Act2/X2 Discrete Stochastic Modelling, June 2001: Solutions

Question 1

(i) Let $X_n = \log S_n$. Then $X_{n+1} = X_n + J_{n+1}$, where each J_n takes values $\log u$ with probability p, $\log d$ with prob $q = 1 - p$.
The condition required is that the J_n are independent.

(ii) $\mathbb{E}(S_{n+1}) = \mathbb{E}(S_n)\mathbb{E}(J_n) = \theta\mathbb{E}(S_n)$, where $\theta = pu + (1-p)d$.
Therefore $\mathbb{E}(S_n) = \theta^n S_0$.

(iii)

a) $\mathbb{E}(X_{n+1}) = \mathbb{E}(X_n) + \mathbb{E}(J_{n+1}) = \mathbb{E}(X_n) + \phi$, where $\phi = p\log u + (1-p)\log d$. Therefore $\mathbb{E}(X_n) = X_0 + n\phi$.
Similarly $\text{Var}\,(X_n) = n\text{Var}\,J = n(p\log^2 u + (1-p)\log^2 d - \phi^2) = np(1-p)\log^2(u/d)$.

b) S_n is approximately lognormally distributed with parameters $\mu = X_0 + n\phi$, $\sigma^2 = np(1-p)\log^2(u/d)$ and therefore $\mathbb{E}(S_n) \approx \exp(\mu + \frac{1}{2}\sigma^2) = S_0 \exp\left(n\phi + \frac{1}{2}np(1-p)\log^2(u/d)\right)$.

(iv)

a) Since $\theta = \frac{1}{2}(1.04) + \frac{1}{2}(0.96) = 1$, the exact value $\mathbb{E}(S_n)$ is S_0.
For the lognormal approximation, $\phi = \frac{1}{2}\log 1.04 + \frac{1}{2}\log 0.96 = -0.0008006$ and $\sigma^2 = \frac{1}{4}n\log^2\left(\frac{1.04}{0.96}\right) = 0.0016017n$, giving the approximate value of $100\exp(2.14 \times 10^{-7}n)$.

b) The expectation of the increment of the random walk $\{\log S_n\}$ is ϕ, which is negative. Thus the walk has a tendency to drift downwards and $\mathbb{P}(\log S_n > \log 125) \to 0$ as $n \to \infty$.

c) We are looking for

$$\mathbb{P}\left(\frac{X_{1000} - \mathbb{E}X_{1000}}{\sqrt{\text{Var}\,(X_{1000})}} > \frac{\log 125 - \log 100 + 0.8006}{\sqrt{1.602}} = 0.809\right).$$

Using the Normal approximation we obtain $\mathbb{P}(Z > 0.809) = 0.209$.

Question 2

(i) [Bookwork] The *state space* of a stochastic process X is the set of values X can take.
A *Markov chain* is a discrete time, discrete state space process X such that $\mathbb{P}(X_{n+m} = j \mid X_n = i, \mathcal{H}_n) = \mathbb{P}(X_{n+m} = j \mid X_n = i)$, where $\mathcal{H}_n$ is the history of the process up until time n.
Two states i and j *communicate* if, starting from i, it is possible to reach j eventually, and vice versa. A communicating class is a group of states which communicate with one another.
A c.c. is *recurrent* in this context if it is closed, i.e. if, having entered the class, the chain cannot then leave it.
The desired result is that the Markov chain eventually enters the one recurrent c.c. and thereafter never leaves it.

(ii)

a) A suitable state space would be {Never married, Married, Formerly married, Community member}, which we can call $\{N, M, F, C\}$.
Diagram.

b) According to the question we have

$$P = \begin{pmatrix} 0.90 & 0.08 & 0 & 0.02 \\ 0 & 0.96 & 0.04 & 0 \\ 0 & 0.10 & 0.80 & 0.10 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

c) CCs are: $\{N\}$ (transient), $\{M, F\}$ (transient), $\{C\}$ (recurrent).

d) The required probability is the probability that the marriage comes to an end by time $t+1$, then the woman enters the religious community by time $t+2$, which is $0.04 \times 0.1 = 0.004$.

e) For all transient states j, the limit $\lim_{n\to\infty} \mathbb{P}(X_n = j \mid X_0 = i)$ is 0 for every i. Since there is only one recurrent state in the state space, it is plain that $\lim_{n\to\infty} \mathbb{P}(X_n = C \mid X_0 = i) = 1$ for each i.

f) Instead of having a fixed transition matrix for all ages, we could have different transition matrices for different age bands.

g) The model leads to the interesting prediction that every woman ultimately joins a religious community, which may be regarded as slightly unreasonable. Introduce an additional state, Dead, to absorb the process in some cases. [Other criticisms are also possible.]

