

Extreme Values in Financial Statistics

Master Thesis Project done at EPFL
towards the French 'Diplôme d'Ingénieur' degree
under the exchange agreement EPFL-ENSEEIH



Thèse de Master
présentée le 15 Août 2015
sous la supervision de
Prof. Stephan Morgenthaler,
Chaire de Statistique Appliquée
École Polytechnique Fédérale de Lausanne
pour l'obtention du Diplôme d'Ingénieur ENSEEIH
en Mathématiques Appliquées et informatique
par

Killian Martin–Horgassan

Lausanne, EPFL, 2015

Mihi cura futuri
— Ovide, Métamorphoses, 13, 363

To my friends and loved ones...

Acknowledgements

First and foremost, I would like to address my wholehearted thanks to Professor Stephan Morgenthaler who accepted to be the supervisor of my master thesis project. His guidance was steady throughout the project. At the same time, he left me free to go into and explore many different directions. As a consequence, I have learnt many things, and the need to work efficiently and set priorities prominent amongst them.

I am most grateful to Doctor Olivier Lévêque (EPFL) for his excellent stochastic calculus course as well as the exercises that go with it. In my opinion, this material is nothing short of the best for anyone who wants to get a grasp of the topic. He was always available whenever I had a question, and I want to thank him for that too.

My most sincere thanks go to Professor Joseph Gergaud, the head of the Applied Mathematics & Computer Science at ENSEEIHT, for allowing me to extend my exchange at EPFL and to pursue my master thesis project there.

Doctor Aurélie Hurault and Mrs Coralie Link, from the Student Exchange teams of ENSEEIHT and EPFL respectively, were there to answer my questions whenever they arose. Their support was very helpful throughout my time as an exchange student at EPFL, and I want to thank them for that.

Finally, I seize this opportunity to tell my friends and my family how much I am beholden to them !

Lausanne, 14 Août 2015

Killian Martin–Horgassan

Preface

I am currently a third-year engineering student at ENSEEIHT, a French 'Grande Ecole d'Ingénieurs' located in Toulouse. Like most students in my class, my first two years of higher education were spent in a 'Classe Prépa Maths Sup-Maths Spé'. There, I began to acquaint myself with advanced mathematics. It was mainly theoretical mathematics, and a good foundation upon which build further knowledge and skills.

Then, I passed a competitive entrance exam and I became an Applied Mathematics & Computer Science engineering student at ENSEEIHT. My training there took a more practical dimension, as the problems we were asked to solve were closer to those encountered in real life. My life took a turn - from the academic perspective at least - when I came across two books : Options, Futures and Other Derivatives by John C. Hull and The Signal and the Noise by Nate Silver. I had always had an interest in Finance, and then I realised that I was woefully ignorant of probability, statistics and stochastic calculus. Not only would I need them in my professional life, they were fascinating branches of Mathematics as well !

I set out to right that wrong and remedy this issue ; during my first semester as an exchange student at EPFL I took courses in those fields. It was not always easy but whatever difficulties I encountered only fed my interest in the subjects. I was given the opportunity to extend my exchange at EPFL to do a master thesis project under the supervision of Prof. Stephan Morgenthaler. I could not quite make up my mind between statistics and finance, and that is why I decided to do my master thesis project on extreme value statistics applied to finance. Extreme value statistics is very different from the kind of statistics I had so far encountered, and from the early days of the project I realised that the topic I had picked was vast. It may be that I shall have only scratched the surface of it at the end of the project. In any case, this project has been a good introduction to extreme value statistics and a incentive to explore further the field of financial mathematics !

Lausanne, 14 Août 2015

Killian Martin-Horgassan

Abstract

The topic of this project is twofold. It is both an introduction to extreme value theory and an application to finance.

First, we give an introduction to extreme value theory by making appear the extreme values distributions from the maxima of sequences of independent identically distributed random variables. This leads us to the two extremal problems and to the Fisher-Tippett-Gnedenko and Von Mises' theorems which provide answers to those problems.

Then, we study the theory of the Black-Scholes model for stock prices. We use a Black-Scholes model with time independent coefficients to simulate the prices over time of five stocks listed on the Paris stock exchange (BNP Paribas, Carrefour, LVMH, Sanofi and Total). Finally, we show that if we consider the stock prices above a high threshold (the 95 %-quantile for instance), we can fit extreme values distributions to the data.

Résumé

Notre sujet a une thématique duale, il est conçu comme une introduction à la théorie des valeurs extrêmes en statistique doublée d'une application au domaine de la finance.

Notre présentons d'abord l'approche en théorie des valeurs extrêmes, suivie d'une découverte des distributions des valeurs extrêmes à travers l'étude des maxima d'échantillons de variables aléatoires indépendantes identiquement distribuées. Cela nous amène aux théorèmes de Von Mises et de Fisher-Tippett-Gnedenko, ils constituent les réponses aux problèmes extrêmes c'est-à-dire le coeur de la théorie des valeurs extrêmes.

Nous nous intéressons alors aux évolutions des prix de 5 actions listées à la bourse de Paris (BNP Paribas, Carrefour, LVMH, Sanofi et Total). Nous étudions d'abord la théorie du modèle de Black-Scholes, nous appliquons ensuite un modèle de Black-Scholes avec coefficients indépendants du temps à nos données. Enfin, nous réalisons la jonction entre les valeurs extrêmes et nos données financières en étudiant, pour chacune des 5 actions, les valeurs des prix dépassant un certain seuil (le 95 %-quantile, par exemple). Nous montrons qu'il est possible de procéder à un ajustement de courbes à l'aide de distributions généralisées des valeurs extrêmes sur ces données.

Contents

Acknowledgements	i
Preface	iii
Abstract (English/Français)	v
List of figures	xi
List of tables	xiii
1 Introduction	1
1.1 A few words to set the scene	1
1.2 Formalising the settings	1
2 Investigating results on the limiting distribution	3
2.1 Playing with the (original) sequence of maxima	3
2.2 Sample following a Normal distribution	3
2.3 Sample following a Cauchy distribution	5
2.4 Sample following an Exponential Distribution	7
2.5 Why does this work ?	8
2.6 What are the limiting distributions in theses cases ?	9
3 The two extremal problems	13
3.1 The extremal limit problem	13
3.2 The domain of attraction problem	13
3.3 Conclusion	14
4 Looking into financial data	15
4.1 Five stocks and their evolution over 15 years	15
4.2 A détour around Stochastic Calculus	18
4.2.1 The Black-Scholes Stochastic Differential Equation	18
4.2.2 Black-Scholes SDE with time-dependent coefficients	21
4.3 Back to the data	22
4.3.1 Observations	30
4.4 Fitting data above a threshold	35
	ix

Contents

4.4.1 Goodness of fit	39
5 Conclusion	43
Bibliography	53
Recommendations on the use of the bibliography	55

List of Figures

2.1	Below, a realisation of the sequence of maxima for i.i.d. standard unit Gaussian RVs	4
2.2	Scatter Plot of the Maxima, $n = 10000$	4
2.3	Maxima against the time steps	4
2.4	Maxima against the time steps and function $n \rightarrow \sqrt{2\log(n)}$	5
2.5	The correction becomes negligible compared to the starting term as n grows large	5
2.6	Below, a realisation of the sequence of maxima for i.i.d. <i>Cauchy</i> (0,1) RVs	6
2.7	Scatter Plot of the Maxima, $n = 10000$	6
2.8	Maxima against the time steps	6
2.9	Maxima against the time steps and function $n \rightarrow \tan(\pi \frac{2-n}{2n})$	7
2.10	Below, a realisation of the sequence of maxima for i.i.d. <i>Exp</i> (1) RVs	7
2.11	Scatter Plot of the Maxima, $n = 10000$	8
2.12	Maxima against the time steps	8
2.13	Maxima against the time steps and function $n \rightarrow \log(n)$	8
4.1	15 years of weekly BNP Stock Price Data	15
4.2	15 years of weekly Carrefour Stock Price Data	15
4.3	15 years of weekly LVMH Stock Price Data	16
4.4	15 years of weekly Sanofi Stock Price Data	16
4.5	15 years of weekly Total Stock Price Data	16
4.6	$\exp(R_t)$ and $1 + R_t$ for BNP Stock Price Data	17
4.7	$\exp(R_t)$ and $1 + R_t$ for Carrefour Stock Price Data	17
4.8	$\exp(R_t)$ and $1 + R_t$ for LVMH Stock Price Data	17
4.9	$\exp(R_t)$ and $1 + R_t$ for Sanofi Stock Price Data	17
4.10	$\exp(R_t)$ and $1 + R_t$ for Total Stock Price Data	18
4.11	Log-gross returns on the stock, the μ 's and the σ^2 's.	24
4.12	The actual stock prices and a 1^{st} simulation of the corresponding geometric Brownian motions.	25
4.13	The actual stock prices and a 2^{nd} simulation of the corresponding geometric Brownian motions.	26
4.14	The actual stock prices and a 3^{rd} simulation of the corresponding geometric Brownian motions.	27

List of Figures

4.15 The actual stock prices and a 4 th simulation of the corresponding geometric Brownian motions.	28
4.16 The actual stock prices and a 5 th simulation of the corresponding geometric Brownian motions.	29
4.17 Cut LVMH stock price time series. In grey the actual prices, in colour five different runs of the simulation	31
4.18 Cut Total stock price time series. In grey the actual prices, in colour five different runs of the simulation	32
4.19 Cut LVMH stock price time series. In grey the actual weekly prices, in colour two runs of the daily simulation	33
4.20 Cut Total stock price time series. In grey the actual weekly prices, in colour two runs of the daily simulation	34
4.21 Data above the 95 %-quantile for the BNP Paribas, Carrefour, LVMH, Sanofi and Total stocks.	35
4.22 Density fitting for the BNP Paribas data - $\hat{\gamma}_{MLE} = 0.052$	36
4.23 Density fitting for the Sanofi data - $\hat{\gamma}_{MLE} = 0.48$	36
4.24 Density fitting for the Carrefour data - $\hat{\gamma}_{MLE} = -0.16$	37
4.25 Density fitting for the Total data - $\hat{\gamma}_{MLE} = -0.34$	37
4.26 Density fitting for the LVMH data - 95 % quantile	38
4.27 Density fitting for the LVMH data - 97.5 % quantile - $\hat{\gamma}_{MLE} = -0.71$	38
4.28 QQ-plots for the BNP Paribas data, in red the regression lines.	39
4.29 QQ-plots for the Carrefour data, in red the regression lines.	40
4.30 QQ-plots for the LVMH data, in red the regression lines.	40
4.31 QQ-plots for the Sanofi data, in red the regression lines.	41
4.32 QQ-plots for the Total data, in red the regression lines.	42

List of Tables

1	Cumulative residuals between $\exp(R_t)$ and $1 + R_t$ for the BNP, Carrefour, LVMH, Sanofi, Total stocks	51
---	---------------------------------------------------------------------------------------------------------------------	----

1 Introduction

1.1 A few words to set the scene

In real life, it is not uncommon to have at one's disposal data about a phenomenon occurring through time. It may be as simple as daily rainfall data in a city for the past two years, or it could be the weekly opening prices of a stock for the past decade.

Most of the time, people would like to use the data at their disposal to make predictions to answer questions, from the prosaic ones such as 'Will it rain tomorrow?' to more consequential ones such as 'Will I make a profit if I cling to my shares today and sell them only tomorrow?'. Of course, those are only vaguely worded questions : it is impossible to answer them satisfactorily without knowing the context, the objectives etc. behind them.

Yet, what these questions have in common is that they focus on the normal 'behaviour' that is to be expected in the future. Depending on the specific issue that is considered, the 'average behaviour' may not be the most interesting thing. For instance, suppose that a government wants to build a network of dams¹. The dams are meant to protect the country from future floods for the next one hundred years, therefore the question that needs to be answered is one of 'worst case event' : "Over the next century, how severe may be the worst flood?".

Extreme events are the kind of events we will be interested in this master thesis project. Although Extreme Value Theory has applications in many fields², we will here apply it more specifically to financial data.

1.2 Formalising the settings

Let $(X_n)_{n \geq 0}$ be a sequence of independent identically distributed random variables with common cumulative distribution function F_X . The sequence of maxima is defined by $M_0 = X_0$ and $\forall n \geq 1$, $M_n = \max_{0 \leq i \leq n}(X_i)$. We would like to determine the limiting distribution of the

¹As was done in The Netherlands beginning in the fifties

²including climate science, seismology, insurance etc.

Chapter 1. Introduction

sequence $(M_n)_{n \geq 0}$.³ This is a matter that will keep us busy quite a long time but the first thing to do is to re-formulate it.

Indeed, let us do a quick and simple computation :

$$\begin{aligned} F_n(t) &= \Pr(\{M_n \leq t\}) \\ &= \Pr(\{\max_{0 \leq i \leq n} (X_i) \leq t\}) \\ &= \Pr(\{X_1 \leq t\} \cap \cdots \cap \{X_n \leq t\}) \\ &= (F_X(t))^n \end{aligned} \tag{1.1}$$

Here we see that little information will be drawn from this result by taking the limit $n \rightarrow +\infty$. The limiting distribution will be degenerate. Indeed, let us consider the upper end-point of F_X ⁴, z^+ . Then,

$$\begin{aligned} \forall z < z^+ \quad \lim_{n \rightarrow \infty} F_n(z) &= 0 \\ \forall z \geq z^+ \quad \lim_{n \rightarrow \infty} F_n(z) &= 1 \end{aligned} \tag{1.2}$$

It turns out we cannot use the limiting distribution directly. A common approach⁵ is to consider a sequence of the maxima, standardized this time (i.e. centred and rescaled).

We will thus consider in all what follows the sequence defined by $(M_n^*)_{n \geq 0} = (\frac{M_n - b_n}{a_n})_{n \geq 0}$ where $(a_n)_{n \geq 0}$ and $(b_n)_{n \geq 0}$ are a sequence of real numbers and positive real numbers respectively. Finding a result on whether such a sequence admits a limiting distributions, and the conditions under which the result holds, will be one of our goals.

The two fundamental problems of extreme value theory More specifically, assuming that there exists a non-degenerate distribution G , what may G be ? That is the **extremal limit problem**. Additionally, what conditions do we have to impose on the common distribution of the random variables making up the sample, F_X , for the sequence $(M_n^*)_{n \geq 0}$ to converge to a non-degenerate distribution function G ? That is the **domain of attraction problem**.

³If we can determine the limiting distribution of the maxima from the data, then we will have a means to make predictions on the occurrence of future extreme events.

⁴that is the smallest z such that $F_X(z)$ be equal to one. For the Normal distribution, z will be $+\infty$, by contrast for a continuous Uniform Distribution $U([a, b])$ it will be b . The definition, properly speaking, of the upper end-point of F_X is the following : $z^+ = \inf\{z : F_X(z) \geq 1\}$.

⁵adopted by the mathematicians that laid the grounds of Extreme Value Theory.

2 Investigating results on the limiting distribution

2.1 Playing with the (original) sequence of maxima

Here, we will generate finite size sequences ($N = 10000$) of independent identically distributed random variables following respectively :

- a standard Normal Distribution $\mathcal{N}(0, 1)$
- a Cauchy Distribution $Cauchy(0, 1)$
- an Exponential Distribution $Exp(1)$

We will compute the sequence of maxima, neither centred nor normalised, and draw the scatter plot as well as the plot of the maxima M_n as a function of the time steps n . We will also draw the $1 - \frac{1}{n}$ -quantiles of the distributions (distributions of the sample, not of the maxima), which we will denote by q_n , as a function of the time steps n . This will lead us to make an interesting observation.

2.2 Sample following a Normal distribution

Computing the quantiles The Normal distribution is a particular case because, unlike in the cases of the Cauchy and the Exponential distribution, there is no explicit form to the cumulative distribution function. We will thus use a "well-known"¹ inequality, holding $\forall t > 0$:

$$\left(\frac{1}{t} - \frac{1}{t^3}\right) \frac{\exp(-\frac{t^2}{2})}{\sqrt{2\pi}} < 1 - \Phi(t) < \frac{1}{t} \frac{\exp(-\frac{t^2}{2})}{\sqrt{2\pi}} \quad (2.1)$$

¹Many textbooks mention it, though it is not necessarily what springs to the mind when thinking about the properties of Gaussian RVs.

