Finance haute fréquence : outils probabilistes, modélisation statistique à travers les échelles et problèmes de trading

Simon Depaule

Novembre 2023

Table des matières

1 Chapitre 1 - What is a good price model for high frequency data

1.1 Order Book and Price Terminologies

Order book : The **Order Book** is a tool that consolidates all the buy and sell orders for a financial asset (e.g., a stock or a derivative) at a given moment. It consists of:

- **Bid Prices :** Prices at which buyers are willing to purchase the asset.
- **Ask Prices :** Prices at which sellers are willing to sell the asset.

Each order includes a **quantity**, representing the volume desired by the buyer or seller. The order book reflects supply and demand, facilitating price discovery.

Different Prices :

1. **Last Traded Price:** The most recent price at which an order was executed on the market.
2. **Mid-Quote Price:** The average between the **best bid** (highest buy price) and the **best ask** (lowest sell price):

$$\text{Mid-Quote} = \frac{\text{Best Bid} + \text{Best Ask}}{2}.$$

   It represents the theoretical price equilibrium between buyers and sellers.

3. **Best Bid:** The highest price at which a buyer is willing to purchase the asset, representing the most aggressive demand.
4. **Best Ask:** The lowest price at which a seller is willing to sell the asset, representing the most competitive offer.
5. **Volume Weighted Average Price (VWAP):** The average price of an asset over a specific period, weighted by the volume traded:

$$\text{VWAP} = \frac{\sum_{i=1}^{n}(P_i \times V_i)}{\sum_{i=1}^{n} V_i},$$

   where P_i is the price of transaction i, and V_i is the corresponding traded volume. VWAP is often used as a benchmark to assess trade execution quality.

Durations : Order durations refer to the length of time an order remains active before it is executed, modified, or canceled. Common durations include:

- **Day Order:** The order remains valid until the end of the trading day if not executed.
- **Good Till Cancelled (GTC):** The order remains active indefinitely until canceled by the investor or executed.
- **Immediate or Cancel (IOC):** The order must be executed immediately, at least partially. Any unexecuted portion is canceled.
- **Fill or Kill (FOK):** The order must be fully executed immediately. If not, the entire order is canceled.
- **Good Till Date (GTD):** The order remains active until a specific date specified by the investor.

1.2 Scale of resolution for models

Why the Brownian Motion (Macroscopic Scale) ?

- Statistical Observation : Empirical data from financial markets shows that price movements exhibit characteristics similar to a Brownian trajectory
- Theoretical Reason : For there to be Absence of Arbitrage Opportunities (AOA) prices must be modeled as semimartingales

Microstructure effects :

- At the microscopic scale, prices are very different from Brownian type sample paths.
- This phenomenon is called **microstructure effect**.
- Therefore, it is quite intricate to model such data.

Levels of resolution for models : We distinguish four levels of resolution for modeling through scales. A good probabilistic/statistical model should provide reasonable dynamics across these levels.

- **L0 - The ultimate level of the order book :** A complex stochastic system in continuous time with discrete values in an appropriate state space which describes all the events of a limit **order book.**
- **L1 - Ultra high frequency level for the price :** One wishes to model all **transaction prices and durations** between these transactions.
- **L2 - Intermediate High Frequency Level :** Here one does not focus on durations but essentially on the **price**, regularly sampled, for example every second or every minute.
- **L3 - Macroscopic Level :** The price is viewed as a continuous semi-martingale.

Remark :

- Most models focus on a stability between levels L2 and L3 only.
- Another issue is the question of the modeled price.

1.3 From the large scales to the fine scales

Idea : We want to estimate the volatility. Normally, it works well to use the sum of squared returns. But this doesn't work with high-frequency data ⇒ Signature Plot

Signature Plot : In the analysis of empirical data, a usual tool is the **signature plot**

- Assume we have price observations P_t at times $t = i/n$, $n \in \mathbb{N}$, $i = 0, \ldots, n$, where $t = 1$ represents, for example, one trading day.
- The signature plot is the function which associates to $k = 1, \ldots, n$ the value:

$$RV_n(k) = \sum_{i=0}^{\lfloor n/k \rfloor - 1} \left(P_{k(i+1)/n} - P_{ki/n} \right)^2 .$$

  - P_t : Price process observed at discrete times
  - k : Subsampling frequency (i.e time intervals of 1 minute, 5 minutes..)
  - $RV_n(k)$: Realized volatility calculated over subsampled intervals of size k
- If P_t is a continuous semi-martingale, as soon as n/k is large enough, $RV_n(k)$ is close to the quadratic variation of the semi-martingale.
- However, in practice, $RV_n(k)$ is very often a decreasing functional, stabilizing for sampling periods larger than 10 minutes (depending on the asset, of course).
- The signature plot allows us to observe how realized volatility behaves as a function of the sampling frequency k

High frequency vs low frequency :

- High frequency prices time series are not of the same nature as low frequency series i.e the scale invariance assumption associated to Brownian type dynamics is not reasonable above some sampling frequency.
- **Goal of the coarse to fine approach to microstructure :** Reconcile these different behaviors across scales starting from the coarse scale, that is the scale where a continuous semi-martingale type dynamics is a relevant model.
  - One starts from a continuous semi-martingale, called efficient (latent) price, and apply a stochastic mechanism to it in order to derive the observed prices at higher frequencies.
  - This mechanism must enable to obtain a suitable dynamics for the observed prices in the high frequencies and its effect has to be negligible in the low frequencies, so that the observed price at the macroscopic scale remains close to the efficient price.

1.4 From L3 to L2 models

1.4.1 Additive microstructure noise approach

Additive microstructure noise approach : We will add some noise to the macrostructure model.

- The **microstructure noise** is defined for all t where the model price exists by

$$\varepsilon_t := \log P_t - \log X_t$$

  where

  - P_t is the model price at time t
  - X_t is the semi-martingale used in order to build the model

- More precisely, in such model, one write the log price P_t observed at time $t = i/n$ as

$$\log P_{i/n} = \log X_{i/n} + \varepsilon_i^n$$

  Such model is said to be additive, if the noise term is centered, stationary and independent of X.

