University of Exeter
Department of Mathematics

Understanding Extremes and Clustering in Chaotic Maps and Financial Returns Data

Sara Ali Alokley

April 2015

Submitted by Sara Ali Alokley, to the University of Exeter as a thesis for the degree of Doctor of Philosophy in Financial Mathematics, April 2015.

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

(signature) .....

Abstract

In this thesis we present a numerical and analytical study of modelling extremes in chaotic dynamical systems. We study a range of examples with different dependency structures, and different clustering characteristics. We compare our analysis to the extreme statistics observed for financial returns data, and hence consider the modelling potential of using chaotic systems for understanding financial returns. As part of the study we use the block maxima approach and the peak over threshold method to compute the distribution parameters that arise in the corresponding extreme value distributions. We compare these computations to the theoretical answers, and moreover we obtain error bounds on the rate of convergence of these schemes. In particular we investigate the optimal block size when applying the block maxima method. Since the time series of observations on a dynamical system have dependency we must therefore go beyond the classic approach of studying extremes for independent identically distributed random variables. This is the main purpose of our study. As part of this thesis, we also study clustering in financial returns, and again investigate the potential of using dynamical systems models. Moreover we can also compare numerical quantification of clustering with theoretical approaches. As further work, we measure the dependency structures in our models using a rescaled range analysis. We also make preliminary investigations into record statistics for dynamical systems models, and relate our findings to record statistics in financial data, and to other models (such as random walk models).

Acknowledgements

In the name of Allah, the Most Gracious, the Most Merciful.

Allah, arouse me to be thankful for Thy favour wherewith Thou hast favoured me and my parents, and to do good that shall be pleasing unto Thee, and include me in (the number of) Thy righteous slaves.

In writing this thesis I am grateful to my supervisor Dr. Mark Holland for his encouragement, expertise, support and invaluable guidance he has provided throughout my period of study in Exeter. His patience and understanding were vital in helping me continue this thesis. I am also grateful for the contribution of my second supervisor Professor Richard Harris.

I would like to thank the College of Engineering, Mathematics and Physical Science and the Business School at the University of Exeter for their support during the studies. I would also like to thank King Faisal University for their financial support.

This thesis is dedicated to the soul of my mother Hesa Albarrak. I would like to thank my beloved family, my father Ali, my husband Dr. Mansour Albarrak and my lovely daughter Hesa for their love and patience through the rough road we have been on together. I would like also to thank my lovely sisters Dr. Mshael, May, Noura and Dr. Alia. I am specially grateful to my uncles Eng. Saleh Albarrak, Dr. Saad Albarrak, Dr. Abdulrahman Albarrak and Ibrahim Albarrak for their support.

Finally I would like to thank my colleagues in Room 201, specially Atiqa and Clare. Also thanks my friends Dr. Reem, Dr. Heba and Aki.

Notation

- We write that $x_n \sim y_n$ if $x_n/y_n \to 1$ as $n \to \infty$, where x_n, y_n are a general sequence.
- We say that $x_n \approx y_n$ if there exists a real constant c_1, c_2 such that $c_1 \leq x_n/y_n \leq c_2$.
- For positive sequences we say that $x_n = O(y_n)$ if there exists a constant $C > 0$ such that $x_n \leq Cy_n$.
- We say that $x_n = o(y_n)$ if $x_n/y_n \to 0$.

Contents

List of Tables

List of Figures

1. Introduction

1.1. Overview

Extremes in financial markets can have huge impacts on both the economy and society. Also, within climate/weather research extreme events are widely studied, e.g. to predict storms and floods. Therefore studying extremes is an active topic in current research. Let us start by describing Extreme Value Theory (EVT) for independent identically distributed (IID) random variables (RVs). We study the limit distribution of Maxima for such a sequence (ξ_n), where the extreme or maxima is given by $M_n = \max(\xi_1, \xi_2, ..., \xi_n)$. The limit distribution takes the form

$$\lim_{n\to\infty} P(a_n(M_n - b_n) \leq x) = G(x), \tag{1.1}$$

for the suitable scaling constants $a_n > 0, b_n$. It is found that this distribution converges (typically) to one of three types of distribution known as Type I, II and III extreme value distribution. We shall describe these distributions in Chapter 2. Going beyond the IID case, we can also study the limit in Equation (1.1) for more general sequences (ξ_n), such as stationary sequences with dependency, see for example Leadbetter et al. [61].

As part of this thesis, we will consider stationary processes generated by discrete dynamical systems. Here, we will consider the dynamical system $(\mathcal{X}, \nu, f)$ where, $\mathcal{X} \subset \mathbb{R}^d$, (some $d \geq 1$), $f : \mathcal{X} \to \mathcal{X}$ is a measurable map and ν is an $f - invariant$ probability measure. Under suitable conditions placed on the dependency and recurrence structures of $(\mathcal{X}, \nu, f)$ we will see that the process M_n (under suitable linear rescaling) converges again to one of the extreme value distributions as in the IID case. In particular to study distributions of maxima in dynamical systems, we consider $M_n = \max(\xi_1, ..., \xi_n)$ with $\xi_n = \phi(f^{n-1})$, and $\phi : \mathcal{X} \to \mathbb{R}$ is a suitable observable function. Moreover, we will investigate how the convergence to $G(x)$ in Equation (1.1) depends on i) the recurrence and dependency structures; ii) the explicit form of the observation ϕ, and iii) the regularity of the measure ν. We will discuss analytical computations in Chapter 2, and numerical investigations in Chapter 3.

Furthermore, for some processes the limit distributions will converge to $G^\theta(x)$ for some $\theta \in [0, 1]$. This value θ is known as the extremal index or clustering parameter which is a measure of clustering in extreme events, see [61]. If $\theta < 1$ then the process exhibits clustering and if $\theta = 1$ this suggests no clustering which is the case for IID observations. Hence, this parameter allows us to understand the dependency structure of the data. In Chapter 4 we will study clustering effects in financial models, and methods to estimate

the extremal index for a range of dynamical system examples. To our knowledge, a numerical investigation into computing the extremal index in dynamical systems has not been discussed in the literature, and hence we make preliminary progress in this direction.

A subject related to extreme value theory is that of *record statistics*, see [1]. In a series of events we say that a record event occurs at time n if its observed value exceeds all the previous values up to that time. Formally, for a sequence of RVs $X_1, X_2, \ldots, X_n$ we say that X_n is a record if

$$X_n > \max(X_1, X_2, ..., X_{n-1}).$$

In Embrechts et al. [26], limit theorems associated with record times for IID RVs (growth of record times, and frequency of records) were established. We investigate numerically this limit behaviour in dynamical system models and compare these results with the IID cases. In financial applications there has been recent work on record statistics in random walk models with IID increments, see for example, [69], [92], [68] and [90]. Within Chapter 6, we investigate the growth rates of record numbers in dynamical systems models, and also extended random walk models with dependent increments. We contrast to the corresponding IID cases. Putting all this together, the purpose of this thesis is to understand the following:

• For dependent processes, such as dynamical systems, under what conditions do we get the same limit distributions for the extremes as in the IID cases?

• For dependent processes, how do numerical schemes such as the block maxima method perform when estimating the distribution parameters?

• Can we use dynamical system models to exhibit clustering, and if so how do the statistical methods in estimating clustering perform on these systems?

• Can we understand extremes, records and clustering in financial returns using dynamical systems models?

1.1.1. Background on financial models

In this section we give a brief overview of classical financial models, e.g. those that lead to the Black-Scholes model for option pricing. The natural asset model, see Higham [51] is given by:

$$\frac{S_{n+1} - S_n}{S_n} = \mu\delta t + \sigma\sqrt{\delta t}X_n, \quad n \in [0, N-1], \tag{1.2}$$

where S_n is the asset price at discrete time n, and N, δt are chosen so that $N\delta t = T$. The constant T is fixed in advance and is the time scale under consideration. The constants μ and σ correspond to the drift and volatility respectively. The left hand side of Equation (1.2) refers to the relative return on the asset over the small time scale δt. Here, (X_n) correspond to a sequence of IID random variables with $E(X_n) = 0$ and unit variance. In the limit $\delta t \to 0$, so that $N \to \infty$ (T fixed), it is shown that the asset prices follow a

log-normal distribution, i.e.

$$\log(S_T/S_0) \sim \mathcal{N}\left((\mu - \frac{1}{2}\sigma^2)T, \sigma^2 T\right).$$

Going to a continuous limit, the natural stochastic model for S_t, $t \in [0, T]$ is the geometric Brownian motion model governed by the stochastic differential equation:

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \tag{1.3}$$

where we interpret these differentials in the Itô calculus framework. In particular, W_t denotes a standard Brownian motion, see [28].

Using these models, a mathematical theory of option pricing was developed, for example by Black-Scholes [11]. In this thesis we do not discuss option pricing, but through the study of extremes and records in asset price movements, there is potential to develop options based upon asset values exceeding (for example) a specified extreme value, or having a specified number of record daily closing prices within a give time period. Indeed, certain exotic options, such as the Lookback option (see [51]) have payoff functions that depend on the maximum value an asset has achieved over a given time period T. It is possible also to have payoff functions giving positive values (say) if a certain number of records are observed over a given time T, etc.

In this thesis, we will focus on the time series X_n, (such as that appearing in Equation (1.2)) and consider scenarios where this time series has a dependency structure. Hence an application in sight is to consider asset models where there is a dependency structure between daily returns. For simplicity, we consider the random walk model

$$S_n = \xi_1 + \ldots \xi_n,$$

and assume a dependency structure for the ξ_i. We remark that the random walk model is not strictly a good model for financial assets, since it can take negative values, and the increments do not reflect proportional changes in S_n. However, for purposes of studying extremes and records, the models are analogous.

In the literature, there is empirical evidence that stock markets returns exhibit dependency or long range dependence see, for example, [17] and [24]. In addition, [12] tested the long range dependence in developed, emerging and in transition economies. They concluded that the evidence of the long range dependence is small in stock market returns. The dependency on the data can be measured by using the Hurst exponent $H \in (0, 1)$ that gives a value $H \leq 0.5$ if the data has short range dependence and $H > 0.5$ if the data has long range dependence. Properties concerning the dependency of the financial returns data can be found in [17] and [24] .

In this thesis we will analyse in detail three examples of dynamical systems with different dependency structures. These examples we call the *Expanding Times 3* map, the *Intermittent* map and the *Hénon* map. We aim to analyse the extreme behaviour of such maps by estimating the tail or the shape parameter, the extremal index, the Hurst exponent

and analyse record statistics. We then compare and contrast our results with extremes in financial returns data. Based on our results the estimated tail index of the financial return data seems to be of type II Fréchet distribution. Also, the estimated extremal index suggests clustering and the estimated H exponent is ≈ 0.5 which suggests independence or short memory. Furthermore, we find that the growth of the number of records in the returns is similar to that observed in the IID case. Interestingly, all these phenomena of the financial returns may be modelled by dynamical systems, such as the Times 3 map, and Hénon map (short memory) and the Intermittent map (with long memory characteristics).

1.2. Thesis outline.

In Chapter 2, we start with a review of extreme value theory and discuss the limit distributions of extremes for IID random variables. We then discuss extremes for more general stationary processes and for processes generated by taking a time series of observations on a discrete dynamical system. The main source on this is Leadbetter et al. [61], and we discuss the conditions that a process must satisfy in order to get convergence to an extreme value distribution. Going further, we make more refined calculations that allow us to extract information on the rate of convergence to an extreme value distribution. This work builds upon recent results derived by Holland and Nicol [53]. In particular we emphasize how the convergence rate depends on the precise analytic form of the observation function. This analysis will be useful for Chapter 3 on applying the block maxima statistical methods.

Within Chapter 2, we also give a background on dynamical systems theory, and explain how the dependency conditions for stationary processes (as developed in [61]) translate to dynamical system applications. These conditions relate to the dependency and recurrence properties of the underlying dynamical system. We also discuss how the convergence to an extreme value distribution depends on the explicit form of the invariant measure as well as the observation function.

In Chapter 3 we review and apply the statistical methods that are used for modelling extremes, such as the block maxima method and the excess over threshold method. We apply these methods to the dynamical systems examples such as Times 3, Intermittent and Hénon maps, and then to the S&P500 returns index and FTSE100 returns index. The problem with the block maxima method is the choice of the block number and size. In a dynamical system setting, and using the recent work of Holland and Nicol [53] we attempt to extract an optimal block size and block number. The idea here is to extend our approaches relative to the works by Faranda et al. [32] and Lucarini et al. [66]. In these references they study the block maxima method to estimate the tail index over a number of block numbers and sizes. Here, we go further and attempt to estimate the rate of convergence to the theoretical tail index. This will build upon the estimates derived by Holland and Nicol [53], and we investigate the results using varying block size and number to minimize the error. A good agreement with the theoretical result is observed for the Expanding Times 3 map, and also for certain parameters of the Intermittent map. We

point out the main difficulties associated with calculating the tail index for the Hénon map.

In Chapter 4, we review and discuss clustering of extremes. We compute numerically the estimated extremal index value in comparison with the theoretical value derived by Freitas et al. [41]. The extremal index is estimated by using the interval estimator that was introduced by Ferro and Segers [34] at different threshold values for the dynamical system examples described in Chapter 3. We discuss the range of thresholds for which we see a good agreement with the theoretical value of the extremal index. Furthermore, we apply this method to the S&P500 and FTSE100 return indices to measure the clustering index (with suitable choice of threshold), and hence decide if there is clustering.

In Chapter 5, we measure the dependency in the data by using the Hurst exponent H that can be estimated by the classic Rescaled Range analysis. We estimate this parameter in the case of IID RVs, and test the methods on our dynamical system examples to see if there is good agreement with the theoretical results. We then apply it to the S&P500 and FTSE100 return indices.

In Chapter 6, we analyse and discuss the occurrence of Record statistics in Dynamical system examples, financial return data in comparison with the IID case (Embrechts et al. [26]) and in the financial prices. Our analysis of its occurrence showed similar behaviour to the Hurst exponent estimated values described in Chapter 5. We also analyse the number of records in biased Random walk, Markov Chain and Random walk with increments generated by Times 3 map and intermittent map. We observe that the number of records in the Random walk model with increments generated by short memory process is close to that of unbiased Random walk.

Finally, in Chapter 7 we conclude our findings and discuss studies which should be investigated in the future.

2. Extremes

2.1. Introduction

The classic Extreme Value Theory (EVT) concerns the study of limit distributions of the maximum values in IID RVs. Such a theory is also applicable to the study of extremes in stationary processes and dynamical systems, under suitable conditions placed on the dependency structures. Under such conditions, the extreme distributions then correspond to those observed for IID random variables. In this Chapter we will give a review of the cases where the limiting distribution behaviour of maxima behaves in the same way as those of IID observations. Furthermore we will discuss the speed of convergence to an extreme value distribution. For IID random variables, this is considered in Leadbetter et al. [61] and recently for dynamical systems in Holland and Nicol [53]. These results will be useful for the numerical estimates of the tail index that will be discussed in Chapter 3.

2.2. Background on Extremes

Let $\xi_1, \xi_2, ...$ be a sequence of IID RVs with distribution function F. Define M_n as the maximum of the first n by:

$$M_n = \max\{\xi_1, \xi_2, ..., \xi_n\}.$$

Extreme value theory is interested in studying the limiting distribution of maxima in n IID RVs. Fisher and Tippett [36] first introduced the limiting distribution of maxima in independent sequence as shown in the following theorem.

Theorem 2.1 *Let (ξ_n) be a sequence of IID RVs. If there exist constants $a_n > 0$, $b_n \in \mathbb{R}$, and some non-degenerate distribution function G such that*

$$\lim_{n\to\infty} P(a_n(M_n - b_n) \leq x) = G(x) \tag{2.1}$$

then, G is a generalized extreme value(GEV) distribution.

The (GEV) distribution can be written as a function of the shape or the tail index ρ and it can be only one of three standard distributions described as follows:

$$G(x) = \begin{cases} \exp\left(-\left(1+\rho\frac{x-\mu}{\sigma}\right)^{\frac{-1}{\rho}}\right) & \text{if } \rho \neq 0; \\ \exp(-\exp(-\frac{x-\mu}{\sigma})) & \text{if } \rho = 0. \end{cases} \tag{2.2}$$

where, $\{x : 1 + \rho(x - \mu)/\sigma > 0\}$ and the parameters $-\infty < \rho < \infty, -\infty < \sigma < \infty, \mu > 0$ known as the shape, scale and location respectively.
The GEV distribution function above can be one of three standard distributions: the Gumbel, Fréchet and Weibull distributions. These distributions correspond to the tail index (shape) parameter ρ as follows:

- Type I(Gumbel) for $\rho = 0$.
- Type II(Fréchet) for $\rho > 0$.
- Type III(Weibull) for $\rho < 0$.

(See, for example, [61] and [26]).

The $P(a_n(M_n - b_n) \leq x)$ can be written as $P(M_n \leq u_n)$, where $u_n = u_n(x) = x/a_n + b_n$. Leadbetter et al. [61] give the following result on the convergence of the $P(M_n \leq u_n)$.

Theorem 2.2 *Suppose (ξ_n) is a sequence of IID RVs and $\tau \in (0, \infty)$. Let (u_n) be a sequence of real numbers such that*

$$\lim_{n\to\infty} n(1 - F(u_n)) = \tau, \tag{2.3}$$

then,

$$\lim_{n\to\infty} P(M_n \leq u_n) = e^{-\tau}. \tag{2.4}$$

2.2.1. Domain of Attraction

In this section we discuss under which conditions we can know if F lies in the domain of attraction of extreme value distribution types. There are some conditions known as *von Mises conditions* which assume the density distribution function f exists as described by Theorem 2.3 below. Moreover, there are some conditions under the topic of regular variations described in Theorem 2.4 below.

Definition 2.1 *(Regular Variation.) A Positive, Lebesgue measurable function L on $(0, \infty)$ is regularly varying at ∞ of index $\alpha \in \mathbb{R}$ if*

$$\lim_{x\to\infty} \frac{L(tx)}{L(x)} = t^\alpha, \; x > 0 \tag{2.5}$$

and write $L \in \mathcal{R}_\alpha$.

If $\alpha = 0$ then L is slowly varying function at ∞. We write $L \in \mathcal{R}_0$.

See for example, [46] and [10].

Examples of slowly varying functions include $\log(x)$, iterated expressions of $\log(x)$, and other expressions such as $e^{(\log x)^\alpha}$, $(\alpha \in (0,1))$. Examples of regularly varying functions (at ∞) with index α include $x^\alpha, x^\alpha \log(1+x)$ and $(x\log(1+x))^\alpha$. Moreover, such functions can display infinite oscillation, see, [26].

Theorem 2.3 *[61] Let (ξ_n) be a sequence of IID RVs with distribution function F absolutely continuous with density f. Then sufficient conditions for F to belong to each of the three possible(Domain of Attraction) are:*

- *Type I: f has a negative derivative f' for all x in some interval $(x_0, x_F), (x_F \leq \infty), f(x) = 0$ for $x \geq x_F$, and*

$$\lim_{t\uparrow x_F} \frac{f'(t)(1-F(t))}{f^2(t)} = -1;$$

- *Type II: $f(x) > 0$ for all $x \geq x_0$ finite, and*

$$\lim_{t\to\infty} \frac{tf(t)}{1-F(t)} = \alpha > 0;$$

- *Type III: $f(x) > 0$ for all x in some finite interval $(x_0, x_F), f(x) = 0$ for $x > x_F$ and,*

$$\lim_{t\uparrow x_F} \frac{(x_F - t)}{1-F(t)} = \alpha > 0.$$

where, $x_F \leq \infty$ is the right endpoint of F, such that

$$x_F = \sup(x; F(x) < 1).$$

Theorem 2.4 *[61] Let (ξ_n) be a sequence of IID RVs with distribution function F. Then sufficient conditions for F to belong to each of the three possible(Domain of Attraction) are:*

- *Type I: if there is some strictly positive function $g(t)$ and*

$$\lim_{t\uparrow x_F} \frac{1-F(t+xg(t))}{1-F(t)} = e^{-x} \qquad \text{for all } x \in \mathbb{R};$$

- *Type II: $x_F = \infty$ and $1 - F \in \mathcal{R}_\alpha$ for each $x > 0$;*
- *Type III: $x_F < \infty$ and $\lim_{h\downarrow 0}(1-F(x_F - xh))/(1-F(x_F - h)) = x^\alpha$ for each $x > 0$.*

where, $g(t)$ could be taken as $g(t) = \int_t^{x_F} (1-F(u))du/(1-F(t))$ for $t < x_F$.

For further details see for example [10], [26], [46] and [61].

The following are examples of distributions which belong to the (domain of attraction) of an extreme value distribution.

Example 2.1 Let (ξ_n) be a sequence of IID standard exponential RVs, $F(x) = 1 - e^{-x}, f(x) = \frac{dF}{dx} = e^{-x}$.
From Theorem 2.3 F belongs to the domain of attraction of a Type I distribution. We show this as follows: First note that $f'(x) < 0$ for all x and

$$\lim_{t \uparrow x_F} \frac{f'(t)(1 - F(t))}{f^2(t)} = \lim_{t \uparrow x_F} \frac{-e^{-t}(1 - (1 - e^{-t})}{(e^{-t})^2} = -1.$$

Choose u_n such that $u_n = -\log \frac{\tau}{n} = -\log \tau + \log n$, we have $(1 - F(u_n)) \sim \tau/n$. Then from Theorem 2.2 we have:

$$P(M_n \leq -\log \tau + \log n) \to e^{-\tau}.$$

By letting $\tau = e^{-x}$ we have $P(M_n \leq x + \log n) = P(M_n - \log n \leq x) \to \exp\{-e^{-x}\}$ with $a_n = 1$, $b_n = \log n$ we obtain equation (2.1).

Example 2.2 Let (ξ_n) be a sequence of IID standard Uniform RVs, $F(x) = x$, $f(x) = \frac{dF}{dx} = 1$ where $x \in [0, 1]$.
From Theorem 2.3 F is in the domain of attraction and belongs to Type III with $\alpha = 1$. Here, $f(x) > 0$ for all $x \in (0, 1)$ and

$$\lim_{t \uparrow x_F} \frac{(x_F - t)}{1 - F(t)} = \lim_{t \uparrow 1} \frac{(1 - t)}{1 - t} = 1.$$

Suppose that $(1 - F(u_n)) \sim \tau/n$ then, we may choose u_n such that $u_n = 1 - \tau/n$ and from Theorem 2.2 we have:

$$P(M_n \leq 1 - \tau/n) \to e^{-\tau}.$$

By letting $\tau = -x$ we have $P(M_n \leq 1 + x/n) = P(n(M_n - 1) \leq x) \to e^x$ with $a_n = n$, $b_n = 1$ and $\alpha = 1$ we obtain equation (2.1).

If the distribution belongs to one of the three types of Domain of Attraction we may obtain the constant a_n and b_n by using the following Corollary.

Corollary 2.1 *[61] The constant a_n, b_n in the convergence $P(a_n(M_n - b_n) \leq x) \to G(x)$ may be taken as:*

- *Type I : $a_n = [g(\gamma)]^{-1}, b_n = \gamma_n$;*
- *Type II : $a_n = \gamma_n^{-1}, b_n = 0$;*
- *Type III : $a_n = (x_F - \gamma_n)^{-1}, b_n = \gamma_n$,*

where, $\gamma_n = F^{-1}(1-1/n) = \inf(x; F(x) \geq 1-1/n)$ and $g(\gamma)$ as defined in Theorem 2.4 before.

Now, we will discuss under which conditions the behaviour of maxima for stationary sequence and Dynamical system is the same as those of independent sequence. We say that the sequence (ξ_n) is stationary if its distribution are invariant under shift of time. Theorem 2.2 above is generalized to stationary sequence (ξ_n) satisfying $D(u_n)$ and $D'(u_n)$ conditions given by Leadbetter et al. [61].

Definition 2.2 $(D(u_n))$ *We say that $D(u_n)$ holds for the sequence (ξ_n) if for any integers p, q and n*

$$1 \leq i_1 < ... < i_p < j_1 < ... < j_q \leq n$$

such that $j_1 - i_p \geq l$ we have

$$|F_{i,j}(u_n) - F_i(u_n)F_j(u_n)| \leq \alpha_{n,l}, \tag{2.6}$$

where $i = i_1, ..., i_p$, $j = j_1, ..., j_q$ and $\alpha_{n,l} \to 0$ as $n \to \infty$ for some sequence $l = l_n = o(n)$.

The $D(u_n)$ condition is a mixing condition, and measures the dependency structure in a stationary sequence. If this condition holds and $\lim_{n\to\infty} n(1-F(u_n)) = \tau$. Then, $\lim_{n\to\infty} \inf P(M_n \leq u_n) = e^{-\tau}$, to guarantee that $\lim_{n\to\infty} P(M_n \leq u_n) = e^{-\tau}$ we need to define the condition $D'(u_n)$. Further discussion on the role of $D(u_n)$ can be found in [61].

Definition 2.3 $(D'(u_n))$ *We say that $D'(u_n)$ holds for the sequence (ξ_n) if*

$$\lim_{k\to\infty} \limsup_{n\to\infty} n \sum_{j=2}^{[n/k]} P(\{\xi_1 > u_n\} \cap \{\xi_j > u_n\}) = 0, \tag{2.7}$$

where $[x]$ denotes the integer part of x.

Note that in the IID case conditions $D(u_n)$ and $D'(u_n)$ hold. In particular the latter condition follows by observing that:

$$P(\{\xi_1 > u_n\} \cap \{\xi_j > u_n\}) = P(\{\xi_1 > u_n\})^2 \sim \frac{\tau^2}{n^2},$$

where to get the right hand side we have used equation (2.3). Condition $D'(u_n)$ then follows. Moreover, if condition $D'(u_n)$ holds then clustering of extremes in stationary sequences is unlikely.

Theorem 2.5 *[61] Let (ξ_n) be a stationary sequence satisfied by $D(u_n)$ and $D'(u_n)$. Suppose that (u_n) is constant. Let $0 \leq \tau \leq \infty$ then*

$$\lim_{n\to\infty} n(1-F(u_n)) \to \tau \tag{2.8}$$

if and only if

$$\lim_{n\to\infty} P(M_n \leq u_n) \to e^{-\tau} \tag{2.9}$$

Idea of proof. We follow the main arguments presented in Leadbetter et al. [61], and at the same time highlight the errors involved between the probability $P(M_n \leq u_n)$ and the limit $e^{-\tau}$.

Suppose that $D(u_n)$ and $D'(u_n)$ conditions hold. Fix k and let $n' = [n/k]$. From

$$(M_{n'} > u_n) = \bigcup_{j=1}^{n'} (\xi_j > u_n),$$

we have

$$\sum_{j=1}^{n'} P(\xi_j > u_n) - \sum_{1 \leq i < j \leq n'} P(\xi_i > u_n, \xi_j > u_n) \leq P(M_{n'} > u_n) \leq \sum_{j=1}^{n'} P(\xi_j > u_n).$$

From the stationarity of (ξ_n) we have

$$\begin{aligned} P(M_{n'} > u_n) &\leq \sum_{j=1}^{n'} P(\xi_j > u_n) \\ &= \sum_{j=1}^{n'} (1 - F(u_n)) \\ &= n'(1 - F(u_n)). \end{aligned} \tag{2.10}$$

This implies that $P(M_{n'} \leq u_n) \geq 1 - n'(1 - F(u_n))$.
Also,

$$\begin{aligned} \sum_{1 \leq i < j \leq n'} P(\xi_i > u_n, \xi_j > u_n) &= \sum_{i=1}^{n'} \sum_{j=i+1}^{n'} P(\xi_i > u_n, \xi_j > u_n) \\ &= \sum_{i=1}^{n'} \sum_{j=i+1}^{n'} P(\xi_1 > u_n, \xi_{j-i} > u_n) \\ &\leq \sum_{i=1}^{n'} \sum_{j=2}^{n'} P(\xi_1 > u_n, \xi_j > u_n) \\ &\leq n' \sum_{j=2}^{n'} P(\xi_1 > u_n, \xi_j > u_n). \end{aligned} \tag{2.11}$$

This implies that $P(M_n \leq u_n) \leq 1 - \sum_{j=1}^{n'} P(\xi_j > u_n) + E_n$,

where $E_n = n' \sum_{j=2}^{n'} P(\xi_1 > u_n, \xi_j > u_n)$.
Hence,

$$1 - n'(1 - F(u_n)) \leq P(M_{n'} \leq u_n) \leq 1 - n'(1 - F(u_n)) + E_n. \tag{2.12}$$

By condition $D'(u_n)$ and since $n' = [n/k]$ we have

$$\lim_{k \to \infty} \limsup_{n \to \infty} [n/k] \sum_{j=2}^{[n/k]} P(\{\xi_1 > u_n\} \cap \{\xi_j > u_n\}) = o(1/k).$$

Now if $n(1-F(u_n)) \to \tau$ holds. Then $n'(1-F(u_n)) \to \tau/k$.
As $n \to \infty$ in equation (2.12) we have

$$1-\tau/k \leq \liminf_{n\to\infty} P(M_{n'} \leq u_n) \leq \limsup_{n\to\infty} P(M_{n'} \leq u_n) \leq 1-\tau/k+o(1/k).$$

By taking the kth power of each term, Leadbetter et al. [61] showed that the $D(u_n)$ condition implies that

$$P(M_n \leq u_n) - P^k(M_{[n/k]} \leq u_n) \to 0 \text{ as } n \to \infty. \tag{2.13}$$

Hence, we have

$$(1-\tau/k)^k \leq \liminf_{n\to\infty} P(M_n \leq u_n) \leq \limsup_{n\to\infty} P(M_n \leq u_n) \leq (1-\tau/k+o(1/k))^k.$$

Letting $k \to \infty$ we have

$$\lim_{n\to\infty} P(M_n \leq u_n) = e^{-\tau}.$$

Now if $P(M_n \leq u_n) \to e^{-\tau}$ holds. Then from equation (2.12) and since $n' = [n/k]$ we have

$$1-P(M_{n'} \leq u_n) \leq n'(1-F(u_n)) \leq 1-P(M_{n'} \leq u_n)+E_n.$$

Since $P(M_n \leq u_n) \to e^{-\tau}$, and equation (2.13) implies that $P(M_{n'} \leq u_n) \to e^{-\tau/k}$.
Letting $n \to \infty$ we have

$$1-e^{-\tau/k} \leq \frac{1}{k}\liminf_{n\to\infty} n(1-F(u_n)) \leq \frac{1}{k}\limsup_{n\to\infty} n(1-F(u_n)) \leq 1-e^{-\tau/k}+o(1/k)$$

By multiplying by k and letting $k \to \infty$ we have that

$$n(1-F(u_n)) \to \tau.$$

This concludes the proof.

Assume that $D'(u_n)$ does not hold and, for each $\tau > 0, u_n(\tau)$ is a sequence satisfying $\lim_{n\to\infty} n(1-F(u_n)) = \tau$ and $D(u_n(\tau))$ holds for each τ. Let $P(M_n \leq u_n(\tau))$ converge for at least one $\tau > 0$ then $\lim_{n\to\infty} P(M_n \leq u_n) = e^{-\theta\tau}$ for all $\tau > 0$, for some fixed θ with $0 \leq \theta \leq 1$. Further discussion on the role of $D'(u_n)$ can be found in [61, p.65]. Here, θ is known as the extremal index which we will discuss in Chapter 4. Furthermore, it was shown in [61] that if $D(u_n)$ and $D'(u_n)$ are satisfied for the stationary sequence (ξ_n), and if $(\tilde{\xi_n})$ is the associated independent sequence, then,

$$\lim_{n\to\infty} P(M_n \leq u_n) = \lim_{n\to\infty} P(\tilde{M_n} \leq u_n) = G(x).$$

where, (M_n) and $(\tilde{M}_n)$ are the maxima from the original sequence and the independent sequence respectively. (in other words, under these two conditions the behaviour of maxima is the same as the associated IID case). This idea is extended to study extremes in dynamical systems under suitable conditions.

Proposition 2.1 *Suppose (ξ_n) is a stationary sequence satisfied by $D(u_n)$ and $D'(u_n)$. Let (u_n) be a sequence such that $n(1-F(u_n) \to \tau$ as $n \to \infty$, and $r \geq 1$ is any fixed integer. Then we have*

$$\left| P(M_n \leq u_n) - \left(1 - \frac{\tau}{k}\right)^k \right| \leq \frac{(2k+1)}{r} + (2k+1)(2r\alpha_{n,l}) + (k-1)\alpha_{n,l} + o(1/k)^k. \quad (2.14)$$

Proof. As before fix k, and let $n' = [n/k]$. For some $m > 0$ with $k < m < n'$ let $I_1 = \{1, 2, ..., n' - m\}$, $I_1^* = \{n' - m + 1, ..., n'\}$. Following Leadbetter et al. [61], and using $D(u_n)$ we have:

$$P(M(I_1) \leq u_n < M(I_1^*)) \leq \frac{1}{r} + 2r\alpha_{n,m}, \quad (2.15)$$

provided $n \geq (2r+1)mk$. In addition we also have

$$|P(M_n \leq u_n) - P^k(M_{n'} \leq u_n)| \leq (2k+1)P(M(I_1) \leq u_n < M(I_1^*)) + (k-1)\alpha_{n,m}. \quad (2.16)$$

From equation (2.12) we have

$$|P(M_{n'} \leq u_n) - (1 - n'(1 - F(u_n)))| \leq E_n, \quad (2.17)$$

as $n \to \infty$ we have $n'(1 - F(u_n)) \to \tau/k$ then

$$|P(M_{n'} \leq u_n) - \left(1 - \frac{\tau}{k}\right)| \leq o(1/k). \quad (2.18)$$

By taking the kth power of each term and using equations (2.15) and (2.16) with $m = l_n$ as specified in the $D(u_n)$ condition we obtain equation (2.14) as required.

2.2.2. Rate of convergence in IID observations

Suppose that $\xi_1, \xi_2, ...$ are IID observations and $(u_n) = u_n(x) = x/a_n + b_n$. Suppose (u_n) satisfies Theorem 2.2 then by writing $\tau_n = n(1 - F(u_n))$ we have

$$\begin{aligned} P(M_n \leq u_n) &= F^n(u_n), \\ &= (1 - (1 - F(u_n))^n, \\ &= \left(1 - \frac{\tau_n}{n}\right)^n, \\ &\to e^{-\tau}. \end{aligned} \quad (2.19)$$

In the above equation, the final convergence result can be further refined by studying the following two approximations:

$$\left(1 - \frac{\tau_n}{n}\right)^n \approx e^{-\tau_n} \quad (2.20)$$

$$e^{-\tau_n} \approx e^{-\tau} \tag{2.21}$$

The error term in equations (2.20) and (2.21) were denoted by $\triangle_n = \left(1 - \frac{\tau_n}{n}\right)^n - e^{-\tau_n}$ and $\triangle'_n = e^{-\tau_n} - e^{-\tau}$ respectively. In [61] the following result is obtained.

Theorem 2.6 *[61] Suppose that (ξ_n) is an IID sequence, and let $\tau_n = n(1-F(u_n))$. Then $P(M_n \leq u_n) - e^{-\tau} = \triangle_n + \triangle'_n$. where*

$$0 \leq -\triangle_n \leq \frac{{\tau_n}^2 e^{-\tau_n}}{2}.\frac{1}{n-1} \leq 0.3.\frac{1}{n-1}. \tag{2.22}$$

If $\tau_n \to \tau$ then $\triangle_n \sim -(\tau^2 e^{-\tau}/2)/n$. Furthermore for $\tau - \tau_n \leq \log 2$ we have

$$\triangle'_n = e^{-\tau}((\tau - \tau_n) + \theta(\tau - \tau_n)^2), \tag{2.23}$$

with $0 < \theta < 1$.

In the following examples we discuss how the rate of convergence depends on $\triangle'_n$, especially for linear scaling sequences $u_n = x/a_n + b_n$.

Example 2.3 Let (ξ_n) be a sequence of IID standard exponential RVs, $F(x) = 1 - e^{-x}$. As in Example 2.1 we choose $u_n = x + \log n$. It was shown in Hall and Wellner [47] that:

$$\sup_x |P(M_n \leq x + \log n) - \exp\{-e^{-x}\}| = \sup_x \left| \left(1 - \frac{e^{-x}}{n}\right)^n - \exp\{-e^{-x}\} \right| \leq \frac{1}{n}(1 + 2/n)e^{-2}.$$

Suppose that $\tau_n(x) = n(1 - F(u_n))$, then it immediately follows that $\tau_n(x) = e^{-x}$ and in particular does not depend on n, i.e. $\tau_n(x) = \tau(x)$.

Example 2.4 Let (ξ_n) be a sequence of IID standard Uniform RVs, $F(x) = x$. As in Example 2.2 we choose $u_n = 1 - \tau(x)/n$ and we have

$$P(M_n \leq x/n + 1) - e^x = \left(1 - \frac{-x}{n}\right)^n - e^x.$$

Using a Taylor theorem argument, for all $0 \leq x \leq n$, we have:

$$0 \leq e^{-x} - \left(1 - \frac{x}{n}\right)^n \leq 0.3.\frac{1}{n-1} \text{ for } n = 1, 2, ...,$$

Again, for this example $\tau_n(x) = \tau(x)$, and is independent of n.

We have shown in the previous Examples 2.3 and 2.4 that we can choose a_n and b_n such that $\tau_n(x) = \tau(x)$ but this is not always the case and so we study this issue in more detail. This requires some further background on regularly and slowly varying functions.

Inversion of regularly varying functions and estimation of Δ'

For a regularly varying function $L(x)$ of index $\alpha \neq 0$, it is shown in Bingham et al. [10] that $L(x) = x^\alpha \ell(x)$, where $\ell(x)$ is a slowly varying function. An important problem is to find an inverse for $L(x)$, and usually only asymptotic expressions are available. This motivates the following definition:

Definition 2.4 *Suppose that $L(x) = x^\alpha \ell(x)$, with $\ell(x)$ slowly varying. We say that $\tilde{L}(x)$ is an asymptotic inverse to $L(x)$ (as $x \to \infty$) if $(L \circ \tilde{L})(x) \sim (\tilde{L} \circ L)(x) \sim x$.*

For simplicity let $\alpha = 1$. Given an explicit representation $L(x) = x\ell(x)$, we can express the asymptotic inverse as $\tilde{L}(x) = x\ell^\sharp(x)$, where $\ell^\sharp(x)$ is called the *de Bruijn conjugate*, see [10]. We have the following simple identity:

$$\ell(x)\ell^\sharp(x\ell(x)) \sim \ell^\sharp(x)\ell(x\ell^\sharp(x)) \sim 1, \quad (x \to \infty). \tag{2.24}$$

The problem we wish to consider is the computation of $\ell^\sharp(x)$ given $\ell(x)$, e.g. $\ell(x) = \log x$. In particular if $\ell(x)$ and $\ell^\sharp(x)$ form a de Bruijn conjugate pair, then the following are also de Bruijn conjugate pairs:

$$\left(\ell(ax), \ell^\sharp(bx)\right), \ \left(a\ell(x), a^{-1}\ell^\sharp(x)\right), \ \left((\ell(x^\alpha))^{1/\alpha}, (\ell^\sharp(x^\alpha))^{1/\alpha}\right), \tag{2.25}$$

for constants $a, b, \alpha > 0$. These identities are useful for computing asymptotic inverses. In particular for more general regularly varying functions of the form $L(x) = x^{\alpha\beta}(\ell(x^\beta))^\alpha$ an asymptotic inverse is given by:

$$\tilde{L}(x) \sim x^{\frac{1}{\alpha/\beta}} \left(\ell^\sharp(x^{\frac{1}{\alpha}})\right)^{\frac{1}{\beta}}. \tag{2.26}$$

In the special case where $\ell(x)$ is a logarithm, a power of a logarithm, or an iterated logarithm it can be shown that $\ell^\sharp(x) = 1/\ell(x)$. E.g. for $\ell(x) = \log x$, we can show this directly by verifying equation (2.24). However, for the slowly varying function $\ell(x) = e^{\sqrt{\log x}}$, it is not true that $\ell^\sharp(x) = 1/\ell(x)$, and inversion methods rely on Lagrange inversion, see [10].

Returning to the estimation of the error Δ', we now attempt to estimate $\tau_n - \tau$ for special sequences a_n and b_n. To illustrate the complexities involved we consider the distribution function $F(x)$ with $1 - F(x) = 1/(x\ell(x))$, where $\ell(x)$ is slowly varying at infinity. As before let $\tau_n(x) = n(1 - F(u_n))$, and consider the linear scaling $u_n = x/a_n + b_n$. Here, we are in the domain of attraction of a Type II distribution with $\tau(x) = x^{-1}$. We now choose a_n and b_n by solving the following equation

$$\left(\frac{x}{a_n} + b_n\right)^{-1} \left(\ell\left(x/a_n + b_n\right)\right)^{-1} = \frac{1}{nx}.$$

The problem being now that an analytic expression for the inverse is not readily available, and hence we cannot have $|\tau_n(x) - \tau(x)|$ identically zero. Hence using techniques of asymptotic inversion of regularly varying functions we show that there is a suitable choice

of a_n and b_n such that this error goes to zero. The error will depend on a_n and b_n, but we do not attempt to optimize.

Since F belongs to a type II distribution then from Corollary 2.1 we may choose $b_n = 0$ and try to solve for a_n so that we have

$$\left(\frac{x}{a_n}\right)^{-1}\left(\ell\left(\frac{x}{a_n}\right)\right)^{-1} = \frac{1}{nx}.$$

If we let $y = 1/a_n$ we obtain the equation:

$$n = y\ell(yx),$$

where in the above x is fixed. Thus we take an asymptotic inverse of this equation and obtain (for $y \to \infty$):

$$y \sim n\ell^\sharp(nx).$$

We note that its inverse is not unique, indeed we could also take $y = n\ell^\sharp(n)$ as an asymptotic inverse. Using this expression, and noting that $\tau(x) = x^{-1}$ we obtain

$$|\tau_n(x) - \tau(x)| = \frac{1}{x}\left|1 - \frac{1}{\ell^\sharp(n)\ell(xn\ell^\sharp(n))}\right| + o(1).$$

As $n \to \infty$, this error goes to zero using the asymptotic properties of the de-bruijn conjugates. Bounds on the speed of convergence will depend on the explicit form of $\ell(x)$ and $\ell^\sharp(x)$, and their asymptotics as $n \to \infty$.

Example 2.5 Suppose that $F(x) = 1 - \frac{1}{x^\alpha \log x}$ where, $\alpha > 0$ and $x > x_0$ for some $x_0 > 0$. Since

$$f(x) = \frac{\alpha}{x^{\alpha+1}\log x} + \frac{1}{x^{(\alpha+1)}(\log x)^2}$$

and from Theorem 2.3 F is in the Domain of Attraction and belongs to type II.

We have that:

$$\lim_{t\to\infty}\frac{tf(t)}{1-F(t)} = \frac{\frac{t\alpha}{t^{\alpha+1}\log t} + \frac{t}{t^{(\alpha+1)}(\log t)^2}}{t^\alpha \log t} = \alpha > 0.$$

From equation (2.3), we have

$$(1 - F(u_n)) \sim \tau/n.$$

Then,

$$\frac{1}{(u_n)^\alpha \log u_n} \sim \frac{\tau}{n}.$$

Since $u_n(x) = x/a_n + b_n$ we attempt to solve the following equation,

$$\left(\frac{x}{a_n} + b_n\right)^\alpha \log\left(\frac{x}{a_n} + b_n\right) \sim n.$$

There is no analytical solution for a_n and b_n so we try to approximate. In addition, from Corollary 2.1 we may choose $b_n = 0$, but it is not easy to find a_n since a_n is a function of

$\gamma_n = F^{-1}(1 - 1/n)$. The latter function does not admit a simple formula for its inverse.

