EXAMINATIONS OF THE HONG KONG STATISTICAL SOCIETY

GRADUATE DIPLOMA, 2016

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* ${}^nC_r$.

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. In a Markov chain model for the progression of a disease, X_n denotes the level of severity in year n, for $n = 0, 1, 2, \ldots$. The state space is $\{1, 2, 3, 4\}$ with the following interpretations: in state 1 the symptoms are under control, states 2 and 3 represent respectively moderate and severe symptoms while state 4 represents a permanent disability.

The transition matrix is $\mathbf{P} = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & 0 & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

(i) Classify the four states as transient or recurrent giving reasons. What does this tell you about the long-run fate of someone with this disease? (3)

(ii) Calculate the 2-step transition matrix. (3)

(iii) State the probability that

(a) a patient whose symptoms are moderate will be permanently disabled one year later, (1)

(b) a patient whose symptoms are under control will have severe symptoms two years later. (1)

(iv) Calculate the probability that a patient whose symptoms are moderate will have severe symptoms four years later. (2)

A new treatment becomes available but only to permanently disabled patients, all of whom receive the treatment. This has a 50% success rate in which case a patient returns to the "symptoms under control" state and is subject to the same transition probabilities as before. A patient whose treatment is unsuccessful remains in state 4 receiving a further round of treatment the following year.

(v) Write out the transition matrix for this new Markov chain and classify the states as transient or recurrent. (2)

(vi) Calculate the stationary distribution of the new chain. The population suffering from the disease currently has no members receiving the new treatment. What proportion of the population will be receiving the new treatment after a long passage of time? (6)

(vii) The annual cost of health care for each patient is zero in state 1, £c in state 2, £$2c$ in state 3 and £$8c$ in state 4, where $c > 0$ is a constant. Calculate the expected annual cost per patient when the system is in steady state. (2)

2. Let X_n for n = 0, 1, 2, … denote the population size in the nth generation of a branching process whose initial population size is 1, i.e. $X_0 = 1$. Each individual in a generation produces a number of offspring, Z, forming the next generation, where $p_i = P(Z = i)$, for i = 0, 1, 2, …, with associated probability generating function (pgf) $G(s)$. The numbers of offspring produced by different individuals are statistically independent of each other.

Let $G_n(z)$ denote the pgf of the total number of individuals in the population in the nth generation for n = 0, 1, 2, … .

(i) Explain why $G_1(s)$, the pgf of X_1, is $G(s)$.

(1)

(ii) Let X_{ni} denote the number of individuals in the nth generation who are descended from the ith member of the first generation. Explain why the pgf of X_{ni} is $G_{n-1}(s)$.

(2)

(iii) If there are $X_1 = x$ individuals in the first generation, prove that

$$E\left[s^{X_n} \mid X_1 = x\right] = \{G_{n-1}(s)\}^x$$

and hence that

$$G_n(s) = G\left(G_{n-1}(s)\right).$$

(7)

(iv) Let $\pi_n = P(X_n = 0)$, for n = 1, 2, 3, …, be the probability that the population has become extinct by the nth generation. Show that π_n is increasing with n. Use the equation for $G_n(s)$ in part (iii) to show that $\pi_n = G(\pi_{n-1})$.

(3)

(v) Let $\pi = \lim_{n\to\infty} \pi_n$ be the probability of ultimate extinction of the population. Deduce that π satisfies the equation $\pi = G(\pi)$. Explain how you know that $\pi = 1$ is a root of this equation.

(2)

(vi) Suppose now that Z has the following distribution.

Number of offspring, i	0	1	2	3
Probability, p_i	0.1	0.2	0.3	0.4

Calculate the probability of ultimate extinction of the population.

(5)

[You may assume that π is given by the smallest non-negative root of the equation $\pi = G(\pi)$.]

3. (i) For a general Markovian queuing system, the equilibrium probability π_n that there are n customers in the system is related to π_0 (the probability that there are no customers in the system) by the formula

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

Write down the definitions of the α_i and β_i as instantaneous transition rates and describe how π_0 may be calculated.

(3)

(ii) A take-away food counter has one server. Customers arrive randomly at a rate of λ per hour. If there is a queue, some customers go elsewhere so that the probability of a potential customer staying for service is $\frac{1}{n+1}$ when there are n customers in the shop already waiting or being served.

Service times are independent but the server is new to the job and tends to make mistakes under pressure, so the rate of service drops to $\frac{\mu}{n+1}$ per hour when there are n customers present, where $\mu > \lambda$.

