Time Series and Forecasting
Forecasting

Maria Eduarda Silva
mesilva@fep.up.pt

School of Economics, University of Porto

LMU, July 1 2022

Forecasting

- Predicting the future as accurately as possible, given all of the information available
  - historical data
  - knowledge of any future events that might impact the forecasts
- What is easier to forecast?
  - daily electricity demand in 3 days time
  - timing of next Halley comet appearance
  - time of sunrise this day next year
  - Google stock price tomorrow
  - Google stock price in 6 months time
  - total sales of drugs in German pharmacies next month

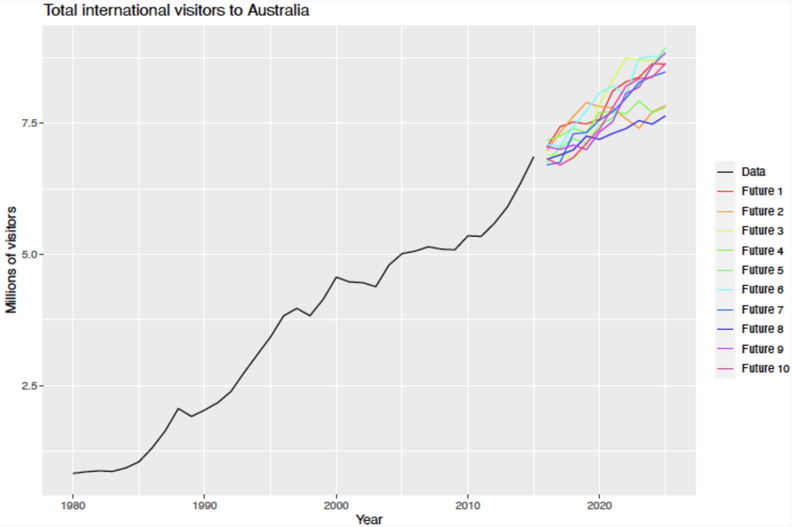
Forecasting

Something is easier to forecast depending on

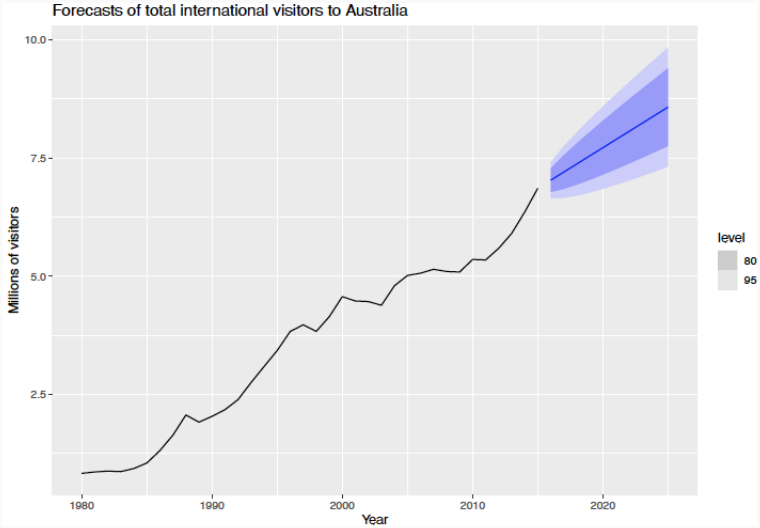
- how well we understand the factors that contribute to it;
- how much data is available;
- how similar the future is to the past;
- there is relatively low natural/unexplainable random variation;
- whether the forecasts can affect the thing we are trying to forecast.

Forecasting is estimating how the sequence of observations will continue into the future

Sample futures



Forecast intervals



Time series models

Time series models use only information on the variable to be forecast

$y_{t+1} = f(y_t, y_{t-1}, y_{t-2}, \ldots, error)$

where t is time and y_t is the quantity of interest at time t like: sales, electricity demand.

ARIMA models and exponential smoothing

- Useful when predictor variables not known or measured
- Useful if prediction of predictor variables difficult
- Does not lead to understanding of the system

Cross-sectional models

Cross-sectional models assume that variable to be forecast is affected by one or more **predictor variables**

$y = f(x_1, x_2, \ldots, error)$

where $x_1, x_2, \ldots$ are variables such as current temperature, GDP, population, time of the day, day of the week, etc

regression models

Mixed models

$y_{t+1} = f(y_t, y_{t-1}, y_{t-2}, \ldots, x_1, x_2, \ldots, error)$
dynamic regression models, panel data models, longitudinal models, transfer function models.

