Time Series

© 2014, A. C. Davison

Ecole Polytechnique Fédérale de Lausanne
EPFL-FSB-MATHAA-STAT
Station 8, 1015 Lausanne, Switzerland
http://stat.epfl.ch

Please signal any errors or omissions by email to Anthony.Davison@epfl.ch.

Course web page

Further material, links to data sets, etc., can be found at the course web page

http://stat.epfl.ch/page-90340-en.html

Planned lecture contents and exercises

1. Introduction (17/2/14)
   - Motivation and examples
   - Basic notions: stationarity; white noise; autoregression; moving average; periodic series
   - Exercises: 1, 2, 3, 4, 5

2. Simple techniques (24/2/14)
   - Plotting; outliers; transformations; differencing
   - Correlogram and partial correlogram. Examples. Box–Ljung test
   - Exercises: 7, 8, 9, 10

3. Periodogram and smoothing (3/3/14)
   - Periodogram: basic properties
   - Smoothing: moving averages; polynomials; local polynomials; splines; loess; STL decomposition
   - Practical 1

4. AR(1) model (10/3/14)
   - Theory
   - Likelihood, including analysis of beaver data
   - Exercises: 12, 14, 15, 16
5. General theory of stationary linear processes (17/3/14)
   - General theory of stationary processes/spectral density and linear filters
   - AR, MA, ARMA models
   - Exercises: 18, 19, 20, 21, 23

6. ARMA models (24/3/14)
   - Difference equations. First example of ACF for AR2 model
   - Computation of ACF for ARMA models, ARMA(1,1) example
   - Practical 2

7. ARIMA modelling (31/3/14)
   - Computation of PACF; ARIMA modelling and example
   - SARIMA modelling and CO2 example
   - Exercises: 25, 26, 27, 28

8. Regression and forecasting (7/4/14)
   - Regression structure, with beaver example. Cochrane–Orcutt algorithm.
   - Forecasting. Bootstrap for prediction uncertainty.
   - Exercises: 29, 30, 31

9. Financial time series (14/4/14)
   - Stylised facts, GARCH models and some extensions.
   - Basic stochastic volatility models
   - Practical 3

10. Frequency domain analysis (28/4/14)
    - Reminders: spectrum, periodogram. Harmonics and leakage. Tapering. FFT.
    - Exercises: 32, 33, 34, 36, 39

11. State space models (5/5/14)
      Comparison with ARIMA models.
    - Exercises: 40, 41; Question 6 of 2010 exam.

12. Multiple series (12/5/14)
    - Multiple time series: cross-correlogram etc.; bivariate spectral theory.
    - Transfer function modelling
    - Models for multiple time series: VAR, VMA, VARMA, factor models.
    - Practical 4
13. Long-range dependence (19/5/14)
   • Long-range dependence
   • More . . .
   • Work on mini-project

14. Other topics/Wrap-up (26/5/14)
   • Threshold models/Discrete time series
   • Hidden Markov models and Gibbs sampler
   • Work on mini-project

**Final mark**

The final mark will be determined as follows:

- there will be a data analysis project, to be undertaken in pairs, involving analysis of a time series data set (to be found by the students concerned, but validated by me), using R. A report on the data is to be handed in (as a PDF file) by 6 June 2014. More about this may be found below;

- a final exam (probably with 5 questions, of which full correct answers to 4 will give full marks);

- let $E$ be the mark for the final exam ($\leq 3$) and $P$ be that for the project ($\leq 2$). Then the mark $M$ for the course is $1 + E + P$ rounded to the nearest half-mark in \{1, 1.5, \ldots, 5.5, 6\}.

**Data analysis project**

**Proposal**

As soon as possible, I would like to see a **brief** proposal about your project, consisting of:

- a description of the data that you plan to analyse, including some suitable plots and maybe some other data exploration;

- an indication of the source of the data set;

- the objectives of your investigation;

- a brief overview of the analyses you anticipate completing.

