Problem Set #1
Business 35150
Winter 2024

The due date for this assignment is January 20, 8:30am (upload on canvas)

Write up the solution clearly, using tables and graphs where appropriate. For example, if the question asks you to calculate or compute some numbers, report these numbers in a table or some other form that clearly displays the results. Do not just copy-paste the unformatted output from running your code. In addition to showing the results, you must also explain and describe your results in words. **Submit everything in one pdf file**. Write your name, and if you worked in a group, the names of **all group members**, on the top of the first page.

This problem set is about the choice of the optimal share of risky assets in a portfolio that maximizes expected utility and the performance of the mean-variance approximation that we discussed in class. The python script you will work with contains a barebones version of the core of optimal asset allocation engines in standard portfolio management and robo-advisor software.

The data file `PS1data.xlsx` posted with this problem set contains a time-series of monthly returns on Treasury bills (`rf`) and monthly returns on the CRSP value-weighted stock index (`rvwind`). Please save this file on your computer in the same directory as the python script (`PS1_2024.ipynb`) that I posted.

1. Have a look at the python script and try to understand what it does. You are welcome to use ChatGPT to get help in understanding the python language. (To illustrate, I posted an example, asking ChatGPT to explain one line of the script, at the end of this problem set.) Then answer the following questions:

   (a) Cell #5: which variable stores the stock index return, which variable stores the T-bill yield?

   (b) Cell #5: Which statistical distribution do the simulated returns (variable `ret`) follow?

   (c) Cell #6: What is the purpose of the multiplication by `12` of the mean and variance estimates?

   (d) Cell #7: What is the purpose of the multiplication of mean and variance by `12/hor`? Why is it not multiplication by `12` as in cell #6? Does it make a difference for the optimal portfolio share calculated in Cell #7 by which factor mean and variance are multiplied?

(e) Cell #8: What do the `W`, `U`, and `meanU` variables in cell #8 represent?

2. Now run all cells (with parameters in cell #3 set to their default values) and look at the output. Specifically, look at the optimal portfolio weight that you get in cells #6 and #7 with 1-month and with long-horizon returns. Why are they different? Possible reasons for the discrepancy include

   - Long-horizon investors need to choose a different portfolio than investors with a short horizon, even when returns are IID?
   - Returns in the data file are not in line with the IID assumption underlying the approximate optimal portfolio formula?
   - The approximation error in the approximate optimal portfolio formula (that is present even in IID data)?

   To find out what the main reason for the discrepancy is, do the following two additional analyses:

   - Look at the numerical evaluation of expected utility done in cell #8. How does the optimal portfolio share according to this analysis compare with those you obtained in cells #6 and #7?
   - Re-run everything with simulated data (i.e., set `sim = 1`).

   Based on all these analyses together, provide an explanation for the discrepancy. (A good way to proceed may be to think about which explanations you can rule out based on all results together).

3. Now start a new Jupyter notebook. For this last problem, you will be able to reuse bits and pieces of the `PS1_2024.ipynb` notebook, but you will have to figure out which ones to copy over to your new notebook. For part of the task, you will also need a few lines of new code. I recommend asking ChatGPT for a suggestion on how to fill in the missing piece (calculation of VaR). Here is the task: Based on the entire historical data in `PS1data.xlsx`, estimate the one-month 95% VaR of a portfolio that invests 40% into Treasury bills and 60% into the stock market index (the portfolio is rebalanced to 40/60 at the end of each month). Explain what this number means. (In addition to your answer, please copy-paste your python code into the document.)

4. Bonus question (optional): Perform the VaR estimation of the previous questions in 5-year rolling windows (instead of the full historical data set) and plot the result. Again, ChatGPT will be useful to get a suggestion on how to implement a rolling 5-year VaR calculation and how to plot the results.

ST

what does this python code do: compW = 1 + comprf + (compret - comprf) @ wmat

The Python code you provided performs the following operations:

1. `1 + comprf`: Adds the value of the variable `comprf` to 1.
2. `(compret - comprf)`: Calculates the difference between the values of the variables `compret` and `comprf`. This subtracts `comprf` from `compret`.
3. `@ wmat`: Performs a matrix multiplication between the result of step 2 (the difference between `compret` and `comprf`) and the matrix `wmat`. This operation uses the `@` operator for matrix multiplication, which is typically available in Python when working with NumPy arrays or other libraries that support matrix operations.

The final result of these operations is assigned to the variable `compW`.

In summary, `compW` is calculated by adding 1 to the value of `comprf`, subtracting `comprf` from `compret`, and then performing matrix multiplication between the result of the subtraction and the matrix `wmat`.

Rege