EXAMINATIONS

20 April 2004 (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 10 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 The stochastic differential equation for geometric Brownian motion is given by

$$dX_t = \rho X_t \, dt + \sigma X_t \, dB_t$$

where B_t is a standard Brownian motion. Show that the solution to the differential equation which also satisfies $X_{t_0} = x$ is

$$X_t = x \exp\left[\sigma\left(B_t - B_{t_0}\right) + \left(\rho - \tfrac{1}{2}\sigma^2\right) \left(t - t_0\right)\right]$$

for $t \geq t_0$. [4]

2 A Box-Jenkins model-fitting procedure suggests that the best fitting model for a set of normalised share price data $x_1, \ldots, x_n$ is ARMA(1,2), with equation

$$X_t = 0.63X_{t-1} + e_t + 0.45e_{t-1} - 0.34e_{t-2},$$

where $\{e_1, e_2, \ldots\}$ is a sequence of uncorrelated, zero-mean random variables with variance σ^2.

(i) Determine whether the model is stationary and invertible. [2]

(ii) Calculate $\gamma_0, \gamma_1, \gamma_2$, the autocovariance function of the fitted model at lags 0, 1, and 2, in terms of σ^2. [4]

[Total 6]

3 The number, $N(t)$, of members of a pension scheme who are receiving benefits at time t, is subject to change of two kinds:

- it increases by 1 when an active member reaches retirement age
- it decreases by 1 when a retired member dies

Assume that retirements occur according to a Poisson process with rate λ and that each retired member, independently, has a probability μdt of dying within the time interval $(t, t + dt)$.

(i) Explain why, under these assumptions, $N(t)$ is a Markov jump process. [1]

(ii) Write down the transition rates of $N(t)$. [1]

(iii) Using the notation $p_n(t) = P(N(t) = n)$, obtain a differential equation satisfied by $p_n(t)$. [2]

(iv) Verify that one solution of the equation in (iii) is given by

$$p_n(t) = \frac{1}{n!} e^{-\lambda/\mu} \left(\frac{\lambda}{\mu}\right)^n, \qquad n = 0, 1, \ldots \quad [2]$$

(v) State what conclusions can be drawn from (iv). [1]
[Total 7]

4 The value, S_t, of a stock exchange index is observed at hourly intervals over the course of a week, generating observations $s_1, s_2, \ldots, s_n$.

Three different models for S_t are being considered:

(a) $S_t = S_0 + \mu t + \sigma B_t$, where B_t is a standard Brownian motion.

(b) $S_t = S_0 e^{\mu t + \sigma B_t}$, where B_t is as in (a).

(c) $\nabla S_t = \mu + \alpha(\nabla S_{t-1} - \mu) + e_t$, where $\{e_t: t = 1, 2, \ldots\}$ is a sequence of uncorrelated, zero-mean random variables.

(i) Outline the processes of model fitting and model validation, including a description of the role of simulation. [3]

(ii) For EACH of the three models above, outline the steps required during the process of fitting the model. [5]

(iii) Choose ONE of the above models and describe TWO tests you would apply as part of the model validation process. [2]
[Total 10]

5 (i) (a) Write down the joint density function of $B(½)$ and $B(1)$, where $\{B(t): t \geq 0\}$ is a standard Brownian motion.

(b) Demonstrate that the conditional distribution of $B(½)$ given that $B(1) = b$ is Normal with mean $½b$ and variance ¼.

[5]

(ii) James runs around a running track at the same time as an "electric hare" which is programmed to go around in exactly one minute. Let $Y(t)$ denote the distance (in metres) by which James is ahead of the hare t minutes after the start and suppose that $Y(t)$ can be modelled as $Y(t) = 5\,B(t)$, where $B(t)$ is as in (i).

(a) If James is leading the hare by 5 metres at time $t = 0.5$, determine the probability that James eventually finishes the race ahead of the hare.

(b) If James finishes the run 5 metres ahead of the hare, determine the probability that he was ahead of the hare at time $t = 0.5$.

[5]
[Total 10]

6 (i) Describe a method for generating a sequence $U_1, U_2, \ldots$ of pseudo-random numbers uniformly distributed in the interval $(0,1)$ and from this a pseudo-random sequence of numbers uniformly distributed over an arbitrary range (a,b), where $a < b$. [3]

(ii) Explain the main advantage of using pseudo-random, as opposed to purely random numbers, for testing the suitability of a model. [1]

(iii) Given a sequence $\{U_n : n = 1,2,\ldots\}$ as in (i), explain how you would simulate

(a) an observation from the Pareto density

$$f(x) = \frac{a}{(1+x)^{a+1}}, \qquad a > 1, x \geq 0$$

(b) a discrete random variable X with probability function given by

$$P(X = i) = \frac{1}{n+2} \text{ for } i = 1,2,\ldots,n, \quad P(X = 0) = \frac{2}{n+2}.$$ [4]

(iv) The Pareto distribution in (iii) is often referred to as a "heavy-tailed distribution". Explain the use of this term and discuss under what circumstances it might offer a suitable model for modelling sizes of claims arriving at an insurance company. [2]
[Total 10]

7 (i) Suppose that $\{X_t : t \geq 0\}$ is a time-homogeneous Markov jump process with generator matrix A and transition matrix $P(t)$. Let π be a (column) vector of probabilities such that $\pi^T A = \mathbf{0}^T$, where $\mathbf{0}$ is a (column) vector whose components are all equal to zero.

(a) Prove that

$$\frac{d}{dt}\pi^T P(t) = \mathbf{0}^T.$$

(b) Explain why π is known as the stationary distribution for X.