(a) Show that the server will be busy for a proportion $\rho(2-\rho)$ of the time where $\rho = \frac{\lambda}{\mu}$.

(6)

[You may use the result that $\sum_{j=1}^{\infty} jx^{j-1} = \frac{1}{(1-x)^2}$, for $|x| < 1$.]

(b) Write down an expression for π_n (n = 0, 1, 2, …), the equilibrium probability that there are n customers in the system. Show that the probability generating function of the number of customers in the system in the steady state is

$$P(z) = \frac{(1-\rho)^2}{(1-\rho z)^2}$$

and hence find the mean and variance of the number of customers in the system in the steady state.

(8)

(c) Show that on average $\lambda\rho$ customers per hour go elsewhere.

(3)

4. In a model for the creation of particles at a time-varying rate, let $N(t)$ denote the number of particles created in $[0, t]$. The probability of a particle being created in the time interval $(t, t+\delta t]$ is $\lambda(t)\delta t + o(\delta t)$, regardless of the number of particles previously created and their creation times. The probability of more than one particle being created in $(t, t+\delta t]$ is $o(\delta t)$.

(i) Write down an expression for the probability that no particles are created in $(t, t+\delta t]$. Show that, for $n \geq 1$, the probability $p_n(t)$ of n particles being created in $[0, t]$ satisfies the differential equation

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

Derive the differential equation for $p_0(t)$.

(6)

(ii) The corresponding probability generating function is defined as

$$G(s,t) = \sum_{n=0}^{\infty} p_n(t)s^n .$$

Using the results in part (i), show that $G(s,t)$ satisfies the partial differential equation

$$\frac{\partial G}{\partial t} = \lambda(t)(s-1)G(s,t),$$

with boundary condition $G(s,0) = 1$.

(5)

(iii) Show that

$$G(s,t) = e^{-\Lambda(t)(s-1)},$$

where $\Lambda(t) = \int_0^t \lambda(u)du$.

(4)

(iv) Deduce the probability mass function $p_n(t)$ of $N(t)$ and identify its distribution.

(5)

5. (a) A forecaster uses simple exponential smoothing to obtain forecasts of a time series whose observed values are $y_1, y_2, \ldots, y_t, \ldots$.

(i) Let L_t denote the smoothed value (the level) of the series at time t and let α be the smoothing constant. Write down

(A) the updating equation for L_t in terms of L_{t-1}, y_t and α,

(B) $\hat{y}_t(h)$, the corresponding forecast at time t for lead time h, for $h \geq 1$.

(2)

(ii) Denoting by e_t the one-step-ahead forecast error, $e_t = y_t - \hat{y}_{t-1}(1)$, for $t \geq 2$, rewrite the updating equation for L_t in terms of L_{t-1}, e_t and α.

(2)

(iii) Assume that the value of L_1 is taken to be equal to y_1. By applying the updating equation iteratively, find an explicit expression for L_t in terms of $y_1, y_2, \ldots, y_t$ and α, in as simple terms as you can.

(4)

(iv) The forecaster wishes to predict the weekly sales of a local newspaper in thousands of copies. The sales for the first four months, i.e. t = 1, 2, 3 and 4, are 13, 15, 12 and 17 respectively. Using the initial value y_1 of the series as the initial smoothed value L_1, and taking the smoothing constant to be 0.3, calculate (to 2 decimal places) the smoothed value and forecast error for each of t = 2, 3, 4.

(4)

(b) The following model defines X_t in terms of W_t and two independent white noise series E_t and A_t, having respective variances σ_E^2 and σ_A^2.

$$W_t = W_{t-1} + E_t$$
$$X_t = W_t + A_t$$

(i) Express the first difference, $Y_t = X_t - X_{t-1}$, in terms of white noise terms only. Deduce that Y_t has autocorrelation function

$$\rho_Y(k) = \begin{cases} \dfrac{-\sigma_A^2}{\sigma_E^2 + 2\sigma_A^2}, & \text{for } k = 1, \\ 0, & \text{for } k \geq 2. \end{cases}$$

(6)

(ii) If the forecaster were to model X_t as a member of the ARIMA(p, d, q) family, what would be the values of p, d and q?

(2)

6. A time series X_t satisfies the model

$$X_t = 10 + A_t + 0.4A_{t-1} - 0.45A_{t-2},$$

where A_t is white noise with variance σ^2.

(i) This model can be described as a member of the ARMA(p, q) family. State the values of p and q and verify that X_t is

(a) stationary,

(b) invertible.

