THE ROYAL STATISTICAL SOCIETY

GRADUATE DIPLOMA EXAMINATION

NEW MODULAR SCHEME

introduced from the examinations in 2009

MODULE 3

SOLUTIONS FOR SPECIMEN PAPER B

THE QUESTIONS ARE CONTAINED IN A SEPARATE FILE

The time for the examination is 3 hours. The paper contains eight questions, of which candidates are to attempt **five**. Each question carries 20 marks. An indicative mark scheme is shown within the questions, by giving an outline of the marks available for each part-question. The pass mark for the paper as a whole is 50%.

The solutions should not be seen as "model answers". Rather, they have been written out in considerable detail and are intended as learning aids. For this reason, they do not carry mark schemes. Please note that in many cases there are valid alternative methods and that, in 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 the questions and solutions, the Society will not be responsible for any errors or omissions.

The Society will not enter into any correspondence in respect of the questions or solutions.

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

© RSS 2008

Graduate Diploma Module 3, Specimen Paper B. Question 1

Part (a)

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

(ii) The auxiliary equation is $z = \theta z^2 + (1 - \theta)$, which has roots $z = 1, (1 - \theta)/\theta$. Hence the general form of the solution of the difference equation is

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

Using the boundary conditions, we find $A + B = 1$ and $A + B[(1 - \theta)/\theta]^N = 0$.

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

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

(iii) The required probability is given by x_a.

Solution continued on next page

Part (b)

(i) The difference equation is $y_i = 1 + \theta y_{i+1} + (1 - \theta)y_{i-1}$ (for $1 \le i \le N - 1$), together with the boundary conditions $y_0 = 0$ and $y_N = 0$.

(ii) A particular solution of the difference equation is of the form $y_i = Ci$ for some constant C. Substituting this form into the difference equation, we find that $C = 1/(1 - 2\theta)$. Hence the general form of the solution of the difference equation is

$$y_i = \frac{i}{1-2\theta} + A + B\left(\frac{1-\theta}{\theta}\right)^i .$$

Using the boundary conditions, we find that $A + B = 0$ and

$$\frac{N}{1-2\theta} + A + B\left(\frac{1-\theta}{\theta}\right)^N = 0 .$$

Hence

$$A = \frac{N}{1-2\theta} \times \frac{1}{\left(\frac{1-\theta}{\theta}\right)^N - 1} = -B$$

and

$$y_i = \frac{i}{1-2\theta} + \frac{N}{1-2\theta} \times \frac{1-\left(\frac{1-\theta}{\theta}\right)^i}{\left(\frac{1-\theta}{\theta}\right)^N - 1} \qquad (0 \le i \le N).$$

(iii) The required expectation is given by y_a.

Graduate Diploma Module 3, Specimen Paper B. Question 2

(i) $P(Y > k) = \phi[(1-\phi)^k + (1-\phi)^{k+1} + (1-\phi)^{k+2} + \ldots] = \dfrac{\phi(1-\phi)^k}{1-(1-\phi)} = (1-\phi)^k.$

$$\therefore P(Y = k+y \mid Y > k)) = \frac{P(Y = k+y)}{P(Y > k)} = \frac{\phi(1-\phi)^{k+y-1}}{(1-\phi)^k} = \phi(1-\phi)^{y-1} = P(Y = y).$$

(ii) The probability that a customer who is being served at time n completes service at time $n+1$ is always ϕ, by (i), irrespective of how long that customer has been waiting for service previously. Hence we have the Markov (or memory-less) property.