[4]

(ii) A telephone call centre receives calls from customers at an average rate of 0.5 per minute. Each call has a random duration which is exponentially distributed with mean 3 minutes, independently of the number or duration of any other calls. Two operators are assigned to handle calls. If a call arrives when both operators are busy, the call is put on hold unless there are already two calls on hold, in which case the new call is lost. When a call ends, one of the calls on hold is immediately put through to the newly free operator.

(a) Identify the five states which are required if this system is to be modelled as a Markov jump process.

(b) Draw the transition diagram for this system.

(c) Write down the generator matrix for the process.

(d) Evaluate the stationary distribution of the system.

[8]
[Total 12]

 PLEASE TURN OVER

8 A restorer of historic sites travels between Bath, Warwick, Stratford and Caernarvon. Having arrived at a site, the restorer stays there for a random number of days, then moves on to one of the others. The restorer's diary for the last two months shows the following time spent in each place:

> Bath (5 days), Warwick (3 days), Caernarvon (7 days), Bath (3 days), Stratford (7 days), Warwick (5 days), Bath (3 days), Stratford (3 days) Caernarvon (8 days), Bath (2 days)

A researcher suggests that the path taken by the restorer forms a Markov chain with state space {Bath, Caernarvon, Stratford, Warwick}.

(i) Suppose $\{X_t : t = 0, 1, 2, \ldots\}$ is a Markov chain with transition matrix P and let $D_{i,n}$ denote the duration of the nth visit to state i, that is, the number of consecutive steps spent by the chain in state i. Show that $D_{i,n}$ is a geometric random variable and is independent of $D_{i,m}$ for $m < n$. [3]

(ii) Discuss, in the light of the data, whether it is likely that this suggestion is satisfied. [2]

(iii) Assuming that the researcher's suggestion is correct, write down an estimate of the transition matrix P of the Markov chain. [2]

(iv) State, giving reasons for your answers, whether the Markov chain model is (a) irreducible, and (b) aperiodic. [2]

(v) The restorer arrives in Warwick on day t_0. Use the estimated transition matrix to calculate the probability that he is in Warwick on day $t_0 + 3$. [3]

[Total 12]

9 A water company is developing a time series model to model the supply of water, X_t, in its reservoirs at the end of month t. The model takes the form

$$X_t = X_{t-1} + R_t - D_t \,,$$

where R_t represents rainfall in month t and D_t represents demand in month t. R_t and D_t are themselves modelled by

$$R_t = \mu_t + \alpha(R_{t-1} - \mu_{t-1}) + e_t$$
$$D_t = \nu - \theta R_t$$

where α, θ, ν are positive constants, μ_t is a deterministic function of time and $\{e_t: t = 1, 2, \ldots\}$ is a sequence of independent Normal random variables with mean 0 and variance σ^2.

(i) (a) Comment on whether it is reasonable to assume that demand is negatively related to rainfall.

It is suggested that μ_t should be cyclical with period 12, i.e. that $\mu_{t+12} = \mu_{12}$ for all t. Comment on whether this suggestion is sensible. [2]

(ii) State, with reasons, whether R_t is stationary. [1]

(iii) (a) Suppose that R_0 is Normally distributed with mean μ_0 and variance σ_R^2. Show that R_t is also Normally distributed for each $t > 0$ and derive an expression for its mean.

(b) Calculate a value of σ_R^2 which ensures that $\mathrm{Var}(R_t) = \sigma_R^2$ for all t. [4]

(iv) Given a sequence of observations $R_1, \ldots, R_{120}, D_1, \ldots, D_{120}$ covering 10 years of data, suggest simple estimators for μ_t $(1 \le t \le 12)$, α, θ and ν. [4]

(v) Assuming that the values of the parameters are known exactly, and that R_{120}, D_{120} and X_{120} are known, derive an estimate $\hat{x}_{120}(1)$ for the supply of water at the end of month 121 and give an expression for $\mathrm{Var}(X_{121} - \hat{x}_{120}(1))$. [3]

[Total 14]

10 An insurance company has an initial capital of u and receives premium income at constant rate c per unit time. Claims, for one unit each, occur according to a Poisson process $\{N_t, t \geq 0\}$ with rate λ. The company's surplus at time t is defined as $S_t = u + ct - N_t$.

(i) A *Lévy process* is defined as a stochastic process with stationary, independent increments. Verify that S_t satisfies this definition. [2]

(ii) For $s > 0$, write down the conditional distribution of $N_{t+s} - N_t$ given N_t. [1]

(iii) Define $Y_t = e^{\beta S_t}$.

(a) Derive an equation which must be satisfied by β if Y_t is to be a martingale.

(b) Show, by means of a sketch or otherwise, that the equation derived in (a) has a negative solution β as long as $c > \lambda$ and a positive solution β if $c < \lambda$.

(c) Comment on whether the condition $c > \lambda$ is likely to be satisfied in practice. [6]

(iv) Let $T = \inf\{t > 0 : S_t < 0 \text{ or } S_t \geq K\}$, where K is a positive constant and assume that $c > \lambda$.

(a) Explain why either $S_T = K$ or $-1 \leq S_T < 0$.

(b) Verify that the conditions of the optional stopping theorem apply to the stopped martingale $Y_{\min(t,T)}$.

(c) Define ψ to be the probability that S goes below zero before it hits K, that is, the probability that $S_T < 0$. Prove that

$$E(Y_T) \geq \psi + (1-\psi)e^{\beta K}$$

(d) Derive a lower bound for ψ using the result in (iv)(c) and state the limit of this lower bound as $K \to \infty$. [6]

[Total 15]

END OF PAPER