EXAMINATIONS

15 September 2003 (pm)

Subject 103 — Stochastic Modelling

Time allowed: Three hours

INSTRUCTIONS TO THE CANDIDATE

1. *Enter all the candidate and examination details as requested on the front of your answer booklet.*

2. *You must not start writing your answers in the booklet until instructed to do so by the supervisor.*

3. *Mark allocations are shown in brackets.*

4. *Attempt all 11 questions, beginning your answer to each question on a separate sheet.*

Graph paper is not required for this paper.

AT THE END OF THE EXAMINATION

Hand in BOTH your answer booklet, with any additional sheets firmly attached, and this question paper.

In addition to this paper you should have available Actuarial Tables and your own electronic calculator.

1 A wheel is divided into 37 equal sections, labelled from 0 to 36. A ball is rolled repeatedly around the wheel, landing in sections $X_1, X_2, \ldots$ Let M_n be the highest outcome achieved from the first n rolls, i.e. $M_n = \max_{1 \le i \le n} X_i$.

(i) Show by general reasoning that M_n is a Markov chain. [1]

(ii) Derive the transition probabilities for M_n. [2]

(iii) Determine whether M_n is irreducible and aperiodic. [1]

(iv) Determine the equilibrium probability distribution of M_n. [1]

[Total 5]

2 (i) Calculate the covariance between the values $X(t)$, $X(t+s)$ taken by a Poisson process $X(t)$ with constant rate λ at the two times t and $t+s$, where $s > 0$. [2]

(ii) Calculate the covariance between the values $B(t)$, $B(t+s)$ taken by a standard Brownian motion $B(t)$ at the two times t and $t+s$, where $s > 0$. [1]

(iii) A share price, S_t, is modelled as $S_t = \exp(Y(t))$, where

$$Y(t) = Y(0) + \sigma B(t) + \kappa(X_1(t) - X_2(t)).$$

Here $B(t)$ is a standard Brownian motion, X_1 and X_2 are rate λ Poisson processes, independent of each other and of $B(t)$, and σ and κ are constants.

(a) Give a definition of a Lévy process.
(b) Show that $Y(t)$ is a Lévy process.
(c) Calculate the covariance of $Y(t)$ and $Y(t+s)$ for $s > 0$. [4]

[Total 7]

3 An ARIMA process X satisfies the recursion

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + e_t + \beta e_{t-1},$$

where e_t is white noise with variance σ^2.

(i) (a) Write down a condition in terms of the roots of an equation for X to be stationary.

(b) Show, in the case where $\alpha_2 = -0.5$, that X is stationary as long as $|\alpha_1| < 1.5$.

[4]

(ii) Derive the spectral density function of X. [3]

[Total 7]

4 Assume that the yield rate Y_t of a certain stock may either take a high constant value y_u or a low constant value y_d, during alternating random exponential periods with respective intensities λ_u, λ_d.

(i) Write down Kolmogorov's forward equation for the probability $P_{u,u}(t)$ that Y is in state y_u at time t, given that it starts in state y_u. [1]

(ii) Show that

$$P_{u,u} = \frac{\lambda_d}{\lambda_u + \lambda_d} + \frac{\lambda_u}{\lambda_u + \lambda_d} e^{-(\lambda_u + \lambda_d)t}. \qquad [2]$$

(iii) Let U_t denote the total amount of time spent in state y_u up until time t. Derive and expression for $E[U_t \mid Y(0) = y_u]$, the expected occupation time in state y_u by time t for the two-state continuous-time Markov chain starting in state y_u.

Hint: Note that $U_t = \int_0^t I_s ds$, where $I_s = \begin{cases} 1 & \text{if } Y_s = y_u \\ 0 & \text{if } Y_s \neq y_u \end{cases}$ and use the previous result. [2]

(iv) Derive an expression for the expected total yield X_t of the stock by time t.
Hint: $X_t = y_u U_t + y_d(t - U_t)$. [2]
[Total 7]

5 Let $X_1, X_2, X_3, \ldots$ be independent, identically distributed random variables with $P(X_1 = +1) = p > ½$ and $P(X_1 = -1) = 1 - p < ½$. Let $S_0 = 0$, $S_n = X_1 + X_2 + \ldots + X_n$ for $n \geq 1$ be the associated random walk.

