EXAMINATIONS

April 2001

Subject 103 — Stochastic Modelling

EXAMINERS' REPORT

1 (i) Given $\mathcal{F}_t$ we know that $N(t + s) - N(t) \sim \text{Poisson}(\lambda s)$.

Hence $E(\theta^{N(t+s)} \mid \mathcal{F}_t) = \theta^{N(t)} e^{(\theta-1)\lambda s}$.

(ii) Now $E(\eta(t + s)\, \theta^{N(t+s)} \mid \mathcal{F}_t) = \eta(t + s)\, \theta^{N(t)} e^{(\theta-1)\lambda s}$, which needs to be equal to $M(t) = \eta(t)\, \theta^{N(t)}$. It follows that $\eta(t) = e^{-(\theta-1)\lambda t}$.

2 (i) The inter-arrival times are much more suitable because they are independent.

(ii) They should be exponentially distributed with the same mean.

Kolmogorov-Smirnov, Anderson-Darling or χ^2 goodness-of-fit test can all be used.

(iii) Successive values should be independent.

Regress X_i on X_{i-1} using ordinary least squares, or fit an AR(1) and test the α_1 parameter for significance (equivalent to Durbin-Watson test).

3 (i) Spectral density $f(\omega) = \frac{1}{2\pi} \Sigma_{-\infty}^{\infty}\ \gamma_k e^{ik\omega}$, or equivalent.

For MA(1), therefore, we have

$$f(\omega) = \frac{\sigma_e^2}{2\pi}(1 + \beta^2 + 2\beta \cos \omega).$$

And for AR(1),

$$f(\omega) = \frac{\sigma_e^2}{2\pi} \frac{1}{1 + \alpha^2 - 2\alpha \cos \omega}.$$

(ii) Clearly from (i) the inverse of the MA(1) is an AR(1), with $\alpha = -\beta$ and with a different value of σ_e^2.

The word "invertible" attached to a MA(1) indicates that the inverse is a stationary AR(1), whereas a non-"invertible" MA(1) has as inverse an AR(1) model which cannot be stationary, such as $X_t = 2X_{t-1} + e_t$.

4 (i) $E(dY_t \mid Y_t = y) = \mu(y)h + o(h)$, $\mathrm{Var}(dY_t \mid Y_t = y) = h + o(h)$

(ii) $E(X_{n+1} - X_n \mid X_n = y) = \theta - \alpha y$, $\mathrm{Var}(X_{n+1} - X_n \mid X_n = y) = \tau^2$.

(iii) $\mu(y)h = (\theta - \alpha y)$ and $h = \tau^2$, so $\mu(y) = (\theta - \alpha y) / \tau^2$.

(iv) The increments of a brownian motion do not depend on its current value, i.e. $\alpha = 0$.

(v) An OU process drifts towards zero, so that $\theta = 0$.

5 (i) Let γ_k denote the autocovariance function of X. Then

$\mathrm{Cov}\,(X_t, e_t) = 0 + \sigma^2 + 0 = \sigma^2$;

$\mathrm{Cov}\,(X_t, e_{t-1}) = \alpha\gamma_0 + 0 + \beta\sigma^2$;

$\gamma_2 = \alpha\gamma_1$

$\gamma_1 = \alpha\gamma_0 + 0 + \beta\,\mathrm{Cov}\,(X_{t-1}, e_{t-1}) = \alpha\gamma_0 + \beta\sigma^2$

$\gamma_0 = \alpha\gamma_1 + \mathrm{Cov}\,(X_t, e_t) + \beta\,\mathrm{Cov}\,(X_t, e_{t-1}) = \alpha^2\gamma_0 + (1 + 2\alpha\beta + \beta^2)\,\sigma^2$,

implying that

$$\gamma_0 = \frac{\sigma^2}{1-\alpha^2}(1 + 2\alpha\beta + \beta^2),$$

$$\rho_1 = \frac{(\alpha+\beta)(1+\alpha\beta)}{1 + 2\alpha\beta + \beta^2},\ \rho_2 = \alpha\rho_1.$$

(ii) Estimate of α is $r_2 / r_1 = 0.8$; estimate of β is given by $\frac{1}{2}(1 + 2\alpha\beta + \beta^2) = (\alpha + \beta)(1 + \alpha\beta)$, or $0.3\beta^2 + 0.84\beta + 0.3 = 0$, with solution $\beta = -1.4 \pm \sqrt{0.96}$.