  - **Pros :** This type of approach is the simplest way to obtain data close to observations of a semi-martingale in the low frequencies, and data very different from observations of a semi-martingale in the high frequencies, the noise term becoming predominant.
  - **Cons :** It models an unobservable quantity (the noise)

Properties wanted for L1 :

- **(i)** Observed prices are discrete. Indeed, market prices usually stay on a **tick grid**
- **(ii)** For many assets, there are quick oscillations of the transaction prices between two values : **bid-ask bounce**
- **(iii)** $(\log P)_t$ is almost surely finite since $(P_t)_{t\geq 0}$ is a jump process so that the number of jumps is finite over every bounded time interval

No L2 → L1 stability in additive microstructure noise :

- Additive microstructure noise models are very convenient to carry on computations, see later, and are often reasonable when considering sampling scales larger than about 5 minutes.
- However, they do not satisfy any of the properties (i), (ii), (iii) and the durations are not modeled. Consequently, they cannot be extended to level L1.

1.4.2 Rounding models

Idea :

- Rounding models are a simple way to accommodate properties (i), (ii), (iii) and the assumption of an underlying semi-martingale efficient price.
- Here the idea is not to focus on the properties of the microstructure noise but directly on the properties of the observed price.
- Thus, it is very natural to consider the model of continuous semi-martingale observed with rounding error.

Rouding models :

- In this model, the **efficient price** is modeled by a **continuous semi-martingale** X_t and the **observed prices** are given by the sample
$$\left(X_{i/n}^{(\alpha_n)},\, i = 0, \ldots, n\right),$$
where
$$X_{i/n}^{(\alpha_n)} = \alpha_n \left\lfloor \frac{X_{i/n}}{\alpha_n} \right\rfloor .$$
- Therefore, $X_{i/n}^{(\alpha_n)}$ is the observation of the efficient price at time i/n, with **rounding error** α_n, corresponding to the **tick size.**
- In this context, prices are discrete and behave as observations from a continuous semi-martingale in the low frequencies, the rounding effect becoming negligible. Furthermore, Properties (ii) and (iii) can be satisfied.

Pros and cons :

- Hence, compared to additive microstructure noise models, rounding models have several nice properties.
- However, the main drawback of these models is that, as additive microstructure noise models, they remain $L3 \to L2$ models and **cannot be extended to level $L1$:** Assume that for any time t, the observed price is given by the rounding value of a semi-martingale. This leads to an observed price with an infinite number of jumps on a finite interval, which is of course hardly acceptable.
- Therefore, the already mentioned drawbacks of $L3 \to L2$ models apply in the rounding case.

1.5 From L3 to L1 models

Suitable properties for a L3 $\to$ L1 model :

- A continuous semi-martingale type behavior at large sampling scales.
- Model for prices and durations (in particular, no notion of sampling frequency is required).
- A clear definition of the price.
- Discrete prices
- Bid-Ask bounce
- Usual stylized facts of returns, durations and volatility. In particular, inverse relation between durations and volatility.
- An interpretation of the model.
- Finite quadratic variation for the price.
- A testable model.
- A useful model, for example for building statistical procedures.

2 Chapter 2 - Optimal Portfolio Liquidation

2.1 Introduction

Problematic : We want to sell a large quantity of a stock (or of several stocks) in one day. How to choose a good way to split this large order in time and volume ?

Naive strategies :

- **Strategy 1 :** Sell everything right now $\Rightarrow$ huge transaction cost due to the influence on the order book. However this cost is known.
- **Strategy 2 :** Sell regularly in the day small amounts of assets $\Rightarrow$ small transaction costs (volumes are much smaller) but the final profit is unknown because of the daily price fluctuations (Volatility risk).

$\Rightarrow$ We need to optimize between transaction costs and volatility risk. $\Rightarrow$ the Almgren and Chriss framework which takes into account the **market impact phenomenon** and emphasizes the importance of having good statistical estimators of market parameters.

2.2 Almgren and Chriss model

Setup :

- We consider we are selling one asset. We have X shares of this asset at $t_0 = 0$ and we want everything to be sold at $t = T$.
- We split $[0, T]$ into N intervals of length $\tau = T/N$ and set $t_k = k\tau,\ k = 0, \ldots, N$.
- A **trading strategy** is a vector $(x_0, \ldots, x_N)$, with x_k the number of shares we still have at time t_k.
- $x_0 = X,\ x_N = 0$ and $n_k = x_{k-1} - x_k$ is the number of assets sold between t_{k-1} and t_k decided at time t_{k-1}.

Price decomposition : The price we have access can moves because of :

- The drift (negligible at the intraday level)
- The volatility
- **The market impact** can be decomposed into
  - A **permanent** market impact
  - A **Temporary** market impact

Permanent impact compenent :

- Market participants see us selling large quantities $\Rightarrow$ they revise their prices down.
- Therefore, the **equilibrium price** of the asset is modified in a **permanent way**.
- Let S_k be the equilibrium price at time t_k:

$$S_k = S_{k-1} + \sigma\tau^{1/2}\xi_k - \tau g\big(n_k/\tau\big),$$

  with ξ_k i.i.d. standard Gaussian and n_k/τ the average trading rate between t_{k-1} and t_k.

Temporary impact component :

- By selling asset, we are **liquidity takers** since we consume the liquidity available in the order book.
- If we sell a large amount of shares, our price per share is significantly worse than when selling only one share.
- We assume this effect is temporary and the liquidity comes back after each period.
- Let $\tilde{S}_k = \big(\sum n_{k,i}p_i\big)/n_k$, with $n_{k,i}$ the number of shares sold at price p_i between t_{k-1} and t_k. We set

$$\tilde{S}_k = S_{k-1} - h\big(n_k/\tau\big).$$

- The term $h(n_k/\tau)$ does not influence the next equilibrium price S_k.

Remark :

- **Liquidity Providers (Makers) :** These are traders who post limit orders or other standing instructions that remain unexecuted until another party is willing to trade at their quoted price. By placing these orders, they add depth and availability of immediate trades to the market, thus providing liquidity.
- **Liquidity Taker :** Being a liquidity taker means you are the one executing trades against already posted orders, thereby consuming the available liquidity rather than providing it.

