THE ROYAL STATISTICAL SOCIETY

2010 EXAMINATIONS – SOLUTIONS

GRADUATE DIPLOMA

MODULE 3

STOCHASTIC PROCESSES AND TIME SERIES

The Society provides these solutions to assist candidates preparing for the examinations in future years and for the information of any other persons using the examinations.

The solutions should NOT be seen as "model answers". Rather, they have been written out in considerable detail and are intended as learning aids.

Users of the solutions should always be aware that in many cases there are valid alternative methods. Also, in the many cases where discussion is called for, there may be other valid points that could be made.

While every care has been taken with the preparation of these solutions, the Society will not be responsible for any errors or omissions.

The Society will not enter into any correspondence in respect of these solutions.

Note. In accordance with the convention used in the Society's examination papers, the notation log denotes logarithm to base *e*. Logarithms to any other base are explicitly identified, e.g. $\log_{10}$.

Graduate Diploma, Module 3, 2010. Question 1

(i) The transition matrix is

$$\begin{pmatrix} 1-\theta & \theta-\phi & \phi \\ 1-\theta & 0 & \theta \\ 0 & 0 & 1 \end{pmatrix}.$$

(ii) $x_0 = 1 + (1 - \theta)x_0 + (\theta - \phi)x_1$

$x_1 = 1 + (1 - \theta)x_0$

(iii) Substituting for $1 + (1 - \theta)x_0$ from the second equation into the first equation in (ii), we obtain

$$x_0 = (1 + \theta - \phi)x_1.$$

Substituting this expression into the second equation in (ii) gives

$$x_1 = 1 + (1 - \theta)(1 + \theta - \phi)x_1$$

so that

$$x_1 = \frac{1}{1-(1-\theta)(1+\theta-\phi)} = \frac{1}{\theta^2 + \phi(1-\theta)}.$$

$$\therefore\ x_0 = (1 + \theta - \phi)x_1 = \frac{1+\theta-\phi}{\theta^2+\phi(1-\theta)},$$ as required.

(iv) Setting $\theta = 1/20$ and $\phi = 1/100$ into the expression for x_0 gives

$$x_0 = \frac{1+(1/20)-(1/100)}{(1/20)^2+(1/100)(19/20)}$$

$$= \frac{1+(4/100)}{(1/400)+(1/100)(19/20)} = \frac{104}{(1/4)+(19/20)} = \frac{260}{3},$$

i.e. 87 to the nearest integer.

Graduate Diploma, Module 3, 2010. Question 2

(i) S is the set of all non-negative integers. The transition probabilities are

$p_{i\,i+1} = \theta \qquad (i \geq 0),$

$p_{i\,0} = 1 - \theta \qquad (i \geq 0).$

(All other transition probabilities are zero.)

(ii) The equations for the stationary distribution, $\pi_j = \sum_{i=0}^{\infty} \pi_i p_{ij}$, reduce to

$$\pi_0 = (1-\theta)\sum_{i=0}^{\infty}\pi_i$$

$$\pi_j = \theta\pi_{j-1} \qquad (j \geq 1),$$

together with the normalisation condition $\pi_0 + \pi_1 + \pi_2 + ... = 1$.

It readily follows that the solution is $\pi_0 = (1-\theta)$ and $\pi_j = \theta^j(1-\theta)$ for $j \geq 1$, which can be written together as $\pi_i = \theta^i(1-\theta)$ for $i \geq 0$, as required.

(iii) For $j > 6$, we have $p_{0j}^{(6)} = 0$ since there cannot be a run of more than 6 correct answers.

For $j = 6$, we have $p_{06}^{(6)} = \theta^6$, since there must be a run of 6 correct answers.

For $j = 0$, we have $p_{00}^{(6)} = 1-\theta$, since a run of no correct answers occurs if and only if the last answer was incorrect.

For $1 \leq j \leq 5$, to obtain a run of exactly j correct answers the last j answers must have been correct and the previous one incorrect. Hence $p_{0j}^{(6)} = (1-\theta)\theta^j$ for $1 \leq j \leq 5$.

