

3.1 Some simple forecasting methods

Some forecasting methods are extremely simple and surprisingly effective. We will use the following four forecasting methods as benchmarks throughout this book.

Average method

Here, the forecasts of all future values are equal to the average (or “mean”) of the historical data. If we let the historical data be denoted by y_1, \dots, y_T , then we can write the forecasts as

$$\hat{y}_{T+h|T} = \bar{y} = (y_1 + \dots + y_T)/T.$$

The notation $\hat{y}_{T+h|T}$ is a short-hand for the estimate of y_{T+h} based on the data y_1, \dots, y_T .

```
meanf(y, h)
# y contains the time series
# h is the forecast horizon
```

Naïve method

For naïve forecasts, we simply set all forecasts to be the value of the last observation. That is,

$$\hat{y}_{T+h|T} = y_T.$$

This method works remarkably well for many economic and financial time series.

```
naive(y, h)
rwf(y, h) # Equivalent alternative
```

Because a naïve forecast is optimal when data follow a random walk (see Section 8.1), these are also called **random walk forecasts**.

Seasonal naïve method

A similar method is useful for highly seasonal data. In this case, we set each forecast to be equal to the last observed value from the same season of the year (e.g., the same month of the previous year). Formally, the forecast for time $T + h$ is written as

$$\hat{y}_{T+h|T} = y_{T+h-m(k+1)},$$

where m = the seasonal period, and k is the integer part of $(h - 1)/m$ (i.e., the number of complete years in the forecast period prior to time $T + h$). This looks more complicated than it really is. For example, with monthly data, the forecast for all future February values is equal to the last observed February value. With quarterly data, the forecast of all future Q2 values is equal to the last observed Q2 value (where Q2 means the second quarter). Similar rules apply for other months and quarters, and for other seasonal periods.

```
snaive(y, h)
```

Drift method

A variation on the naïve method is to allow the forecasts to increase or decrease over time, where the amount of change over time (called the **drift**) is set to be the average change seen in the historical data. Thus the forecast for time $T + h$ is given by

$$\hat{y}_{T+h|T} = y_T + \frac{h}{T-1} \sum_{t=2}^T (y_t - y_{t-1}) = y_T + h \left(\frac{y_T - y_1}{T-1} \right).$$

This is equivalent to drawing a line between the first and last observations, and extrapolating it into the future.

```
rwf(y, h, drift=TRUE)
```

Examples

Figure 3.1 shows the first three methods applied to the quarterly beer production data.

```
# Set training data from 1992 to 2007
beer2 <- window(ausbeer, start=1992, end=c(2007, 4))
# Plot some forecasts
autoplot(beer2) +
  autolayer(meanf(beer2, h=11),
    series="Mean", PI=FALSE) +
  autolayer(naive(beer2, h=11),
    series="Naïve", PI=FALSE) +
  autolayer(snaive(beer2, h=11),
    series="Seasonal naïve", PI=FALSE) +
  ggtitle("Forecasts for quarterly beer production") +
  xlab("Year") + ylab("Megalitres") +
  guides(colour=guide_legend(title="Forecast"))
```

