EXAMINATIONS OF THE ROYAL STATISTICAL SOCIETY

GRADUATE DIPLOMA, 2013

MODULE 3 : Stochastic processes and time series

Time allowed: Three hours

Candidates should answer **FIVE** *questions.*

All questions carry equal marks.
The number of marks allotted for each part-question is shown in brackets.

Graph paper and Official tables are provided.

Candidates may use calculators in accordance with the regulations published in the Society's "Guide to Examinations" (document Ex1).

The notation log *denotes logarithm to base* ***e***.
Logarithms to any other base are explicitly identified, e.g. $\log_{10}$.

Note also that $\binom{n}{r}$ *is the same as* ${}^{n}C_{r}$.

This examination paper consists of 12 printed pages.
This front cover is page 1.
Question 1 starts on page 2.

There are 8 questions altogether in the paper.

1. (a) In the context of a time-homogeneous Markov chain with states $i = 0, 1, 2, \ldots$, define the n-step transition probability $p_{ij}^{(n)}$. Derive the Chapman-Kolmogorov equations for the $p_{ij}^{(n)}$ and express them in terms of transition matrices. Indicate clearly at which step(s) in your derivation you make use of

(i) time-homogeneity,

(ii) the Markov property.

(6)

(b) In the following extract from a child's book-game, played with a fair 6-sided die, the heroine Crystil is ambushed by her arch-enemy Dr Zarg. The reader rolls the die with the outcomes in the table below. He or she continues to do this until Dr Zarg flees into the night and the reader turns to page 51 of the book.

Score on die	Outcome
1	*Dr Zarg stuns Crystil with his microtaser.*
2	*Dr Zarg wounds Crystil with his astromace.*
3	*Crystil temporarily blinds Dr Zarg with a thunderflash. Add 1 to your next die roll.*
4	*Crystil injures Dr Zarg with her rapier of truth. Add 2 to your next roll.*
5 or 6	*Dr Zarg flees into the night. Turn to page 51.*

(i) Treating a score of more than 6 as 6, set up the transition matrix for a Markov chain model of this situation, with four states: next roll normal, add 1 to next roll, add 2 to next roll, Dr Zarg flees into the night.

(4)

(ii) Calculate the matrix of 2-step ahead transition probabilities. Deduce the probability that it will take more than two further rolls of the die before Dr Zarg flees into the night if

(A) Dr Zarg has just stunned Crystil,

(B) Crystil has just injured Dr Zarg with her rapier of truth.

(5)

(iii) Calculate the expected number of throws needed for Dr Zarg to flee into the night so that the reader can turn to page 51.

(5)

2. (i) A generalised birth and death process is a continuous time Markov chain whose state space is the set of all non-negative integers. The instantaneous transition rates are as follows.

	Transition	Rate
Birth	$n \to n+1$	α_n
Death	$n \to n-1$	β_n

Explain why, assuming it exists, the equilibrium probability distribution π_n, for n = 0, 1, 2, …, must satisfy

$$(\alpha_n + \beta_n)\pi_n = \alpha_{n-1}\pi_{n-1} + \beta_{n+1}\pi_{n+1}, \text{ for } n = 1, 2, \ldots .$$

(3)

State the corresponding equation for n = 0 and deduce that

$$\pi_n = \frac{\alpha_{n-1}\alpha_{n-2}\ldots\alpha_0}{\beta_n\beta_{n-1}\ldots\beta_1}\pi_0, \text{ for } n = 1, 2, 3, \ldots .$$

Explain how you would evaluate π_0.

(7)

(ii) Consider a newspaper stand at which potential customers arrive randomly at a maximum rate of 16 per minute. The sight of a queue puts many people off and, if there are n customers being served or waiting to be served, only a proportion $(n + 1)^{-1}$ of arriving customers join the queue. The newspaper seller can serve on average 20 customers per minute with independent exponentially distributed service times.

(a) Formulate this as a birth and death process, and show that, in the steady state, the number of customers at the newspaper stand follows a Poisson distribution.

(5)

(b) Calculate the number of customers using the newspaper stand per hour on average.

(5)

3. Consider an M/G/1 queue in equilibrium with arrival rate λ and suppose that the service time, X, has a probability density function $f(x)$ whose mean is $\frac{1}{\mu}$ and whose moment generating function is $M(\theta)$.

The probability that j customers arrive during the service time of an arbitrarily chosen customer is denoted by π_j and the corresponding probability generating function is $\Pi(z)$.

(i) By conditioning on X, write down an expression for π_j and hence show that $\Pi(z) = M(\lambda(z-1))$.

(6)

(ii) You are given that the probability generating function of N_S, the number of customers left behind in the system by a customer who has just completed service, is

$$P(z) = \frac{(z-1)p_0\Pi(z)}{z - \Pi(z)}$$

where p_0 is the probability that N_S is 0. Prove that $p_0 = 1-\rho$, where $\rho = \frac{\lambda}{\mu}$.