The transition probabilities are

$p_{01} = \theta, \quad p_{00} = 1-\theta,$

$p_{j\,j-1} = \phi(1-\theta), \quad p_{j\,j+1} = \theta(1-\phi), \quad p_{jj} = 1-\theta-\phi+2\theta\phi \; (j \ge 1).$

(iii) The equations for the stationary distribution are $\pi_j = \sum_i \pi_i p_{ij}$ $(j \ge 0)$, which in the present case ($\theta = 1/4$, $\phi = 1/2$) reduce to

$\pi_0 = (3/4)\pi_0 + (3/8)\pi_1, \quad \pi_1 = (1/4)\pi_0 + (1/2)\pi_1 + (3/8)\pi_2,$

$\pi_j = (1/8)\pi_{j-1} + (1/2)\pi_j + (3/8)\pi_{j+1} \quad (j \ge 2),$

which further simplify to the equations given in the question, i.e.

$\pi_0 = (3/2)\pi_1, \quad \pi_1 = (1/2)\pi_0 + (3/4)\pi_2, \quad \pi_j = (1/4)\pi_{j-1} + (3/4)\pi_{j+1} \; (j \ge 2).$

The difference equation for $j \ge 1$ has as its general solution

$\pi_j = A_1\alpha_1^j + A_2\alpha_2^j \quad (j \ge 1).$

α_1 and α_2 are the roots of the equation $3\alpha^2 - 4\alpha + 1 = 0$, i.e. 1 and 1/3. So the general solution is $\pi_j = A_1 + A_2(1/3)^j$ $(j \ge 1)$.

Since the distribution must satisfy the normalisation condition $\sum_j \pi_j = 1$, we must have $A_1 = 0$ and hence $\pi_j = A_2(1/3)^j$ $(j \ge 1)$. Using the normalisation condition again,

$$\pi_0 + A_2\sum_{j=1}^{\infty}\left(\frac{1}{3}\right)^j = 1, \quad \text{i.e. } \pi_0 + A_2/2 = 1.$$

We also now have $\pi_1 = A_2/3$ and $\pi_0 = (3/2)\pi_1$. Thus $A_2 = 1$, and finally

$\pi_0 = 1/2, \quad \pi_j = (1/3)^j \quad (j \ge 1).$

Graduate Diploma Module 3, Specimen Paper B. Question 3

(i) For $n = 0$, we have $p_0(t + h) = p_0(t) + p_1(t)[\mu h + o(h)] + o(h)$.

Hence $dp_0/dt = \mu p_1(t)$.

For $n \geq 1$, we have

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

Hence

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

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

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

i.e.

$$\partial G/\partial t = -(\lambda + \mu)z\partial G/\partial z + \lambda z^2\partial G/\partial z + \mu\,\partial G/\partial z = (\lambda z - \mu)(z - 1)\partial G/\partial z.$$

(iii) Differentiating the equation of part (ii) with respect to z,

$$\partial^2 G/\partial z\partial t = \lambda(z - 1)\partial G/\partial z + (\lambda z - \mu)\partial G/\partial z + (\lambda z - \mu)(z - 1)\partial^2 G/\partial z^2.$$

Setting $z = 1$, and using the fact that $m(t) = \partial G/\partial z|_{z=1}$, we obtain

$$dm/dt = (\lambda - \mu)m(t).$$

(iv) The equation of part (iii) is a first order ordinary differential equation whose general solution is of the form

$$m(t) = A\exp[(\lambda - \mu)t]$$

where the constant A is determined by the initial condition $m(0) = a$.

Thus $A = a$ and

$$m(t) = a\exp[(\lambda - \mu)t] \quad (t \geq 0).$$

Graduate Diploma Module 3, Specimen Paper B. Question 4

(i) The detailed balance equations are $(\lambda/n)\pi_{n-1} = \mu\pi_n \quad (n \ge 1)$.

