EXAMINATIONS OF THE ROYAL STATISTICAL SOCIETY

GRADUATE DIPLOMA, 2017

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.

© RSS 2017

1. In a Markov chain model for the state of a machine in a factory, X_n denotes the state in week n, where the state space is $\{0, 1, 2, \ldots, k\}$. The state k represents a machine that is in perfect working order while in state 0 the machine has broken down. The intermediate states $k - 1, \ldots, 3, 2, 1$ represent successively worse states of operation. If in any week the machine is operating in state i, where $1 \le i \le k$, it has a probability p of deteriorating by one state, where $0 < p < 1$. Otherwise its condition is unchanged. A machine that has broken down remains in that state.

(i) Write down the transition matrix of this Markov chain and classify the states as either transient or recurrent.

(4)

(ii) Let w_i be the expected number of weeks until a machine currently in state i breaks down. For $1 \le i \le k$ write down a recurrence relationship for w_i. For a machine in perfect working order deduce that the expected number of weeks until break down is $\dfrac{k}{p}$.

(5)

Suppose now that when the machine reaches state 0 it is sent away for repair. Any machine being repaired has a probability $\frac{2}{3}$ of returning in state k the following week and resuming operation, independently of how many weeks it has been away for repair.

(iii) Write down the transition matrix of the Markov chain for this new situation and classify the states as either transient or recurrent.

(2)

(iv) Write down a set of equations satisfied by the stationary distribution π_i for $i = 0, 1, \ldots, k$ and deduce expressions for the probabilities π_i in terms of p and k.

(6)

(v) The machine becomes more costly to operate as it deteriorates; specifically, the cost per week of operating in state i is $1.1^{k-i} \times C$ for $1 \le i \le k$ and the cost of repair is R. Show that the long-term average cost per week of running and repairing the machine is given by

$$\frac{1}{3p+2k}\left[3pR + 20C\left(1.1^k - 1\right)\right].$$

(3)

2. In a branching process, the numbers of offspring produced by different individuals are statistically independent of each other and these numbers all have probability generating function (pgf) $G(s)$. Let X_n denote the size of generation n and take $X_0 = 1$.

Let $G_n(s)$ denote the pgf of X_n.

(i) Prove that $G_{n+1}(s) = G_n(G(s))$ for $n = 0, 1, 2, \ldots$.

(5)

Consider the case where $G(s) = \dfrac{1+2s}{4-s}$.

(ii) Verify that the equation in part (i) is satisfied by

$$G_n(s) = \frac{n-(n-3)s}{3-n(s-1)}, \quad n = 1, 2, \ldots .$$

(4)

(iii) Find $E(X_n)$ and $\mathrm{Var}(X_n)$.

(6)

(iv) Show that $E(X_n \mid X_n \neq 0) = 1 + \dfrac{n}{3}$.

(3)

(v) Is extinction certain for this population? Justify your answer.

(2)

3. (i) A queuing system is described as M/M/k, FIFO, with arrival rate λ and mean service time $\frac{1}{\mu}$. Explain what this implies about the queue.

(4)

(ii) For this queue the equilibrium probability distribution of N, the number in the system, is given by

$$p_n = P(N = n) = \begin{cases} \frac{(k\rho)^n}{n!} p_0, & \text{for } 0 \le n \le k-1, \\ \frac{k^k}{k!} \rho^n p_0, & \text{for } n \ge k, \end{cases}$$

where $\rho = \frac{\lambda}{k\mu} < 1$.

For the case k = 2 evaluate p_0 and hence, or otherwise, show that the proportion of customers who are served immediately on arrival is

$$P_a = \frac{(1+2\rho)(1-\rho)}{1+\rho}.$$

(5)

(iii) Suppose instead that you have two independent M/M/1 queues, both with mean service time $\frac{1}{\mu}$. Customers arrive at rate λ but now are equally likely to choose either queue. Write down the arrival rate for each queue and show that, in this case, the proportion of all customers who are served immediately on arrival is $P_b = 1-\rho$, where ρ is defined in part (ii). Show that P_a always exceeds P_b. (6)

(iv) A small supermarket has two self-service scanners where customers can pay for their purchases. Customers arrive at these randomly at an average rate of 20 per hour and have independent exponentially distributed service times taking, on average, 5 minutes each to complete their purchases. Currently there is a single queue for both scanners but a new manager is considering two further options:

(a) relocating one of the scanners to the opposite end of the shop so that there will be separate queues for each machine with customers assumed equally likely to enter either queue;

(b) selling one scanner leaving only one to serve all the customers.

