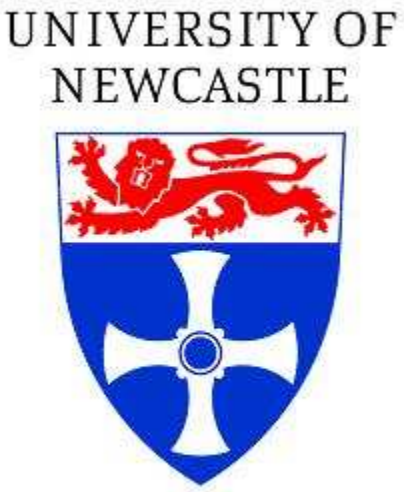
Stability of Random Dynamical Systems and Applications

Carlos G. Pacheco

A Thesis submitted for the degree of

Doctor of Philosophy



School of Mathematics and Statistics

Supervisor: Dr. Jordan Stoyanov

Contents

Abstract

In different branches of science, we deal with models which are both, stochastic (or random) and evolving dynamically in time. We consider general structures called Random Dynamical Systems consisting of sequences of random elements appropriately parameterized. The random elements can be random variables, finite or infinite dimensional random vectors or sequences of stochastic processes or even functionals of stochastic processes. Our models allow to analyse a superposition of random effects.

One of the most important problems is to find the limit behaviour of such systems. There are several different ways to look for the limit, e.g. in week sense or in strong sense. The existence of any kind of a limit is understood as a stability property of the system under study. In such a case we say that the random system obeys or reaches that specific kind of regularity in the limit and also that the system is stabilising in limit. In the Thesis we consider different kinds of stability/regularity, such as: ergodicity of Markov processes describing random walks in random environment, the limit-processes of sequences of stochastic processes based on random sums of arbitrary random variables and the convergence of a sequence of algorithms when solving an integral equation. An illustration in an applied area is also given.

Although the type of stability/regularity seems different for the above models, it will be clear that they actually come all from the same general concept - the existence of limits of the probabilistic laws describing the evolution of random systems. In other words, since we deal with general random elements, they are governed by specific probabilistic laws, then under some conditions, functions or functionals of the random elements have limits, in a specified way. Any such a limit is a regularity property of the system.

Besides their theoretical nature, the models considered in the Thesis and the obtained results are directly related to problems in other areas. One such an area is statistical physics, where one of the fundamental problems is to analyse the behaviour of large ensembles of randomly moving particles. Another area of possible applications is Financial mathematics. We consider a model of financial time series that resembles the so-called Binomial model. Under some conditions, it is known that the Binomial model converges to a functional of the Brownian Motion. This involves the Cox-Ross-Rubinstein framework and the Black-Scholes model. Similarly, under appropriate scaling, the models we study converge to the so-called Generalized Hyperbolic Lévy Motion.

In summary, this Thesis is about stability properties of random dynamical systems.

Symbol Index

$a.s.$	almost surely
$ch.f.$	characteristic function
$d.f.$	distribution function
$f.d.d.$	finite dimensional distributions
$r.v.$	random variable
(A, ν, γ)	The characteristic triplet of a Lévy process
$\lvert x \rvert$	The absolute value of x
$\lfloor x \rfloor$	The largest interger below of x
$\mathcal{B}_S$	Borel sigma-algebra of the space S
$C(S)$	Continuous functions from the space S to $\mathbb{R}$
$C_b(S)$	Bounded continuous functions from the space S to $\mathbb{R}$
C	Space of continuous functions
$C_{\rightarrow}$	Space of increasing continuous functions
$C_{\rightrightarrows}$	Space of strictly increasing continuous functions
D	Space of *cádlág* functions (left-continuous with right-hand limits)
$D_{\rightarrow}$	Space of increasing *cádlág* functions
$D_{\rightrightarrows}$	Space of strictly increasing *cádlág* functions
J_1	The Skorohod topology in D
$\mathcal{L}(X)$	The law of the random element X
$(\Omega, \mathcal{F}, P)$	Probability space Ω with sigma algebra $\mathcal{F}$ and probability measure P
$\mathbb{R}$	The set of real numbers $(-\infty, \infty)$
$X \stackrel{d}{=} Y$	X and Y have the same law
$X \sim f$	The density of the random variables X is f
$X \stackrel{f.d.d}{\rightarrow} Y$	Convergence of the finite dimensional ditribution of X to those of Y
$X \stackrel{d}{\rightarrow} Y$	Convergence in distribution of X to Y
X_{t-}	$\lim X_{t-\varepsilon}$ for $\varepsilon > 0$ and $\varepsilon \rightarrow 0$

1 Introduction

In the following section we summarize general ideas, notion, questions and results, that are treated in this Thesis, underlying the inter-relation of the topics.

1.1 Stability of random systems

In real life there are several phenomena which evolve randomly in time, essentially this means that we can not predict the exact future outcomes. Nevertheless, we always guess that it might be possible to observe some patterns and regularity. Because it is important for us to explain and understand the uncertainty, apparently randomness of the phenomena, we appeal for adequate models that help to pursue our aims. This is how we arrive to the necessity of studying and analyzing, from general point of view, the *limit behaviour* of random elements that arise from a *random system*. The random and dynamical systems that we work with can be analyzed as schemes which consist of an infinite sequence of transformations or functions of collections of random quantities. Thus, our main goal is to study different stochastic dynamical models and explore the existence of any sort of *stability* which can be described in an appropriate way. Typically, a stability property arises as a result of a limiting procedure. For example, we may think of stability when a system reaches a state or position which in some sense we can call an equilibrium. This would mean that the system remains unchanged or "almost" invariant.

The following is a list of models, each being both stochastic and dynamical.

Random walks

Law of large numbers

Markov chains and ergodic distributions

Central limit theorem

Domain of attraction in limit theorems

Extreme value theory and other transformations

Functional limit theorems

Monte-Carlo techniques

Random walks. Consider a sequence of random variables $X_1, X_2, ...$ To analyze the *sums* of random variables is a fundamental problem in probability. Suppose we are interested in the new sequence $Y_1, Y_2, ...$ defined as

$$Y_n = \sum_{i=1}^{n} X_i \text{ for } n = 1, 2, ...$$

This is the so-called *random walk* which is another collection, or sequence of random variables based on the original sequence $\{X_i, i = 1, 2, ...\}$. Thus the sums form a random system obtained by transformations of random variables. We are interested to study the limit behaviour of the sequence $\{Y_n, n = 1, 2, ...\}$, i.e. we want to describe the variable Y_n when $n \to \infty$. It could be the case that the variable Y_n explodes to ∞ or $-\infty$. In such a case, the system can not be stable.

Law of large numbers. Sometimes, instead of analyzing the bare sum we can apply a different transformation that describes a new random system that brings stability. For instance, suppose we are interested to study the average value rather that the bare sum Y_n, i.e.

$$S_n = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

From probability theory we know that if X is a random variable (r.v.) and $\{X_i, i = 1, 2, ...\}$ are independent and identically distributed (i.i.d.) r.v. like X, and the expectation $E(|X|)$ is finite, then the average S_n tends to $E(X)$ as $n \to \infty$ almost surely (a.s.). The above result is the well-known *strong law of large numbers*. We give this example to illustrate, vaguely, the idea of a random system that brings stability in the form of convergence a.s. Another similar example of convergence a.s. is the so-called *laws of iterated logarithm*. Also Birkhoff´s individual ergodic Theorem, which is related to the next point, can be regarded as a strong law of large numbers for Markov chains; and this is a important result for the Monter-Carlo techniques, also mentioned below.

Markov chains and ergodic distributions. The existence of a stationary distribution for a Markov chains is another example. The main interest is to find limit distribution, or the so-called *ergodic* distribution of the Markov chain. The Markov chain, indexed with

discrete time, defines a random system (see Stenflo (2001)), and the existence of the ergodic distribution defines its stability property.

In this dissertation we think of a random system in the following unified way. We start with a collections of random elements, e.g. random events, random variables, random vectors, stochastic processes. Then we define a new sequence of random elements as a collection based on transformations of the original random elements. Or we say that the random system is generated by the sequential transformations of a collection of random elements.

That is, given a collection of random elements $X_1, X_2, ...$ we can obtain a new sequence of random elements of the form

$$Y_n = H_n(X_1, X_2...), n = 1, 2, ...$$

where H_n are transformations of $X_1, X_2, ...$

Central limit theorem. The *central limit theorem (CLT)* is another fundamental result exhibiting a stability property. It states the following. Let $X_1, X_2, ...$ be a sequence of random variables which are i.i.d. with $E(X_1)$ and $Var(X_1)$ finite. Then, as $n \to \infty$, the distribution of the centered and normalized sums,

$$S_n = \frac{1}{\sqrt{nVar(X_1)}} \sum_{i=1}^{n} (X_i - E(X_1)),$$

converges to the standard normal distribution $N(0, 1)$.

Domain of attraction. Similar to the kind of stability related to the Central limit theorem is the following concept. We say that a r.v. X is in the *domain of attraction of* the r.v. Y, or that X *obeys a central limit theorem,* if there exist a sequence of real numbers $\{a_n, n \geq 1\}$ and a sequence of positive numbers $\{b_n, n \geq 1\}$ such that

$$\frac{1}{b_n} \sum_{i=1}^{n} X_i - a_n \xrightarrow{d} Y \text{ as } n \to \infty.$$

Here $X_1, X_2, ...$ are independent copies of X. Under some conditions on the tails of the distribution (The tails of the distribution of a r.v. X is refered to be $P(|X| > \alpha)$), the

so-called *heavy-tailed distributions*, it is possible to find the domain of attraction. It is known as the *stable laws* (a classical reference is Gnedenko and Kolmogorov (1954), see also Meerschaert and Scheffler (2001), Whitt (2002), Lamperti (1977)). It is also still possible to have a stability property, e.g. after relaxing the independence condition for the r.v.s $X_1, X_2, ...$(see Whitt (2002)).

Extreme value theory and other transformations. Summation of random variables is a transformation that, under some conditions, brings stability. However, it is possible to prove other limit theorems for other types of transformations, such as *multiplication (products)*, taking a *minimum* or a *maximum*, or even a combination of them, see Whitt (2002). This involves the random quantities T_n, Max_n and Min_n, where

$$T_n = X_1 \cdot ... \cdot X_n,\ Max_n = \max\{X_1, ..., X_n\},\ Min_n = \min\{X_1, ..., X_n\}.$$

For these cases it possible to describe a random system that reaches stability in some sense. For example, under some conditions there are sequences of real numbers $\{a_n, n \geq 1\}$ and positive numbers $\{b_n, n \geq 1\}$ such that each random sequence $a_n (T_n)^{b_n}$, $a_n Max_n + b_n$ and $a_n Min_n + b_n$ converges in distributions to a specific random variable. The case of Max_n and Min_n are widely studied in the so-called EVT (extreme value theory). The *Gumble, Fréchet and Weibull distributions* arise as the limit in this type of random systems. The products T_n have also several applications and it has been studied the possible stability of a system that it can bring; an extended treatment can be found in Galambos and Simonelli (2004).

In all the examples mentioned above, the random systems obey stability as a specific limit of transformed sequences of r.v.s. There is even a more general theory dealing with random elements in functional spaces. The next examples of stochastic dynamical systems are also treated in this Thesis.

Functional limit theorems. Consider the following construction. Let $X_1, X_2, ...$ be i.i.d. r.v.s. with mean μ and variance σ^2, finite. Let us define the following sequence of stochastic processes

$$S_n(t) = \frac{1}{\sqrt{\sigma^2 n}} \sum_{i=1}^{\lfloor tn \rfloor} (X_i - \mu),\ t \geq 0,\ n = 1, 2, ...,$$

where $\lfloor x \rfloor$ stands for the integer part of the real number x. There is a result (see Billingsley (1999)) called the *Donsker´s theorem,* saying that this sequence of stochastic processes converges in distribution (carefully and formally defined) to the *Brownian motion.* This kind of results are called *functional limit theorems* because this is convergence in the space of functions. In this case, with a sequence of r.v.s $X_1, X_2, ...$ and a sequence of transformations we obtain a sequence of stochastic processes. The stability property is reached when the limit as $n \to \infty$, specified in some sense, of this sequence stochastic processes is another stochastic process. The construction of the Brownian motion using the so-called *Haar* and *Schauder functions* is another example. Although the construction is different from simple sums, it is still based on transformations of a countable family of r.v.s. In this construction the limit of the random elements is also the Brownian Motion (see Karatzas and Shreve (1991), p.56) which specifies the stability of this random system.

Monte-Carlo techniques. A topic related to stability is the so-called Monte-Carlo technique used for example for estimating an integral of a deterministic function. The algorithm is fed by random quantities to approximate the integral of the function with a sequence of other random quantities produced by the iteration when running the algorithm. Again, the stability of the random system is reached when approximating the exact value of the integral.

Also, we have to mention that there is a theory for stability of stochastic differential equations. This concerns mainly diffusions, which are stochatic differential equations driven by the Brownian Motion, and has its origins in the classical theory of stability of deterministic differential equations (see Mao (1997) or Oksendal (2003)).

1.2 Main topics treated in this work and contributions

Let us describe briefly the structure of the dissertation also mentioning about the main new developments which can be considered as a contribution to the area of stochastic dynamical systems.

This Thesis comprises different models of random systems that reach some sort of stability. The work can be summarized as follows.

In Chapter 2, we consider several models of random motions in random environment. This requires us to study classes of discrete-time Markov chains with values in the interval $(0, 1)$. These chains are ergodic and the important problem is to find explicitly the ergodic distribution.

In Chapter 3, we study models based in random sums of r.v.s. We establish limit theorems for some stochastic processes that generalized the Continuous Time Random Walks (CTRW). This will help to propose an economic model base on discrete lattices path-structure to study financial series. We also study the moment determinancy of random sums.

Chapter 4 is entirely devoted to pricing theory. We use models of Chapter 3 for pricing contingent claims.

Chapter 5 is dedicated to statistical inference. This also has a flavor of stability property of a random system.

In Chapter 6, we propose a numerical algorithm for solving specific classes of integral equations.

We have included appendices at the end of Chapters 2 and 3 to recall notions and results that we use througout, but also they are meant to be a compact self-contained summaries of the general theory in this area. The same can be said about Sections 3.3, 4.1 and 4.2.

The contributions made in this work can be identified in the text as those results, propositions, which are numbered.

The Propositions of Chapter 2 describe properties of the ergodic distribution or determine the ergodic distribution for some Markov chains, specifically described.

The Propositions of Chapter 3 establish the limit of sequences of specific continuous-time

stochastic processes based on random sums of r.v.s. In the same chapter, we have a result that establishes the moment determinancy of random sums of r.v.s.

In Chapter 4 section 4.3, after exposing the general theory of pricing in Sections 4.1 and 4.2, it is described a general mechanism for pricing using the models of Chapter 3.

In Section 5.1 we have a proposition that describes a consistent estimator of the diffusion coefficient of a Geometric Brownian motion. Also stated in the Section, we have a proposition dealing with the so-called order statistics. In Section 5.2 it is suggested a procedure for estimating parameters of a model treated on Chapters 3 and 4.

In Chapter 6, it is proposed a technique, based on Monte-Carlo techniques, for solving integral equations.

The above, together with some important definitions and remarks, represent the contribution of this work.

Successfully, we answer interesting questions for different stochastic dynamical models which are either new or extensions of models treated by other authors.

2 Random Motion in Random Environment

The Markov chains considered in this section arise naturally as extensions of the following simple model.

A particle is moving from one position to another position in the interval $(0,1)$. We denote by X_n its position at time $n = 1, 2, ...$ and let at time n, $X_n = x$. Then at time $n+1$ the particle moves towards 1 with probability p and towards 0 with probability $q = 1-p$. If the particle goes up, the new position is uniformly distributed in the interval $(x, 1)$, while, if it goes down, the position is uniformly distributed in the interval $(0, x)$.

Several variations of this Markov models have been studied recently. Among the available sources we mention here Diaconis and Freedman (1999), Stoyanov and Pirinsky (2000), Letac (2002) and Iacus and Negri (2003).

In Stoyanov and Pirinsky (2000) it is found the ergodic distribution of similar models by using the Fréchet-Shohat theorem. Here we follow different procedure allowing us to find the ergodic distribution in similar models.

2.1 Models, questions of interest and some results

In the following discussion, X is a random variable (r.v.) with distribution function (d.f.) F and density f, and we write $X \sim F$ or $X \sim f$.

Let us think of the motion of a physical particle, or the movements of a stock price, or the ratio of some variables, along discrete time, say $n = 1, 2, ...$ The notion of the particle is random, in one or another sense, which will be specified. After rescaling all values, we can consider all variables involved to be in the interval $(0, 1)$. Starting from $X_0 \in (0, 1)$, the position of the particle at time n is denoted by X_n. Hence X_n, $n = 0, 1, ...$ is a discrete-time stochastic process. If at time n, $X_n = x$, $x \in (0, 1)$, the further motion is described as follows: with probability p_n the particle goes up, i.e. $X_{n+1} > X_n$, while, with probability $q_n = 1 - p_n$, the particle moves down, so $X_{n+1} < X_n$.

Now we need to introduce two functions. For any $x \in (0, 1)$, and n, let

$$f_n(u \mid x),\ u \in (x, 1) \text{ and } g_n(u \mid x),\ u \in (0, x)$$

be probability density functions.

We define random motion of the particle, from position $X_n = x$, by associating its up-movements with $f_n(\cdot \mid x)$ and down-movements with $g_n(\cdot \mid x)$. That is

$$(X_{n+1} \mid X_n = x) \sim f_n(\cdot \mid x) \text{ on } (x,1), \text{ (if going up) with probability } p_n, \text{ while}$$
$$(X_{n+1} \mid X_n = x) \sim g_n(\cdot \mid x) \text{ on } (0,x), \text{ (if going down) with probability } q_n. \tag{1}$$

Since both, the direction to move, and the position of the particle are random, we use the term random motion in random environment.

Thus, the motion of the particle is described by the random sequence $\{X_n, n = 0,1,2,...\}$, which is a non-homogeneous Markov chain with discrete time and state space $(0,1)$.

In general, the properties of this Markov chain depend on the probabilities $\{p_0, p_1, ...; q_0, q_1, ... : q_n = 1 - p_n\}$ and the densities $\{f_n(\cdot \mid x),\ n = 0,1,2,...\}$ and $\{g_n(\cdot \mid x),\ n = 0,1,2,...\}$. Obviously, there are many ways to choose these paramenters of the model. Only after specifing them, we can look at the behaviour of the chain as $n \to \infty$.

2.1.1 Specific setting

Suppose that $h = (h(x), x \in (0,1))$ is a density function and let $Y_1, Y_2, ...$ be a sequence of independent r.v.s. each with density h.

Let us assume that for any $x \in (0,1)$ and any $n = 1,2,...$, $f_n(\cdot \mid x)$ and $g_n(\cdot \mid x)$ are specified by the same density h on $(0,1)$ rescaled repectively on the subintervals $(x,1)$ and $(0,x)$ as follows. If at time n, $X_n = x$, $x \in (0,1)$, we take $f_n(\cdot \mid x)$ to be the density of the r.v. $(1-x)Y_n + x$ on $(x,1)$, and similarly, $g_n(\cdot \mid x)$ the density of the r.v. $x - xY_n$ on $(0,x)$.

The densities $f_n(\cdot \mid x)$ and $g_n(\cdot \mid x)$ are "similar" to h (they have the same shape), but instead of being on $(0,1)$ they are defined on $[x,1)$ and $(0,x)$, respectively. We easily see that

$$f_n(u \mid x) = \left(\frac{1}{1-x}\right) h\left(\frac{u-x}{1-x}\right),\ u \in (x,1), \text{ and } g_n(u \mid x) = \left(\frac{1}{x}\right) h\left(\frac{x-u}{x}\right),\ u \in (0,x).$$

By (1), the model for X_{n+1} is:

$$X_{n+1} = \begin{cases} X_n + (1 - X_n)Y_n & \text{with probability } p_n \\ X_n - X_n Y_n & \text{with probability } q_n. \end{cases} \tag{2}$$

Moreover, if $B_n,\ n = 0, 1, 2, ...$ are Bernoulli r.v.s, i.e. $B_n = 1$ with probability p_n, and $B_n = 0$ with probability q_n, then

$$X_{n+1} = X_n(1 - Y_{n+1}) + Y_{n+1}B_n,\ n = 0, 1, 2, ...; X_0 = x_0.$$

Under the above assumptions on $\{Y_n\}$ and $\{B_n\}$ the random sequence $\{X_n,\ n = 0, 1, 2, ...\}$ is a Markov chain. This is referred to be the *iterated function system* representation of the Markov chain, see Stenflo (2001) and Diaconis and Freedman (1999). The Markov chain defined above can be considered as a typical example of a random dynamical system.

By using the recursive formula for X_{n+1}, we can find the following iteration formula:

$$X_n = X_0 \prod_{i=1}^{n}(1 - Y_i) + \sum_{k=1}^{n-1} B_k \left\{ Y_k \prod_{j=k+1}^{n} (1 - Y_j) \right\} + B_n Y_n, \tag{3}$$

Here, to recall, $Y_1, ..., Y_n, B_1, ..., B_n$ are all mutually independent r.v.s defined as above.

2.1.2 Particular cases

An interesting particular case is when Y_n is a uniform r.v. and B_n a Bernoulli r.v. with constant p_n, i.e. $p_n = p$ for all $n = 1, 2, ...$, where $p \in (0, 1)$. This means that from its current position $X_n = x$, the particle moves up with probability p and down with probability $q = 1 - p$. This model was studied in Stoyanov and Pirinsky (2000), and it is found that the ergodic distribution of the Markov chain X_n is beta distribution $\beta(p, q)$ with $q = 1 - p$.

Recall that a r.v. ξ is beta-distributed with parameters a and b, $a > 0$, $b > 0$, and we write $\xi \sim \beta(a, b)$, if the density of ξ is

$$p(x) = \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1 - x)^{b-1},\ x \in (0, 1).$$

Here $\Gamma(\cdot)$ stands for the classical Euler-gamma function. And the moments of ξ are given by

$$E\left[\xi^{k+1}\right] = \prod_{i=0}^{k} \frac{a + i}{a + b + i},\quad k = 0, 1, 2, ... \tag{4}$$

One of our goals is to find the ergodic distributions of the Markov chain X_n for other choices of Y_n and B_n.

Figure 1: Simulation of Case 1

Figure 2: Simulation of Case 1

To have an idea of how the Markov chain behaves we present the simulation of one path. Using equation (2) where B_n ∼Bernoulli(0,1) with parameter $p = .6$ and $Y_n \sim \beta(4, 500)$ for $n = 0, 1, 2, ...$, we start the Markov Chain at $X_0 = .1$ and let it run up to $10,000$ points. Additionally, we show the histogram for the values of the path (Figures 1 and 2, respectively). Birkhoff´s individual ergodic theorem (In the Appendix) justifies why the histogram is an approximation for the ergodic distribution, so, the histogram gives a glance of how the ergodic distribution looks like.