For simplicity assume $\alpha = 1$. An initial guess might try with $a_n = \frac{1}{n}$, $b_n = 0$ but for this choice $|\tau_n(x) - \tau(x)|$ does not go to zero. Using techniques of asymptotic inversion we try $a_n = \frac{\log n}{n}$ and $b_n = 0$. The former choice for a_n follows from the fact that the de-bruijn conjugate of $\log x$ is $1/\log x$ (when $x \to \infty$). We have

$$\begin{aligned}
\left|\frac{n}{u_n \log u_n} - \frac{1}{x}\right| &= \left|\frac{\log n}{x \log\left(\frac{nx}{\log n}\right)} - \frac{1}{x}\right|, \\
&= \left|\frac{1}{x}\left(\frac{\log n}{\log n + \log x - \log\log n} - 1\right)\right|, \\
&= \left|\frac{1}{x}\left(\left(1 + \frac{\log x}{\log n} - \frac{\log\log n}{\log n}\right)^{-1} - 1\right)\right|, \\
&= \left|\frac{1}{x}\left(\frac{\log x}{\log n} - \frac{\log\log n}{\log n}\right)\right| + O\left(\frac{(\log\log n)^2}{(\log n)^2}\right). \qquad (2.27)
\end{aligned}$$

Hence,

$$|\tau_n(x) - \tau(x)| \leq C\frac{\log\log n}{\log n},$$

where the constant C depends on x. Thus for fixed x and as $n \to \infty$ the error of order is

$$\frac{\log\log n}{\log n}.$$

For different choice of a_n, b_n, this bound might be improved. Indeed to minimize the error $\triangle_n'$ it then becomes a general theoretical problem in asymptotic analysis to choose the best a_n and b_n. Indeed, as mentioned in Leadbetter et al. [61] the speed or rate of convergence of $P(M_n \leq u_n) \to e^{-\tau}$ will depend on the choice of a_n and b_n, and so different distribution types will give rise to different speeds of convergence.

2.2.3. Excess over a threshold.

One approach of modelling maxima is by looking at the data that exceed threshold value. Where, the distribution of exceeding threshold value can be described by the Generalized Pareto (GP) distribution. See for example, [26].

Definition 2.5 *Let X be a RV with distribution function F. The distribution over the threshold u is given by:*

$$F_u(x) = P(X - u \leq x | X > u) = \frac{F(x+u) - F(u)}{1 - F(u)} \qquad (2.28)$$

where, $0 \leq x \leq x_F - u$.

Theorem 2.7 *[25] We can find a positive measurable function $\beta(u)$ such that,*

$$\lim_{u \to x_F} \sup_{0 \leq x < x_F - u} |F_u(x) - GP(x)| = 0$$

if and only if F in the Domain of Attraction of $G(x)$.

Where, F_u has been defined before in equation (2.28), and $GP(x)$ is the Generalized Pareto distribution that is given by:

$$GP(x) = \begin{cases} 1 - (1 + \rho x/\beta)^{-\frac{1}{\rho}}, & \rho \neq 0 \\ 1 - exp(-x/\beta), & \rho = 0, \end{cases} \tag{2.29}$$

where, $\beta > 0$ and $x \geq 0$ if $\rho \geq 0$ and, $0 \leq x \leq -\frac{\beta}{\rho}$ if $\rho < 0$. Here, ρ and β are the shape and the scale parameter respectively.

The theorem above goes to Pickands III [78] and, Balkema and De Haan [2].

Remark. The limit distribution of normalized maxima can be modelled by a GEV distribution whereas, the excess distribution over thresholds can be modelled by a GP distribution.

In applications there are two methods used for modelling maxima, block maxima method and excess over a threshold method. The latter method needs an estimation of the threshold value as we will discuss in the next chapter. The choice of the appropriate threshold can be estimated by the mean excess plot approach which is based on the linearity of the mean excess function plot of the GP distribution. This approach is widely used in the literature see, for example, [26], [15] and [25]. The mean excess function is defined as the following:

Definition 2.6 *Let X be an RV and u a threshold. Then the mean excess function is given by:*

$$e(u) = E[X - u | X > u]. \tag{2.30}$$

As shown for example in [26] and [25] the mean excess function of a GP distribution with $\rho < 1$ and β is given by:

$$e(u) = E[X - u | X > u] = \frac{\beta + \rho\, u}{1 - \rho}, \tag{2.31}$$

where, $\beta + \rho\, u > 0$.

As we can see from equation (2.31) the mean excess function of a GP distribution is a linear function of u. We will use this idea to estimate the appropriate threshold u in the

data examples in the next chapter.

2.2.4. Review of probabilistic methods in dynamical systems

Convergence to an extreme value distribution in a dynamical system is studied under certain mixing and recurrence conditions. Furthermore, conditions are also required on the regularity (smoothness) of the underlying invariant probability. In particular we will study (one-dimensional) systems where this measure is absolutely with respect to the Lebesgue measure. In Chapter 3 we will investigate numerically the convergence to an extreme value distribution for a range of dynamical system examples, each having different dependency structures, and different recurrence structures. Moreover we will study systems in dimension two, and consider situations where the invariant measure is not absolutely continuous with respect to volume. In this section we outline the main dynamical system models, and discuss precisely their dependency structures, their recurrence properties, and the properties of their invariant measures.

Let X be a set and suppose Σ is a $\sigma - algebra$ of subsets of X.

Definition 2.7 *(A measure) Suppose that* $\nu : \Sigma \longrightarrow \mathbb{R}^+ \cup \{\infty\}$ *is a function which satisfies the following conditions:*

- $\nu(\emptyset) = 0$
- *if* $A_n \cap A_m = \emptyset$ *for* $n \neq m$ *then,*

$$\nu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \nu(A_n)$$

where $\{A_n\}_{n=1}^{\infty}$ *is a countable collection of pairwise disjoint sets in* Σ.

Then, ν *is a measure, and* (X, Σ, ν) *is a measure space.*

If $\nu(X) = 1$, then $\nu(X)$ and (X, Σ, ν) are called a probability measure and a probability space respectively. We now define *Leb* measure in $\mathbb{R}$. This generalizes to higher dimensions. In $\mathbb{R}^2$ and $\mathbb{R}^3$ this is denoted as area and (resp.) volume.

Definition 2.8 *(Lebesgue(Leb) measure) The Leb measure on* $\mathbb{R}$ *of a subset* A *is defined as*

$$Leb(A) = \inf\left(\sum_{i=1}^{\infty}(b_i - a_i) : A \subset \cup_{i=1}^{\infty}[a_i, b_i]\right)$$

where $\{[a_i, b_i]\}_i$ *are disjoint.*

See for example, [29].

Now we will define what is known by absolutely continuous measure. For our main examples, we will mainly consider such measures.

Definition 2.9 *(Absolutely continuous measure) Let ν, m be two probability measures on (X, Σ). We say that ν is absolutely continuous with respect to m or we write it as $\nu \ll m$, for each $B \in \Sigma$ with $m(B) = 0$ we have $\nu(B) = 0$. Furthermore we say that ν is absolutely continuous with respect to m if $\nu = hm$ where h is the density and defined by $\nu = \int_B h\, dm$.*

See for example, [20]. A measure which is not absolutely continuous is (for example) the Dirac measure δ_b on a subset B. This measure gives the value 1 if $b \in B$, and 0 if $b \notin B$.

Definition 2.10 *(Invariant measure) Suppose that (X, Σ, ν) is a probability space, $f : X \longrightarrow X$. We say that ν is an $f-$invariant measure, if $\nu(f^{-1}B) = \nu(B)$ for each $B \in \Sigma$.*

See [20].

Consider for example the doubling map $f : [0, 1] \to [0, 1]$ that is defined by:

$$f(x) = \begin{cases} 2x & x \leq \frac{1}{2}, \\ 2x - 1 & x > \frac{1}{2}. \end{cases} \tag{2.32}$$

Then $\nu = \text{Leb}$ is the (absolutely continuous) f-invariant measure. Recall that ν is f-invariant if $\nu(f^{-1}B) = B$ for all $B \in \Sigma$.

Definition 2.11 *(Ergodic measure) Suppose that (X, Σ, ν) is a probability space. We say that ν is an ergodic measure if $f^{-1}(B) = B$ implies $\nu(B) = 0$ or 1.*

Definition 2.12 *(SBR measure) Let ν be a $f-$invariant measure $\nu : \Sigma \longrightarrow [0, 1]$. Suppose there exists a set $B \subset I$ of positive Lebesgue measure such that for any continuous function $\phi : I \longrightarrow \mathbb{R}$ has*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \phi(f^n(x)) = \int \phi\, d\nu \text{ for every } x \in B.$$

Then ν is called a Sinai-Bowen-Ruelle (SBR) measure.

See [20]. A known example of a map that has an SBR measure is the doubling map above. See for example [96]. An example in dimension 2 is the Hénon map, see chapter 3.

Definition 2.13 *(Markov map) Suppose we have an interval $I = [0, 1]$ and divide this interval into a finite number of closed sub intervals $I_i = [x_i, x_{i+1}], i = 0, ..., k-1$ where, $0 = x_0 < x_1 < ... < x_k = 1$. Let $f : I \to I$ be a C^1 map, monotone on each of the open intervals $intI_i = (x_{i-1}, x_i)$ with the following property:*

- *(piecewise expanding) If there exist $\gamma > 1$ such that $|f'(x)| \geq \gamma$ for all $x \in I_i (i = 1, ..., k)$;*
- *(Markov property)If $f(intI_i) \cap int(I_j) \neq$ then $f(intI_i) \supset intI_j$. where, $i, j = 1, ..., k$.*

See [79].

It is proven for example in [20] that for a Markov map there exists an $f-$*invariant* probability measure which is absolutely continuous with respect to Lebesgue measure as stated in the following Theorem:

Theorem 2.8 *If $f : I \to I$ is a Markov map and $\{I_i\}_{i=1}$ is the corresponding partition. Then, there is an $f-$invariant probability measure ν on the Borel sets of I which is absolutely continuous with respect to the Lebesgue measure.*

In our application, we will consider a Markov expanding map. Further details on these definitions can be found, for example, in [20] and [89].

Mixing and decay of correlations

There are a variety of forms or definitions of the mixing property in dynamical systems see for example [89] and [77].

Definition 2.14 *(Mixing) We say that f is strong mixing if and only if*

$$\nu(f^{-n}A \cap B) \to \nu(A)\nu(B) \text{ as } n \to \infty,$$

for all $A, B \in \Sigma$ where, ν is an invariant probability measure.

Definition 2.15 *We say f has decay of correlation in Banach space $\mathcal{B}$ if $\forall \psi_1, \psi_2 \in \mathcal{B}$, we have the correlation function $C_n(\psi_1, \psi_2, \nu) \to 0$ as $n \to \infty$.*

$$C_n(\psi_1, \psi_2, \nu) = \left| \int \psi_1 . \psi_2 \circ f^n d\nu - \int \psi_1 d\nu \int \psi_2 d\nu \right|.$$

E.g. the space $\mathcal{B}$ could be space of Lipschitz continuous functions. If ψ_1 and ψ_2 are characteristic functions then the definition of decay of correlations can be reduced to the definition of mixing. See for example, [9].

Definition 2.16 *(Recurrent) A point x for f is recurrent if for any $\epsilon > 0$, there exist $k \geq 1$ such that*

$$|f^k(x) - x| < \epsilon.$$

See for example, [22]. An important problem is knowing how the integer k depends on ϵ. For example, the map $x \to x + \alpha \mod 1$, with $\alpha \notin \mathbb{Q}$, is such that every point $x \in [0, 1]$ has a dense orbit. For example, it can be shown that $|f^k(x) - x| < 1/k$ for an infinite sequence of values $k \in \mathbb{N}$ (see [77]).

Properties of dynamical systems:

In this section we present a theoretical discussion on the properties of dynamical systems. We give examples with different correlation (dependency) structures, and different recur-

rence and measure properties. In our study we will investigate three examples of chaotic maps and here we briefly outline some of their properties:

1. Expanding Times 3 map.

   For $f : [0,1] \to [0,1]$, the map defined by:

   $$f(x) = 3x \mod 1, \tag{2.33}$$

   - f admits an invariant probability measure $\nu \ll m$ is absolutely continuous with respect to Lebesgue measure m. The map f is Markov and expanding.
   - f has exponential decay of correlation (Mixing) for $\mathcal{B}$ the class of Lipschitz continuous function[1]. That is there exists $C > 0$ and $\theta < 1$ such that

     $$C_n(\psi_1, \psi_2, \nu) \leq C\|\psi_1\|_{\text{Lip}}\|\psi_2\|_{L^\infty}\theta^n.$$

     See, for example [9].

2. Intermittent map.

   For $f : [0,1] \to [0,1]$, the map defined by:

   $$f(x) = \begin{cases} x(1 + 2^\beta x^\beta) & x \leq \frac{1}{2}, \\ 2x - 1 & x > \frac{1}{2}. \end{cases} \tag{2.34}$$

   Where, $\beta \in (0,1)$,

   - f admits an invariant probability measure $\nu \ll m$ is absolutely continuous with respect to Lebesgue measure m. As shown for example in [62] and [52].
   - f is Mixing, and there exists $C > 0$ such that

     $$C_n(\psi_1, \psi_2, \nu) \leq C\|\psi_1\|_{\text{Lip}}\|\psi_2\|_{L^\infty}n^{1-1/\beta}.$$

     See [94]. Furthermore, it was shown in [7] that for $\beta \in (0, 0.5)$ this map has short range dependence and for $\beta \in (0.5, 1)$ it has long range dependence for $\mathcal{B}$ the class of Lipschitz continuous function. Short or long range dependence refers to summability (or not) of correlation which we will discuss in Chapter 5.

3. Hénon map.

   This map is defined by:

   $$f\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 - ax^2 + y \\ bx \end{bmatrix} \tag{2.35}$$

[1] See Appendix A for the definition of Lipschitz continuous.

where, $a = 1.4, b = 0.3$.
For a positive Lebesgue measure set of parameters $(a, b) \in \mathbb{R}$ the following hold, see [5] and [96].

- f admits an SBR measure ν. See, [96] and [82].
- f has exponential decay of correlation (Mixing) for $\mathcal{B}$ the class of Lipschitz continuous functions. In particular there exists $C > 0$ and $\theta < 1$ such that

$$C_n(\psi_1, \psi_2, \nu) \leq C\|\psi_1\|_{\text{Lip}}\|\psi_2\|_{\text{Lip}}\theta^n.$$

  See, [95] and [5]. Notice that we must have both $\psi_1, \psi_2 \in \mathcal{B}$.

However, it is open question if these statements are true for $a = 1.4, b = 0.3$ so our investigation is numerical.

2.3. Extremes in Dynamical Systems.

Consider a dynamical system $(\mathcal{X}, \nu, f)$ where, $\mathcal{X} \subset \mathbb{R}$, $f : \mathcal{X} \to \mathcal{X}$ is a measurable map and ν is an $f - invariant$ probability measure. Given an observable function $\phi : \mathcal{X} \to \mathbb{R}$ achieving a global maxima at $\tilde{x} \in \mathcal{X}$, we assume that it has the representation $\phi(x) = \psi(\text{dist}(x, \tilde{x}))$, where $\psi : \mathbb{R}^+ \to \mathbb{R}$, and $\text{dist}(\cdot, \cdot)$ is a distance function or appropriate metric (e.g. Euclidean distance). In this chapter and in Chapter 3 we will focus on the case where $\tilde{x}$ is a non periodic point. Let ξ_1, ξ_2.... be a stationary sequence and again let $M_n = \max(\xi_1, ..., \xi_n)$ where we have here $\xi_n = \phi(f^{n-1})$. The invariant measure ensures that the sequence is a stationary sequence.

What we want to show in dynamical systems is that the corresponding result holds.

$$n\nu\{x : \phi(x) \geq u_n\} \to \tau \Leftrightarrow \nu(M_n \leq u_n) \to e^{-\tau}, \tag{2.36}$$

where $\nu\{x : \phi(x) \geq u_n\}$ corresponds to $(1 - F(u_n))$ as described in Theorem 2.2 and 2.3 i.e. taking $P = \nu$

$$\begin{aligned} F(u) &= \nu(\xi_1 \leq u) \\ &= \nu(x : \phi(x) \leq u) \\ &= \nu(x : x \geq \phi^{-1}(u)). \end{aligned} \tag{2.37}$$

It was shown for example, in [16], [37], [45] and [55] that the convergence of maxima has the same behaviour as maxima of an associated IID sequence under the conditions $D'(u_n)$ and $D_2(u_n)$. $D'(u_n)$ condition is the same as the one defined before. $D_2(u_n)$ condition was first introduced by Collet [16] and is given below.

Definition 2.17 ($D_2(u_n)$). *We say that $D_2(u_n)$ holds for the sequence $\xi_1, \xi_2, ...$ if for any integers l, t and n*

$$|\nu(\{\xi_1 > u_n\} \cap \{M_{l,t} \leq u_n\}) - \nu(\xi_1 > u_n)\nu(M_l \leq u_n)| \leq \gamma(n,t). \tag{2.38}$$

Where $\gamma(n,t)$ is non increasing in t for each n and $n\gamma(n,t_n) \to 0$ as $n \to \infty$ for some sequence $t_n = o(n)$, $t_n \to \infty$, and, $M_{l,t} = \max(\xi_t, \xi_{t+1}, ..., \xi_{t+l})$.

Theorem 2.9 *[37] Suppose (u_n) is such that $n\nu\{x : \phi(x) \geq u_n\} \to \tau$ as $n \to \infty$ for some $\tau \geq 0$. If $D_2(u_n)$ and $D'(u_n)$ hold then,*

$$\nu(M_n \leq u_n) \to e^{-\tau} \text{ as } n \to \infty.$$

This theorem is similar to Theorem 2.2 but is, now applicable to dynamical systems. Furthermore, it was shown in [38], [45], [32], [66] and [54] that under some observable functions ϕ which we will define later with absolutely continuous invariant probability measure and some mixing conditions the limit distribution of extremes converge to the three types of the GEV distributions. For a wide range of dynamical systems it is possible to derive general conditions on the recurrence and decay of correlations that ensure that we have convergence to a GEV distribution, see for example [54]. Furthermore, under similar conditions Holland and Nicol [53] studied the speed of convergence of $\nu(M_n \leq u_n)$ to $G(u)$. These conditions are given below:

- (H1) (Decay of correlations). There exists a monotonically decreasing sequence $\Theta(j) \to 0$ such that for all ψ_1 Lipschitz continuous and all $\psi_2 \in L^\infty$:

$$\left| \int \psi_1 . \psi_2 \circ f^j d\nu - \int \psi_1 d\nu \int \psi_2 d\nu \right| \leq \Theta(j) \|\psi_1\|_{\text{Lip}} \|\psi_2\|_{L^\infty}.$$

  Where $\| \, . \, \|_{Lip}$ denotes the Lipschitz norm.

- (H1s) (Decay of correlations: strong). There exists a monotonically decreasing sequence $\Theta(j) \to 0$ such that for all ψ_1, ψ_2 Lipschitz continuous:

$$\left| \int \psi_1 . \psi_2 \circ f^j d\nu - \int \psi_1 d\nu \int \psi_2 d\nu \right| \leq \Theta(j) \|\psi_1\|_{\text{Lip}} \|\psi_2\|_{\text{Lip}}.$$

- (H2a)(Strong quantitative recurrence rates). There exist numbers $\gamma, \alpha > 0$ such that

$$\tilde{g}(n) \sim n^\gamma \Rightarrow \nu(E_n) \leq \frac{C}{n^\alpha}.$$

  Where, $\tilde{g} : N \longrightarrow \mathbb{R}$ is a monotonically increasing function and, $E_n := \{x \in \mathcal{X} : \text{dist}(x, f^j(x)) \leq \frac{1}{n}, \text{ for some } j \in [1, \tilde{g}(n)]\}$. If $x \in \bigcap_{n=1}^\infty E_n \Rightarrow x$ recurrent.

- (H2b)(Weak quantitative recurrence rates). For some $\tilde{\gamma} > 1, \alpha > 0$:

$$\tilde{g}(n) \sim (\log n)^{\tilde{\gamma}} \Rightarrow \nu(E_n) \leq \frac{C}{n^{\alpha}}.$$

Condition (H1) can be considered as the $D_2(u_n)$ condition by taking $\psi_1 = 1_{\{\xi_1 > u_n\}}$ and $\psi_2 = 1_{\{M_l \leq u_n\}}$ as mentioned in [37]. For hyperbolic dynamical systems, condition (H1s) is required over and above (H1), and further approximation methods must be used to ensure condition $D_2(u_n)$ holds, see for example [45]. Condition (H2) is a condition on the recurrence rate and together with condition (H1)/ H1s) can be used to verify the $D'(u_n)$ condition.

2.3.1. Rate of convergence in dynamical systems

Under the above suitable conditions Holland and Nicol [53] estimated the difference between $\nu(M_n \leq u_n)$ and $G_{\sqrt{n}}(u)$ where, $G_{\sqrt{n}}(u) = (1 - \tau_n(u)/\sqrt{n})^{\sqrt{n}}$ and $\tau_n(u) = n\nu(x : \phi(x) > u/a_n + b_n)$ as presented in the following Theorem.

Theorem 2.10 *Suppose that $f : \mathcal{X} \to \mathcal{X}$ is a map with ergodic measure ν with density in $L^{1+\delta}(m)$ for some $\delta > 0$, and ν absolutely continuous with respect to m. We have the following cases. (1) Suppose that $\Theta(n) = O(\theta_0^n)$ for some $\theta_0 < 1$ and (H1),(H2b) hold. Then for all $\epsilon > 0$ and $\nu - a.e.\, \tilde{x} \in \mathcal{X}$ we have that*

$$|\nu(M_n \leq u_n) - G_{\sqrt{n}}(u)| \leq C_1 \frac{(\log n)^{1+\epsilon}}{\sqrt{n}} + \frac{C_2}{n^{\alpha-\epsilon}}, \tag{2.39}$$

where C_1, $C_2 > 0$ are constants independent of n, but dependent on $\tilde{x}$.
(2) Suppose that $\Theta(n) = O(n^{-\zeta})$ for some $\zeta > 0$ and (H1),(H2b) hold. Then for all $\epsilon > 0$ and $\nu - a.e.\, \tilde{x} \in \mathcal{X}$ we have that

$$|\nu(M_n \leq u_n) - G_{\sqrt{n}}(u)| \leq C_1 n^{-\frac{1}{2}+\kappa} + C_2 n^{-\alpha+\kappa}, \;\; \textit{with } \kappa = \epsilon + \frac{2(1+2\delta)}{\zeta\delta} \tag{2.40}$$

where C_1, $C_2 > 0$ is a constant independent of n, but dependent on $\tilde{x}$.

Remark. If ζ or δ are close to zero then the error estimate of equation (2.40) is of little use since it is large.

This theorem gives an estimate of the error rates of $\nu(M_n \leq u_n) \to G_{\sqrt{n}}(u)$ under the above suitable conditions. In the block Maxima method which is a common method used for modelling maxima (we will describe this method in detail later). We have $n = pq + r$ with $0 \leq r < p$ where, p is the block number and q is the block size, by writing $n = p(n)q(n)$. In order to obtain an observation that behaves as an independent observation we may take a gap of length t between each block where, $t = g(n)$.

Proposition 2.2 *Suppose that $f : \mathcal{X} \to \mathcal{X}$ is ergodic with respect to a measure ν which has a density $h \in L^{1+\delta(m)}$ for some $\delta > 0$. Suppose that (H1) holds. Then for $\nu - a.e.\, \tilde{x} \in \mathcal{X}$, all p, q such that $n = pq + r$, and $t < p$ we have*

$$|\nu(M_n \leq u_n) - (1 - p\nu(\xi_1 > u_n)^q| \leq \mathcal{E}_n, \tag{2.41}$$

where for any $\epsilon > 0$ *and* $\delta_1 = \delta/(1+2\delta)$:

$$\mathcal{E}_n = \max(qt, p)\nu(\xi_1 \geq u_n) + qp^2(\nu(\xi_1 > u_n))^2 + C_1 pq^2 \Theta(t)^{\delta_1 - \epsilon} + pq \sum_{j=2}^{t} \nu(\xi_1 > u_n, \xi_j > u_n). \tag{2.42}$$

Apart from the proof of this Proposition being related to the proof of Theorem 2.5 as equations (2.12) and (2.13) with $k = \sqrt{n}$. Under this Proposition Holland and Nicol [53] proved Theorem 2.10 and found that the optimal block number and size is $p \approx q \approx \sqrt{n}$. In our application we will investigate this choice for a range of dynamical system examples.

As we mentioned before the three types of the GEV distribution can be obtained by choosing the suitable observable function as described by Freitas and Freitas [37]. The three types of GEV distributions and the corresponding tail index ρ can be generated from the following three observable functions maximized at $\tilde{x}$.

$$\phi_1(x) = -\log(\text{dist}(x, \tilde{x})) \to \text{Type I.} \tag{2.43}$$

The corresponding shape: $\rho = 0$

$$\phi_2(x) = (\text{dist}(x, \tilde{x}))^{-\frac{1}{\alpha}} \to \text{Type II.} \tag{2.44}$$

The corresponding shape: $\rho = 1/(\alpha D(\tilde{x}))$

$$\phi_3(x) = C - (\text{dist}(x, \tilde{x}))^{\frac{1}{\alpha}} \to \text{Type III.} \tag{2.45}$$

The corresponding shape: $\rho = -1/(\alpha D(\tilde{x}))$.

Where, $\alpha > 0$, and $\alpha, C \in \mathbb{R}$. Here, *dist* is the distance between x and $\tilde{x}$ and $D(\tilde{x})$ is the local dimension of a measure ν at $\tilde{x}$ that is given by

$$D(\tilde{x}) = \lim_{r \to 0} \frac{\log \nu(B(\tilde{x}, r))}{\log r} \tag{2.46}$$

(if the limit exists). The upper and lower local dimension at the point $\tilde{x}$ are defined respectively by

$$\overline{D}(\tilde{x}) = \limsup_{r \to 0} \frac{\log \nu(B(\tilde{x}, r))}{\log r}, \quad \underline{D}(\tilde{x}) = \liminf_{r \to 0} \frac{\log \nu(B(\tilde{x}, r))}{\log r}. \tag{2.47}$$

Where, $B(\tilde{x}, r)$ is a ball of radius r centred on $\tilde{x}$. If $\overline{D} = \underline{D} = D$ then we say that the local dimension exists at $\tilde{x}$. Further details on the dimension theory in Dynamical system can be found for example in [93], [30] and [76].
A description on the meaning of these observables can be found in [37] and [66].

In IID RVs we discussed the convergence of probabilities of the form $P(M_n \leq u_n)$, where $u_n = u/a_n + b_n$. We showed that a_n and b_n can be obtained if F is in the Domain of Attraction. See, Theorem 2.3. Similarly, for dynamical systems $F(u_n)$ corresponds to the measure $\nu\{x : \phi(x) \leq u_n\}$. As before we consider linear scaling $u_n = u/a_n + b_n$, and investigate the limiting behaviour of $n\nu(\phi > u_n)$, and in particular if it converges to $\tau(u)$, for some $\tau(u)$. In the following, we suppose that ν is absolutely continuous with respect to the Lebesgue measure, and for simplicity we assume the dimension of the system is one.

For the observable $\phi_1(x) = -\log(\text{dist}(x, \tilde{x}))$ we have that

$$\nu(\phi_1(x) \geq u_n) = \nu(x : \text{dist}(x, \tilde{x}) \leq e^{-u_n})$$

If e^{-u_n} is small, then $\nu(x : \text{dist}(x, \tilde{x}) \leq e^{-u_n}) \sim 2h(\tilde{x})e^{-u_n}$. Here, h is the probability density $d\nu(\tilde{x}) = h(\tilde{x})dx$.
If we let $u_n = u + \log n$ then this implies that

$$\nu\{x : \phi_1(x) \geq u_n\} \sim 2h(\tilde{x})e^{-u}/n.$$

Thus we get a type I behaviour for the limit.

Similarly, for the observable $\phi_2(x) = (\text{dist}(x, \tilde{x}))^{-\frac{1}{\alpha}}$. we have that

$$\nu(x : \text{dist}(x, \tilde{x})^{-\frac{1}{\alpha}} \geq u_n) = \nu(x : \text{dist}(x, \tilde{x}) \leq u_n^{-\alpha}) \sim 2h(\tilde{x})u_n^{-\alpha}/n.$$

Let holds, $u_n = u/a_n + b_n, b_n = 0, a_n = n^{-1/\alpha}$ so $u_n = u/n^{-1/\alpha}$. This implies that

$$\nu\{x : \phi_2(x) \geq u_n\} \sim \frac{2h(\tilde{x})u_n^{-\alpha}}{n}/n.$$

Furthermore, we consider the following observable function

$$\phi_4(x) = (\text{dist}(x, \tilde{x}))^{-\frac{1}{\alpha}}(-\log(\text{dist}(x, \tilde{x}))) \rightarrow \text{Type II.} \tag{2.48}$$

For simplicity consider the $\alpha = 1$ case. We can write $\phi_4(x) = \psi(\text{dist}(x, \tilde{x}))$, with $\psi(y) = y^{-1}(-\log y)$ and we take $y \rightarrow 0$. Using the techniques of Section 2.2.2, an asymptotic inverse is given by $\psi^{-1}(y) = (-y \log y)^{-1}$. We have the following:

$$\nu\{x : \phi_4(x) \geq u_n\} \sim 2h(\tilde{x})\psi^{-1}(u_n),$$

and we choose a_n, b_n so that $\psi^{-1}(u_n) = u^{-1}/n$, i.e.

$$\frac{u}{a_n} + b_n = \psi(1/(un)) = nu(\log n + \log u).$$

Since we take $n \to \infty$ with u fixed, we can choose $b_n = 0$, and $a_n = 1/(n \log n)$. We therefore have:

$$\nu\{x : \phi_4(x) \geq u(n \log n)\} \sim 2h(\tilde{x})u^{-1}/n. \tag{2.49}$$

When $\alpha \neq 1$, we again get the same asymptotic as in the Type II case above, but in this case we choose a scaling sequence of the form $a_n = n^{-1/\alpha}\ell(x)$. The function $\ell(x)$ is slowly varying and can be estimated using equation (2.26) in Section 2.2.2.

Consider the case now where we have a more general SRB measure ν, whose local dimension at $\tilde{x} \in \mathcal{X}$ is $D(\tilde{x})$. In this case equation (2.46) gives us bounds on the behaviour of $\nu\{x : \phi(x) \geq u_n\}$ for a given sequence u_n. If we are to obtain convergence to an extreme value distribution of Types I-III we need to recover asymptotics as described above. We will show however, that equation (2.46) gives only upper and lower bounds on the limiting behaviour of $\nu\{\phi(x) \geq u_n\}$ if we take linear scaling sequences.

First of all equation (2.46) gives the following: $\forall \epsilon > 0$, $\exists r_0 > 0$, such that for all $r < r_0$ we have

$$r^{D+\epsilon} \leq \nu(B(\tilde{x}, r)) \leq r^{D-\epsilon}. \tag{2.50}$$

Moreover, for general SRB measures it is known that

$$\liminf_{r\to 0} \frac{\nu(B(\tilde{x}, r))}{r^D} = 0, \ \limsup_{r\to 0} \frac{\nu(B(\tilde{x}, r))}{r^D} = \infty.$$

In the case of measures ν absolutely continuous with respect to Lebesgue measure (and so $D \in \mathbb{N}$), we have

$$\liminf_{r\to 0} \frac{\nu(B(\tilde{x}, r))}{r^D} = C_D h(\tilde{x}),$$

where C_D is a uniform constant, and $h(\tilde{x})$ is the Radon-Nikodym derivative of ν.

If $\phi(x) = \psi(\text{dist}(x, \tilde{x})$, then we have that

$$(\psi^{-1}(u_n))^{D+\epsilon} \leq \nu\{x : \phi(x) \geq u_n\} \leq (\psi^{-1}(u_n))^{D-\epsilon}. \tag{2.51}$$

If $\psi(y) = -\log y$, then we might expect a Type I convergence. However for a linear scaling sequence, we cannot get the same asymptotic behaviour as in the case of absolutely continuous invariant measures. If indeed $\nu\{\phi(x) \geq u_n\} \to \tau(u)$, then using equation (2.51) we obtain bounds on the (now nonlinear) sequence u_n as follows:

$$\lim_{n\to\infty} n(e^{-u_n})^{D+\epsilon} \leq \tau, \implies u_n \geq \frac{u + \log n}{D + \epsilon}, \tag{2.52}$$

$$\lim_{n\to\infty} n(e^{-u_n})^{D-\epsilon} \geq \tau, \implies u_n \leq \frac{u + \log n}{D - \epsilon}, \tag{2.53}$$

where we assumed Type I behaviour with $\tau(u) = e^{-u}$.

As discussed in Leadbetter et al. [61] Theorem 2.6, and [53] on the convergence of $\nu(M_n \leq u_n)$ to $G(u)$ where, $G(u)$ is one of the type I-III distributions, in the next chapter we will investigate this convergence for particular examples, such as the Times 3 map, Intermittent map and Hénon map. We review below the theoretical convergence results. First of all, we have the following corollary which is a simple consequence of Theorem 2.6.

Corollary 2.2 *Let $(f, \mathcal{X}, \nu)$ be an ergodic dynamical system, then:*

$$|\nu(M_n \leq u_n) - G(u)| \leq |\nu(M_n \leq u_n) - G_{\sqrt{n}}(u)| + \triangle_{\sqrt{n}} + \triangle'_n. \tag{2.54}$$

We prove this result as follows. First we have:

$$|\nu(M_n \leq u_n) - G(u)| \leq |\nu(M_n \leq u_n) - G_{\sqrt{n}}(u)| + |G_{\sqrt{n}}(u) - G(u)|,$$

where

$$\begin{aligned} |G_{\sqrt{n}}(u) - G(u)| &= |(1 - \tau_n(u)/\sqrt{n})^{\sqrt{n}} - G(u)| \\ &\leq |(1 - \tau_n(u)/\sqrt{n})^{\sqrt{n}} - e^{-\tau_n(u)}| + |e^{-\tau_n(u)} - G(u)|. \end{aligned}$$

By Applying Theorem 2.6 we obtain that

$$|G_{\sqrt{n}}(u) - G(u)| \leq \triangle_{\sqrt{n}} + \triangle'_n.$$

To estimate $|\nu(M_n \leq u_n) - G_{\sqrt{n}}(u)|$ we just appeal to Theorem 2.10. For the Times 3 map, and the Intermittent map we have the following results, see [53, Section 3].

Proposition 2.3 *Suppose $(f, [0,1], Leb)$ is the times d map, $d = 3, 4, ...$ For the observation $\phi_1(x) = -log|x - \tilde{x}|$, we have for $Leb - a.e.\, \tilde{x} \in [0,1]$ and all $\epsilon > 0$:*

$$|\nu(M_n \leq u + \log n) - e^{-2e^{-u}}| \leq C \frac{(\log n)^{1+\epsilon}}{\sqrt{n}}. \tag{2.55}$$

Where, $C(\tilde{x}) > 0$ is a uniform constant dependent on $\tilde{x}$.

Proposition 2.4 *Suppose $(f, [0,1], \nu)$ is the intermittent map defined by*

$$f(x) = \begin{cases} x(1 + 2^\beta x^\beta) & x \leq \frac{1}{2}, \\ 2x - 1 & x > \frac{1}{2}. \end{cases}$$

with, $\beta < 1/20$. For the observation $\phi_1(x)$ defined before, we have for $\nu - a.e.\, \tilde{x} \in [0,1]$ and for all $\epsilon > 0$:

$$|\nu(M_n \leq u_n) - e^{-2h(\tilde{x})e^{-u}}| \leq C n^{-0.5+k}, \text{ with } k = \epsilon + 10\beta. \tag{2.56}$$

Where, $C(\tilde{x}) > 0$ is a constant independent of n.

For hyperbolic dynamical systems, the discussions presented earlier concerning local dimension estimates make it difficult to estimate Δ'_n. Hence we obtain the following result concerning the Hénon map. See [53, Section 4].

Proposition 2.5 *Suppose $(f, [-1,1]^2, \nu)$ is the Hénon map dynamical system, where ν is the ergodic measure. Then there exist $\alpha > 0$ such that*

$$|\nu(M_n \leq u_n) - G_{\sqrt{n}}(u)| \leq \frac{C}{n^\alpha}. \tag{2.57}$$

In this latter example, despite the potential non-convergence of the limit $\tau_n(u) := n\nu\{\phi(x) > u_n\}$, it is feasible to study how this quantity scales (or oscillates) as $n \to \infty$. We study this quantity in Chapter 3.

3. Numerical Estimate of the Tail Index

3.1. Introduction

There are two known approaches of modelling maxima or extremes. The block maxima method is the old and widely used model. Here the block maxima consists of dividing the data into numbers of blocks with equal size or length. The problem with this method is the optimal block size and number. Recently, in a dynamical system Faranda et al. [32] and Lucarini et al. [66] presented a numerical convergence of the block maxima to the extreme value distribution. In their study they fixed the length of the series and varied the block number. As we discussed in the previous chapter the theoretical optimal block size and number in the dynamical system has been investigated recently by Holland and Nicol [53]. In this chapter we extended and used the same numerical analysis described by [32] and [66] in a wide range of dynamical system examples. We also compare the result with the optimal block size and number as given in [53].

The second method is the excess over high threshold. In this method extremes correspond to data that exceed a specified level or threshold value. In this chapter, the block maxima method and the excess over threshold methods are described in Sections 3.2 and 3.3 respectively. The block maxima method is applied to the Dynamical systems examples described in the previous chapter (expanding Times 3, Intermittent and Hénon maps) and this is for a range of block numbers, in Section 3.4. The excess over threshold method together with the mean excess function are applied to the expanding Times 3 and Hénon maps, in Section 3.4. Furthermore we applied both methods to the financial returns data from the Standard and poor 500 (S&P500) return index and Financial Times Stock Exchange 100 (FTSE100) return index in Section 3.5.

3.2. Block Maxima

In the application the block maxima method is used to fit the GEV distribution to the maximum blocks. The algorithm works as follows: divide your data $\xi_1, ..., \xi_n$ into p blocks of size q, here $n = pq$. For each block calculate the maximum $M_q i = \max(\xi_{(i-1)q+1}, ..., \xi_{qi})$, $i = 1, ..., p$. Then fit the GEV distribution to $\{M_{q1}, M_{q2}, ..., M_{qp}\}$. There are different ways of fitting GEV distribution, such as, the Maximum Likelihood (ML) and Probability Weighted Moment(PWM) estimators see, [56]. A discussion about these estimators can be found in [26]. In our data examples we used the ML estimator. Let $\{M_{q1}, ..., M_{qp}\}$ be our sample data from the block maxima method. Assume this sample data are IID from

GEV distribution that is given by equation (2.2). Then the likelihood function can be written as:

$$L(\rho, \mu, \sigma; M_{q1}, ..., M_{qp}) = \prod_{i=1}^{p} g_{\rho,\mu,\sigma}(M_q i), \tag{3.1}$$

where, $g_{\rho,\mu,\sigma}$ is the density of the GEV distribution. The log likelihood function is given by taking the logarithm of the likelihood function above

$l(\rho, \mu, \sigma; M_{q1}, ..., M_{qp}) = \sum_{i=1}^{p} \ln g_{\rho,\mu,\sigma}(M_q i)$

$$= -p \ln \sigma - (1 + 1/\rho) \sum_{i=1}^{p} \ln(1 + \rho \frac{M_{qi} - \mu}{\sigma}) - \sum_{i=1}^{p} (1 + \rho \frac{M_{qi} - \mu}{\sigma})^{-1/\rho}. \tag{3.2}$$

See, [26] and [25].
The estimated parameter can be found by solving the maximize problem of the log likelihood function $(\frac{\partial l}{\partial \rho}, \frac{\partial l}{\partial \mu}, \frac{\partial l}{\partial \sigma}) = 0$. It was shown in [32] that $a_n = 1/\sigma$ and $b_n = \mu$. In the application the result is obtained using the MATLAB Statistics Toolbox functions such as (gevfit) see for example, [71]. The ML estimator is efficient for $\rho > -1/2$. In general, the Financial data has $\rho \geq 0$, see [26, p. 319]. So in our application the ML estimator can be used to estimate the parameters of the GEV distribution. It is difficult to find the optimal choice of the block number and size. For example, if the block size is large this leads to small bias in estimation of the parameters. Whereas, if the block number is large this leads to more block maxima and small variance. However, an accurate result can be obtained by taking very large data. Leadbetter et al. [61] studied the rate of convergence to GEV distribution. They found that the fit can be better by taking $n = 10^3 - 10^6$. Furthermore, in a dynamical system recently, Holland and Nicol [53] studied the rate or speed of convergence of GEV distributions described before in Theorem 2.10. They found that optimal choice for the block number and size is by taking $p \approx q \approx \sqrt{n}$. (Their results do not hold if $\tilde{x}$ is periodic, also they focused on the observable $\phi_1(x)$. For other observations the error could be much larger, see Example 2.5) . Furthermore, a numerical study of the block maxima to GEV distribution for the observables $\phi_i(x)$, $i = 1, 2, 3$ can be found in [32]. They found that an appropriate choice for block size and number is by taking $p > 10^3$ and $q > 10^3$.

In our application we apply the block maxima method to S&P500 return index, FTSE100 return index and to examples of dynamical systems such as Times 3 map and Hénon map. Moreover, in dynamical system examples we plot the estimated shapes versus a range of block number and size and compare the result with the theoretical shape value. The theoretical shape value for the observable ϕ_1, ϕ_2 and ϕ_3 is $\rho = 0, \rho = 1/(\alpha D(\tilde{x}))$ and $\rho = -1/(\alpha D(\tilde{x}))$ respectively. A discussion on the theoretical value of the shape parameter can be found in [66]. As we discussed before the block maxima method has its drawback in that it needs very large data to get an accurate result. Another method used to model maxima is excess over threshold method which is described in the next section.

3.3. Excess over Threshold

Consider we have an observation $\xi_1, ..., \xi_n$ with distribution function F and threshold u. Theorem 2.7 shows that the data over high threshold u can be modelled by a GP distribution. Denote the data that exceed the threshold u by $\tilde{\xi_1}, ..., \tilde{\xi_N}$, where, $1 < N < n$. Compute $M_i = \tilde{\xi_i} - u$, so our data will be $M_1, ..., M_N$ then, fit the GP distribution to this data.

Similar to the block maxima method the Maximum likelihood is one method used to fit the GP distribution. An alternative method is the probability weighted moment(PWM). In our application we use the ML estimator to estimate the parameters of the GP distribution. The results obtained using the MATLAB Statistics Toolbox functions such as (gpfit) see for example, [72]. The log likelihood function is given by:
$l(\rho, \beta; M_1, ..., M_N) = \sum_{i=1}^{N} \ln gp_{\rho,\beta}(M_i)$

$$= -N \ln \beta - (1 + 1/\rho) \sum_{i=1}^{N} \ln(1 + \rho \frac{M_i}{\beta}), \tag{3.3}$$

where $gp_{\rho,\beta}$ is the density of the GP distribution.
See, [26] and [25].

