Charles University in Prague
Faculty of Mathematics and Physics

DOCTORAL THESIS

Mgr. Lenka Slámová, M.Sc.

Generalized stable distributions and their applications

Department of probability and mathematical statistics

Supervisor: Prof. Lev Klebanov, DrSc.
Study branch: 4M5 Econometrics and operational research

Prague 2014

Univerzita Karlova v Praze
Matematicko–fyzikální fakulta

DIZERTAČNÍ PRÁCE

Mgr. Lenka Slámová, M.Sc.

Zobecněná stabilní rozdělení a jejich aplikace

Katedra pravděpodobnosti a matematické statistiky

Školitel: Prof. Lev Klebanov, DrSc.
Studijní obor: 4M5 Ekonometrie a operační výzkum

Praha 2014

Acknowledgements

Most of all, I owe my gratitude to my supervisor, prof. Lev Klebanov, DrSc., for the interesting topic of my dissertation, for his patient guidance and valuable comments and advices during my research. I am also thankful to prof. Svetlozar Rachev from Stony Brook University for the opportunity to spend a very fruitful year at the Department of applied mathematics and statistics. My thanks go also to Dr. Aaron Kim from Stony Brook University, for his time and our valuable discussions.

I would like to thank my family and friends for their moral support throughout my studies, for their encouragement and understanding.

This work was partially supported by the grants P402/12/G097 and P203/12/0665 of the Grant Agency of the Czech Republic and further by the grant SVV-2014-260105. The support by Mobility fund of Charles University in Prague and Karel Urbánek endowment fund is gratefully acknowledged.

I declare that I carried out this doctoral thesis independently, and only with the cited sources, literature and other professional sources. I understand that my work relates to the rights and obligations under the Act No. 121/2000 Coll., the Copyright Act, as amended, in particular the fact that the Charles University in Prague has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 paragraph 1 of the Copyright Act.

In Prague, 20.10.2014 Lenka Slámová

Název práce: Zobecněná stabilní rozdělení a jejich aplikace
Autor: Mgr. Lenka Slámová, MSc.
Katedra: Katedra pravděpodobnosti a matematické statistiky
Vedoucí doktorské práce: Prof. Lev Klebanov, DrSc.

Abstrakt: Tato práce se zabývá různými zobecněními vlastnosti striktní stability s důrazem na diskrétní rozdělení s určitou formou stability. Zavedeny jsou tři možné definice diskrétní stability, následovány studií vlastností několika speciálních případů diskrétních stabilních rozdělení. Náhodná normalizace, která je použita v definici diskrétní stability, funguje i v případě spojitých náhodných veličin. Záměnou klasické normalizace v definici stability za náhodnou normalizaci je zaveden nový koncept ležérní stability. Jsou prezentovány příklady jak spojitých, tak diskrétních ležérně stabilních rozdělení. Disrétní stabilní rozdělení mají využití v diskrétních modelech, které vykazují těžké chvosty. Využití těchto rozdělení je ukázáno na modelu hodnocení vědecké práce a na modelu pro finanční časové řady. Metoda odhadu parametrů diskrétních stabilních rozdělení je v práci rovněž prezentována.

Klíčová slova: diskrétní stabilní rozdělení, ležérní stabilita, diskrétní aproximace stabilních rozdělení

Title: Generalized stable distributions and their applications
Author: Mgr. Lenka Slámová, MSc.
Department: Department of probability and mathematical statistics
Supervisor: Prof. Lev Klebanov, DrSc.

Abstract: This thesis deals with different generalizations of the strict stability property with a particular focus on discrete distributions possessing some form of stability property. Three possible definitions of discrete stability are introduced, followed by a study of some particular cases of discrete stable distributions and their properties. The random normalization used in the definition of discrete stability is applicable for continuous random variables as well. A new concept of casual stability is introduced by replacing classical normalization in the definition of stability by random normalization. Examples of casual stable distributions, both discrete and continuous, are given. Discrete stable distributions can be applied in discrete models that exhibit heavy tails. Applications of discrete stable distributions on rating of scientific work and financial time series modelling are presented. A method of parameter estimation for discrete stable family is also introduced.

Keywords: discrete stable distribution, casual stability, discrete approximation of stable distribution

Contents

1. Introduction

Stability in probability theory refers to a property of probability distributions when a sum of normalized, independent and identically distributed (i.i.d.) random variables has the same distribution (up to scale and shift) no matter how many summands we consider. Random variables with this property are called stable and they form a wide class of probability distributions. Except for one particular case, the Gaussian distribution, all stable distributions are heavy tailed. The classical stability refers to stability under summation but the concept can be extended onto other systems as well. Stability under maxima (or max-stability) leads to heavy tailed distributions called generalized extreme value distributions; stability under random summation where the number of summands is a random variable leads to heavy tailed ν-stable distributions. Stability of discrete systems is a topic that has not been studied as extensively as others but here also the discrete stable distributions exhibit heavy tails. It was believed that the stability property and heavy tails are linked together. In this thesis we study different generalizations of the concept of stability with main focus on discrete distributions and we show that some generalized forms of stability lead to distributions with exponential tails.

Introduced by Paul Lévy in (Lévy, 1925), stable distributions are a generalization of Gaussian distribution in several ways. The theory of stable distributions was developed in monographs by Lévy (1937) and Khintchine (1938), and further extended in the work by Gnedenko and Kolmogorov (1949) and Feller (1970). There exist few equivalent definitions of stable distributions. Paul Lévy defined stable distributions by specifying their characteristic function. For that he used the Lévy-Khintchine representation of infinitely divisible distributions. Second definition is connected to the "stability" property – a sum of stable random variables is again a stable random variable, a well known property of Gaussian random variables. Third is the generalized central limit theorem – stable distributions appear as a limit of sums of independent and identically distributed random variables without the standard assumption of the central limit theorem about finite variance. This result generalizes the central limit theorem and is due to Gnedenko and Kolmogorov (1949). Gaussian distribution is a special (limit) case of stable distributions, the only stable law with finite variance. Recent and extensive overview of the theory of stable random variables can be found in Zolotarev (1986), Uchaikin and Zolotarev (1999) and Samorodnitsky and Taqqu (1994).

The problem of sums of random number of random variables and the asymptotic behaviour of these random sums was first considered by Robbins (1948), who showed that a random sum of independent Gaussian random variables converges to a mixture of normal laws. General asymptotic theory of random sums of random variables was introduced in Gnedenko (1983). The analogies of stable and infinitely divisible distributions under random summation were first introduced by Klebanov et al. (1985). Specifically, the number of summands was assumed to be geometrically distributed. Klebanov et al. (1987) investigated more general summation schemes, but general theory was obtained later – Klebanov and Rachev (1996) generalized the theory of random stability for arbitrary distribution of the number of summands ν by introducing ν-infinitely divisible and ν-stable distributions. They described all random variables ν admitting an analogy of the Gaussian distribution under the random summation of ν random variables. Independently, Bunge (1996) and Gnedenko and Korolev (1996) obtained almost equivalent results by introducing random stability and random summation schemes. Finally, Klebanov et al. (2012) considered ν-stable distributions generated by summations with rational generating functions and introduced a new class of distributions with generating functions connected to Chebyshev polynomials.

The special case of geometric stable distributions has received a lot of attention thanks to their applicability in the construction of financial models. A nice overview of the results connected to the geometric stable laws was given by Kozubowski and Rachev (1999). Special cases of geometric stable distributions such as Linnik distribution or Mittag-Leffler distribution have been considered separately in literature with no connection to geometric stability. Linnik distribution was introduced by Linnik (1953). Linnik used these distributions as a tool in a study of characteristic functions. Devroye (1990) showed that Linnik random variable is a mixture of stable and exponential random variables. Anderson (1992) studied multivariate case of the Linnik distribution and a method of parameters estimation; Anderson and Arnold (1993) studied discrete time stochastic processes with stationary Linnik distribution. Kotz and Ostrovskii (1996) showed a mixture representation for the densities of Linnik distribution. Erdoğan and Ostrovskii (1998) considered the asymmetric version of Linnik distribution that coincides with the geometric stable distribution of Klebanov et al. (1985). The Mittag-Leffler distribution was introduced in Pillai (1990) and studied further in Lin (1998) where the tail behaviour and moments of the distribution were investigated.

All distributions we mentioned until now are continuous distributions. In many practical applications continuous distributions are often preferred over discrete distributions because they offer more flexibility. There are however cases of practical applications where one need to describe heavy tails in discrete data. Citations of scientific papers (first observed by Price (1965)), word frequency (Zipf (1949)) and population of cities are all well known examples of discrete data with power tails. A simple discrete power law distribution was introduced by Zipf (1949) and relied on the zeta function (therefore called Zipf or zeta distribution).

Another possibility is to consider discrete variants of stable and ν-stable distributions. The notion of discrete stability for lattice random variables on non-negative integers was introduced in Steutel and van Harn (1979). They introduced so called binomial thinning operator $\odot$ for normalization of discrete random variables. That means that instead of standard normalization aX by a constant $a \in (0,1)$, they consider $a \odot X = \sum_{i=1}^{X} \epsilon_i$, where ϵ_i are i.i.d. random variables with Bernoulli distribution with parameter a. As opposed to the standard normalization, this thinning operation conserves the integral property of a discrete random variable X. Together with a study of discrete self-decomposability they obtained the form of generating function of such discrete stable distributions. By considering only non-negative discrete random variables, they obtained a discrete version of α-stable distributions that are totally skewed to the right. Moreover, the construction allows the index of stability α only smaller or equal to one. Devroye (1993) studied three classes of discrete distributions connected to stable laws, one of them being the discrete stable distribution. Devroye (1993) derived distributional identities for these distributions offering a method for generating random samples. Christoph and Schreiber (1998) studied discrete stable distributions more into details, offering formulas for the probabilities as well as their asymptotic behaviour. They showed that the discrete stable distribution belongs to the domain of normal attraction of stable distribution totally skewed to the right with index of stability smaller than one. The non-existence of a closed form formula of the probability mass function and non-existence of moments implies that the classical parameter estimation procedures such as maximum likelihood and method of moments cannot be applied. Marcheselli et al. (2008) and Doray et al. (2009) suggested some methods of parameter estimation of the discrete stable family based on the empirical characteristic function or on the empirical probability generating function.

Discrete stable distributions in limit sense on the set of all integers were introduced in Klebanov and Slámová (2013). Two new classes of discrete distributions were introduced, generalizing the definition of discrete stable distribution of Steutel and van Harn (1979) on random variables on the set of all integers. It was shown that the newly introduced symmetric discrete stable distribution can be considered a discrete analogy of symmetric α-stable distribution with index of stability $\alpha \in (0, 2]$, whereas the introduced discrete stable distribution for random variables on $\mathbb{Z}$ can be viewed as a discrete analogy of α-stable distribution with index of stability $\alpha \in (0, 1) \cup \{2\}$ and with skewness β. Slámová and Klebanov (2012) gave two distributional identities for the symmetric discrete stable and discrete stable random variables, allowing for simple random generator. Possible estimation procedures for the class of discrete stable laws were also considered.

Discrete distribution connected to geometric stable distributions were also studied. For example Pillai and Jayakumar (1995) introduced a discrete analogy of the Mittag-Leffler distribution that is also geometrically infinitely divisible. In this thesis we will show, that similarly as Mittag-Leffler distribution is a special case of geometric stable distribution, the discrete Mittag-Leffler distribution is a special case of the discrete analogy of ν-stable distribution. Discrete Linnik distribution was introduced in Devroye (1993).

The aim of this thesis is to study different generalizations of the strict stability property with a particular focus on discrete distributions with some form of stability property. The starting point of the thesis are discrete stable distributions introduced in Steutel and van Harn (1979). Their definition of discrete stability is a simple generalization of the classical stability property where they consider only one type of thinning operator. The classical stability property can be formulated in several equivalent ways and our aim is to study generalizations of these equivalent definitions for the discrete case. We propose three definitions of discrete stability for random variables on non-negative integers. The main focus is on the first definition that generalizes the definition of Steutel and van Harn (1979) by allowing the thinning operator to be an arbitrary distribution satisfying certain condition. We introduce also the symmetric and asymmetric variant of discrete stable distribution. The definition of discrete stability on all integers, similarly as in Klebanov and Slámová (2013), is possible only in the limit sense. The generalized definition of discrete stability showed applicable for continuous random variables as well. If we replace the classical normalization in the definition of stability by a random normalization, as is done in the case of discrete stability, a generalized definition of stability for continuous random variables is obtained. We call this generalized stability property “casual stability”. We study this property and we show that many probability distributions are in fact casual stable, including geometric distribution, negative binomial distribution, gamma distribution, tempered stable distribution and Laplace distribution. This is a surprising result as it shows that even distributions with exponential tails are in some sense stable.

Outline of the thesis

The thesis is divided into two parts. Part I gathers the theoretical results about different generalizations of stability property, Part II includes some application of discrete stable distributions introduced in Part I.

Chapter 2 serves as a review of definitions and basic results about stable distributions following Samorodnitsky and Taqqu (1994); discrete stable distributions following the papers by Steutel and van Harn (1979) and Klebanov and Slámová (2013) and ν-stable distributions following Klebanov and Rachev (1996). Chapter 3 follows Slámová and Klebanov (2014a) and focus on three discrete approximations of stable distributions. By approximating the characteristic function of stable distribution discrete distributions with the same tail behaviour are obtained. These distributions coincide with the discrete stable distributions introduced in Klebanov and Slámová (2013). We show that by truncating and tempering the tails we obtain yet another discrete approximation of stable distributions with Gaussian and exponential tails. Another discrete distribution is obtained by discretizing the Lévy measure of stable distribution. In Chapter 4 three possible definitions of discrete stability for non-negative integer-valued random variables are given. These definitions consider different approaches to introducing discrete stability, each of them being a discrete version of a different definition of stability in the usual sense. The first definition generalizes the approach taken by Steutel and van Harn (1979) and considers a general thinning operator to normalize the sum of discrete random variables. The second definition takes the opposite path and uses a general so called portlying operator to normalize discrete random variables. The last approach combines the two definitions and as it turns out includes the previous two definitions. Examples of the thinning and portlying operators for which a positive discrete stable random variable exists are provided. Chapter 5 is dedicated to study of analytical properties of discrete stable distributions in the first sense. The study is focused mainly on the class of distributions connected to modified geometric thinning operator and we give results on characterizations, probabilities, moments, limiting distributions and asymptotic behaviour for positive and symmetric discrete stable random variables. The last Section of this Chapter gives also some results on properties of positive discrete stable random variables with Chebyshev thinning operator. Casual stable distributions are studied in Chapter 6, following Klebanov and Slámová (2014). Formal definitions are given and it is shown that many distributions, both continuous and discrete, are casual stable. Of course positive stable and positive discrete stable random variables are casual stable, but surprisingly also their tempered variants that have exponential tails are casual stable as well. A limit theorem for convergence to casual stable random variable is also given. In Chapter 7 a discrete analogy of ν-stable distributions is introduced. The ν-discrete stable distribution is defined in a similar way ν-stable distributions were defined in Klebanov and Rachev (1996); and a special case of these distributions connected to geometric summation scheme is studied.

The second Part of this thesis is dedicated to applications of discrete stable distributions. Chapter 8 focuses on rating of scientific work and a model for number of paper citations leading to discrete stable distribution is introduced. Chapter 9 concerns a method of estimation of parameters of the discrete stable family. Similarly as stable laws, discrete stable distributions are defined through characteristic function and do not posses a probability mass function in a closed form. This inhibit the use of classical estimation methods such as maximum likelihood and other approach has to be applied. This method departs from the $\mathcal{H}$-method of maximum likelihood suggested by Kagan (1976) where the likelihood function is replaced by a function called informant (which is, essentially, a score function), an approximation of the likelihood function in some Hilbert space. For this method only some functionals of the distribution are required, such as a probability generating function or a characteristic function. This method is adopted for the case of discrete stable distributions and in a simulation study the performance of this method is shown. Chapter 10 is devoted to a financial application of discrete stable distributions. A new NGARCH model with tempered discrete stable and approximated discrete stable distributed innovations is introduced. The performance of this new model on market data of index S&P 500 and options on this index is considered. The Chapter is concluded by the study of option traders' market sentiment incorporated in the option prices. By comparing the relatively calm recent period in March 2014 with a volatile period at the beginning of the financial crisis in September 2008, and by studying the possible large losses expressed via Value at Risk, the data indicate that the market is still heavy-tailed with a possibility of big losses, however this risk decreased significantly since 2008.

Part I

Generalized definitions of stability

2. Mathematical preliminaries

This Chapter serves as an introductory text to the theory of stable and discrete stable distributions. We give the definitions and summarize known results from the literature.

2.1 Stable distributions

First introduced in 1920s and 1930s by Lévy and Khintchine, the theory of stable distributions is very wast and there exist many publications covering the topic. The classical references include Gnedenko and Kolmogorov (1954), Feller (1970) Zolotarev (1986) and Samorodnitsky and Taqqu (1994). Stable distributions can be defined in several equivalent ways. We will focus only on strictly stable distributions.

Definition 2.1. A random variable X is said to have a strictly stable distribution if for any positive numbers a, b, there exists a positive number c such that

$$aX_1 + bX_2 \overset{d}{=} cX, \tag{2.1}$$

where X_1 and X_2 are independent copies of X and where $\overset{d}{=}$ denotes equality in distribution. A random variable X is called symmetric stable if its distribution is symmetric, i.e. X and $-X$ have the same distribution.

Feller (1970) showed that for a stable random variable X there exists a number $\alpha \in (0, 2]$ such that the number c in (2.1) satisfies

$$c^\alpha = a^\alpha + b^\alpha.$$

The number α is called *index of stability.* Second equivalent definition relates to the so called *stability property* of a stable random variable, where a normalized sum of independent copies of X have then same distribution as X.

Definition 2.2. A random variable X is said to be strictly stable, if for any $n \geq 2$ there exists a constant $a_n > 0$ such that

$$X \overset{d}{=} a_n \sum_{i=1}^{n} X_i, \tag{2.2}$$

where $X_1, \ldots, X_n$ are independent copies of X.

This definition can be rewritten in a slightly modified form.

Definition 2.3. A random variable X is said to be strictly stable, if for any $n \geq 2$ there exists a constant $A_n > 0$ such that

$$A_n X \overset{d}{=} \sum_{i=1}^{n} X_i, \tag{2.3}$$

where $X_1, \ldots, X_n$ are independent copies of X.

The equivalence of Definition 2.3 and Definition 2.1 was shown, for example, in Feller (1970). As Feller (1970) also showed, the index of stability α satisfies $a_n = n^{-1/\alpha}$ and $A_n = n^{1/\alpha}$.

The proof that the following definition of stable distribution is equivalent to the previous definitions is more complicated and can be found for example in Gnedenko and Kolmogorov (1954) or Klebanov (2003). The proof is based on the infinite divisibility of stable distribution and the Lévy-Khintchine representation of infinitely divisible distribution.

Definition 2.4. A random variable X is said to have a stable distribution if there is a quadruple of parameters $(\alpha, \beta, \sigma, \mu)$ with $\alpha \in (0, 2]$, $\beta \in [-1, 1]$, $\sigma > 0$ and $\mu \in \mathbb{R}$, such that the characteristic function of X given as $f(t) = \mathrm{E}e^{\mathrm{i}tX}$ has the following form:

$$f(t) = \begin{cases} \exp\left\{-\sigma^\alpha |t|^\alpha \left(1 - \mathrm{i}\beta \mathrm{sign}(t) \tan \frac{\pi\alpha}{2}\right) + \mathrm{i}\mu t\right\}, & \alpha \neq 1, \\ \exp\left\{-\sigma^\alpha |t| \left(1 + \mathrm{i}\beta \frac{2}{\pi} \mathrm{sign}(t) \log |t|\right) + \mathrm{i}\mu t\right\}, & \alpha = 1. \end{cases} \tag{2.4}$$

X is strictly stable if (2.4) holds with $\mu = 0$.

Since every stable distribution can be characterized by the four parameters $\alpha, \beta, \sigma, \mu$, we will denote stable distribution and random variable by $\mathrm{S}(\alpha, \beta, \sigma, \mu)$. The parameter α is the index of stability, β is the skewness parameter, σ is the scaling parameter and μ is the location parameter. If $\beta > 0$ the distribution is skewed to the right, while it is skewed to the left if $\beta < 0$ and symmetric if $\beta = 0$. If $\beta = 1$ or $\beta = -1$ we say that X is totally skewed to the right or to the left respectively. The support of stable distribution is $\mathbb{R}$, unless $\beta = \pm 1$ and $\alpha < 1$. In that case the support of the distribution is a half-line.

All stable distributions are absolutely continuous and have densities, however only in some special cases the densities are known in closed form. These include the Gaussian distribution $\mathrm{S}(2, 0, \sigma, \mu)$, the Cauchy distribution $\mathrm{S}(1, 0, \sigma, \mu)$ and the Lévy distribution $\mathrm{S}(\frac{1}{2}, 1, \sigma, \mu)$.

The last definition of stability states that stable distributions appear as the only limits of normalized sums of independent and identically distributed (i.i.d.) random variables.

Definition 2.5. A random variable X is said to have a stable distribution if it has a domain of attraction, i.e. if there exists a sequence of i.i.d. random variables $Y_1, Y_2, \dots,$ and a sequence of positive numbers $\{a_n\}$ and real numbers $\{b_n\}$, such that

$$\frac{Y_1 + Y_2 + \cdots + Y_n}{a_n} + b_n \xrightarrow{d} X.$$

The notation $\xrightarrow{d}$ denotes here and further in the thesis convergence in distribution. The random variable is strictly stable if the above holds with $\{b_n\}$ identically equal to zero for all n.

The proof of the equivalence relationship between this Definition 2.5 and Definition 2.2 can be found, for example, in Gnedenko and Kolmogorov (1954). One of the implications is straighforward as it is enough to take Y_is to be independent copies of X. The definition of domain of normal attraction is of special interest.

Definition 2.6. We say the the random variables Y_is belong to the domain of normal attraction of random variable X when $a_n = n^{1/\alpha}$.

2.1.1 Properties

In this Section we provide some basic properties of stable laws. The proofs can be found in Samorodnitsky and Taqqu (1994).

Property 2.7. *Let X_1 and X_2 be independent random variables with $X_i \sim \mathrm{S}(\alpha, \beta_i, \sigma_i, \mu_i)$, $i = 1, 2$. Then $X_1 + X_2 \sim \mathrm{S}(\alpha, \beta, \sigma, \mu)$, with*

$$\sigma = (\sigma_1^\alpha + \sigma_2^\alpha)^{1/\alpha}, \quad \beta = \frac{\beta_1 \sigma_1^\alpha + \beta_2 \sigma_2^\alpha}{\sigma_1^\alpha + \sigma_2^\alpha}, \quad \mu = \mu_1 + \mu_2.$$

Property 2.8. *Let $X \sim \mathrm{S}(\alpha, \beta, \sigma, \mu)$. Let a be a real constant. Then*

$$\begin{aligned} aX &\sim \mathrm{S}(\alpha, \operatorname{sign}(a)\beta, |a|\sigma, a\mu), & \alpha \neq 1, \\ aX &\sim \mathrm{S}(1, \operatorname{sign}(a)\beta, |a|\sigma, a\mu - \tfrac{2}{\pi} a (\log |a|) \sigma \beta), & \alpha = 1. \end{aligned}$$

Property 2.9. *Let $\alpha \in (0, 2)$. Then $X \sim \mathrm{S}(\alpha, \beta, \sigma, 0)$ if and only if $-X \sim \mathrm{S}(\alpha, -\beta, \sigma, 0)$.*

Property 2.10. *Let $X \sim \mathrm{S}(\alpha, \beta, \sigma, \mu)$. Then X is symmetric if and only if $\beta = 0$ and $\mu = 0$.*

Property 2.11. *Let X be $\mathrm{S}(\alpha, \beta, \sigma, 0)$ with $\alpha \in (0, 2) \setminus \{1\}$. Then there exist two i.i.d. random variables Y_1 and Y_2 with common distribution $\mathrm{S}(\alpha, 1, \sigma, 0)$ such that*

$$X \stackrel{d}{=} \left(\frac{1 + \beta}{2} \right)^{1/\alpha} Y_1 - \left(\frac{1 - \beta}{2} \right)^{1/\alpha} Y_2.$$

Property 2.12. *Let $X \sim \mathrm{S}(\alpha, 1, \sigma, 0)$. Then the Laplace transform of X, defined as $L(u) = \mathrm{E}e^{-uX}$, takes the following form:*

$$L(u) = \begin{cases} \exp\left\{ -\frac{\sigma^\alpha}{\cos \frac{\pi\alpha}{2}} u^\alpha \right\}, & \alpha \neq 1, \\ \exp\left\{ \sigma \frac{2}{\pi} u \log u \right\}, & \alpha = 1. \end{cases}$$

The following property describes the tail behaviour of stable distribution. The tails are heavy with index α, meaning that they decay polynomially with rate α.

Property 2.13. *Let $X \sim \mathrm{S}(\alpha, \beta, \sigma, \mu)$ with $0 < \alpha < 2$. Then*

$$\begin{aligned} \lim_{\lambda \to \infty} \lambda^\alpha \mathbb{P}(X > \lambda) &= C_\alpha \frac{1 + \beta}{2} \sigma^\alpha, \\ \lim_{\lambda \to \infty} \lambda^\alpha \mathbb{P}(X < -\lambda) &= C_\alpha \frac{1 - \beta}{2} \sigma^\alpha, \end{aligned}$$

where $C_\alpha = \frac{1-\alpha}{\Gamma(2-\alpha)\cos(\pi\alpha/2)}$ if $\alpha \neq 1$, and $C_\alpha = 2/\pi$ if $\alpha = 1$.

The heaviness of tails leads to non-existence of finite moments of order $p < \alpha$.

Property 2.14. *Let $X \sim \mathrm{S}(\alpha, \beta, \sigma, 0)$ with $\alpha \in (0, 2)$. Then*

$$\begin{aligned} &\mathrm{E}|X|^p < \infty, && \textit{for any} \quad 0 < p < \alpha, \\ &\mathrm{E}|X|^p = \infty, && \textit{for any} \quad p \geq \alpha. \end{aligned}$$

2.2 Discrete stable distributions

We can rewrite the Definition 2.1 with $c = 1$ as $X \stackrel{d}{=} aX_1 + (1 - a^\alpha)^{1/\alpha} X_2$. To define discrete stability, one need a different normalization as the normalization by a (where $a < 1$) does not preserve integers. Steutel and van Harn (1979) thus introduced binomial thinning operator $\odot$ on non-negative integer-valued random variables as

$$p \odot X = \sum_{i=1}^{X} \varepsilon_i,$$

where ε_i are i.i.d. Bernoulli random variables with $\mathbb{P}(\varepsilon_i = 1) = 1 - \mathbb{P}(\varepsilon_i = 0) = p$.

Definition 2.15 (Steutel and van Harn (1979))**.** A non-negative integer-valued random variable X is said to be positive discrete stable with parameter γ if there exists a constant $a \in (0, 1)$ such that

$$a \odot X_1 + (1 - a^\gamma)^{1/\gamma} \odot X_2 \stackrel{d}{=} X, \tag{2.5}$$

where X_1, X_2 are independent copies of X.

This definition can be equivalently reformulated to obtain an analogy of the Definition 2.2, stating that X is positive discrete stable random variable, if for all $n \in \mathbb{N}$ there is a constant $p_n \in (0, 1)$ such that

$$X \stackrel{d}{=} \sum_{i=1}^{n} p_n \odot X_i, \tag{2.6}$$

where $X_1, X_2, \dots$ are independent copies of X.

Remark 2.16. Steutel and van Harn (1979) used term *discrete stable random variable* instead of positive discrete stable but the former term will be reserved for the general case of random variables on $\mathbb{Z}$. This terminology was used in Slámová and Klebanov (2012) and already adopted by Barabesi and Pratelli (2014).

Steutel and van Harn (1979) showed that the probability generating function of a positive discrete stable random variable X, defined as $\mathcal{P}(z) = \mathrm{E} z^X$, takes the following form:

$$\mathcal{P}(z) = \exp\left\{-\lambda(1 - z)^\gamma\right\}, \quad \gamma \in (0, 1],\ \lambda > 0.$$

Poisson random variable is a special case of positive discrete stable random variable with $\gamma = 1$. Discrete stable distribution is infinitely divisible and it is in fact compound Poisson distribution. We can write a positive discrete stable random variable X with parameters γ and λ as a random sum

$$X \stackrel{d}{=} \sum_{i=1}^{N} Y_i,$$

where N is a Poisson random variable with parameter λ and $Y_1, Y_2, \dots$ are i.i.d. random variables with probability generating function

$$\mathcal{P}_Y(z) = 1 - (1 - z)^\gamma.$$

Discrete distribution with this probability generating function is called Sibuya distribution.

Devroye (1993) showed a useful stochastic representation of a positive discrete stable random variable.

Theorem 2.17. *A positive discrete stable random variable with parameters γ, λ is distributed as a Poisson random variable with parameter $\lambda^{1/\gamma} S_\gamma$, where S_γ is a positive stable random variable with Laplace transform* $\exp\{-u^\gamma\}$.

Klebanov and Slámová (2013) introduced a generalization of the definition (2.6) by considering integer-valued random variables – both symmetric and asymmetric. The symmetric discrete stable distribution was defined as follows.

Definition 2.18. A symmetric integer-valued random variable X is said to be symmetric discrete stable if there exists a sequence $\{p_n\} \in (0,1)$, $p_n \searrow 0$, such that

$$X \stackrel{d}{=} \lim_{n\to\infty} \sum_{i=1}^{n} \tilde{X}_i(p_n), \quad \text{where} \quad \tilde{X}(p) = \sum_{j=1}^{|X|} \varepsilon_j, \tag{2.7}$$

where ε_j are i.i.d. random variables with two-sided Bernoulli distribution, i.e. $\mathbb{P}(\varepsilon = \pm 1) = p$ and $\mathbb{P}(\varepsilon = 0) = 1 - 2p$, and $X_1, X_2, \dots$ are independent copies of X.

We showed that this definition leads to a distribution with probability generating function

$$\mathcal{P}(z) = \exp\left\{-\lambda\left(1 - \frac{1}{2}\left(z + \frac{1}{z}\right)\right)^\gamma\right\}, \quad \gamma \in (0,1],\, \lambda > 0.$$

The characteristic function takes form $f(t) = \exp\{-\lambda(1-\cos t)^\gamma\}$. It was shown that symmetric discrete stable distribution is a discrete analogy of the symmetric α-stable distribution with $\alpha = 2\gamma$ in the following sense: let X be a symmetric discrete stable random variable and let $a > 0$. Consider a random variable $X^a = aX$ taking values in $a\mathbb{Z} = \{0, \pm a, \pm 2a, \dots\}$ and let $\lambda \approx a^{-2\gamma}$. Then $X^a \xrightarrow{d} \mathrm{S}(2\gamma, 0, \tau, 0)$. Slámová and Klebanov (2012) gave a stochastic representation of a symmetric discrete stable random variable X.

Theorem 2.19. *A symmetric discrete stable random variable with parameters γ, λ is distributed as a compound Poisson random variable with intensity $\lambda^{1/\gamma}\mathrm{S}_\gamma$ and jumps taking values ± 1 with the same probability, and where S_γ is a positive stable random variable with Laplace transform* $\exp\{-u^\gamma\}$.

The asymmetric discrete stable distribution was also defined by Klebanov and Slámová (2013).

Definition 2.20. An integer-valued random variable X is said to be discrete stable if there exist sequences $\{p_n^{(1)}\}, \{p_n^{(2)}\} \in (0,1)$, $p_n^{(1)} \downarrow 0$, $p_n^{(2)} \downarrow 0$, such that

$$X \stackrel{d}{=} \lim_{n\to\infty} \sum_{i=1}^{n} \bar{X}_i\left(p_n^{(1)}, p_n^{(2)}\right), \quad \text{where} \quad \bar{X}_i\left(p_n^{(1)}, p_n^{(2)}\right) = \sum_{j=1}^{X_i^+} \varepsilon_j^{(i)}(p_n^{(1)}) - \sum_{j=1}^{X_i^-} \epsilon_j^{(i)}(p_n^{(2)}), \tag{2.8}$$

where $\varepsilon_j^{(i)}(p)$, $\epsilon_j^{(i)}(p)$ are i.i.d. random variables with Bernoulli distribution, i.e. $\mathbb{P}(\varepsilon_j^{(i)}(p) = 1) = \mathbb{P}(\epsilon_j^{(i)}(p) = 1) = 1 - \mathbb{P}(\varepsilon_j^{(i)}(p) = 0) = 1 - \mathbb{P}(\epsilon_j^{(i)}(p) = 0) = p$, and $X_1, X_2, \dots$ are independent copies of X.

Remark 2.21. Barabesi and Pratelli (2014) used the name asymmetric discrete stable for this distribution.

This definition leads to a distribution with probability generating function

$$\mathcal{P}(z) = \exp\left\{-\lambda_1 (1-z)^\gamma - \lambda_2 (1-1/z)^\gamma\right\}, \quad \gamma \in (0,1],\ \lambda_1, \lambda_2 > 0.$$

The case of $\gamma = 1$ corresponds to a distribution known as Skellam distribution, as it was introduced by Skellam (1946) as a difference of two independent Poisson random variables with parameters λ_1 and λ_2 respectively.

It was shown that discrete stable distribution is a discrete analogy of the α-stable distribution with $\alpha = \gamma$ for $\gamma \in (0,1)$ and $\alpha = 2$ for $\gamma = 1$. It follows directly from the probability generating function that a discrete stable random variable can be represented as a difference of two positive discrete stable random variables.