Profit and Loss :

- The result of the sale of the asset is

$$\sum_{k=1}^{N} n_k\tilde{S}_k = XS_0 + \sum_{k=1}^{N}(\sigma\tau^{1/2}\xi_k - \tau g(n_k/\tau))x_k - \sum_{k=1}^{N} n_k h(n_k/\tau)$$

- The trading cost $C = XS_0 - \sum_{k=1}^{N} n_k\tilde{S}_k$ is equal to

$$\underbrace{\sum_{k=1}^{N}(\sigma\tau^{1/2}\xi_k)x_k}_{\text{Vol. cost}} + \underbrace{\sum_{k=1}^{N}\tau g(n_k/\tau)x_k}_{\text{Perm. Impact cost}} + \underbrace{\sum_{k=1}^{N} n_k h(n_k/\tau)}_{\text{Temp. Impact cost}}.$$

Mean-Variance Analysis :

- Consider a static strategy (fully known in t_0), which is in fact optimal in this framework. We have

$$\mathbb{E}[C] = \sum_{k=1}^{N} \tau x_k g\big(n_k/\tau\big) \ + \sum_{k=1}^{N} n_k h\big(n_k/\tau\big), \quad \mathrm{Var}[C] = \sigma^2 \sum_{k=1}^{N} \tau x_k^2.$$

- In order to build optimal trading trajectories, we will look for strategies minimizing

$$\mathbb{E}[C] + \lambda\, \mathrm{Var}[C],$$

with λ a **risk aversion parameter.**

2.3 Naive Strategies

Assumptions :

- **Permanent impact :** We will assume a **linear permanent impact** $g(v) = \gamma v$. Hence, if we sell n shares, the price per share decreases by γn. Thus

$$S_k = S_0 + \sigma \sum_{j=1}^{k} \tau^{1/2} \xi_j - \gamma(X - x_k)$$
$$\sum_{k=1}^{N} \tau x_k g\big(n_k/\tau\big) = \gamma \sum_{k=1}^{N} x_k(x_{k-1} - x_k) = \tfrac{1}{2}\gamma X^2 - \tfrac{1}{2}\gamma \sum_{k=1}^{N} n_k^2.$$

- **Temporary impact :** We will assume an **affine temporary impact** $h(n_k/\tau) = \varepsilon + \eta(n_k/\tau)$ where ε represents a fixed cost (fees + bid-ask spread).

Regular strategy (Regular liquidation) :

- Take $n_k = X/N$, $x_k = (N-k)X/N$, for $k = 1, \dots, N$.
- We easily get

$$\mathbb{E}[C] = \tfrac{1}{2}\gamma X^2 + \varepsilon X + \tilde{\eta}\tfrac{X^2}{T},$$
$$\mathrm{Var}[C] = \tfrac{\sigma^2}{3} X^2 T \big(1 - \tfrac{1}{N}\big)\big(1 - \tfrac{1}{2N}\big).$$

- We can show this strategy has the **smallest expectation**. However, **the variance can be very big** if T is large.

Immediate selling : Sell everything at t_0

- Take $n_1 = X$, $n_2 = \cdots = n_N = 0$, $x_1 = \cdots = x_N = 0$.
- We get

$$\mathbb{E}[C] = \varepsilon X + \frac{\eta X^2}{\tau}, \quad \mathrm{Var}[C] = 0.$$

- This strategy has the **smallest variance**. However, if τ is small, the **expectation can be very large.**

2.4 Optimal Strategies

Optimization program : The trader wants to minimize

$$\begin{aligned} U(C) &= \mathbb{E}[C] + \lambda \operatorname{Var}[C] \\ &= \frac{1}{2}\gamma X^2 + \varepsilon X + \frac{\tilde{\eta}}{\tau}\sum_{k=1}^{N}(x_{k-1} - x_k)^2 \; + \; \lambda\sigma^2 \sum_{k=1}^{N} \tau x_k^2. \end{aligned}$$

Solution :

- It is shown that the solution can be written $x_0 = X$ and for $j = 1, \ldots, N$:

$$x_j = \frac{\sinh(K(T - t_j))}{\sinh(KT)} X, \quad n_j = \frac{2\sinh(K\tau/2)}{\sinh(KT)} \cosh(K(T - j\tau + \tau/2))X,$$

  where K satisfies $\frac{2}{\tau^2}(\cosh(K\tau) - 1) = \tilde{K}$.

- If $\lambda = 0$, then $\tilde{K} = K = 0$ and so $n_j = \tau/T = X/N$. We retrieve the strategy with minimal expected cost.

Remarks :

- The solution is time homogeneous : if we compute the optimal strategy in t_k, we obtain the values between t_k and T of the optimal strategy computed in t_0.
- In this approach, we obtain an efficient frontier of trading.
- The optimal trajectories are **very sensitive to the volatility** parameter. It is therefore important to obtain accurate volatility estimates.
- The Almgren and Chriss framework can be extended in dimension n (if we sell several assets). In that case, correlation parameters come into the picture.
- **Trading tactics :** Only the total quantities which have to be executed in each time window are provided by the Almgren-Chriss approach. The way to deal with them inside each window (the trading tactic) is an intricate issue : Should we use market orders only or a combination with limit orders ? When should we trade inside each window ?

3 Chapter 3 : Model with uncertainty zones and associated statistical procedures for volatility, correlation and lead-lag

3.1 Modelling ultra high frequency data

Properties of the microstructure model wanted :

- Model for prices AND durations
- Discrete prices
- Bid-Ask bounce
- Stylized facts of returns, durations and volatility. In particular, inverse relation between durations and volatility.
- A diffusive behavior at large sampling scales.
- No hesitation about the price.
- Finite quadratic variation for the microstructure noise.
- A useful model.

Proposed answer : Model with uncertainty zones. In this model, prices and durations are functionals of some hitting times of an underlying continuous semi-martingale

3.2 Model with uncertainty zones

Aversion for price changes :

- In an idealistic framework, transactions would occur when the **efficient price** crosses the tick grid.
- In practice, **uncertainty** about the efficient price and aversion for price changes of market participants.
- The price changes only when market participants are convinced that the efficient price is far from the last traded price.
- We introduce a parameter η quantifying this **aversion for price changes**.

Notation :

- X_t : Efficient price
- α : tick size, a : ask, b : mid, $m = \frac{a+b}{2}$ midpoint
- t_i : time of the i-th transaction with price change
- P_{t_i} : transaction price at time t_i
- $L_i = \frac{|P_{t_{i+1}} - P_{t_i}|}{\alpha}$: size of the i-th price jump
- Explanatory variables process for the size of the jumps : χ_t
- Uncertainty zones : $U_k = [0, \infty) \times (d_k, u_k)$ with

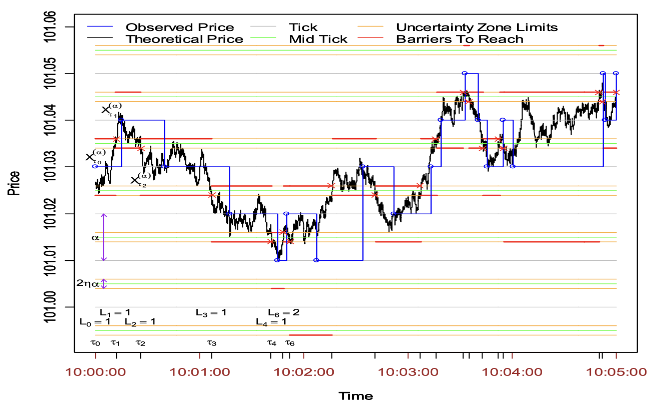
$$d_k = (k + 1/2 - \eta)\alpha \quad \text{and} \quad u_k = (k + 1/2 + \eta)\alpha.$$

- $U = 2\eta\alpha < \alpha$: Uncertainty regeion around the mid-point m
- τ_i : i-th exit time of an uncertainty zone.

