MMAT 5310 Midterm

There are 8 questions in this midterm. There are 100 points in total.

1. Refer to the data set data_q1.csv, answer the following questions.
   a. Construct a table with the headers namely, Strategy, N(Buy), Buy > 0, and Buy (the average) as in Brock et al (1992). (5 points)
   b. Plot the returns for each strategy in a single Box Plot. (5 points)
   c. Conduct an ANOVA test with H_0 being that the two strategies have the same population mean return with
   $$r_{i,j} = \mu + \alpha_i + \epsilon_{i,j}.$$
   Write down the conclusion based on the p-value, and find the values of μ (the overall average), α_i. We map α_1 to refer to Strategy A. (5 points)

2. An analyst conducts a test for day-of-the-week effect using the data set data_q2.csv, with the ANOVA model
   $$r_{i,j} = \mu + \alpha_i + \epsilon_{i,j},$$
   where α_i is the day-of-the-week effect. Answer the following questions.
   a. Summarize the mean and standard deviation of the return for each day of the week in a table. (5 points)
   b. Conduct the ANOVA test, then conclude whether the day-of-the-week effect is observed. (5 points)
   c. Predict the daily return on a Tuesday with a 95% confidence interval using the resulting ANOVA model in (c). (5 points)

3. Use the auto_claim data set in the library ExamPAData for this question.
   a. Plot a scatter plot using ggplot for CLM_AMT5 against INCOME. (5 points)
   b. Fit a linear regression model for predicting the mean of CLM_AMT5 with MAX_EDUC being PhD, AGE at 30, and GENDER being female with 95% confidence interval. (5 points)
   c. Fit a logistic model for predicting the probability of at least one claim (based on CLM_FREQ5) for MAX_EDUC being PhD, AGE at 30, and GENDER being female. (5 points)
   d. Based on the model in (c), plot the ROC curve and compare with the random prediction. (5 points)

4. Use the data factor.csv and ret.csv provided for this question.
   a. Regress the excess return on mktrf, smb, and hml with pooled estimates of variance. Test whether the coefficient of hml on IBM is equal to zero at 5% significance level. (5 points)
   b. Consider the factor cma. Find the excess return on this factor (namely the cma return minus the risk free return, rf). Regress this excess return on mktrf, smb, and hml. Conclude whether this factor generates abnormal returns (you may refer to Cremers and Weinbaum 2010). (5 points)

5. Consider the data file return2.csv. Both return and weight values are in percentage. Construct the return attribution using the BHB model. Give a brief description for the results. (10 points)

6. Use the data set portfolio.rds for this question.
   a. On any day, if we index it as t, then the Close column in this file is S_t, CloseDifference is $S_t - S_{t-1}$, and Div is D_t. Find the buy-and-hold return and name it column B where the buy-and-hold daily return is defined as $r_t = (S_t + D_t)/S_{t-1} - 1$. (5 points)
   b. The column M and W are daily returns of two strategies. Use ANOVA model to determine whether the three strategies, namely B, M, and W, have the same daily return. (5 points)

7. In this question, a comparison of trading strategies based on closing prices from the Heikin Ashi candlestick and based on the official closing prices is conducted. In the Heikin Ashi candlestick, the closing price for any trading day is defined to be sum of opening price, closing price, high price, and low price for that trading day divided by four. Use the data set Nikkei.csv which was directly downloaded from Factiva. We test a long-only moving-average strategy as follows. For any trading day, if 5-day moving average is strictly above 150-day moving average, the daily return of the next trading day is assumed to be the daily return of this strategy. Otherwise, this strategy will take zero as daily return for that trading day. Take the following steps as a guide to finish the hypothesis test on whether running this moving-average strategy using the Heikin Ashi closing price is different from using the original closing price. (15 points)

    a. Load the data into R. (The first three lines should be ignored.)
    b. You may make sure the date is using an appropriate data type and select only the Date, Close, Open, High, Low for easier handling.
    c. Define a new column called HClose as in the definition of Heikin Ashi candlestick. (Basically adding the four columns, namely, Close, Open, High, Low, and then divided by four.) Call this column HClose.
    d. Define new columns for 5-day, 150-day moving average for Close/HClose separately. (4 new columns in total) Hint: you may use the rollmean function in zoo, and make sure it is sorted appropriately by column Date.
    e. Define new columns for signals based on Close/HClose separately. (2 new columns in total) You may use 1 to represent buy signals and 0 for neutral signals.
    f. Define a new column for forward daily return on each trading day. You may use log return and use the original closing price (not using Heikin Ashi closing) for this calculation.
    g. Define two new columns for daily returns for the two strategies, namely, using the Heikin Ashi closing price versus using the original closing price to calcuate moving average.
    h. Focus on Date and the two columns defined in step g. Make the data set "tidy" using pivot_longer. Filter out any NA values.
    i. Define two more columns. First, define Year, which is the calendar year of that trading day, and Month, which is the calendar month of that trading day, both columns as factors (using as.factor). Hint: you may consider the year and month functions in lubridate.
    j. Run the ANOVA with the two types of trading strategies as the main treatment and Year and Month as blocking factors.
    k. Interpret the ANOVA results and conclude.

8. We will consider a record (csv files) of cash flows relating to investment in the Inflation-linked bonds due in 2024 (Issue No. 03GB2406R; Stock Code 4246)

    More info: https://www.hkgb.gov.hk/tc/retail/iBond_Rates.html.

    The file hsbc2.csv contains the cash flows of investing this bond off-the-run. Find the daily IRR and annualize based on 365-day daily compounding. (5 points)

End of Midterm