An introduction to univariate time-series

Susan Thomas

August 12, 2009

Course Syllabus

Main topics

- Stationary stochastic processes: building and estimating linear models for the mean of a time series.
- Forecasting the mean of stationary stochastic processes
- Non-stationary processes: detecting and testing models of stationarity in the time series.
- Modelling volatility of a time series.
- Switching models, modelling cycles in time series.
- Non-parametric time series modelling.

Course modalities

Course materials

- Main reference book for the class: James Hamilton, "Time Series Analysis".
- The statistical package used for the class will be `R`.
- The typesetting package used for project submissions will be LaTeX. Submissions will be of the PDF format. `Word` files will not be accepted for evaluation.

Course Evaluation

- Student performance evaluation is based on:
  1. Project (45%)
  2. Quizzes (5%), and
  3. Final exam (50%)

Project modalities

- Project group size will be two people.
- The scope and format of the project submissions should be similar to the types of papers that get published in *Applied Financial Economics*.
- There should be a clearly defined economic question to the project.
  The sole focus should *not* be on the econometric technique.
- Replications of papers already published, based on univariate time series analysis, using Indian data is permitted as long as the application has some innovation.
  This could be involve either a new period of data, or a modified version of the methodology.

Dates for the project

- The project proposal should be fixed and submitted by 1^{st} September.
  The project proposal should include:
  - The question the project will address. If possible, it should include the reference paper that you used in putting the proposal together.
  - The data that you will be analysing for the project.

  Depending upon whether I find the topics passable or not, the projects can either be fixed for the group. Or I will assign a project to the group.
- There will be a mid-semester review involving a submission (of 4-5 pages) and a class presentation. Tentative date of presentation: 29^{th} September 2009.
- There will be a final presentation and paper submission. Date to be decided later.

Dates for the project

- The project proposal should be fixed and submitted by 1^{st} September.
  The project proposal should include:
  - The question the project will address. If possible, it should include the reference paper that you used in putting the proposal together.
  - The data that you will be analysing for the project.

  Depending upon whether I find the topics passable or not, the projects can either be fixed for the group. Or I will assign a project to the group.
- There will be a mid-semester review involving a submission (of 4-5 pages) and a class presentation. Tentative date of presentation: 29^{th} September 2009.
- There will be a final presentation and paper submission. Date to be decided later.

Introduction to Time Series Analysis

The importance of being a time series

- A time series is a variable is indexed by time. Therefore, if x is the variable, a time series is denoted by $(x_0, x_1, x_2, \ldots, x_{t-1}, x_t, x_{t+1}, \ldots, x_T)$.
- Each period, there is an information shock, ϵ_t.
  This is also called the “innovation” to the x_t variable.
- Therefore, the time series x_t grows as a cumulation of innovations or information shocks.
  It can grow as an explicit function of time (*deterministic trend*),
  as a function of the latest innovation (*stochastic trend*),
  or both.
- The relevant **information set,** I_t becomes:

$$I_t \sim (x_{t-1}, x_{t-2}, \ldots, t, \epsilon_{t-1}, \epsilon_{t-2}, \ldots)$$

The importance of being a time series

- A time series is a variable is indexed by time. Therefore, if x is the variable, a time series is denoted by $(x_0, x_1, x_2, \ldots, x_{t-1}, x_t, x_{t+1}, \ldots, x_T)$.
- Each period, there is an information shock, ϵ_t.
  This is also called the "innovation" to the x_t variable.
- Therefore, the time series x_t grows as a cumulation of innovations or information shocks.
  It can grow as an explicit function of time (*deterministic trend*),
  as a function of the latest innovation (*stochastic trend*),
  or both.
- The relevant **information set,** I_t becomes:

$$I_t \sim (x_{t-1}, x_{t-2}, \ldots, t, \epsilon_{t-1}, \epsilon_{t-2}, \ldots)$$

