EXAMINATIONS

11 September 2000 (pm)

Subject 103 — Stochastic Modelling

Time allowed: Three hours

INSTRUCTIONS TO THE CANDIDATE

1. *Write your surname in full, the initials of your other names and your Candidate's Number on the front of the answer booklet.*
2. *Mark allocations are shown in brackets.*
3. *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 and this question paper.

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

1 There are N individuals in a population, some of whom have a certain infection that spreads as follows. Contacts between two members of this population occur in accordance with a Poisson process having rate λ. When a contact occurs, it is equally likely to involve any of the $P = \binom{N}{2}$ pairs of individuals in the population.

If a contact involves an infected and a non-infected individual, then with probability p the noninfected individual becomes infected. Once infected, an individual remains infected throughout. Let $X(t)$ denote the number of infected members of the population at time t.

(i) State whether $X(t)$, $t > 0$ is a continuous-time Markov jump process. If so, write down its state space and transition rates; if not, explain how it can be expressed in terms of a different process which *is* Markov. [2]

(ii) Derive an expression for the expected time until all members are infected, starting from a single infected individual [2]

[Total 4]

2 During a long motorway journey a child amuses himself by noting down, at the end of each minute, the lane in which the car is travelling. The motorway has three lanes and the journey lasts N minutes.

(i) Describe how to fit a three-state time-homogeneous Markov chain model to the data, writing down formulae for the estimates of the transition probabilities. [2]

(ii) Describe one test which could be applied to determine whether the process possesses the Markov property. [2]

[Total 4]

3 A standard Ornstein-Uhlenbeck process may be defined as a stationary zero-mean Gaussian process $\{U_t : t \in \mathbf{R}\}$ with autocovariance function given by

$$\mathrm{Cov}(U_t, U_s) = \frac{\tau^2}{2\theta} e^{-\theta|t-s|} \qquad (s, t \in \mathbf{R}).$$

(i) Show that the process $\{U_n : n = 1, 2, ...\}$, obtained by observing U only at integer times, is a first-order autoregression. [2]

(ii) Derive expressions for the parameters α and σ^2 of the autoregression in terms of θ and τ^2. [2]

[Total 4]

4 (i) State Itô's lemma as it applies to a stochastic process $\{X_t : t \geq 0\}$ and a function $f(X_t)$ which is not explicitly dependent on t. [2]

(ii) Apply Ito's Lemma with $f(x) = x^4$ to calculate the stochastic differential $d(B_t^4)$, where B_t is standard Brownian motion. [3]

(iii) Hence express the Itô integral $\int_0^t B_s^3 \, dB_s$ in terms of B_t and of an ordinary integral involving B_s. [2]

[Total 7]

5 Consider a homogeneous Markov chain with state space $S = \{1, 2, 3\}$ and transition matrix

$$P = \begin{pmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{3}{4} & \frac{1}{4} & 0 \end{pmatrix}.$$

(i) Calculate the 3-step transition matrix. [2]

(ii) Calculate, for each of the following initial conditions, the probability that the chain will be in state 3 when it is observed at time $n = 3$ given that:

(a) the chain is in state 1 at time zero

(b) the chain is in state 1 at time zero and in state 2 at time 1

(c) the probabilities of being in states 1, 2, and 3 at time zero are given by $\frac{14}{31}, \frac{9}{31}$ and $\frac{8}{31}$ respectively [4]

(iii) How would your answers to (a), (b) and (c) change if the time of the observation were $n = 300$ instead of $n = 3$? [2]

[Total 8]

6 The daily closing price of a share is observed every trading day for a year, yielding a sequence of values $\{s_1, ..., s_n\}$. A model is required for the purposes of predicting future variability of the share price. The model suggested is a Brownian one.

(i) Explain briefly, on purely theoretical grounds, which of the two models

I: $S_t = \mu + \alpha t + \sigma B_t$

II: $\log(S_t) = \mu + \alpha t + \sigma B_t$

you would expect to provide a better fit. [1]

(ii) Refer to Figures 1 and 2 below.

(a) Explain briefly whether your chosen model appears to provide a good fit to the data.

(b) State *one* of the tests you could carry out on the data to ascertain whether the model fits adequately. [3]

(iii) (a) Describe how a Lévy process model differs from a Brownian model.

