THE ROYAL STATISTICAL SOCIETY

HIGHER CERTIFICATE EXAMINATION

NEW MODULAR SCHEME

introduced from the examinations in 2007

MODULE 7

SPECIMEN PAPER

AND SOLUTIONS

The time for the examination is 1½ hours. The paper contains four questions, of which candidates are to attempt **three**. Each question carries 20 marks. An indicative mark scheme is shown within the questions, by giving an outline of the marks available for each part-question. The pass mark for the paper as a whole is 50%.

The solutions should not be seen as "model answers". Rather, they have been written out in considerable detail and are intended as learning aids. For this reason, they do not carry mark schemes. Please note that in many cases there are valid alternative methods and that, in cases where discussion is called for, there may be other valid points that could be made.

While every care has been taken with the preparation of the questions and solutions, the Society will not be responsible for any errors or omissions.

The Society will not enter into any correspondence in respect of the questions or solutions.

Note. In accordance with the convention used in all the Society's examination papers, the notation log denotes logarithm to base *e*. Logarithms to any other base are explicitly identified, e.g. $\log_{10}$.

1. (i) For each of the following types of time series decomposition models, derive an expression for the difference between the trend and the seasonally adjusted series in terms of the trend and irregular components. Show all your working.

(*A*) Additive decomposition.

(4)

(*B*) Multiplicative decomposition.

(5)

(*C*) Pseudo-additive decomposition.

(5)

(ii) Interpret each of the derived quantities for (i)(*A*), (*B*) and (*C*) in words.

(3)

(iii) Describe how you would determine an appropriate choice of decomposition model for a given time series.

(3)

2. Consider the following events that have occurred for a particular time series. Which part of the model decomposition (trend, seasonal or irregular) should the following be assigned to in the model decomposition?

(i) Increase in sales of clothes shop due to an unusually cold winter.

(1)

(ii) Increase in sales of clothes shop in the run up to Christmas.

(1)

(iii) Fall in sales of clothes shop due to an economic recession.

(1)

(iv) Decrease in sales in April due to the timing of the Easter holiday.

(1)

(v) For each of the effects in the four cases (i) to (iv), what type of prior adjustment, temporary or permanent, would you use? Explain your answers.

(10)

(vi) Would you use the original, seasonally adjusted, trend or irregular estimates to estimate any impact of particular effects? What do you consider the most appropriate way of carrying out an analysis? Explain your answer.

(6)

3. (i) Suppose a population of commodities can be divided into a set of exhaustive and non-overlapping domains. Prove that the population Laspeyres price index (for time t with base period 0) can be expressed in terms of the domain Laspeyres price indices and domain values (or turnovers).

(7)

(ii) Prove that the population Paasche price index (for time t with base period 0) can be expressed in terms of the domain Paasche price indices and domain values.

(7)

(iii) State whether the following price indices can or cannot be expressed in terms of their domain price indices and domain values.

(*A*) Fisher price index.

(*B*) Geometric Laspeyres price index.

(*C*) Tornqvist price index.

(*D*) Walsh price index.

(2)

(iv) Using the domain data from the table below, calculate the Laspeyres price index for the whole population.

Domain	*Domain Laspeyres price index*	*Domain base period value*
Food	1.06	\$50
Clothing	0.97	\$20
Fuel	1.40	\$30
Accommodation	1.20	\$60

(4)

4. (i) State the main reason for chain-linking a Laspeyres index.

(2)

(ii) Define the term *reference period*.

(1)

(iii) Use the data in the table below to answer the following.

(a) Calculate the Laspeyres price index for period 2 using period 0 as the base period.

(4)

(b) Calculate a chain-linked Laspeyres price index, by introducing period 1 as a new base period and linking it on to the index series from period 1.

(5)

(c) State the value of the chain-linked index at period 0 when the index is referenced to period 1.

(2)

(d) Calculate the Geometric Laspeyres price index for period 2 using period 0 as the base period.

(4)

(iv) Explain under what circumstances Geometric Laspeyres price indices can give misleadingly low numbers.

(2)

Commodity	**Period 0**		**Period 1**		**Period 2**	
	Price	*Quantity*	*Price*	*Quantity*	*Price*	*Quantity*
White bread	20p	10	25p	8	30p	2
Brown bread	30p	4	32p	5	45p	7
Rye bread	60p	1	55p	2	40p	5

SOLUTIONS

Question 1

Part (i)

(A) For the additive decomposition, we have (in obvious notation) $O_t = T_t + S_t + I_t$.

$$\therefore\ SA_t - T_t \ = \ \left\{O - S_t\right\} - T_t \ = \ \left\{T_t + S_t + I_t - S_t\right\} - T_t \ = \ T_t + I_t - T_t \ = \ I_t \ .$$

(B) For the multiplicative decomposition, $O_t = T_t \times S_t \times I_t$.

$$\therefore\ SA_t - T_t \ = \ \frac{O_t}{S_t} - T_t \ = \ \frac{T_t \times S_t \times I_t}{S_t} - T_t \ = \ \left(T_t \times I_t\right) - T_t \ = \ T_t\left(I_t - 1\right).$$

(C) For the pseudo-additive case,

$$O_t = T_t + \left\{T_t \times (S_t - 1)\right\} + \left\{T_t \times (I_t - 1)\right\} = T_t (S_t + I_t - 1) \ .$$

$$\therefore\ SA_t - T_t \ = \ \left\{O_t - T_t\left(S_t - 1\right)\right\} - T_t \ = \ \left\{T_t \times I_t\right\} - T_t \ = \ T_t\left(I_t - 1\right).$$

Part (ii)

We can interpret the differences between the trend and the seasonally adjusted series as follows.