Statistical Forecasting

- Thing to be forecasted: a random variable y_{T+h}
- $\mathcal{F} : y_1, \ldots, y_T$ represents the observations (what we know)
- $y_{T+h}|\mathcal{F}$ means *the random variable y_{T+h} given what we know in $\mathcal{F}$*
- Forecast distribution: the distribution of the random variable *given what we know* $y_{T+h}|\mathcal{F}$
- The point forecast is the mean (or median) of $y_{T+h}|\mathcal{F}$
- The forecast variance is $\mathrm{var}(y_{T+h}|\mathcal{F})$
- A predicton interval or interval forecast is a range of values of y_{T+h} with high confidence

Prediction and Forecasting

- Forecasting method: algorithm which produces a point forecast

Prediction and Forecasting

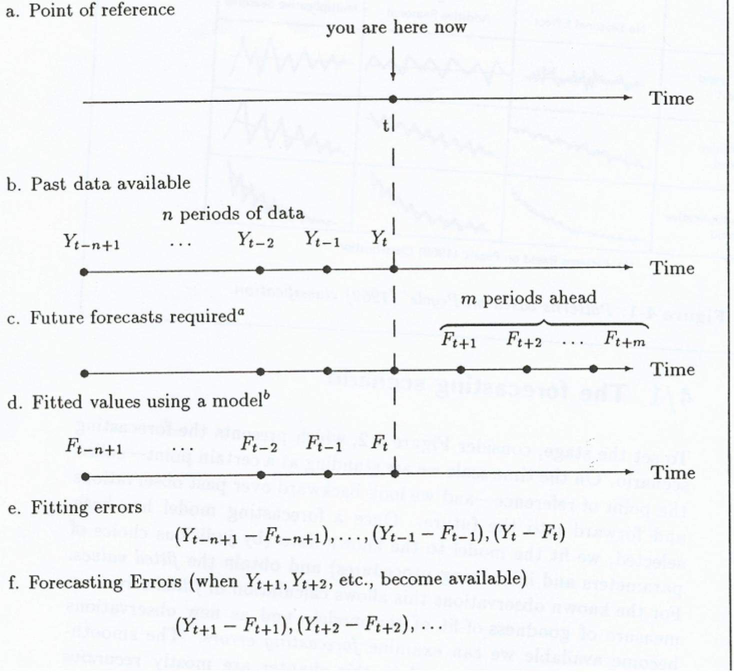
- Forecasting method: algorithm which produces a point forecast
- Statistical model: is a Data Generation Process may be used to

Prediction and Forecasting

- Forecasting method: algorithm which produces a point forecast
- Statistical model: is a Data Generation Process may be used to
  - construct a probability distribution for y_{n+h} (from where one can obtain a point forecast)

Prediction and Forecasting

- Forecasting method: algorithm which produces a point forecast
- Statistical model: is a Data Generation Process may be used to
  - construct a probability distribution for y_{n+h} (from where one can obtain a point forecast)
  - construct confidence intervals for the forecasts


a. Point of reference
you are here now
Time
t
b. Past data available
n periods of data
Y_{t-n+1} $\dots$ Y_{t-2} Y_{t-1} Y_t
Time
c. Future forecasts required[a]
m periods ahead
F_{t+1} F_{t+2} $\dots$ F_{t+m}
Time
d. Fitted values using a model[b]
F_{t-n+1} $\dots$ F_{t-2} F_{t-1} F_t
Time
e. Fitting errors
$(Y_{t-n+1} - F_{t-n+1}), \dots, (Y_{t-1} - F_{t-1}), (Y_t - F_t)$
f. Forecasting Errors (when Y_{t+1}, Y_{t+2}, etc., become available)
$(Y_{t+1} - F_{t+1}), (Y_{t+2} - F_{t+2}), \dots$

Strategy

- Split the series into (i) training set (ii) test set

Strategy

- Split the series into (i) training set (ii) test set
- Choose a method/model

Strategy

- Split the series into (i) training set (ii) test set
- Choose a method/model
- Start the forecasting procedure /Estimate the model using the training set

Strategy

- Split the series into (i) training set (ii) test set
- Choose a method/model
- Start the forecasting procedure /Estimate the model using the training set
- If the residuals of the model are not white noise they may contain important information for the forecasts

Strategy

- Split the series into (i) training set (ii) test set
- Choose a method/model
- Start the forecasting procedure /Estimate the model using the training set
- If the residuals of the model are not white noise they may contain important information for the forecasts
- A model which fits the training set well will not necessarily forecast well

Strategy

- Split the series into (i) training set (ii) test set
- Choose a method/model
- Start the forecasting procedure /Estimate the model using the training set
- If the residuals of the model are not white noise they may contain important information for the forecasts
- A model which fits the training set well will not necessarily forecast well
- Produce forecasts for the test set to obtain the forecasting accuracy measures

Strategy

- Split the series into (i) training set (ii) test set
- Choose a method/model
- Start the forecasting procedure /Estimate the model using the training set
- If the residuals of the model are not white noise they may contain important information for the forecasts
- A model which fits the training set well will not necessarily forecast well
- Produce forecasts for the test set to obtain the forecasting accuracy measures
- Optimize parameters or model

Strategy

- Split the series into (i) training set (ii) test set
- Choose a method/model
- Start the forecasting procedure /Estimate the model using the training set
- If the residuals of the model are not white noise they may contain important information for the forecasts
- A model which fits the training set well will not necessarily forecast well
- Produce forecasts for the test set to obtain the forecasting accuracy measures
- Optimize parameters or model
- Obtain forecasts

