MBA Forecasting Course
Access Week

"The farther backward you can look, the farther forward you are likely to see."
–Winston Churchill

"Forecasting is like driving a car forward by looking in the rear-view mirror. If it hasn't happened before, you won't see it in the model."
–unknown

1 Overview

This lecture will start with a course overview: how the lectures will be organized and how course grade will be determined.

The game of "Texas Hold'em Poker" will be used to highlight the features of forecasting in a business setting.

Next we will introduce the basic idea of a time series, deterministic and stochastic. The students should be able to give examples. The basic "Approaches" to time series analysis and forecasting are explained. This is followed by a basic introduction to R.

The first forecasting technique, Exponential Smoothing, is presented. We start with a basic model, then add a trend and end with a series that has a trend and a seasonal cycle.

2 Forecasting is not Fortune Telling

- First might think of a Crystal Ball or Tarot Cards.
- Think Poker Cards.
- With optimal forecasting there is still uncertainty.
- Purpose is to efficiently use available information.
- Only useful if you can take advantage of better information. For example
  - Contingency plans.
  - Flexibility in the production process.
  - The right amount of hedging, for uncertainty.
- Suppose you had "different" types of cards.

3 What is a time series?

A set of observations indexed by time.

$$\ldots, x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots$$

- Deterministic

$$\begin{array}{rl} \ldots, 1, 2, 3, 4, 5, 6, 7, 8, \ldots & x_t = t \\ \ldots, 1, 4, 9, 16, 25, 36, 49, 64, \ldots & x_t = t^2 \end{array}$$

- Random, let $z_t \sim iidN(0, 1)$

$$\begin{array}{rl} x_t = z_t & x_t \sim iidN(0, 1) \\ x_t = \mu + z_t & x_t \sim iidN(\mu, 1) \\ x_t = \mu + \sigma z_t & x_t \sim iidN(\mu, \sigma^2) \end{array}$$

- Random with dependence, let $z_t \sim iidN(0, 1)$

$$x_t = \rho x_{t-1} + \sigma z_t \qquad x_t \sim \mathrm{AR}(1)$$

4 Examples of time series

- GDP for the United States (quarterly, annual)
- Interest rates (from traded bonds)
- Stock price of Microsoft (traded stock)
- Price of natural gas (market price quotes)
- Sales of hot water heaters (monthly)

5 Methodology

For time series we will want to determine the following:

1. What model explains the observed series? (`IDENTIFY`)
2. How do we estimate this model? (`ESTIMATE`)
3. How do we (use the estimated model to) forecast the observed series? (`FORECAST`)

6 Basic Approaches to time series

1. Smoothing.
   - No formal model.
   - Averages of past values.
   - Fairly common. Easy to use and not bad in some situations.
2. ARIMA models.
   - Most widely used time series models.
   - Typically start by transforming the observed series $W_t = f(X_t)$.
   - Fit a model with AR(p) and MA(q) components
   $$W_t = \mu + \phi_1 W_{t-1} + \ldots + \phi_p W_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q}.$$
   - X_t may be a vector. Will be able to use other variables to help forecast. VAR models in week 5.
3. Regression models with time dependent errors.
   - Suppose you have the model
   $$y_t = \alpha + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \epsilon_t.$$
   A linear regression model where the variables are time series.
   - ϵ_t is typically not "nice".
   - The error process typically has dependence. It may follow its own ARMA model.
   - How do we <u>estimate</u> this type of a model?
   - How do we <u>forecast</u> with this type of a model?

7 Let's talk about R and R Markdown

1. More than one way to perform any task
2. I will give you one approach to the computer code
3. Libraries
   - Install on you computer
   - Need to load in your program.

Let's do an example (in the directory "Simple_Expo_Smoothing") "NG_plot_SES.Rmd"

- Read some data into R.
- Look at it to make sure it was entered correctly.
- Look at some of its summary statistics.
- Plot the series.

8 More about R

- It is very helpful to include comments
  # your comment is a comment
- R is case sensitive
- It is very helpful to include titles for programs, graphs, regressions, forecasts, etc.

9 Points to note in "NG_plot_SES.Rmd"

- The function ts() creates a time series from a table.
- ts(data_from_excel, start=c(1997,1), frequency = 12, names =c("Natural Gas Price"))
- Frequency (12 monthly, 1 annual, 4 quarterly)
- Print your series to make sure it is correct in R
- Print summary statistics to learn about your data.
- Plot your series.
- Online help. ?function It is very extensive: equivalent to an entire shelf full of manuals

10 Smoothing

We will consider three types of smoothing

1. Exponential smoothing – nontrending data.

2. Holt smoothing – for data with a trend.

3. Holt-Winter's smoothing – for seasonal data with a trend.

- Smoothing is not a model-based method.

- Think of smoothing as an algorithm to plot a smooth curve through the data.

- There are reasonable situations to use smoothing

  1. Very few observations.
  2. Huge number of variables to forecast.
  3. Slowly evolving mean with measurement error.

- Smoothing is fairly easy to use. Does not require much training.

10.1 When is it optimal?

Answer: Slowly evolving mean with measurement error

Suppose

$$\mu_t = \mu_{t-1} + u_t$$

where $u_t \sim iid(0, \sigma_u^2)$. The mean follows a random walk.

The observed series is

$$x_t = \mu_t + \epsilon_t$$

where $\epsilon_t \sim iid(0, \sigma_\epsilon^2)$.

It is assumed that u_t and ϵ_t are uncorrelated.

The expectation is that $\sigma_u \ll \sigma_\epsilon$.

