ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS

EXAMINATIONS OF THE ROYAL STATISTICAL SOCIETY

HIGHER CERTIFICATE IN STATISTICS, 2017

MODULE 7 : Time series and index numbers

Time allowed: One and a half hours

Candidates should answer **THREE** *questions.*

Each question carries 20 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 8 printed pages.
This front cover is page 1.
Question 1 starts on page 2.

There are 4 questions altogether in the paper.

1. (a) Explain how moving averages can be used to extract the seasonal and trend components from a quarterly time series. What change in the procedure is required if the seasonal effect is multiplicative rather than additive? (5)

(b) The table below shows the quarterly sales figures x_t of a certain product for the years 2014 to 2016 in units of £1000.

Quarter		Sales (×£1000)	5-point MA	EWMA
2014	Q1	494		334.0
	Q2	313		325.6
	Q3	383		348.6
	Q4	482	401.4	401.9
2015	Q1	535	401.4	455.2
	Q2	294	443.4	390.7
	Q3	313	457.8	359.6
	Q4	593	407.4	453.0
2016	Q1	554	427.0	493.4
	Q2	283		
	Q3	392		
	Q4	605		

MA: Moving Average
EWMA: Exponentially Weighted Moving Average

(i) Write down the formula for a 5-point simple symmetric moving average of x_t.

Explain which other estimates for the 5-point MA can be calculated from the data in the table, and give their values correct to 1 decimal place.

What unfortunate property does this 5-point moving average have when extracting seasonal and trend components from quarterly data? Illustrate your answer with an example from the table. Describe a 2×4 moving average and explain why it might be preferred to both a 4-point and a 5-point simple moving average.

(7)

(ii) Write down the formula for the exponentially weighted moving average of x_t with smoothing parameter 0.4.

Setting your initial smoothed value equal to 334, the average sales figure for 2013, use this moving average to calculate, as far as the method allows, the missing estimates in the column headed EWMA. Give your answers correct to 1 decimal place.

Suppose now that the initial smoothed value is taken to be the actual value 494 for Q1 2014. Re-calculate the estimate for Q2 2014 and comment on the effect this will have on the estimated trend values in the short term and in the long term.

(6)

(c) Describe one advantage and one disadvantage of increasing the length of a moving average when estimating a trend. (2)

2. A stationary time series X_t follows an ARIMA(p, q) model described by the following equation.

$$X_t = \theta_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t + \theta_1 \varepsilon_{t-1}.$$

(i) State the values of p and q. (1)

(ii) What assumptions are usually made about the errors, ε_t? (3)

(iii) Show that for this model the expected value of X_t is given by

$$\mu_X = \frac{\theta_0}{1-\phi_1-\phi_2}.$$

Explain why the stationarity of X_t is needed in your argument. (3)

(iv) Write down the equation which can be used to calculate a 1-step ahead forecast from the fitted model. (3)

An analyst has fitted two models, A and B, to a time series x_t of 184 daily values of the concentration of a chemical in the waste from an industrial process. For each model, the tables below contain information about the estimated parameters and a Box-Ljung lack-of-fit test.

(v) For each model identify the orders of the ARMA model which has been fitted and write down its equation. (4)

(vi) Calculate the respective expected values of A and B and comment. (2)

(vii) What null hypothesis is being tested here by the Box-Ljung test? Describe, and briefly justify, the conclusions the analyst would draw from the two Box-Ljung tests. (4)

Coefficients of model A:

ar1	ar2	intercept
1.2244	-0.2939	12.4014

Box-Ljung test

```
data:  residuals from  model A
X-squared = 24.154, df = 10, p-value = 0.007203
```

Coefficients of model B:

ar1	ar2	ma1	intercept
0.6136	0.2933	0.8219	12.2508

Box-Ljung test

```
data:  residuals from  model B
X-squared = 7.992, df = 9, p-value = 0.535
```

3. In the construction of certain indices using a basket of commodities, 0 denotes the base period, t is the current period, p represents price, q represents quantity and i identifies the commodities contributing to the index.

(i) Prove that the ratio of the Laspeyres volume index to the Laspeyres price index is equal to the ratio of the Paasche volume index to the Paasche price index, when these indices are calculated from the same data.

(6)

(ii) Starting from the Laspeyres volume index formula expressed as a function of prices and quantities, show that this index can be expressed as a weighted arithmetic mean of volume relatives and state what the weights are.

(5)

(iii) Show that the Paasche volume index can be expressed as a weighted harmonic mean of volume relatives and state what the weights are.

(5)

(iv) The formula for the Walsh price index is

$$P_W(0, t) = \frac{\sum_i p_{ti}\sqrt{q_{0i}q_{ti}}}{\sum_i p_{0i}\sqrt{q_0 q_{ti}}}.$$

Show that this index can also be expressed as a weighted arithmetic mean of price relatives and state what the weights are.

(2)

(v) State one advantage and one disadvantage that the Walsh price index has when compared to the Laspeyres price index.

(2)

4. (a) The price per unit and quantity sold of commodity i of a group of commodities are respectively p_{0i} and q_{0i} at time 0 and p_{ti} and q_{ti} at time t.

(i) State the formula for the Paasche price index, $P_P(0, t)$.

What interpretation can you put on the numerator and denominator of $P_p(0, t)$?

(3)

(ii) State the formulae for the corresponding Laspeyres and Fisher price indices, $P_L(0, t)$ and $P_F(0, t)$. Explain why the Fisher index is sometimes described as 'ideal'. Are there any reasons why the Laspeyres price index might still be used in preference?

(4)

(iii) Suppose you have data available for three periods (0, 1 and 2) from which you have calculated Laspeyres and Paasche indices. Prove that the chain-linked Fisher index (referenced to period 0 and linked at period 1) is the geometric mean of the chain-linked Laspeyres index and the chain-linked Paasche index (both being referenced to period 0 and linked at period 1).

(6)

(b) Information about a retail company's sales have been grouped according to the type of outlet in the following table.

Type of Outlet	*2013 value (£ millions)*	*2014 value (£ millions)*	*2013 Laspeyres volume index (base period 2013)*	*2014 Laspeyres volume index (base period 2014)*
Neighbourhood stores	176	198	101.3	100.5
Out-of-town stores	298	295	98.2	100.1
Online	79	106	103.2	104.5

(i) Calculate the chain-linked Laspeyres volume index for 2015 (referenced to 2013 and linked at 2014). Give your answer correct to 1 decimal place.

(4)

(ii) If the Paasche volume indices for 2014 (based on 2013) and 2015 (based on 2014) are 100.3 and 102.2 respectively, calculate the chain-linked Fisher index for 2015 (referenced to 2013 and linked at 2014). Give your answer correct to 1 decimal place.

(3)

BLANK PAGE

BLANK PAGE

BLANK PAGE