Homework Assignment #3

Due Date: April 29 (campus class) and April 30 (Weekend class)
Note: All tests are based on the 5% significance level.

Remark: The goal of this assignment is to provide students a chance to analyze financial time series from different categories. Real financial time series used in this assignment seem to contain possible outliers. But those outlying data points might be due to jumps and time-varying volatility, two common features of financial data. We shall discuss volatility modeling and heavy-tail distributions next. The number of outliers should reduce substantially after that.

1. Consider the monthly simple returns of the Decile 10 portfolio of CRSP from 1961 to 2014. The data are in the file `m-dec12910-6114.txt`.
   (a) Build a time series model for the return series. You should perform model checking. Write down the fitted model.
   (b) Define a dummy variable for January, i.e.
   \[ I_{t}^{(\text{jan})} = \begin{cases} 
   1 & \text{if } t \text{ corresponds to January} \\
   0 & \text{otherwise} 
   \end{cases} \]
   Build a regression model with time series errors the return series, using \( I_{t}^{(\text{jan})} \) as the explanatory variable. Write down the fitted model.
   (c) Is the regression of part (b) model adequate? Why?
   (d) Compare the models in parts (a) and (b). Which model do you prefer using the in-sample fit? Why?
   (e) Which model do you prefer using out-of-sample predictions? You may answer the question by using the last 60 data points for prediction.

2. Consider the daily VIX index from January 3, 2006 to April 20, 2016. You may use the following command from `quantmod` to download the data
   ```r
   require(quantmod)
   getSymbols('''VIX''',from='''2006-01-03''',to='''2016-04-20''')
   vix <- as.numeric(VIX[,6])
   ```
(a) Build a time series model for the VIX index, including model checking.
(b) Write down the fitted model.
(c) Obtain 1-step to 5-step ahead point predictions of the VIX index at the forecast origin April 20, 2016.

3. Consider the monthly simple returns of short-term bond portfolio from January 1961 to December 2015. The data are obtained from CRSP and in the file \texttt{bond6m-6115.txt} (column 3). For numeric stability, let \( r_t = 100 \times (\text{bond-return}) \).

(a) Build a time series model for the \( r_t \) series, including model checking.
(b) Write down the fitted model.
(c) There are several large residuals in the fitted model. Identify the two largest residuals as outliers. Refine the model by handling the two largest outliers.
(d) Use the refined model to compute 1-step to 4-step ahead point predictions. [You need to provide values of the explanatory variables for prediction. To this end, use the following commands]

\[
\text{xx <- matrix(0,4,2)}
\text{predict(modname,4,newxreg=xx)}
\]

where \texttt{modname} is the name of the model you used in part (c).

4. Consider the daily CDS spreads of JP Morgan from July 20, 2004 to September 19, 2014. The period includes the financial crisis of 2008 so that the CDS spreads vary substantially. The data are in the file \texttt{d-cdsJPM.txt} (column 2). Since the spreads are small, we consider the time series \( x_t = 100 \times (\text{spreads}) \). In addition, sample ACF of \( x_t \) shows strong serial dependence. Therefore, we analyze the differenced series \( y_t = (1 - B) x_t \).

(a) Build a time series model for \( y_t \). Write down the fitted model.
(b) Is the model obtained in part (a) adequate? Why?
(c) To improve the fit, identify sequentially the largest four outliers of the fitted model. Write down the fitted model with the four largest outliers included.
(d) Let \( a_t \) be the residuals of the model in part (c) and \( \rho_i \) be the lag-\( i \) ACF of \( a_t \). Test \( H_0 : \rho_1 = \cdots = \rho_{10} = 0 \) versus \( H_a : \rho_i \neq 0 \) for some \( 1 \leq i \leq 10 \). Draw your conclusion.

5. Consider, again, the daily CDS spreads of JP Morgan in Problem 4. Now, consider \( z_t = \log(x_t) \). Let \( d_t = (1 - B) z_t \).

(a) Build a time series model for \( d_t \). Write down the fitted model.
(b) Is the model obtained in part (a) adequate? Why?
(c) To improve the fit, identify sequentially the largest two outliers of the fitted model. Write down the fitted model with the two largest outliers included.

(d) Let $a_t$ be the residuals of the model in part (c) and $\rho_i$ be the lag-$i$ ACF of $a_t$. Test $H_0 : \rho_1 = \cdots = \rho_{10} = 0$ versus $H_a : \rho_i \neq 0$ for some $1 \leq i \leq 10$. Draw your conclusion.

(e) Comment on the effect of the log transformation by comparing models of Problems 4 and 5.