EXAMINATIONS

8 April 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 10 questions, beginning your answer to each question on a separate sheet.*

Graph paper is not required for this paper.

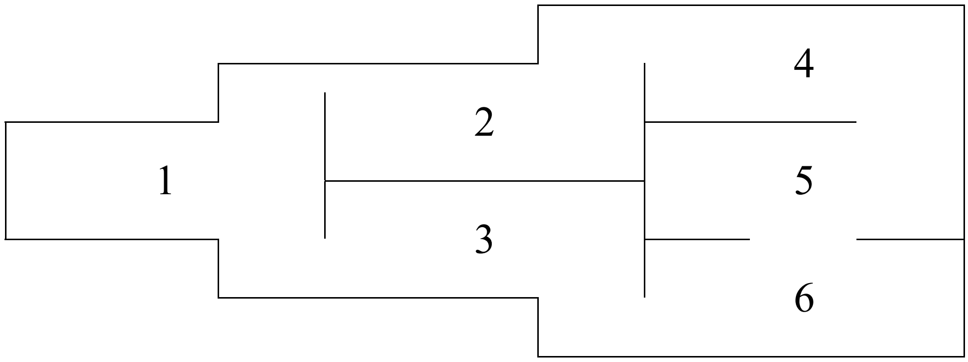
AT THE END OF THE EXAMINATION

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.

© Faculty of Actuaries
© Institute of Actuaries

1 A science student is observing a mouse in the following simple maze:



During each period under observation, the mouse moves to any adjacent, accessible compartment with equal probability. Successive moves have the Markov property. Let X_t be the compartment number that the mouse is in at time t.

(i) Draw a transition diagram, including on your diagram the probabilities of the possible transitions. [2]

(ii) Evaluate the probability that a mouse which starts in compartment 1 at time 0 is also in compartment 1 (a) at time 1, (b) at time 2, and (c) at time 3. [2]

(iii) State with reasons whether:

(a) the Markov chain X possesses a stationary distribution

(b) $P(X_t = j \mid X_0 = i)$ converges to some limit π_j as $t \to \infty$ [2]

[Total 6]

2 An index of salaries, S_t, and an index of prices, P_t, are modelled as being related to each other in the following way:

$$\nabla \ln S_t = \theta_S + \alpha_1 \nabla \ln S_{t-1} + \alpha_2 \nabla \ln P_{t-1} + e_{S,t}$$
$$\nabla \ln P_t = \theta_P + \beta_1 \nabla \ln S_{t-1} + \beta_2 \nabla \ln P_{t-1} + e_{P,t}$$

where $e_{S,t}$ and $e_{P,t}$ are two independent zero-mean white noise processes, with variances σ_1^2 and σ_2^2 respectively, and θ_S and θ_P are constants.

(i) Explain why the above model is written in terms of $\nabla\ln P$ and $\nabla\ln S$ instead of just P and S. [2]

(ii) Comment on whether it is reasonable that $\nabla\ln S_t$ should be affected by $\nabla\ln P_{t-1}$ and that $\nabla\ln P_t$ should be affected by $\nabla\ln S_{t-1}$. [1]

(iii) Use matrix notation to express $(\nabla\ln S_t, \nabla\ln P_t)^T$ as a Vector Autoregression and identify the order p of the VAR(p) process. [2]

(iv) Suppose the parameters of the model have been estimated. Describe the use of sensitivity analysis in determining the validity of the model. [2]

[Total 7]

3 (i) Let Y_t denote Brownian motion with drift μ and variance rate σ^2, starting at 0. Write down the expectation $E(Y_t)$ and variance $\text{Var}(Y_t)$ of this process. [1]

(ii) Let X_t denote a biased random walk which moves up or down by D at time intervals of size h, i.e.

$$X_t = \sum_{i=1}^{\left\lfloor \frac{t}{h} \right\rfloor} Z_i \,,$$

where $Z_1, Z_2, \ldots$ are independent random variables, each with distribution $P[Z_i = D] = p$, $P[Z_i = -D] = 1 - p$. Calculate the expectation and variance of X_{nh} and derive an expression for the expectation and variance of X_t. [3]

[Notation: $\lfloor x \rfloor$ denotes the integer part of x.]

(iii) (a) Use the approximation $\lfloor x \rfloor \approx x$ to derive conditions on p and D such that the expectation and variance of the process X_t approximate for large t those of a Brownian motion with drift μ and variance rate σ^2.