2.3 ν-stable distributions

The general theory of stability of sums of random number of random variables was introduced in Klebanov and Rachev (1996). They introduced the concept of ν-infinitely divisible and ν-stable distributions. Their approach generalized the class of geometric infinitely divisible and geometric stable distributions that were introduced in Klebanov et al. (1985). Special cases of geometric stable distributions are Mittag-Leffler distribution and Linnik distribution.

We review here the definitions and results obtained by Klebanov et al. (1985), followed by Klebanov and Rachev (1996) and Klebanov et al. (2012) . Let $X_1, X_2, \dots$ be a sequence of i.i.d. random variables. Let us assume that $\{\nu_p, p \in \Delta\}, \Delta \subset (0,1)$, is a family of non-negative integer-valued random variables that are independent of the sequence $\{X_i, i = 1, 2, \dots\}$. We will further assume that $\mathrm{E}\nu_p$ exists and that $\mathrm{E}\nu_p = 1/p$ for all $p \in \Delta$. We remind that a random variable Y is called infinitely divisible if for all $n \in \mathbb{N}$ there exists a sequence of i.i.d. random variables $X_1, \dots, X_n$, such that

$$Y \stackrel{d}{=} X_1 + X_2 + \dots + X_n.$$

The definition of ν-infinitely divisible distribution is as follows.

Definition 2.22. A random variable Y is called ν-infinitely divisible if for any $p \in \Delta$ there exists a sequence of i.i.d. random variables $\{X_j^{(p)}, j = 1, 2, \dots\}$ independent of ν_p such that

$$Y \stackrel{d}{=} \sum_{j=1}^{\nu_p} X_j^{(p)}. \tag{2.9}$$

A special case of infinite divisible distribution is a Gaussian distribution, and the ν-Gaussian analogy is of particular interest as well.

Definition 2.23. A random variable X is called strictly ν-Gaussian, if $\mathrm{E}X = 0, \mathrm{E}X^2 < \infty$ and for all $p \in \Delta$

$$X \stackrel{d}{=} p^{1/2} \sum_{j=1}^{\nu_p} X_j,$$

where $X_1, X_2, \dots$ are independent copies of X, moreover independent of $\{\nu_p, p \in \Delta\}$.

Let us denote by $\mathcal{P}_p$ the probability generating function of ν_p and by $\mathfrak{P}$ a semigroup with operation of superposition $\circ$ generated by the family $\{\mathcal{P}_p, p \in \Delta\}$.

Theorem 2.24. *A strictly ν-Gaussian random variable X exists if and only if the semigroup $\mathfrak{P}$ is commutative.*

Let us consider the following system of functional equations (that appears in the proof of Theorem 2.24):

$$\varphi(t) = \mathcal{P}_p(\varphi(t)) \quad \text{for all} \quad p \in \Delta \tag{2.10}$$

with initial conditions

$$\varphi(0) = 1,$$
$$\varphi'(0) = -1.$$

The equation (2.10) is called Poincaré equation. If the semigroup $\mathfrak{P}$ is commutative, then this system has a unique solution (for proof, see for example Gnedenko and Korolev (1996)).

The following theorem shows a one-on-one map relationship between infinitely divisible distributions and ν-infinitely divisible distributions.

Theorem 2.25. *A characteristic function g is ν-infinitely divisible if and only if it is representable in the form of*

$$g(t) = \varphi(-\log f(t)), \tag{2.11}$$

where $f(t)$ is an infinitely divisible characteristic function and φ is a standard solution of (2.10).

The previous theorem is a corner stone for defining strictly ν-stable distribution.

Definition 2.26. A function $g(t)$ is called a ν-stable (strictly ν-stable) characteristic function if it admits representation (2.11) in which φ is a standard solution of (2.10) and $f(t)$ is the characteristic function of a stable (strictly stable) distribution.

Example 2.27 (Classical summation scheme)**.** If we consider ν_p to be deterministic, namely $\nu_p = \frac{1}{p}$ with probability 1 and $p \in \Delta = \{1/n, n \in \mathbb{N}\}$, we obtain the classical stability. We have $\mathcal{P}_p(z) = z^{1/p}$, and obviously $\mathfrak{P}$ is commutative as $\mathcal{P}_{p_1} \circ \mathcal{P}_{p_2}(z) = z^{1/(p_1 p_2)}$. The Poincaré equation (2.10) takes form

$$\varphi(t) = \varphi^n(t/n),$$

which has solution $\varphi(t) = \exp\{-t\}$. Therefore via the one-on-one map (2.11), for strictly stable characteristic function f we obtain strictly ν-stable characteristic function $g(t) = f(t)$.

Example 2.28 (Geometric summation scheme)**.** Assume that ν_p is a geometric random variable with parameter p: $\mathbb{P}(\nu_p = k) = p(1-p)^{k-1}$, $k \in \mathbb{N}$. We have $\mathcal{P}_p(z) = pz/(1-(1-p)z)$. We can easily see that $\mathcal{P}_{p_1} \circ \mathcal{P}_{p_2}(z) = p_1 p_2/(1 - (1 - p_1 p_2)z)$ and therefore the semigroup generated by $\{\mathcal{P}_p\}$ is commutative. The equation (2.10) takes form

$$\varphi(t) = \frac{p_0 \varphi(p_0 t)}{1 - (1 - p_0)\varphi(p_0 t)}.$$

The solution of this equation is the Laplace transform of exponential distribution,

$$\varphi(t) = \frac{1}{1+t}.$$

Therefore the geometric stable distribution has characteristic function

$$g(t) = \frac{1}{1 - \log f(t)},$$

where $f(t)$ is the characteristic function of stable distribution. As a special case we obtain ν-analogy of Gaussian distribution with characteristic function $1/(1 + at^2)$ which correspond to the Laplace distribution.

Example 2.29 (Chebyshev summation scheme)**.** This example was obtained by Klebanov et al. (2012). Let us assume that the number of summands $\nu_p, p \in \Delta$, with $\Delta = \{1/n^2, n \in \mathbb{N}\}$, comes from a distribution with probability generating function

$$\mathcal{P}_p(z) = \frac{1}{T_{1/\sqrt{p}}(z)}, \quad p \in \Delta,$$

where $T_n(x)$ is the Chebyshev polynomial of degree n, defined as $T_n(x) = \cos(n \arccos(x))$. Chebyshev polynomials are commutative, as $T_n \circ T_m(x) = T_m \circ T_n(x)$, therefore the semigroup $\mathfrak{P}$ is commutative as well. The Poincaré equation (2.10) has a solution

$$\varphi(t) = \frac{1}{\cosh(\sqrt{2t})},$$

satisfying the initial conditions. The ν-Gaussian random variable has characteristic function

$$g(t) = \frac{1}{\cosh(at)},$$

corresponding to hyperbolic secant distribution.

3. Discrete approximations of stable distributions

Stable distributions are on the rise in financial applications since Mandelbrot (1963) noted that Gaussian distribution does not provide a good fit for financial returns that exhibit leptokurtic behaviour and heavy tails. We say that a distribution has heavy tails if the variance of the distribution is not finite. However, the infinite variance of stable distributions and the fact that financial returns have heavier tails on a short time scale and almost Gaussian on a long scale brings into question the appropriateness of the stable model of returns. Grabchak and Samorodnitsky (2010) studied this paradox and suggested that a more appropriate model for financial returns is obtained by tempering of tails of a heavy tailed distribution to get a distribution that behaves like the original heavy tailed in the middle but whose tails are exponential. They show, using the pre-limit theorem by Klebanov et al. (1999), that the sum of a large number of independent and identically distributed random variables behave as a stable random variable even though the tails of the random variables are not heavy.

Stable distributions with exponentially tempered tails have been considered in the literature under different names – truncated Lévy flights (Koponen (1995)), CGMY model (Carr et al. (2002)) and finally tempered stable distributions (Rosiński (2007)). Tempered stable distributions appear by exponential tilting of the Lévy measure of stable distributions. The resulting distributions have finite moments of any order and exponential tails. Another variant of stable distribution with Gaussian tails was introduced in Menn and Rachev (2009). The smoothly truncated stable density corresponds to that of a Gaussian distribution on the tails and to that of a stable distribution in the centre, with additional conditions on the smoothness of the distribution function.

In this Chapter, we focus on discrete approximations of stable distributions with heavy, exponential and Gaussian tails and therefore offer an alternative to the stable, tempered stable and smoothly truncated stable distributions respectively. In the first and second Sections, we introduce two approximations of the stable characteristic function leading to discrete distributions that were introduced in Klebanov and Slámová (2013). Also we study two different approximations with Gaussian and exponential tails, appearing as a result of truncation and tempering the heavy tails of the discrete stable distributions. In the third Section, we study a discrete approximation resulting from discretizing the Lévy measure of stable distributions. We obtain a discrete distribution that allows for the tail index to be an arbitrary positive number (in this case it is not index of stability, because these distributions are not stable).

This Chapter contains results from Slámová and Klebanov (2014a).

3.1 First approximation

Let us first consider the case of symmetric α-stable distributions. Their characteristic functions are given by the following formula

$$f(t) = \exp\left\{-\sigma^{\alpha}|t|^{\alpha}\right\},$$

with $\alpha \in (0,2]$ being the index of stability and $\sigma > 0$ being the scaling parameter. For arbitrary t we write $|t|^\alpha = (t^2)^\gamma$, where $\gamma = \alpha/2$. Let us use the following approximation. We have

$$t^2 = \lim_{a\to 0} \frac{2}{a^2}(1-\cos(at)),$$

therefore let us write $t^2 \approx \frac{2}{a^2}(1-\cos(at))$, as $a \to 0$. Hence the characteristic function of symmetric α-stable distribution can be approximated as

$$\log f(t) = -\sigma^\alpha |t|^\alpha \approx \log g(t,a) = -\sigma^{2\gamma}\frac{2^\gamma}{a^{2\gamma}}(1-\cos(at))^\gamma,$$

for small values of a.

Lemma 3.1. *The function*

$$g(t,a) = \exp\left\{-\sigma^{2\gamma}\frac{2^\gamma}{a^{2\gamma}}(1-\cos(at))^\gamma\right\}$$

is a characteristic function of a distribution given on the lattice $a\mathbb{Z} = \{0, \pm a, \pm 2a, \dots\}$ for any positive a.

Proof. We can rewrite $g(t,a)$ as $g(t,a) = \exp\{-\lambda(1-h(t,a))\}$, where

$$h(t,a) = 1-(1-\cos(at))^\gamma = \sum_{k=1}^{\infty}\binom{\gamma}{k}(-1)^{k-1}\cos(at)^k.$$

The series coefficients are positive for $\gamma \in (0,1]$, moreover $h(0,a) = 1$, $h(t,a)$ is periodic with period $2\pi a$, hence the function $h(t,a)$ is a characteristic function of a random variable on $a\mathbb{Z}$. Therefore $g(t,a)$ is a characteristic function of compound Poisson random variable with intensity of jumps λ and jumps in $a\mathbb{Z}$ with characteristic function $h(t,a)$. □

It is clear that

$$\lim_{a\to 0} g(t,a) = f(t),$$

and therefore $g(t,a)$ can be considered as a discrete approximation of $f(t)$ for a sufficiently small a. This distribution with $a = 1$ was introduced in Klebanov and Slámová (2013) by considering a discrete analogy of the stability property $X = n^{-1/\alpha}(X_1 + X_2 + \cdots + X_n)$ and called symmetric discrete stable (SDS) distribution.

It is obvious from the construction of γ-symmetric discrete stable distribution that it belongs to the domain of normal attraction of 2γ-stable distribution. From the known characterization of the domain of attraction of stable distributions (see, for example, Ibragimov and Linnik (1971)), a SDS random variable must satisfy the following tail assumptions as $x \to \infty$

$$\lim_{x\to\infty} x^{2\gamma}\mathbb{P}(|X| > x) = \begin{cases} \lambda\frac{a^{2\gamma}}{2^\gamma}\frac{1}{\Gamma(1-2\gamma)\cos(\pi\gamma)}, & \text{if } \gamma \neq \frac{1}{2}, \\ \lambda\frac{a^{2\gamma}}{2}\frac{2}{\pi}, & \text{if } \gamma = \frac{1}{2}. \end{cases} \tag{3.1}$$

So far we have introduced a discrete approximation of the symmetric α-stable distribution that has the same tail behaviour. Another approximation leading to a distribution with Gaussian tails can be obtained in the following way. Let us consider a function

$$g(t,a,M) = \exp\left\{-\lambda\sum_{k=1}^{M}(-1)^k\binom{\gamma}{k}\cos(at)^k + \lambda\sum_{k=1}^{M}(-1)^k\binom{\gamma}{k}\right\}.$$

For sufficiently large values of M this function can be considered an approximation of the function $g(t,a)$, as

$$\lim_{M\to\infty} g(t,a,M) = g(t,a).$$

So it is also a discrete approximation of symmetric α-stable distribution, as

$$\lim_{a\to 0}\lim_{M\to\infty} g(t,a,M) = f(t).$$

However, we cannot exchange the order of the limits.

Both characteristic functions $g(t,a)$ and $g(t,a,M)$ are infinitely divisible as they correspond to compound Poisson distributions. The distributions of jumps are given by $h(t,a) = 1-(1-\cos(at))^\gamma$ and $h(t,a,M) = 1-\sum_{k=1}^{M}(-1)^k\binom{\gamma}{k}\cos(at)^k + \sum_{k=1}^{M}(-1)^k\binom{\gamma}{k}$ respectively. It can be verified that these distributions have no mass at 0 and the second distribution has truncated jumps in absolute value larger than aM. We will call the distribution given by characteristic function $g(t,a,M)$ approximate symmetric discrete stable distribution.

The characteristic function is an entire function hence the tails of the approximate symmetric discrete stable distribution behave like $o(\exp(-bx))$, as $x\to\infty$, for all $b>0$ by the Raikov's theorem (see, for example, Linnik (1964)). The approximate symmetric discrete stable distribution thus belongs to the domain of normal attraction of Gaussian distribution and as such has finite variance. It follows from the pre-limit theorem of Klebanov et al. (1999) that for not too large values of n the sum $S_n = n^{-1/2\gamma}(X_1+\cdots+X_n)$ behaves like symmetric α-stable distribution with $\alpha = 2\gamma$. This property is due to the truncation of the bigger jumps, therefore the distribution behaves like stable distribution in the middle, and like Gaussian on the tails.

3.2 Second approximation

In the previous Section, we introduced a discrete approximation of symmetric α-stable distribution. Here we give a discrete approximation of α-stable distribution with index of stability $\alpha\in(0,1)$ and skewness $\beta\in[-1,1]$. The characteristic function of strictly α-stable distribution with skewness parameter β and scale parameter $\sigma>0$ is given by

$$f(t) = \exp\left\{-\sigma^\alpha|t|^\alpha\left(1-\mathrm{i}\beta\,\mathrm{sign}(t)\tan\frac{\pi\alpha}{2}\right)\right\}.$$

We can rewrite this as

$$\log f(t) = -\lambda_1(-\mathrm{i}t)^\alpha - \lambda_2(\mathrm{i}t)^\alpha,$$

where

$$\lambda_1 = \frac{\sigma^\alpha}{\cos\frac{\pi\alpha}{2}}\frac{1+\beta}{2},\quad \lambda_2 = \frac{\sigma^\alpha}{\cos\frac{\pi\alpha}{2}}\frac{1-\beta}{2}.$$

We use the following approximation: $\mathrm{i}t\approx(1-e^{-\mathrm{i}at})/a$ as $a\to 0$ and $-\mathrm{i}t =\approx (1-e^{\mathrm{i}at})/a$ as $a\to 0$, therefore the characteristic function of α-stable distribution can be approximated by a characteristic function of a discrete distribution as

$$\log f(t)\approx\log g(t,a) = -\frac{\lambda_1}{a^\alpha}(1-e^{\mathrm{i}at})^\alpha - \frac{\lambda_2}{a^\alpha}(1-e^{-\mathrm{i}at})^\alpha,\quad \text{as}\quad a\to 0.$$

This distribution for $a = 1$ was introduced in Klebanov and Slámová (2013) as discrete stable distribution and it was shown there that $g(t,a)$ is a characteristic function only for $\alpha \in (0,1]$. From the construction of the approximation we see that

$$\lim_{a \to 0} g(t,a) = f(t).$$

Discrete stable distribution has therefore the same behaviour of tails as α-stable distribution and it is again infinitely divisible.

We can obtain yet another discrete approximation of α-stable distribution with exponential tails by tempering the tails of discrete stable distribution. Because discrete stable distribution is a compound Poisson distribution with intensity $\lambda_1 + \lambda_2$ and distribution of jumps with characteristic function

$$h(t,a) = 1 - \frac{\lambda_1}{\lambda_1+\lambda_2}(1 - e^{\mathrm{i}at})^{\alpha} - \frac{\lambda_2}{\lambda_1+\lambda_2}(1 - e^{-\mathrm{i}at})^{\alpha},$$

the Lévy-Khintchine representation of discrete stable characteristic function takes the following form

$$\log g(t,a) = \int_{-\infty}^{\infty} \left(e^{\mathrm{i}atx} - 1\right) \nu(\mathrm{d}x),$$

where $\nu(\mathrm{d}x)$ is the Lévy measure,

$$\nu(\mathrm{d}x) = (\lambda_1 + \lambda_2) \sum_{k=-\infty}^{\infty} p_k \delta_{ak}(\mathrm{d}x),$$

where

$$p_k = \begin{cases} \frac{\lambda_1}{\lambda_1+\lambda_2}(-1)^{k+1}\binom{\alpha}{k}, & k > 0, \\ \frac{\lambda_2}{\lambda_1+\lambda_2}(-1)^{k+1}\binom{\alpha}{|k|}, & k < 0, \\ 0, & k = 0. \end{cases}$$

and δ_k is the Dirac measure, i.e. $\delta_x(A) = 1$ if $x \in A$ and 0 otherwise. The classical idea leading to tempered infinitely divisible distribution consists of exponential tempering of the corresponding Lévy measure (Rosiński (2007)). We will use tempering function of the form $q(x) = e^{-\theta_1 x}\mathbf{1}_{x>0} + e^{-\theta_2|x|}\mathbf{1}_{x<0}$. The tempered infinitely divisible distribution is then obtained by multiplying the Lévy measure by this tempering function. As a result we obtain a distribution with characteristic function

$$\begin{aligned}\log g(t,a,\theta_1,\theta_2) = -\lambda_1 \left(1 - e^{\mathrm{i}at}e^{-\theta_1}\right)^{\alpha} - \lambda_2 \left(1 - e^{-\mathrm{i}at}e^{-\theta_2}\right)^{\alpha} \\ + \lambda_1 \left(1 - e^{-\theta_1}\right)^{\alpha} + \lambda_2 \left(1 - e^{-\theta_2}\right)^{\alpha}.\end{aligned}$$

This characteristic function is an analytic function in the strip $\mathrm{Im}(t) \in (-\theta_2, \theta_1)$ and by the Raikov's theorem (see Linnik (1964)) the tails are $O(\exp(-bx))$, as $x \to \infty$, for all $b > 0$. Therefore the tempered discrete stable distribution belong to the domain of normal attraction of Gaussian distribution. By the pre-limit theorem of Klebanov et al. (1999) we can show that for not too large values of n, the normalized sum $S_n = n^{-1/\alpha}(X_1 + \cdots + X_n)$ behaves like α-stable distribution.

3.3 Third approximation

Another way to find a discrete approximation of strictly stable distributions is to discretize its Lévy (or spectral) measure. Stable distribution is an infinitely divisible distribution and as such has a Lévy-Khintchine representation of its characteristic function. This representation takes the following form (see, for example, (Zolotarev, 1986, § 34) or Samorodnitsky and Taqqu (1994))

$$\log f(t) = P \int_0^\infty \left(e^{\mathrm{i}tx} - 1 - \mathrm{i}tx\mathbf{1}_{|x|\leq 1}\right) \frac{\mathrm{d}x}{x^{1+\alpha}} + Q \int_{-\infty}^0 \left(e^{\mathrm{i}tx} - 1 - \mathrm{i}tx\mathbf{1}_{|x|\leq 1}\right) \frac{\mathrm{d}x}{|x|^{1+\alpha}}, \tag{3.2}$$

where $P, Q > 0$, $0 < \alpha < 2$. The Lévy measure is

$$U(\mathrm{d}x) = \frac{P}{x^{1+\alpha}}\mathbf{1}_{x>0}(x)\mathrm{d}x + \frac{Q}{|x|^{1+\alpha}}\mathbf{1}_{x<0}(x)\mathrm{d}x.$$

Every infinitely divisible random variable is a limit of compound Poisson random variables, and the Lévy measure $U(\mathrm{d}x)$ express the intensity of jumps of size x. The term $\mathrm{i}tx\mathbf{1}_{|x|\leq 1}$ is the compensation of the small jumps that ensures that the integral converges. If we discretize the Lévy measure we obtain discrete infinitely divisible distribution. This can be achieved by discretizing the integrals in (3.2), as follows

$$\begin{aligned} \log f(t) \approx \log g(t,a) = P \sum_{k=1}^{\infty} \left(e^{\mathrm{i}tak} - 1\right) \frac{a}{(ak)^{1+\alpha}} \\ + Q \sum_{k=-\infty}^{-1} \left(e^{\mathrm{i}tak} - 1\right) \frac{a}{|ak|^{1+\alpha}}, \quad \text{as} \quad a \to 0. \end{aligned} \tag{3.3}$$

Here we omit the compensation of small jumps. The distribution given by characteristic function $g(t,a)$ is still infinitely divisible, with the Lévy measure given by

$$V(\mathrm{d}x) = \sum_{k=-\infty}^{\infty} P\frac{a}{(ak)^{1+\alpha}}\mathbf{1}_{k>0}(k)\delta_{ak}(\mathrm{d}x) + Q\frac{a}{|ak|^{1+\alpha}}\mathbf{1}_{k<0}(k)\delta_{ak}(\mathrm{d}x), \quad k \in \mathbb{Z}.$$

As a result, we obtain a distribution with characteristic function

$$\log g(t,a) = \frac{1}{a^\alpha}\left(P\mathrm{Li}_{1+\alpha}\left(e^{\mathrm{i}at}\right) + Q\mathrm{Li}_{1+\alpha}\left(e^{-\mathrm{i}at}\right) - (P+Q)\zeta(1+\alpha)\right),$$

where $\mathrm{Li}_{1+\alpha}(x)$ is the polylogarithm function and $\zeta(1+\alpha)$ is the Zeta function. It is interesting to note that g is a characteristic function for all positive values of α. We can rewrite this with $\sigma = (P+Q)$ and $\beta = \frac{P}{P+Q}$ to obtain

$$g(t,a) = \exp\left\{\frac{\sigma}{a^\alpha}\left(\beta\mathrm{Li}_{1+\alpha}\left(e^{\mathrm{i}at}\right) + (1-\beta)\mathrm{Li}_{1+\alpha}\left(e^{-\mathrm{i}at}\right) - \zeta(1+\alpha)\right)\right\}.$$

By truncating the series in (3.3) we obtain yet another approximation of α-stable distribution by an entire characteristic function.

4. On definitions of discrete stability

In the previous Chapter we introduced a possible approach to obtain discrete analogies of stable distributions. We obtained three discrete distributions by approximation of the characteristic function of stable distribution or of its Lévy measure. These distributions are discrete approximations of the stable distributions and it is not clear what properties they share with the stable distributions – by the construction it is obvious they have the same tail behaviour, but it is not clear whether they share other properties as the stability property, self-similarity, infinite divisibility and others. In this Chapter we define three new classes of discrete probability distributions by generalizing the stability property for discrete random variables.

The strict stability property of continuous random variables can be defined in several ways (see Definitions 2.1 – 2.3). We say that a random variable X is strictly stable if one of the following holds

$$X \stackrel{d}{=} a_n \sum_{i=1}^{n} X_i, \tag{4.1}$$

$$A_n X \stackrel{d}{=} \sum_{i=1}^{n} X_i, \tag{4.2}$$

$$cX \stackrel{d}{=} aX_1 + bX_2, \tag{4.3}$$

where $X_1, X_2, \ldots, X_n$ are independent copies of X and a_n, A_n, a, b and c are positive constants. If we want to define a discrete analogy of stability we have to reconsider the normalization by the constants a_n, A_n, and a, b and c, as the normalized random variables are not necessarily integer-valued. We may consider the following modification. Consider for example the first definition and let us assume that X is non-negative integer-valued random variable. We may write

$$X = \underbrace{1 + 1 + \cdots + 1}_{X \text{ times}}, \quad \text{and} \quad pX = \underbrace{p + p + \cdots + p}_{X \text{ times}},$$

where we normalize X by a constant $p \in (0, 1)$. Instead we can consider a thinning operator $p \odot X$, where

$$p \odot X = \underbrace{\varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_X}_{X \text{ times}},$$

where ε_i are i.i.d. Bernoulli random variables with $\mathrm{E}\varepsilon_i = p$, i.e.

$$\varepsilon_i = \begin{cases} 1, & \text{with probability } p, \\ 0, & \text{with probability } 1 - p. \end{cases}$$

In the following Sections we introduce three different definitions of discrete stability generalizing the definitions of strict stability 2.1 – 2.3 for the case of non-negative integer-valued random variables.

4.1 On first definition of discrete stable distributions

In this Section we give a definition of discrete stability that generalizes the first definition of strict stability (4.1) for discrete random variables. The multiplication by a constant a_n can be understood as a normalization of the sum $\sum X_i$, or normalization of the individual summands X_i. In the case of discrete random variables one can not use this normalization as it violates the integral property of the summands. We need to find a different normalization that maintains the integral property. One possibility is to use the binomial thinning operator

$$\tilde{X}(\alpha) = \alpha \odot X = \sum_{i=1}^{X} \epsilon_i, \quad \text{where} \quad \mathbb{P}(\epsilon_i = 1) = 1 - \mathbb{P}(\epsilon_i = 0) = \alpha,$$

instead of $a_n X_i$. This normalization was used in Steutel and van Harn (1979) to define discrete stability on $\mathbb{N}_0$. One can generalize this definition of discrete stability by considering a general normalization, or "thinning" operator.

Definition 4.1. Let $X, X_1, X_2, \ldots, X_n, \ldots$ denote a sequence of independent and identically distributed (i.i.d.) non-negative integer-valued random variables. Assume that for every $n \in \mathbb{N}$ there exists a constant $p_n \in (0,1)$ such that

$$X \stackrel{d}{=} \sum_{i=1}^{n} \tilde{X}_i(p_n), \quad \text{where} \quad \tilde{X}_i(p_n) = p_n \odot X_i = \sum_{j=1}^{X_i} \varepsilon_j^{(i)}(p_n), \tag{4.4}$$

and $\varepsilon_j^{(i)}(p_n)$ are i.i.d. non-negative integer-valued random variables. Then we say that X is *positive discrete stable random variable in the first sense.*

This definition is rather general as it offers a flexibility on the choice of the "thinning" distribution of random variables ε. This flexibility is however limited as a positive discrete stable random variable exists only for some choice of the thinning distribution. A question is therefore how to describe the family of thinning distributions for which a positive discrete stable random variables exists.

Let us denote the probability generating functions of the random variables X and $\varepsilon(p_n)$ by $\mathcal{P}(z) = \mathrm{E}[z^X]$ and $\mathcal{Q}_{p_n}(z) = \mathrm{E}[z^{\varepsilon(p_n)}]$ respectively. There is an equivalent definition of positive discrete stability in terms of those probability generating functions.

Proposition 4.2. *A random variable X is positive discrete stable if and only if for all $n \in \mathbb{N}$ there exists a constant $p_n \in (0,1)$ such that*

$$\mathcal{P}(z) = \mathcal{P}^n(\mathcal{Q}_{p_n}(z)). \tag{4.5}$$

Proof. It follows from the definition (4.4) that X is positive discrete stable if and only if

$$\mathcal{P}(z) = \left[\mathcal{P}_{\tilde{X}}(z)\right]^n.$$

The probability generating function of $\tilde{X}$ can be computed in the following way.

$$\begin{aligned}\mathcal{P}_{\tilde{X}}(z) = \mathrm{E}\left[z^{\tilde{X}}\right] &= \sum_{k=0}^{\infty} \mathbb{P}(X = k)\mathrm{E}\left[z^{\sum_{j=1}^{X} \varepsilon_j(p_n)} | X = k\right] \\ &= \sum_{k=0}^{\infty} \mathbb{P}(X = k)\left(\mathrm{E}\left[z^{\varepsilon_1(p_n)}\right]\right)^k \\ &= \mathcal{P}(\mathcal{Q}_{p_n}(z)).\end{aligned}$$

Hence X is positive discrete stable if and only if its probability generating function satisfy the relation

$$\mathcal{P}(z) = \mathcal{P}^n(\mathcal{Q}_{p_n}(z)).$$

□

Remark 4.3. It follows from the definition that a positive discrete stable random variable X is infinitely divisible: for every $n \in \mathbb{N}$ there exist random variables $Y_1, Y_2, \dots, Y_n$ such that

$$X \stackrel{d}{=} Y_1 + Y_2 + \cdots + Y_n.$$

This obviously holds for $Y_i = \tilde{X}_i(p_n)$.

Further denote by $\mathbb{Q}$ a semigroup generated by the family of probability generating functions $\{\mathcal{Q}(z) = \mathcal{Q}_{p_n}(z), n \in \mathbb{N}\}$ with operation of superposition $\circ$. It can be shown that a superposition of two probability generating functions is again a probability generating function.

Lemma 4.4. *If $\mathcal{Q}_1(z)$ and $\mathcal{Q}_2(z)$ are two probability generating functions of two random variables with values in $\mathbb{N}_0$, then their superposition*

$$\mathcal{Q}_1 \circ \mathcal{Q}_2(z) := \mathcal{Q}_1(\mathcal{Q}_2(z))$$

is also a probability generating function of some random variable with values in $\mathbb{N}_0$.

Proof. Let N be a random variable with values in $\mathbb{N}_0$ with probability generating function $\mathcal{Q}_1$ and $X_1, X_2, \dots$ i.i.d. random variables with values in $\mathbb{N}_0$ with probability generating function $\mathcal{Q}_2$. Define a new random variable S by

$$S = \sum_{i=1}^{N} X_i.$$

Then S is a random variable with values in $\mathbb{N}_0$. Its probability generating function can be computed using the fundamental formula of conditional expectation as follows

$$\begin{aligned}
\mathcal{Q}_S(z) &= \mathrm{E}\left[z^S\right] = \mathrm{E}\left[z^{\sum_{i=1}^N X_i}\right] = \sum_{n=0}^{\infty} \mathbb{P}(N=n)\mathrm{E}\left[z^{\sum_{i=1}^N X_i} | N=n\right] \\
&= \sum_{n=0}^{\infty} \mathbb{P}(N=n)\mathrm{E}\left[z^{\sum_{i=1}^n X_i}\right] = \sum_{n=0}^{\infty} \mathbb{P}(N=n)\left[\mathrm{E}z^{X_1}\right]^n \\
&= \sum_{n=0}^{\infty} \mathbb{P}(N=n)\left[\mathcal{Q}_2(z)\right]^n = \mathcal{Q}_1(\mathcal{Q}_2(z)).
\end{aligned}$$

So the superposition $\mathcal{Q}_1 \circ \mathcal{Q}_2(z)$ is a probability generating function of random variable S with values in $\mathbb{N}_0$. □

Now we show that the semigroup $\mathbb{Q}$ must be commutative.

Theorem 4.5. *Let X be a positive discrete stable random variable. Then the semigroup $\mathbb{Q}$ must be commutative.*

Proof. Let us denote $G(z) = \log \mathcal{P}(z)$. Then (4.5) is equivalent to

$$G(z) = nG(\mathcal{Q}_{p_n}(z)), \quad n \in \mathbb{N}. \tag{4.6}$$

Let $G(z)$ be a solution of (4.6). Then for all $n \in \mathbb{N}$ it must hold

$$\mathcal{Q}_{p_n}(z) = G^{-1}\left(\frac{1}{n}G(z)\right).$$

It follows from here that for all $n_1, n_2 \in \mathbb{N}$

$$\begin{aligned}\mathcal{Q}_{p_{n_1}}\left(\mathcal{Q}_{p_{n_2}}(z)\right) &= G^{-1}\left(\frac{1}{n_1}G\left(G^{-1}\left(\frac{1}{n_2}G(z)\right)\right)\right)\\ &= G^{-1}\left(\frac{1}{n_1}\frac{1}{n_2}G(z)\right)\\ &= \mathcal{Q}_{p_{n_2}}\left(\mathcal{Q}_{p_{n_1}}(z)\right),\end{aligned}$$

which means that $\mathbb{Q}$ is commutative. □

Similarly as for the classical stable distribution, we can show that the constants p_n have to take form $p_n = n^{-1/\gamma}$ for some $\gamma > 0$.