Now we present the cases we work with. One of the extensions is to assume that Y_n follow beta-distribution, i.e. the density $h = \mathcal{L}(Y_n)$ is not just the uniform.

Case 1. Suppose $\mathcal{L}(Y_n) = \beta(1, b)$ for some $b > 0$, and $p_n = p \in (0, 1)$ for all $n = 1, 2, ...$

Case 2. Suppose $\mathcal{L}(Y_n) = \beta(a, a + 1)$ for some $a > 0$, and $p_n = \frac{1}{2}$ for all $n = 1, 2, ...$

In the next two cases the probability of going up decreases as the particle goes up, and the probability of going down decreases as it goes down. In other words, if the particle moves closer to 1, then, for the next step, it is more likely that it moves towards 0; or towards 1 if the particle is closer to 0.

Case 3. Take $\mathcal{L}(Y_n) = \beta(1, b)$ for some $b > 0$, and let now the probability p_n be random and uniformly distributed on the interval $(0, 1 - X_n)$ for $n = 1, 2, ...$

Case 4. Assume that $\mathcal{L}(Y_n) = \beta(a, a + 1)$ for some $a > 0$, and let again p_n be random and uniformly distributed on the interval $(0, 1 - X_n)$for $n = 1, 2, ...$

In the next case it can be seen how the probability of going up increases when the particle

moves up over the time, and the probability of going down increases when it moves down, having the two boundaries $\{0\}$ and $\{1\}$ as absorbing points.

Case 5. Here $\mathcal{L}(Y_n) = \beta(a, b)$ for some a, $b > 0$, and p_n is assumed to be random and uniformly distributed on the interval $(0, X_n)$ for $n = 1, 2, ...$

Let us mention that, for the case when Y_n is uniform in $(0, 1)$ and $p_n = p$ (included in Case 1 or 2), Stoyanov and Pirinsky (2000) has found the ergodic distribution by studying the limit of the moments of X_n and applying Fréchet-Shohat theorem.

The recursive formula (2) can be used to find the moments of X_n. Indeed, from (2) we can obtain the first moment:

$$E\left[X_n\right] = E\left[X_0\right]\left(\frac{b}{a+b}\right)^n + p\frac{a}{a+b}\frac{\left(1-\left(\frac{b}{a+b}\right)^n\right)}{1-\frac{b}{a+b}}.$$

Taking the limit as $n \to \infty$, we conclude that the first moment of the ergodic distribution is equal to p.

However, to find $E\left[X_n^k\right]$ and the limits as $n \to \infty$ for $k = 2, 3, ...$ using the above recursive formula, is not an easy task.

Instead, we derive an equation called a *distributional equation*, for the limiting (ergodic) random quantity. From it, we obtain the moments of the ergodic distribution. Since the ergodic distributions have bounded support in the interval $(0, 1)$, it is uniquely determined by its moments. For Case 2 (in Proposition 3) we find directly the ergodic distribution using the distributional equation.

It is important to recall that in general there are distributions with unbounded support which are different but have the same moments (we will come back to this in Chapter 3). This property is called M-indeterminancy, see Stoyanov (1997).

The following proposition helps for all the cases cited above.

Proposition 1 *Assume that the distribution of Y_n is $\beta(a, b)$ with $a, b > 0$, and let $p_n = p$ for all $n = 1, 2, ...,\ p \in (0, 1)$. Then $\lim_{n\to\infty} X_n \stackrel{d}{=} X$, for some r.v. X with $E[X] = p$ and $E[X^2] = \frac{a+1+2bp}{a+1+2b}p$. Moreover, the moments of the X satisfy the following recursive formula:*

$$E(X^n) = \frac{1}{1-p\frac{b}{b+a}}\ p\ \sum_{i=1}^{n}\binom{n}{i}E\left[X^{n-i}(1-Y)^{n-i}Y^i\right]$$

Proof. First we show that, as $n \to \infty$, the variables X_n converge in distribution to some r.v., say X. Then we find that X is related to the r.v.s. Y and B, the generic independent r.v. such that $Y_n \stackrel{d}{=} Y$ and $B_n \stackrel{d}{=} B$. These three variables satisfy a relation which will be the distributional equation for X. The distribution of X, or its law $\mathcal{L}$, is the desired ergodic distribution.

Let P be the one-step transition probability function of the time-homogeneous Markov chain $\{X_n\}$. Let us show that there is an invariant measure. Suppose that $X_n = x$, $x \in I = (0,1)$, and let $P(x,A) = P(X_{n+1} \in A \mid X_n = x)$ with $A \in \mathcal{B}_I$.

For a function $f \in C_b(I)$, continuous in I and bounded, and any $x \in I$, we have the following sequences of relations:

$$\begin{aligned}\int_I f(y)P(x,dy) &= E\left[f(x(1-Y)+YB)\right] \\ &= E\left[f(x(1-Y)+YB) \mid B=1\right]P(B=1) \\ &\quad + E\left[f(x(1-Y)+YB) \mid B=0\right]P(B=0) \\ &= \int_0^1 \left(f(x(1-y)+y)p + f(x(1-y))q\right) F_{\beta(a,b)}(dy).\end{aligned}$$

The conditions on the integrand and the dominated convergence theorem imply that

$$\int_I f(y)P(x_n,dy) \to \int_I f(y)P(x,dy) \text{ as } n \to \infty \ ,$$

for $\{x_n, n \geq 1\}$, a convergent sequence of numbers in I such that $x_n \to x$ as $n \to \infty$. This means that P satisfies the weak-Feller property (see Appendix: Markov Chains). The first theorem in the Appendix assures that there exists an invariant measure, say μ, corresponding to the kernel P. Moreover, $\lim_{n\to\infty} X_n \stackrel{d}{=} X$, where X is a r.v. with distribution μ, and the three variables X, Y, B satisfy the following distributional equation:

$$X \stackrel{d}{=} X(1-Y) + YB. \tag{5}$$

We are going to use this relation to find the moments of X.

Notice first that in the relation $X_{n+1} = X_n(1-Y_n) + Y_nB_n$, the variables Y_n and B_n are generated independently of the value X_n. Since $Y_n \stackrel{d}{=} Y$, $B_n \stackrel{d}{=} B$, we can consider X, Y and

B to be independent. Now, using the linearity of the expectation, the independence of the random quantities involved, and the fact that $E[B] = p$, $E[Y] = \frac{a}{a+b}$, we easily find that $E[X] = p$. The next is to find the second moment of X. We have

$$E\left[X^2\right] = E\left[(X(1-Y)+YB)^2\right]$$

$$= E[X^2]E[(1-Y)^2] + 2E[X]E[(1-Y)Y]E[B] + E[Y^2]E[B^2].$$

Since $E[Y(1-Y)] = \frac{ab}{(a+b)(a+b+1)}$, $E[Y^2] = \frac{a(a+1)}{(a+b)(a+b+1)}$, $E[(1-Y)^2] = \frac{b(b+1)}{(a+b)(a+b+1)}$, and $E[B^2] = p$, we find $E[X^2] = \frac{\frac{2b}{a+1}p+1}{\frac{2b}{a+1}+1}p = \frac{a+1+2bp}{a+1+2b}p$.

The last statement can be obtained from the equation (5) and the Newton's binomial formula:

$$E[X^n] = E[(X(1-Y)+YB)^n] = E\left[\sum_{i=0}^{n}\binom{n}{i}(X(1-Y))^{n-i}Y^i\right].$$

■

Now we can proceed to the detailed analysis of the cases cited before.

In the next proof we make use of the following properties of beta-distribution (see Gupta and Nadarajah (2004)). If a r.v. ξ is $\beta(a,b)$-distributed, the following property can be easily derived using the Beta function,

$$E[\xi^{k+1-i}(1-\xi)^i] = \frac{a+k-i}{a+k+b}E[\xi^{k-i}(1-\xi)^i] \text{ for } k = 1,2,... \text{ and } i = 0,1,...k. \quad (6)$$

Proposition 2 (Case 1) *Suppose that the distribution of Y_n is $\beta(1,b)$ and $p_n = p \in (0,1)$ for all $n = 1,2,....$ Then, $X_n \xrightarrow{d} X$ as $n \to \infty$, where the r.v. $X \sim \beta(bp, bq)$.*

Proof. We proceed by induction, showing that

$$E\left[X^{k+1}\right] = \frac{bp+k}{b+k}E\left[X^k\right], k = 0,1,... \quad (7)$$

From Proposition 1 we have $E[X] = p$ and $E[X^2] = \frac{bp+1}{b+1}p$, and hence

$$E\left[X^2\right] = \frac{bp+1}{b+1}E[X].$$

Thus we have the first inductive step. Now, we assume that the moment relation (7) holds for all $i \leq k$, and the goal is to show that (7) is true for $k+1$.

Indeed, since $X \stackrel{d}{=} X(1-Y)+YB$, we have

$$E\left[X^{k+1}\right] = E\left[(X(1-Y)+YB)^{k+1}\right]$$

$$= \sum_{i=0}^{k+1} \binom{k+1}{i} E(X^i) E\left[Y^{k+1-i}(1-Y)^i\right] E(B^{k+1-i})$$

(by (6))

$$= \sum_{i=0}^{k} \frac{(k+1)k!}{(k+1-i)(k-i)!i!} E(X^i) \frac{(1+k-i)}{1+k+b} E\left[Y^{k-i}(1-Y)^i\right] E(B^{k+1-i})$$

$$+E(X^{k+1})E((1-Y)^{k+1})$$

(since (7) holds for $i \leq k$)

$$= \frac{(k+1)}{k+1+b}\left(\sum_{i=0}^{k} \frac{k!}{(k-i)!i!} E(X^i)E(Y^{k-i}(1-Y)^i)E(B^{k-i}) - (1-p)E(X^k)E((1-Y)^k)\right)$$

$$+\frac{bp+k}{b+k}\frac{b}{b+k+1}E(X^k)$$

$$= \frac{(k+1)}{k+1+b}\left(E(X^k) - (1-p)\frac{b}{b+k}E(X^k)\right) + \frac{bp+k}{b+k}\frac{b}{b+k+1}E(X^k)$$

$$= \frac{bp+k}{b+k}E(X^k).$$

From (7) we easily conclude that

$$E\left[X^{k+1}\right] = \prod_{i=0}^{k} \frac{bp+i}{b+i}, \quad k = 0, 1, 2, ...$$

Which corresponds to the moments of $\beta(bp, bq)$ as in (4), we conclude that $X \sim \beta(bp, bq)$.

■

The proof of the next result is not based on the moments, instead we work directly with the distributions.

Proposition 3 (Case 2) *We assume now that for any n, the distribution of Y_n is $\beta(a, a+1)$. Then the Markov chain X_n $n = 1, 2, ...$ has an ergodic distribution which is the uniform distribution on $(0, 1)$.*

Proof. The idea is as follows. Let us assume that the ergodic variable X is $U(0,1)$. Then we show that the distribution of the r.v. $X(1-Y)+YB$, is also $U(0,1)$. Finally we use the fact that the distributional equation (5) has only one solution/distribution. Thus we want to show that

$$P(X(1-Y)+YB \le x) = P(X \le x) = x \ \text{ for each } x \in (0,1)\,.$$

Indeed, take any $x \in (0,1)$. Then by the total probability formula,

$$P(X(1-Y)+YB \le x)$$

$$= P(X(1-Y)+YB \le x \mid B=0)P(B=0)+$$

$$P(X(1-Y)+YB \le x \mid B=1)P(B=1)$$

$$= \frac{1}{2}(P(X(1-Y) \le x) + P(X(1-Y)+Y \le x))$$

$$= \frac{1}{2}\left(\int_0^1 P\left(X < \frac{x}{1-y}\right) f_Y(y)dy + \int_0^1 P\left(X < \frac{x-y}{1-y}\right) f_Y(y)dy\right)$$

(using the fact that $X \sim U(0,1)$, by assumption)

$$= \tfrac{1}{2}\left(\int_0^{1-x} \tfrac{x}{1-y} \tfrac{y^{a-1}(1-y)^a}{\int_0^1 y^{a-1}(1-y)^a dy}dy + \int_{1-x}^1 \tfrac{y^{a-1}(1-y)^a}{\int_0^1 y^{a-1}(1-y)^a dy}dy\right)$$
$$+\tfrac{1}{2}\left(\int_0^x \tfrac{(x-y)}{1-y} \tfrac{y^{a-1}(1-y)^a}{\int_0^1 y^{a-1}(1-y)^a dy}dy\right)$$

$$= \tfrac{1}{2}\left(x\int_0^{1-x} \tfrac{y^{a-1}(1-y)^{a-1}}{\frac{1}{2}\int_0^1 y^{a-1}(1-y)^{a-1}dy}dy + \int_{1-x}^1 \tfrac{y^{a-1}(1-y)^a}{\int_0^1 y^{a-1}(1-y)^a dy}dy\right)$$
$$+\tfrac{1}{2}\left(\int_0^x \tfrac{(x-y)}{1-y} \tfrac{y^{a-1}(1-y)^a}{\int_0^1 y^{a-1}(1-y)^a dy}dy\right)$$

$$= \tfrac{1}{2}\left(2x\left(1 - \int_{1-x}^1 \tfrac{y^{a-1}(1-y)^{a-1}}{\int_0^1 y^{a-1}(1-y)^{a-1}dy}dy\right) + \int_{1-x}^1 \tfrac{y^{a-1}(1-y)^a}{\int_0^1 y^{a-1}(1-y)^a dy}dy\right)$$
$$+\tfrac{1}{2}\left(x\int_0^x \tfrac{y^{a-1}(1-y)^{a-1}}{\int_0^1 y^{a-1}(1-y)^a dy}dy - \int_0^x \tfrac{y^a(1-y)^{a-1}}{\int_0^1 y^{a-1}(1-y)^a dy}dy\right)$$

$$= \tfrac{1}{2}\left(2x\left(1 - \int_{1-x}^1 \tfrac{y^{a-1}(1-y)^{a-1}}{\int_0^1 y^{a-1}(1-y)^{a-1}dy}dy\right) + \int_{1-x}^1 \tfrac{y^{a-1}(1-y)^a}{\int_0^1 y^{a-1}(1-y)^a dy}dy\right)$$
$$+\tfrac{1}{2}\left(x\int_0^x \tfrac{y^{a-1}(1-y)^{a-1}}{\frac{1}{2}\int_0^1 y^{a-1}(1-y)^{a-1}dy}dy - \int_0^x \tfrac{y^a(1-y)^{a-1}}{\int_0^1 y^{a-1}(1-y)^a dy}dy\right).$$

Since

$$\int_{1-x}^{1} \frac{y^{a-1}(1-y)^{a-1}}{\int_0^1 y^{a-1}(1-y)^{a-1}dy}dy = \int_0^x \frac{y^{a-1}(1-y)^{a-1}}{\int_0^1 y^{a-1}(1-y)^{a-1}dy}$$

and

$$\int_{1-x}^{1} \frac{y^{a-1}(1-y)^{a}}{\int_0^1 y^{a-1}(1-y)^{a}dy}dy = \int_0^x \frac{y^{a}(1-y)^{a-1}}{\int_0^1 y^{a-1}(1-y)^{a}dy}dy,$$

these terms cancel out each other and the whole expression equals x. Hence the distribution of $X(1-Y)+YB$ is effectively $U(0,1)$. ■

Remark. In general, for arbitrary a, b and p, the ergodic distribution of the Markov chain $\{X_n, n = 0, 1, 2...\}$ may be different from beta-distribution. If however, the parameters of the Markov chain, say a', b', p' are close to the values a, b, p in the previous cases, then the distribution of X (i.e. the limiting distribution of the Markov Chain or, its ergodic distribution) would be "close" to the distribution $\beta(\frac{2b}{a+1}p, \frac{2b}{a+1}q)$. If we plug the parameters of Cases 1 and 2 into $\beta(\frac{2b}{a+1}p, \frac{2b}{a+1}q)$, we obtain the corresponding ergodic distributions.

For Case 3, 4 and 5 we have that p_n is a function of X_n, so B_n is not independent of X_n. We are going to analyze these cases.

Similarly to Proposition 1 we can prove that, as $n \to \infty$, $X_n \xrightarrow{d} X$ for some r.v. X. In this case p_n is a continuous function x, where $x = X_n$, on $x \in [0, 1]$. Actually, $p_n = 1 - X_n$. As in the proof of Proposition 1, we can find the first and the second moments of the stationary distribution as follows:

$$E[X] = E[X(1-Y)+YB] = E[X]E[1-Y] + E[Y]E[B],$$

hence $E[X] = \frac{1}{2}$. Then

$$E[X^2] = E\left[(X(1-Y)+YB)^2\right]$$

$$= E[X^2]E[(1-Y)^2] + 2E[XB]E[(1-Y)Y] + E[Y^2]E[B^2]$$

$$= E[X^2]E[(1-Y)^2] + 2E[X]E[1-X]E[(1-Y)Y] + E[Y^2]E[B^2].$$

Therefore $E[X^2] = \frac{1}{2}\frac{a+b+1}{a+2b+1}$.

Now we can use these observations in order to prove the following result.

Proposition 4 (Case 3) *Suppose the distribution of Y_n is $\beta(1, b)$ and let $p_n = 1 - X_n$ for $n = 1, 2, ...$ Then $X_n \xrightarrow{d} X$ as $n \to \infty$, where $X \sim \beta(\frac{b}{2}, \frac{b}{2})$.*

Proof. As in Proposition 2, we use induction arguments. We verify that the moments of the stationary distribution are equal to the moments of the $\beta(\frac{b}{2}, \frac{b}{2})$-distribution. This will follow from the recurrent relation

$$E\left[X^{k+1}\right] = \frac{\frac{b}{2}+k}{b+k} E\left[X^k\right], \ k = 0, 1, ... \tag{8}$$

We have just stated that $E[X^2] = \frac{(\frac{b}{2}+1)}{(b+1)}\frac{1}{2} = \frac{(\frac{b}{2}+1)}{(b+1)}E[X]$, so the first induction step is checked. Notice that, for any i and j, we have

$$E[X^i B^j] = E[X^i B^j \mid B = 1]P(B = 1) + E[X^i B^j \mid B = 0]P(B = 0),$$
$$E[X^i](1 - E[X]) = E[X^i] \times \frac{1}{2} = E[X^i]E[B^j].$$

From these relations we see that the induction proof remains the same as in Proposition 2 when $p = \frac{1}{2}$. ■

Remark. Similarly to Case 2, in Case 4 the invariant distribution is $U(0, 1)$ because $P(B = 1) = 1 - E[X] = \frac{1}{2}$, so the proof of Proposition 3 remains the same.

Remark. For Case 5, if $X_n = 1$, $P(B_n = 1) = 1$, then $X_{n+1} = X_n(1 - Y_n) + Y_n B_n = 1$; but if $X_n = 0$, then $X_{n+1} = 0$. Hence $\{0\}$ and $\{1\}$ are absorbing points and therefore $\delta_{\{0\}}$ and $\delta_{\{1\}}$ are invariant measures (where δ_{x_0} stands for the delta-Dirac at point x_0).

Remark. In Rao (1984) and Meyn & Tweedie (1996), it is pointed out that the ergodic distribution of Markov chains often comes as the solution of an integral equation. In our cases we have that the ergodic distribution is the one with distribution function F that solves the integral equation

$$F(x) = \int_0^1 P\left((u(1 - Y) + YB) < x\right) F(du), \ x \in (0, 1),$$

where $Y \sim Beta$ and $B \sim$ Bernoulli. Therefore, we can see the ergodic distribution of the Markov chain under study as the unique solution of an integral equation. If the distribution function F has density f, the integral equation can be seen as

$$\int_0^x f(u)du = \int_0^1 P\left((u(1 - Y) + YB) < x\right) f(u)du. \tag{9}$$

All this motivated us to include a chapter showing the use of Monte-Carlo techniques to approximate solutions of integral equation.

2.2 Appendix: Markov Chains

In this section we state some basic notation and results from ergodic theory which we use in this Chapter (see Hernández-Lerma & Lasserre (2003) and Meyn & Tweedie (1996) for details).

Let $X = \{X_n, n = 0, 1, ...\}$ be a sequence of r.v.s. on the probability space $(\Omega, \mathcal{F}, P)$ with values in the measurable space $(I, \mathcal{B}_I)$. We say that X is a discrete-time homogeneous Markov chain (MC) if

$$P(X_{n+1} \in B \mid X_0, ..., X_{n-1}, X_n) = P(X_{n+1} \in B \mid X_n) \text{ for all } B \in \mathcal{B},\ n = 0, 1, ...$$

This is called the Markov property. It means that conditional on its present value, the future is independent of the past.

An essential characteristic of the MC X is its $1 - step$ transition probability function

$$P(x, B) := P(X_{n+1} \in B \mid X_n = x)\ x \in I,\ B \in \mathcal{B}_I.$$

It is also called a probability kernel, or simply a kernel.

Definition A probability measure ν on I is said to be an invariant measure for P, and also for the MC X, if

$$\nu(B) = \int_I P(x, B)\nu(dx) \text{ for all } B \in \mathcal{B}_I.$$

A set $B \in \mathcal{B}_I$ is invariant for P if $P(x, B) = 1$ whenever x is in B. An invariant measure ν is said to be ergodic if $\nu(B) = 0$ or $\nu(B) = 1$ for every invariant set B.

Useful criteria to determine if a MC admits invariant measures are the weak- and the strong-Feller properties. We say that the transition probability function P is weak-Feller if for every sequence $\{x_n\}$ in I such that $x_n \to x \in I$ as $n \to \infty$, and every bounded continuous function f on I, the following holds:

$$\int f(y)P(x_n, dy) \to \int f(y)P(x, dy) \text{ as } n \to \infty.$$

Similarly, P is called strong-Feller if for every bounded measurable function f on I,

$$\int f(y)P(x_n, dy) \to \int f(y)P(x, dy).$$

Theorem Let I be a compact metric space, and P a weak-Feller transition probability function of a Markov chain on I. Then P admits an invariant measure.

In the following result we use the notation

$$P^{(n)}f(x) = E\left[f\left(X_n\right) \mid X_0 = x\right].$$

Theorem(Birkhoff´s individual ergodic theorem) Let μ be an invariant measure for the Markov chain with transition probability function P. For every function $f \in L_1(\mu)$, there exists a function $f^* \in L_1(\mu)$ such that

$$(i)\ P^{(n)}f \to f^* \text{ as } n \to \infty\ \mu - a.s. \text{ and } (ii)\ \int_I f^* d\mu = \int_I f d\mu.$$

3 Models based on Random Sums of Random Variables

Motivation. In different areas such as biology, physics, finance, sociology we have to study or analyze variables or characteristics of phenomena which are random in nature and evolve over time, i.e. these are stochastic dynamical phenomena. The observations taken over time form *time series* and the process that generates these observations is called a *stochastic process* or a *random process*. If in general the time series shows any presence of a pattern, we say that there is *statistical regularity* on the series. This is what we called stability property of a random dynamical system. Whitt (2002) presents simulations of probabilistic models to show intuitively the meaning of the statistical regularity. In Probability theory, any statistical regularity is a reflection of the existence of a limit of some rescaled stochastic process. The rescaling corresponds to a specific transformation of random quantities that define the random system.

