Ec 5043 Prof. Le Wang
Managerial Economics II

Homework #2

February 11, 2020

Instruction: Do all the following empirical exercises using R. Turn in your answer with tables and graphs, if any, (along with your program and output files appended at the end of document). Refer to the output file whenever appropriate when discussing your results.

Note that for all simulation exercises, set the seed number to 123456 to ensure the reproducibility of your results.

1. Question 1. [Covariance and Correlation]

Use the data **wage2.csv** for this question.

1. Using R, calculate the covariance and correlation between individual wages (**wage**) and his/her education (**educ**). **This question is corresponding to the definition and the implementation of covariance and correlation on the slides.**
2. Is the correlation in 1) statistically different from zero? Interpret the results. Does this result make sense? **This question is corresponding to test statistics of zero correlation.**
3. **[Optional]** Suppose that individual wages (**wage**) are determined by education (**educ**) and inate ability (**IQ**) as follows

$$wage = a \cdot education + b \cdot IQ$$

Using the covariance formula that we learned from class, can you write the down the covariance between education and wages in terms of the following: **This question is corresponding to the properties of covariance, specifically Property 5.**

(a) The direct relationship between wages and education
(b) The indirect relationship between wages and education through ability (IQ) .

2. Question 2. [Combined Forecast]

Use the data **forecast.csv** for this question. There are threee variables of interset. The original object of interest, y, and forecasts of $\boldsymbol{y}$ from two different models, **forecast1** and **forecast2**. Use what we learned from class (**Application 3: Optimal Forecast Combination**), create a combined forecat that could minimize the variance of forecast error. Does it work? Comment on the reason why it may or may not work.

3. Question 3. [Portfolio Diversification and Nonlinear Dependence]

Download Apple (AAPL) and Blackrock MuniHol (MFL) stocks. These two stocks are negatively correlated with each other. Analyze how you can construct an optimal porfolio by assigning the weights that minimize the risk or variance of the portfolio. The following website has collected all the pair-wise correlations between stocks, and you can easily search all the top 20 stocks that are positively or negatively correlated with a particular stock.

`http://www.market-topology.com/correlation/`

1. Download them using **quantmod in R.** Pick the data after 2018-01-01.
2. Calculate the daily returns for each stock (omit the first observation that is missing).
3. Create the best portfolio with optimal weights that could minimize the risks. **This question is corresponding to Application 2: Portfolio Diversification on the slides.**
4. Using entropy, calculate the dependence between the two stocks. **This question is corresponding to Entropy and Predictability on the slides.**

4. Question 4. [Principal Component Analysis]

Run the following code in R to generate data for our analysis. Then employ the Principal Component Analysis to reduce the dimensions of the data. Explain your process and thoughts.

```
set.seed(123456)
library(MASS)
Sigma <- matrix(c(10,3,4, 3,2,1, 4, 1,2),3,3)
data <- mvrnorm(n = 1000, rep(0, 3), Sigma)
```