Faculty of Actuaries **Institute of Actuaries**

REPORT OF THE BOARD OF EXAMINERS ON THE EXAMINATIONS HELD IN

April 2002

Subject 103 — Stochastic Modelling

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.

K Forman
Chairman of the Board of Examiners

11 June 2002

1 (i) Y_k is a martingale, and the values $Y_1, \ldots, Y_k$ are determined by $X_1, \ldots, X_k$, so

$$E[Y_{k+1} \mid X_1, X_2, \ldots, X_k] = E[Y_{k+1} \mid Y_1, Y_2, \ldots, Y_k] = Y_k.$$

Now $E[Y_{k+1}|X_1, X_2, \ldots, X_k] = E\left[\alpha_{k+1}.\exp\sum_{j=1}^{k+1} X_j \mid X_1, X_2, \ldots, X_k\right]$

$$= \alpha_{k+1}.\, E\left[e^{X_{k+1}}.\exp\sum_{j=1}^{k} X_j \mid X_1, X_2, \ldots, X_k\right]$$

$$= \alpha_{k+1}.E\left[e^{X_{k+1}}\right].\frac{Y_k}{\alpha_k}$$

We therefore require

$$\alpha_{k+1} = \frac{\alpha_k}{E\left[e^{X_{k+1}}\right]} = \alpha_k \exp\left[-\mu_{k+1} - \frac{1}{2}\sigma_{k+1}^2\right]$$

(ii) The solution depends critically on the moment generating function of a Normal variable. Non-normal variables have different mgfs, so the answer obtained would be different.

Answers to Question 1 were disappointing considering how straightforward it was, suggesting that candidates lacked practice at applications of martingales.

2 (i) $\dfrac{d}{db}(be^{-cb}) = (1 - cb)e^{-cb} = 0$ when $b = 1/c$.

Check that this is a maximum: $\dfrac{d^2}{db^2}(be^{-cb}) = (-2c + c^2 b)e^{-cb} = -ce^{-1}$ when $b = 1/c$. This is negative, so X is indeed maximised at $b = 1/c$, giving a maximum value of e^{-1}/c.

(ii) Itô's Lemma has a number of possible forms:

$df(X_t) = f'(X_t)dX_t + ½ f''(X_t)(dX_t)^2$, or

$$df(X_t, t) = \frac{\partial f}{\partial t}(X_t, t)dt + \frac{\partial f}{\partial x}(X_t, t)dX_t + ½\frac{\partial^2 f}{\partial x^2}(X_t, t)(dX_t)^2$$

are OK, with $(dX_t)^2 = dt$.

$\frac{d^2}{db^2}(be^{-cb}) = (-2c + c^2b)e^{-cb}$. Therefore

$$dX_t = (1 - cB_t)\ e^{-cB_t}\, dB_t + ½(-2c + c^2B_t)\, e^{-cB_t}\, dt.$$

(iii) Quite the reverse. If B is large and negative X takes enormously negative values. X is not in the least stationary.

A common cause of lost marks was failing to check that the turning point was a maximum.

3 (i) A linear trend means that the line of best fit to the data plotted against time would have a non-zero slope or that there is evidence from the observations that there is an underlying tendency for the values to increase or decrease with time at a constant rate.

Seasonal variation is another deterministic component of the mean which causes $E(X_t)$ to depend on the remainder when t is divided by the period, d; to spot it from the data, look for recurring patterns in the data or check the sample ACF.

(ii) **Either** use moving averages: set $Y_t = ½(X_t + X_{t-1})$, which has had the seasonal variation smoothed out.

Or use seasonal differencing: set $Y_t = \nabla_2 X_t = X_t - X_{t-2}$ (from the Box-Jenkins armoury)

Or use any linear filter $Y_t = \sum_j a_j X_{t-j}$ as long as $\sum_{\text{even } j} a_j = \sum_{\text{odd } j} a_j$ (any such filter does answer the question, though it may look very strange)

Or method of seasonal means: estimate a mean for the even-numbered observations and another for the odd-numbered ones, then subtract these from the corresponding observations to obtain a set of residuals, which can then be analysed.

