Problem Set #3
Business 35150
Winter 2024

The due date for this assignment is February 17, 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.

The data file `PS3data1.xlsx` contains monthly time-series of stock index (`rvwind`), a Treasury bond portfolio return (`rtreas`), and T-bill returns (`rf`). Please save the data files in the same directory as the python script (`PS3_2024.ipynb`) that I posted.

In this problem, we investigate the empirical properties of portfolios constructed using the rule

$$\omega_t = \frac{\mathbb{E}_t[r_{t+1}] - r_{f,t} + \frac{1}{2}\sigma^2}{\gamma\sigma^2} \tag{1}$$

that we discussed in the lecture. Recall that the case $\gamma = 1$ is the growth-optimal portfolio. Based on this portfolio weight, we calculate log portfolio returns every month using the approximation we introduced in the lecture

$$r_{p,t+1} \approx r_{f,t} + \omega_t(r_{t+1} - r_{f,t}) + \frac{1}{2}\omega_t(1-\omega_t)\sigma^2. \tag{2}$$

Compared to the lecture notes, I made one modification here: a time-varying risk-free rate. The log risk-free rate $r_{f,t}$ is the T-bill log return during month $t+1$, i.e, in the spreadsheet in the same row as the risky asset return r_{t+1}. The reason why it has a t subscript and not $t+1$ in the formula above is that the return is already known at the end of period t (it's the yield of the T-bill at that time) even though the return is realized in period $t+1$.

We are interested in what happens to the wealth of an investor who starts at the beginning of initial period with log wealth of zero (= wealth level of \$1) and invests in the portfolio according to the corresponding rule for ω_t in (1) above.

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. Answer the following questions:

(a) To estimate the numerator in the expression for the portfolio weight ω_t in equation (1) above, we assume that stock index returns are conditionally log-normally distributed, so $\mathbb{E}_t[r_{t+1}] - r_{f,t} + \frac{1}{2}\sigma^2 = \log\{\mathbb{E}_t[\exp(r_{t+1} - r_{f,t})]\}$. Moreover, we assume that $\mathbb{E}_t[\exp(r_{t+1} - r_{f,t})]$ is constant. This means that we can use the log of the sample average of $\exp(r_{t+1} - r_{f,t})$ over the relevant time span in the data set as an estimate of $\mathbb{E}_t[r_{t+1}] - r_{f,t} + \frac{1}{2}\sigma^2$. There are two locations in the code where this calculation takes place. Where? Copy-paste the relevant lines from the code into your answer.

(b) Cell #6: `w = np.tile(m.reshape(-1, 1), (1, G)) / np.tile(gamma*v, (T-burn, 1))`. What is this line doing?

(c) Cell #6: Suppose the settings are `oos = True`. If you replaced `m = m[burn-1:-1]` with `m = m[burn:]`, would the portfolio return constructed in the next cell still be "out-of-sample"?

(d) Cell #9: The figure created in this cell plots the values in the array `cumlogrp`, which is the cumulative log return, yet the y-axis label of the figure says "log wealth." Is the y-axis mislabeled or is this correct? Explain why.

2. Now run all cells in the script up to cell #9 with parameters set for in-sample analysis, i.e., `oos = False` (or run all cells, but just ignore the output of cell #10 for now). Run it first for stocks (`stocks = True`), save the results, then for bonds (`stocks = False`).

   (a) Look at the figures that plot the time-series of log wealth. Focus on the ending values of log wealth for each value of γ. Is the ranking at the end the result what you would have expected? Why?

   (b) Look at figures that show the portfolio weights for different γ values. What does it mean when portfolio weights $\omega_t > 1$? How would you describe an investment strategy with this property?

3. The investment strategies examined in the in-sample analysis in question (2.) could not have been implemented by any actual investor because we constructed the portfolio weight ω_t with an estimate of the numerator in this ratio that used the full time-series of data until the end to construct the sample average of $\exp(r_{t+1} - r_{f,t})$. Obviously, an investor in the 1950s, in the early parts of the data set, would not have had this information. (The variance σ^2 is also estimated based on full sample information, but this is more minor issue because variances generally can be estimated more precisely than expected returns. We'll ignore this issue for the variance.) Let's therefore do an out-of-sample analysis that constructs portfolios in a more realistic way. Rerun the script (again up to cell #9) with parameters now set for out-of-sample analysis (set `oos = True`), again for stocks and then for bonds.

(a) Look the plots of log wealth levels. Focus at the values in the final time period. Is the ranking in terms of γ the same or different from what you found in the in-sample analysis? If there are differences, what's your interpretation of these differences?

(b) Is there an investment strategy (i.e., a value of γ) that you can recommend universally for any investor, at any point in time? If yes, why? If not, why not?

(c) Compare the ending levels of log wealth for the bond portfolios with the highest ending wealth levels with those for the stock portfolios with the highest ending wealth levels. To get more perspective on this, compute the average stock index and bond portfolio return. Looking at these average returns and then comparing the ending levels of log wealth for bonds and stocks, is there anything surprising here? What is the explanation?

4. Now we turn to cell #10 to examine the role of the portfolio log return approximation that we used in cell #7. As we discussed in class, this approximation implicitly assumes that the investor continuously rebalances the portfolio within each month at extremely high frequency to have it always at the weight ω_t. Let's now instead assume that the investor rebalances only once per month. We calculate the weights ω_t exactly as you did in in the in-sample analysis in (1.) using stock index returns. But instead of using the portfolio return approximation, we now calculate the portfolio return as

$$R_{p,t+1} = R_{f,t} + \omega_t(R_{t+1} - R_{f,t}). \tag{3}$$

Note that these are all simple, not log returns, in this expression. We then compound these returns over time. Have a look at the code in cell #10 and make sure you understand how the above portfolio return calculation and compounding is implemented.

(a) Now set `oos = True` and run all the entire script, including cell #10, for stocks (not bonds). Do you encounter any problems or weird results in the output of cell #10? (e.g., look at the log wealth plot produced by cell #10.)

(b) What is going on that causes the problem? [Hint: modify the code to calculate the minimum monthly return corresponding to each `gamma` value in the array `rpx`.] Is this just a technical problem in numerical calculations or is this a real financial problem that the investor faces?