Extreme Values in Financial Statistics

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

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 ENSEEIHT
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

[To do later]

Lausanne, 14 Août 2015 K. M–H.

Preface

[To do later]

Lausanne, 14 Août 2015 K. M–H.

Abstract

Résumé

Contents

List of Figures

List of Tables

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 dams[1]. 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 fields[2], 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

[1] As was done in The Netherlands beginning in the fifties

[2] including climate science, seismology, insurance etc.

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 $n \longrightarrow +\infty$. The limiting distribution will be degenerate. Indeed, let us consider the upper end-point of F_X[4], z^+. Then,

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

It turns out we cannot use the limiting distribution directly. A common approach[5] is to consider a sequence of the maxima, centred and normalised.

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***.

[3]If 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.

[4]that 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 : $z^+ = \inf\{z : F_X(z) \geq 1\}$.

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

2 The two extremal problems

2.1 The extremal limit problem

2.1.1 Introduction

Our approach We are here interested in proving a result of convergence in distribution $Y_n \xrightarrow[n\to+\infty]{d} Y$. 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, $F_{Y_n}(x) \xrightarrow[n\to+\infty]{} F_Y(x)$. Anyone who has dabbled in Statistics knows that in the context of discrete random variables $(Y_n)_{n\geq 0}$, Y it is easier to show the convergence of the probability mass functions i.e. that $\forall x, f_{Y_n}(x) \xrightarrow[n\to+\infty]{} f_Y(x)$ 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 :

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

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.**

2.1.2 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 $\{x : 1+\gamma x > 0$ and $\gamma \in \mathbb{R}$ is the **E**xtreme **V**alue **I**nded 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 $Y = Y_\gamma$. $Y_n \xrightarrow[n\to+\infty]{d} Y \iff \forall g$ continuous, bounded and real-valued functions, $\mathrm{E}(g(Y_n)) \xrightarrow[n\to+\infty]{} \int_{-\infty}^{+\infty} g(v)\,\mathrm{d}G_\gamma(v)$.

$$\begin{aligned} &\Pr(\{M_n^* \le x\}) = F_X(x)^n \\ \implies &\mathrm{E}(g(\frac{M_n - b_n}{a_n})) = n\int_{-\infty}^{+\infty} g(\frac{x-b_n}{a_n}) F_X(x)^{n-1}\,\mathrm{d}F_X(x) \end{aligned} \tag{2.1}$$

$F_X : \mathbb{R} \to [0,1]$, then if F_X is assumed to be continuous[1], 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.

$$\begin{aligned} F_X(x) &= Q(1-\frac{v}{n}) \\ \implies \mathrm{d}F_X(x) &= -\frac{1}{n}\mathrm{d}v \\ \implies n\int_{-\infty}^{+\infty} g(\frac{x-b_n}{a_n}) F_X(x)^{n-1}\,\mathrm{d}F_X(x) &= n\int_n^0 g(\frac{U(\frac{n}{v})-b_n}{a_n})(1-\frac{v}{n})^{n-1}(-\frac{1}{n})\,\mathrm{d}v \\ \implies \mathrm{E}(g(\frac{M_n-b_n}{a_n})) &= \int_0^n g(\frac{U(\frac{n}{v})-b_n}{a_n})(1-\frac{v}{n})^{n-1}\,\mathrm{d}v \end{aligned} \tag{2.2}$$

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 $\int_0^n g(\frac{U(\frac{n}{v})-b_n}{a_n})(1-\frac{v}{n})^{n-1}\,\mathrm{d}v$. If we take the limit $n \to +\infty$, in particular :

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

A limit can be obtained for $\mathrm{E}(g(\frac{M_n-b_n}{a_n}))$ when for some sequence of positive numbers $(a_n)_{n\ge0}$, $\frac{U(\frac{n}{v})-b_n}{a_n}$ is convergent $\forall\, v \ge 0$[2]

[1] An assumption that will be satisfied in all what follows

[2] Beirlant underlines that if we take $v = 1$, we get the idea that taking $(b_n)_{n\ge0} = (U(n))$ will guarantee it works.

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

Proposition : The possible limits in $(\mathscr{C})$ are given by $ch_\gamma(u) = c\int_1^u v^{\gamma-1}\,\mathrm{d}v = c\frac{u^\gamma-1}{\gamma}$, 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.

$\forall u,\ v > 0$,

$$\frac{U(xuv)-U(x)}{a(x)} = \frac{U(xuv)-U(xu)}{a(xu)}\frac{a(xu)}{a(x)} + \frac{U(xu)-U(x)}{a(x)} \tag{2.3}$$

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

$\forall u,\ v > 0$,

$$\begin{aligned}
\frac{a(xuv)}{a(x)} &= \frac{a(xuv)}{a(xv)}\frac{a(xv)}{a(x)} \\
\Longrightarrow \frac{a(xuv)}{a(x)} &= \frac{a((xv)u)}{a(xv)}\frac{a(xv)}{a(x)} \\
&\xrightarrow[x\to+\infty]{} g(uv) = g(u)g(v)
\end{aligned} \tag{2.4}$$

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

$$\begin{aligned}
\frac{a(xuv}{a(x)} &= \frac{(xu)^\gamma}{x^\gamma}\frac{l(xu)}{l(x)} \\
&= u^\gamma \frac{l(xu)}{l(x)}
\end{aligned} \tag{2.5}$$

For this quantity to converge to u^γ, $\frac{l(xu)}{l(x)}$ must converge to . u being a positive, 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.

