

Forecasting: principles and practice

Exercises: Set 1

21 October 2013

Before doing any exercises in R, load the **fpp** package using `library(fpp)`.

1. This question involves reading in data from a csv file and converting it to a ts object.
 - (a) Download the data from robjhyndman.com/data/tute1.csv. Open the file 'tute1.csv' in Excel and review its contents. You should find four columns of information. Columns B through D each contain a quarterly data series, labelled Sales, AdBudget and GDP. Sales contains the quarterly sales for a small company over the period 1981–2005. AdBudget is the advertising budget and GDP is the gross domestic product. All series have been adjusted for inflation.
 - (b) Import the file `tute1.csv` into R and convert the data to time series:

```
tute1 <- ts(import[,-1], start=1981, frequency=4)
```

Why did we remove the first column?
 - (c) Construct time series plots of each of the three series

```
plot(tute1)
```
 - (d) What features do you notice about each of the series AdBudget, Sales and GDP?
2. Use the Dow Jones index (data set `dowjones`) to do the following:
 - (a) Produce a time plot of the series.
 - (b) Produce forecasts using the drift method and plot them.
 - (c) Show that the graphed forecasts are identical to extending the line drawn between the first and last observations.
 - (d) Try some of the other benchmark functions to forecast the same data set. Which do you think is best? Why?
3. For each of the following series, make a graph of the data with forecasts using the most appropriate of the four benchmark methods: mean, naive, seasonal naive or drift.
 - (a) Annual bituminous coal production (1920–1968). Data set `bicoal`.
 - (b) Price of chicken (1924–1993). Data set `chicken`.
 - (c) Monthly total of people on unemployed benefits in Australia (January 1956–July 1992). Data set `dole`.
 - (d) Monthly total of accidental deaths in the United States (January 1973–December 1978). Data set `usdeaths`.
 - (e) Quarterly production of bricks (in millions of units) at Portland, Australia (March 1956–September 1994). Data set `bricksq`.
 - (f) Annual Canadian lynx trappings (1821–1934). Data set `lynx`.

In each case, do you think the forecasts are reasonable? If not, how could they be improved?