Homework 3

Nick Climaco

February 12, 2024

Chapter 5: The Forecaster's Toolbox

Exercise 1

Produce forecasts for the following series using whichever of NAIVE(y), SNAIVE(y) or RW(y ~ drift()) is more appropriate in each case:

Australian Population (global_economy) Bricks (aus_production) NSW Lambs (aus_livestock) Household wealth (hh_budget). Australian takeaway food turnover (aus_retail).

Exercise 2

Use the Facebook stock price (data set gafa_stock) to do the following:

Produce a time plot of the series. Produce forecasts using the drift method and plot them. Show that the forecasts are identical to extending the line drawn between the first and last observations. Try using some of the other benchmark functions to forecast the same data set. Which do you think is best? Why?

Exercise 3

Apply a seasonal naïve method to the quarterly Australian beer production data from 1992. Check if the residuals look like white noise, and plot the forecasts. The following code will help.

```r
# Extract data of interest
recent_production <- aus_production |>
  filter(year(Quarter) >= 1992)

# Define and estimate a model
fit <- recent_production |> model(SNAIVE(Beer))
#Look at the residuals

fit |> gg_tsresiduals()
```

```
# Look a some forecasts
fit |> forecast() |> autoplot(recent_production)
```

What do you conclude?

4) Repeat the previous exercise using the Australian Exports series from `global_economy` and the Bricks series from `aus_production`. Use whichever of NAIVE() or SNAIVE() is more appropriate in each case.

Exercise 4

Repeat the previous exercise using the Australian Exports series from global_economy and the Bricks series from aus_production. Use whichever of NAIVE() or SNAIVE() is more appropriate in each case

Exercise 7

For your retail time series (from Exercise 7 in Section 2.10):

a. Create a training dataset consisting of observations before 2011 using

```
{r}
myseries_train <- myseries |>
  filter(year(Month) < 2011)
```

b. Check that your data have been split appropriately by producing the following plot.

```
autoplot(myseries, Turnover) +
  autolayer(myseries_train, Turnover, colour = "red")
```

c. Fit a seasonal naïve model using SNAIVE() applied to your training data (myseries_train).

```
fit <- myseries_train |>
  model(SNAIVE())
```

d. Check the residuals.

```
fit |> gg_tsresiduals()
```

Do the residuals appear to be uncorrelated and normally distributed?

e.Produce forecasts for the test data

```
fc <- fit |>
  forecast(new_data = anti_join(myseries, myseries_train))
fc |> autoplot(myseries)
```

f. Compare the accuracy of your forecasts against the actual values.

```
fit |> accuracy()
fc |> accuracy(myseries)
```

g.How sensitive are the accuracy measures to the amount of training data used?