$$\begin{aligned}
\frac{U(xuv)-U(x)}{a(x)} &= \frac{U(xuv)-U(xu)}{a(xu)}\frac{a(xu)}{a(x)} + \frac{U(xu)-U(x)}{a(x)} \\
&\xrightarrow[x\to+\infty]{} h_\gamma(uv) = h_\gamma(v)u^\gamma + h_\gamma(u)
\end{aligned} \tag{2.6}$$

- If $\gamma = 0$

$$\begin{aligned}
h_0(uv) &= h_0(u) + h_0(v) \\
\Longrightarrow \exists c \in \mathbb{R} &: h_0(u) = c\log(u)
\end{aligned} \tag{2.7}$$

- If $\gamma \neq 0$

$$h_\gamma(uv) = h_\gamma(v)u^\gamma + h_\gamma(u) \tag{2.8}$$

And by symmetry,

$$h_\gamma(uv) = h_\gamma(u)v^\gamma + h_\gamma(v) \tag{2.9}$$

Hence,

$$\begin{aligned} & h_\gamma(u)v^\gamma + h_\gamma(v) = h_\gamma(v)u^\gamma + h_\gamma(u) \\ \iff & h_\gamma(u)(v^\gamma - 1) = h_\gamma(v)(u^\gamma - 1) \\ \implies & \exists d : h_\gamma(u) = d(u^\gamma - 1) \end{aligned} \tag{2.10}$$

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,

$$\frac{U(xu) - U(x)}{a(x)} \xrightarrow[x \to +\infty]{} h(u) \implies h(u) = ch_\gamma(u) \tag{2.11}$$

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 $(\mathscr{C})$ holds, with $b_n = U(n)$ and $a_n = a(n)$.

$$\begin{aligned} \mathrm{E}(g(\frac{M_n - b_n}{a_n})) &= \int_0^n g(\frac{U(\frac{n}{v}) - b_n}{a_n})(1 - \frac{v}{n})^{n-1}\,\mathrm{d}v \\ &= \int_0^n g(\frac{U(n\frac{1}{v}) - U(n)}{a(n)})(1 - \frac{v}{n})^{n-1}\,\mathrm{d}v \\ &\xrightarrow[x \to +\infty]{} \int_0^{+\infty} g(h_\gamma(\frac{1}{v}))\exp(-v)\,\mathrm{d}v \end{aligned} \tag{2.12}$$

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

2.1.3 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 $u \in \mathbb{R}$, $v \in]0, +\infty[$.

$$
\begin{aligned}
& \mathrm{d}u = -\frac{1}{v}\mathrm{d}v \\
\implies & \int_0^{+\infty} g(u)\exp(-v)\,\mathrm{d}v = \int_0^{+\infty} g(u)\exp(-\exp(-u))(-v)\,\mathrm{d}u \\
\implies & \int_0^{+\infty} g(u)\exp(-v)\,\mathrm{d}v = \int_0^{+\infty} g(u)\exp(-\exp(-u))(-\exp(-u))\,\mathrm{d}u \\
\implies & \int_0^{+\infty} g(u)\exp(-v)\,\mathrm{d}v = \int_{-\infty}^{+\infty} g(u)(-\exp(-u)\exp(-\exp(-u)))\,\mathrm{d}u \\
\implies & \int_0^{+\infty} g(u)\exp(-v)\,\mathrm{d}v = \int_{-\infty}^{+\infty} g(u)\,\mathrm{d}(\exp(-\exp(-u)))
\end{aligned}
\tag{2.13}
$$

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

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

$$
\begin{aligned}
& u = h_\gamma(\frac{1}{v}) = \frac{(\frac{1}{v})^\gamma - 1}{\gamma} \\
\iff & (\frac{1}{v})^\gamma = \gamma u + 1 \\
\iff & \frac{1}{v} = (\gamma u + 1)^{\frac{1}{\gamma}} \\
\iff & v = (\gamma u + 1)^{-\frac{1}{\gamma}}
\end{aligned}
\tag{2.14}
$$

$$
\begin{aligned}
& \mathrm{d}v = -\frac{1}{\gamma}\gamma(\gamma u + 1)^{-\frac{1}{\gamma}-1} mathrmdu \\
\implies & \mathrm{d}v = -(\gamma u + 1)^{-\frac{\gamma+1}{\gamma}}\,\mathrm{d}u
\end{aligned}
$$

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

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

$$
\begin{aligned}
\int_0^{+\infty} g(u)\exp(-v)\,\mathrm{d}v &= \int_{+\infty/-\gamma^{-1}}^{-\gamma^{-1}/-\infty} g(u)\exp(-(1+\gamma u)^{-\frac{1}{\gamma}})(-(1+\gamma u)^{-\frac{\gamma+1}{\gamma}})\,\mathrm{d}u \\
&= \int_{-\gamma^{-1}/-\infty}^{+\infty/-\gamma^{-1}} g(u)(1+\gamma u)^{-\frac{\gamma+1}{\gamma}}\exp(-(1+\gamma u)^{-\frac{1}{\gamma}})\,\mathrm{d}u
\end{aligned}
\tag{2.15}
$$

$$\mathrm{d}(\exp(-(1+\gamma u)^{-\frac{1}{\gamma}})) = (1+\gamma u)^{-\frac{\gamma+1}{\gamma}} \exp(-(1+\gamma u)^{-\frac{1}{\gamma}})\mathrm{d}u \tag{2.16}$$

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

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

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

Finally, to sum it up :

- <u>$\gamma = 0$:</u> **Gumbel distribution**
  $G_\gamma(u) = \exp(-\exp(-u)),\ u \in \mathbb{R}$

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

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

2.2 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[3] 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 them. This is the historical approach, and a rather straightforward one, still by no means are alternative approaches uninteresting[4].