Question 3

(i) a) The model is $(1 - \alpha_1 B - 0.4B^2)X = (1 + 0.3B)e$.

b) For stationarity we need the roots of $1 - \alpha_1 z - 0.4z^2 = 0$ to lie outside the unit circle. The roots are $1.25(\alpha_1 \pm \sqrt{\alpha_1^2 + 1.6})$. The upper root will be greater than 1 if $\sqrt{\alpha_1^2 + 1.6} > 0.8 - \alpha_1$, ie if $\alpha_1 > -0.6$. By symmetry, the lower root will be less than -1 if $\alpha_1 < 0.6$. Therefore the appropriate range of values is $-0.6 < \alpha_1 < 0.6$.

For X to be I(1) we need ∇X to be stationary. We need $1 - B$ to be a factor of $(1 - \alpha_1 B - 0.4B^2)$: this means that $\alpha_1 = 0.6$ is the single permissible value, leaving the equation $(1 + 0.4B)\nabla X = (1 + 0.3B)e$, which certainly means that ∇X is stationary.

c) Where X is stationary it is ARIMA$(2, 0, 1)$. In the case $\alpha_1 = 0.6$, the model is ARIMA$(1, 1, 1)$.

d) Least squares estimation involves minimising $\sum e_t^2$, but e_t cannot be deduced from the values of the observable quantities $\{X_t\}$, since $e_t = X_t - \alpha_1 X_{t-1} - 0.4X_{t-2} - 0.3e_{t-1}$, an equation involving e_{t-1}. Once a value is specified for e_0, all other e_t can be deduced. Backforecasting is a method for obtaining an estimate of e_0 by using forecasting techniques on the time-reversed process $\{X_{n-t}\}$.

(ii) a) Yule-Walker equations: $\rho_1 = 0.1 + 0.2\rho_1$ and $\rho_2 = 0.1\rho_1 + 0.2$. These imply that $\rho_1 = \frac{1}{8}$ and that $\rho_2 = \frac{17}{80}$.

b) We need to minimise $\sum_{t=3}^{n} (Y_t - \mu - 0.1(Y_{t-1} - \mu) - 0.2(Y_{t-2} - \mu))^2$. Setting the derivative equal to 0 we have

$$0.7(n-2)\hat{\mu} = 0.7 \sum_{3}^{n-2} Y_t + Y_n + 0.9Y_{n-1} - 0.3Y_2 - 0.2Y_1.$$

Thus $\hat{\mu}$ is not quite equal to $\overline{Y}$.

c) $\hat{Y}_n(1) = \hat{\mu} + 0.1(Y_n - \hat{\mu}) + 0.2(Y_{n-1} - \hat{\mu})$ and $\hat{Y}_n(2) = \hat{\mu} + 0.1(\hat{Y}_n(1) - \hat{\mu}) + 0.2(Y_n - \hat{\mu})$.

Question 4

(i) (Bookwork) The matrix form of the C-K equations is $P^{(m+n)} = P^{(m)}P^{(n)}$.
By definition $P^{(1)} = P$.
We therefore have $P^{(n+1)} = PP^{(n)}$; it follows by mathematical induction that $P^{(n)} = P^n$.

(ii)

$$P\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad P\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} .3 \\ 0 \\ -.3 \end{pmatrix}, \quad P\begin{pmatrix} 3 \\ -5 \\ 3 \end{pmatrix} = \begin{pmatrix} .6 \\ -1 \\ .6 \end{pmatrix},$$

so that the corresponding eigenvalues are 1, 0.3 and 0.2.

