MSc Mathematical and Computational Finance

Statistics and Financial Data Analysis - Problem Sheet 7 & 8

Please restrict answers to Questions 1 and 2 to 2 pages, remembering to write your main observations and conclusions, using data analysis and graphs to support your findings. This is a combined problem sheet so the deadline for submission will be extended to Thursday 3rd of December mid-day.

1. Load the `'WheatExportData.pkl'` file by importing the pandas library `'import pandas as pd'` and by using the command

   ```
   df = pd.read_pickle('WheatExportData.pkl')
   ```

   This will return the data to a dataframe `df`. This data contains the weekly US wheat export figures. Build a model taking into account potential seasonality, autocorrelation and heteroskedasticity of the data. Give appropriate plots and output to demonstrate your choice of models, justifying parameter choice and why it best fits the data. You may wish to experiment with the function `'seasonal_decompose'` from the `'statsmodels'` package.

2. Load the `'USDJPYVolATM.csv'` file by importing the pandas library `'import pandas as pd'` and by using the command

   ```
   df_vol = pd.read_csv('USDJPYVolATM.csv')
   ```

   This will return the data to a dataframe `df_vol`. This data contains daily at-the-money (ATM) volatility level expressed in % of the USDJPY FX rate for various expiries. Plot the volatility term structure for the dates `'2016-01-04'`,`'2017-12-04'` and `'2018-12-31'`. Analyse the correlation between the changes in volatlity of various expiries and propose a principal component analysis for capturing the moves of the volatility term-structure curve.

3. Consider the model:
   $$X_t = \phi \max(-X_{t-1}, 0) + \sigma Z_t$$
   where Z is a white noise process and ϕ and σ are constants.

   (a) Given a stream of observations X_i's, write a script which will fit this model using regression.

   (b) Write a script which, given X_0, ϕ and σ, simulates a path of X of length k. For $X_0 = 0$, $\sigma = 1$ and $\phi = 2$, discuss the long-term behaviour of this process.

(c) Using your script from part (b), for $\sigma = 1$ and $\phi = 0.9$, discuss the long-term behaviour of X for a range of starting values X_0. Do you think this process is stationary? If so, give an approximation of its unconditional distribution.

(d) Using a simulated path of X as observations, consider the result of fitting an appropriate ARMA model. Describe the distribution of residuals, and how they are related to X_{t-1}.

4. Consider an ARCH(1) model, where

$$\sigma_t^2 = \alpha_0 + \alpha_1 Z_{t-1}^2$$

and

$$Z_t | F_{t-1} \sim N(0, \sigma_t^2)$$

(a) Assuming Z is a stationary process, derive closed form expressions for $\mathbf{E}[Z_t^2 | F_{t-1}]$ and $E[Z_t^4 | F_{t-1}]$, and hence for

$$E[Z_t^2], \qquad E[Z_t^4], \qquad E[Z_t^2 \; Z_{t-1}^2] \quad \text{and} \quad E[Z_t^2 \; Z_{t-2}^2]$$

Explain how this relates to ARCH(1) model with the autorcorrelation function of Z^2.

(b) Give a recursive formula for $\mathbf{E}[\sigma_{t+h}^2 | F_t]$.

(c) Show that

$$Var[\phi Z_{t+1} + Z_{t+2} | F_t] = \phi^2 Var[Z_{t+1} | F_t] + Var[Z_{t+2} | F_t]$$

Are Z_{t+1} and Z_{t+2} independent?

(d) If the mean equation corresponds to an AR(1) process, that is, $X_t = \phi_1 X_{t-1} + Z_t$, show that

$$Var[X_{t+h} | F_t] = \sum_{k=1}^{h} \phi_1^{2(h-k)} E[\sigma_{t+k}^2 | F_t]$$

This can be used to estimate prediction bounds for X. Is X normally distributed?