3.3.1. Threshold selection

The problem with the excess over threshold method is to find the optimal threshold value. In the literature a number of approaches have been introduced for estimating the threshold value. A review of extreme value threshold estimation can be found in [83]. We have discussed in the previous Chapter that if (X_n) is a sequence of IID RVs with distribution function $F(x) = P(X > u_n)$ in the domain of attraction then $u_n = x/a_n + b_n$. But the problem is if F is unknown. Examples of methods used to investigate this problem are the threshold stability plots, which plot the parameter estimated using Maximum likelihood estimator versus a range of thresholds assuming the data has GP distribution as illustrated in [15]. Other methods under the Bayesian approach such as [3] and [74] used the Gamma distribution below the threshold and the GP distribution above. In addition [14] used the normal distribution below the threshold and GP distribution above the threshold. As discussed in [83] the disadvantage of this approach is that they are not much used in practise. A natural choice of the threshold value is as a high quantile but how high this quantile should be is a matter for debate. The quantile function of the distribution function F is given by

$$F^{\leftarrow}(t) = \inf\{x \in \mathbb{R} : F(x) \geq t\}, \;\; 0 < t < 1. \tag{3.4}$$

Where, $x_t = F^{\leftarrow}(t)$ is the $t-$ quantile of F. See [26].
As we mentioned in Chapter 2, see equation (2.31), the linearity of the mean excess plot is one tool for choosing the appropriate threshold value. Assuming extremes over threshold

have GP distribution see Definition 2.6. This method is widely used in the literature see for example [26], [15] and [25]. The empirical mean excess plot is defined by:

$$e_n(u) = \frac{\sum_{i=1}^{n}(\xi_i - u)}{\sum_{i=1}^{n} I_{(\xi_i > u)}} \tag{3.5}$$

We plot $e_n(u)$ versus a range of threshold u then we choose u that corresponds to the linear part of the function. Once we choose the appropriate threshold the shape can be estimated by using the ML estimator. As pointed out by Embrechts et al. [26, p. 356] "The reader should never expect a unique choice of u to appear". Also, according to Embrechts et al. [25, p.280] the threshold value can be chosen if the plot shows evidence of a kink below the threshold value and straightening out above it.
Furthermore, if the data has extremal index ($\theta < 1$) this means that the data cluster over the threshold so the data accept dependency, then the ML estimator cannot be applied to estimate the parameters. One method may deal with this problem is by declustering the data see for example, [15, pp. 99] and [25, p. 303]. (A description of the extremal index will be presented in Chapter 4.) In this chapter we will choose the threshold based on the mean excess plot and we may consider other methods to find the optimal threshold value in our future study see, [83]. Furthermore, according to [65] and [67] the excess over threshold method in dynamical systems gives the same results obtained by the block maxima under the distance observable function. Moreover, they stated that there is no need for the mixing condition assumption to apply to this method. In addition, they did not consider the problem with the optimal threshold value. In our application of the dynamical system examples we assume that $\tilde{x}$ is not a periodic point hence the data is unlikely to cluster above the threshold.

3.4. Data Examples

In this section we estimated the shape parameter using the block Maxima method versus a range of block numbers for large observations. We applied this for three maps, Times 3, Intermittent and Hénon maps. Furthermore, we applied the excess over Threshold method to Time 3 and Hénon map.

3.4.1. Expanding Times 3 map

For $f : [0, 1] \to [0, 1]$, the map is defined by:

$$f(x) = 3x \mod 1. \tag{3.6}$$

In this section we examine the convergence to GEV under the observables ϕ_1, ϕ_2 and ϕ_3 that is given by equations (4.15), (4.16) and (2.45). Set $\alpha = 2$ and choose $\tilde{x}$ at random (according to the Uniform distribution). By the ergodic theorem almost all $\tilde{x} \in (0, 1)$ is non-periodic and so we expect Theorem 2.10 to apply. In our simulation results we display

for a representative $\tilde{x}$. We start from an initial condition, we choose $x_0 = \sqrt{2}$ and, we iterate the map up to $n = 10^6$ iterations. (The choice of the initial value did not affect the estimated value of the shape parameter. Here x_0 could be any point generated from the standard uniform distribution on the interval $(0, 1)$) Then we fit the GEV distribution by using the ML estimator with $p = q = \sqrt{n}$ as illustrated in Figure 3.1 . We expect the true shape parameter value under the observables ϕ_1, ϕ_2 and ϕ_3 is $0, 1/(\alpha D(\tilde{x})) = 0.5$ and $1/(\alpha D(\tilde{x})) = -0.5$ respectively see, Section 3.2. The times 3 map is a one dimensional map, hence we set $D(\tilde{x}) = 1$. As presented in the following Table, we found that the results verified the theoretical shape value under the three observable functions.

Observable	$\hat{\rho}$	Theoretical value
ϕ_1	0.0069	0
ϕ_2, $\alpha = 2$	0.5143	0.5
ϕ_3, $\alpha = 2$	-0.5048	- 0.5

Table 3.1. Estimation of the shape parameter of the Times 3 map under the observables ϕ_1, ϕ_2 and ϕ_3 using the block Maxima method with $p = q = \sqrt{n}$ for $n = 10^6$.

Furthermore, for fixed n (we set $n = 10^8$) we estimated the shapes parameter for a range of block numbers under the observables ϕ_1, ϕ_2 and ϕ_3. We start from a minimum suitable block number p then we increase the block number (consequently, the block size q decreases $q = n/p$). Moreover, we compare the result with the theoretical shape parameter as illustrated in Figure 3.2. We found that an improvement of the convergence to the theoretical value increases with $p \geq 10^3$. This result is in agreement with the result obtained in [32]. Furthermore, we observed a good convergence with $p = q = \sqrt{n} = 10^4$ this result verified the theoretical result obtained by [53]. Since this map is a short memory process we can ignore the gap $t = g(n) < p$. In this case $t = \log n$ is small hence we expect that this does not affect the result.

Also, we consider the case when the observable function of the form

$$\phi_4(x) = (\text{dist}(x, \tilde{x}))^{-\frac{1}{\alpha}}(-\log(\text{dist}(x, \tilde{x}))). \tag{3.7}$$

where, $\alpha > 0$. As before and under this observable function we set $n = 10^8$ with $\alpha = 2$ and we estimated the shapes for a range of block number. The result is illustrated in Figure 3.2.d. The result we observe is a poor convergence in comparison with other observable function (ϕ_1, ϕ_2 and ϕ_3). This might be because the error rate for this observable is larger than the other observables. Also, we found that for $\alpha \leq 1$ the ML function does not converge.

We now analyse error rate as given in Chapter 2, and optimize block size.

By setting $n = pq$ with $t < p$ and since $\nu(\xi_1 > u_n) \sim \frac{\tau}{n}$ then the error rate expression in

Equation(2.42) can be written as

$$\mathcal{E}_n(p) = \max\left(\frac{nt}{p}, p\right)\frac{C}{n} + p\frac{C^2}{n} + C_1\frac{n^2}{p}\Theta(t)^{\delta_1-\epsilon} + n\tilde{C}, \tag{3.8}$$

where, $\tilde{C} = \sum_{j=2}^{t}\nu(\xi_1 > u_n, \xi_j > u_n)$. This error can be minimized as follows:
We may choose t such that $n^2\Theta(t) \to 0$ faster than $1/n$ and it is shown in [53] that $n\tilde{C} \to 0$. Hence we have,

$$\mathcal{E}_n(p) = \frac{t}{p} + \frac{p}{n}. \tag{3.9}$$

Minimize over p:

$$\mathcal{E}_n'(p) = \frac{-t}{p^2} + \frac{1}{n} \tag{3.10}$$

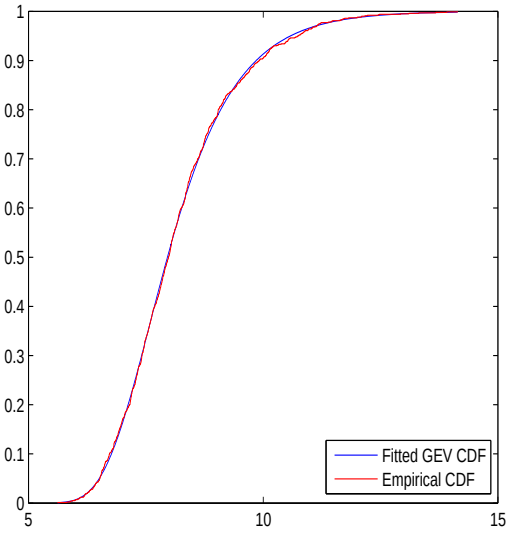
By solving $\mathcal{E}_n'(p) = 0$ we have $p = \sqrt{tn}$ hence, $\mathcal{E}_n(p) = 2\sqrt{\frac{t}{n}}$. which gives the first term on the right hand side of equation (2.39). Consequently, in this example with $n = 10^8$ the error rate will be $\approx \frac{4}{10^4} = 0.0004$, assuming that $t = (\log n)^{\tilde{\gamma}}$ for some $\tilde{\gamma} > 1$. In applications, see [53], and usually $\tilde{\gamma} = 2$ would suffice.

Now we apply the excess over threshold method under the observables ϕ_1, ϕ_2 with $\alpha = 2$ and this for $n = 5000$. Here, we assume that the data has GP distribution. We estimate the parameter by using the ML estimator. We choose the threshold value u based on the plot of the mean excess which is given in Figure 3.3. Here, we choose u when the mean excess plot is linear see Definition 2.6. The results are illustrated in the following Table.

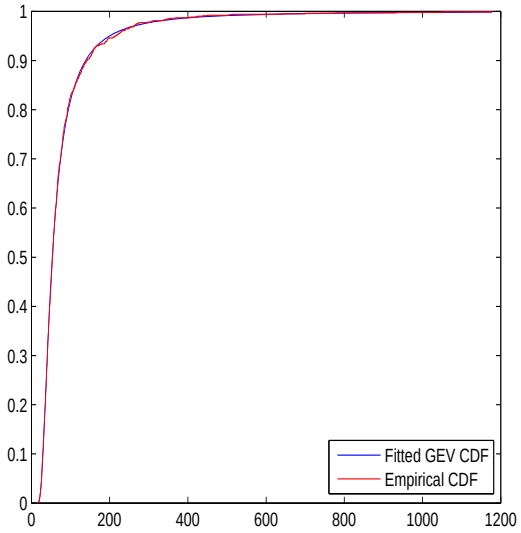
Observable	Threshold	Number of exceedances	Quantile	$\hat{\rho}$
ϕ_1	1.82	1614	0.68	-0.0027
	2.9	500	0.90	0.0302
ϕ_2, $\alpha = 2$	2.57	1555	0.70	0.5004
	5.7	279	0.95	0.5003
	10	100	0.98	0.7015

Table 3.2. The estimated shape parameter using the excess over threshold method for the Times 3 map under the observables ϕ_1, ϕ_2 and for $n = 5000$. The threshold values are chosen based on the mean excess plot.

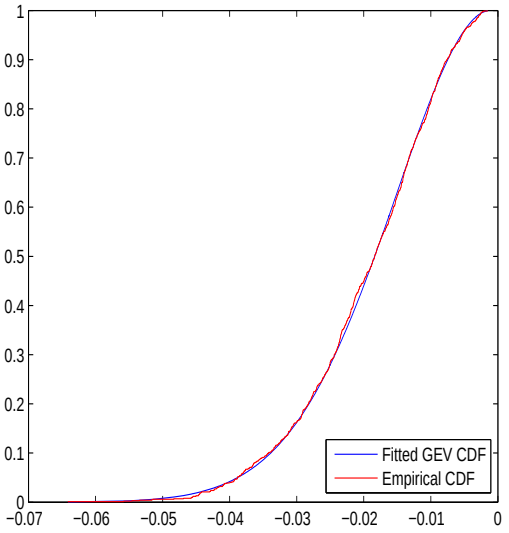
Furthermore, we plot the error $|\rho - \hat{\rho}|$ at different threshold values by using the excess over threshold method, where ρ is the theoretical shape parameter and $\hat{\rho}$ is the estimated shape parameter as illustrated in Figure 3.4. For the observable ϕ_1 we plot $|0-\hat{\rho}|$ versus a range of threshold values. For the observable ϕ_2 with $\alpha = 2$ we plot $|0.5-\hat{\rho}|$ versus a range of threshold values. In both observables the result is consistent with the mean excess plot. For example for the observable ϕ_1 by using the mean excess plot we chose $u = 1.82$; this value represents the smallest error in comparison with other threshold values. Similarly, for the observable ϕ_2 one of the thresholds we chose in the mean excess plot was $u = 5.7$ this value represents a very small error.



(a) ϕ_1

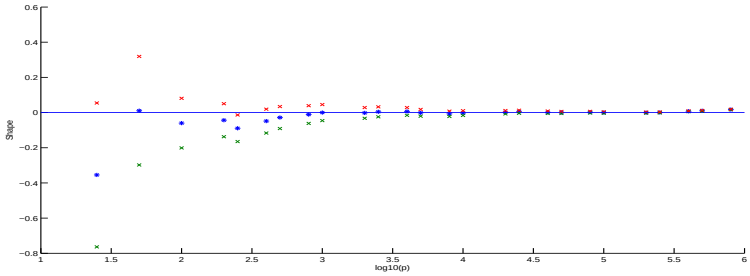


(b) ϕ_2

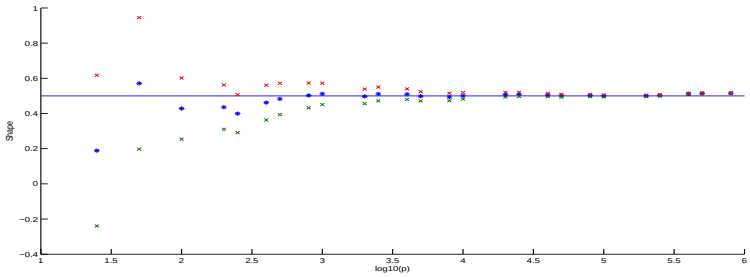


(c) ϕ_3

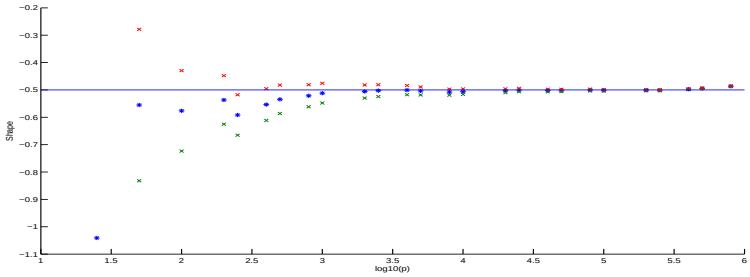
Figure 3.1. Times 3 map, $n = 10^6$ for: (a) ϕ_1 observable empirical distribution of block maxima and fitted GEV distribution. (b) ϕ_2 observable empirical distribution of block maxima and fitted GEV distribution (c) ϕ_3 observable empirical distribution of block maxima and fitted GEV distribution with $\alpha = 2, C = 0$.



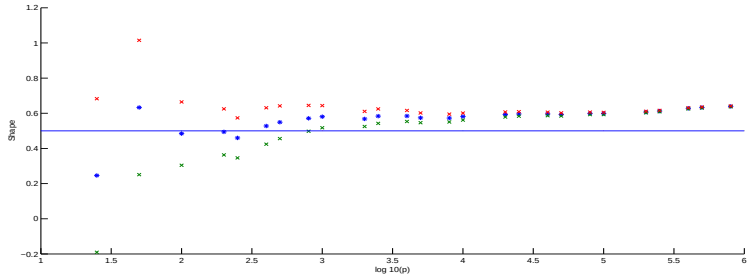
(a) ϕ_1



(b) $\phi_2, \alpha = 2$

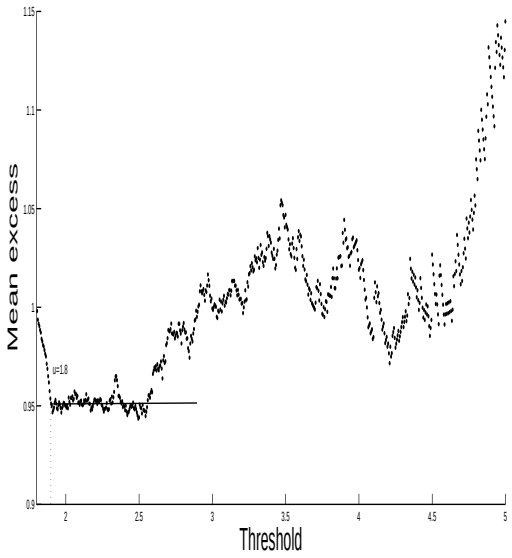


(c) $\phi_3, \alpha = 2$

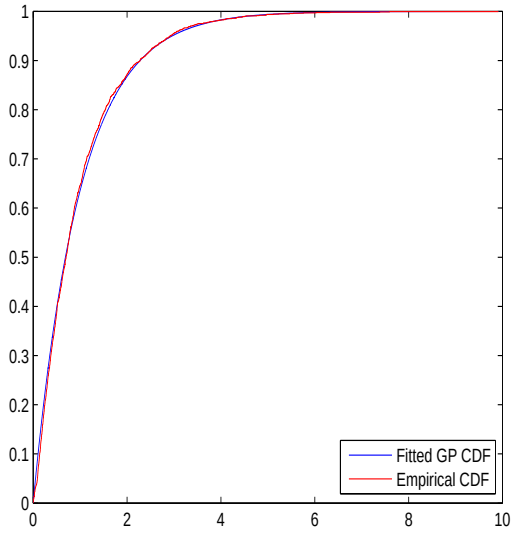


(d) $\phi_4, \alpha = 2$

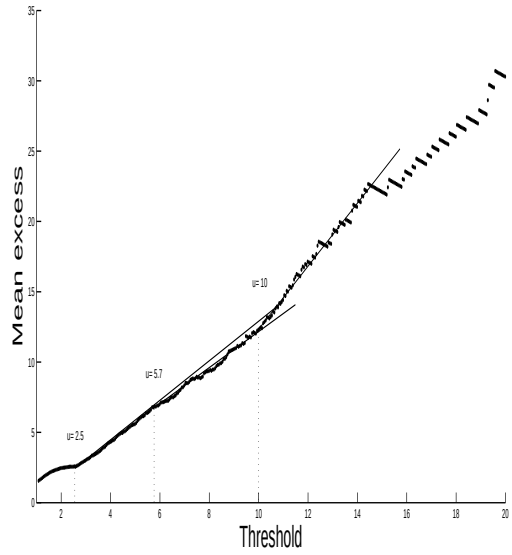
Figure 3.2. The estimated shape parameter(* line) with 95% confidence interval(x line) versus $\log_{10}(p)$ for Times 3 map,$n = 10^8$ (a) ϕ_1 observable (b) ϕ_2 observable with $\alpha = 2$. (c) ϕ_3 observable with $\alpha = 2$ and $C = 0$. (d) ϕ_4 observable with $\alpha = 2$. The straight line corresponds to the theoretical shape value.



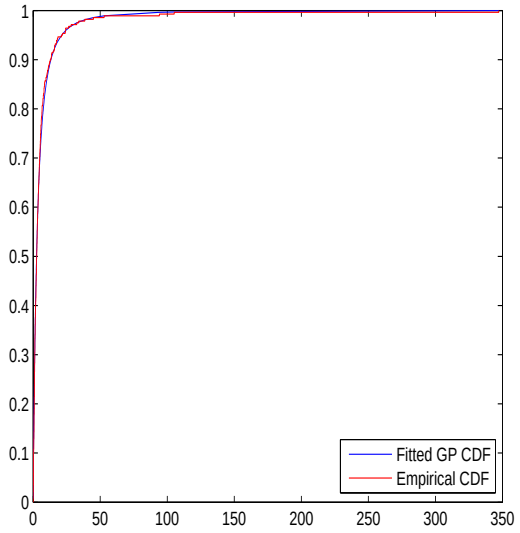
(a) ϕ_1



(b) ϕ_1

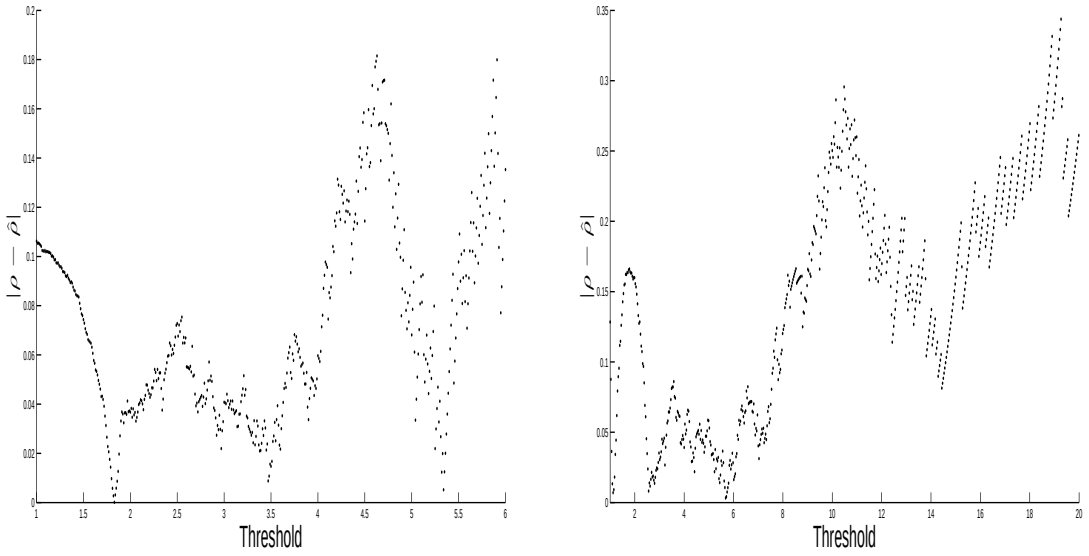


(c) ϕ_2



(d) ϕ_2

Figure 3.3. Times 3 map, observable ϕ_1: (a) The mean excess plot (b) Empirical distribution of excess over threshold $u = 1.8$ and fitted GP distribution. Times 3 map, observable ϕ_2 with $\alpha = 2$: (c) The mean excess plot (d) Empirical distribution of excess over threshold $u = 5.7$ and fitted GP distribution.



(a) ϕ_1 (b) ϕ_2

Figure 3.4. The error $|\rho - \hat{\rho}|$ of the Times 3 map at different threshold values by using the excess over threshold method, where ρ is the theoretical shape parameter and $\hat{\rho}$ is the estimated shape parameter (a) Observable ϕ_1 (b) Observable ϕ_2 with $\alpha = 2$.

3.4.2. Hénon map

The Hénon map is defined by:

$$f \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 - ax^2 + y \\ bx \end{bmatrix}, \tag{3.11}$$

where, $a = 1.4, b = 0.3$. The Hénon map is first introduced by Hénon [50]. This is a two dimensional map with a strange attractor $\Lambda \subset \mathbb{R}^2$. A rigorous proof that a strange attractor exists is unknown for these parameters. However for b small, a strange attractor Λ is known to exist, see [4]. Moreover for those parameters, there exists an SRB measure ν. As for the Times 3 map we can try to estimate the shapes parameter for a range of block numbers under the observables ϕ_1, ϕ_2 and ϕ_3. We take an initial value $x_0 = (x_0, y_0)$ using a (uniform) random number generator, and similarly for the value $\tilde{x}$. However, for certain observable types we will set $\tilde{x} = (x_t, y_t)$, where t is the number of iteration steps taken, very large relative to the time series length under consideration. In fact we set $t = 10^8$.

- Case 1: Here we consider the distance function:

  $$\text{dist}_1((u, v), (x, y)) := |u - x|.$$

  Thus we study the evolution of the first coordinate under iteration. For $\tilde{x}$, it suffices to take this value in $\mathbb{R}$, and in fact we take it from a uniform distribution on $[0, 1]$. (See Figure 3.5)

- Case 2: Here we consider the distance function:

  $$\text{dist}_2((u, v), (x, y)) := \left((u - x)^2 + (v - y)^2\right)^{1/2}.$$

  In this case the observable function depends on both coordinates in $\mathbb{R}^2$. For t large we will set $\tilde{x} = (x_t, y_t)$. In fact we set $t \gg n$, where $n = 10^6$ is the time series length.

Remark. When the observable depends only on the first coordinate, it suffices to take any $\tilde{x} \in [0, 1]$, and we expect the results of Chapter 2 to apply for typical (probability one) choice. However, in Case 2, if we picked $\tilde{x} \in \mathbb{R}^2$ at random, then with Lebesgue-probability one we would have $\nu(B(\tilde{x}, r)) = 0$ for all r sufficiently. This is a consequence of the fact that Λ has zero area. Hence we should take $\tilde{x} = (x_t, y_t)$ for t large, so that we are in a good approximation of having $\tilde{x}$ in (or very close to) Λ. The value t should be larger than the time series length under consideration.

From the results, we obtained a better convergence for the one dimensional case. For case 1, and under the observable ϕ_1 we expect to find $\rho = 0$. This is verified by our results. We set $\alpha = 2, C = 0$ for the observables ϕ_2 and ϕ_3. We expect that $\rho = 0.5$ and

-0.5 respectively and again good convergence is observed for these values see, Figure 3.6. Interestingly, in case 1 we observed a good convergence with $p = q = \sqrt{n}$. We should remark that since we only study the evolution of the first coordinate, then we expect our results to be consistent with those studies in one-dimension. Hence the local dimension constant D will not feature. In fact, if the conjectured SRB measure exists, then the conditional measures of ν on unstable (horizontal) manifolds will be absolutely continuous with respect to Lebesgue measure. For case 2, we set $\alpha = 4, C = 0$ for the observables ϕ_2 and ϕ_3. See, Figure 3.7. We do not record the theoretical values, as these are unknown and depend further on estimates of the local dimension D.
Since the map has short memory we may ignore the gap t and in Case 1. The conditional measure on unstable manifold is absolutely continuous with respect to Lebesgue measure, and so in Caes 1 we expect the convergence rate to be in agreement with Theorem 2.10. However in case 2 we might expect bad convergence due to the fact that the SRB is not absolutely continuous, and also due to the fact that we do not know if $\tilde{x} = (x_t, y_t)$ is on the attractor Λ. Recall that the convergence of $n(1-F(u_n)) \to \tau$ corresponds to studying the convergence of $n\nu(x : \phi(x) > u_n)$. From the discussion of Chapter 2 this might not converge to τ as $n \to \infty$ due to fluctuations in $\nu(B(\tilde{x}, r))$ as $r \to 0$.

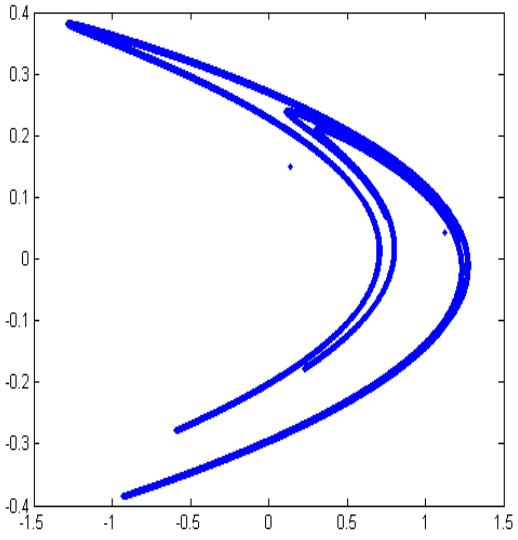
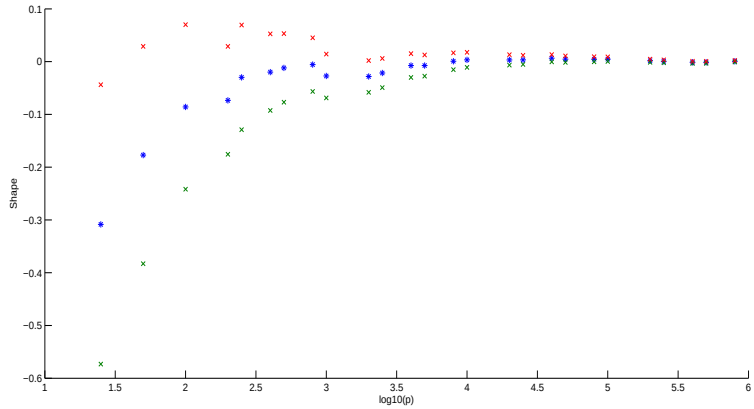
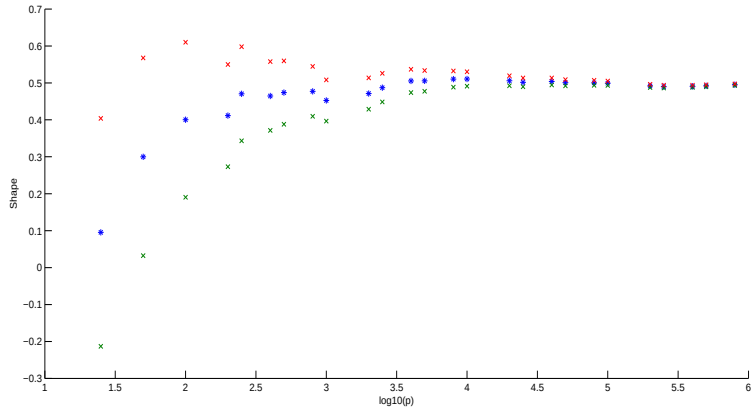


Figure 3.5. Attractor of the Hénon map

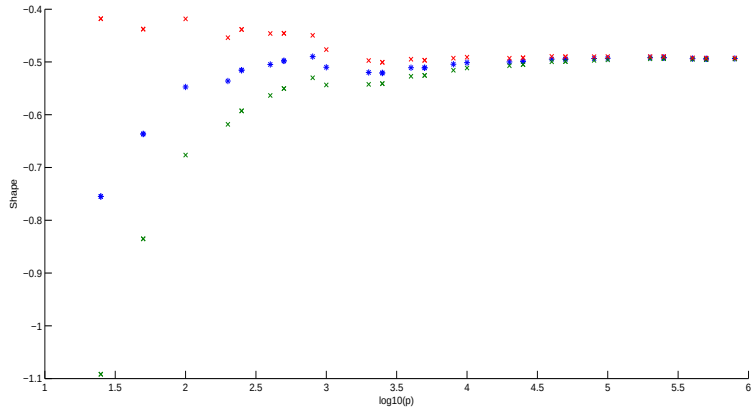
We now apply the excess over threshold method under the observables ϕ_1, ϕ_2 as in case 1 (here, $\tilde{x}$ is a typical random number) with $\alpha = 2$ and this for $n = 5000$. Similar to the Times 3 map we choose the threshold value u based on the plot of the mean excess which is given in Figure 3.8. The results are illustrated in the following Table 3.3.



(a) ϕ_1

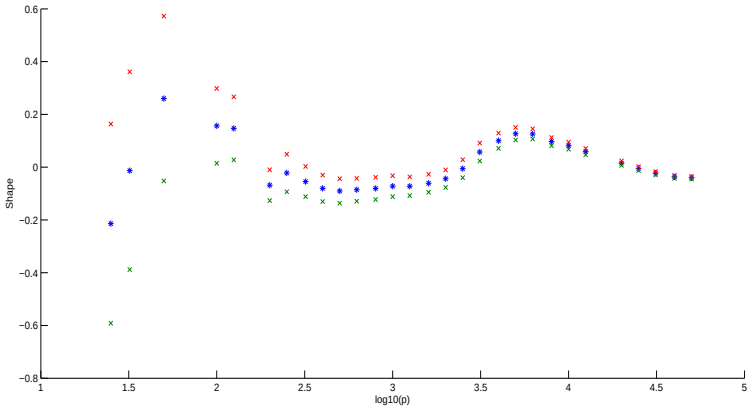


(b) $\phi_2, \alpha = 2$

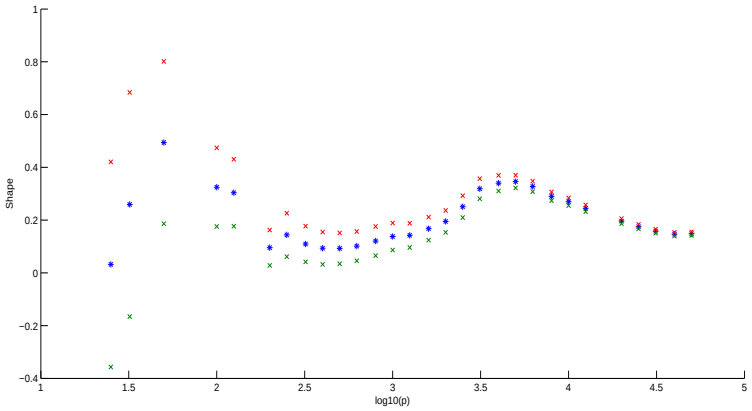


(c) $\phi_3, \alpha = 2, C = 0$

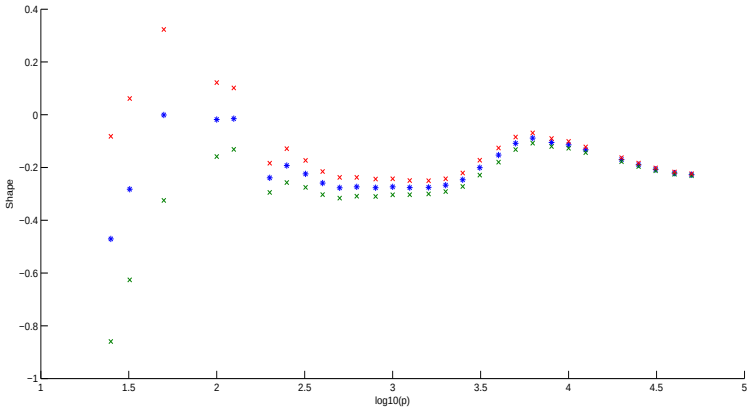
Figure 3.6. (Case 1) The estimated shape parameter(* line) with 95% confidence interval(x line) versus $\log_{10}(p)$ for the Hénon map, $n = 10^8$ where $\tilde{x}$ is a random number generated from a uniform distribution on $(0, 1)$. (a) ϕ_1 observable (b) ϕ_2 observable (c) ϕ_3 and $\alpha = 2$,$C = 0$.



(a) ϕ_1

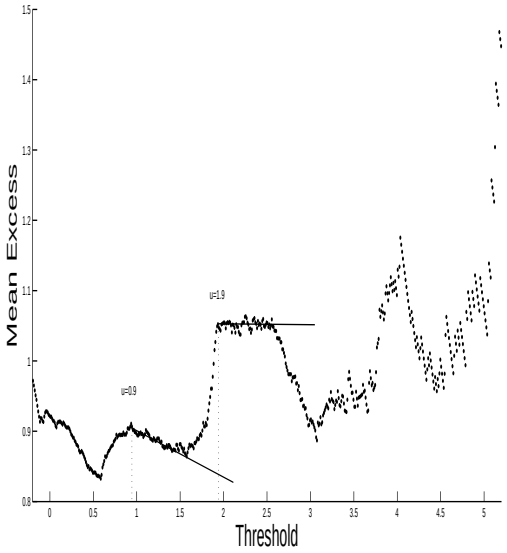


(b) $\phi_2, \alpha = 4$

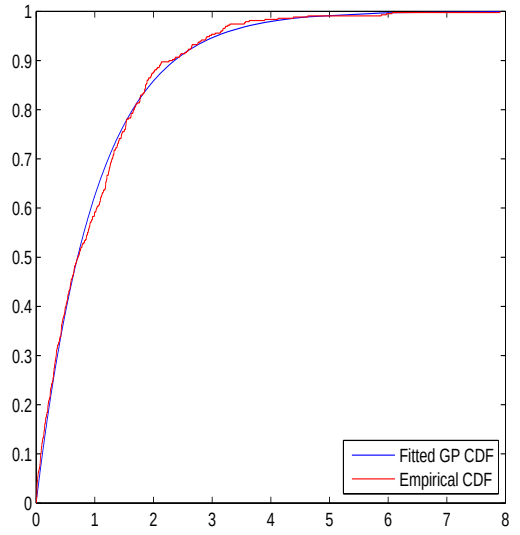


(c) $\phi_3, \alpha = 4, C = 0$

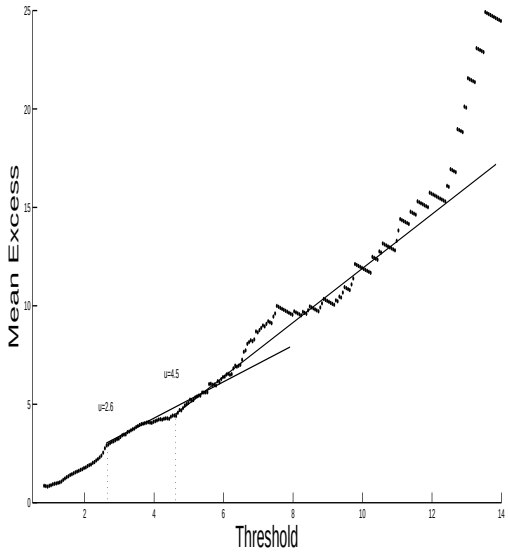
Figure 3.7. (Case 2) The estimated shape parameter(* line) with 95% confidence interval(x line) versus $\log_{10}(p)$ for the Hénon map, $n = 10^6, t = 10^8$ where $\tilde{x} = (x_t, y_t)$. (a) ϕ_1 observable (b) ϕ_2 observable (c) ϕ_3 and $\alpha = 4$,$C = 0$.



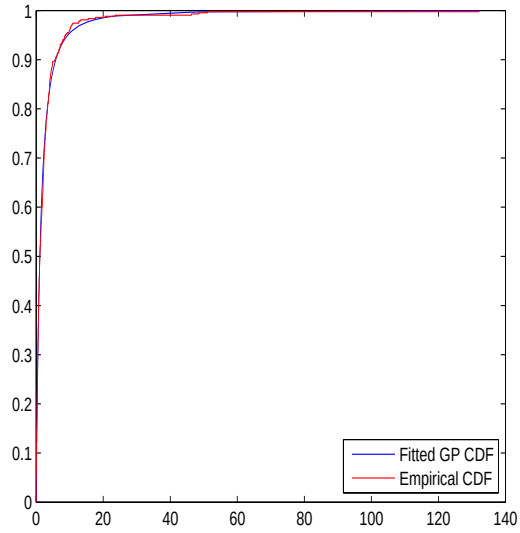
(a) ϕ_1



(b) ϕ_1



(c) ϕ_2



(d) ϕ_2

Figure 3.8. Hénon map, observable ϕ_1:(a) The mean excess plot (b) Empirical distribution of excess over threshold $u = 1.9$ and fitted GP distribution. Times 3 map, observable ϕ_2 with $\alpha = 2$:(c) The mean excess plot (d) Empirical distribution of excess over threshold $u = 2.6$ and fitted GP distribution.

Observable	Threshold	Number of exceedances	Quantile	$\hat{\rho}$
ϕ_1	0.9	1486	0.70	0.0303
	1.9	429	0.90	0.0061
ϕ_2, $\alpha = 2$	2.6	424	0.92	0.5151
	4.5	162	0.97	0.5838

Table 3.3. The estimated shape parameter using the excess over threshold method for the Hénon map under the observables ϕ_1, ϕ_2 and for $n = 5000$. The threshold values are chosen based on the mean excess plot.

3.4.3. Intermittent map.

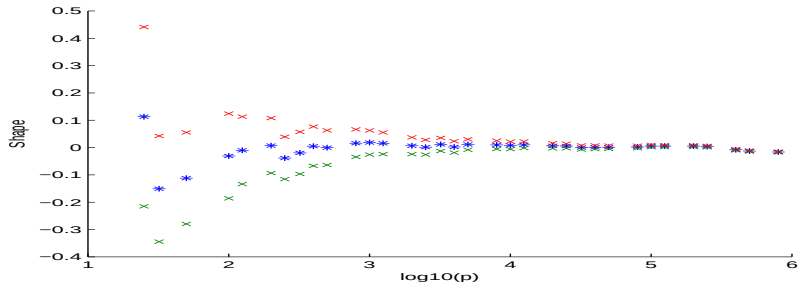
For $f : [0, 1] \to [0, 1]$, the map is defined by:

$$f(x) = \begin{cases} x(1 + 2^{\beta} x^{\beta}) & x \leq \frac{1}{2}, \\ 2x - 1 & x > \frac{1}{2}. \end{cases} \tag{3.12}$$

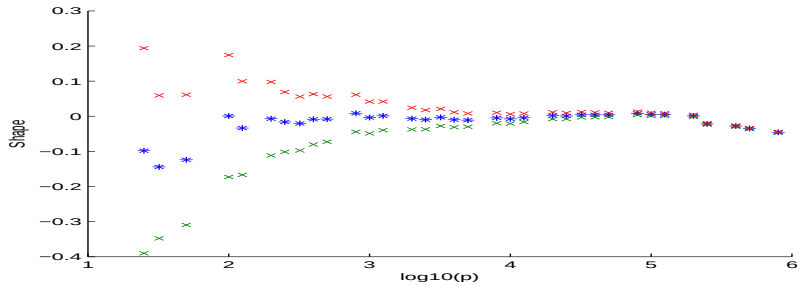
Where, $\beta \in (0, 1)$, see [62].

Similar to the previous maps we estimated the shape parameter versus range of block number p under the three observables described before. For the observable ϕ_2 we set $\alpha = 2$ and for the observable ϕ_3 we set $\alpha = 3$ with $\tilde{x} = 0.8491$. We repeat this for $\beta = 0.1, 0.3, 0.5$ and 0.7. We do this for $n = 10^8$ iterations starting from initial value ($x_0 = 0.1712$). The results under the observable ϕ_1, ϕ_2 and ϕ_3 are illustrated in Figure 3.9, 3.10 and 3.11.

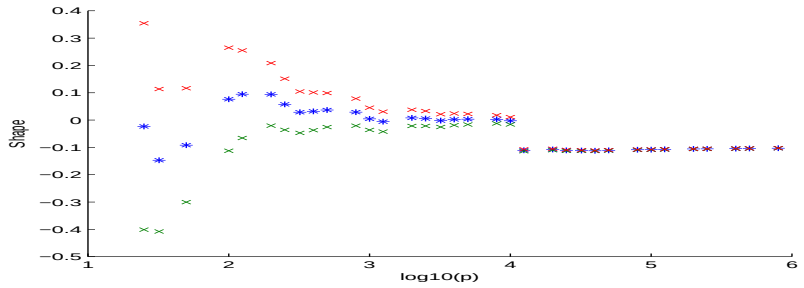
We found a good convergence to GEV distribution types for $\beta \in (0.0.3]$. On the other hand, for $\beta \in [0.3, 1)$ we observed a poor convergence to GEV distributions types. Since the map has absolutely continuous measure for $\beta \in (0, 1)$ see, [62] and [52]. Furthermore, as we mentioned in Chapter 2 for $\beta \in (0, 0.5)$ this map has short range dependence and for $\beta \in (0.5, 1)$ it has long range dependence see, [7] and [8]. Therefore, the cause of bad convergence for $\beta \in [0.3, 1)$ might be related to the dependency structure of this map.