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

[3] F_X being the distribution of the X_i of the sample.
[4] In particular, conditions based on the sole behaviour of F_X can be formulated.

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 $xr(x) \xrightarrow[x\to+\infty]{} \alpha > 0$, then $F_X \in \mathscr{D}(\Phi_\alpha)$.
2. If $x^+ < +\infty$ and $(x^+ - x)r(x) \xrightarrow[x\to x^+]{} \alpha > 0$, then $F_X \in \mathscr{D}(\Psi_\alpha)$.
3. If $r(x)$ is ultimately positive in the neighbourhood of x^+, is differentiable on that neighbourhood and is such that $\frac{\mathrm{d}r}{\mathrm{d}x}(x) \xrightarrow[x\to x^+]{} 0$, then $F_X \in \mathscr{D}(\Delta)$.

2.3 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*.

3 Investigating results on the limiting distribution

3.1 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 $\frac{1}{n}$-quantiles of the distributions (distributions of the sample, not of the maxima) as a function of the time steps n. This will lead us to make an interesting observation.

3.1.1 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 $\forall\, t > 0$:

$$(\frac{1}{t} - \frac{1}{t^3})\frac{\exp(-\frac{t^2}{2})}{\sqrt{2\pi}} < 1 - \Phi(t) < \frac{1}{t}\frac{\exp(-\frac{t^2}{2})}{\sqrt{2\pi}} \tag{3.1}$$

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

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

$$1 - \Phi(t) \sim_{t \to +\infty} \frac{1}{t} \frac{\exp(-\frac{t^2}{2})}{\sqrt{2\pi}} \tag{3.2}$$

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

$$\begin{aligned}
& F_X(q_{\frac{1}{n}}) = 1 - \frac{1}{n} \\
\iff \quad & \frac{1}{q_{\frac{1}{n}}} \frac{\exp(-\frac{q_{\frac{1}{n}}^2}{2})}{\sqrt{2\pi}} = \frac{1}{n} \\
\iff \quad & \log(q_{\frac{1}{n}}) + \log(\exp(-\frac{q_{\frac{1}{n}}^2}{2})) + \log(\sqrt{2\pi}) = \log(n)
\end{aligned} \tag{3.3}$$

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

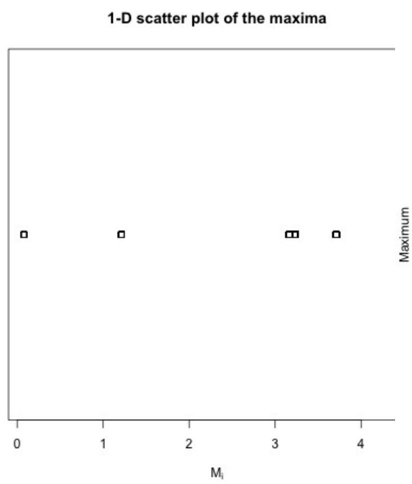

Figure 3.2 – Scatter Plot of the Maxima, n = 10000

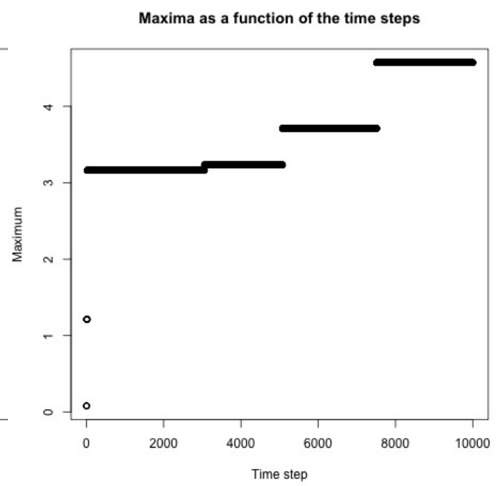

Figure 3.3 – Maxima against the time steps

[2]See the fourth of the figures below

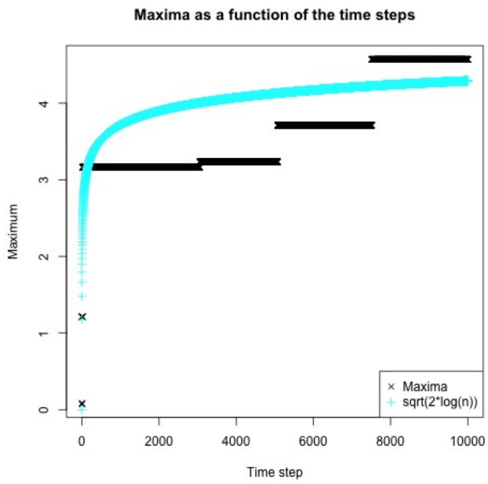

Figure 3.4 – Maxima against the time steps and function $n \to \sqrt{2\log(n)}$

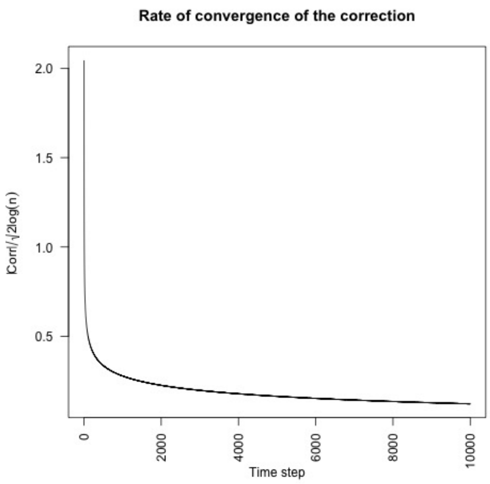

Figure 3.5 – The correction becomes negligible compared to the starting term as n grows large

