Lund University
Centre for Mathematical Sciences
Mathematical Statistics

# Financial Statistics, fall 2023

## Examination – Project

There are two different tracks in this examination project:

**A**: Solve assignment 1a, 2 & 3

**B**: Solve assignment 1b, 2 & 4

The solutions to the assignments are to be documented in a written report and orally presented at a (zoom) seminar. There will be **four** seminars, alternating between track **A** and **B** respectively. The time for the seminars will be published on the Canvas course page.

Your written report must be submitted via email the day before the seminar (at least 24h in advance), and we also need you to sign up in the Canvas calender at the seminar of your preference. Please removing your booking ASAP if you cannot participate!

The project can be solved alone or (preferably) working in groups of two (but not more than two). Groups are expected to do more as their joint work capacity is far greater than that of a student working alone. *Sharing of solutions between groups is strictly prohibited.* The University regulations[1] makes it clear that both taking advantage of the work of another group as well as providing another group with your solutions are considered a reportable offence. Any questions should therefore be posed directly to (and only to) the examiner.

The written report shall include a description of what you have done, discussing choices made, models used etc. as well as figures presenting the result and code (in a separate zip file). Data can be downloaded from the canvas page, but you are free to download additional data if you find those data relevant.

Remember to keep an open mind when solving the problems!

Erik Lindström email: `erik.lindstrom@matstat.lu.se`

**Good luck!**

[1] `https://www.medarbetarwebben.lu.se/forska-och-utbilda/stod-till-utbildning/disciplinarenden`

## 1 GAS models, track A & B (10p)

We saw during the lectures on the Generalized Autoregressive Score framework that it can recover the Gaussian GARCH(1,1) volatility model when assuming that returns are Gaussian with time varying variance.

**Track A. Derive the Student-$t$ GAS model**, assuming that the innovations are Student-$t$ distributed with time varying variance.

**Track B. Derive dynamic correlation matrices.** Another generalization is to consider the case of a zero mean, unit variance bivariate Gaussian distribution.

$$r_t \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right)$$

Derive an expression for a updating the correlation parameter $\rho$ as new observations become available.

## 2 Calibration of implied volatility, track A & B (20p)

You are asked to estimate the (implied) volatility from European Call options. We assume that the model of interest is the standard Black & Scholes model (for more information, see Chapter 9 in the course book).

The issue you are facing is that the liquidity is poor, and you have at most two available contracts at any time. Hence, you are asked to use (=compare) the *non-linear least squares* and the *non-linear filtering* approaches considered during the lectures.

You have at time $t$ observations of the call option(s), $i = 1, \ldots$

$$c_t^{\text{Market}}(S_t, K_{i,t}, \tau_t) = c_t^{\text{B\&S}}(S_t, K_{i,t}, \tau_t, r_f, \sigma_t^2) + \eta_{i,t}$$

where $K_{i,t}$ are the Strikes at time $t$, $\tau$ is the time to maturity, $r_f = 0.04$ the risk free rate and $\sigma_t^2$ the unknown volatility.

Your task is to estimate $\sigma_t^2$, $t = 1, \ldots, T$, motivating any assumptions you use in the modelling process. Present your results numerically and graphically.

The data is found in `ProjectOptionData.mat`

## 3 Modelling using SDEs, Track A, (30p)

News tend to focus on doom and gloom, but much in the world is constantly improving. One example of this is the return of bluefin tuna (lat: **Thunnus thynnus**) in the Öresund Strait during late spring, summer and early fall. They feed on garfish (lat: **Belone belone**) and mackerel (lat: Scomber scombrus) during the summer months as can be seen in Figure 1.

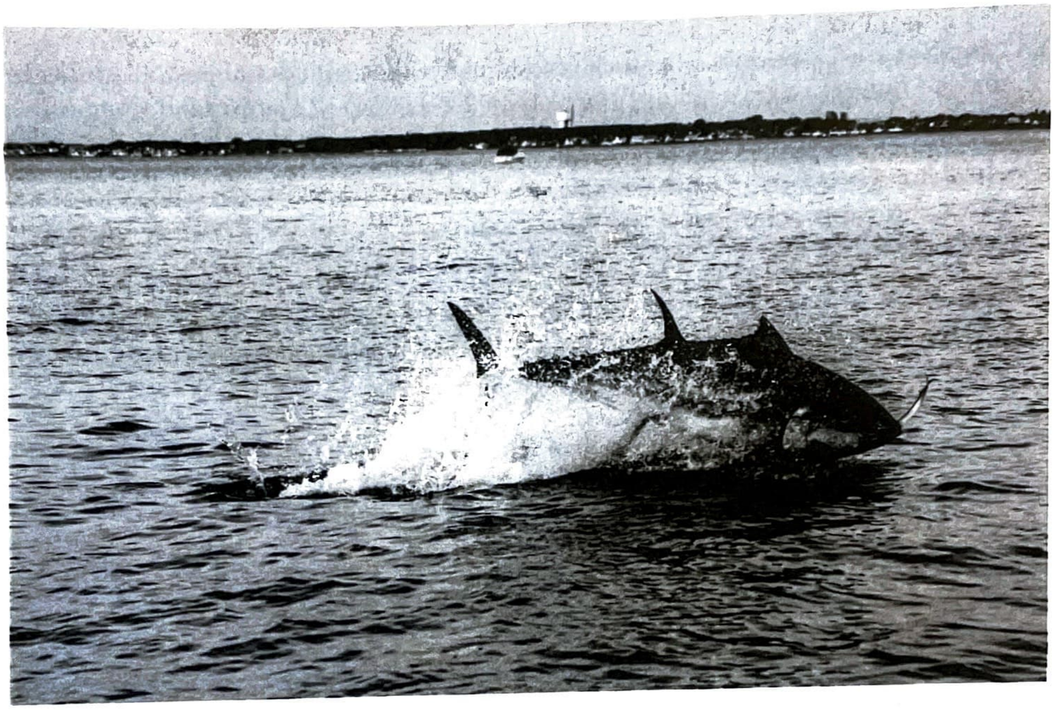

Figure 1: Bluefin tuna chasing garfish above and below the surface in Öresund, summer 2023.