So, as $X_0 = 0$ by convention, i.e. with probability 1, we have

$$P(X_6 = j) = p_{0j}^{(6)} = \begin{cases} 1-\theta & j = 0 \\ (1-\theta)\theta^j & j = 1, ..., 5 \\ \theta^6 & j = 6 \\ 0 & j > 6 \end{cases} .$$

Solution continued on next page

(iv) $$E(\text{winnings}) = \sum_{j=1}^{6} (\text{winnings} = j) P(X_6 = j)$$

$$= \sum_{j=1}^{6} 50 \times 2^j P(X_6 = j)$$

$$= 50 \sum_{j=1}^{6} 2^j P(X_6 = j)$$

$$= 50 \times \left\{ 2^1 (1-\theta)\theta + 2^2 (1-\theta)\theta^2 + 2^3 (1-\theta)\theta^3 + 2^4 (1-\theta)\theta^4 + 2^5 (1-\theta)\theta^5 + 2^6 \theta^6 \right\}$$

Inserting $\theta = \frac{1}{2}$ gives

$$50 \times \left\{ \left(2^1 \times \tfrac{1}{4}\right) + \left(2^2 \times \tfrac{1}{8}\right) + \left(2^3 \times \tfrac{1}{16}\right) + \left(2^4 \times \tfrac{1}{32}\right) + \left(2^5 \times \tfrac{1}{64}\right) + \left(2^6 \times \tfrac{1}{64}\right) \right\}$$

$$= \quad 50 \times \left\{ \tfrac{1}{2} + \tfrac{1}{2} + \tfrac{1}{2} + \tfrac{1}{2} + \tfrac{1}{2} + 1 \right\}$$

= 175 (£).

<u>Graduate Diploma, Module 3, 2010. Question 3</u>

<u>Part (i)</u>

(a) The difference equation is $x_i = \theta x_{i+1} + (1-\theta)\, x_{i-1}$ (for $1 \le i \le N-1$), together with the boundary conditions $x_0 = 1$ and $x_N = 0$.

(b) The auxiliary equation is $z = \theta z^2 + (1-\theta)$. The roots of this are $z = 1$ and $z = (1-\theta)/\theta$. Hence (provided $\theta \ne ½$ – see part (ii)(b) below) the general form of the solution of the difference equation is

$$x_i = A + B\left(\frac{1-\theta}{\theta}\right)^i .$$

The boundary conditions give $A + B = 1$ and $A + B\left(\frac{1-\theta}{\theta}\right)^N = 0$. Hence

$$A = \frac{\left(\frac{1-\theta}{\theta}\right)^N}{\left(\frac{1-\theta}{\theta}\right)^N - 1} \qquad \text{and} \qquad B = \frac{-1}{\left(\frac{1-\theta}{\theta}\right)^N - 1}.$$

Thus the solution is

$$x_i = \frac{\left(\frac{1-\theta}{\theta}\right)^N - \left(\frac{1-\theta}{\theta}\right)^i}{\left(\frac{1-\theta}{\theta}\right)^N - 1} \qquad \text{(for } 0 \le i \le N\text{).}$$

(c) The required probability is simply given by x_a.

<u>Part (ii)</u>

(a) We let $N \to \infty$ in the result of part (i)(c).

If $0 < \theta < ½$, then $\left(\frac{1-\theta}{\theta}\right)^N \to \infty$ as $N \to \infty$ and it follows that $x_a \to 1$.

If $½ < \theta < 1$, then $\left(\frac{1-\theta}{\theta}\right)^N \to 0$ as $N \to \infty$. Hence $x_a \to \left(\frac{1-\theta}{\theta}\right)^a$.

Thus the probability that Player A is eventually ruined is 1 if $0 < \theta < ½$ and is $\left(\frac{1-\theta}{\theta}\right)^a$ if $½ < \theta < 1$.

Solution continued on next page

(b) In the case $\theta = ½$, the auxiliary equation of part (i)(b) has the repeated root $z = 1$. Hence the general form of the solution of the difference equation is $x_i = A + Bi$. The boundary conditions give $A = 1$ and $A + BN = 0$, so $B = -1/N$.

Thus the solution is $x_i = 1 - (i/N)$ (for $0 \leq i \leq N$).

The probability that Player A is ruined is $x_a = 1 - (a/N)$.

If Player B has unlimited resources, in the limit as $N \to \infty$ we find that the probability that Player A is ruined is 1.

