6 (i) (a) $\frac{dm_1}{dt} = \frac{d}{dt} E_x X_t = \frac{1}{dt} E_x[dX_t] = \alpha E_x[b - X_t] = \alpha(b - m_1)$, where E_x denotes conditional expectation given $X_0 = x$. This derivation uses the fact that the increments of Brownian motion have expectation equal to zero.

(b) $\frac{d}{dt}[e^{\alpha t} m_1(t)] = \alpha b e^{\alpha t}$, implying that

$m_1(t) = e^{-\alpha t}[x + b(e^{\alpha t} - 1)] = b + (x - b)e^{-\alpha t}$.

(ii) (a) $dY_t = 2X_t dX_t + (dX_t)^2 = 2X_t[\alpha(b - X_t)dt + \sigma\sqrt{X_t}dB_t] + \sigma^2 X_t dt$

$= [2\alpha b + \sigma^2]X_t dt - 2\alpha X_t^2 dt + 2\sigma X_t^{3/2} dB_t$

(b) $\frac{d}{dt} m_2(t) = [2\alpha b + \sigma^2]m_1(t) - 2\alpha m_2(t)$. Again we have used the fact that Brownian increments have mean zero.

(c) We do not need to solve the equation, but just to note that since dm_2/dt tends to 0, this implies that $2\alpha \lim_{t\to\infty} m_2(t) = [2\alpha b + \sigma^2] \lim_{t\to\infty} m_1(t) = 2\alpha b^2 + b\sigma^2$. Therefore

$\lim_{t\to\infty} E[X_t^2 \mid X_0 = x] = b^2 + \frac{b\sigma^2}{2\alpha}$, from which we deduce that

$\lim_{t\to\infty} \mathrm{Var}[X_t \mid X_0 = x] = \frac{b\sigma^2}{2\alpha}$.

This question was relatively poorly answered, although much of it is based on the standard theory of Ordinary Differential Equations.

7 (i) Taking covariances with X_{t-k} for $k \geq 1$ in (1) gives

$$Cov(X_t, X_{t-k}) = \alpha_1 Cov(X_{t-1}, X_{t-k}) + \alpha_2 Cov(X_{t-2}, X_{t-k}) + \cdots + \alpha_p Cov(X_{t-p}, X_{t-k}),$$

which gives the Yule-Walker equations since, by definition, $\gamma_k = Cov(X_t, X_{t-k})$ for $0 \leq k \leq p$.

For $k = 0$, there is an extra term which accounts for $Cov(X_t, e_t) = \sigma^2$.

(ii) A diagnostic test is based on the partial ACF and uses the fact that, for an AR(2) process, the partial autocorrelations, ϕ_k, are zero for $k > 2$.

The values of ϕ_k are estimated by the partial ACF, $\hat{\phi}_k$, and for $k > 2$ the asymptotic variance of $\hat{\phi}_k$ is $1/n$. Using a normal approximation, values of the sample partial ACF outside the range $\pm 2/\sqrt{n}$ indicate that the AR(2) model may be inadequate.

(iii) (a) The process can be written in terms of the backward shift operator as $(1-0.6B-0.3B^2)X_t = e_t$.

Hence the characteristic polynomial is $1-0.6z-0.3z^2$ with roots $\dfrac{0.6 \pm \sqrt{(0.6)^2+1.2}}{-0.6}$, i.e. the roots are $-1 \pm \sqrt{156}/6$.

Since both roots lie outside the unit circle, the process can be stationary.

(b) X_t is not Markov since the conditional distribution of X_{k+1} given the history up to time k depends on X_{k-1} as well as on X_k.

(c) The Yule-Walker equations in this case yield

$$\gamma_0 = 0.6\gamma_1 + 0.3\gamma_2 + 1 \qquad (3)$$
$$\gamma_1 = 0.6\gamma_0 + 0.3\gamma_1 \qquad (4)$$
$$\gamma_2 = 0.6\gamma_1 + 0.3\gamma_0. \qquad (5)$$

From (4) we have

$$0.7\gamma_1 = 0.6\gamma_0 \Rightarrow \gamma_1 = \frac{6\gamma_0}{7} \qquad (6)$$

and substituting into (5) we get

$$\gamma_2 = \frac{36}{70}\gamma_0 + \frac{3}{10}\gamma_0 = \frac{57}{70}\gamma_0. \qquad (7)$$

Inserting the last two equations into (3) we obtain

$$\gamma_0 = \frac{36}{70}\gamma_0 + \frac{171}{700}\gamma_0 + 1$$

which gives

$$\left(1 - \frac{36}{70} - \frac{171}{700}\right)\gamma_0 = 1 \Rightarrow \gamma_0 = \frac{700}{169}.$$

Then (6) and (7) yield resp.

$$\gamma_1 = \frac{6\gamma_0}{7} = \frac{600}{169}, \quad \gamma_2 = \frac{570}{169}.$$

The examiners were pleased to note the high quality of answers to this question. It appears that the theoretical principles of Time Series analysis are well understood.