(b) Comment on the values of h for which such an approximation might be appropriate.

[4]
[Total 8]

4 (i) Given a pseudo-random number U uniformly distributed over [0,1], obtain an expression in terms of U and θ for a non-negative pseudo-random variable X which has density function

$$f(x) = \theta e^{-\theta x}$$ [2]

(ii) A sequence of simulated observations is required from the density function

$$g(x) = k(\theta)\frac{e^{-\theta x}}{1+x}, \quad x > 0$$

where θ is a non-negative parameter and $k(\theta)$ is a constant of integration not involving x.

(a) Describe a procedure that applies the Acceptance-Rejection method to obtain the required observations.

(b) Derive an expression involving θ and $k(\theta)$ for the expected number of pseudo-random variables required to generate a single observation from the density g using this method.

[6]
[Total 8]

5 (i) Let $W(t)$ be defined by $W(t) = -4B(kt)$, where $B(t)$ is a standard Brownian motion.

(a) Calculate the value of k which gives $W(t)$ the same expectation and covariance function as $B(t)$.

(b) Prove that, for this value of k, W is a standard Brownian motion.

[4]

(ii) Prove that $\exp[2B(t)-2t]$ is a martingale. [5]

[Total 9]

6 A No-Claims Discount system operated by a motor insurer has four levels:

Level 1: 0% discount
Level 2: 25% discount
Level 3: 40% discount
Level 4: 50% discount

The rules for moving between these levels are as follows:

Following a claim-free year, move to the next higher level, or remain at level 4.

Following a year with one or more claims:

move back **one** level, or remain at level 1, if, in the year before the most recent year, there were no claims;

move back **two** levels, or move to level 1 or remain at level 1 if, in the year before the most recent year, there was one or more claims.

For a given policyholder the probability of no claims in a year is 0.8.

(i) Let $X(t)$ denote the state, either 1, 2, 3 or 4, of the policyholder in year t. Explain why $\{X(t)\}_{t=1}^{\infty}$ is **not** a Markov chain. [2]

(ii) (a) By increasing the number of states, define a new stochastic process $\{Y(t)\}_{t=1}^{\infty}$ which **is** Markov and is such that $Y(t)$ indicates the discount level for the policyholder in year t.

(b) Write down the transition matrix for the Markov chain $\{Y(t)\}_{t=1}^{\infty}$.

(c) Calculate the long-run probability that the motorist is in discount level 3.

[8]

[Total 10]

7 The medical insurance division of a large insurance company models each policyholder's state of health as a three-state Markov jump process, the states being H (healthy), S (sick) and D (dead). The instantaneous transition rates between the states are $\sigma_{HS} = \sigma$, $\sigma_{SH} = \rho$, $\sigma_{HD} = \mu$, $\sigma_{SD} = \nu$.

(i) Write down the generator matrix of the Markov jump process. [2]

(ii) The policyholder pays contributions at rate C when in state H and receives benefits at rate B when in state S. No death benefit is payable. The company uses the model to set the ratio of contributions to benefits. Without doing any calculations, explain in general terms how this can be done. [2]

(iii) A trainee believes that the model is too simplistic. For each of the trainee's suggestions below, comment on whether following the suggestion would be likely to improve the model's predictive power:

(a) The transition rates should depend on the age of the policyholder.

(b) The transition rates should vary according to the time of year.

(c) σ_{SH} and σ_{SD} should also depend on the duration of the sickness to date. [3]

(iv) Outline the principal difficulty in fitting a model with parameters dependent on all the factors in part (iii). [1]

(v) Assume that several years of quarterly claims data are available. Describe a test to determine whether the model with annually time-varying transition rates, as in (iii)(b), is a better fit to the data than the model with constant transition rates. [2]

[Total 10]

8 A branch of a bank has three cash dispensers. If at time t a cash dispenser is working, the probability that it will break down in $(t, t + dt)$ is independent of the state of the other cash dispensers and is equal to

$$\alpha\, dt + o(dt).$$

When a cash dispenser breaks down, repair work begins immediately. The time taken to repair a broken machine is exponentially distributed with mean $1/\beta$. There are enough repair teams to repair all three cash dispensers at the same time, if necessary.

Define X_t as the number of machines not working at time t.

(i) Write down the state space for X_t. [1]

(ii) (a) If a cash dispenser is working at time 0, prove that the time until its first breakdown is exponentially distributed with mean $1/\alpha$.