Dynamics :

- $d \log X_u = a_u\, du + \sigma_{u_-}\, dW_u$.
- $\mathbb{P}[L_i = s \mid \mathcal{F}_{\tau_i}] = \phi_s(\chi_{\tau_i})$.
- $\tau_{i+1} = \inf\{t > \tau_i : X_t = X_{\tau_i}^{(\alpha)} \pm \alpha(L_i - \frac{1}{2} + \eta)\}$,

  with $X_{\tau_i}^{(\alpha)}$ the value of X_{τ_i} rounded to the nearest multiple of α.
- $\tau_i \leq t_i < \tau_{i+1}$ and $P_{t_i} = X_{\tau_i}^{(\alpha)}$.



Interpretation of η :

- Quantifies the aversion for price changes (with respect to the tick size) of market participants.
- In the UHF, the order book can not "follow" the efficient price and is reluctant to price changes. Reluctancy measured by η.
- $2\eta\alpha$ represents the **implicit spread** of a large tick asset (see later).
- A small $\eta\,(< 1/2)$ means that for market participants, the tick size is too big and conversely.

Estimation of η :

- A **continuation** is a price variation whose direction is the same as the one of the preceding variation.
- An **alternation** is a price variation whose direction is opposite to the one of the preceding variation.
- $N^c = \#$ continuations. $N^a = \#$ alternations.
- Estimator $\hat{\eta}$:

$$\hat{\eta} = \frac{N^c}{2N^a}.$$

Intuitions avout η :

- $2\eta\alpha$ as an implicit unobservable spread.
- If η is too small $\rightarrow$ uncertainty zone is small $\rightarrow$ strong mean reversion of the observed price $\rightarrow$ decreasing signature plot. It means that tick size is too big.
- $\eta \sim \frac{1}{2} \rightarrow$ last traded price can be seen as a sampled Brownian motion $\rightarrow$ no microstructure effects $\rightarrow$ flat signature plot $\rightarrow$ uncertainty zone ≈ 1 tick $\rightarrow$ optimal.

The following linear model describes well the spread:

$$\eta\alpha \sim \frac{\sigma}{\sqrt{M}} + kS$$

where M is the number of trades and σ is a price process volatility.

Predicting consequences of tick value changes :

- α too small encourages free-riding (directional HFT) and traditional market makers cannot fix their quotes.
- α too large implies price sloppiness. Moreover, it favors speed (race to the top of book) $\Rightarrow$ High investments in infrastructure.
- What happens to η if one changes the tick value?
- How to obtain the following **optimal** situation:
  - Spread ~ 1 tick
  - η close to $1/2$
  - Cost of market orders = cost of limit orders = 0.

3.3 Volatility estimation

Retrieving the efficient price : The **efficient price** is given by

$$X_{\tau_i} = P_{t_i} - \alpha\left(\frac{1}{2} - \eta\right)\text{sign}(P_{t_i} - P_{t_{i-1}})$$
$$\widehat{X}_{\tau_i} = P_{t_i} - \alpha\left(\frac{1}{2} - \widehat{\eta}\right)\text{sign}(P_{t_i} - P_{t_{i-1}})$$

where :

- t_i : Observation time
- τ_i : Exit time
- P_t : Observed price
- X_t : Efficient price
- $\hat{\eta}$: Estimator of the aversion for price changes

Remark : We use here an estimation of η. η describes the distribution of high frequency tick returns :

- **Small** : Tick size large. (Uncertainty zone small, which means that we have a strong (average) reversion of the observed price and thus, a decreasing signature plot.)
- $\mu \sim \frac{1}{2}$: Tick size in some sense optimal. (Last traded price can be seen as a Brownian Motion, thus no microstructure effects. Flat signature plot and uncertainty zone = 1 tick.)

Theorem (Estimation of the volatility) : Let

$$\widehat{\text{RV}}_t = \sum_{i=1}^{N_{\alpha,t}} \left(\log(\widehat{X}_{\tau_i}) - \log(\widehat{X}_{\tau_{i-1}})\right)^2.$$

We have

$$\alpha^{-1}\left(\widehat{\text{RV}}_t - \text{RV}_t\right) \xrightarrow{\mathcal{L}_s} \gamma_t \int_0^t \nu_u dW_{\theta_u},$$

where W is a Brownian motion independent of B and θ_u, γ_u and ν_u depend on X_u, σ_u, and explanatory variables, involving for example the order book.

3.4 Covariation estimation

Framework :

- We now consider two correlated assets :

$$\begin{aligned} \mathrm{d}\log X_t &= \mu_t^X \,\mathrm{d}t + \sigma_t^X \,\mathrm{d}W_t \\ \mathrm{d}\log Y_t &= \mu_t^Y \,\mathrm{d}t + \sigma_t^Y \,\mathrm{d}B_t \\ \mathrm{d}\langle W, B\rangle_t &= \rho_t \,\mathrm{d}t \end{aligned}$$

- We want to estimated the integrated covariance

$$\int_0^1 \rho_t \sigma_t^X \sigma_t^Y \, dt$$

- We will face two main difficulties :
  - Asynchronicity of the data
  - Microstructure effects

Usual estimator of the covariation :

- We start with the usual case **without asynchronicity or microstructure effects**. We observe $(X_{i/n}, Y_{i/n})$, $i = 0, \ldots, n$. Let

$$\Delta_i^n X = \log X_{i/n} - \log X_{(i-1)/n}.$$

- An estimator of the integrated covariance with accuracy $n^{-\frac{1}{2}}$ is given by

$$\widehat{c}_n = \sum_{i=1}^{n} \Delta_i^n X \Delta_i^n Y$$

- **Case of constant volatilities and correlation :** When the correlation and volatility parameters are supposed to be constant $\rho_t = \rho, \sigma_t^X = \sigma^X, \sigma_t^Y = \sigma^Y$, an estimator of the correlation ρ with accuracy $n^{-\frac{1}{2}}$ is given by

$$\frac{\widehat{c}_n}{\sqrt{\sum_{i=1}^n (\Delta_i^n X)^2 \sum_{i=1}^n (\Delta_i^n Y)^2}}.$$