In this chapter, we study the statistical regularity of time series generated by specific type of stochastic processes. These processes are a general version of the CTRW, we call them **Generalized Continuous Time Random Walks**, GCTRW for short. One of the motivations is based on a model of random sums of random variables. It is important to note that the behaviour of the model resembles a discrete scheme, even though the time parameter is continuous.

3.1 The compound Poisson process

The compound Poisson process is usually defined as

$$S_t = \sum_{i=1}^{N_t} X_i,\ t \geq 0,$$

where $X_1, X_2, ...$ is a sequence of i.i.d. r.v. and $N_t,\ t \geq 0$, is a Poisson process.

In this chapter we want to study limits of continuous-time stochastic processes defined similarly as the compound Poisson process. Consider the following sequence of compound

Poisson processes:

$$S_t^{(n)} = \sum_{i=1}^{N_t^{(n)}} X_i^{(n)},\ n = 1, 2, ...,\ t \geq 0.$$

Here $N_t^{(n)},\ t \geq 0$, is a homogeneous Poisson process with parameter n, and $X_i^{(n)}$ $i = 1, 2, ...$ are Bernoulli independent r.v.s (also independent of the Poisson process). We assume that $X_i^{(n)} = \pm\frac{1}{n}$ each with probability $\frac{1}{2}$. It can be proved that for any fixed t, $t \geq 0$, the limit $\lim_{n\to\infty} S_t^{(n)}$ exits and it is normally distributed $N(0, t)$. Moreover, the sequence of stochastic processes $\{S^n\}$ converge (in an appropriate sense) to a Brownian Motion.

In the above model, the Poisson process $N^{(n)}$ corresponds to a counting process with exponentially distributed *interarrival times*.

Our goal is to study a similar but more general case when $X_i^{(n)}$ are arbitrary r.v.s which in some sense tend to be "small" and $N^{(n)}$ is an arbitrary counting process, tending to be "big". We combine the two random effects, one "small" and one "big", to obtain in the limit a proper random or non-random quantity.

3.2 Generalized continuous time random walks

After the above discussion, we turn to stochastic processes that behave like the Compound Poisson process. We have in mind the following description: Suppose that at time t the process has value S_t and stays there for some random time τ (called interarrival time), then it moves up or down according to the value of a r.v. $X \in \mathbb{R}$. Hence at time $t + \tau$ the value of the process will be $S_{t+\tau} = S_t + X$. When the random time $\tau \sim Exp$, the process S is a Compound Poisson process. If τ is a positive r.v. with arbitraty distribution, then the non-negative and non-decreasing process $N^{(n)}$ is a renewal process, or counting process in which case the process S is usually called Continuous-time Random Walk (CTRW), see Scalas (2005) for a general review. Other names for S are *Increment process* (Korolyuk and Limnios (2004)) or *Renewal-Reward processes* (Grimmett and Stirzarker (2004) and Whitt (2002)).

Generally τ represents the time when the process remains idled, and therefore it is a positive variable.

Definition The CTRW process is defined as follows:

$$S_t = \sum_{i=1}^{N_t} X_i,\ t \geq 0,$$

where $N_t = \max\left\{k : \sum_{i=1}^{k} \tau_i \leq t\right\}$, $X_1, X_2, ...$ are r.v.s with values in $\mathbb{R}$ and $\tau_1, \tau_2, ...$ are positive r.v.s (the interarrival times) and $S_0 = 0$. The process N_t, $t \geq 0$ is called the *Inverse process* to the sum $\sum_{i=1}^{n} \tau_i$ (see Whitt (2002) p.201), and it is worth to notice that we have

$$N_t = \min\left\{k : \sum_{i=1}^{k} \tau_i > t\right\} - 1.$$

The classical Compound Poisson process is a particular example when the interarrival times are exponential r.v.s.

Processes of this kind have been studied by many authors. There is a huge literature, books and articles, some of them are included in our list of references. Mitov and Nadarajah (2004) deals with the limit behaviour in distribution of random sums, i.e. for fixed time t. Lin and Stoyanov (2002) and Gut (2003) study the so-called moment determinacy of the distribution of these random sums. Asmussen (2002) uses this model applied to ruin theory. Eberlein and Hammerstein (2004) analyze different aspects and applications of this model in mathematical finance. Weron (2002) uses this model for pricing options.

Here we will study a more general model in which $\tau_1, \tau_2, ...$ are arbitrary r.v.s. in $\mathbb{R}$, so we allow τ_i to take also negative values. We have the following definition.

Definition (GCTRW) Suppose $\{X_i, i = 1, 2, ...\}$ and $\{\tau_i, i = 1, 2, ...\}$ are sequences of r.v.s. Then following process

$$S = \left\{S_t = \sum_{i=1}^{N_t - 1} X_i,\ t \geq 0\right\},$$

is said to be a Generalized Continuous-time Random Walks (GCTRW).
Here $N = \{N_t,\ t \geq 0\}$ is the inverse process triggered by $\{\tau_i\}_{i=1}^{\infty}$ and defined as follows:

$$N_t = \min\left\{k : \sum_{i=1}^{k} \tau_i > t\right\},\ t \geq 0.$$

Notice that N_t+1 corresponds to the first passage time of the continuous-time stochastic process $\left\{\sum_{i=1}^{\lfloor t \rfloor} \tau_i, t \geq 0\right\}$, which is a simple CTRW.

It is well-known that under some conditions and a proper type of convergence, the sequence of Compound Poisson processes has a limit which is a Brownian Motion. More results are available in the literature about the limit of the scaled CTRW. Under some assumptions the limit of scaled CTRW is a Lévy process (see Whitt (2002), Becker-Kern et al (2004)). This, however, is not always the case as shown by Meerschaert and Scheffler (2001).

Our goal now is to find conditions under which a sequence of GCTRW, after appropriate rescaling, has a proper limit-process.

We study the limit behaviour of the following sequence of GCTRW:

$$S_t^{(n)} = \sum_{i=1}^{N_t^{(n)}} X_i^{(n)},\ t \geq 0,\ n = 1, 2, ...$$

Here $\left\{X_i^{(n)}, i, n \in \mathbb{N}\right\}$ is a double array of r.v.s, and $\left\{N_t^{(n)}, n \in \mathbb{N}\right\}$ is a sequence of inverse processes, where for each n the process $N_t^{(n)}$ is triggered by the sum of $\left\{\tau_i^{(n)}, i \in \mathbb{N}\right\}$.

We cite some results related to this problem.

Korolyuk and Limnios (2004) work with a general version of the Compound Poisson process in risk theory, they prove prove weak convergence of the scaled version to a Lévy process.

In Becker-Kern et al (2004), Meerschaert and Scheffler (2004), CTRW is used to model the motion of particles. Assuming that the r.v.s. involved have a domain of attraction to stable distributions, they prove convergence in distribution of the sequence of CTRW.

In Chapter VIII of Jacod and Shiryaev (2003), using the triplet of *Characteristics of a Semimartingales,* it is shown how some processes, that include the GCTRW, converge to a process with *independent increments*. This is a more general class of stochastic processes than the so-called *Lévy Processes*.

We are going to show that, under appropriate conditions, the sequence of processes GCTRW converges to a Lévy process of the type of subordination of a stable process.

3.3 Convergence of stochastic processes

We have to make precise the concept of convergence of stochastic processes.

We are given a probability space $(\Omega, \mathcal{F}, P)$, where all random variables or processes are defined. If X is a random element on this space, then the values of X can be real numbers, $n-$dimensional vectors or functions. In this section S is the space of continuous functions C or the space of cádlág functions D which are subspaces of the space of all functions $\mathbb{R}^{[0,\infty)}$.

If X is a random element and A any measurable set in S, $P(\{\omega : X(\omega) \in A\})$ is the probability measure on S or the law of X (Notation: $\mathcal{L}(X)$). For a r.v. X, $F(x) = P_X((-\infty, x]) = P(\{\omega : X(\omega) \in (-\infty, x]\})$ is a d.f., and for a random vector X in $\mathbb{R}^k$ the d.f. is

$$F_X(x_1, ..., x_k) = P_X((-\infty, x_1] \times ... \times (-\infty, x_k])$$

$$= P(\{\omega : X_1(\omega) \in (-\infty, x_1], ..., X_k(\omega) \in (-\infty, x_k]\}).$$

The d.f.s are used to define convergence in distribution of r.v.s and random vectors. If X is a stochastic process, then the probability measure P_X is a more abstract object. It is necessary to extend the concept of the convergence of the d.f.s for stochastic processes

Suppose $X^{(1)}, X^{(2)}, ...$ is a sequence of r.v.s. (or k-dimensional random vectors), $X^{(n)}$ has a d.f. F_n, and let X be another r.v. (or k-dimensional random vector) with d.f. F. We say that $\left\{X^{(n)}\right\}$ convergences in distribution to X (notation: $X^{(n)} \xrightarrow{d} X$) if $F_n(x) \to F(x)$ as $n \to \infty$ for all $x \in \mathbb{R}$ (or $\mathbb{R}^k$), which are points of continuity of F.

Theorem Let $X^{(1)}, X^{(2)}, ...$ a sequence of r.v.s and $F_1, F_2, ...$ their d.f.s The following two statements are equivalent:
(1) $E\left[f(X^{(n)})\right] \to E\left[f(X)\right]$ as $n \to \infty$ for all bounded continuous functions $f : \mathbb{R} \to \mathbb{R}$;
(2) $F_n(x) \to F(x)$ *as* $n \to \infty$ for any x which is a point of continuity of F.

Based on this result, the concept of convergence in distribution can be extended to cover more general sequences of random elements when it is not straightforward to talk about d.f.s.

Definition Given a sequence of random elements $X^{(1)}, X^{(2)}, ...$ in the space S endowed

with a topology which generates a sigma algebra. We say that $X^{(n)} \xrightarrow{d} X$ as $n \to \infty$ if

$$E\left[f(X^{(n)})\right] \to E\left[f(X)\right] \text{ as } n \to \infty$$

for all bounded continuous functions f from S to $\mathbb{R}$.

This can be written as follows:

$$\int_S f(\omega)P_n(d\omega) \to \int_S f(\omega)P(d\omega) \text{ as } n \to \infty,$$

where $P, P_1, P_2, ...$ are the corresponding probability measures generated by $X, X^{(1)}, X^{(2)}, ...$ This is called convergence in distribution or weak convergence.

We can think of the probability measures on one space S as a space of elements with a topology. Given a measure μ on S, an $\epsilon-$ neighborhood of μ is the set of all measures ν on S such that

$$\left|\int_S f(\omega)\mu(d\omega) - \int_S f(\omega)\nu(d\omega)\right| < \epsilon$$

for a finite collection of continuous functions $f : S \to \mathbb{R}$. These sets topologize the space of measures on S. When S is separable and complete, this topology can be defined by the so-called *Prohorov metric* (see Billingsley (1999)). So, when we talk about the weak-convergence of a sequence of random elements, we actually consider the convergence of the probability measures of the corresponding random elements in the topologized space of measures.

These concepts allow us to determine convergence in distribution for stochastic processes. Let $\left\{X^{(n)},\ n \geq 1\right\}$ be a sequence of stochastic processes. We have that $X^{(n)} \xrightarrow{d} X$ as $n \to \infty$, if

$$\int_S f(\omega)P_n(d\omega) \to \int_S f(\omega)P(d\omega),\ n \to \infty,$$

for all bounded continuous function f on S. Here P_n is the probability measure corresponding to $X^{(n)}$, $n = 1, 2, 3, ...$ and P of X. It is also said that the measures P_n converges weakly to P and this is denoted by $P_n \Longrightarrow P$.

The following result provides a method to verify weak convergence.

Theorem (Billingsley (1999), p58) Let $\left\{X^{(n)},\ n \geq 1\right\}$ be a sequence of stochastic processes. We have that $X^{(n)} \xrightarrow{d} X$ as $n \to \infty$, if the corresponding sequence of measures $\{P_n,\ n \geq 1\}$ is relatively compact and the finite dimensional distributions of $X^{(n)}$ converge to those of X.

For a stochastic process X, the *finite dimensional distributions* are

$$\left\{F(t_1, ..., t_k;\ x_1, ..., x_k),\ t_1, ..., t_k \geq 0,\ \ (x_1, ..., x_k) \in \mathbb{R}^k\right\} =$$

$$\left\{P\left(\{\omega : X_{t_1}(\omega) \in (-\infty, x_1], ..., X_{t_k}(\omega) \in (-\infty, x_k]\}\right),\ t_1, ..., t_k \geq 0,\ \ (x_1, ..., x_k) \in \mathbb{R}^k\right\}.$$

Another way to verify the relative compactness is to use the property called tightness. A family of probability measures $\{P_n,\ n \geq 1\}$ on S is *tight* if for every $\epsilon > 0$ there exists a compact set $K \subset S$ such that $P_n(K) > 1 - \epsilon$ for all n.

It is important to mention that if S is separable and complete, then relatively compactness and tightness are equivalent properties. There are different ways to prove tightness, see Billingsley (1999).

Another way to prove week convergence is to use the characteristic triplet of semi-martingales (see Jacod and Shiryaev (2003)).

The following is an important result for convergence of continuous function of random elements (see Whitt (2002), p.85 or Billingsley (1999), p.20).

Theorem (Continuous-Mapping Theorem) Consider a sequence of random elements $\left\{X^{(n)}, n \geq 1\right\}$ in the space S, endowed with a topology that generates the sigma algebra. Let X be another random element in S such that $X^{(n)} \xrightarrow{d} X$ as $n \to \infty$. If g is a continuous function from S to another topological space S^*, then

$$g(X^{(n)}) \xrightarrow{d} g(X^{(n)}) \text{ as } n \to \infty.$$

3.3.1 Example of convergence of compound Poisson processes

Consider the following example of a sequence of stochastic processes in the space D that converges in distribution to a Brownian motion which, as is well-known is in the space C.

Let, for each n, $n = 1, 2, ...$, $\left\{X_i^{(n)}, i = 1, 2, ...\right\}$ be a sequence of independent Bernoulli r.v.s of the following form:

$$P\left(X_i^{(n)} = \pm\frac{1}{\sqrt{n}}\right) = \frac{1}{2},\ i = 1, 2, ...,\ n = 1, 2, ...;$$

Let $\left\{N^{(n)}, n \geq 1\right\}$ be a sequence of Poisson processes, $N^{(n)}$ with parameter n, they all are independent of $\left\{X_i^{(n)}, i = 1, 2, ...\right\}$ for each $n = 1, 2, ...$

We want to analyze the limit of the sequence of *compound Poisson processes*

$$S^{(n)} = \left\{ S_t^{(n)} = \sum_{i=1}^{N_t^{(n)}} X_i^{(n)},\ t \geq 0 \right\},\ n = 1, 2, ...$$

Proposition 5 *For the sequence* $\left\{S^{(n)},\ n = 1, 2, ...\right\}$ *defined above, we have convergence of the finite dimensional distributions to the Brownian motion* B*, i.e.*

$$S^{(n)} \overset{f.d.d}{\to} B.$$

This is a well-known result (see Billingsley (1999), p.154). We will present a proof of it using the *Linderberg´s CLT* (in the Appendix "Sums of random variables") and the following proposition.

Proposition 6 *Let* N *be a Poisson random variable with parameter* λ *and* $X_1, X_2, ...$ *be independent Bernoulli r.v.s, where, for fixed* $\Delta > 0$*, and* $p \in (0, 1)$*,*

$$P\left(X_i = \Delta\right) = p \text{ and } P\left(X_i = -\Delta\right) = 1 - p = q,\ i = 1, 2, ...$$

Then

$$\sum_{i=1}^{N} X_i \overset{d}{=} \Delta(N_1^* - N_2^*),$$

where N_1^* *and* N_2^* *are independent Poisson r.v.s with parameters* $p\lambda$ *and* $q\lambda$*, respectively.*

Proof. The proof is just checking that the ch.f. of $\sum_{i=1}^{N} X_i$ is equal to the ch.f. of $\Delta(N_1^* - N_2^*)$, in which case the distributions of the r.v.s are the same. ■

Proof. (of Proposition 5) Without loss of generality, fix $t = 1$. From Proposition 6 and the fact that a sum of independent Poisson r.v.s is again Poisson, we have the following:

$$S_1^{(n)} \stackrel{d}{=} \sum_{i=1}^{n} \frac{1}{\sqrt{n}}(N_i^{(n)} - M_i^{(n)}),$$

where $N_i^{(n)}$ and $M_i^{(n)}$ $i = 1, 2, ...$ are independent Poisson r.v.s each with parameter $\frac{1}{2}$. Now we use the Linderberg´s theorem to this sum of r.v.s. The first two conditions can be easily checked.

Following the notation in the Linderberg´s theorem, we define, for $\epsilon \geq 0$, and any n, i,

$$E_{n,i,\varepsilon} := \int_{|x|>\varepsilon} x^2 d\mu_{n,i}(x) = \sum_{\{x:|x|>\varepsilon \text{ and } x\in\mathbb{S}_n\}} x^2 P\left[\frac{1}{\sqrt{n}}(N_i^{(n)} - M_i^{(n)}) = x\right].$$

Here $\mu_{n,i}(x)$ is the probability measure of the random quantity $\frac{1}{\sqrt{n}}(N_i^{(n)} - M_i^{(n)})$ and $\mathbb{S}_n := \left\{s : s = \frac{m}{\sqrt{n}}, m \in \mathbb{Z}\right\}$, which is a discrete subset of the real line, is the set of values of the r.v. $\frac{1}{\sqrt{n}}(N_i^{(n)} - M_i^{(n)})$. By the symmetry property of the r.v. $\frac{1}{\sqrt{n}}(N_i^{(n)} - M_i^{(n)})$, we obtain

$$\begin{aligned} E_{n,i,\varepsilon} &:= 2 \sum_{\{x:x>\varepsilon \text{ and } x\in\mathbb{S}_n\}} x^2 P\left[\frac{1}{\sqrt{n}}(N_i^{(n)} - M_i^{(n)}) = x\right] \\ &= 2 \sum_{\{x:x>\varepsilon \text{ and } x\in\mathbb{S}_n\}} x^2 P\left[N_i^{(n)} - M_i^{(n)} = x\sqrt{n}\right]. \end{aligned}$$

Using Proposition 6 again, we have

$$P\left[N_i^{(n)} - M_i^{(n)} = m\right] = P\left[\sum_{i=1}^{N} X_i = m\right],$$

where N is a Poisson r.v. with parameter 1, and X_i, $i = 1, 2, ...$ are Bernoulli independent r.v.s taking values 0 or 1 each with probability $\frac{1}{2}$. Then, by the total probability formula, we have that

$$\begin{aligned} P\left[\sum_{i=1}^{N} X_i = m\right] &= 2\sum_{k\geq m} P\left[\sum_{i=1}^{N} X_i = m \mid N = k\right] P[N = k] \\ &= 2\sum_{k\geq 0} P\left(\sum_{i=1}^{m+2k} X_i = m\right) P[N = m + 2k] \end{aligned}$$

$$= 2\sum_{k\geq 0}\binom{m+2k}{m}\left(\frac{1}{2}\right)^{m+2k}\left(\frac{e^{-1}}{(m+2k)!}\right)$$
$$= 2\frac{1}{m!}\left(\frac{1}{2}\right)^{m}e^{-1}\sum_{k\geq 0}\frac{1}{(2k)!} \leq 2\frac{1}{m!}\left(\frac{1}{2}\right)^{m}.$$

From this we have that

$$E_{n,i,\varepsilon} \leq 2\sum_{m>\varepsilon\sqrt{n}} m^2\left(2\frac{\left(\frac{1}{2}\right)^m}{m!}\right) = 2\sum_{m>\varepsilon\sqrt{n}} m(m-1+1)\left(2\frac{\left(\frac{1}{2}\right)^m}{m!}\right)$$
$$= \sum_{m>\varepsilon\sqrt{n}}\frac{\left(\frac{1}{2}\right)^{m-2}}{(m-2)!} + 2\sum_{m>\varepsilon\sqrt{n}}\frac{\left(\frac{1}{2}\right)^{m-1}}{(m-1)!}$$
$$= \left(\frac{\left(\frac{1}{2}\right)^{r-2}}{(r-2)!}+2\right)\sum_{m>\varepsilon\sqrt{n}}\frac{\left(\frac{1}{2}\right)^{m-1}}{(m-1)!}$$

(here r is the first integer greater than $\varepsilon\sqrt{n}$)

$$= \left(\frac{\left(\frac{1}{2}\right)^{\left[\varepsilon\sqrt{n}\right]+1-2}}{(\left[\varepsilon\sqrt{n}\right]+1-2)!}+2\right)\left(2-\frac{\left(1-\left(\frac{1}{2}\right)^{\left[\varepsilon\sqrt{n}\right]}\right)}{1-\frac{1}{2}}\right)$$
$$= \left(\frac{\left(\frac{1}{2}\right)^{\left[\varepsilon\sqrt{n}\right]-1}}{(\left[\varepsilon\sqrt{n}\right]-1)!}+2\right)\left(\frac{1}{2}\right)^{\left[\varepsilon\sqrt{n}\right]-1},$$

where $[\varepsilon\sqrt{n}]$ is the integer part of $\varepsilon\sqrt{n}$. Therefore

$$\lim_{n\to\infty}\sum_{i=1}^{k(n)} E_{n,i,\varepsilon} \leq \lim_{n\to\infty} n\left(\frac{\left(\frac{1}{2}\right)^{\left[\varepsilon\sqrt{n}\right]-1}}{(\left[\varepsilon\sqrt{n}\right]-1)!}+1\right)\left(\frac{1}{2}\right)^{\left[\varepsilon\sqrt{n}\right]} = 0.$$

Since the three conditions in the Lindeberg´s theorem are satisfied, we can conclude that $S_1^{(n)} \xrightarrow{d} Z$ as $n\to\infty$, where Z is a r.v. normally distributed with mean 0 and variance 1. A similar conclusion is true for times t different from 1: $\lim_{n\to\infty} S_t^{(n)}$ is a normally distributed r.v. with mean 0 and variance t.

Recalling now that the process $S^{(n)}$ has independent increments, it is straightforward to obtain the convergence of the finite dimensional distributions of $S^{(n)}$ to the corresponding finite dimensional distributions of the Brownian motion.

■

3.4 Limit of the scaled GCTRW

Our goal is to study the limit of the scaled GCTRW (recall Defintion (GCTRW))

$$S = \left\{ S_t = \sum_{i=1}^{N_t - 1} X_i,\ t \geq 0 \right\},$$

where

$$N_t = \min \left\{ k : \sum_{i=1}^{k} \tau_i > t \right\},\ t \geq 0.$$

The first thing we can say is that the trajectories of this processes are functions in the space $D([0,\infty))$ and in general, they do not have independent increments and, they are not Markov processes.