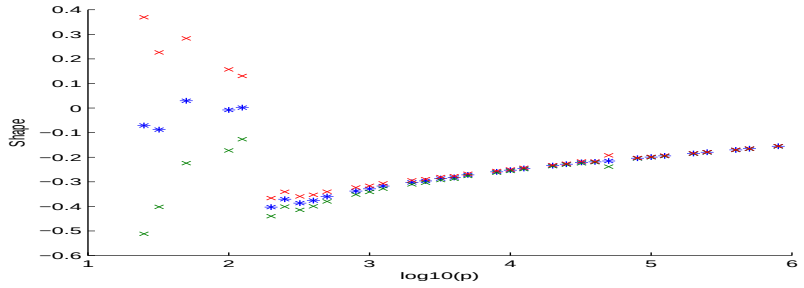
(a) $\phi_1, \beta = 0.1$



(b) $\phi_1, \beta = 0.3$

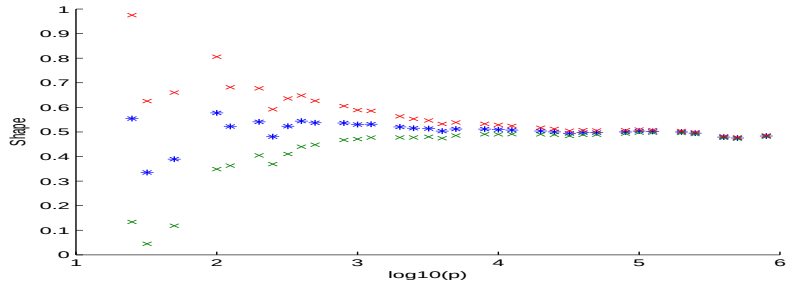


(c) $\phi_1, \beta = 0.5$

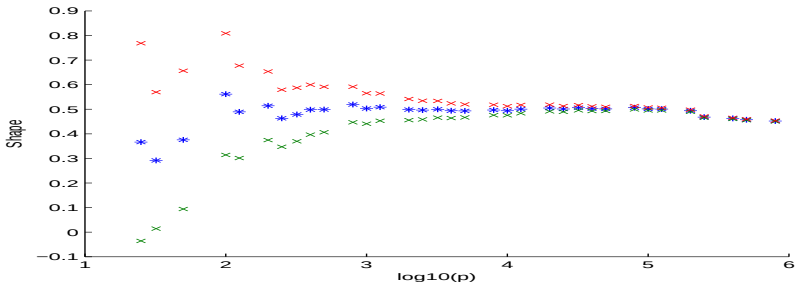


(d) $\phi_1, \beta = 0.7$

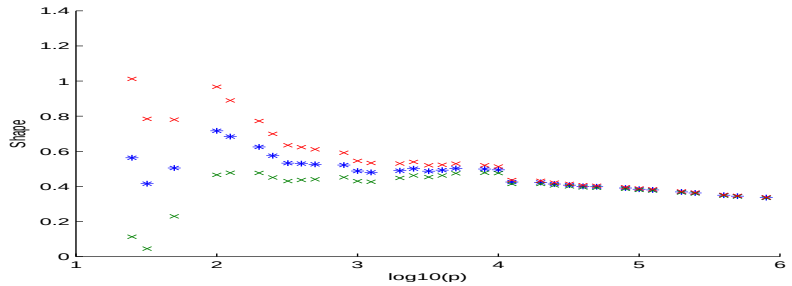
Figure 3.9. The estimated shape parameter(* line) with 95% confidence interval(x line) versus $\log_{10}(p)$ for Intermittent map,ϕ_1 observable with $n = 10^8$ (a) $\beta = 0.1$ (b) $\beta = 0.3$ (c) $\beta = 0.5$ (d) $\beta = 0.7$.



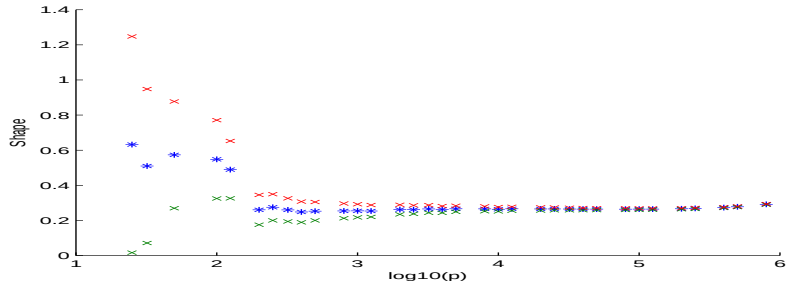
(a) $\phi_2, \beta = 0.1$



(b) $\phi_2, \beta = 0.3$

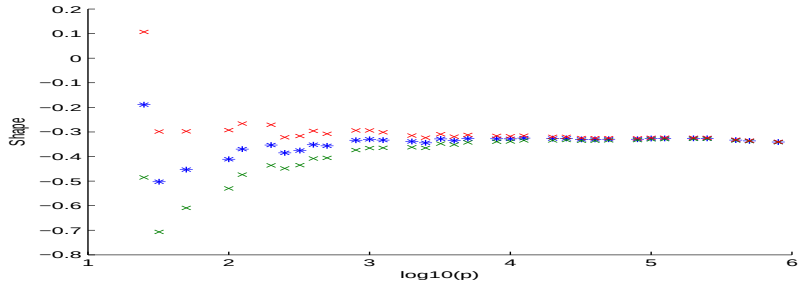


(c) $\phi_2, \beta = 0.5$

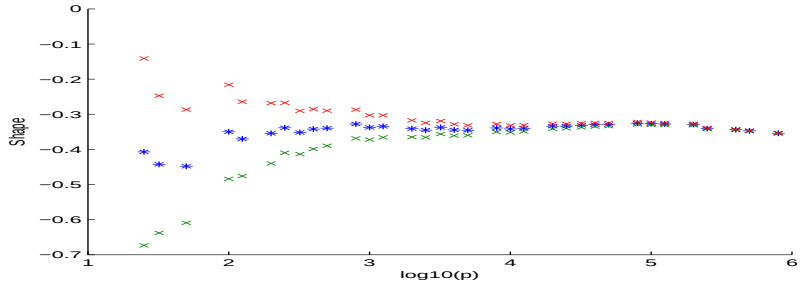


(d) $\phi_2, \beta = 0.7$

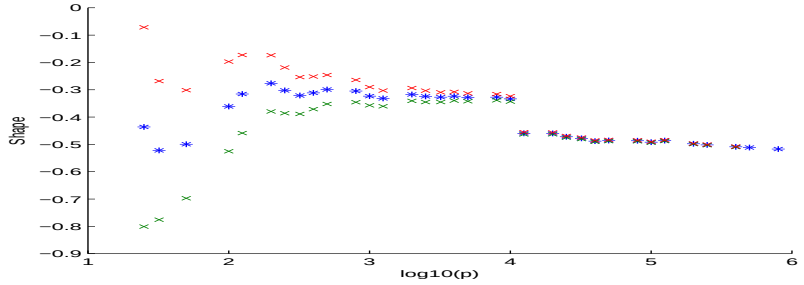
Figure 3.10. The estimated shape parameter(* line) with 95% confidence interval(x line) versus $\log_{10}(p)$ for Intermittent map,ϕ_2 observable with $n = 10^8, \alpha = 2$ (a) $\beta = 0.1$ (b) $\beta = 0.3$ (c) $\beta = 0.5$ (d) $\beta = 0.7$.



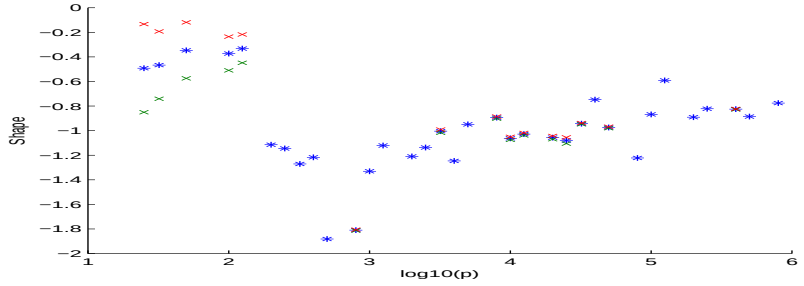
(a) $\phi_3, \beta = 0.1$



(b) $\phi_3, \beta = 0.3$



(c) $\phi_3, \beta = 0.5$



(d) $\phi_3, \beta = 0.7$

Figure 3.11. The estimated shape parameter(* line) with 95% confidence interval(x line) versus $\log_{10}(p)$ for Intermittent map,ϕ_3 observable with $n = 10^8, \alpha = 3$ (a) $\beta = 0.1$ (b) $\beta = 0.3$ (c) $\beta = 0.5$ (d) $\beta = 0.7$.

3.5. Analysis of Financial Returns

3.5.1. Data description

In our study we consider a time series of daily logarithmic returns R_n of the Standard and Poor 500 (S&P500) index and Financial Times Stock Exchange 100 (FTSE100) index. The S&P 500 index is a capital weighted index of the 500 companies representing all the major industries in the United States of America. Each stock selected in the index is proportionate to its market value. It is also widely used as a bench mark for the U.S stock market. The FTSE 100 represents the performance of the companies in the UK; it comprises the largest 100 companies listed on the London Stock exchange. It is also a capital weighted average used as a benchmark for the performance of the UK stock markets. As we introduced in the introduction one model that can describe the stock market behaviour is the random walk with independent increments or returns given by the following equation:

$$X_n = X_{n-1} + \xi_n \tag{3.13}$$

where, for example $X_0 = 0$ and the jump sizes ξ_n are IID RVs. The daily stock prices S_n can be described by the geometric random walk as mentioned in Chapter 1. In particular a discrete model is relevant, and modelling daily fluctuations can be changed to modelling fluctuations over any time period, for example monthly or yearly. Then, the daily logarithmic returns is given by $R_n = \log \frac{S_n}{S_{n-1}}$. In this section we will analyse the limit distribution of maxima in the financial return data. Here we consider $M_n = \max\{R_1, R_2, ..., R_n\}$. If the stock prices follow the RW model we expect the return to behave as those of IID observations. Even, if the return is dependent for example, modelled by chaotic map it may behave as IID observations as well.

3.5.2. Extremes of the returns

As in the previous examples, we estimated the shape parameter versus range of block number. For S&P500 return index we fit the GEV distribution by using ML estimator for the period January 1950 to September 2014 that consists of $n = 16000$ observations (see, Figure 3.12 .a). The results of the shape values are in the range (0,0.5) for most block numbers and this suggests that the fitted distribution is of Type II distribution. However, some block numbers suggest $\rho = 0$ and hence a Gumbel fit can not be ruled out. Further statistical analysis can be performed (e.g. via suitable hypothesis test) to decide the precise fit. Also we consider the case when $p = q \approx \sqrt{n}$, here we choose $p = q = 126$. We found that the estimated shape parameter is $\hat{\rho} = 0.1995$ with (0.0614, 0.3376) confidence interval.

Furthermore, we applied the excess over threshold method see, Figure 3.13 .b. We fit the GP distribution and we choose the threshold values based on the mean excess plot. The results are illustrated in Table 3.4. We found that the estimated shape parameter agrees with the shape parameter obtained by the block maxima method with $p = q = 126$.

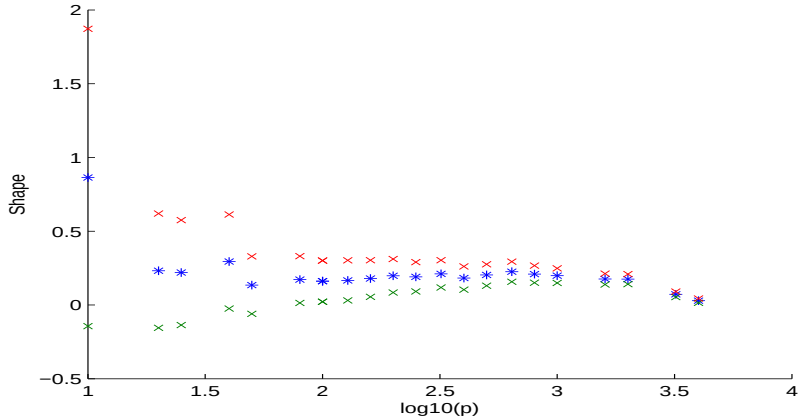
Similarly, for the FTSE100 index we estimated the shape parameter versus range of block

number for the period from January 1984 to September 2014 that consist of $n = 8000$ observations (see, Figure 3.12 .b). Similar to the S& P500 index we found that the shape values are in the range (0,0.5). Also, when $p = q \approx \sqrt{n}$, here we choose $p = q = 89$. We found that the estimated shape parameter is $\hat{\rho} = 0.3420$ with (0.1313, 0.5527) confidence interval.
Furthermore, we applied the excess over threshold method see, Figure 3.13 .d. As illustrated in Table 3.4 we found that the estimated shape parameters are on the confidence interval range of the shape parameter obtained by the block maxima method with $p = q = 89$.
An application of extreme value theory in the financial returns data can be found for example, in [17], [25] and [43]. A common feature found in all was that the tail index is positive which indicates that the tail index is of type II.

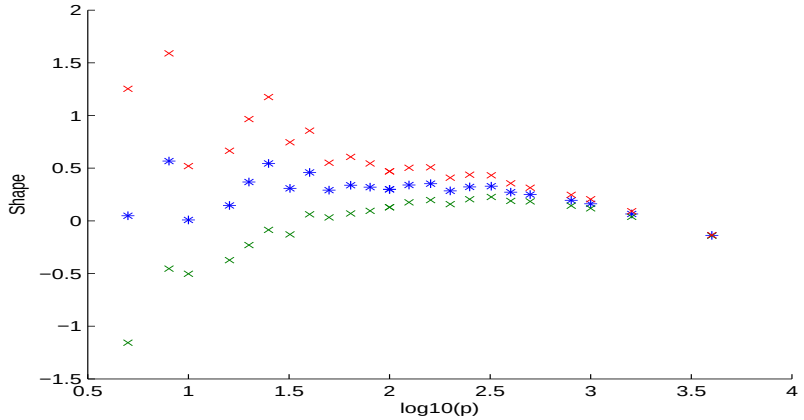
In finance Extreme Value Theory is used for measuring risk and methods include Value at Risk(VaR) and Expected Short fall(ES) analysis. These are functions of the EVD parameters. Risk management is a very important task either in investment or in insurance. Risk management modellers face a challenging task when trying to account for rare or extreme events. Therefore, studying extreme value theory can help to improve the performance of risk managers, see [25]. Furthermore, understanding the behaviour of extreme value distribution in dynamical system examples might help in modelling financial returns data. For example under which observable might we obtain the same type of distribution and convergence as those of dynamical systems examples?

Index	Threshold	Number of exceedances	Quantile	$\hat{\rho}$
S&P500	0.0054	3797	0.75	0.1363
	0.015	735	0.95	0.2012
FTSE100	0.008	1499	0.85	0.1846
	0.011	893	0.90	0.2203

Table 3.4. The estimated shape parameter using the excess over threshold method for the S&P500 and FTSE100 log returns indices. The threshold values are chosen based on the mean excess plot.

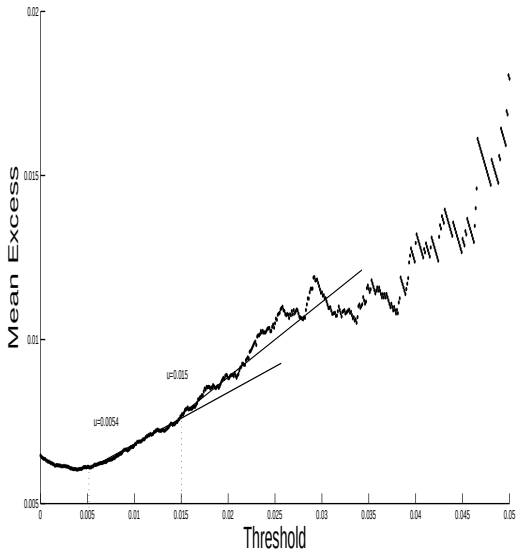


(a) S&P500 index

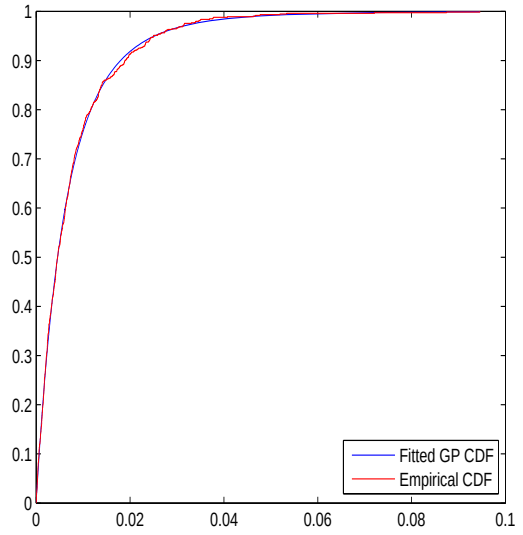


(b) FTSE100 index

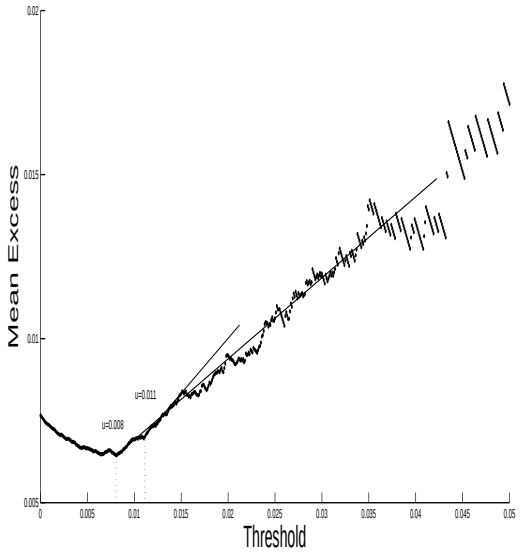
Figure 3.12. The estimated shape parameter(* line) with 95% confidence interval(x line) versus $\log_{10}(p)$ (a)S&P500 return index for the period from January 1950 to September 2014.($n = 16000$) (b) FTSE100 return index for the period from January 1984 to September 2014. ($n = 8000$).



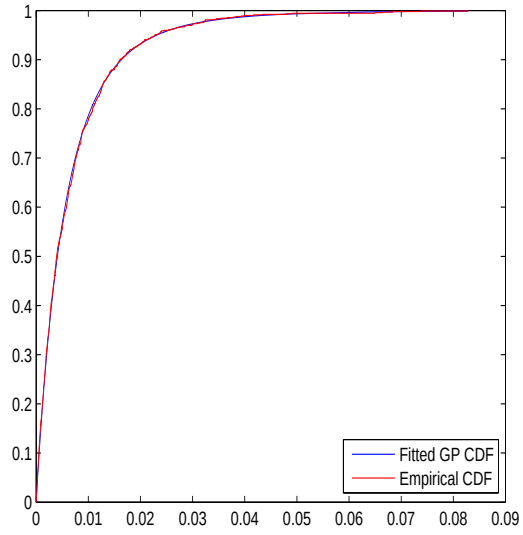
(a) S&P500



(b) S&p500



(c) FTSE100



(d) FTSE100

Figure 3.13. S&P500 index log return data (a)The mean excess function (b) Empirical distribution of excess over threshold $u = 0.015$ and fitted GP distribution. FTSE100 index log return data : (c) The mean excess function (d) Empirical distribution of excess over threshold $u = 0.011$ and fitted GP distribution.

3.6. Summary and Remarks

The limit distribution of maxima in IID observations converge in distribution to the GEV Distribution which can be one of only three types Gumbel, Fréchet and Weibull distributions. Stationary sequences satisfy some mixing conditions (i.e D and D' conditions) the distribution of maxima is also converged to the GEV Distribution as in the IID case. Similarly, in dynamical systems and under some mixing conditions around non periodic points the distribution of maxima converge to the GEV distribution. In modelling maxima there are two known approaches that are the block maxima and the excess over threshold methods. The problem with the block maxima method is the optimal block size and number. In the dynamical system Holland and Nicol [53] proved that under certain conditions the optimal block number and size for n observations is $p \approx q \approx \sqrt{n}$.

In this chapter we studied an optimization of the block maxima method using the same algorithm that is used in Faranda et al. [32] and Lucarini et al. [66] to a range of dynamical system examples. We considered the three dynamical system examples. The times 3 map, intermittent map with different values of β and the Hénon map. We found a good convergence obtained for the Times 3 map, the intermittent map with $\beta \in (0, 0.3)$ on the other hand a poor convergence with $\beta \in (0.3, 1]$. In the Hénon map we found a better convergence observed under the observable functions in one-dimension (the first (x) coordinate only). Furthermore, the choice of $p \approx q \approx \sqrt{n}$ is optimal in the Times 3 map, Intermittent map with $\beta \in (0, 03)$ and Hénon map in one dimensional.
Furthermore, we applied the excess over threshold method to the Times3 map and to the Hénon map in the one dimensional case. The results are in agreement with the result obtained by the block maxima method. Then, we applied these methods to the Financial return data of the S&P500 and FTSE100 indices. We found that the limit distribution of maxima is of type II. In comparison with the dynamical system examples we could generate the same distribution of the maxima using the observable ϕ_2.

Finally, we may say that in order to obtain a good convergence to a GEV distribution in a dynamical system we require

- the systems satisfy suitable mixing conditions and/or conditions on the rate of decay of correlations, analogous to $D(u_n)$;
- the systems satisfy conditions on their recurrence over short time scales, e.g. similar to $D'(u_n)$;
- the system has an absolutely continuous invariant measure;
- the choice of the observable function to be sufficiently smooth (e.g. regularly varying);
- the system has non periodic behaviour.

We will see in the next chapter for stationary sequence that D' failed the distribution of maxima converge to GEV distribution but with respect to the parameter θ which is known as the extremal index or clustering parameter. Similarly, in the dynamical system it is observed that under some mixing conditions and around repelling periodic point the distribution of maxima converges to the GEV distribution with respect to θ.

4. Clustering of Extremes

4.1. Introduction

We have discussed in the previous chapter that the convergence of maxima in stationary sequence and dynamical system under some suitable conditions have similar behaviour to that of IID sequence. For a stationary sequence that does not satisfy the D' condition the limit distribution of maxima converge to the GEV distributions with respect to a parameter θ is known as the extremal index. The dependency or clustering of extreme in stationary sequence can be understood by estimating this extremal index value $\theta \in [0, 1]$. The extremal index is a measure of clustering or dependency in time series observations. The smaller the value of θ the more clustering we have. Leadbetter et al. [61] showed that θ^{-1} is the mean cluster size, the number of exceedance high threshold in a cluster. For example, if $\theta = 0.5$ then the mean cluster size is 2. Recently, in a dynamical systems similar behaviour is observed for observables maximized at a periodic point.

In this chapter, we give a short review and discussion of the extremal index in stationary sequences and dynamical systems, in Section 4.2 and 4.3 respectively. We apply the method of Ferro and Segers [34], Ferro [33] to estimate the extremal index for a range of threshold values and we compare the results with the theoretical extremal index given recently by Freitas et al. [41]. The method is presented in Section 4.4. We estimated the extremal index of the dynamical system examples defined in the previous Chapter, in Section 4.5. Furthermore, we estimated the extremal index of the S&P500 and FTSE100 return indices, in Section 4.6.

4.2. Background on the Extremal Index

In this section we will give a brief review of the extremal index. First we will introduce some definitions about the point processes and how this is related to Extreme value theory.

It is known that the point process of exceedances threshold converge to poisson process in IID and stationary sequences satisfy D and D' conditions see, for example [61] and [26]. The point process can be defined using the Dirac delta measure δ_x, $x \in E$ as: $\delta_x = 1$ if $x \in B$ and $\delta_x = 0$ if $x \notin B$, where E is a state space and $B \subset E$. For a sequence

$(x_i)_{i\geq 1}$ in E,

$$m(B) = \sum_{i=1}^{\infty} \delta_{x_i}(B) = \sum_{i:x_i \in B} 1 = card\{i : x_i \in B\} = N(B).$$

Simply, $N(B)$ counts the number of $x_i \in B$.

Definition 4.1 *[26] Let N be a point process that is satisfied the following conditions:*

- *For $B \in \Sigma$ and $k \geq 0$,*

$$P(N(B) = k) = \begin{cases} e^{-\mu(B)} \frac{(\mu(B))^k}{k!} & \text{if } \mu(B) < \infty, \\ 0 & \text{if } \mu(B) = \infty, \end{cases}$$

- *For any $m \geq 1$, if $B_1, \ldots, B_m$ are mutually disjoint sets in Σ then $N(B_1), \ldots, N(B_m)$ are independent RVs.*

Then N is a Poisson process or a Poisson random measure with mean measure μ. Here Σ is a σ-algebra of the Borel sets generated by the open sets.

Definition 4.2 *[26](poisson process) A point process $(N(t))_{t>0}$ is a Poisson process with intensity $\lambda > 0$ if :*

- *N(0)=0.*
- *It has independent increments.*
- *N(t) is poisson distributed with mean λt*

$$P(N(t) = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}, k = 0, 1, \ldots.$$

Let (ξ_n) be a sequence of RVs and u a threshold value then the point process of exceedances this threshold value is given by

$$N_n(.) = \sum_{i=1}^{n} \delta_{n^{-1}i}(.) I(\xi_i > u), \quad n = 1, 2, \ldots,$$

with state space $E = (0, 1]$ counts the number of observation that exceed a level or threshold value. Further details on this topic can be found in [61] and [26].

Extremal index:

Let (ξ_n) be stationary sequence and M_n be the corresponding maximum values. Let, $(\hat{\xi_n})$ be the associated IID sequence, and $(\hat{M_n})$ the corresponding maximum values. Then there

exists $\theta \in (0,1]$ such that,

$$\lim_{n\to\infty} P((\hat{M}_n - a_n)/b_n \leq x) = G(x), \tag{4.1}$$

if and only if

$$\lim_{n\to\infty} P((M_n - a_n)/b_n \leq x) = G^{\theta}(x). \tag{4.2}$$

θ is called the extremal index of the sequence (ξ_n). See [61].

Definition 4.3 *[61] Let (ξ_n) be a strictly stationary sequence. Then we say that the sequence (ξ_n) has extremal index $\theta \in [0,1]$ if for every $\tau > 0$ there exist a sequence (u_n) such that,*

$$(a) \lim_{n\to\infty} n(1 - F(u_n)) \to \tau \tag{4.3}$$

$$(b) \lim_{n\to\infty} P(M_n \leq u_n) \to e^{-\theta\tau} \tag{4.4}$$

From theorems (2.2) and (2.5) we can see that IID sequence and stationary sequence satisfied $D(u_n)$ and $D'(u_n)$ conditions $\theta = 1$. Leadbetter et al [61, p. 65] showed that if $D'(u_n)$ does not hold then $\lim_{n\to\infty} P(M_n \leq u_n) \to e^{-\theta\tau}$. Similarly, this idea is extended to the dynamical systems under some conditions around the periodic point, as we will see in the next Section.

Example 4.1 (Max-Auto-Regressive) Suppose $(U_n)_{n\geq 1}$ is a sequence of independent unit-Fréchet distribution RVs. For $0 < \theta \leq 1$ let

$$\begin{aligned} \xi_1 &= U_1/\theta, \\ \xi_n &= \max\{(1-\theta)\xi_{n-1}, U_n\}. \end{aligned} \tag{4.5}$$

Where, $n \geq 2$. Then (ξ_n) is a Max-Auto-Regressive process with extremal index θ. See [34].

4.3. Extremal Index in Dynamical Systems

Recently, in dynamical systems Freitas et al. [41] stated a set of conditions around a periodic point in order of extreme to converge to extreme value distribution with the extremal index θ. Their approach is checking the periodicity and the mixing of the systems. The periodicity means for example, the weather is periodic with period 365. The point x is a periodic point of period p if $f^p(x) = x$. See [22].

Definition 4.4 *Condition $SP_{p,\theta}(u_n)$ holds for $\xi_0, \xi_1, \xi_2, ...$ if*

$$\lim_{n\to\infty} \sup_{1\leq j<p} P(\xi_j > u_n | \xi_0 > u_n) = 0, \lim_{n\to\infty} P(\xi_p > u_n | \xi_0 > u_n) \to (1-\theta) \tag{4.6}$$

and,

$$\lim_{n\to\infty} \sum_{i=0}^{[\frac{n-1}{p}]} P(\xi_0 > u_n, \xi_p > u_n, \xi_{2p}, ..., \xi_{ip} > u_n) = 0, \tag{4.7}$$

where, $p \in \mathbb{N}$ *and* $\theta \in [0,1]$.

This condition is called the summable periodicity of period p. If this condition holds for $i, s, l \in \mathbb{N} \cup \{0\}$, then the following events are defined: $Q_{p,i}(u) = \{\xi_i > u, \xi_{i+p} \leq u\}$, $Q^*_{p,i}(u) = \{\xi_i > u\}$ $Q_{p,i}(u)$ and $\wp_{p,s,l}(u) = \bigcap_{i=s}^{s+l-1} Q^c_{p,i}(u)$

Definition 4.5 *($D^p(u_n)$condition) Assume that $SP_{p,\theta}(u_n)$ holds. The condition $D^p(u_n)$ will be said to hold for the sequence $\xi_0, \xi_1, ...$ if for any integers l, t and n*

$$|P(Q_{p,0}(u_n) \cap \wp_{p,0,l}(u_n)) - P(Q_{p,0}(u_n))P(\wp_{p,0,l}(u_n))| \leq \alpha(n,t), \tag{4.8}$$

where, $\alpha(n,t)$ is non increasing in t for each n and $n\alpha(n,t) \to 0$ as $n \to \infty$ for some sequence $t_n = o(n)$.

Definition 4.6 *($D'_p(u_n)$ condition) Assume that $D^p(u_n)$ holds. The condition $D'_p(u_n)$ will be said to hold for the sequence $\xi_0, \xi_1, ...$ if there exist a sequence $(k_n)_{n\in\mathbb{N}}$ such that $k_n \to \infty$ and $k_n t_n = o(n)$.*

$$\lim_{n\to\infty} n \sum_{j=1}^{[n/k_n]} P(Q_{p,0}(u_n) \cap (Q_{p,j}(u_n)) = 0. \tag{4.9}$$

$D'_p(u_n)$ condition is the same as $D'(u_n)$ condition which was defined before. Freitas et al. [41] found that under $D^p(u_n)$ and $D'_p(u_n)$ conditions we can extend the extreme value distribution with the extremal index in dynamical systems as stated in the following Theorems:

Theorem 4.1 *Let $\xi_0, \xi_1, ...$ be a stationary sequence satisfying $D^p(u_n)$ and $D'_p(u_n)$ for some $p \in \mathbb{N}$ and $\theta \in (0,1)$. Suppose that (u_n) such that $\lim_{n\to\infty} n(1 - F(u_n)) \to \tau$ for some $\tau \geq 0$.*

Then

$$\lim_{n\to\infty} P(M_n \leq u_n) = \lim_{n\to\infty} P(\wp_{p,0,n}(u_n)) = e^{-\theta\tau.} \tag{4.10}$$

In the application we estimated the extremal index around $\tilde{x}$ where $\tilde{x}$ is a repelling periodic point which means that $|(f^p)'(\tilde{x})| > 1$. We expect that our numerical results will agree with the following theorem.

Theorem 4.2 *Suppose that (u_n) such that $\lim_{n\to\infty} n(1 - F(u_n)) \to \tau$ for some $\tau \geq 0$. and $\tilde{x}$ is a repelling periodic point of period p, with*

$$\theta = \theta(\tilde{x}) = 1 - \left|\frac{1}{(f^p)'(\tilde{x})}\right|. \tag{4.11}$$

Assume that $D^p(u_n)$ and $D'_p(u_n)$ hold. Then

$$\lim_{n\to\infty} P(M_n \leq u_n) = \lim_{n\to\infty} P(\wp_{p,0,n}(u_n)) = e^{-\theta\tau}.$$

Remark. Recently, Freitas et al. [40] developed some conditions around a periodic point in order for the point process to converge to the poisson process with intensity θ. Indeed, the rate or speed of convergence of the extremal index under the block maxima method was studied in [39].

4.4. Methodology

4.4.1. Estimating the extremal index: Interval estimator

There are several approaches used to estimate the extremal index such as, the Blocks estimator that is a result of Hsing et al. [57]. The idea of this estimator is similar to the block maxima method described in Chapter 2. The block estimator divides the data into p blocks of size q where, $n = pq$. Then for each block the maximum is computed as $M_q^{(i)} = \max(\xi_{(i-1)q+1}, ..., \xi_{iq})$, then the extremal index can be estimated by the following equation

$$\hat{\theta}_n = \frac{\sum_{i=1}^{p} I_{(M_q^{(i)} > u)}}{\sum_{i=1}^{n} I_{(\xi_i > u)}} = \frac{P}{N}. \tag{4.12}$$

Where, P is the number of blocks with at least one exceedance of the threshold value u and N is the number of observations that exceed a threshold value u. With this method a better result can be obtained by taking maxima of the sliding block $p = n-q+1$ instead of taking the joint block $p = n/q$ see, for example [80]. Another approach is known as Run's estimator based on the O'Brien [75] result. The extremal index by the run's method can be estimated by the following equation

$$\hat{\theta}_n = \frac{\sum_{i=1}^{n-r} I_{A_{i,n}}}{N}. \tag{4.13}$$

Where, $A_{i,n} = (\xi_i > u, \xi_{i+1} \leq u_n, ..., \xi_{i+r} \leq u)$. and N as defined above. Further illustration on these methods can be found in [85] and [26]. The problem with these methods is that they require a choice of the run length and the block size parameter. Ferro and Segers [34] introduced the interval estimator that does not require a choice of such parameters. The interval estimator is based on the limit result of the number of exceedances threshold, and to the convergence of the inter-exceedance times (time between two successive exceedances) of threshold u by the sequence (ξ_n). The number of exceedances threshold is the point process of exceedances. Moreover, the inter exceedance time in poisson process has an exponential distribution and from this idea the interval estimator comes. In our study we will use the interval estimator described by Ferro and Segers [34] and Ferro [33] who stated the following theorem.

Theorem 4.3 *Let (ξ_n) be a stationary sequence and $(u_n)_{n\geq 1}, (r_n)_{n\geq 1}$ be a sequence of thresholds and positive integers respectively.*

- *(a) $r_n \to \infty$ and $r_n(1 - F(u_n)) \to \tau$.*
- *(b) $P(M_{r_n} \leq u_n) \to e^{(-\theta\tau)}$.*

  (a) and (b) for some $\tau \in (0, \infty)$ and $\theta \in [0, 1]$. For $1 \leq k \leq l$ and $u \in \mathbb{R}$. Let $\mathcal{F}_{k,l}(u)$ be the $\sigma-$ field generated by the events $(\xi_i > u), k \leq i \leq l$.

$$\alpha_{n,q}(u) = \max_{1\leq k\leq n-q} \sup |P(B|A) - P(B)|, \tag{4.14}$$

  where, the supremum is over all $A \in \mathcal{F}_{1,k}(u)$ with $P(A) > 0$ and all $B \in \mathcal{F}_{k+q,n}(u)$.

- *(c) For all $c > 0$ and there are positive integers $q_n = o(r_n)$ such that $\alpha_{cr_n,q_n}(u_n) = o(1)$, then*

$$P((1 - F(u_n))T(u_n) > t) \to \theta e^{(-\theta t)} \quad \text{for } t > 0.$$

  Where, $T(u_n)$ is distributed as $\min\{n \geq 1, \xi_{n+1} > u\}$, and $\xi_1 > u$

Consider your observations are represented as a sequence $(\xi_n)_{n\geq 1}$. Let $N = N_n(u) = \sum_{i=1}^{n} I(\xi_i > u)$ be the number of observations exceeding threshold u by the sequence (ξ_n). Let $1 \leq S_1 \leq ... \leq S_N \leq n$ be the observed exceedance times and, $T_i = S_{i+1} - S_i$ be the time between two successive exceedances. Where, $i = 1, ..., N - 1$ then the extremal index $\hat{\theta}$ can be estimated using the following equations:

$$\hat{\theta}_n(u) = \frac{2\big(\sum_{i=1}^{N-1} T_i\big)^2}{(N-1)\sum_{i=1}^{N-1} T_i^2} \tag{4.15}$$

$$\hat{\theta}^*_n(u) = \frac{2\big(\sum_{i=1}^{N-1} (T_i - 1)\big)^2}{(N-1)\sum_{i=1}^{N-1} (T_i - 1)(T_i - 2)} \tag{4.16}$$

The following equation is a combination of 4.15 and 4.16 to make the estimator value lie in $[0, 1]$.

$$\tilde{\theta}_n(u) = \begin{cases} 1 \wedge \hat{\theta}_n(u) & \text{if} \max\{T_i : 1 \leq i \leq N-1\} \leq 2 \\ 1 \wedge \hat{\theta}^*_n(u) & \text{if} \max\{T_i : 1 \leq i \leq N-1\} > 2 \end{cases} \tag{4.17}$$

$\tilde{\theta}_n(u)$ is defined as the interval estimator for the extremal index see, [34] and [33]. In the application we estimated the extremal index value by using this interval estimator for the financial return data and dynamical system example as the next section shows.

Threshold estimator:

On the estimation of the threshold value Smith and Weissman [85] stated that " There is no theoretical reason why the threshold for estimating θ has to be the same as the threshold for estimating the tail of F " This means that in IID data or if the data has no

clustering of extreme (extremal index $\theta = 1$) the methods used in the literature to find the appropriate threshold value to estimate the tail index can be applied to estimate the extremal index, such as the mean excess plot which was described in Chapter 2. As we mentioned in Section 3.3.1 a common choice of the threshold value can be obtained from the relation $u_n(\tau) = F^{\leftarrow}(1 - \tau/n)$, where $F^{\leftarrow}$ is the quantile function of the distribution function F see, equation (3.4). In fact, the optimal threshold value depends on the rate of convergence to θ, and according to Robert et al. [80] there are no general known theory about this rate of convergence. In [80] they estimated the extremal index by using the interval estimator based on the default value of $\tau = 1$. In our application we estimated the extremal index for a wide range of threshold values then looked at the part of u where the plot does not change. We consider the dynamical systems examples we described in the previous chapter, such as, the Expanding Times 3 map, the intermittent map, Hénon map and in the financial return data. For such examples we compare the result with the theoretical value.

4.5. Data Examples

In this section, the interval estimator is applied to the Times 3 map, Intermittent map, Hénon map and to the S&P500 index and FTSE100 index return data. The estimated result of the interval estimator is given at numbers of threshold u. The results is obtained using our MATLAB code, see Appendix B.

Blocks and the Interval estimators:

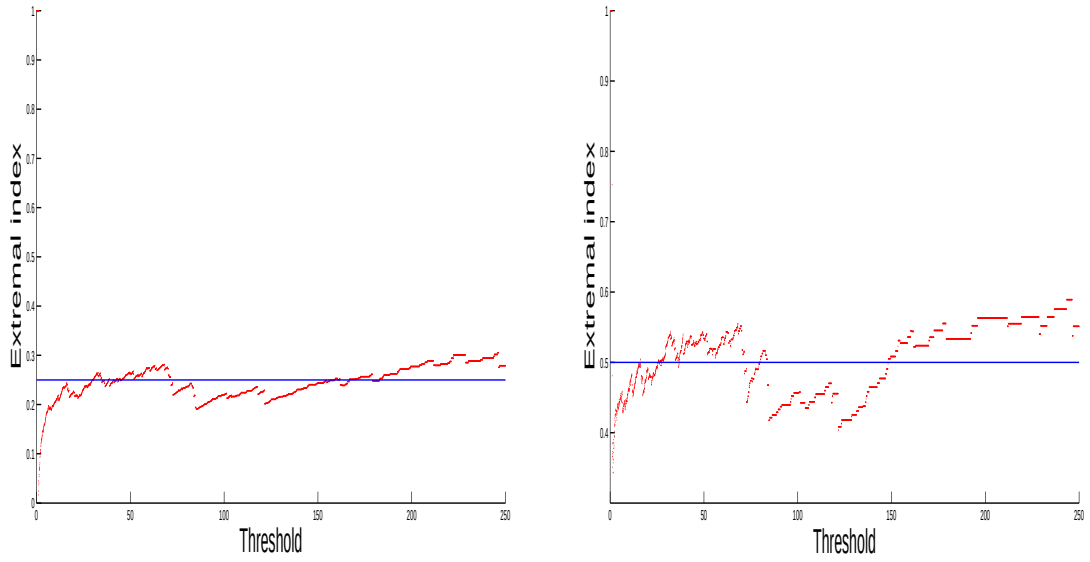
In this section we applied the Blocks estimator and the Interval estimator to the Times 3 map under the observable ϕ_1 with $\tilde{x} = 0$. The theoretical extremal index is ≈ 0.6667 as we will discuss in the next section. We start from an initial value and we iterate the map up to $n = 10^6$. In the Block estimator we choose $p = q = \sqrt{n}$. (as we discussed in the previous chapter this value gives an optimal block number and size in most application to fit the GEV distribution, hence we expected that this choice is an appropriate choice). We compare the result with the Interval estimator as illustrated in the following Table. Since, in both methods the choice of the threshold value is largely debated, we estimated the extremal index here with $u_n(\tau) = F^{\leftarrow}(1-\tau/n)$, where $F^{\leftarrow}$ is the quantile function and we set $\tau \approx 1$. We estimated the extremal index with $u = .99, .999$, and 0.9999 quantiles. The results in the block estimator are close to the theoretical value observed with $u = 0.9999$ quantile. Whereas, in the interval estimator a close result was observed with $u = .99$ and $u = .999$. In order to obtain good results, both methods require a large threshold but block estimator requires a larger threshold value.

Furthermore, we examine the extremal index of the Max Auto-regressive process that is described in Example (4.1) by using the interval estimator method. For this example we

compute the extremal index for a sequence of length $n = 5000$ with $\theta = 0.25$ and $\theta = 0.5$ at different threshold values. As illustrated in Figure 4.1, we found that for $\theta = 0.5$ the estimated extremal index in the range $(0.4, 0.6)$ and for $\theta = 0.25$ the estimated extremal index is in the range $(0.2, 0.3)$. Therefore, we may say that the result is consistent with the theoretical extremal index value.

Threshold	Interval estimator	Block estimator
0.99	0.6698	0.1
0.999	0.5854	0.4750
0.9999	0.7529	0.6800

Table 4.1. Estimation of θ of the Times 3 map with $\tilde{x} = 0$ for $n = 10^6$, using the interval estimator and the block estimator with $p = q = \sqrt{n}$.



(a) $\theta = 0.25$. (b) $\theta = 0.5$.

Figure 4.1. The estimated extremal index from Max-autoregressive at different threshold u for $n = 5000$. (a) $\theta = 0.25$. (b) $\theta = 0.5$.

4.5.1. Expanding Times 3 map.

Here we applied the interval estimator method of the extremal index to the Times 3 map $f(x) = 3x \mod 1$, under the observable $\phi(x) = -\log d(x, \tilde{x})$. Where, $\tilde{x}$ here is a repelling periodic point. In Freitas et al. [41] they showed that the true value of θ is given by equation (4.11) as $\theta = \theta(\tilde{x}) = 1 - |\frac{1}{(f^p)'(\tilde{x})}|$. See, Theorem(4.2). They defined the extremal index as

$$\theta = \lim_{n\to\infty} P(\xi_p \leq u_n | \xi_0 > u_n). \tag{4.18}$$

See Definition(4.4). This means that the probability that exceedance does not occur at time p given that exceedance has occurred at time 0 is θ. Alternatively, O'Brien [75]

described the extremal index θ under some suitable mixing conditions as

$$\theta = \lim_{n\to\infty} P(M_{1,s_n} \leq u_n | \xi_1 > u_n) \tag{4.19}$$

where, $M_{1,s_n} = \max(\xi_2, \xi_3, ..., \xi_{s_n})$ and (s_n) is a sequence of positive integers satisfying $s_n = o(n)$ and (u_n) is a sequence of real numbers such that $n(1 - F(u_n))$ is $O(1)$. This implies that if

$$P(M_{1,s_n} > u_n | \xi_1 > u_n) \to 0 \tag{4.20}$$

then, $\theta = 1$. Equation (4.19) of the extremal index can be written as

$$\theta = \lim_{s\to\infty} \lim_{u\to\infty} P(\xi_2 \leq u, \xi_3 \leq u, ..., \xi_s \leq u | \xi_1 > u) \tag{4.21}$$

(See, for example,[84] and [85]) The extremal index in Equation (4.18) depends only on two random variables ξ_p and ξ_0. On the other hand, in Equation (4.19) we should investigate the whole sequence $\xi_2, \xi_3, ..., \xi_s$.

In this example we estimate the extremal index when $\tilde{x} = 0$ and $\tilde{x} = 0.3$. Assuming that the observations satisfy $D^p(u_n)$ and $D'_p(u_n)$ conditions as defined before and $\tilde{x}$ is a repelling periodic point with period p. The corresponding true value of θ can be found from equation (4.11). For $\tilde{x} = 0$ we have $|(f)'(\tilde{x}) = 3| > 1$, then $\tilde{x}$ is a repelling periodic point with $p = 1$ and $\theta = 1 - |\frac{1}{(f^p)'(\tilde{x})}| = 0.6667$. Similarly, with $\tilde{x} = 0.3$ we have $\theta = 1 - |\frac{1}{(f^p)'(\tilde{x})}| = 1 - |\frac{1}{(f^4)'(0.3)}| = 0.9815$.