Chapter 2. Investigating results on the limiting distribution

From there, it is easy to see that the following holds :

$$1 - \Phi(t) \sim_{t \rightarrow +\infty} \frac{1}{t} \frac{\exp(-\frac{t^2}{2})}{\sqrt{2\pi}} \quad (2.2)$$

When n grows large, the $1 - \frac{1}{n}$ -quantile grows very large so it is valid to replace $1 - \Phi(t)$ by its equivalent in the equation satisfied by the quantiles :

$$\begin{aligned} F_X(q_n) &= 1 - \frac{1}{n} \\ \Leftrightarrow \quad \frac{1}{q_n} \frac{\exp(-\frac{q_n^2}{2})}{\sqrt{2\pi}} &= \frac{1}{n} \\ \Leftrightarrow \quad \log(q_n) + \log(\exp(-\frac{q_n^2}{2})) + \log(\sqrt{2\pi}) &= \log(n) \end{aligned} \quad (2.3)$$

This equation cannot be solved analytically, we will resolve it iteratively. The starting point is $\log(n) = \frac{t_0^2}{2}$, which gives us $t_0 = \sqrt{(2\log(n))}$. If we then run the Newton-Raphson algorithm, we see that the corrections to t_0 from the next iterations are small enough that we can keep t_0 as solution.².

Figure 2.1 – Below, a realisation of the sequence of maxima for i.i.d. standard unit Gaussian RVs



Figure 2.2 – Scatter Plot of the Maxima, $n = 10000$



Figure 2.3 – Maxima against the time steps

²See the fourth of the figures below



Figure 2.4 – Maxima against the time steps and function $n \rightarrow \sqrt{2\log(n)}$



Figure 2.5 – The correction becomes negligible compared to the starting term as n grows large

2.3 Sample following a Cauchy distribution

Computing the quantiles Let X_1, \dots, X_n be i.i.d. RVs $\sim \text{Cauchy}(0, 1)$. The distribution function is $F_X(t) = \frac{1}{\pi} \arctan(x) - \frac{1}{2}$. The $1 - \frac{1}{n}$ -quantiles satisfy the equation :

$$\begin{aligned}
 F_X(q_n) &= 1 - \frac{1}{n} \\
 \Leftrightarrow \frac{\arctan(q_n)}{\pi} + \frac{1}{2} &= 1 - \frac{1}{n} \\
 \Leftrightarrow \frac{\arctan(q_n)}{\pi} &= \frac{2-n}{n} \\
 \Leftrightarrow q_n &= \tan\left(\frac{\pi}{2} \frac{2-n}{n}\right)
 \end{aligned} \tag{2.4}$$

Figure 2.6 – Below, a realisation of the sequence of maxima for i.i.d. *Cauchy*(0,1) RVs

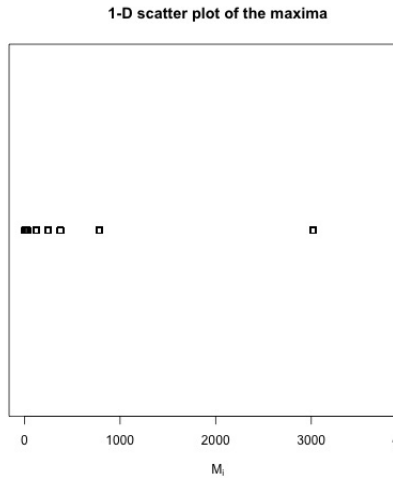


Figure 2.7 – Scatter Plot of the Maxima, $n = 10000$

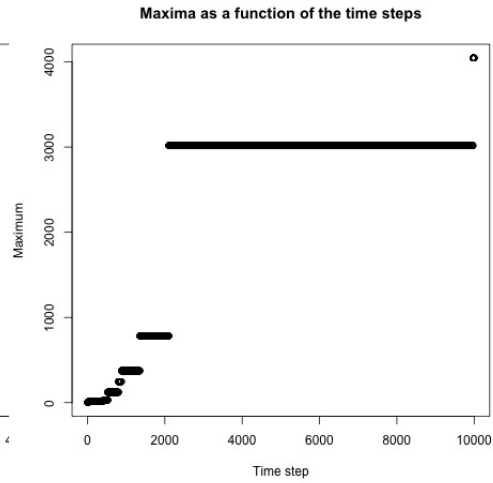


Figure 2.8 – Maxima against the time steps

2.4. Sample following an Exponential Distribution

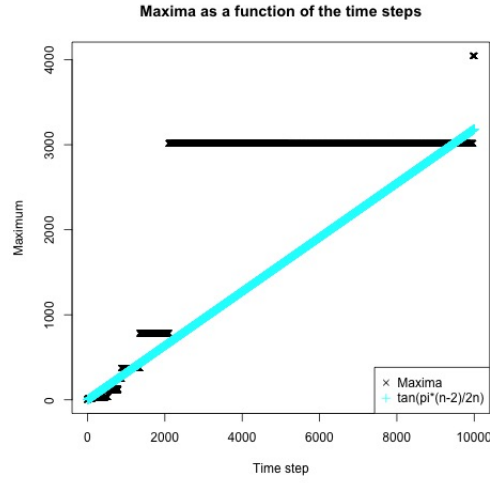


Figure 2.9 – Maxima against the time steps and function $n \rightarrow \tan(\pi \frac{2-n}{2n})$

$n \rightarrow \tan(\pi \frac{2-n}{2n})$ is roughly linear in n . We know that if $|x| < \frac{\pi}{2}$, $\tan(x) = x + \frac{x^3}{3} + \frac{2x^5}{15} + \dots$ so if we take the first order approximation, we get $\tan(x) \approx x$ which confirms what the plot seems to suggest.³

2.4 Sample following an Exponential Distribution

Computing the quantiles Let X_1, \dots, X_n be i.i.d. RVs $\sim \text{Exponential}(\lambda)$. The distribution function is $F_X(t) = 1 - \exp(-\lambda t)$. The $1 - \frac{1}{n}$ -quantiles satisfy the equation :

$$\begin{aligned} F_X(q_n) &= 1 - \frac{1}{n} \\ \Leftrightarrow 1 - \exp(-\lambda q_n) &= 1 - \frac{1}{n} \\ \Leftrightarrow q_n &= \frac{1}{\lambda} \log(n) \end{aligned} \tag{2.5}$$

Figure 2.10 – Below, a realisation of the sequence of maxima for i.i.d. $\text{Exp}(1)$ RVs

³Of course the expansion is valid as for a number of observations n greater than 1, $-1 < \frac{2}{n} - 1 < 1$ and thus $|\frac{\pi}{2} \frac{2-n}{n}| = |\frac{\pi}{2} (\frac{2}{n} - 1)| < \frac{\pi}{2}$.



Figure 2.11 – Scatter Plot of the Maxima, $n = 10000$
 Figure 2.12 – Maxima against the time steps



Figure 2.13 – Maxima against the time steps and function $n \rightarrow \log(n)$

2.5 Why does this work ?

(The underlying idea) Why have we made a link between the $\frac{1}{n}$ -quantiles of the common distribution of the X_i and the sequence of the M_n ? Actually, the link is that if the sample is made up of independent realizations, then M_n is an estimate of the $1 - \frac{1}{n}$ quantile.

The idea behind this is as follows, that the maxima will get closer to the upper end-point of the distribution of the X_i , F_{X_i} . That is also the case for the $1 - \frac{1}{n}$ -quantiles of F_{X_i} . Of course, it is only an intuition !

2.6 What are the limiting distributions in these cases ?

Preliminary remark In what follows we are using theoretical results exposed in the next chapter, chapter 3, concerning the three possible limiting distributions as well as Von Mises' theorem.

Fréchet limiting distribution Let us assume that a random variable X follows a *Cauchy*(0, 1) distribution, $x^+ = +\infty$. The limiting distribution can only be either a Gumbel or a Fréchet-type distribution.

$$f_X(x) = \frac{1}{\pi} \frac{1}{1+x^2} \quad (2.6)$$

$$F_X(x) = \frac{1}{\pi} \arctan(x) + \frac{1}{2} \quad (2.7)$$

$$\begin{aligned} r(x) &= \frac{f_X(x)}{1 - F_X(x)} \\ &= \frac{\frac{1}{1+x^2}}{\frac{\pi}{2} - \arctan(x)} \\ &= \frac{\frac{1}{1+x^2}}{\frac{\pi}{2} - (\frac{\pi}{2} - \arctan(\frac{1}{x}))} \\ &= \frac{1}{(1+x^2) \arctan(\frac{1}{x})} \\ &= \frac{1}{(1+x^2)(\frac{1}{x} + o(\frac{1}{x^2}))} \\ &= \frac{1}{\frac{1+x^2}{x} + o(1)} \\ \Rightarrow xr(x) &= \frac{x^2}{x^2 + 1 + o(1)} \\ \Rightarrow xr(x) &\xrightarrow{x \rightarrow +\infty} 1 \end{aligned} \quad (2.8)$$

Finally, by Von Mises' theorem, we can conclude that the limiting distribution for the standardized maxima of a *Cauchy*(0, 1) sample is a Fréchet distribution.

Fréchet distribution Let us assume that a random variable X follows a *Exp*(1) distribution, $x^+ = +\infty$. The limiting distribution here again can only be either a Gumbel or a Fréchet-type distribution.

$$f_X(x) = \lambda \exp(-\lambda x) \quad (2.9)$$

$$F_X(x) = 1 - \exp(-\lambda x) \quad (2.10)$$

$$\begin{aligned} r(x) &= \frac{f_X(x)}{1 - F_X(x)} \\ &= \frac{\lambda \exp(-\lambda x)}{1 - (1 - \exp(-\lambda x))} \\ &= \frac{\lambda \exp(-\lambda x)}{\exp(-\lambda x)} \\ &= \lambda \\ \Rightarrow \frac{dr}{dx}(x) &= 0 \end{aligned} \quad (2.11)$$

Finally, by Von Mises' theorem, we can conclude that the limiting distribution for the standardized maxima of an $\text{Exp}(1)$ sample is a Gumbel distribution.

Gumbel distribution - bis Let us assume that a random variable X follows a $\mathcal{N}(0, 1)$ distribution, $x^+ = +\infty$. The limiting distribution here again can only be either a Gumbel or a Fréchet-type distribution. It turns out that in that case, the limiting distribution is a Gumbel-type distribution. The computation is simple, as shown below :

$$\begin{aligned} \frac{dr}{dx}(x) &= \frac{\frac{df}{dx}(x)(1 - F_X(x)) - f(x)(-f(x))}{(1 - F_X(x))^2} \\ &= \frac{f(x)^2 + \frac{df}{dx}(x)(1 - F_X(x))}{(1 - F_X(x))^2} \end{aligned} \quad (2.12)$$

It yields an indeterminate form. The key to solve the issue here is to use an expansion of the cumulative distribution function of a standard Gaussian random variable :

$$\Phi(x) \approx \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \left(x + \frac{x^3}{3} + \frac{x^5}{15}\right)$$

$$\begin{aligned} \frac{dr}{dx}(x) &= \frac{\frac{\exp(-x^2)}{2\pi} - \frac{x}{2\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) + \frac{x \exp(-x^2)}{2\pi} \left(x + \frac{x^3}{3} + \frac{x^5}{15}\right)}{\left(\frac{1}{2} - \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \left(x + \frac{x^3}{3} + \frac{x^5}{15}\right)\right)^2} \\ &= \frac{\frac{\exp(-x^2)}{2\pi} - \frac{x}{2\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) + \frac{x \exp(-x^2)}{2\pi} \left(x + \frac{x^3}{3} + \frac{x^5}{15}\right)}{\frac{1}{4} - \frac{\exp(-x^2)}{\sqrt{2\pi}} \left(x + \frac{x^3}{3} + \frac{x^5}{15}\right) + \frac{\exp(-x^2)}{2\pi} \left(x + \frac{x^3}{3} + \frac{x^5}{15}\right)^2} \end{aligned} \quad (2.13)$$

Now we can see that if we take the limit when x goes to $+\infty$, we get a 0. By Von Mises' theorem, we can conclude that the limiting distribution for the standardized maxima of a $\mathcal{N}(0, 1)$ sample is a Gumbel distribution.

2.6. What are the limiting distributions in these cases ?

Weibull distribution Cauchy, Normal and Exponential distributions all have an infinite upper end-point, thus we will never get a Weibull distribution as limiting distribution. Let us assume that a random variable X follows a $Unif([0, 1])$ distribution, $x^+ = 1 < +\infty$.

$$f_X(x) = 1 \quad (2.14)$$

$$F_X(x) = x \quad (2.15)$$

$$\begin{aligned} r(x) &= \frac{f_X(x)}{1 - F_X(x)} \\ &= \frac{1}{1 - x} \\ \Rightarrow (x^+ - x)r(x) &= (x^+ - x)\frac{1}{1 - x} \\ \Rightarrow (1 - x)r(x) &= (1 - x)\frac{1}{1 - x} = 1 \\ \Rightarrow (x^+ - x)r(x) &\xrightarrow{x \rightarrow x^+} 1 > 0 \end{aligned} \quad (2.16)$$

Finally, by Von Mises' theorem, we can conclude that the limiting distribution for the standardized maxima of a $Unif([0, 1])$ sample is a Weibull distribution.

3 The two extremal problems

3.1 The extremal limit problem

The answer to the extremal limit problem It turns out that all possible non-degenerate limiting distributions i.e. all extreme values distribution make up a one-parameter family $G_\gamma(x) = \exp(-(1 + \gamma x)^{-\frac{1}{\gamma}})$, where the support of G is the set $\{x : 1 + \gamma x > 0\}$ and $\gamma \in \mathbb{R}$ is the Extreme Value Index or EVI. The three sub-cases are the following :

- $\gamma = 0$: Gumbel distribution
 $G_\gamma(u) = \exp(-\exp(-u)), u \in \mathbb{R}$
- $\gamma > 0$: Fréchet distribution
 $G_\gamma(u) = \exp(-(1 + \gamma u)^{-\frac{1}{\gamma}}), u \in]-\gamma^{-1}, +\infty[$
- $\gamma < 0$: Weibull distribution
 $G_\gamma(u) = \exp(-(1 + \gamma u)^{-\frac{1}{\gamma}}), u \in]-\infty, -\gamma^{-1}[$

For the derivations and proofs relative to this section, please report to the appendix.

3.2 The domain of attraction problem

Definition The domain of attraction of an extreme value distribution family (i.e. Gumbel, Fréchet-type or Weibull-type) is the set of distribution functions F_X^1 such that the sequence of standardized maxima $(M_n^*)_{n \geq 0}$ will converge in distribution to that extreme value distribution family.

Remark There are many approaches to characterize the domains of attraction of the extreme value distribution families. We have decided to use Von Mises' theorem to characterize

¹ F_X being the distribution of the X_i of the sample.

Chapter 3. The two extremal problems

them. This is the historical approach, and a rather straightforward one, still by no means are alternative approaches uninteresting².

Hazard function Let X be a random variable with probability density function/mass function f_X and distribution function F_X , then we define the hazard function r as follows :

$$r(x) = \frac{f_X(x)}{1-F_X(x)}.$$

A few preliminary notations Φ_α , Ψ_α , Δ are respectively the symbols used to denote a Fréchet, a Weibull and a Gumbel distributions, with :

- $\Phi_\alpha(x) = \exp(-x^\alpha)$
- $\Psi_\alpha(x) = \exp(-|x|^\alpha)$ (let us bear in mind that this is a notation, due to historical reasons).
- $\Delta(x) = \exp(-\exp(-x))$

Von Mises' theorem

1. If $x^+ = +\infty$ and $xr(x) \xrightarrow{x \rightarrow +\infty} \alpha > 0$, then $F_X \in \mathcal{D}(\Phi_\alpha)$.
2. If $x^+ < +\infty$ and $(x^+ - x)r(x) \xrightarrow{x \rightarrow x^+} \alpha > 0$, then $F_X \in \mathcal{D}(\Psi_\alpha)$.
3. If $r(x)$ is ultimately positive in the neighbourhood of x^+ (with $x^+ \leq +\infty$), is differentiable on that neighbourhood and is such that $\frac{dr}{dx}(x) \xrightarrow{x \rightarrow x^+} 0$, then $F_X \in \mathcal{D}(\Delta)$.