3.4.1 The stable conditions

We impose the following conditions on the GCTRW, which we call the *stable conditions:*. There exists a sequence of increasing positive numbers $\{c_i, i = 1, 2, ...\}$ and sequences of real numbers $\{a_i, i = 1, 2, ...\}$ and $\{b_i, i = 1, 2, ...\}$ such that:

(i) for r.v.s $X_1, X_2, ...$, i.i.d., $\sum_{i=1}^{n} \frac{X_i}{c_n} + a_n$ converges to a stable r.v., and,

(ii) for r.v.s $\tau_1, \tau_2, ...$, i.i.d, and independent from $\{X_i\}$, $\sum_{i=1}^{n} \frac{\tau_i}{c_n} + b_n$ converges to a stable r.v.

That is, X_1 and τ_1 belong to the domain of attraction of stable laws. Notice that both share the same scaling factor c_n, this is crucial for our analysis.

Now, consider the following sequence of GCTRW:

$$S^{(n)} = \left\{ S_t^{(n)} = \sum_{i=1}^{N_t^{(n)} - 1} X_i^{(n)}, t \geq 0 \right\},\ n = 1, 2, ...$$

with

$$X_i^{(n)} \stackrel{d}{=} \frac{X_1}{c_n} + \frac{a_n}{n} \text{ and } \tau_i^{(n)} \stackrel{d}{=} \frac{\tau_1}{c_n} + \frac{b_n}{n}.$$

With these assumptions we can find the process-limits of the following random walks in continuous time (simple CTRWs):

$$X^{(n)} = \left\{ X_t^{(n)} = \sum_{i=1}^{\lfloor nt \rfloor} X_i^{(n)},\ t \geq 0 \right\} \text{ and } Z^{(n)} = \left\{ Z_s^{(n)} = \sum_{i=1}^{\lfloor nt \rfloor} \tau_i^{(n)},\ s \geq 0 \right\}.$$

Under the stable conditions it is known that $X^{(n)} \xrightarrow{d} X$ and $Z^{(n)} \xrightarrow{d} Z$ in the J_1 topology, where X and Z are stable processes (see Whitt (2002)).

Proposition 7 *For the random sequences of processes* $\left\{S^{(n)}\right\}$ *and the conditions described above, we have that*

$$S^{(n)} \xrightarrow{f.d.d} S \text{ as } n \to \infty,$$

where S *is the stable process* X *subordinated to* Y*, and* Y *is the first passage time of the stable process* Z*.*

Proof. Let us show first that $S_t^{(n)} \xrightarrow{d} X_{Y_t}$, as $n \to \infty$, for each $t \geq 0$. We know that for each $t \geq 0$, $X_t^{(n)} \xrightarrow{d} X_t$ as $n \to \infty$. Also, for fixed t and any positive number r

$$\lim_{n\to\infty} P\left(\frac{N_t^{(n)} - 1}{n} \leq r\right) = \lim_{n\to\infty} P\left(\min\left\{k : \sum_{i=1}^{k} \tau_i^{(n)} > t\right\} \leq nr + 1\right) \tag{10}$$

$$= \lim_{n\to\infty} P\left(\sum_{i=0}^{k} \tau_i^{(n)} > t, \text{ for some } k = 1, 2, ..., nr + 1\right)$$

$$= \lim_{n\to\infty} P\left(\sum_{i=0}^{\lfloor ns+1 \rfloor} \tau_i^{(n)} > t, \text{ for some } s \in [0, r]\right)$$

$$= P\left(Z_s > t, \text{ for some } s \in [0, r]\right).$$

This corresponds to the probability that the first passage time of the stable process $\{Z_s, s > 0\}$ to the set (t, ∞) occurs before time r. Hence, for fixed, t, we have that $\frac{N_t^{(n)}}{n} \xrightarrow{d} Y_t$ as $n \to \infty$. By the Transfer Theorem (in the Appendix "Random sums of random variables"),

$$S_t^{(n)} \xrightarrow{d} X_{Y_t} \text{ as } n \to \infty \text{ for each } t \geq 0.$$

Since $N_s^{(n)} \leq N_t^{(n)}$ always, for $s < t$ and for all n, then $Y_s \leq Y_t$. This means that Y is an increasing process; the process

$$S = \{X_{Y_t}, t \geq 0\} \tag{11}$$

is well defined as a subordination.

If $s < t$, we have that

$$S_t^{(n)} = \sum_{i=1}^{N_t^{(n)} - N_s^{(n)}} X_i^{(n)} + \sum_{i=1}^{N_s^{(n)}} X_i^{(n)} \text{ for any } n$$

and, by the Transfer Theorem again,

$$\lim_{n\to\infty} S_t^{(n)} = X_{Y_t - Y_s} + X_{Y_s} \text{ for all } s \text{ and } t,\ s < t.$$

Since $Y_t - Y_s$ is a stopping time (see Appendix "First passage time of a Lévy process"), we can apply the *Strong Markov property* (see Applebaum (2004), p.83) to see that

$$X_{Y_s} \stackrel{d}{=} X_{Y_t - Y_s + Y_s} - X_{Y_t - Y_s},\ \text{thus}$$

$$S_t \stackrel{d}{=} X_{Y_t - Y_s} + X_{Y_s}.$$

Then, it is possible to show that $\left(S_{t_1}^{(n)}, ..., S_{t_k}^{(n)}\right) \stackrel{d}{\to} (S_{t_1}, ..., S_{t_k})$ as $n \to \infty$ for any finite collection of times $t_1, ..., t_k,\ k = 2, 3, ...$ This proves converges of the corresponding finite dimensional distributions. ■

3.4.2 Particular cases for the stable conditions

From the previous section we know that the limit of the scaled GCTRW is the subordination S of the form X_Y. If Y is a Lévy process, then S is a Lévy process, because X is a Lévy process (Sato (1999), p.197). We want to find conditions to assure that Y, i.e. the first passage time of the stable process Z, is a Lévy process.

Throughout this section we denote by $S_\alpha(\sigma, \beta, \mu)$ the stable distribution that characterizes Z, and (A_Z, ν_Z, γ_Z) its triplet.

Remark. We know that if $P(Z_s \leq Z_{s-} \text{ for every } s > 0) = 1$ and $P(\limsup_{s\to\infty} Z_s = \infty) = 1$, then Y is a Lévy process with strictly increasing trajectories (see Theorem (First Passage) in the Appendix). The first condition is equivalent to $\nu_Z((0, \infty)) = 0$. We can analyze two cases, when $\alpha = 2$ or $\alpha \in (0, 2)$.

1) If $\alpha = 2$, then the process Z is the Brownian Motion with drift μ and variance $2\sigma^2$ (Whitt (2002), p.111). It means that $\nu_Z((-\infty, \infty)) = 0$. Hence, for this case, if $\mu > 0$ (positive

drift) then with probability one the process goes to ∞, and we conclude that Y is a Lévy process.

2) If $\alpha \in (0,2)$, then $A = 0$ and

$$\nu_Z(B) = \int_{B\cap(-\infty,0)} \frac{\lambda_1}{x^{\alpha+1}} dx + \int_{B\cap(-\infty,0)} \frac{\lambda_2}{(-x)^{\alpha+1}} dx, \text{ for all measurable } B \subset \mathbb{R}$$

(Sato (1999), p.80). So, for Y to be a Lévy process we need $\lambda_2 = 0$ and $\lambda_1 > 0$. In terms of the stable parameters, we have that $\nu_Z((0,\infty)) = 0$ if and only if $\beta = -1$ (Sato (1999), Def. 14.16, p.87).

Remark. From previous Remark we conclude that it is possible for Y to be a Lévy process if the first and the second moments of τ are finite, depending on how the scaling is taken in order to obtain a positive drift. If we do not have this condition for the first and second moments, we need the following condition in terms of the "tails" of τ (Whitt (2002), Theorem 4.5.1):

$$\frac{P(\tau > t)}{P(\tau < -t) + P(\tau > t)} \to 0 \text{ when } t \to \infty.$$

For example, this is satisfied if τ is truncated from above but not from below.

Now we present the following example.

Example. Consider the following sequence of GCTRW

$$S^{(n)} = \left\{ S_t^{(n)} = \sum_{i=1}^{N_t^{(n)}} X_i^{(n)}, t \geq 0 \right\}, n = 1, 2, ... \tag{12}$$

where

$$P\left(X_i^{(n)} = \pm\frac{\Delta_0}{\sqrt{n}} + \frac{\gamma_0}{n}\right) = \frac{1}{2}, \text{ and } P\left(\tau_i^{(n)} = \pm\frac{\Delta_1}{\sqrt{n}} + \frac{\gamma_1}{n}\right) = \frac{1}{2}$$

for all i and n with $\gamma_0, \Delta_0, \Delta_1$ real numbers and $\gamma_1 > 0$.

These r.v.s belong to the domain of attraction of a Normal law, and therefore they satisfy the stable conditions.

Proposition 8 *For the sequence constructed in the example above, we have that*

$$S^{(n)} \xrightarrow{d} S \text{ as } n \to \infty \text{ in the topology } J_1,$$

where S is the Generalized Hyperbolic Lévy process, and $S_t \sim GH$ with parameters $\left(-\frac{1}{2},\ \sqrt{\left(\frac{\gamma_1}{\Delta_1\Delta_0}\right)^2+\left(\frac{\gamma_0}{\Delta_0^2}\right)^2},\ \frac{\gamma_0}{\Delta_0^2},\ \frac{\Delta_0}{\Delta_1}t,\ 0\right)$.
Because the first paramter of the GH is $\lambda=-\frac{1}{2}$, this corresponds to the NIG process.

Proof. By Proposition 7 we know that $S^{(n)} \overset{f.d.d.}{\to} S$, where S is the stable process X subordinated to Y, and Y is the first passage time of the stable process Z. We know also that $X^{(n)} \overset{d}{\to} X$ and $Z^{(n)} \overset{d}{\to} Z$ in the J_1 topology, the processes X and Z are both Brownian motions with drifts γ_0 and γ_1, and variances Δ_0^2 and Δ_1^2, respectively. By the results in Appendix-"First passage time of Lévy processes" (see Theorem (First Passage)) we know that Y is a strictly increasing Lévy process. We can also prove that $\frac{N^{*(n)}}{n} \overset{f.d.d.}{\to} Y$ as $n\to\infty$, where $N^{*(n)}$ is the monotone increasing version of $N^{(n)}$, i.e. $N^{*(n)}$ is built by joining the points of $N^{(n)}$ with lines. Then, by Theorem (Convergence in $D_{\rightrightarrows}$) (in Appendix "The Skorohod Topology J_1 and convergence"), $\frac{N^{*(n)}}{n} \overset{d}{\to} Y$ as $n\to\infty$ in the J_1 topology. Since

$$S_t^{(n)} = \sum_{i=1}^{N_t^{(n)}} X_i^{(n)} = \sum_{i=1}^{\left\lfloor n\frac{N_t^{*(n)}}{n}\right\rfloor} X_i^{(n)},$$

$S^{(n)}$ is a composition of $X^{(n)}$ and $\frac{N^{*(n)}}{n}$. Theorem (Continuous Composition-Maps) (also in the Appendix "The Skorohod Topology J_1 and convergence") implies that this composition is a continuous map, because $X\in C$ and $Y\in D_{\to}$. By the Continuous-mapping Theorem,

$$S^{(n)} \overset{d}{\to} S \text{ as } n\to\infty \text{ in the topology } J_1.$$

Let us find the parameters of S. Since $S_t = X_{Y_t}$ for each $t\geq 0$, we have that the conditional distribution of S_t given Y_t is normal, namely

$$S_t \mid Y_t \sim N(\gamma_0 Y_t, \Delta_0^2 Y_t).$$

It is known (see Appendix "First passage time of Lévy processes") that Y is driven by a GIG distribution and $Y_t \sim GIG\left(-\frac{1}{2}, \frac{t}{\Delta_1}, \frac{\gamma_1}{\Delta_1}\right)$, for $t\geq 0$. This allows us to derive (see Appendix "Some distributions") that

$$S_t \sim GH$$

with parameters as stated. Therfore, S is a Generalized Hyperbolic motion. ■

Let us mention that Eberlein and Hammerstein (2004) have presented a similar result, though they do not specify a construction for $N_t^{(n)}$. They assume that $\left\{N^{(n)}n \geq 1\right\}$ is a sequence of positive integer-valued r.v.s such that $\frac{N^{(n)}}{n} \xrightarrow{d} Y$ as $n \to \infty$, with $Y \sim GIG$.

3.5 On the problem of moments for random sums

In this section we study the distributions of random sums of r.v.s and their properties expressed in terms of the moments. This is related to the classical problem of moments, see Stoyanov (1997, 2004).

In general, we assume that a r.v. ξ has finite all moments $m_k = E(\xi^k)$, $k = 1, 2, ...$ and we ask the question: Is the distribution $\mathcal{L}(\xi)$ uniquely determined by its moment sequence $\{m_k\}$? If "yes", we say that ξ, and also $\mathcal{L}(\xi)$, is unique, or M-determinate. If "no", there are different distributions with the same moments as ξ, and in this case the distribution $\mathcal{L}(\xi)$ is non-unique, or M-indeterminate.

Several important distributions, such as Normal, Exponential, Poisson, are M-determinate. Striking examples of distributions which are M-indeterminate are the lognormal distribution and the cube of a normal distribution. Details can be seen in Stoyanov (1997).

In previous sections, we have studied random sums of random quantities and their limits. It is natural to analyze the moment determinancy of these limits. This kind of problems were considered recently, see Lin and Stoyanov (2002) and Gut (2003).

Consider the random sum $T_N = \sum_{i=1}^{N} X_i$, where X_i $i = 1, 2, ...$ is a sequence of i.i.d. r.v.s and N is a positive integer-valued r.v., which is independent of $\{X_i\}$. Each r.v. is either M-determinate or M-indeterminate. What can we say about the moment determinacy of the random sum T_N?

One of the conclusions from the above cited papers is that if X_i is M-indeterminate or N is M-indeterminate, then the random sum T_N is M-indeterminate.

However another conclusion is quite non-trivial: There exist random sums such that both X_i and N are M-determinate, while T_N is becoming M-indeterminate.

Here we want to derive conclusions about the moment determinacy of sequences of

random sums $S^{(n)} = \sum_{i=1}^{N^{(n)}} X_i^{(n)}$, where $N^{(n)}$ is a sequence of integer-valued r.v. and $\left\{X_i^{(n)}, i \geq 1\right\}$ $n = 1, 2, ...$, is a triangular array. Below are two examples, in the first one the limit is M-indeterminate, while in the second, the limit is M-determinate.

Proposition 9 *(1) There exist a triangular array of r.v.s* $\left\{X_i^{(n)},\ i = 1, 2, ...\right\},\ n = 1, 2, ...$ *and a sequence of interger-valued r.v.s* $N^{(n)}$ *such that* $S^{(n)} = \sum_{i=1}^{N^{(n)}} X_i^{(n)}$ *is M-determinate for each fixed* n*. However* $\lim_{n\to\infty} e^{S^{(n)}}$ *is M-indeterminate.*
(2) Contrary, There exist a triangular array of r.v.s $\left\{X_i^{(n)},\ i = 1, 2, ...\right\},\ n = 1, 2, ...$ *and a sequence of interger-valued r.v.s* $N^{(n)}$ *such that* $S^{(n)} = \sum_{i=1}^{N^{(n)}} X_i^{(n)}$ *is M-indeterminate for each fixed* n*. However,* $\lim_{n\to\infty} S^{(n)}$ *is M-determinate.*

Proof. (1) Consider the r.v.s $X_i^{(n)} \overset{d}{=} \frac{B}{\sqrt{n}}$, where B is a symmetric Bernoulli r.v., values ± 1 each with probability $\frac{1}{2}$, and let $N^{(n)} = n$. Clearly, for each n, these r.v.s are M-determinate. The transformation $e^{S^{(n)}}$ is M-determinate for any fixed n (Stoyanov (2002), Hausdorff moment problem). However, we know that the limit $\lim_{n\to\infty} e^{S^{(n)}}$ exits and has lognormal distribution which is M-inderterminate.
So, it is possible to have a converging sequence $\left\{S^{(n)},\ n = 1, 2, ...\right\}$ of random sums such that it is M-determinate for each n, but the limit $\lim_{n\to\infty} S^{(n)}$ is M-indeterminate.

(2) Consider the r.v.s $X_i^{(n)} \overset{d}{=} \frac{Y^3}{\sqrt{n}}$, where Y is a Normal r.v. with zero mean and, again, let $N^{(n)} = n$. It is known that Y^3 is M-indeterminate (Stoyanov (2004)). It is easy to see that the sum $S^{(n)} = \sum_{i=1}^{N^{(n)}} X_i^{(n)}$ is also M-indeterminate for each n. The r.v. $\frac{Y^3}{\sqrt{n}}$ has finite mean and variance, then by the Central limit theorem we conclude that $\lim_{n\to\infty} S^{(n)}$ is a normal r.v. which is M-determinate.
Therefore, there exits a sequence of random sums $\left\{S^{(n)},\ n = 1, 2, ...\right\}$ with $S^{(n)}$ M-indeterminate for each n and the limit $\lim_{n\to\infty} S^{(n)}$ being M-determinate.

■

3.6 Appendix

Here we give a compact review on different aspects of stochastic processes and distributions that we use throughout this chapter. We recall concepts and results of models such as *Lévy processes* and random sums, which are studied in the chapter.

3.6.1 Lévy, Stable and Self-similar processes, and Subordinations

There is a large class of stochastic process whose paths belong to the space D, this is the class of *Lévy processes*. They have been widely studied and applied in different areas. Basic references are Sato (1999) and Bertoin (1996).

To define a Lévy process we need the notion of stochastic continuity.

Definition A stochastic process $S = (S_t, t \geq 0)$ is said to be stochastically continuous or continuous in probability if, for every $t \geq 0$ and $\varepsilon > 0$,

$$\lim_{s \to t} P(|S_s - S_t| > \varepsilon) = 0.$$

Definition A stochastic process $S = (S_t, t \geq 0)$ on $\mathbb{R}$ is a Lévy process if the following conditions are satisfied.

(1) $S_0 = 0$ a.s.

(2) For any times $t_0, t_1, ..., t_n$, $n \geq 2$ such that $0 \leq t_0 \leq t_1 \leq ... \leq t_n$, the r.v.s $S_{t_0},\ S_{t_1} - S_{t_0},\ S_{t_2} - S_{t_1}, ...,\ S_{t_n} - S_{t_{n-1}}$ are independent (i.e. the process has independent increments).

(3) The distribution of $S_{t+s} - S_s$ does not depend on s (temporally homogeneous).

(4) The process is stochastically continuous.

Notice that a Lévy process is described in terms of its finite dimensional distributions. The Kolmogorov Extension of Measures Theorem ensures the existence of these mathematical objects. We consider the Lévy processes as elements of the space D. It is possible to prove that every Lévy process has a unique cádlág modification (see Protter (2004), p.4). If the process satisfies conditions (1), (2) and (4) above then it is called an *Additive process*.

An important concept related to the Lévy processes is the so-called infinite divisibility.

Definition A probability measure μ on $\mathbb{R}$ is infinitely divisible if, for any positive integer m, there is a probability measure μ_m on $\mathbb{R}$ such that

$$\mu = \mu_m * \mu_m * ... * \mu_m = (\mu_m)^{*m}$$

(Equivalently: μ is the m-fold convolution of μ_m).

A r.v. X in $\mathbb{R}$ with probability measure μ is a infinitely divisible if for any integer n we have the equality $X \stackrel{d}{=} X_1 + ... + X_n$, where $X_1, ..., X_n$ are i.i.d. r.v.s.

Any infinitely divisible r.v. has a special form of its ch.f., it is called the Lévy-Khintchine formula (see Sato (1999), p. 37).

Theorem If X is an infinitely divisible r.v. with measure μ, then its ch.f. $\Psi(\theta) = E(e^{i\theta X})$ is

$$\psi(\theta) = E(e^{i\theta S_1}) = \exp\left[-\frac{1}{2}A\theta^2 + i\gamma\theta + \int_{\mathbb{R}} \left(e^{i\theta x} - 1 - i\theta x 1_{\{|x|\leq 1\}}(x)\right) \nu(dx)\right]$$

where A is nonnegative number, ν is a measure on $\mathbb{R}$ satisfying $\nu(\{0\}) = 0$ and $\int_{\mathbb{R}} \min\left(|x|^2, 1\right) \nu(dx) < \infty$, and $\gamma \in \mathbb{R}$. This representation in terms of (A, ν, γ) is unique. Conversely, for any choice (A, ν, γ) satisfying the conditions above, there exits an infinitely divisible distribution μ having $E(e^{i\theta X})$ as its ch.f.

Let $S = (S_t, t \geq 0)$ be a Lévy process. It is easy to see that for every t, S_t is an infinitely divisible r.v. Moreover, if $\psi(\theta)$ is the ch.f. of X_1 then $(\psi(\theta))^t$ is the ch.f. of S_t. In fact, if X is an infinitely divisible r.v. then the construction of a process via $\left(E(e^{i\theta X})\right)^t$ is a Lévy process. There is a one-to-one correspondence between the Lévy processes and the infinitely divisible distributions (see Sato (1999), Sato (2001), Bertoin (1996)). The parameters (A, ν, γ) are called the *generating triplet* of the Lévy process.

The following is a more general result related to additive processes. It is called the Lévy-Khintchine representation.

Theorem If $S = (S_t), t \geq 0$ is an additive process, then, for any $t \geq 0$, the law of S_t is infinitely divisible.

Let us mention only that there is a useful characterization of Lévy processes telling us that a stochastically continuous temporally homogeneous Markov process starting from zero

at time zero is a Lévy process. Details can be seen in (Sato (1999), p. 57).

An important class of Lévy processes are the so-called self-similar processes.

Definition A stochastic process S is said to be self-similar with index $H > 0$ (Hurst parameter) if, for all $a > 0$,

$$\{S_{at} : t \geq 0\} \stackrel{d}{=} \left\{a^H S_t : t \geq 0\right\}.$$

The Brownian Motion is an example of a self-similar process with index $H = \frac{1}{2}$. Also we know that if $Z = \{Z_t,\ t \geq 0\}$ is a stationary process, then $S_t = t^H Z_{\ln t}$, $t > 0$ defines a self-similar process with parameter H (see Whitt (2002), p.97).

Another important class of Lévy processes, related to the self-similar processes, is the one generated by stable laws.

Definition A r.v. X is said to have a stable law if for any positive numbers a_1 and a_2, there is a real number b, and positive number c such that

$$a_1 X_1 + a_2 X_2 \stackrel{d}{=} b + cX,$$

where X_1 and X_2 are independent copies of X.