3.1.2 Sample following a Cauchy distribution

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

$$
\begin{aligned}
& F_X(q_{\frac{1}{n}}) = 1 - \frac{1}{n} \\
\iff \quad & \frac{\arctan(q_{\frac{1}{n}})}{\pi} + \frac{1}{2} = 1 - \frac{1}{n} \\
\iff \quad & \frac{\arctan(q_{\frac{1}{n}})}{\pi} = \frac{2-n}{n} \\
\iff \quad & q_{\frac{1}{n}} = \tan\left(\frac{\pi}{2}\frac{2-n}{n}\right)
\end{aligned}
\tag{3.4}
$$

Figure 3.6 – Below, a realisation of the sequence of maxima for i.i.d. *Cauchy(0,1)* RVs

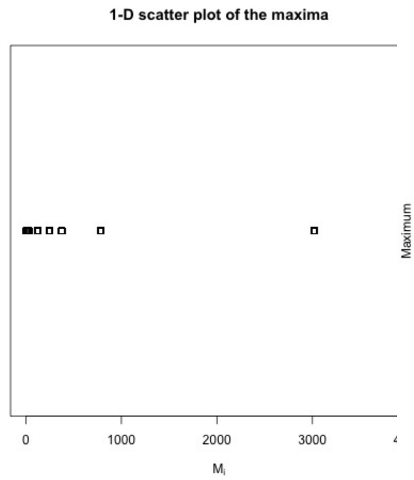

Figure 3.7 – Scatter Plot of the Maxima, n = 10000

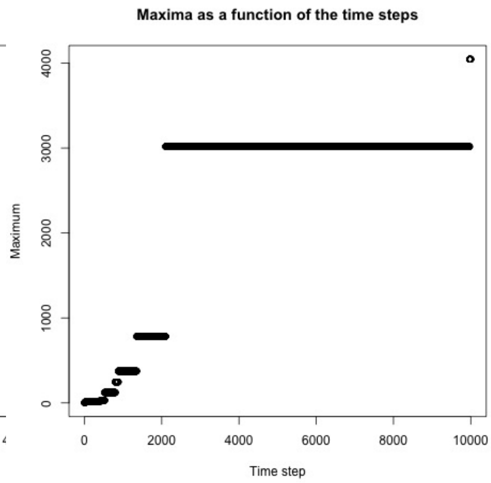

Figure 3.8 – Maxima against the time steps

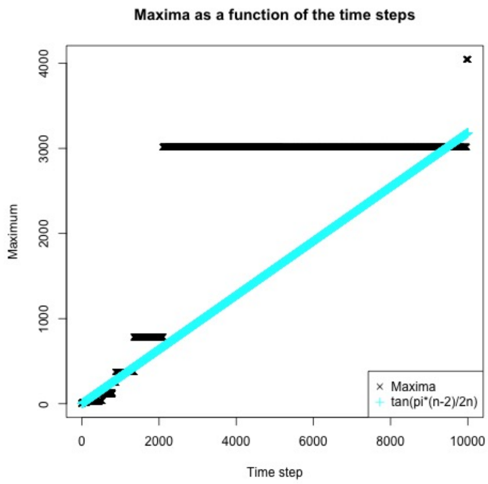

Figure 3.9 – Maxima against the time steps and function $n \to \tan(\pi\frac{2-n}{2n})$

3.1.3 Sample following an Exponential Distribution

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

$$\begin{aligned} & F_X(q_{\frac{1}{n}}) = 1 - \frac{1}{n} \\ \iff \quad & 1 - \exp(-\lambda q_{\frac{1}{n}}) = 1 - \frac{1}{n} \\ \iff \quad & q_{\frac{1}{n}} = \frac{1}{\lambda}\log(n) \end{aligned} \tag{3.5}$$

Figure 3.10 – Below, a realisation of the sequence of maxima for i.i.d. *Exp(1)* RVs