Theorem 4.6. *Let X be a positive discrete stable random variable in the first sense. Then there exists $\gamma > 0$ such that p_n in (4.4) takes form*

$$p_n = n^{-1/\gamma}.$$

Proof. The proof follows (Uchaikin and Zolotarev, 1999, §2.4) where a similar statement for stable distributions is proved. From the definition it follows that for every $n \geq 2$ we have $X \stackrel{d}{=} \sum_{i=1}^{n} \tilde{X}_i(p_n)$ where $X_1, X_2, \ldots$ are independent copies of X. Then

$$X \stackrel{d}{=} p_2 \odot X_1 + p_2 \odot X_2,$$

therefore also

$$X \stackrel{d}{=} p_2 \odot (p_2 \odot X_1 + p_2 \odot X_2) + p_2 \odot (p_2 \odot X_3 + p_2 \odot X_4).$$

But the operation $\odot$ is associative: $p \odot (p \odot X) = p^2 \odot X$. Let us denote $Y = p_2 \odot X_1 + p_2 \odot X_2$. Then (using result from proof of Proposition 4.2)

$$\mathcal{P}_{p_2 \odot Y}(z) = \mathcal{P}_Y(\mathcal{Q}_{p_2}(z)) = \mathcal{P}^2(\mathcal{Q}_{p_2}(\mathcal{Q}_{p_2}(z))) = \mathcal{P}^2(\mathcal{Q}_{p_2^2}(z)),$$

because $\mathbb{Q}$ is commutative. Therefore

$$X \stackrel{d}{=} p_2^2 \odot X_1 + p_2^2 \odot X_2 + p_2^2 \odot X_3 + p_2^2 \odot X_4$$

and similarly for every $n = 2^k$

$$X \stackrel{d}{=} p_2^k \odot X_1 + p_2^k \odot X_2 + \cdots + p_2^k \odot X_n. \tag{4.7}$$

On the other hand, we have

$$X \stackrel{d}{=} p_n \odot X_1 + p_n \odot X_2 + \cdots + p_n \odot X_n. \tag{4.8}$$

Comparing (4.7) with (4.8), with $n = 2^k$, we have $p_n = p_2^k$. Hence

$$\log p_n = k \log p_2 = \frac{\log n}{\log 2} \log p_2 = \log n^{\log p_2/\log 2}.$$

So we obtain that

$$p_n = n^{-1/\gamma_2}, \quad \gamma_2 = -\log 2/\log p_2 > 0, \quad n = 2^k, k = 1, 2, \dots .$$

In a similar way, starting with sums with 3 terms $X \stackrel{d}{=} p_3 \odot X_1 + p_3 \odot X_2 + p_3 \odot X_3$, we get

$$p_n = n^{-1/\gamma_3}, \quad \gamma_3 = -\log 3/\log p_3 > 0, \quad n = 3^k, k = 1, 2, \dots .$$

And in general case,

$$p_n = n^{-1/\gamma_m}, \quad \gamma_m = -\log m/\log p_m > 0, \quad n = m^k, k = 1, 2, \dots .$$

But for $m = 4$ we obtain both $\gamma_4 = -\log 4/\log p_4$ and $\log p_4 = -1/\gamma_2 \log 4$. Hence $\gamma_4 = \gamma_2$. By induction we conclude that $\gamma_m = \gamma$ for all m and therefore

$$p_n = n^{-1/\gamma}, \quad \text{for all} \quad n \geq 2.$$

□

The question is how to extend the definition of discrete stability to contain not only random variables on $\mathbb{N}_0$, but also on the whole integers $\mathbb{Z}$. It is obvious that the sum in definition of $\tilde{X}$ does not make sense for random variables that can achieve negative values. One possibility is to take the positive and negative part of X separately and consider again the same thinning operator. We can, however, obtain a wider class of distributions if we assume a different thinning operator than in Definition 4.1.

Definition 4.7. Let $X, X_1, X_2, \dots, X_n, \dots$ denote a sequence of independent and identically distributed (i.i.d.) integer-valued random variables. Assume that for every $n \in \mathbb{N}$ there exists a constant $p_n \in (0, 1)$ such that

$$X \stackrel{d}{=} \lim_{n\to\infty} \sum_{i=1}^{n} \bar{X}_i(p_n), \quad \text{where} \quad \bar{X}_i(p_n) = \sum_{j=1}^{X_i^+} \varepsilon_j^{(i)}(p_n) - \sum_{j=1}^{X_i^-} \epsilon_j^{(i)}(p_n), \tag{4.9}$$

$\varepsilon_j^{(i)}(p_n), \epsilon_j^{(i)}(p_n)$ are i.i.d. integer-valued random variables, and X^+ and X^- are the positive and negative part of X, respectively (i.e. $X^+ = X$ if $X \geq 0$ and 0 otherwise, $X^- = -X$ if $X < 0$ and 0 otherwise). Then we say that X is *discrete stable random variable in the limit sense.*

The main difference is that we do not assume the random variables ε, ϵ to be non-negative. The definition of discrete stability is only in the limit sense, not the algebraic one where we have equivalence in distribution in (4.9) instead of the limit.

Let us denote again the probability generating function of the random variables X and $\varepsilon(p_n))$ (and also $\epsilon(p_n)$) by $\mathcal{P}(z) = \mathrm{E}[z^X]$ and $\mathcal{R}_{p_n}(z) = \mathrm{E}[z^{\varepsilon(p_n)}] = \mathrm{E}[z^{\epsilon(p_n)}]$ respectively. We denote by $\mathcal{P}_1$ the generating function of the sequence $\{a_k = \mathbb{P}(X = k), k = 1, 2, \dots\}$ and by $\mathcal{P}_2$ the generating function of the sequence $\{b_k = \mathbb{P}(X = k), k = -1, -2, \dots\}$. We denote $\mathcal{P}_0 = \mathbb{P}(X = 0)$. It is obvious that the generating function of X^+ is $\mathcal{P}_0 + \mathcal{P}_1(z)$, and the generating function of X^- is $\mathcal{P}_2(z)$. There is an equivalent definition of discrete stability in the limit sense in terms of those generating functions.

Proposition 4.8. *A random variable X is discrete stable in the limit sense if and only if for all $n \in \mathbb{N}$ there exists a constant $p_n \in (0,1)$ such that*

$$\mathcal{P}(z) = \lim_{n\to\infty} \left[\mathcal{P}_0 + \mathcal{P}_1(\mathcal{R}_{p_n}(z)) + \mathcal{P}_2\left(\mathcal{R}_{p_n}(1/z)\right)\right]^n . \tag{4.10}$$

Proof. It follows from the definition (4.9) that X is discrete stable if and only if

$$\mathcal{P}(z) = \lim_{n\to\infty} \left[\mathcal{P}_{\bar{X}}(z)\right]^n .$$

The probability generating function of $\bar{X}$ can be computed in the following way.

$$\begin{aligned}
\mathcal{P}_{\bar{X}}(z) = \mathrm{E}\left[z^{\bar{X}}\right] &= \sum_{k=-\infty}^{\infty} \mathbb{P}(X=k)\mathrm{E}\left[z^{\sum_{j=1}^{X^+}\varepsilon_j(p_n) - \sum_{j=1}^{X^-}\epsilon_j(p_n)} | X = k\right] \\
&= \sum_{k=0}^{\infty} \mathbb{P}(X=k)\left(\mathrm{E}\left[z^{\varepsilon_1(p_n)}\right]\right)^k + \sum_{k=-\infty}^{-1} \mathbb{P}(X=k)\left(\mathrm{E}\left[z^{-\epsilon_1(p_n)}\right]\right)^{-k} \\
&= \mathcal{P}_0 + \mathcal{P}_1(\mathcal{R}_{p_n}(z)) + \mathcal{P}_2(\mathcal{R}_{p_n}(1/z)).
\end{aligned}$$

Hence X is discrete stable if and only if its probability generating function satisfies the relation

$$\mathcal{P}(z) = \lim_{n\to\infty} \left[\mathcal{P}_0 + \mathcal{P}_1(\mathcal{R}_{p_n}(z)) + \mathcal{P}_2(\mathcal{R}_{p_n}(1/z))\right]^n .$$

□

It is important to note that we do not define discrete stability property in the algebraic sense as we defined it for the non-negative integer-valued random variables. This also leads to the fact that we have no condition on the thinning operator $\mathcal{R}$ similar to Theorem 4.5.

In the following Subsections we introduce some examples of commutative semigroups $\mathbb{Q}$ leading to different positive discrete stable random variables. We will also give corresponding examples of discrete stable distributions in the limit sense. The proofs of the results will be provided in Chapter 5.

4.1.1 Binomial thinning operator

Assume that the probability generating function $\mathcal{Q}$ is that of Bernoulli distribution with parameter $p \in (0,1)$, i.e. we have $\mathcal{Q}(z) = pz + (1-p)$. It is easy to verify that the semigroup $\mathbb{Q}$ generated by probability generating functions of this form is commutative, as

$$\mathcal{Q}_{p_1}\left(\mathcal{Q}_{p_2}(z)\right) = p_1 p_2 z + (1 - p_1 p_2).$$

This operator was used in Steutel and van Harn (1979) to define discrete stable distribution on $\mathbb{N}_0$ and it was showed there that it leads to a distribution with probability generating function given by

$$\mathcal{P}(z) = \exp\left\{-\lambda(1-z)^{\gamma}\right\}, \quad \gamma \in (0,1],\ \lambda > 0. \tag{4.11}$$

To obtain a generalization of this distribution on $\mathbb{Z}$ we can consider two-sided binomial thinning operator defined as $\mathcal{R}(z) = (1-p)+pqz+p(1-q)z^{-1}$, where $q \in [0,1]$. This thinning operator leads to a distribution on $\mathbb{Z}$ with probability generating function given by

$$\mathcal{P}(z) = \exp\left\{-\lambda\left(\frac{1+\beta}{2}\right)\left(1-qz-(1-q)\frac{1}{z}\right)^{\gamma} - \lambda\left(\frac{1-\beta}{2}\right)\left(1-q\frac{1}{z}-(1-q)z\right)^{\gamma}\right\},$$

with $\lambda > 0, \gamma \in (0,1], \beta \in [-1,1], q \in [0,1]$. We can see that for $\beta = 1$ and $q = 1$ the distribution reduces to positive discrete stable (4.11).

4.1.2 Thinning operator of geometric type

A generalization of the previous example can be obtained if we consider $\mathcal{Q}$ to be the probability generating function of modified geometric distribution with parameters $p \in (0,1)$ and $\kappa \in [0,1)$. Consider a function

$$\mathcal{Q}(z) = \left(\frac{(1-p)+(p-\kappa)z^m}{(1-p\kappa)-\kappa(1-p)z^m}\right)^{\frac{1}{m}}, \quad \begin{array}{l} \{0 \le \kappa < 1,\ 0 < p < 1,\ m = 1\} \\ \text{or} \\ \{0 < p < \kappa < 1, \quad m \in \mathbb{N},\ m > 1\}. \end{array} \tag{4.12}$$

Lemma 4.9. *The function $\mathcal{Q}(z)$ is a probability generating function.*

Proof. To verify that $\mathcal{Q}(z) = \sum_{n=0}^{\infty} q_n z^n$ is a probability generating function we have to show that the generating sequence $\{q_n, n = 0,1,\dots\}$ is a probability mass function, i.e. $\sum_n q_n = 1$, and $0 \le q_n \le 1$. We see that $\sum_n q_n = \mathcal{Q}(1) = 1$. We expand $\mathcal{Q}$ into a power series to obtain the generating series $\{q_n, n = 0,1,\dots\}$. We will treat the case of $m = 1$ and $m > 1$ separately.

Let first $m = 1$. Then we obtain $\mathcal{Q}(z) = \sum_{n=0}^{\infty} q_n z^n$, with

$$\begin{aligned} q_0 &= \frac{1-p}{1-p\kappa}, \\ q_n &= p\ \kappa^{n-1}\frac{(1-p)^{n-1}(1-\kappa)^2}{(1-p\kappa)^{n+1}}, \quad n \ge 1. \end{aligned}$$

We can easily verify that for $0 < \kappa < 1$ and $0 < p < 1$, $\{q_n\}$ is a probability mass function and thus $\mathcal{Q}$ is a probability generating function.

Let $m > 1$. We obtain

$$\mathcal{Q}(z) = \sum_{n=0}^{\infty} q_n z^{mn},$$

where the coefficients q_n are given as

$$q_n = \sum_{j=0}^{n}\left(\frac{1-p}{1-p\kappa}\right)^{1/m+n-j}\left(\frac{p-\kappa}{1-p}\right)^{j}\kappa^{n-j}\binom{1/m+n-j-1}{n-j}\binom{1/m}{j}, \quad n \in \mathbb{N}_0.$$

This can be reduced to

$$q_n = \kappa^n \left(\frac{1-p}{1-p\kappa}\right)^{1/m+n} \binom{1/m+n-1}{n} {}_2\mathrm{F}_1\left(\{-1/m,-n\}, 1-1/m-n, \frac{(\kappa-p)(1-p\kappa)}{(1-p)^2\kappa}\right).$$

It follows from the properties of the hypergeometric ${}_2\mathrm{F}_1$ function that $0 \le q_n \le 1$ if and only if

$$0 \le \frac{(\kappa-p)(1-p\kappa)}{(1-p)^2\kappa} \le 1.$$

This is fulfilled if and only if $0 \le p \le \kappa \le 1$. However, if $k=1$ or $p=k$ or $p=0$ we obtain a degenerate distribution. From here if follows that $\{q_n\}$ is a probability mass function if and only if $0 < p < \kappa < 1$. □

The distribution given by the probability generation function $\mathcal{Q}$ with $m=1$ is sometimes called modified geometric distribution (Phillips (1978)) or zero-modified geometric distribution (Johnson et al. (2005)). This distribution is obtained as a mixture of a degenerate distribution and geometric distribution: let U be a degenerate random variable identically equal to zero, and let V be a geometrically distributed random variable with parameter $b \in (0,1]$. Let $q \in (0,1)$ and denote $Z = qU + (1-q)V$. Then the probability generating function of the mixture Z is given as

$$\mathcal{Q}(z) = q + (1-q)\frac{bz}{1-(1-b)z}.$$

We can reparametrize this distribution, by putting

$$q = \frac{1-p}{1-p\kappa} \quad \text{and} \quad b = \frac{1-\kappa}{1-p\kappa}$$

with $p \in (0,1)$ and $\kappa \in [0,1)$. Then the probability generating function takes form (4.12) with $m=1$.

The parameter m specifies the lattice of the distribution. We will denote the distribution with probability generating function $\mathcal{Q}$ by $\mathcal{G}(p,\kappa,m)$. If $m=1$, we will write simply $\mathcal{G}(p,\kappa)$.

Lemma 4.10. *The function $\mathcal{Q}(z)$ can be decomposed as*

$$\mathcal{Q}(z) = S^{-1} \circ B_p \circ S(z), \quad \textit{where} \quad S(z) = \frac{(1-\kappa)z^m}{1-\kappa z^m}, \quad B_p(z) = pz + 1 - p. \tag{4.13}$$

Proof. The decomposition can be verified by computation, as

$$S^{-1}(y) = \left(\frac{y}{(1-\kappa)+\kappa y}\right)^{1/m}.$$

□

The function $B_p(z)$ is the probability generating function of the Bernoulli distribution. In previous Subsection we showed, that B_p generates a commutative semigroup. Using the decomposition (4.13) it is easy to see that the semigroup $\mathbb{Q}$ is commutative, as

$$\begin{aligned}\mathcal{Q}_{p_1}(\mathcal{Q}_{p_2}(z)) &= S^{-1} \circ B_{p_1} \circ S \circ S^{-1} \circ B_{p_2} \circ S(z) \\ &= S^{-1} \circ B_{p_1} \circ B_{p_2} \circ S(z)\end{aligned}$$

and we already showed that $B_{p_1} \circ B_{p_2}(z) = B_{p_1 p_2}(z)$.

If we choose $m = 1$ and $\kappa = 0$ the modified geometric distribution reduces to the Bernoulli distribution. We can modify the operator $\mathcal{Q}$ and consider two-sided thinning operator of geometric type. This can be done by considering

$$\mathcal{R}(z) = S^{-1} \circ B_p \circ S^{(2)}(z) \quad \text{where} \quad S^{(2)}(z) = qS(z) + (1-q)S(z^{-1}) \quad q \in [0,1]$$

instead of $\mathcal{Q}(z)$. We will denote two-sided modified geometric distribution by $2\mathcal{G}(p, \kappa, q, m)$. We see that $\mathcal{Q}$ is obtained from $\mathcal{R}$ by considering $q = 1$.

We will study discrete stable distributions with $\mathcal{G}$ thinning operator (of geometric type) more into details in Chapter 5. It will be shown there that this choice of thinning operator in Definition 4.1 leads to a distribution with probability generating function given by

$$\mathcal{P}(z) = \exp\left\{-\lambda\left(\frac{1-z^m}{1-\kappa z^m}\right)^\gamma\right\}, \quad \lambda > 0,\ \gamma \in (0,1],\ \kappa \in [0,1),\ m \in \mathbb{N}. \tag{4.14}$$

4.1.3 Thinning operator of Chebyshev type

Let us consider a function of the following form

$$\mathcal{Q}(z) = \frac{2\left(b + T_p\left(\frac{(1+b)z-2b}{2-(1+b)z}\right)\right)}{(1+b)\left(1 + T_p\left(\frac{(1+b)z-2b}{2-(1+b)z}\right)\right)}, \tag{4.15}$$

where $p \in (0,1)$ and $b \in (-1,1)$ and $T_p(x) = \cos(p \arccos x)$.

Remark 4.11. The function T_n for $n \in \mathbb{N}$ is called Chebyshev polynomial. It belongs to the class of orthogonal polynomials. There is an extensive literature about Chebyshev polynomials, see for example Rivlin (1974). Chebyshev polynomials are commutative, $T_n \circ T_m(x) = T_m \circ T_n(x)$; they have the nesting property, $T_n \circ T_m(x) = T_{mn}(x)$. This holds true also for $T_p(x)$ with $p \in (0,1)$, defined as $T_p(x) = \cos(p \arccos x)$. However, in this case T_p is not a polynomial any more.

The function $\mathcal{Q}(z)$ can be decomposed in the following way:

$$\mathcal{Q}(z) = R^{-1} \circ T_p\ circR(z), \quad \text{where} \quad R(z) = \frac{(1+b)z-2b}{2-(1+b)z}, R^{-1}(y) = \frac{2(b+y)}{(1+b)(1+y)}. \tag{4.16}$$

Lemma 4.12. *The function $\mathcal{Q}(z)$ is a probability generating function.*

Proof. Let us consider only the case of $b = 0$ and $p = \frac{1}{n}$, $n \in \mathbb{N}, n \geq 2$. Then we can rewrite the function $\mathcal{Q}(z)$ in the following form

$$\mathcal{Q}(z) = \frac{2\cos\left(\frac{1}{n}\arccos\frac{z}{2-z}\right)}{1+\cos\left(\frac{1}{n}\arccos\frac{z}{2-z}\right)}.$$

Using the exponential and logarithmic forms of cos and arccos functions $\cos(x) = (e^{\mathrm{i}x}+e^{-\mathrm{i}x})/2$ and $\arccos(x) = \frac{\pi}{2}+\mathrm{i}\log\left(\mathrm{i}x+\sqrt{1-x^2}\right)$ we can rewrite $\cos(\frac{1}{n}\arccos y)$ into the following form

$$\begin{aligned}\cos(\tfrac{1}{n}\arccos y) &= \tfrac{1}{2}\left(e^{\mathrm{i}\frac{\pi}{2n}-\log\left(\mathrm{i}y+\sqrt{1-y^2}\right)^{1/n}} + e^{-\mathrm{i}\frac{\pi}{2n}+\log\left(\mathrm{i}y+\sqrt{1-y^2}\right)^{1/n}}\right)\\ &= \tfrac{1}{2}e^{\mathrm{i}\frac{\pi}{2n}}\left(\mathrm{i}y+\sqrt{1-y^2}\right)^{-1/n} + \tfrac{1}{2}e^{-\mathrm{i}\frac{\pi}{2n}}\left(\mathrm{i}y+\sqrt{1-y^2}\right)^{1/n}\\ &= \tfrac{1}{2}\frac{e^{\mathrm{i}\frac{\pi}{n}}+(\mathrm{i}y+\sqrt{1-y^2})^{2/n}}{e^{\mathrm{i}\frac{\pi}{2n}}(\mathrm{i}y+\sqrt{1-y^2})^{1/n}}.\end{aligned}$$

Hence $\mathcal{Q}(z)$ simplifies into (we use substitution $y = \frac{z}{2-z}$)

$$\begin{aligned}\mathcal{Q}(z) &= 2\frac{e^{\mathrm{i}\frac{\pi}{n}}+(\mathrm{i}y+\sqrt{1-y^2})^{2/n}}{\left[e^{\mathrm{i}\frac{\pi}{2n}}+(\mathrm{i}y+\sqrt{1-y^2})^{1/n}\right]^2}\\ &= 2\frac{1+(y-\mathrm{i}\sqrt{1-y^2})^{2/n}}{\left[1+(y-\mathrm{i}\sqrt{1-y^2})^{1/n}\right]^2}\\ &= \frac{2}{1+\frac{2}{(y-\mathrm{i}\sqrt{1-y^2})^{1/n}+(y-\mathrm{i}\sqrt{1-y^2})^{-1/n}}}.\end{aligned}$$

So for $z \in (0,1]$ we have

$$\mathcal{Q}(z) = \frac{2}{1+\frac{2}{\left(\frac{z}{2-z}-2\mathrm{i}\frac{\sqrt{1-z}}{2-z}\right)^{1/n}+\left(\frac{z}{2-z}-2\mathrm{i}\frac{\sqrt{1-z}}{2-z}\right)^{-1/n}}}.$$

We have to show that $\mathcal{Q}(z)$ is a real function of z. Let $x = \frac{z}{2-z}$, $y = -2\frac{\sqrt{1-z}}{2-z}$ and $u = x+\mathrm{i}y = r(\cos\phi+\mathrm{i}\sin\phi)$. Then using Moivre's formula

$$\begin{aligned}&\left(\frac{z}{2-z}-2\mathrm{i}\frac{\sqrt{1-z}}{2-z}\right)^{1/n}+\left(\frac{z}{2-z}-2\mathrm{i}\frac{\sqrt{1-z}}{2-z}\right)^{-1/n}\\ &\qquad= r^{1/n}(\cos(\phi/n)+\mathrm{i}\sin(\phi/n))+r^{-1/n}(\cos(\phi/n)-\mathrm{i}\sin(\phi/n)).\end{aligned}$$

This number is real if and only if $r=1$. But

$$r = ||x+\mathrm{i}y|| = \sqrt{x^2+y^2} = \frac{z^2+4(1-z)}{(2-z)^2} = 1.$$

We conclude that for $z \in (0,1]$ the function $\mathcal{Q}(z)$ is real valued. Moreover $\mathcal{Q}(1) = 1$. To complete the proof we need to show that $\mathcal{Q}(z)$ is a power series with nonnegative coefficients expressing probabilities.

We denote $\mathcal{Q}$ related to the parameter p by $\mathcal{Q}_p(z)$. The inverse function of $\mathcal{Q}_p(z)$ is

$$\mathcal{Q}_p^{-1}(y) = \frac{2T_n\left(\frac{y}{2-y}\right)}{1+T_n\left(\frac{y}{2-y}\right)}.$$

This follows from the decomposition (4.16), $\mathcal{Q}_p(z) = R^{-1} \circ T_p \circ R(z)$, where $R(z) = \frac{z}{2-z}$ and from the fact that the inverse function of $T_p(x)$ is $T_n(x)$. This can be verified easily from the definition $T_p(x) = \cos(p \arccos x)$. For $n \in \mathbb{N}$ is T_n the Chebyshev polynomial.

Consider first the simple case of $n = 2$. We know that $T_2(x) = 2x^2 - 1$ (see, for example, Rivlin (1974)). Therefore

$$\mathcal{Q}_p^{-1}(y) = 1 + \frac{4}{y} - \frac{4}{y^2}.$$

We may inverse this function again to obtain

$$\mathcal{Q}_p(z) = \mathcal{Q}_{1/2}(z) = \frac{-2 + 2\sqrt{2-z}}{1-z},$$

for $z < 1$. The power series expansion is now easy to obtain

$$\mathcal{Q}_{1/2}(z) = \sum_{m=0}^{\infty} \frac{\sqrt{2}}{2^{m+1}} (-1)^m \binom{\frac{1}{2}}{m+1} {}_2F_1\left(1, \tfrac{1}{2} + m, 2 + m, \tfrac{1}{2}\right) z^m.$$

It can be verified that the coefficients

$$p_m = \frac{\sqrt{2}}{2^{m+1}} (-1)^m \binom{\frac{1}{2}}{m+1} {}_2F_1\left(1, \tfrac{1}{2} + m, 2 + m, \tfrac{1}{2}\right)$$

are all positive as $\binom{\frac{1}{2}}{m+1}$ is positive for m even and negative for m odd and the hypergeometric function ${}_2F_1(1, \frac{1}{2} + m, 2 + m, \frac{1}{2})$ is always positive for $m \geq 0$. Therefore $\mathcal{Q}_{1/2}(z)$ is a probability generating function.

Now we will show by induction that $\mathcal{Q}_p(z)$ is a probability generating function for all p of the form $p = 1/2^k$, with $k \in \mathbb{N}$. We already showed that it is true for $p = \frac{1}{2}$. Let us assume $\mathcal{Q}_p(z)$ is a probability generating function for $p = \frac{1}{2^k}$, $k \geq 1$. Because of the nesting property of T_p we have $T_{p/2} = T_p \circ T_{1/2}$, therefore we may write

$$\begin{aligned} \mathcal{Q}_{p/2}(z) &= R^{-1} \circ T_{p/2} \circ R(z) = R^{-1} \circ T_p \circ T_{1/2} \circ R(z) = \\ &= R^{-1} \circ T_p \circ R \circ R^{-1} \circ T_{1/2} \circ R(z) \\ &= \mathcal{Q}_p \circ \mathcal{Q}_{1/2}(z) \end{aligned}$$

By induction assumption $\mathcal{Q}_p(z)$ is a probability generating function, as well as $\mathcal{Q}_{1/2}(z)$. The composition of two probability generating function is a probability generating function itself, therefore we conclude that $\mathcal{Q}_{p/2}(z)$ is probability generating function.

□

We denote the probability distribution given by the probability generating function (4.15) by $\mathcal{T}(p, b)$.

Proposition 4.13. *Let* $\varepsilon \sim \mathcal{T}(p, b)$. *Then* $\mathrm{E}\varepsilon = p^2$.

Proof. We compute the expectation of ε using the property of probability generating functions as $\mathrm{E}\varepsilon = \mathcal{Q}'(1)$. By deriving $\mathcal{Q}(z)$ we obtain

$$\mathcal{Q}'(z) = \frac{2(1-b)\frac{\mathrm{d}}{\mathrm{d}z}T_p(u(z))}{(1+b)(1+T_p(u(z)))^2},$$

where

$$\begin{aligned} u(z) &= \frac{(1+b)z-2b}{2-(1+b)z}, \\ \frac{\mathrm{d}}{\mathrm{d}z}T_p(u(z)) &= \frac{\mathrm{d}}{\mathrm{d}u}T_p(u)u'(z), \\ u'(z) &= \frac{2(1+b)(1-b)}{(2-(1+b)z)^2}. \end{aligned}$$

Using the relation between Chebyshev polynomials of the first and second kind (see Erdélyi et al. (1953a)) we obtain

$$\frac{\mathrm{d}}{\mathrm{d}u}T_p(u) = pU_{p-1}(u) = p\frac{\sin(p\arccos u)}{\sin(\arccos u)}.$$

Putting all together and setting $z = 1$, $u = u(1) = 1$ we obtain

$$\mathcal{Q}'(1) = \frac{4(1-b)p^2\frac{1+b}{1-b}}{4(1+b)} = p^2.$$

□

The semigroup $\mathbb{Q}$ generated by probability generating functions of this form is commutative. From the decomposition (4.16) follows tat

$$\begin{aligned} \mathcal{Q}_{p_1}\left(Q_{p_2}(z)\right) &= R^{-1}\circ T_{p_1}\circ R\circ R^{-1}\circ T_{p_2}\circ R(z) \\ &= R^{-1}\circ T_{p_1}\circ T_{p_2}\circ R(z). \end{aligned}$$

But

$$\begin{aligned} T_{p_1}\circ T_{p_2}(x) &= \cos\left(p_1\arccos\left(\cos\left(p_2\arccos x\right)\right)\right) \\ &= \cos\left(p_1p_2\arccos x\right) \\ &= T_{p_2}\circ T_{p_1}(x). \end{aligned}$$

We will study discrete stable distributions with Chebyshev type ($\mathcal{T}$) thinning operator more into details in Chapter 5. It will be shown there that this choice of thinning operator in Definition 4.1 leads to a distribution with probability generating function given by

$$\mathcal{P}(z) = \exp\left\{-\lambda\left(\arccos\frac{(1+b)z-2b}{2-(1+b)z}\right)^{\gamma}\right\}, \quad \gamma\in(0,2],\ \lambda>0,\ b\in(-1,1). \tag{4.17}$$

4.2 On second definition of discrete stable distributions

In this Subsection we give a definition of discrete stability that generalizes the second definition of strict stability (4.2) for discrete random variables. The constant A_n in (4.2) takes form $A_n = n^{1/\alpha}$ for some $0 < \alpha \leq 2$. Hence the product $A_n X$ is generally not integer-valued and we have to find a different normalization. Compared to the normalization used in previous Subsection we need a "portlying" normalization rather than thinning, therefore we will look for distributions with expected value bigger than 1.

Definition 4.14. Let $X, X_1, X_2, \ldots, X_n, \ldots$ denote a sequence of independent and identically distributed non-negative integer-valued random variables. Assume that for every $n \in \mathbb{N}$ there exists a constant $p_n > 0$ such that

$$\hat{X}(p_n) \stackrel{d}{=} \sum_{i=1}^{n} X_i, \quad \text{where} \quad \hat{X}(p_n) = \sum_{j=1}^{X} \varepsilon_j(p_n), \tag{4.18}$$

and $\varepsilon_j(p_n)$ are i.i.d. non-negative integer-valued random variables. Then we say that X is *positive discrete stable random variable in the second sense.*

Let us denote the probability generating functions of the random variables X and $\varepsilon(p_n)$ by $\mathcal{P}(z) = \mathrm{E}[z^X]$ and $\mathcal{Q}_{p_n}(z) = \mathrm{E}[z^{\varepsilon(p_n)}]$ respectively. There is an equivalent definition of positive discrete stability in the second sense in terms of those probability generating functions.

Proposition 4.15. *A random variable X is positive discrete stable in the second sense if and only if for all $n \in \mathbb{N}$ there exists a constant $p_n > 0$ such that*

$$\mathcal{P}(\mathcal{Q}_{p_n}(z)) = \mathcal{P}^n(z). \tag{4.19}$$

Proof. It follows from the definition (4.18) that X is positive discrete stable in the second sense if and only if

$$\mathcal{P}_{\hat{X}}(z) = \mathcal{P}^n(z).$$

The probability generating function of $\hat{X}$ can be computed in the same way as in Proposition 4.2. We obtain

$$\mathcal{P}_{\hat{X}}(z) = \sum_{k=0}^{\infty} \mathbb{P}(X = k) \left(\mathrm{E}\left[z^{\varepsilon_1(p_n)}\right]\right)^k = \mathcal{P}(\mathcal{Q}_{p_n}(z)).$$

Hence X is positive discrete stable in the second sense if and only if its probability generating function satisfy the relation

$$\mathcal{P}(\mathcal{Q}_{p_n}(z)) = \mathcal{P}^n(z). \tag{4.20}$$

□

Further denote by $\mathbb{Q}$ a semigroup generated by the family of probability generating functions $\{\mathcal{Q}(z) = \mathcal{Q}_{p_n}(z), n \in \mathbb{N}\}$ with operation of superposition. We show that the semigroup $\mathbb{Q}$ must be commutative.

Theorem 4.16. *Let X be positive discrete stable random variable in the second sense. Then the semigroup $\mathbb{Q}$ must be commutative.*

Proof. Let us denote $G(z) = \log \mathcal{P}(z)$. Then (4.19) is equivalent to

$$nG(z) = G(\mathcal{Q}_{p_n}(z)), \quad n \in \mathbb{N}. \tag{4.21}$$

Let $G(z)$ be a solution of (4.21). Then for all $n \in \mathbb{N}$ it must hold

$$\mathcal{Q}_{p_n}(z) = G^{-1}\left(nG(z)\right).$$

It follows from here that for all $n_1, n_2 \in \mathbb{N}$

$$\begin{aligned}\mathcal{Q}_{p_{n_1}}\left(\mathcal{Q}_{p_{n_2}}(z)\right) &= G^{-1}\left(n_1 G\left(G^{-1}\left(n_2 G(z)\right)\right)\right)\\ &= G^{-1}\left(n_1 n_2 G(z)\right)\\ &= \mathcal{Q}_{p_{n_2}}\left(\mathcal{Q}_{p_{n_1}}(z)\right),\end{aligned}$$

which means that $\mathbb{Q}$ is commutative. □

In the following Subsections we introduce some examples of commutative semigroups $\mathbb{Q}$ leading to several possible distributions that are discrete stable in the second sense.

4.2.1 Degenerate portlying operator

Assume that the probability generating function $\mathcal{Q}(z) = z^n$, i.e. the portlying distribution is a degenerate one taking only one value n. It is obvious that the semigroup $\mathbb{Q}$ is then commutative. This choice of $\mathcal{Q}$ leads to a distribution with probability generating function $\mathcal{P}(z) = z$, i.e. a degenerate distribution localized at point 1. We are dealing with a simple summation $n = \sum_{i=1}^{n} 1$.