In this case we take the positive square root to ensure invertibility.

6 (i) Expectation is £2.4m. Safety loading is

$$\rho = \frac{C - \alpha\mu}{\alpha\mu} = \frac{(400/12) \times 80{,}000 - 2{,}000 \times 1{,}200}{2{,}000 \times 1{,}200} = 11.11\%$$

where α denotes the mean arrival rate of claims and μ the mean claim size.

(ii) Conditions for validity of diffusion approximation are $\frac{u}{\mu}$ large, ρ small, $\rho\frac{u}{\mu}$ moderate, where u is the reserve.

In this case $\frac{u}{\mu}$ is large, ρ is not particularly small and $\rho\frac{u}{\mu} = \frac{2\times 10^7 \times 11.1 \times 10^{-2}}{1{,}200}$ = 1,852, clearly too large. We conclude that the diffusion approximation is not appropriate.

(iii) Decide on a quantum of time, which may be a month or may be smaller. For each time period generate a Poisson variate to indicate the number of claims received and, conditional on this, a Normal variate with appropriate mean and variance to represent the total sum claimed. Subtract this from the total premium income over the period (deterministic), using the resulting quantity as the increment of the surplus process. Run the simulation for an extended period of time, stopping if/when it goes below zero. A large number of simulations should be performed, with the probability of ruin being estimated using standard techniques based on the Binomial distribution.

The importance of reproducibility is for sensitivity analysis. The estimated probability may depend heavily on the values assumed for mean and standard deviation of the claim size, or on other numerical parameters. It is necessary to vary the initial assumptions and run the simulation again, just to ensure that conclusions are not substantially changed if the parameter values used do not adequately reflect the actual conditions experienced.

7 (i) $P\{S_t \le x\} = P\{\exp(\mu t + \sigma B_t) \le x\} = P\{\mu t + \sigma\sqrt{t}N \le \ln(x)\} = \Phi\left(\frac{\ln(x) - \mu t}{\sigma\sqrt{t}}\right)$, where Φ denotes the standard Normal distribution function.

(ii) We have to find m so that $P\{S_t \le m\} = P\{e^{\mu t + \sigma B_t} \le m\} = P\{\mu t + \sigma B_t \le \ln(m)\}$ $= P\{\sigma B_t \le \ln(m) - \mu t\} = \frac{1}{2}$. Since σB_t is a symmetric normal variable, its median is 0 and the last equation can only be satisfied if $\ln(m) - \mu t = 0$ and $m = e^{\mu t}$.

The expectation is $ES_t = Ee^{(\mu t + \sigma B(t))} = Ee^{\mu t} e^{\sigma\sqrt{t}N} = e^{\mu t}\, e^{\frac{\sigma^2}{2}t} = e^{\left(\mu + \frac{\sigma^2}{2}\right)t}$.

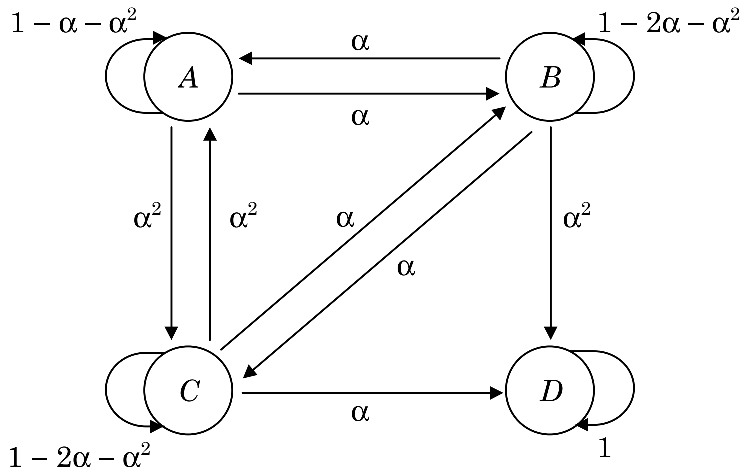
(iii) By the same token, $E(S_t \mid F_s) = S(s)\ E(e^{\mu(t-s)+\sigma(B(t)-B(s))}) = S(s)\ e^{\left(\mu+\frac{\sigma^2}{2}\right)(t-s)}$.