(b) Outline the difficulties would you encounter in practice if you were fitting a Lévy process model to the data provided. [3]

[Total 7]

Figure 1: the share price S_n

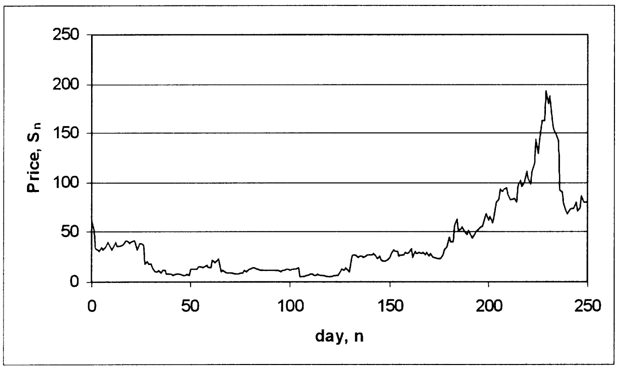
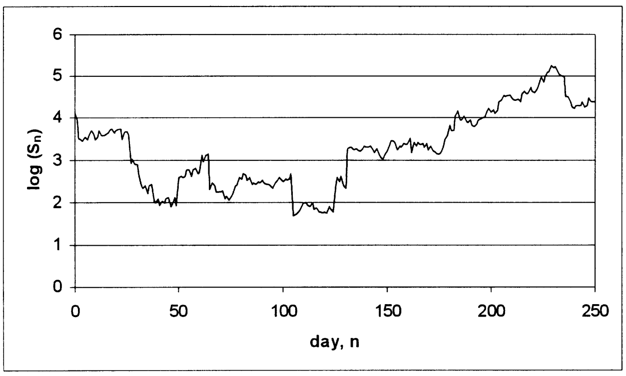


Figure 2: log-transformed share price log S_n



7 A client wishes to model the behaviour of a stochastic process $\{X_t : t \geq 0\}$ which represents the average annual return for a particular class of asset. After a number of observations the client has determined that $\text{Corr}(X_t, X_{t-1}) = 0.7$ and $\text{Corr}(X_t, X_{t-2}) = 0.5$. He thinks that one of the two models

I: $X_t = \mu + 0.7(X_{t-1} - \mu) + 0.5(X_{t-2} - \mu) + e_t$

II: $X_t = \mu + e_t + 0.7e_{t-1} + 0.5e_{t-2}$

will be best, but cannot decide which. He has simulated both processes from time $t = 1$ to time $t = 200$, but has not obtained the results he expected, so is seeking your advice.

(i) (a) Outline a suitable method of simulating a second-order autoregression, assuming you have access to a reliable stream $\{u_k : k \geq 0\}$ of pseudo-random numbers uniformly distributed over the range [0, 1].

(b) Explain why might it be desirable to ensure that the stream $\{u_k\}$ can be re-used if necessary. [4]

(ii) State why neither of the suggested models is suitable. [1]

(iii) (a) Derive the lag-1 and lag-2 autocorrelations, ρ_1 and ρ_2, of a second-order autoregressive process

$$X_t = \mu + \alpha_1 (X_{t-1} - \mu) + \alpha_2 (X_{t-2} - \mu) + e_t .$$

(b) Find values of the parameters α_1 and α_2 which would provide a suitable AR(2) model for $\{X_t : t = 0, 1, 2, \ldots\}$. [5]

[Total 10]

8 Consider a time-homogeneous Markov jump process $\{X(t) : t \geq 0\}$ with two states denoted by 0, 1, and transition rates $\sigma_{0,1} = \lambda$, $\sigma_{1,0} = \mu$.

(i) State Kolmogorov's forward equation for the probability $P_{0,0}(t)$ that X is in state 0 at time t, given that it starts in state 0. [1]

(ii) Show that $P_{0,0}(t) = \dfrac{\mu}{\lambda+\mu} + \dfrac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)t}$. [3]

(iii) Let O_t denote the total amount of time spent in state 0 up until time t, which may be expressed as $O_t = \int_0^t I_s ds$, where $I_s = \begin{cases} 1 & \text{if } X_s = 0 \\ 0 & \text{if } X_s \neq 0 \end{cases}$. Derive, using the result in part (ii), an expression for $\mathbf{E}[O_t \mid X(0) = 0]$, the expected occupation time in state 0 by time t for the two-state continuous-time Markov chain starting in state 0. [2]

(iv) Write down the expected occupation time in state 1 by time t for the two-state continuous-time Markov chain starting in state 0. [1]

(v) A health insurance scheme labels members as "healthy" (state 0) or "unhealthy" (state 1) at any time. When in state 0, members pay contributions at rate α; when in state 1 they receive benefit at rate β. Expenses amount to a constant γ per member per unit time.

(a) Explain how the above model can be used to calculate α in terms of β and γ.

(b) State the assumptions which you make in applying the model.