Letting $\rho = \lambda/\mu$, it follows that $\pi_n = (\rho^n/n!)\pi_0 \quad (n \ge 0)$. To satisfy the normalisation condition $\sum_{n=0}^{\infty}\pi_n = 1$, we therefore require that $\sum_{n=0}^{\infty}\left(\rho^n/n!\right)\pi_0 = 1$, i.e. $e^{\rho}\pi_0 = 1$ on summing the exponential series. Thus the condition can always be satisfied by taking $\pi_0 = e^{-\rho}$. So the equilibrium distribution is a Poisson distribution with mean ρ for all values of the parameters λ and μ, i.e.

$$\pi_n = e^{-\rho}\rho^n/n! \quad (n \ge 0).$$

(ii) The mean rate at which customers <u>join</u> the queue is

$$\sum_{n=0}^{\infty}\frac{\pi_n\lambda}{n+1} = \sum_{n=0}^{\infty}\frac{e^{-\rho}\rho^n}{n!}\frac{\lambda}{n+1} = \mu e^{-\rho}\sum_{n=0}^{\infty}\frac{\rho^{n+1}}{(n+1)!} = \mu e^{-\rho}\left(e^{\rho}-1\right) = \mu\left(1-e^{-\rho}\right).$$

(iii) The probability θ that a randomly arriving customer joins the queue is given by

$$\theta = \sum_{n=0}^{\infty}\pi_n\frac{1}{n+1}.$$

Comparing this with the expressions in part (ii), we find

$$\theta = \mu(1 - e^{-\rho})/\lambda = (1 - e^{-\rho})/\rho.$$

(iv) In equilibrium, let the random variable X denote the number of customers in the queue when a customer arrives. Let B denote the event that the arriving customer joins the queue. Using the basic conditional probability formula,

$$P(X = n \mid B) = P(B \mid X = n)P(X = n)/P(B)$$

$$= \frac{\frac{1}{n+1}\frac{e^{-\rho}\rho^n}{n!}}{\frac{1-e^{-\rho}}{\rho}} = \frac{1}{e^{\rho}-1}\frac{\rho^{n+1}}{(n+1)!} \quad (n \ge 0).$$

This is a truncated Poisson distribution, with the first term of the Poisson distribution missing.

<u>Graduate Diploma Module 3, Specimen Paper B. Question 5</u>

(i) $\phi(z) = 1 + \sum_{k=1}^{p} \phi_k z^k$; $\theta(z) = 1 + \sum_{i=1}^{q} \theta_i z^i$. The model may be written as $\phi(L)Y_t = \theta(L)\varepsilon_t$.

(ii) The infinite moving average expression is $Y_t = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}$ where, for the series to be convergent [in mean square], the coefficients ψ_i must satisfy the condition $\sum_{i=0}^{\infty} \psi_i^2 < \infty$. The process $\{Y_t\}$ is stationary if and only if such an infinite moving average expression exists.

(iii) The stationarity criterion is that the roots of the autoregressive characteristic equation $\phi(z) = 0$ are outside the unit circle in the complex plane.

We now have the special case $Y_t = \frac{1}{2} Y_{t-1} + \varepsilon_t + \varepsilon_{t-1} + \frac{1}{4} \varepsilon_{t-2}$ $(-\infty < t < \infty)$.

(iv) The autoregressive characteristic equation is $1 - \frac{1}{2} z = 0$, whose root $z = 2$ lies outside the unit circle. Hence the process is stationary.

(v) The model may be rewritten as $(1 - \frac{1}{2} L)Y_t = (1 + L + \frac{1}{4} L^2)\varepsilon_t$. Hence

$$y_t = (1 + L + \tfrac{1}{4}L^2)(1 - \tfrac{1}{2}L)^{-1}\varepsilon_t$$

$$= (1 + L + \tfrac{1}{4}L^2)(1 + \tfrac{1}{2}L + \tfrac{1}{4}L^2 + ...)\varepsilon_t$$

$$= (1 + \tfrac{3}{2}L + L^2 + \tfrac{1}{2}L^3 + ...)\varepsilon_t$$

$$= \varepsilon_t + \tfrac{3}{2}\varepsilon_{t-1} + \sum_{i=2}^{\infty} \left(\tfrac{1}{2}\right)^{i-2} \varepsilon_{t-i} \ .$$

