PMTS - Assignment C/SMTS - Assignment 2

Deadline: Monday, 28\textsuperscript{th} April, 2014

Total marks: [25]. Marks are shown in boxes [ ]. There are 2 questions in this assignment.

1. Consider the ARIMA(0, 1, 2) model specified by the equation

\[ Y_t = Y_{t-1} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2}, \quad t \in \mathbb{Z}. \]

Suppose that observed process values \( y_1, y_2, \ldots, y_T \) have been obtained. In terms of some, or all, of the \( \{y_t\} \) and/or the \( \{\epsilon_t\} \), give expressions for:

(a) \( \hat{y}_T(1) \), the one step ahead forecast (there is no need to present this in its infinite autoregressive form), [2]
(b) the forecast error for \( \hat{y}_T(1) \), [2]
(c) the forecast error variance for \( \hat{y}_T(1) \). [1]
2. A businesswoman is trying to assess the performance of her hotels in a particular city. Throughout a 14 year period running from January 1995 to December 2008, the same set of hotels were in full operation and the average number of occupied rooms for each month was recorded.

A statistician, who is hired to analyze the data and to make recommendations, enters the data into S+ and carries out an initial exploratory analysis.

> occ.rts <- rts(occupancy, start = 1995, frequency = 12, units = "months")
> ts.plot(occ.rts, xlab = "", ylab = "", main = "monthly hotel occupancies from Jan. 1995 to Dec. 2008")
> occ.zero.rts <- log(occ.rts)
> occ.one.rts <- occ.rts^(1/4)

Both log and quartic root transformations, via the use of log() and (1/4) respectively, have been applied to the data.

(a) Describe the main patterns in the original data set. What do you suppose is the main objective in experimenting with the log and quartic root transformations? Do you anticipate that stationarity would be achieved after either one of the two transformations is applied to the data? Justify your answer. [4]

...continued
Further procedures were carried out in S+ which are presented below.

```r
> d12.occ.one.rts <- diff(occ.one.rts, lag = 12)
> acf(d12.occ.one.rts)
> acf(d12.occ.one.rts, type = "partial")
> d1d12.occ.one.rts <- diff(d12.occ.one.rts)
> acf(d1d12.occ.one.rts)
> acf(d1d12.occ.one.rts, type = "partial")
> library(MASS)
> occ.arima <- arima(occ.one.rts, order = c(5, 1, 5),
                  seasonal = list(order = c(1, 1, 1), period = 12))
```

![Graph of Series: d1d12.occ.one.rts](image)

...continued
(b) i) Specify all of the transformations that have been carried out in moving from the original series to the one that is involved in the above plots. [2]

ii) What is the final objective in carrying out the transformations and has it been achieved? [3]

iii) In the fitting of the ARIMA$(p, d, q) \times (P, D, Q)_s$ model, specify the values for $p, d, q, P, D, Q$ and $s$ that have been used. [3]

iv) Why do you suppose that the particular ARIMA$(p, d, q) \times (P, D, Q)_s$ model that was fitted had been identified as a candidate on the basis of the above plots? Has this model been applied to the original data or a transformed version? Justify your answer. [4]
To help with reservations and capacity planning, some forecasts are generated for a sequence of months beyond the end of December 2008.

```r
> occ.fore$pred
     1  2   3  4  5    6    7    8
2009: 5.359920 5.266575 5.270371 5.399333 5.386050 5.586014 5.798706 5.831726
2010: 5.400613 5.289506 5.300641 5.460555 5.437911 5.612417 5.843924 5.884442
     9 10 11 12
2009: 5.479201 5.479202 5.270598 5.449428
2010: 5.505306 5.505456 5.321794 5.488347

> start deltat frequency
2009 0.08333333 12
```

```r
> occ.fore$se
     1  2   3  4  5    6
2009: 0.02341118 0.02386111 0.02415319 0.02462358 0.02497149 0.02572200
2010: 0.02875636 0.02908746 0.02925700 0.02930547 0.02930650 0.02931096
     7 8 9 10 11 12
2009: 0.02572279 0.02572620 0.02578343 0.02593030 0.02632106 0.02637647
2010: 0.02932429 0.02934449 0.02952897 0.02964241 0.02977046 0.02990751

> start deltat frequency
2009 0.08333333 12
```

(c) Let \( \{Y_t\} \) denote the original series and \( \{Z_t\} \) denote the transformed series prior to any differencing.

i) Write down a 99% prediction interval for the value of \( Z_t \) in June 2010. [2]

ii) Write down a 99% prediction interval for the value of \( Y_t \) in June 2010. [1]

iii) Suppose that the maximum capacity across the hotels is 1150. What should the statistician advise the businesswoman about the likelihood that the hotel group will be able to meet demand in June 2010? [1]

Important Note:

- Please read the current version of the Mathematics & Statistics Coursework Policy. Copies can be obtained from the course website, or in hardcopy from the programme administrator.