Definitions and notations

- Aim: given a time series with n observations, $y_1, y_2, \ldots, y_T$ forecast $y_{T+1}, y_{T+2}, \ldots, y_{T+h}$

Definitions and notations

- Aim: given a time series with n observations, $y_1, y_2, \ldots, y_T$ forecast $y_{T+1}, y_{T+2}, \ldots, y_{T+h}$
- T is the origin, $T+h$ is the forecast horizon and h is the number of steps-ahead

Definitions and notations

- Aim: given a time series with n observations, $y_1, y_2, \ldots, y_T$ forecast $y_{T+1}, y_{T+2}, \ldots, y_{T+h}$
- T is the origin, $T + h$ is the forecast horizon and h is the number of steps-ahead
- $\hat{y}_T(h)$ denotes the (point) forecast of y_{T+h}, $h = 1, 2, \ldots$

Definitions and notations

- Aim: given a time series with n observations, $y_1, y_2, \ldots, y_T$ forecast $y_{T+1}, y_{T+2}, \ldots, y_{T+h}$
- T is the origin, $T+h$ is the forecast horizon and h is the number of steps-ahead
- $\hat{y}_T(h)$ denotes the (point) forecast of y_{T+h}, $h = 1, 2, \ldots$
- $\hat{y}_T(h)$ is a function of $\underline{y} = (y_1, y_2, \ldots, y_T)$, $\hat{y}_T(h) = g(\underline{y})$

Definitions and notations

- Aim: given a time series with n observations, $y_1, y_2, \ldots, y_T$ forecast $y_{T+1}, y_{T+2}, \ldots, y_{T+h}$
- T is the origin, $T+h$ is the forecast horizon and h is the number of steps-ahead
- $\hat{y}_T(h)$ denotes the (point) forecast of y_{T+h}, $h = 1, 2, \ldots$
- $\hat{y}_T(h)$ is a function of $\underline{y} = (y_1, y_2, \ldots, y_T)$, $\hat{y}_T(h) = g(\underline{y})$
- Forecast error at i steps-ahead $e_n(i) = y_{n+i} - \hat{y}_n(i)$

Criteria to evaluate forecasts

- Scale dependent measures

Criteria to evaluate forecasts

- Scale dependent measures
  - Mean Error, **ME**, $(1/h)\sum_{i=1}^{h} e_n(i)$

Criteria to evaluate forecasts

- Scale dependent measures
  - Mean Error, **ME**, $(1/h)\sum_{i=1}^{h} e_n(i)$
  - Mean Absolute Error, **MAE**, $(1/h)\sum_{i=1}^{h} |e_n(i)|$

Criteria to evaluate forecasts

- Scale dependent measures
  - Mean Error, **ME**, $(1/h)\sum_{i=1}^{h} e_n(i)$
  - Mean Absolute Error, **MAE**, $(1/h)\sum_{i=1}^{h} |e_n(i)|$
  - Mean Squared Error, **MSE**, $(1/h)\sum_{i=1}^{h}(e_n(i))^2$

Criteria to evaluate forecasts

- Scale dependent measures
  - Mean Error, **ME**, $(1/h)\sum_{i=1}^{h} e_n(i)$
  - Mean Absolute Error, **MAE**, $(1/h)\sum_{i=1}^{h} |e_n(i)|$
  - Mean Squared Error, **MSE**, $(1/h)\sum_{i=1}^{h}(e_n(i))^2$

- Scale independent measures:

Criteria to evaluate forecasts

- Scale dependent measures
  - Mean Error, **ME**, $(1/h)\sum_{i=1}^{h} e_n(i)$
  - Mean Absolute Error, **MAE**, $(1/h)\sum_{i=1}^{h} |e_n(i)|$
  - Mean Squared Error, **MSE**, $(1/h)\sum_{i=1}^{h}(e_n(i))^2$

- Scale independent measures:
  - Mean Absolute Percentual Error, **MAPE**, sensible for $y_t >> 0$ and when natural zero exists $(1/h)\sum_{i=1}^{h} |e_n(i)/y_{n+i}| \times 100$

Criteria to evaluate forecasts

- Scale dependent measures
  - Mean Error, **ME**, $(1/h)\sum_{i=1}^{h} e_n(i)$
  - Mean Absolute Error, **MAE**, $(1/h)\sum_{i=1}^{h} |e_n(i)|$
  - Mean Squared Error, **MSE**, $(1/h)\sum_{i=1}^{h}(e_n(i))^2$

- Scale independent measures:
  - Mean Absolute Percentual Error, **MAPE**, sensible for $y_t >> 0$ and when natural zero exists $(1/h)\sum_{i=1}^{h} |e_n(i)/y_{n+i}| \times 100$
  - Mean Absolute Scaled Error, **MASE** (Hyndman and Koehler, 2006) Define the scaled error as:

$$q_n(i) = \frac{e_n(i)}{1/(n-1)\sum_{i=2}^{n} |y_i - y_{i-1}|}$$

MASE=mean($|q_n(i)|$)