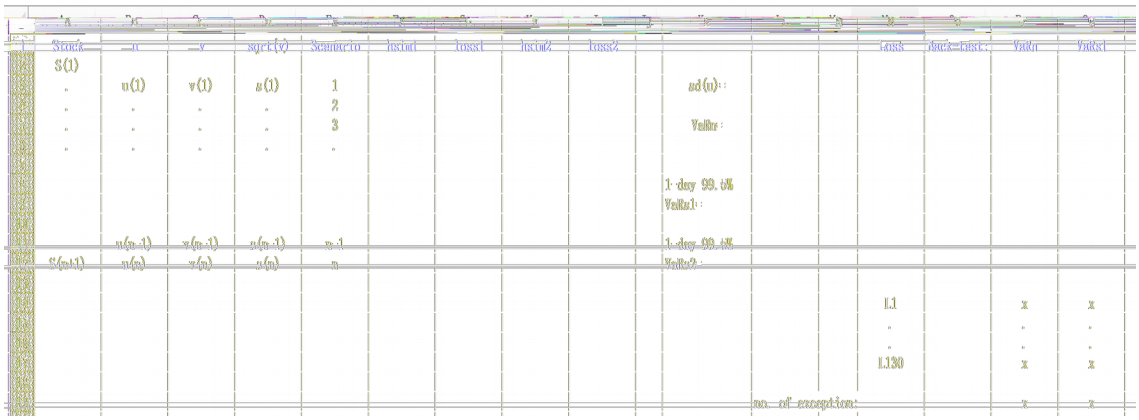
RMSC4002 2017/18 1st term Assignment 2

Q1.

This is a continuation of Q2 of assignment 1. In Q2 of assignment 1, you chose one stock to fit a EWMA model. Copy the Excel sheet containing the stock price, u, *v* and *sqrt(v)* to a new sheet. Suppose we buy 500 shares of this stock on the last day (31/12/2016). We want to compute the 1-day 99.5% VaR using historical simulation and normal model. The following provides a template for this Excel sheet.



(a) Set up the historical simulation hsim1 in column F using the formula $\hat{v}(i) = v_n \times v_i / v_{i-1}$ and column H hsim2 using the formula at the bottom of page 4 in chapter 3, i.e. $\hat{v}(i) = v_n \times \frac{v_{i-1} + (v_i - v_{i-1})\,\sigma_{n+1}/\sigma_i}{v_{i-1}}$, for *i=1,…,n*. Compute the loss for each scenario and sort the loss in column G and I respectively.

(b) Compute the 1-day 99.5% VaRs1 and VaRs2 using hsim1 and hsim2 in L9 and L12.

(c) Compute the sd. of u in L3 and compute the 1-day 99%VaR using normal model in L5.

(d) Compute the actual loss of holding the 500 shares based on the price of the most recent 130 days in column N.

(e) Perform the back-testing and find out the number of exceptions of the VaR in L5, L9 and L12 in columns P to R using the most current 250 days. What is your conclusion from these results?

Q2

The file “credit.csv” contains 7 columns and 690 records of credit application in a bank:

Column	Attribute Information:	name	value
1	Age:	Age	continuous.
2	Mean time at address:	Address	continuous.
3	Mean time with employers:	Employ	continuous.

4	Time with bank:	Bank	continuous.
5	Monthly housing expense:	House	continuous.
6	Savings account balance:	Save	continuous.
7	Result:	Result	binary

The first 6 columns are continuous while the last column is the result of the credit application (1=accept, 0=reject).

(a) Read in the dataset "credit.csv" and save it in d.

(b) Using the last 5 digits of your student ID as random seed, partition d into training dataset d1 and testing dataset d2 as follow:

```
> set.seed(xxxxx)                  # use the last 5 digits of your student id
> id<-sample(1:690,size=600)       # generate random row index for d1
> d1<-d[id,]                       # training dataset
> d2<-d[-id,]                      # testing dataset
```

(c) Using d1 and the glm() function in R, fit a logistic regression of Result on other variables. Exclude the insignificant variables step by step to arrive at a final model and save the output in lreg. Produce the classification table for this logistic regression on the training dataset d1.

(d) Using the predict() function in R to produce the classification table on the testing dataset d2.

(e) Compute and compare the misclassification rate for training and testing dataset.

(f) Produce the lift chart for training dataset d1 as on p.11 of Chapter 5, using cumulative percentage of success vs. the proportion (1:n)/n. (i.e. the second graph on p.11).

(g) Plot a similar lift chart for the testing dataset d2 on the same graph in (f). [Hint: use lines() instead of plot() will add the line on the same graph in (f).]

(h) Compare and comment on these two lines in (f) and (g).

You need to submit three files via eLearning system: asg2-1.xls, asg2-2.r and asg2-2.doc. asg2-1.xls is the Excel file for Q1 and asg2-2.r is the R codes and asg2-2.doc is the output for Q2. The R codes should be **fully commented** as in my notes and ready to execute without bugs. You should also put a hard copy of your asg2-2.doc in Q2 in the drop-box.

The University places very high importance on honesty in academic work submitted by students, and adopts a policy of zero tolerance on cheating and plagiarism. See the following link for details:

http://www.cuhk.edu.hk/policy/academichonesty/

Submit your files on or before November 13, 2017
and put a hard copy in the drop-box on or before November 14, 2017.