- **Case of a constant correlation but non constant volatilities :** In the case where the volatility parameters are no longer constant, one can consider

$$\begin{aligned} \widehat{\rho}_n &= \frac{2}{\pi} \frac{\widehat{c}_n}{\widehat{a}_n} \\ \widehat{a}_n &= \sum_{i=1}^{n-1} \left| \Delta_{i+1}^n X \Delta_i^n Y \right| \end{aligned}$$

  Indeed, $\widehat{a}_n$ is an estimator of

$$\frac{2}{\pi} \int_0^1 \sigma_s^X \sigma_s^Y \, \mathrm{d}s.$$

Previous tick scheme :

- Suppose X is observed at times $(T^{X,i}), i = 1, \ldots$, and Y at times $(T^{Y,i}), i = 1, \ldots$.
- Construct:
$$\overline{X}_t = X_{T^{X,i}}, \text{ for } t \in [T^{X,i}, T^{X,i+1})$$
and
$$\overline{Y}_t = Y_{T^{Y,i}}, \text{ for } t \in [T^{Y,i}, T^{Y,i+1}).$$
- Given h, the previous tick covariation estimator is:
$$V_h = \sum_{i=1}^{m} \left(\log \overline{X}_{ih} - \log \overline{X}_{(i-1)h}\right) \left(\log \overline{Y}_{ih} - \log \overline{Y}_{(i-1)h}\right) \tag{1.1}$$
- **Issue:**
  - For non-synchronous data, if no price change (jump) occurs for X or Y within $](i-1)h, ih[$, this estimator equals zero.
  - The challenge lies in the asynchronicity of X and Y. While interpolation over intervals of length h is a potential solution, the choice of h heavily influences results and introduces significant biases.
  - A new estimator is required, based solely on the original data, without prior synchronization or interpolation. Such an estimator would avoid biases and dependencies on h, ensuring robustness and accuracy.

Remark :

- Eeach increment in (1.1),
$$(P^1_{t_i} - P^1_{t_{i-1}})(P^2_{t_i} - P^2_{t_{i-1}}),$$
contributes to the sum when and only when both P^1 and P^2 'jump' together during the interval $]t_{i-1}, t_i]$ of length h, thus all the other occasions – when at most one of the two prices jumps – are ignored.
- Such occasions of zero increment will become dominant if h becomes *finer*.
- On the other hand, *coarser* h may not be able to capture rapid movements of processes – multiple jumps that may have occurred during h – so that the realized covariance estimator may fail to reflect such microscopic movements (which are crucial for variance–covariance estimation).
- In other words, large h leads to inefficient use of data.

Epps Effect: Empirical observation highlighting the dependence of correlation estimators for high-frequency stock returns on the sampling frequency. Specifically, **correlations tend to decrease as the sampling frequency increases.**

A convergent estimator under asynchronicity (Hayashi-Yoshida Estimator) :

- Let $I_i^X = (T^{X,i}, T^{X,i+1}]$ and $I_j^Y = (T^{Y,j}, T^{Y,j+1}]$.
- The Hayashi-Yoshida estimator is

$$U_n = \sum_{i,j} \Delta X(I_i^X)\Delta Y(I_j^Y)\mathbf{1}_{\{I_i^X \cap I_j^Y \neq \emptyset\}}.$$

- This estimator does not need any selection of h and is convergent if the arrival times are independent from the price.
- Nevertheless, it is not robust to microstructure effects.

Remarks :

- The Hayashi-Yoshida estimator is designed to compute the covariance of two assets when their prices are observed asynchronously (non-synchronous data). This is particularly important in high-frequency finance, where prices are updated irregularly and independently.
- **Non-Synchronous Data :** Unlike simple covariance estimators, the Hayashi-Yoshida estimator accounts for the fact that P^X and P^Y are not observed at the same times. This is done by summing up the products of price changes whenever their observation intervals overlap.
- **Indicator Function:** The indicator $\mathbf{1}_{\{I_i^X \cap I_j^Y \neq \emptyset\}}$ ensures that only overlapping intervals contribute to the covariance estimate.
- **Avoiding Bias:** By using overlapping increments rather than interpolating the missing data, the estimator avoids introducing biases caused by imputation.
- The Hayashi-Yoshida estimator is free of any synchronization of the original data, hence free from any problems caused by it, especially, biases of Epps type. The estimator is shown to have consistency as the observation frequency tends to infinity.

Theorem : In the model with uncertainty zones, the Hayashi-Yoshida estimator is a consistent estimator of the covariation provided one uses the estimated values of the efficient prices.

3.5 Lead-Lag estimation

Observation from practitioners in finance :

- Some assets are leading some other assets.
- This means that a “lagger” asset may partially reproduce the behavior of a “leader” asset.
- This common behavior is unlikely to be instantaneous. It is subject to some time delay called “lead-lag”.

3.5.1 A toy model : The Bachelier Model

Bachelier Model :

- For $t \in [0,1]$, and $(B^{(1)}, B^{(2)})$ such that $\langle B^{(1)}, B^{(2)} \rangle_t = \rho t$, set

$$X_t := x_0 + \sigma_1 B_t^{(1)}, \quad \widetilde{Y}_t := y_0 + \sigma_2 B_t^{(2)},$$

- Define $Y_t := \widetilde{Y}_{t-\theta}$, $t \in [\theta, 1]$. Our **lead-lag model** is given by the bidimensional process (X_t, Y_t).
- We have

$$\begin{cases} X_t = x_0 + \sigma_1 B_t^{(1)} \\ Y_t = y_0 + \rho\sigma_2 B_{t-\theta}^{(1)} + \sigma_2 \sqrt{1-\rho^2} W_{t-\theta} \end{cases}.$$

Intuitive estimator :

- Assume the data arrive at regular and synchronous time stamps in the Bachelier model, *i.e.*, we have data

$$(X_0, Y_0), (X_{\Delta_n}, Y_{\Delta_n}), (X_{2\Delta_n}, Y_{2\Delta_n}), \ldots, (X_1, Y_1),$$

  and suppose $\theta = k_0 \Delta_n$, $k_0 \in \mathbb{Z}$.
- Let

$$C_n(k) := \sum_i \big(X_{i\Delta_n} - X_{(i-1)\Delta_n}\big)\big(Y_{(i+k)\Delta_n} - Y_{(i+k-1)\Delta_n}\big).$$