For the three possible options (including the existing setup) calculate, where possible, the proportion of customers who will be served immediately on arrival and advise the manager appropriately.

(5)

4. (i) In a Poisson process for incidents occurring in time at rate λ let $N(a, b]$ denote the number of incidents in the interval $(a, b]$. One of the conditions for a Poisson process is that

$$P\big(N(t,\, t+\delta t]=1\big)=\lambda\delta t+o(\delta t)\,.$$

What is meant by the term $o(\delta t)$ in this equation? State the corresponding condition on $P\big(N(t,\, t+\delta t]=0\big)$ and the independence assumption required to define the Poisson process.

(3)

(ii) Using the conditions in part (i) show that $p_0(t)$, the probability that no incident occurs up to time t, is given by

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

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

(6)

(iii) Now suppose that there are k independent Poisson processes having rates $\lambda_1, \lambda_2, \ldots, \lambda_k$ and that $N_i(a, b]$ denotes the number of incidents in $(a, b]$ in the ith process. Hence $N(a, b]=\sum_{i=1}^{k} N_i(a, b]$ denotes the total number of incidents in $(a, b]$. Show that N satisfies the conditions in part (i) and hence find the distribution of the time to the first incident throughout all k processes.

(11)

5. The stationary time series X_t satisfies

$$X_t = 3 + 0.6X_{t-1} + A_t - A_{t-1},$$

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

(i) Show that X_t has mean $\mu = 7.5$.

(2)

(ii) Show that $\text{Cov}(X_t, A_t) = \sigma^2$ and that $\text{Var}(X_t) = 1.25\sigma^2$, indicating where you have used the stationarity of X_t.

(6)

(iii) Calculate ρ_1, the lag one autocorrelation of X_t, and deduce that the auto-correlation function (acf) is given by $\rho_k = -0.2 \times 0.6^{k-1}$, for $k = 1, 2, 3, \dots$.

(7)

(iv) Write down, or calculate, the mean and acf of the stationary time series model

$$Y_t = 0.6Y_{t-1} + E_t,$$

where E_t is a white noise series with variance τ^2.

(2)

(v) A statistician, studying realisations of X_t and Y_t, has plotted each series against time. Suggest, with reasons, two ways in which the two plots will differ visually.

(3)

6. 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

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

(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.

(5)

(ii) Deduce that $f(\omega) = f(-\omega)$, for $-\pi \le \omega \le \pi$.

(1)

(iii) Show that for a white noise series A_t with variance σ^2, $f(\omega)$ is constant. What does this tell you about the likely appearance of realisations of white noise?

(3)

(iv) Calculate ρ_k for the time series model $Y_t = A_t - \alpha A_{t-3}$, where $|\alpha| < 1$, and show that its spectral density function is

$$f(\omega) = \frac{\sigma^2}{2\pi} \{1 + \alpha^2 - 2\alpha \cos(3\omega)\} .$$

(6)

(v) For the case $\alpha > 0$ identify the values of ω in the range $-\pi \le \omega \le \pi$ where the spectral density of Y_t takes its maximum and express that maximum value as a function of α.

(3)

(vi) The special case $\omega = \pi$ provides some insight into the spectral density at high frequencies. Describe how the parameter α affects $f(\omega)$ for high frequencies. What does this tell you about how typical realisations of Y_t will compare with white noise?

(2)

7. In the plots **on the next page** the time series **`Prices`** consists of 200 daily commodity prices, x_t for t = 1, ..., 200, in coded units. Figures 1 and 2 are time series plots of, respectively, **`Prices`** and the series **`diff1`** of first differences, $\Delta x_t = x_t - x_{t-1}$. Figures 3 and 4 are plots of the sample autocorrelation function (acf) and partial autocorrelation function (pacf) of Δx_t.

(i) Referring to Figures 1 and 2, explain why the series has been differenced once.

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

(3)

(ii) In Figures 3 and 4, the horizontal dotted lines are at $\pm\frac{2}{\sqrt{199}}$. What role do these play in the identification of possible models for Δx_t?

In the light of Figures 3 and 4, comment on the following ARIMA models as candidates for $y_t = \Delta x_t$:

(a) white noise;

(b) AR(p);

(c) MA(q).