Graduate Diploma, Module 3, 2010. Question 4

(i) The state space is the set of all non-negative integers. The instantaneous transition rates are as follows.

transition	rate	
$i \to i+1$	λ	$(i \geq 0)$
$i \to i-1$	μ	$(i \geq 1)$

(ii) For $n = 0$ we have

$$p_0(t+h) = p_0(t)\left[1 - \lambda h + o(h)\right] + p_1(t)\left[\mu h + o(h)\right] + o(h)$$

and so

$$dp_0 / dt = -\lambda p_0(t) + \mu p_1(t).$$

For $n \geq 1$ we have

$$p_n(t+h) = p_n(t)\left[1 - \lambda h - \mu n h + o(h)\right] + p_{n-1}(t)\left[\lambda h + o(h)\right] + p_{n+1}(t)\left[\mu(n+1)h + o(h)\right] + o(h)$$

and so

$$dp_n / dt = -(\lambda + \mu n)\, p_n(t) + \lambda p_{n-1}(t) + \mu(n+1)\, p_{n+1}(t).$$

(iii) Multiplying the nth forward equation by z^n and summing over n,

$$\sum_{n=0}^{\infty} \frac{dp_n}{dt} z^n = -\sum_{n=0}^{\infty} (\lambda + \mu n)\, p_n(t) z^n + \sum_{n=1}^{\infty} \lambda p_{n-1}(t) z^n + \sum_{n=0}^{\infty} \mu(n+1)\, p_{n+1}(t) z^n ,$$

i.e.

$$\frac{\partial G}{\partial t} = -\lambda G - \mu z \frac{\partial G}{\partial z} + \lambda z G + \mu \frac{\partial G}{\partial z} ,$$

i.e.

$$\frac{\partial G}{\partial t} = (z-1)\left(\lambda G - \mu \frac{\partial G}{\partial z}\right), \quad \text{as required.}$$

Solution continued on next page

(iv) The suggested solution is $G(z,t) = \exp[\theta(t)(z-1)]$ where $\theta(t) = \frac{\lambda}{\mu}(1-\mathrm{e}^{-\mu t})$ and we verify this by substituting it into the left-hand and right-hand sides of the partial differential equation (as in part (iii)).

Left-hand side:

$$\frac{\partial G}{\partial t} = (z-1)\frac{d\theta}{dt}\exp[\theta(t)(z-1)] = (z-1)\,\lambda e^{-\mu t}\exp[\theta(t)(z-1)].$$

Right-hand side:

$$(z-1)\left(\lambda G - \mu\frac{\partial G}{\partial z}\right) = (z-1)(\lambda - \mu\theta(t))\exp[\theta(t)(z-1)]$$

$$= (z-1)\,\lambda e^{-\mu t}\exp[\theta(t)(z-1)].$$

These are equal, so this is indeed a solution of the partial differential equation.

Since $p_0(0) = 1$, the initial condition for the equation must be $G(z, 0) = 1$ and this is indeed satisfied by the solution (since $\theta(0) = 1$).

$G(z, t)$ is the generating function for $N(t)$. Its form is that for the Poisson distribution with mean $\theta(t)$, so this is the distribution of $N(t)$.

[Solution continues on next page]

(i) The state space is the set of all non-negative integers. The instantaneous transition rates are as follows.

transition	rate	
$i \to i+1$	λ	$(i \geq 0)$
$1 \to 0$	α	
$i \to i-1$	β	$(i \geq 2)$

(ii) The detailed balance equations are $\lambda\pi_0 = \alpha\pi_1$ and, for $n \geq 2$, $\lambda\pi_{n-1} = \beta\pi_n$.

Thus $\pi_1 = (\lambda/\alpha)\pi_0$ and $\pi_n = (\lambda/\beta)\pi_{n-1}$ $(n \geq 2)$. Using this relation recursively, we find

$$\pi_n = (\lambda/\beta)^{n-1}\pi_1 = (\lambda/\alpha)(\lambda/\beta)^{n-1}\pi_0 \quad (n \geq 1).$$

Using the normalisation condition $\sum_{n=0}^{\infty}\pi_n = 1$, we require that

$$\left\{1 + (\lambda/\alpha)\sum_{n=1}^{\infty}(\lambda/\beta)^{n-1}\right\}\pi_0 = 1.$$

This can only be satisfied it the geometric series $\sum_{n=1}^{\infty}(\lambda/\beta)^{n-1}$ converges, and the necessary and sufficient condition for this is $\lambda < \beta$.