4.2.2 Geometric portlying operator

Let us consider now geometric distribution with parameter $p \in (0,1)$ with probability generating function

$$\mathcal{Q}(z) = \frac{pz}{1-(1-p)z}.$$

Such distribution generates a commutative semigroup $\mathbb{Q}$, as

$$\begin{aligned}\mathcal{Q}_{p_1}\left(\mathcal{Q}_{p_2}(z)\right) &= \frac{p_1 p_2 z}{1-(1-p_2)z-(1-p_1)p_2 z} = \frac{p_1 p_2 z}{1-z+p_1 p_2 z}\\ &= \mathcal{Q}_{p_2}\left(\mathcal{Q}_{p_1}(z)\right).\end{aligned}$$

Proposition 4.17. *Let X be an integer-valued random variable with probability generating function*

$$\mathcal{P}(z) = \exp\left\{-\lambda\left(1-\frac{1}{z}\right)^{\gamma}\right\}.$$

Then X is positive discrete stable in the second sense.

Proof. Let $\mathcal{Q}(z) = \frac{pz}{1-(1-p)z}$ and set p so that $p^{-\gamma} = n$. Then

$$\begin{aligned}\log \mathcal{P}(\mathcal{Q}(z)) &= -\lambda\left(1-\frac{1-(1-p)z}{pz}\right)^{\gamma} = -\lambda\left(\frac{pz-1+(1-p)z}{pz}\right)^{\gamma}\\ &= -\lambda p^{-\gamma}\left(1-\frac{1}{z}\right)^{\gamma} = n\log\mathcal{P}(z).\end{aligned}$$

Hence by Proposition 4.15 the random variable X is positive discrete stable in the second sense. □

It is important to note that the probability generating function $\mathcal{P}(z)$ defines a non-positive integer-valued random variable.

4.2.3 Portlying operator of Chebyshev type

Consider a probability generating function

$$\mathcal{Q}(z) = \frac{1}{T_n\left(\frac{1}{z}\right)}, \quad n \in \mathbb{N}, \tag{4.22}$$

where $T_n(x)$ is the Chebyshev polynomial, $T_n(x) = \cos(n \arccos x)$. Klebanov et al. (2012) showed that the function $\mathcal{Q}(z) = \mathcal{Q}_n(z)$ is indeed a probability generating function of a random variable with values in $\mathbb{N}$. The semigroup $\mathbb{Q}$ generated by the family $\{\mathcal{Q}_n(z), n \in \mathbb{N}\}$ is commutative. We have $\mathcal{Q}(z) = R^{-1} \circ S \circ R(z)$, where $R(z) = \frac{1}{z}$ and $S(x) = T_n(x)$. Hence

$$\begin{aligned}\mathcal{Q}_{n_1}\left(\mathcal{Q}_{n_2}(z)\right) &= R^{-1} \circ T_{n_1} \circ R \circ R^{-1} \circ T_{n_2} \circ R(z) = R^{-1} \circ T_{n_1} \circ T_{n_2} \circ R(z) \\ &= R^{-1} \circ T_{n_2} \circ T_{n_1} \circ R(z) = \mathcal{Q}_{n_2}\left(\mathcal{Q}_{n_1}(z)\right),\end{aligned}$$

because Chebyshev polynomials are commutative.

Theorem 4.18. *Consider the following function*

$$\mathcal{P}(z) = \left(\frac{1-\sqrt{1-z^2}}{z}\right)^M, \quad M \in \mathbb{N}. \tag{4.23}$$

Then $\mathcal{P}$ is a probability generating function of a random variable on $\mathbb{N}$. Moreover if X is an integer-valued random variable with probability generating function $\mathcal{P}$ then X is positive discrete stable in the second sense.

Proof. Let us show first that $\mathcal{P}(z)$ is a probability generating function. We will consider only the case $M = 1$. For $M > 1$ the result will follow as $\mathcal{P}(z) = \mathcal{P}_1^M(z)$, where $\mathcal{P}_1(z) = \frac{1-\sqrt{1-z^2}}{z}$, and integer power of a probability generating function is a probability generating function of a sum of i.i.d. random variables. It is obvious that $\mathcal{P}(1) = 1$. We can write $\mathcal{P}(z)$ as

$$\begin{aligned}\mathcal{P}(z) &= \frac{1}{z}\left(1-\sqrt{1-z^2}\right) = \frac{1}{z} - \sum_{k=0}^{\infty}(-1)^k \binom{\frac{1}{2}}{k} z^{2k-1} \\ &= \sum_{k=1}^{\infty}(-1)^{k-1}\binom{\frac{1}{2}}{k} z^{2k-1}.\end{aligned}$$

The coefficients of the series are all positive, because the binomial coefficient $\binom{\frac{1}{2}}{k}$ involves $(k-1)$ negative factors.

Now let us show that X is positive discrete stable in the second sense. Let $\mathcal{Q}(z)$ be as in (4.22). Then

$$\mathcal{P}(\mathcal{Q}(z)) = T_n\left(\frac{1}{z}\right) - \sqrt{T_n^2\left(\frac{1}{z}\right) - 1}.$$

We can use the explicit expression of Chebyshev polynomial to obtain

$$T_n\left(\frac{1}{z}\right) = \frac{(1+\sqrt{1-z^2})^n + (1-\sqrt{1-z^2})^n}{2z^n}$$

and

$$\sqrt{T_n^2\left(\frac{1}{z}\right) - 1} = \frac{(1+\sqrt{1-z^2})^n - (1-\sqrt{1-z^2})^n}{2z^n}.$$

From here we see that

$$\mathcal{P}(\mathcal{Q}(z)) = \mathcal{P}(z)^n.$$

Hence by Proposition 4.15 the random variable X is positive discrete stable in the second sense. □

In the proof of the theorem we showed that

$$\mathcal{P}(z) = \sum_{k=1}^{\infty} (-1)^{k-1} \binom{\frac{1}{2}}{k} z^{2k-1},$$

so the probabilities $\mathbb{P}(X = k)$ are given as $(-1)^{k-1}\binom{\frac{1}{2}}{k}$ for all odd $k > 0$ and 0 otherwise.

Remark 4.19. The probability distribution with generating function (4.23) for $M = 1$ is known (see (Feller, 1968, §XI.3)) as a distribution of the first passage time of a random walk through $+1$. Let us consider a sequence of Bernoulli trials $X_1, X_2, \dots$ with probability $p = 1/2$, i.e. $\mathbb{P}(X_i = 1) = 1 - \mathbb{P}(X_i = -1) = 1/2$ and denote $S_n = X_1 + X_2 + \cdots + X_n$, $S_0 = 0$. Then the random walk S_n passes through $+1$ for the first time at time m if

$$S_1 \le 0, \dots S_{m-1} \le 0, \quad S_m = 1.$$

The probability of this event is given by the probability generating function (4.23).

In continuous case we have a similar result. The first passage time of a Brownian motion through a level $a > 0$ has Lévy distribution, a special case of stable distribution with $\alpha = 1/2$.

The discrete stable distribution with probability generating function (4.23) with $M = 1$ can be considered a discrete analogy of Lévy distribution as is shown in the following Theorem.

Theorem 4.20. *Discrete stable random distribution in the second sense with probability generating function*

$$\mathcal{P}(z) = \frac{1-\sqrt{1-z^2}}{z}$$

belongs to the domain of normal attraction of stable distribution $\mathrm{S}\left(\frac{1}{2}, 1, 1, 0\right)$*, i.e. Lévy distribution.*

Proof. Let $X_1, X_2, \dots, X_n$ be i.i.d. positive discrete stable random variables in the second sense with probability generating function $\mathcal{P}(z)$. The characteristic function of X_1 is equal to $f(t) = \mathcal{P}\left(e^{\mathrm{i}t}\right)$. Denote

$$S_n = \frac{1}{n^2} \sum_{i=1}^{n} X_i.$$

Then the characteristic function of S_n is equal to

$$f_n(t) = f^n\left(t/n^2\right) \longrightarrow \exp\left\{-\sqrt{2}(-\mathrm{i}t)^{1/2}\right\}, \quad \text{as} \quad n \to \infty.$$

Moreover $(-\mathrm{i}t)^{1/2} = \frac{1}{\sqrt{2}}|t|^{1/2}\left(1 - \mathrm{i}\,\mathrm{sgn}(t)\right)$. □

Consider now a slightly different setting with portlying operator with probability generating function

$$\mathcal{Q}(z) = \left(\frac{1}{T_n\left(\frac{1}{z^m}\right)}\right)^{1/m}, n, m \in \mathbb{N}. \tag{4.24}$$

As was noted in Klebanov et al. (2012), $\mathcal{Q}$ is a probability generating function of a random variable with values in $m\mathbb{N}$.

Theorem 4.21. *Let X be an integer-valued random variable with probability generating function*

$$\mathcal{P}(z) = \frac{1 - \sqrt{1 - z^{2m}}}{z^m}, \quad m \in \mathbb{N}.$$

Then X is positive discrete stable in the second sense.

Proof. We have

$$\mathcal{P}(\mathcal{Q}(z)) = T_n\left(\frac{1}{z^m}\right) - \sqrt{T_n\left(\frac{1}{z^m}\right)^2 - 1},$$

and using results from the proof of Theorem 4.18,

$$\mathcal{P}(\mathcal{Q}(z)) = \frac{1 - \sqrt{1 - z^{2m}}}{z^m}.$$

□

4.3 On third definition of discrete stable distributions

In this Section we give a definition of discrete stability that generalizes the third definition of strict stability (4.3) for discrete random variables. As it turns out, this definition is a combination of the two previous definitions.

Definition 4.22. Let X, X_1 and X_2 be independent and identically distributed non-negative integer-valued random variables. Assume that for any positive numbers p_1 and p_2 there exists a positive number p such that

$$\tilde{X}(p) \stackrel{d}{=} \tilde{X}_1(p_1) + \tilde{X}_2(p_2), \quad \text{where} \quad \tilde{X}(p) = \sum_{j=1}^{X} \varepsilon_j(p) \tag{4.25}$$

and $\varepsilon_j(p)$ are i.i.d. non-negative integer-valued random variables. Then we say that X is *positive discrete stable random variable in the third sense.*

Let us denote the probability generating functions of the random variables X and $\varepsilon(p)$ by $\mathcal{P}(z) = \mathrm{E}[z^X]$ and $\mathcal{Q}_p(z) = \mathrm{E}[z^{\varepsilon(p)}]$ respectively. Let us again denote the semigroup generated by $\{\mathcal{Q}_p, p \in \Delta\}$ with operation of superposition by $\mathbb{Q}$. There is an equivalent definition of positive discrete stability in the third sense in terms of those probability generating functions, following directly from the Definition.

Proposition 4.23. *A random variable X is positive discrete stable in the third sense if and only if for any positive numbers p_1 and p_2 there exists a positive number p such that*

$$\mathcal{P}(\mathcal{Q}_p(z)) = \mathcal{P}(\mathcal{Q}_{p_1}(z))\mathcal{P}(\mathcal{Q}_{p_2}(z)). \tag{4.26}$$

We can show that every random variable positive discrete stable in the first sense is also positive discrete stable in the third sense.

Theorem 4.24. *Let X be positive discrete stable in the first sense. Then X is positive discrete stable in the third sense. Moreover* (4.25) *holds with*

$$p^\gamma = p_1^\gamma + p_2^\gamma.$$

Proof. Let X be positive discrete stable in the first sense, and let $X_1, X_2, \dots$ be independent copies of X. Then the semigroup $\mathbb{Q}$ is commutative, $p \in \Delta = (0,1)$ and for any $n \geq 2$ there exists a constant $p_n \in (0,1)$ such that

$$X \stackrel{d}{=} \sum_{i=1}^{n} p_n \odot X_i.$$

From Theorem 4.6 we know that $p_n = n^{-1/\gamma}$. Let $p_1, p_2 \in \Delta$. Then for all $n_1, n_2 \geq 2$

$$p_1 \odot X_1 + p_2 \odot X_2 \stackrel{d}{=} \sum_{i=1}^{n_1} p_1 p_{n_1} \odot X_i + \sum_{j=n_1+1}^{n_1+n_2} p_1 p_{n_2} \odot X_j.$$

If p_1^γ, p_2^γ are rational, then we can find n_1, n_2, p such that

$$p_1 p_{n_1} = p p_{n_1+n_2},$$
$$p_2 p_{n_2} = p p_{n_1+n_2},$$

or equivalently

$$p_1^\gamma = p^\gamma \frac{n_1}{n_1+n_2},$$
$$p_2^\gamma = p^\gamma \frac{n_2}{n_1+n_2}.$$

But then, with $n = n_1 + n_2$

$$p_1 \odot X_1 + p_2 \odot X_2 = \sum_{i=1}^{n} p p_n \odot X_i = p \odot X.$$

Moreover p_1, p_2, p satisfy the relationship $p_1^\gamma + p_2^\gamma = p^\gamma$. By continuity argument it follows that (4.25) hold for any choice of p_1, p_2 with p such that $p_1^\gamma + p_2^\gamma = p^\gamma$. □

Under some additional conditions we may show that the opposite statement holds true as well.

Theorem 4.25. *Let X be positive discrete stable in the third sense and assume that the semigroup $\mathbb{Q}$ is commutative, $\Delta = (0,1)$ and that there exists a constant $\gamma > 0$ such that*

$$p^\gamma = p_1^\gamma + p_2^\gamma.$$

Then X is positive discrete stable in the first sense.

Proof. We may show this by induction. Because X is positive discrete stable in the third sense, we have for $p_1 = p_2 = 2^{-1/\gamma}$ that

$$X \stackrel{d}{=} \tilde{X}_1(p_2) + \tilde{X}_2(p_2).$$

Let $n \geq 2$ and let us assume that

$$X \stackrel{d}{=} \sum_{i=1}^{n} \tilde{X}_i(p_n), \quad \text{with} \quad p_n = n^{-1/\gamma}.$$

Denote $Y = \sum_{i=1}^{n} \tilde{X}_i(p_n)$ and let $p = \left(\frac{n}{n+1}\right)^{1/\gamma}$. Because X is positive discrete stable in the third sense, $Y \stackrel{d}{=} X$ and $p^\gamma + p_{n+1}^\gamma = 1$, we have

$$X \stackrel{d}{=} \tilde{Y}(p) + \tilde{X}_{n+1}(p_{n+1}).$$

The probability generating function of the right-hand side is

$$\begin{aligned}
\mathcal{P}_Y(\mathcal{Q}_p(z))\mathcal{P}(\mathcal{Q}_{p_{n+1}}(z)) &= \mathcal{P}^n(\mathcal{Q}_{p_n}(\mathcal{Q}_p(z)))\mathcal{P}(\mathcal{Q}_{p_{n+1}}(z)) \\
&= \mathcal{P}^n(\mathcal{Q}_{p_n p}(z))\mathcal{P}(\mathcal{Q}_{p_{n+1}}(z)) \\
&= \mathcal{P}^{n+1}(\mathcal{Q}_{p_{n+1}}(z)),
\end{aligned}$$

because $\mathbb{Q}$ is commutative and $pp_n = p_{n+1}$. Therefore

$$X \stackrel{d}{=} \sum_{i=1}^{n+1} \tilde{X}_i(p_{n+1}).$$

□

Example 4.26 (Binomial thinning operator)**.** Let us consider the case of the binomial thinning operator with probability generating function $\mathcal{Q}(z) = (1-p) + pz$. Then a random variable X with probability generating function $\mathcal{P}(z) = \exp\{-\lambda(1-z)^\gamma\}$ is positive discrete stable in the third sense, as (4.26) holds if

$$p^\gamma = p_1^\gamma + p_2^\gamma.$$

Example 4.27 (Modified geometric thinning operator)**.** We can verify that the positive discrete stable random variable in the first sense with modified geometric thinning operator is also positive discrete stable in the third sense. Let X be a positive discrete random variable in the first sense with probability generating function $\mathcal{P}(z) = \exp\left\{-\lambda\left(\frac{1-z}{1-\kappa z}\right)^{\gamma}\right\}$. Then

$$\mathcal{P}\left(\mathcal{Q}_p(z)\right) = \exp\left\{-\lambda p^{\gamma}\left(\frac{1-z}{1-\kappa z}\right)^{\gamma}\right\}.$$

Thus (4.26) holds if

$$p^{\gamma} = p_1^{\gamma} + p_2^{\gamma}.$$

Example 4.28 (Chebyshev thinning operator)**.** In the same manner we see that a positive discrete stable random variable in the first sense with Chebyshev thinning operator X is positive discrete stable in the third sense. Let $\mathcal{P}$ be as in (4.17) and $\mathcal{Q}$ as in (4.15). We have

$$\mathcal{P}(\mathcal{Q}_p(z)) = [\mathcal{P}(z)]^{p^{\gamma}}.$$

Therefore again (4.26) holds if

$$p^{\gamma} = p_1^{\gamma} + p_2^{\gamma}.$$

Example 4.29 (Chebyshev portlying operator)**.** Now let's look at an example with Chebyshev portlying operator with probability generating function $\mathcal{Q}_n(z) = 1/T_n(1/z)$. Then a random variable X with probability generating function $\mathcal{P}(z) = \left(1-\sqrt{1-z^2}\right)/z$ is positive discrete stable in the third sense, as (4.26) holds if

$$n = n_1 + n_2.$$

5. Properties of discrete stable distributions

In previous Chapter we introduced several variants of discrete stability. Distributions, that are discrete stable in the first sense, form the widest and most interesting class of distributions, and we will study them more into details in this Chapter. We will focus mainly on the distributions with thinning operator of geometric type, but we will give some results on distributions with thinning operator of Chebyshev type as well.

5.1 Positive discrete stable random variables with $\mathcal{G}$ thinning operator

To remind the definition, a non-negative integer-valued random variable X is said to be positive discrete stable in the first sense, if

$$X \stackrel{d}{=} \sum_{j=1}^{n} \tilde{X}_j, \quad \text{where} \quad \tilde{X}_j = \sum_{i=1}^{X_j} \varepsilon_i^{(j)}, \tag{5.1}$$

where $X_1, X_2, \dots$ are independent copies of X and $\varepsilon_i^{(j)}$ are i.i.d. non-negative integer-valued random variables. Throughout this Section we will assume that the random variables $\varepsilon_i^{(j)}$ come from modified geometric distribution $\mathcal{G}(p, \kappa, m)$ with probability generating function $\mathcal{Q}$ of the form

$$\mathcal{Q}(z) = \left(\frac{(1-p) + (p-\kappa)z^m}{(1-p\kappa) - \kappa(1-p)z^m} \right)^{\frac{1}{m}}, \quad \begin{array}{l} \{0 \le \kappa < 1,\ 0 < p < 1,\ m = 1\} \\ \text{or} \\ \{0 < p < \kappa < 1, \quad m \in \mathbb{N}, m > 1\}. \end{array} \tag{5.2}$$

We remind that $\mathcal{Q}(z)$ can be decomposed as $\mathcal{Q}(z) = S^{-1} \circ B_p \circ S(z)$, where $B_p(z) = pz + (1-p)$ and

$$S(z) = \frac{(1-\kappa)z^m}{1-\kappa z^m}, \quad S^{-1}(y) = \left(\frac{y}{(1-\kappa) + \kappa y} \right)^{\frac{1}{m}}.$$

Theorem 5.1. *A non-negative integer-valued random variable X is positive discrete stable with $\mathcal{G}$ thinning operator if and only if $\mathcal{Q}$ takes form* (5.2) *and the probability generating function $\mathcal{P}(z) = \mathrm{E}z^X$ is given as*

$$\mathcal{P}(z) = \exp\left\{ -\lambda \left(\frac{1-z^m}{1-\kappa z^m} \right)^{\gamma} \right\} \quad \text{with} \quad \gamma \in (0,1],\ \lambda > 0,\ \kappa \in [0,1),\ m \in \mathbb{N}. \tag{5.3}$$

Proof. Let $h(z) = \log \mathcal{P}(z)$. From Proposition 4.2 it follows that X is positive discrete stable if and only if $h(z) = nh(\mathcal{Q}(z))$ for all n. Set

$$h(z) = -\lambda \left(\frac{1-z^m}{1-\kappa z^m} \right)^{\gamma}$$

and select γ such that $1/p^\gamma = n$. We see that

$$\frac{1-z^m}{1-\kappa z^m} = 1 - \frac{(1-\kappa)z^m}{1-\kappa z^m} = 1 - S(z).$$

Therefore, using the decomposition of $\mathcal{Q}(z)$,

$$\begin{aligned} nh(\mathcal{Q}(z)) &= -\lambda n\left(1 - S(\mathcal{Q}(z))\right)^\gamma = -\lambda n\left(1 - B_p(S(z))\right)^\gamma \\ &= -\lambda n\left(p - pS(z)\right)^\gamma = -\lambda n p^\gamma (1 - S(z))^\gamma \\ &= -\lambda \left(\frac{1-z^m}{1-\kappa z^m}\right)^\gamma = h(z). \end{aligned}$$

□

The parameter m determines the size of the lattice of the distribution. We will denote positive discrete stable random variable (and associated distribution) by $\mathrm{PDS}^m(\gamma,\lambda,\kappa)$. In the case when m is omitted we will understand that $m=1$. If moreover κ is omitted, we will understand that $\kappa = 0$, in which case the discrete stable distribution reduces to the discrete stable distribution as it was introduced in Steutel and van Harn (1979). In Figure A.1 the probabilities of $\mathrm{PDS}(\gamma,\lambda,\kappa)$ random variables are shown for different values of parameters. The probabilities were obtained using the classical inverse Fourier transform theorem (see, for example, Lachout (2004)) and the fast Fourier transform algorithm.

The characteristic function is given as

$$f(t) = \exp\left\{-\lambda\left(\frac{1-e^{\mathrm{i}tm}}{1-\kappa e^{\mathrm{i}tm}}\right)^\gamma\right\}.$$

The case of $\gamma = 1$ is a special one as it leads to a distribution with finite variance and exponential tails. As a simple corollary we obtain Poisson distribution by taking $\kappa = 0$ and $\gamma = 1$.

5.1.1 Characterizations

In this Subsection we present several characterizations of positive discrete stable random variables.

Theorem 5.2. *Let $\gamma \in (0,1)$ be a given parameter. Let $X, X_1, X_2, \dots$ be i.i.d. non-negative integer-valued random variables and Y be a non-negative integer-valued random variable, independent of the sequence $X_1, X_2, \dots$. Then X is positive discrete stable* $\mathrm{PDS}(\gamma,\lambda)$ *random variable if and only if*

$$X \overset{d}{=} \sum_{j=1}^{Y} Y^{-1/\gamma} \odot X_j, \quad \textit{where} \quad p \odot X = \sum_{i=1}^{X} \varepsilon_i(p) \tag{5.4}$$

and $\varepsilon_i(p)$ are i.i.d. Bernoulli random variables with probability generating function $\mathcal{Q}_p(z) = 1 - p + pz$.

Proof. First let us show that if X is PDS(γ, λ) then it has the representation (5.4). Let $\mathcal{P}(z)$ be the probability generating function of X. The probability generating function of the right-hand side of (5.4) can be computed in the following way.

$$\begin{aligned}\mathrm{E}\left[z^{\sum_{j=1}^{Y} Y^{-1/\gamma}\odot X_j}\right] &= \mathrm{E}\left[\mathrm{E}\left[z^{\sum_{j=1}^{Y} Y^{-1/\gamma}\odot X_j}|Y\right]\right] = \mathrm{E}\left[\mathcal{P}_X^Y\left(\mathcal{Q}_{Y^{-1/\gamma}}(z)\right)\right]\\ &= \mathrm{E}\left[\exp\left\{-\lambda Y\left(1-\mathcal{Q}_{Y^{-1/\gamma}}(z)\right)^{\gamma}\right\}\right] = \mathrm{E}\left[\exp\left\{-\lambda Y Y^{-1}\left(1-z\right)^{\gamma}\right\}\right]\\ &= \exp\left\{-\lambda\left(1-z\right)^{\gamma}\right\} = \mathcal{P}(z).\end{aligned}$$

The proof of the inverse statement is more complicated and relies on the method if intensively monotone operators. The condition (5.4) can be translated into the form of probability generating functions as

$$\mathcal{P}(z) = \sum_{k=0}^{\infty} \mathbb{P}(Y=k) \prod_{j=1}^{k} \mathcal{P}\left(\mathcal{Q}_{k^{-1/\gamma}}(z)\right). \tag{5.5}$$

Put $G(z) = \log \mathcal{P}(z)$ and $h(z) = G(z)/(1-z)^{\gamma}$. Then we can rewrite (5.5) as

$$\begin{aligned}h(z) &= (1-z)^{-\gamma} \sum_{k=0}^{\infty} \mathbb{P}(Y=k) \sum_{j=1}^{k} \left(1-\mathcal{Q}_{k^{-1/\gamma}}(z)\right)^{\gamma} h\left(\mathcal{Q}_{k^{-1/\gamma}}(z)\right)\\ &= (1-z)^{-\gamma} \sum_{k=0}^{\infty} \mathbb{P}(Y=k) \left(1-z\right)^{\gamma} h\left(\mathcal{Q}_{k^{-1/\gamma}}(z)\right)\\ &= \sum_{k=0}^{\infty} \mathbb{P}(Y=k) h\left(\mathcal{Q}_{k^{-1/\gamma}}(z)\right).\end{aligned} \tag{5.6}$$

Let A be an operator acting on $g \in C[0,1]$ such that

$$(Ag)(z) = \begin{cases} \sum_{k=0}^{\infty} \mathbb{P}(Y=k) g\left(\mathcal{Q}_{k^{-1/\gamma}}(z)\right), & z<1\\ g(0), & z=1.\end{cases}$$

We can verify that A is an intensively monotone operator (see Kakosyan et al. (1984)) and that $Ag \in C[0,1]$. It is clear that $Aa = a$ for all constant functions a. It follows from (Kakosyan et al., 1984, Theorem 1.1.2) that the only solution of (5.6) is identically equal to a constant. Hence $h(z) = -\lambda$ and

$$\mathcal{P}(z) = \exp\left\{-\lambda(1-z)^{\gamma}\right\}.$$

□

Theorem 5.3. *Let $\gamma, \gamma' \in (0,1]$ and assume that $\gamma' \le \gamma$. Let S_γ be a γ-stable random variable with Laplace transform $\exp\{-u^{\gamma}\}$. Then*

$$\mathrm{PDS}(\gamma', \lambda, \kappa) \stackrel{d}{=} \mathrm{PDS}\left(\gamma'/\gamma, \lambda^{1/\gamma}\mathrm{S}_\gamma, \kappa\right).$$

Proof. The characteristic function of the right-hand side can be computed as

$$\begin{aligned}
\mathrm{E}\left[\exp\left\{it\mathrm{PDS}\left(\gamma'/\gamma, \lambda^{1/\gamma}\mathrm{S}_\gamma, \kappa\right)\right\}\right] &= \mathrm{E}\left[\exp\left\{-\lambda^{1/\gamma}\mathrm{S}_\gamma\left(\frac{1-e^{\mathrm{i}t}}{1-\kappa e^{\mathrm{i}t}}\right)^{\gamma'/\gamma}\right\}\right] \\
&= \exp\left\{-\lambda\left(\frac{1-e^{\mathrm{i}t}}{1-\kappa e^{\mathrm{i}t}}\right)^{\gamma'}\right\} \\
&= \mathrm{E}\left[\exp\left\{it\mathrm{PDS}\left(\gamma', \lambda, \kappa\right)\right\}\right].
\end{aligned}$$

□

The following Corollary can be applied for simulations of positive discrete stable random variables.

Corollary 5.4. *Let $Y, Y_1, Y_2, \ldots$ be a sequence of i.i.d. random variables with geometric distribution, $\mathbb{P}(Y = n) = (1-\kappa)\kappa^{n-1}, n \geq 1$. Let N be a random variable, independent of the sequence $Y_1, Y_2, \ldots,$ with Poisson distribution with random intensity $\lambda^{-1/\gamma}\mathrm{S}_\gamma$, where S_γ is a γ-stable random variable with Laplace transform $\exp\{-u^\gamma\}$. Then*

$$\sum_{j=1}^{N} Y_j$$

has the same distribution as a positive discrete stable random variable $\mathrm{PDS}(\gamma, \lambda, \kappa)$.

Proof. Let $X = \sum_{j=1}^{N} Y_j$. Then X is a compound Poisson random variable with random intensity $\lambda^{1/\gamma}\mathrm{S}_\gamma$ and jumps $Y_1, Y_2, \ldots$ with characteristic function

$$g(t) = \frac{(1-\kappa)e^{\mathrm{i}t}}{1-\kappa e^{\mathrm{i}t}}.$$

The characteristic function of a compound Poisson random variable with intensity τ and characteristic function of jumps $h(t)$ is $\exp\{-\tau(1-h(t))\}$. Therefore X is in fact $\mathrm{PDS}(1, \lambda^{1/\gamma}\mathrm{S}_\gamma, \kappa)$. We thus obtain the result from the previous Theorem 5.3 with $\gamma' = \gamma$. □

5.1.2 Moments

Theorem 5.5. *Let X be $\mathrm{PDS}(\gamma, \lambda, \kappa)$ random variable with $\gamma = 1$ and $\kappa > 0$. Then the n-th factorial moment can be computed using the following formula*

$$\mathrm{E}\left[(X)_n\right] = \frac{\kappa^n}{(1-\kappa)^n} n! \sum_{s=0}^{n-1} \frac{1}{(s+1)!}\binom{n-1}{s}\frac{\lambda^{s+1}}{\kappa^{s+1}}. \tag{5.7}$$

Proof. Let $\mathcal{P}(z)$ be the probability generating function of X. The n-th factorial moment of discrete random variable can be computed as the value of the n-th derivative of the probability generating function at point 1, i.e.

$$\mathrm{E}\left[(X)_n\right] = \frac{\mathrm{d}^n}{\mathrm{d}z^n}\mathcal{P}(z)\Big|_{z=1}.$$

Since $\mathcal{P}(z) = \exp\{g(z)\}$, with

$$g(z) = -\lambda\left(1 - (1-\kappa)\frac{z}{1-\kappa z}\right),$$

we compute the n-th derivative using the Bruno's formula (Faa di Bruno (1857))

$$\frac{\mathrm{d}^n}{\mathrm{d}z^n}\mathcal{P}(z)\Big|_{z=1} = \sum_{k=1}^{n} \mathcal{P}(1) B_{n,k}(g'(1), g''(1), \dots, g^{(n-k+1)}(1)),$$

where $B_{n,k}(x_1, \dots, x_{n-k+1})$ is the Bell's polynomial,
(5.8)

$$B_{n,k}(x_1,\dots,x_{n-k+1}) = \sum_{i_1,\dots,i_{n-k+1}} \frac{n!}{i_1! i_2! \dots i_{n-k+1}!}\left(\frac{x_1}{1!}\right)^{i_1}\left(\frac{x_2}{2!}\right)^{i_1}\cdots\left(\frac{x_{n-k+1}}{(n-k+1)!}\right)^{i_{n-k+1}},$$

where we sum over all possible combinations such that $i_1 + 2i_2 \cdots + (n-k+1)i_{n-k+1} = n$ and $i_1 + i_2 \cdots + i_{n-k+1} = k$. By differentiating the function $g(z)$ we obtain

$$g^{(i)}(1) = i!\lambda\frac{\kappa^{i-1}}{(1-\kappa)^i}.$$

Plugging that into the Bell's polynomial we obtain

$$\begin{aligned}
B_{n,k}\left(g'(1), g''(1), \dots, g^{(n-k+1)}(1)\right) &= \sum_{i_1,\dots,i_{n-k+1}} \frac{n!}{i_1! i_2! \dots i_{n-k+1}!}\prod_{j=1}^{n-k+1}\left(\frac{g^{(j)}(1)}{j!}\right)^{i_j} \\
&= \sum_{i_1,\dots,i_{n-k+1}} \frac{n!}{i_1! i_2! \dots i_{n-k+1}!}\prod_{j=1}^{n-k+1}\left(\frac{\lambda\kappa^{j-1}}{(1-\kappa)^j}\right)^{i_j} \\
&= \sum_{i_1,\dots,i_{n-k+1}} \frac{n!}{i_1! i_2! \dots i_{n-k+1}!}\frac{\lambda^k\kappa^n}{\kappa^k(1-\kappa)^n} \\
&= \frac{\lambda^k\kappa^n}{\kappa^k(1-\kappa)^n} B_{n,k}(1!, 2!, \dots, (n-k+1)!) \\
&= \frac{\lambda^k\kappa^n}{\kappa^k(1-\kappa)^n}\binom{n}{k}\binom{n-1}{k-1}(n-k)!.
\end{aligned}$$

Hence the n-th factorial moment is

$$\begin{aligned}
\mathrm{E}\left[(X)_n\right] &= \sum_{k=1}^{n}\frac{\lambda^k\kappa^n}{\kappa^k(1-\kappa)^n}\binom{n}{k}\binom{n-1}{k-1}(n-k)! \\
&= \frac{\kappa^n}{(1-\kappa)^n}\sum_{k=1}^{n}\frac{\lambda^k}{\kappa^k}\frac{n!}{k!}\binom{n-1}{k-1}.
\end{aligned}$$

The result follows from here by setting $s = k-1$. □

5.1.3 Probabilities

In the next Theorem we show connection between the probabilities of a positive discrete stable random variable and moments of a tempered stable random variable.

Theorem 5.6. *Let X be a* PDS(γ, λ) *random variable with* $\gamma < 1$. *Let Y be a tempered stable random variable with characteristic function* $f_Y(t) = \exp\{-(\lambda^{1/\gamma} - \mathrm{i}t)^\gamma + \lambda\}$. *Then we can write the probabilities* $\mathbb{P}(X = k)$ *as*

$$\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^{k/\gamma}}{k!} \mathrm{E} Y^k.$$

Before we proceed to the proof of the Theorem, we state a simple Lemma.