(iii) Set $Y = \nabla(\log X) = \nabla(a + bt + Z_t) = b + \nabla Z_t$.

Since it is stated that Z is $I(1)$ it follows that ∇Z is stationary.

Question 3 was generally well answered.

4 (i) In terms of the backwards shift operator we have

$$(1 + 2\alpha B - \alpha^2 B^2)Y = Z.$$

We must find the values of α such that the roots of the polynomial $1 + 2\alpha x - \alpha^2 x^2$ lie outside the unit circle.

The roots are $\frac{1}{\alpha}\left(1 \pm \sqrt{2}\right)$, so we require that $\frac{\sqrt{2}+1}{|\alpha|} > 1$ and $\frac{\sqrt{2}-1}{|\alpha|} > 1$, in other words that $|\alpha| < \sqrt{2}-1$.

(ii) $Y_t = -2\alpha Y_{t-1} + \alpha^2 Y_{t-2} + Z_t$

$$\text{Cov}[Y_t, Y_t] = \gamma_0 = -2\alpha\gamma_1 + \alpha^2\gamma_2 + \sigma^2 \qquad (1)$$

$$\text{Cov}[Y_t, Y_{t-1}] = \gamma_1 = -2\alpha\gamma_0 + \alpha^2\gamma_1 \qquad (2)$$

$$\text{Cov}[Y_t, Y_{t-2}] = \gamma_2 = -2\alpha\gamma_1 + \alpha^2\gamma_0 \qquad (3)$$

From (2); $\gamma_1 = -\dfrac{2\alpha\gamma_0}{1-\alpha^2}$ (4)

Substitute for γ_1 from (4) into (3)

$$\gamma_2 = 2\alpha \cdot \frac{2\alpha\gamma_0}{1-\alpha^2} + \alpha^2\gamma_0 = \gamma_0 \cdot \left(\frac{5\alpha^2 - \alpha^4}{1-\alpha^2}\right) \qquad (5)$$

substitute for γ_1 from (4) and γ_2 from (5) into (1)

$$\Rightarrow \qquad \gamma_0 = \frac{\sigma^2\left(1-\alpha^2\right)}{\left(1+\alpha^2\right)\left(1-6\alpha^2+\alpha^4\right)} \qquad (6)$$

substitute for γ_0 from (6) into (4) and (5) to find γ_1 and γ_2

$$\Rightarrow \qquad \gamma_1 = \frac{-2\alpha\sigma^2}{\left(1+\alpha^2\right)\left(1-6\alpha^2+\alpha^4\right)}$$

and

$$\gamma_2 = \frac{\left(5\alpha^2-\alpha^4\right).\sigma^2}{\left(1+\alpha^2\right)\left(1-6\alpha^2+\alpha^4\right)}$$

(Alternative form for the denominator: $1 - 5\alpha^2 - 5\alpha^4 + \alpha^6$.)

Generally well answered, although the exact range of permitted values for α in (i) caused difficulties.

5 (i) If $r_1, r_2, r_3, \ldots$ were all close to 1, that would indicate a need for differencing. That is not the case here.

(ii) The ACF of an AR(1) is geometrically decreasing. That is approximately the case here, so AR(1) is not completely unreasonable, but we need a plot of the sample PACF to check.

(iii) There is a fairly significant departure from a white noise process. The model does not appear to fit.

(iv) (a) $\hat{x}_{30}(1) = 0.49541 + 0.5118 x_{30}$,
$\hat{x}_{30}(2) = 0.49541 + 0.5118 \hat{x}_{30}(1) = 0.74896 + 0.2619 x_{30}$.

[Numerical values are acceptable here: x_{30} is in fact equal to 1.01, so the forecasts would be $\hat{x}_{30}(1) = 1.012$, $\hat{x}_{30}(2) = 1.014$: a margin of error is acceptable, as x_{30} must be read off the graph.]

(b) The forecasts are unreliable:

There are only 30 observations in the series, so the confidence intervals would be quite wide in any case.

In addition, we have seen that the model is inadequate, casting further doubt on them.

Many candidates provided interesting and sensible comments on the data provided, though the problems caused by the small sample size were not generally recognised.

