Programming Exercises for Lecture 3
Submission deadline: **May-14-2024**

1. Estimate the Fama-French (FF, Fama and French, 1993; Fama and French, 2015) factor models for Apple stock. The FF factor models can be viewed as an extension of the CAPM and they are also widely used in research of empirical finance. In addition to the risk premium of market portfolio $E\left[R_m - R_f\right]$, the FF factor models include other factors which were considered to have more explanatory power on variation of the cross-section of expected stock return.

   The FF three factor model is given by:

$$E\left[R_i - R_f\right] = \beta E\left[MKT\right] + \beta_{sm} E\left[SMB\right] + \beta_{hl} E\left[HML\right], \tag{1}$$

   where $MKT = R_m - R_f$ is the excess return of market portfolio, SMB is the difference between returns of small cap and large cap companies' stocks, and HML is the difference between returns of value and growth companies' stocks. The FF five factor model is given by:

$$E\left[R_i - R_f\right] = \beta E\left[MKT\right] + \beta_{sm} E\left[SMB\right] + \beta_{hl} E\left[HML\right] + \beta_{rw} E\left[RMW\right] + \beta_{ca} E\left[CMA\right], \tag{2}$$

   where RMW is the difference between stock returns of companies with robust and weak profitability, and CMA is the difference between stock returns of companies with low and high investments. We can estimate factor loadings β, β_{sm}, β_{hl}, β_{rw} and β_{ca} by fitting the following multiple linear regressions:

$$\begin{aligned} R_i - R_f &= \alpha + \beta MKT + \beta_{sb} SMB + \beta_{hl} HML + u, & (3)\\ R_i - R_f &= \alpha + \beta MKT + \beta_{sb} SMB + \beta_{hl} HML + \beta_{rw} RMW + \beta_{ca} CMA + u. & (4)\end{aligned}$$

   a. Import `data_exercise_3c.csv` with `pandas` function `read_csv`. The data include monthly adjusted close price of Apple stock (`AAPL_Adj_Close`), and monthly MKT (`Mkt_RF`), SMB, HML, RMW, CMA, the momentum factor MOM and monthly U.S. risk free interest rate (`RF`), from April-2011 to Feb-2022. Data of these factors and monthly U.S. risk free interest rate are shown in percentage (%).

   b. Transfer dates in column `Date` to ISO-8601 format with method `to_datetime`. Calculate monthly simple and excess returns of Apple stock (with adjusted close prices). Assign the monthly simple and excess returns in new columns called `AAPL_sr` and `AAPL_rp`. Notice that the first element in these new columns should be `nan`.

c. Calculate summary statistics of simple returns and excess returns of Apple stocks and these factors.

d. Use the excess returns calculated from the monthly adjusted close price of Apple to estimate the CAPM and print the estimation results. What are the estimates of α and β?

e. Use the excess returns calculated from the monthly adjusted close price of Apple to estimate the FF three factor model with the linear regression model of (3) and print the estimation results. What are the estimates of α and factor loadings?

f. Use the excess return calculated from the monthly adjusted close price of Apple to estimate the FF five factor model with the linear regression model of (4) and print the estimation results. What are the estimates of α and factor loadings?

g. Momentum factor MOM (Carhart, 1997), the difference between returns of winner and loser stocks, is also often considered as an important variable on explaining variation of the cross-section of expected stock return. Adding this factor into (2), the FF five factor model becomes the Fama-French-Carhart six factor model. Estimate α and factor loadings of this model.

h. For the CAPM and factor models, sometimes we pay more attentions on the estimate of α than the estimates of factor loadings. The risk premium predicted by the CAPM or factor models in (1) and (2) can be viewed as a benchmark. The estimate of α is the risk premium beyond such a benchmark and often is used as a performance measure of the stock. The higher the estimate of α, the more the possibility that the stock can beat the market.

   Compare the estimate of α in d. to g., what do you find?

2. Use data `wage2` for this problem, and estimate linear regressions with an *intercept*.

   a. Run a simple regression of `IQ` on `educ` to obtain the estimated slope $\tilde{\delta}_1$ and intercept $\tilde{\delta}_0$.

   b. Run the simple regression of `log(wage)` on `educ`, and obtain the estimated slope $\tilde{\beta}_1$ and intercept $\tilde{\beta}_0$.

   c. Run the multiple regression of `log(wage)` on `educ` and `IQ`, and obtain the estimated slopes $\hat{\beta}_1$ and $\hat{\beta}_2$ and intercept $\hat{\beta}_0$, respectively.

d. Verify that $\tilde{\beta}_1 = \hat{\beta}_1 + \hat{\beta}_2\tilde{\delta}_1$.

e. Verify whether the result is applied to the *intercept term*, i.e., whether $\tilde{\beta}_0 = \hat{\beta}_0 + \hat{\beta}_2\tilde{\delta}_0$ also holds.

3. Consider using data `wage1` and the following regression:

$$\texttt{log(wage)} = \beta_0 + \beta_1\texttt{educ} + \beta_2\texttt{exper} + \beta_3\texttt{tenure} + u \tag{5}$$

Confirm the partialling out interpretation of the OLS estimates for β_1 by explicitly doing the partialling out with the data. This first requires regressing `educ` on `exper` and `tenure` and saving the residuals $\hat{r}_1$. Then, regress `log(wage)` on $\hat{r}_1$. Compare estimate of the coefficient for $\hat{r}_1$ with that for β_1 in (5). Are they the same? Is this result applied to estimating the *intercept term* β_0?