Lemma 5.7. *Let S_γ be γ-stable random variable with Laplace transform* $L(u) = \mathrm{E}e^{-uS_\gamma} = \exp\{-u^\gamma\}$ *and density function $p(x)$. Let $\theta > 0$. Let Y be a random variable with density function*

$$p_Y(x) = e^{-\theta x} p(x) / L(\theta).$$

Then Y is a tempered stable random variable with characteristic function

$$f(t) = \exp\{-(\theta - \mathrm{i}t)^\gamma + \theta^\gamma\}.$$

Proof. We may compute the characteristic function of Y as follows:

$$\begin{aligned} f_Y(t) = \mathrm{E}e^{\mathrm{i}tY} &= \int_0^\infty e^{\mathrm{i}tx} p_Y(x) \mathrm{d}x = \int_0^\infty e^{\mathrm{i}tx} e^{-\theta x} p(x)/L(\theta) \mathrm{d}x \\ &= e^{\theta^\gamma} \int_0^\infty \exp\{-(\theta - \mathrm{i}t)x\}\, p(x) \mathrm{d}x \\ &= e^{\theta^\gamma} L(\theta - \mathrm{i}t) = \exp\left\{-(\theta - \mathrm{i}t)^\gamma + \theta^\gamma\right\}. \end{aligned}$$

□

Now we can prove the Theorem.

Proof of Theorem 5.6. It follows from Theorem 5.3 that a positive discrete stable random variable PDS(γ, λ) is a Poisson random variable with random intensity $\lambda^{1/\gamma} S_\gamma$, where S_γ is a γ-stable random variable with Laplace transform $L(u) = \exp\{-u^\gamma\}$ and density function $p(x)$. Therefore the probabilities $\mathbb{P}(X = k)$ can be computed as

$$\begin{aligned} \mathbb{P}(X = k) &= \int_0^\infty e^{-\lambda^{1/\gamma} s} \frac{(\lambda^{1/\gamma} s)^k}{k!} p(s) \mathrm{d}s \\ &= \frac{\lambda^{k/\gamma}}{k!} L(\lambda^{1/\gamma}) \int_0^\infty s^k e^{-\lambda^{1/\gamma} s} p(s)/L(\lambda^{1/\gamma}) \mathrm{d}s. \end{aligned}$$

But $e^{-\lambda^{1/\gamma} s} p(s)/L(\lambda^{1/\gamma})$ is a density function of a tempered stable random variable Y with characteristic function $f(t) = \exp\{-(\lambda^{1/\gamma} - \mathrm{i}t)^\gamma + \lambda\}$. Therefore

$$\begin{aligned} \mathbb{P}(X = k) &= \frac{\lambda^{k/\gamma}}{k!} L(\lambda^{1/\gamma}) \int_0^\infty s^k p_Y(s) \mathrm{d}s \\ &= \frac{\lambda^{k/\gamma}}{k!} L(\lambda^{1/\gamma}) \mathrm{E} Y^k. \end{aligned}$$

□

Theorem 5.8. *Let X be a* PDS$(\gamma, \lambda, \kappa)$ *random variable with $\gamma = 1$ and $\kappa > 0$. Then the probability $\mathbb{P}(X = m)$ for $m \geq 1$ can be computed using the following formula*

$$\mathbb{P}\left(X = m\right) = e^{-\lambda} \sum_{s=0}^{m-1} \frac{\lambda^{s+1}}{(s+1)!} \binom{m-1}{s} \kappa^{m-s-1} (1-\kappa)^{s+1}. \tag{5.9}$$

Proof. We compute the probabilities by expanding the probability generating function into power series.

$$\begin{aligned}
\mathcal{P}(z) &= \exp\left\{-\lambda\left(1 - (1-\kappa)\frac{z}{1-\kappa z}\right)\right\} \\
&= e^{-\lambda} + e^{-\lambda} \sum_{n=1}^{\infty} \frac{\lambda^n}{n!} (1-\kappa)^n \frac{z^n}{(1-\kappa z)^n} \\
&= e^{-\lambda} + e^{-\lambda} \sum_{n=1}^{\infty} \sum_{j=0}^{\infty} \frac{\lambda^n}{n!} (1-\kappa)^n \kappa^j \binom{n+j-1}{j} z^{n+j} \\
&= e^{-\lambda} + e^{-\lambda} \sum_{n=1}^{\infty} \sum_{m=n}^{\infty} \frac{\lambda^n}{n!} (1-\kappa)^n \kappa^{m-n} \binom{m-1}{m-n} z^m \\
&= e^{-\lambda} + e^{-\lambda} \sum_{m=1}^{\infty} \sum_{n=1}^{m} \frac{\lambda^n}{n!} (1-\kappa)^n \kappa^{m-n} \binom{m-1}{n-1} z^m \\
&= e^{-\lambda} + e^{-\lambda} \sum_{m=1}^{\infty} \sum_{s=0}^{m-1} \frac{\lambda^{s+1}}{(s+1)!} (1-\kappa)^{s+1} \kappa^{m-s-1} \binom{m-1}{s} z^m.
\end{aligned}$$

The probabilities $\mathbb{P}(X = m)$ are obtained from this results as the coefficients of the probability generating function by z^m, as $\mathcal{P}(z) = \sum_{m=0}^{\infty} \mathbb{P}(X = m) z^m$. □

Corollary 5.9. *Let X be* PDS$(\gamma, \lambda, \kappa)$ *random variable with $\gamma = 1$ and $\kappa > 0$. Then the probability $\mathbb{P}(X = m)$ for $m \geq 1$ can be expressed in the following ways*

$$\mathbb{P}\left(X = m\right) = e^{-\lambda} \lambda (1-\kappa) \kappa^{m-1} \, {}_1F_1\left(1-m, 2, \frac{\beta-1}{\beta}\lambda\right)$$

and

$$\mathbb{P}\left(X = m\right) = e^{-\lambda} \lambda (1-\kappa) \kappa^{m-1} \frac{1}{m} L_{m-1}^{(1)}\left(\frac{\beta-1}{\beta}\lambda\right),$$

where ${}_1F_1(a, b, z)$ is the Kummer confluent hypergeometric function and $L_n^{(\alpha)}(z)$ is the generalized Laguerre polynomial.

Proof. The first assertion follows directly from (5.9). The second assertion follows from the relation between Laguerre polynomial and Kummer confluent hypergeometric function (see for example (Erdélyi et al., 1953b, pp. 268)), stating that

$$L_n^{(\alpha)}(z) = \binom{n+\alpha}{n} {}_1F_1(-n, \alpha+1, z).$$

□

5.1.4 Continuous analogies

Let us consider a random variable $X^a = aX$, with $X \sim \mathrm{PDS}(\gamma, \lambda, \kappa)$ and $a > 0$. Then X^a takes values in $a\mathbb{N}_0 = \{0, a, 2a, \cdots\}$. We study the limit behaviour of X^a as $a \to 0$ with $\kappa \to 1$.

Theorem 5.10. *Let X be a positive discrete stable random variable with parameters γ, λ and κ and let $X^a = aX$ with $a > 0$. Let $\kappa = 1 - ac$. Then*

$$f^a(t) = \exp\left\{-\lambda\left(\frac{1-e^{\mathrm{i}at}}{1-\kappa e^{\mathrm{i}at}}\right)^{\gamma}\right\} \longrightarrow \varphi(t) = \exp\left\{-\lambda\left(\frac{-\mathrm{i}t}{c-\mathrm{i}t}\right)^{\gamma}\right\}, \quad \textit{as } a \to 0.$$

Proof. The limit characteristic function can be computed in a straightforward way. We have

$$\frac{1-e^{\mathrm{i}at}}{1-\kappa e^{\mathrm{i}at}} = \frac{1-e^{\mathrm{i}at}}{1-e^{\mathrm{i}at}+ace^{\mathrm{i}at}} \approx \frac{-\mathrm{i}at}{-\mathrm{i}at+ace^{\mathrm{i}at}}, \quad \text{as} \quad a \to 0.$$

Hence we have

$$\varphi(t) = \lim_{a\to 0} \exp\left\{-\lambda\left(\frac{-\mathrm{i}at}{-\mathrm{i}at+ace^{\mathrm{i}at}}\right)^{\gamma}\right\} = \exp\left\{-\lambda\left(\frac{-\mathrm{i}t}{-\mathrm{i}t+c}\right)^{\gamma}\right\}.$$

□

Next we show that discrete stable distribution on $\mathbb{N}_0$ can be considered a discrete analogy of stable distribution with index of stability $\alpha = \gamma$ and skewness parameter $\beta = 1$.

Theorem 5.11. *Let X be a positive discrete stable random variable with parameters γ, λ and κ and let $X^a = aX$ with $a > 0$. Let $\lambda = b/a^{\gamma}$. Then*

$$f^a(t) = \exp\left\{-\lambda\left(\frac{1-e^{\mathrm{i}at}}{1-\kappa e^{\mathrm{i}at}}\right)^{\gamma}\right\}$$

$$\longrightarrow \varphi(t) = \exp\left\{-\sigma|t|^{\gamma}\left(1-\mathrm{i}\,\mathrm{sign}(t)\tan\left(\frac{\pi\gamma}{2}\right)\right)\right\}, \quad \textit{as } a \to 0,$$

where $\sigma = \frac{b}{(1-\kappa)^{\gamma}}\cos\left(\frac{\pi\gamma}{2}\right)$.

Proof. We have

$$\frac{1-e^{\mathrm{i}at}}{1-\kappa e^{\mathrm{i}at}} = \frac{1-e^{\mathrm{i}at}}{1-\kappa+\kappa\left(1-e^{\mathrm{i}at}\right)} \approx \frac{-\mathrm{i}at}{(1-\kappa)-\kappa\mathrm{i}at} \quad \text{as} \quad a \to 0.$$

Hence

$$-\lambda\left(\frac{1-e^{\mathrm{i}at}}{1-\kappa e^{\mathrm{i}at}}\right)^{\gamma} \approx -\frac{b}{a^{\gamma}}\left(\frac{-\mathrm{i}at}{(1-\kappa)-\kappa\mathrm{i}at}\right)^{\gamma} \quad \text{as} \quad a \to 0$$

$$\to -\frac{b}{(1-\kappa)^{\gamma}}(-\mathrm{i}t)^{\gamma} \quad \text{as} \quad a \to 0.$$

Finally we notice that

$$(-\mathrm{i}t)^{\gamma} = |t|^{\gamma}(-\mathrm{i}\,\mathrm{sign}(t))^{\gamma} = |t|^{\gamma}\cos(\pi\gamma/2)\left(1-\mathrm{i}\,\mathrm{sign}(t)\tan(\pi\gamma/2)\right).$$

□

5.1.5 Asymptotic behaviour

In this Subsection we show that the tails of discrete stable PDS$(\gamma, \lambda, \kappa)$ distribution are heavy with tail index γ.

Proposition 5.12. *The discrete stable distribution* PDS$(\gamma, \lambda, \kappa)$ *belongs to the domain of normal attraction of α-stable distribution with characteristic function*

$$g(t) = \exp\left\{-\frac{\lambda}{(1-\kappa)^\gamma}\cos(\pi\gamma/2)\,|t|^\gamma\left(1 - \mathrm{i}\,\mathrm{sign}(t)\tan\left(\frac{\pi\gamma}{2}\right)\right)\right\}.$$

Proof. Let $X_1, X_2, \ldots, X_n$ be i.i.d. PDS$(\gamma, \lambda, \kappa)$ random variables with characteristic function

$$f(t) = \exp\left\{-\lambda\left(\frac{1-e^{\mathrm{i}t}}{1-\kappa e^{\mathrm{i}t}}\right)^\gamma\right\}.$$

Let us denote S_n the normalized sum

$$S_n = \frac{X_1 + X_2 + \cdots + X_n}{n^{1/\gamma}}.$$

Then the characteristic function of S_n is given as

$$\mathrm{E}\left[e^{\mathrm{i}tS_n}\right] = f^n\left(\frac{t}{n^{1/\gamma}}\right) = \exp\left\{-\lambda\left(\frac{1-e^{\mathrm{i}t/n^{1/\gamma}}}{1-\kappa e^{\mathrm{i}t/n^{1/\gamma}}}\right)^\gamma\right\}.$$

We use the Taylor expansion of exp to obtain

$$\begin{aligned}\log \mathrm{E}\left[e^{\mathrm{i}tS_n}\right] &= -\lambda n\left(\frac{-\mathrm{i}t}{(1-\kappa)n^{1/\gamma}} + O(t^2/n^{2/\gamma})\right)^\gamma \\ &= -\frac{\lambda}{(1-\kappa)^\gamma}(-\mathrm{i}t)^\gamma(1+O(n^{-2/\gamma}))^\gamma, \quad \text{as} \quad n \to \infty.\end{aligned}$$

Hence

$$g(t) = \lim_{n\to\infty}\mathrm{E}\left[e^{\mathrm{i}tS_n}\right] = \exp\left\{-\frac{\lambda}{(1-\kappa)^\gamma}(-\mathrm{i}t)^\gamma\right\}.$$

We can rewrite the exponent using

$$(-\mathrm{i}t)^\gamma = |t|^\gamma(-\mathrm{i}\,\mathrm{sign}(t))^\gamma = |t|^\gamma\cos(\pi\gamma/2)\,(1 - \mathrm{i}\,\mathrm{sign}(t)\tan(\pi\gamma/2)).$$

□

5.2 Discrete stable random variables with $\mathcal{G}$ thinning operator

In this Section we will study more into detail the discrete stable distribution in the limit sense with two-sided modified geometric thinning operator, as defined in Section 4.1. To remind the definition, an integer-valued random variable X is said to be discrete stable in the limit sense, if

$$X \overset{d}{=} \lim_{n\to\infty}\sum_{i=1}^{n}\bar{X}_i(p_n), \quad \text{where} \quad \bar{X}_i(p_n) = \sum_{j=1}^{X_i^+}\varepsilon_j^{(i)} - \sum_{j=1}^{X_i^-}\epsilon_j^{(i)}, \tag{5.10}$$

where $X_1, X_2, \ldots$ are independent copies of X and $\varepsilon_j^{(i)}, \epsilon_j^{(i)}$ are i.i.d. integer-valued random variables. Throughout this Section we will assume that the random variables $\varepsilon_j^{(i)}, \epsilon_j^{(i)}$ come from two-sided modified geometric distribution $2\mathcal{G}(p, \kappa, m, q)$ with probability generating function $\mathcal{R}$. We remind that the probability generating function $\mathcal{R}$ is given as

$$\mathcal{R}(z) = S^{-1} \circ B_p \circ S^{(2)}(z), \tag{5.11}$$

where

$$\begin{aligned} S(z) &= \frac{(1-\kappa)z^m}{1-\kappa z^m}, \\ S^{-1}(y) &= \left(\frac{y}{1-\kappa(1-y)}\right)^{\frac{1}{m}}, \\ B_p(z) &= 1-p+pz, \end{aligned}$$

and finally

$$S^{(2)}(z) = qS(z) + (1-q)S(z^{-1}).$$

Theorem 5.13. *An integer-valued random variable X is discrete stable in the limit sense with two-sided modified geometric thinning operator, if and only if $\mathcal{R}(z)$ takes form* (5.11) *and the probability generating function $\mathcal{P}(z) = \mathrm{E}z^X = \sum_{k=-\infty}^{\infty} \mathbb{P}(X=k)z^k$ takes form*

$$\begin{aligned} \mathcal{P}(z) = \exp\Bigg\{ &-\lambda\left(\frac{1+\beta}{2}\right)\left(1 - q\frac{(1-\kappa)z^m}{1-\kappa z^m} - (1-q)\frac{(1-\kappa)z^{-m}}{1-\kappa z^{-m}}\right)^{\gamma} \\ &-\lambda\left(\frac{1-\beta}{2}\right)\left(1 - (1-q)\frac{(1-\kappa)z^m}{1-\kappa z^m} - q\frac{(1-\kappa)z^{-m}}{1-\kappa z^{-m}}\right)^{\gamma}\Bigg\} \end{aligned} \tag{5.12}$$

with $\gamma \in (0,1]$, $\lambda > 0$, $\kappa \in [0,1), \beta \in [-1,1], q \in [0,1]$.

Proof. We have shown in Proposition 4.8 that a random variable X is discrete stable in the limit sense if and only if

$$\mathcal{P}(z) = \lim_{n\to\infty} \left[\mathcal{P}_0 + \mathcal{P}_1(\mathcal{R}(z)) + \mathcal{P}_2(\mathcal{R}(1/z))\right],$$

where $\mathcal{P}_1$ is the generating function of the sequence $\{p_1, p_2, \ldots\}$ with $p_k = \mathbb{P}(X=k)$ and $\mathcal{P}_2$ is the generating function of the sequence $\{q_1, q_2, \ldots\}$ with $q_k = \mathbb{P}(X=-k)$. Let us assume that $\mathcal{P}_1$ and $\mathcal{P}_2$ take the following form

$$\mathcal{P}_i(z) = \mathcal{P}_i(1) - \lambda_i\left(\frac{1-z^m}{1-\kappa z^m}\right)^{\gamma} + o\left(\left(\frac{1-z^m}{1-\kappa z^m}\right)^{\gamma}\right), \quad i = 1,2, \tag{5.13}$$

with $\gamma \in (0,1]$. We notice that

$$\frac{1-z^m}{1-\kappa z^m} = 1 - S(z).$$

This simplifies the computation, as $1 - S\left(\mathcal{R}(z)\right) = 1 - \left(1 - p + pS^{(2)}(z)\right) = p\left(1 - S^{(2)}(z)\right)$ and similarly for $1 - S(\mathcal{R}(1/z))$.

We can now compute the limit

$$\mathcal{P}(z) = \lim_{n\to\infty} \left[\mathcal{P}_0 + \mathcal{P}_1\left(\mathcal{R}(z)\right) + \mathcal{P}_2\left(\mathcal{R}(1/z)\right)\right]^n .$$

Let $p = n^{-1/\gamma}$. Then

$$\begin{aligned}\mathcal{P}(z) &= \lim_{n\to\infty} \left[1 - \lambda_1 \left(1 - S(\mathcal{R}(z))\right)^\gamma - \lambda_2 \left(1 - S(\mathcal{R}(1/z))\right)^\gamma\right]^n \\ &= \lim_{n\to\infty} \left[1 - \lambda_1 p^\gamma \left(1 - S^{(2)}(z)\right)^\gamma - \lambda_2 p^\gamma \left(1 - S^{(2)}(1/z)\right)^\gamma\right]^n \\ &= \exp\left\{-\lambda_1 \left(1 - q\frac{(1-\kappa)z^m}{1-\kappa z^m} - (1-q)\frac{(1-\kappa)z^{-m}}{1-\kappa z^{-m}}\right)^\gamma \right. \\ &\qquad \left. -\lambda_2 \left(1 - q\frac{(1-\kappa)z^{-m}}{1-\kappa z^{-m}} - (1-q)\frac{(1-\kappa)z^{m}}{1-\kappa z^{m}}\right)^\gamma\right\}.\end{aligned}$$

By setting $\lambda = \lambda_1 + \lambda_2$ and $\beta = \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2}$, we obtain the desired result. □

We will denote discrete stable distribution (and random variable) by $\mathrm{DS}^m(\gamma, \beta, \lambda, q, \kappa)$. The parameter m specifies the size of the lattice of the distribution. If we omit m then it is understood that $m = 1$. If κ is omitted we will understand that $\kappa = 0$. If moreover q is omitted we will understand that $q = 1$. In this case the probability generating function (5.12) reduces to

$$\exp\left\{-\lambda\left(\frac{1+\beta}{2}\right)(1-z)^\gamma - \lambda\left(\frac{1-\beta}{2}\right)(1-1/z)^\gamma\right\}$$

which corresponds to the discrete stable distribution introduced in Klebanov and Slámová (2013). In the case of $\beta = 1$ and $q = 1$, the $\mathrm{DS}(\gamma, 1, \lambda, 1, \kappa)$ random variable correspond to positive discrete stable random variable $\mathrm{PDS}(\gamma, \lambda, \kappa)$. In Figure A.2 the probabilities of $\mathrm{DS}(\gamma, \beta, \lambda, 1, \kappa)$ random variables are shown for different values of parameters. The probabilities were again obtained using the classical inverse Fourier transform theorem and the fast Fourier transform algorithm.

Remark 5.14. A discrete stable random variable $X \sim \mathrm{DS}(\gamma, \beta, \lambda, q, \kappa)$ is infinitely divisible, as for all $n \in \mathbb{N}$,

$$X = Y_1 + Y_2 + \cdots + Y_n, \quad \text{where} \quad Y_i \sim \mathrm{DS}(\gamma, \beta, \lambda/n, q, \kappa),\ i = 1, \ldots, n.$$

For the sake of simplicity we will denote

$$g(z) = \left(1 - q\frac{(1-\kappa)z^m}{1-\kappa z^m} - (1-q)\frac{(1-\kappa)z^{-m}}{1-\kappa z^{-m}}\right)^\gamma, \tag{5.14}$$

$$h(z) = g\left(z^{-1}\right) = \left(1 - (1-q)\frac{(1-\kappa)z^m}{1-\kappa z^m} - q\frac{(1-\kappa)z^{-m}}{1-\kappa z^{-m}}\right)^\gamma. \tag{5.15}$$

Then the probability generating function of a $\mathrm{DS}(\gamma, \beta, \lambda, q, \kappa)$ random variable can be written simply as

$$\mathcal{P}(z) = \exp\left\{-\lambda\left(\frac{1+\beta}{2}\right)g(z) - \lambda\left(\frac{1-\beta}{2}\right)h(z)\right\}.$$

5.2.1 Properties

Discrete stable distribution shares many interesting properties with stable distributions. In this Subsection we show that analogies of Properties 2.7 – 2.11 hold also for discrete stable distributions.

Property 5.15. *Let X_1 and X_2 be independent random variables with $X_i \sim \mathrm{DS}(\gamma, \beta_i, \lambda_i, q, \kappa)$, $i = 1, 2$. Then $X_1 + X_2 \sim \mathrm{DS}(\gamma, \beta, \lambda, q, \kappa)$, with*

$$\lambda = \lambda_1 + \lambda_2, \quad \beta = \frac{\beta_1\lambda_1 + \beta_2\lambda_2}{\lambda_1 + \lambda_2}.$$

Proof. Using the notation (5.14)–(5.15), the probability generating function of X_i, $i = 1, 2$, is

$$\mathcal{P}_i(z) = \exp\left\{-\lambda_i \left(\frac{1+\beta_i}{2}\right) g(z) - \lambda_i \left(\frac{1-\beta_i}{2}\right) h(z)\right\}.$$

The probability generating function of X_1+X_2 is a product of the single probability generating functions. Therefore

$$\begin{aligned}
\log \mathcal{P}_{X_1+X_2}(z) = &- \lambda_1 \left(\frac{1+\beta_1}{2}\right) g(z) - \lambda_1 \left(\frac{1-\beta_1}{2}\right) h(z) \\
&- \lambda_2 \left(\frac{1+\beta_2}{2}\right) g(z) - \lambda_2 \left(\frac{1-\beta_2}{2}\right) h(z) \\
= &- (\lambda_1 + \lambda_2)\frac{1}{2}\left(1 + \frac{\lambda_1\beta_1 + \lambda_2\beta_2}{\lambda_1 + \lambda_2}\right) g(z) \\
&- (\lambda_1 + \lambda_2)\frac{1}{2}\left(1 - \frac{\lambda_1\beta_1 + \lambda_2\beta_2}{\lambda_1 + \lambda_2}\right) h(z) \\
= &- \lambda \left(\frac{1+\beta}{2}\right) g(z) - \lambda \left(\frac{1-\beta}{2}\right) h(z),
\end{aligned}$$

where $\lambda = \lambda_1 + \lambda_2$ and $\beta = (\beta_1\lambda_1 + \beta_2\lambda_2)/(\lambda_1 + \lambda_2)$. □

Analogy of the Property 2.8 holds only for positive discrete stable random variables.

Property 5.16. *Let $X \sim \mathrm{PDS}(\gamma, \lambda, \kappa)$. Let $a \in (0, 1)$. Then $\tilde{X}(a) \sim \mathrm{PDS}(\gamma, a^\gamma\lambda, \kappa)$.*

Proof. The probability generating function of $\tilde{X}(a)$ is equal to

$$\exp\left\{-\lambda\left(1 - S(\mathcal{Q}_a(z))\right)^\gamma\right\} = \exp\left\{-\lambda a^\gamma \left(1 - S(z)\right)^\gamma\right\}.$$

□

Property 5.17. *Let $X \sim \mathrm{DS}(\gamma, \beta, \lambda, q, \kappa)$. Then $-X \sim \mathrm{DS}(\gamma, -\beta, \lambda, q, \kappa)$.*

Proof. This follows from the fact that $g(z^{-1}) = h(z)$, where we use the notation (5.14)–(5.15). Then the probability generating function of $-X$ is given as

$$\mathcal{P}(z^{-1}) = \exp\left\{-\lambda \left(\frac{1+\beta}{2}\right) h(z) - \lambda \left(\frac{1-\beta}{2}\right) g(z)\right\},$$

and this is the probability generating function of $\mathrm{DS}(\gamma, -\beta, \lambda, q, \kappa)$. □

Property 5.18. *Let $X \sim \mathrm{DS}(\gamma, \beta, \lambda, q, \kappa)$. Then X is symmetric if and only if $q = 1/2$ or $\beta = 0$.*

Proof. A discrete random variable is symmetric if and only if $\mathcal{P}(z) = \mathcal{P}(z^{-1})$. Using the notation (5.14)–(5.15), and the fact that $g(z^{-1}) = h(z)$, it follows that a discrete stable random variable is symmetric if and only if

$$-\lambda\left(\frac{1+\beta}{2}\right) g(z) - \lambda\left(\frac{1-\beta}{2}\right) h(z) = -\lambda\left(\frac{1+\beta}{2}\right) h(z) - \lambda\left(\frac{1-\beta}{2}\right) g(z).$$

But this holds true if and only if $\beta = 0$ or $g(z) = h(z)$. The latter condition is satisfied only if $q = 1/2$. □

Property 5.19. *Let X be $\mathrm{DS}(\gamma, \beta, \lambda, q, \kappa)$. Then there exist two i.i.d. random variables Y_1 and Y_2 with common distribution $\mathrm{DS}(\gamma, 1, \lambda, 1, \kappa)$ such that*

$$X \stackrel{d}{=} \bar{Y}_1\left(\left(\frac{1+\beta}{2}\right)^{1/\gamma}\right) - \bar{Y}_2\left(\left(\frac{1-\beta}{2}\right)^{1/\gamma}\right).$$

Proof. Let $Y_1, Y_2 \sim \mathrm{DS}(\gamma, 1, \lambda, 1, \kappa)$. Their probability generating function is

$$\mathcal{P}(z) = \exp\left\{-\lambda\left(\frac{1-z}{1-\kappa z}\right)^{\gamma}\right\}.$$

Moreover, the probability generating function of $\bar{Y}_i(p)$ is obtained in closed form, as Y_i are in fact positive discrete stable random variables. So we have

$$\mathcal{P}_{\bar{Y}_i(p)} = \mathcal{P}(\mathcal{R}_p(z))$$

Similarly as in the Proof of Theorem 5.13 we can compute that

$$\mathcal{P}(\mathcal{R}_p(z)) = \exp\left\{-\lambda\left(1 - S(\mathcal{R}_p(z))\right)^{\gamma}\right\} = \exp\left\{-\lambda p^{\gamma}\left(1 - q\frac{(1-\kappa)z}{1-\kappa z} - (1-q)\frac{(1-\kappa)z^{-1}}{1-\kappa z^{-1}}\right)^{\gamma}\right\}.$$

The probability generating function of the difference $\bar{Y}_1(p_1) - \bar{Y}_2(p_2)$ is computed as

$$\mathcal{P}(\mathcal{R}_{p_1}(z))\mathcal{P}(\mathcal{R}_{p_2}(1/z)).$$

Putting all together we obtain the desired result. □

5.2.2 Continuous analogies

Let us consider a random variable $X^a = aX$, with $X \sim \mathrm{DS}(\gamma, \beta, \lambda, q, \kappa)$ and $a > 0$. Then X^a takes values in $a\mathbb{Z} = \{0, \pm a, \pm 2a, \cdots\}$. We show that the limit distribution of X^a is α-stable distribution with index of stability γ and skewness β. We study the limit behaviour of X^a as $a \to 0$ and $q \to 1/2$.

Theorem 5.20. *Let X be a discrete stable random variable with parameters γ, β, λ, q and $\kappa = 0$. Let $X^a = aX$ with $a > 0$ and let $2q - 1 \approx a$ as $a \to 0$. Then*

$$\begin{aligned} f^a(t) = \exp\Bigg\{-\lambda\left(\frac{1+\beta}{2}\right)\left(1 - qe^{\mathrm{i}at} - (1-q)e^{-\mathrm{i}at}\right)^{\gamma} - \\ - \lambda\left(\frac{1-\beta}{2}\right)\left(1 - qe^{-\mathrm{i}at} - (1-q)e^{\mathrm{i}at}\right)^{\gamma}\Bigg\} \\ \longrightarrow \varphi(t) = \exp\left\{-\lambda\cos\frac{\pi\gamma}{2}|t|^{\gamma}\left(1 - \mathrm{i}\beta\,\mathrm{sign}(t)\tan\frac{\pi\gamma}{2}\right)\right\}, \quad \textit{as } a \to 0. \end{aligned}$$

Proof. We may rewrite the characteristic exponent of $f^a(t)$ as

$$\log f^a(t) \approx -\lambda\left(\frac{1+\beta}{2}\right)((2q-1)(-\mathrm{i}at))^\gamma - \lambda\left(\frac{1-\beta}{2}\right)((2q-1)(\mathrm{i}at))^\gamma, \quad \text{as} \quad a \to 0$$

and because $q \approx (1+a)/2$ we have

$$\approx -\lambda\left(\frac{1+\beta}{2}\right)(-\mathrm{i}t)^\gamma - \lambda\left(\frac{1-\beta}{2}\right)(\mathrm{i}t)^\gamma.$$

To complete the proof it is enough to notice that $(-\mathrm{i}t)^\gamma = |t|^\gamma\left(\cos\frac{\pi\gamma}{2} - \mathrm{i}\sin\frac{\pi\gamma}{2}\right)$ and $(\mathrm{i}t)^\gamma = |t|^\gamma\left(\cos\frac{\pi\gamma}{2} + \mathrm{i}\sin\frac{\pi\gamma}{2}\right)$. □

Remark 5.21. It can be shown that the case of $\kappa > 0$ leads to a similar result, the limit distribution is again α-stable with index of stability γ and skewness β.

Proof. □

5.3 Symmetric discrete stable random variables with $\mathcal{G}$ thinning operator

In the previous Section we studied the general case of discrete stable distribution in the limit sense. The symmetric version of such distribution is special case with interesting properties and we will therefore study it more into details in this Section. The symmetric discrete stable distribution in the limit sense is obtained by considering the symmetric two-sided modified geometric thinning operator $2\mathcal{G}(a, \kappa, \frac{1}{2}, m)$.

Theorem 5.22. *A symmetric integer-valued random variable X is symmetric discrete stable with symmetric two-sided $\mathcal{G}$ thinning operator if and only if the thinning operator takes form* (5.11) *with $q = 1/2$ and the probability generating function $\mathcal{P}(z) = \mathrm{E}z^X \sum_{k=-\infty}^{\infty} \mathbb{P}(X = k)z^k$ takes form*

$$\mathcal{P}(z) = \exp\left\{-\lambda\left(1 - \frac{1-\kappa}{2}\left(\frac{z^m}{1-\kappa z^m} + \frac{z^{-m}}{1-\kappa z^{-m}}\right)\right)^\gamma\right\} \tag{5.16}$$

with parameters $\gamma \in (0, 1]$, $\lambda > 0$, $\kappa \in [0, 1)$ and $m \in \mathbb{N}$.

Proof. The proof follows from the proof of Theorem 5.13. In the symmetric case we have $\mathcal{P}_1(z) = \mathcal{P}_2(z)$, therefore $\lambda_1 = \lambda_2$ and moreover $q = 1/2$. The probability generating function (5.12) thus reduces to (5.16). □

We will denote symmetric discrete stable distribution (and also random variable) by $\mathrm{SDS}^m(\gamma, \lambda, \kappa)$. In case when m is omitted we will understand that $m = 1$. If κ is omitted we will understand that $\kappa = 0$, in which case the symmetric discrete stable distribution reduces to the symmetric discrete stable distribution as it was introduced in Klebanov and Slámová (2013). In Figure A.3 the probabilities of $\mathrm{SDS}(\gamma, \lambda, \kappa)$ random variables are shown for different values of parameters. The probabilities were again obtained using the classical inverse Fourier transform theorem and the fast Fourier transform algorithm.

The characteristic function is given as

$$f(t) = \exp\left\{-\lambda\left(1-(1-\kappa)\frac{\cos(tm)-\kappa}{\kappa^2-2\kappa\cos(tm)+1}\right)^{\gamma}\right\}.$$

The case of $\gamma = 1$ is a special one as it leads to a distribution with finite variance and exponential tails.

