IEOR-E4709: Data Analysis for Financial Engineering

Instructor: Dr. Agostino Capponi

Instructor: Dr. Agostino Capponi (who am I?)

Textbook: No required textbook. Recommended textbooks are listed in the syllabus. The book by John Rice is a good source for the first part of the course (MLE, hypothesis testing, confidence intervals). Montgomery et al. is a good reference for linear regression. The books by Tsay and Hamilton are good textbooks for the time series analysis part. The lecture notes by John Cochrane on Time series for macroeconomics and Finance are an excellent resource for time series analysis. All these resources are good complements to the material covered in class. However, a thorough knowledge and understanding of my lecture notes would suffice to perform well in the class.

Course will be administered on Courseworks. All teaching material will be posted on Courseworks (keep a close eye on Courseworks daily). The syllabus is also on Courseworks (please read it carefully).

Attendance is mandatory. My lecture notes will be posted on Courseworks but they are a complement, rather than a substitute, of class attendance.

Grade distribution:

- Homework (25%). Will cover class material. Biweekly homework assignments. No late homework accepted. Collaboration allowed, but write your solution independently.
- Midterm Exam (35%). Closed book, closed notes, no collaboration allowed.
- Final Exam (40%). Closed book, closed notes, no collaboration allowed. It will be cumulative and include all material covered since the beginning of the semester.

Final Grade:

Based on total weighted points earned during semester: 89% → A-, 79%→ B-, 69%→ C- (guaranteed). Discretion will be applied and final grade may depend on the overall performance of the class.

Course topics: (see syllabus). Feel free to ask questions in class.

LECTURE 1: Data Analysis for Financial Engineering

- What is Data Analysis for Financial Engineering?

  Study of quantitative methods to make statistical inferences from data, provide low-dimensional representation of data, and analyze time series of data. It develops mathematical and statistical tools to accomplish these tasks with a focus on financial data (stock returns, transaction costs, bond prices, etc…).

  Aims at answering questions such as

  - How to perform point estimation, for instance how to estimate the mean and variance of a stock price? How to perform interval estimation and provide a range of values which contains the average stock return with high probability?
  - Can we estimate the average returns of US hedge funds?
  - How to describe the dynamics of stock volatility over time?
  - How to model the dynamics of a stock price, taking into account its dependence on past prices, and which econometric models can be best used to account for correlation between volatility and asset prices?
  - Can we filter out noise from prices? How to use stochastic filtering techniques to estimate hidden states from noisy observations?

- Course objectives: Develop analytical tools to

  - Understand the pros and cons of different estimators for the mean and variance of the stock price process (maximum likelihood, methods of moments).
  - Test hypotheses such as "the return process of stock prices is Gaussian with zero mean", or "the return price of a stock price comes from a *t*-distribution with *n* degrees of freedom".
  - Use principal component analysis to extract the leading risk factors of interest rate models or stock prices in the S&P500.
  - Analyze financial time series, and use statistical tools to estimate autocorrelation and intertemporal dependence.
  - Find the model which best fits a time series of stock price data, and use it to price options.
  - Construct parsimonious linear models, in which a set of explanatory variables (stock prices, stock volatility, GDP, etc…) is used to predict default risk of an institution.
  - Kalman filtering techniques for estimating hidden states from noisy signals.

- Which is the target audience?

    - Engineers and applied mathematicians who want to develop sophisticated quantitative skills for solving data analysis problems arising in the financial industry.
    - Students interested in pursuing Wall-Street careers in risk management, data science, or quantitative trading
- Ph.D. students interested in research. This class would be foundational.