3.3 Conclusion

Fisher-Tippett-Gnedenko theorem The Fisher-Tippett-Gnedenko theorem, also known as the **extremal theorem**, states that if the sequence of standardized maxima converges in distribution to a non-degenerate distribution, then this distribution belongs to one of the three aforementioned extreme value distribution families. The theorem thus provides an answer to the *extremal limit problem*. Von Mises' theorem, encountered in the previous section, provides a complementary answer, that to the *domain of attraction problem*.

²In particular, conditions based on the sole behaviour of F_X can be formulated.

4 Looking into financial data

4.1 Five stocks and their evolution over 15 years

We have chosen to study five stocks listed on the Paris Stock Exchange : BNP Paribas, Carrefour, LVMH, Sanofi and Total stocks.¹ The evolution of the stock prices has been studied over the past 15 years, on a weekly basis. We first draw the data itself, then the net returns and the gross log returns on the stocks²

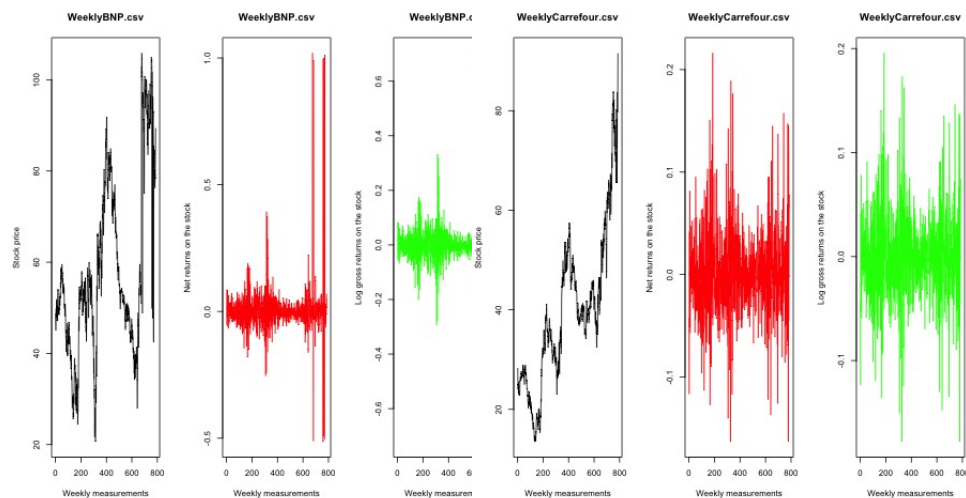


Figure 4.1 – 15 years of weekly BNP Stock Price Data

Figure 4.2 – 15 years of weekly Carrefour Stock Price Data

¹We have chosen companies positioned on different domains, otherwise, information from different stock might more easily be redundant.

²Both quantities are widely used in Finance.

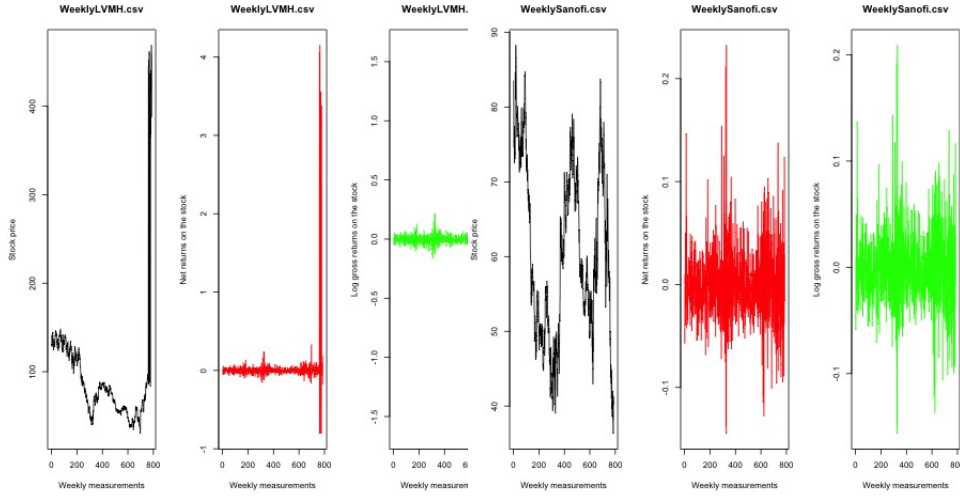


Figure 4.3 – 15 years of weekly LVMH Stock Price Data Figure 4.4 – 15 years of weekly Sanofi Stock Price Data

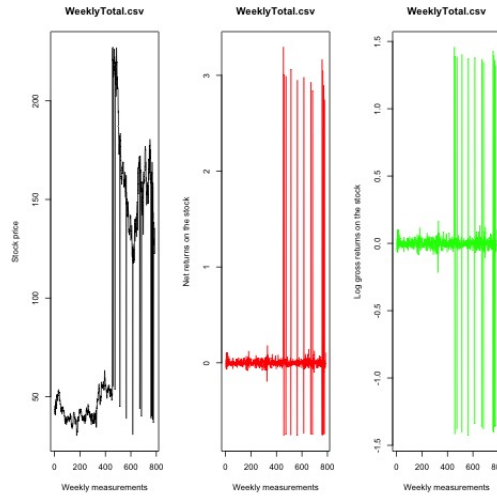


Figure 4.5 – 15 years of weekly Total Stock Price Data

Let X_t be the price of a stock at time t , the gross return at time $t + 1$ is defined as the ratio $\frac{X_{t+1}}{X_t}$, the net return at time $t + 1$ is defined as the ratio $r_t = \frac{X_{t+1} - X_t}{X_t}$ and the log gross return at time $t + 1$ is defined as the log of the gross return at time $t + 1$ i.e. $R_t = \log(\frac{X_{t+1}}{X_t})$. The latter two quantities are of particular interest in Finance.

Let us observe that the relationship between R_t , X_t and X_{t+1} can be rewritten as $X_{t+1} = \exp(R_{t+1})X_t$. An approximation would be to take $X_{t+1} = (1 + R_{t+1})X_t$ by taking the expansion of the exponential, cut at order 1. Below are the plots of the quantities $\exp(R_t)$ and $1 + R_t$ for the five stocks previously considered. As we can see from the value of the residuals (of order

4.1. Five stocks and their evolution over 15 years

$1E-15^3$), this is in practice a very good approximation !



Figure 4.6 – $\exp(R_t)$ and $1 + R_t$ for BNP Stock Price Data

Figure 4.7 – $\exp(R_t)$ and $1 + R_t$ for Carrefour Stock Price Data

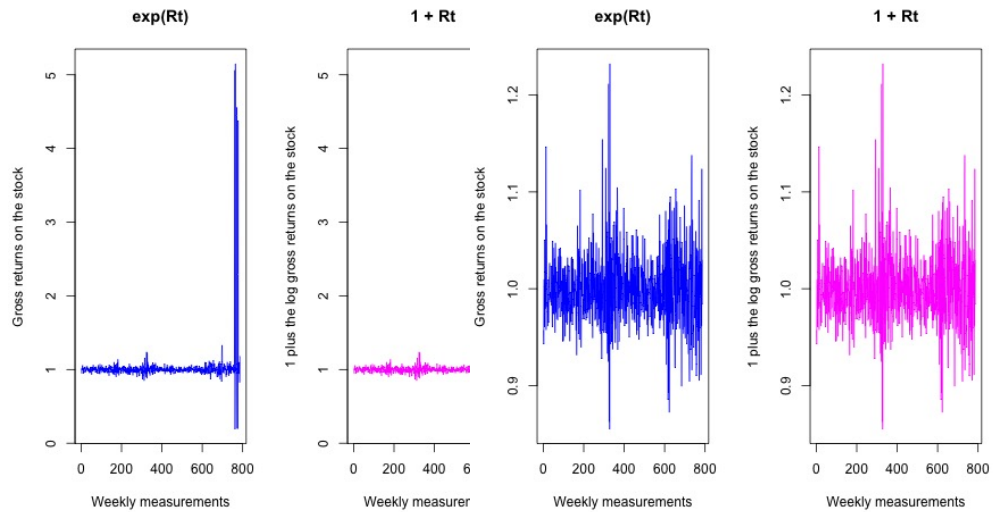


Figure 4.8 – $\exp(R_t)$ and $1 + R_t$ for LVMH Stock Price Data

Figure 4.9 – $\exp(R_t)$ and $1 + R_t$ for Sanofi Stock Price Data

³report to the appendix 'Additional figures' for the details.



Figure 4.10 – $\exp(R_t)$ and $1 + R_t$ for Total Stock Price Data

4.2 A détour around Stochastic Calculus

4.2.1 The Black-Scholes Stochastic Differential Equation

Presentation It is customary to model the evolution of stock prices by a stochastic process $(S_t)_{t \geq 0}$ satisfying the Black-Scholes stochastic differential equation⁴, which reads as follows :

$$dS_t = \mu S_t dt + \sigma S_t dB_t \quad (4.1)$$

where $(B_t)_{t \geq 0}$ is a standard Brownian Motion with respect to a filtration $(\mathcal{F}_t)_{t \geq 0}$. Let us observe that the stochastic differential equation above is the Black-Scholes SDE with time independent coefficients : both the drift $\mu \in \mathbb{R}$ and the volatility $\sigma > 0$ are constant.⁵ An initial condition must be specified : $S_0 = s_0 > 0$.

Resolution - existence of a solution to the Black-Scholes SDE Let us consider the following generic stochastic differential equation :

$$\begin{cases} dX_t = f(X_t)dt + g(X_t)dB_t \\ X_0 = x_0 \end{cases}$$

⁴Actually, a first model could be as simple as using arithmetical random walks. Two problems would however arise. The first is that if we used an ARW we could get negative stock prices, the second problem is that stocks selling at low prices usually have low price increments while stocks selling at high prices usually have larger price increments. In the Black-Scholes model, the price increment at time $t + 1$ is made up of a component proportional to the price at time t and a stochastic component. The stochastic component is also proportional to the price at time t . Both problems are thus solved. The process $(S_t)_{t \geq 0}$ is continuous and when S_t reaches 0, the infinitesimal increment which is $S_t(\mu dt + \sigma dB_t)$ is also 0. Therefore, S_t cannot go below 0.

⁵In the next subsection, we will deal with the Black-Scholes SDE with time dependent coefficients.

where $(B_t)_{t \geq 0}$ is a standard Brownian Motion with respect to a filtration $(\mathcal{F}_t)_{t \geq 0}$, $x_0 \in \mathbb{R}$, $f, g : \mathbb{R} \rightarrow \mathbb{R}$ are Lipschitz functions. Then, by a theorem from Stochastic Calculus⁶, we know that there exists a unique process that is continuous and adapted to the filtration $(\mathcal{F}_t)_{t \geq 0}$.

Here, f and g are respectively the functions $x \rightarrow \mu x$ and $x \rightarrow \sigma x$. These functions are Lipschitz functions, therefore we know the SDE admits a solution. And fortunately, in the case of the Black-Scholes equation, the solution can be made explicit⁷).

Resolution - step 1 Let us set $S_t = \phi_t Z_t$ where ϕ_t is the (deterministic) solution to the ordinary differential equation :

$$\begin{cases} d\phi_t = \mu \phi_t dt \\ \phi_0 = 1 \end{cases}$$

Solving this ODE is elementary and yields the solution $\phi_t = \exp(\mu t)$. Now, let us differentiate X_t under its form as a product $S_t = \phi_t Z_t$.

$$\begin{aligned} d(S_t) &= d(\phi_t Z_t) \\ &= \phi_t dZ_t + Z_t d\phi_t + d\langle \phi, Z \rangle_t \\ &= \phi_t dZ_t + \mu \phi_t Z_t dt \\ &= \phi_t dZ_t + \mu S_t dt \end{aligned} \tag{4.2}$$

where the infinitesimal quadratic covariation between ϕ and Z is zero as ϕ has bounded variations. If we compare the last right-hand term of the series of equations just above to the original Black-Scholes SDE, we see that $\sigma S_t dB_t = \phi_t dZ_t$. Hence, $dZ_t = \sigma Z_t dB_t$. Here, we must be very careful as this differential equation does not integrate as it would in the settings of 'usual' calculus, in particular, integrating it to $\log(Z_t) - \log(Z_0) = \sigma(B_t - B_0)$ is totally wrong !

Resolution - step 2 Now, let us set $Y_t = \log(Z_t)$. Using Ito-Doeblin's formula, we get :

$$d(\log(Z_t)) = \frac{1}{Z_t} dZ_t + \frac{1}{2} \left(\frac{-1}{Z_t^2} \right) d\langle Z \rangle_t \quad (\star)$$

⁶readers wanting to get ahold of an excellent course on Stochastic Calculus are advised to refer to Dr. L  v  que's course. It can be found at <http://ipg.epfl.ch/~leveque/>

⁷This is not always the case : solving SDEs in general is not that simple, and finding an explicit solution is not guaranteed in the general case, even though we know one exists.

→ How to compute $d \langle Z \rangle_t$?

$$\begin{aligned}
 S_t &= \phi_t Z_t \\
 \Rightarrow Z_t &= \frac{S_t}{\phi_t} \\
 \Rightarrow dZ_t &= \frac{1}{\phi_t} dS_t + S_t d\left(\frac{1}{\phi_t}\right) + d \langle \frac{1}{\phi}, S \rangle_t \\
 &= \frac{1}{\phi_t} dS_t - \frac{S_t}{\phi_t^2} d(\phi_t) \\
 &= \frac{1}{\phi_t} dS_t - \frac{\mu S_t}{\phi_t} dt \\
 &= \frac{1}{\phi_t} (\mu S_t dt + \sigma S_t dB_t) - \frac{\mu S_t}{\phi_t} dt \\
 &= \frac{\sigma S_t}{\phi_t} dB_t
 \end{aligned} \tag{4.3}$$

where the infinitesimal quadratic covariation between $\frac{1}{\phi}$ and Z is zero as $\frac{1}{\phi}$ has bounded variations, in the third equality from the top.

Hence, using the Isometry formula, we have that :

$$\begin{aligned}
 \langle Z \rangle_t &= \int_0^t \frac{\sigma^2 S_s^2}{\phi_s^2} ds \\
 &= \int_0^t \sigma^2 Z_s^2 ds \\
 \Rightarrow d \langle Z \rangle_t &= \sigma^2 Z_t^2 dt
 \end{aligned}$$

Back to (★), we now have :

$$\begin{aligned}
 d(\log(Z_t)) &= \frac{1}{Z_t} dZ_t + \frac{1}{2} \left(\frac{-1}{Z_t^2} \right) \sigma^2 Z_t^2 dt \\
 \Leftrightarrow \frac{1}{Z_t} dZ_t &= d(\log(Z_t)) + \frac{\sigma^2}{2} dt
 \end{aligned} \tag{4.4}$$

Resolution - step 3 Combining the previous equation with $dZ_t = \sigma Z_t dB_t$, we get :

$$\begin{aligned}
 \sigma dB_t &= d(\log(Z_t)) + \frac{\sigma^2}{2} dt \\
 \Rightarrow \log(Z_t) - \log(Z_0) &= -\frac{\sigma^2}{2}(t-0) + \sigma(B_t - B_0) \\
 \Rightarrow \log(Z_t) - \log\left(\frac{S_0}{\phi_0}\right) &= -\frac{\sigma^2}{2}t + \sigma B_t \\
 \Rightarrow \log(Z_t) &= \log(s_0) - \frac{\sigma^2}{2}t + \sigma B_t \\
 \Rightarrow Z_t &= s_0 \exp\left(-\frac{\sigma^2}{2}t + \sigma B_t\right)
 \end{aligned} \tag{4.5}$$

Finally, by remembering that $S_t = \phi_t Z_t = \exp(\mu t) Z_t$, we get :

$$\forall t \geq 0, S_t = s_0 \exp\left((\mu - \frac{\sigma^2}{2})t + \sigma B_t\right)$$

The stochastic process $(S_t)_{t \geq 0}$, made explicit above, that is solution to the Black-Scholes equation is generally called Geometric Brownian Motion in the literature.

