Midterm Exam Instructions

Be verbose. Explain clearly your reasoning, methods, and results in your written work.
No code is necessary, but including it in your answer could result in partial credit.

Written answers are worth the amounts stated.
Total points available: 120 (includes 20 points of extra credit).

General Guidelines

1. **Format:** Answers should be formatted as a PDF. You may convert your Python notebook, if you use one, directly to PDF. Alternatively, use an editor like Word and print to PDF
2. **Restate the question:** Include the question number and restate each question before your answer.
3. **Submission:** When finished, upload your PDF to Canvas along with your code (optional).
4. **Repository:** Do not check code or answers into your repository until after the exam is completed by all.
5. **Data:** Data for problems are available in the repository named by the question number.
6. **Permitted resources:**
    - You may use your notes and the internet **for coding syntax help only.**
    - **You may not use an LLM** – turn off any coding helper you may have installed.
    - **You may not work with other students** – all work must be your own.
7. **Honor Code:** All students will be held to the **Duke Community Standard.**

Duke's Community Standard:

Duke University is a community dedicated to scholarship, leadership, and service and to the principles of honesty, fairness, respect, and accountability. Citizens of this community commit to reflect upon and uphold these principles in all academic and non-academic endeavors, and to protect and promote a culture of integrity.
To uphold the Duke Community Standard:

1. I will not lie, cheat, or steal in my academic endeavors,
2. I will conduct myself honorably in all my endeavors; and
3. I will act if the Standard is compromised.

Midterm Exam Questions

1. (20 pts)

Using `problem1.csv`:

a. (2) Calculate the mean, standard deviation, skewness, and kurtosis of the data.

b. (3) Given a choice between a normal distribution and a t-distribution, which one would you choose to model the data? Why?

c. (4) Fit both distributions and prove or disprove your choice in part (b).

d. (4) Calculate the 5% and 1% VaR and ES for each distribution.

e. (3) Calculate the 5% and 1% VaR and ES for the data using historical simulation.

f. (4) Compare the results in parts (d) and (e). How does this line up with your choice in part (b) and results in part (c)?

---

2. (20 pts)

Using `problem2.csv`:

You and your team have done research into the speed at which correlations and variances change through time. You have found that the speed of change is different for correlations versus variances. Correlations are slower moving but variances update faster.

a. (5) Given that you have decided to use an exponentially weighted correlation and variance estimator, and will use a different lambda for each, should you choose a higher or lower lambda for the correlation estimator? Why?

b. (5) Given your choice in part (a), and possible λ values of 0.94 and 0.97, calculate the exponentially weighted correlation for the input data.

c. (5) Given your choice in part (a), and possible λ values of 0.94 and 0.97, calculate the exponentially weighted variance for the input data.

d. (5) Combine the results in a final covariance matrix.

---

3. (20 pts)

Using `problem3.csv`:

You are given the input covariance matrix and need to use it for risk analysis.

a. (5) Calculate the eigenvalues of the covariance matrix. What do you see?

b. (5) Calculate the nearest PSD matrix using Higham's algorithm. How does this change the eigenvalues?

c. (5) Calculate the nearest PSD matrix using the Near PSD method of Rebonato and Jackel. How does this change the eigenvalues?

d. (5) Compare the results in parts (b) and (c). Which method do you prefer and why?

---

4. (60 pts)

Using the price data in `problem4.csv`:

You hold a portfolio with a current value of $100,000.

a. (5) Calculate the number of shares (fractions of shares are OK) of each of these stocks in your portfolio so that each stock has an equal weight.

b. (10) Calculate the daily returns of each stock using arithmetic returns. Show the first 5 rows and last 5 rows of the return data.

c. (15) Remove the mean from each series. Fit both a normal and a t-distribution to the returns of each stock. Show the parameters of best fit for each stock.

d. (30) Calculate the 1% VaR and ES for each stock and the portfolio as $ values. Use a Gaussian copula to model the dependence structure between the stocks. Calculate the 1% VaR and ES using a historical simulation as well. Present the results in a table and compare the results. Which method do you prefer and why?