(5)

(ii) Calculate the mean and variance of X_t.

(4)

(iii) If $\hat{X}_t(h)$ denotes the minimum mean square error forecast of X_{t+h} at time t, explain why $\hat{X}_t(h) = 10$, for $h \geq 3$, and express both $\hat{X}_t(1)$ and $\hat{X}_t(2)$ in terms of forecast errors a_t and a_{t-1}. Explain how a_t and a_{t-1} would be calculated.

(6)

(iv) Calculate the h-step-ahead forecast error variance for each of $h = 1, 2$ and $h \geq 3$. Find also a 90% prediction interval for X_{t+h} for $h \geq 3$ and show that it has constant width.

(5)

7. (i) A linear model for a time series X_t can be described using the model

$$\phi(B)X_t = \theta(B)A_t$$

where $\phi(B)$ and $\theta(B)$ are characteristic polynomials in the backshift, or lag, operator B, and where A_t is white noise. Explain how you would decide whether X_t is

(a) stationary,

(b) invertible.

(2)

(ii) The time series X_t satisfies

$$X_t = 1.8X_{t-1} - 0.8X_{t-2} + A_t + 0.6A_{t-1}.$$

Investigate

(a) stationarity,

(b) invertibility,

of X_t. Classify X_t as a member of the ARIMA(p, d, q) family, i.e. identify p, d, and q.

(5)

(iii) The estimation method known as 'conditional least squares' requires the calculation of the errors a_t for t = 1, 2, ..., n from the time series data x_t as a function of unknown parameters $\phi_1, ..., \phi_p$ and $\theta_1, ..., \theta_q$. Illustrate this using the model given in part (ii), by writing out the difference equation for calculating a_t with its initial values. Explain briefly how these are used in the subsequent calculations to find the conditional least squares estimates. Why is this method known as *conditional* least squares?

(5)

(iv) A linear time series model whose characteristic polynomials have a common root is said to be redundant since it is equivalent to a simpler model. Show that the model

$$Z_t = 2.8Z_{t-1} - 2.6Z_{t-2} + 0.8Z_{t-3} + A_t - 0.4A_{t-1} - 0.6A_{t-2}$$

is redundant, simplify it and correctly classify it as a member of the ARIMA(p, d, q) family.

(4)

(v) Generalise this to a general ARIMA(p, d, q) model with a common root ω. In what way, if at all, will

(a) the errors,

(b) the conditional sum of squares,

depend on ω? Justify your answer. What problem would trying to fit a redundant model, such as that in part (iv), pose at the model fitting stage?

(4)

8. A scientist studying plant nutrition is considering a time series x_t consisting of 187 successive observations of daily nitrogen intake in a variety of wheat. Figure 1 **on the next page** shows a plot of this series. Figures 2 and 3 **on the next page** are plots of the sample autocorrelation function (acf) and partial autocorrelation function (pacf) for the series.

(i) In Figures 2 and 3, acf or pacf values outside the horizontal lines are significantly different from zero (at the 5% level). Comment on the following ARIMA models as candidates for x_t.

(a) White noise

(b) MA(2)

(c) AR(2)

(6)

(ii) The edited computer output **on the second page following** shows some results obtained by fitting three ARIMA models to the nitrogen intake data. The software used fits an ARIMA(p, d, q) model to x_t by first calculating y_t by taking dth differences of x_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 have been fitted and write down explicitly their model equations.

(4)

(iii) Treating each of the three models in turn, consider briefly whether their parameter estimates are statistically significant. What do you conclude about the suitability of these models?

(5)

(iv) What do you learn from the values of

(a) the log likelihood,

(b) the AIC,

about the suitability of each of the models?

(2)

(v) A colleague of the scientist claims that Figure 1 shows signs of non-stationarity and proposes taking first differences. A quick calculation shows that the standard deviation of x_t is 1.29 while $y_t = x_t - x_{t-1}$ has standard deviation 1.53. Discuss the colleague's proposal in the light of these standard deviations, Figure 2 and the model fitting results.