If S is to be a martingale, the conditional expectation must be equal to $S(s)$.

This will happen if $\mu = -\frac{1}{2}\sigma^2$.

From part (ii) we see that for this stock with initial value 1, the median of the distribution at time t goes to 0 exponentially fast for large t hence, a very bad investment!

8 (i) Transition Graph



(ii) All transition probabilities must lie in [0,1].
Now $1 - 2\alpha - \alpha^2 \le 1 - \alpha - \alpha^2 \le 1$ for $\alpha \ge 0$, so it suffices to ensure that $1 - 2\alpha - \alpha^2 \ge 0$ i.e. $\alpha \le \sqrt{2} - 1$. So the range of possible values of α is $[0, \sqrt{2} - 1]$.

(iii) The chain is not irreducible since D is a trap state.
The chain is aperiodic by inspection.

(iv) A stationary probability distribution, if it exists, must obey

$$\begin{aligned}
(1-\alpha-\alpha^2)\,\pi_A + \alpha\pi_B + \alpha^2\pi_C &= \pi_A \\
\alpha\pi_A + (1-2\alpha-\alpha^2)\,\pi_B + \alpha\pi_C &= \pi_B \\
\alpha^2\pi_A + \alpha\pi_B + (1-2\alpha-\alpha^2)\,\pi_C &= \pi_C \\
\alpha^2\pi_B + \alpha\pi_C + \pi_D &= \pi_D
\end{aligned}$$

The last equation implies $\pi_B = \pi_C = 0$, and this in turn shows that $\pi_A = 0$.
Hence the stationary probability distribution is $\pi = (0, 0, 0, 1)^T$.

It is unique: there is just one recurrent class and it is aperiodic. (Or point out that there is no other solutions to the equations.)

(v) With $\alpha = 0.1$, the transition matrix is

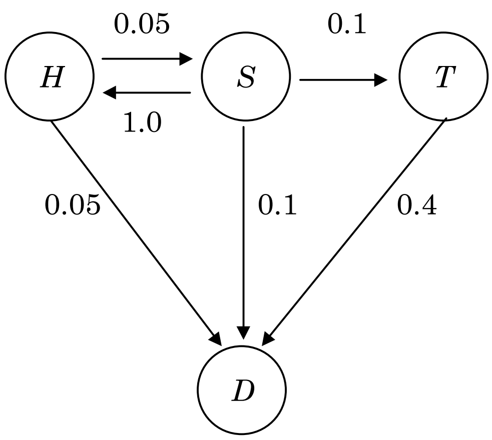
$$\begin{pmatrix} 0.89 & 0.1 & 0.01 & 0 \\ 0.1 & 0.79 & 0.1 & 0.01 \\ 0.01 & 0.1 & 0.79 & 0.1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Its square is

$$\begin{pmatrix} 0.8022 & 0.169 & 0.0268 & 0.002 \\ 0.169 & 0.6441 & 0.159 & 0.0279 \\ 0.0268 & 0.159 & 0.6342 & 0.18 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

the relevant entries being the last column.

9 (i) Transition Graph:



(ii) KFE: $P'(t) = P(t)\,A$,

$$A = \begin{pmatrix} -0.1 & 0.05 & 0 & 0.05 \\ 1.0 & -1.2 & 0.1 & 0.1 \\ 0 & 0 & -0.4 & 0.4 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

(iii) The probability of staying in state H for 10 years is $\int_{10}^{\infty} 0.1e^{-0.1x}\,dx = e^{-1}$.