(iii) Let V be the square matrix whose columns are the three eigenvectors, Λ the diagonal matrix whose entries are the eigenvalues. Then $PV = V\Lambda$, so that $P = V\Lambda V^{-1}$ and $P^n = V\Lambda^n V^{-1}$, where $\Lambda^n = \mathrm{diag}(1, 0.3^n, 0.2^n)$.

(iv) **Either** Let $n \to \infty$ in the above equation, giving

$$\lim_{n\to\infty} P^n = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} V^{-1}.$$

The limit of P^n thus has three identical rows, each being equal to the top row of V^{-1}. Standard matrix inversion techniques reveal this to be $\left(\frac{5}{16}, \frac{6}{16}, \frac{5}{16}\right)^T$.
Or According to the Ergodic Theorem: if X is irreducible and aperiodic then there exists a solution π to $\pi^T P = \pi^T$, and this is the limiting distribution of X.
Verification that X is irreducible and aperiodic.
We have $0.5\pi_1 + 0.25\pi_2 + 0.2\pi_3 = \pi_1$ and $0.5\pi_3 + 0.25\pi_2 + 0.2\pi_1 = \pi_3$, implying that $\pi_1 = \pi_3$ and that $\pi_2 = 1.2\pi_3$. Hence the solution is $\pi = \left(\frac{5}{16}, \frac{6}{16}, \frac{5}{16}\right)^T$.

(v) a) $q_{ij}(t) = \mathbb{P}(Y(t) = j \mid Y(0) = i) = \sum_n \mathbb{P}(N(t) = n)\mathbb{P}((X_n = j \mid X_0 = i) = \sum_n e^{-\lambda t}(\lambda t)^n/n! \times (V\Lambda^n V^{-1})_{ij}$.
This is the (i,j)th element of the matrix $e^{-\lambda t} V \sum_n \frac{(\lambda t)^n}{n!}\Lambda^n V^{-1}$.
The matrix $e^{-\lambda t}\sum_n \frac{(\lambda t)^n}{n!}\Lambda^n$ is a diagonal matrix, with (i,i)th entry $e^{-\theta t}\sum_n \frac{(\theta t)^n}{n!}\lambda_i^n = e^{-\theta(1-\lambda_i)t}$.

b) $\lim_{t\to\infty} \Lambda_t = \lim_{n\to\infty} \Lambda^n = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$,
implying that the limiting distribution of Y is the same as for X.

Question 5

(i) Yes. The fact that the sample ACF is close to 1 for all small k is a sure sign.

(ii) Looking at the ACF, we might try MA(2); from the PACF we might think of AR(3). The principle of parsimony indicates a model with not many parameters: MA(2) or ARMA(1,1) would do.

(iii) An ARIMA(1,1,1) model is being fitted to X.
The equation of the fitted model is

$$X_t - X_{t-1} = 0.0766 + 0.3229(X_{t-1} - X_{t-2}) + e_t - 0.8546e_{t-1}.$$

The constant term is not significantly different from 0; the others are.

(iv) Various possibilities: (2 each)
Residuals vs fitted values, aimed at detecting a relationship between fitted value and innovation variance (heteroscedasticity). A cone-shaped pattern of dots would indicate a poor fit.
Histogram of residuals, or normal probability plot. The purpose of this is to test the assumption that the innovations are normally distributed. A skewed histogram, or one exhibiting significant kurtosis, might indicate lack of normality.
(Partial) correlogram of residual time series. Any values of the sample (partial) ACF which lie outside the confidence bounds might be taken as evidence of inadequate fit.
Periodogram of the residuals. This would be expected to be roughly flat; substantial deviations indicate hidden periodicities.
Other things which may get a mark include: the Ljung-Box chi-square statistic, tests based on runs, anything else that makes sense.

(v)

a) LCG will do fine: $w_n = aw_n + c(\text{mod}M)$, $u_n = w_n/M$.

b) Polar method, for example: $z = \sqrt{-2\log u}\sin(2\pi v)$, where u and v are uniform.

c) In cell C156 enter the formula `=sigma*SIMNORM()`;
in cell B156 the formula should be `=mu+alpha1*(B155-mu)+C156+beta1*C155`;
finally, in cell A156 enter the formula `=A155+B156`.