4.2.2 Black-Scholes SDE with time-dependent coefficients

Presentation In the simple Black-Scholes SDE, the drift μ and the volatility σ were time-independent constants. Let us now consider a more general version of the Black-Scholes SDE :

$$\begin{cases} dS_t = \mu(t)S_t dt + \sigma(t)S_t dB_t \\ S_0 = s_0 > 0 \end{cases}$$

where $(B_t)_{t \geq 0}$ is a standard Brownian Motion with respect to a filtration $(\mathcal{F}_t)_{t \geq 0}$, μ, σ two continuous functions such that there exists $K_1 > 0$, $K_2 > 0$, such that $\forall t \geq 0, |\mu(t)| \leq K_1$, $K_2 \leq |\sigma(t)| \leq K_1$.

Resolution - existence of a solution to the generalized Black-Scholes SDE Let us consider the following generic stochastic differential equation :

$$\begin{cases} dX_t = f(t, X_t)dt + g(t, X_t)dB_t \\ X_0 = x_0 \end{cases}$$

where $(B_t)_{t \geq 0}$ is a standard Brownian Motion with respect to a filtration $(\mathcal{F}_t)_{t \geq 0}$, $x_0 \in \mathbb{R}$, $f, g : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$ are jointly continuous in (t, x) and Lipschitz in x . Then, by a theorem from Stochastic Calculus, we know that there exists a unique solution $(X_t)_{t \geq 0}$ to the SDE. in the case of the generalized Black-Scholes SDE, the conditions are met and we can thus conclude that it admits a unique solution. The solution can be made explicit here too, fortunately !

Resolution Let us set $Y_t = \log(S_t)$, we then have $dY_t = \frac{1}{S_t} dS_t - \frac{1}{2} \frac{1}{S_t^2} d\langle S \rangle_t$. If we remember that $dS_t = \mu(t)S_t dt + \sigma(t)S_t dB_t$ and apply the Isometry formula, we get that $d\langle S \rangle_t = \sigma(t)^2 S_t^2 dt$. We thus get :

$$\begin{aligned}
 \sigma dY_t &= \frac{1}{S_t} dS_t - \frac{1}{2} \frac{1}{S_t^2} d\langle S \rangle_t \\
 &= \frac{1}{S_t} dS_t - \frac{1}{2} \sigma(t)^2 dt \\
 &= \frac{1}{S_t} (\mu(t) S_t dt + \sigma(t) S_t dB_t) - \frac{1}{2} \sigma(t)^2 dt \\
 &= (\mu(t) - \frac{1}{2} \sigma(t)^2) dt + \sigma(t) dB_t \tag{4.6} \\
 \Rightarrow Y_t &= y_0 + \int_0^t (\mu(s) - \frac{1}{2} \sigma(s)^2) ds + \int_0^t \sigma(s) dB_s \\
 \Rightarrow Y_t &= \log(s_0) + \int_0^t (\mu(s) - \frac{1}{2} \sigma(s)^2) ds + \int_0^t \sigma(s) dB_s \\
 \Rightarrow S_t &= s_0 \exp(\int_0^t (\mu(s) - \frac{1}{2} \sigma(s)^2) ds + \int_0^t \sigma(s) dB_s)
 \end{aligned}$$

Let us observe that the solution found in the case of the generalized Black-Scholes SDE is coherent with the solution found for the simple Black-Scholes SDE⁸.

4.3 Back to the data

Fitting the data to the model We are now going to assume the stock prices $S_t^{(i)}$ ⁹ we have been studying follow the (simple) Black-Scholes model that is that their prices can be modeled by a geometric brownian motion. We will determine the corresponding μ_i 's and σ_i 's. To do that, we will make explicit the gross return at time $t + 1$:

$$\begin{aligned}
 \frac{S_{t+1}}{S_t} &= \frac{s_0 \exp((\mu - \frac{\sigma^2}{2})(t+1) + \sigma B_{t+1})}{s_0 \exp((\mu - \frac{\sigma^2}{2})t + \sigma B_t)} \\
 \Rightarrow \frac{S_{t+1}}{S_t} &= \exp(\mu - \frac{\sigma^2}{2}) \exp(\sigma(B_{t+1} - B_t)) \tag{4.7}
 \end{aligned}$$

⁸Just set functions μ and σ equal to constants μ and σ and we are back with the Geometric Brownian Motion previously found.

⁹The notations are the following :

- $S_t^{(1)}$ for the BNP stock.
- $S_t^{(2)}$ for the Carrefour stock.
- $S_t^{(3)}$ for the LVMH stock.
- $S_t^{(4)}$ for the Sanofi stock.
- $S_t^{(5)}$ for the Total stock.

A first observation that we could make is that as $x \rightarrow \exp(\sigma x)$ is a Borel-measurable function (because it is continuous) and the increments of the brownian motion are independent, $\frac{S_t}{S_{t-1}} = \exp(\mu - \frac{\sigma^2}{2}) \exp(\sigma(B_t - B_{t-1}))$ and $S_t = s_0 \exp((\mu - \frac{\sigma^2}{2})t + \sigma B_t) = s_0 \exp((\mu - \frac{\sigma^2}{2})t + \sigma(B_t - B_0))$ are independent $\forall t \geq 0$.¹⁰ Let us now consider the log-gross returns $R_t = (\mu - \frac{\sigma^2}{2}) + (\sigma(B_t - B_{t-1}))$. It is thus easy to see that R_t has a gaussian distribution $\mathcal{N}((\mu - \frac{\sigma^2}{2}), \sigma^2)$. Therefore, to find the μ_i 's and σ_i 's, we just have to compute the mean and the variance of the empirical log-gross returns which we respectively denote by μ_{EMP} and σ_{EMP}^2 .

$$\begin{cases} \sigma^2 &= \sigma_{EMP}^2 \\ \mu &= \mu_{EMP} + \frac{\sigma_{EMP}^2}{2} \end{cases} \quad (4.8)$$

¹⁰This is the formal justification behind the well-known fact that under the Black-Scholes model, the gross return at a time step is independent of the price of the stock at that date.

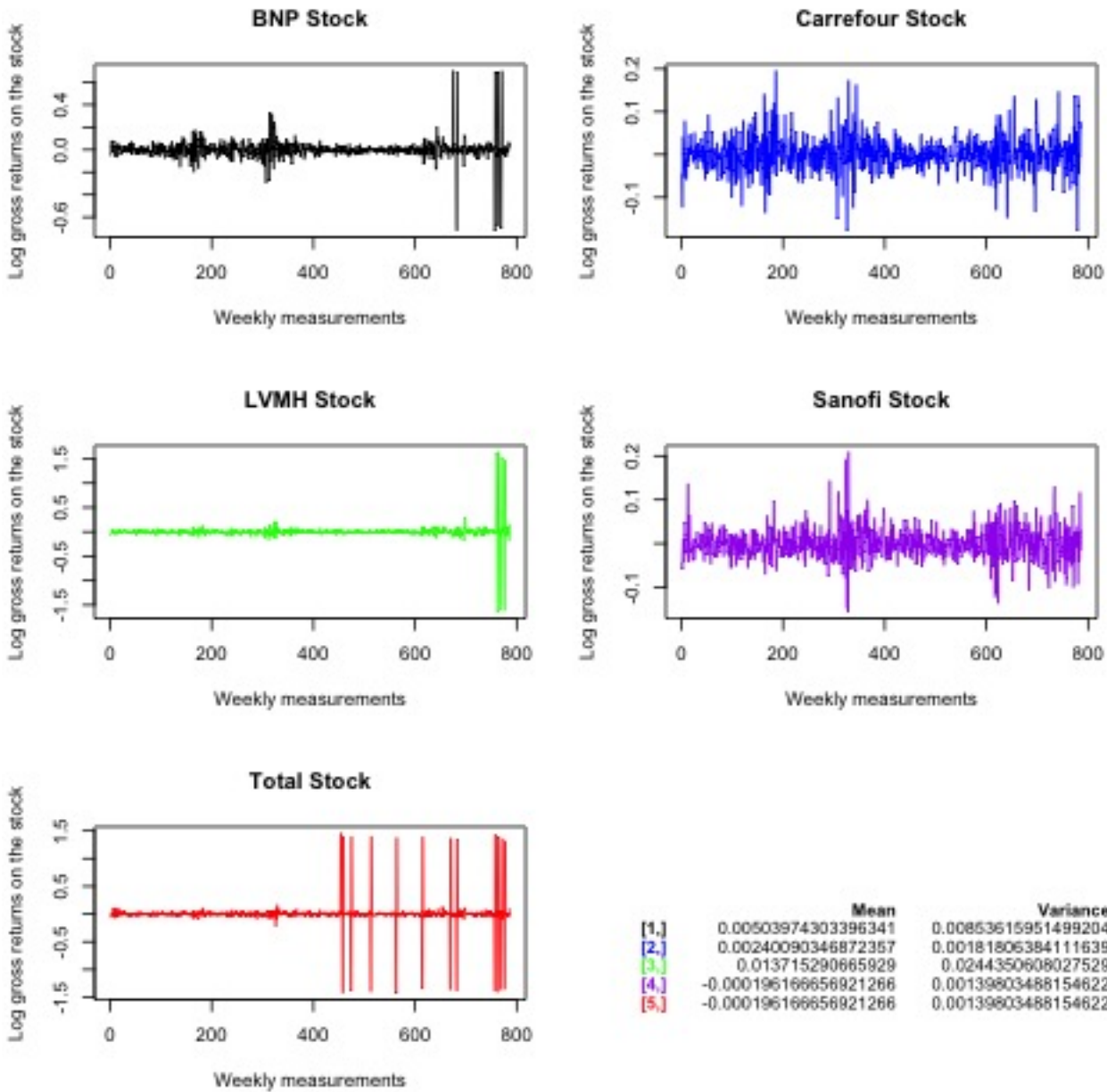


Figure 4.11 – Log-gross returns on the stock, the μ 's and the σ^2 's.

Now that we have the μ 's and the σ^2 's for the stocks, we will simulate geometric Brownian motions with these parameters, and compare the results to the original data. We will run several, say 5, simulations and make a few observations. Below the corresponding plots can be found with the **yellow line representing each time the simulated geometric Brownian motion** :

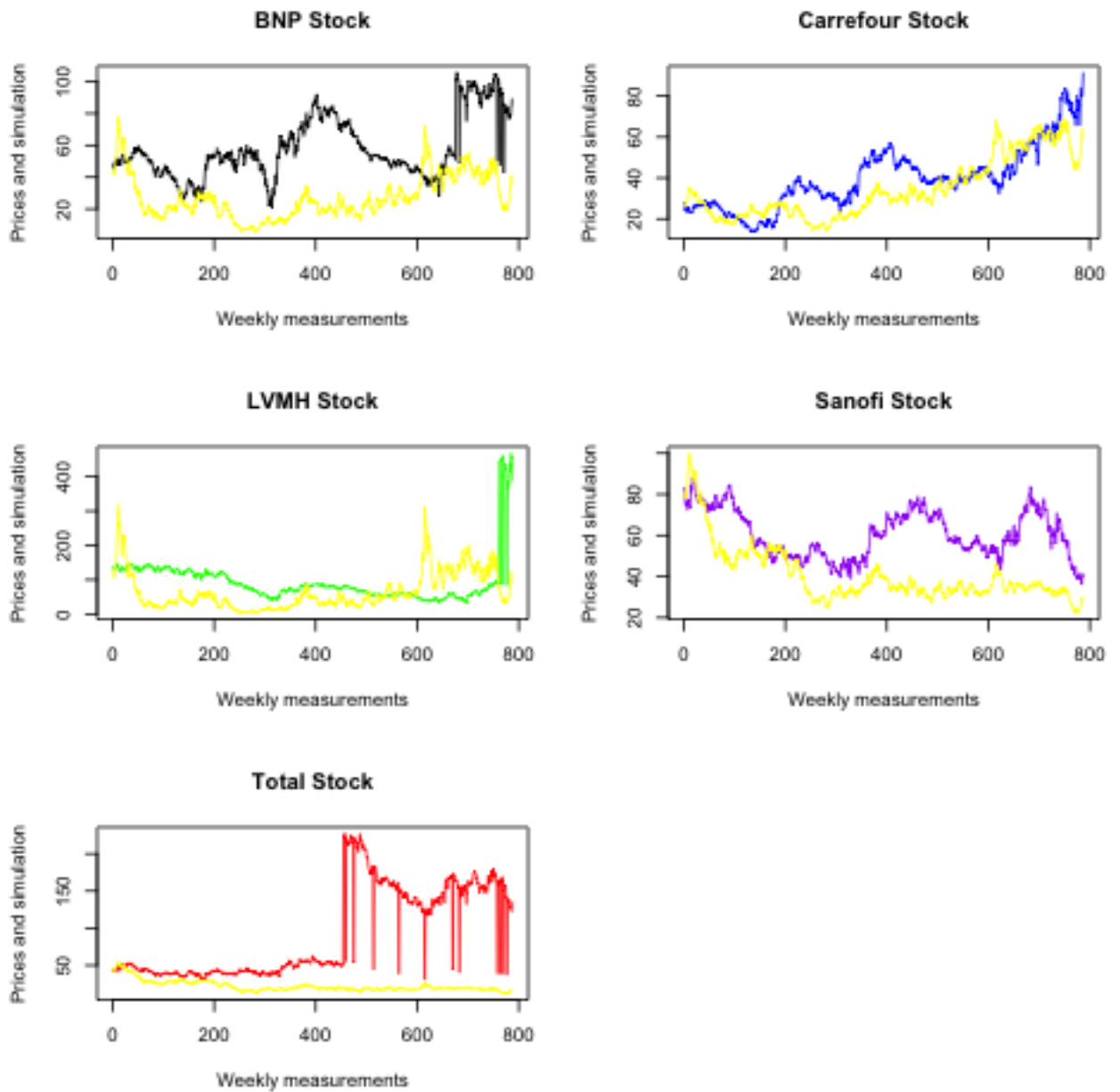


Figure 4.12 – The actual stock prices and a 1st simulation of the corresponding geometric Brownian motions.

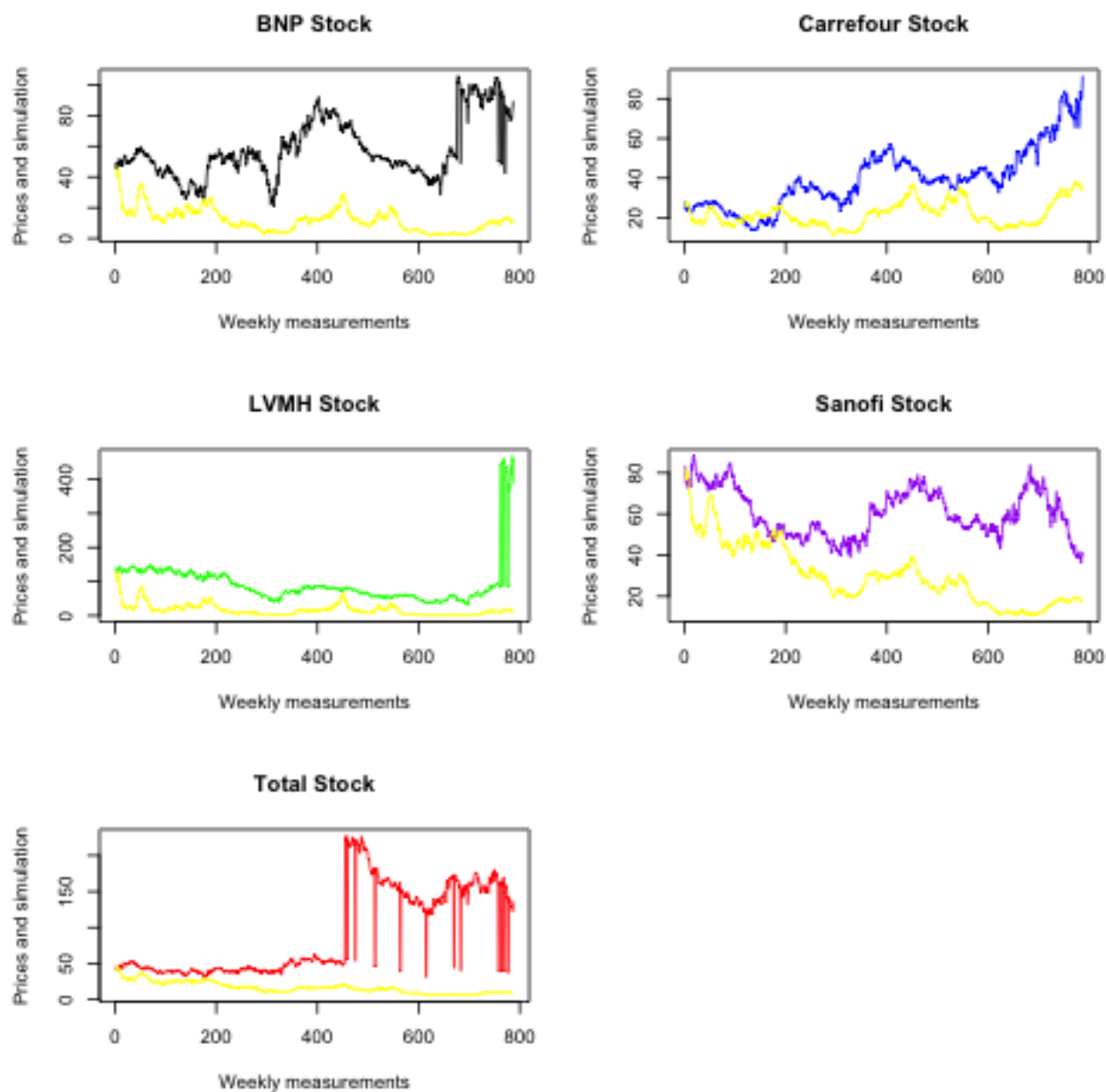


Figure 4.13 – The actual stock prices and a 2^{nd} simulation of the corresponding geometric Brownian motions.