It turns out that the following relation holds for a stable r.v.:

$$a_1^\alpha + a_2^\alpha = c^\alpha \text{ with } \alpha \in (0, 2].$$

The number α is called the index of the stable law. It can be proved that a stable r.v. is infinitely divisible, and its law depends on four parameters, and its ch.f. is given by

$$\ln E(e^{i\theta X}) = \begin{cases} -\sigma^\alpha |\theta|^\alpha (1 - i\beta(sign\theta)\tan(\frac{\pi\alpha}{2})) + i\mu\theta \text{ if } \alpha \neq 1 \\ -\sigma |\theta| (1 - i\beta\frac{\pi}{2}(sign\theta)\ln(|\theta|)) + i\mu\theta \text{ if } \alpha = 1. \end{cases}$$

A standard notation is $X \sim S_\alpha(\sigma, \beta, \mu)$ (see Whitt (2002), p.111). The r.v. $\xi \sim N(\mu, \sigma^2)$ is a stable r.v. $S_2(\sigma, 0, \mu)$. The Cauchy r.v. is stable $S_1(\sigma, 0, \mu)$. The Lévy distribution is stable $S_{1/2}(\sigma, 1, \mu)$.

We say that a stochastic process $S = (S_t, t \geq 0)$ is a stable motion if its increments are stable r.v.s.

Except the cases when $\alpha = 1$ and $\beta \neq 1$, the stable process S is a self-similar process, and it satisfies the following property:

$$S_{at} \stackrel{d}{=} a^H S_t \text{ for any } t \geq 0 \text{ and } a > 0, \text{ where } H = \frac{1}{\alpha}.$$

Of importance is the concept called subordination. Roughly speaking, a subordination is a change of the time parameter for a stochastic process.

Definition Let $X = (X_t, t \geq 0)$ be a stochastic process, and consider a striclty increasing stochastic process $Y = (Y_t, t \geq 0)$. The process S defined as

$$S = \{X_{Y_t},\ t \geq 0\}$$

is called a subordination of X to the subordinator Y.

If X and Y are Lévy processes, then S is also a Lévy process (see Sato (1999), p.197).

3.6.2 First passage time of Lévy processes

Let $X = (X_t, t \geq 0)$ be a Lévy process with characteristics triplet (A, ν, γ). Consider the following two properties:

1) $P(X_t \leq X_{t-} \text{ for every } t > 0) = 1$;

2) $P(\limsup_{t \to \infty} X_t = \infty) = 1$.

Condition 1) is equivalent to having $\nu((0, \infty)) = 0$. This means that the process does not have positive jumps (see Sato (1999), p.345).

With these conditions it is possible to determine a process called the *first passage time process*, denoted by $R = \{R_x, x \geq 0\}$, defined by

$$R_x(\omega) = \inf\{t > 0 : X_t(\omega) > x\},\ \omega \in \Omega,\ x \geq 0.$$

For each $x \geq 0$, R_x is a stopping time (See Bertoin (1996), p.22).

Theorem (First Passage) The process $R = (R_t, t \geq 0)$ is a Lévy process whose sample paths are a.s. strictly increasing. We also have the following formula:

$$E[e^{-uR_x}] = e^{-x\psi^{-1}(u)},$$

where $\psi(w) = \frac{1}{2}Aw^2 + \gamma w + \int_{(-\infty,0)}(e^{wx} - 1 - wx1_{[-1,0)}(x))\nu(dx)$, and ψ^{-1} is its inverse function (see Sato (1999), p.346).

If X is a Brownian motion drift γ and variance σ^2, then $\nu = 0$, therefore $\Psi(w) = \frac{1}{2}\sigma^2 w^2 + \gamma w$ and the inverse is $\Psi^{-1}(u) = \frac{1}{\sigma^2}\left(\sqrt{\gamma^2 + 2\sigma^2 u} - \gamma\right)$, see Sato (1999), Example 46.6, p.350. This corresponds to an inverse Gaussian distribution (IG), which is a particular case of the Generalized Inverse Gaussian distribution (GIG). see Shoutens (2003), p.53. With the above parameters of the Brownian motion we have that the first passage process R is a Lévy process driven by the $GIG\left(\frac{-1}{2},\ \frac{x}{\sigma},\ \frac{\gamma}{\sigma}\right)$ distribution, i.e. $R_x \sim GIG\left(\frac{-1}{2},\ \frac{x}{\sigma},\ \frac{\gamma}{\sigma}\right)$ for $x \geq 0$.

3.6.3 Sums of random variables

Recall that self-similar processes and stable distributions arise naturally as the limit of sequences of random elements. Consider a sequence of i.i.d. r.v.s $X_1, X_2, ...$ in $\mathbb{R}$. Define the following sequence of processes in the space D :

$$\left\{ Z_t^{(n)} = \frac{1}{b_n}\left(\sum_{i=1}^{\lfloor tn \rfloor} X_i - a_n\right),\ t \geq 0 \right\},\ n = 1, 2, ...,$$

for some sequences of real numbers $\{a_n,\ n \geq 0\}$ and $\{b_n,\ n \geq 0\}$. Similarly we can define the object

$$\left\{ S_t^{(s)} = \frac{1}{b(s)}\left(\sum_{i=1}^{\lfloor ts \rfloor} X_i - a(s)\right),\ t \geq 0 \right\}, s \geq 0.$$

for real functions $a(s)$ and $b(s)$, $s \geq 0$, assuming $b(s)$ positive. The following result by Lamperti shows that there is a strong implication in the convergence of $S^{(s)}$, as $s \to \infty$.

Theorem If $S^{(s)} \overset{f.d.d.}{\to} S$ as $s \to \infty$, then S is self-similar with some index H. Moreover, the normalizing function $b(s)$ is a regularly varying function with index H (see Whitt (2002), p.98)

Recall that a Borel measurable function $b(s) : \mathbb{R}^+ \to \mathbb{R}^+$ is said *regularly varying function* with index $\rho \in \mathbb{R}$, and we write $b \in RV(\rho)$, if

$$\lim_{s\to\infty} \frac{b(as)}{b(s)} = a^\rho, \ a > 0.$$

If $\rho = 0$, the function b is called *slowly varying* (see Bingham et al (1998), or Meerschaert and Scheffler (2001), p95). For example, the function $c(s) = cs^\rho$, $s \geq 0$, is in $RV(\rho)$, for any $\rho \in \mathbb{R}$.

Stable r.v.s also appear in a natural way. We say that a r.v. X *belongs to the domain of attraction of* one stable distributions $S_\alpha(\sigma, \beta, \mu)$ if for the r.v.s $X_1, X_2, ...$, independent copies of X, there exit sequences of real numbers $\{a_n, \ n \geq 0\}$ and positive numbers $\{b_n, \ n \geq 0\}$ such that

$$\frac{\sum_{i=1}^n X_i - a_n}{b_n} \overset{d}{\to} X \ \text{ as } \ n \to \infty.$$

Here $X \sim S_\alpha(\sigma, \beta, \mu)$. Under wide conditions on X we have this kind of stability. Let $F(x) = P(X \leq x)$ be the d.f. of X and $G^c(x)$ the *tails* of X, i.e.

$$G^c(x) = P(|X| > x) = 1 - F(x) + F(-x).$$

Then, the following *general central limit theorem* holds.

Definition (see Whitt (2002), p.114) The r.v. X belongs to the domain of attraction of a stable law $S_\alpha(\sigma, \beta, \mu)$, with $\alpha \in (0, 2)$ (i.e. $\alpha \neq 2$), if and only if the tail $G^c(x)$ is a regularly varying function with index equal to $-\alpha$ and

$$\frac{1 - F(x)}{G^c(x)} \to \frac{1+\beta}{2} \text{ as } x \to \infty.$$

The next result is the well-known Lindeberg´s theorem.

Theorem (Lindeberg´ Theorem) Let $X_i^{(n)}$ $i = 1, 2, ..., k(n)$ be independent r.v.s and an interger-valued function k such that $k(n) \to \infty$ as $n \to \infty$. Consider the following sum of random variables:

$$S^{(n)} = \sum_{i=1}^{k(n)} X_i^{(n)}, \ n = 1, 2, ...$$

and suppose that the following conditions are satisfied:

1) $E(X_i^{(n)}) = 0$ for all $i = 1, ..., n$ and $n = 1, 2, ...$

2) $\sum_{i=1}^{k(n)} E\left(\left(X_i^{(n)}\right)^2\right) = 1$, for each n;

3) (Lindeberg´s condition)

$$\lim_{n\to\infty} \sum_{i=1}^{k(n)} E_{n,i,\varepsilon} = 0 \text{ for all } \varepsilon > 0,$$

where $E_{n,i,\varepsilon} := \int_{|x|>\varepsilon} x^2 \mu_{n,i}(dx)$ and $\mu_{n,i} = \mathcal{L}(X_i^{(n)})$.

Then $S^{(n)} \xrightarrow{d} Z$ as $n \to \infty$, where $Z \sim N(0, 1)$.

A similar statement is the Lyapunov Central limit theorem, in which the key Lyapunov´s condition is expressed in terms of the moment of order 3 (see Pollard (1984), p.51)

3.6.4 Random sums of random variables

Theorem (Transfer Theorem) Consider the random sum

$$S^{(n)} = \sum_{i=1}^{N^{(n)}} X_i^{(n)}, \ n = 1, 2, ...$$

where $\left\{X_i^{(n)}, \ i, n \in \mathbb{N}\right\}$ is a double array of r.v.s, and $\left\{N^{(n)}, \ n = 1, 2, ...\right\}$ is a sequence of integer-valued r.v.s, all of them independent of each other. Suppose that there exits an increasing sequence $\{k_n, \ n = 1, 2, ...\}$ of natural numbers such that

$$\sum_{i=1}^{k_n} X_i^{(n)} \xrightarrow{d} X \ \text{ and } \ \frac{N^{(k_n)}}{k_n} \xrightarrow{d} Y, \text{ as } n \to \infty.$$

Then $S^{(n)} \xrightarrow{d} Z$, as $n \to \infty$, where Z is a r.v. with ch.f.

$$E(e^{i\theta Z}) = \int \left(E(e^{i\theta X})\right)^u \mu_Y(du).$$

(see Gnedenko and Korolev (1996), p. 98)

3.6.5 The Skorohod J_1 topology and convergence

For studying convergence in distribution (weak convergence) of a sequence of stochastic process in functional spaces C or D, we need to provide these spaces with a topology so to be able to talk about continuous functions.

Given two functions in C, say $f(t)$ and $g(t)$, $t \in [0, \infty)$, the following function defines a metric in C (therefore a topology)(it is called uniform, or Kolmogorov metric):

$$\rho_C(f, g) := \sup_{t \in [0,\infty)} |f(t) - g(t)| .$$

If $f(t)$ and $g(t)$, $t \in [0, \infty)$ are functions in D, the following defines a metric in D :

$$\rho_D(f, g) := \inf \left\{ \varepsilon > 0 : \exists \lambda \in \Lambda : \sup_{t \in [0,\infty)} |f(t) - g(\lambda(t))| + \sup_{t \in [0,\infty)} |t - \lambda(t)| \leq \varepsilon \right\}$$

where Λ is the set of all strictly increasing functions $\lambda(t)$, $t \in [0, \infty)$, which are continuous, onto and with $\lambda(0) = 0$.

This metric, devised by Skorohod, defines a topology, it is called Skorohod topology and usually denoted by J_1.

The topology J_1 in D coincides with the topology in C generated by the above metric.

With the topology J_1 we have the following results.

Theorem (Convergence in $D_{\rightrightarrows}$) Let $\{X_n,\ n \geq 1\}$ be a sequence of stochastic processes in the space $D_{\rightrightarrows}$ (see Symbols Index). If $X^{(n)} \overset{f.d.d.}{\to} X$, as $n \to \infty$, for a stochastically continuous process X, then $X^{(n)} \overset{d}{\to} X$ as $n \to \infty$ in the J_1 topology (Bingham (1971)).

Theorem (Continuous Composition-Maps) Consider the product space $D \times D_{\to}$, endowed with the product topology of J_1. Let $\{(x_n, y_n),\ n \geq 1\}$ be a sequence of functions in $D \times D_{\to}$ such that $(x_n, y_n) \to (x, y)$ as $n \to \infty$ in the J_1 topology. It is true that if $(x, y) \in C \times D_{\to}$ or $(x, y) \in D \times C_{\rightrightarrows}$ then, $x_n \circ y_n \to x \circ y$ as $n \to \infty$ in the J_1 topology.

This means that the composition function is a continuous mapping under these restrictions (Whitt (2002), Theorem 13.2.2, p.430).

3.6.6 Some distributions

We present some distributions of random variables that we use (see Eberlein and Hammerstein (2004) or Shoutens (2003)). These distributions are infinitely divisible and therefore they can generate Lévy processes.

Let $K_\nu(x)$ be the so called *Bessel function* of the third kind:

$$K_\nu(z) = \frac{1}{2}\int_0^\infty u^{v-1}e^{-\frac{1}{2}z(u+\frac{1}{u})}du, \quad z>0 \text{ and } \nu \in \mathbb{R}.$$

The *Generalized Inverse Gaussian (GIG)* density function $f_{GIG}(x)$ for a positive r.v. X, denoted by $GIG\left(\lambda,\chi,\psi\right)$, is defined by

$$f_{GIG}(x) = \frac{(\psi/\chi)^\lambda}{2K_\lambda(\psi\chi)}x^{\lambda-1}\exp\left[-\frac{1}{2}(\psi^2 x+\chi^2/x)\right],\ x>0.$$

And we have the following restrictions for the parameters: if $\lambda>0$, then $\chi\geq 0$ and $\psi>0$; if $\lambda=0$, then $\chi>0$ and $\psi>0$; if $\lambda<0$, then $\chi>0$ and $\psi\geq 0$. When $\lambda=-\frac{1}{2}$, it is the Inverse Gaussian distributions (IG), which corresponds to the distribution of the first time when a Brownian motion with positive drift surpasses a specific level (see Shoutens (2003)). The density of the IG r.v. (denoted $IG(\chi,\psi)$) is

$$f_{IG}(x) = \frac{\chi}{\sqrt{2\pi}}e^{\psi\chi}x^{3/2}\exp\left[-\frac{1}{2}\left(\frac{\chi^2}{x}+\psi^2 x\right)\right], x>0.$$

Notice that if $X\sim IG(\chi,\psi)$, then $aX\sim IG(\sqrt{\alpha}\chi,\frac{\psi}{\sqrt{\alpha}})$.

A real-valued r.v. X has *Generalized Hyperbolic Distribution(GH)* if its density is given by

$$\begin{aligned} &f_{GH}(x;\lambda,\alpha,\beta,\delta,\mu)\\ &= a(\lambda,\alpha,\beta,\delta,\mu)\left[\delta^2+(x-\mu)^2\right]^{(\lambda-\frac{1}{2})/2}K_{\lambda-\frac{1}{2}}\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)e^{\beta(x-\mu)},\ x\in\mathbb{R}. \end{aligned}$$

Here

$$a(\lambda,\alpha,\beta,\delta,\mu) = \frac{(\alpha^2-\beta^2)^{\frac{\lambda}{2}}}{\sqrt{2\pi}\alpha^{\lambda-\frac{1}{2}}\delta^\lambda K_v(\delta\sqrt{\alpha^2-\beta^2})},$$

with $\alpha > 0$ (the shape), $0 \leq |\beta| < \alpha$ (the skewness), $\mu \in \mathbb{R}$ (location), $\delta > 0$ (scaling) and $\lambda \in \mathbb{R}$. It is denoted by $GH(\lambda, \alpha, \beta, \delta, \mu)$.

The GH distribution can be generated as a mixture of Normal and GIG in the following way. Let X and Y be r.v.s such that $X \mid Y \sim N(\mu + \beta Y, Y)$ and $Y \sim GIG\left(\lambda, \delta, \sqrt{\alpha^2 - \beta^2}\right)$. Then $X \sim GH(\lambda, \alpha, \beta, \delta, \mu)$, (see Eberlein and Hammerstein (2004) or Bingham and Kiesel (2004), p. 69.). It is also known that the Normal distribution $N(\mu, \sigma^2)$ is the limit of the Generalized Hyperbolic distribution as $\delta \to \infty$ and $\frac{\delta}{\alpha} \to \sigma^2$ (see Cont and Tankov (2004), p.126).

The *Lévy distribution* is a stable law $S_{1/2}(\sigma, 1, \mu)$ and the form of the density is

$$f_L(x) = \left(\frac{\sigma}{2\pi}\right)^{1/2} \frac{1}{(x-\mu)^{3/2}} \exp\left[\frac{1}{2}\frac{-\sigma}{x-\mu}\right], \; x > 0.$$

This corresponds to the distribution of the first time when a Brownian motion with no drift hits a specific barrier (see Woyczynski (2001), p.243)

4 Pricing Theory

One of the fundamental problems in the financial industry is to suggest a monetary value (price) of a family of contracts called *contingent claims* written on other securities, usually risky assets. This problem is called *option pricing.* There is a general theory of stochastic financial modelling, answering questions related to pricing contingent claims, or determining the "fair" price of a contract (see Bingham and Kiesel (2004), Hull (1997) for an extended description).

A European call option is a common example of a contingent claim. The "fair" value, or fair price, of such an option is the price paid by the buyer of the contract with his/her specific rights in the future.

In a *financial market* we deal with *financial securities (shares, futures, options, etc),* which explicitly describe the rights and obligations of the traders. The whole concept of pricing relies on the idea of "avoiding arbitrage", equivalently to "replicating the contract". We precise this later in detail. If there is a portfolio of financial securities that mimics the profit or loss of the contract, then both of them should have the same value because that avoids arbitrage (making money from nothing). Hence the price of the contract is the value of the portfolio at the moment of buying the contract. We can summarize this concept in the following principle,

- Two apparently different financial instruments with the same payoff (profit or loss) at maturity for all possible market scenarios have the same value at the beginning.

A first attempt to formalize this is to use models over discrete time.

4.1 Discrete models

Suppose that, in a financial market, there are n assets (stock securities). Let $S_i(t)$ denote the stock price of security $i = 1, ..., n$ at time t, $t = 0, 1, ..., T$ (possibly infinite time). There is also a free-risk security asset (bank account, or bond) whose price at time t is $B(t)$, $t = 0, 1, ..., T$. For each i and t, $S_i(t)$ is a random variable, so the vector

$$\mathbf{S}(t) = (S_1(t), ..., S_n(t)), \ t = 0, 1, ..., T$$

defines a multidimensional discrete-time stochastic process $\mathbf{S}$ on some underlying probability space $(\Omega, \mathcal{F}, P)$.

Investors (traders) invest money in financial securities making *portfolios*, which are strategies to invest in the market. Formally a portfolio is a stochastic process θ on the probability space $(\Omega, \mathcal{F}, P)$ defined by the vector $\theta(t) = (\theta_0(t), \theta_1(t), ..., \theta_n(t))\,,\ t = 0, 1, ..., T$, where $\theta_i(t)$, for $i = 1, ..., n$, is the number of units invested in security i and $\theta_0(t)$ is the ammount of money invested in the risk-free security $B(t)$ at time t.

It is assumed that the market has a nested structure which represents the flow of information revealed to the investors, so we need to consider the so-called filtration, denoted by $\mathbb{F} = \{\mathcal{F}_0 \subset \mathcal{F}_1 \subset ... \subset \mathcal{F}_T \subset \mathcal{F}\}$. Here $\mathcal{F}_t$ represents the "information" at time t, which is available to all traders.

Definition We say that a stochastic process $X(t),\ t = 0, 1, ..., T$ is adapted to the filtration $\mathbb{F}$ if for any t, $X(t)$ is measurable with respect to $\mathcal{F}_t$. We say that X is $\mathbb{F}$-predictable if $X(t)$ is measurable with respect to $\mathcal{F}_{t-1}$ for $t = 1, ..., T$.

The market is provided with a filtration $\mathbb{F}$ and the stock price process $\mathbf{S}$ and the risk-free asset B are adapted to $\mathbb{F}$. Moreover, the portfolio θ is $\mathbb{F}$-predictable. Given a portfolio θ we define the *value process* $V(t)$ at time t, $t = 0, 1, ..., T$ as follows:

$$V(t) = \theta_0(t)B(t) + \sum_{i=1}^{n} \theta_i(t)S_i(t)\ ,\ t = 0, 1, ..., T.$$

This function $V(t)$ is the monetary value of the investment at time t when performing, or following the strategy θ.

Now we have the concept of self-financing portfolio and the important concept of no-arbitrage (see Pliska (1997) or Kijima (2003)).

Definition (Self-financing portfolio) A portfolio θ is said to be self-financing trading strategy if

$$V(t) = \theta_0(t+1)B(t) + \sum_{i=1}^{n} \theta_i(t+1)S_i(t)\ ,\ t = 0, 1, ..., T-1.$$

This means that the values of the process just before and after any transaction are equal at any time, i.e. no money is added to or withdrawn from the strategy during the time, but just reallocated.

Definition (Arbitrage Opportunity) An arbitrage opportunity in the market $(\mathbf{S}, B)$ is the existence of some self-financing strategy θ such that at time t, $V(0) = 0$,, while at maturity T, $V(T) \geq 0$ and $E(V(T)) > 0$.

This means that without risking money there is a possibility to gain money (the conditions imply that $P(V(T) > 0) > 0$). Another term is "free lunch".

Remark (Efficient market) It is well agreed that theoretically a market does not accept arbitrage opportunities. If a market accepts arbitrage, then the players and investors would move their strategies to make some profit from such opportunities. Even if this happens in a market, this does not last long, and eventually the market moves to "stable" conditions with no-arbitrage. Then we say that markets are efficient, so no free lunch.

As mentioned previously, one of our goals is to be able to price contingent claims, written on the securities ($\mathbf{S}$,B).

Definition A **European contingent claim** is a r.v. X on $(\Omega, \mathcal{F}, P)$, representing a payoff at maturity T. The **price** at time $t \leq T$ should be such that the extended market avoids arbitrage. Essentially, this means that given a price process $\{p(t), t = 0, 1, ..., T\}$ for the European contingent claim X, the payoff at maturity T is $p(T) = X$. Then the *extended market* $(\mathbf{S}, B, p)$, which is the original market with added the new security, the price process $p(t)$, does not allow arbitrage opportunities. Such a price process p is called **admissible** price (see Definition 1.21 or Definition 5.29 in Föllmer (2002)).

For convinience, we use just "contingent claim" instead of European contingent claim, with the understanding that there are other types of claims in the market.

It might be the case that there are many price processes $p(t)$, $t = 0, 1, ..., T$, satisfying the

above conditions. The important question is when does there exit a unique price process? Or, how to choose one of these prices processes, if there are many? The following concepts help to clarify such situation.

Defintion(Martingale measure) A probability measure Q is called a Martingale measure on the measurable space $(\Omega, \mathcal{F})$ if two conditions are satisfied:

i) Q is equivalent to P, the original probability measure of the market, and

ii) $E_Q\left(\frac{S_i(t)}{B(t)} \mid \mathcal{F}_s\right) = \frac{S_i(s)}{B(s)}$ for all s, t with $s \leq t$ for $i = 1, ..., n$.