5.3.1 Characterizations

Theorem 5.23. *Let $\gamma, \gamma' \in (0,1]$ and assume that $\gamma' \le \gamma$. Let S_γ be a γ-stable random variable with Laplace transform $\exp\{-u^\gamma\}$. Then*

$$\mathrm{SDS}(\gamma', \lambda, \kappa) \stackrel{d}{=} \mathrm{SDS}\left(\gamma'/\gamma, \lambda^{1/\gamma} S_\gamma, \kappa\right).$$

Proof. The proof of the Theorem is done in the same way as the proof of Theorem 5.3. □

Corollary 5.24. *Let $Y, Y_1, Y_2, \ldots$ be a sequence of i.i.d. random variables with two-sided geometric distribution, $\mathbb{P}(Y = \pm n) = \frac{1}{2}(1-\kappa)\kappa^{n-1}, n \ge 1$. Let N be a random variable, independent of the sequence $Y_1, Y_2, \ldots$, with Poisson distribution with random intensity $\lambda^{-1/\gamma} S_\gamma$, where S_γ is a γ-stable random variable with Laplace transform $\exp\{-u^\gamma\}$. A random variable X is symmetric discrete stable $\mathrm{SDS}(\gamma, \lambda, \kappa)$ if and only if*

$$X \stackrel{d}{=} \sum_{j=1}^{N} Y_j.$$

Proof. Let $X = \sum_{j=1}^{N} Y_j$. Then X is a compound Poisson random variable with random intensity $\lambda^{1/\gamma} S_\gamma$ and jumps $Y_1, Y_2, \ldots$ with characteristic function

$$g(t) = \frac{1}{2}\frac{(1-\kappa)e^{\mathrm{i}t}}{1-\kappa e^{\mathrm{i}t}} + \frac{1}{2}\frac{(1-\kappa)e^{-\mathrm{i}t}}{1-\kappa e^{-\mathrm{i}t}}.$$

The characteristic function of a compound Poisson random variable with intensity τ and characteristic function of jumps $h(t)$ is $\exp\{-\tau(1-h(t))\}$. Therefore X is in fact $\mathrm{SDS}(1, \lambda^{1/\gamma} S_\gamma, \kappa)$. We thus obtain the result from the previous Theorem 5.23 with $\gamma' = \gamma$. □

5.3.2 Probabilities

Theorem 5.25. *Let X be $\mathrm{SDS}(\gamma, \lambda)$ random variable. Then*

$$\mathbb{P}(X = k) = \sum_{i=|k|}^{\infty}\sum_{j=0}^{\infty}(-1)^{i+j}\binom{\gamma j}{i}\frac{\lambda^j}{j!}\frac{1}{2^i}\binom{i}{\frac{i+k}{2}}, \quad k \in \mathbb{Z}.$$

In case $\gamma = 1$ this simplifies to

$$\mathbb{P}(X = k) = e^{-\lambda} I_k(\lambda), \quad k \in \mathbb{Z}.$$

where I_k is the modified Bessel function of the first kind.

Proof. The generating function of a discrete random variable taking values in $\mathbb{Z}$ is a power series, with coefficients equal to probabilities, i.e.

$$\mathcal{P}_X(z) = \sum_{k=-\infty}^{\infty} \mathbb{P}(X = k) z^k.$$

(Note that this series converges only for $\varepsilon < |z| \leq 1$). Thus expanding (5.16) with $\kappa = 0$ into a power series we obtain the probabilities. We use Taylor expansion of exponential function, binomial expansion and interchange of sums.

$$\exp\left\{-\lambda\left[1 - \frac{1}{2}\left(z + \frac{1}{z}\right)\right]^{\gamma}\right\} = \sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\sum_{l=0}^{i}(-1)^{i+j}\binom{\gamma j}{i}\binom{i}{l}\frac{\lambda^j}{j!}\frac{1}{2^i}z^{2l-i} =$$

change of notation $k = 2l - i$ and interchange of sums

$$= \sum_{k=-\infty}^{\infty}\sum_{i=|k|}^{\infty}\sum_{j=0}^{\infty}(-1)^{i+j}\binom{\gamma j}{i}\binom{i}{\frac{i+k}{2}}\frac{\lambda^j}{j!}\frac{1}{2^i}z^k.$$

From this the first result follows. Taking $\gamma = 1$ the first binomial coefficient $\binom{j}{i}$ turns 0 for $j < i$ and we have, for $k \geq 0$,

$$\begin{aligned}\mathbb{P}(X = k) &= \sum_{i=k}^{\infty}\sum_{j=i}^{\infty}(-1)^{i+j}\binom{j}{i}\binom{i}{\frac{i+k}{2}}\frac{\lambda^j}{j!}\frac{1}{2^i} = \\ &= e^{-\lambda}\sum_{l=0}^{\infty}(\lambda/2)^{k+2l}\frac{1}{\Gamma(l+1)\Gamma(l+k+1)} = \\ &= e^{-\lambda} I_k(\lambda).\end{aligned}$$

□

5.3.3 Continuous analogies

Let us consider a case of random variable $X^a = aX$, with $X \sim \mathrm{SDS}(\gamma, \lambda, \kappa)$ and $a > 0$. Then X^a takes values in $a\mathbb{Z} = \{0, \pm a, \pm 2a, \cdots\}$. We study the limit behaviour of X^a as $a \to 0$ with $\kappa \to 1$.

Theorem 5.26. *Let X be a symmetric discrete stable random variable with parameters γ, λ and κ and let $X^a = aX$ with $a > 0$. Let $\kappa = 1 - ac$. Then*

$$f^a(t) = \exp\left\{-\lambda\left(1 - (1-\kappa)\frac{\cos(at) - \kappa}{\kappa^2 - 2\kappa\cos(at) + 1}\right)^{\gamma}\right\}$$

$$\longrightarrow \varphi(t) = \exp\left\{-\lambda\left(\frac{t^2}{t^2 + c^2}\right)^{\gamma}\right\}, \quad \textit{as } a \to 0.$$

Proof. The limit characteristic function can be computed in a straightforward way. We have

$$\begin{aligned}\left(1 - (1-\kappa)\frac{\cos(at) - \kappa}{\kappa^2 - 2\kappa\cos(at) + 1}\right) &= \left(1 + ac\frac{1 - \cos(at) - ac}{2(1-ac)(1-\cos(at)) + a^2c^2}\right) \\ &\approx \left(1 + \frac{act^2/2 - c^2}{t^2 - a\ ct^2 + c^2}\right) \quad \text{as} \quad a \to 0\end{aligned}$$

Hence we have

$$\varphi(t) = \lim_{a\to 0} \exp\left\{-\lambda\left(1+\frac{act^2/2-c^2}{t^2-act^2+c^2}\right)^{\gamma}\right\} = \exp\left\{-\lambda\left(\frac{t^2}{t^2+c^2}\right)^{\gamma}\right\}.$$

□

Next we show that symmetric discrete stable is a discrete analogy of symmetric stable distribution with index of stability $\alpha = 2\gamma$.

Theorem 5.27. *Let X be a symmetric discrete stable random variable with parameters γ, λ and κ and let $X^a = aX$ with $a > 0$. Let $\lambda = b/a^{2\gamma}$. Then*

$$f^a(t) = \exp\left\{-\lambda\left(1-(1-\kappa)\frac{\cos(at)-\kappa}{\kappa^2-2\kappa\cos(at)+1}\right)^{\gamma}\right\} \longrightarrow \varphi(t) = \exp\left\{-\sigma|t|^{2\gamma}\right\}, \quad \text{as } a \to 0,$$

where $\sigma = \frac{b}{2^\gamma}\frac{(1+\kappa)^\gamma}{(1-\kappa)^{2\gamma}}$.

Proof. We have

$$\begin{aligned} 1-(1-\kappa)\frac{\cos(at)-\kappa}{\kappa^2-2\kappa\cos(at)+1} &= (1+\kappa)\frac{1-\cos(at)}{\kappa^2-2\kappa\cos(at)+1} \\ &\approx \frac{(1+\kappa)}{2}\frac{a^2t^2}{(1-\kappa)^2+\kappa a^2t^2} \quad \text{as} \quad a \to 0 \end{aligned}$$

Hence

$$\begin{aligned} -\lambda\left(1-(1-\kappa)\frac{\cos(at)-\kappa}{\kappa^2-2\kappa\cos(at)+1}\right)^{\gamma} &\approx -\frac{b}{a^{2\gamma}}\left(\frac{(1+\kappa)}{2}\frac{a^2t^2}{(1-\kappa)^2+\kappa a^2t^2}\right)^{\gamma} \quad \text{as} \quad a \to 0 \\ &\to -\frac{b}{2^\gamma}\frac{(1+\kappa)^\gamma}{(1-\kappa)^{2\gamma}}|t|^{2\gamma} \quad \text{as} \quad a \to 0. \end{aligned}$$

□

5.3.4 Moments

In this Subsection we give a formula for factorial moments of $\mathrm{SDS}(1, \lambda, \kappa)$ distribution and show that fractional moments of $\mathrm{SDS}(\gamma, \lambda, \kappa)$ of non-integer order up to 2γ exists.

Theorem 5.28. *Let X be $\mathrm{SDS}(\gamma, \lambda, \kappa)$ random variable with $\gamma = 1$ and $\kappa > 0$. Then the n-th factorial moment can be computed using the following formula*

(5.17)

$$\mathrm{E}\left[(X)_n\right] = \frac{1}{(1-\kappa)^n}\sum_{k=1}^{n}\frac{\lambda^k}{2^k}B_{n,k}\left(0, 2!(\kappa-1), 3!(\kappa^2+1), \ldots, (n-k+1)!(\kappa^{n-k}-(-1)^{n-k+1})\right),$$

where $B_{n,k}$ is the Bell's polynomial (5.8).

Proof. The proof is analogous to the proof of Theorem 5.5 and therefore is omitted. □

Theorem 5.29. *Let $X \sim \mathrm{SDS}(\gamma, \lambda, \kappa)$ with $0 < \gamma < 1$. Then*

$$\begin{aligned} \mathrm{E}|X|^r &< \infty, \quad \textit{for any} \quad 0 < r < 2\gamma, \\ \mathrm{E}|X|^r &= \infty, \quad \textit{for any} \quad r \geq 2\gamma. \end{aligned}$$

Proof. The moments of non-integer order $\mathrm{E}|X|^r$ for any $0 < r < 2$ can be computed using the following formula (see for example (Klebanov, 2003, Lemma 2.2)):

$$\mathrm{E}|X|^r = c_r \int_0^\infty (1 - \mathrm{Re}(f(t))) \frac{\mathrm{d}t}{t^{r+1}},$$

with

$$c_r = -\frac{r}{\Gamma(1-r)\cos(\pi r/2)}$$

and where $f(t)$ is the characteristic function of the distribution of X. Since SDS is a symmetric distribution, the characteristic function of X is real, and equal to

$$f(t) = \exp\left\{-\lambda \left(\frac{(1-\cos(t))(1+\kappa)}{\kappa^2 - 2\kappa\cos(t) + 1}\right)^\gamma\right\}.$$

We may thus compute the moments.

$$\begin{aligned}
\mathrm{E}|X|^r &= c_r \int_0^\infty \left[1 - \exp\left\{-\lambda \left(\frac{(1-\cos(t))(1+\kappa)}{\kappa^2 - 2\kappa\cos(t) + 1}\right)^\gamma\right\}\right] \frac{\mathrm{d}t}{t^{1+r}} \\
&= c_r \int_0^1 \left[1 - \exp\left\{-\lambda \left(\frac{(1-\cos(t))(1+\kappa)}{\kappa^2 - 2\kappa\cos(t) + 1}\right)^\gamma\right\}\right] \frac{\mathrm{d}t}{t^{1+r}} \\
&\quad + c_r \int_1^\infty \left[1 - \exp\left\{-\lambda \left(\frac{(1-\cos(t))(1+\kappa)}{\kappa^2 - 2\kappa\cos(t) + 1}\right)^\gamma\right\}\right] \frac{\mathrm{d}t}{t^{1+r}}.
\end{aligned}$$

Using the limit comparison test we see that the first integral converges for $r < 2\gamma$ and diverges for $r \geq 2\gamma$, and the second integral converges for all $r > 0$. □

5.3.5 Asymptotic behaviour

In this Subsection we show that the tails of symmetric discrete stable $\mathrm{SDS}(\gamma, \lambda, \kappa)$ distribution are indeed heavy with tail index 2γ.

Proposition 5.30. *The symmetric discrete stable distribution* $\mathrm{SDS}(\gamma, \lambda, \kappa)$ *belongs to the domain of normal attraction of symmetric* α*-stable distribution with characteristic function*

$$g(t) = \exp\left\{-\frac{\lambda}{2^\gamma}\frac{(1+\kappa)^\gamma}{(1-\kappa)^{2\gamma}}|t|^{2\gamma}\right\}.$$

Proof. Let $X_1, X_2, \ldots, X_n$ be i.i.d. $\mathrm{SDS}(\gamma, \lambda, \kappa)$ random variables with characteristic function

$$f(t) = \exp\left\{-\lambda \left(\frac{(1-\cos(t))(1+\kappa)}{\kappa^2 - 2\kappa\cos(t) + 1}\right)^\gamma\right\}.$$

Let us denote S_n the normalized sum

$$S_n = \frac{X_1 + X_2 + \cdots + X_n}{n^{1/2\gamma}}.$$

Then the characteristic function of S_n is given as

$$\mathrm{E}\left[e^{\mathrm{i}tS_n}\right] = f^n\left(\frac{t}{n^{1/2\gamma}}\right) = \exp\left\{-\lambda \left(\frac{(1-\cos(t/n^{1/2\gamma}))(1+\kappa)}{\kappa^2 - 2\kappa\cos(t/n^{1/2\gamma}) + 1}\right)^\gamma\right\}.$$

We use the Taylor expansion of cos to obtain

$$\begin{aligned}\log \mathrm{E}\left[e^{\mathrm{i}tS_n}\right] &= -\lambda n \left(\frac{t^2}{2n^{1/\gamma}} \frac{1+\kappa}{(1-\kappa)^2} + O(n^{-3/2\gamma}) \right)^{\gamma} \\ &= -\frac{\lambda}{2^\gamma} \frac{(1+\kappa)^\gamma}{(1-\kappa)^{2\gamma}} |t|^{2\gamma} (1 + O(n^{-3/2\gamma}))^\gamma, \quad \text{as} \quad n \to \infty.\end{aligned}$$

Hence

$$g(t) = \lim_{n\to\infty} \mathrm{E}\left[e^{\mathrm{i}tS_n}\right] = \exp\left\{ -\frac{\lambda}{2^\gamma} \frac{(1+\kappa)^\gamma}{(1-\kappa)^{2\gamma}} |t|^{2\gamma} \right\}.$$

□

Theorem 5.31. *Let* $X \sim \mathrm{SDS}(\gamma, \lambda, \kappa)$ *with* $0 < \gamma < 1$. *Then*

$$\lim_{x\to\infty} x^{2\gamma} \mathbb{P}(|X| > x) = \begin{cases} \frac{\lambda}{2^\gamma} \frac{(1+\kappa)^\gamma}{(1-\kappa)^{2\gamma}} \frac{1}{\Gamma(1-2\gamma)\cos(\pi\gamma)}, & \text{if } \gamma \neq \frac{1}{2}, \\ \frac{\lambda}{2^\gamma} \frac{(1+\kappa)^\gamma}{(1-\kappa)^{2\gamma}} \frac{2}{\pi}, & \text{if } \gamma = \frac{1}{2}. \end{cases} \tag{5.18}$$

Proof. We apply (Ibragimov and Linnik, 1971, Theorem 2.6.7.): $\mathrm{SDS}(\gamma, \lambda, \kappa)$ distribution belongs to the domain of normal attraction of $\mathrm{S}(\alpha, \beta, c, \mu)$ with $\alpha = 2\gamma$, $\beta = 0$, $c = \lambda/2^\gamma(1+\kappa)^\gamma(1-\kappa)^{-2\gamma}$ and $\mu = 0$, hence the tail functions of $\mathrm{SDS}(\gamma, \lambda, \kappa)$ are given as

$$\begin{aligned} F(x) &= (c_1 + \alpha_1(x))|x|^{-\alpha}, \quad \text{for} \quad x < 0, \\ 1 - F(x) &= (c_2 + \alpha_2(x))x^{-\alpha}, \quad \text{for} \quad x > 0, \end{aligned}$$

where $\alpha_i(x) \to 0$ as $|x| \to \infty$. The constants c_1, c_2 satisfy following conditions:

$$\begin{aligned} \beta &= (c_1 - c_2)/(c_1 + c_2), \\ c &= \begin{cases} \Gamma(1-\alpha)(c_1 + c_2)\cos(\pi\alpha/2), & \text{if} \quad \alpha \neq 1, \\ \frac{\pi}{2}(c_1 + c_2), & \text{if} \quad \alpha = 1. \end{cases} \end{aligned}$$

We can easily see that for $\alpha \neq 1$ we have

$$c_1 = c_2 = \frac{1}{2} \frac{\lambda}{2^\gamma} \frac{(1+\kappa)^\gamma}{(1-\kappa)^{2\gamma}} \frac{1}{\Gamma(1-2\gamma)\cos(\pi\gamma)},$$

and for $\alpha = 1$ we have

$$c_1 = c_2 = \frac{\lambda}{2^\gamma} \frac{(1+\kappa)^\gamma}{(1-\kappa)^{2\gamma}} \frac{1}{\pi}.$$

Hence

$$\begin{aligned} \lim_{x\to\infty} x^{2\gamma} \mathbb{P}(|X| > x) &= \lim_{x\to\infty} x^{2\gamma} (F(-x) + 1 - F(x)) \\ &= \lim_{x\to\infty} x^{2\gamma} \left[(c_1 + \alpha_1(-x))x^{-2\gamma} + (c_2 + \alpha_2(x))x^{-2\gamma} \right] \\ &= 2c_1. \end{aligned}$$

□

5.3.6 Asymptotic expansion of probabilities

In this Subsection we give an asymptotic expansion of the probabilities of the symmetric discrete stable distribution with $\kappa = 0$. The following result is an adaptation of the approach used in Christoph and Schreiber (1998) for positive discrete stable random variables.

Theorem 5.32. *Let $X \sim \mathrm{SDS}(\gamma, \lambda)$, with $0 < \gamma < 1$. Then for any fixed integer m and $n \to \infty$*

$$\mathbb{P}(X = n) = \frac{2^{-n}}{\pi} \sum_{j=1}^{m} \frac{(-1)^{j+1}}{j!} \lambda^j \sin(\gamma j \pi) B(\gamma j + 1, n - \gamma j) + O(n^{-\gamma(m+1)-1}), \tag{5.19}$$

where $B(x, y) = \Gamma(x)\Gamma(y)/\Gamma(x + y)$ is the Beta function. Moreover

$$\mathbb{P}(X = n) = \frac{2^{-n}}{\pi} \sum_{j=1}^{[(\gamma+1)/\gamma]} \frac{(-1)^{j+1}}{j!} \lambda^j \Gamma(\gamma j + 1) \sin(\gamma j \pi) n^{-\gamma j - 1} + O(n^{-\gamma-2}). \tag{5.20}$$

Proof. Using the stochastic representation of $\mathrm{SDS}(\gamma, \lambda)$ random variable as a compound Poisson random variable with random intensity (Slámová and Klebanov (2012)) we have

$$\mathbb{P}(X = n) = \int_0^\infty e^{-s} I_n(s) p_\gamma^\lambda(s) \mathrm{d}s,$$

where $I_n(s)$ is the modified Bessel function of the first kind and $p_\gamma^\lambda(s)$ is the density function of the random variable S_γ^λ with characteristic function

$$g(t) = \exp\left\{-\lambda |t|^\gamma \exp(-\mathrm{i}\,\mathrm{sgn}(t)\gamma\pi/2)\right\}.$$

The density function $p_\gamma^\lambda(s)$ has the following series representation (Christoph and Wolf (1992)):

$$p_\gamma^\lambda(s) = \frac{1}{\pi} \sum_{j=1}^{m} \frac{(-1)^{j+1}}{j!} \lambda^j \Gamma(\gamma j + 1) \sin(\gamma j \pi) s^{-\gamma j - 1} + A_m(s), \tag{5.21}$$

for any $m \geq 0$, where $A_m(s) = O(s^{-\gamma(m+1)-1})$ as $s \to \infty$. We may compute the probability as

$$\mathbb{P}(X = n) = \frac{1}{\pi} \sum_{j=1}^{m} \frac{(-1)^{j+1}}{j!} \lambda^j \Gamma(\gamma j + 1) \sin(\gamma j \pi) \int_0^\infty e^{-s} I_n(s) s^{-\gamma j - 1} \mathrm{d}s + \int_0^\infty e^{-s} I_n(s) A_m(s) \mathrm{d}s.$$

We approximate the modified Bessel function $I_n(s)$ by the first term of its infinite series representation $\Gamma(n+1)^{-1}(s/2)^n$. Then the first integral turns into

$$\int_0^\infty e^{-s} I_n(s) s^{-\gamma j - 1} \mathrm{d}s \approx \frac{1}{2^n} \frac{\Gamma(n - \gamma j)}{\Gamma(n + 1)}, \quad \text{as} \quad n \to \infty.$$

The remainder term is obtained by computing the integral with $j = m + 1$ and by approximating the ratio of two Gamma functions for large n using the Stirling's formula

$$\frac{\Gamma(n - \gamma j)}{\Gamma(n + 1)} = n^{-\gamma j} \left(n^{-1} + O\left(n^{-2}\right)\right), \text{ as } n \to \infty. \tag{5.22}$$

If we set $m = [(\gamma + 1)/\gamma]$ and apply (5.22) on all terms in (5.19), we obtain (5.20). □

5.4 Positive discrete stable random variables with $\mathcal{T}$ thinning operator

The $\mathcal{G}$ thinning operator (of geometric type) used to define discrete stable distributions in the previous Sections is not the only possibility. As was showed in Chapter 4 we can consider also a $\mathcal{T}$ thinning operator (of Chebyshev type) given by the following probability generating function

$$\mathcal{Q}(z) = \left(\frac{2\left(b + T_p\left(\frac{(1+b)z^m - 2b}{2-(1+b)z^m}\right)\right)}{(1+b)\left(1 + T_p\left(\frac{(1+b)z^m-2b}{2-(1+b)z^m}\right)\right)} \right)^{1/m}, \tag{5.23}$$

where $p \in (0,1)$, $b \in (-1,1)$ and $m \in \mathbb{N}$, and $T_p(x) = \cos\left(p \arccos x\right)$.

Theorem 5.33. *A non-negative integer-valued random variable X is positive discrete stable with $\mathcal{T}$ thinning operator if and only if its probability generating function is given as*

$$\mathcal{P}(z) = \exp\left\{ -\lambda \left(\arccos \frac{(1+b)z^m - 2b}{2-(1+b)z^m} \right)^{\gamma} \right\} \quad \textit{with} \quad \gamma \in (0,2],\ \lambda > 0,\ b \in (-1,1),\ m \in \mathbb{N}. \tag{5.24}$$

Proof. Let $h(z) = \log \mathcal{P}(z)$. From Proposition 4.2 it follows that X is positive discrete stable if and only if $h(z) = nh(\mathcal{Q}(z))$ for all n, where $\mathcal{Q}$ is as in (5.23). Set

$$h(z) = -\lambda \left(\arccos \frac{(1+b)z^m - 2b}{2-(1+b)z^m} \right)^{\gamma}$$

and select γ such that $1/p^\gamma = n$. Then

$$\begin{aligned} nh(\mathcal{Q}(z)) &= -\lambda n \left(\arccos \frac{(1+b)\mathcal{Q}(z)^m - 2b}{2-(1+b)\mathcal{Q}(z)^m} \right)^{\gamma} \\ &= -\lambda n \left(\arccos T_p\left(\frac{(1+b)z^m - 2b}{2-(1+b)z^m} \right) \right)^{\gamma} \\ &= -\lambda n \left(p \arccos \frac{(1+b)z^m - 2b}{2-(1+b)z^m} \right)^{\gamma} \\ &= h(z). \end{aligned}$$

□

We will denote the discrete stable distribution with Chebyshev thinning operator $\mathcal{T}$ and with parameters $\gamma \in (0,2], \lambda > 0,\ b \in (-1,1)$ and $m \in \mathbb{N}$, by $\mathcal{T}\mathrm{PDS}(\gamma, \lambda, b, m)$. If m is omitted then $m = 1$. If moreover b is omitted we will understand that $b = 0$.

5.4.1 Characterizations

Theorem 5.34. *Let $\gamma' \in (0,2]$ and $\gamma \in (0,1]$ and assume that $\gamma' \leq 2\gamma$. Let S_γ be a γ-stable random variable with Laplace transform $\exp\{-u^\gamma\}$. Then*

$$\mathcal{T}\mathrm{PDS}(\gamma', \lambda, b) \stackrel{d}{=} \mathcal{T}\mathrm{PDS}\left(\gamma'/\gamma, \lambda^{1/\gamma}\mathrm{S}_\gamma, b\right).$$

Proof. For sake of simplicity we will do the proof only for the case $b = 0$. The case $b \neq 0$ can be proved in the same way. The probability generating function of $X \sim \mathcal{T}\text{PDS}\left(\gamma'/\gamma, \lambda^{1/\gamma}\text{S}_\gamma\right)$ is computed as

$$\mathcal{P}(z) = \text{E}z^X = \text{E}\exp\left\{-\lambda^{1/\gamma}\text{S}_\gamma\left(\arccos\frac{z}{2-z}\right)^{\gamma'/\gamma}\right\}$$

and using the Laplace transform formula for S_γ we have

$$\mathcal{P}(z) = \exp\left\{-\lambda\left(\arccos\frac{z}{2-z}\right)^{\gamma'}\right\}.$$

This is the probability generating function of $\mathcal{T}\text{PDS}(\gamma', \lambda)$. □

Corollary 5.35. *Let $Y, Y_1, Y_2, \dots$ be a sequence of i.i.d. random variables with probability generating function*

$$\mathcal{P}(z) = 1 - \tfrac{1}{\pi}\arccos\frac{(1+b)z - 2b}{2-(1+b)z}.$$

Let N be a random variable, independent of the sequence $Y_1, Y_2, \dots$, with Poisson distribution with random intensity $\lambda^{1/\gamma}\pi\text{S}_\gamma$, where $\gamma \in (0,1]$ and S_γ is a γ-stable random variable with Laplace transform $\exp\{-u^\gamma\}$. A random variable X is positive discrete stable $\mathcal{T}\text{PDS}(\gamma, \lambda, b)$ if and only if

$$X \stackrel{d}{=} \sum_{j=1}^{N} Y_j.$$

Proof. Let $X = \sum_{j=1}^{N} Y_j$. Then X is a compound Poisson random variable with random intensity $\lambda^{1/\gamma}\pi\text{S}_\gamma$ and jumps $Y_1, Y_2, \dots$ with characteristic function

$$g(t) = 1 - \tfrac{1}{\pi}\arccos\frac{(1+b)e^{\text{i}t} - 2b}{2-(1+b)e^{\text{i}t}}.$$

The characteristic function of a compound Poisson random variable with intensity τ and characteristic function of jumps $h(t)$ is $\exp\{-\tau(1-h(t))\}$. Therefore X is in fact $\mathcal{T}\text{PDS}(1, \lambda^{1/\gamma}\text{S}_\gamma, b)$. We thus obtain the result from the previous Theorem 5.34 with $\gamma' = \gamma$. □

5.4.2 Continuous analogies

Let us consider a $\mathcal{T}$ positive discrete stable random variable $X \sim \mathcal{T}\text{PDS}(\gamma, \lambda, b)$. We are interested in the limit distribution of a random variable $X^a = aX$, where $a \downarrow 0$. We show that the limit distribution is in fact α-stable with index of stability $\alpha = \gamma/2$ and with skewness $\beta = 1$.

Theorem 5.36. *Let X be a random variable with probability generating function*

$$\mathcal{P}(z) = \exp\left\{-\lambda\left(\arccos\frac{(1+b)z-2b}{2-(1+b)z}\right)^{\gamma}\right\}, \quad \gamma \in (0,2],\ \lambda > 0,\ b \in (-1,1).$$

Let $X^a = aX$ and assume that $\lambda = \frac{\sigma}{a^{\gamma/2}}$. Then the characteristic function of X^a converges pointwise to the characteristic function of α-stable distribution,

$$f^a(t) = \exp\left\{-\lambda\left(\arccos\frac{(1+b)e^{\mathrm{i}at}-2b}{2-(1+b)e^{\mathrm{i}at}}\right)^{\gamma}\right\}$$
$$\longrightarrow \exp\left\{-\sigma 2^{\gamma}\cos\frac{\pi\gamma}{4}\left(\frac{1+b}{1-b}\right)^{\gamma/2}|t|^{\gamma/2}\left(1-\mathrm{i}\,\mathrm{sign}(t)\tan\frac{\pi\gamma}{4}\right)\right\}.$$

Proof. For sake of simplicity we will do the proof only for $b=0$. The characteristic function of X^a can be approximated as

$$\log f^a(t) = -\lambda\left(\arccos\frac{e^{\mathrm{i}at}}{2-e^{\mathrm{i}at}}\right)^{\gamma}$$
$$\approx -\lambda\left(\arccos\frac{1+\mathrm{i}at}{1-\mathrm{i}at}\right)^{\gamma}, \quad \text{as} \quad a \to 0.$$

Moreover $\arccos(z) \approx \sqrt{2}\sqrt{1-z}$ as $z \to 1$. We have

$$1-\frac{1+\mathrm{i}at}{1-\mathrm{i}at} = \frac{-2\mathrm{i}at}{1-\mathrm{i}at}.$$

Put together we obtain

$$\log f^a(t) \approx -\frac{\sigma}{a^{\gamma/2}}\left(2\sqrt{\frac{-\mathrm{i}at}{1-\mathrm{i}at}}\right)^{\gamma} \quad \text{as} \quad a \to 0$$
$$\to -\sigma 2^{\gamma}(-\mathrm{i}t)^{\gamma/2}, \quad \text{as} \quad a \to 0.$$

Moreover we have $(-\mathrm{i}t)^{\gamma/2} = \cos\frac{\pi\gamma}{4}|t|^{\gamma/2}\left(1-\mathrm{i}\,\mathrm{sign}(t)\tan\frac{\pi\gamma}{4}\right)$. The proof is therefore completed. □

6. Casual stable random variables

The difference between classical definition of stability and the definition of discrete stability in the first sense lies in the different approach to normalization of the sum $\sum_{i=1}^{n} X_i$. Definition of discrete stability in the first sense assumes a specific choice of a thinning operator that is applied on single summands. Assume a non-negative integer-valued random variable X with cumulative distribution function (c.d.f.) $F(x)$, with probability generating function

$$\mathcal{P}(z) = \mathrm{E}z^X = \int_0^\infty z^x \mathrm{d}F(x). \tag{6.1}$$

We normalize X by transforming this probability generating function using the thinning operator $\mathcal{Q}(z)$ into

$$\mathcal{P}(\mathcal{Q}(z)) = \int_0^\infty [\mathcal{Q}(z)]^x \, \mathrm{d}F(x).$$

The normalized random variable is denoted $\tilde{X}$ and its probability generating function is $\mathcal{P}(\mathcal{Q}(z))$. We replaced the probability generating function z of a degenerate random variable concentrated at point 1 into a probability generating function $\mathcal{Q}(z)$. A similar approach can be considered for the case of continuous random variables using Laplace transforms instead of the probability generating functions. We remind that a Laplace transform of a random variable X is defined as $L(s) = \mathrm{E}e^{-sX}$. Consider first a non-negative random variable X with c.d.f. $F(x)$. Its Laplace transform has the form

$$L(s) = \int_0^\infty e^{-sx} \mathrm{d}F(x),$$

or

$$L(s) = \int_0^\infty \left[e^{-s}\right]^x \mathrm{d}F(x). \tag{6.2}$$

Similarly as in the discrete case where we apply the thinning operator $\mathcal{Q}$ by replacing the probability generating function of a degenerate distribution z by a different probability generating function $\mathcal{Q}(z)$ in (6.1), we can replace the Laplace transform of a degenerate random variable e^{-s} in (6.2) by another Laplace transform of a distribution concentrated on positive semi-axis, say $g(s)$. This g-normalized random variable X will be denoted $\tilde{X}$ and its Laplace transform $L_g(s)$ takes form

$$L_g(s) = \int_0^\infty [g(s)]^x \, \mathrm{d}F(x).$$

It is important to note that for discrete random variables there is a one-on-one correspondence between the probability generating function and its Laplace transform, $L(s) = \mathcal{P}(e^{-s})$ and the two normalizations by thinning operator $\mathcal{Q}$ and by the Laplace transform $g(s)$ are therefore identical.

This Chapter contains results from Klebanov and Slámová (2014) where the casual stability was introduced and some examples of casual stable distributions were studied. Here we review the results and provide additional examples.

6.1 A general definition of stability

Let $X, X_1, X_2, \dots$ be i.i.d. non-negative random variables with Laplace transform $L(s)$. We would like to define a new type of stability with this new type of g-normalization where instead of the classical normalization in the definition of strict stability

$$X \stackrel{d}{=} a_n \sum_{i=1}^{n} X_i$$

we use the random normalization with function $g(s)$ and define casual stability as

$$X \stackrel{d}{=} \sum_{i=1}^{n} \tilde{X}_i,$$

with $\tilde{X}_i$ having Laplace transform $\int_0^\infty [g(s)]^x \mathrm{d}F(x) = L(-\log g(s))$.

Definition 6.1. Let X be a non-negative random variable with Laplace transform $L(s)$. Assume that for every $n \in \mathbb{N}$ there exists a Laplace transform $g_n(s)$ such that

$$L^n(-\log g_n(s)) = L(s). \tag{6.3}$$

Then we say that a random variable X is casual stable.