(6)

(iii) The edited computer output **on the second page following** shows some results obtained by fitting three ARIMA models to the commodity price series. The software used fits an ARIMA(p, d, q) model to x_t by first taking dth differences of x_t to get y_t and then fitting the model in the form

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \theta_0 + A_t + \sum_{i=1}^{q} \theta_i A_{t-i} \,,$$

where A_t is white noise.

State which models for x_t have been fitted and write down explicitly their model equations in terms of y_t. Consider briefly whether their parameter estimates are statistically significant.

What role do the horizontal dotted lines play in the plots of the residual acf?

(8)

(iv) What do you learn about the suitability of each of the models from

(a) the plot of the residual acf,

(b) the AIC value?

(3)

Diagrams and computer output are on the next two pages

Figure 1

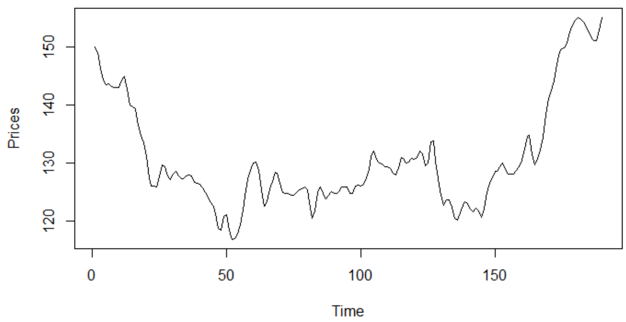


Figure 2

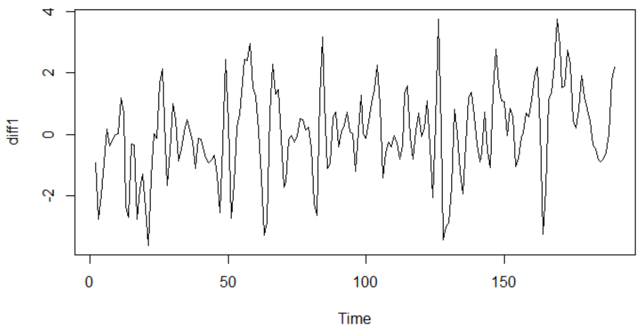


Figure 3

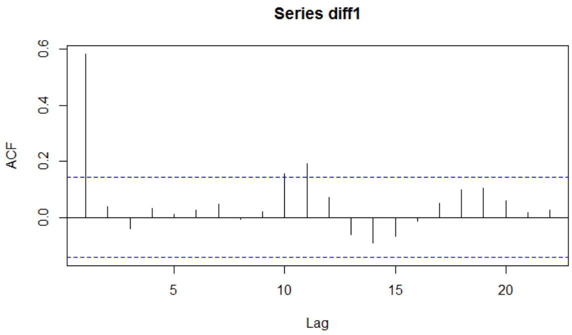
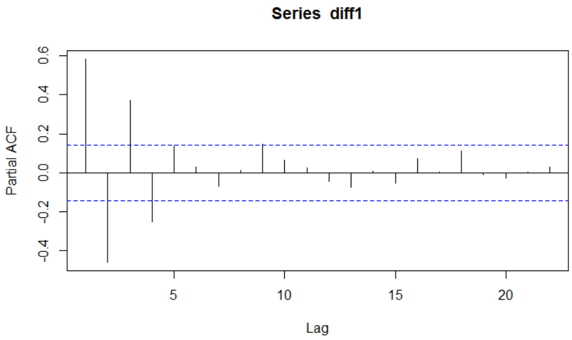


Figure 4

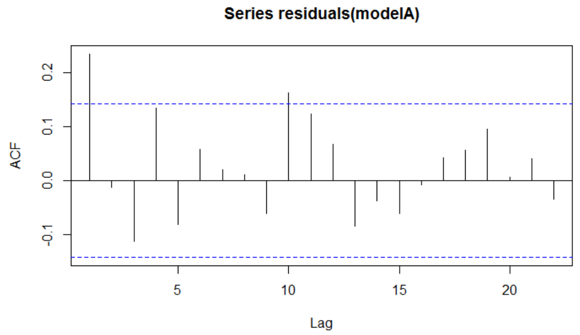


Model A

```
arima(x = Prices, order = c(0, 1, 1))

       Coefficients:
                ma1
             0.8417
       s.e.  0.0327

       sigma^2 estimated as 0.9808
       log likelihood = -266.96,   aic = 535.93
```