The importance of being a time series

- A time series is a variable is indexed by time. Therefore, if x is the variable, a time series is denoted by $(x_0, x_1, x_2, \ldots, x_{t-1}, x_t, x_{t+1}, \ldots, x_T)$.
- Each period, there is an information shock, ϵ_t.
  This is also called the "innovation" to the x_t variable.
- Therefore, the time series x_t grows as a cumulation of innovations or information shocks.
  It can grow as an explicit function of time (*deterministic trend*),
  as a function of the latest innovation (*stochastic trend*),
  or both.
- The relevant **information set,** I_t becomes:

$$I_t \sim (x_{t-1}, x_{t-2}, \ldots, t, \epsilon_{t-1}, \epsilon_{t-2}, \ldots)$$

Notation for the time series

- A time series variable: x_t.
- A time series: $(x_t, x_{t-1}, x_{t-2}, \ldots, x_1, x_0)$.
- T is the standard number of observations in a time series.

Notation on time series operators

- *Lag* values of x_t: x_{t-1} is x_t lagged by one. This is done using the *lag operator*, denoted by $L(x_t)$.
  - Example 1: $x_t - \alpha x_{t-1} = (1 - \alpha L)x_t$.
  - Example 2: $(1 - \beta L)^2 x_t = x_t - 2\beta x_{t-1} + \beta^2 x_{t-2}$.
- The *backward lag* is the same as a value that *leads* in time: x_{t+1}.
  This is denoted as $B(x_t)$.
  It is also called the *forward operator*.
- Example 1: $x_t - \alpha x_{t+1} = (1 - \alpha B)x_t$.

Notation on time series operators

- *Differencing* a series: create a series which is the differences between adjacent terms in a time series, x_t. This is denoted as Δx_t.
  - Example 1: $y_t = \Delta x_t = x_t - x_{t-1} = (1 - L)x_t$. The "order of the differencing" here is one.
  - Example 2: differencing of order two

$$
\begin{aligned}
z_t &= \Delta y_t = && \Delta^2 x_t \\
z_t &= (1 - L)y_t = && (1 - L)(1 - L)x_t \\
z_t &= y_t - y_{t-1} = && x_t - 2x_{t-1} + x_{t-2}
\end{aligned}
$$

Notation on time series, continued.

- A time series variable x_t can be typically modelled as:

  $$x_t = f(I_t) + \epsilon_t$$

  where ϵ_t is the innovation or the information shock to the series at every point t, and
  where I_t could be t, or indicator functions of time such as dummies for the day of the week if it is a daily time series, etc.
- We refer to the equation generating the x_t series as the *Data Generating Process* or DGP.

Notation continued.

- $E(\epsilon_t) = 0$ with the standard meaning that on expectation ϵ_t will be zero.
- $E(\epsilon_t^2)$ has the standard meaning of the variance of ϵ_t.
- A third statistic used in time series econometrics is

$$E(\epsilon_t \epsilon_{t-s})$$

This is *autocovariance* of ϵ_t at a lag of s. It is analagous to $E(xy)$ except that x, y are both values of the time series random variable at different lags.
Example: For a lag of 2

t	$x = \epsilon_t$	$y = epsilon_{t-2}$
0	ϵ_0	.
1	ϵ_1	.
2	ϵ_2	ϵ_0
3	ϵ_3	ϵ_1
4	ϵ_4	ϵ_2
.	.	.
T	ϵ_T	ϵ_{T-2}

References

Textbooks

- *Box, G.E.P. and G.M.Jenkins*, (1994), **Time series analysis: forecasting and control**, Prentice Hall International, Inc., New Jersey.
- *Hamilton, J.D.* (1994), **Time series analysis**, Princeton University Press, Princeton, New Jersey.
- *Maddala, G.S. and In-Moo Kim* (1999), **Unit roots, cointegration, and structural change**, Cambridge University Press, Cambridge.
- *Bannerjee, A. et.al.*, (1993), **Cointegration, error correction and the econometric analysis of non-stationary data**, Oxford University Press, Oxford.

Lectures will also refer to specific papers.