(vi) In general, $\text{Var}(Y_t) = \sum_{i=0}^{\infty} \psi_i^2 \sigma^2$. Thus, in the present case,

$$\text{Var}(Y_t) = \left[1 + (3/2)^2 + \sum_{i=2}^{\infty} (1/4)^{i-2}\right]\sigma^2$$

$$= [1 + (9/4) + (4/3)]\sigma^2 = 55\sigma^2/12.$$

Graduate Diploma Module 3, Specimen Paper B. Question 6

(i) The acf for the raw data dies away slowly, more or less linearly, which indicates the presence of trend. The purpose of taking first differences is to produce a series that is stationary.

Inspection of the acf of the differenced data reveals that it dies away quickly, which suggests that the differenced data constitute a stationary series. Approximate 95% limits for checking which of the autocorrelations or partial autocorrelations differ significantly from 0 are at $\pm 2/\sqrt{167}$, i.e. at ± 0.155. There is a clear cut-off point at lag 1 in the acf, which suggests an MA(1) model for the differenced data, i.e. an ARIMA(0,1,1) model, and this is the most obvious model to try to fit. The cut-off point at lag 2 in the pacf suggests an AR(2) model for the differenced data, i.e. an ARIMA(2,1,0) model. An ARIMA(1,1,0) model might also be tried.

(ii) An ARIMA(0,1,1) model has been fitted. The equation of the fitted model is

$$Y_t = Y_{t-1} + \varepsilon_t - 0.3434\varepsilon_{t-1}.$$

(iii) If the model is an adequate one, the residuals should behave like white noise. In particular, the acf should be zero, subject to sampling errors.

The Box-Pierce statistics are functions of the autocorrelations r_τ at lag τ of the residuals. The statistic at lag k is based on the sum of the squares of the first k autocorrelations, i.e. on $\sum_{\tau=1}^{k} r_\tau^2$, and under the white noise hypothesis has a chi-squared distribution. If the associated p-values are not significant, this suggests that the fitted model is an adequate one.

In the present case, the model appears adequate, as none of the p-values is anywhere near significant.

(iv) The one-step ahead forecast error is a white noise term, whose variance is estimated by the residual mean square 0.00005335. The values of the Y_t and the ε_t are very plausibly Normally distributed (using the central limit theorem and the fact that the exchange rates are averages over a month). Hence an approximate 95% prediction interval for the average exchange rate for January 2008 is given by $0.4846 \pm 1.96\sqrt{0.00005335}$, i.e. 0.4846 ± 0.0143, i.e. the interval is (0.4703, 0.4989).

The observed value of 0.4785 lies well within this interval, so there is nothing particularly surprising about the discrepancy between the observed and forecast values.

Graduate Diploma Module 3, Specimen Paper B. Question 7

(i) The updating equations are as follows.

$$L_t = \alpha\left(Y_t / I_{t-p}\right) + (1-\alpha)\left(L_{t-1} + B_{t-1}\right)$$
$$B_t = \gamma\left(L_t - L_{t-1}\right) + (1-\gamma)B_{t-1}$$
$$I_t = \delta\left(Y_t / L_t\right) + (1-\delta)I_{t-p}$$

(ii) $\hat{y}_T(h) = (L_T + hB_T)\, I_{T-p+h.}$

(iii) (a) $\hat{y}_T(1) = (4168.42 + 29.22)(0.665) = 2791.43$

(b) $\hat{y}_T(12) = [4168.42 + (12)(29.22)](1.391) = 6286.01$

(iv) For January 1995, the values are as follows.