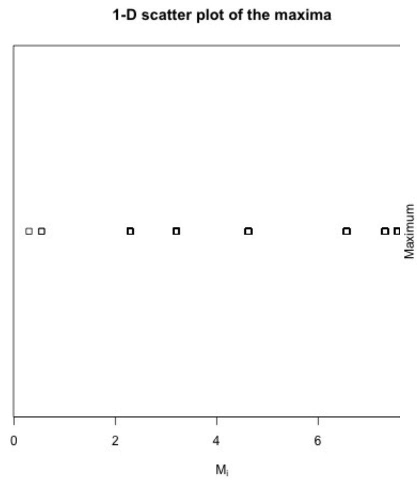

Figure 3.11 – Scatter Plot of the Maxima, n = 10000

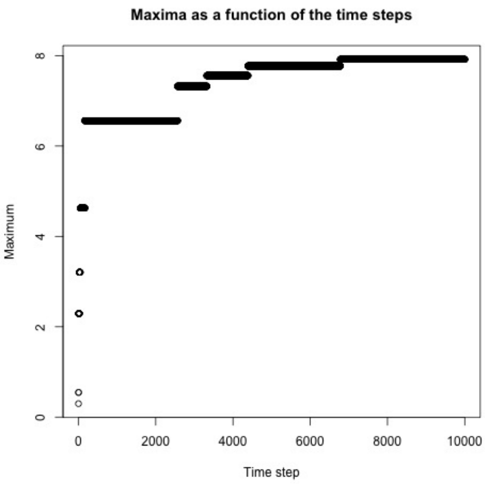

Figure 3.12 – Maxima against the time steps

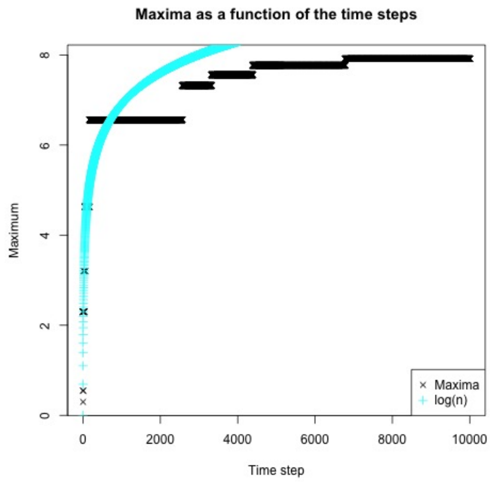

Figure 3.13 – Maxima against the time steps and function $n \to \log(n)$

3.1.4 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 $\frac{1}{n}$-quantiles are a good proxy for the M_n.

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 $\frac{1}{n}$-quantiles of F_{X_i}. Of course, it is only an intuition !

3.2 What are the limiting distributions in theses cases ?

Gumbel 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.

$$f_X(x) = \frac{1}{\pi}\frac{1}{1+x^2} \tag{3.6}$$

$$F_X(x) = \frac{1}{\pi}\arctan(x) + \frac{1}{2} \tag{3.7}$$

$$\begin{aligned}
r(x) &= \frac{f_X(x)}{1-F_X(x)} \\
&= \frac{\frac{1}{1+x^2}}{\frac{\pi}{2}-\arctan(x)} \\
&= \frac{\frac{1}{1+x^2}}{\frac{\pi}{2}-(\frac{\pi}{2}-\arctan(\frac{1}{x}))} \\
&= \frac{1}{(1+x^2)\arctan(\frac{1}{x})} \\
&= \frac{1}{(1+x^2)(\frac{1}{x}+o(\frac{1}{x^2}))} \\
&= \frac{1}{\frac{1+x^2}{x}+o(1)} \\
\implies xr(x) &= \frac{x^2}{x^2+1+o(1)} \\
\implies xr(x) &\xrightarrow[x\to+\infty]{} 1
\end{aligned} \tag{3.8}$$

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

$$f_X(x) = \lambda\exp(-\lambda x) \tag{3.9}$$

$$F_X(x) = 1-\exp(-\lambda x) \tag{3.10}$$

$$\begin{aligned} r(x) &= \frac{f_X(x)}{1 - F_X(x)} \\ &= \frac{\lambda \exp(-\lambda x)}{1 - (1 - \exp(-\lambda x))} \\ &= \frac{\lambda \exp(-\lambda x)}{\exp(-\lambda x)} \\ &= \lambda \\ \implies \frac{\mathrm{d}r}{\mathrm{d}x}(x) &= 0 \end{aligned} \tag{3.11}$$

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

Fréchet 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 Fréchet-type distribution.[3]

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, $x^+ = 1 < +\infty$.

$$f_X(x) = 1 \tag{3.12}$$

$$F_X(x) = x \tag{3.13}$$

$$\begin{aligned} r(x) &= \frac{f_X(x)}{1 - F_X(x)} \\ &= \frac{1}{1-x} \\ \implies (x^+ - x) r(x) &= (x^+ - x)\frac{1}{1-x} \\ \implies (1-x) r(x) &= (1-x)\frac{1}{1-x} = 1 \\ \implies (x^+ - x) r(x) &\xrightarrow[x \to x^+]{} 1 > 0 \end{aligned} \tag{3.14}$$

[3] The computation is more involved in that case than with the $Exp(1)$ sample.

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.

4 Looking into real-world data

4.1 Five real-world 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 stocks[2]

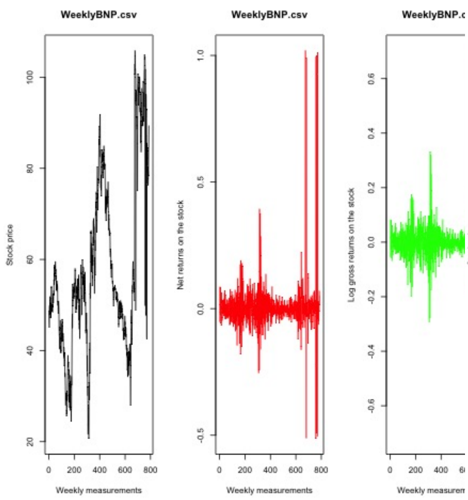

Figure 4.1 – 15 years of weekly BNP Stock Price Data

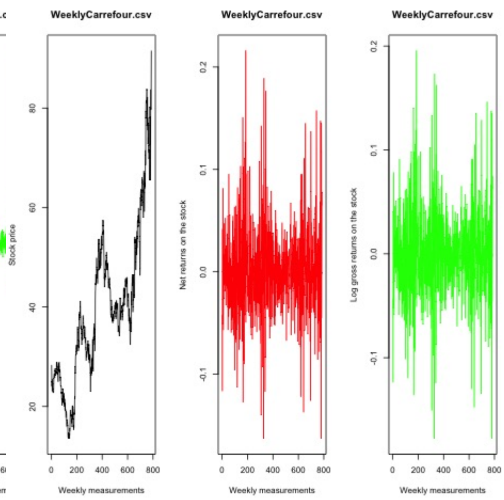

Figure 4.2 – 15 years of weekly Carrefour Stock Price Data

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

[2] Both quantities are widely used in Finance.