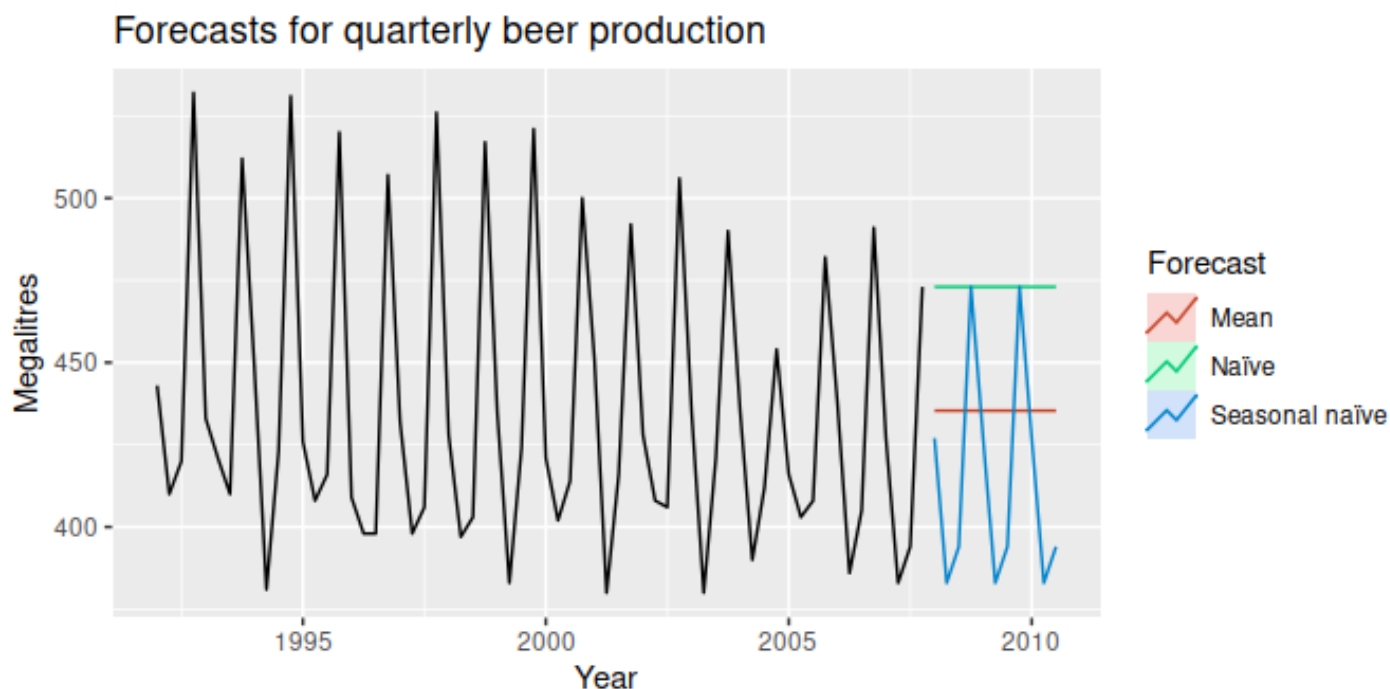


Figure 3.1: Forecasts of Australian quarterly beer production.

In Figure 3.2, the non-seasonal methods are applied to a series of 200 days of the Google daily closing stock price.

```
autoplot(goog200) +  
  autolayer(meanf(goog200, h=40),  
    series="Mean", PI=FALSE) +  
  autolayer(rwf(goog200, h=40),  
    series="Naïve", PI=FALSE) +  
  autolayer(rwf(goog200, drift=TRUE, h=40),  
    series="Drift", PI=FALSE) +  
  ggtitle("Google stock (daily ending 6 Dec 2013)") +  
  xlab("Day") + ylab("Closing Price (US$)") +  
  guides(colour=guide_legend(title="Forecast"))
```

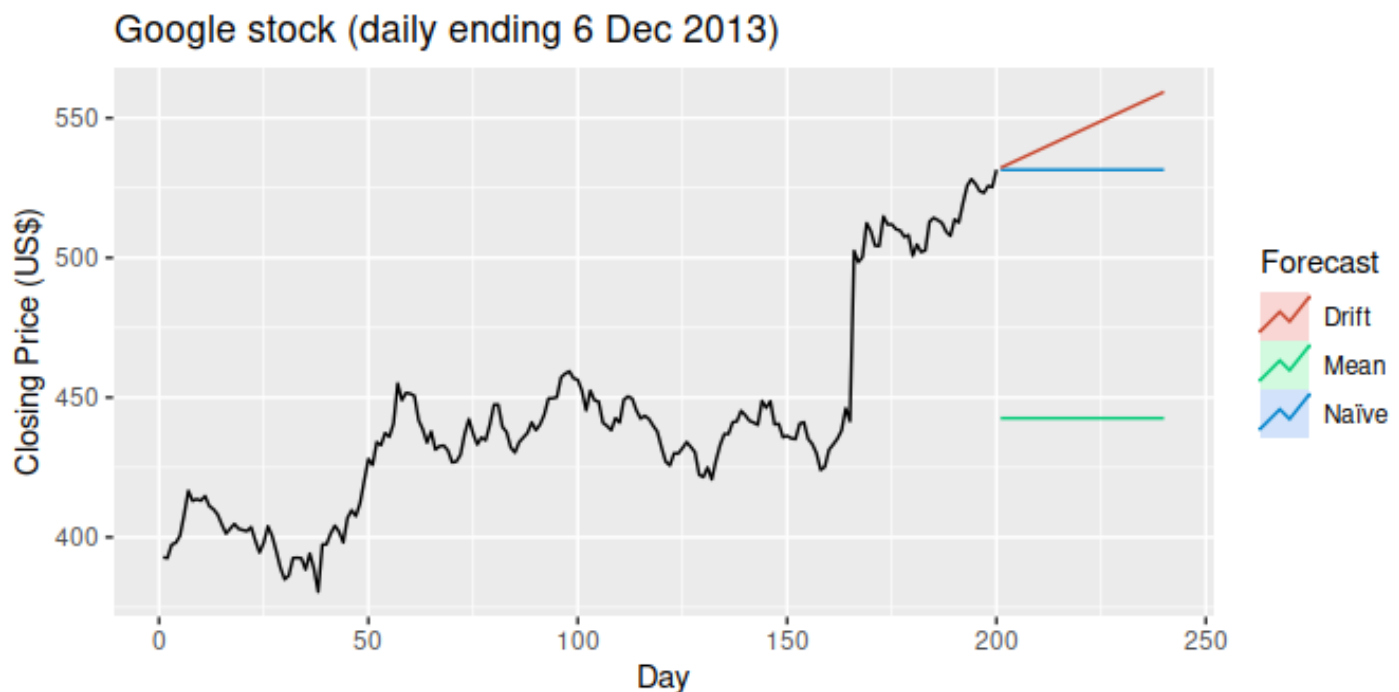


Figure 3.2: Forecasts based on 200 days of the Google daily closing stock price.

Sometimes one of these simple methods will be the best forecasting method available; but in many cases, these methods will serve as benchmarks rather than the method of choice. That is, any forecasting methods we develop will be compared to these simple methods to ensure that the new method is better than these simple alternatives. If not, the new method is not worth considering.