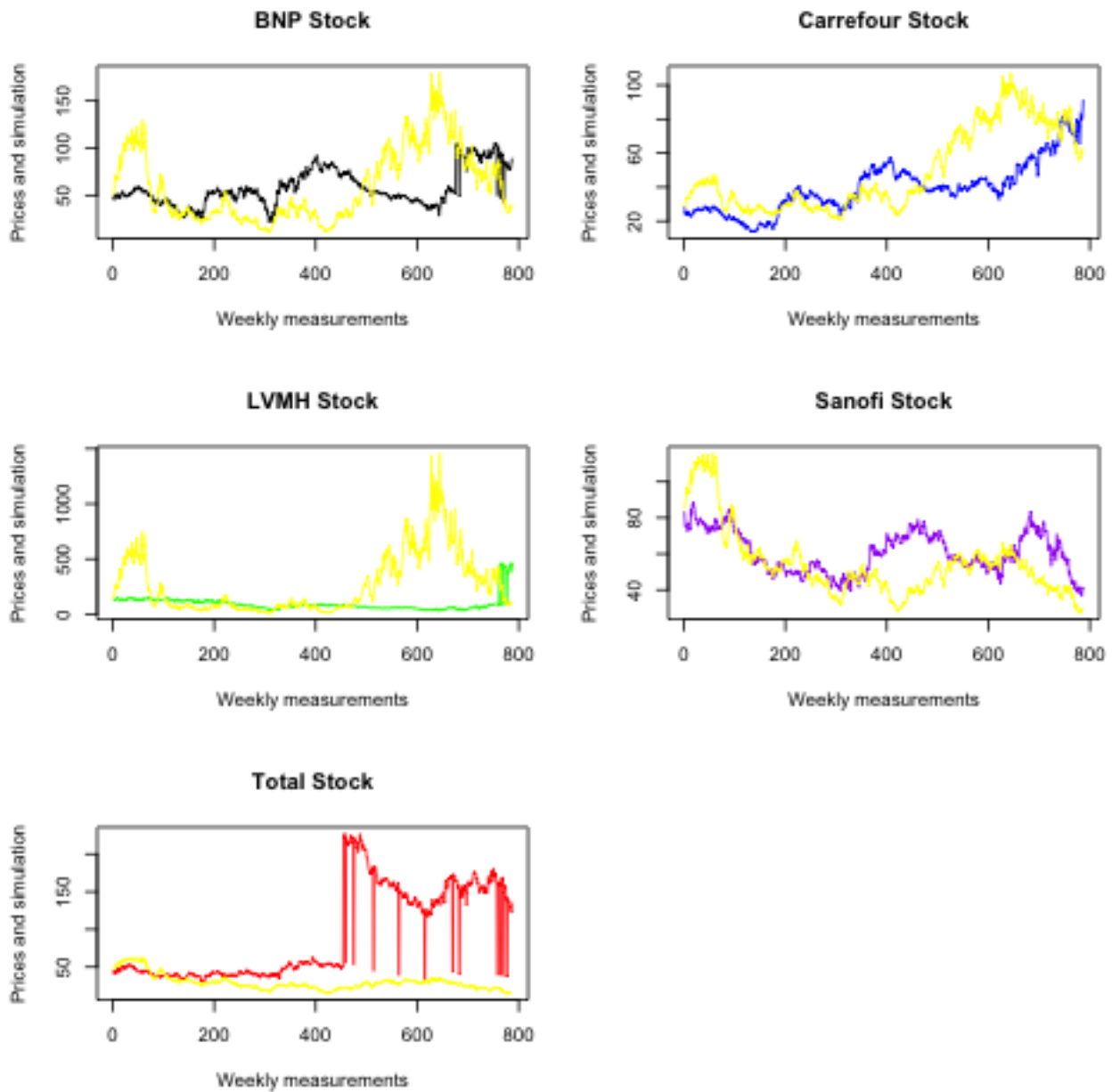


Figure 4.14 – The actual stock prices and a 3rd simulation of the corresponding geometric Brownian motions.

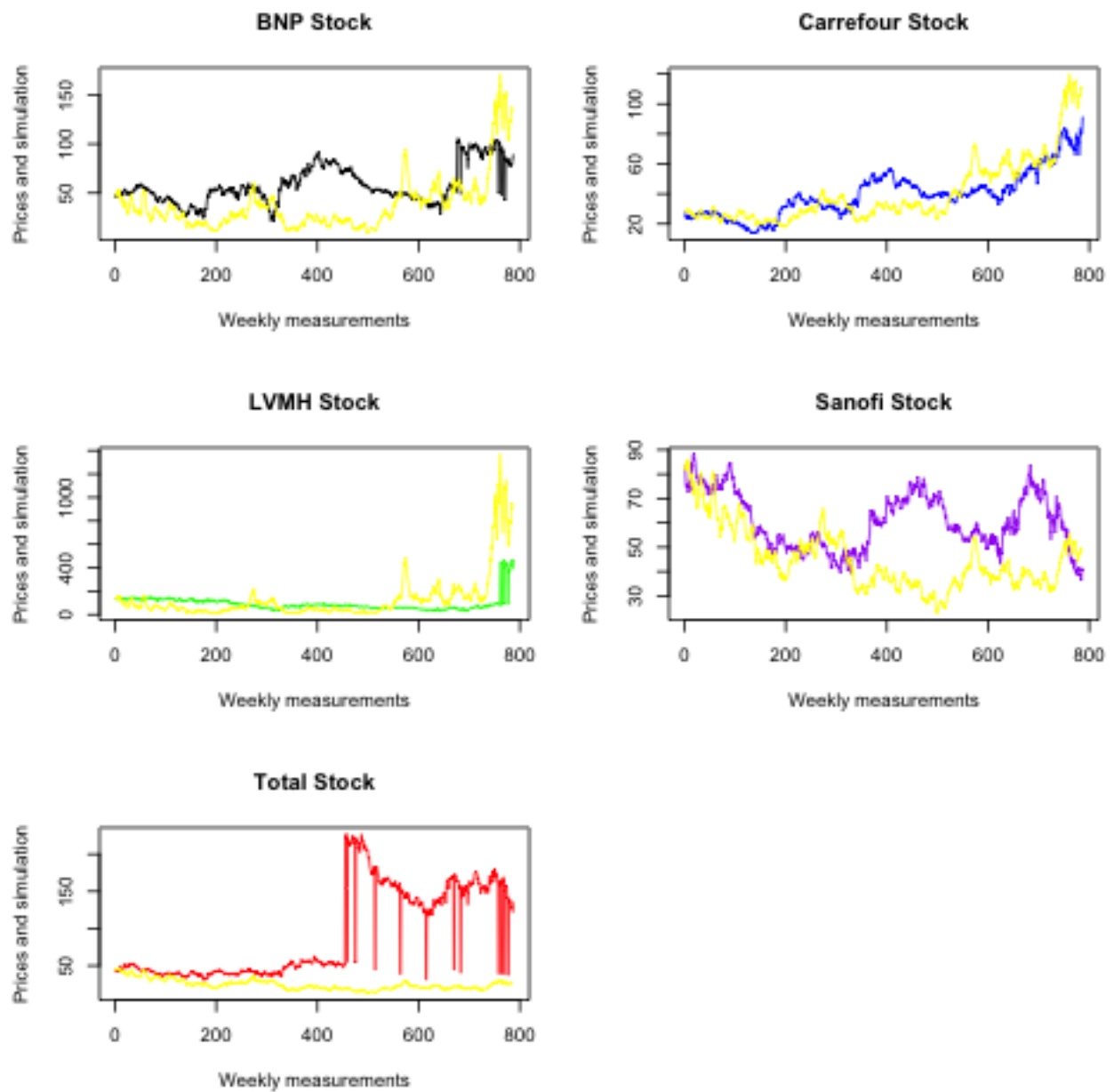


Figure 4.15 – The actual stock prices and a 4th simulation of the corresponding geometric Brownian motions.

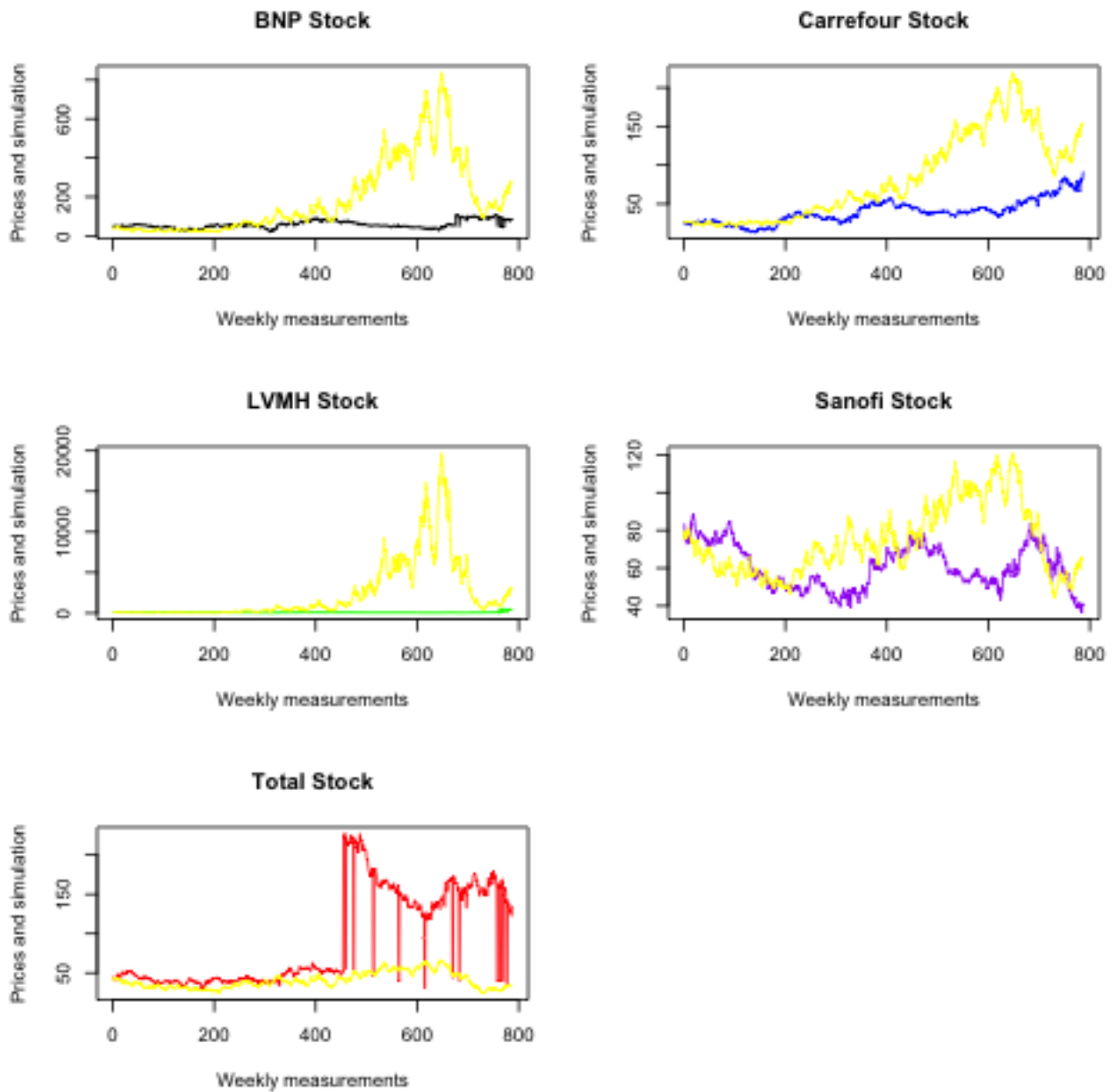


Figure 4.16 – The actual stock prices and a 5th simulation of the corresponding geometric Brownian motions.

4.3.1 Observations

As we can see, using a geometric Brownian motion to model stock prices yields rather good results though the last simulation displays a case of (gross) overestimation of the prices. However, let us note that this is true for the BNP, Carrefour and Sanofi stocks. If we have a look at the Total stock, we see that from week 500 or so the stock price varies quickly widely. The geometric Brownian motion is unable to cope with this behaviour, all five simulations are outright failures. There is a similar issue with the LVMH stock at the end of the series after a rather flat behaviour for most of the duration of the observations, which may explain why the simulations are poor in the case of that stock too.

As a consequence, we have decided to cut the LVMH and Total data so as to keep only the part of the time series where the stock prices are not too volatile, and then apply our simulation to the cut time series. As we can see below, the result is better¹¹

¹¹The LVMH stock, though, is still more volatile than the Total stock. That explains why the simulation is poorer for the former than for the latter.