Assuming this is the case, we have

$$\left\{1 + \frac{\lambda/\alpha}{1-(\lambda/\beta)}\right\}\pi_0 = 1.$$

$$\therefore \left\{\frac{1-\frac{\lambda}{\beta}+\frac{\lambda}{\alpha}}{1-\frac{\lambda}{\beta}}\right\}\pi_0 = 1.$$

$$\therefore \frac{\alpha\beta - \alpha\lambda + \beta\lambda}{\alpha\beta - \alpha\lambda}\pi_0 = 1.$$

Thus $\pi_0 = \dfrac{\alpha(\beta-\lambda)}{\alpha(\beta-\lambda)+\beta\lambda}$

(this is the required long-term proportion of the time that the queue is empty, i.e. the cashier is free) and, for $n \geq 1$,

$$\pi_n = \frac{\alpha(\beta-\lambda)}{\alpha(\beta-\lambda)+\beta\lambda}\left(\frac{\lambda}{\beta}\right)^{n-1}.$$

(iii) As there are now at least 2 customers at the checkout, the service rate is β. Given the Markov assumption, the residual service time of the customer currently being served and the service times for the other $k-1$ customers are all exponentially distributed with parameter β.

The length of time that the newly arrived customer has to wait before his service starts is the sum of k independent identically distributed random variables, each exponentially distributed with parameter β. This sum has the gamma distribution with parameters β and k.

<u>Graduate Diploma, Module 3, 2010. Question 6</u>

(i) The autoregressive characteristic equation is $1 - (1/2)z + (3/64)z^2 = 0$, which has roots $z = 8/3, 8$. Both the roots are greater than one in modulus, so the stationarity condition is satisfied.

(ii) Multiplying through by $Y_{t-\tau}$ in the model equation and taking expectations gives

$$\gamma_t = \tfrac{1}{2}\gamma_{t-1} - \tfrac{3}{64}\gamma_{t-2} \qquad (\tau \geq 1).$$

Dividing through by γ_0, we obtain the Yule-Walker equations

$$\rho_t = \tfrac{1}{2}\rho_{t-1} - \tfrac{3}{64}\rho_{t-2} \qquad (\tau \geq 1).$$

(iii) Taking $\tau = 1$, and using both $\rho_0 = 1$ and the symmetry condition $\rho_{-1} = \rho_1$, we obtain $\rho_1 = \tfrac{1}{2} - \tfrac{3}{64}\rho_1$ which gives $\rho_1 = 32/67$.

(iv) The general solution of the difference equation of part (ii) is of the form $\rho_\tau = A_1\alpha_1^{\tau} + A_2\alpha_2^{\tau}$ $(\tau \geq -1)$, where A_1 and A_2 are arbitrary constants and α_1 and α_2 are the roots of the auxiliary equation

$$\alpha^2 - \tfrac{1}{2}\alpha + \tfrac{3}{64} = 0.$$

The roots of the auxiliary equation are [the inverses of the roots of the characteristic equation of part (i)] 3/8 and 1/8. Hence the general solution is $\rho_\tau = A_1(3/8)^\tau + A_2(1/8)^\tau$. Using the conditions $\rho_0 = 1$ and $\rho_1 = 32/67$, we obtain

$$A_1 + A_2 = 1 \qquad \text{and} \qquad (3/8)A_1 + (1/8)A_2 = 32/67.$$

Hence $A_1 = 189/134$ and $A_2 = -55/134$, and the solution is as stated in the question.

Graduate Diploma, Module 3, 2010. Question 7

(i) Stationary time series do not show any trend or periodic variation, and do not have any systematic change in variance.

For a strictly stationary series, the observation at time t, X_t, has the following property: the joint distribution of $X_{t_1}, ..., X_{t_k}$ is the same as that of $X_{t_1+\tau}, ..., X_{t_k+\tau}$ for all t_1, …, t_k and all τ.

For a weakly stationary series, $E(X_t) = \mu$ (a constant) and $\text{Cov}(X_t, X_{t+\tau})$ is a function of the lag τ alone – the autocovariance at lag τ, often denoted by γ_τ.

(ii) An MA(1) process $\{X_t\}$ is a process that satisfies the model equation

$$X_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1},$$

in which $\{\varepsilon_t\}$ is a white noise process with mean 0. A white noise process is a stationary process whose terms are uncorrelated with each other.

Taking expectations in the model equation gives $E(X_t) = \mu$ for all t.

Taking variances in the model equation gives

$$\text{Var}(X_t) = \text{Var}(\varepsilon_t) + \theta^2\text{Var}(\varepsilon_{t-1}) = (1 + \theta^2)\sigma^2,$$

where σ^2 is the variance of the white noise process.