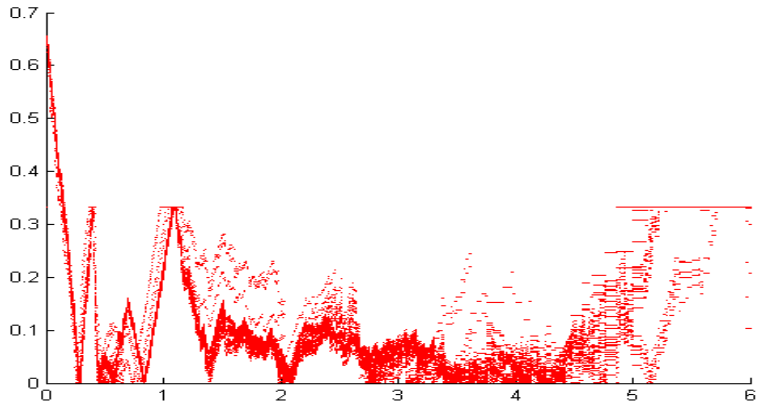
We estimate the extremal index for $n = 2 \times 10^4, 4 \times 10^4, 8 \times 10^4$, and 10^5 at thresholds $u \in [0, 7]$ when $\tilde{x} = 0$ and 0.3, as illustrated in Figures 4.3 and 4.4 respectively. Good convergence was observed for $u \in [3, 5]$ as n increased, however we cannot draw the same conclusion for other thresholds. Choosing the best threshold is a challenging statistical problem. As the result shows for $\tilde{x} = 0.3$ the extremal index is close to one. This verifies the theoretical extremal index value around this point. Moreover, this implies that extremes above threshold may have similar behaviour to that of IID observations. Whereas, when $\tilde{x} = 0$ extremes accept clustering we observed that $\theta < 1$. As we increase n the appropriate threshold value can be found by looking at the part of u where the plot does not change and this corresponds to $u \in [3, 5]$ for $\theta \approx 0.6667$.

In addition, with $\tilde{x} = 0.3$ and, $n = 10^4$ we found that $\theta \approx 1$, this corresponds to the 0.95 quantile $u = 3.7115$ and, to the 0.99 quantile $u = 5.3042$ see, Figure 4.5.a On the other hand, with $\tilde{x} = 0$ and, $n = 10^4$ we found $\theta \approx 0.7361$, this corresponds to the 0.95 quantile $u = 2.9525$ and, $\theta \approx 0.6387$ for the 0.99 quantile $u = 4.5793$ see, Figure 4.5.b.

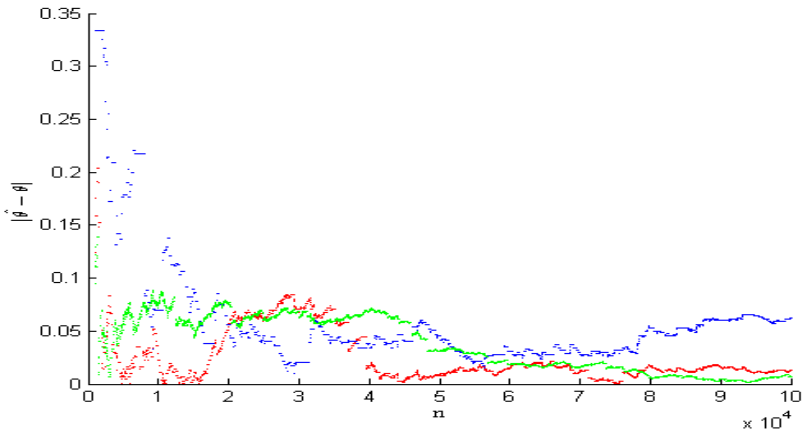
As we discussed before the appropriate threshold value can be estimated from the mean excess plot to fit the GP distribution. Here, we examine the behaviour of the mean excess plot of the Times 3 map with $\tilde{x} = 0.3$ and $\tilde{x} = 0$ for $n = 10^4$ as illustrated in Figures 4.5.c and 4.5.d respectively. For $\tilde{x} = 0.3$ we choose $u = 1.9$, this value corresponds to $\theta = 1$ and this agrees with the true value of the extremal index. On the other hand, for

$\tilde{x} = 0$ it is not obvious how to choose the appropriate threshold value based on the plot because it is linear from the beginning as $u \in [0, 3]$. However, we find that for $u = 3.5$ which corresponds to a large kink in the plot where the extremal index is $\theta \approx 0.6898$, this value is close to the theoretical extremal index value.

Furthermore, we plot $|\hat{\theta} - \theta|$ versus a range of thresholds u, for Times 3 map with $\tilde{x} = 0$. Where, the corresponding theoretical value is $\theta = 0.6667$ and $\hat{\theta}$ is the estimated extremal index value. We do this for a range of n values. We set $n = 1500, 1750, 2000, ..., 10000$. The results are shown in Figure 4.2.a. We found that $|\hat{\theta} - \theta| \approx 0$ for $u \in [3, 4]$. Moreover, we plot $|\hat{\theta} - \theta|$ with fixed thresholds $u = 3, 4$ and 5 versus a range of n values. We set $n = 1000, 1010, 1020..., 100000$, as illustrated in Figure 4.2.b. We found that $|\hat{\theta} - \theta| < 0.1$ with $u = 3, 4, 5$. We observed a better convergence with $u = 4$. For $u = 3$ we found that as n increases the $|\hat{\theta} - \theta| \to 0$.

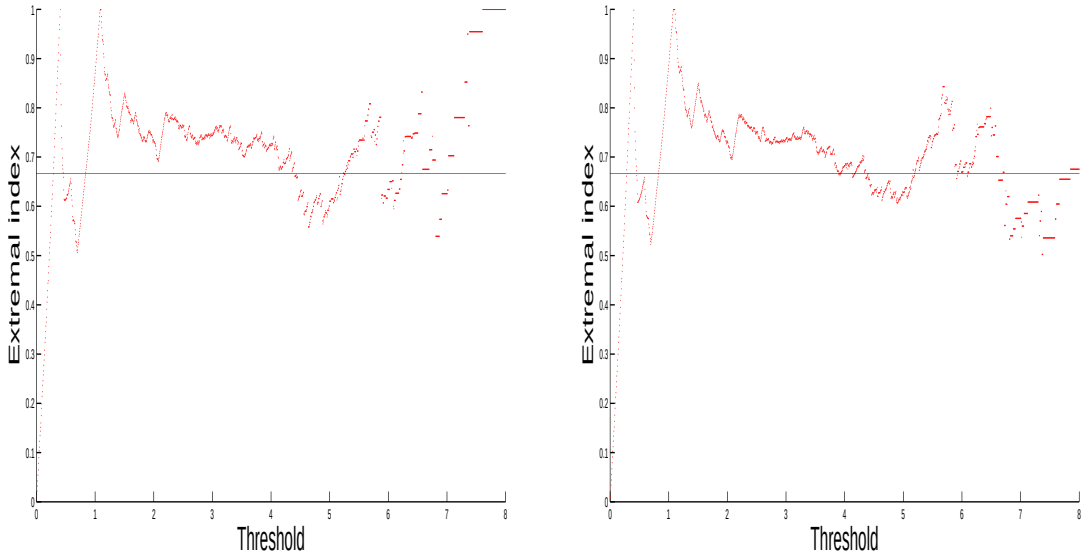


(a)



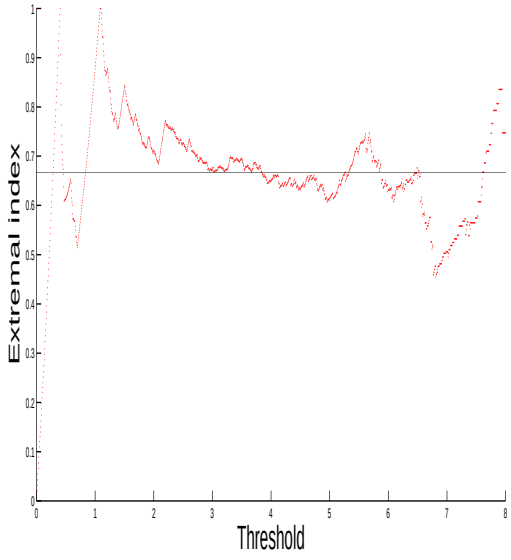
(b)

Figure 4.2. The Times 3 map with $\theta = 2/3$: (a) $|\hat{\theta} - \theta|$ plot for $n = 1500, 1750, 2000, 2250, ..., 10000$ at different threshold u. (b) $|\hat{\theta} - \theta|$ plot versus $n = 1000, 1010, 1020, 1030, ..., 100000$ where the green line corresponds to $u = 3$, the red line correspond to $u = 4$ and the blue line corresponds to $u = 5$.

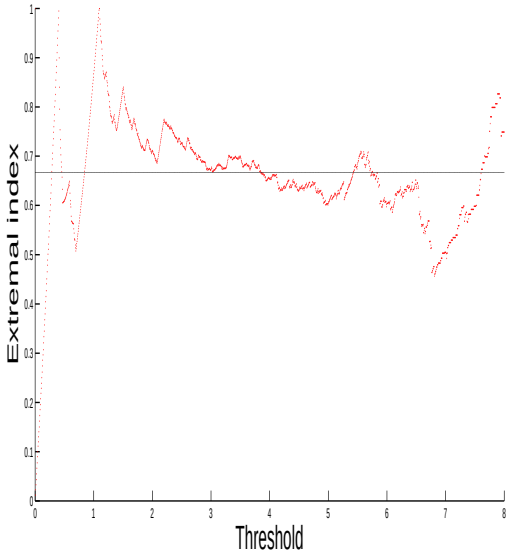


(a)

(b)

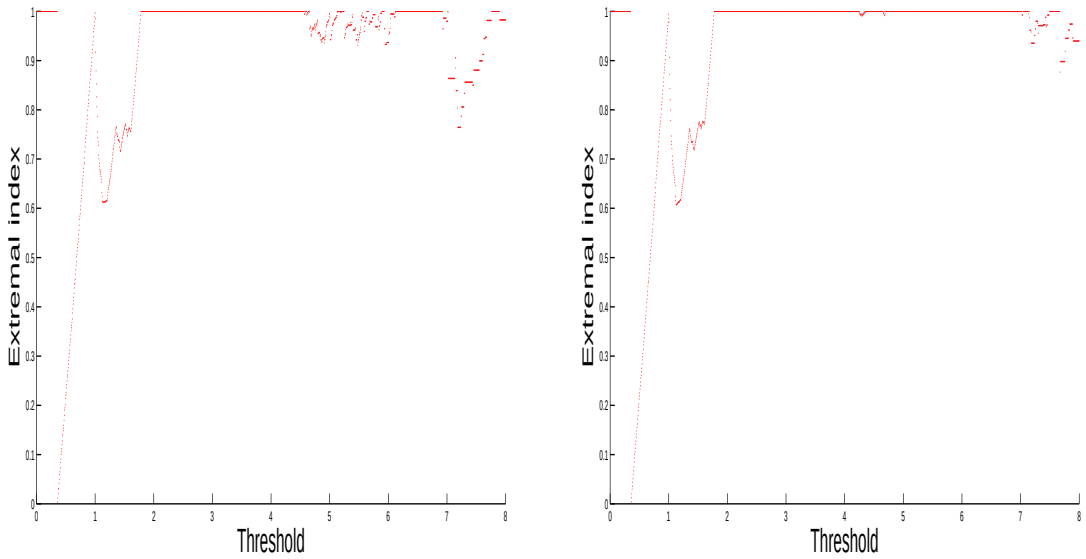


(c)



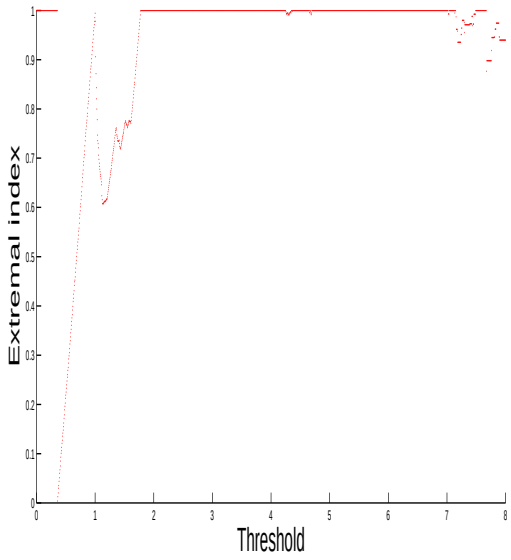
(d)

Figure 4.3. The estimated extremal index versus number of thresholds $u \in [0, 8]$ for Times 3 map with $\tilde{x} = 0$ (a) $n = 2 \times 10^4$ (b) $n = 4 \times 10^4$ (c) $n = 8 \times 10^4$ and (d) $n = 10^5$. The black straight line corresponds to the theoretical extremal index value $\theta = 2/3$.

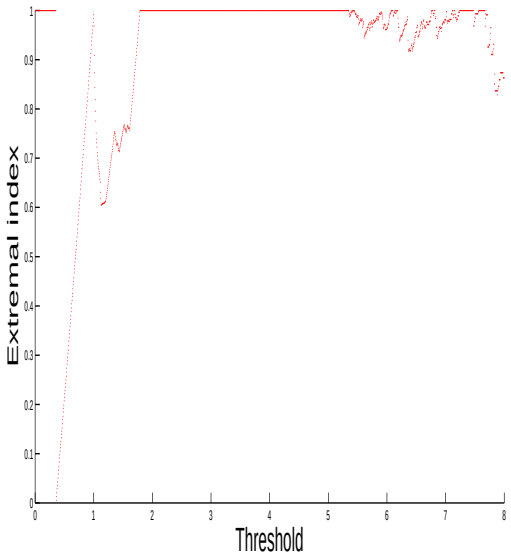


(a)

(b)

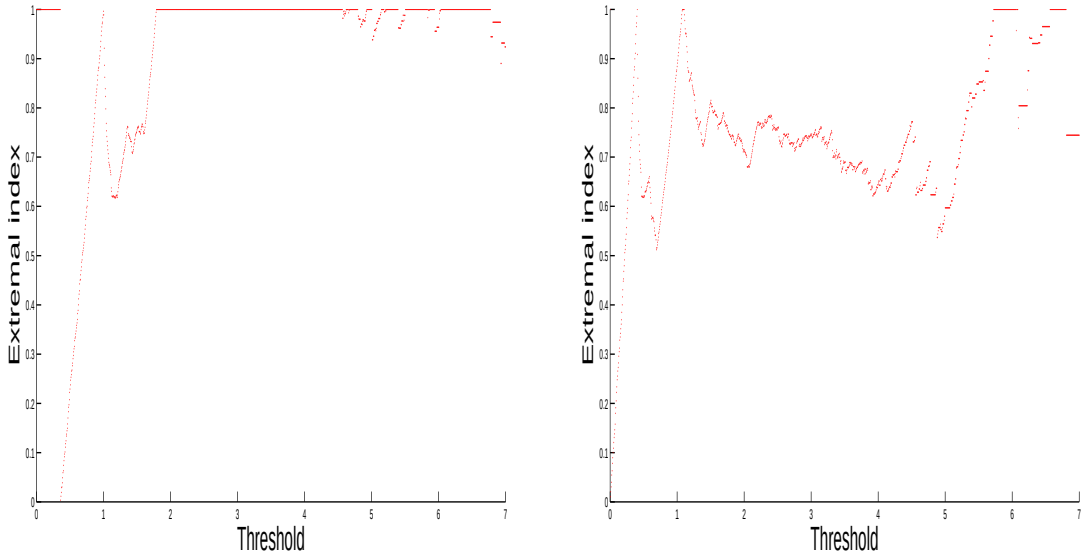


(c)

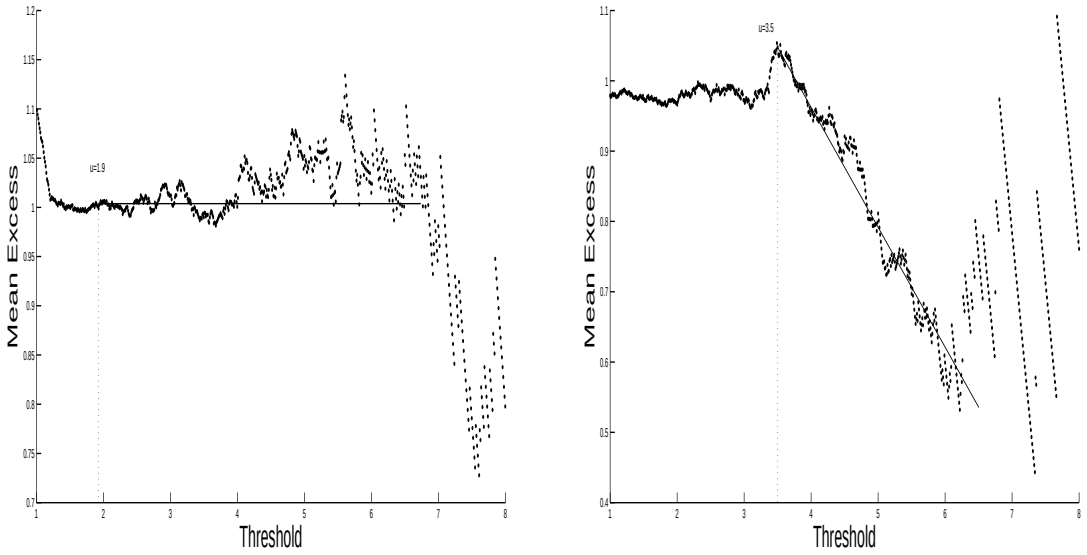


(d)

Figure 4.4. The estimated extremal index versus number of thresholds $u \in [0, 8]$ for the Times 3 map with $\tilde{x} = 0.3$ (a) $n = 2 \times 10^4$ (b) $n = 4 \times 10^4$ (c) $n = 8 \times 10^4$ and (d) $n = 10^5$. The theoretical extremal index value $\theta \approx 1$.



(a) (b)



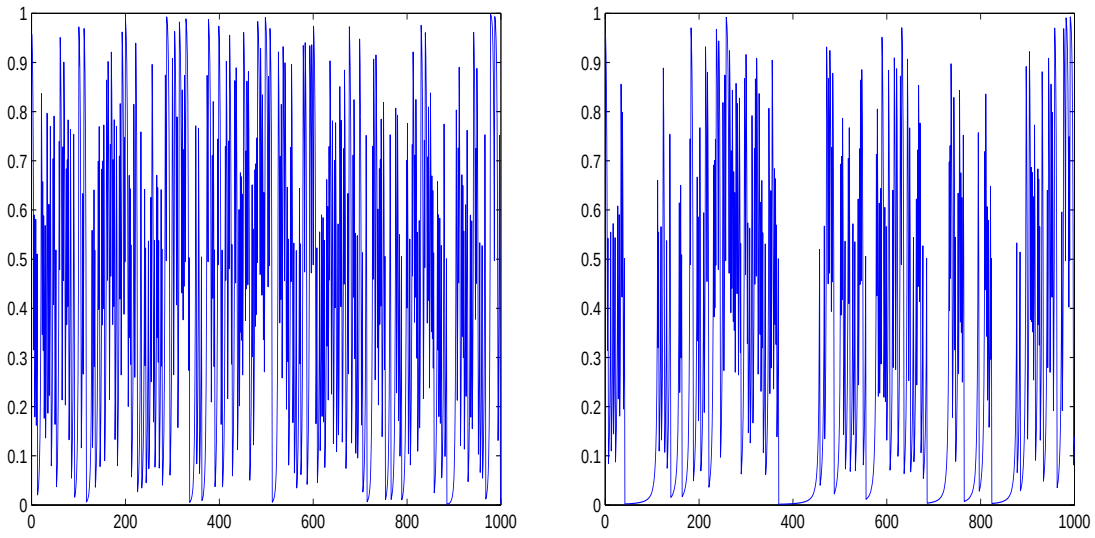
(c) $\tilde{x} = 0.3$. (d) $\tilde{x} = 0$.

Figure 4.5. (a) The estimated extremal index versus number of thresholds $u \in [0, 7]$ for the Times 3 map with $\tilde{x} = 0.3$ and $n = 10^4$. (b) The estimated extremal index versus number of thresholds $u \in [0, 7]$ for the Times3 map with $\tilde{x} = 0$ and $n = 10^4$. (c) The mean excess plot for the Times 3 map with $\tilde{x} = 0.3$. (d) The mean excess plot for the Times 3 map with $\tilde{x} = 0$.

4.5.2. Intermittent map.

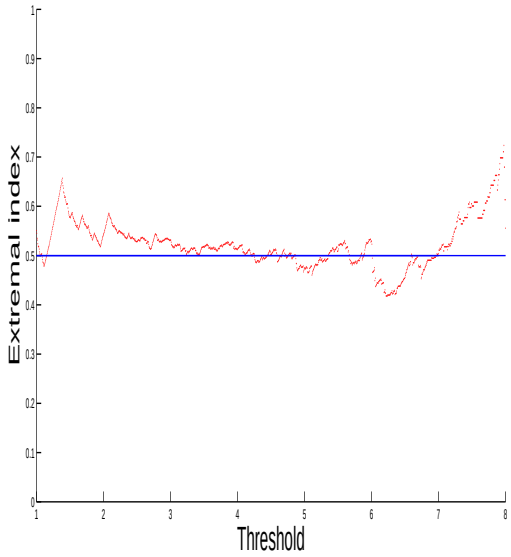
In this section we consider the Intermittent map that is defined by Equation (3.12) under the observable ϕ_1. As in the Times 3 map we started from an initial condition and we iterate the map up to $n = 10^5$ iterations. We estimated the extremal index versus a range of threshold values. As we discussed before, in [41] they showed that the true value of θ is given by $\theta = \theta(\tilde{x}) = 1 - |\frac{1}{(f^p)'(\tilde{x})}|$. Assuming that the observations satisfy $D^p(u_n)$ and $D'_p(u_n)$ conditions around $\tilde{x}$ and, $\tilde{x}$ is a repelling periodic point with period p.

We found that $\tilde{x} = 1$ is a repelling periodic point with $p = 1$, $(|(f)'(\tilde{x}) = 2| > 1)$. The corresponding true value of θ can be obtained from equation (4.11). Hence, we have $\theta = 1 - |\frac{1}{(f^p)'(\tilde{x})}| = 0.5$. We found that when $\beta \in (0, 0.3]$ the result is in agreement with the theoretical extremal index value at most threshold values. Whereas, with $\beta \in [0.5, 1)$ a good agreement can be achieved with high threshold value.
Moreover, we found that for $\tilde{x} = 0$, $(|(f)'(\tilde{x}) = 1|)$, then we expect that $\theta = 1 - |\frac{1}{(f^p)'(\tilde{x})}| \approx 0$. We estimated the extremal index when $\beta = 0.1, 0.3, 0.5$ and 0.7. for $\tilde{x} = 1$ and 0. The results are illustrated in Figure 4.7 and 4.8 respectively. As the result when $\tilde{x} = 0$ we observed that $\theta \to 0$ as we increase β. Here $\tilde{x} = 0$ is known as a neutral fixed point and at this point the intermittent map spend lot of time near zero. See for example [87] and [21]. This can be explained by Figure 4.6. In this Figure we iterate the map with $\beta = 0.3 \text{and} 0.7$ and we choose $n = 1000$.
Furthermore, we consider the case when β is very small(close to 0) with $\tilde{x} \approx 1/3$. Here, We may say that $\tilde{x} = 1/3$ is a repelling periodic point with $p = 2$, $(|(f)'(\tilde{x}) = 4| > 1)$. Hence, we have $\theta = 1 - |\frac{1}{(f^2)'(1/3)}| \approx 0.75$. We observed a good agreement with the theoretical value can be achieved with a high threshold, see Figure 4.9.

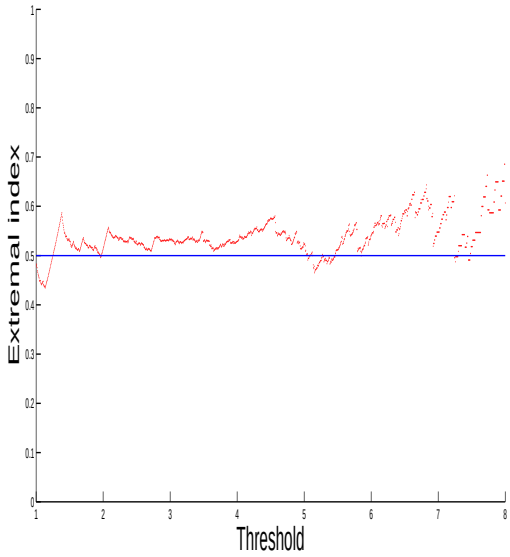


(a) $\tilde{x} = 0, \beta = 0.3$ (b) $\tilde{x} = 0, \beta = 0.7$

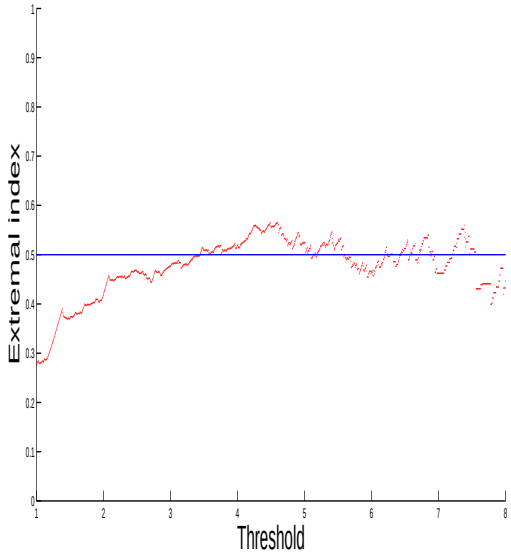
Figure 4.6. 1000 iterations of the intermittent map with (a) $\beta = 0.3$ and (b)$\beta = 0.7$.



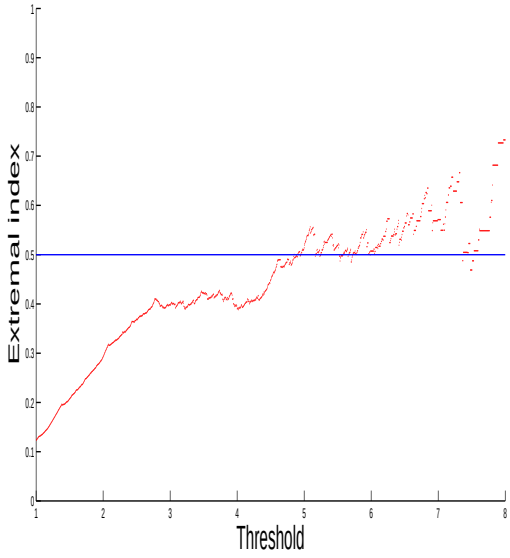
(a) $\tilde{x} = 1, \beta = 0.1$



(b) $\tilde{x} = 1, \beta = 0.3$

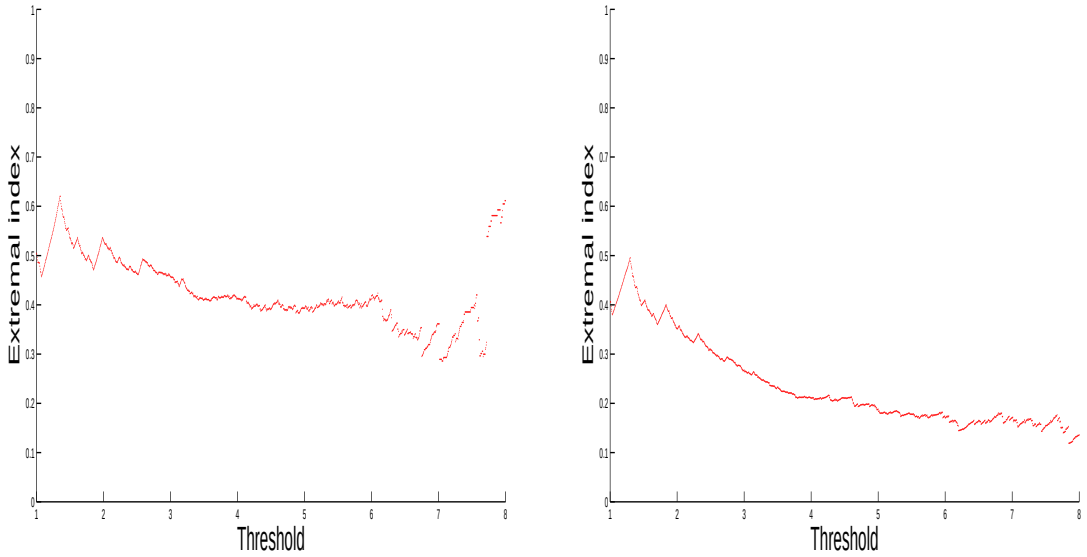


(c) $\tilde{x} = 1, \beta = 0.5$



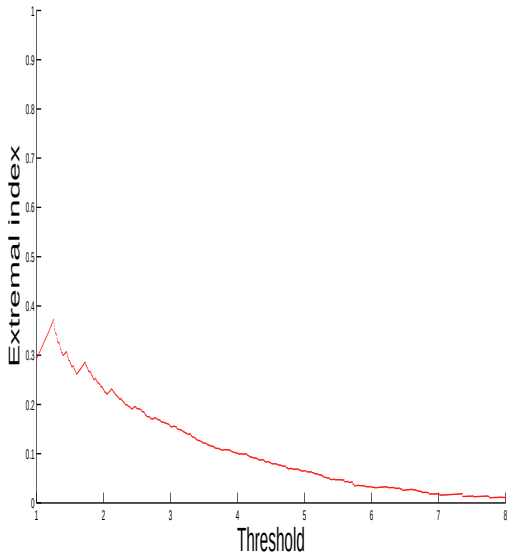
(d) $\tilde{x} = 1, \beta = 0.7$

Figure 4.7. The estimated extremal index versus number of thresholds $u \in [0, 8]$ for the Intermittent map with $\tilde{x} = 1$(a) $\beta = 0.1$ (b) $\beta = 0.3$ (c) $\beta = 0.5$ and (d) $\beta = 0.7$. The straight line corresponds to the theoretical extremal index value $\theta \approx 1$

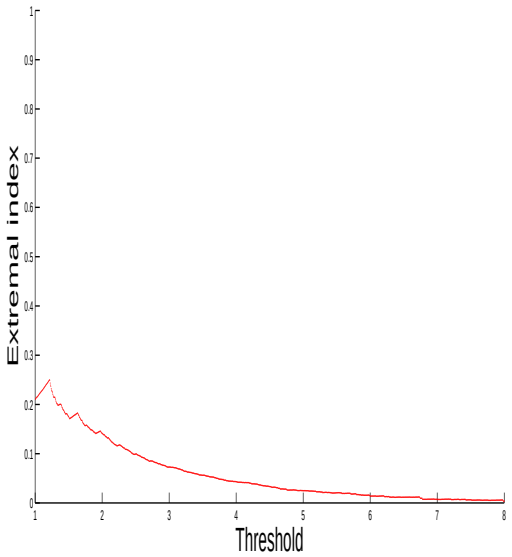


(a) $\tilde{x} = 0, \beta = 0.1$

(b) $\tilde{x} = 0, \beta = 0.3$



(c) $\tilde{x} = 0, \beta = 0.5$



(d) $\tilde{x} = 0, \beta = 0.7$

Figure 4.8. The estimated extremal index versus number of thresholds $u \in [0, 8]$ for the Intermittent map with $\tilde{x} = 0$ and (a) $\beta = 0.1$ (b) $\beta = 0.3$ (c) $\beta = 0.5$ and (d) $\beta = 0.7$. The straight line corresponds to the theoretical extremal index value $\theta \approx 0$

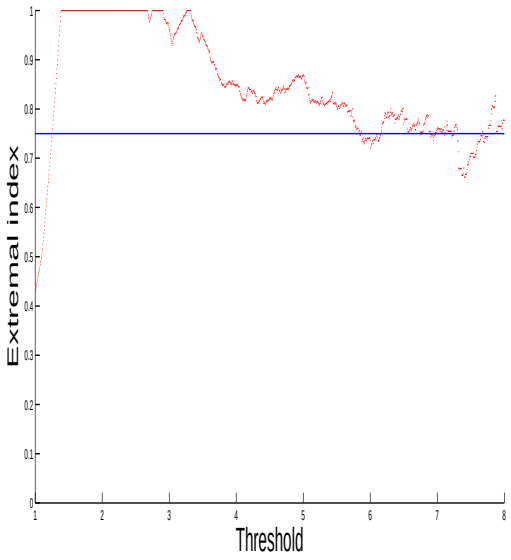
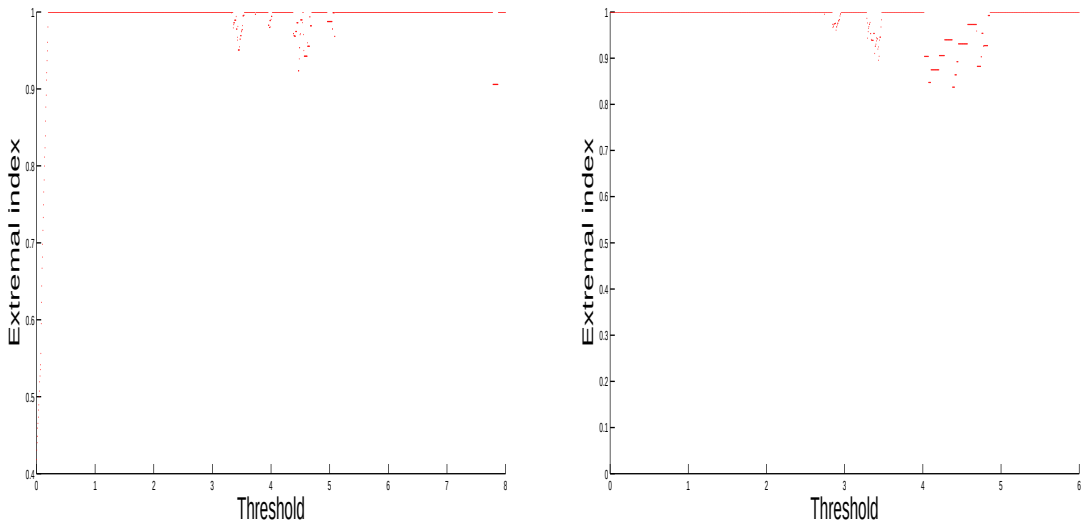


Figure 4.9. The estimated extremal index from the intermittent map with $\tilde{x} = 1/3$ at different thresholds u for $n = 10^5$.

4.5.3. Hénon map

In this example we estimate the extremal index value for the Hénon map under the observable ϕ_1. We expect the extremal index value to be close to 1 for the non periodic point since it is difficult to find a periodic point for this map. The result is illustrated in the following Figure 4.10. We investigate the two cases described in Section (3.4.2). We found that in all considered cases the estimated extremal index value is ≈ 1.



(a) (b)

Figure 4.10. The estimated extremal index of the Hénon map at a range of thresholds u under the observable ϕ_1 for $n = 10^4$. (a) (Case 1) $\tilde{x}$ and x_0 are typical random numbers generated from the Uniform distribution. (b)(Case 2) $x_0 = (0.1386, 0.1493), \tilde{x} = (-0.3051, 0.3016)$.

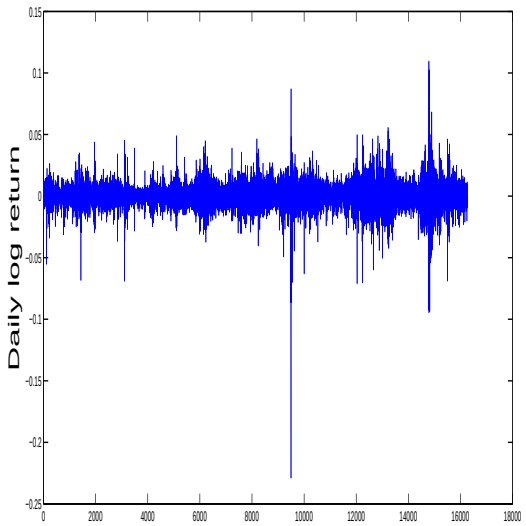
4.6. Extremal Index of Financial Returns

In this example we estimated the extremal index θ, using the interval estimator of Ferro and Segers [34], Ferro [33]. We estimated θ versus a range of thresholds u in S&P500 and FTSE100 log return indices described in the previous Chapter. The result is illustrated in Figure 4.11.

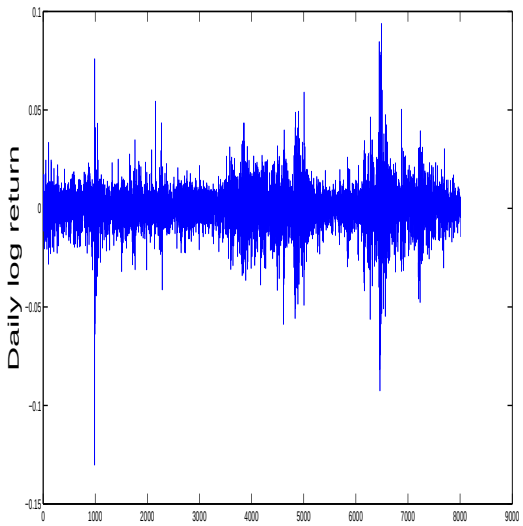
In the S&P500 return index $\theta \approx 0.3740$, this corresponds to the 0.95 quantile for $u = 0.0144$ and, $\theta \approx 0.3051, 0.4379$ this corresponds to the 0.99 and 0.999 quantiles for $u = 0.0255, 0.0489$ respectively. In addition, the result shows a stable region of the plot of $\theta \approx 0.4$. This result is in agreement with Galbraith and Zernov [42] who used the interval estimator and Hamidieh et al. [48] who used the Max-Spectrum based estimator which is based on the re sampling and the scaling parameter in the block maxima approach. They stated that the Max-Spectrum estimator gives a close result to that of the interval estimator depending on the choice of the stable range. However, Longin [64], who used the block method estimator and Finkenstadt and Rootzn [35] who used the point estimator found that $\theta \approx 0.7$.

In the FTSE100 return index, the estimated $\theta \approx 0.3878$. This corresponds to the 0.95 quantile for $u = 0.0159$ and, $\theta \approx 0.2711, 0.7506$, this corresponds to the 0.99 and 0.999 quantiles for $u = 0.0282, 0.0535$. In addition, the result shows a stable region of the plot of $\theta \approx 0.3$. This result is in agreement with Robert et al. [80] who used the sliding block method and the interval estimator to the negative log returns of FTSE100 index. However, Laurini and Tawn [60] estimated the extremal index using the two threshold estimator for the negative and positive log return of FTSE100 index. They found that θ is mostly 0.9 throughout the whole period of the positive log return and slightly differs in the negative log return.

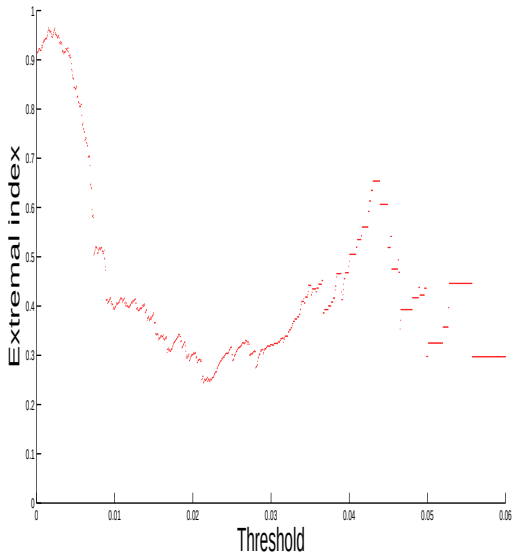
In addition, the estimated threshold value by using the mean excess plot is illustrated in Section 3.5. For the S&P500 index the estimated threshold values were 0.0054 and 0.15 and the corresponding extremal index values are 0.8149 and 0.3861 respectively. Similarly for the FTSE100 index the estimated threshold values were 0.011 and 0.008and the corresponding extremal index values are 0.6397 and 0.8518 respectively. An accurate result is not obvious but we may say that in both S&P500 index and FTSE100 index extremes accept clustering.



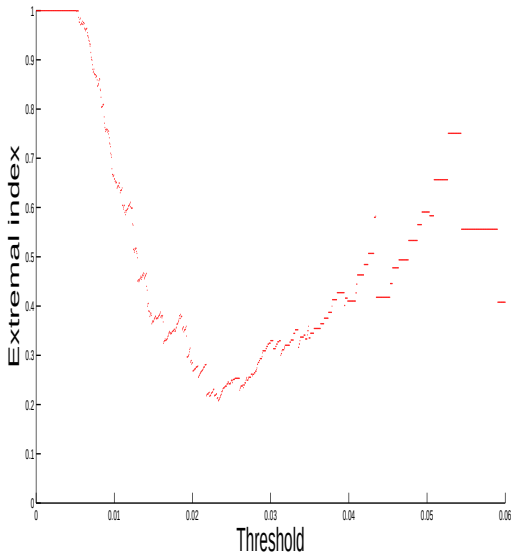
(a) S&P500 return index.



(b) FTSE100 return index.



(c) S&P500 extremal index.



(d) FTSE100 extremal index.

Figure 4.11. (a) Daily log returns of the S&P500 index for the period from January 1950 to September 2014, that consists of ($n = 16278$) observations. (b) Daily log returns of the FTSE100 index for the period from January 1984 to September 2014 that consists of ($n = 8010$) observations. (c)The estimated extremal index of S&P500 index return at a range of threshold u (d)The estimated extremal index of FTSE100 index return at a range of threshold u

4.7. Summary and Remarks

We have discussed in this Chapter the extension of the extreme value theory in dynamical systems examples. Leadbetter et al. [61] showed that in stationary sequence if D' condition does not hold then extreme value distribution has similar behaviour as IID observations but with respect to the clustering parameter θ. A similar idea holds in the Dynamical system around periodic point, as described recently in [41].

In this chapter, we estimated the extremal index for a range of dynamical system examples, by using the interval estimator method given by Ferro and Segers [34], and Ferro [33] for a range of threshold values. We consider the Times 3 map around the periodic point $\tilde{x} = 0, 0.3$. We found that the result is in agreement with the theoretical extremal index value with threshold u value large but not too large. Moreover, we consider the Intermittent map with different value of β around $\tilde{x} = 1$. We found a poor convergence observed as $\beta \to 1$. also we test the case when β is very small (≈ 0) around the point $\tilde{x} = 1/3$. We found that a good agreement was achieved as we increase u.

Interestingly, we also examined the behaviour of the extremal index for the intermittent map around the neutral fix point $\tilde{x} = 0$. We observed that as $\beta \to 1$ the estimated extremal index $\theta = 0$. This result is in agreement with the idea that the intermittent map spend lots of time near zero as $\beta \to 1$ (in other words, we get more clustering as we increase β around this point), see Thaler [87]. For the Hénon map, since it is difficult to find a periodic point we estimated the extremal index for the cases described in Chapter 2. We found that the extremal index is ≈ 1 for the two cases.

Similarly, for the financial returns in S&P500 and FTSE100 indices we observe that the extremal index satisfies $\theta < 1$. In comparison with the dynamical systems examples with observables maximized at a periodic point, we may say that the financial return data in our case is not IID data. If it is IID then we expect the extremal index to be equal to one (since no clustering occurs for IID). Furthermore, the extremal index for many financial returns data has value strictly less than one, see for example [80],[42] and [48]. Hence, the Geometric random walk or the Geometric Brownian motion prices models discussed in Chapter 3 which assume that the returns are IID, might not be a reliable model to describe the prices or the returns. Other models of the financial returns data such as ARCH and GARCH processes, the extremal index satisfies $\theta < 1$ see for example [25] and [26]. In these models the volatility changes in time and therefore we might investigate the volatility clustering in a future study.