8 (i) The equation is

$$X_t = \mu + \alpha(X_{t-1} - \mu) + e_t + \beta e_{t-1}.$$

The parameters are α (the autoregressive parameter), β (the moving average parameter), the mean level μ and the innovation standard deviation σ.

(ii) A time series process is (weakly) stationary if the mean of the process, $m_t = E(X_t)$, does not vary with time and the covariance of the process, $Cov(X_t, X_s)$ depends only on the time difference $t-s$ and not on the particular values t, s.

For the model in (1) to be stationary, $|\alpha| < 1$ is needed.

(iii) For the method of moments, we calculate the theoretical ACF ρ_1, ρ_2 in terms of the parameters α, β. Then we find the sample ACF, say r_1, r_2 from the data. Subsequently we obtain estimates for α, β by equating ρ_1 with r_1 and ρ_2 with r_2.

The value of σ^2 is estimated using the calculated value of γ_0 and the sample variance, whereas an estimate for μ is the sample mean $\overline{x}$.

(iv) (a) Using the given values we obtain the forecasts

$$\hat{x}_{25}(1) = 9.12 + 0.71(14.82) + 0.17(-1.98) = 19.306$$

and

$$\hat{x}_{25}(2) = 9.12 + 0.71(19.306) = 22.827.$$

(b) For exponential smoothing the equation is

$$\hat{x}_{25}(1) = \hat{x}_{24}(1) + \alpha\left(x_{25} - \hat{x}_{24}(1)\right) = 12.97 + 0.3(14.82 - 12.97) = 13.525.$$

(v) Exponential smoothing is simple to apply and does not suffer from problems of over-fitting. If the data appear fairly stationary but are not especially well fitted by any of the Box-Jenkins methods, exponential smoothing is likely to produce more reliable results. More advanced versions of exponential smoothing can cope with varying trends and multiplicative variation.

Many candidates omitted to mention σ as a parameter in part (i). Marks for this question were not quite as good as for Q7, indicating that the practical aspects of Time Series analysis are less well understood than the theoretical ones.

9 (i) The Markov model implies that holding times are exponentially distributed.

(ii) The generator matrix is as follows (in minutes then, equivalently, in hours):

$$\begin{array}{c} \\ W \\ A \\ M \\ S \\ H \end{array} \begin{array}{c} \begin{array}{ccccc} W & A & M & S & H \end{array} \\ \begin{bmatrix} -\frac{1}{15} & \frac{1}{15} & 0 & 0 & 0 \\ 0 & -\frac{1}{20} & \frac{3}{400} & \frac{1}{400} & \frac{1}{25} \\ 0 & 0 & -\frac{1}{30} & 0 & \frac{1}{30} \\ 0 & 0 & 0 & -\frac{1}{180} & \frac{1}{180} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

$$\begin{array}{c} \\ W \\ A \\ M \\ S \\ H \end{array} \begin{array}{c} \begin{array}{ccccc} W & A & M & S & H \end{array} \\ \begin{bmatrix} -4 & 4 & 0 & 0 & 0 \\ 0 & -3 & 0.45 & 0.15 & 2.4 \\ 0 & 0 & -2 & 0 & 2 \\ 0 & 0 & 0 & -\frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

(a) The equations are as follows (first if t is in minutes, then in hours)

$$\frac{d}{dt} p_{WM}(t) = -\frac{1}{30} p_{WM}(t) + \frac{3}{400} p_{WA}(t)$$
$$\frac{d}{dt} p_{WA}(t) = -\frac{1}{20} p_{WA}(t) + \frac{1}{15} p_{WW}(t)$$

$$\frac{d}{dt} p_{WM}(t) = -2 p_{WM}(t) + 0.45 p_{WA}(t)$$
$$\frac{d}{dt} p_{WA}(t) = -3 p_{WA}(t) + 4 p_{WW}(t)$$

(b) First note that $p_{WW}(t) = e^{-t/15}$. Then try inserting the given formula in the second equation above:

$$LHS = \frac{d}{dt} p_{WA}(t) = -\frac{1}{5} e^{-t/20} + \frac{4}{15} e^{-t/15}$$

and

$$RHS = -\frac{1}{20} p_{WA}(t) + \frac{1}{15} p_{WW}(t) = -\frac{1}{5} e^{-t/20} + \frac{1}{5} e^{-t/15} + \frac{1}{15} e^{-t/15},$$

which is equal to the RHS, as required.

We should also check that the formula gives $p_{WA}(0) = 0$, which it does.

(c) $\frac{d}{dt}\left[e^{t/30} p_{WM}(t) \right] = \frac{3}{400} e^{t/30}.4\left[e^{-t/20} - e^{-t/15} \right]$ with $p_{WM}(0) = 0$ implies that $e^{t/30} p_{WM}(t) = 0.9 - 1.8e^{-t/60} + 0.9e^{-t/30}$, which simplifies to $p_{WM}(t) = 0.9e^{-t/30} - 1.8e^{-t/20} + 0.9e^{-t/15}$.

