Quantitative finance

Interviews preparation

Vivien Tisserand

Abstract

This is a summary of interview questions that I found and digressed on in quantitative finance. It is a mixed of applied mathematics and computer science.
I am not found of brainteasers, they are a poor way to assess for a candidate's hability to be an asset for the teams. This work smoothly transitioned to a sort of *vademecum* in applied mathematics : through several questions, it goes through several techniques that are easy to forget with time. I myself refer to it quite often when I forget about the way to solve an arithmetico-geometric sequence or the general solution of a second-order differential equation.

Contents

Chapter 1

Statistics

1.1 Estimating the support of an uniform law

Suppose that we have $x_1, \ldots, x_n$ observations from an uniform law $X \sim \mathcal{U}[0, \theta]$, where θ is an unknown parameter that we want to estimate. Give at least two estimators for θ and compare them.

- **Method of moments:**

  Having a look at the first order moment, it appears that $\mathbb{E}[X] = \dfrac{\theta}{2}$. Taking the empirical counter-party of this theoretical quantity, we have $\hat{\theta}^{\text{MM}} = \dfrac{2}{n}\sum_{i=1}^{n} x_i$.
  By applying the strong law of large numbers and the continuous mapping theorem, $\hat{\theta}^{\text{MM}} \xrightarrow{a.s.} \theta$. Thus this estimator is consistent.
  We want asymptotic results on the convergence of this estimator. Before using the CLT, we have to check for the existence of a second-order moment.

$$
\begin{aligned}
\mathbb{E}[X^2] &= \int_{\mathbb{R}} x^2 f(x)\,\mathrm{d}x \\
&= \int_0^\theta x^2 \frac{1}{\theta}\,\mathrm{d}x \\
&= \left[\frac{1}{3\theta}x^3\right]_0^\theta \\
&= \frac{\theta^2}{3} < +\infty
\end{aligned}
$$

  Thus, we have $\mathbb{V}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \dfrac{\theta^2}{12}$. So, $\mathbb{V}[2X_1] = \dfrac{\theta^2}{3}$.

  By applying the central limit theorem, we have :

$$
\sqrt{n}(\hat{\theta}^{\text{MM}} - \theta) \xrightarrow{(d)} \mathcal{N}\left(0, \frac{\theta^2}{3}\right).
$$

  We have to evaluate the risk of this estimator, that we write as the sum of the squared bias and the variance :

  $\text{MSE}(\hat{\theta}^{\text{MM}}) = \mathbb{E}[(\hat{\theta}^{\text{MM}} - \theta)^2] = \mathbb{E}[(\hat{\theta}^{\text{MM}} - \mathbb{E}[\hat{\theta}^{\text{MM}}])^2] + \mathbb{E}[\hat{\theta}^{\text{MM}} - \theta]^2 = \mathbb{V}[\hat{\theta}^{\text{MM}}] + (\mathbb{E}[\hat{\theta}^{\text{MM}}] - \theta)^2.$

We have $\mathbb{E}[\hat{\theta}^{\mathrm{MM}}] = 0$ and $\mathbb{V}[\hat{\theta}^{\mathrm{MM}}] = \dfrac{1}{n^2} n \mathbb{V}[2X_1] = \dfrac{\theta^2}{3n}$.

Thus,

$$\mathrm{MSE}(\hat{\theta}^{\mathrm{MM}}) = \frac{\theta^2}{3n}.$$

- **Maximum likelihood:**

  Let's write the likelihood of this model:

$$\begin{aligned} L((X_1, \dots, X_n), \theta) &= \prod_{i=1}^{n} f_X(X_i) \\ &= \prod_{i=1}^{n} \frac{1}{\theta} \mathbb{1}_{[0,\theta]}(X_i) \\ &= \frac{1}{\theta^n} \prod_{i=1}^{n} \mathbb{1}_{[0,\theta]}(X_i). \end{aligned}$$

  And this function is maximized by choosing the smallest θ such that all of the X_i lie in $[0, \theta]$, that is $\hat{\theta}^{\mathrm{MLE}} = \max_{1 \le i \le n} X_i$.

  To check the consistency of this estimator, we will have a look at its convergence (in probability). Let $\theta \in \Theta$ and $\varepsilon > 0$:

$$\begin{aligned} \mathbb{P}_\theta(|\hat{\theta}^{\mathrm{MLE}} - \theta| \ge \varepsilon) &= \mathbb{P}_\theta(\hat{\theta}^{\mathrm{MLE}} \ge \theta + \varepsilon) + \mathbb{P}_\theta(\hat{\theta}^{\mathrm{MLE}} \le \theta - \varepsilon) \\ &= 0 + \mathbb{P}_\theta(\max_{1 \le i \le n} X_i \le \theta - \varepsilon) \\ &= \prod_{i=1}^{n} \mathbb{P}_\theta(X_i \le \theta - \varepsilon) \\ &= \left(1 - \frac{\varepsilon}{\theta}\right)^n \underset{n \to +\infty}{\longrightarrow} 0. \end{aligned}$$

  Thus, $\hat{\theta}^{\mathrm{MLE}} \xrightarrow{\mathbb{P}} \theta$: this estimator is consistent.