5. Measuring Dependency

5.1. Introduction and Background

The famous Black and Scholes [11] model assumed that the stock market returns are independent and modelled by the Brownian Motion and this supports the idea of the efficient market hypothesis[1]. However, empirical evidence showed that stock markets returns accept dependency as Greene and Fielitz [44] when they analyse the New York Stock Exchange. Therefore, there has been wide interest in studying the dependency behaviour of the stock market returns. The classic Rescaled Range R/S analysis is a widely used method to estimate the Hurst parameter $H \in (0, 1)$ as a measure of dependency. This parameter is used as a measure of long memory process if $H > 0.5$ and to short memory process if $H < 0.5$. Then the R/S method is the range of partial sums of deviation of a time series from its mean which was first introduced by Hurst [58]. Then it became widely used in finance and economics after Mandelbrot [70] applied it to financial data. It is used to estimate the Hurst value H as a measure of long range dependence in the time series data. It is very effective in large sample but it is sensitive to the short range dependence. Lo [63] drives the modified R/S statistics to avoid this problem. In [86] they look at Lo'S method (modifiedR/S statistic) and strongly advise not to use this method for testing for long range dependence. The Variance Rescaled analysis V/S is suggested by Cajueiro and Tabak [13] as an improvement and alternative method to the classic R/S analysis. In [12], they tested the long range dependence in developed, emerging and transition economies. by using the classic R/S and V/S methods in equity returns and volatility and found that the estimate value of the Hurst parameter for R/S method is in the range (0.5 - 0.66) and for V/S is in the range (0.4- 0.62). Furthermore, they conclude that the evidence of the long range dependence is small in equity return. He and Qian [49] compare the R/S to the V/S method by means of Monte Carlo simulation. They conclude that they cannot judge which method is better but in general the R/S analysis gives a value < 0.5 if the theoretical value < 0.5 and around 0.5 if the theoretical value around 0.5 and > 0.5 if the theoretical value > 0.5. Furthermore, R/S analysis gives better estimation in real world application.

In this chapter we measure the dependency in the data by using the R/S analysis. We test this method in IID observations generated from Gaussian and Uniform distribution. Also, we test this method in our dynamical system examples described in the previous

[1] The efficient market hypothesis, which assumes that market returns are unpredictable see for example [31].

chapters such as, Intermittent map for several different values of the parameter $\beta \in (0,1)$, Times 3 map, Hénon map. We then apply to the S&P500 and FTSE100 returns indices. We employ the R/S analysis as described in [6] and [12] to estimate the Hurst exponent as a measure of dependency. We show that the estimated value of the H exponent of the financial returns data suggests short range dependence, but we cannot rule out clustering effects (as is evident in Chapter 4). This is consistent with what we see in dynamical system models, e.g. Times 3 map but with observable maximized around a periodic point.

A long range dependence can be defined using the autocorrelation function as:

Definition 5.1 *Let (X_t) be a stationary process with mean $\mu = E(X_t)$ and $C(\tau)$ the autocorrelation that is given by:*

$$C(\tau) = \frac{E[(X_{t+\tau} - \mu)(X_t - \mu)]}{E[(X_t - \mu)^2]},$$

then we say that the process has long range dependence or long memory if

$$C(\tau) \sim c\tau^{-\alpha} \text{ as } \tau \to \infty. \tag{5.1}$$

Where, $0 < \alpha < 1$ and $c > 0$ is constant, and so $\sum_{\tau=0}^{\infty} C(\tau) = \infty$, (the correlation decay to zero so slowly).

Remark. On the other hand, the process has short range dependence if for $\alpha > 1$, and $\sum_{\tau=1}^{\infty} C(\tau) < \infty$ (the correlation decay to zero very fast).

Here, $C(\tau) \sim c\tau^{-(2-2H)}$ as $\tau \to \infty$ where, $H = 1 - \alpha/2 \in (0.5, 1)$ is the Hurst exponent, see [6].

- For $0 < H < 1/2$, $\sum_{\tau=1}^{\infty} C(\tau) < \infty$ the process has short range dependence if $C(\tau) \sim c\alpha^{-\tau}$ as $\tau \to \infty$ where, $\alpha > 1$.(the correlation decay to zero very fast as exponential decay).

- For $1/2 < H < 1$, $\sum_{\tau=1}^{\infty} C(\tau) = \infty$ the process has long range dependence if $C(\tau) \sim c\tau^{-\alpha}$ as $\tau \to \infty$ where, $0 < \alpha < 1$.(the correlation decay to zero very slowly as a power function).

Definition 5.2 *[29] A Brownian motion or the Wiener process is a stochastic process $(W_t)_{t\geq 0}$ with the following properties:*

- *W_t is a continuous function of t with probability 1 and $W_0 = 0$;*

- *for every $t \geq 0$ and $s > 0$, the increment $W_{t+s} - W_t$ has normal distribution with mean zero and variance s; thus,*

$$P(W_{t+s} - W_t \leq x) = \frac{1}{\sqrt{2\pi s}} \int_{-\infty}^{x} \exp\left(\frac{-u^2}{2s}\right) du.$$

- *For any* $0 \leq t_0 \leq t_1 \leq ... \leq t_2 m$, *the increments* $W_{t_2} - W_{t_1}, W_{t_4} - W_{t_3}, ..., W_{t_{2m}} - W_{t_{2m-1}}$, *are independent.*

The Brownian motion is a special case of a general process called Fractional Brownian motion that is defined as:

Definition 5.3 *[29] The Fractional Brownian motion of index* $-H(0 < H < 1)$ *is defined to be a Gaussian process* $X : [0, \infty) \to \mathbb{R}$ *on some probability space such that:*

- X_t *is a continuous function of* t *with probability 1 and* $X_0 = 0$*;*
- *for every* $t \geq 0$ *and* $s > 0$ *the increment* $X_{t+s} - X_t$ *has normal distribution with mean zero and variance* s^{2H}*, so that*

$$P(X_{t+s} - X_t \leq x) = \frac{1}{s^H \sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(\frac{-u^2}{2s^{2H}}\right) du.$$

The Fractional Brownian motion is a Brownian motion with $H = 0.5$. The increments of the Brownian motion are independent on the other hand the increments of the fractional Brownian motion accept dependency. Further details on this process can be found for example in [29] and [81]. The following Theorems give the H exponent of the Brownian motion and fractional Brownian motion processes.

Theorem 5.1 *[6] If* X_t *is such that* X_t^2 *is ergodic and* $\frac{1}{\sqrt{t}} \sum_{s=1}^{t} X_s$ *converge weakly to a Brownian motion then*

$$\frac{(R_\tau / S_\tau)}{\sqrt{\tau}} \to_d \xi$$

where, ξ *is a continuous RV.*

Theorem 5.2 *[6] If* X_t *is such that* X_t^2 *is ergodic and* $t^{-H} \sum_{s=1}^{t} X_s$ *converge weakly to a Fractional Brownian motion then*

$$\frac{(R_\tau / S_\tau)}{\tau^H} \to_d \xi$$

where, ξ *is a continuous RV.*

5.2. Measuring the Hurst Exponent

In this section we give a brief description of the classic R/S analysis. We will apply this method to estimate the Hurst exponent in independent processes, dynamical system examples and in S&P500 and FTSE100 returns indices.

The R/S analysis is described by the following:

Consider we have a time series data $\{x_1, x_2, ..., x_n\}$ of length n, and a sub period data

$\{x_1, x_2, ..., x_\tau\}$ of length $\tau < n$. Then the Hurst exponent can be calculated by the following steps:

- Calculate the mean
$$\bar{x} = \frac{1}{\tau}\sum_{t=1}^{\tau} x_t, \tag{5.2}$$

- Calculate the standard deviation
$$S_\tau = \sqrt{\frac{1}{\tau}\sum_{t=1}^{\tau}(x_t - \bar{x})^2}, \tag{5.3}$$

- Compute the range R
$$R_\tau = \max_{1\le t\le\tau}\left(\sum_{k=1}^{t}(x_k - \bar{x})\right) - \min_{1\le t\le\tau}\left(\sum_{k=1}^{t}(x_k - \bar{x})\right), \tag{5.4}$$

- Finally, the rescaled range is given by $(R/S)_\tau$,

- Repeat the above steps by increasing the value of τ and plot $Log(R/S)_\tau$ versus $Log(\tau)$. Then H corresponds to the slope of the straight line which can be described as
$$Log(R_\tau/S_\tau) \approx Log(c) + HLog(\tau). \tag{5.5}$$

Where, c is constant. The V/S analysis is similar to that of R/S analysis. In the V/S method the range is given by:

$$R_\tau = \sum_{t=1}^{\tau}\left(\sum_{k=1}^{t}(x_k - \bar{x})\right)^2 - \frac{1}{\tau}\left(\sum_{t=1}^{\tau}\sum_{k=1}^{t}(x_k - \bar{x})\right)^2, \tag{5.6}$$

Then, the Rescaled Variance is given by:

$$(V_\tau/S_\tau) = \frac{\sum_{t=1}^{\tau}(\sum_{k=1}^{t}(x_k - \bar{x}))^2 - \frac{1}{\tau}(\sum_{t=1}^{\tau}\sum_{k=1}^{t}(x_k - \bar{x}))^2}{\tau S_\tau^2}. \tag{5.7}$$

Plot the $Log(V_\tau/S_\tau)$ versus $Log(\tau)$. Then H corresponds to the slope of the straight line divided by 2.
See [6] and [12].

Example 5.1 (Hurst in IID RVs) Here we apply the R/S analysis to estimate the H exponent in IID RVs. We generated 10^3 samples of length $n = 5000$ drawn from Gaussian and Uniform distribution. We take the average of H exponent over 10^3 samples. We start with a time span of length 100 then we increase this by 10. The results are illustrated in Table 5.1. We found that $H \approx 0.5$.

Similarly, we will estimate the Hurst exponent in dynamical system examples as presented in the next Section.

5.3. Computing the Hurst Parameter in Dynamical Systems

In this section we estimate the Hurst exponent by using the R/S method for the Times 3 map, Intermittent map and Hénon map under the observable $\phi_1(x) = -\log \text{dist}(x, \tilde{x})$ where, $\tilde{x}$ is a typical random number.[2] Starting with an initial value, generated from the uniform distribution. We iterate the map up to $n = 5000$ iterations. Then we take the average of the H exponent over 10^3 samples for different initial value. The results are illustrated in Table 5.1.

A Numerical estimate for H in the Times 3 map gives $H = 0.5165$, and this is close to the theoretical value of $H = 0.5$ (since this map has short range dependence).

For the Intermittent map we found that the value of the H exponent increases as we increase β. We observed that $H \to 1$ as $\beta \to 1$. We also found that H around 0.5 when $\beta < 0.5$. It was shown in [73] that this map under suitable scaling of Birkhoff sums converges to the Brownian motion for $\beta \in (0, 0.5)$. The results are in agreement with the results obtained by Bhansali et al. [7]. They found that the intermittent map with $\beta \in (0, 0.5)$ has short range dependence and with $\beta \in (0.5, 1)$ has long range dependence. Moreover, the estimated value of the H exponent can explain the poor convergence to extreme value distribution when $\beta \in (0.5, 1)$, see Section 3.4.3.

It was shown in [95] that the Hénon map has an exponential decay of correlation. Therefore, we expect to find $H \approx 0.5$. For this map we considered the observable function in (one dimensional). We started with the initial value $x_0 = (x_0, y_0)$ where x_0 is a random number generated from the Uniform distribution where, $y_0 \leq 0.1$. [3] As before the result obtained by taking the average of the H exponent over 10^3 samples for different initial value. We found that $H \approx 0.4930$. This verifies that the Hénon map has an exponential decay of correlation or short range dependence.

From the results, we may use the estimated value of the H exponent as a measure of the decay of correlation or speed of mixing and then for checking $D(u_n)$ condition or the dynamical system version such that (H1) or $D_2(u_n)$ conditions defined in Chapter 2 which can help to investigate the convergence to extreme value distribution.

[2] Here, the estimated value of H exponent does not depend on the observable function.
[3] The range of the initial value here is chosen as given in [22] Exercise (5.10).

	Mean(H)
Gaussian RVs	0.5160
Uniform RVs	0.5143
Times 3 map	0.5165
Intermittent map	
$\beta = 0.1$	0.5216
$\beta = 0.3$	0.5334
$\beta = 0.5$	0.5971
$\beta = 0.7$	0.6951
$\beta = 0.9$	0.7865
$\beta = 0.99$	0.8173
Hénon map	0.4930

Table 5.1. Estimation of the mean Hurts exponent in IID RVs and dynamical system examples, using the R/S analysis. The average of the H exponent is computed over 10^3 samples of length $n = 5000$ for different initial value.

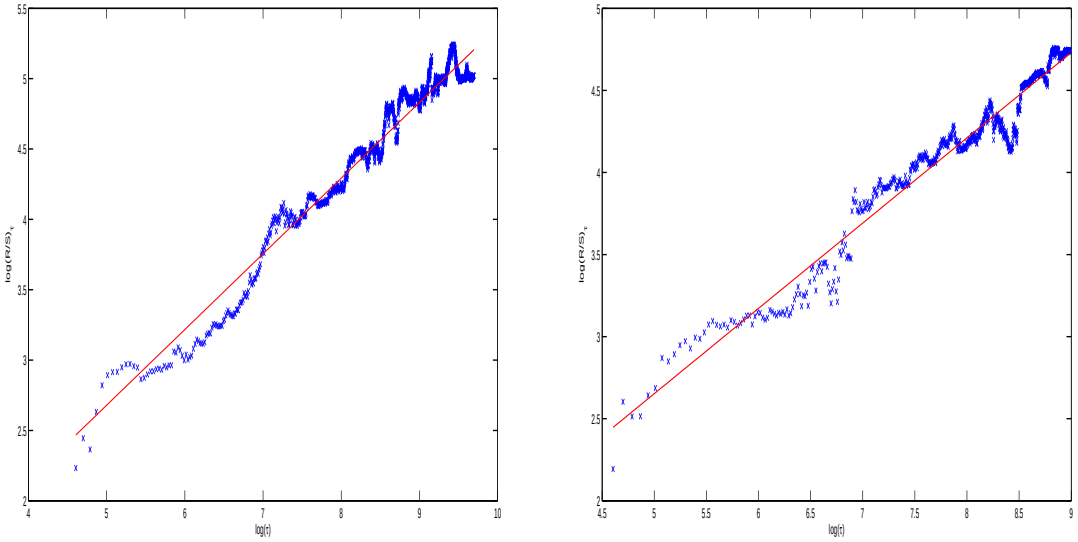
5.4. Computing the Hurst in Financial Returns Data

In this section we estimated the Hurst exponent by using the R/S analysis to the data described in Chapter 2. The results are illustrated in Table 5.2 and Figure 5.1 [4]. We found that the Hurst value is around 0.5 from both indices. This result suggests independent or short range dependence. Our results are in agreement with the result obtained by Cajueiro and Tabak [12] for both indices. Also, Couillard and Davison [19] applied the R/S analysis to the returns of the S&P 500 index. They found ($H \approx 0.54$) and they conclude that the Brownian motion cannot be rejected as a model for the price index (i.e a model with independent returns).

Index	H	standard error	R^2	adjusted R^2
S&P500	0.5377	0.1095	0.9532	0.9532
FTSE100	0.5196	0.0979	0.9554	0.9554

Table 5.2. Estimation of the Hurst exponent for daily log return of S&P 500 and FTSE100 indices, using the R/S method.

[4] The stander error, R^2 and adjusted R^2 are measures of the goodness of fit. See B Appendix for the Matlab code.



(a) S&P500 (b) FTSE100

Figure 5.1. (a) The R/S plot of the daily log returns of the S&P500 index. (b) The R/S plot of the daily log returns of the FTSE100 index.

6. Investigation into Record Behaviour of Asset Returns

6.1. Introduction and Background

Record statistics were well studied many years ago for IID RVs see for example [1]. Record statistics is defined as the following:

Definition 6.1 *Consider a sequence of RVs* (X_n)*, we say an (upper record) occurs at time* n *if*

$$X_n > \max\{X_1, X_2, ..., X_{n-1}\}$$

(Lower records if $X_n < \min\{X_1, X_2, ..., X_{n-1}\}$*). Record times* L_{N_n} *are time where we observe the* N_n^{th} *record value at time* n*. for example,*

$$L_{N_n} = \min\{i : X_i > X_{L_{N_n-1}}\}$$

Then, $X_{L_{N_n}} = R_{N_n}$ *the record value.*

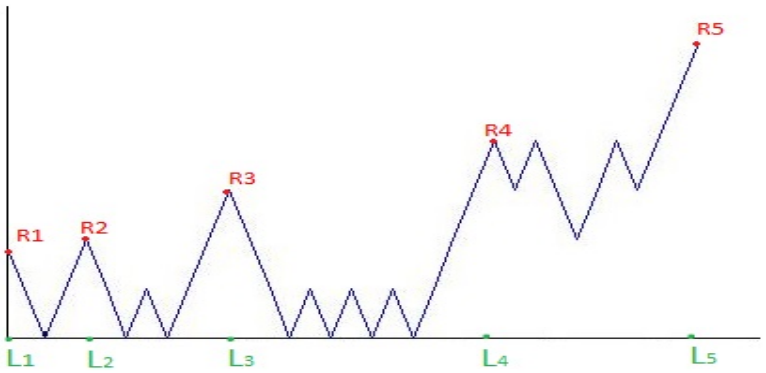


Figure 6.1. Record times

In this chapter we are interested in answering the following question:

Does the growth of record times for dynamical systems and financial returns coincide with the theory for IID process ?

Record times can be used as a tool to understand the dependency structure of the financial returns data. For example, if record times behave as record times of the IID observation then we might obtain the same limit of extreme value distribution as the

IID case. In our study we will look at upper records and we will call it strong record if $X_n > \max\{X_1, X_2, ..., X_{n-1}\}$ and weak record if $X_n \geq \max\{X_1, X_2, ..., X_{n-1}\}$. We attempt to analyse the behaviour of records in a sequence (ξ_i) of IID Rvs. Then we will analyse the process generated by a dynamical system where, $\xi_i = \phi(f^{i-1})$. We will compare the results with the financial returns data. Recently, there has been a wide interest in studying records in a random walk process because of its wide use in several fields. See for example, [69], [92] and [68]. An application to the S&P500 prices index can be found in [92] and [90]. In this chapter, we consider the limit theorem associated with record times for IID RVs (growth of record times, the frequency of records) see, [26] to analyse the behaviour of records in dynamical system examples and to the Financial returns and prices data. For the prices we assume that $X_n = \xi_1 + \xi_2 ... + \xi_n$ where, (ξ_i) are IID with $P(\xi_i = +1) = p, P(\xi_i = -1) = q = 1 - p$ (Random Walk model). Then we compare this model to the case where (ξ_i) form a dependent sequence, such as a Markov chain model and dynamical systems.

In this chapter, we start with a review of Embrechts et al. [26] results for record statistics in IID RVs. The results of Embrechts et al. [26] are tested numerically on IID observations, in Section 6.2. We also apply their result to the Times 3 map, Intermittent map with different values of β and to the financial returns in comparison with the IID case, in Section 6.3 and 6.4 respectively. In Section 6.4, we discuss the main finding of Majumdar and Ziff [69], Wergen et al. [92] and Majumdar et al. [68] for the random walk model and we consider an example of a simple random walk with dependent increments or jumps to analyse the expected number of records. Finally we discuss the growth of record number in the Financial prices.

6.2. Records for IID RVs

Assume that (ξ_i) are IID RVs with continuous distribution function F. Let N_n be the number of records observed up to time n. This number is given by:

$$N_n = \sum_{i=1}^{n} I_i \tag{6.1}$$

where, I_i is an indicator function defined by:

$$I_i = \begin{cases} 1 & \text{if we observe a record ,} \\ 0 & \text{otherwise.} \end{cases} \tag{6.2}$$

The probability of the ith record or record rate is given by :

$$P_i = P(\xi_i > \max\{\xi_1, \xi_2, ..., \xi_{i-1}\}) = \frac{1}{i}.$$

Hence, the expected or the mean number of records up to time n is :

$$\begin{aligned} E(N_n) &= \sum_{i=1}^{n} E(I_i) \\ &= \sum_{i=1}^{n} (1.P(I_i = 1) + 0.P(I_i = 0)) \\ &= \sum_{i=1}^{n} \frac{1}{i} \\ &= H_n \approx \ln n + \gamma \text{ as } n \to \infty \end{aligned} \tag{6.3}$$

Where, H_n is a harmonic number and $\gamma = 0.577215$ is the Euler-Mascheroni constant. See [1].
This indicates that the expected number of records grows as $\ln n$ for large n. Moreover, this result is independent of the underlying distribution f. Embrechts et al.[26] state a theorem on the limit distribution of record times and record number for IID RVs as presented in the following Theorem:

Theorem 6.1 *(Limit result for the growth of record times) Suppose ξ_n are IID with continuous distribution function. If L_{N_n} denotes the N_n^{th} record time then*

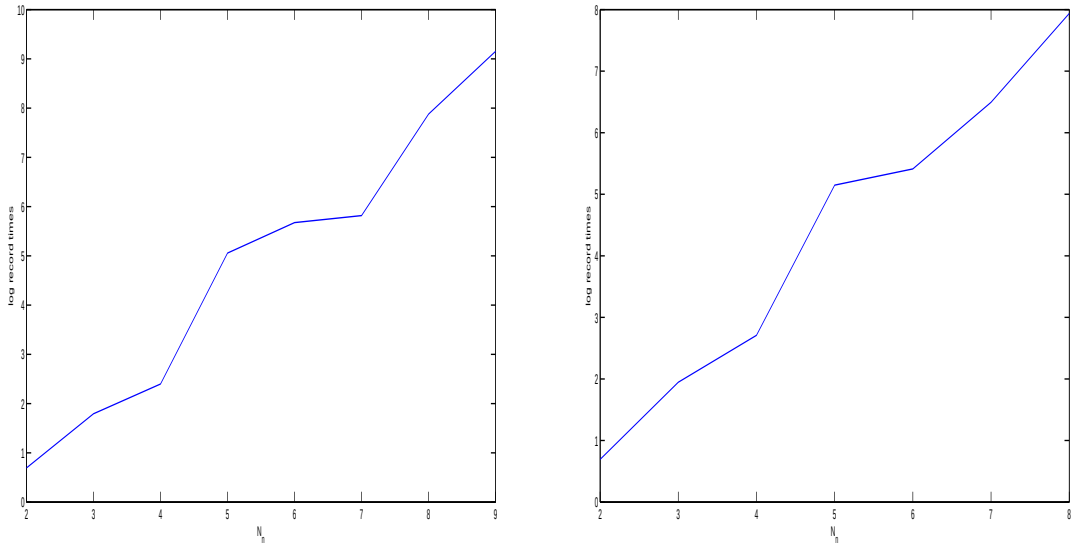
$$\lim_{N_n \to \infty} {N_n}^{-1} \ln L_{N_n} = 1 \ a.s. \tag{6.4}$$

In other words, this theorem says for large number N_n of records observed the ratio of the logarithm of the time of the N_n^{th} record L_{N_n} to the number of records converge to 1 for IID observations. We test this theory of records for a sequence of IID RVs drawn from Gaussian, Uniform, log normal and heavy tailed distributions when the number of observations is $n = 10^4$ as illustrated in Figure 6.2. In these examples we found that the expected number of record times is around 8 and $\log L_{N_n}/n$ seems to converge to one. However, to verify that we need a very large number of record times for a large number of observations as well. Furthermore, [26] state a theorem on the limit behaviour of the growth number of records in a given period of time for IID RVs.

Theorem 6.2 *(Limit result for the frequency of records) Suppose X_n are IID with continuous distribution function. If $N(1, n]$ denotes the number of records observed up to time n then*

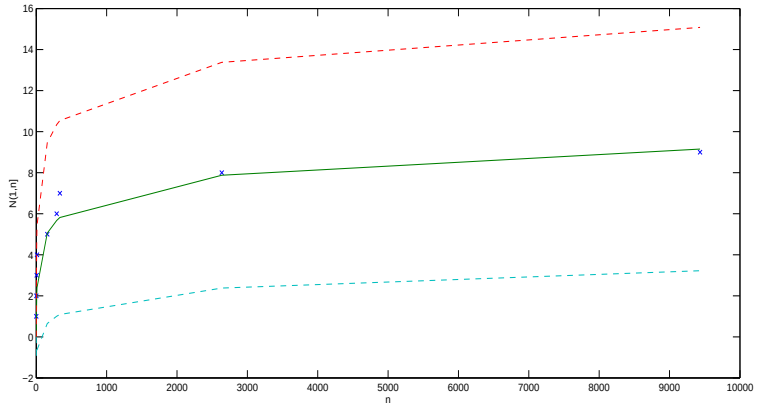
$$\lim_{n \to \infty} (\ln n)^{-1} N(1, n] = 1 \ a.s. \tag{6.5}$$

This theorem suggests that the number of record in the interval $(1, n]$ grow as $\ln n$ for large n. We aim to use this theorem as a tool for testing dependency by analysing the behaviour of the number of records. We first apply this Theorem to test the limit behaviour of the record number in Gaussian and Uniform observations with ($n = 10^4$). As illustrated in Figure 6.3, we found that the number of records grows as $\ln n$ with 95% confidence bands and the observed number of records is ≈ 9. This is close to the number of records given by equation (6.3) for $n = 10^4$.

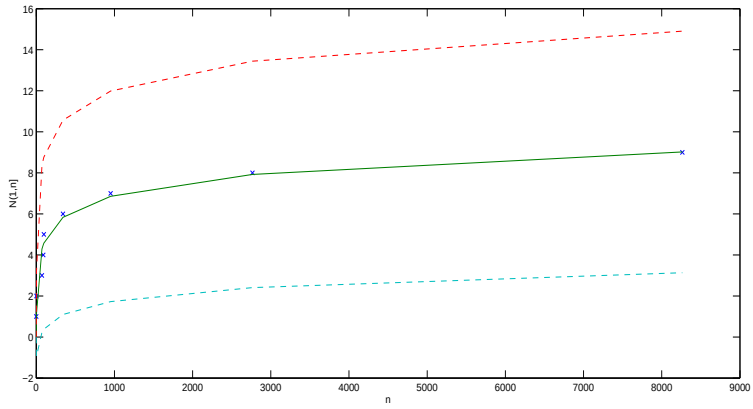


(a) (b)

Figure 6.2. The logarithm of record times to the number of records observed in IID RVs ($n = 10^4$) generated from (a) Gaussian distribution (b) Uniform distribution.



(a)



(b)

Figure 6.3. The number of records $N(1, n], n = 10^4$ from IID (a) Gaussian RVs (b) Uniform RVs. The solid line (middle) is $\ln n$. The dashed lines are the 95% confidence bands.

Remark. In Chapter 4 we remarked that the point process of exceeding threshold in stationary sequence under $D(u_n), D'(u_n)$ condition converges to Poisson process see for example, [61] and [26]. Indeed, similar behaviour has been observed in dynamical system see, for example [40]. Also, Leadbetter et al. [61] proved that the point process of record times in stationary sequence satisfies similar conditions, and converges to a Poisson process. To our knowledge, no study has investigated the convergence of the point process of record times in dynamical systems.

Extremes and records:

It was shown in [1] that the limit distribution of normalized record values such that for constants a_n and b_n we have $P(a_n(X_{L_N} - b_n)) \to \tilde{G}$ where, $\tilde{G} = \Phi(-\log(-\log G(x)))$ here $\Phi(x)$ is a standard normal distribution and $G(x)$ is the GEV distribution that was defined in Chapter 2. There are three possible distributions of normalized record values given below:

1. $$\Phi(x) = \int_{-\infty}^{x} \frac{e^{-x^2/2}}{\sqrt{2\pi}}\, dx.$$

2. $$\Phi_\alpha(x) = \begin{cases} 0 & x \leq 0, \\ \Phi(\log(x)^\alpha) & x > 0. \end{cases}$$

3. $$\tilde{\Phi}_\alpha = \begin{cases} \Phi(\log(-x)^{-\alpha}) & x < 0\ , \\ 1 & x \geq 0. \end{cases}$$

Where, $\alpha > 0$ and 1, 2 and 3 correspond to type I, II and III (GEV) distribution respectively.

6.3. Records in Dynamical Systems

In this section we apply Theorem 6.2 to the dynamical system examples defined in the previous chapters such as the Times 3 map and Intermittent map with different values of β. Furthermore, we will apply this to the S&P500 return index and FTSE100 return index. We find that the growth of record number agrees with the result obtained in Chapter 5 for the H exponent. In the IID case, $H = 0.5$ and $N(1, n] \sim \ln n$. If $H > 0.5$ we also observe that $N(1, n]$ does not grow as $\ln n$ and this implies dependency. However, if $H = 0.5$ and $N(1, n] \sim \ln n$ this does not imply independency.

6.3.1. Expanding Times 3 map

Here we apply Theorem 6.2 to the Times 3 map defined in Chapter 3, see Figure 6.4. The number of records in the interval $(1, n], n = 10^4$ grows as $\ln n$. This behaviour is similar to that of IID case. This also confirms the result obtained by the H exponent in Chapter 5.

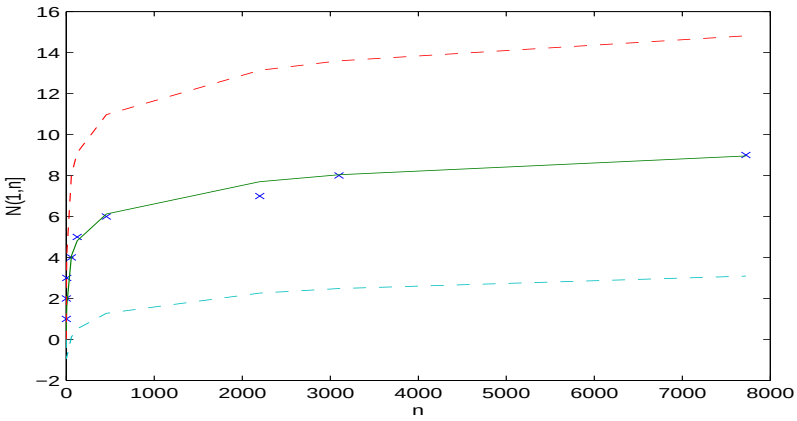


Figure 6.4. The number of records $N(1, n], n = 10^4$ from Times 3 map. The solid line (middle) is $\ln n$. The dashed lines are the 95% confidence bands.

6.3.2. Intermittent map

Similar to the previous maps we apply Theorem 6.2 to the Intermittent map with $\beta = 0.1, 0.5$ and 0.9. See Figures 6.5, 6.6 and 6.7. The results suggest that for the small value of β the map behaves as IID observations. Similar behaviour suggested by the H exponent value was observed in Chapter 5.

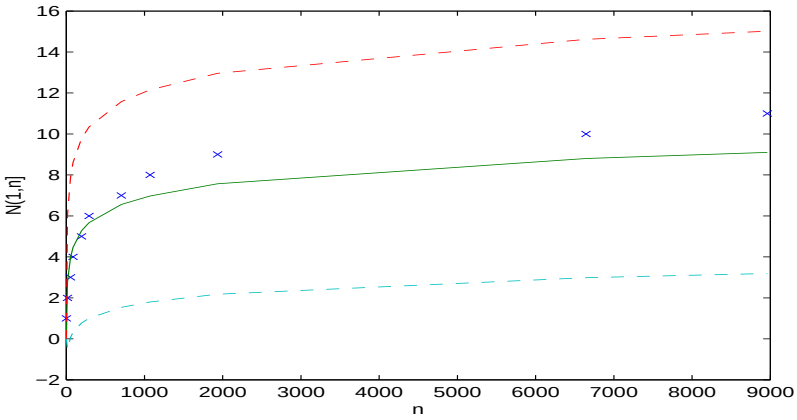


Figure 6.5. The number of records $N(1, n], n = 10^4$ from Intermittent map with $\beta = 0.1$ .The solid line (middle) is $\ln n$. The dashed lines are the 95% confidence bands.

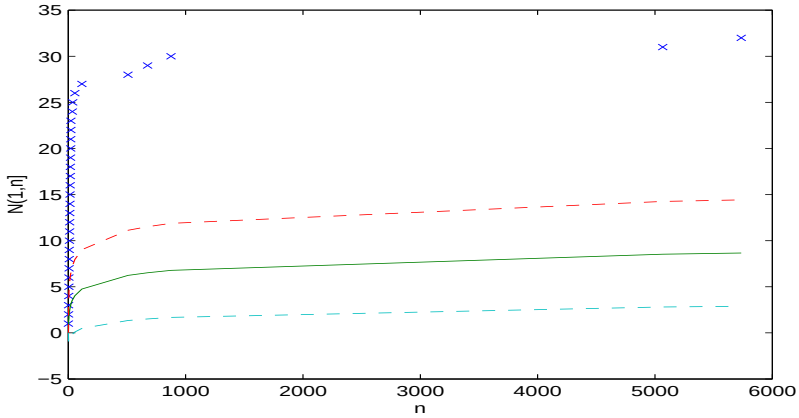


Figure 6.6. The number of records $N(1,n], n = 10^4$ from Intermittent map with $\beta = 0.5$ .The solid line (middle) is $\ln n$. The dashed lines are the 95% confidence bands.

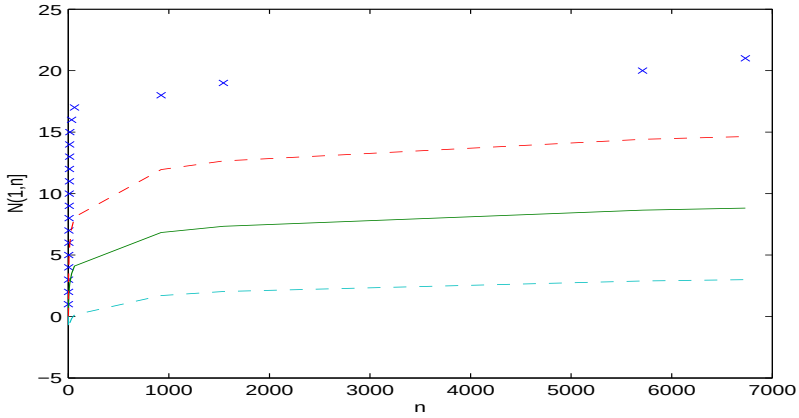


Figure 6.7. The number of records $N(1,n], n = 10^4$ from Intermittent map with $\beta = 0.9$ .The solid line (middle) is $\ln n$. The dashed lines are the 95% confidence bands.

6.4. Records of Financial Returns

As with the previous dynamical systems examples we apply Theorem 6.2 to the Financial returns data. We found that the observed number of records in the S&P500 returns index during the period January 1950 to September 2014, ($n = 16278$) is 16, see Figure 6.8. On the other hand, the observed number of records in the FTSE100 returns index during the period January 1984 to September 2014, ($n = 8010$) is 9, as we observed in the IID case where the expected number of records is described by equation(6.3), see Figure 6.9. However, the number of records in the S&P500 is larger than what is observed in the FTSE100 and the IID case. This agrees with the H exponent value observed in the S&P500 that is slightly larger than the FTSE100 as illustrated in Section 5.4. We also investigated the growth of record number in smaller sized observations ($n \approx 5000$) and the results were the same as the IID case in both indices. Our results are in agreement with the recent work of Wergen [90] who analysed record statistics of the return for the S&P500 returns data described in [92]. They found that the number of records observed differs from the IID case. In addition, they found that for short period it gives a closer result to that of IID observation.

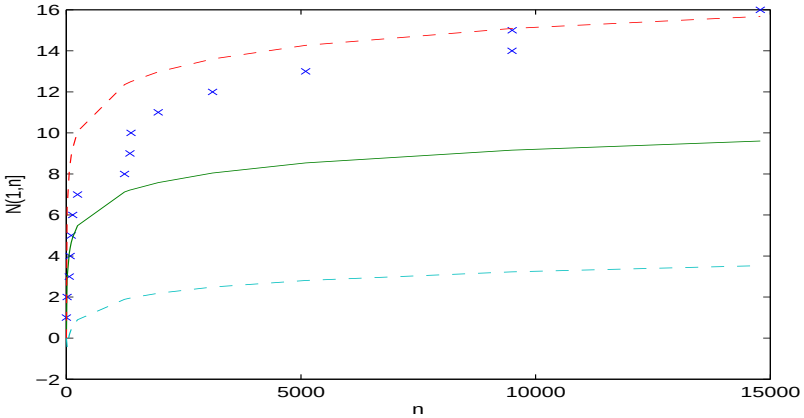


Figure 6.8. The number of records $N(1, n]$, $n = 16278$ of the S&P500 return index for the period from January 1950 to September 2014. The solid line (middle) is $\ln n$. The dashed lines are the 95% confidence bands.

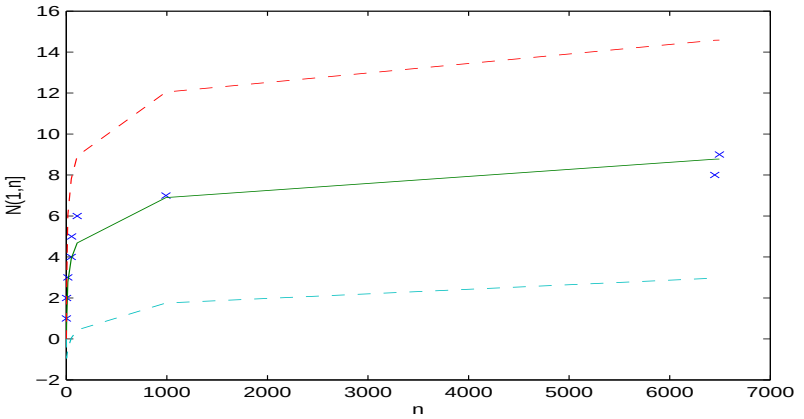


Figure 6.9. The number of records $N(1, n]$, $n = 8010$ of the FTSE100 return index for the period from January 1984 to September 2014. The solid line (middle) is $\ln n$. The dashed lines are the 95% confidence bands.

6.4.1. Records in random walk models

As we mentioned in the introduction, recently there has been an interest in studying record statistics in the random walk process because the RW model has a wide application in various subjects such as weather, economics and finance. As we discussed in Chapter 3 one model of the stock price index is the geometric random walk. In this section we will give a brief review of the results of record statistics in the RW process. We will consider an example of a simple RW with jump probability p and we will show how the number of records observed was affected by the probability p. In addition, we will apply Theorem 6.1 and, we will compare the result with the IID case.

Majumdar and Ziff [69] computed the probability of N records in n steps for the RW process defined by $X_0 = 0$, $X_i = X_{i-1} + \xi_i$ where ξ_i are IID RVs drawn from continuous and symmetric distribution $\phi(\xi)$ that is given by:

$$P(N_n, n) = \binom{2n - N + 1}{n} 2^{-2n+N-1}. \tag{6.6}$$

Where, $N \leq n$ and the expected number of records is given by:

$$E[N_n] \sim \frac{2}{\sqrt{\pi}}\sqrt{n} \quad \text{as } n \to \infty. \quad \text{(continuous jump)}. \tag{6.7}$$

They found that this result is independent of the jump distribution $\phi(\xi)$.

They also considered the case when the jump distribution is discrete such that $\xi \in \{+1, -1\}$ and they found that the expected number of records is $1/\sqrt{2}$ of that in the continuous case which is given by

$$E[N_n] \sim \frac{\sqrt{2n}}{\sqrt{\pi}} \quad \text{as } n \to \infty. \quad \text{(discrete jump)}. \tag{6.8}$$

Then Wergen et al. [92] followed Majumdar and Ziff [69] with Gaussian jump distribution but by adding a drift c in their model that is defined by $X_0 = 0, X_i = X_{i-1} + \xi_i + c$ where ξ_i are IID RVs drawn from Gaussian distribution. They computed the expected number of records for this model:

$$E[N_n] \sim \frac{2}{\sqrt{\pi}}\sqrt{n} + \frac{c}{\sqrt{2}\sigma} \text{ as } n \to \infty. \tag{6.9}$$

Moreover, their results were compared to sample data taken from the S&P500 index. They compare the number of records obtained from the data to that of biased random (RW with additional drift c); they determined the drift c and the standard deviation σ from the data with Monte Carlo simulation and found that the number of records is close to that of biased random walk. Furthermore, the number of records for the same data after detrending (subtract the fitted line from the logarithm stock prices) was compared to that of unbiased random walk model, and the result is close to that of unbiased random walk. A detailed

analysis of biased RW can be found in Majumdar et al. [68]. In addition, Wergen [91] computed record statistics in AR(1) with $X_0 = 0$, $X_i = \alpha X_{i-1} + \xi_i$ where, $\alpha \in (0, 1)$ and ξ_i is IID RVs from continuous symmetric jump distribution. Moreover, in the GARCH(1.1) with $X_0 = 0$ and $X_i = X_{i-1} + \sigma_i \xi_i$ where $\sigma_i^2 = \alpha_0 + \alpha_1 \xi_{i-1}^2 \sigma_{i-1}^2 + \beta_1 \sigma_{i-1}^2$, the result of these models were computed analytically and found that the AR(1) model gives similar results for the symmetric RW when α is close to 1, and when α is close to zero the mean number of records close to that of IID RVs. But for the GARCH(1,1) model with Gaussian distribution the expected number of records is close to that obtained from symmetric random walk. Furthermore, for heavy tailed distribution the number of records decreases when the tail becomes heavier but independent of the estimated parameters $(\alpha, \alpha_0, \beta)$. Both models do not describe the record statistics correctly in comparison with the data observed.

RW example:

Here we give an example of record statistics in a simple discrete time RW process that is given by:

$$\begin{aligned} X_0 &= 0 \\ X_n &= \xi_1 + \xi_2 + ... + \xi_n, \end{aligned} \tag{6.10}$$

Where, $\xi_i \in \{-1, +1\}$ and $P(\xi_i = +1) = p; P(\xi_i = -1) = q = (1 - p), 0 < p < 1$.
When, $p = q = \frac{1}{2}$ refer to unbiased(drift-less) RW and when $p \neq \frac{1}{2}$ to biased RW (with drift).
For unbiased RW we have $E[\xi_i] = 0$, this corresponds to $p = 0.5$.
Since,

$$X_n = \sum_{i=1}^{n} \xi_i$$

and for each ξ_i we have

$$E[\xi_i] = (+1)p + (-1)q = 2p - 1.$$

Then,

$$E[X_n] = \sum_{i=1}^{n} E[\xi_i] = n(2p - 1).$$

Thus, $E[X_n] = 0$ when $p = 1/2$.

For this example we compute the average number of records over 100 paths of equation (6.10) for $n = 10^3$ steps at different values of p, $p \in (0, 1)$. The result is illustrated in Figure 6.10. We found that the $E[N_n]$ when $p = 0.5$ grows as $\sqrt{n}$; this agrees with the result obtained by Majumdar and Ziff [69] for unbiased random walk. We also observed that for large value of p close to one we have

$$\frac{E[N_n]}{n} \to 1,$$

and for small value of p close to zero we have

$$\frac{E[N_n]}{n} \to 0.$$

Furthermore, in Figure 6.11 we investigate the rescaled mean number of strong records $E[N_n]/\sqrt{n}$ as a function of p at different steps length $n = 10^3$, 5000 and 10^4. We found that for $p > 0.5$ the rescaled mean number of records increases as we increase n. For $p = 0.5$ (unbiased RW) we found that $E[N_n]/\sqrt{n} = 0.7937$ when $n = 1000$, $E[N_n]/\sqrt{n} = 0.7938$ when $n = 5000$ and $E[N_n]/\sqrt{n} = 0.7877$ when $n = 10000$. In all considered cases the $E[N_n]/\sqrt{n}$ is close to the result obtained by Majumdar and Ziff [69] that is given by equation (6.8) ($(E[N_n]/\sqrt{n}) \sim 0.7979$).

We also investigate the limit distribution of record times for one sample path Theorem 6.1 as illustrated in Figure 6.12. We do not observe that $\ln N_n/n$ converges to one and it seems to converge to zero as $n \to \infty$. We also observe consecutive record times illustrated in the third Figure of Figure 6.12.