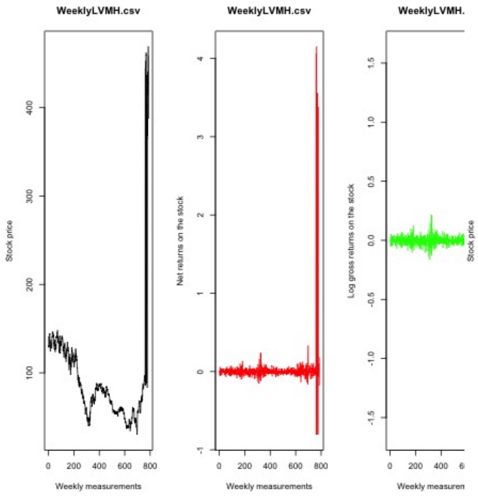

Figure 4.3 – 15 years of weekly LVMH Stock Price Data

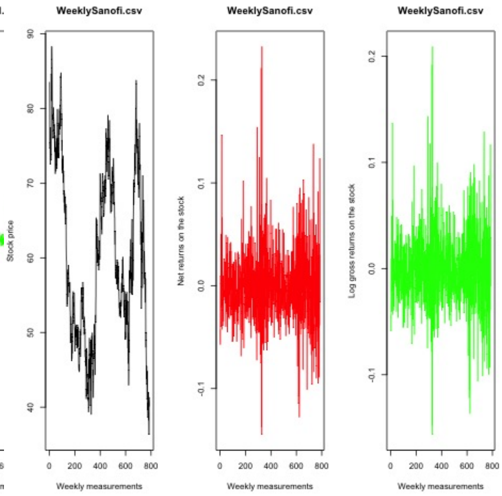

Figure 4.4 – 15 years of weekly Sanofi Stock Price Data

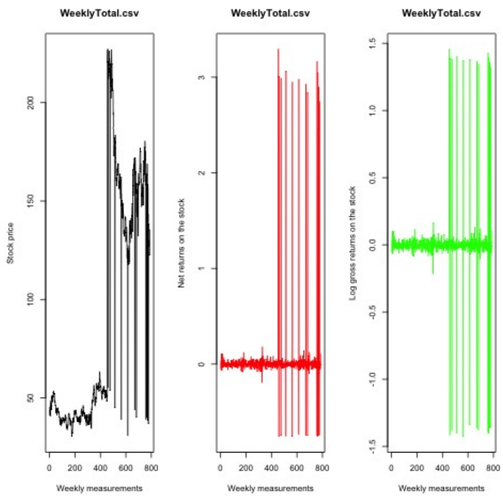

Figure 4.5 – 15 years of weekly Total Stock Price Data

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, this is in practice a very good approximation !