Bear in mind that this will count for 40% of the course mark, and the number of credits for this course is 4, so this is much less work than a semester project.

The proposal should be made by **7 April 2014 at the very latest**, and the earlier the better—then I can give feedback, and you can start work.

**Data**

You could look through books in the library, check out links on the course web page, or look at the many times series data sets available in R. It is best if the data comprise a few hundred equally-spaced observations, but you should try to find something that interests you.
Report

The report should be typed in either English or French. Do not include hand-written material unless you want to write out equations and don’t know \LaTeX. Some notes on report-writing can be found at \url{http://stat.epfl.ch/page-34953.html}, though your report should be much shorter than for a semester project.

The structure should be

**Introduction** Briefly state the purpose of the analysis, what type of data were analysed, what methods were used, and what the results were.

**Description** of the methods used, using your own words. Give the key elements only: you can refer to the lecture notes and to books, but should give careful references (to pages and equations etc.) so I can easily look them up. It is not enough simply to give a list of sources at the end of the work: references should be mentioned in the text, and only those mentioned in the text should be listed at the end.

**Discussion** of the results in more detail. Include crucial plots and tables only, make sure they are legible, and that their axis labels and captions are clear and informative; each graph should tell the reader a clear and coherent story. The text can give more details, if they are needed, and show where the graph/table fits into the overall picture. The discussion should show how the statistical methods shed light on the data and the underlying problem. Tables should have appropriate numbers of digits.

**Conclusions** carrying the take-away message from your analysis, in your own words. Again, this should be related to the data. Convince the reader that you know what you did and are aware of its strengths and limitations. Sketch what more you might do, if you had more time.

**Marking scheme**

<table>
<thead>
<tr>
<th></th>
<th>Correctness</th>
<th>Incorrect, many errors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accurate, appropriate use of statistical tools</td>
<td>10 9 8 7 6 5</td>
</tr>
<tr>
<td></td>
<td>Incorrect, many errors</td>
<td>4 3 2 1 0</td>
</tr>
<tr>
<td>Score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interpretation</td>
<td>Correct, appropriate</td>
<td>5 4 3</td>
</tr>
<tr>
<td></td>
<td>Little or none</td>
<td>2 1 0</td>
</tr>
<tr>
<td>Score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discussion</td>
<td>Thoughtful, imaginative</td>
<td>5 4 3</td>
</tr>
<tr>
<td></td>
<td>Banal</td>
<td>2 1 0</td>
</tr>
<tr>
<td>Score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality of writing</td>
<td>Good grammar and punctuation, including mathematics</td>
<td>5 4 3</td>
</tr>
<tr>
<td></td>
<td>Poor grammar and punctuation</td>
<td>2 1 0</td>
</tr>
<tr>
<td>Score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graphics and tables</td>
<td>Clearly labelled, well-chosen, good captions and discussion, appropriate numbers of digits</td>
<td>10 9 8 7 6 5</td>
</tr>
<tr>
<td></td>
<td>Poorly labelled, no discussion, unmotivated, unedited output</td>
<td>4 3 2 1 0</td>
</tr>
<tr>
<td>Score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Referencing</td>
<td>Full, accurate, and detailed references given</td>
<td>5 4 3</td>
</tr>
<tr>
<td></td>
<td>Inadequate citation of sources</td>
<td>2 1 0</td>
</tr>
<tr>
<td>Score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Originality and scope</td>
<td>Original choice of topic, wide range of tools/ideas</td>
<td>10 9 8 7 6 5</td>
</tr>
<tr>
<td></td>
<td>Standard choice of dataset, limited range of tools/ideas</td>
<td>4 3 2 1 0</td>
</tr>
<tr>
<td>Score</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Grand total (max 50) [4]
Time Series: Bibliography

There are many books on statistical time series analysis, in addition to the many more in other fields. There are also two main statistical journals devoted to the area: *Journal of Time Series Analysis*, and *Journal of Forecasting*, plus journals of finance and econometrics/economics plus general statistical journals in which time series articles appear. So plenty of reading is available!

References