To obtain the ACF, first consider

$$\gamma_1 = E[(\varepsilon_t + \theta\varepsilon_{t-1})(\varepsilon_{t-1} + \theta\varepsilon_{t-2})] = \theta\sigma^2$$ (the only non-zero contribution is from $E[\theta\varepsilon_{t-1}\varepsilon_{t-1}]$)

and

$$\gamma_\tau = E[(\varepsilon_t + \theta\varepsilon_{t-1})(\varepsilon_{t-\tau} + \theta\varepsilon_{t-\tau-1})] = 0 \quad \text{for } \tau \geq 2.$$

The autocorrelations ρ_τ are now given by $\rho_\tau = \gamma_\tau/\gamma_0$. Hence

$$\rho_1 = \theta/(1 + \theta^2),$$

$$\rho_\tau = 0 \qquad (\tau \geq 2).$$

Solution continued on next page

(iii) A partial autocorrelation coefficient measures the correlation between observations *k* steps apart that is not accounted for by the autocorrelations in between.

For an MA(1) process, the PACF is damped cosine or exponential decay.

(iv) For MA(1), the ACF has a cut-off point at lag 1. It is 0 after lag 1, apart from sampling variation. The PACF does not have a clear cut-off point.

(v) Series B seems to best match these requirements for being MA(1). It has a cut-off point in the ACF at lag 1 which suggests an MA(1) model. Perhaps there is a cut-off point at lag 2 in the PACF, which would suggest an AR(2) model, though this is not so clear-cut, and we would in any case prefer the more parsimonious model.

Series C appear to be white noise, as there do not appear to be any values of the ACF or PACF that would differ significantly from zero.

(Note: approximate 95% probability limits for the sample autocorrelations and partial autocorrelations when the underlying process values are zero are at $\pm 2/\sqrt{50}$, i.e. ± 0.28.)

Series A has an exponential decay with alternating sign for the ACF, and the PACF has a cut-off point at lag 1. It appears to be AR(1) with a negative coefficient. (A first estimate of the AR parameter is –0.84.)

<u>Graduate Diploma, Module 3, 2010. Question 8</u>

(i) (a) The updating equations are as follows.

$$L_t = \alpha Y_t + (1-\alpha)(L_{t-1} + B_{t-1})$$
$$B_t = \gamma(L_t - L_{t-1}) + (1-\gamma)B_{t-1}$$

(b) $\hat{y}_T(h) = L_T + hB_T$.

(ii) At any time t, the "fitted (forecast)" value in the table is the forecast value of Y_t as made at the previous time point (t – 1), i.e. it is $\hat{y}_{t-1}(1)$ as defined in the question (that is, the forecast at time t – 1 for lead time 1), and we have $\hat{y}_{t-1}(1) = L_{t-1} + B_{t-1}$.

The "residual (error)" value is the difference between the observed value ("Temp" in the table) and the fitted value, i.e. it is $Y_t - \hat{y}_{t-1}(1)$.

For t = 10 we have

"fitted" = 23.774 + 0.357 = 24.131

"residual"” = 24.1 – 24.131 = –0.031.

(iii) Starting from the data at time t = 10, we want

$\hat{y}_{10}(1)$ = 24.109 + 0.340 = 24.449,
$\hat{y}_{10}(2)$ = 24.109 + (2 × 0.340) = 24.789,
$\hat{y}_{10}(3)$ = 24.109 + (3 × 0.340) = 25.129.

(iv) For t = 11, using the new observed value 24.6, we have

Level: L_t = (0.7)(24.6) + (0.3)(24.109 + 0.340) = 24.555,
Trend: B_t = (0.8)(24.555 – 24.109) + (0.2)(0.340) = 0.425.

Solution continued on next page

(v) For any time point t, the "deviation" e_t is the same as the "residual" as described in part (ii), so that $e_t = Y_t - \hat{y}_{t-1}(1)$. If the historical series runs from $t = 1$ to $t = T$,

$$\text{MAD} = \sum_{t=1}^{T} |e_t|/T \,,$$
$$\text{MSD} = \sum_{t=1}^{T} e_t^2/T \,.$$

In the present case,

MAD = (0.060 + 0.167 + … + 0.031)/10 = 0.087,
MSD = (0.004 + 0.028 + … + 0.001)/10 = 0.011.

(vi) Given some appropriately chosen initial values for the level and trend, for any given set of values of α and γ (each between 0 and 1 inclusive, of course), the numerical values of all the quantities in the table may be calculated for each point in the historical series. By looking at a grid of values of α and γ or by carrying out a formal optimisation, the values of α and γ that minimise the MAD or MSD may be found as the best set of values to use.