(6)

[Hint. Consider $\lim_{z\to 1} P(z)$ and use l'Hôpital's Rule, which says that if each of the two functions $f(x)$ and $g(x)$ tends to zero as $x \to a$ then the limit of their ratio, $\frac{f(x)}{g(x)}$ as $x \to a$, is equal to the limit of the ratio of their derivatives, $\frac{f'(x)}{g'(x)}$ as $x \to a$.]

(iii) Define $P_Q(z)$ to be the probability generating function of N_Q, the number in the queue waiting for service immediately after a customer has departed. Show that

$$P_Q(z) = p_0 + \frac{P(z) - p_0}{z}$$

and express $P_Q(z)$ as a function of λ, μ, z and $M(.)$.

(8)

4. (i) State the assumptions for a Poisson process for incidents occurring in time at rate λ, and use them to show that $p_0(t)$, the probability that there is no incident by time t, starting from time zero, is given by

$$p_0(t) = e^{-\lambda t}.$$

Deduce that the time to the first incident has an exponential distribution.

(10)

(ii) Now suppose that the rate of the process, instead of being constant, varies with time t and is in fact equal to t.

(a) Show that, for this process,

$$p_0(t) = e^{-\frac{1}{2}t^2}$$

and find the expected time to the first incident.

(6)

[You may use, without proof, the result that $\int_0^\infty \sqrt{u}\, e^{-u}\, du = \frac{\sqrt{\pi}}{2}$.]

(b) If the first incident is known to have occurred at time s, show that the probability that no incidents occur in $(s, s + t]$ is given by

$$p_0(t \mid s) = \exp\left(-st - \tfrac{1}{2}t^2\right).$$

Deduce that the independence property of the intervals between incidents, which is known to hold for the constant rate process, fails for this variable rate process.

(4)

5. Define the autocorrelation function, ρ_k, k = 0,1,2, ..., of a stationary time series, X_t.

(2)

The time series W_t satisfies

$$W_t = 2.6W_{t-1} - 2.24W_{t-2} + 0.64W_{t-3} + A_t - 1.3A_{t-1} + 0.4A_{t-2}$$

where A_t is a white noise series with variance σ^2.

(i) Show that this model can be expressed as an ARIMA(1, 1, 1) model and verify that $Z_t = W_t - W_{t-1}$ is stationary and invertible.

(8)

(ii) Calculate $\text{Cov}(Z_t, A_t)$ and show that $\text{Var}(Z_t) = 1.25\sigma^2$.

(5)

(iii) For Z_t, show that $\rho_k = 0.8\rho_{k-1}$ for $k \geq 2$, evaluate ρ_1 and sketch ρ_k, k= 2, 3, ...

(5)

6. Figure 1 of Appendix A shows the UK's daily gas consumption x_t (t = 1, …, 366), in coded units for a year. In addition figures 2 to 4 are plots of the sample autocorrelation function (acf) and partial autocorrelation function (pacf) for the series Δx_t of first differences and the series $\Delta\Delta_7 x_t$ obtained by seasonally differencing Δx_t at lag 7.

(i) Referring to figures 1 and 2, explain why the series has been differenced and then seasonally differenced.

In subsequent model fitting the analyst did not include a constant in any models. Why was this?

(4)

(ii) In the light of figures 3 and 4, comment on the following seasonal ARIMA models as candidates for $y_t = \Delta\Delta_7 x_t$.

(a) White noise.

(b) MA(1).

(c) MA with parameters at lags 1 and 7.

(d) AR with parameters at lags 1 and 7.

(7)

(iii) In the extracts from computer output shown in Appendix B, seasonal ARIMA models have been fitted to the data. In each case identify the member of the seasonal ARIMA class and write down explicitly the equation fitted. Compare the two models from the point of view of

(a) how well they fit the data,

(b) model adequacy.

(9)