  In order to estimate the risk of this estimator, we have to look at the law that the maximum of n independent uniform laws follows. This is done by looking at the cumulative distribution function. Let $x \in [0, \theta]$:

$$\begin{aligned} \mathbb{P}_\theta(X_{(n)} \le x) &= \mathbb{P}_\theta\left(\bigcap_{i=1}^{n} X_i \le x\right) \\ &= \prod_{i=1}^{n} \mathbb{P}_\theta(X_i \le x) \\ &= \left(\frac{x}{\theta}\right)^n. \end{aligned}$$

Thus,

$$F_{X_{(n)}} = \begin{cases} 0 & \text{if } x < 0 \\ \left(\frac{x}{\theta}\right)^n & \text{if } 0 \leq x \leq \theta \\ 1 & \text{if } x > \theta \end{cases}$$

This cdf as smooth as we need to take its derivative: that will be the density we were looking for:

$$f_{X_{(n)}}(x) = n\frac{x^{n-1}}{\theta^n}\mathbb{1}_{[0,\theta]}(x)$$

Let's compute the bias and the variance.

$$\mathbb{E}[\hat{\theta}^{\text{MLE}}] = \int_{\mathbb{R}} x f_{X_{(n)}}(x)\,\mathrm{d}x = \int_0^\theta \frac{n}{\theta^n}x^n\,\mathrm{d}x = \frac{n}{\theta^n}\left[\frac{x^{n+1}}{n+1}\right]_0^\theta = \frac{n}{n+1}\theta.$$

Then, the bias is : $B(\hat{\theta}^{\text{MLE}}) = \frac{n}{n+1}\theta - \theta = -\frac{1}{n+1}\theta \neq 0$. We can introduce a corrected estimator that we will consider too : $\hat{\theta}^{\text{MLE}}_{\text{corr}} = \frac{n+1}{n}\hat{\theta}^{\text{MLE}}$, such that $\mathbb{E}[\hat{\theta}^{\text{MLE}}_{\text{corr}}] = \theta$: an unbiased estimator.

Then, we have

$$\mathbb{E}[(\hat{\theta}^{\text{MLE}})^2] = \int_{\mathbb{R}} x^2 f_{X_{(n)}}(x)\,\mathrm{d}x = \int_0^\theta \frac{n}{\theta^n}x^{n+1}\,\mathrm{d}x = \frac{n}{\theta^n}\left[\frac{x^{n+2}}{n+2}\right]_0^\theta = \frac{n}{n+2}\theta^2.$$

And

$$\text{MSE}(\hat{\theta}^{\text{MLE}}) = \mathbb{E}[(\hat{\theta}^{\text{MLE}} - \theta)^2] = \mathbb{E}[(\hat{\theta}^{\text{MLE}})^2] - 2\theta\mathbb{E}[\hat{\theta}^{\text{MLE}}] + \theta^2$$

Thus,

$$\text{MSE}(\hat{\theta}^{\text{MLE}}) = \frac{n}{n+2}\theta^2 - 2\frac{n}{n+1}\theta^2 + \theta^2 = \frac{2\theta^2}{(n+1)(n+2)}.$$

And

$$\text{MSE}(\hat{\theta}^{\text{MLE}}_{\text{corr}}) = \left(\frac{n+1}{n}\right)^2\mathbb{E}[(\hat{\theta}^{\text{MLE}})^2] - 2\frac{n+1}{n}\theta\mathbb{E}[(\hat{\theta}^{\text{MLE}} + \theta^2 = \frac{\theta^2}{n(n+1)}.$$

- **Maximum a posteriori:**

  We write the likelihood of the model in terms of θ :

$$L((X_1,\dots,X_n),\theta) = \frac{1}{\theta^n}\prod_{i=1}^n \mathbb{1}_{[0,\theta]}(X_i) = \frac{1}{\theta^n}\mathbb{1}_{[X_{(n)},=\infty[}(\theta).$$

Remark : the set such that $L(.,\theta) > 0$ is $[0,\theta]$: it depends on θ, thus the model is not regular. Keep that in mind when dealing with Fisher information for instance.

1. **Flat prior:**

   We apply the definition for a Bayesian estimator with a prior density π_0 :

$$\begin{aligned}\hat{\theta}^{\mathrm{B}} &= \frac{\displaystyle\int_{\Theta} \theta L(x,\theta)\pi_0(\theta)\,\mathrm{d}\lambda(\theta)}{\displaystyle\int_{\Theta} L(x,\theta)\pi_0(\theta)\,\mathrm{d}\lambda(\theta)} \\ &= \frac{\displaystyle\int_{X_{(n)}}^{+\infty} \theta^{-n+1}\,\mathrm{d}\theta}{\displaystyle\int_{X_{(n)}}^{+\infty} \theta^{-n}\,\mathrm{d}\theta} \\ &= \frac{n-1}{n-2} X_{(n)}.\end{aligned}$$

   Bias and MSE are not computed there for sanity reasons.

2. **Jeffreys prior:**

   The density function of this prior is proportional to the squareroot of the determinant of the Fisher information matrix.

   Thus we need to compute this quantity for this model, with n observations (as it is not regular, $I_n \neq nI_1$) :

$$I_n(\theta) = \mathbb{E}\left[\frac{\partial \log L_n(\theta)}{\partial \theta}^2\right]$$

   We have $I_n(\theta) = \mathbb{E}[(-n/\theta)^2] = \dfrac{n^2}{\theta^2}$

   (If we had taken the expectancy of the second-order derivative of the log-likelihood, we would not have had the same result has the model is not regular.)

   This gives us the noninformative prior (Jeffreys) : $\pi_0(\theta) \propto \theta^{-1}$.