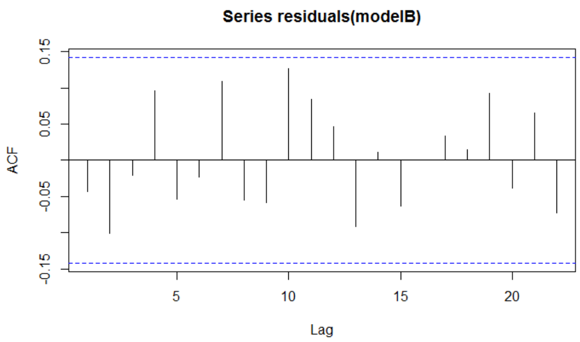


Model B

```
arima(x = Prices, order = c(0, 1, 2))

       Coefficients:
                ma1     ma2
             1.2045  0.4266
       s.e.  0.0678  0.0705

       sigma^2 estimated as 0.8577
       log likelihood = -254.5,   aic = 513
```

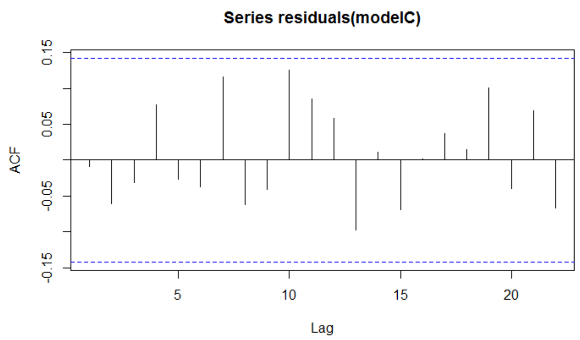


Model C

```
arima(x = Prices, order = c(1, 1, 1))

       Coefficients:
                ar1     ma1
             0.3047  0.7674
       s.e.  0.0760  0.0432

       sigma^2 estimated as 0.9077
       log likelihood = -259.74,   aic = 523.47
```



8. The updating equations of the Holt-Winters forecasting procedure for a time series y_t with multiplicative seasonal variation of period p are given by

$$\ell_t = \alpha\left(\frac{y_t}{s_{t-p}}\right) + (1-\alpha)(\ell_{t-1} + b_{t-1})$$

$$b_t = \gamma(\ell_t - \ell_{t-1}) + (1-\gamma)b_{t-1}$$

$$s_t = \delta\left(\frac{y_t}{\ell_t}\right) + (1-\delta)s_{t-p}$$

where α, γ and δ are the smoothing constants.

(i) Explain what properties of the series are described by the three quantities ℓ_t, b_t and s_t. Why might Holt-Winters be preferable to a forecasting method based on fixed parameter regression? (4)

(ii) Write down expressions for

(a) the forecast $\hat{y}_t(h)$ at time t for lead time h, for $1 \le h \le p$,

(b) the residual error at time t. (2)

The daily sales of a new campus coffee shop were recorded over five weeks and analysed using the Holt-Winters method with multiplicative seasonal variation of period 7 and smoothing constants α = 0.25, γ = 0.15 and δ = 0.2. The table below shows the sales and other quantities that have been calculated daily for Week 5.

Day	*Sales (£)*	ℓ_t	b_t	s_t
Monday	972	720.15	5.96	1.18
Tuesday	738	725.05	5.80	1.02
Wednesday	828	760.69	10.27	1.00
Thursday	846	789.34	13.03	1.02
Friday	1080	848.07	19.89	1.13
Saturday	576	818.14	12.41	0.83
Sunday	486	790.44	6.40	0.70

(iii) Showing your working, calculate forecasts of sales for

(a) Monday in Week 6,

(b) Friday in Week 7. (4)

(iv) The sales on the Monday of Week 6 turned out to be £891. Given this fact, calculate the values of ℓ_t, b_t and s_t for this day. What now is your forecast for the Friday of Week 7? (8)

(v) The mean absolute percentage error is an accuracy measure defined as

$$\text{MAPE} = \frac{100}{n}\sum_{t=1}^{n}\frac{|e_t|}{y_t}, \quad \text{where } e_t = y_t - \hat{y}_{t-1}(1).$$

Why would MAPE be of particular relevance when the multiplicative form of Holt-Winters forecasting is appropriate? (2)

BLANK PAGE