MSc Mathematical and Computational Finance

Statistics and Financial Data Analysis - Problem Sheet 6

Please restrict answers to Questions 1, 2 and 3 to 2 pages each, remembering to write your main observations and conclusions, using data analysis and graphs to support your findings.

1. For the stocks **GS, GOOG, WMT, BP**, consider the use of ARMA models for the volume traded each day for the period `'2012-01-04'- '2019-11-04'`. You should choose and fit a model for each stock, checking whether the necessary assumptions are satisfied. (You may need to choose not to use all the data available, if there are unusual circumstances for any of these stocks.).

   Extra for experts: For your chosen model for each stock, give forecasts for volumes traded over the coming week, including prediction limits.

2. From the FRED website, download the monthly rate for the 3-Month Treasury Bill (**TB3MS**) from:

   `https://fred.stlouisfed.org/series/TB3MS`

   by saving the data as CSV file. Upload the CSV file into a dataframe, setting the index to be the dates column. Analyse the time series for the period `'2000-01-01'-'2019-01-01'` and propose appropriate time series models which would be suitable for this data, accounting for any heteroskedasticity. Provide clear evidence in the form of data, graphs and statistical tests in order to back up your arguments and conclusions.

3. The one month **LIBOR** rates for **GBP** and **USD** can be found in the FRED website and have been already downloaded onto the CSV files: `'GBPRate.csv'` and `'USDRate.csv'`. Read these files into a dataframe, ensuring you clean the data as it contains non-numerical inputs on some dates. For the following analysis consider only data over the period: `'2015-11-13'-'2020-11-13'`.

   Remember that the relationship between the FX Forward F rate for GBPUSD maturing at time t, the FX 'spot' rate for GBPUSD today S_0 and rates for USD (called the 'domestic' rate) and GBP (called the 'foreign' rate) over a period $dt = (t - t_0)$ (always expressed as a fraction of the year), is defined theoretically as:

   $$F_t = S_{t_0} e^{(r_d - r_f)(t - t_0)}$$

   By downloading the **GBPUSD** data from yahoo finance using the ticker `GBPUSD=X` and **aligning the dates of all datasets** accordingly:

(a) Construct the historic realised 1M carry rate Z_{real} as follows: look up the '30 day ahead' FX spot rate as the realised forward rate in 30 days, $F_{t_{30}}$, by shifting GBPUSD dataframe `df_FX` by 30 days, creating a new dataframe `df_FX_30` using the command:

```
df_FX_30 = df_FX.shift(-30)
```

and hence create the realised carry rate using $dt = (t_{30} - t_0) = 30/365$:

$$Z_{real} = log(F_{t_{30}}/S_{t_0})/dt$$

(b) Create the predicted carry rate using the LIBOR rates for the next 30 days (use as a proxy for the given 1M Libor rates), where r_d is the USD rate and r_f is the GBP rate:

$$Z_{pred} = \tilde{r}_d - \tilde{r}_f$$

(c) Can you find a relation between the Z_{real} and Z_{pred}? And if so, is there a trading strategy you can propose which would take advantage of this relation? Ensure you can back-up your arguments with rigorous data analysis, using statistical tests, graphs and interpretation of results.

4. In time series we often need to update our estimates given new observations. Consider an AR(1) model:

$$(X_t - m) = \phi_1(X_{t-1} - m) + Z_t$$

for $Z \sim WN(0, \sigma^2)$, to be fitted using the method of moments/Yule–Walker approach. It is inefficient to recalculate coefficients directly. By considering the quantities

$$S^T_{xx'} = \sum_{t \leq T} X_t X_{t-1}, \quad S^T_x = \sum_{t \leq T} X_t \quad S^T_{xx} = \sum_{t \leq T} X_t^2$$

find a simple algorithm to iteratively calculate $m^T, \hat{\phi}^T_1, (s^2)^T$ and $\hat{\rho}^T_1$, where m denotes the estimate of m given the first T observations, and similarly for $\hat{\phi}^T_1, (s^2)^T, \hat{\rho}^T_1$.