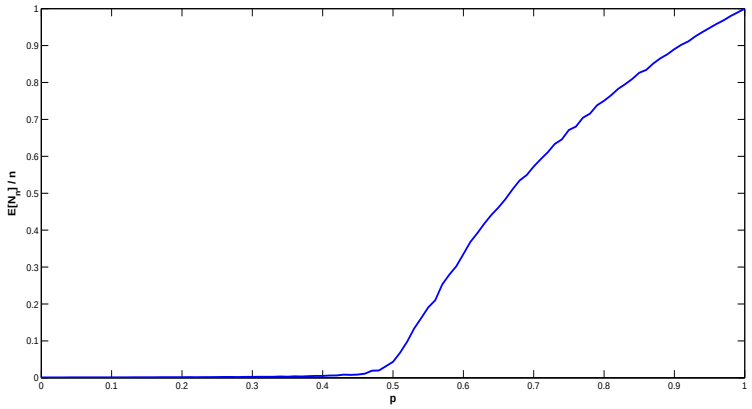


Figure 6.10. The mean number of records to the number of steps $n = 10^3$ as a function of $p \in (0, 1)$ for a biased random walk. The average is computed over 10^2 samples.

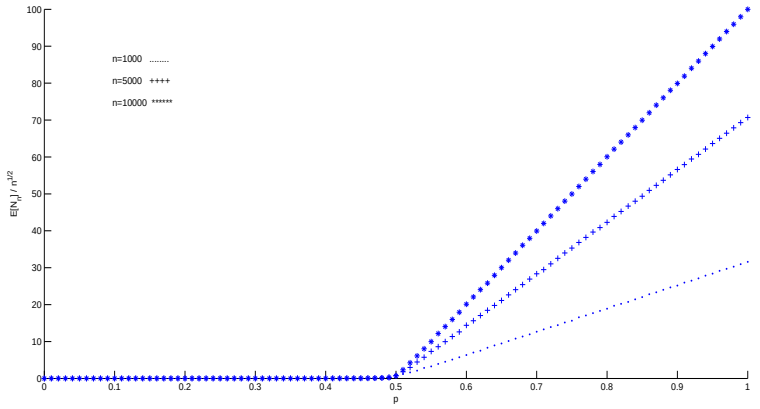


Figure 6.11. Rescaled mean number of records $E[N_n]/\sqrt{n}$ with different number of steps $(.)n = 1000$, $(+)n = 5000$ and $(*)n = 10000$ as a function of $p \in (0,1)$ (biased RW). For each n the average is computed over 10^2 samples. The $E[N_n]/\sqrt{n}$ with $p = 0.5$ for $n = 1000$ is $= 0.7937$, for $n = 5000$ is $= 0.7938$ and for $n = 10000$ is $= 0.7877$.

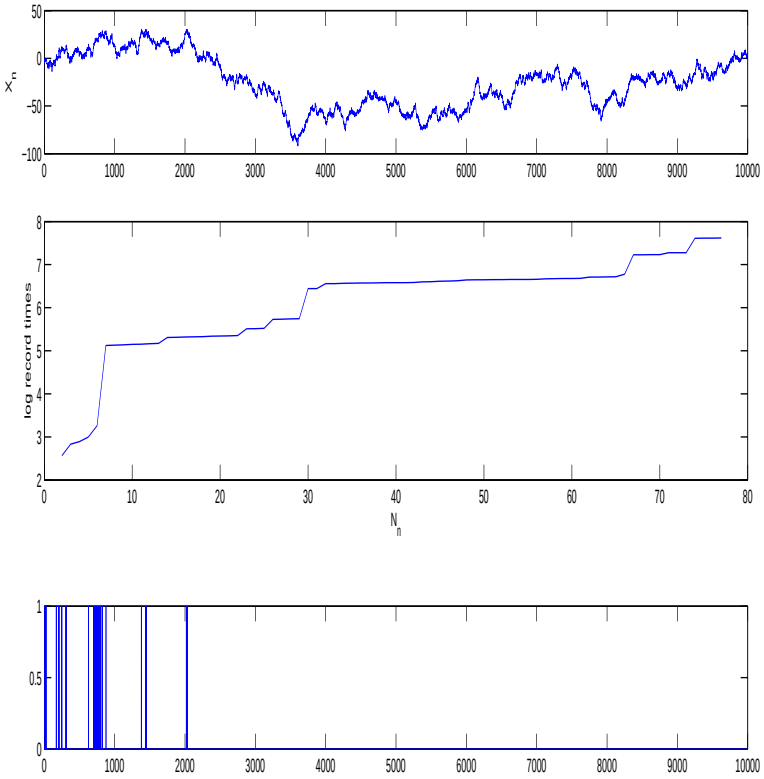


Figure 6.12. Growth of record times for a random walk. (Top) Random walk simulation. (Middle) Growth of record time as a function of the number of records observed. (Bottom) An indicator function that gives 1 if we observe a record and zero otherwise.

6.4.2. Records in Markov Chain models

In this section we consider a simple Markov Chain process (ξ_i) where, ξ_i is weakly dependent. As in the previous section, we will show how $E[N_n]$ of records vary with the probability p.

Definition 6.2 *A Markov Chain $\{\xi_n\}$ is a stochastic process in which the future is independent of the past, but depends on the present. With Markov property*

$$P(\xi_{n+1} = j|\xi_n = i, \xi_{n-1} = i_{n-1}, ..., \xi_0 = i_0) = P(\xi_{n+1} = j|\xi_n = i)$$

$$P(\xi_{n+1} = j|\xi_n = i) = T_{ij}$$

where, T_{ij} is known as the transition matrix of the chain.

This model can be described by the following diagram and equation:

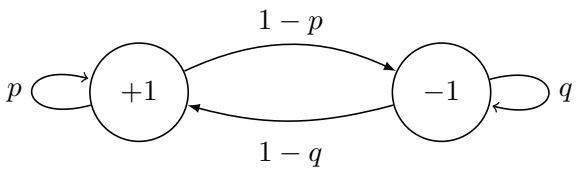


Figure 6.13. Transition diagram.

$$\begin{aligned} X_0 &= 0 \\ X_n &= \xi_1 + \xi_2 + ... + \xi_n, \end{aligned} \tag{6.11}$$

where $\xi_i \in \{-1, +1\}$ with initial distribution $P^{(0)} = (P(\xi_1 = +1)\ P(\xi_1 = -1)) = (1/2\ 1/2)$ and

$$P(\xi_{i+1} = +1/\xi_i = +1) = p$$

$$P(\xi_{i+1} = -1/\xi_i = +1) = 1 - p$$

$$P(\xi_{i+1} = +1/\xi_i = -1) = 1 - q$$

$$P(\xi_{i+1} = -1/\xi_i = -1) = q$$

or $T = \begin{bmatrix} p & 1-p \\ 1-q & q \end{bmatrix}$.

In this example we attempt to analyse the behaviour of the expected number of records to the number of steps n as a function of p and q. The idea here is similar to that for the biased RW. We take the average over 100 sample paths of equation (6.11) with different values of p and q for, $n = 10^3$ steps. The result is illustrated in Figure 6.14.

We observed that

$$\frac{E[N_n]}{n} \approx C_1/\sqrt{n} \text{ when } p = q = 0.5.$$

This result is close to that of records observed in the unbiased RW. We also observed that for large p, q close to one we have

$$\frac{E[N_n]}{n} \to 1,$$

and for small value of p, q close to zero we have

$$\frac{E[N_n]}{n} \to 0.$$

This indicates that the number of records depends on the choice of p and q.

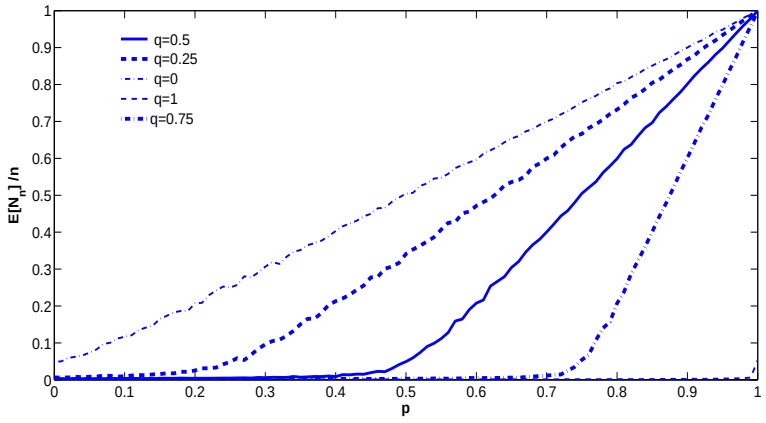


Figure 6.14. The mean number of records to the number of steps n as a function of p for different values of $q = 0, 0.25, 0.5, 0.75$ and 1 for, $n = 10^3$.

6.4.3. Random walk models with dependent increments

In this section we investigate the behaviour of the expected number of records in random walk with dependent increments or jumps.

First we consider the case of the Markov Chain process but with a $p = q$ unbiased Markov Chain. For $p = q$ we have that

$$\lim_{n\to\infty} E(\xi_n) = 0.$$

This can be illustrated by the following:

$$E(\xi_n) = 1P_1^{(n)} + (-1)P_2^{(n)},$$

where $P^{(n)} = \left(P_1^{(n)} \; P_2^{(n)}\right) = P^{(0)}T^n$ is the distribution over the states after n steps. The eigenvalues of T are 1 and a value $|\lambda| < 1$. Hence, for any vector $P^{(0)} \in \mathbb{R}^2$ we have $P^{(0)}T^n \to \tilde{P}$ with $\tilde{P}T \to \tilde{P}$ and the convergence rate of $|P^{(0)}T^n - \tilde{P}|$ is of the order $|\lambda|^n$. In this case we have $\tilde{P} = (1/2,\ 1/2)$.

See for example [59].

For this process we compute the rescaled expected number of records $E[N_n]/\sqrt{n}$ with different initial values and for each initial value we take $p = q = 0.1, 0.3, 0.45, 0.01$ and 0.49 where, $n = 10^4$ steps and we compute the average number of records over 10^2 samples. The result is illustrated in the following Tables. We found that $E[N_n]/\sqrt{n}$ is independent of the initial value. For strong records we observed that $E[N_n]/\sqrt{n} \sim \sqrt{2/\pi}$. This result is similar to that observed in the unbiased RW (discrete jump). Whereas, we found that the expected number of weak records is two times that of the strong records.

P^0	$E[N_n]/\sqrt{n}$				
p	0.1	0.3	0.45	0.01	0.49
$(\frac{1}{100}\ \frac{99}{100})$	1.5523	1.5802	1.5458	1.5354	1.5915
$(\frac{4}{10}\ \frac{6}{10})$	1.3109	1.7542	1.755	1.6573	1.6085
$(\frac{6}{10}\ \frac{4}{10})$	1.434	1.5455	1.5874	1.6212	1.5915
$(\frac{1}{10}\ \frac{9}{10})$	1.3661	1.6888	1.5725	1.7082	1.6085
$(\frac{9}{10}\ \frac{1}{10})$	1.9554	1.5455	1.7208	1.7783	1.4691

Table 6.1. Weak Record: unbiased Markov Chain model with different initial values For each initial value we computed the Rescaled mean number of records $E[N_n]/\sqrt{n}$ at $p = 0.1, 0.3, 0.45, 0.01$ and 0.49. with $n = 10^4$ steps, the average number of records is computed over 10^2 samples.

P^0	$E[N_n]/\sqrt{n}$				
p	0.1	0.3	0.45	0.01	0.49
$(\frac{1}{100}\ \frac{99}{100})$	0.7830	0.7795	0.7259	0.7475	0.7786
$(\frac{4}{10}\ \frac{6}{10})$	0.8262	0.7581	0.8027	0.7838	0.8464
$(\frac{6}{10}\ \frac{4}{10})$	0.8216	0.8021	0.8865	0.8078	0.7413
$(\frac{1}{10}\ \frac{9}{10})$	0.7899	0.7756	0.7951	0.7137	0.7823
$(\frac{9}{10}\ \frac{1}{10})$	0.7449	0.7502	0.8717	0.7539	0.7377

Table 6.2. Strong record: unbiased Markov Chain model with different initial values For each initial value we computed the Rescaled mean number of records $E[N_n]/\sqrt{n}$ at $p = 0.1, 0.3, 0.45, 0.01$ and 0.49. with $n = 10^4$ steps, the average number of records is computed over 10^2 samples.

Now we consider the process (ξ_i) where, ξ_i is generated by the Times 3 map and the intermittent map with different values of β. We examine the behaviour of the following process.

$$X_n = \xi_1 + \xi_2 + ... + \xi_n, \tag{6.12}$$

where,

$$\phi(x) = \begin{cases} +1 & \text{if } x \leq 1/2 \\ -1 & \text{if } x > 1/2. \end{cases} \tag{6.13}$$

Here $\xi_i = \phi(f^{i-1}x)$ where, f can be considered as the Time 3 map defined by equation (3.6) or the Intermittent map defined by equation (3.12).

We found that the number of records in this process with increments generated by the Times 3 map and Intermittent map with β close to zero, is slightly higher than the unbiased RW. In addition, for the process with increment generated by Intermittent map the number of records increases as β increases.

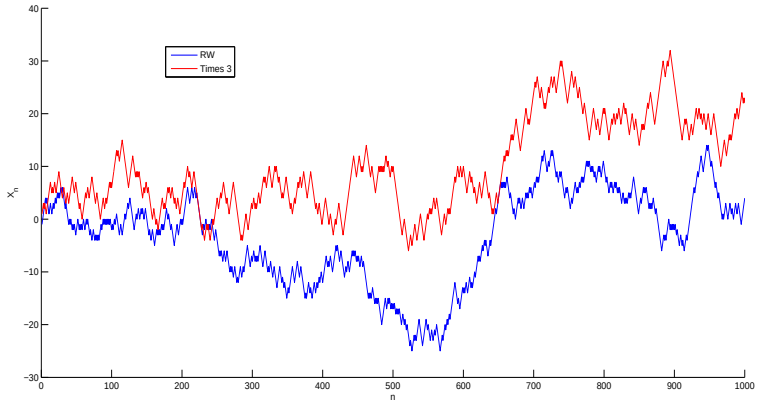


Figure 6.15. (blue) A numerical simulation of unbiased RW with $(n = 1000)$.(Red) A numerical simulation of RW with increments generated by Times 3 map with $(n = 1000)$.

Now let

$$\xi_i = \begin{cases} +1 - \mu & \text{if } x \leq 1/2 \\ -1 - \mu & \text{if } x > 1/2. \end{cases} \tag{6.14}$$

Where, μ is the mean of ξ:

$$\mu = \frac{1}{n} \sum_{i=0}^{N-1} \phi(f^{i-1}x).$$

In our application we let $n \to \infty$ such that $(n = 10^6)$. We computed μ for the Times 3 map and the Intermittent map with different value of β as illustrated in the following Table. For the Times 3 map we found that $\mu \to 0$ as n becomes large.

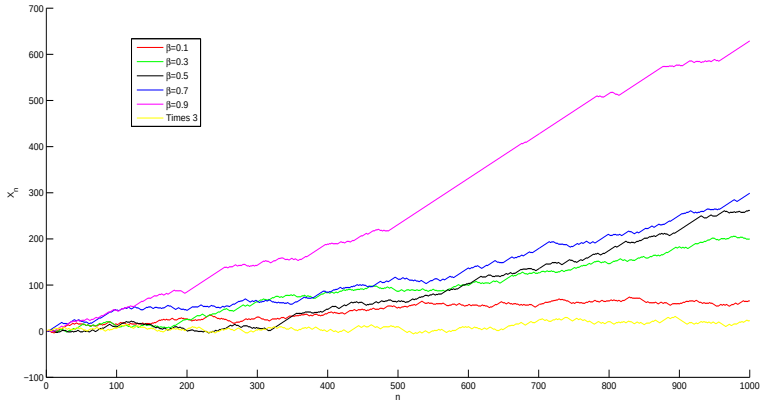


Figure 6.16. (yellow) A numerical simulation of RW with increments generated by Times 3 map with ($n = 1000$). A numerical simulation of RW with increments generated by Intermittent map with (n=1000) for different values of β (red) corresponds to $\beta = 0.1$, (green) corresponds to $\beta = 0.3$, (black) corresponds to $\beta = 0.5$, (blue) corresponds to $\beta = 0.7$ and (pink) corresponds to $\beta = 0.9$.

β	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
μ	0.044	0.0933	0.1435	0.2046	0.274	0.3528	0.4488	0.5354	0.6406

Table 6.3. In this table we computed $\mu = \frac{1}{n}\sum_{i=0}^{N-1}\phi(f^{i-1}x)$ for the Intermittent map for different value of β for typical x initial value and ($n = 10^6$).

We found that

$$E(\xi_i) = E(\phi) - \mu = 0.$$

Moreover, we observed that the number of records grows as $\sqrt{n}$ for the RW process with increments generated by the times 3 map and the Intermittent map with (no μ) as given by equation 6.14. This result is close to the unbiased RW defined by Majumdar and Ziff [69]. Consequently we expected the following results to hold about the mean record number $E(N_n)$ up to n steps for the following process as $n \to \infty$.

- Unbiased Markov Chain process ($p = q$) is:

$$\frac{E[N_n]}{\sqrt{n}} \approx C_1.$$

- Random walk process with dependent increments generated by the Times 3 map is:

$$\frac{E[N_n]}{\sqrt{n}} \approx C_2.$$

- Random walk process with dependent increments generated by the Intermittent map (no μ) is:

$$\frac{E[N_n]}{\sqrt{n}} \approx C_{3,\beta}.$$

For the above process we estimated numerically the rescaled expected number of strong and weak records for $n = 10^5$ steps over 10^2 sample paths. The results are illustrated in the following Tables.

We found that for unbiased Markov Chain process the rescaled number of strong records is close to that of rescaled number of strong records in unbiased RW. For weak records it seems to be two times that of strong records. For random walk process with dependent increments generated by the Times 3 map it is slightly higher than those of unbiased RW and Markov chain models. For the random walk model with increments generated by the intermittent map we observed that as we increase β the rescaled number of strong and weak records increase. We also examine the rescaled number of strong and weak records for the random walk process with increments generated by the Gaussian distribution(Majumdar and Ziff [69] model) see equation (6.7). We found that $\frac{E[N_n]}{\sqrt{n}} \approx 1.1255$. Here, we should point out that there is no clustering effect in our analysis of the expected number of records of the RW models with increments generated by the Times 3 map and the intermittent map.

Model	$\frac{E[N_n]}{\sqrt{n}}$
Unbiased RW with $\xi_i \in \{-1, 1\}$	1.5573
Unbiased Markov Chain	1.6268
RW with ξ_i generated from Gaussian distribution	1.1255
RW with ξ_i generated from the Times 3 map	1.6809
RW with ξ_i generated from the Intermittent map	
$\beta = 0.1$	1.2061
$\beta = 0.3$	1.8262
$\beta = 0.5$	2.4803
$\beta = 0.7$	6.3568
$\beta = 0.9$	31.6046

Table 6.4. Weak Record : The estimated rescaled mean number of records $\frac{E[N_n]}{\sqrt{n}}$ of the RW models for typical x initial value and $(n = 10^5)$ over 10^2 samples.

Model	$\frac{E[N_n]}{\sqrt{n}}$
Unbiased RW with $\xi_i \in \{-1, 1\}$	0.7867
Unbiased Markov Chain	0.8011
RW with ξ_i generated from Gaussian distribution	1.1257
RW with ξ_i generated from the Times 3 map	1.0810
RW with ξ_i generated from the Intermittent map	
$\beta = 0.1$	1.1381
$\beta = 0.3$	2.0390
$\beta = 0.5$	2.4803
$\beta = 0.7$	6.3568
$\beta = 0.9$	39.1714

Table 6.5. Strong Record: The estimated rescaled mean number of records $\frac{E[N_n]}{\sqrt{n}}$ of the RW models for typical x initial value and ($n = 10^5$) over 10^2 samples.

6.4.4. Records of Financial prices

In this section we apply the limit distribution of record times Theorem 6.1, to the S&P500 and FTSE100 price indices as illustrated in Figures 6.17, 6.18, 6.19 and 6.20. We do not observe that $\ln N_n/n$ converges to one and we observe consecutive record times. We found that the number of weak and strong records observed in the S&P500 price index for $n = 16279$ is 1145 and 1118 respectively. On the other hand, the number of weak and strong records observed in the FTSE100 price index for $n = 8011$ is 351 and 331 respectively.

The number of records observed in both indices is larger than the expected number of records in the RW with independent increments (unbiased RW) or RW with increments generated by short memory process. We have observed that the growth of record number of the S&P500 returns index seems not to grow as $\log n$. Whereas, the growth of record numbers of the FTSE100 returns index grow as $\log n$, see Section 6.4. Moreover, the H exponent obtained for the S&P500 returns index is higher than that of the FTSE100 returns index see, Section 5.4. We also found that the financial returns data exhibit clustering see, Section 4.6. All these facts might influence the number of records observed in the financial prices.

As mentioned in [90] the number of records in the financial data is sensitive to its length. For example if we want to compare our results we expected the following:

Short data might be described by the unbiased random walk or by random walk with increments generated by short memory such as the Times 3 map or by the intermittent map with no μ and $\beta < 0.5$. On the other hand long data might be modelled by random walk with increments generated by long memory process such as the Intermittent map.

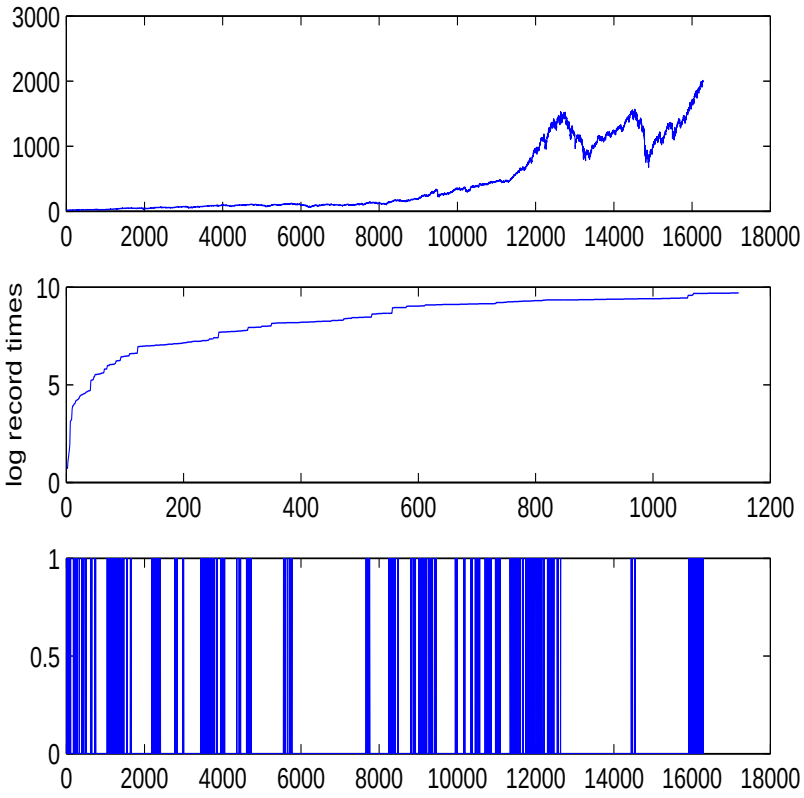


Figure 6.17. Weak record: (Top) The S&P500 price index for the period from January 1950 to September 2014. (Middle) Growth of record time as a function of the number of records observed. (Bottom) An indicator function that gives 1 if we observe a record and zero otherwise.

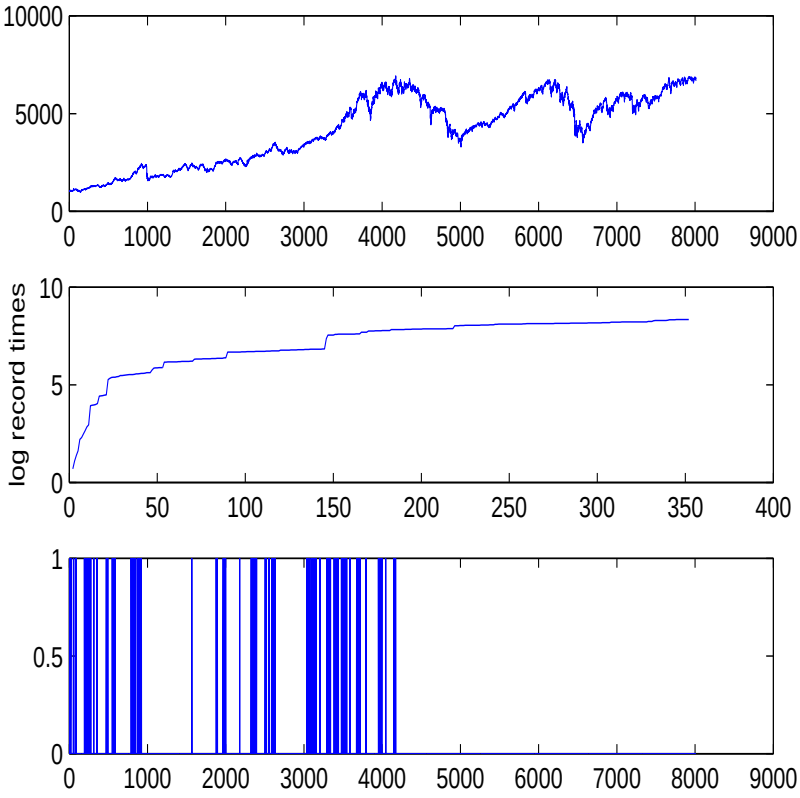


Figure 6.18. Weak record: (Top) The FTSE100 price index for the period from January 1984 to September 2014. (Middle) Growth of record time as a function of the number of records observed. (Bottom) An indicator function that gives 1 if we observe a record and zero otherwise.

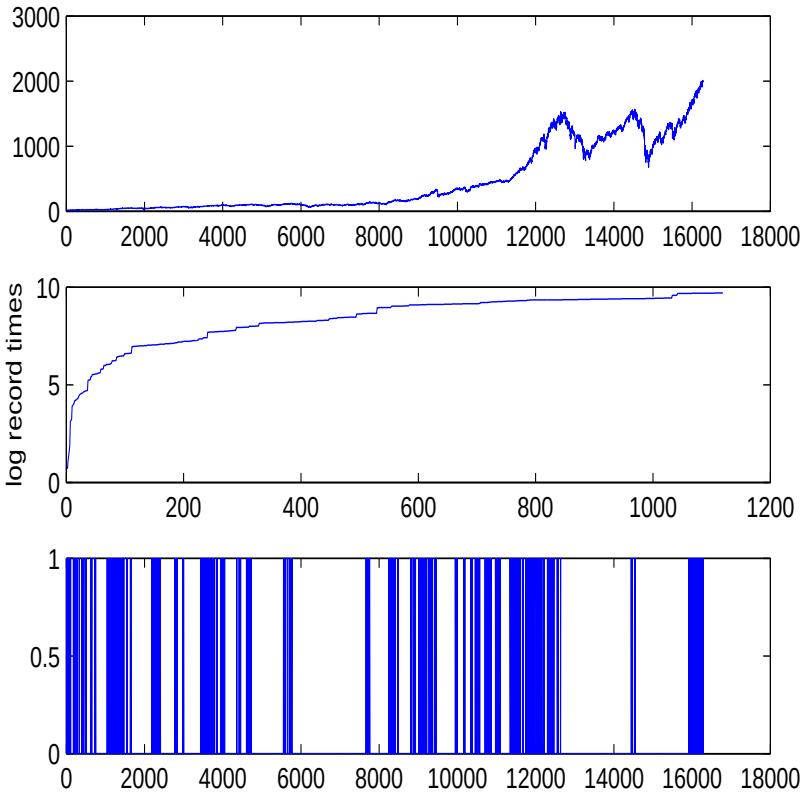


Figure 6.19. Strong record: (Top) The S&P500 price index for the period from January 1950 to September 2014. (Middle) Growth of record time as a function of the number of records observed. (Bottom) An indicator function that gives 1 if we observe a record and zero otherwise.

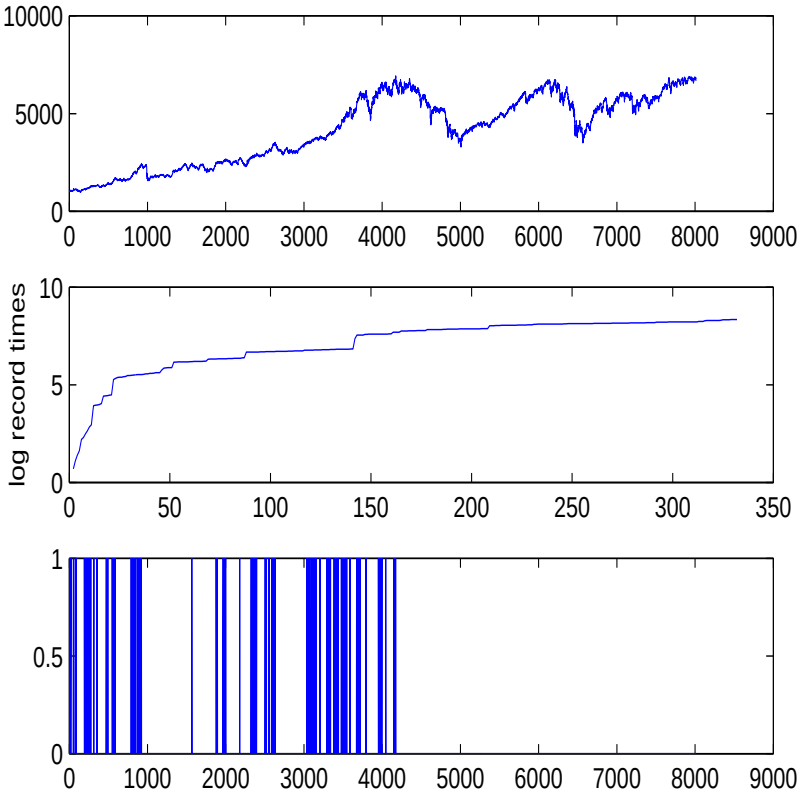


Figure 6.20. Strong record: (Top) The FTSE100 price index for the period from January 1984 to September 2014. (Middle) Growth of record time as a function of the number of records observed. (Bottom) An indicator function that gives 1 if we observe a record and zero otherwise.

7. Conclusion

This thesis is devoted to understand extremes and clustering in dynamical system models and financial returns data. In Chapter 2, we started by giving a review of the extreme value distributions that govern the statistics of IID RVs, and for a time series of observations on a dynamical system. We also discussed under which conditions the distributions of extremes have similar behaviour to that of IID observations, especially for dependent processes. We also discussed the rate of convergence to an extreme value distributions in IID case, and also in the dynamical system case building upon the recent study of Holland and Nicol [53]. In particular we showed how the convergence rate depended on the form of the observation function.

In Chapter 3 we presented a numerical study of modelling extremes in dynamical system examples, and in the S&P500 and FTSE100 returns indices. For modelling extremes we used the block maxima and the excess over threshold methods. Numerical results for the dynamical system examples agreed with the theoretical optimal block size and number described recently by Holland and Nicol [53] under some mixing and recurrence conditions. This assumes the system has an absolutely continuous invariant measure, and moreover the convergence rate depended on the functional form of the observable function. For the Times 3 map we showed how the convergence rate depended on the form of the observable function. Moreover, we found that the estimated tail index of the financial returns is of Type II distribution and this corresponds to the observable $\phi_2(x) = (\text{dist}(x, \tilde{x}))^{-\frac{1}{\alpha}}$ in our dynamical system examples.

In Chapter 4, we studied clustering in the dynamical systems examples and in the financial returns data by estimating the extremal index value. We found that our numerical results agreed with the theoretical extremal index described recently in [41] for a range of threshold values. Moreover our analysis of the financial returns data seems to exhibit clustering.

In Chapter 5, we measured dependency in the dynamical systems examples and in the financial returns by using the rescaled range analysis to estimate the H exponent. Based on our results it is difficult to distinguish between independent and short memory process. We also found that the H exponent in both returns indices was around 0.5 and it was slightly higher for the S&P500 index.

In Chapter 6, we have discussed the limit distribution of record times and number in IID RVs as given in [26]. In IID observations the number of records grow as $\ln n$. Similar behaviour is observed in dynamical system examples such as Times 3 map and the intermittent map as $\beta \to 0$. Hence we may say that this behaviour holds not only in IID observations but also in short memory process. This result can be used as a test of long

range dependency in the data. In addition, we showed that in the Markov chain and random walk process the mean number of records depends on the choice of the probability parameters. We also analysed the behaviour of record number in our returns data. We found that the number of records in the FTSE100 returns is close to that observed in the IID observations. Whereas, in the S&P500 returns the observed number is slightly higher. Furthermore, our analysis of the financial prices showed that the number of records observed is larger than that observed in the unbiased random walk. This result is also noticed in the recent work of Wergen et al.[92] and Wergen [90]. A question we might ask is, can the random walk model with dependent returns (dynamical systems examples) explain the behaviour of the financial prices? The RW model with increments generated by the intermittent map with the suitable parameter β might agree with the number of records observed in financial prices. An interesting problem to consider in the future is the volatility clustering, Large changes in the prices followed by large changes (the prices show clustering). The volatility clustering is related to the autocorrelation function. As discussed for example in, [23], [17], [18] and [88] the absolute returns or the square returns have a slow decay of correlation and are used as measure of volatility clustering. We plan to examine this fact in our future study.

Based on our analysis of extremes in dynamical systems examples and under some conditions we can conclude the following:

- In modelling extremes we can control the estimated value of the tail index based on the observable function $\phi_i(x), i = 1, 2, 3$ as described in Chapter 2.
- Clustering depends on the periodicity of the systems. For example to guarantee that the system has extremal index 1, choose $\tilde{x}$ as a non repelling periodic point. Here $\tilde{x}$ corresponds to the point where the observable is maximized.
- Using the intermittent map short memory or long memory processes can be generated and we might control the expected number of records.

A. Appendix

In Chapter 2 we called Σ as a $\sigma - algebra$ that is defined as the following:

Definition A.1 *[89] Suppose that X is a set and Σ is a collection of subset of X. Σ is called a $\sigma - algebra$ if :*

1. $\emptyset \in \Sigma$,
2. *if* $B \in \Sigma$ *then* $X/B \in \Sigma$,
3. *if* $B_n \in \Sigma, n \geq 1$ *then* $\bigcup_{n=1}^{\infty} B_n \in \Sigma$.

Definition A.2 $(a.e)$ *We say that a property holds almost everywhere* $(a.e)$ *if the set of points on which the property does not hold has measure zero.*

See, [26].

Definition A.3 *A function f defined on X is said to be integrable if $\int |f| \, d\nu < \infty$. The space of integrable functions is defined to be*
$L^1(X, \mathcal{B}, \nu) = (f : X \to \mathbb{R} |\ f$ *is measurable,* $\int |f| \, d\nu < \infty)$.
For $p \geq 1$ we defined the L^p space to be:
$L^p(X, \mathcal{B}, \nu) = (f : X \to \mathbb{R} |\ f$ *is measurable,* $\int |f|^p \, d\nu < \infty)$.

See, [89].

Definition A.4 *(Lipschitz continuity) We say that $f : [a, b] \to \mathbb{R}$ is Lipschitz continuity if there exist a constant C such that*

$$|f(x_1) - f(x_2)| \leq C|x_1 - x_2|$$

for all $x_1, x_2 \in [a, b]$.

See, [27].

- $\| f \|_p = (\int |f|^p d\nu)^{\frac{1}{p}} < \infty$.
- $\| f \|_\infty = \lim_{p \to \infty} \| f \|_p$.
- $\| f \|_\infty = \inf(c \geq 0 : |f(x) \leq c$ for almost every x).
- $\| f \|_{Lip} = \sup \left(\frac{|f(x_1) - f(x_2)|}{|x_1 - x_2|} : x_1 \neq x_2 \right)$.

See, [89].

B. Appendix

The MATLAB codes used in this thesis is presented below:

B.1. Chapter 3 Matlab codes:

```
% Block Maxima function for the times 3 map %
function  [shape,CI,CII]  = BMT3(p,q)
%p=block size.
%q=block number.
%script for times3 map
u=sqrt(2);%initial value
n=10^8;%n=p*q;
for i=1:n;
u=mod(3*u,1);
y(i)=u; %value at step i
end
xtilde(1,1)=0%0.8491;%rand(1,1)
alpha=2;
C=0;
z=(abs(y-xtilde)).^(-1/alpha);
%z=-log(abs(y-xtilde));
%z=C-(abs(y-xtilde)).^(1/alpha);
x=reshape(z,p,q); %reshape(X,p,q) returns the p-by-q matrix whose  are taken
     columnwise from X.


    for i=1:q
        m(i)=max(x(1:p,i));%returns the maximum values.
    end
 param = gevfit(m);   % gevfit(m) returns maximum likelihood estimates of
    the parameters.
 shape  = param(1);   % Tail or shape index parameter of the GEV
    distribution.
 [param,paramCI] = gevfit(m);
shape
kCI  = paramCI(:,1); %confidence interval of the shape parameter.
CI=kCI(1,1);
CII=kCI(2,1);
end
```

```
%Block Maxima script for the Times 3 map %
for v=2:6
    for w=1:2
    n=10^8;
```

```
    q=(10^v)/(2^w);
    p=n/q;
[shape,CI,CII]=BMT3(p,q);
hold on
plot(log10(q),shape,'*',log10(q),CI,'x',log10(q),CII,'x');
hold off
    end
end
for m=2:5
for k=0:3
    q=(10^m)*2^k ;
    n=10^8;
    p=n/q;
[shape,CI,CII]=BMT3(p,q);
hold on
plot(log10(q),shape,'*',log10(q),CI,'x',log10(q),CII,'x');
hold off
end
end
```

```
% Block Maxima function for the HENON map in one-dimensional%
function  [shape,CI,CII]  = BMHENON1D(p,q)
%p=block size.
%q=block number.
a=1.4;
b=0.3;
x(1)= 0.1386;%initial value rand(1,2).
y(1)= 0.1493;%
n=10^6;
alpha=4;
C=0;
t=10^8;%number of iterations
n=10^6;%number of observations
for i=1:t;
    x(i+1)=1-a*(x(i))^2+y(i);
    y(i+1)=b*x(i);
end
 xtilde=x(t); %or rand(1,1)but on the attractor area
for j=1:n; %The observabel functions.
    z(j)=-log(abs(x(j)-xtilde)); %%Type 1
    %z(j)=(abs(x(j)-xtilde)).^(-1/alpha); %% Type 2
    %z(j)=C -(abs(x(j)-xtilde)).^(1/alpha); %% Type 3
end
f=reshape(z,p,q);%reshape(X,p,q) returns the p-by-q matrix whose elements
    are taken columnwise from X.
    for i=1:q
        m(i)=max(f(1:p,i));%returns the maximum values.
    end
 param = gevfit(m);   %param = gevfit(m) returns maximum likelihood
    estimates of the parameters.
 shape  = param(1);   % Tail or shape index parameter of the GEV distrib
 [param,paramCI] = gevfit(m);

shape
kCI  = paramCI(:,1);%confidence interval of the shape parameter.
```

```
CI=kCI(1,1);
CII=kCI(2,1);
end
```

```
% Block Maxima script for the HENON map in one-dimensional%
for m=2:3
for k=0:3
    q=(10^m)*2^k;
    n=10^6;
    p=n/q;
[shape,CI,CII]=BMHENON1D(p,q);
hold on
plot(log10(q),shape,'*',log10(q),CI,'x',log10(q),CII,'x');
hold off
    end
end
 for b=0:1
  for   c=3:5
     q=(10^b)*(5^c);
     n=10^6;
     p=n/q;

     [shape,CI,CII]=BMHENON1D(p,q);
 hold on
 plot(log10(q),shape,'*',log10(q),CI,'x',log10(q),CII,'x');
 hold off
 end
 end
for v=2:5
    for w=1:2

    n=10^6;
    q=(10^v)/(2^w);
    p=n/q;
    [shape,CI,CII]=BMHENON1D(p,q);
hold on
plot(log10(q),shape,'*',log10(q),CI,'x',log10(q),CII,'x');
hold off
end
end
q=1600;
n=10^6;
 p=n/q;
   [shape,CI,CII]=BMHENON1D(p,q);
hold on
plot(log10(q),shape,'*',log10(q),CI,'x',log10(q),CII,'x');
hold off
```

```
% Block Maxima function for the HENON map in two-dimensional%
function  [shape,CI,CII]  = BMHENON2D(p,q)
%p=block size.
%q=block number.
a=1.4;
b=0.3;
xo=[0.1386   0.1493] ; %rand(1,2)initial value
```

```
x(1)=xo(1);
y(1)=xo(2);
alpha=4;
C=0;
t=10^8;%number of iterations
n=10^6;%number of observations
for i=1:t;
    x(i+1)=1-a*(x(i))^2+y(i);
    y(i+1)=b*x(i);
end
 xtilde=[x(t) y(t)];
for j=1:n;%The observabel functions.
     z(j)= -0.5*log( (x(j)- xtilde(1)).^2+(y(j)- xtilde(2)).^2 ); %%Type 1
     %z(j)=sqrt( (x(j)- xtilde(1)).^2+(y(j)- xtilde(2)).^2 ).^(-1/alpha); %%
     Type 2
     %z(j)=C-sqrt( (x(j)- xtilde(1)).^2+(y(j)- xtilde(2)).^2 ).^(1/alpha);
    %% Type 3
end
f=reshape(z,p,q);%reshape(X,p,q) returns the p-by-q matrix whose elements
    are taken columnwise from X.

    for i=1:q
        m(i)=max(f(1:p,i));%returns the maximum values.
    end
 param = gevfit(m);   %param = gevfit(m) returns maximum likelihood
    estimates of the parameters.
 shape  = param(1);   % Tail or shape index parameter of the GEV distrib
 [param,paramCI] = gevfit(m);
shape
kCI  = paramCI(:,1);%confidence interval of the shape parameter.
CI=kCI(1,1);
CII=kCI(2,1);
end
```

```
% Block Maxima script for the HENON map in two-dimensional%%%%
for m=2:3
for k=0:3
    q=(10^m)*2^k
    n=10^6;
    p=n/q;
[shape,CI,CII]=BMHENON2D(p,q);
hold on
plot(log10(q),shape,'*',log10(q),CI,'x',log10(q),CII,'x');
hold off
    end
end
 for b=0:1
  for   c=3:5
      q=(10^b)*(5^c);
      n=10^6;
      p=n/q;

      [shape,CI,CII]=BMHENON2D(p,q);
 hold on
 plot(log10(q),shape,'*',log10(q),CI,'x',log10(q),CII,'x');
```

```
 hold off
 end
 end
for v=2:5
    for w=1:2

    n=10^6;
    q=(10^v)/(2^w);
    p=n/q;
    [shape,CI,CII]=BMHENON2D(p,q);
hold on
plot(log10(q),shape,'*',log10(q),CI,'x',log10(q),CII,'x');
end
end
q=1600;
n=10^6;
 p=n/q;
   [shape,CI,CII]=BMHENON2D(p,q);
hold on
plot(log10(q),shape,'*',log10(q),CI,'x',log10(q),CII,'x');
hold off
```

```
% Block Maxima function for the Intermittent map %
function  [shape,CI,CII]  = BMINT(p,q)
%p=block size.
%q=block number.
 n=10^8;
 beta=0.1;%\beta \in(0,1)
 y(1,1)=0.1712;%rand(1,1)initial value
for j=2:n;
    if y(j-1)<=0.5;
        y(j)=y(j-1)*(1+2^beta*y(j-1)^beta);
    else
        y(j)=2*y(j-1)-1;
    end
end
xtilde(1,1)= 0.8491;%rand(1,1)
alpha=3;% alpha>0;
C=0;
%z=(abs(y-xtilde)).^(-1/alpha);
 z=-log(abs(y-xtilde));
 %z=C -(abs(y-xtilde)).^(1/alpha);
f=reshape(z,p,q);%reshape(X,p,q) returns the p-by-q matrix whose elements
    are taken columnwise from X.
    for i=1:q
        m(i)=max(f(1:p,i));%returns the maximum values.
    end
 param = gevfit(m);%param = gevfit(m) returns maximum likelihood estimates
    of the parameters.
 shape  = param(1);   % Tail or shape index parameter k.
 [param,paramCI] = gevfit(m);
shape
kCI  = paramCI(:,1);%confidence interval of the shape paramete
CI=kCI(1,1);
CII=kCI(2,1);
```

```
end
```

```
%Block Maxima script for the Intermittent map %
for v=2:6
    for w=1:2
    n=10^8;
    q=(10^v)/(2^w);
    p=n/q;
[shape,CI,CII]=BMINT(p,q);
hold on
plot(log10(q),shape,'*',log10(q),CI,'x',log10(q),CII,'x');
hold off
    end
end
for m=2:5
for k=0:3
    q=(10^m)*2^k ;
    n=10^8;
    p=n/q;
[shape,CI,CII]=BMINT(p,q);
hold on
plot(log10(q),shape,'*',log10(q),CI,'x',log10(q),CII,'x');
hold off
end
end
```

```
%Threshold Exceedance(POT)for the Times3 map%
clear;
u= sqrt(2);%rand(1,1);
npts=5000;%number of observations
xtilde(1,1)= 0.8491;%rand(1,1)
alpha=2;
C=0;
for i=1:npts;
u=mod(3*u,1);
y(i)=u; %value at step i
    z(i)=-log(abs(y(i)-xtilde)); %%Type 1
    %z(i)=(abs(y(i)-xtilde)).^(-1/alpha); %% Type 2
    %z(i)=C -(abs(y(i)-xtilde)).^(1/alpha); %% Type 3
end
U=quantile(z,0.99)
%U=1.9;% Threshold (we estimated the threshold value from the Mean Excess
    script file).
f = z(z>U) - U;
length(f)
%compute the values that exceed the threshold value u.
%Fitting the GP Distribution Using Maximum Likelihood
param = gpfit(f); %gpfit(X) returns maximum likelihood estimates of the
    parameters
    %of the two-parameter generalized Pareto (GP) distribution given the
    data in X.
shape  = param(1)% Tail index (shape) parameter
scale  = param(2)   % Scale parameter
[F,zi] = ecdf(f);%[F,zi] = ecdf(f) returns the empirical cumulative
    distribution function (cdf),
```

```
                %F, evaluated at the points in zi,using the data in the
    vector f.
plot(zi,gpcdf(zi,shape,scale),'-');
hold on; stairs(zi,F,'r'); hold off;
legend('Fitted GP CDF','Empirical CDF','location','southeast');
[nll,acov] = gplike(param, f);
stdErr = sqrt(diag(acov))
```

```
% Mean Excess plot for the Henon map under the observable \phi_i,i=1,2,3.
clear;
u= sqrt(2);%initial value
npts=5000;
for i=1:npts;
u=mod(3*u,1); %times3 map
y(i)=u; %value at step i
end
xtilde=rand(1);
alpha=3;
C=0;
for i=1:npts;
    z(i)=-log(abs(y(i)-xtilde)); %%Type 1
    %z(i)=(abs(y(i)-xtilde)).^(-1/alpha); %% Type 2
    %z(i)=C -(abs(y(i)-xtilde)).^(1/alpha); %% Type 3
end
xmax=max(z);
 xmin=min(z);
for u=xmin:0.05:xmax
f = z(z>u) - u;
me=mean(f);
hold on
plot(u,me,'o')
hold off
end
```

```
%Threshold Exceedance(POT)for the HENON map%
clear;
a=1.4;
b=0.3;
x(1)=0;%initial value
y(1)=0;
npts=5000;%number of observations
xtilde=0.155;%rand(1,1)
alpha=3;
C=0;
for i=1:npts;
    x(i+1)=1-a*(x(i))^2+y(i);
    y(i+1)=b*x(i);
    %z(i)=-log(abs(x(i)-xtilde)); %%Type 1
    z(i)=(abs(x(i)-xtilde)).^(-1/alpha); %% Type 2
    %z(i)=c -(abs(x(i)-xtilde)).^(1/alpha); %% Type 3
end
U=2;% Threshold (we estimated the threshold value from the Mean Excess
    script file).
%U=quantile(z,.93)
f = z(z>U) - U; %compute the values that exceed the threshold value u.
```

```
length(f)

param = gpfit(f); %gpfit(X) returns maximum likelihood estimates of the
    parameters
    %of the two-parameter generalized Pareto (GP) distribution given the
    data in X.
shape  = param(1)% Tail index (shape) parameter
scale  = param(2)   % Scale parameter
[F,zi] = ecdf(f);%[F,zi] = ecdf(f) returns the empirical cumulative
    distribution function (cdf),
                %F, evaluated at the points in zi,using the data in the
    vector f.
plot(zi,gpcdf(zi,shape,scale),'-');
hold on; stairs(zi,F,'r'); hold off;
legend('Fitted GP CDF','Empirical CDF','location','southeast');
[nll,acov] = gplike(param, f);
stdErr = sqrt(diag(acov))
```

```
% Mean Excess plot for the Henon map under the observable \phi_i,i=1,2,3.
clear;
a=1.4;
b=0.3;
x(1)=0;%initial value.
y(1)=0;
npts=5000;
xtilde=0.155;%rand(1,1).
alpha=4;
C=0;
for i=1:npts;
    x(i+1)=1-a*(x(i))^2+y(i);
    y(i+1)=b*x(i);
   z(i)=-log(abs(x(i)-xtilde)); %%Type 1
    %z(i)=(abs(x(i)-xtilde)).^(-1/alpha); %% Type 2
    %z(i)=C -(abs(x(i)-xtilde)).^(1/alpha); %% Type 3
end
 xmax=max(z);
 xmin=min(z);
for u=xmin:0.05:xmax
f = z(z>u) - u;
me=mean(f);
hold on
plot(u,me,'o')
hold off
end
```

