Time Series – Tutorial 3

Dr. Kevin Burke

Question 1

The `hours` data contains monthly average hours worked per week in the U.S. manufacturing sector for July 1982 through June 1987.

a) Load in the data using `data(hours)`.

b) Plot the series with monthly labels and comment.

c) Plot the ACF and comment.

d) Carry out a classical decomposition using `decompose`, save the result as `hoursdecom` and plot.

e) Extract the seasonal component from `hoursdecom` and plot with monthly labels.

f) Extract the random component from `hoursdecom` and comment on its ACF. Note: you will have to add the argument `na.action=na.omit` to the ACF since `decompose` creates `NA`s in the series.

Question 2

The `wages` data contains monthly average hourly wages (USD) for workers in the U.S. textile industry for July 1981 - June 1987.

a) Plot the series with monthly labels and comment.

b) Fit two models: M1 - linear trend, M2 - quadratic trend. Compare the models in terms of adjusted R-squared.

c) Superimpose both fitted models on the time plot from part (a) using different colours.

d) Produce a timeplot of the residuals for both models. Do they look random?

e) Carry out a `runs` test on the residuals (this tests if a sequence of values are independent).

f) Plot the ACF for the residuals.

Question 3

The `beersales` data contains monthly U.S. beer sales in millions of barrels for the period July 1975 - December 1990.

a) Plot the series with monthly labels and comment.

b) Carry out a classical decomposition and plot.

c) Fit a seasonal effects model to this data. Interpret the regression output.

d) Produce a timeplot of the residuals from part (c).

e) Fit a seasonal effects model with quadratic overall trend. Interpret the regression output.

f) Produce a timeplot of the residuals from part (e).

g) Plot the beer sales series again but set `xlim=c(1975, 2010)` in the plot. Now use the predict function to plot the model from part (e) over this time period. A version of the following code should work for you:

```
tinew <- seq(1975, 2000, by=1/12)
newdata <- data.frame(cbind(ti=tinew, months=season(beersales)))
newdata$months <- factor(newdata$months,
                         labels=levels(season(beersales)) )
pred <- predict(model, newdata=newdata)
lines(x=tinew, y=pred, col=2, lty=2)
```

Question 4

The `winnebago` data contains monthly units sales of recreational vehicles from Winnebago Inc., from November 1966 through February 1972.

a) Plot the series with monthly labels and comment.

b) Apply `decompose` and comment. Extract the seasonal component and plot this separately with monthly labels.

c) Create a new series by applying the natural logarithm: `logwinnebago <- log(winnebago)`. Plot this series and comment (the high variance later in time should be reduced).

d) Fit a linear trend model with seasonal effects to the logged data data.

e) Produce a timeplot of the residuals and comment.

f) Plot the ACF of the residuals and comment. Also, carry out a `runs` test of independence.

g) Investigate normality of residuals.

h) Investigate the presence of non-constant variance by plotting predicted values against residuals.

Question 5

Consider the `oilfilters` data which contains monthly sales of John Deere oil filters.

a) Plot the series with monthly labels.

b) Plot the ACF for this series.

c) Carry out a classical decomposition using `decompose`, save the result as `oildecom` and plot. Extract the seasonal component from `oildecom` and plot with monthly labels.

d) Fit a seasonal effects model to the data. Superimpose this fitted model on the time plot from part (a).

e) For the above model, investigate the residuals in terms of normality, non-constant variance and autocorrelation.

f) Fit a harmonic trend model with one cos-sin pair using `harmonic(oilfilters,m=1)`. Superimpose this fitted model on the time plot from part (a).

g) Repeat part (f) but increase the order of the harmonic component (change `m`) to achieve a good fit. Decide on a model based on the adjusted R-squared.

Question 6

Let Y_t be a process with both linear and seasonal trend which can be written as:

$$Y_t = \beta_0 + \beta_1 t + s_t + e_t$$

where $s_t = s_{t-12}$, i.e., the s_t value repeats every 12 time units (assuming monthly seasonality here).

a) Consider the seasonal differenced series

$$W_t = \nabla_{12} Y_t = Y_t - Y_{t-12}.$$

Show that this series is stationary by calculating the mean and autocovariance functions. Also calculate the autocorrelation.

b) Consider how the situation would change if

$$Y_t = \beta_0 + \beta_1 t + s_t + X_t$$

where X_t is a zero-mean stationary process with autocovariance function γ_k^*.

c) Consider the series

$$Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + s_t + e_t$$

(i.e., *quadratic* and seasonal trend). Show that $W_t = \nabla_{12} Y_t = Y_t - Y_{t-12}$ is non-stationary.

d) Following on from part (c), show that $\nabla W_t = W_t - W_{t-1}$ is stationary. Note: do not work out the autocovariance function here - just use the fact that the sum of stationary series is stationary.

Question 7

In Lecture3 a seasonal effects model was fitted to the `tempdub` data. We will now consider differencing instead.

a) Plot the ACF for `tempdub`. Comment.

b) Create the following seasonal differenced series: `sdifftemp <- diff(tempdub,lag=12)`.

c) Plot the ACF for the differenced series.

d) Investigate normality of the differenced series.

Question 8

The `hours` data was considered in Question 1.

a) Difference this series. Plot the ACF.

b) Seasonal difference this series. Plot the ACF.

c) Now both difference and seasonal difference this series. Plot the ACF.

Question 9

In Question 3 we considered the `beersales` data. Part (g) showed the issue with assuming a model for the trend.

a) Seasonal and quadratic trends were apparent in this series. Thus, apply both a difference and seasonal difference to the series. Plot the ACF.

b) Investigate normality of the differenced series.