This means that the processes $\frac{S_i}{B} = \left(\frac{S_i(t)}{B(t)},\ t \geq 0\right)$, $i = 1, ..., n$, called discounted processes are martingales under Q. We denote by $\mathcal{P}$ the set of all martingale measures in a particular market.

Theorem (Fundamental theorem of asset pricing) The market is free of arbitrage if and only if $\mathcal{P}$ is not empty (Föllmer (2002), Theorem 5.17, p. 217).

The next theorem shows the relation between martingale measures and the discounted stock price processes (Föllmer (2002) Theorem 5.30 jointly with Definition 5.29, p. 225).

Theorem (Risk-Neutral valuation formula) Let X be a contingent claim, and p an admissible price process for X. Then there is no arbitrage in the extended market if and only if there exits a martingale measure $Q \in \mathcal{P}$ such that

$$p(t) = E_Q\left(\frac{X}{B(t)} \mid \mathcal{F}_t\right)\ ,\ t = 0, 1, ..., T.$$

So far we have a characterization of admissible prices for contingent claims using martingale measures, and we know that if the market is free of arbitrage, then there exits at least one martingale measure. We want to know under which conditions there exists just one admissible price process.

Definition (Replicating portfolio) A contingent claim X is said to be **attainable** if there exists a self-financing trading strategy θ of the original securities, called **replicating portfolio**, such that at maturity T, $V(T) = X$. In this case the portfolio process θ is said to generate, or replicate, the contingent claim X.

It means that we get the same payoff at time T investing in X or investing in the original securities using the replicating portfolio θ. It is also known that there exits a unique replicating portfolio for a contingent claim (Proposition 3.1.1 in Musiela and Rutkowski(1997), page 73). The process V is a candidate for the prices process p because it is unique and it gives the same payoff at time T. Indeed, the process V avoids arbitrage, i.e. V is an admissible prices process. The following result clarifies the relation of replicating portfolios and martingale measures.

Theorem If the contingent claim X is attainable, then the price process

$$p(t) = E_Q\left(\frac{X}{B(T)} \mid \mathcal{F}_t\right) \ , \ t = 0, 1, ..., T$$

is the same for every martingale measure Q in $\mathcal{P}$, and $V(t) = p(t)$ for $t = 0, 1, ..., T$ (Föllmer (2002) Theorem 5.33, p. 228).

Therefore, if the contingent claim is attainable, there is no ambiguity for the price process using a martingale measure, all of them yield the same price process, which corresponds to the value of the replicating portfolio.

To conclude this section we state the important theorem of complete markets (Föllmer (2002) Theorem 5.39, p. 232). The binomial model to price options is an example of this result.

Theorem In an arbitrage-free market there exits just one martingale measure if and only if all contingent claims are attainable. In such a case the market is called complete.

4.2 Continuous time models

In the previous section we have considered a discrete-time model. A more realistic or convenient assumption is to deal with continuous-time models. The approach to price a claim in continuous time is similar that in discrete-time models, again using the concept of self-financial strategies.

Now we have the stochastic integral involved which is one of the most important tools for this frame.

An important example of a model in continuous time is a functional of the Brownian Motion. The assumption that the prices are driven by the Geometric Brownian motion led to the famous Black-Scholes model for option pricing. A more general class of models in continuous time is based on Lévy processes.

Here we summarize some notions for financial pricing using models in continuous time.

Suppose that a market consists of risky-assets $\mathbf{S}$ and one free-risk security B. We define B and $\mathbf{S}$ to be positive *semimartingales* on a standard continuous-time filtered space $(\Omega, \mathcal{F}, \mathbb{F}, P)$,

$$\mathbf{S} = \{(S_1(t), ..., S_n(t)),\ t \geq 0\} \text{ and } B = \{B(t),\ t \geq 0\}.$$

The financial market $(B, \mathbf{S})$ is also called an *economy* $\mathcal{E}$.

The filtration $\mathbb{F} = \{\mathcal{F}_t,\ t \leq T\}$ satisfies the "usual conditions": (a) $\mathbb{F}$ is right continuous (i.e. $\mathcal{F}_t = \cap_{u>t}\mathcal{F}_u$); (b) $\mathcal{F}_0$ contains all the null-sets of $\mathcal{F}$. It is convenient to assume that $\mathcal{F}_0$ is trivial (i.e. if $A \in \mathcal{F}_0$ either $P(A) = 0$ or $P(A) = 1$).

Once again, the procedure for pricing requires to define what is called a *portfolio,* which is a rule for investing in the economy $\mathcal{E}$.

A **trading strategy** is a predictable stochastic process $\theta = \{(\theta_0(t), \theta_1(t), ..., \theta_n(t)),\ t \geq 0\}$ (predictability is slightly more general than the left continuity of the sample paths of the process, see Meyer (2000), p.132 or Definition 4.23 in Hunt and Kennedy (2000), p.79 or Protter (1990)). The variable $\theta_i(t)$ is the number of units invested in security i, $i = 1, ..., n$ and $\theta_0(t)$ is the money invested in the risk-free security $B(t)$ at time t.

We consider a *self-financial portfolio* , i.e. a trading strategy without possibility of infusions of new funds or withdrawals of cash during the time of the investment.

Definition We call **Self-Financial Trading Strategy** θ for the economy $\mathcal{E}$ a stochastic process $\theta = (\theta_0(t), \theta_1(t), ..., \theta_n(t))$ satisfying the following conditions:

i) θ is $\mathcal{F}-$ predictable;

ii) θ has the self-financial property,

$$\theta(t) \cdot (B(t), \mathbf{S}(t)) = \theta(0) \cdot (B(0), \mathbf{S}(0)) + \int_0^t \theta(u) d(B(u), \mathbf{S}(u)).$$

We have the notation $V(t) = \theta(t) \cdot (B(t), \mathbf{S}(t))$.

Note that the integral $\int_0^t \theta(u) d(B(u), \mathbf{S}(u))$ is well-defined since the process $(B, \mathbf{S})$ is a semimartingale and θ is predictable (see Theorem 4.31 in Jacod and Shiryaev (1987), p.46 or Protter (1990)).

Once again, we deal with markets which are free of arbitrage (no "free-lunch").

Definition (Arbitrage Opportunity) We say that the economy $\mathcal{E}$ admits an arbitrage opportunity at time T if there exists a self-financial strategy θ such that $\theta(0) \cdot (B(0), \mathbf{S}(0)) = 0$, $\theta(T) \cdot (B(T), \mathbf{S}(T)) \geq 0$, and $P(\theta(T) \cdot (B(T), \mathbf{S}(T)) > 0) > 0$ (P-a.s.). If there are not such strategies, the economy is free of arbitrage.

Definition (Martingale Measure) A martingale measure for the economy $\mathcal{E}$ is a measure Q on Ω, equivalent to P, such that the process

$$\mathbf{S}^* := \mathbf{S}/B = \{(S_1(t)/B(t), ..., S_n(t)/B(t)),\ t \geq 0\}$$

is a martingale under Q. And the space of martingale measures is denoted by $\mathcal{P}$.

We have the following useful result (Mel´nikov et al (2002), p.32 or Bingham and Kiesel (2004) Theorem 6.1.1, p.234).

Theorem If $\mathcal{P}$ is not empty, then the economy $\mathcal{E}$ is free of arbitrage.

A **contingent claim** with maturity T is a bounded $\mathcal{F}_T-$measurable r.v. X. As in the Definition of European Contingent claim in the previous section, the price of a contingent claim should avoid arbitrage. Provided that $\mathcal{P}$ is not empty, for any $Q \in \mathcal{P}$, the process

$$p(t) = E_Q\left(\frac{X}{B(T)} \mid \mathcal{F}_t\right)$$

is a martingale, with $p(T) = X(T)$. Then, with the price process p for the contingent claim X the extended market is free of arbitrage, and therefore it is an admissible price process. So, every martingale measure defines a price process for a contingent claim, but not every admissible price process is defined with a martingale measure.

A contingent claim is *attainable* if there exists a self-financial strategy that replicates the payoff at time T (i.e. $V(T) = X$). It is known that a replicating strategy, when it exits, is unique (see Proposition 10.1.1 in Musiela and Rutkowski (1997), p.234). Then $V(t)$ defines an admissible price process and it is unique. This means that

$$V(t) = E_Q\left(\frac{X}{B(T)} \mid \mathcal{F}_t\right) \text{ for all } Q \in \mathcal{P}.$$

Now we have the following (see Theorem 3.3 in Mel´nikov et al (2002), p.33).

Theorem Assume $\mathcal{P}$ is not empty. Then there exist a unique martingale measure in the economy $\mathcal{E}$ if and only if all contingent claims are attainable.

Such a market, as already mentioned above, is called **complete**.

So, if the market has just one martingale measure there is a unique fair price for every claim. The classical example is when $\mathbf{S}$ is modelled using a diffusion driven by the Brownian Motion, for example the Black-Scholes scheme.

Remark One of the most useful results for discrete-time scheme is the Risk-Neutral valuation formula stated in previous section. It describes a characterization of the space of admissible price processes using martingale measures. Unfortunately, in the continuous setting, with the definition of arbitrage opportunity, it is not possible to have a general Risk-Neutral valuation formula for continuous-time. Nevertheless, using a stronger concept

of arbitrage opportunity it is true that the existence of martingale measures in the market is equivalent to the non-existence of arbitrage opportunities. It follows from such a result that a price process p for a contingent claim X is admissible if and only if

$$p(t) = E_Q\left(\frac{X}{B(T)} \mid \mathcal{F}_t\right), \, t \geq 0$$

for some martingale measure Q in $\mathcal{P}$ (see Delbaen and Schachermayer (1994)).

We conclude this section with the stronger concept of arbitrage opportunity that leads to say that an efficient market has a non-empty space of matingale measures (see Delbaen and Schachermayer (1994), Musiela and Rutkowski (1997) or Bingham and Kiesel (2004)).

Definition (No free luch with vanishing risk) An economy $\mathcal{E}$ is called NFLVR (no free lunch with vanishing risk) if there are no sequences of simple self-financial strategies $\{\theta^m, \; m \geq 1\}$ and $\mathcal{F}_T-$ measurable non negative r.v. X such that:

1) $V_T^m \geq \delta_m$ for a sequence of positive numbers δ_m converging to 0,

2) $\lim_{m\to\infty} V_T = X$ and

3) $P\{X > 0\} > 0$.

The meaning of this is that we can not use simple strategies to approximate a position in the market in which there is no risk to loose but there is a chance of making profit.

4.3 Applications of GCTRW in finance modelling

In finance one of the basic models used for prices of securities is the so-called Binomial Model (or Cox-Ross-Rubinstein for the whole framework). Despite its simplicity (or perhaps due to it), it is possible to price uniquely any contingent claim under the robustness of risk-neutral valuation. Equivalently, there is possibility of perfect hedging (replicating payoff). This property is called uniqueness of the Martingale Measure. Such property also holds for continuous-time models driven by the Brownian Motion. The Black-Scholes framework for pricing uses the Geometric Brownian Motion (particular diffusion model) to model stock-prices. Strikingly, we can see that "in the limit", pricing within the discrete time framework of Cox-Ross-Rubinstein is equivalent to pricing in the continuous-time Black-Scholes scheme.

When the lattice-partition of the time of the discrete model becomes finer, then the limit in probability turns out to be a model in continuous time.

The Brownian motion has been a useful model for pricing. Under non-arbitrage opportunities, diffusions models define a complete market, so there is only one martingale measure to use for pricing. Nevertheless it has been pointed out that these models do not fit perfectly the financial time series observed in real markets. Thus, other models, based on Lévy processes, have been suggested and widely analyzed. A popular model for the stock price of financial securities is a stochastic process of the form

$$\mathbf{S}_t = \mathbf{S}_0 e^{\mathbf{Y}_t},\ t \geq 0,$$

where S_0 are the initial stock price, at time $t = 0$, and $\mathbf{Y}_t$ a $\mathbb{R}^n-$valued Lévy process.

Many particular cases of Lévy processes have been studied. One family that seems rich enough to model financial series is the Generalized Hyperbolic motion. This model has been strongly suggested in finance by authors like Ole E. Barndorff-Nielsen and Ernst Eberlein (see Eberlein (2001) or Prause (1999), for instance).

We would like to propose a model that exploits the simplicity and practicality of a binomial lattice and at the same time hold a good level of aggreement with financial series.

As a matter of fact, we know that the transactions in the financial market occur in a discrete manner, so we can think of discrete times. Between two consecutive transactions there is certain amount of random time. This also means that the number of transactions varies randomly. Those number of transactions reflect how prone traders are for buying or selling assets.

We suggest that the willingness for buying or selling is driven accordingly by people´s whimsical perception of the economy. We imagine that people have some perception of the state of the economy, as soon as that perception "hits a level" that they consider critical, investors may be more or less prone to sell or buy. This describes an internal mechanism that triggers the number of transactions.

Furthermore, prices increase and decrease accordingly to specific rules and policies of financial houses. In general those rules determine how much the price can change when

dealers interact. Sellers and buyers are not allowed to propose a big changes of value from previous transaction. Instead, they have to push the prices with "small" bets.

The following model for the stock prices, based on GCTRW discussed in the previous Chapter (see Definition (GCTRW), reflects the above ideas.

Suppose that an asset/stock price is modelled by the stochastic process

$$S_t = S_0 e^{Z_t},\ t \geq 0,$$

where S_0 is the price at time 0 and S_t the price at time t. The process $Z = \{Z_t,\ t \geq 0\}$ is a GCTRW. Recall the example of the section "Particular cases for stable conditions" in the previous Chapter. The process Z is defined as

$$Z_t = \sum_{i=1}^{N_t - 1} X_i \quad \text{with} \quad N_t = \min\left\{ k : \sum_{i=1}^{k} \tau_i > t \right\}, \tag{13}$$

where

$$X_i = \begin{cases} \varepsilon_0 & \text{with probability } p_0 \\ -\varepsilon_0 & \text{with probability } 1 - p_0 \end{cases} \quad \text{and} \quad \tau_i = \begin{cases} \varepsilon_1 & \text{with probability } p_1 \\ -\varepsilon_1 & \text{with probability } 1 - p_1 \end{cases}$$

with $p_0 \in (0,1)$, $p_1 \in (\frac{1}{2}, 1)$ and $\varepsilon_0, \varepsilon_1$ positive numbers. The r.v. X_i represents the change of the price in transaction i, and the integer-valued r.v. N_t is the number of transaction up to time t. The number of transactions is determined by the random walk $\sum_{i=1}^{k} \tau_i$. When this random walk crosses a specified level, then N_t is triggered.

Remark From the previous Chapter we know about the statistical regularity of this process. With the condition that $p_1 > \frac{1}{2}$, Proposition 8 tells us that the process Z behaves like a particular case of Generalized Hyperbolic process when it is observed at high frequency. The model we obtain is a Normal Inverse Gaussian process, which corresponds to the GH distribution with parameter $\lambda = -\frac{1}{2}$. On the other hand, we know that the general form of GH distribution has been widely reported as a good model for financial series. In Prause (1999), these models are used for financial series. It is interesting to mention that, in this document, estimations of the GH distribution leads to values of λ close to $-\frac{1}{2}$ for the Deutsche Bank returns. Using the so-called *Minimal Kolmogorov Distance*, this author obtains an

estimation of -0.5002 for λ, which corresponds to the NIG distribution. Furthermore, it is known that the normal distribution comes from the Generalized Hyperbolic distribution (see Appendix "Some distributions"). This suggests that the Brownian Motion behaviour observed in financial series can be a result of extreme behavior in the Generalized Hyperbolic motion, therefore in the GCTRW. We suggest that the movement of prices in financial markets are driven by a hidden process, which associated with the state of the economy. Such a process releases as an outcome the desire of people to buy or sell assets, i.e. the number of transactions in the market is triggered by a process larking from behind. With our particular setting we actually suggest that such a hidden process is close to a Brownian motion.

This model, based on a GCTRW, deals with a continuous-time subordination of the Brownian motion. Subordinations of diffusions have been considered in the literature (see, for instance, Prigent (2003), Proposition 3.3.13, p.337).

The discrete path-structure resembles a binomial lattice. We want to exploit this facts for pricing contingent claims.

We will show how model (13) is equivalent to model (12) of Section 3.4.2. Now, using model (12), we show the simulation of 3 paths (Figure 3) with $\Delta_0 = 0.2$, $\gamma_0 = 1$, $\Delta_1 = 0.3$, $\gamma_1 = 0.1$ and $n = 10,000$ (in Matlab).

Also we show the simulation of one path with the same parameters as before but now with $n = 100,000$, Figure 4, then the increments of every 100 observations, Figure 5, and the histogram of these increments, Figure 6. We can observe the high pick in the histogram corresponging to the presence of "jumps", which can be seen in the graph of the increments.

In order to price a contingent claim we have to define the market, or the economy. In this case it consists of the risk-free security $B_t = e^{rt}$ (which is the price of a bond with interest rate r, assuming $B_0 = 1$) and the asset S_t.

Recall that a martingale measure Q is such that $\frac{S_t}{B_t}$, $t \geq 0$ is a martingale, i.e.

$$E_Q\left(\frac{S_t}{B_t} \mid \mathcal{F}_s\right) = \frac{S_s}{B_s}, \text{ a.s. for } s \leq t.$$

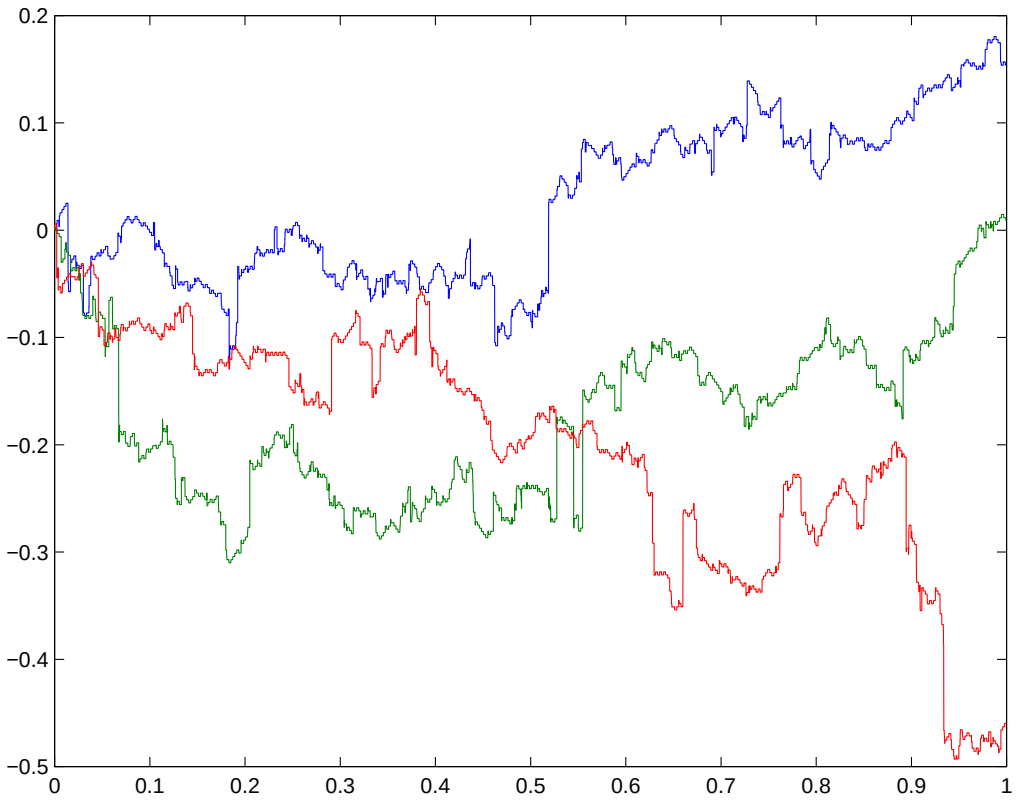


Figure 3: Simulation of 3 paths

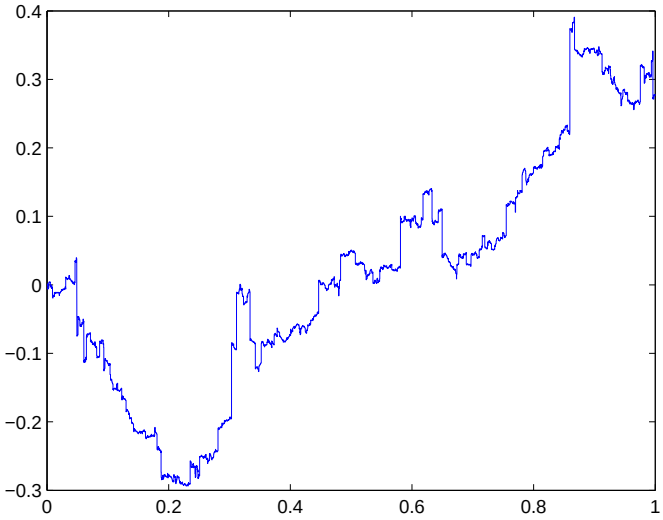


Figure 4: Simulation with $n = 100,000$

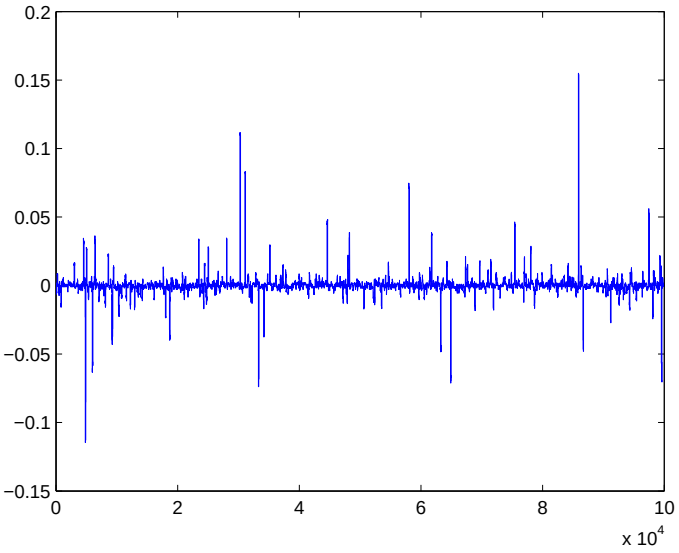


Figure 5: Increments

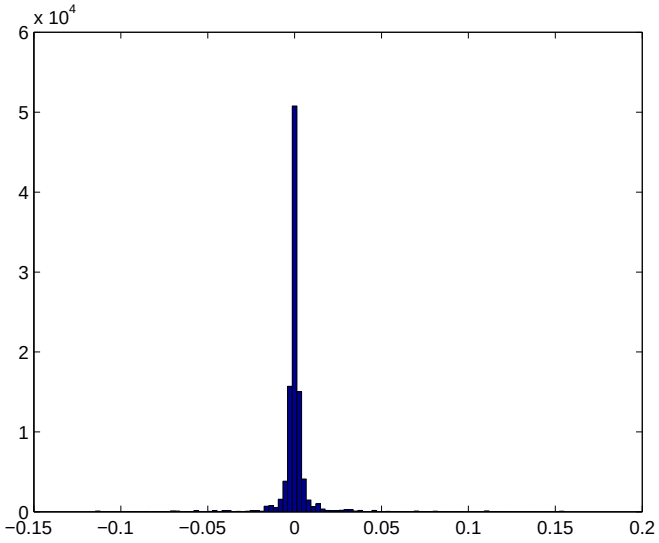


Figure 6: Histogram

With this equation and using the properties of S_t we obtain

$$E_Q\left(\frac{S_t}{B_t} \mid \mathcal{F}_s\right) = E_Q\left(\frac{S_0 e^{\sum_{i=1}^{N_s} X_i} e^{\sum_{i=1}^{N_t - N_s} X_i}}{e^{r(t-s)} e^{rs}} \mid \mathcal{F}_s\right) = \frac{S_s}{B_s} E_Q\left(\frac{e^{\sum_{i=1}^{N_t - N_s} X_i}}{e^{r(t-s)}} \mid \mathcal{F}_s\right).$$

Then, for Q to be a martingales measures, the last expression must equal $\frac{S_s}{B_s}$. So, a necessary condition for the martingale property of Q is that

$$E_Q\left(e^{\sum_{i=1}^{N_t - N_s} X_i} \mid \mathcal{F}_s\right) = e^{r(t-s)}. \tag{14}$$

The r.v.s X_i and τ_i can be expressed as symmetric Bernoulli r.v.s, i.e.

$$P(X_i = \pm 2\varepsilon_0\sqrt{p_0 q_0} + \varepsilon_0(2p_0 - 1)) = \frac{1}{2}, \text{ and}$$

$$P(\tau_i = \pm 2\varepsilon_1\sqrt{p_1 q_1} + \varepsilon_1(2p_1 - 1)) = \frac{1}{2}.$$

According to Proposition 8 in the previous chapter, under high frequency, the process $\sum_{i=1}^{N_t - N_s} X_i$ approximates a Normal Inverse Gaussian Lévy motion with specific parameters that can be derived from the two expressions. With the notation in Proposition 8 we set $\Delta_j = 2\varepsilon_j\sqrt{p_j q_j}$ and $\gamma_j = \varepsilon_j(2p_j - 1))$, $j = 0, 1$.