Fit a Lotka-Volterra model to the tuna $\{T\}$ and garfish $\{G\}$ data provided. Assume that the sampling is done on a biyearly basis $\Delta t = 6/12$, as the number of fish are recorded during migration to and away from the Öresund strait. The stochastic Lotka-Volterra model is given by

$$dG_t = (\alpha G_t - \beta G_t T_t)\,dt + \sigma_G G_t dW_t^{(G)} \tag{1}$$
$$dT_t = (\delta G_t T_t - \gamma T_t)\,dt + \sigma_T T_t dW_t^{(T)} \tag{2}$$

Fit the model parameters using the data in `TunaGarfish.mat`. The *first* seminar on Track A will use the data series $G_1$ & $T_1$, the *second* seminar on Track A will use the data series $G_2$ & $T_2$.

Make sure to validate that the parameters you have obtained are reasonable.

## 4 Portfolio Optimization, Track B (30p)

Portfolio allocation is fundamental task in finance. You have data from January 1st, 2019 to December 4th, 2023.

The task is to balance three separate asset classes: **stocks, bitcoins** and **investment grade corporate bonds**.

- NYA The Nasdaq Composite is a stock market index that includes almost all stocks listed on the Nasdaq stock exchange (more than 2500 stocks). The Nasdaq Composite is a capitalization-weighted index, `https://en.wikipedia.org/wiki/Nasdaq_Composite`. The data is workdays only.
- BTC-USD This is bitcoin prices quoted in USD. The data is recorded for every weekday, not only workdays.
- LQD The iShares iBoxx $ Investment Grade Corporate Bond ETF seeks to track the investment results of an index composed of U.S. dollar-denominated, investment grade corporate bonds. The data is workdays only.

The data is found in `NYA.csv`, `BTC-USD.csv` and `LQD.csv`.

You can read and plot the data in `Matlab` by using the commands

```
BTCtable=readtable('BTC-USD.csv')
NASDAQtable=readtable('NYA.csv')
LQDtable=readtable('LQD.csv')
```

Next, we need to find those days where prices exist for all three asset classes

```
% Find common dates
[Dates,IDates]=intersect(BTCtable.Date,NASDAQtable.Date);
```

Next, we compute prices and corresponding returns. Finally, we plot the returns

```
Prices=[NASDAQtable.Close LQDtable.Close BTCtable.Close(IDates)];
Returns=log(Prices(2:end,:))-log(Prices(1:end-1,:));

figure
Asset={'NASDAQ','LQD','BTC'}
for d=1:3
  subplot(3,1,d)
  plot(NASDAQtable.Date(2:end),Returns(:,d))
  title(Asset{d})
end
```

This results in the following Figure

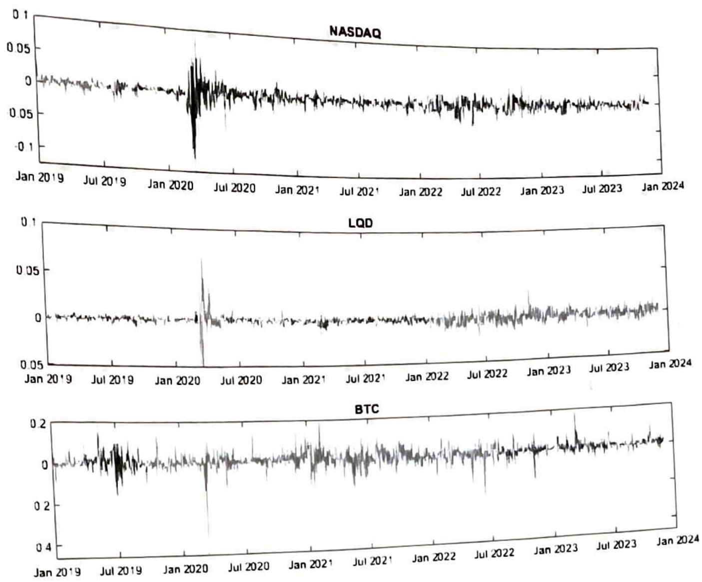


Figure 2: Asset prices from 2019 to 2023.

## Modelling task

We are interested in solving the Markowitz mean variance portfolio allocation problem

$$\max_{\omega \in \Omega} \mu^T \omega - \frac{\gamma}{2} \omega^T \Sigma^{-1} \omega \tag{3}$$

$$s.t.\ \mathbf{1}^T \omega = 1 \tag{4}$$

where $\mu$ is the conditional expected value and $\Sigma$ is the conditional covariance and $\Omega$ is the set of feasible strategies.

**Finding $\gamma$** Start by estimating the unconditional mean and covariance, and use those to to find a reasonable value for risk aversion parameter $\gamma$. The exact value is not important, but the process of finding it is.

You can use $\gamma = 2$ or $\gamma = 3$ henceforth if you cannot find a strategy for selecting a reasonable value for the risk aversion parameter.

**Statistical modelling** Use the data from 2019-2021 to find a suitable model for the asset classes. Document the modelling choices made.

**Evaluation** Test your statical model on the remaining data set by using it within the Markowitz mean variance framework. Record the performance of the portfolio, and compare it to a "buy and hold" strategy derived by fitting a static mean and covariance to the data.

Feel free to modify the Markowitz framework any way you see fit.