Appendices A and B are on the next 2 pages

Appendix A to Question 6

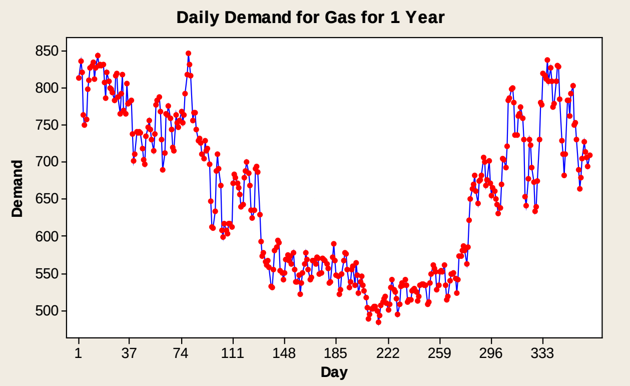

Figure 1

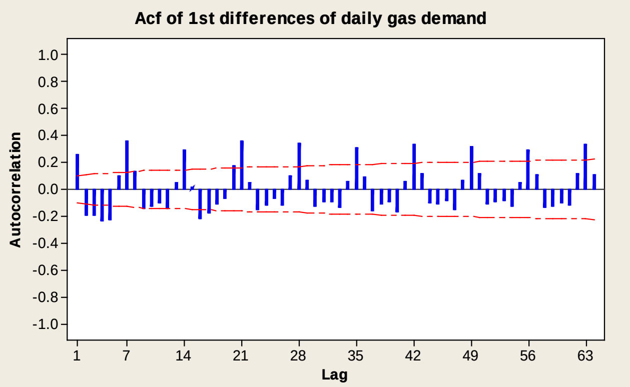

Figure 2

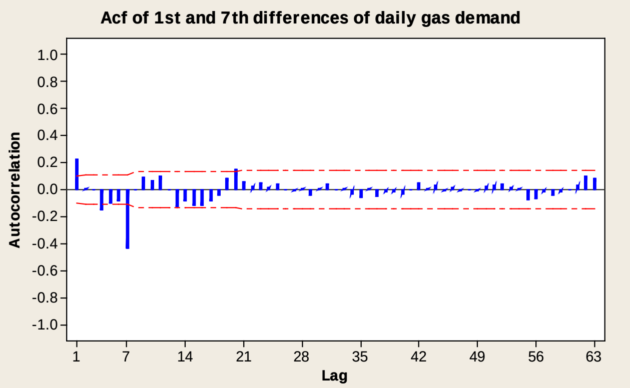

Figure 3

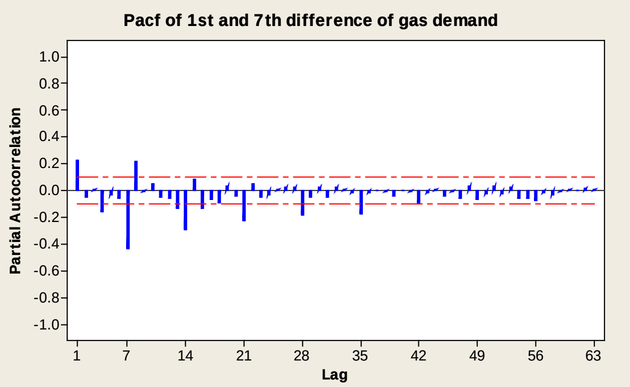

Figure 4

Appendix B to Question 6

ARIMA Model 1: Demand

```
Final Estimates of Parameters

Type        Coef  SE Coef      T      P
MA   1  -0.2399   0.0515  -4.66  0.000

Differencing: 1 regular, 1 seasonal of order 7
Number of observations:  Original series 366, after differencing 358
Residuals:    SS =  204064 (backforecasts excluded)
              MS =  572  DF = 357

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag             12      24      36      48
Chi-Square    98.3   120.6   125.8   130.6
DF              11      23      35      47
P-Value      0.000   0.000   0.000   0.000
```

ARIMA Model 2 : Demand

```
Final Estimates of Parameters

Type        Coef  SE Coef      T      P
MA   1  -0.3033   0.0507  -5.98  0.000
SMA  7   0.9523   0.0224  42.58  0.000

Differencing: 1 regular, 1 seasonal of order 7
Number of observations:  Original series 366, after differencing 358
Residuals:    SS =  101792 (backforecasts excluded)
              MS =  286  DF = 356

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag             12      24      36      48
Chi-Square    18.7    40.2    49.5    56.9
DF              10      22      34      46
P-Value      0.045   0.010   0.041   0.130
```

7. The spectral density function $f(\omega)$ of a stationary time series model having autocovariance function $\{\gamma_k\}$ (k = 0, 1, 2, …) may be written as

$$f(\omega) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} \gamma_k e^{-i\omega k}, \text{ for } -\pi < \omega < \pi.$$

(i) Stating any general properties of the autocorrelation function that you assume, show that $f(\omega)$ may be written equivalently as

$$f(\omega) = \frac{\gamma_0}{2\pi}\left\{1 + 2\sum_{k=1}^{\infty} \rho_k \cos \omega k\right\},$$

where ρ_k is the autocorrelation function of the time series. Deduce that $f(\omega) = f(-\omega)$, for $-\pi < \omega < \pi$.

(7)

(ii) Evaluate $f(\omega)$

(a) for a white noise series A_t with variance σ^2,

(b) for the time series $Y_t = A_t + 0.6A_{t-4}$.

(7)

(iii) In each case (a) and (b), sketch $g(\omega) = 2\pi\dfrac{f(\omega)}{\gamma_0}$ for $0 < \omega < \pi$ and comment on what the shape of this function tells you about the characteristics of the series.

(6)

8. (i) The time series X_t is second order stationary with mean μ, variance γ_0 and autocorrelation function (acf) ρ_k, for $k \ge 0$. Explain what is meant by the phrase *second order stationary*.

(3)

(ii) Find expressions, in terms of μ, γ_0 and ρ_k, for the mean, variance and acf of $Y_t = aX_t + bX_{t-1}$ where a and b are constants and deduce that Y_t is also second order stationary.

(9)

(iii) Deduce also that the time series $V_t = X_t - 2X_{t-1} + X_{t-2}$ is second order stationary.

(2)

(iv) If the acf of X_t is α^k, for $k \ge 0$, find the value(s) of ω such that $X_t - \omega X_{t-1}$ is white noise.

(6)

BLANK PAGE

BLANK PAGE

BLANK PAGE