Using these parameters, the limit the left-hand-side of (14) turns out to be the expectation of the exponential of a Normal Inverse Gaussian-r.v., which can be obtained by means of the ch.f. We end up with the following expression:

$$e^{\left(\frac{\gamma_1}{\Delta_1}-\sqrt{\left(\frac{\gamma_1}{\Delta_1}\right)^2-2\left(\gamma_0+\frac{1}{2}\Delta_0^2\right)}\right)\frac{t-s}{\Delta_1}}=e^{r(t-s)}.$$

We plug back the original parameters to obtain the following relation in terms of the probabilities p_0 and p_1 to specify a martingale measure:

$$\frac{-1}{2\varepsilon_1\sqrt{p_1(1-p_1)}}\left(\sqrt{-2\left(\varepsilon_0(2p_0-1)+2\varepsilon_0^2p_0(1-p_0)\right)+\left(\frac{2p_1-1}{2\sqrt{p_1(1-p_1)}}\right)^2}-\frac{2p_1-1}{2\sqrt{p_1(1-p_1)}}\right)=r.$$

Hence, if we choose a pair (p_0, p_1) that satisfy the above relation, we can price a contigent claim using the GCTRW, with the understanding that these probabilities do not correspond to a martingale measure of the GCTRW, but they appear in the limit, i.e. under high frequency. However, the discrete path-structure of the GCTRW can be used to price contigent claims in a correct manner when taking limits.

If the terminal payoff of a contingent claim with maturity T is $f_T(\{S_t, 0 \le t \le T\})$ (this includes path-dependent contracts), we can calculate the price of the contigent claim by using the martingale measure (p_0, p_1) in the following way:

$$\text{Fair Price} = \lim_{n\to\infty} \frac{E_Q(f_T(\{S_t^{(n)}, 0 \le t \le T\}))}{B_T}.$$

Here $S_t^{(n)}$ is defined with the symmetric-Bernoulli representation of the r.v.s X_i and τ_i, in order to apply the results in the previous chapter.

We have to mention that the concept of embedding a random number of shocks for pricing is suggested in different works, see for example Rachev and Rüschendorf (1994), Weron (2002), Benhamou (2002), Prigent (2003). Although the way it is suggested is considerably different.

5 Statistical Inference

In this Chapter we want to deal with some aspects of statistical inference. In many cases we model phenomena with stochastic processes of a specific structure, e.g. diffusions, Lévy processes, GCTRW; depending on some unknown parameters. Using such models, and having data available, we try to identify or estimate, the unknow parameters.

We want to find the unknown parameters θ of a continuous-time stochastic process X_t, $t \in [0,T]$. The solution of the problem depends on the available data.

The data (measurements) can be taken at equally spaced times $\frac{k}{n}T,\ k = 1,2,...,n$, at arbitrary (non-random) times $0 < t_1 < t_2 < ... < t_n < T$, or, at n random times.

5.1 Random sampling data

Suppose we want to estimate the diffusion coefficient σ^2 of the Brownian motion $X_t = \sigma W_t,\ t \geq 0$, where W_t is a standard Brownian motion.

The following is well-known. Let $W = (W_t,\ t \in [0,T])$ be a standard Brownian motion. If $\left\{t^n_{(1)},...,t^n_{(n)}\right\},\ n \geq 1$ are ascendingly ordered fixed times in $[0,T]$, such that $\max\limits_{2\leq k\leq n}\left|t^n_{(k)} - t^n_{(k-1)}\right| \to 0$ as $n \to \infty$, then

$$\sum_{k=2}^{n}\left(W_{t^n_{(k)}} - W_{t^n_{(k-1)}}\right)^2 \to T \text{ a.s.}$$

The sum $\sum_{k=2}^{n}\left(W_{t^n_{(k)}} - W_{t^n_{(k-1)}}\right)^2$ is called quadratic variation of the Brownian Motion (see Oksendal (2003)).

Suppose that we have n observations of prices for the same asset. The observations come at random times $\tau_1,...,\tau_n$ in the interval $[0,T]$, according to a Uniform $(0,T)$. If $\tau_{(1)},...,\tau_{(n)}$ are the order statistics, then $\left\{S_{\tau_{(1)}},...,S_{\tau_{(n)}}\right\}$ represents the data in chronological order. We assume that the asset prices follow the following Geometric Brownian motion

$$dS_t = \mu S_t dt + \sigma S_t dW_t, t \geq 0, \text{ and } S_0 > 0,$$

where W_t is the standard Brownian motion.

We know that the solution of this stochastic differential equation is

$$S_t = S_0 \exp\left[\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma W_t\right] \tag{15}$$

(see Mao (1997, p.300)).

We construct an estimator for the volatility σ^2.

Proposition 10 *As $n \to \infty$,*

$$\widehat{\sigma}_n^2 = \frac{1}{T}\sum_{k=2}^{n}\left(\log\left(\frac{S_{\tau_{(k)}^n}}{S_{\tau_{(k-1)}^n}}\right)\right)^2 \to \sigma^2 \ a.s.$$

Proof. Without loss of generality assume that $T = 1$. By equation (15) we have

$$\widehat{\sigma}_n^2 = \sum_{k=2}^{n}\left(\left(\mu - \frac{\sigma^2}{2}\right)\left(\tau_{(k)}^n - \tau_{(k-1)}^n\right)\right)^2 + \sum_{k=2}^{n}\left(\sigma W_{\tau_{(k)}^n} - \sigma W_{\tau_{(k-1)}^n}\right)^2. \tag{16}$$

Gupta and Nadarajah (2004) have found that $\tau_{(k+1)}^n - \tau_{(k)}^n$ has distribution Beta with parameters 1 and n for all $1 \leq k \leq n$. This means that $\tau_{(k+1)}^n - \tau_{(k)}^n \xrightarrow{d} 0$ (because mean and variance tend to zero). Since it converges to a constant, the convergence is also in probability (see Grimmet and Stirzaker (2001), p.310). It implies that there exists a a.s.-convergent subsequence (see Grimmet and Stirzaker (2001), p.314). The r.v. $\tau_{(k+1)}^n - \tau_{(k)}^n$ can not increase for bigger n but decrease more, so the convergence is also a.s. for the whole sequence. This means that

$$\sum_{k=2}^{n}\left(\sigma W_{\tau_{(k)}^n} - \sigma W_{\tau_{(k-1)}^n}\right)^2 \to \sigma^2 \text{ a.s. as } n \to \infty.$$

An easy calculation shows that

$$E\left(\sum_{k=2}^{n}\left(\left(\mu - \frac{\sigma^2}{2}\right)\left(\tau_{(k)}^n - \tau_{(k-1)}^n\right)\right)^2\right) \to 0, \text{ as } n \to \infty.$$

This means that the first sum in equation (16) converges to 0 in mean, this implies convergence in probability and, again, by the nature of the order statistics, it is also convegence a.s. So, we can conclude that $\widehat{\sigma}_n^2 \to \sigma^2$, a.s. as $n \to \infty$.

■

Related to this we have the following proposition about order statistics.

Proposition 11 *For each $n = 1, 2, ...$, suppose we have random times $\{\tau_1^n, ..., \tau_n^n\}_{n\geq 1}$ which are i.i.d. r.v. with common density function $f(x)$ and distribution function $F(x)$ which is strictly monotone on $(0, \infty)$.*

Let $\left\{\tau_{(1)}^n, ..., \tau_{(n)}^n\right\}_{n\geq 1}$ be the order statistics of the sample of size n, then

$$\left|\tau_{(k+1)}^n - \tau_{(k)}^n\right| \to 0 \text{ a.s. for all } 1 \leq k \leq n, \text{ when } n \to \infty.$$

Proof. By the same arguments of the last proof it is enough to prove convergence in distribution.

Mood (1982) has the following result: For $k \leq n$ such that $\tau_{(k)}^n \leq F^{-1}(p) \leq \tau_{(k+1)}^n$, we have that the asymptotic distribution of $\tau_{(k)}^n$ and $\tau_{(k+1)}^n$ is Normal with mean $F^{-1}(p)$ and variance $\frac{p(1-p)}{f(F^{-1}(p))^2}\frac{1}{n}$. Then as $n \to \infty$ the variance goes to 0 and, $\tau_{(k)}^n$ and $\tau_{(k+1)}^n$ converge to $F^{-1}(p)$. Hence $\left|\tau_{(k)}^n - \tau_{(k+1)}^n\right| \xrightarrow{d} \delta_0$. ■

5.2 Inference for GCTRW

In Chapter 3 we have defined the Generalized Continuous Time Random Walks, and in Chapter 4 we use them to propose a financial model and its application in pricing theory. In this section we want discuss on estimation of the parameters for these models.

The following is an heuristic idea of how to estimate parameters of the GCTRW.

Suppose we have n observation $\{Z_{t_0}, Z_{t_1}, ..., Z_{t_n}\}$ at times $t_0 < t_1 < ... < t_n$ of the GCTRW-model defined in Chapter 4, in Section"Application of the GCTRW in Financial Modelling". We want to find estimators for the parameters $\varepsilon_0,\ p_0,\ \varepsilon_1$ and p_1.

We recall the definition of the model. The process Z is given by

$$Z_t = \sum_{i=1}^{N_t - 1} X_i \text{ with } N_t = \min\left\{k : \sum_{i=1}^{k} \tau_i > t\right\},\ t \geq 0,$$

where

$$X_i = \begin{cases} \varepsilon_0 & \text{with probability } p_0 \\ -\varepsilon_0 & \text{with probability } 1 - p_0 \end{cases}$$

and

$$\tau_i = \begin{cases} \varepsilon_1 & \text{with probability } p_1 \\ -\varepsilon_1 & \text{with probability } 1 - p_1. \end{cases}$$

Here $p_0 \in [0,1]$, $p_1 \in (\frac{1}{2}, 1]$ and $\varepsilon_0, \varepsilon_1$ are positive numbers.

For convenience we assume that the times $t_0, t_1, ..., t_n$ are equally spaced in time, and $t_i - t_{i-1} = \rho$ for $i = 1, ..., n$.

If we have high frequency of the data, then Proposition 8 tells us that the increments

$$I(Z_{t_i}) = Z_{t_i} - Z_{t_{i-1}},\ i = 1, 2, ..., n,$$

approximate samples from a NIG r.v., call it S_ρ.

The last part of the proof of Proposition 8 tells us that S_ρ can be seen as the mixture of a Normal r.v. X and a GIG r.v. Y in the following form:

$$S_\rho \mid Y \sim N(\gamma_0 Y, \Delta_0^2 Y),\ \text{and } Y \sim GIG\left(-\frac{1}{2}, \frac{\rho}{\Delta_1}, \frac{\gamma_1}{\Delta_1}\right),$$

where $\Delta_j = 2\varepsilon_j \sqrt{p_j q_j}$ and $\gamma_j = \varepsilon_j (2p_j - 1))$, $j = 0, 1$.

Recall that (see Appendix "Some distributions")

$$\gamma_0 Y \sim IG\left(\frac{\sqrt{\gamma_0}\rho}{\Delta_1}, \frac{\gamma_1}{\sqrt{\gamma_0}\Delta_1}\right) \text{ and } \Delta_0^2 Y \sim IG\left(\frac{\Delta_0 \rho}{\Delta_1}, \frac{\gamma_1}{\Delta_0 \Delta_1}\right).$$

We split the data into m sets, assuming for simplicity that each set has the same number of observations, say $k = \frac{n}{m}$.

Then, if k is considerable large, classical estimators $\widehat{\mu}$ and $\widehat{\sigma}^2$ for a Normal r.v. can be used to sample from the non-observable r.v.s $\gamma_0 Y$ and $\Delta_0^2 Y$. That is, given the m sets of data with k observations in each one, we build the estimators $\widehat{\mu}$ and $\widehat{\sigma}^2$ for each set. We denote them by $\widehat{\mu}_i$ and $\widehat{\sigma}_i^2$, $i = 1, ..., m$. These quantities represent the m "synthetic"-samples from $\gamma_0 Y$ and $\Delta_0^2 Y$, respectively.

Datta (2005) is one of the authors who has discussed the likelihood estimators of the parameters of a IG distribution. Given samples $\{s_i,\ i = 1, ..., n\}$ from $IG(\chi, \psi)$, the likelihood estimators for χ and ψ are

$$\widehat{\chi} = \sqrt{\sum_{i=1}^{n} \frac{1}{s_i} + \frac{n}{\sum_{i=1}^{n} s_i}} \text{ and } \widehat{\psi} = \frac{n}{\sum_{i=1}^{n} s_i} \sqrt{\sum_{i=1}^{n} \frac{1}{s_i} + \frac{n}{\sum_{i=1}^{n} s_i}},\ \text{respectively.}$$

Then, we can use these estimators and the synthetic-samples $\widehat{\mu}_i$ and $\widehat{\sigma}_i^2$, $i = 1, ..., m$, to obtain estimators of the unknown parameters. That is, we compute the estimations

$\widehat{\chi_0} := \widehat{\chi}\,(\{\widehat{\mu}_i, i = 1, ..., m\})$,
$\widehat{\psi_0} := \widehat{\psi}\,(\{\widehat{\mu}_i, i = 1, ..., m\})$,
$\widehat{\chi_1} := \widehat{\chi}\,(\{\widehat{\sigma}_i^2, i = 1, ..., m\})$ and
$\widehat{\psi_1} := \widehat{\psi}\,(\{\widehat{\sigma}_i^2, i = 1, ..., m\})$.
These are estimations of the quantities

$$\frac{\sqrt{\gamma_0}\rho}{\Delta_1},\ \frac{\gamma_1}{\sqrt{\gamma_0}\Delta_1},\ \frac{\Delta_0\rho}{\Delta_1}\ \text{ and }\ \frac{\gamma_1}{\Delta_0\Delta_1},\ \text{respectively.}$$

The value ρ is known.

We can sustitute the variables $\Delta_0,\ \gamma_0,\ \Delta_1$ and γ_1 to have expressions in term of $\varepsilon_0,\ p_0,\ \varepsilon_1$ and p_1. Thus we arrive at the following system of four equations with four unknown variables:

$$\widehat{\chi_0} = \frac{\rho\sqrt{\varepsilon_0(2p_0-1)}}{2\varepsilon_1\sqrt{p_0(1-p_0)}},\quad \widehat{\psi_0} = \frac{(2p_1-1)}{\sqrt{\varepsilon_0(2p_0-1)}\sqrt{p_1(1-p_1)}}$$
$$\widehat{\chi_1} = \frac{\rho 2\varepsilon_0\sqrt{p_0(1-p_0)}}{2\varepsilon_1\sqrt{p_0(1-p_0)}},\quad \widehat{\psi_1} = \frac{(2p_1-1)}{4\varepsilon_0\sqrt{p_0(1-p_0)}\sqrt{p_1(1-p_1)}}.$$

The solution of this system gives estimatiors for the parameters $\varepsilon_0,\ p_0,\ \varepsilon_1$ and p_1 of the GCTRW.

6 Numerical Methods for Integral Equations

Motivation. In this section we developed a numerical procedure based on probabilistic tools to solve integral equations. Again, we can see how stability (different from that in previous Chapters) arises from a random system that receives random quantities to evolve. The motivation for treating this topic came from the Chapter 2 when we were trying to determine the ergodic distribution of a Markov chain. It turned out, the density of the ergodic distribution is the solution of a specific integral equation, which is also the case of other Markov chains. Unfortunately, the methods developed in this Chapter cannot be applied to the integral equation of Chapter 2.

An integral equation is an equation that involves integrals of an unknown function. It is of interest to find the solution, i.e. the function that satisfies the integral equation. We cite examples of integral equations.

In each of the cases below, find $f : \mathbb{R} \to \mathbb{R}$ satisfying the equation:

1) $0 = -x + \int_0^x e^{x-y} f(y)dy$;

2) $f(x) = e^{-x} - \frac{1}{2} + \frac{1}{2}e^{-x-1} + \frac{1}{2}\int_0^1 (x+1)e^{-xy} f(y)dy$;

3) $f(x) = x + \int_0^x (y-x)f(y)dy$.

If the limits of the integral are fixed, as in 2), the equation is called Fredholm integral equation. If the limits are variables, as in 1) and 3), it is called Volterra integral equation. If the unknown function $f(x)$ is only under the integral sign, as in 1), the equation is said to be of the first kind, otherwise the equation is said to be of the second kind, as in 2) and 3). These are common examples of integral equations.

There is vast literature about integral equations, and several methods to find an analytic solution or numerical approximation to the solution. Here we want to propose a numerical method to find an approximation to the unknown function.

The **first restriction** we impose is the following, we will work with integral equations where the integrand is the unknown function $f(x)$ multiplying a known function $K(x, y)$, as in 1), 2) and 3). The known function $K(x, y)$ is called the kernel.

When the kernel is multiplying the unknown function it is said that the integral equation

defines a linear operator.

Like in 1), 2) and 3), we deal with equations of the form

$$0 = \varphi(x) + \int K(x,y)f(y)dy$$

or

$$f(x) = \varphi(x) + \int K(x,y)f(y)dy$$

where $K(x,y)$ is known and $f(y)$ is the unknown function. The other case is when the unknown function is inside the kernel, as in

$$f(x) = x + \int_0^x \sin(x - f(y))dy.$$

We are not working with equations of this form.

The **second restriction** is that we consider $\varphi \neq 0$. We will see why this condition is necessary. This is why we cannot use the method for the case of the Markov chain of Chapter 2.

So the integral equation we are trying to solve has following the general form

$$\alpha f(x) = \varphi(x) + \int_{a(x)}^{b(x)} K(x,y)f(y)dy \tag{17}$$

where $\alpha = 0$ or 1, and $\varphi(x) \neq 0$. The functions $a(x)$ and $b(x)$ can be constants.

The idea of the method relies on the following concept. Suppose that the solution of the integral equation $f(x)$ is constant, i.e. $f(x) = c$ for all x. Then, for x_0 in the domain of $f(x)$, we can write the equation (17) as

$$\alpha c = \varphi(x_0) + \int_{a(x_0)}^{b(x_0)} K(x_0,y)cdy = \varphi(x_0) + c\int_{a(x_0)}^{b(x_0)} K(x_0,y)dy.$$

It is just a linear equation for c, so can solve to find the value of the constant, that is

$$c = \frac{\varphi(x_0)}{\alpha - \int_{a(x_0)}^{b(x_0)} K(x_0,y)dy}.$$

The idea of the method is to give a sequence of stepwise functions that are constant within intervals, then we can turn the integral equation into a simple linear equations. At each step a new stepwise function is found and they approximate the real solution $f(x)$.

6.1 Numerical approximation

The method we propose is based on the classical Monte-Carlo method to approximate integrals when the integrand is known. The idea is to use an iteration technique, where the approximation is within a finite interval of the effective domain of the unknown function $f(x)$. At each step we have new finer partition for the interval, such partition comes from the generation of random numbers in the interval.

We can summarize the method as follows:

Consider an interval I contained in the domain of $f(x)$. We want to approximate $f(x)$ in the interval I.

Step 1: We start with a point randomly chosen in I. We assume that the function is a constant, then the integral equation is turned into a simple linear equation and we can solve it to find the unknown constant.

Step 2: We generate a new random point in I. We assume that the function is a stepwise function in the following way. With the solution of the linear equation of the previous step, we assume that $f(x)$ takes that value in a subinterval of I. And for the other part of I the function takes a new unknown value. Again, this turns the integral equation into a linear equation, and we can find the new unknown value. Then we repeat this step.

The iteration is based on the generated random numbers in I. We can generate numbers from the uniform distribution on the interval to cover homogeneously.

6.1.1 Approximation for Fredholm equations

Suppose that we want to find an approximation on the interval $[r, l]$ for the solution $f(x)$ of a Fredholm integral equation of the second kind:

$$f(x) = \varphi(x) + \int_a^b K(x, y) f(y) dy, \tag{18}$$

where $\varphi(x)$ and $K(x, y)$ are known functions. We take $[r, l]$ such that $[a, b] \subset [r, l]$.

Step 1: We start with a point $x_0 \in [r, l]$. From (18) we know that $f(x)$ satisfies

$$f(x_0) = \varphi(x_0) + \int_a^b K(x_0, y) f(y) dy.$$

Now we assume that $f(x) = y_0$ for $x \in [a, b]$, which implies that

$$y_0 = \varphi(x_0) + y_0 \int_a^b K(x_0, y)dy.$$

Therefore

$$y_0 = \frac{\varphi(x_0)}{1 - \int_a^b K(x_0, y)dy}.$$

With this we are suggesting that $f(x) \approx y_0$ for $x \in [r, l]$.

Step 2: We generate a random number x_1 on $[r, l]$. From (18) we know that

$$f(x_1) = \varphi(x_1) + \int_a^b K(x_1, y)f(y)dy.$$

We assume that the solution is a stepwise function.