(i) State a necessary condition which must be satisfied by the constants θ and c if the process

$$M_n = e^{\theta S_n - cn}$$

is to be a martingale. [1]

(ii) Use the condition in (i) above to solve for θ as a function of c. [2]

(iii) Let T_1 be the first time that S_n hits 1, i.e. $T_1 = \min\{n : S_n = 1\}$. State the conditions on c and on $\theta(c)$ under which it is valid to use the Optional Stopping Theorem to evaluate $E(M_{T_1})$. [2]

(iv) Derive the moment generating function $E(e^{-cT_1})$ for $c > 0$. [3]
[Total 8]

 PLEASE TURN OVER

6 (i) Calculate the autocovariance function $\{\gamma_k : k \geq 0\}$ and the autocorrelation function $\{\rho_k : k \geq 0\}$ for the m^{th} order Moving Average process

$$X_t = \mu + \frac{1}{m+1}(e_t + e_{t-1} + \ldots + e_{t-m}),$$

where $\{e_t : t \geq 0\}$ is a sequence of uncorrelated, zero-mean random variables with common variance σ_e^2. [4]

(ii) Explain whether or not the process is invertible in the case where $m = 2$. [4]

[Total 8]

7 The data set plotted in Figure 1a represents the number of applications, x_t, for travel insurance received by an insurance company's web site, measured for a total of 60 consecutive months. Figure 1b displays the logarithm of the same data set, $y_t = \ln(x_t)$.

(i) A statistician decides, on the basis of these plots, to fit a linear time series model to y_t rather than to x_t. State, giving reasons for your answer, whether you agree with this decision. [2]

(ii) Explain what is meant by the terms "seasonal variation" and "linear trend". Outline **one** method of compensating for seasonal variation and linear trend in a data set which exhibits both. [3]

(iii) The data set z_t is a seasonally adjusted, detrended version of y_t. Figures 2a and 2b below display respectively the sample autocorrelation function and sample partial autocorrelation function of z_t.

(a) Explain what feature of the graphs enables you to conclude that it is reasonable to fit a stationary model to the data.

(b) Suggest values of p, d and q such that an ARIMA(p, d, q) model is likely to provide a good fit to the data set z_t. Give reasons for your suggestion. [3]

(iv) Explain how to produce a forecast $\hat{x}_{60}(1)$ for the value of x_{61} given the Box-Jenkins forecast $\hat{z}_{60}(1)$ for the value of z_{61}. [2]

[Total 10]