B.2. Chapter 4 Matlab codes:

```
%Estimation the Extremal index for the Times3 map%
function  [e4]  = exT3( U )
n=10^4; % number of observations.
u=sqrt(2); %initial value.
for i=1:n
u=mod(3*u,1); % times3 map.
x(i)=u;       %value at step i.
```

```
end
xtilde=0;%0.3
y=-log(abs(x-xtilde));
%Computing the Extremal index.
for j=1:n
if y(j)> U
S(j)=j; %The exceedance times.
end
end
S(S==0)=[];
N=length(S);%The number of observations exceeding threshold.
m=N-1;
for k=1:m
    T(k)=S(k+1)-S(k);%Interexceedance times
end
A=sum(T.^2);
Theta2=(2*sum(T)^2)/((N-1)*A);
for J=1:m
    t(J)=T(J)-1;
    tt(J)=(T(J)-1)*(T(J)-2);
end
Theta3=(2*(sum(t))^2)/((N-1)*sum(tt));
e2=Theta2;
e3=Theta3;
if (max(T) <= 2)
    e4= abs(min(1,e2));
elseif (max(T)> 2)
        e4=abs(min(1,e3));
end
end
```

```
%The Extremal index of the Times3 map at range of threshold values%
for U=0:0.005:7 % threshold range.
[e4]=exT3(U);
hold on
plot(U,e4,'r:','LineWidth',1);
hold off
end
```

```

% This function compute the extremal index of the Max-Auto-Regressive
% process
function [e4] = exmaxaut( U )

n=5000;% number of observations
theta=0.5;% theta \in (0,1]
w= rand(1,n).^(-1);
y(1,1)=w(1,1)./theta;

for i=2:n
    x(1,1)=y(1,1);
    x(1,i)=max((1-theta).*x(1,i-1),w(1,i));
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
for j=1:n

if x(j)> U

S(j)=j; %The exceedance times.
end
end
S(S==0)=[];
N=length(S);%The number of observations exceeding threshold.
m=N-1;
for k=1:m
    T(k)=S(k+1)-S(k);%Interexceedance times
end

A=sum(T.^2);

%Theta1=(2*m*(n^2))/((N^2)*A);

 Theta2=(2*sum(T)^2)/((N-1)*A);

for J=1:m
    t(J)=T(J)-1;
    tt(J)=(T(J)-1)*(T(J)-2);
end

Theta3=(2*(sum(t))^2)/((N-1)*sum(tt));

%e1=Theta1;
e2=Theta2;
e3=Theta3;

if (max(T) <= 2)

    e4= min(1,e2);

elseif (max(T)> 2)
        e4=min(1,e3);
end
end
```

```
% The Extremal index function of the HENON map - one dimensional%%%
function  [e4]  = exHENON1D( U )
n=10^4;% number of observations.
a=1.4;
b=0.3;
%Case(1.1)
% x(1)= 0;
% y(1)=0;
% xtilde= 0.155;%rand(1,1)
%Case(1.2)
 x(1)= 0.1386;% The initial value is x_0=(x(1),y(1))as rand(1,2).
 y(1)=0.1493;
 xtilde= -0.3051;%xtilde=x_t where for example (t=10^8)iteration number >> n
    .

```

```
for i=1:n;
    x(i+1)=1-a*(x(i))^2+y(i);
    y(i+1)=b*x(i);
    z(i)=-log(abs(x(i)-xtilde)); %%Type 1
end
%Computing the Extremal index.
for j=1:n
if z(j)> U
S(j)=j; %The exceedance times.
end
end
S(S==0)=[];
N=length(S);%The number of observations exceeding threshold.
m=N-1;
for k=1:m
    T(k)=S(k+1)-S(k);%Interexceedance times
end
A=sum(T.^2);
 Theta2=(2*sum(T)^2)/((N-1)*A);
for J=1:m
    t(J)=T(J)-1;
    tt(J)=(T(J)-1)*(T(J)-2);
end
Theta3=(2*(sum(t))^2)/((N-1)*sum(tt));
e2=Theta2;
e3=Theta3;
if (max(T) <= 2)
    e4= abs(min(1,e2));
elseif (max(T)> 2)
        e4=abs(min(1,e3));
end
end
```

```
%The Extremal index script of the Henon map one-dimensional at range of
    threshold values%
for U=0:0.005:7 % threshold range
[e4]=exHENON1D(U);
hold on
plot(U,e4,'r:','LineWidth',1);
hold off
end
```

```
% The Extremal index function of the HENON map two-dimensional%
function  [e4]  = exHENON2D( U )
a=1.4;
b=0.3;
xo=[0.1386   0.1493] ; %rand(1,2)initial value
x(1)=xo(1);
y(1)=xo(2);
alpha=4;
C=0;
t=10^8;%number of iterations
n=10^4;%number of observations
for i=1:t;
    x(i+1)=1-a*(x(i))^2+y(i);
```

```
    y(i+1)=b*x(i);
end
 xtilde=[x(t) y(t)];
%xtilde=[-0.3051   0.3016];
%The observabel function.
for j=1:n;
     z(j)= -0.5*log( (x(j)- xtilde(1)).^2+(y(j)- xtilde(2)).^2 ); %%Type 1
     %z(j)=sqrt( (x(j)- xtilde(1)).^2+(y(j)- xtilde(2)).^2 ).^(-1/alpha); %%
     Type 2
     %z(j)=C-sqrt( (x(j)- xtilde(1)).^2+(y(j)- xtilde(2)).^2 ).^(1/alpha);
    %% Type 3
end
%Computing the Extremal index.
for j=1:n
if z(j)> U
S(j)=j; %The exceedance times.
end
end
S(S==0)=[];
N=length(S);%The number of observations exceeding threshold.
m=N-1;
for k=1:m
    T(k)=S(k+1)-S(k);%Interexceedance times
end
A=sum(T.^2);
 Theta2=(2*sum(T)^2)/((N-1)*A);
for J=1:m
    t(J)=T(J)-1;
    tt(J)=(T(J)-1)*(T(J)-2);
end
Theta3=(2*(sum(t))^2)/((N-1)*sum(tt));
e2=Theta2;
e3=Theta3;
if (max(T) <= 2)
    e4= abs(min(1,e2));
elseif (max(T)> 2)
        e4=abs(min(1,e3));
end
end
```

```
%The Extremal index script of the Henon map two-dimensionalat range of
    threshold values%
for U=0:0.005:6 % threshold range
[e4]=exHENON2D(U);
hold on
plot(U,e4,'r:','LineWidth',1);
hold off
end
```

```
%Estimation the Extremal index for the Intermittent map%
function  [e4]  = exINT( U )
n=10^4;% number of observations.
beta=0.7;%\beta \in (0,1).
y(1,1)=  0.1712; %initial value ,rand(1,1).
for j=2:n;
```

```
    if y(j-1)<=0.5;
        y(j)=y(j-1)*(1+2^beta*y(j-1)^beta);
    else
        y(j)=2*y(j-1)-1;
    end
end
xtilde(1,1)= 0;%1,1/3.
z=-log(abs(y-xtilde));
%Computing the Extremal index.
for j=1:n
if z(j)> U
S(j)=j; %The exceedance times.
end
end
S(S==0)=[];
N=length(S);%The number of observations exceeding threshold.
m=N-1;
for k=1:m
    T(k)=S(k+1)-S(k);%Interexceedance times
end
A=sum(T.^2);
 Theta2=(2*sum(T)^2)/((N-1)*A);
for J=1:m
    t(J)=T(J)-1;
    tt(J)=(T(J)-1)*(T(J)-2);
end
Theta3=(2*(sum(t))^2)/((N-1)*sum(tt));
e2=Theta2;
e3=Theta3;
if (max(T) <= 2)
    e4= abs(min(1,e2));
elseif (max(T)> 2)
        e4=abs(min(1,e3));
end
end
```

```
%The Extremal index of the Intermittent map at range of threshold values.
for U=0:0.005:7 % threshold range
[e4]=exINT(U);
hold on
plot(U,e4,'r:','LineWidth',1);
hold off
end
```

```
%Estimation of the Extremal index by using the Block method for the Times 3
%map with xtilde=0.
%Here we choose p=q=sqrt(n), and we estimated the threshold value  from the
    quantile function.
clear;
%p=block size.
%q=block number.

u=sqrt(2);%initial value
q=1000;
p=q;
```

```
n=p*q;%n=p*q;

for i=1:n;
u=mod(3*u,1);
y(i)=u; %value at step i
end

xtilde(1,1)= 0;
z=-log(abs(y-xtilde));
x=reshape(z,p,q); %reshape(X,p,q) returns the p-by-q matrix whose elements
                  % are taken columnwise from X.
    for i=1:q
        m(i)=max(x(1:p,i));%returns the maximum values.
    end

U=quantile(z,.9999);%Threshold
%The block method.
for j=1:n
if z(j)> U
S(j)=j; %The exceedance times.
end
end
S(S==0)=[];
N=length(S);%The number of observations
%exceeding threshold.
for k=1:q
if m(k)> U
f(k)=k; %The exceedance times.
end
end
f(f==0)=[];
NM=length(f);%The number of blocks
%in which there is at least one exceedance
Theta=NM/N
```

B.3. Chapter 5 Matlab codes:

```
% The H exponent function of the Times 3 map%
function  [H]  = HT3
N=5000;%number of observations.
u=rand(1);
xtilde=rand(1);
for i=1:N;
u=mod(3*u,1);
y(i)=u; %value at step i
x(i)=-log(abs(y(i)-xtilde));%observabel function.
end
%Estimation the H exponent.
for n=100:10:N
   X(n)=log(n);
    xbar(n)=mean(x(1:n));
   S(n)=std(x(1:n));
for t=1:n
    diff(t)= (x(t)-xbar(n));
```

```
y=cumsum(diff(1:t));
end
R(n)=max(y)-min(y);
Q(n)=R(n)/S(n);
end
X(X==0)=[];
Q(Q==0)=[];
plot(X,log(Q),'x')
hold on
p=polyfit(X,log(Q),1);% finds the coefficients of a polynomial p of degree 1
    that fits the data.
H=p(1);
yfit=polyval(p,X); %returns the value of a polynomial p evaluated at X, yfit
   =p(1)X.
plot(X,yfit,'r-');
end
```

```
% The H exponent function of the Henon map%
function  [H]  = Hhenon
N=5000;%number of observations.
a=1.4;
b=0.3;
x(1)=rand(1,1);%initial value rand(1,2).
y(1)= rand(1,1)/10;
for i=1:N;
    x(i+1)=1-a*(x(i))^2+y(i);
    y(i+1)=b*x(i);
end
xtilde=rand(1,1);
for j=1:N;
    z(j)=-log(abs(x(j)-xtilde));
end
for n=100:10:N
   X(n)=log(n);
    xbar(n)=mean(z(1:n));
    S(n)=std(z(1:n));
for t=1:n
     diff(t)= (z(t)-xbar(n));
yy=cumsum(diff(1:t));
end
R(n)=max(yy)-min(yy);
Q(n)=R(n)/S(n);
end
X(X==0)=[];
Q(Q==0)=[];
plot(X,log(Q),'x')
hold on
p=polyfit(X,log(Q),1);% finds the coefficients of a polynomial p of degree 1
    that fits the data.
H=p(1);
yfit=polyval(p,X); %returns the value of a polynomial p evaluated at X, yfit
   =p(1)X.
plot(X,yfit,'r-');
```

```
% The H exponent function of the Intermittent map%
```

```
function   [H]   = HINT
N=5000;%number of observations.
xtilde=rand(1,1);
 beta=0.1;%\beta \in(0,1)
 u(1,1)= rand(1,1);%initial value
for j=2:N;
    if u(j-1)<=0.5;
        u(j)=u(j-1)*(1+2^beta*u(j-1)^beta);
    else
        u(j)=2*u(j-1)-1;
    end
end
x=-log(abs(u -xtilde));
for n=100:10:N
   X(n)=log(n);
    xbar(n)=mean(x(1:n));
    S(n)=std(x(1:n));
for t=1:n
     diff(t)= (x(t)-xbar(n));
y=cumsum(diff(1:t));
end
R(n)=max(y)-min(y);
Q(n)=R(n)/S(n);
end
X(X==0)=[];
Q(Q==0)=[];
plot(X,log(Q),'x')
hold on
p=polyfit(X,log(Q),1);% finds the coefficients of a polynomial p of degree 1
     that fits the data.
H=p(1);
yfit=polyval(p,X);%returns the value of a polynomial p evaluated at X, yfit=
    p(1)X.
plot(X,yfit,'r-');


%calculate goodness of fit

%  yresid = log(Q)- yfit;
%  SSresid = sum(yresid.^2);
%  SStotal = (length(log(Q))-1) * var(log(Q));
%  stderror = sqrt(SSresid /(length(log(Q))-2));

end
```

```
% The H exponent function of the S&P500(FTSE100) index returns%
x= importdata('LRSP.txt')%importdata('LRFTSE.txt')
N=length(x)
for n=100:10:N
   X(n)=log(n);
    xbar(n)=mean(x(1:n));
    S(n)=std(x(1:n));
for t=1:n
     diff(t)= (x(t)-xbar(n));
y=cumsum(diff(1:t));
```

```
end
R(n)=max(y)-min(y);
Q(n)=R(n)/S(n);
end
X(X==0)=[];
Q(Q==0)=[];
plot(X,log(Q),'x')
hold on
p=polyfit(X,log(Q),1);%Fit a linear model to the data.
H=p(1)
yfit=polyval(p,X);
plot(X,yfit,'r-')
%calculate goodness of fit
 E = log(Q)- yfit;
 SSE = sum(E.^2);
 SST = (length(log(Q))-1) * var(log(Q));
 stderror = sqrt(SSE /(length(log(Q))-2))
 Rsquare = 1 - SSE/SST
 adjRsquare = 1 - SSE/SST * (length(log(Q))-1)/(length(log(Q))-2)
```

B.4. Chapter 6 Matlab codes:

```
%Growth of record times(MATLAB script).
%---The growth of record times from gaussian rvs.
clear;
t=10000;
%the number of observations
x=randn(1,t);
for i=1:t
    DD(i)=max(x(1:i))-x(i);
end
for j=1:t
    if DD(j) == 0
        j;
        logl(j)= log(j);
    end
end
logl(logl==0)=[];
n=[2:length(logl)+1];
plot(n,logl)
xlabel('N');
ylabel('log LN');
%% ---The growth of record times from Uniform rvs.
clear;
t=10000;   %the number of observations
x=rand(1,t);
for i=1:t
    DD(i)=max(x(1:i))-x(i);
end
for j=1:t
    if DD(j) == 0
        j;
        logl(j)= log(j);
    end
```

```
end
logl(logl==0)=[];
n=[2:length(logl)+1];
plot(n,logl)
xlabel('N');
ylabel('log LN');
```

```
%limit result for the frequency of records in IID RVs.
clear;
x=rand(1,10000);% random number drawn from the standard normal distribution.
     or, you can use rand for random number drawn from the uniform
    distribution.
for i=1:10000
DD(i)= max(x(1:i))-x(i);
if DD(i) == 0;
        i;% record time.
    end
end
l=zeros(1,10000);
for j=1:10000
      if DD(j) == 0;
           l(j)=1;
    end
end
t=find(l);%returns the linear indices corresponding to
              %the nonzero entries of the array l.
NR=length(t)% Number of records observed(L(t)).
Lt=[1:NR];
y1=1.96*sqrt(log(t))+log(t);
y2=log(t)-1.96*sqrt(log(t));
plot(t,Lt,'x',t,log(t),'-',t,y1,'--',t,y2,'--')
ylabel('N(1,n]');% Number of records up to time n.
xlabel('n')% Times
```

```
%Embrechts results applied to the RW.
clear;
n=10000;
tvals=[0:n];
Y=2*floor(2*rand(1,n))-1;
s=zeros(1,n);
for i=2:n
    s(1,1)=Y(1,1);
    s(1,i)=s(1,i-1)+Y(1,i);
end
for i=1:n
     DD(i)=max(s(1:i))-s(i);
      if DD(i) == 0;
             i;
             logl(i)= log(i);
      end
end
subplot(3,1,2)
logl(logl==0)=[];
t=[2:length(logl)+1];
plot(t,logl)
```

```
ylabel('log record times');
subplot(3,1,1);plot(tvals,[0,s])
logl=zeros(1,n);
for j=1:n
      if DD(j) == 0;
           logl(j)=1;
      end
end
subplot(3,1,3)
plot(logl)
```

```
%Markov Chain
% this M-file plot the expected number of records
% versus the probability p in Markov Chain.

function y =MCENR(p)
P0=[1/2 1/2];%initial distribution(the state of the system at time Zero).
q=0.5;
n=1000;
m=100;
S=zeros(n,m);
R=zeros(n,m);
l=zeros(n,m);
X=zeros(n,m);
T=[p 1-p
    1-q q];  %transition matrix

for j=1:m;
for i=1:n;
       Q(i,j)=rand(1,1);
      Tn=T^i;
    Pn=[P0*Tn];
    if Q(i,j)<=Pn(1)
        X(i,j)=1;
    else
        X(i,j)=-1;
    end
 S(i,j)=sum(X(:,j));

R(i,j)=max(S(1:i,j))-S(i,j);
if R(i,j) == 0;     % we find that the record occur when R=0 and from this
    we can find the record times
                    %which correspond to the record values.
             i;
 l(i,j)=i;            %Gives record times.
     end

end
for k=1:length(l);
    NR(k,j)=l(k,j);

end
f=find(NR(:,j));% return the non zero element which correspond to record
    times
NofR(j)=length(f);%Gives the numbers of records for each asset path
```

```
end
y=mean(NofR)/n;

end
```

```
%Random Walk
% this M-file computed the rescaled expected number of records
% versus the probability p in Random Walk. For unbiased RW set p=0.5.
function y =RWENRM(p)
n=10000;%number of observations.
m=100;%sample paths
S=zeros(n,m);
R=zeros(n,m);
l=zeros(n,m);
X=zeros(n,m);

for j=1:m;
for i=2:n;

 if rand <= p
X(i,j)=1;
else
X(i,j)=-1;
end
S(i,j)=sum(X(:,j));
%%%%%%%%%%%%%%%%%%%%%%%%
% R(i,j)=max(S(1:i,j))-S(i,j);
% if R(i,j) == 0;      % we find that the  weak record occur when R=0.
%%%%%%%%%%%%%%%%%%%%%%%%
R(i,j)=max(S(1:i-1,j))-S(i,j);
if R(i,j) < 0;      % we find that the strong record occur when R<0 and from
    this we can find the record times
                    %which correspond to the record values.
             i;
 l(i,j)=i;             %Gives record times.
     end

end
for k=1:length(l);
    NR(k,j)=l(k,j);

end
f=find(NR(:,j));% return the non zero element which correspond to record
    times
NofR(j)=length(f);%Gives the numbers of records for each asset path
end
y=mean(NofR)/sqrt(n);
end
```

```
function y=RENINTM
%RW with increments generated by the Times 3 map.
n=100000;%number of observations.
m=100;%sample paths
S=zeros(n,m);
R=zeros(n,m);
```

```
l=zeros(n,m);
X=zeros(n,m);

    beta=0.9;%\beta \in(0,1)
    mu=0.6406;
for j=1:m;
    y(1,1)=rand(1,1);%initial value
for i=2:n;

    if y(i-1)<=0.5;
        y(i)=y(i-1)*(1+2^beta*y(i-1)^beta);
    else
        y(i)=2*y(i-1)-1;
    end

 if y(i) <= 0.5
X(i,j)=1-mu;
else
X(i,j)=-1-mu;
end
S(i,j)=sum(X(:,j));
%%%%%%%%%%%%%%%%%%%%%%%%
% R(i,j)=max(S(1:i,j))-S(i,j);
% if R(i,j) == 0;     % we find that the  weak record occur when R=0.
%%%%%%%%%%%%%%%%%%%%%%%%

R(i,j)=max(S(1:i-1,j))-S(i,j);
if R(i,j) < 0;     % we find that the strong record occur when R<0 and from
    this we can find the record times
                    %which correspond to the record values.
            i;
 l(i,j)=i;            %Gives record times.
     end

end
for k=1:length(l);
    NR(k,j)=l(k,j);

end
f=find(NR(:,j));% return the non zero element which correspond to record
    times
NofR(j)=length(f);%Gives the numbers of records for each asset path
end
y=mean(NofR)/sqrt(n)
end
```

```
clear;
%RW with increments generated by the Times 3 map.
% This M-file computed the rescaled expected number of records
n=1000;%number of observations.
m=100;%sample paths
S=zeros(n,m);
R=zeros(n,m);
l=zeros(n,m);
X=zeros(n,m);
```

```

for j=1:m;
    u=rand(1,1);%initial value
for i=2:n;
u=mod(3*u,1);
y(i)=u; %value at step i
 if y(i) <= 0.5
X(i,j)=1;
else
X(i,j)=-1;
end
S(i,j)=sum(X(:,j));
%%%%%%%%%%%%%%%%%%%%%%%%
% R(i,j)=max(S(1:i,j))-S(i,j);
% if R(i,j) == 0;     % we find that the  weak record occur when R=0.
%%%%%%%%%%%%%%%%%%%%%%%%
R(i,j)=max(S(1:i-1,j))-S(i,j);
if R(i,j) < 0;     % we find that the strong record occur when R<0 and from
    this we can find the record times
                    %which correspond to the record values.
            i;
 l(i,j)=i;             %Gives record times.
     end

end
for k=1:length(l);
    NR(k,j)=l(k,j);

end
f=find(NR(:,j));% return the non zero element which correspond to record
    times
NofR(j)=length(f);%Gives the numbers of records for each asset path
end
y=mean(NofR)/sqrt(n)
```

```
%Embrechts results applied to the financial prices data
clear;
s= importdata('SSP.txt');
%s= importdata('SFTSE.txt');
n=length(s);
for i=2:n

% DD(i)=max(s(1:i,j))-s(i,j);
% if DD(i) == 0; % we find that the  weak record occur when DD=0.

     DD(i)=max(s(1:i-1))-s(i);
      if DD(i) < 0;% we find that the  strong record occur when DD<0.
             i;
             logl(i)= log(i);
      end
end
subplot(3,1,2)
logl(logl==0)=[];
ll=length(logl)
t=[2:ll+1];
```

```
plot(t,logl)
ylabel('log record times');
subplot(3,1,1);plot(s)
logl=zeros(1,n);
for j=2:n
    % if DD(j) == 0;
       if DD(j) < 0;
            logl(j)=1;
       end
end
subplot(3,1,3)
plot(logl)
```

Bibliography

[1] B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja. *Records*. Wiley series in probability and statistics Probability and statistics section. New York: Wiley, 1998, xviii, 312 p.

[2] A. A. Balkema and L. De Haan. "Residual life time at great age". In: *The Annals of Probability* (1974), pp. 792–804.

[3] C. N. Behrens, H. F. Lopes, and D. Gamerman. "Bayesian analysis of extreme events with threshold estimation". In: *Statistical Modelling* 4.3 (2004), pp. 227–244.

[4] M. Benedicks and L. Carleson. "The dynamics of the Hnon map". In: *Annals of Mathematics* (1991), pp. 73–169.

[5] M. Benedicks and L.-S. Young. "Markov extensions and decay of correlations for certain Hnon maps". In: *Astrisque* 261 (2000), pp. 13–56.

[6] J. Beran. *Statistics for long-memory processes*. Monographs on statistics and applied probability. 94014144 Bibliography, pp. 262-301 and indexes. New York: Chapman Hall, 1994, x, 315p.

[7] R. J. Bhansali, M. P. Holland, and P. S. Kokoszka. "Intermittency, long-memory and financial returns". In: *Long Memory in Economics*. Springer, 2007, pp. 39–68.

[8] R. J. Bhansali and M. Holland. "Frequency analysis of chaotic intermittency maps with slowly decaying correlations". In: *Statistica Sinica* 17.1 (2007), p. 15.

[9] R. Bhansali, M. Holland, and P. Kokoszka. "Chaotic maps with slowly decaying correlations and intermittency". In: *Fields Institute Communications* 44 (2004), pp. 99–126.

[10] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*. Vol. 27. Cambridge university press, 1989.

[11] F. Black and M. Scholes. "The pricing of options and corporate liabilities". In: *The journal of political economy* (1973), pp. 637–654.

[12] D. O. Cajueiro and B. M. Tabak. "Testing for long-range dependence in world stock markets". In: *Chaos, Solitons Fractals* 37.3 (2008), pp. 918–927.

[13] D. O. Cajueiro and B. M. Tabak. "The rescaled variance statistic and the determination of the Hurst exponent". In: *Mathematics and Computers in Simulation* 70.3 (2005), pp. 172–179.

[14] J. Carreau and Y. Bengio. "A hybrid Pareto model for asymmetric fat-tailed data: the univariate case". In: *Extremes* 12.1 (2009), pp. 53–76.

[15] S. Coles. *An introduction to statistical modeling of extreme values*. Vol. 208. Springer, 2001.

[16] P. Collet. "Statistics of closest return for some non-uniformly hyperbolic systems". In: *Ergodic Theory and Dynamical Systems* 21.2 (2001), pp. 401–420.

[17] R. Cont. “Empirical properties of asset returns: stylized facts and statistical issues”. In: (2001).

[18] R. Cont. “Volatility clustering in financial markets: empirical facts and agent-based models”. In: *Long memory in economics*. Springer, 2007, pp. 289–309.

[19] M. Couillard and M. Davison. “A comment on measuring the Hurst exponent of financial time series”. In: *Physica A: Statistical Mechanics and its Applications* 348.0 (2005), pp. 404–418.

[20] W. De Melo and S. Van Strien. *One-dimensional dynamics*. Vol. 25. Citeseer, 1993.

[21] J. Dedecker, H. G. Dehling, and M. Taqqu. “Weak convergence of the empirical process of intermittent maps in L2 under long-range dependence”. In: *arXiv preprint arXiv:1311.5873* (2013).

[22] R. L. Devaney. “An Introduction to Chaotic Dynamical Systems”. In: (2003).

[23] Z. Ding and C. W. Granger. “Modeling volatility persistence of speculative returns: a new approach”. In: *Journal of econometrics* 73.1 (1996), pp. 185–215.

[24] Z. Ding, C. W. Granger, and R. F. Engle. “A long memory property of stock market returns and a new model”. In: *Journal of empirical finance* 1.1 (1993), pp. 83–106.

[25] P. Embrechts, R. Frey, and A. McNeil. *Quantitative risk management*. Vol. 10. Princeton Series in Finance. Princeton: Princeton University Press, 2005.

[26] P. Embrechts, C. Klppelberg, and T. Mikosch. *Modelling extremal events for insurance and finance*. Applications of mathematics, New York: Springer, 1997, xv, 645 p.

[27] K. Eriksson, D. Estep, and C. Johnson. *Applied Mathematics: Body and Soul: Volume 1: Derivatives and Geometry in IR3*. Vol. 1. Springer, 2004.

[28] A. Etheridge. *A course in financial calculus*. Cambridge University Press, 2002.

[29] K. Falconer. *Fractal geometry: mathematical foundations and applications*. John Wiley Sons, 2013.

[30] K. J. Falconer and K. Falconer. *Techniques in fractal geometry*. Vol. 16. Wiley Chichester (W. Sx.), 1997.

[31] E. F. Fama. “Efficient capital markets: A review of theory and empirical work*”. In: *The Journal of Finance* 25.2 (1970), pp. 383–417.

[32] D. Faranda et al. “Numerical convergence of the block-maxima approach to the Generalized Extreme Value distribution”. In: *Journal of statistical physics* 145.5 (2011), pp. 1156–1180.

[33] C. A. Ferro. “Statistical Methods for Cluster of Extreme Values”. PhD thesis. 2003.

[34] C. A. Ferro and J. Segers. “Inference for clusters of extreme values”. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 65.2 (2003), pp. 545–556.

[35] B. Finkenstadt and H. Rootzn. *Extreme values in finance, telecommunications, and the environment*. CRC Press, 2003.

[36] R. A. Fisher and L. H. C. Tippett. “Limiting forms of the frequency distribution of the largest or smallest member of a sample”. In: *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 24. Cambridge Univ Press, pp. 180–190.

[37] A. C. M. Freitas and J. M. Freitas. "On the link between dependence and independence in extreme value theory for dynamical systems". In: *Statistics Probability Letters* 78.9 (2008), pp. 1088–1093.

[38] A. C. M. Freitas, J. M. Freitas, and M. Todd. "Hitting time statistics and extreme value theory". In: *Probability Theory and Related Fields* 147.3-4 (2010), pp. 675–710.

[39] A. C. M. Freitas, J. M. Freitas, and M. Todd. "Speed of convergence for laws of rare events and escape rates". In: *arXiv preprint arXiv:1401.4206* (2014).

[40] A. C. M. Freitas, J. M. Freitas, and M. Todd. "The compound Poisson limit ruling periodic extreme behaviour of non-uniformly hyperbolic dynamics". In: *Communications in Mathematical Physics* 321.2 (2013), pp. 483–527.

[41] A. C. M. Freitas, J. M. Freitas, and M. Todd. "The extremal index, hitting time statistics and periodicity". In: *Advances in Mathematics* 231.5 (2012), pp. 2626–2665.

[42] J. W. Galbraith and S. Zernov. "Extreme dependence in the NASDAQ and SP 500 composite indexes". In: *Applied Financial Economics* 19.13 (2009), pp. 1019–1028.

[43] M. Gilli. "An application of extreme value theory for measuring financial risk". In: *Computational Economics* 27.2-3 (2006), pp. 207–228.

[44] M. T. Greene and B. D. Fielitz. "Long-term dependence in common stock returns". In: *Journal of financial Economics* 4.3 (1977), pp. 339–349.

[45] C. Gupta, M. Holland, and M. Nicol. "Extreme value theory and return time statistics for dispersing billiard maps and flows, Lozi maps and Lorenz-like maps". In: *Ergodic Theory and Dynamical Systems* 31.5 (2011), p. 1363.

[46] L. Haan. *On regular variation and its application to the weak convergence of sample extremes*. Mathematisch Centrum, 1970.

[47] W. Hall and J. A. Wellner. "The rate of convergence in law of the maximum of an exponential sample". In: *Statistica Neerlandica* 33.3 (1979), pp. 151–154.

[48] K. Hamidieh, S. Stoev, and G. Michailidis. "On the estimation of the extremal index based on scaling and resampling". In: *Journal of Computational and Graphical Statistics* 18.3 (2009), pp. 731–755.

[49] L.-Y. He and W.-B. Qian. "A Monte Carlo simulation to the performance of the R/S and V/S methodsStatistical revisit and real world application". In: *Physica A: Statistical Mechanics and its Applications* 391.14 (2012), pp. 3770–3782.

[50] M. Henon. "A two-dimensional mapping with a strange attractor". In: *Communications in Mathematical Physics* 50.1 (1976), pp. 69–77.

[51] D. J. Higham. *An introduction to financial option valuation : mathematics, stochastics and computation*. Desmond J. Higham. ill. ; 25cm. Cambridge: Cambridge University Press, 2004, xxi, 276p.

[52] M. Holland. "Slowly mixing systems and intermittency maps". In: *Ergodic Theory and Dynamical Systems* 25.01 (2005), pp. 133–159.

[53] M. Holland and M. Nicol. "Speed of convergence to an extreme value distribution for non-uniformly hyperbolic dynamical systems". In: *Preprint* (2014).

[54] M. Holland, P. Rabassa, and A. Sterk. "On the convergence to an extreme value distribution for non-uniformly hyperbolic dynamical systems". In: *arXiv preprint arXiv:1404.3941* (2014).

[55] M. P. Holland et al. “Extreme value laws in dynamical systems under physical observables”. In: *Physica D: Nonlinear Phenomena* 241.5 (2012), pp. 497–513.

[56] J. Hosking, J. R. Wallis, and E. F. Wood. “Estimation of the generalized extreme-value distribution by the method of probability-weighted moments”. In: *Technometrics* 27.3 (1985), pp. 251–261.

[57] T Hsing, J Hsler, and M. Leadbetter. “On the exceedance point process for a stationary sequence”. In: *Probability Theory and Related Fields* 78.1 (1988), pp. 97–112.

[58] H. E. Hurst. “Long-term storage capacity of reservoirs”. In: *Trans. Amer. Soc. Civil Eng.* 116 (1951), pp. 770–808.

[59] P. W. Jones and P. Smith. *Stochastic processes : an introduction*. 2nd. Chapman Hall/CRC texts in statistical science series. Boca Raton, FL: Chapman Hall/CRC, 2010, x, 221 p.

[60] F. Laurini and J. A. Tawn. “New estimators for the extremal index and other cluster characteristics”. In: *Extremes* 6.3 (2003), pp. 189–211.

[61] M. Leadbetter, G Lindgren, and H Rootz en. *Extremes and Related Properties of Random Sequences and Processes*. 1983.

[62] C. Liverani, B. Saussol, and S. Vaienti. “A probabilistic approach to intermittency”. In: *Ergodic Theory and Dynamical Systems* 19.03 (1999), pp. 671–685.

[63] A. W. Lo. *Long-term memory in stock market prices*. Tech. rep. National Bureau of Economic Research, 1989.

[64] F. M. Longin. “From value at risk to stress testing: The extreme value approach”. In: *Journal of Banking Finance* 24.7 (2000), pp. 1097–1130.

[65] V. Lucarini, D. Faranda, and J. Wouters. “Universal behaviour of extreme value statistics for selected observables of dynamical systems”. In: *Journal of statistical physics* 147.1 (2012), pp. 63–73.

[66] V. Lucarini et al. “Extreme value theory for singular measures”. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 22.2 (2012), p. 023135.

[67] V. Lucarini et al. “Towards a general theory of extremes for observables of chaotic dynamical systems”. In: *Journal of statistical physics* 154.3 (2014), pp. 723–750.

[68] S. N. Majumdar, G. Schehr, and G. Wergen. “Record statistics and persistence for a random walk with a drift”. In: *Journal of Physics A: Mathematical and Theoretical* 45.35 (2012), p. 355002.

[69] S. N. Majumdar and R. M. Ziff. *Universal Record Statistics of Random Walks and L’evy Flights*. 2008.

[70] B. Mandelbrot. “Statistical methodology for nonperiodic cycles: from the covariance to Rs analysis”. In: *Annals of Economic and Social Measurement, Volume 1, number 3*. NBER, 1972, pp. 259–290.

[71] Mathworks. *Modelling Data with the Generalized Extreme Value Distribution*. 2014.

[72] Mathworks. *Modelling Tail Data with the Generalized Pareto Distribution*. 2014.

[73] I. Melbourne and M. Nicol. “Almost sure invariance principle for nonuniformly hyperbolic systems”. In: *Communications in Mathematical Physics* 260.1 (2005), pp. 131–146.

[74] F. F. Nascimento, D. Gamerman, and H. F. Lopes. “A semiparametric Bayesian approach to extreme value estimation”. In: *Statistics and Computing* 22.2 (2012), pp. 661–675.

[75] G. L. O’Brien. “Extreme values for stationary and Markov sequences”. In: *The Annals of Probability* (1987), pp. 281–291.

[76] Y. B. Pesin. *Dimension theory in dynamical systems: contemporary views and applications*. University of Chicago Press, 2008.

[77] K. Petersen. *Ergodic theory*. Vol. 2. Cambridge University Press, 1989.

[78] J. Pickands III. “Statistical inference using extreme order statistics”. In: *the Annals of Statistics* (1975), pp. 119–131.

[79] M. Pollicott and M. Yuri. *Dynamical systems and ergodic theory*. Vol. 40. Cambridge University Press, 1998.

[80] C. Y. Robert, J. Segers, and C. A. Ferro. “A sliding blocks estimator for the extremal index”. In: *Electronic Journal of Statistics* 3 (2009), pp. 993–1020.

[81] S. Rostek. *Option Pricing in Fractional Brownian Markets*. 2009.

[82] D. Ruelle. “a Strange Attractor?” In: *Notices of the AMS* 53.7 (2006).

[83] C. Scarrott and A. MacDonald. “A review of extreme value threshold es-timation and uncertainty quantification”. In: *REVSTATStatistical Journal* 10.1 (2012), pp. 33–60.

[84] R. L. Smith. “The extremal index for a Markov chain”. In: *Journal of Applied Probability* (1992), pp. 37–45.

[85] R. L. Smith and I. Weissman. “Estimating the extremal index”. In: *Journal of the Royal Statistical Society. Series B (Methodological)* (1994), pp. 515–528.

[86] V. Teverovsky, M. S. Taqqu, and W. Willinger. “A critical look at Lo’s modified R/S statistic”. In: *Journal of statistical Planning and Inference* 80.1 (1999), pp. 211–227.

[87] M. Thaler. “Estimates of the invariant densities of endomorphisms with indifferent fixed points”. In: *Israel Journal of Mathematics* 37.4 (1980), pp. 303–314.

[88] J.-J. Tseng and S.-P. Li. “Asset returns and volatility clustering in financial time series”. In: *Physica A: Statistical Mechanics and its Applications* 390.7 (2011), pp. 1300–1314.

[89] P. Walters. *An introduction to ergodic theory*. Vol. 79. Springer Science Business Media, 2000.

[90] G. Wergen. “Modeling record-breaking stock prices”. In: *Physica A: Statistical Mechanics and its Applications* 396 (2014), pp. 114–133.

[91] G. Wergen. “Records in stochastic processestheory and applications”. In: *Journal of Physics A: Mathematical and Theoretical* 46.22 (2013), p. 223001.

[92] G. Wergen, M. Bogner, and J. Krug. “Record statistics for biased random walks, with an application to financial data”. In: *Physical Review E* 83.5 (2011). PRE, p. 051109.

[93] L.-S. Young. “Dimension, entropy and Lyapunov exponents”. In: *Ergodic Theory and Dynamical Systems* 2.01 (1982), pp. 109–124.

[94] L.-S. Young. “Recurrence times and rates of mixing”. In: *Israel Journal of Mathematics* 110.1 (1999), pp. 153–188.

[95] L.-S. Young. “Statistical properties of dynamical systems with some hyperbolicity”. In: *Annals of Mathematics* (1998), pp. 585–650.

[96] L.-S. Young. “What are SRB measures, and which dynamical systems have them?” In: *Journal of statistical physics* 108.5 (2002), pp. 733–754.