

CSC425 – Time series analysis and forecasting

Homework 1

Due on Friday January 25th, 2013 before midnight

Total points: 22

Reading assignment (pdf documents posted under week 1 and 2)

- Chapter 1 and chapter 2: sections 2.1, 2.2, 2.3
- Review course documents posted under week 1 and 2.

For SAS users

- Review SAS lab tutorial and lecture notes
- Review notes on SAS procedures posted under SAS resources in the course website under the Documents link.
- Useful resources are at <http://www.ats.ucla.edu/stat/sas/> and http://support.sas.com/documentation/cdl_main/

For R users

- Review lab exercise using R, and the example `R_ibm6708.docx` posted under week 1.
- Review notes on R posted under SAS resources in the course website under the Documents link.
- Useful resources are at <http://www.statmethods.net/> and <http://www.ats.ucla.edu/stat/R/>
- Install the following packages that will be used throughout the course:

IMPORTANT INFORMATION FOR R USERS

The following packages will be used in our course and need to be added to R: fBasics, fSeries, tseries, fGarch. I recommend you install the complete package **Rmetrics**. This can be installed in R using the following two commands:

```
> source("http://www.rmetrics.org/Rmetrics.R")
> install.Rmetrics()
```

PROBLEMS

Problem 1 [8 points]

Consider the daily closing prices for Apple stock prices from January 2000 to December 2012. The data file is `apple_m_2000_2012.csv` and contains dates (date) and daily prices (price). Note that the data are separated by commas. The simple returns are computed in the codes below.

Code to import data for Problem 1

R and SAS users: Save the datafile on your computer. Change the work directory in SAS/R to the folder containing the data. **Be careful** when copying code from this file to SAS or R since the quote symbols “ ” copied from a Word document may not be recognized in SAS and R.

SAS code (review SAS codes under week 1 and 2, and week 1 lab exercise)

```

PROC IMPORT datafile= "apple_d_2000_2012.csv" out = apple dbms=csv
replace;
delimiter=',';
getnames = yes;
RUN;

* DATA step to create new variable "return";
DATA apple;
* create copy of dataset;
set apple;
* return computed as p_t/p_(t-1)-1 ;
return = price/lag(price)-1;

* print out dataset to verify data were entered correctly;
proc print;
run;

```

R code (review R code under week 1 and 2, and week 2 lab exercise - e.g. the file R_ibm6708.docx under week 1)

```

# import libraries for TS analysis
library(tseries)
library(fBasics)
library(zoo)
# create dataset myd
myd= read.table('apple_d_2000_2012.csv', header=T, sep=',')
price = myd$Price
# create time series object "ts"
ts = zoo(price,as.Date(as.character(myd$Date),format="%m/%d/%Y"))
return = lag(ts)/ts-1;

# the following code fixes problem with dates.
return=c(0, coredata(return))
return =
zoo(return,as.Date(as.character(myd$Date),format="%m/%d/%Y"))

# strip off the dates and just create a simple numeric vector
ret = coredata(return);

```

- Compute a new variable **lnreturn** for log returns by using the formula $\ln\text{return} = \log(\text{ret}+1)$.
- Compute the sample mean, standard deviation, skewness, excess kurtosis, minimum, and maximum of the log-return series.
- Based on the descriptive statistics values, the histogram and the qqplot, what can you say about the distribution of the log returns?
- Test the null hypothesis of perfect symmetry at 5% significance level.
- Test the null hypothesis of excess kurtosis equal to zero (normal tails) at 5% significance level.
- Test the null hypothesis that the log returns are normally distributed. Use the Jarque-Bera test.

Problem 2 [12 points]

The file CPIUS12.csv contains monthly inflation data from January 1991 to November 2012. The Inflation rate is calculated from the Consumer Price Index (CPI-U) which is compiled by the Bureau of Labor Statistics, and can be downloaded from the St Louis Federal Reserve Bank website. A CPI of 195 indicates 95% inflation since 1982.

The commonly used inflation rate is the change in the Consumer Price Index from a year earlier. Thus the inflation rate of 1.77% observed in November 2012 indicates the change of CPI index from the previous November 2011.

The data file contains three variables: **date** (month/day/year), **cpi** and inflation rate (**inflation**). Use the code below (SAS or R) to import the data into the software package of your choice. See examples of analysis posted under Week 2. *[Try to write the code to import the data by yourself....if you run into problems, post a message at D2L – I'll post the code there!]*

R and SAS users: Save the datafile on your computer. Change the work directory in SAS/R to the folder containing the data.

Questions

- a) Analyze distribution of inflation rates using a histogram and a quantile plot.
- b) Describe the distribution of inflation rates in terms of symmetry/skewness and tails.
- c) Test the hypothesis of normality using the Jarque Bera test at 5% level
- d) Is the result of the test consistent with your answers for points a) and b)?
- e) Create a time plot for the inflation rate data. Make sure the plot is correctly labeled and nicely titled. Analyze the time trend displayed by the plot, and discuss if data show any striking pattern, such as trends or seasonality?
- f) Compute and plot the first 15 lags of ACF and discuss if the series shows evidence of serial correlation.
- g) Compute the Ljung Box test to evaluate if the inflation rates are serially correlated.
- h) Evaluate the largest order of the serial correlations that are significantly different from zero.

“Reflection” Problem [2 pts]

Post a message on the discussion board reflecting on the topics in week 2. Indicate the assignment in this module you found to be the easiest, the one you found to be the hardest, and why. I created a new Thread called “Week 2 Comments”. Click on the Discussions link, and select the forum. Add your comments to week 2 thread.

Submission instructions

Submit the homework at the Course Web <http://d2l.depaul.edu> under the Dropbox link.

1. Submit a word document with your answers and the sas/r code in a separate file.
2. Keep a copy of all your submissions!
3. If you have questions about the homework, email me BEFORE the deadline.
4. The assignment will lose 10% of the points per day, after the due date.
5. Assignments submitted FIVE days after the deadline will not be accepted.