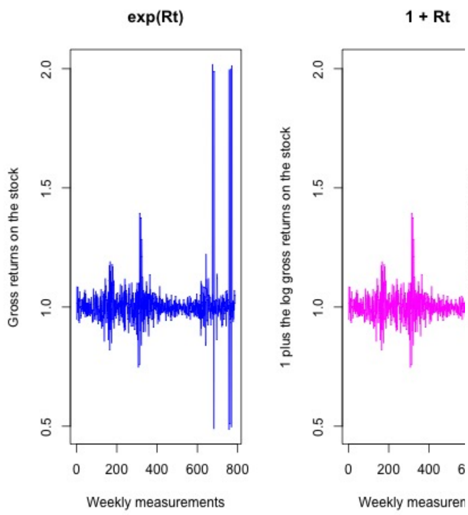

Figure 4.6 – $\exp(R_t)$ and $1+R_t$ for BNP Stock Price Data

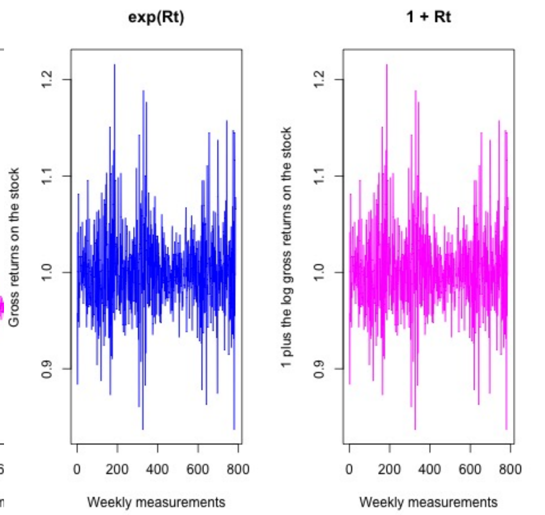

Figure 4.7 – $\exp(R_t)$ and $1+R_t$ for Carrefour Stock Price Data

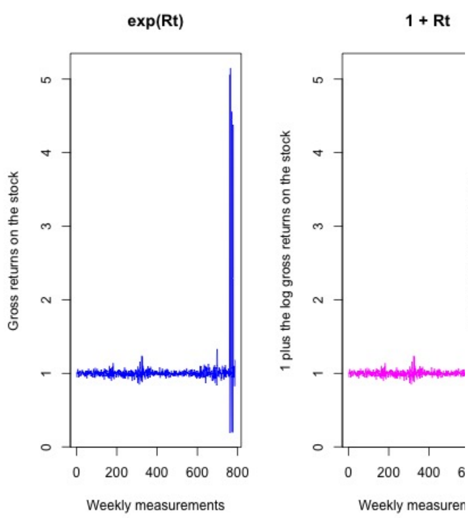

Figure 4.8 – $\exp(R_t)$ and $1+R_t$ for LVMH Stock Price Data

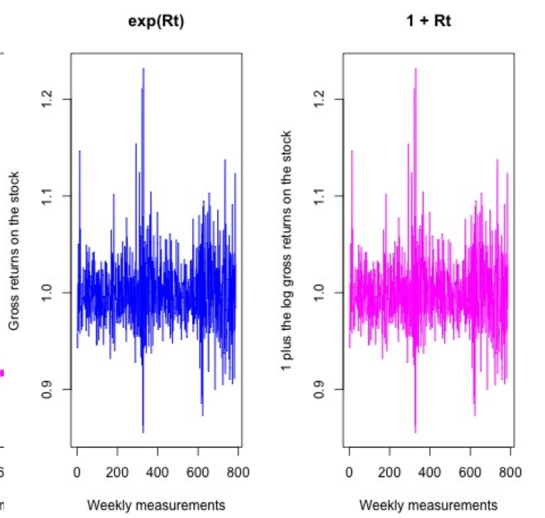

Figure 4.9 – $\exp(R_t)$ and $1+R_t$ for Sanofi Stock Price Data

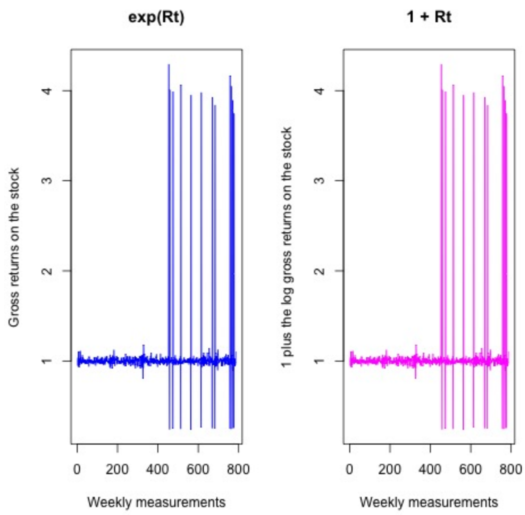

Figure 4.10 – $\exp(R_t)$ and $1 + R_t$ for Total Stock Price Data

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

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 equation, which reads as follows :

$$dS_t = \mu S_t dt + \sigma S_t dB_t \tag{4.1}$$

where $(B_t)_{t\geq 0}$ is a standard Brownian Motion with respect to a filtration $(\mathscr{F}_t)_{t\geq 0}$. 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.[3] An initial condition must be specified : $S_0 = s_0 > 0$.

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

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

$$\begin{cases} \mathrm{d}X_t = f(X_t)\mathrm{d}t + g(X_t)\mathrm{d}B_t \\ X_0 = x_0 \end{cases}$$

where $(B_t)_{t\geq 0}$ is a standard Brownian Motion with respect to a filtration $(\mathscr{F}_t)_{t\geq 0}$, $x_0 \in \mathbb{R}$, $f, g : \mathbb{R} \longrightarrow \mathbb{R}$ are Lipschitz functions. Then, by a theorem from Stochastic Calculus[4], we know that there exists a unique process that is continuous and adapted to the filtration $(\mathscr{F}_t)_{t\geq 0}$. Here, f and g are respectively the functions $x \longrightarrow \mu x$ and $x \longrightarrow \sigma t$. 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 explicit[5]).

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

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

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 $S_t = \phi_t Z_t$.

$$\begin{aligned} \mathrm{d}(S_t) &= \mathrm{d}(\phi_t Z_t) \\ &= \phi_t \mathrm{d}Z_t + Z_t \mathrm{d}\phi_t + \mathrm{d} < \phi, Z >_t \\ &= \phi_t \mathrm{d}Z_t + \mu\phi_t Z_t \mathrm{d}t \\ &= \phi_t \mathrm{d}Z_t + \mu S_t \mathrm{d}t \end{aligned} \tag{4.2}$$

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 $\sigma S_t \mathrm{d}B_t = \phi_t \mathrm{d}Z_t$. Hence, $\mathrm{d}Z_t = \sigma Z_t \mathrm{d}B_t$. 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 <u>totally wrong !</u>

Resolution - step 2 Now, let us set $Y_t = \log(Z_t)$. Using Ito-Doeblin's formula, we get :
$d(\log(Z_t)) = \frac{1}{Z_t}\mathrm{d}Z_t + \frac{1}{2}(\frac{-1}{Z_t^2})\mathrm{d} < Z >_t \; (\star)$

[4] readers wanting to get ahold of an excellent course on Stochastic Calculus are advised to refer to Dr. Lévêque's course. It can be found at http://ipg.epfl.ch/~leveque/

[5] This 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.