- Heuristically, we have

$$C_n(k) \approx \Delta_n^{-1} \mathbb{E}\big[(X_\cdot - X_{\cdot-\Delta_n})(Y_{\cdot+k\Delta_n} - Y_{\cdot+(k-1)\Delta_n})\big] + \Delta_n^{1/2} \xi^n.$$

- Moreover,

$$\Delta_n^{-1} \mathbb{E}\big[(X_\cdot - X_{\cdot-\Delta_n})(Y_{\cdot+k\Delta_n} - Y_{\cdot+(k-1)\Delta_n})\big] = \begin{cases} 0 & \text{if } k \neq k_0, \\ \rho\sigma_1\sigma_2 & \text{if } k = k_0. \end{cases}$$

- Thus we can (asymptotically) detect the value k_0 that defines θ in the very special case $\theta = k_0 \Delta_n$ by maximizing in k the contrast sequence

$$k \mapsto |C_n(k)|.$$

3.5.2 The Lead Lag Model

The Lead-Lag model (Assumptions) : Let $\theta > 0$ (for simplicity, extensions are quite straightforward) and set

$$\mathbb{F}^\theta = (\mathcal{F}_t^\theta)_{t\geq 0}, \quad \text{with } \mathcal{F}_t^\theta = \mathcal{F}_{t-\theta}.$$

- We have
$$X = X^c + A, \quad Y = Y^c + B.$$
- $(X_t^c)_{t\geq 0}$ is a continuous $\mathbb{F}^\theta$-local martingale, and $(Y_t^c)_{t\geq 0}$ is a continuous $\mathbb{F}^\theta$-local martingale.
- $\exists v_n \to 0$, $v_n^{-1} \max\big\{\sup |I_n^{X,i}|, \sup |I_n^{Y,i}|\big\} \to 0$.
- The $T^{X,i}$ are $\mathbb{F}^{v_n}$-stopping times and the $T^{Y,i}$ are $\mathbb{F}^{\theta+v_n}$-stopping times.

Estimator :

- We set
$$U_n(\theta) = \sum_{i,j} \Delta X(I_i^X)\Delta Y(I_j^Y)\mathbf{1}_{\{I_i^X \cap (I_j^Y)_{-\theta} \neq \emptyset\}},$$
with $(I_j^Y)_{-\theta} = \big(T^{Y,j} - \theta, T^{Y,j+1} - \theta\big]$.
- Eventually, $\hat{\theta}_n$ is defined as a solution of
$$\big|U_n(\hat{\theta}_n)\big| = \max_{\theta \in \mathcal{G}^n} \big|U_n(\theta)\big|,$$
where $\mathcal{G}^n$ is a sufficiently fine grid.

4 Chapter 4 : Tick values and regulation

4.1 Tick value, tick size and spread

- **Exchange rule :** There exists a price grid for orders.
- **Tick value :** smallest price increment. Dimension: currency of the asset.
  - Subject to changes by the exchange.
  - In some markets, the spacing of the grid can depend on the price.

Notion of tick size :

- In practice: the tick value is given little consideration. What is important is the **tick size**
- **Tick size** qualifies the traders' **aversion** to price movements of one tick.
- In the US, The SEC has required a whole report about the tick size issue in the context of high frequency trading and a special roundtable was devoted to this topic on 5 February 2013.
- In Europe, one of the main conclusions of the Foresight report is the crucial need for proper tick sizes to regulate high frequency trading.

Large tick asset :

- **Large tick stocks** are such that the bid-ask spread is almost always equal to one tick
- **Small tick stocks** have spreads that are typically a few ticks.
- This leads to the following questions :
  - Small tick assets : spread is a good proxy for the tick size.
  - If spread $\simeq$ 1 tick $\Rightarrow$ **How to quantify the tick size ?**
  - Tick value change $\Rightarrow$ **What happens to the microstructure?** Can we define an **optimal tick value?**

How to assess the quality of the tick value ?

- **Signature plot :**

$$RV_n(k) = \sum_{i=1}^{\lfloor n/k \rfloor - 1} \left(P_{k(i+1)/n} - P_{ki/n}\right)^2$$

For $\frac{n}{k}$ large enough it is close to the quadratic variation. However, in practice, it often decreases due to microstructural effects. For a good tick value, the signature plot remains flat, and the process resembles a continuous semimartingale.

Market making strategy :

- **Market makers :** prove liquidity, they send **limit orders** $\Rightarrow$ delayed execution. Pocket the spread. $\exists$ volatility risk.

- **Market takers :** impatient traders. Send **market orders** $\Rightarrow$ immediate execution. **Pay the spread**. No volatility risk.

- **Wyart et al.:** consider a simple market making strategy. Its average P&L per trade is

$$\text{P\&L} = \frac{S}{2} - \frac{c}{2}\sigma_1,$$

with c depending on the assets but of order $1 \sim 2$.

- One can show that :
  - Market and limit orders will have the same average cost equal to zero
  - Market maker's P&L $= 0$

- Therefore:

$$S \sim c\sigma_1.$$

This relationship is very well satisfied on market data.

5 Chapter 6 : Rough Volatility

5.1 Some elements about volatility modeling

Main classes of volatility models :

- Prices are often modeled as **continuous semi-martingales** of the form

$$dP_t = P_t(\mu_t dt + \sigma_t dW_t).$$

- The volatility process σ_s is the most important ingredient of the model. Practitioners consider essentially three classes of volatility models:
  - Deterministic volatility (Black and Scholes 1973),
  - Local volatility (Dupire 1994),
  - Stochastic volatility (Hull and White 1987, Heston 1993, Hagan et al. 2002, ...).
- In terms of regularity, in these models, the volatility is either very smooth or with a smoothness similar to that of a Brownian motion.
- To allow for a wider range of smoothness, we can consider the **fractional Brownian motion** in volatility modeling.

Definition (Fractional Brownian Motion) : The fractional Brownian motion (fBm) with Hurst parameter H is the only process W^H to satisfy:

- **Self-similarity**: $(W^H_{at}) \overset{\mathcal{L}}{=} a^H (W^H_t)$.
- **Stationary increments**: $(W^H_{t+h} - W^H_t) \overset{\mathcal{L}}{=} (W^H_h)$.
- **Gaussian process** with $\mathbb{E}[W^H_1] = 0$ and $\mathbb{E}[(W^H_1)^2] = 1$.