This definition is restrictive – the corresponding random variable has to be infinitely divisible and the Laplace transform also cannot be arbitrary. The casual stability can be interpreted in the following way. Let us suppose that we have a (discrete or continuous) flow of particles passing through a filter. Each particle of the flow may generate some other particles or just disappear, according to a probability distribution with Laplace transform $g_n(s)$. The casual stability means that the distribution of n such flows after passing through the filter has the same distribution as the initial flow before passing through the filter. Alternatively, we may say that an additive system is "randomly similar" to its initial element, so the system is "randomly self-similar".

The contribution of this definition of casual stability is that it encompasses classical positive stable random variables, as well as positive discrete stable random variables and positive tempered stable random variables. Other examples of casual stable distributions are gamma distribution, geometric distribution, negative binomial distribution. In the following Subsections we show that these distributions are casual stable by deriving the form of the normalizing function $g_n(s)$ and by showing that it is a Laplace transform.

6.1.1 Stable distribution

Let X be a positive stable random variable with Laplace transform $L(s) = \exp\{-s^\alpha\}$ and $\alpha \in (0,1)$. With the choice of $g_n(s) = \exp\{-a_n s\}$, which corresponds to a generate random variable concentrated at point a_n, we have just ordinary normalization, and corresponding casual stable distribution coincide with stable distribution totally skewed to the right ($\beta = 1$) with index of stability $\alpha \in (0,1)$. To show this let $g_n(s) = \exp\{-a_n s\}$ with $a_n = n^{-1/\alpha}$. Then

$$L^n(-\log g_n(s)) = \exp\left\{-n a_n^\alpha s^\alpha\right\} = L(s),$$

therefore X is casual stable.

6.1.2 Discrete stable distribution

Let us consider the modified geometric thinning operator from Section 5.1 where

$$\mathcal{Q}_n(z) = \left(\frac{(1-p) + (p-\kappa)z^m}{(1-p\kappa) - \kappa(1-p)z^m} \right)^{1/m},$$

with $p = p_n$. The corresponding normalizing Laplace transform is obtained as $g_n(s) = \mathcal{Q}_n(e^{-s})$.

As we know, this thinning operator leads to positive discrete stable random variables with probability generating function

$$\mathcal{P}(z) = \exp\left\{ -\lambda \left(\frac{1-z^m}{1-\kappa z^m} \right)^{\gamma} \right\}.$$

We can easily see that the resulting distribution is casual stable as well, as $L(s) = \mathcal{P}(e^{-s})$. So with $p_n = n^{-1/\gamma}$

$$L^n(-\log g_n(s)) = \mathcal{P}^n(\mathcal{Q}_n(e^{-s})) = \mathcal{P}(e^{-s}) = L(s).$$

In the same way we can show that discrete stable distribution with the Chebyshev thinning operator, studied in Section 5.4, is also casual stable.

6.1.3 Tempered stable distribution

This example is most interesting as it shows that some tempered stable random variables are also casual stable. In particular we show that Inverse Gaussian distribution is casual stable. Let us consider a positive stable random variable with index of stability $\alpha = 1/m, m \in \mathbb{N}$, with Laplace transform $L(s) = \exp\{-s^\alpha\}$. We shift the Laplace transform θ units to the right and we obtain tempered stable random variable (as in Lemma 5.7) with Laplace transform

$$L(s) = \exp\left\{ -(s+\theta)^\alpha + \theta^\alpha \right\}.$$

It appears that this distribution is casual stable with normalizing function g_n of the form

$$g_n(s) = \exp\left\{ \theta - \left[\frac{1}{n}(s+\theta)^\alpha + \frac{n-1}{n}\theta^\alpha \right]^{1/\alpha} \right\}.$$

Lemma 6.2. *The function $g_n(s)$ is a Laplace transform of a probability distribution of a probability distribution if $1/\alpha \in \mathbb{N}$.*

Proof. In the case when $1/\alpha \in \mathbb{N}$ we can rewrite the function $g_n(s)$ as a product of $1/\alpha$ Laplace transforms, and as such it is a Laplace transform itself. Namely

$$\begin{aligned} g_n(s) &= \exp\left\{ \theta - \theta \left(\frac{n-1}{n} \right)^{1/\alpha} \sum_{k=0}^{1/\alpha} \binom{1/\alpha}{k} \frac{1}{(n-1)^k} \left(\frac{\theta+s}{\theta} \right)^{k\alpha} \right\} \\ &= \prod_{k=0}^{1/\alpha} \exp\left\{ \theta^{1-k\alpha} \left(\frac{n-1}{n} \right)^{1/\alpha} \binom{1/\alpha}{k} \frac{1}{(n-1)^k} \left(\theta^{k\alpha} - (\theta+s)^{k\alpha} \right) \right\} \\ &= \prod_{k=0}^{1/\alpha} h_k(s). \end{aligned}$$

For every k, $h_k(s)$ is a Laplace transform of a tempered stable random variable, as $0 < k\alpha \leq 1$. □

Now we can show that tempered stable random variable with Laplace transform L is casual stable.

$$\begin{aligned} L^n(-\log g_n(s)) &= \exp\left\{-\left[(s+\theta)^\alpha + (n-1)\theta^\alpha\right] + n\theta^\alpha\right\} \\ &= \exp\left\{-(s+\theta)^\alpha + \theta^\alpha\right\} \\ &= L(s). \end{aligned}$$

For $\theta \to 0$ we obtain the classical case of normalization – g_n is Laplace transform of a degenerate distribution at point $a_n = n^{-1/\alpha}$. As a particular case we find that Inverse Gaussian distribution (the case $\alpha = 1/2$) is casual stable.

6.1.4 Tempered discrete stable distribution

Let us consider distribution with Laplace transform

$$L(s) = \exp\left\{-\lambda\left(1 - ae^{-s}\right)^\gamma + \lambda(1-a)^\gamma\right\}, \quad a \in (0,1].$$

This distribution is a tempered version of the discrete stable distribution with binomial thinning operator. We can show that tempered positive discrete stable random variable is again casual stable. The normalizing function g_n takes form

$$g_n(s) = \frac{1}{a}\left[1 - \left(\frac{1}{n}\left(1 - ae^{-s}\right)^\gamma + \frac{n-1}{n}\left(1-a\right)^\gamma\right)^{1/\gamma}\right].$$

Let us assume that $1/\gamma \in \mathbb{N}$. Then we can expand the power $1/\gamma$ using binomial expansion and after some computations we obtain that

$$g_n(s) = \sum_{k=0}^{1/\gamma} \binom{1/\gamma}{k} \left(\frac{n-1}{n}\right)^{1/\gamma} \frac{1}{(n-1)^k} h_k(s),$$

where

$$h_k(s) = \frac{1}{a}\left[1 - (1-a)^{1-\gamma k}\left(1 - ae^{-s}\right)^{\gamma k}\right].$$

We see that $g_n(s)$ is a linear combination of functions $h_k(s)$. If those functions are Laplace transforms of some probability distribution, so is $g_n(s)$. Let us denote $\mathcal{P}_k(z) = h_k(-\log z)$. We have $\gamma k \le 1$. For $\gamma k = 1$ we see that $\mathcal{P}_k(z) = z$ and it is a probability generating function. For $\gamma k < 1$ we may expand the power γk into an infinite power series $\mathcal{P}_k(z) = \sum_{j=0}^\infty p_j z^j$. The coefficients before z^j are

$$\begin{aligned} p_0 &= \frac{1}{a}\left(1 - (1-a)^{1-\gamma k}\right), \\ p_j &= -\frac{(1-a)^{1-\gamma k}}{a}\binom{\gamma k}{j}(-1)^j a^j, \quad j \ge 1. \end{aligned}$$

Because $\gamma k < 1$, the binomial coefficients in p_j change sign. Therefore, for $a \in (0,1]$ the coefficients are all positive, moreover $\mathcal{P}_k(1) = 1$ and $\mathcal{P}_k(z)$ are therefore probability generating functions and $h_k(s)$ are Laplace transforms.

We can obtain similar result for the tempered discrete stable distribution with modified geometric thinning operator. The Laplace transform of such distribution takes form

$$L(s) = \exp\left\{-\lambda\left(\frac{1 - ae^{-s}}{1 - \kappa e^{-s}}\right)^\gamma + \lambda\left(\frac{1-a}{1-\kappa}\right)^\gamma\right\}, \quad a \in (0,1].$$

6.1.5 Geometric distribution

We can show that geometric distribution with Laplace transform

$$L(s)=\frac{1-a}{1-ae^{-s}},\quad a\in[0,1)$$

is casual stable. We look for a normalizing function $g_n(s)$ such that $L^n(-\log g_n(s))=L(s)$ and such that g_n is a Laplace transform. By simple rearranging we deduce that function

$$g_n(s)=\frac{1}{a}\left[1-(1-a)^{1-1/n}\left(1-ae^{-s}\right)^{1/n}\right]$$

satisfies the equation (6.3). Also we can verify that $g_n(s)$ is a Laplace transform. Series expansion gives us

$$\left(1-ae^{-s}\right)^{1/n}=\sum_{k=0}^{\infty}(-1)^k\binom{1/n}{k}a^ke^{-sk},$$

and hence $g_n(s)$ can be written as a sum of Laplace transforms and therefore is a Laplace transform.

The normalizing distribution can be derived in other way as well. The probability generating function of the normalizing distribution can be rewritten as

$$\mathcal{P}_n(z)=\frac{1-c(1-bz)^{\gamma}}{1-c(1-b)^{\gamma}},$$

where $\gamma=1/n$, $b=a$ and $c=(1-a)^{1-1/n}$. This distribution is modified scaled Sibuya distribution. The scaled Sibuya distribution was introduced by Christoph and Schreiber (2000) and its probability generating function takes form $1-c(1-z)^{\gamma}$ with exponent $\gamma\in(0,1]$ and scale parameter $c\in(0,1]$. It appears as a mixture of the classical Sibuya distribution and distribution concentrated at 0, with weights c and $1-c$ respectively. If we denote the probabilities of scaled Sibuya distribution by p_k, $k=0,1,\ldots$, we can create a new distribution by multiplying the probabilities by b^k and normalizing them. We therefore obtain a distribution with probabilities

$$q_k=\frac{b^kp_k}{\sum_{k=0}^{\infty}b^kp_k}=\frac{b^kp_k}{1-c(1-b)^{\gamma}}.$$

The probability generating function of modified scaled Sibuya distribution is then

$$\mathcal{Q}(z)=\sum_{k=0}^{\infty}q_kz^k=\frac{\sum_{k=0}^{\infty}p_k(bz)^k}{1-c(1-b)^{\gamma}}=\frac{1-c(1-bz)^{\gamma}}{1-c(1-b)^{\gamma}}.$$

6.1.6 Negative binomial distribution

Consider now negative binomial distribution with Laplace transform

$$L(s)=\left(\frac{1-a}{1-ae^{-s}}\right)^k,\quad a\in[0,1),\ k>0.$$

We see right away that the normalizing function $g_n(s)$ takes the same form as for geometric distribution, i.e.

$$g_n(s)=\frac{1}{a}\left[1-(1-a)^{1-1/n}\left(1-ae^{-s}\right)^{1/n}\right].$$

Therefore negative binomial distribution is also casual stable.

6.1.7 Gamma distribution

Consider a random variable X with gamma distribution with Laplace transform

$$L(s) = \frac{1}{(1+bs)^{\gamma}},$$

with parameters $b > 0$ and $\gamma > 0$. We show that the random variable X is casual stable. We must have $L^n(-\log g_n(s)) = L(s)$ for any integer $n > 1$. From here we find that

$$g_n(s) = \exp\left\{\frac{1}{b}\left(1-(1+bs)^{1/n}\right)\right\}.$$

For every $n > 1$ the function $g_n(s)$ is a Laplace transform of tempered stable distribution: with $\alpha = 1/n$ and $\theta = 1/b$, we may rewrite g_n as

$$g_n(s) = \exp\left\{-b^{\alpha-1}\left[(\theta+s)^{\alpha}-\theta^{\alpha}\right]\right\}.$$

Therefore, X is a casual stable random variable.

6.2 Casual stability of random variables of arbitrary sign

Let X be a random variable taking values on the whole real line, with c.d.f. $F(x)$. The characteristic function $f(t)$ of X may be written as

$$f(t) = \int_{-\infty}^{\infty} e^{\mathrm{i}tx}\mathrm{d}F(x) = \int_0^{\infty}(e^{-\mathrm{i}t})^{|x|}\mathrm{d}(1-F(-x)) + \int_0^{\infty}(e^{\mathrm{i}t})^{x}\mathrm{d}F(x).$$

The "random normalization" in the case of random variables on the whole real line is done in the following way. We replace e^{it} in the second integral by a characteristic function $g(t)$ but in the first integral we replace e^{-it} by $g(-t)$. We obtain the characteristic function of g-normalized random variable $f_g(t)$ in the form

$$f_g(t) = \int_0^{\infty}(g(-t))^{|x|}\mathrm{d}(1-F(-x)) + \int_0^{\infty}(g(t))^{x}\mathrm{d}F(x).$$

The definition of casual stability in this case is obvious now.

Definition 6.3. Let X be a random variable with characteristic function $f(t)$. Assume that for every $n \in \mathbb{N}$ there exists a characteristic function $g_n(s)$ such that

$$[f_{g_n}(t)]^n = f(t). \tag{6.4}$$

Then we say that a random variable X is casual stable. If a random variable X is discrete and moreover it is casual stable then we say it is discrete stable in algebraic sense.

The verification of the fact that a random variable with values on the whole real line is casual stable is generally more difficult than for the case of positive random variables. However, we can prove, for example, that the Laplace distribution is casual stable. The same fact holds for Linnik distribution as well.

6.2.1 Laplace distribution

Consider Laplace distribution with c.d.f. $F(x)$ and characteristic function

$$f(t) = \frac{1}{1+a^2t^2},$$

where $a > 0$. We can rewrite $f(t)$ as

$$f(t) = \frac{1}{2}\frac{1}{1-\mathrm{i}at} + \frac{1}{2}\frac{1}{1+\mathrm{i}at} = \int_0^\infty [e^{\mathrm{i}t}]^x \mathrm{d}F(x) + \int_0^\infty [e^{-\mathrm{i}t}]^{|x|}\mathrm{d}(1-F(-x)).$$

If we replace $e^{\mathrm{i}t}$ by $g_n(t)$ and $e^{-\mathrm{i}t}$ by $g_n(-t)$ and if we moreover assume that g_n is symmetric, i.e. $g_n(t) = g_n(-t)$, we obtain g_n-normalized characteristic function

$$f_{g_n}(t) = \frac{1}{1-a\log g_n(t)}.$$

If the function $g_n(t)$ satisfying $[f_{g_n}(t)]^n = f(t)$ is a characteristic function, then Laplace distribution is casual stable. By simple computation we obtain

$$g_n(t) = \exp\left\{\frac{1}{a}\left[1-\left(1+a^2t^2\right)^{1/n}\right]\right\}.$$

To show that $g_n(t)$ is a characteristic function it suffices to consider only the case of $a = 1$ as the general case is obtained by rescaling and taking a positive power. We thus have to show that $\exp\{1-(1+t^2)^\alpha\}$ is a characteristic function for $\alpha \in (0,1)$. We may rewrite $g_n(t)$ in the following way

$$g_n(t) = \exp\left\{1 - c_\alpha \int_0^\infty \left(1-e^{-x(1+t^2)}\right)\frac{\mathrm{d}x}{x^{1+\alpha}}\right\},$$

where

$$c_\alpha = 1/\int_0^\infty \left(1-e^{-x}\right)\frac{\mathrm{d}x}{x^{1+\alpha}} = -1/\Gamma(-\alpha) > 0 \quad \text{for} \quad \alpha \in (0,1).$$

We may approximate the integral by integral sums. Let $T \in \mathbb{R}_+$ and $m \in \mathbb{N}$ be fixed and denote $x_k = Tk/m$ for $k = 1, \ldots, m$. Then we may approximate $g_n(t)$ as

$$\begin{aligned} g_n(t) &\approx \exp\left\{1 - c_\alpha \lim_{T\to\infty}\lim_{m\to\infty}\sum_{k=1}^m \left(1-e^{-x_k(1+t^2)}\right)\frac{T/m}{x_k^{1+\alpha}}\right\} \\ &= e \times \lim_{T\to\infty}\lim_{n\to\infty}\prod_{k=1}^m \exp\left\{c_\alpha \frac{T}{m}(e^{-x_k(1+t^2)}-1)x_k^{1+\alpha}\right\}. \end{aligned}$$

De Finetti's theorem (see (Lukacs, 1970, § 5.4)) states that a characteristic function $f(t)$ is infinitely divisible if and only if it can be written as

$$f(t) = \lim_{m\to\infty} \exp\left\{a_m(h_m(t)-1)\right\},$$

where a_m are positive constants and $h_m(t)$ are characteristic functions. We see that $g_n(t)$ takes a form of a product of functions of this kind and $h_m(t)$ are characteristic functions of Gaussian distribution. We conclude that $g_n(t)$ is a characteristic function.

6.2.2 Linnik distribution

Linnik distribution is a special case of geometric stable distribution (belonging to the class of ν-stable distributions). Its characteristic function is given as

$$f(t) = \frac{1}{1 + a^\alpha |t|^\alpha}, \quad \alpha \in (0, 2].$$

It is a symmetric distribution and Laplace distribution is a special case with $\alpha = 2$. In a similar, but technically more demanding way, we can show that Linnik distribution is also casual stable.

6.2.3 Symmetric geometric distribution

We can introduce a discrete analogy of Laplace distribution if we consider a discrete probability distribution derived from geometric distribution with characteristic function

$$f(t) = \frac{1}{2}\frac{1-a}{1-ae^{\mathrm{i}t}} + \frac{1}{2}\frac{1-a}{1-ae^{-\mathrm{i}t}} = (1-a)\frac{1 - a\cos(t)}{1 - 2a\cos(t) + a^2}.$$

We can show that this distribution is casual stable and therefore symmetric discrete stable in algebraic sense. To show that discrete Laplace distribution is casual stable we need to find the normalizing function $g_n(t)$. We will assume once again that g_n is symmetric, i.e. $g_n(t) = g_n(-t)$. We replace both $e^{\mathrm{i}t}$ and $e^{-\mathrm{i}t}$ by $g_n(t)$. Then the g_n-normalized characteristic function is $f_{g_n}(t) = (1-a)/(1-ag_n(t))$ and we must have

$$\frac{(1-a)^n}{(1-ag_n(t))^n} = (1-a)\frac{1 - a\cos(t)}{1 - 2a\cos(t) + a^2}.$$

From here we get the normalizing function $g_n(t)$ as

$$g_n(t) = \frac{1}{a}\left[1 - \left((1-a)^{n-1}\frac{1 - 2a\cos(t) + a^2}{1 - a\cos(t)}\right)^{1/n}\right].$$

We can show that g_n is a characteristic function. We start by rewriting $g_n(t)$ as

$$g_n(t) = \frac{1}{a}\left[1 - (1 - h_n(t))^{1/n}\right],$$

where

$$h_n(t) = 1-(1-a)^{n-1}\frac{1 - 2a\cos(t) + a^2}{1 - a\cos(t)} = \frac{1 - (1-a)^{n-1}(1+a^2) + a\ (2(1-a)^{n-1} - 1)\cos(t)}{1 - a\cos(t)}.$$

To prove that $g_n(t)$ is a characteristic function it is sufficient to show that $h_n(t)$ is a characteristic function. Then, by expanding the $1/n$ power into power series, we see that g_n is a sum of characteristic functions . Using power expansion of $1/(1 - a\cos(t))$ through powers of $\cos(t)$ we find that the coefficient before $\cos^k(t)$ for $k = 0$ is $1 - (1-a)^{n-1}(1+a^2) > 0$ (for sufficiently large n and $a \in (0,1)$) and for $k \geq 1$ is

$$a^k(1-a)^{n-1}(1-a^2) > 0.$$

We see that the function h_n is a convex combination of characteristic functions $\cos^k(t)$ for $k = 0, 1, 2, \ldots$. Therefore g_n is a characteristic function.

We conclude that the symmetric geometric distribution is casual stable, and therefore it is symmetric discrete stable in the algebraic sense. The tails of this distribution are exponential, and the limit distribution (when the size of the lattice goes to zero) is Laplace distribution.

6.2.4 Symmetric negative binomial distribution

In a similar way as we introduced symmetric geometric distribution we can introduce symmetric negative binomial distribution with characteristic function

$$f(t) = \left(\frac{1}{2}\frac{1-a}{1-ae^{\mathrm{i}t}} + \frac{1}{2}\frac{1-a}{1-ae^{-\mathrm{i}t}}\right)^k = \left((1-a)\frac{1-a\cos(t)}{1-2a\cos(t)+a^2}\right)^k, \quad k > 0.$$

As it turns out, the normalizing characteristic function takes the same form as in the case of the symmetric geometric distribution, i.e.

$$g_n(t) = \frac{1}{a}\left[1 - \left((1-a)^{n-1}\frac{1-2a\cos(t)+a^2}{1-a\cos(t)}\right)^{1/n}\right].$$

The symmetric negative binomial distribution is therefore casual stable and discrete stable in the algebraic sense as well.

6.3 Convergence to casual stable distribution

In this Section we give a limit theorem for convergence to casual stable distribution in the case of non-negative random variables. The general case can be considered as well but the formulations appear to be more complicated.

A simple example of this theorem is the convergence to Poisson distribution. Let $X_1, X_2, \ldots$ be i.i.d. random variables with probability generating function $\mathcal{P}(z)$. Let us assume that X_1 has finite first moment and the probability generating function takes form

$$\mathcal{P}(z) = 1 - c(1-z) + o((1-z)).$$

Let us now consider g_n-normalized random variables $\tilde{X}_i$, where $g_n(s)$ is the Laplace transform of binomial thinning operator with probability generating function $\mathcal{Q}(z) = 1 - p + pz$, i.e. $g_n(s) = \mathcal{Q}(e^{-s}) = 1 - p + pe^{-s}$, with $p = 1/n$. Then the probability generating function of the g_n-normalized random variables $\tilde{X}_i$ takes form

$$\mathcal{P}(\mathcal{Q}(z)) = 1 - cp(1-z) + o((1-z)).$$

Let us denote $S_n = \sum_{i=1}^n \tilde{X}_i$. Then the probability generating function of the sum S_n converges to the probability generating function of the Poisson distribution:

$$\mathcal{P}_{S_n}(z) = \mathcal{P}^n(\mathcal{Q}(z)) = \left(1 - c\frac{1}{n}(1-z) + o((1-z))\right)^n$$
$$\longrightarrow \exp\{-c(1-z)\} \quad \text{as} \quad n \to \infty.$$

This convergence result can be generalized as in the following theorem. Let $L(s)$ be a Laplace transform of a non-negative random variable, which is casual stable with Laplace transform $g_n(s)$, so

$$L(s) = L^n(-\log g_n(s)), \quad \forall n \in \mathbb{N}.$$

Theorem 6.4. *Suppose that* $h(s)$ *is a Laplace transform such that* $\sup_{s>0}|h(s)-L(s)|/s^a<\infty$ *for some positive* a. *Suppose also that*

$$\sup_{s>0}\frac{ns^a}{|g_n^{-1}(\exp(-s))|^a}\to 0,\quad as\quad n\to\infty.$$

Let $X_1, X_2, \dots$ *be a sequence of i.i.d. random variables with Laplace transform* $h(s)$. *Then*

$$\sum_{i=1}^{n}\tilde{X}_i\xrightarrow{d}Y,$$

where $\tilde{X}_i$ *is a* g_n*-normalized random variable, and* Y *is a random variable with Laplace transform* $L(s)$.

Proof. Let us introduce the following distance between two probability distributions:

$$\lambda_a(U,V)=\sup_{s>0}\frac{|L_U(s)-L_V(s)|}{s^a},$$

where U and V are random variables with Laplace transforms L_U and L_V respectively. The convergence in the distance λ_a means weak convergence of distributions. We can show the convergence of $S_n=\sum_{i=1}^n\tilde{X}_i$ to Y in the following way. We remind that Y is casual stable with the normalizing function g_n, i.e. $L(s)=L^n(-\log g_n(s))$.

$$\begin{aligned}\lambda_a(S_n,Y)&=\sup_{s>0}\frac{|h^n(-\log g_n(s))-L(s)|}{s^a}=\sup_{s>0}\frac{|h^n(-\log g_n(s))-L^n(-\log g_n(s))|}{s^a}\\&\le\sup_{s>0}\frac{n|h(-\log g_n(s))-L(-\log g_n(s))|}{s^a}=\sup_{y>0}\frac{n|h(y)-L(y)|}{|g_n^{-1}(\exp(-y))|^a}\\&=\sup_{y>0}\frac{ny^a}{|g_n^{-1}(\exp(-y))|^a}\frac{|h(y)-L(y)|}{y^a}\to 0\quad\text{as}\quad n\to\infty.\end{aligned}$$

□

This result is a particular case of a result on convergence to infinitely divisible distribution in scheme of series (see, for example, Petrov (1975) or Lukacs (1970)).

For discrete distributions it might be more convenient to work with probability generating functions instead of the Laplace transforms. A probability generating function $\mathcal{P}(z)$ of a distribution with Laplace transform $L(s)$ is given as $\mathcal{P}(z)=L(-\log z)$. Let $\mathcal{P}(z)$ be a probability generating function of a casual stable distribution with normalizing probability generating function (or thinning operator) $\mathcal{Q}_n(z)$. The previous Theorem can be reformulated using the probability generating functions in the following way.

Corollary 6.5. *Suppose that* $\mathcal{R}(z)$ *is a probability generating function such that* $\sup_{z\in(0,1)}|\mathcal{R}(s)-\mathcal{P}(z)|/(1-z)^a<\infty$ *for some positive* a. *Suppose also that*

$$\sup_{z\in(0,1)}\frac{n(1-z)^a}{|1-\mathcal{Q}_n^{-1}(z)|^a}\to 0,\quad as\quad n\to\infty.$$

Let $X_1, X_2, \dots$ *be a sequence of i.i.d. random variables with probability generating function* $\mathcal{R}(z)$. *Then*

$$\sum_{i=1}^{n}\tilde{X}_i\xrightarrow{d}Y,$$

where $\tilde{X}_i$ is a normalized random variable X_i by the thinning operator $\mathcal{Q}_n$, and Y is a random variable with probability generating function $\mathcal{P}(z)$.

Let us give some examples of application of this Theorem.

6.3.1 Convergence to gamma distribution

Consider a random variable Y with gamma distribution with Laplace transform

$$L(s) = \frac{1}{(1+bs)^{\gamma}},$$

with parameters $b > 0$ and $\gamma > 0$. We showed that the random variable Y is casual stable with normalizing function

$$g_n(s) = \exp\left\{\frac{1}{b}\left(1-(1+bs)^{1/n}\right)\right\}.$$

Suppose now that $h(s)$ is a Laplace transform of a random variable X_1 such that

$$\sup_{s>0}\frac{|h(s)-L(s)|}{s^a} < \infty$$

for some $a > 1$. We have

$$\sup_{s>0}\frac{ns^a}{|g_n^{-1}(\exp(-s))|^a} = \sup_{s>0}\frac{ns^a b^a}{((1+bs)^n-1)^a} =$$
$$= \sup_{s>0}\left(\frac{n^{1/a}}{\sum_{k=1}^{n}\binom{n}{k}b^{k-1}s^{k-1}}\right)^a \leq \frac{1}{n^{a-1}} \to 0 \quad \text{as} \quad n\to\infty.$$

All conditions of Theorem 6.4 are met. So we may say that

$$\sum_{i=1}^{n}\tilde{X}_i(n) \xrightarrow{d} Y,$$

where $\tilde{X}_i(n)$ are i.i.d. random variables with Laplace transform $h(-\log g_n(s))$.

6.3.2 Convergence to negative binomial distribution

Consider geometric distribution with Laplace transform

$$L(s) = \frac{1-a}{1-ae^{-s}},$$

with parameter $a > 0$. We showed that geometric distribution is casual stable with normalizing function

$$g_n(s) = \frac{1}{a}\left[1-(1-a)^{1-1/n}\left(1-ae^{-s}\right)^{1/n}\right].$$

Let us assume that $X_1, X_2, \ldots$ are Poisson random variables with Laplace transform

$$h(s) = \exp\left\{-\lambda(1-e^{-s})\right\}.$$

Then the sum of g_n-normalized random variables has Laplace transform $h^n(-\log g_n(s))$. But

$$\begin{aligned}
h^n(-\log g_n(s)) &= \exp\left\{-\lambda n\left(1-\frac{1-(1-a)^{1-1/n}\left(1-ae^{-s}\right)^{1/n}}{a}\right)\right\} \\
&= \exp\left\{-\lambda n\frac{a-1}{a}\left[1-\exp\left(\frac{1}{n}\log\frac{1-ae^{-s}}{1-a}\right)\right]\right\} \\
&\approx \exp\left\{\lambda\frac{a-1}{a}\log\frac{1-ae^{-s}}{1-a}\right\} \quad \text{as} \quad n\to\infty \\
&= \left(\frac{1-a}{1-ae^{-s}}\right)^{\lambda\left(\frac{1}{a}-1\right)}.
\end{aligned}$$

So the g_n-normalized Poisson distribution converges to negative binomial distribution.

6.3.3 Convergence to discrete stable distribution

Let Y be a positive discrete stable random variable with probability generating function

$$\mathcal{P}(z) = \exp\left\{-\left(\frac{1-z}{1-\kappa z}\right)^{\gamma}\right\}.$$

We showed that discrete stable distribution is casual stable with the $\mathcal{G}$ thinning operator with probability generating function

$$\mathcal{Q}_n(z) = \frac{(1-p)+(p-\kappa)z}{(1-p\kappa)-\kappa(1-p)z}, \quad p = n^{-1/\gamma}.$$

Let $X_1, X_2, \ldots$ be i.i.d. random variables with probability generating function $\mathcal{R}(z)$ and assume that $\mathcal{R}$ is such that

$$\sup_{z\in(0,1)} \frac{|\mathcal{R}(z)-\mathcal{P}(z)|}{(1-z)^a}$$

for some $a > \gamma$. Furthermore we have

$$\begin{aligned}
\sup_{z\in(0,1)} \frac{n(1-z)^a}{|1-\mathcal{Q}_n^{-1}(z)|^a} &= \sup_{z\in(0,1)} n(1-z)^a\frac{|1-\kappa(p(1-z)-z)|^a}{p^a(1-\kappa)^a(1-z)^a} \\
&= \sup_{z\in(0,1)} n^{1-a/\gamma}\frac{\left|1-\kappa(n^{-1/\gamma}(1-z)-z)\right|^a}{(1-\kappa)^a} \\
&\leq n^{1-a/\gamma}\left(\frac{1-\kappa n^{-1/\gamma}}{1-\kappa}\right)^a \to 0 \quad \text{as} \quad n\to\infty
\end{aligned}$$

for $a > \gamma$. So, based on the convergence Theorem 6.4 and its Corollary 6.5, we may say that

$$S_n = \sum_{i=1}^{n} \tilde{X}_i \xrightarrow{d} Y.$$

7. ν-discrete stable distributions

In Chapters 4 and 5 we considered discrete stability – stability in sense of summation of a deterministic number of discrete random variables. We have shown that discrete stable distributions are discrete analogies of classical stable distributions. In this Chapter we will study generalization of this concept when instead of the sum of a deterministic number of discrete random variables we consider a random number of summands. The concept of stability under random summation was introduced independently in Klebanov and Rachev (1996) (under the name of ν-stability) and Bunge (1996).

Let $X_1, X_2, \ldots$ denote a sequence of i.i.d. discrete random variables. Let $\{\nu_p, p \in \Delta\}$ be a family of discrete random variable with values in $\mathbb{N}$, independent of $\{X_j, j \in \mathbb{N}\}$. Further we will assume that $\mathrm{E}\nu_p$ exists and $\mathrm{E}\nu_p = 1/p$, for all $p \in \Delta$. Let us denote by $\mathcal{P}_p$ the probability generating function of ν_p and by $\mathfrak{P}$ the semigroup with operation of superposition generated by the family $\{\mathcal{P}_p(z), p \in \Delta\}$.

Our aim is to obtain some form of stability property for the random sum

$$\sum_{j=1}^{\nu_p} X_j.$$

We start by reminding that there is a one-on-one mapping between infinitely divisible distributions and ν-infinitely divisible distributions (Theorem 2.25). As discrete stable distributions in the first sense are all infinitely divisible distributions, we define ν-discrete stable distribution in a similar way as the ν-stable distributions (Definition 2.26).

Definition 7.1. A function $g(t)$ is called a ν-discrete stable characteristic function if it admits representation

$$g(t) = \varphi(-\log f(t)), \tag{7.1}$$

in which φ is a standard solution of the system of equations

$$\begin{aligned} \varphi(t) &= \mathcal{P}_p(\varphi(pt)), \\ \varphi(0) &= 1, \quad \varphi'(0) = -1, \end{aligned} \tag{7.2}$$

and $f(t)$ is the characteristic function of a discrete stable distribution in the first sense.

Distribution or random variable with ν-discrete stable characteristic function will be called ν-discrete stable distribution and ν-discrete stable random variable respectively.

It was shown in Gnedenko and Korolev (1996) that the Poincaré equation (7.2) has a unique solution if and only if the semigroup $\mathfrak{P}$ is commutative. We give here two examples of commutative summation schemes, already introduced in Chapter 2, for which (7.2) has a unique solution.