6 (i) The generator is

$$\begin{pmatrix} -\alpha - 4\lambda & 3\lambda & \lambda & \alpha \\ 5\lambda & -\alpha - 7\lambda & 2\lambda & \alpha \\ 0 & 0 & -2\alpha & 2\alpha \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

(ii) The required event is that either no jump out of A has taken place by t or that a jump to D has taken place. This has probability

$$e^{-(\alpha+4\lambda)t} + (1 - e^{-(\alpha+4\lambda)t})\,\frac{\alpha}{\alpha+4\lambda}.$$

(iii) The backward equations state $P'(t) = AP(t)$.

We have

$$\frac{d}{dt}p_{PD}(t) = -2\alpha p_{PD}(t) + 2\alpha$$

(iv) The solution to $p'_{PD}(t) = -2\alpha p_{PD}(t) + 2\alpha$ with $p_{PD}(0) = 0$ is $p_{PD}(t)$ $= 1 - e^{-2\alpha t}$.

Well done, on the whole. Where there were difficulties they may have been due to practising more with the time-inhomogeneous version of the equation rather than the time-homogeneous one.

7 (i) **Either** $P(X > x) = P(U < e^{-4x}) = e^{-4x}$, so that $f_X(x) = 4e^{-4x}$,
Or probability distribution function $F(x) = 1 - e^{-\lambda x}$ has inverse $F^{-1}(y) = -\frac{1}{\lambda}\log(1 - y)$. Hence the inversion method reads:

(1) Generate y from $U(0, 1)$.

(2) Return $x = -\frac{1}{\lambda}\log(1 - y)$.

$-\frac{1}{\lambda}\log y$ is as good as $-\frac{1}{\lambda}\log(1 - y)$ here, since $1 - Y$ is also $U(0, 1)$.

(ii) (a) Obtain numbers $y_1, y_2, \ldots, y_n$ by the above procedure, being outcomes of random variables $Y_1, Y_2, \ldots, Y_n$ independent exponentially distributed with parameter $\lambda = 4$.

Put $t_j = y_1 + y_2 + \ldots + y_j$ and return

$$x_t = 0 \text{ if } t < t_1,$$
$$x_t = j \text{ if } t_j \le t < t_{j+1}.$$

(b) Take the value of the Poisson process at time 1.

(iii) Calculate $F(i)$ for each $i = 0, 1, \ldots$. Then

(1) Generate u from $U(0, 1)$.

(2) Return the smallest i such that $u \le F(i)$.

(iv) The method in (ii) requires an average of four uniform r.v.s per Poisson r.v.; the method in (iii) requires only one. Since the distribution function needs to be calculated only once, (iii) should be much more efficient.

Very poorly answered, considering that it deals only with simulating standard exponential and Poisson random variables. Even the inverse distribution function method was largely misunderstood. There must be the suspicion that many candidates are not getting as far as Unit 7 in the Core Reading.

8 (i) First choose a class of model which might be supposed, for physical reasons, to provide a reasonable fit to the data. Identify the parameters of the model.

Next estimate the values of the parameters from the data.

See if the observed data match the pattern which would be expected if the model were accurate and if the parameters had the values given by their estimates. If not, the model should be revised.

(ii) (a) Exponential distribution in each case, with rate σ in H, ρ in S.

(b) The time spent in state H before the next visit to S has mean σ^{-1}.

Therefore a reasonable estimate for σ is the reciprocal of the mean length of each visit: $\hat{\sigma}$ = (Number of transitions from H to S)/(Total time spent in state H up until the last transition from H to S), although it would be equally valid to use the Maximum Likelihood Estimator, which is (Number of transitions from H to S)/(Total time spent in state H).

Similarly for $\hat{\rho}$.

(c) Testing whether the successive holding times are independent exponential variables would be best, and any procedure which does test this is acceptable. Something like using the χ^2 goodness-of-fit test on the even-numbered holding times, then again on the odd-numbered ones, springs to mind, but there may be other, equally reasonable, answers.

(iii) (a) For a time-inhomogenous model the transition rates σ and ρ are functions of t.

It is certainly possible to improve the fit by using a time-inhomogenous model in this instance.