Proposition :

- For all $\varepsilon > 0$, W^H is $(H-\varepsilon)$-Hölder a.s.
- The absolute moments of the increments of the fBm satisfy

$$\mathbb{E}[|W^H_{t+h} - W^H_t|^q] = K_q h^{Hq}.$$

- If $H > 1/2$, the fBm exhibits long memory in the sense that

$$\mathrm{Cov}[W^H_{t+1} - W^H_t, W^H_1] \sim \frac{C}{t^{2-2H}}.$$

Mandelbrot-van Ness representation : We have

$$W^H_t = \int_0^t \frac{dW_s}{(t-s)^{\frac{1}{2}-H}} + \int_{-\infty}^0 \left(\frac{1}{(t-s)^{\frac{1}{2}-H}} - \frac{1}{(-s)^{\frac{1}{2}-H}} \right) dW_s.$$

About option data :

- Classical stochastic volatility models generate reasonable dynamics for the volatility surface.
- However, they do not allow to fit the **volatility surface**, in particular the **term structure of the ATM skew**:

$$\psi(\tau) := \frac{\partial}{\partial k}\sigma_{\text{BS}}(k,\tau)\bigg|_{k=0},$$

  where k is the log-moneyness and τ the maturity of the option.
- **The skew is well-approximated by a power-law function of time to expiry τ. In contrast, conventional stochastic volatility models generate a term structure of ATM skew that is constant for small τ.**
- Models where the volatility is driven by a fBm generate an ATM volatility skew of the form

$$\psi(\tau) \sim \tau^{H-\frac{1}{2}},$$

  at least for small τ.

5.2 Building the Rough FSV model

The RFSV model : These empirical findings suggest we model the log-volatility as a fractional Brownian motion:

$$\sigma_t = \sigma e^{\nu W_t^H}.$$

Multiscaling in finance :

- An important property of volatility time series is their multiscaling behavior, see Mantegna and Stanley (2000) and Bouchaud and Potters (2003).
- This means one observes essentially the same law whatever the time scale.
- In particular, there are periods of high and low market activity at different time scales.
- Very few models reproduce this property, see multifractal models.

5.3 Application of the RFSV model : Volatility prediction

Prediction of a fractional Brownian Motion :

- There is a nice prediction formula for the fractional Brownian motion. For $H < \frac{1}{2}$,

$$\mathbb{E}\left[W^H_{t+\Delta} \middle| \mathcal{F}_t\right] = \frac{\cos(H\pi)}{\pi} \Delta^{H+\frac{1}{2}} \int_{-\infty}^{t} \frac{W^H_s}{(t-s+\Delta)(t-s)^{H+\frac{1}{2}}} \, ds.$$

- We apply the previous formula to the prediction of the log-volatility:

$$\mathbb{E}\left[\log \sigma^2_{t+\Delta} \middle| \mathcal{F}_t\right] = \frac{\cos(H\pi)}{\pi} \Delta^{H+\frac{1}{2}} \int_{-\infty}^{t} \frac{\log \sigma^2_s}{(t-s+\Delta)(t-s)^{H+\frac{1}{2}}} \, ds.$$

- After a simple change of variable, the prediction of the log-volatility can be written:

$$\mathbb{E}\left[\log(\sigma^2_{t+\Delta}) \middle| \mathcal{F}_t\right] \sim \frac{\cos(H\pi)}{\pi} \int_{0}^{1} \frac{\log(\sigma^2_{t-\Delta u})}{(u+1)u^{H+\frac{1}{2}}} \, du.$$

The only time scale that appears in the above regression is the horizon Δ. As it is known by practitioners:

> *If trying to predict volatility one week ahead, one should essentially look at the volatility over the last week. If trying to predict the volatility one month ahead, one should essentially look at the volatility over the last month.*

Proposition (Conditional distribution of the fractional Brownian motion) : In law,

$$W^H_{t+\Delta} | \mathcal{F}_t \sim \mathcal{N}\left(\mathbb{E}\left[W^H_{t+\Delta} | \mathcal{F}_t\right], c\Delta^{2H}\right),$$

with

$$c = \frac{\sin\left(\pi\left(\frac{1}{2} - H\right)\right)\Gamma\left(\frac{3}{2} - H\right)^2}{\pi\left(\frac{1}{2} - H\right)\Gamma\left(2 - 2H\right)}.$$

Prediction of the variance : Therefore, our predictor of the variance writes:

$$\mathbb{E}\left[\sigma^2_{t+\Delta} \middle| \mathcal{F}_t\right] = \exp\left(\mathbb{E}\left[\log\left(\sigma^2_{t+\Delta}\right) \middle| \mathcal{F}_t\right] + 2\nu^2 c \Delta^{2H}\right).$$

5.4 The microstructural foundations of rough volatility

Summary of what we have seen and objectives :

- **We know that: Volatility is rough!**
- On any asset, using any reasonable volatility proxy/statistical method (realized volatility, realized kernels, uncertainty zones, Garman-Klass, implied volatility, power variations, autocorrelations, Whittle, ...), one concludes that volatility is rough.
- It cannot be just coincidence...
- We want to show that typical behaviors of market participants at the high frequency scale naturally lead to rough volatility.
- Our modeling tool : **Hawkes processes**.

Definition (Hawkes Process) :

- A **Hawkes process** $(N_t)_{t\geq 0}$ is a self-exciting point process, whose intensity at time t, denoted by λ_t, is of the form

$$\lambda_t = \mu + \sum_{0<J_i<t} \phi(t - J_i) = \mu + \int_{(0,t)} \phi(t-s)dN_s,$$

where μ is a positive real number, ϕ a regression kernel, and the J_i are the points of the process before time t.

Order flow and volatility :

- Thus, it is nowadays classical to model the order flow (**number of trades**) thanks to Hawkes processes.
- It is known from financial economics theory that the order flow is essentially the same thing as the **integrated volatility** (variance) if the time scale is large enough :

$$N_t \approx \int_0^t \sigma^2(s)ds.$$

Two main reaons for the popularity of Hawkes processes :

- These processes represent a very **natural and tractable extension of Poisson processes**. In fact, comparing point processes and conventional time series, **Poisson processes** are often viewed as the counterpart of **iid random variables**, whereas **Hawkes processes** play the role of **autoregressive processes.**
- Another explanation for the appeal of Hawkes processes is that it is often **easy to give a convincing interpretation to such modeling**. To do so, the branching structure of Hawkes processes is quite helpful.