(iv) First transition from H must be to S or D, each equally likely. If to D, then it is certain that no terminal illness will occur; otherwise, the probability of avoiding a terminal illness is d_S.

From S similarly, except that the transition probabilities are to H with prob. $\frac{1.0}{1.2} = \frac{5}{6}$, to D or to T, each with prob. $\frac{0.1}{1.2} = \frac{1}{12}$. Once in T it is not possible to avoid terminal illness.

Solving the above equations, $d_S = \frac{1}{12} + \frac{5}{6} \times \frac{1}{2}(1 + d_S)$, implying that $d_S = \frac{6}{7}$, $d_H = \frac{13}{14}$.

(v) The Markov property implies that the time spent in state T has exponential distribution. The rate is 0.4 per year, so the expectation is 2.5 years.

The expected time spent in terminal illness given current health is $\mathbb{P}$(ever hit $T \mid X_0 = H) \times 2.5$ years $= \frac{2.5}{14}$ years.

10 (i) $(1 - B)\ Y = \beta B\ (1 - B)\ I$,
with solution $Y = \beta BI$ + const

(ii) We have the vector equation

$$\begin{pmatrix} X_n \\ I_n \end{pmatrix} = \begin{pmatrix} 0 & (1+\pi)\beta \\ 0 & 1+\alpha \end{pmatrix} \begin{pmatrix} X_{n-1} \\ I_{n-1} \end{pmatrix} + \begin{pmatrix} \text{const} \\ 0 \end{pmatrix} + \begin{pmatrix} e_n^{(1)} \\ e_n^{(2)} \end{pmatrix},$$

which clearly represents a vector AR(1).

(iii) I is not stationary: the condition for an AR(1) to be stationary is that the autoregressive parameter is less than 1 in absolute value.

I is not $I(1)$, either, since $\nabla I = e^{(2)} + \alpha BI$ which, as already stated, is not stationary.

Z is therefore neither $I(0)$ nor $I(1)$.

(iv) The equation for the sum of squares is

$$SS = \sum_{t=2}^{n} (e_t^{(2)})^2 = \sum_{t=2}^{n} (I_t - (1+\alpha)\ I_{t-1})^2.$$

Differentiating,

$$0 = -2 \sum_{t=2}^{n} I_{t-1}(I_t - (1+\alpha)\ I_{t-1}),$$

implying that

$$\hat{\alpha} = \frac{\Sigma_{t=2}^{n} I_{t-1}(I_t - I_{t-1})}{\Sigma_{t=2}^{n} I_{t-1}^2}.$$

(v) (a) First we need to obtain $N(0,1)$ variates Z_1 and Z_2. The core reading mentions two methods: either

$$Z_1 = \sqrt{-2\ln U_1}\,\sin(2\pi U_2), \qquad Z_2 = \sqrt{-2\ln U_1}\,\cos(2\pi U_2)$$

or

$$Z_1 = V_1\sqrt{\frac{-2\ln S}{S}}, \qquad Z_2 = V_2\sqrt{\frac{-2\ln S}{S}},$$

where $V_i = 2U_i - 1$, $S = \sqrt{V_1^2 + V_2^2}$ and any values of U_1 and U_2 which give $S > 1$ are rejected.

Now define $E_1 = \sigma_1 Z_1$, $E_2 = \rho\sigma_2 Z_1 + \sqrt{1-\rho^2}\;\sigma_2 Z_2$.

(b) Sensitivity analysis applies mostly to the initial assumptions. Values for the parameters α, β, π, σ_1, σ_2 and ρ must be assumed, but may not correspond exactly to the actual situation. The head of household should investigate whether making small changes to the values used will make large differences to the conclusions.

(vi) The revised model still meets the requirements set down at the start of the problem. It is likely to prove more tractable in that ln I is now a simple random walk with drift, and is therefore $I(1)$.