(b) If the age profile is represented by a density function $f(a)$; then the overall average rate at which a healthy employee falls sick is $\sigma = \int f(a)\sigma(a)da$, roughly constant for all t. The same of course applies to the overall average rate of recovery. (It is not necessary to write down the integral to obtain full marks: any explanation which covers the basic principle will suffice.)

Generally good answers. A number of candidates answered part (i) in the context of this particular model rather than in general terms and so were only awarded a reduced number of marks.

9 (i)

	Level at start of this year after:			
Level at start of prev yr	*0 claims in previous yr*	*1 claims in previous yr*	*2 claims in previous yr*	*3 or more claims in previous yr*
5	4	5	5	5
4	3	5	5	5
3	2	4	5	5
2	1	3	4	5
1	1	1	2	5

For each policyholder, the number of claims in each year has a Poisson (0.25) distribution. So

$P(0 \text{ claims}) = e^{-0.25} = 0.7788$

$P(1 \text{ claim}) = 0.25\, e^{-0.25} = 0.1947$

$P(2 \text{ claims}) = 0.25^2 \,.\, \dfrac{e^{-0.25}}{2} = 0.0243$

$P(3 \text{ or more claims}) = 1 - (0.7788 + 0.1947 + 0.0243) = 0.0022$

$$\text{Transition matrix } P = \begin{pmatrix} 0.9735 & 0.0243 & 0 & 0 & 0.0022 \\ 0.7788 & 0 & 0.1947 & 0.0243 & 0.0022 \\ 0 & 0.7788 & 0 & 0.1947 & 0.0265 \\ 0 & 0 & 0.7788 & 0 & 0.2212 \\ 0 & 0 & 0 & 0.7788 & 0.2212 \end{pmatrix}$$

(ii) In order to be in level 1 in year 3 the policyholder requires two consecutive claim-free years. The probability of this is $0.7788^2 = 0.6065$.

A similar argument can be used for the probability of being in level 3 in year 3, but it may be simpler to calculate the whole vector of probabilities x_3.

$$\begin{aligned} x_1 &= \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \end{pmatrix} \\ x_2 &= \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \end{pmatrix}.P = \begin{pmatrix} 0 & 0.7788 & 0 & 0.1947 & 0.0265 \end{pmatrix} \\ x_3 &= \begin{pmatrix} 0 & 0.7788 & 0 & 0.1947 & 0.0265 \end{pmatrix}.P \\ &= \begin{pmatrix} 0.6065 & 0 & 0.3033 & 0.0396 & 0.0506 \end{pmatrix} \end{aligned}$$

Probability of being in level 3 is 30.33%

(iii) (a) The required conditions are that the chain is irreducible and aperiodic.

(b) Irreducibility: level i can be reached from level j in $|j - i|$ steps;
Aperiodicity: $p_{ii} > 0$ for some i.

(c) The stationary distribution π will not depend on the starting position.

Require $\begin{pmatrix} \pi_1 & \pi_2 & \pi_3 & \pi_4 & \pi_5 \end{pmatrix} P = \begin{pmatrix} \pi_1 & \pi_2 & \pi_3 & \pi_4 & \pi_5 \end{pmatrix}$

This gives the following equations:

$$0.9735\,\pi_1 + 0.7788\pi_2 = \pi_1 \qquad (1)$$

$$0.0243\pi_1 + 0.7788\pi_3 = \pi_2 \qquad (2)$$

$$0.1947\pi_2 + 0.7788\pi_4 = \pi_3 \qquad (3)$$

$$0.0243\pi_2 + 0.1947\pi_3 + 0.7788\pi_5 = \pi_4 \qquad (4)$$

and

$$\pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 = 1 \qquad (5)$$

Solving the simultaneous equations:

from (1) $\Rightarrow \pi_2 = \dfrac{1-0.9735}{0.7788}\pi_1 = 0.0340\pi_1$;

substitute for π_2 into (2) $\Rightarrow \pi_3 = \dfrac{0.0340-0.0243}{0.7788}\pi_1 = 0.01244\pi_1$;

substitute for π_2 and π_3 into (3)