In the additive case, the difference is the Irregular component.

In the multiplicative and pseudo-additive cases, this difference is not proportional to the Irregular component, though in periods of stable (flat) trend this may be a reasonable approximation. Note also that the greater the level of the trend (T) the greater the difference between it and the product Trend × Irregular ($T \times I$).

Part (iii)

An appropriate decomposition model can be found by fitting different ARIMA models to the time series, under an additive and multiplicative framework. The standard information criteria can then be used to determine the model of best fit. For example, we could use the Akaike Information Criterion or the Bayesian Information Criterion. A low value would indicate a preferred model.

Alternatively, a graphical interpretation of how the seasonal factors are changing over time may give an indication of an appropriate model decomposition. For example, if the Seasonal × Irregular factors are changing rapidly and the seasonal factor path cannot cope with the movement in the seasonal factors, this may indicate a poor model decomposition choice. In practice, we would expect that for a multiplicative decomposition the amplitude is directly related to the level of the time series.

Question 2

(i) Irregular.

(ii) Seasonal.

(iii) Trend.

(iv) Seasonal.

(v) For case (i), a temporary prior correction as it is a one-off event.

For case (ii), there are two possible answers. Seasonal adjustment processes are reasonably robust, so should be able to cope with gradual changes in the seasonal factors over time; depending on the significance of the impact, this is an example of moving seasonality. Alternatively, a temporary prior correction could be applied.

For case (iii), a temporary prior correction to improve the estimation of the seasonal factor.

For case (iv), a permanent prior correction as this is a systematic calendar-related effect and it should be removed from the final seasonally adjusted estimates.

(vi) The most appropriate approach is to use a regression-ARIMA methodology, or a variant of this, on the original estimates before seasonal adjustment. This way, any estimated impacts for particular effects have not been distorted or impacted on by the seasonal adjustment process. For example, if the irregular component is used to estimate effects then some of the potential impacts may already have been captured, or distorted, as part of the seasonal adjustment process.

Question 3

(i) For domain d the Laspeyres price index is

$$P_{L;d}(0,t) = \frac{\sum_{i \in d} R_{0ti} v_{0i}}{v_{0d}}$$

where $R_{0ti} \left[= \frac{p_{ti}}{p_{0i}} \right]$ is the price relative for commodity i,

$v_{0i} \left[= p_{0i} q_{0i} \right]$ is the base period value for commodity i,

$v_{0d} = \sum_{i \in d} v_{0i}$ is the total base period value for domain d.

For the population U the Laspeyres price index is

$$P_L(0,t) = \frac{\sum_{i \in U} R_{0ti} v_{0i}}{\sum_{i \in U} v_{0i}} = \frac{\sum_{d \in U} v_{0d} \sum_{i \in d} R_{0ti} \frac{v_{0i}}{v_{0d}}}{\sum_{d \in U} \sum_{i \in d} v_{0i}} = \frac{\sum_{d \in U} v_{0d} P_{L;d}(0,t)}{\sum_{d \in U} v_{0d}}$$

(ii) For domain d the Paasche price index is

$$P_{P;d}(0,t) = \frac{v_{td}}{\sum_{i \in d} \frac{v_{ti}}{R_{0ti}}}$$

where v_{ti} is the current period value for commodity i,

$v_{td} = \sum_{i \in d} v_{ti}$ is the total current period value for domain d.

For the population U the Paasche price index is

$$P_P(0,t) = \frac{\sum_{i \in U} v_{ti}}{\sum_{i \in U} \frac{v_{ti}}{R_{0ti}}} = \frac{\sum_{d \in U} \sum_{i \in d} v_{ti}}{\sum_{d \in U} \frac{v_{td}}{v_{td}} \sum_{i \in d} \frac{v_{ti}}{R_{0ti}}} = \frac{\sum_{d \in U} v_{td}}{\sum_{d \in U} \frac{v_{td}}{P_{P;d}(0,t)}}$$

(iii) (A) No. (B) Yes. (C) No. (D) No.

(iv) $$\frac{(50 \times 1.06) + (20 \times 0.97) + (30 \times 1.40) + (60 \times 1.20)}{50 + 20 + 30 + 60} = \frac{186.4}{160} = 1.165$$

<u>Question 4</u>

(i) Chain-linking introduces up-to-date quantity data (or value data or value shares), making a Laspeyres index more reflective of a changing economy.

(ii) The reference period is the period for which the index is forced to take a specific value (usually 1 or 100).

Part (iii)

(a) $$P_L(0,2)=\frac{(30\times10)+(45\times4)+(40\times1)}{(20\times10)+(30\times4)+(60\times1)}=\frac{520}{380}=1.368$$

(b) $$P_L(0,1)=\frac{(25\times10)+(32\times4)+(55\times1)}{(20\times10)+(30\times4)+(60\times1)}=\frac{433}{380}=1.139$$

$$P_L(1,2)=\frac{(30\times8)+(45\times5)+(40\times2)}{(25\times8)+(32\times5)+(55\times2)}=\frac{545}{470}=1.160$$

$$\text{Chain-linked}=P_L(0,1)\times P_L(1,2)=1.139\times1.160=1.321$$

(c) $$\text{Re-referenced value}=\frac{1}{P_L(0,1)}=\frac{1}{1.139}=0.878$$

(d) $$P_{GL}(0,2)=\left(\frac{30}{20}\right)^{\frac{10}{15}}\left(\frac{45}{30}\right)^{\frac{4}{15}}\left(\frac{40}{60}\right)^{\frac{1}{15}}=1.310\times1.114\times0.973=1.42$$

(iv) A commodity with a small price relative, unless it has a very small base weight, will dominate the calculation.