EXAMINATIONS OF THE ROYAL STATISTICAL SOCIETY

HIGHER CERTIFICATE IN STATISTICS, 2016

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. (i) An additive trend and seasonal decomposition for a quarterly time series $y(t)$ may be represented as

$$y(t) = T(t) + C(t) + R(t),$$

where $T(t)$ and $R(t)$ are respectively the trend component and the random or irregular component. Explain why the constraint $C(t) = C(t - 4)$ should be applied to $C(t)$ and state the usual 'standardising' condition imposed on the sum $C(1) + C(2) + C(3) + C(4)$.

(3)

Analyst A applies an additive trend and seasonal decomposition to some sales data for a cake manufacturing company. The table below gives the data plus partial results of using a simple 3-point moving average to smooth the sales series along with the corresponding errors for each quarter, Q1 to Q4.

Period	Sales (£K) $y(t)$	Smoothed value, $y^*(t)$	Error
Q1 2012	241		
Q2 2012	283	268.33	14.67
Q3 2012	281	296.33	–15.33
Q4 2012	325	297.67	27.33
Q1 2013	287	296.00	–9.00
Q2 2013	276	288.00	–12.00
Q3 2013	301	309.67	–8.67
Q4 2013	352	325.00	27.00
Q1 2014	322	335.67	–13.67
Q2 2014	333		
Q3 2014	351		
Q4 2014	410		

(ii) Calculate the missing smoothed values and errors for Q2 and Q3 of 2014.

(3)

[Note: candidates should not attempt to calculate smoothed values for the 'end-points', Q1 2012 or Q4 2014.]

(iii) What weakness does the choice of a 3-point moving average have when smoothing quarterly data?

(1)

(iv) Calculate estimates of the seasonal components for Q1 to Q4.

(4)

Question 1 continued on the next page

(v) A linear regression model fitted to the smoothed values yields

$$y^*(t) = 254.99 + 8.72t,$$

where t is the number of quarters from Q1 2012 with $t = 1$ for Q1 2012. Use the additive trend and seasonal model from part (i) to find A's forecasts of sales for Q3 and Q4 of 2015.

(3)

(vi) Analyst B fits a linear regression model of $y(t)$ on t and an indicator variable for Q4, obtaining the following output.

```
Coefficients:
                  Estimate     Std. Error
(Intercept)       242.282        8.006
     t              9.157        1.122
     Q4            46.798        8.947
```

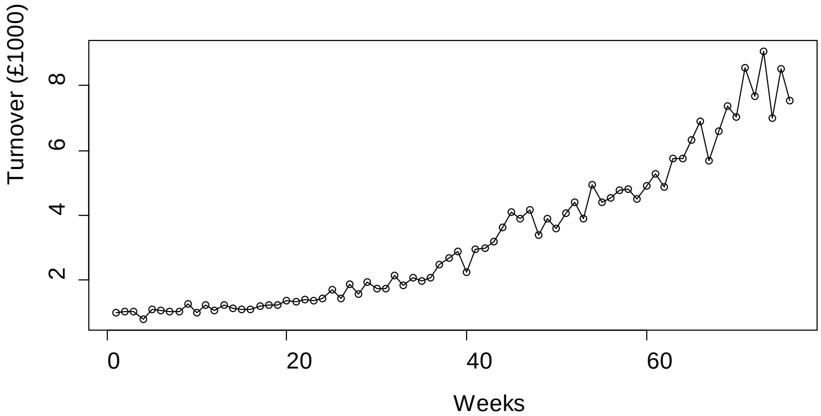
Discuss whether the explanatory variables, t and Q4, should be retained in the model.

(4)

(vii) Calculate forecasts of sales for Q3 and Q4 of 2015 from B's model and compare them with those from A's model as calculated in part (v).

(2)

2. A time series of the weekly total turnover in thousands of pounds, w_t, of a small retail start-u



(i) Describe the series in terms of its trend, seasonality and any other noteworthy features.

(3)

(ii) In analysing this series a forecaster calculates logarithms (base e) of the data, $x_t = \log(w_t)$, and then takes first differences, $y_t = x_t - x_{t-1}$. Suggest reasons why these choices were made.