(3)

Diagrams and computer output are on the next two pages

Figure 1

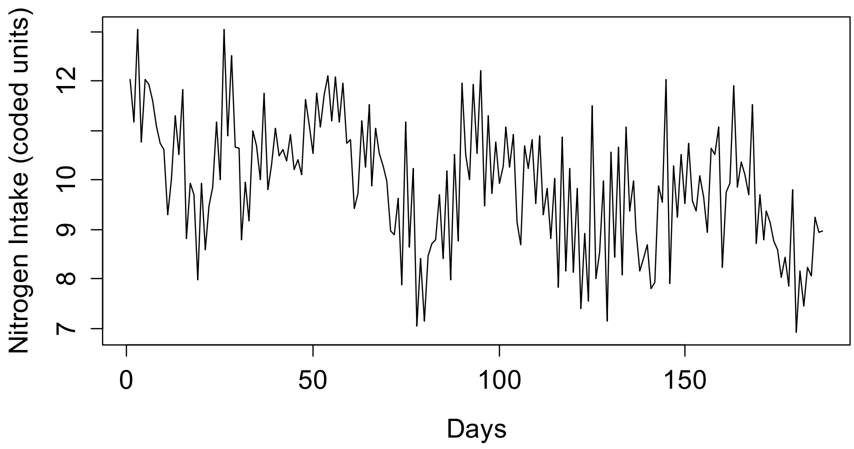


Figure 2

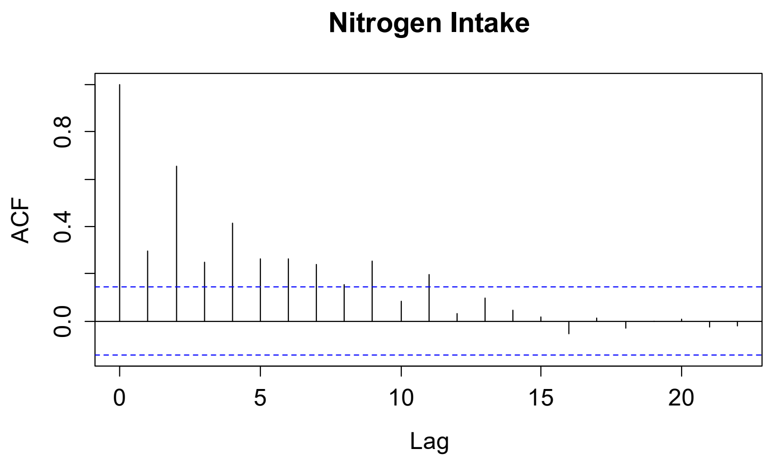
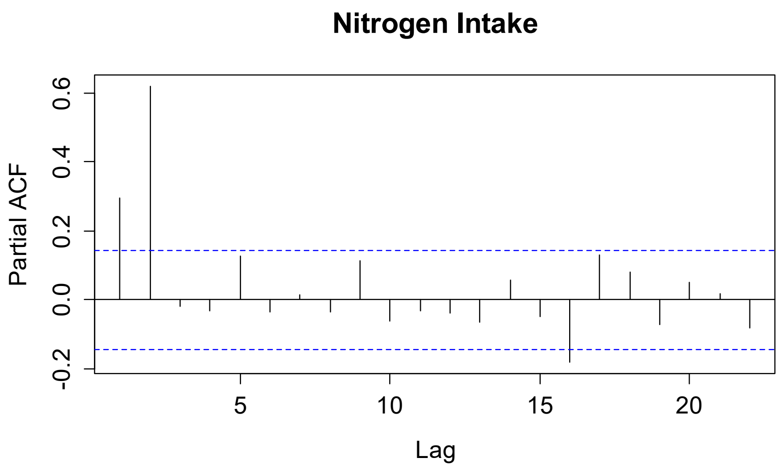


Figure 3



Model 1

```
Model:
arima(x = nitrogen intake, order = 2, 0, 0)

Coefficients:
           ar1      ar2   intercept
        0.1168   0.6282      9.9283
  s.e.  0.0565   0.0567      0.2667

sigma^2 estimated as 0.9148:   log likelihood = -257.57,   AIC = 523.15
```

Model 2

```
Model:
arima(x = nitrogen intake, order = 3, 0, 0)

Coefficients:
           ar1      ar2      ar3   intercept
        0.1145   0.6278   0.0037      9.9284
  s.e.  0.0727   0.0572   0.0739      0.2678

sigma^2 estimated as 0.9148:   log likelihood = -257.57,   AIC = 525.14
```

Model 3

```
Model:
arima(x = nitrogen intake, order = 2, 0, 1)

Coefficients:
           ar1      ar2      ma1   intercept
        0.1201   0.6272  -0.0055      9.9285
  s.e.  0.0902   0.0605   0.1149      0.2678

sigma^2 estimated as 0.9148:   log likelihood = -257.57,   AIC = 525.14
```

BLANK PAGE