Requirements for reports

1. The questions should be addressed in the same order they appear in the assignment. The text of the question MUST be retained and placed before each answer. The working language is English.

2. The answer to a particular question may take a form of a plot, formula etc followed by a brief explanation and **a conclusion**. All your conclusions MUST be justified numerically, i.e., by some estimates, values of a test statistic, plots, etc. For example, it is not sufficient to say that the distribution of the returns is not normal, you must refer to some numerical measures, plots etc which prove that. The answers do not need to be lengthy but, again, they MUST be convincing in mathematical and statistical sense, i.e., in terms of some quantitative measures. Note that I pay much attention to the conclusions, so try to make it as clear as possible.

3. Each student MUST use a unique data set. It is your responsibility to make sure that no one else is using the same data. **Check the list on my Google Drive and fill in your name and data you are going to use**. Here is a link to that document https://docs.google.com/document/d/1La7Xs32Yehch5irPh6xed5jje-N2tE1pdkVK4ZhRz7s/edit?usp=sharing.

4. When submitting your report, the subject of your e-mail must be `Quantitative Methods, assignment title: your name`. Otherwise, your report may get lost or not be processed properly and on time!

5. The due date on Linear Models assignment is October 7, 2017.

6. **Late submission**: 25% off for each week after the due date.

7. The answer to a question MUST contain code in `R` or some other language which can be placed in the appendix of the report.

8. Failures to comply with the above rules may reduce your grade for the assignment.

Lecture slides

Slides on linear models (AR, MA, etc) https://www.dropbox.com/s/hkteqo4n3zzry8c/LINEAR%20MODELS.pdf?dl=0

ARCH/GARCH slides https://www.dropbox.com/s/1c4vts3keeq5wge/GARCH.pdf?dl=0

Data sources for the assignment

1. Use `getSymbols` command of `quantmod` package to download prices for some stock or commodity (oil, gold, wheat, etc) from Federal Reserve Economic Data repository http://research.stlouisfed.org/fred2/, Yahoo Finance or Google Finance. You may want to try the following commands if download does not start:

   `options(download.file.method="libcurl")` or `options(download.file.method="wget")` or `options(download.file.method="wininet")`

- Clearly state in your report what kind of data you are using (daily, monthly etc).
- Check for the missing data and remove the respective entries from the dataset, if any. You may use the following script as an example:
```
GOLD=getSymbols('GOLDAMGBD228NLBM', src='FRED', auto.assign=FALSE)
GOLD = na.omit(GOLD)
```
See also Section 1.3.3 of [1].
- If you did find the missing data, add a comment on that.

Assignment on ARMA-GARCH

1. Stationary AR models. Get a time series for some stock using `getSymbols` command:

   (a) Compute and plot the log price x_t and the log return r_t. Comment on the two plots (how volatile the data are, volatility clustering, outliers etc).

   (b) Compute and plot the first 12 lags of ACF of x_t. Comment on the plot. Based on the ACF, is there a unit root in x_t dataset? Why?

   (c) Consider the time series for r_t. Perform the Ljung-Box test for $m = 12$. Draw a conclusion and justify it with the statistical language, i.e., in terms of the critical region or p-value.

   (d) Use the command `ar(`r_t`,method='mle',order.max=20)` to specify the order of an AR model for r_t. Use the PACF and AIC criteria (`ar()` and `pacf()` commands). Compare both approaches.

   (e) Build an AR model for r_t. Plot the time series of the residuals, ACF and p-values of the Ljung-Box test (command `tsdiag()`). Perform the Ljung-Box test of the residuals by hand adjusting the degrees of freedom for the number of the model parameters (see [2], p.66). Is the model adequate? Why? Refine the model by eliminating all estimates with t-ratio less than 1.645 and check the new model as described above. Is the new model adequate? Why? Write down the final model.

   (f) Does the model imply existence of a cycle? Why? If the cycles are present, compute the average length of these cycles.

   (g) Use the fitted AR model to compute 1-step to 4-step ahead forecasts of r_t at the forecast origin corresponding to the last observed date of the time series. Also, compute the corresponding 95% interval. Plot these results.

2. Consider a MA model for r_t:

   (a) Choose the order of such model. Support your choice with the ACF plot.

   (b) Build the model. Refine it by removing coefficients estimates with t-ratio less than 1.645. Write down the fitted model.

   (c) Compute the Ljung-Box statistic of the residuals of the fitted MA model. Is there serial correlation in the residuals? Why?

   (d) Consider the in-sample fits of the AR model of Problem 1 and the MA model. Which model is preferred? Why?

(e) Use `backtest` at some forecast origin with horizon h = 1 to compare the two models. Indicate clearly the parameters of such backtesting (the estimation and forecasting subsamples, forecast origin and so on). Which model is preferred? Why?

3. Yet again, focus on the log return series r_t of the asset from Problem 1. Build an ARMA model including

   (a) Choosing the order of the model,

   (b) Writing down the model,

   (c) Checking the model for adequacy by analyzing the residuals,

   (d) Backtesting and comparing the model with those of Problems 1 and 2.

4. Consider the daily range (daily high minus daily low) of a "blue chip" stock (Apple, Coca-Cola etc.) for the last 4 years. Compute the first 100 lags of ACF of this series. Is there evidence of long-range dependence? Explain! If the range series has long memory, build an AFRIMA model for the data.

5. Consider the log return series r_t of the asset from Problem 1.

   (a) Build an appropriate ARMA model.

   (b) Test the residuals for the ARCH effect.

   (c) Fit an ARMA-GARCH Gaussian model to the data.

   (d) Check the model by analyzing standardized residuals.

   (e) Rebuild and check the model using Student t innovations.

   (f) Build and check an ARMA-APACRH model (order=2).

   (g) Make and plot forecasts based on the above models.

   See examples in [2], Chapter 2.

References

[1] J. Verzani, Using R for Introductory Statistics, Second Edition, Chapman & Hall/CRC The R Series, Taylor & Francis, 2014.
URL https://books.google.ru/books?id=O86uAwAAQBAJ

[2] R. S. Tsay, An Introduction to Analysis of Financial Data with R, 1st Edition, Wiley Publishing, 2012.