(2)

The forecaster then fits an ARIMA(p, d, q) model to y_t and includes a constant term. The computer output below gives the results from the model-fitting software.

```
Coefficients:
           ar1       ar2    intercept
       -0.7983   -0.2276     0.0278
s.e.    0.1121    0.1126     0.0057
```

(iii) State the orders p, d and q of the fitted model and write out the equation for y_t in terms of past y_{t-i} and an error series, ε_t, whose properties you should define.

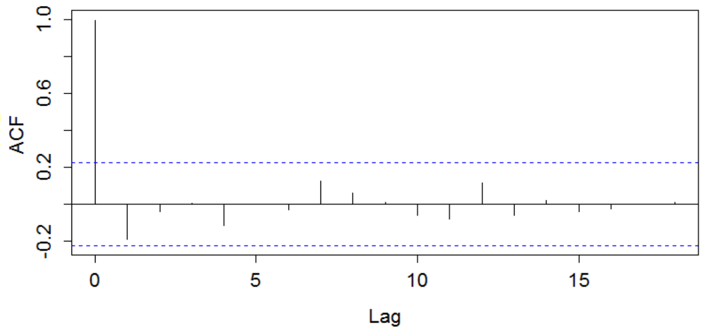
(6)

(iv) State the corresponding orders p, d and q of the ARIMA model for x_t. What property of the series x_t would have motivated the forecaster to include a constant term in the model for y_t?

(2)

Question 2 continued on the next page

(v) The forecaster has also plotted the residual autocorrelations from the fitted model.



Explain why this would have been done and what we learn from this plot.

(3)

(vi) Use the equation you wrote down in part (iii) to find an expression for a 1-step-ahead forecast of y_{77}. Describe briefly how you would then deduce a forecast of the original untransformed series.

(4)

3. 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 Laspeyres volume index, $Q_L(0,t)$. What interpretation can you put on the numerator and denominator of $Q_L(0,t)$?

(3)

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

(1)

(iii) Show that

$$Q_L(0,t) = \frac{\sum_i p_{ti}q_{ti}}{\sum_i p_{0i}q_{0i}} \Bigg/ P_P(0,t).$$

(4)

(iv) What term is used to describe $P_P(0,t)$ when used in this way to find $Q_L(0,t)$?

(1)

(v) The table below contains data on the production of a manufacturer of fish products for the years 2008 and 2014. Using the separate Paasche price index for each product group, calculate a Laspeyres volume index of the total production over all three groups of products in 2014 using 2008 as the base period.

(6)

Product group	*2008 sales (£ thousand)*	*2014 sales (£ thousand)*	*P_P(2008, 2014) for product group*
Frozen fillets	245	269	98.6
Dressed fish	173	166	97.3
Fish fingers	111	105	101.2

(vi) The company also knows that its sales for the category *Other seafood* were worth £33 000 in 2008 but the corresponding figure for 2014 is unavailable. However, sales by weight were 2750 kg in 2008 and 3130 kg in 2014. Calculate a Laspeyres volume index for the combined sales of fish products and other seafood.

(5)

4. (a) (i) 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. The Laspeyres price index for this group of commodities at time t based on time 0 is defined as

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

Show that $P_L(0,t)$ may be written as a weighted average of price relatives, $R_{0ti} = \frac{p_{ti}}{p_{0i}}$. What interpretation do the weights have?

(4)

(ii) The commodities sold by a company may be divided into two mutually exclusive groups A and B. Show that the overall Laspeyres price index for the company can be written as a weighted average of their respective price indices $P_{L,A}(0,t)$ and $P_{L,B}(0,t)$, stating the interpretation of the weights.

(5)

(b) Suppose a country's economic output can be classified into three sectors according to the following table.

Economic sector	*2012 value (€ billions)*	*2013 value (€ billions)*	*2013 Laspeyres volume index (base period 2012)*	*2014 Laspeyres volume index (base period 2013)*
Agriculture	30	31	100.2	99.5
Manufacturing	50	48	98.3	98.5
Services	60	75	102.9	105.4

(i) Calculate the Laspeyres volume indices for the whole economy in 2013 using 2012 as the base period, and in 2014 using 2013 as the base period.

(6)

(ii) Link the two index numbers you have just calculated to give a chain-linked Laspeyres volume index for 2014, referenced to 2012.

(3)

(iii) State the main reason why chain-linking might be preferred to direct calculation of $Q_L(2012, 2014)$.

(2)

BLANK PAGE