(iii) (a) The expected length of time spent in state W is 15 mins, after which there is a transition to state A with probability 1.

(b) The other equations are:

$$T_A = 20 + 0.15\,T_M + 0.05\,T_S$$

$$T_M = 30$$

$$T_S = 180$$

(c) Solving these equations gives $T_A = 20 + 4.5 + 9 = 33.5$ and $T_W = 48.5$ mins

A number of students suggested that $p_{WW}(t) = 1 - p_{WA}(t)$, which works in the two-state case. In this example, however, it is only possible to state that $p_{WW}(t) + p_{WA}(t) + p_{WM}(t) + p_{WS}(t) + p_{WH}(t) = 1$, which is not the same.
It was disappointing that not many candidates made substantial progress with solving the differential equations, but the last part of the question was in general well done.

10 (i) The Wiener process can be defined as $W_t = \mu t + \sigma B_t$. In this case $\sigma = 1$.

(ii) We need to show that

$$E\left[S_t \mid S_s\right] = S_s, \qquad 0 < s < t.$$

as well as proving that $E[|S_t|] < \infty$.

But $W(t) \sim N(\mu t, t)$, so that $M_t(x) = e^{\mu x + tx^2/2}$

We have

$$E\left[S_t \mid S_s\right] = E\left[e^{-2\mu W(t)} \mid S_s\right]$$

$$= E\left[e^{-2\mu\left(W(t) - W(s) + W(s)\right)} \mid S_s\right]$$

$$= e^{-2\mu W(s)} E\left[e^{-2\mu(W(t-s))}\right] \qquad (1)$$

using stationarity and independence of the increments.

From the definition of a Wiener process with drift above, we have

$$W(t-s) \sim N(\mu(t-s), (t-s)), \qquad (2)$$

so

$$E\left[e^{-2\mu(W(t-s))}\right] = M_{t-s}(-2\mu),$$

where M_{t-s} is the moment generating function of the normal distribution in (2). But for this distribution we know that $M_{t-s}(x) = e^{\mu(t-s)x + (t-s)x^2/2}$.

Therefore

$$M_{t-s}(-2\mu) = e^{-2\mu^2(t-s) + (t-s)(2\mu)^2/2} = 1;$$

now (1) shows that $E\left[S_t \mid S_s\right] = S_s$, as required.

Finally, check the expectation: $S_t > 0$, so

$$E[|S_t|] = E[S_t] = E[e^{-2\mu W(t)}] = M_t(-2\mu) = e^{\mu t(-2\mu) + t(-2\mu)^2/2} < \infty.$$

(iii) (a) $P[S(t) < a] = P[e^{-2\mu(B(t) + \mu t)} < a] = P[-2\mu B(t) < b + 2\mu^2 t]$, where $b = \log a$. If $\mu > 0$, this becomes

$$P\left[B(t) > -\frac{b}{2\mu} - \mu t\right] = \Phi\left[\mu\sqrt{t} + \frac{b}{2\mu\sqrt{t}}\right],$$

whereas in the case $\mu < 0$ we have

$$P\left[B(t) < -\frac{b}{2\mu} - \mu t\right] = \Phi\left[|\mu|\sqrt{t} + \frac{b}{2|\mu|\sqrt{t}}\right],$$

and in both cases the RHS tends to 1 as $t \to \infty$.

(b) $P[T_a > t] \le P[S(t) > a] \to 0$ as $t \to \infty$.

By definition, $S(T_a) = a$. Therefore $E[S(T_a)] = a$.

(c) This is not a contradiction because the conditions of the optional stopping theorem are not satisfied. Neither $S(t)$ nor T_a is bounded above, even though $S(t)$ is a convergent martingale.

(iv) (a) In this case T_a is only finite if $S(t)$ hits a, which is not certain. However, as above it is certain that $S(t) \to 0$ almost surely.

Therefore

$$S(T_a) = \begin{cases} a & \text{if } T_a < \infty \\ 0 & \text{if } T_a = \infty \end{cases}.$$

It follows that $E[S(T_a)] = aP[T_a < \infty]$.

(b) Now the optional stopping theorem applies, since $S(t \wedge T_a)$ is bounded below by 0 and above by a.

We may deduce that

$$1 = S(0) = aP[T_a < \infty], \quad \text{i.e. that } P[T_a < \infty] = \frac{1}{a}.$$

In part (ii) a large number of candidates did not even attempt to prove that $E(|S_t|) < \infty$ *: this condition is a requirement for S to be a martingale and should not be omitted. However, most candidates had a good idea of how to prove that S satisfied the conditional expectation condition.*

Parts (iii) and (iv) attracted at best sketchy answers. The examiners were unsure whether this was due to pressure of time or to lack of familiarity with applications of the optional stopping theorem.

END OF EXAMINERS' REPORT