STATS 211 - Fall 2025, Session 2
Final Project

1 Logistics

This project is worth 20% of your course grade. You should work in groups (**three students, if possible**), starting no later than Week 3 of the session.

IMPORTANT: I strongly recommend to regularly discuss with the instructors the progress you are making. Do not start working on the project towards the end of the session!

2 Project overview

The project consists in modeling financial data (e.g. market index, currency exchange rates, stock prices, etc.) using a t distribution with unknown parameters. The parameters have to be estimated using Bayesian methods. In particular, suitable priors should be implemented and the posterior distributions for each parameter have to be sampled using a Metropolis within Gibbs sampler.
Once fitted the model to the data, two risk measures have to be estimated: Value at Risk (VaR) and Expected Shortfall (ES).

Deliverables After finishing the project, you are expected to submit a report file (each team only needs to submit one copy) and prepare a presentation recording on Zoom:

- Report (required, but not assessed): Generally, in this report, you should clearly describe the contribution of each team member to the project, explain the methodology implemented (i.e. sampler, VaR and ES, etc.), present the data and the results of the analysis, as well as conclusions and insights.
- Presentation slides: You will submit the slides for the presentation. In the slides, each team should clearly present their work, including the problem being solved, the ideas and solutions proposed to solve the problem, the data results got from the experiments, etc.
- Recording link of Presentation on Zoom: Each team will prepare a recorded zoom presentation. In the presentation, each team member will present his/her own piece of work. The recommended length of the presentation for each team is 10-15 minutes. Please send over your shared zoom recording link to the instructor by email, together with two above files (i.e. the report and the slides).

Timeline

- Week 2: teams are finalized and the project is released
- Week 8: the report, the slides and the recording link have to be submitted via email. The exact deadline will be announced around Week 6.

3 Project details

Title: Estimation of Value-at-Risk and Expected Shortfall Using Metropolis within Gibbs Sampling Method

Project outline: This project will give you a hands-on experience with statistical modeling, risk assessment, and programming, which are valuable skills in many fields.

Objective: Estimate the Value-at-Risk and Expected Shortfall of a market index (or any other suitable financial data) using Metropolis within Gibbs to model the index returns as a t distribution.

Data: To obtain a suitable data set, follow the below steps:

- Choose a market index (e.g., S&P 500, Dow Jones). Alternatively, other financial data can be chosen (e.g. stock closing prices, exchange rates, etc.).
- Collect historical daily closing prices (3 years of data).
- Calculate daily returns and transform them into log returns.

Note: data can be downloaded from several websites, such as Yahoo Finance, Google Finance, Investing.com and Quandl.

Steps

1. Understand the Concepts
    - Review and understand the basics of Metropolis within Gibbs sampling, VaR, and ES.
    - Understand how to model data with a t distribution.
2. Data Preparation
    - Import and clean the market index data.
    - Calculate the daily returns to be used in the analysis.
3. Metropolis within Gibbs Sampling for Parameter Estimation
    - Implement the Metropolis within Gibbs sampling in R (or Python or other software) to estimate the three parameters of the t distribution of index returns.
    - Ensure proper convergence diagnostics.
4. VaR Calculation
    - Use the estimated parameters to calculate the VaR for the index.

5. ES Calculation
    - Use the estimated parameters to calculate the ES for the index.
6. Report
    - Document the methodology, implementation, and results.
    - Discuss the findings and any implications.

Guidance

- Understand each step: Make sure to understand how Metropolis, Gibbs and Metropolis within Gibbs sampling works and how it is applied to estimate the parameters of a t distribution. The key insight, often called the "trick" or "hierarchy", is to recognize that a t distribution can be expressed as a scale mixture of Normal distributions. Thus, a random variable, say Y, following a t distribution, can be written as

$$\begin{aligned} Y|\lambda &\sim N(\mu, \sigma^2/\lambda) \\ \lambda &\sim \text{Gamma}(\nu/2, \nu/2), \end{aligned}$$

  so that $Y \sim t_\nu(\mu, \sigma^2)$, obtained integrating out the latent variable λ.
- Check assumptions: Verify the assumptions of the models and methods you are using and understand the implications if they are violated.
- Validate your model: Use diagnostic plots to check the convergence of the sampler.
- Interpret results carefully: Understand what the VaR and ES are telling you about the risk of the index.
- Do not limit your output to point estimates, but give measures of uncertainty as well (i.e. credible or confidence intervals).

Some references

- *Bayesian Data Analysis*, by Andrew Gelman, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin, 2013 Chapman and Hall/CRC, New York, 3rd Edition.
- *Monte Carlo Statistical Methods*, by Christian Robert and George Casella, 1999 Springer Science, Media New York, 1st Edition.
- *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan,* by John K. Kruschke, 2015, Academic Press / Elsevier, 2nd Edition.
- *Econometric Modeling of Value-at-Risk*, by T. Angelidis and S. Degiannakis, 2009, Nova Science Publisher, Inc.

In addition to the above readings, students are strongly encouraged to access online resources to build their knowledge on relevant statistical, computational and risk management concepts.

Note: The use of AI is strongly encouraged. However, if you do so, you will need to declare in a statement what have you used it for in the project.