

CSC425 – Time series analysis and forecasting

Homework 4

Due on Friday March 1st, 2013

Total Points: 27

Reading assignment:

1. Chapter 2 sections 2.6, 2.7, 2.8, 2.9 on ARIMA and seasonal models. Read Chapter 3, sections 3.1 – 3.5 on volatility models
2. Review course documents posted under week 6 and 7.
3. Forecasting a Seasonal ARMA Process. Document *Steelshipments_SARIMAmodel.pdf* posted on Week 7.
4. Papers posted under week 7, especially
 - a. Cont, R. (2001) Empirical properties of asset returns: stylized facts and statistical issues. Quantitative Finance, Vol. 1, No. 2, 223-236.
 - b. Rydberg, T.H. (2001) Realistic Statistical Modelling of Financial Data. International Statistical Review 68:3, 233-258
 - c. Extracts from the book "Market Models" discussing empirical properties of GARCH processes

Problem 1 [just for fun!] – NO CREDIT

I noticed some students had questions regarding the SARIMA notation. Write the model expression for the following SARIMA models for the $\{X_t\}$ time series.

Example: SARIMA(2,0,0)(1,0,0)[12]

The notation indicates a regular AR(2) model and a seasonal AR(1)[12] model

$$(1-\phi_1 B - \phi_2 B^2)(1-\Phi_1 B^{12})X_t = a_t$$

- a) SARIMA(2,1,0)(1,1,0)[4]
- b) SARIMA(1,0,1)(0,0,1)[12]
- c) SARIMA(0,1,3)(0,1,1)[4]

Problem 2 [15 pts]

Consider the quarterly earnings per share of the FedEx stock from the first quarter of 2000 to the fourth quarter of 2010 fiscal year. The data were obtained from the Fedex website. Here is the code to enter the data in SAS: (HINT: similar to the house starts analysis of week 6)

SAS code

```
data fedex;
infile "fedex_earnings2010.txt" delimiter="09"x missover firstobs=2;
title 'Fedex Quarterly earnings';
input earn date yyq6.;
format date yyq6.;
run;
```

R code

```
# IMPORT DATA
myd=read.table("fedex_earnings2010.txt",header=T)
```

```
# quarterly earnings variable
earn = myd[,1]
#CREATE TIME SERIES OBJECT "ts" of quarterly earnings
ts = ts(earn, start=c(2000, 1) , frequency=4)
```

- a) Analyze the distribution of the quarterly earnings. Are the data normally distributed? Provide appropriate statistics, tests and graphs to support your conclusions.
- b) Is there any evidence of a seasonal effect in the data? Explain your answer.
- c) Is the quarterly earnings time series stationary? Provide statistics, or graphs to support your conclusions.
- d) Compute the Dickey-Fuller test for the quarterly earnings for $p=1, 3$, and 5 . Write down the test hypotheses, and the test statistics and analyze the results of the test
- e) Take the first difference of the Fedex earnings, and analyze its autocorrelation function. Is the differenced time series stationary?
- f) Is you de-trend and de-seasonalize the fedex quarterly earnings, do you obtain a stationary time series?
- g) Apply an airline model $SARIMA(0,1,1)(0,1,1)_4$ to estimate the Fedex quarterly earnings.
 - a. Write down the estimated expression of the fitted model
 - b. Are all model coefficients significant?
 - c. Analyze the residuals to check the model validity
- h) Apply a seasonal model $SARIMA(0,1,0)(0,1,1)_4$ with no $MA(1)$ regular component to estimate the Fedex quarterly earnings. (If you are not sure how to do this, post a message on the discussion forum – I'll give you some help!)
 - a. Write down the estimated expression of the fitted model
 - b. Are all model coefficients significant?
 - c. Analyze the residuals to check the model validity
- i) Which model is a better fit for the data? Explain what evidence you use to select the best model.
- j) Compute forecasts for Fedex quarterly earnings for the next 4 quarters (2011Q1-Q4) using the selected model.
- k) The released quarterly earnings for Fedex in the first two quarters of the 2011 fiscal year are as follows: 2011Q1: \$1.20, 2011Q2: \$0.89. Compare the forecasts of your selected model with the actual earnings. Do the prediction intervals contain the actual values?

Problem 3 [10 pts]

The file `ndx_0512.csv` contains (NDX) Nasdaq index closing prices and log returns from January 2005 to December 2012. The data file contains dates (date), daily prices (price) and log returns (rtn). (Use the SAS or R code for S&P500 index data analyses used as examples in week 7). Analyze the log returns following the steps below.

1. Is there evidence of serial correlations in the log returns? Use autocorrelations and 5% significance level to answer the question.
2. If there is serial correlation in the log returns, fit an appropriate AR model.
3. Is there evidence of ARCH effects in the log returns? Use appropriate tests at 5% significance level to answer this question.
4. Analyze the PACF of the squared log returns, and describe patterns.
5. Fit a GARCH(1,1) model for the NDX log returns using a normal distribution for the innovations. Perform model checking and write down the fitted model.
6. Find a GARCH(1,1) model for the NDX log returns using a the t-distribution for the innovations. Perform model checking and write down the fitted model.

7. What model provides the best fit for the data? Explain.
8. Use the selected model to compute up to 7 step-ahead forecasts of the simple returns and its volatility.

“Reflection” Problem [2 pts]

Post a message on the discussion board reflecting on the topics in hwork 4. 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 “Hwork 4 Comments”.

Submission instructions

Submit the homework at the Course Web page <http://d2l.depaul.edu>.

1. Keep a copy of all your submissions!
2. Submit the SAS &R code along with your answers. You can zip the SAS/R files with your word document.
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 three days after the deadline will not be accepted.