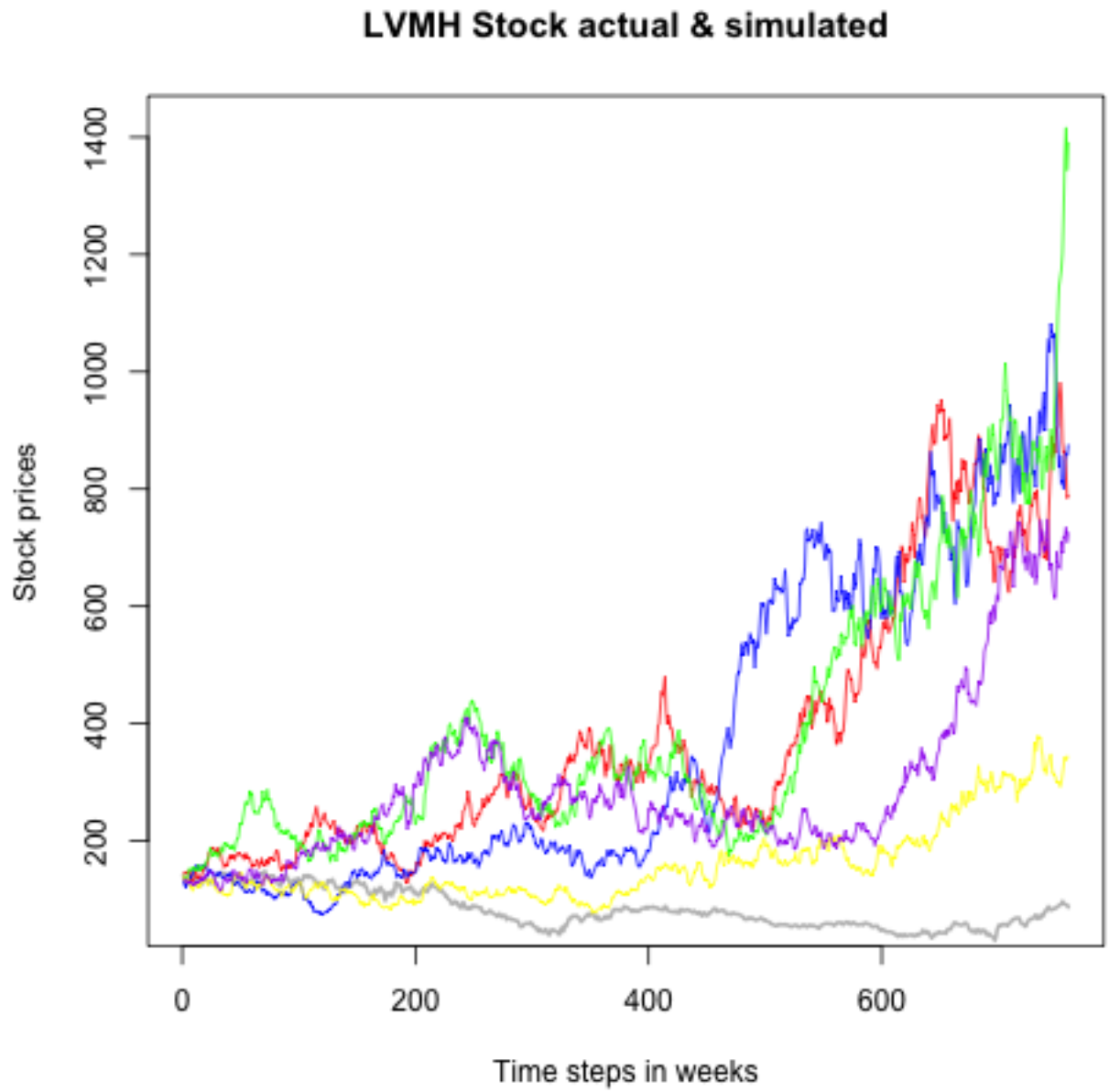


Figure 4.17 – Cut LVMH stock price time series. In **grey** the actual prices, in colour five different runs of the simulation

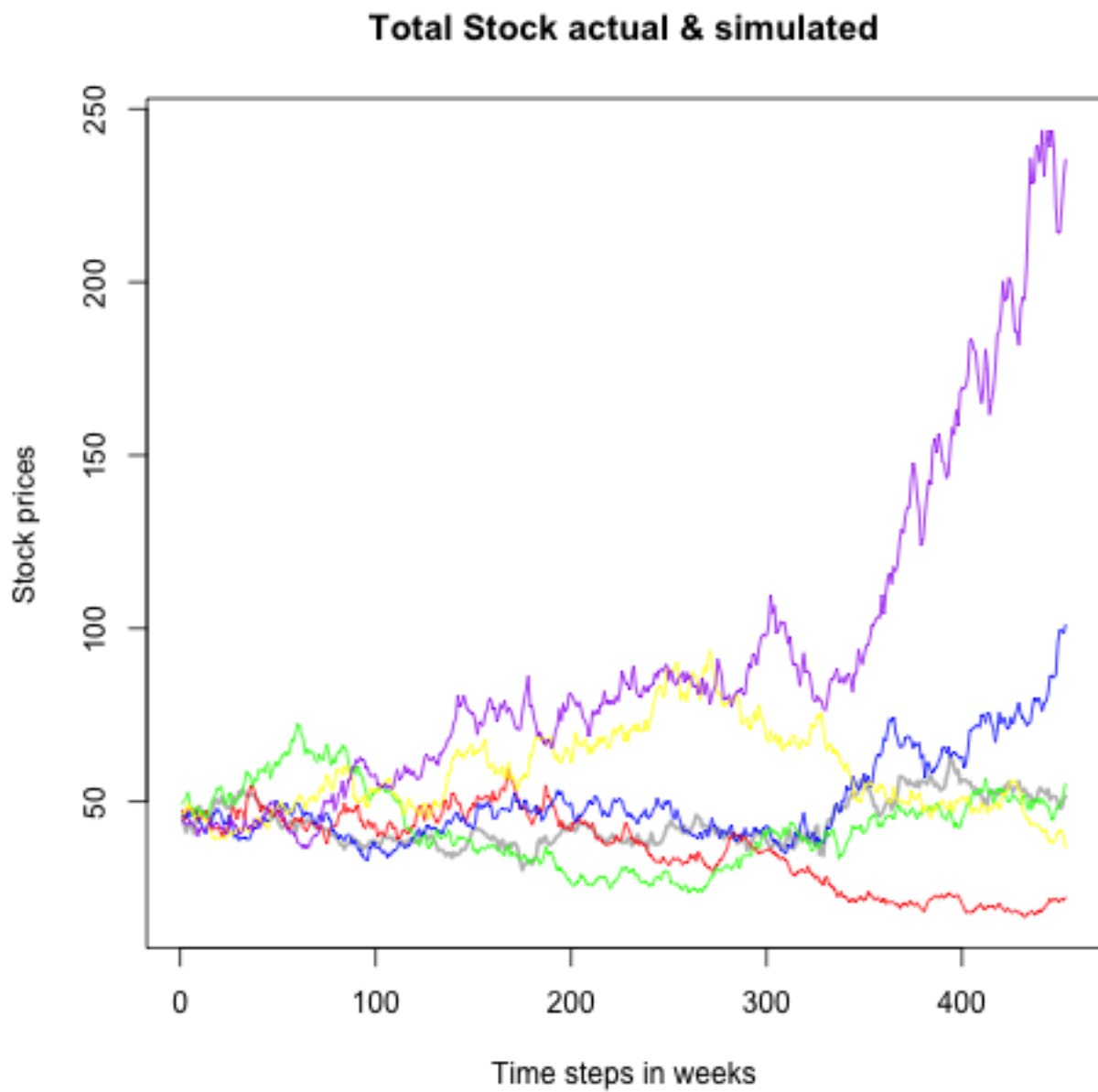


Figure 4.18 – Cut Total stock price time series. In **grey** the actual prices, in colour five different runs of the simulation

We can even get a better simulation if we choose a smaller time increment for it : the actual data is weekly, but in our simulation we can choose to take a daily time increment if we so desire. As shown below, it improves a bit the quality of the simulation.

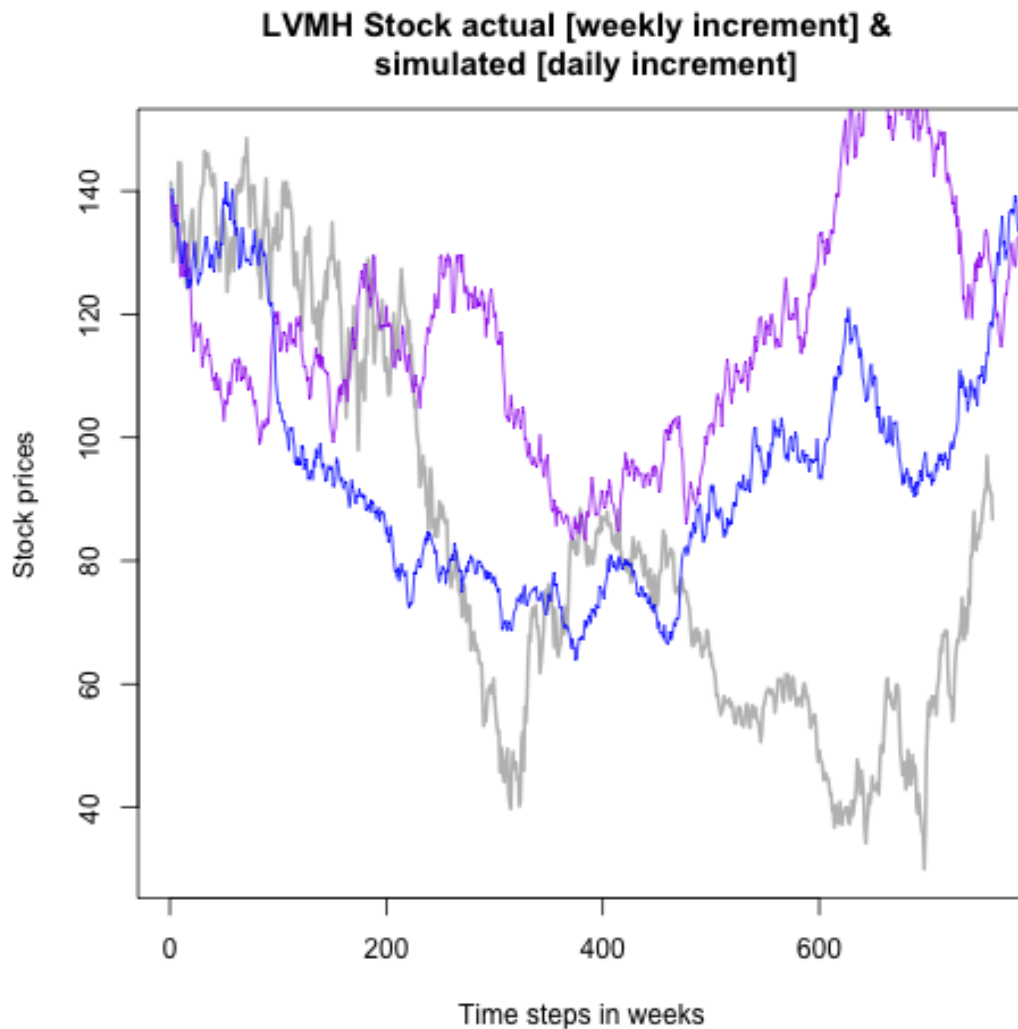


Figure 4.19 – Cut LVMH stock price time series. In **grey** the actual weekly prices, in colour two runs of the daily simulation

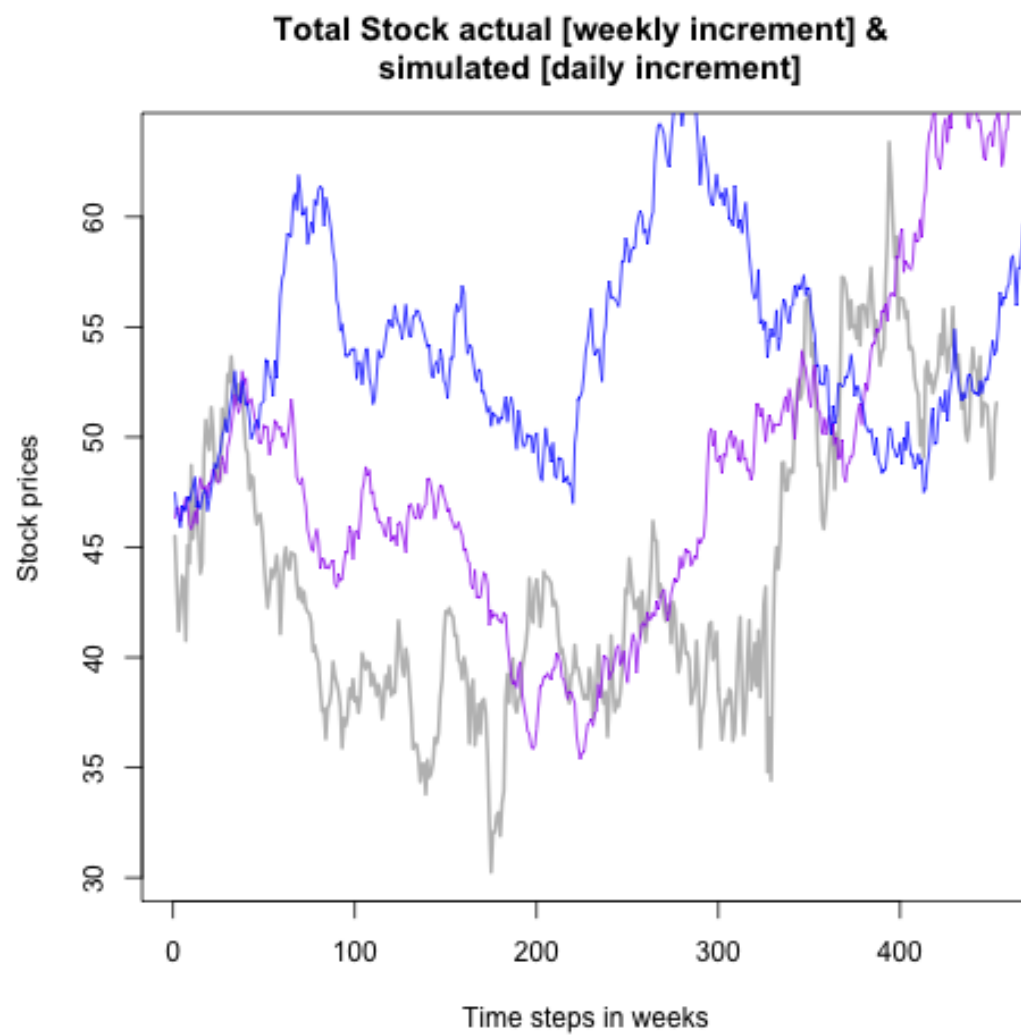


Figure 4.20 – Cut Total stock price time series. In **grey** the actual weekly prices, in colour two runs of the daily simulation

4.4 Fitting data above a threshold

Finally, we are going to try and fit a generalised extreme value distribution to our data. For each stock, we keep only the data that is above the empirical 95 %-quantile. We then call the function *gev.fit* and *gev.diag* from the R package *ismev*¹². We are particularly interested in the density plot thus obtained, as they will allow us to conclude as to the distribution fitting the data.

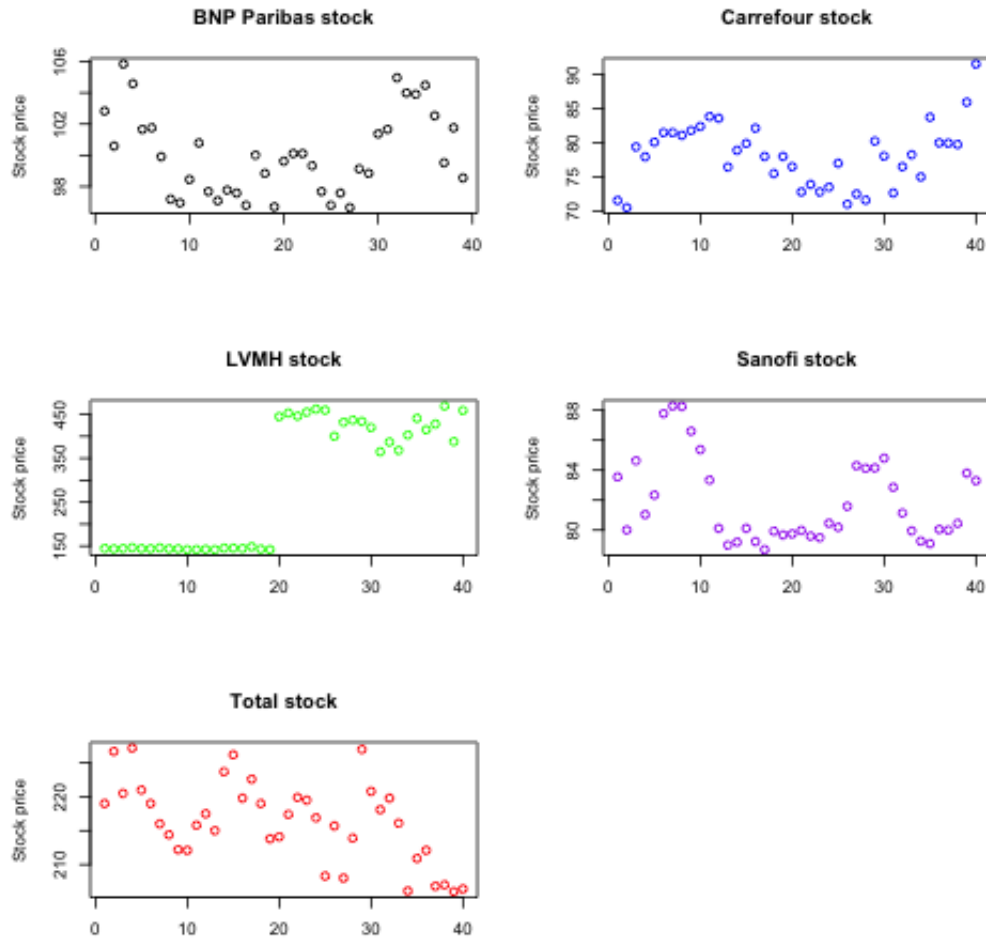


Figure 4.21 – Data above the 95 %-quantile for the BNP Paribas, Carrefour, LVMH, Sanofi and Total stocks.