If $x_1 \notin [a, b]$ we assume that

$$f(x) = \begin{cases} y_0, & \text{if } x \in [a, b] \\ y_1, & \text{if } x \notin [a, b]. \end{cases}$$

Then we have

$$y_1 = \varphi(x_1) + y_0 \int_a^b K(x_1, y)dy.$$

If $x_1 \in [a, b]$ and, $x_0 < x_1$ we assume that

$$f(x) = \begin{cases} y_0, & \text{if } x \in [a, x_{0+}\frac{x_1 - x_0}{2}] \\ y_1, & \text{if } x \in [x_{0+}\frac{x_1 - x_0}{2}, b]. \end{cases}$$

This means that we split the interval $[a, b]$ into subintervals. In those subintervals, we assume that the solution takes the values that we have found and a new unknown one. If $x_1 < x_0$, we still can do the same. Then we have

$$y_1 = \varphi(x_1) + y_0 \int_a^{x_{0+}\frac{x_1 - x_0}{2}} K(x_1, y)dy + y_1 \int_{x_{0+}\frac{x_1 - x_0}{2}}^b K(x_1, y)dy,$$

and we find y_1,

$$y_1 = \frac{\varphi(x_1) + y_0 \int_a^{x_{0+}\frac{x_1 - x_0}{2}} K(x_1, y)dy}{1 - \int_{x_{0+}\frac{x_1 - x_0}{2}}^b K(x_1, y)dy}.$$

Again this suggests that $f(x_1) \approx y_1$. After finding y_1 we can generate another random number x_2 and start again this Step 2, spliting with a new stepwise function to find y_2. This gives the iteration.

The idea is that this sequence of stepwise functions will converge to the solution of the integral equation on $[r, l]$.

6.1.2 Approximation for Volterra equations

Now we want to find an approximation on the interval $[r, l]$ for the solution $f(x)$ of a Volterra integral equation of the second kind:

$$f(x) = \varphi(x) + \int_a^x K(x, y) f(y) dy. \tag{19}$$

Again, $\varphi(x)$ and $K(x, y)$ are known functions. We take $a \leq r$.

Step 1: We start with a point $x_0 \in [r, l]$. From (19) we know that $f(x)$ satisfies

$$f(x_0) = \varphi(x_0) + \int_a^{x_0} K(x_0, y) f(y) dy.$$

We assume that $f(x) = y_0$ for $x \in [a, l]$, which implies that

$$y_0 = \varphi(x_0) + y_0 \int_a^{x_0} K(x_0, y) dy.$$

Therefore $y_0 = \frac{\varphi(x_0)}{1 - \int_a^{x_0} K(x_0, y) dy}$. As for the Fredholm equation we suggest that $f(x_0) \approx y_0$.

Step 2: We generate a random number x_1 on $[r, l]$. From (19) we know that

$$f(x_1) = \varphi(x_1) + \int_a^{x_1} K(x_1, y) f(y) dy.$$

We assume that the solution $f(x)$ is a stepwise function.

If $x_1 < x_0$ we assume that

$$f(x) = \begin{cases} y_1, & \text{if } x \in [a, x_0] \\ y_0, & \text{if } x \in [x_0, l]. \end{cases}$$

Thus we have

$$y_1 = \varphi(x_1) + y_1 \int_a^{x_1} K(x_1, y) dy,$$

and solving for y_1 yields

$$y_1 = \frac{\varphi(x_1)}{1 - \int_a^{x_1} K(x_1, y)dy}.$$

If $x_0 < x_1$, we take

$$f(x) = \begin{cases} y_0, & \text{if } x \in [a, x_0] \\ y_1, & \text{if } x \in [x_0, l] \end{cases}$$

which leads to

$$y_1 = \varphi(x_1) + y_0 \int_a^{x_0} K(x_1, y)dy + y_1 \int_{x_0}^{x_1} K(x_1, y)dy.$$

Then y_1 is the solution of a linear equation and

$$y_1 = \frac{\varphi(x_1) + y_0 \int_a^{x_0} K(x_1, y)dy}{1 - y_1 \int_{x_0}^{x_1} K(x_1, y)dy}.$$

And we can do the same when generating a new random number. At each iteration we have that $f(x_n) \approx y_n$. We expect to have a better approximations after more steps.

Remark. From the last two descriptions of the method, we can see that $\varphi(x)$ must be different from 0, otherwise the solution of the linear equations would be always 0.

6.1.3 Convergence of the algorithm

We present a heurtistic argument for the convergence of the algorithms presented above.

Let us write the integral equation in the form

$$\alpha f(x) = T(f(x)), \tag{20}$$

where $f(x)$ is the unknown function. and T is an operator involving an integral in the following way

$$T(f(x)) = \varphi(x) + \int_{a(x)}^{b(x)} K(x, y)f(y)dy.$$

Assume that there exits a unique solution $f : \mathbb{R} \to \mathbb{R}$ of (20), which is a Borel-measurable function. Suppose we want to approximate the function $f(x)$ in the finite interval $[r, l]$.

After n steps of the algorithm we find a stepwise function $f_n(x)$ which satisfies equation (20) for a single point $x_n \in [r, l]$, i.e.

$$\alpha f_n(x_n) = T(f_n(x_n)).$$

Notice that at step $n+1$, we add a new random number $x_{n+1} \in [r, l]$, so we construct a new stepwise function $f_{n+1}(x)$ that satisfy equation (20) on the point x_{n+1}. The form of the function is

$$f_{n+1}(x) = y_{n+1} I_{\{V(x_{n+1})\}}(x) + f_n(x) I^c_{\{V(x_{n+1})\}}(x), \tag{21}$$

where y_{n+1} is the value of the stepwise function in the subinterval $V(x_{n+1})$, $I_{\{V(x_{n+1})\}}(x)$ is the indicator function of the neighborhood $V(x_{n+1})$, and $I^c_{\{V(x_{n+1})\}}(x)$ the indicator function on the complement of $V(x_{n+1})$, i.e. $[r.l] \setminus V(x_{n+1})$. Equation (21) comes from the fact that we built the step function $f_{n+1}(x)$ from $f_n(x)$, basically we assume that $f_{n+1}(x)$ is a unknown constant in $V(x_{n+1})$ and the known step function $f_n(x)$ is in $[r, l] \setminus V(x_{n+1})$.

By proposition (11) the lengths of the neighborhoods $V(x_{n+1})$ tend to zero. Then, we can see from equation (21) that for n "big enough"

$$\alpha f_{n+1}(x_k) \approx \alpha f_k(x_k) = T(f_k(x_k))$$

for $k < n+1$ but "k not too far from $n+1$". Hence for "big" n, equation (20) holds for point $x = x_{n+1}$ and also $f_{n+1}(x_k) \approx T(f_{n+1}(x_k))$ for a collection of points $\{x_k,\ k < n+1\}$ previously sampled from $[r, l]$.

Since the lengths of the neighborhoods $V(x_{n+1})$ go to zero when $n \to \infty$, then the sequence of sampled points will cover the interval $[r, l]$ with probability one, therefore

$$\lim_{n\to\infty} f_n(x) = f(x) \text{ for almost all } x \text{ in } [r, l].$$

6.2 Examples

We present two examples. The programs are written in Maple and can be downloaded from my website.

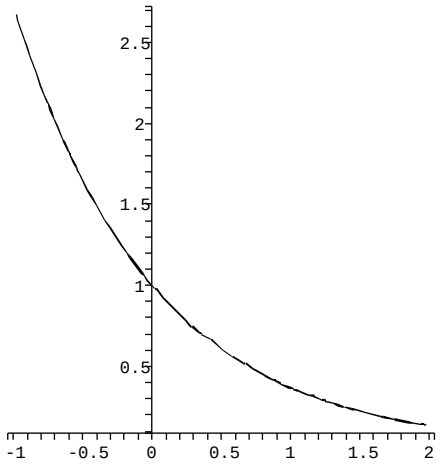


Figure 7: Approximation for the Fredholm equation (exp(-x))

Example 1 Consider the following Fredholm integral equation of the second kind

$$f(x) = e^{-x} - \frac{1}{2} + \frac{1}{2}e^{-x-1} + \frac{1}{2}\int_0^1 (x+1)e^{-xy} f(y)dy.$$

We know that the solution of this equation is $f(x) = e^{-x}$, and we want to see how well the algorithm works.

We start with $x_0 = 0$, and we generate random numbers which are uniformly distributed on the interval $[-1, 2]$. We have performed 100 iterations and plot just the last 70 iterations (Figure 7). This is because we omit the "first" iterations since they are not as accurate as the last ones. Since we know the exact solution we also plot the error in percentage of the real value, i.e. we plot $100 \times \left(\frac{f(x_n)-y_n}{f(x_n)}\right)$ for $n = 1, ..., 100$ (Figure 8). This gives an idea of the rate of convergence.

Example 2 Consider the following Volterra integral equation of the second kind

$$f(x) = x + \int_0^x (y-x)f(y)dy.$$

We know that the solution of this equation is $f(x) = \sin(x)$, as before, we want to test the algorithm.

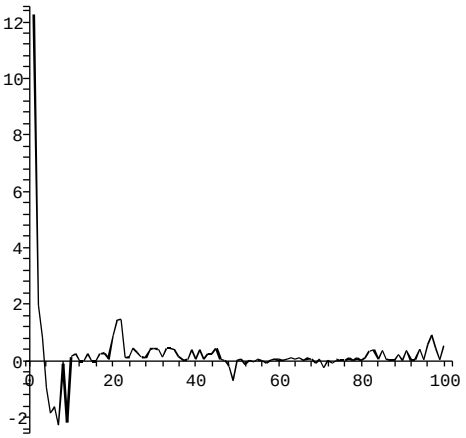


Figure 8: Percentage of the error (exp(-x))

We start with $x_0 = 9$ and generate random numbers uniformly distributed on $[0, 10]$. We generate 400 iterations and plot the last 200. As the graph (Figure 9) shows, the values close to 10 are not very good approximations, this is because in the algorithm the values closer to 0 collect more information from previous iterations than those values closer to 10. Additionally we plot the percentage of the error, i.e. $100 \times \left(\frac{f(x_n)-y_n}{f(x_n)}\right)$ for $n = 51, ..., 400$ (Figure 10).

Remark. The first problem we can highlight is the following. As the iteration runs we have to calculate integrals on smaller intervals each time, this represents a limitation computationally.

Remark. It would be interesting to extend our procedure for solving integral equations to a more general class where the unknown function is inside the kernel, i.e.

$$\int K(x, y, f(y))dy.$$

Such an example is the equation $f(x) = x + \int_0^x \sin(x - f(y))dy$.

Remark. Also notice that we use a specific construction for the stepwise functions. We can modify the way the partition is made, as well as the random number generation.

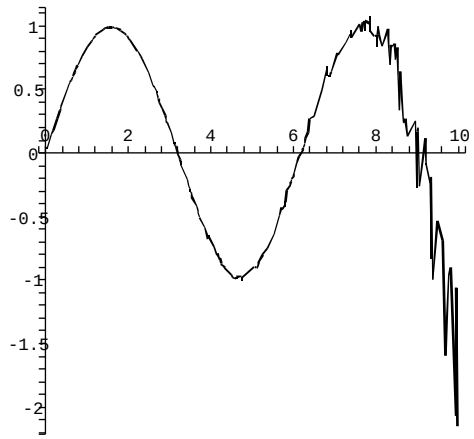


Figure 9: Approximation for the Volterra equation (sin(x))

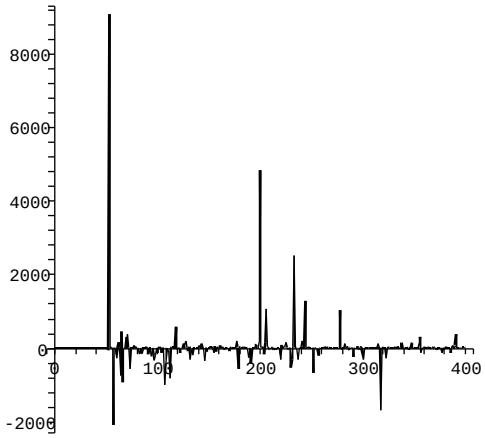


Figure 10: Percentage of the error (sin(x))

References

[1] Amussen, S. (2002). *Ruin Probability*. World Scientific, Singapore.

[2] Applebaum, D. (2004). *Lévy Processes and Stochastic Calculus*. Cambridge University Press, Cambridge.

[3] Atkinson, K. E. (1997). *The Numerical Solution of Integral Equations of the Second Kind*. Cambridge University Press, Cambridge.

[4] Becker-Kern, P., Meerschaert, M. M. and Scheffler, H.P. (2004). Limit theorems for continuous-time random walks with two times scales. *J. Appl. Probab.* **41**, 455–466.

[5] Benhamou, E. (2002). *Option pricing with Lévy Processes using Lévy exponent*, German Finance Association Annual Meeting, the former EconWPA (downloalable).

[6] Bertoin, J. (1996). *Lévy Processes*. Cambridge University Press, Cambridge.

[7] Billingsley, P (1999). *Convergence of Probability Measures*, 2nd ed. John Wiley & Sons, New York.

[8] Bingham, N.H. (1971). Limit Theorems for Occupation Times of Markov Processes. Z. Wahrscheinlichkeitstheorie verw. Geb. **17**, 17–22.

[9] Bingham, N.H., Goldie, C. and Tuegels, J. (1989). *Regular Variation* (Encyclopedia Math. Its Appl. **27**). Cambridge University Press, Cambridge.

[10] Bingham, N.H. and Kiesel, R. (2004). *Risk-Neutral Valuation,* 2nd ed. Springer, London.

[11] Borovkov, K.A. (2003). *Elements of Stochastic Modelling.* World Scientific, Singapore.

[12] Casella, G. and Berger, R.L. (1990). *Statistical Inference*. Duxbury Press, Belmont.

[13] Cont, R. and Tankov, P. (2004). *Financial Modelling with Jump Processes*. Chapman & Hall, Boca Raton (FL).

[14] Dana, R. and Jeanblanc, M. (2002). *Financial Markets in Continuous time*. Spinger, Berlin.

[15] Datta, G.S. (2005). An alternative derivation of the distributions of the maximum likelihood estimators of the parameters in an inverse Gaussian distribution. *Biometrika* **92**, 975–977.

[16] Delbaen F. and Schachermayer, W. (1994). Arbitrage and free lunch with bounded risk for unbounded continuous processes. *Mathematical Finance* **4**, 343–348.

[17] Diaconis, P. and Freedman, D. (1999). Iterated random functions. *SIAM Rev.* **41**, 45–76.

[18] Dudley, R. M. (1989). *Real Analysis and Probability*. Wadsworth & Brooks, Belmont (CA).

[19] Duffie, D. (2001). *Dynamic Asset Pricing Theory,* 3rd ed. Princeton University Press, Princeton (NJ).

[20] Eberlein, E. (2001). Application of Generalized Hyperbolic Lévy Motions to Finance. In: *Lévy Processes: Theory and Applications*, O. E. Barndorff-Nielsen, T. Mikosch and S. I. Resnick (Eds.), Birkhäser, Boston, 319–336.

[21] Eberlein, E. and Hammerstein, E. A. v. (2004). Generalized Hyperbolic and Inverse Gaussian distributions: Limiting cases and Approximation of Processes. In: *Seminar on Stochastic Analysis, Random Fields and Applications IV*, Progress in Probability 58, R.C.Dalang, M. Dozzi, F. Russo (Eds.), Birkhäser Verlag , 221–264.

[22] Föllmer, H. and Leukert, P. (1999). Efficient Hedging: Cost versus Short Risk, *Finance and Stochastics* **4**, 117–146.

[23] Föllmer, H. and Leukert, P. (1999), Quantile hedging, *Finance and Stochastics* **3**, 251–273.

[24] Föllmer, H. and Schied, A. (2002). *Stochastic Finance: An introduction in Discrete Time*. Walter de Gruyter, Berlin.

[25] Gnedenko, B. and Kolmogorov, A. N. (1954). *Limit Distributions for Sums of Independent Random Variables*. Addison-Wesley Publishing Company (Russian ed. 1949).

[26] Gnedenko, B. and Korolev, V. (1996). *Random Summation: Limit Theorems and Applications*. CRC Press, Boca Raton.

[27] Gut, A. (2003). On the moment problem for random sums, *J. Appl. Probab.* **40**, 797–802.

[28] Grimmet, G. and D. Stirzaker (2001). *Probability and Random Processes*. Oxford University Press, NewYork.

[29] Gupta, A. and Nadarajah, S. (2004). Order Statistics and Records. In: *Handbook of Beta Distribution and its Applications*, Gupta, A. and Nadarajah, S. (Eds), Macel Deckker, New York, 89–96.

[30] Hernandéz-Lerma, O. and Lasserre, J.B. (2003). *Markov Chains and Invariant Probabilities*. Birkhäuser, Basel.

[31] Hunt, P. and Kennedy, J. (2001). *Financial Derivatives in Theory and Practice*. John Wiley, NewYork.

[32] Iacus, S. M. and Negri, I. (2003). Estimating unobservable signal by Markovian noise induction. *Statistical Methods and Applications* **12**, 153–167.

[33] Ito, K. (2004). *Stochastic Processes*. Springer, Berlin.

[34] Jacod, J. and Shiryaev, A.N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. Springer, Berlin.

[35] Jerri, A. J. (1999). *Introduction to Integral Equations with Applications*. John Wiley & Sons, New York.

[36] Kallenberg, O. (2002). *Foundations of Modern Probability*. Springer, New York.

[37] Kallsen, J. (2002). Utility-based derivative pricing in incomplete markets. In: *Mathematical finance, Bachelier Congress (Paris)*, Springer, Berlin, 313–338.

[38] Kijima, M. (2003). *Stochastic Processes with Application to Finance*. Chapman & Hall, London and CRC, Boca Raton.

[39] Korolyuk, V. S. and Limnios, N. (2004). Lévy approximation of increment processes with Markov switching, *Stochastics and Stochastics Reports* **76**, 383–394.

[40] Kruglov, V. M. (1976). Convergence of distributions of sums of a random number of independent random variables to the normal distribution, *Bull. Moscow Univ. Ser. 1,* Math. and Mech. **31**, 5–12.

[41] Lamperti, J. (1977). *Stochastic Processes*. Springer, Berlin

[42] Letac, G. (2002). Donkey walk and Dirichlet distribution, *Statistics & Probability Letters* **57**, 17–22.

[43] Lin, G. D., and Stoyanov, J. (2002). On the moment determinacy of the distribution of compound geometric sums, *J. Appl. Probab.* **39**, 545–554.

[44] Mao, X. (1997). *Stochastic Differential Equations & Applications*. Horwood Publishing, Chichester.

[45] Meerschaert, M. M. and Scheffler, H.P. (2001). *Limit Distributions for Sums of Independent Random Vectors: Heavy Tails in Theory and Practice*. John Wiley & Sons, New York.

[46] Meerschaert, M. M. and Scheffler, H.P. (2004). Limit theorems for continuous-time random walks with infinite mean waiting times, *J. Appl. Probab.* **41**, 623–638.

[47] Mel´nikov, A., Volkov, S. and Nechaev, M. (2001). *Mathematics of Financial Obligation.* AMS ser. vol. **212**, Providence, (RI).

[48] Meyn, S.P. and Tweedie, R.L. (1996). *Markov Chains and Stochastic Stability*, 3rd Edition. Springer, London.

[49] Mitov, K., and Nadarajah, S. (2004). Beta Distributions in Stochastic Processes. In: *Handbook of Beta Distribution and its Applications*, Gupta, A. and Nadarajah, S. (Eds.), Macel Deckker, New York, 165–202.

[50] Mood, A., Graybill, F.A., and Boes, D.C. (1982). *Introduction to the Theory of Statistics*. McGraw- Hill, New York.

[51] Musiela, M. and Rutkowski, M. (1997). *Martingale Methods in Financial Modeling*. 2nd ed. Springer-Verlag. Heidelberg.

[52] Oksendal, B. (2003). *Stochastic Differential Equations,* 6th ed. Springer, Berlin.

[53] Pacheco C., and Stoyanov, J. (2006). More random walks in random environment with beta-ergodic distributions. In preparetion.

[54] Pliska, S. (1997). *Introduction to Mathematical Finance*. Blackwell, Oxford.

[55] Pollard, D. (1984). *Convergence of Stochastic Processes*. Springer, Berlin.

[56] Prause, K. (1999). *The Generalized Hyperbolic Model: Estimation, Financial Derivatives, and Risk Measure*. PhD Thesis, Universität Freiburg.

[57] Prigent, J.L. (2003). *Weak Convergence of Financial Markets*. Springer, Berlin.

[58] Protter, P.E. (2004). *Stochastic Integration and Differential Equation*. Springer, Berlin.

[59] Rachev, S. T. and Rüschendorf, L. (1994). Models for option prices, *Theory Probab. Appl.* **39** , 120–152.

[60] Rao, M.M. (1984). *Probability Theory with Application.* Academic Press, Orlando (FL).

[61] Scalas, E. (2005). Five Years of Continuous-time Random Walks in Econophysics. Economic Working Paper Archive EconWPA.

[62] Sato, K.I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Cambridge.

[63] Sato, K. (2001). Basic Results in Lévy processes. In: *Lévy Processes: Theory and Applications*, O. E. Barndorff-Nielsen, T. Mikosch and S. I. Resnick (Eds.), Birkhäser, Boston, 3–39.

[64] Shiryaev, A. N. (1995). *Probability,* 2nd ed. Springer, New York.

[65] Stenflo, O. (2001). Ergodic theorems for Markov chains represented by iterated function systems, *Bull. Polish Acad. Sci. Math.* **49**, 27–43.

[66] Stoyanov, J. (1997). *Counterexamples in Probability*, 2nd edn. Wiley, Chichester.

[67] Stoyanov, J. (2002). Moment problems related to the solution of stochastic differential equations. *Lecture Notes Control & Inform. Science (Springer)* **280**, 480–469.

[68] Stoyanov, J. (2004). Stieltjes classes for moment-indeterminate probability distributions. *J. Appl. Probab.* **41A**, 281–294.

[69] Stoyanov, J. and Pirinsky, C. (2000). Random motions, classes of ergodic Markov chains and beta distributions. *Statistics & Probability Letters* **50**, 293–304.

[70] Tong, H. (1990). *Non-linear Time Series*. Oxford University Press, New York.

[71] Weron, R. (2002). Pricing European Options on instruments with a constant dividend yield: the randomized discrete-time approach. *Probability and Mathematical Statistics* **22**, 407–418.

[72] Willmot, G. E. and Sheldon Lin, X. (2001). *Lundberg Approximations for Compound Distributions with Insurance Applications*, Springer, New York.

[73] Whitt, W. (2002). *Stochastic-Processes Limits*. Springer, New York.

[74] Woyczynski, W. A. (2001). Lévy processes in the Physical Sciences. In: *Lévy Processes: Theory and Applications*, O. E. Barndorff-Nielsen, T. Mikosch and S. I. Resnick (Eds.), Birkhäser, Boston, 241–266.