(c) Discuss whether they are likely to be satisfied in practice. [4]

[Total 11]

9 Suppose that the evolution of the price of an asset follows the lognormal model $\log(S_t) = Y_t = y + \mu t + \alpha B_t$ where B_t denotes the standard Brownian motion and μ is a negative drift. The asset will be liquidated at the stopping time $T_\alpha = \inf\{t : Y_t = a\}$ when its value reduces to e^a, where a is some number less than y. Consider now the exponential $V_t = \exp(uY_t - c(u)t)$.

(i) Derive the condition on $c(u)$ under which $\{V_t : t \geq 0\}$ is a martingale. [3]

(ii) State the optional stopping theorem and explain how it is used. [3]

(iii) Derive the moment generating function $f(y, v) = \mathbf{E}[e^{-vT_u} \mid Y_0 = y]$ of the bankruptcy time for positive v by applying the optional stopping theorem to the martingale V_t. [5]

[Total 11]

10 Consider a survival model with two states "alive" (A) and "dead" (D), with time-dependent transition rate from A to D equal to $\mu(t) = \mu t$. The time parameter, t, represents the age of the individual under consideration.

(i) Calculate the transition probability $P_{AA}(s, t)$, defined by

$$P_{AA}(s, t) = \mathbf{P}(X(t) = A \mid X(s) = A).$$ [2]

(ii) Show, by making use of the formula

$$\mathbf{E}(X) = \int_0^\infty \mathbf{P}[X \geq x]\, dx$$

for a positive random variable X, that the expected future lifespan of an individual aged s is

$$\mathbf{E}[R_s] = \frac{1}{\sqrt{\mu}} \frac{1 - G(s\sqrt{\mu})}{g(s\sqrt{\mu})},$$

where G is the standard Gaussian probability distribution function and g is its density. [4]

(iii) It is desired to calibrate the above model so that the expected future lifespan of an individual aged 70 is 6 years. Derive an approximation to the corresponding value of μ, using the double inequality

$$\frac{1}{x} - \frac{1}{x^3} \leq \frac{1 - G(x)}{g(x)} \leq \frac{1}{x}.$$ [5]

(iv) A company wishes to test the validity of the above model. They assume that the true force of mortality from age 70 onwards is of the form $\mu(t) = a + bt$ and intend to test whether $a = 0$. The testing method will be to simulate one sample of size 1000 when $a = 0$ and another when $a \neq 0$, then to see which most resembles the data which the company has collected.

Explain how to simulate a value from the proposed distribution, for arbitrary values of a and b. [5]

[Total 16]

11 The movements of a consumer price index are to be subjected to time series analysis with the aim of forecasting future behaviour. The index is calculated monthly.

(i) Explain whether you would expect to fit a model which included (a) a trend term, (b) a seasonal effect. [3]

The values $\{x_t : 1 \le t \le n\}$ are the residuals which remain once any trend or seasonal variations have been removed. An ARIMA(1, 1, 1) model is to be fitted to the $\{x_t\}$.

(ii) (a) Assuming the ARIMA(1, 1, 1) model is correct, write down an equation for X_{n+1} in terms of the white noise process $\{e_t: 1 \le t \le n + 1\}$ and the observations $\{x_t: 1 \le t \le n\}$.

(b) State the parameters of the model. [2]

(iii) The Box-Jenkins procedure defines the k-step-ahead forecast for X to be

$$\hat{x}_n(k) = \mathbf{E}(X_{n+k} | x_n, x_{n-1}, \ldots, x_1).$$

(a) Derive the 1-step-ahead and 2-step-ahead forecasts for X for the ARIMA(1, 1, 1) model, assuming that the values of the parameters and the value of e_0 are known exactly.

(b) Evaluate the prediction variance $\text{Var}(X_{n+1} - \hat{x}_n(1))$, again assuming that the values of the parameters are known. [5]

(iv) The most elementary form of the technique known as exponential smoothing produces at time n a 1-step-ahead forecast x_n^* defined by

$$x_n^* = x_n + \xi(x_{n-1}^* - x_n),$$

for some $\xi \in (0, 1)$ which may be chosen by the user.

Show that, for particular values of the autoregressive and moving average parameters, the Box-Jenkins forecasts above coincide with the forecasts produced by exponential smoothing. [2]

(v) (a) State whether the ARIMA(1, 1, 1) model is $I(0)$, $I(1)$ or neither.

(b) Discuss whether there is a difference between an $I(0)$ model and an $I(1)$ model in terms of the conditional distribution of X_{n+k} given $\{x_t : 1 \le t \le n\}$ for large values of k. [3]

(vi) It is suggested that a salaries index might be cointegrated with the consumer price index.

(a) Explain what is meant by the suggestion.

(b) Comment on whether it is a reasonable suggestion. [3]

[Total 18]