Poisson cluster representation :

- Under the assumption $\|\phi\|_1 < 1$, where $\|\phi\|_1$ denotes the L^1 norm of ϕ, **Hawkes processes** can be represented as a **population process where migrants arrive according to a Poisson process with parameter μ.**
- Then each migrant gives birth to children according to a non-homogeneous Poisson process with intensity function ϕ, these children also giving birth to children according to the same non-homogeneous Poisson process, see Hawkes (74).
- Now consider, for example, the classical case of buy (or sell) market orders. Then **migrants can be seen as exogenous orders, whereas children are viewed as orders triggered by other orders.**

Stability condition :

- The assumption $\|\phi\|_1 < 1$ is crucial in the study of Hawkes processes.
- If one wants to get a **stationary intensity with finite first moment**, then the condition $\|\phi\|_1 < 1$ is required
- This condition is also necessary in order to obtain **classical ergodic properties** for the process.
- For these reasons, this condition is often called the **stability condition in the Hawkes literature.**

Degree of endogeneity of the market :

- The **degree of endogeneity of the market** is defined by $\|\phi\|_1 < 1$
- The parameter $\|\phi\|_1$ corresponds to the average number of children of an individual.
- $\|\phi\|_1^2$ corresponds to the average number of grandchildren of an individual, ...
- Therefore, if we call cluster the descendants of a migrant, then the **average size of a cluster** is given by

$$\sum_{k\geq 1} \|\phi\|_1^k = \frac{\|\phi\|_1}{1-\|\phi\|_1)}$$

- Thus, the average proportion of endogenously triggered events is

$$\frac{\|\phi\|_1/(1-\|\phi\|_1)}{(1+\|\phi\|_1/(1-\|\phi\|_1))} = \|\phi\|_1$$

Unstable Hawkes processes :

- Our aim is to study the behavior at large time scales of so-called **nearly unstable Hawkes processes**, which correspond to these estimations of $\|\phi\|_1$, close to 1.
- This will give us insights on the properties of the integrated volatility.
- Furthermore, we want to take into account another stylized fact: The function ϕ has typically a power law tail:

$$\phi(x) \sim_{x \to +\infty} \frac{K}{x^{1+\alpha}},$$

  with α of order 0.5–0.7.
- This memory effect is likely due to metaorders splitting.

Microstructural foundations for the RFSV model :

- **It is clearly established that there is a linear relationship between cumulated order flow and integrated variance.**
- **Consequently, the "derivative" of the order flow corresponds to the spot variance.**
- **Thus, endogeneity of the market together with order splitting lead to a superposition effect, which explains (at least partly) the rough nature of the observed volatility.**
- **Near instability together with a tail index $\alpha \sim 0.6$ correspond to $H \sim 0.1$.**
- **In fact, one can show that rough volatility is just a consequence of the no statistical arbitrage principle.**

6 Annales

6.1 Sujet 2023 - 2024

(2 pts) Give the CAPM equations for expectation and variance of returns, providing accurate definitions of all quantities. What does the CAPM tell us about the relationship between risk and returns in finance? How can you prove the validity of CAPM in practice? (Give main mathematical elements without entering in details)

(2 pts) Consider you have access to interest rates curves between 2000 and 2023, one curve every month, with 10 maturity points for each curve, from 1 year to 10 years. Describe how you would analyze the results of a PCA applied to this dataframe and what you expect to see.

(2 pts) Consider two one-dimensional Itô processes X and Y driven by two correlated Brownian motions W and B with constant correlation ρ and (stochastic) volatilities σ_t^X and σ_t^Y. We observe X and Y over $[0, 1]$ at times i/n, $i = 0, \ldots, n$ with $n \geq 0$. Give an estimator of the parameter ρ and provide a sketch of proof of its consistency as n goes to infinity.

(2 pts) How would you demonstrate that rough volatility models are superior to conventional stochastic volatility models? (1 page maximum, several answers are possible)

(2 pts) Using Hawkes processes, explain the connection between market microstructure, market impact, and rough volatility (1 page maximum).

6.2 Sujet 2021 - 2022

1. What is the market portfolio? In the CAPM model, at equilibrium, what is the portfolio of the agents made of? Give the CAPM equation for the expected return of an asset. How can one test statistically for the validity of the CAPM?

2. Define the Ridge estimator and provide its closed form formula (without proof). What is the interest of Ridge estimator compared to ordinary least squares? What does Lasso mean? Define the Lasso estimator. What is the interest of Lasso estimator compared to Ridge? How do we choose the regularization parameter (λ) of the Ridge and Lasso estimators? Give one example of a situation where Ridge or Lasso estimators are useful in finance.

3. In the PCA of a cloud of data points in $\mathbb{R}^p$ given by the $n \times p$ array X, how does one build the best sub-vector space of $\mathbb{R}^p$ with dimension k to project the array?

4. State the Marcenko-Pastur theorem (no need to give the density of the Marcenko-Pastur law, just call it L). How can this result be used in finance?

5. In the Black-Scholes model, how does the option price evolve with the drift? Connect this with statistical estimation of the drift from historical data.

6. You work for a brokerage company, with access to an accurate database about past transactions, order flows, and order books. Your mission is to build an algorithm enabling you to buy in 8 hours 600,000 Hermes stocks and 300,000 LVMH stocks. Which mathematical tools, models, and statistical methods do you use to build it (give mathematical formulas when needed)? How do you implement and calibrate the models? What do you measure on data in the construction process? How do you run some testing procedures? (2 pages maximum!)

7. How do you assess the quality of the tick value for a given asset?

8. From a time series of prices (high frequency data over several years), how would you show that volatility is rough? What are the advantages of rough models compared to classical models used for options? What are the microstructural foundations of rough volatility? Summarize how this can be shown mathematically.

6.3 Sujet 2021 - 2022 Rattrapages

1. State the Markowitz optimization problem and solve it.
2. Define the Ridge estimator and provide its closed-form formula with **proof**. What is the interest of the Ridge estimator compared to ordinary least squares? Define the Lasso estimator. What is the interest of Lasso estimator compared to Ridge? How do we choose the regularization parameter (λ) of the Ridge and Lasso estimators? Give one **example** of a situation where Ridge or Lasso estimators are useful in finance.
3. State the Marcenko-Pastur theorem (no need to give the density of the Marcenko-Pastur law, just call it L). How can this result be used in finance?
4. In the Black-Scholes model, how does the option price evolve with the drift? Connect this with statistical estimation of the drift from historical data.
5. How do you assess the quality of an order book model? Give detailed explanations and examples.
6. How do you assess the quality of the tick value for a given asset?
7. From a time series of prices (high-frequency data over several years), how would you show that volatility is rough? What are the advantages of rough models compared to classical models used for options? What are the microstructural foundations of rough volatility? Summarize how this can be shown mathematically.