The pattern followed by the LVMH data is most dissimilar to that followed by the other stocks. It suggests that the fitting may yield poor results in that case.

¹²It is based on maximum-likelihood fitting for the extreme value distribution. We have to rely on numerical resolution as there is no explicit solution to the system of maximum likelihood equations.

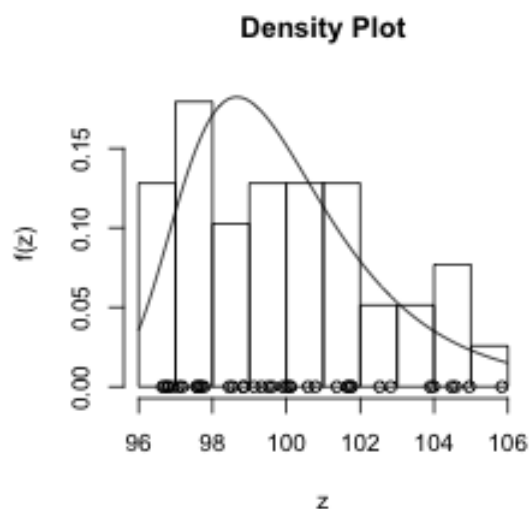


Figure 4.22 – Density fitting for the BNP Paribas data - $\hat{\gamma}_{MLE} = 0.052$

BNP Paribas stock

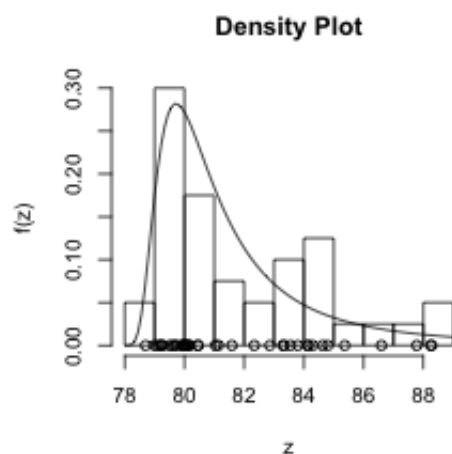


Figure 4.23 – Density fitting for the Sanofi data - $\hat{\gamma}_{MLE} = 0.48$

Sanofi stock

The curves are skewed to the right (it is particularly obvious in the case of the Sanofi stock). It corresponds to the case $\gamma > 0$ in the generalised extreme value distribution representation G_γ or, equivalently, to a Fréchet-type distribution. The values of the estimate of γ are in both cases positive, which confirms that which a naked-eye observation seemed to suggest.

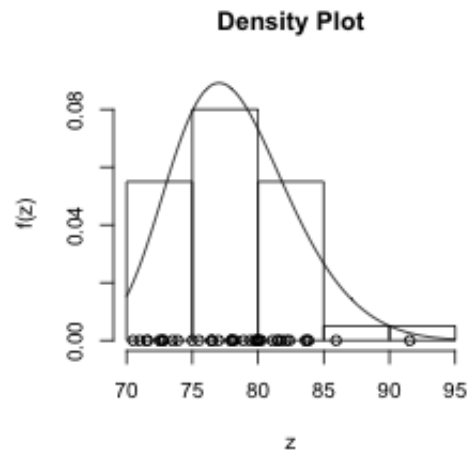


Figure 4.24 – Density fitting for the Carrefour data - $\hat{\gamma}_{MLE} = -0.16$

Carrefour stock

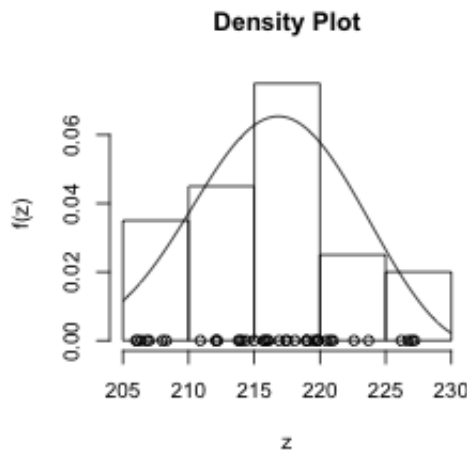


Figure 4.25 – Density fitting for the Total data - $\hat{\gamma}_{MLE} = -0.34$

Total stock

The values of the estimate of γ are in both cases negative here. Both the Carrefour and the Total above-the-threshold data thus fit a Weibull distribution (corresponds to the case $\gamma < 0$ in the generalised extreme value distribution representation G_γ). Let us note that the shapes of the curves alone might have been misleading !

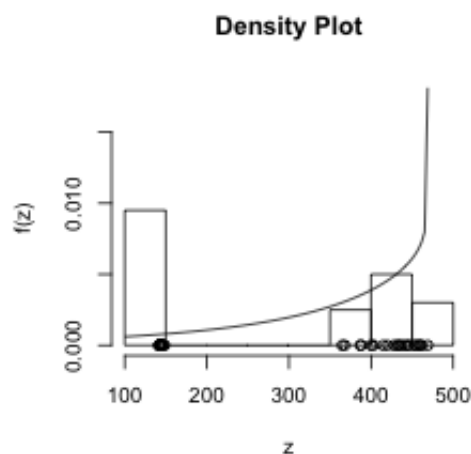


Figure 4.26 – Density fitting for the LVMH data - 95 % quantile

LVMH stock - 95 % quantile

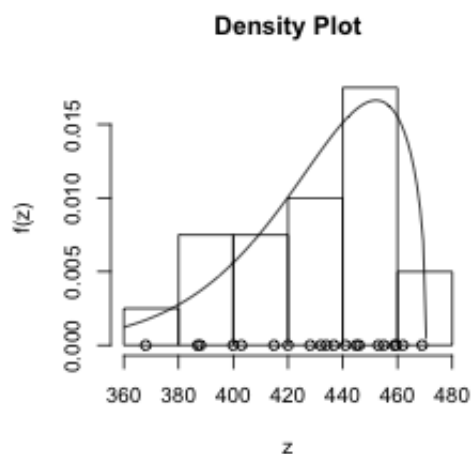


Figure 4.27 – Density fitting for the LVMH data - 97.5 % quantile - $\hat{\gamma}_{MLE} = -0.71$

LVMH stock - 97.5 % quantile

As we suspected the structure of the above the 95 %-quantile data for the LVMH stock makes the fitting arduous. To say the least, it is hard to recognize an extreme value distribution. If we take the data above the 97.5 %-quantile, however, the shape of the curve is that of a generalised extreme value distribution with $\gamma < 0$ that is a Weibull-type distribution. The estimate $\hat{\gamma}_{MLE} = -0.71$ supports that observation.

4.4.1 Goodness of fit

When assessing the goodness of fit of a cumulative distribution function to the data, in extreme value settings, it is customary to use Quantile-Quantile plots¹³ also called QQ-plots. We expect to see a straight line if the cumulative distribution function we have computed fits the data well. In practice, we have plotted regression lines to make appear those "straight lines", and the closer the actual data is to those lines, the better the fit. Below are the QQ-plots for the fitting cumulative distribution functions previously computed. For each stock, two QQ-plots have been plotted, one for the entire set of data and the other for the data above a high quantile.

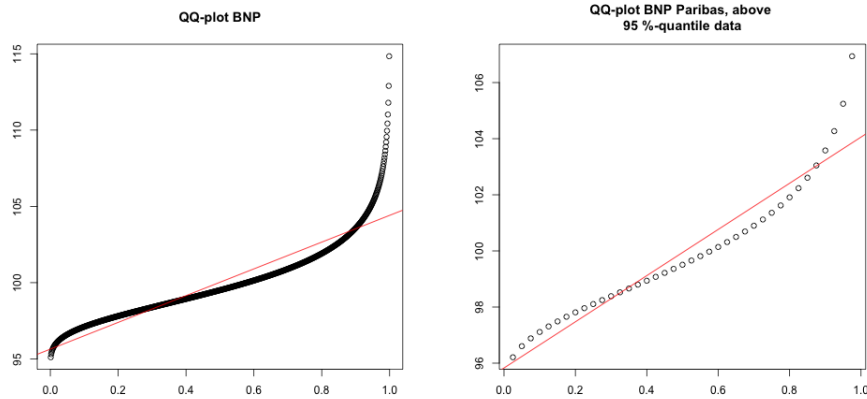


Figure 4.28 – QQ-plots for the BNP Paribas data, in **red** the regression lines.

¹³First we sort the data into the $X_{(i)}$ so as to order the sample, and then we plot the points of coordinates $(X_{(i)}, \frac{n-i+1}{n+1})$ where n is the size of the sample.

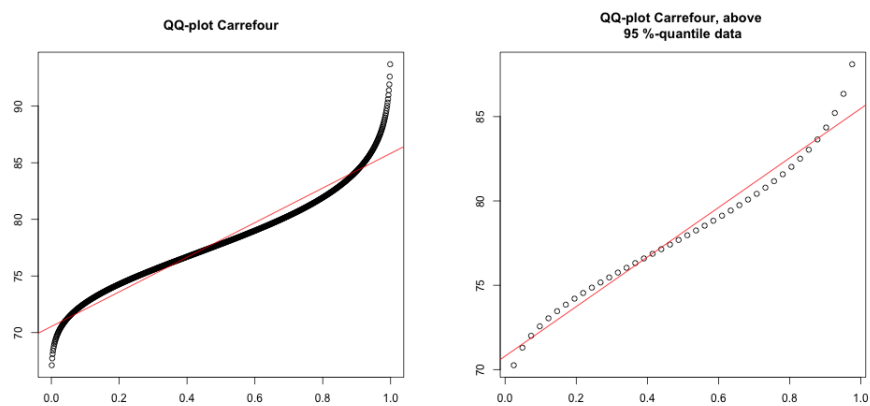


Figure 4.29 – QQ-plots for the Carrefour data, in **red** the regression lines.

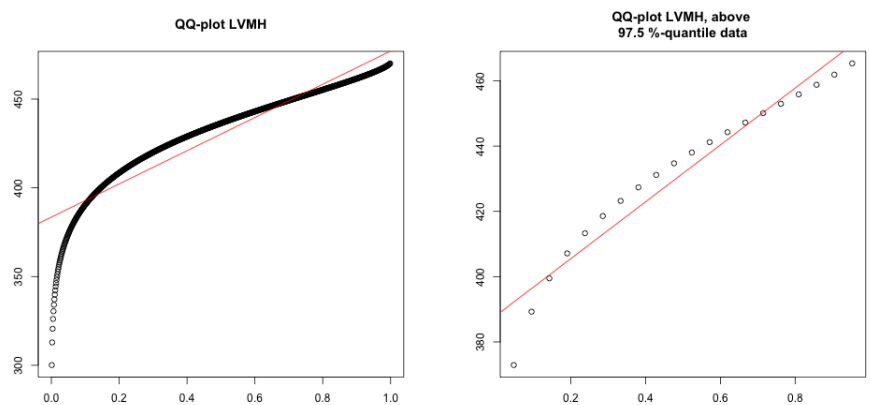


Figure 4.30 – QQ-plots for the LVMH data, in **red** the regression lines.

4.4. Fitting data above a threshold

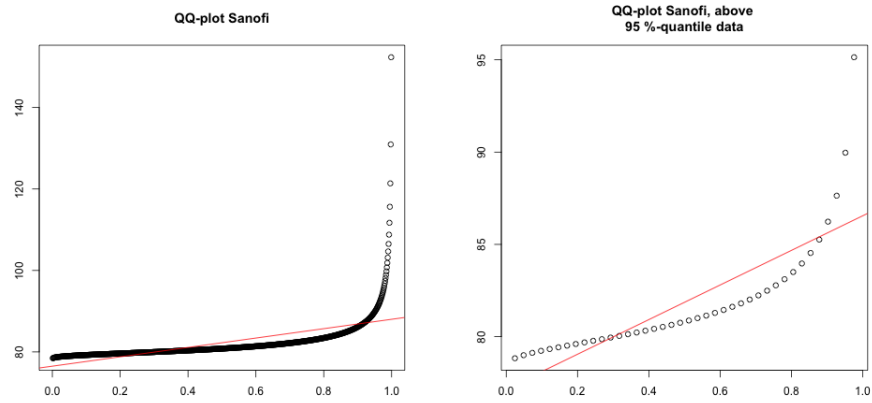


Figure 4.31 – QQ-plots for the Sanofi data, in **red** the regression lines.

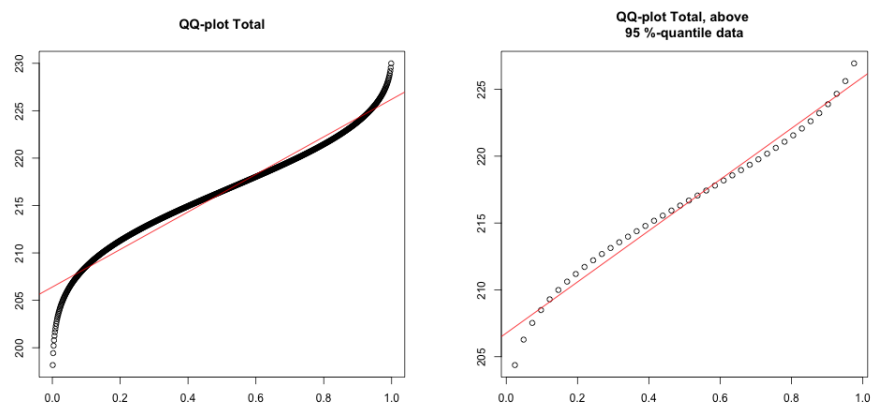


Figure 4.32 – QQ-plots for the Total data, in **red** the regression lines.

Chapter 4. Looking into financial data

For each stock, when we consider the QQ-plot for the whole of the data set, the fit is rather poor¹⁴. However, when we consider the QQ-plots for the data taken only above a high quantile, the goodness of fit is much higher, something which is particularly obvious in the case of the BNP Paribas, Carrefour, LVMH and Total stocks though it is a bit less so for the Sanofi stock.

The important lesson to be drawn here is that we were right to try and fit the data above a high quantile with extreme values distributions, something which the use of QQ-plots has confirmed.

¹⁴Though the "poorness of fit" is not the same for all the stocks : the fit is better for the BNP Paribas, Carrefour and Total stocks than for the LVMH and Sanofi stocks.

5 Conclusion

Statistics usually revolves around the behaviour of the mean. The central limit theorem is an example of this. Extreme value theory takes a very different approach, as it is the study of the behaviour of the maxima. The central limit theorem finds its extreme value counterpart in the Fisher-Tippett-Gnedenko and Von Mises' theorems. Together they make up the answer to the extremal problems : knowing what the possible limiting distributions are and knowing under which conditions there is convergence to those distributions.

Using the Black-Scholes with time independent coefficients is actually a rather good model for the simulation of stock prices, on the condition that the actual stock is not too volatile. As we have been able to see, should the actual stock display huge jumps for instance, the simulation would eventually be a poor approximation of the actual prices. Beyond the scope of the project, let us mention that stochastic jump processes are usually considered a better, more sophisticated model for stock prices. This is something we now set out to study more in-depth.

A result of note bridging the gap between on the one hand extreme value theory and finance on the other hand is that if we consider the stock prices above a high quantile (*e.g.* 95 % or 97.5 %), the empirical distributions fit extreme values distributions.

From a more personal perspective, I must say that this project has been a wonderful experience. It was the opportunity to get back to R, to do more stochastic calculus and to get started in extreme value theory. Last but not least, following a course in class, with lecture notes that were originally conceived as such is not always easy. Working on one's own on reference books, the scope of which is often much broader than what one would need is another experience entirely however. This project has taught me precious lessons, and I will not forget them.

Derivations and proofs

The two extremal problems

The extremal limit problem

Introduction

Hereafter are set forth the derivations for the part on the limiting distributions.