The model implies

$$x_{T+h} = \mu_{T+h} + \epsilon_{T+h}$$

Just substitute in the equation of μ_t repeatedly to get that

$$x_{T+h} = \mu_T + \sum_{s=1}^{h} u_{T+s} + \epsilon_{T+h}.$$

The u and the ϵ have zero mean so the estimate (forecasted value) is based on μ_T. Unfortunately, this is unknown hence we must estimate it. To get the estimate (forecasted value) use this algorithm

1. $\hat{\mu}_1 = x_1$
2. $\hat{\mu}_t = \alpha x_t + (1-\alpha)\hat{\mu}_{t-1}$ for $t = 2, 3, \ldots, T$.
3. Future values will be estimated with $\hat{\mu}_T$, i.e. $\hat{x}_{T+h} = \hat{\mu}_T$

where

- α is called the smoothing parameter.
- $\alpha \in [0, 1]$.
- Smaller values of α put less weight on the current observation. In the limit as α approaches zero, every value is estimated to be x_1.
- Larger values of α put more weight on the current observation and less weight on past observations. In the limit as α approaches one the estimate is simply the last observation (which may not be a bad estimate).
- As a general rule of thumb, α should take values on the interval $[.05, .35]$.
- It is possible to estimate σ_ϵ^2 with the residuals.

10.2 Exponential weights

To see where the name comes from consider

$$\begin{aligned}
\hat{\mu}_t &= \alpha x_t + (1-\alpha)\hat{\mu}_{t-1} \\
&= \alpha x_t + (1-\alpha)\left(\alpha x_{t-1} + (1-\alpha)\hat{\mu}_{t-2}\right) \\
&= \vdots \\
&= \sum_{s=0}^{t-1} \alpha(1-\alpha)^s x_{t-s}
\end{aligned}$$

The recursive structure is very convenient. You do not need to go through the entire data set every time there is new data.

Let's do an example (in the directory "Simple_Expo_Smoothing") "NG_plot_SES.Rmd"''

10.3 Holt Smoothing

- Appropriate for a model with slowly evolving mean and a time trend.
- $x_t = \mu_t + \beta_t t + \epsilon_t$

- $\mu_t = \mu_{t-1} + u_t$
- $\beta_t = \beta_{t-1} + v_t$
- Each of these series ϵ_t, u_t and v_t are iid, have zero means and are independent of each other.

You again follow an algorithm to obtain forecasts

1. **Initialize**:
   (a) $\hat{x}_2 = x_2$
   (b) $\hat{\beta}_2 = x_2 - x_1$
2. **Update**:
   (a) $\hat{x}_t = \alpha x_t + (1-\alpha)\left(\hat{x}_{t-1} + \hat{\beta}_{t-1}\right)$
   (b) $0 < \alpha < 1$
   (c) $\hat{\beta}_t = \gamma\left(\hat{x}_t - \hat{x}_{t-1}\right) + (1-\gamma)\hat{\beta}_{t-1}$
   (d) $0 < \gamma < 1$
   (e) for $t = 3, 4, \ldots, T$.
3. **Forecast**:
   (a) $\hat{x}_{T+h} = \hat{x}_T + \hat{\beta}_T h$

- α smoothes the level.
- γ smoothes the slope on the time trend.

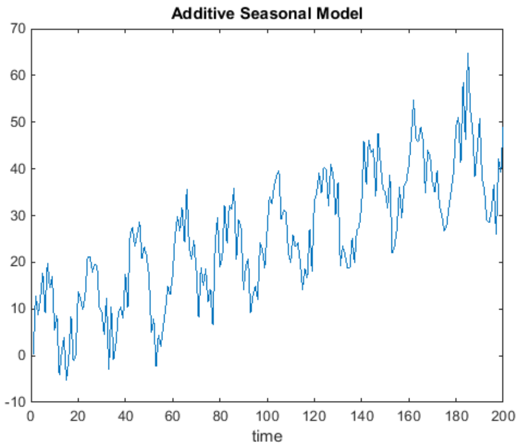
Let's look at an example (in the directory "Holt_Smoothing") "Annual_Hot_Water.Rmd"

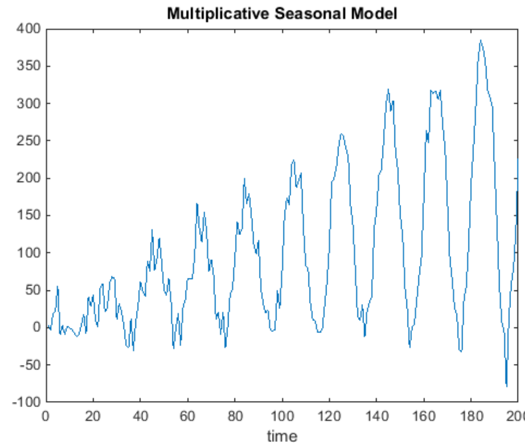
10.4 Holt-Winter's Smoothing for seasonal effects

There are two forms

1. Multiplicative $x_t = (\mu_t + \beta_t t)\, s_t + \epsilon_t$
2. Additive $x_t = \mu_t + \beta_t t + s_t + \epsilon_t$.

Figure 1: Two types of seasonal models. For the additive model the amplitude is constant. For the multiplicative model the amplitude changes with the trend of the series.





- Again you follow an algorithm. Initialize, Update, Forecast.
- These follow a set of recursive equations.
- In general there will be three different weights that you can set.
- Let's do an example (in the directory "Holt_Winters_Smoothing") "Quarterly_Sales.Rmd"