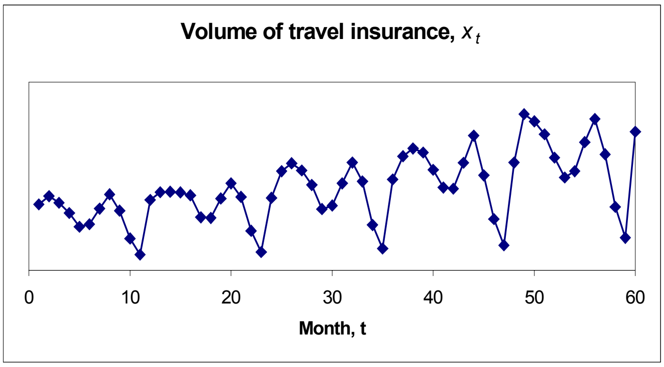

Figure 1a

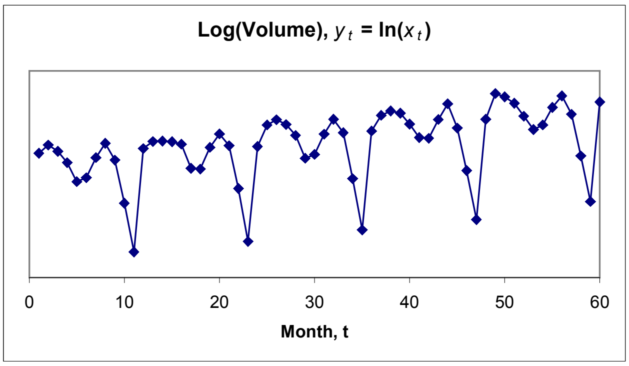

Figure 1b

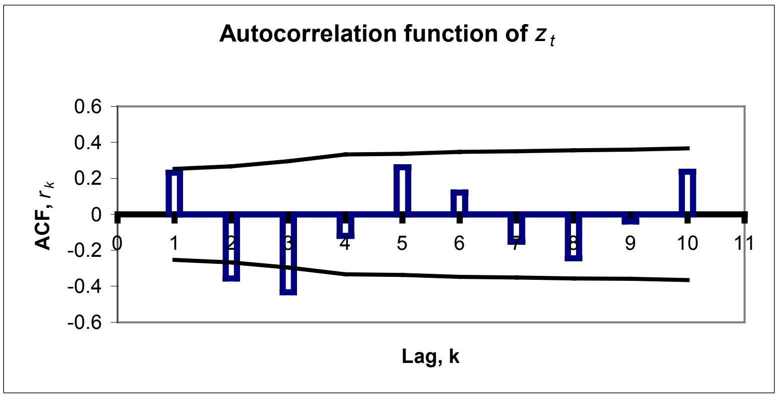

Figure 2a

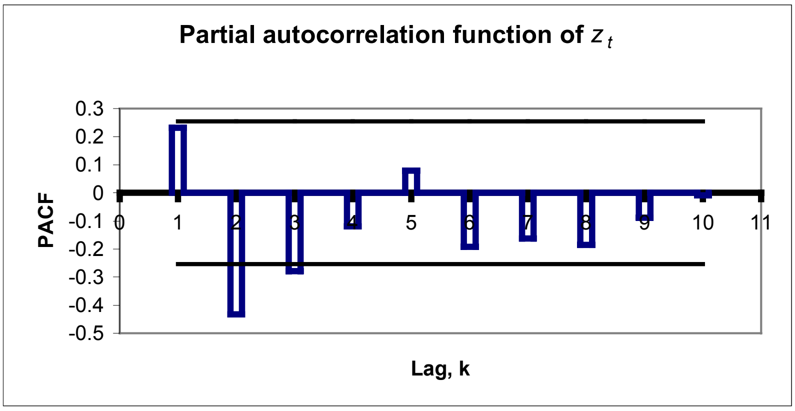

Figure 2b

PLEASE TURN OVER

8 A no-claims discount scheme has six classes of discount numbered from 0 (no discount) to 5 (maximum discount). A claim-free year results in a move to the next higher discount status (or in the retention of the maximum discount status); similarly a year with one or more claims results in a move to the next lower discount class (or in the retention of the no-discount status).

(i) Assuming that the probability of a claim-free period is α, where $0 < \alpha < \frac{1}{2}$, write down the transition graph and the transition matrix of the discrete-time Markov chain which models this scheme. [3]

(ii) (a) Write down the equations obeyed by the stationary probability distribution $\pi = (\pi_0, \pi_1, \ldots, \pi_5)$ of the Markov chain.

(b) Solve successively for $\pi_1, \pi_2, \pi_3, \ldots, \pi_5$ in terms of π_0.

(c) Derive an expression, in terms of α only, for the value of π_0.

[5]

(iii) Explain how this analysis can be used to help the insurance company to set its premium levels. [2]

[Total 10]

9 A model of mortality after retirement suggests that the age at death, T, of an individual aged 65 at retirement is a random variable satisfying

$$P(T > t) = \exp\left(-\int_{65}^{t} \mu(x)dx\right),$$

where $\mu(x)$, the force of mortality, is given by $\mu(x) = a + bx$, and a and b are positive parameters whose values are known.

(i) Determine a formula involving a uniform pseudo-random variable U for simulating the age at death, T, of a single individual currently aged 65. [5]

(ii) As part of a simulation exercise, a trainee is required to consider a sample of 10 individuals now aged 65 and to generate a single observation of Y, the number of individuals in that sample who will still be alive at age 75.

(a) Describe a method for doing this which does not involve 10 repetitions of the method in (i).

(b) Explain the main advantages of the method in (a) when a large number of simulated copies of the random variable Y are needed. [4]

(iii) The trainee is asked to investigate the effect of a reduction in the mortality rate by repeating the whole simulation for different values of a and b. State with reasons whether the trainee should use the same sequence of pseudo-random numbers as before or whether a different sequence would be preferable. [1]

[Total 10]

10 (i) Define a geometric Brownian motion and write down a stochastic differential equation which it satisfies. [2]

(ii) Using the substitution $Y_t = e^{\alpha t}\left(X_t - c\right)$, or otherwise, find a solution for the stochastic differential equation

$$dX_t = -\alpha(X_t - c)dt + \sigma dB_t$$

with the initial condition $X_0 = x_0$, where B_t is standard Brownian motion. [5]

(iii) Write down a condition satisfied by the stationary density function, $\pi(x)$, of a diffusion process, assuming that such a stationary density exists. [2]

(iv) Verify that the stationary distribution of the diffusion process in (i) is Normal, and identify the mean and variance. [2]

(v) When the equation in (iii) is applied to the geometric Brownian motion, the general solution is found to be

$$\pi(x) = \frac{A}{x} + Bx^k,$$

where $k = -2 + 2\mu/\sigma^2$ and A and B are arbitrary constants. Show that there is no probability density function $\pi(x)$ which solves the equation in (iii) in the case of the geometric Brownian motion and explain the implication of this. [2]

[Total 13]

11 A continuous-time process with three states is observed from time 0 up until the time of the 20th transition. The results may be summarised as follows:

			No. of transitions from state i to:		
State, i	*No. of visits to state i*	*Minutes spent in state i*	*State 1*	*State 2*	*State 3*
1	8	48	–	3	5
2	4	160	1	–	3
3	8	240	7	1	–

(i) Describe the stages of model fitting and model verification in the modelling process. [2]

(ii) Suppose that a Markov jump process model is to be fitted to the data set above. List all the parameters of the model and discuss the assumptions made when such a model is fitted to a data set. [4]

(iii) Estimate the parameters of the model in (ii) above and write down the estimated generator matrix. [4]

(iv) Suggest **one** test which could be applied as part of the model verification process. State the null hypothesis, H_0, identify the test statistic and name its distribution under H_0. [2]

(v) The 20th transition of the observed process takes it into state 1. Use the estimated parameter values to give point estimates of the times until the 21st and 22nd transitions. [3]

[Total 15]