$\Rightarrow \pi_4 = \dfrac{0.01244-0.1947\times 0.0340}{0.7788}\pi_1 = 0.00747\pi_1$;

substitute for π_2, π_3 and π_4 into (4)

$$\Rightarrow \pi_5 = \frac{0.00747 - 0.1947 \times 0.01244 - 0.0243 \times 0.0340}{0.7788} \pi_1 = 0.00541\pi_1$$

and substituting for π_2, π_3, π_4 and π_5 into (5) $\Rightarrow \pi_1 = 0.9440$

(and $\pi_2 = 0.0321$, $\pi_3 = 0.0117$, $\pi_4 = 0.0071$ and $\pi_5 = 0.0051$).

(iv) A chi-squared goodness-of-fit test is best here.

Very good answers on the whole. Some confusion was caused by the fact that the states were presented in the reverse of the standard order, but most candidates coped with this pretty well.

10 (i) A Lévy process Y_t can be decomposed as $Y_t = y_0 + \mu t + \sigma B_t + N_t$, where μt is a deterministic component, σB_t a continuous random component (Brownian motion) and N_t a purely discontinuous component which may be regarded as a compound Poisson process, independent of the Brownian component.

(ii) (a) $E(M_{t+s} \mid \mathcal{F}_t) = \exp(-2ab - 2b^2(t+s) + 2bB_t)\, E(e^{2b(B_{t+s}-B_t)} \mid \mathcal{F}_t)$. The independent increment property implies that this conditional expectation is $\exp(2b^2 s)$.

Therefore $E(M_{t+s} \mid \mathcal{F}_t) = \exp(-2ab - 2b^2(t+s) + 2bB_t + 2b^2 s) = M_t$.

Ought to check that $E(|M_t|) < \infty$. Since $M_t \geq 0$, $E(|M_t|) = E(M_t) = e^{-2ab}$.

$M_t \geq 0$ by definition. Since B_t is continuous, it follows that M_t is continuous. Further, $M_0 < 1$. Therefore M_t cannot exceed 1 without first passing through 1, which does not happen until time T.

(b) The optional stopping theorem states that, for any martingale Y and stopping time T adapted to the same filtration, $EY_T = Y_0$ if T is bounded or Y is bounded or $Y_{t \wedge T}$ is bounded.

The last of these conditions holds in this case.

We conclude that $P(B \text{ hits } a + bt) = P(M \text{ hits } 1) = E(M_T) = M_0 = e^{-2ab}$.

(iii) ct is deterministic, $-X_t$ is a compound Poisson process with constant jump height $-k$ and the multiplier of the Brownian component is $\sigma = 0$.

(iv) Since $X_t/k \sim P(\lambda)$, we have $E(-X_t) = -\lambda kt$, $\text{Var}(-X_t) = k^2\lambda t$. Therefore $E(S_t) = s_0 + (c - k\lambda)t$, $\text{Var}(S_t) = k^2\lambda t$.

(v) (a) In order for the mean and variance to match we require $\mu = c - k\lambda$, $\sigma = k\sqrt{\lambda}$.

(b) $c > k\lambda$ is the condition for premium income to outstrip outgoings on average.

(c) $s_0 + \mu t + \sigma B_t = 0$ if and only if $B_t = -\dfrac{s_0}{\sigma} - \dfrac{\mu}{\sigma}t$.

From above, the probability that B ever hits the line $a + bt$ is e^{-2ab}. Therefore the required approximation to the probability of ruin is $\exp\left(-2\dfrac{\mu s_0}{\sigma^2}\right) = \exp\left(-2\dfrac{(c-k\lambda)s_0}{k^2\lambda}\right)$ if $c > k\lambda$.

(d) The difference is that S has a discontinuous component, whereas S^* approximates this with a continuous one.
The difference will be significant if s_0 is small and $c - k\lambda$ large, as then S^* will be much less likely to hit 0 than S; in the opposite case the approximation may be quite reasonable.

Most candidates were not able to attempt every part of this question. Very few made the connection between part (ii)(b) and part (v)(c). There seems to be a general lack of confidence when dealing with martingales or Lévy processes.