$\rightarrow$ How to compute $\mathrm{d} < Z >_t$?

$$
\begin{aligned}
& S_t = \phi_t Z_t \\
\implies & Z_t = \frac{S_t}{\phi_t} \\
\implies & \mathrm{d}Z_t = \frac{1}{\phi_t}\mathrm{d}S_t + S_t d(\frac{1}{\phi_t}) + \mathrm{d} < \frac{1}{\phi}, S >_t \\
& = \frac{1}{\phi_t}\mathrm{d}S_t - \frac{S_t}{\phi_t^2} d(\phi_t) \\
& = \frac{1}{\phi_t}\mathrm{d}S_t - \frac{\mu S_t}{\phi_t}\mathrm{d}t \\
& = \frac{1}{\phi_t}(\mu S_t \mathrm{d}t + \sigma S_t \mathrm{d}B_t) - \frac{\mu S_t}{\phi_t}\mathrm{d}t \\
& = \frac{\sigma S_t}{\phi_t}\mathrm{d}B_t
\end{aligned}
\tag{4.3}
$$

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 :

$$
\begin{aligned}
& < Z >_t = \int_0^t \frac{\sigma^2 S_s^2}{\phi_s^2} \mathrm{d}s \\
& = \int_0^t \sigma^2 Z_s^2 \mathrm{d}s \\
\implies & \mathrm{d} < Z >_t = \sigma^2 Z_t^2 \mathrm{d}t
\end{aligned}
$$

Back to ($\star$), we now have :

$$
\begin{aligned}
& \mathrm{d}(\log(Z_t)) = \frac{1}{Z_t}\mathrm{d}Z_t + \frac{1}{2}(\frac{-1}{Z_t^2})\sigma^2 Z_t^2 \mathrm{d}t \\
\iff & \frac{1}{Z_t}\mathrm{d}Z_t = d(\log(Z_t)) + \frac{\sigma^2}{2}\mathrm{d}t
\end{aligned}
\tag{4.4}
$$

Resolution - step 3 Combining the previous equation with $\mathrm{d}Z_t = \sigma Z_t \mathrm{d}B_t$, we get :

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

Finally, by remembering that $S_t = \phi_t Z_t = \exp(\mu t) Z_t$, we get :
$\forall t \geq 0, S_t = s_0 \exp((\mu - \frac{\sigma^2}{2})t + \sigma B_t)$
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 :
$\begin{cases} \mathrm{d}S_t = \mu(t) S_t \mathrm{d}t + \sigma(t) S_t \mathrm{d}B_t \\ S_0 = s_0 > 0 \end{cases}$
where $(B_t)_{t\geq 0}$ is a standard Brownian Motion with respect to a filtration $(\mathscr{F}_t)_{t\geq 0}$, 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 :
$\begin{cases} \mathrm{d}X_t = f(t, X_t)\mathrm{d}t + g(t, X_t)\mathrm{d}B_t \\ X_0 = x_0 \end{cases}$
where $(B_t)_{t\geq 0}$ is a standard Brownian Motion with respect to a filtration $(\mathscr{F}_t)_{t\geq 0}$, $x_0 \in \mathbb{R}$, $f, g : \mathbb{R}_+ \times \mathbb{R} \longrightarrow \mathbb{R}$ 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 $\mathrm{d}Y_t = \frac{1}{S_t}\mathrm{d}S_t - \frac{1}{2}\frac{1}{S_t^2}\mathrm{d} < S >_t$. If we remember that $\mathrm{d}S_t = \mu(t) S_t \mathrm{d}t + \sigma(t) S_t \mathrm{d}B_t$ and apply the Isometry formula, we get that $\mathrm{d} < S >_t = \sigma(t)^2 X_t^2 \mathrm{d}t$. We thus get :

$$
\begin{aligned}
\sigma \mathrm{d}Y_t &= \frac{1}{S_t}\mathrm{d}S_t - \frac{1}{2}\frac{1}{S_t^2}\mathrm{d} < S >_t \\
&= \frac{1}{S_t}\mathrm{d}S_t - \frac{1}{2}\sigma(t)^2\mathrm{d}t \\
&= \frac{1}{S_t}(\mu(t)S_t\mathrm{d}t + \sigma(t)S_t\mathrm{d}B_t) - \frac{1}{2}\sigma(t)^2\mathrm{d}t \\
&= (\mu(t) - \frac{1}{2}\sigma(t)^2)\mathrm{d}t + \sigma(t)dBt \qquad (4.6)\\
\implies Y_t &= y_0 + \int_0^t (\mu(s) - \frac{1}{2}\sigma(s)^2)\,\mathrm{d}s + \int_0^t \sigma(s)\,\mathrm{d}B_s \\
\implies Y_t &= \log(s_0) + \int_0^t (\mu(s) - \frac{1}{2}\sigma(s)^2)\,\mathrm{d}s + \int_0^t \sigma(s)\,\mathrm{d}B_s \\
\implies S_t &= s_0 \exp(\int_0^t (\mu(s) - \frac{1}{2}\sigma(s)^2)\,\mathrm{d}s + \int_0^t \sigma(s)\,\mathrm{d}B_s)
\end{aligned}
$$

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 SDE[6].

4.3 Back to the data

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

Bibliography

- Modelling Extremal Events, Embrechts, Kluppelberg and Milkosch
- 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!