Level: $L_t = 0.2(2367/0.665) + 0.8(4168.42 + 29.22) = 4069.99$

Trend: $B_t = 0.1(4069.99 - 4168.42) + 0.9(29.22) = 16.46$

Index: $I_t = 0.2(2367/4069.99) + 0.8(0.665) = 0.648$

Fitted: From (iii), $\hat{y}_T(1) = 2791.43$

Residual: Sales – Fitted = 2367 – 2791.43 = –424.43

(v) Given some appropriately chosen initial values for the level and the trend and for the first twelve seasonal indices, for any given set of values of the smoothing constants α, γ and δ (each between 0 and 1 inclusive, of course) the numerical values of all the quantities in the table may be calculated for each month in the series. The sum of squares of the residuals (or some other appropriate function of the residuals) may be used as a measure of how well the Holt-Winters method with the chosen values of α, γ and δ performs. By looking at a grid of values of α, γ and δ or by carrying out a formal optimisation, the values that minimise the sum of squares of the residuals may be found as the best set of values to use.

Graduate Diploma Module 3, Specimen Paper B. Question 8

(i) Using the fact that $\rho_0 = 1$ and the symmetry property of the autocorrelation function ($\rho_\tau = \rho_{-\tau}$), we may write

$$f(\omega) = \frac{1}{2\pi}\sum_{\tau=-\infty}^{\infty} \rho_\tau e^{-i\omega\tau} = \frac{1}{2\pi}\left[1 + \sum_{\tau=1}^{\infty} \rho_\tau \left(e^{-i\omega\tau} + e^{i\omega\tau}\right)\right] = \frac{1}{2\pi}\left[1 + 2\sum_{\tau=1}^{\infty} \rho_\tau \cos\omega\tau\right]$$

Since $\cos(\omega\tau) = \cos(-\omega\tau)$, it follows that $f(\omega) = f(-\omega)$ $(-\pi \le \omega \le \pi)$.

(ii) $f(\omega)$ gives the relative weight of oscillations of frequency ω for the process. More specifically, for $0 \le \omega_1 < \omega_2 \le \pi$, $2\int_{\omega_1}^{\omega_2} f(\omega)d\omega$ gives the proportion of the process variance that is accounted for by oscillations with frequencies in the interval (ω_1, ω_2).

(iii) For a white noise process $\{\varepsilon_t\}$, $\rho_\tau = 0$ (for $\tau \ne 0$), so that $f(\omega) = 1/(2\pi)$ (for $-\pi \le \omega \le \pi$). White noise has a flat spectrum – oscillations of all frequencies are equally represented in the process.

(iv) $\text{Var}(Y_t) = E(Y_t^2) = E[(\varepsilon_t - \theta\varepsilon_{t-1})^2] = (1 + \theta^2)\sigma^2$, where σ^2 is the white noise variance.

$\text{Cov}(Y_t, Y_{t-1}) = E[(\varepsilon_t - \theta\varepsilon_{t-1})(\varepsilon_{t-1} - \theta\varepsilon_{t-2})] = -\theta\sigma^2$.

For $\tau \ge 2$, $\text{Cov}(Y_t, Y_{t-\tau}) = E[(\varepsilon_t - \theta\varepsilon_{t-1})(\varepsilon_{t-\tau} - \theta\varepsilon_{t-\tau-1})] = 0$.

Hence for the MA(1) process,

$$\rho_1 = -\frac{\theta}{1+\theta^2}, \qquad \rho_\tau = 0 \text{ for } \tau \ge 2,$$

It follows from the formula of part (i) that

$$f(\omega) = \frac{1}{2\pi}\left(1 - \frac{2\theta}{1+\theta^2}\cos\omega\right) \qquad (0 \le \omega \le \pi).$$

(v) If $\theta > 0$, $f(\omega)$ is a monotonic increasing function of ω $(0 \le \omega \le \pi)$. Higher frequency fluctuations of short period predominate. The process is more "jagged" than white noise.

If $\theta < 0$, $f(\omega)$ is a monotonic decreasing function of ω $(0 \le \omega \le \pi)$. Lower frequency fluctuations of long period predominate. The process is smoother than white noise.