(b) If all three cash dispensers are working at time 0, derive the distribution of the time until the first breakdown. [5]

(iii) Define $P_m(t)$ as the probability that m machines are not working at time t. Show that the forward equations for the process X imply that:

$$P_0'(t) = -3\alpha P_0(t) + \beta P_1(t)$$

$$P_3'(t) = -3\beta P_3(t) + \alpha P_2(t)$$

and for $m = 1, 2$,

$$P_m'(t) = -((3-m)\alpha + m\beta)P_m(t) + (4-m)\alpha P_{m-1}(t) + (m+1)\beta P_{m+1}(t).$$

[3]

(iv) At time 0 all three cash dispensers are working. Show that the forward equations have a solution

$$P_m(t) = \binom{3}{m}\theta(t)^m[1-\theta(t)]^{3-m}, \quad m = 0,1,2,3$$

where $\theta(t)$ satisfies a differential equation which you should identify. [5]

[Total 14]

9 In a simple discrete-time model for the price of a share, the change in price at time t, X_t, is assumed to be independent of anything that has happened before time t and to have distribution:

$$X_t = \begin{cases} 1 & \text{with probability } p \\ -1 & \text{with probability } q \\ 0 & \text{with probability } r = 1 - p - q \end{cases},$$

where $p, q, r > 0$.

Let $S_0 = m$ (where m is a positive integer) be the original price of the share, $S_n = S_0 + \sum_{t=1}^{n} X_t$ be the price after n time units and define

$$Y_n = (q/p)^{S_n}$$

(i) Show that $\{Y_n : n \geq 1\}$ is a martingale and that for any positive integer n,

$$E(Y_n) = (q/p)^m$$ [4]

(ii) Let T be the time until the share price reaches either 0 or N for the first time, where N is an integer greater than m.

(a) Show that $|Y_n| \leq C$ for $n \leq T$, for some constant C.

(b) Write down an expression for $E(Y_T)$.

[2]

(iii) Assuming that $p \neq q$, calculate the probability that, starting from m, the share price reaches 0 before it reaches N, i.e. $P(S_T = 0 \mid S_0 = m)$. [3]

(iv) In the case $p = q$ you may assume that $P(S_T = 0 \mid S_0 = m) = (N-m)/N$. Now define $Z_n = S_n^2 - 2np$.

(a) Prove that $\{Z_n : n \geq 1\}$ is a martingale.

(b) Show that the expected value of the time until absorption, T, is given by

$$E(T \mid S_0 = m) = \frac{m(N-m)}{2p}.$$

[5]
[Total 14]

10 The price of a stock through time, $\{X_t : t \geq 0\}$ is thought to be modelled by the following relationship:

$$X_t = 1.7X_{t-1} - 0.4X_{t-2} - 0.3X_{t-3} + e_t - 0.7e_{t-1} + 0.12e_{t-2}$$

For this model,

(i) Write the equation in terms of the backward shift operator B in the form

$$\phi(B)(1 - B)^d X_t = \theta(B)e_t,$$

where $\phi(B)$ and $\theta(B)$ are polynomials in B. [3]

(ii) Identify the values of p, d and q for which X is an ARIMA(p, d, q) process. [1]

(iii) Explain whether the process $\{X_t : t \geq 0\}$ is stationary. [1]

(iv) For the value of d from (ii), put $W_t = (1 - B)^d X_t$. Explain why the model can be written in the equivalent form

$$W_t = \sum \psi_i e_{t-i}$$

and calculate ψ_i for $i = 0, 1, 2$. [3]

(v) Another representation of the model is

$$\pi(B)W_t = e_t$$

where $\pi(B) = 1 - \sum_{i=1}^{\infty} \pi_i B^i$. Calculate π_i for $i = 1, 2$. [1]

(vi) Define two vector-valued stochastic processes Y and Z as

$$Y_t = (X_t, X_{t-1}, X_{t-2})^T, \quad Z_t = (X_t, X_{t-1}, X_{t-2}, X_{t-3}, X_{t-4})^T.$$

Explain which, if any, of the processes $\{X_t : t \geq 0\}$, $\{Y_t : t \geq 0\}$ and $\{Z_t : t \geq 0\}$ has the Markov property. [2]

(vii) Given a set of observations $(x_1, x_2, \ldots, x_n)$ from an ARIMA(1, 1, 1) process with unknown parameter values, outline the main steps that need to be taken so that one can obtain forecasts for future values of the process. [3]

[Total 14]