Our approach We are here interested in proving a result of convergence in distribution $Y_n \xrightarrow[n \rightarrow +\infty]{d} Y$. Anyone who has dabbled in probability knows that one of the classic ways to prove such a result is to use the definition and show that $\forall x$ in which F_Y is continuous, $F_{Y_n}(x) \xrightarrow[n \rightarrow +\infty]{} F_Y(x)$. Anyone who has dabbled in statistics knows that in the context of discrete random variables $(Y_n)_{n \geq 0}$, it is easier to show the convergence of the probability mass functions i.e. that $\forall x, f_{Y_n}(x) \xrightarrow[n \rightarrow +\infty]{} f_Y(x)$ and conclude by Scheffé's lemma that $(Y_n)_{n \geq 0}$ converges in distribution to Y . In the context of Extreme Value Theory, though, we will use a result based on the convergence of expectations. It is the Helly-Bray theorem :

$$Y_n \xrightarrow[n \rightarrow +\infty]{d} Y \iff \forall g \text{ continuous, bounded and real-valued functions, } E(g(Y_n)) \xrightarrow[n \rightarrow +\infty]{} E(g(Y)).$$

Remark Our approach here is based upon the notes taken while reading "Chapter 2 : The Probabilistic side of extreme Value Theory" of Statistics of Extremes, Theory and Applications by Beirlant *et alii*. **We thus want to underline that the approach here is not our own, and that we are not engaging in plagiarism but giving account of the work done on our bibliographic sources.**

Solving the problem - step 1

The answer to the extremal limit problem It turns out that all possible non-degenerate limiting distributions i.e. all extreme values distribution make up a one-parameter family $G_\gamma(x) = \exp(-(1 + \gamma x)^{-\frac{1}{\gamma}})$, where the support of G is the set $\{x : 1 + \gamma x > 0 \text{ and } \gamma \in \mathbb{R} \text{ is the}$

Chapter 5. Conclusion

Extreme Value Inded or EVI. This is what we set out to show here.

In the context of our problem $Y_n = M_n^* = \frac{M_n - b_n}{a_n}$ and $Y = Y_\gamma$. $Y_n \xrightarrow[n \rightarrow +\infty]{d} Y \iff \forall g$ continuous, bounded and real-valued functions, $E(g(Y_n)) \xrightarrow[n \rightarrow +\infty]{} \int_{-\infty}^{+\infty} g(v) dG_\gamma(v)$.

$$\begin{aligned} \Pr(\{M_n^* \leq x\}) &= F_X(x)^n \\ \implies E(g(\frac{M_n - b_n}{a_n})) &= n \int_{-\infty}^{+\infty} g(\frac{x - b_n}{a_n}) F_X(x)^{n-1} dF_X(x) \end{aligned} \quad (1)$$

$F_X : \mathbb{R} \rightarrow [0, 1]$, then if F_X is assumed to be continuous¹, by the Intermediate Value Theorem there exists in particular a solution to the equation $F(x) = 1 - \frac{v}{n}$. This is equivalent to the equation $x = U(\frac{n}{v}) = Q(1 - \frac{v}{n})$ where U is the tail quantile function of F_X and Q is the quantile function of F_X .

$$\begin{aligned} F_X(x) &= Q(1 - \frac{v}{n}) \\ \implies dF_X(x) &= -\frac{1}{n} dv \\ \implies n \int_{-\infty}^{+\infty} g(\frac{x - b_n}{a_n}) F_X(x)^{n-1} dF_X(x) &= n \int_n^0 g(\frac{U(\frac{n}{v}) - b_n}{a_n}) (1 - \frac{v}{n})^{n-1} (-\frac{1}{n}) dv \\ \implies E(g(\frac{M_n - b_n}{a_n})) &= \int_0^n g(\frac{U(\frac{n}{v}) - b_n}{a_n}) (1 - \frac{v}{n})^{n-1} dv \end{aligned} \quad (2)$$

where $x = U(\frac{n}{v}) = Q(1 - \frac{v}{n})$ goes from $-\infty$ to $+\infty$. Indeed,

- for $v = n$, $x = U(1) = Q(0) = -\infty$
- for $v = 0^+$, $x = U(+\infty) = Q(1) = +\infty$

which explains the change of bounds in the integral in the third equation. We are now working on $\int_0^n g(\frac{U(\frac{n}{v}) - b_n}{a_n}) (1 - \frac{v}{n})^{n-1} dv$. If we take the limit $n \rightarrow +\infty$, in particular :

- the integral will be taken between 0 and $+\infty$
- $(1 - \frac{v}{n})^{n-1} \xrightarrow[n \rightarrow +\infty]{} (\exp(v))^{-1} = \exp(-v)$

A limit can be obtained for $E(g(\frac{M_n - b_n}{a_n}))$ when for some sequence of positive numbers $(a_n)_{n \geq 0}$, $\frac{U(\frac{n}{v}) - b_n}{a_n}$ is convergent $\forall v \geq 0$ ²

¹An assumption that will be satisfied in all what follows

²Beirlant underlines that if we take $v = 1$, we get the idea that taking $(b_n)_{n \geq 0} = (U(n))$ will guarantee it works.

A preparatory phase A condition that must be imposed is the condition (\mathcal{C}) : "For some positive function a and $\forall u > 0$, $\frac{U(xu)-U(x)}{a(x)} \xrightarrow{x \rightarrow +\infty} h(u)$ where h is not identically equal to 0".

Proposition : The possible limits in (\mathcal{C}) are given by $ch_\gamma(u) = c \int_1^u v^{\gamma-1} dv = c \frac{u^\gamma-1}{\gamma}$, where $c > 0$, $\gamma \geq 0$ and $h_0(u)$ is interpreted as $\log(u)$.

The case $c > 0$ can be reduced to $c = 1$ by incorporating c into a .

$\forall u, v > 0$,

$$\frac{U(xuv) - U(x)}{a(x)} = \frac{U(xuv) - U(xu)}{a(xu)} \frac{a(xu)}{a(x)} + \frac{U(xu) - U(x)}{a(x)} \quad (3)$$

If (\mathcal{C}) holds then $\frac{a(xu)}{a(x)}$ converges to a $g(u)$.

$\forall u, v > 0$,

$$\begin{aligned} \frac{a(xuv)}{a(x)} &= \frac{a(xuv)}{a(xv)} \frac{a(xv)}{a(x)} \\ \implies \frac{a(xuv)}{a(x)} &= \frac{a((xv)u)}{a(xv)} \frac{a(xv)}{a(x)} \\ \xrightarrow{x \rightarrow +\infty} g(uv) &= g(u)g(v) \end{aligned} \quad (4)$$

We recognize the Cauchy Functional Equation, if a function g satisfies this equation, then $g : \mathbb{R}_+ \rightarrow \mathbb{R}$ is of the form u^γ , where γ is a real number. Now, if we write $a(x) = x^\gamma l(x)$,

$$\begin{aligned} \frac{a(xuv)}{a(x)} &= \frac{(xu)^\gamma l(xu)}{x^\gamma l(x)} \\ &= u^\gamma \frac{l(xu)}{l(x)} \end{aligned} \quad (5)$$

For this quantity to converge to u^γ , $\frac{l(xu)}{l(x)}$ must converge to 1. u being a positive number, this is equivalent to the fact that l must be a *slowly varying* function. Hence, $a(x) = x^\gamma l(x)$ is a *regularly varying* function with index of regular variation γ .

$$\begin{aligned} \frac{U(xuv) - U(x)}{a(x)} &= \frac{U(xuv) - U(xu)}{a(xu)} \frac{a(xu)}{a(x)} + \frac{U(xu) - U(x)}{a(x)} \\ \xrightarrow{x \rightarrow +\infty} h_\gamma(uv) &= h_\gamma(v)u^\gamma + h_\gamma(u) \end{aligned} \quad (6)$$

- If $\gamma = 0$

$$\begin{aligned} h_0(uv) &= h_0(u) + h_0(v) \\ \implies \exists c \in \mathbb{R} : h_0(u) &= c \log(u) \end{aligned} \quad (7)$$

- If $\gamma \neq 0$

$$h_\gamma(uv) = h_\gamma(v)u^\gamma + h_\gamma(u) \quad (8)$$

And by symmetry,

$$h_\gamma(uv) = h_\gamma(u)v^\gamma + h_\gamma(v) \quad (9)$$

Hence,

$$\begin{aligned} h_\gamma(u)v^\gamma + h_\gamma(v) &= h_\gamma(v)u^\gamma + h_\gamma(u) \\ \iff h_\gamma(u)(v^\gamma - 1) &= h_\gamma(v)(u^\gamma - 1) \\ \implies \exists d : h_\gamma(u) &= d(u^\gamma - 1) \end{aligned} \quad (10)$$

where d is a constant, if we take $d = \frac{1}{c\gamma}$, we get $ch_\gamma(u) = \frac{u^\gamma - 1}{\gamma}$

To conclude,

$$\frac{U(xu) - U(x)}{a(x)} \xrightarrow{x \rightarrow +\infty} h(u) \implies h(u) = ch_\gamma(u) \quad (11)$$

for some constant c , with the auxiliary function a regularly varying with index γ .

Back to the search of an explicit form for the limiting distributions Let us assume that the condition (\mathcal{C}) holds, with $b_n = U(n)$ and $a_n = a(n)$.

$$\begin{aligned} E(g(\frac{M_n - b_n}{a_n})) &= \int_0^n g(\frac{U(\frac{n}{v}) - b_n}{a_n})(1 - \frac{v}{n})^{n-1} dv \\ &= \int_0^n g(\frac{U(n\frac{1}{v}) - U(n)}{a(n)})(1 - \frac{v}{n})^{n-1} dv \\ &\xrightarrow{x \rightarrow +\infty} \int_0^{+\infty} g(h_\gamma(\frac{1}{v})) \exp(-v) dv \end{aligned} \quad (12)$$

The last integral must be equal to $\int_{-\infty}^{+\infty} g(u) dG_\gamma(u)$

Solving the problem - step 2 : expliciting the G_γ in earnest

Case $\gamma = 0$ We make the following change of variables : $u = h_0(\frac{1}{v}) = \log(\frac{1}{v}) = -\log(v)$.
 $u = \log(v) \iff v = \exp(-u)$, with $u \in \mathbb{R}$, $v \in]0, +\infty[$.

$$\begin{aligned}
 du &= -\frac{1}{v} dv \\
 \implies \int_0^{+\infty} g(u) \exp(-v) dv &= \int_0^{+\infty} g(u) \exp(-\exp(-u))(-v) du \\
 \implies \int_0^{+\infty} g(u) \exp(-v) dv &= \int_0^{+\infty} g(u) \exp(-\exp(-u))(-\exp(-u)) du \\
 \implies \int_0^{+\infty} g(u) \exp(-v) dv &= \int_{-\infty}^{+\infty} g(u)(-\exp(-u) \exp(-\exp(-u))) du \\
 \implies \int_0^{+\infty} g(u) \exp(-v) dv &= \int_{-\infty}^{+\infty} g(u) d(\exp(-\exp(-u)))
 \end{aligned} \tag{13}$$

Finally, $\gamma > 0 \implies \underline{G_\gamma(u) = \exp(-\exp(-u))}$

Case $\gamma \neq 0$ We make the following change of variables :

$$\begin{aligned}
 u &= h_\gamma(\frac{1}{v}) = \frac{(\frac{1}{v})^\gamma - 1}{\gamma} \\
 \iff (\frac{1}{v})^\gamma &= \gamma u + 1 \\
 \iff \frac{1}{v} &= (\gamma u + 1)^{\frac{1}{\gamma}} \\
 \iff v &= (\gamma u + 1)^{-\frac{1}{\gamma}}
 \end{aligned} \tag{14}$$

$$\begin{aligned}
 dv &= -\frac{1}{\gamma} \gamma (\gamma u + 1)^{-\frac{1}{\gamma}-1} du \\
 \implies dv &= -(\gamma u + 1)^{-\frac{\gamma+1}{\gamma}} du
 \end{aligned}$$

v varies from 0^+ to $+\infty$,

- If $\gamma > 0$, $u = \frac{(\frac{1}{v})^\gamma - 1}{\gamma}$, varying from $-\infty$ to $-\gamma^{-1}$.
- If $\gamma < 0$, $u = -\frac{v^{-\gamma} + 1}{-\gamma}$, varying from $-\gamma^{-1}$ to $-\infty$.

$$\begin{aligned}
 \int_0^{+\infty} g(u) \exp(-v) dv &= \int_{+\infty/-\gamma^{-1}}^{-\gamma^{-1}/-\infty} g(u) \exp(-(1 + \gamma u)^{-\frac{1}{\gamma}})(-(1 + \gamma u)^{-\frac{\gamma+1}{\gamma}}) du \\
 &= \int_{-\gamma^{-1}/-\infty}^{+\infty/-\gamma^{-1}} g(u) (1 + \gamma u)^{-\frac{\gamma+1}{\gamma}} \exp(-(1 + \gamma u)^{-\frac{1}{\gamma}}) du
 \end{aligned} \tag{15}$$

$$d(\exp(-(1 + \gamma u)^{-\frac{1}{\gamma}})) = (1 + \gamma u)^{-\frac{\gamma+1}{\gamma}} \exp(-(1 + \gamma u)^{-\frac{1}{\gamma}}) du \quad (16)$$

Finally, the last integral in (2.15) is equal to,

- If $\gamma > 0$, $\int_{-\gamma^{-1}}^{+\infty} g(u) d(\exp(-(1 + \gamma u)^{-\frac{1}{\gamma}}))$
- If $\gamma < 0$, $\int_{-\infty}^{-\gamma^{-1}} g(u) d(\exp(-(1 + \gamma u)^{-\frac{1}{\gamma}}))$

Finally, to sum it up :

- $\gamma = 0$: Gumbel distribution
 $G_\gamma(u) = \exp(-\exp(-u)), u \in \mathbb{R}$
- $\gamma > 0$: Fréchet distribution
 $G_\gamma(u) = \exp(-(1 + \gamma u)^{-\frac{1}{\gamma}}), u \in]-\gamma^{-1}, +\infty[$
- $\gamma < 0$: Weibull distribution
 $G_\gamma(u) = \exp(-(1 + \gamma u)^{-\frac{1}{\gamma}}), u \in]-\infty, -\gamma^{-1}[$

Additional figures

Chapter 4 : Looking into financial data

Stock	Cumulative residuals between $\exp(R_t)$ and $1 + R_t$
BNP	3.22E-15
Carrefour	1.44E-15
LVMH	4.75E-15
Sanofi	1.55E-15
Total	3.86E-15

Table 1 – Cumulative residuals between $\exp(R_t)$ and $1 + R_t$ for the BNP, Carrefour, LVMH, Sanofi, Total stocks



Bibliography

- Modelling Extremal Events, Embrechts, Kluppelberg and Mikosch
- Statistics of Extremes, Theory and Application, Beirlant, Goegebeur, Segers and Teugels
- An Introduction to Statistical Modeling of Extreme Values, Coles
- Introduction to Scientific Programming and Simulation Using R, Jones, Maillardet and Robinson

Recommendations on the use of the bibliography

First and foremost, it must be said that **all the books included in the Bibliography section are good books** (I will focus on the books dealing with Extreme Value Theory ; the book on R programming is a very fine one though, with a nice hands-on approach). However, **they may not be intended for the same uses and the same users**. I am giving my opinion on those books in case someone chances upon the repository and they want to learn more about Extreme Value Theory. **It is only my opinion, nothing more.**

- An Introduction to Statistical Modelling of Extreme Values, by Stuart Coles (Springer) is not surprisingly an introductory book; most of the time, at best a sketch of the proof is given. It may be best used in conjunction with a traditional course given by a teacher that will go more in-depth themselves. I would not recommend it as a primary source for a master thesis project, however.
- Statistics of Extremes, Theory and Applications, by Beirlant, Goegebeur, Segers and Teugels (Wiley) is a balanced book, going enough into details. It is thus fit as a source for a master thesis project. The only thing going against it is the following : many figures to illustrate what is being explained is a good thing, but it would have been nice to detail a bit more the derivations of the proofs, at times.
- Modelling extremal events for Insurance and Finance, by Embrechts, Kluppelberg and Mikosch (Springer) is nothing short of a summa. It is extremely detailed and thorough. I would not recommend it, however, as a source for a master thesis project as it would be a bit too ambitious to use it for that purpose. If you have more time and are ready for an involved reading, all the more so if you are interested in the financial applications of Extreme Value Theory, search for a companion no further!