Geometric summation scheme A typical example of a commutative semigroup $\mathfrak{P}$ is a semigroup generated by geometric distribution leading to a geometric summation scheme. Let ν_p be a geometric random variable with parameter p, that means

$$\mathbb{P}(\nu_p = k) = p(1-p)^{k-1}, \quad k \geq 1.$$

Probability generating function of a geometric random variable is given by the following formula
$$\mathcal{P}_p(z) = \frac{pz}{1-(1-p)z}.$$
We can easily see that the semigroup generated by $\{\mathcal{P}_p, p \in \Delta\}$ is commutative, as
$$\mathcal{P}_{p_1} \circ \mathcal{P}_{p_2}(z) = \frac{p_1 p_2 z}{1-(1-p_1 p_2)z} = \mathcal{P}_{p_2} \circ \mathcal{P}_{p_1}(z).$$
The Poincaré equation (7.2) takes the following form
$$\varphi(x) = \frac{p\varphi(px)}{1-(1-p)\varphi(px)}, \quad p \in \Delta,$$
and it has a unique standard solution in the form
$$\varphi(x) = \frac{1}{1+x}. \tag{7.3}$$

Chebyshev summation scheme Another example of a commutative semigroup $\mathfrak{P}$ is a semigroup generated by distributions with probability generating functions of the form
$$\mathcal{P}_p(z) = \frac{1}{T_{1/\sqrt{p}}\left(\frac{1}{z}\right)},$$
for $p \in \Delta = \{1/n^2, n \in \mathbb{N}\}$, where $T_n(z) = \cos(n \arccos z)$ is the Chebyshev polynomial. Klebanov et al. (2012) showed that $\mathcal{P}_p(z)$ is truly a probability generating function. We can easily verify, that $\{\mathcal{P}_p, p \in \Delta\}$ generates a commutative semigroup, as
$$\mathcal{P}_{p_1} \circ \mathcal{P}_{p_2}(z) = \frac{1}{T_{n_1}\left(T_{n_2}\left(\frac{1}{z}\right)\right)} = \frac{1}{T_{n_1 n_2}\left(\frac{1}{z}\right)}.$$
The Poincaré equation (7.2) takes the following form
$$\varphi(x) = \frac{1}{T_{1/\sqrt{p}}\left(\frac{1}{\varphi(px)}\right)}, \quad p \in \Delta.$$
The following function satisfies this equation
$$\varphi(x) = \frac{1}{\cosh\left(\sqrt{2x}\right)}. \tag{7.4}$$

7.1 ν-positive discrete stable random variables

In this Section we will study ν-positive discrete stable random variables. By Definition 7.1, a random variable is ν-positive discrete stable, if its characteristic function $g(t)$ takes form
$$g(t) = \varphi(-\log f(t)),$$

where f is the characteristic function of discrete stable distribution,

$$f(t) = \exp\left\{-\lambda\left(\frac{1-e^{\mathrm{i}t}}{1-\kappa e^{\mathrm{i}t}}\right)^{\gamma}\right\}.$$

We can show that for ν-positive discrete stable random variables an analogy of the stability property for sums of random number of random variables holds.

Theorem 7.2. *Let X be ν-positive discrete stable random variable with probability generating function*

$$\mathcal{P}(z) = \varphi\left[\lambda\left(\frac{1-z}{1-\kappa z}\right)^{\gamma}\right].$$

Then

$$X_1 \stackrel{d}{=} \sum_{i=1}^{\nu_p} \tilde{X}_i\left(p^{1/\gamma}\right), \quad \text{where} \quad \tilde{X}_i\left(p^{1/\gamma}\right) = \sum_{j=1}^{X_i} \varepsilon_j^{(i)}\left(p^{1/\gamma}\right), \tag{7.5}$$

a $\varepsilon_j^{(i)}(a)$ are i.i.d. $\mathcal{G}(a,\kappa)$ random variables with probability generating function

$$\mathcal{Q}(z) = \frac{(1-a)+(a-\kappa)}{(1-\kappa a)-\kappa(1-a)z}, \quad a \ = p^{1/\gamma},$$

$\{X_j, j\in\mathbb{N}\}$ is a sequence of i.i.d. copies of X_1, independent of $\{\nu_p, p\in\Delta\}$.

Proof. Let us compute the probability generating function of $Y = \sum_{i=1}^{\nu_p} \tilde{X}_i\left(p^{1/\gamma}\right)$ denoted by $\mathcal{P}_Y(z)$. We have $\mathcal{P}_Y(z) = \mathcal{P}_p(\mathcal{P}(\mathcal{Q}(z)))$. But

$$\begin{aligned}
\mathcal{P}(\mathcal{Q}(z)) &= \varphi\left[\lambda\left(\frac{1-\mathcal{Q}(z)}{1-\kappa\mathcal{Q}(z)}\right)^{\gamma}\right] \\
&= \varphi\left[\lambda\left(p^{1/\gamma}\frac{1-z}{1-\kappa z}\right)^{\gamma}\right] \\
&= \varphi\left[p\lambda\left(\frac{1-z}{1-\kappa z}\right)^{\gamma}\right].
\end{aligned}$$

Moreover φ is a solution of (7.2), therefore

$$\begin{aligned}
\mathcal{P}_p\left(\mathcal{P}(\mathcal{Q}(z))\right) &= \mathcal{P}_p\left(\varphi\left[p\lambda\left(\frac{1-z}{1-\kappa z}\right)^{\gamma}\right]\right) \\
&= \varphi\left[\lambda\left(\frac{1-z}{1-\kappa z}\right)^{\gamma}\right].
\end{aligned}$$

So we have shown that $\mathcal{P}_Y(z) = \mathcal{P}(z)$, therefore X_1 has the same distribution as Y. □

Example 7.3 (Geometric positive discrete stable random variables)**.** Let us consider the geometric summation scheme. Then the standard solution of Poincaré equation takes form

$$\varphi(x) = \frac{1}{1+x}.$$

The geometric positive discrete stable characteristic function thus takes form

$$g(t) = \frac{1}{1+\lambda\left(\frac{1-e^{\mathrm{i}t}}{1-\kappa e^{\mathrm{i}t}}\right)^{\gamma}}.$$

Example 7.4 (Chebyshev positive discrete stable random variables)**.** Let us consider the Chebyshev summation scheme. Then the standard solution of Poincaré equation takes form

$$\varphi(x) = \frac{1}{\cosh\left(\sqrt{2x}\right)}.$$

The Chebyshev positive discrete stable characteristic function thus takes form

$$g(t) = \frac{1}{\cosh\left(\sqrt{2\lambda\left(\frac{1-e^{it}}{1-\kappa e^{it}}\right)^{\gamma}}\right)}.$$

7.2 Properties of geometric positive discrete stable random variables

In this Section we will study more into details a special case of ν-positive discrete stable random variables, where ν_p is geometrically distributed, i.e.

$$\mathbb{P}(\nu_p = k) = p(1-p)^{k-1}, \quad kgeq1.$$

7.2.1 Geometric Poisson random variables

The probability generating function of geometric PDS(1,λ, κ) distribution is

$$\mathcal{P}(z) = \frac{1}{1+\lambda\frac{1-z}{1-\kappa z}}.$$

The distribution with $\kappa = 0$ is called geometric Poisson as PDS(γ, λ, κ) distribution with $\gamma = 1$ and $\kappa = 0$ is Poisson distribution. The geometric Poisson distribution coincides with geometric distribution with parameter $1/(1+\lambda)$. Therefore, if X is a geometric Poisson random variable with probability generating function $\mathcal{P}(z) = \left(1+\lambda(1-z)\right)^{-1}$ we have

$$\mathbb{P}(X = k) = \left(\frac{\lambda}{1+\lambda}\right)^{k}\frac{1}{1+\lambda}.$$

For $\kappa > 0$ the probabilities $\mathbb{P}(X = k)$ are also in closed form.

Proposition 7.5. *Let X be a geometric PDS(1,λ, κ) random variable with probability generating function $\mathcal{P}(z) = \left(1+\lambda\frac{1-z}{1-\kappa z}\right)^{-1}$. Then*

$$\mathbb{P}(X = k) = (1-\kappa)\lambda\frac{(\kappa+\lambda)^{k-1}}{(1+\lambda)^{k+1}}.$$

In the next Theorem we study the limit distribution of random variables $X^a = aX$, with X being geometric Poisson random variable, $a > 0$ and $a \to 0$. In this case the random variable X^a takes values in $a\mathbb{N} = \{0, a, 2a, \dots\}$ and the limit distribution will be therefore a continuous analogy of the geometric Poisson distribution. We show that the limit distribution of geometric Poisson random variable with $\kappa = 0$ is exponential distribution.

Theorem 7.6. *Let X be a geometric Poisson random variable with characteristic function*

$$f(t) = \frac{1}{1+\lambda(1-e^{\mathrm{i}t})}.$$

Let us consider a random variable $X^a = aX$, with $a > 0$ and characteristic function denoted by $f^a(t)$. Let $\lambda = \frac{\sigma}{a}$. Then

$$f^a(t) = \frac{1}{1+\lambda(1-e^{\mathrm{i}a\ t})} \longrightarrow \varphi(t) = \frac{1}{1-\sigma \mathrm{i}\, t}, \quad as \quad a \to 0.$$

Proof. The limit of the characteristic function $f^a(t)$ can be computed in a straightforward way. We make an approximation $1 - e^{\mathrm{i}at} \approx -\mathrm{i}at$, as $a \to 0$. Then

$$\varphi(t) = \lim_{a\to 0}\left(1+\lambda(1-e^{\mathrm{i}at})\right)^{-1} = \left(1-\sigma \mathrm{i}\, t\right)^{-1}.$$

□

7.2.2 Geometric positive discrete stable random variables

A geometric positive discrete stable random variable X has probability generating function taking the following form

$$\mathcal{P}_X(z) = \frac{1}{1+\lambda\left(\frac{1-z}{1-\kappa z}\right)^{\gamma}}, \quad \gamma \in (0,1],\ \lambda > 0,\ \kappa \in [0,1).$$

We will denote positive geometric discrete stable distribution and random variable with parameters γ, λ and κ by geo-PDS$(\gamma, \lambda, \kappa)$. In the case when κ is omitted we will understand $\kappa = 0$.

This distribution with $\kappa = 0$ (corresponding to the binomial thinning operator) was introduced in Pillai and Jayakumar (1995) as discrete Mittag-Leffler distribution and generalized in Devroye (1993) as discrete Linnik distribution. Pillai and Jayakumar (1995) and Devroye (1993) showed several interesting properties of the distribution, namely:

- geo-PDS(γ, λ) distributions is geometrically infinitely divisible and therefore infinitiely divisible.
- geo-PDS(γ, λ) distribution belongs to the domain of normal attraction of γ-stable distribution.
- Let $X \sim$ geo-PDS(γ, λ). The limit distribution of a random variable $X^a = aX$, with $a = \frac{1}{n}, n \in \mathbb{N}$ and $a \to 0$, is Mittag-Leffler distribution.
- geo-PDS$(\gamma, 1)$ distribution satisfies the following distributional identity:

  $$\text{geo-PDS}(\gamma, 1) \stackrel{d}{=} \text{Poisson}(V^{1/\gamma} S_\gamma),$$

  where V is an exponential random variable with mean 1 and S_γ is γ-stable random variable with Laplace transform $\exp\{-u^\gamma\}$.

We can obtain similar results for the generalized geometric positive discrete stable distribution geo-PDS$(\gamma, \lambda, \kappa)$ with $\kappa > 0$.

Limit distribution In the following theorem we show that the limit distribution of geo-PDS distribution is geometric stable distribution.

Theorem 7.7. *Let X be geometric positive discrete stable random variable with characteristic function*

$$f(t) = \left(1 + \lambda \left(\frac{1 - e^{\mathrm{i}t}}{1 - \kappa e^{\mathrm{i}t}}\right)^{\gamma}\right)^{-1}.$$

Let us consider a random variable $X^a = aX$, with $a > 0$ and characteristic function denoted by $f^a(t)$. Let $\lambda = \frac{\sigma}{a^\gamma}$. Then

$$\begin{aligned} f^a(t) &= \left(1 + \lambda \left(\frac{1 - e^{\mathrm{i}at}}{1 - \kappa e^{\mathrm{i}at}}\right)^{\gamma}\right)^{-1} \\ &\longrightarrow \varphi(t) = \left(1 + \frac{\sigma}{(1-\kappa)^\gamma} \cos(\pi\gamma/2)|t|^\gamma \left(1 - \mathrm{i}\,\mathrm{sign}(t)\tan(\pi\gamma/2)\right)\right)^{-1}, \quad \textit{as} \quad a \to 0. \end{aligned}$$

Proof. The proof is almost identical to the proof of Theorem 5.11. □

Characterizations We show that geo-PDS random variable has the same distribution as PDS random variable with random intensity.

Theorem 7.8. *Let V be an exponential random variable with mean 1 and let S_γ be γ-stable random variable with Laplace transform* $\exp\{-u^\gamma\}$. *Then*

$$\textit{geo}\text{-PDS}(\gamma, \lambda, \kappa) \overset{d}{=} \mathrm{PDS}(1, \lambda^{1/\gamma} V^{1/\gamma} S_\gamma, \kappa).$$

Proof. The probability generating function of $X \sim \mathrm{PDS}(1, \lambda, \kappa)$ is equal to

$$\mathrm{E}z^X = \exp\left\{-\lambda\left(\frac{1-z}{1-\kappa z}\right)\right\}.$$

Therefore we compute the probability generating function of $X \sim \mathrm{PDS}(1, \lambda^{1/\gamma} V^{1/\gamma} S_\gamma, \kappa)$ as

$$\begin{aligned} \mathrm{E}z^X &= \mathrm{E}\exp\left\{-\lambda^{1/\gamma} V^{1/\gamma} S_\gamma \left(\frac{1-z}{1-\kappa z}\right)\right\} \\ &= \mathrm{E}\exp\left\{-\lambda V \left(\frac{1-z}{1-\kappa z}\right)^{\gamma}\right\}. \end{aligned}$$

The Laplace transform $\mathrm{E}e^{-uV}$ of exponential random variable V with mean 1 is $(1+u)^{-1}$. Hence

$$Ez^X = \left(1 + \lambda\left(\frac{1-z}{1-\kappa z}\right)^{\gamma}\right)^{-1},$$

and this is the probability generating function of geo-PDS$(\gamma, \lambda, \kappa)$ random variable. □

7.3 ν-symmetric discrete stable distribution

By Definition 7.1, a random variable is ν-symmetric discrete stable, if its characteristic function $g(t)$ takes form

$$g(t) = \varphi(-\log f(t)),$$

where f is the characteristic function of symmetric discrete stable distribution,

$$f(t) = \exp\left\{-\lambda\left(1-(1-\kappa)\frac{\cos t-\kappa}{\kappa^2-2\kappa\cos t+1}\right)^{\gamma}\right\}.$$

Part II

Applications of discrete stable distributions

8. Rating of scientific work

It is very often to base the rating of scientific work on the number of citations of corresponding paper, author or journal, in which the paper was published (so-called impact factor). The problem of rating scientific work based on the number of citations is very actual as it seems that the number of citations does not fully mimic the qualities of the paper and therefore citation index is not a good method for ranking of scientists. Price (1965) noted that the number of citations has power tails based on a large dataset of papers and their citations. Several models that capture the power tails and try to explain them have been proposed in the literature. Simkin and Roychowdhury (2007) argue that the process of citing is very random and authors "pick" several papers and cite them in their work or they copy citations from previous publications in their field. This leads to an apparent disproportion in the citations, where recent papers get more citations and the number of citations decreases with age of the paper. This does not hold true for so called "sleeping beauties in science" – old works that are not cited for a long time and later one author cites them and create a subsequent increase of citations. Peterson et al. (2010) proposed a model where papers are either cited directly or indirectly through a list of references in a newer paper. They show that the indirect mechanism of citing leads to power tails. It is apparent that the works that get the most citations ("sleeping beauties" or the indirectly cited works) are the essential publications in the field.

In this Chapter we describe a model, in which the number of citations in one field of science follows discrete stable distribution with the modified geometric thinning operator. The first approach was proposed in Klebanov and Slámová (2014) and assumes randomness in the publishing and citation processes. We will consider also a limit approach to this model that is based on Theorem 6.4 about convergence to a casual stable distribution.

8.1 The model

At first let us consider the simplest model of paper publication with one author. We consider only the case when there is at least one publication at the beginning (in opposite case there will be no citations at all). Let us denote by q the probability that a paper and all following papers of the author will be rejected. Then the probability to have exactly k published papers is given as the probability of $k-1$ acceptances, given as $(1-q)^{k-1}$ (we assume one paper was already published) and one rejection, with probability q. Hence the probability is $q(1-q)^{k-1}$. In other words, the number of published papers has geometric distribution with probability generating function

$$\frac{qz}{1-(1-q)z}.$$

Suppose now that every published paper generates some citations. The probability of the paper to be cited depends on the number of its previous citations (Price (1976) called this cumulative advantage process). Consider a paper having $k-1$ citations and let the probability that this paper will not be cited again be p/k, where p is the probability is the probability that the paper will not be cited at all. Therefore the probability that the paper will be cited exactly k times is

$$\frac{p}{k}\prod_{j=1}^{k-1}\left(1-\frac{p}{j}\right) = p(1-p)_{k-1}/k!,$$

where $(a)_n = a(a+1)\cdots(a+n-1)$ is the Pochhammer symbol. Therefore, the probability generating function of the distribution of citations of one paper is

$$\sum_{k=1}^{\infty}\frac{p(1-p)_{k-1}}{k!}z^k = 1-(1-z)^p.$$

This is well-known Sibuya distribution with parameter p. So the number of citations of one publication has Sibuya distribution with parameter p. The probability generating function of the distribution of the number of citations of all papers coming from one author is obtained as a superposition of the probability generating functions of the Sibuya distribution of the number of citations and of the geometric distribution of the number of published papers; that is

$$1-\left(1-\frac{qz}{1-(1-q)z}\right)^p, \tag{8.1}$$

with parameters $q \in (0,1]$ and $p \in (0,1]$. It is easy to see that the corresponding distribution has a heavy tail (its limit behaviour is of order $1/k^p$, as $k \to \infty$).

Suppose now that we are interested in the distribution of the number of citations in some field of science. It is natural to assume that the number of scientists having publications in this field has Poisson distribution with parameter λ. Then the probability generating function of the number of all publications in the field is a superposition of the Poissonian probability generating function with probability generating function (8.1); that is

$$\exp\left\{-\lambda\left(\frac{1-z}{1-(1-q)z}\right)^p\right\}.$$

This is the probability generating function of discrete stable distribution in the first sense with the modified geometric thinning operator with parameters $\gamma = p \in (0,1]$, $\lambda > 0$ and $\kappa = 1-q$.

We can easily see that the distribution does not have finite first moment for $p < 1$. It has mode at zero and finite median. So, from any empirical data we will see that the empirical mean is much larger than the empirical median. Also, many citations will belong to a (relatively) small number of publications, while the main part of publications will have small number of citations. This big difference between scientist is explained in our model just by random nature of the publication and citation processes. Therefore, the ranking of scientist, scientific institutions or journals may not be based on the citation number. Such ranking will often produce principal misunderstandings of what is essential in science and lead also to random mistakes.

8.2 Limit approach

We can consider a limit approach based on the convergence Theorem 6.4. Let X denote the number of published papers of an author (it is obvious that X takes non-negative integer values) and let us denote by $\mathcal{R}$ the probability generating function of X. Let us assume that each paper is not cited at all with some positive probability $q \in (0,1)$ (we exclude the degenerate case when the probability of not being cited is equal to 1) or it has a random number of citations with geometric distribution with parameter $b \in (0,1]$. Thus the number of citations is a mixture of degenerate and geometric distributions, and its probability generating function thus takes form

$$\mathcal{Q}(z) = q + (1-q)\frac{bz}{1-(1-b)z}.$$

We can reparametrize this distribution, by putting

$$q = \frac{1-p}{1-p\kappa} \quad \text{and} \quad b = \frac{1-\kappa}{1-p\kappa}$$

with $p \in (0,1)$ and $\kappa \in [0,1)$. Then the probability generating function takes form

$$\mathcal{Q}(z) = \frac{(1-p)+(p-\kappa)z}{(1-\kappa p)-\kappa(1-p)z}.$$

This is the probability generating function of the modified geometric thinning operator from Subsection 4.1.2. The number of citations of all papers coming from one author is thus a random variable given as

$$\tilde{X} = \sum_{i=1}^{X} \epsilon_i,$$

where ϵ_i is the number of citations of the i-th paper. Its probability generating function takes form $\mathcal{R}(\mathcal{Q}(z))$.

Let us now assume that we have n authors with $n \to \infty$ and let $p = n^{-1/\gamma}$. We showed in Section 6.3, using the Theorem 6.4 about convergence to causal stable distribution, that the number of citations coming from n authors,

$$S_n = \sum_{i=1}^{n} \tilde{X}_i,$$

converges in distribution to a random variable with discrete stable distribution with probability generating function

$$\mathcal{P}(z) = \exp\left\{-\lambda\left(\frac{1-z}{1-\kappa z}\right)^{\gamma}\right\}.$$

This result coincides with the model from previous Section.

9. Statistical inference for discrete stable distributions

Discrete stable distributions studied in Chapter 5 form a discrete generalization of stable distributions and as such have similar properties as their continuous counterparts. They are expressed through a probability generating function and the probability mass function does not generally have a closed form formula and no moments exist. This fact inhibits the use of classical statistical methods of estimation such as maximum likelihood or method of moments.

Kagan (1976) introduced an analogue of the maximum likelihood method by studying an "approximation" of the likelihood function on a finite-dimensional Hilbert space $\mathcal{H}$. Instead of the likelihood function a function called "informant" as an operator in the Hilbert space $\mathcal{H}$ is introduced. Kagan (1976) showed that the behaviour of the resulting estimator is analogous to that of classical maximum likelihood estimator and many properties such as consistency and asymptotic normality are conserved.

In this Chapter we adapt and optimize this method for the case of discrete stable distributions. We compare the results of this method with the $k-L$ procedure that was introduced by Feuerverger and McDunnough (1981) that uses k fixed points to fit the empirical characteristic function with the theoretical one.

The approximated maximum likelihood (AML further on) method is described in Section 1, where we also summarize the known results about the properties of the estimator. In Section 2 we adapt the AML method to the case of discrete stable distributions with binomial thinning operator and in Section 3 we give an overview of the results of a simulation study and show the quality of the AML estimator.

This Chapter contains results from Slámová and Klebanov (2014b).

9.1 Approximated maximum likelihood method

In this Section we describe a method of estimation that was introduced by Kagan (1976), and further extended in (Gerlein and Kagan, 1979). The proposed method is a very general approach that can be used in cases when the distribution is not defined through density function or the probability mass function and instead only some functionals of the distribution are given as functions of parameters (eg. characteristic function, probability generating function etc.).

Let $\{P_\theta, \theta \in \Theta\}$ be a family of probability distributions on a measurable space $(\mathcal{X}, \mathcal{M})$, where the parametric space $\Theta \subset \mathbb{R}^d$, and $X \sim P_\theta$. In the maximum likelihood estimation one assumes the existence of a density function $p(x,\theta)$ and of a function

$$\mathbf{J}(x,\theta) = \left(\frac{\frac{\partial p}{\partial \theta_i}(x,\theta)}{p(x,\theta)} \right)_{i=1,\dots,d}.$$

The maximum likelihood estimator θ^* of the parameter θ, given a set of n observations $x_1, \dots, x_n$, is a solution of $\sum_{j=1}^n \mathbf{J}(x_j,\theta) = \mathbf{0}$. The Fisher information matrix is given as

$$\mathbf{I}(\theta) = \mathrm{E}_\theta(\mathbf{J}(X,\theta)\mathbf{J}(X,\theta)^T).$$

However, if the density does not exist or it is not known in a closed form, this method cannot be used.

Consider a linear space $\mathcal{L}_k$ generated by a set of complex valued functions $\{\varphi_0(x), \dots, \varphi_k(x)\}$, $\varphi_0 \equiv 1$ on space $\mathcal{X}$ with inner product denoted by $(\cdot,\cdot)_\theta$ and defined by $\big(\varphi(X), \psi(X)\big)_\theta = \mathrm{E}_\theta \left[\varphi(X)\overline{\psi}(X)\right]$, where $\overline{\psi}$ denotes complex conjugate of ψ. The functions $\varphi_i, i = 1, \dots, k$ are such that $\mathrm{E}_\theta \varphi_i(X)\overline{\varphi_i}(X) < \infty, \theta \in \Theta, i = 1, \dots, k$. We assume we know functionals of our distribution, namely, for $i, j = 0, \dots, k$

$$\begin{aligned} \pi_i(\theta) &= \big(1, \varphi_i(X)\big)_\theta = \mathrm{E}_\theta \overline{\varphi_i}(X), \\ \pi_{ij}(\theta) &= \big(\varphi_i(X), \varphi_j(X)\big)_\theta = \mathrm{E}_\theta \varphi_i(X)\overline{\varphi_j}(X), \end{aligned}$$

are known as functions of the parameter θ. Further, we will use the following notation: $\varphi(x) = (\varphi_i(x), i = 0, \dots, k)$, $\pi(\theta) = (\pi_i(\theta), i = 0, \dots, k)$ and $\mathbf{\Pi}(\theta) = (\pi_{ij}(\theta), i, j = 0, \dots, k)$.

The method is an analogue of the maximum likelihood method in the sense that it approximates the undefined function $\mathbf{J}(x, \theta)$ by its projection onto the linear space $\mathcal{L}_k$. Gerlein and Kagan (1979) call this method $\mathcal{H}$-method of maximum likelihood, where $\mathcal{H}$ is the Hilbert space $\mathcal{L}_k$.

The projection of $\mathbf{J}(x, \theta)$ will be denoted $\widehat{\mathbf{J}}(x, \theta)$ and as part of the linear space $\mathcal{L}_k$ takes the following form

$$\widehat{\mathbf{J}}(x, \theta) = \mathbf{c}^T(\theta)\varphi(x),$$

where $\mathbf{c}(\theta) = (c_{ij}(\theta), i = 0, \dots, k; j = 1, \dots, d)$.

We compute the approximation of the Fisher information matrix as

$$\begin{aligned} \widehat{\mathbf{I}}(\theta) = \|\widehat{\mathbf{J}}(X, \theta)\|^2 &= \left(\widehat{\mathbf{J}}(X, \theta), \widehat{\mathbf{J}}(X, \theta)\right)_\theta = \mathrm{E}_\theta \left[\widehat{\mathbf{J}}(X, \theta)\widehat{\mathbf{J}}^*(X, \theta)\right] \\ &= \mathbf{c}^T(\theta)\mathrm{E}_\theta \left[\varphi(X)\varphi^*(X)\right] \overline{\mathbf{c}}(\theta) = \mathbf{c}^T(\theta)\mathbf{\Pi}(\theta)\overline{\mathbf{c}}(\theta). \end{aligned} \tag{9.1}$$

Since $\widehat{\mathbf{J}}$ is a projection of $\mathbf{J}$ onto $\mathcal{L}_k$, the following orthogonality condition has to hold for all $i = 1, \dots, d$ and $j = 0, \dots, k$:

$$\left(J_i(X, \theta) - \widehat{J}_i(X, \theta), \varphi_j(X)\right)_\theta = \mathrm{E}_\theta \left[\left(J_i(X, \theta) - \widehat{J}_i(X, \theta)\right) \overline{\varphi_j}(X)\right] = 0. \tag{9.2}$$

From this set of equalities the form of the unknown matrix $\mathbf{c}(\theta)$ retrieves, as is shown in the following lemma.

Lemma 9.1. *If an inverse of the matrix $\mathbf{\Pi}(\theta)$ exists then $\mathbf{c}(\theta) = \mathbf{\Pi}^{-1}(\theta)\nabla\pi(\theta)$.*

Proof. It follows from the orthogonality condition that for all $i = 1, \dots, d$ and $j = 0, \dots, k$

$$\mathrm{E}_\theta \left[J_i(X, \theta)\overline{\varphi_j}(X)\right] = \mathrm{E}_\theta \left[\widehat{J}_i(X, \theta)\overline{\varphi_j}(X)\right].$$

The left-hand side equals

$$\mathrm{E}_\theta \left[J_i(X, \theta)\overline{\varphi_j}(X)\right] = \int \frac{\frac{\partial p(x,\theta)}{\partial \theta_i}}{p(x, \theta)} \overline{\varphi_j}(x)p(x, \theta)\mathrm{d}x = \frac{\partial}{\partial \theta_i} \int \overline{\varphi_j}(x)p(x, \theta)\mathrm{d}x = \frac{\partial \pi_j(\theta)}{\partial \theta_i}.$$

The right hand side can be rewritten as

$$\mathrm{E}_\theta \left[\widehat{J}_i(X, \theta)\overline{\varphi_j}(X)\right] = \mathrm{E}_\theta \left[\mathbf{c}_{\cdot i}^T(\theta)\varphi(X)\overline{\varphi_j}(X)\right] = \sum_{m=0}^{k} c_{mi}(\theta)\pi_{mj}(\theta).$$

Put together, we obtain

$$\frac{\partial \pi_j(\theta)}{\partial \theta_i} = \sum_{m=0}^{k} c_{mi}(\theta)\pi_{mj}(\theta), \quad i = 1, \ldots, d; j = 0, \ldots, k.$$

If we use a matrix notation, $\nabla\pi(\theta) = \mathbf{\Pi}(\theta)\mathbf{c}(\theta)$. Hence if the inverse of $\mathbf{\Pi}(\theta)$ exists, then the matrix $\mathbf{c}(\theta) = \mathbf{\Pi}^{-1}(\theta)\nabla\pi(\theta)$. □

The maximum likelihood estimator θ^* of the parameter θ is obtained as the solution of $\sum_m \mathbf{J}(x_m, \theta) = \mathbf{0}$. The AML estimator $\widehat{\theta}^*$ of the parameter θ is obtained in a very similar way; instead of $\mathbf{J}$ we consider its approximation $\widehat{\mathbf{J}}$. Hence we are trying to find a solution of the set of equations

$$\sum_{m=1}^{n} \widehat{\mathbf{J}}(x_m, \theta) = \mathbf{0}, \tag{9.3}$$

or equivalently

$$\sum_{m=1}^{n}\sum_{i=0}^{k} c_{ij}(\theta)\varphi_i(x_m) = 0, \quad j = 1, \ldots, d.$$

The following properties of the AML estimator were shown in (Kagan, 1976).

Theorem 9.2. *The AML estimator $\widehat{\theta}^*$, that is a solution of* (9.3)*, is consistent and asymptotically normal*

$$\sqrt{n}\left(\widehat{\theta}^* - \theta\right) \to \mathcal{N}(0, \widehat{\mathbf{I}}^{-1}(\theta)).$$

Remark 9.3. We can see that the AML estimator is not asymptotically efficient in the classical sense. However, the approximated Fisher information matrix converges to the theoretical Fisher information matrix as k goes to infinity: $\lim_{k\to\infty} \widehat{\mathbf{I}}(\theta) = \mathbf{I}(\theta)$. This follows from the monotonicity property of the approximated Fisher information that was shown in Kagan (1976). The Theorem 9.2 shows that with k going to infinity, we can achieve very high asymptotic efficiency but for the price of computation speed as the computational complexity grows with higher values of k.

9.2 Estimating parameters of discrete stable distributions

The method described in previous Section is very general and can be used for many distributions where classical approaches fail due to the lack of a closed form of density function or probability mass function. The AML method is applicable if we know only some functionals of the distribution such as characteristic function or probability generating function, which is the case of stable distributions or discrete stable distributions. In this Section we apply this method to the case of discrete stable distributions with binomial thinning operator and describe the algorithm of estimation. We will consider the positive discrete stable distribution with probability generating function

$$\mathcal{P}(z) = \exp\left\{-\lambda(1-z)^\gamma\right\}$$

and symmetric discrete stable distribution with probability generating function

$$\mathcal{P}(z) = \exp\left\{-\lambda\left(1 - \frac{1}{2}\left(z + \frac{1}{z}\right)\right)^\gamma\right\}.$$

In both cases the parametric space is $\Theta = (0, 1] \times (0, \infty)$.

The first step is to choose the functionals $\pi_i(\theta)$ and $\pi_{i,j}(\theta)$ and therefore the set of functions $\mathcal{L}_k$. For discrete stable distributions a natural choice is the probability generating function $\mathcal{P}(z)$. We define $\varphi_i(x) = z_i^x$ with $z_i \in \mathbb{C}$, for $i = 1, \ldots, k$, and $z_0 = 1$. Then $\pi_i(\theta) = \mathrm{E}_\theta(\overline{z_i}^X) = \mathcal{P}(\overline{z_i})$. The choice of z_is can be done in an optimal way by maximizing the determinant of the approximated Fisher information matrix and thus obtaining optimal estimate in sense of efficiency. We thus need to solve the following optimization problem

$$\max_{\mathbf{z} \in \mathcal{A}} \quad |\widehat{\mathbf{I}}(\theta)|, \tag{9.4}$$

where $\mathbf{z} = (z_1, \ldots, z_k)^T$. The set $\mathcal{A}$ is the domain of definition of the approximated Fisher information matrix $\widehat{\mathbf{I}}(\theta)$. The approximated Fisher information matrix reduces from (9.1) to (9.5) thanks to Lemma 9.1:

$$\widehat{\mathbf{I}}(\theta) = \nabla\pi(\theta)^T \, \overline{\Pi(\theta)^{-1}} \, \overline{\nabla\pi(\theta)} \tag{9.5}$$

The probability generating function of PDS distribution is defined for $|z| \leq 1$. The set $\mathcal{A}$ is therefore given by $\mathcal{A} = \{\mathbf{z} \in \mathbb{C}^k : |z_i| \leq 1, i = 1, \ldots, k\}$. It turns out that the optimal solution $\mathbf{z}$ of (9.4) is such that $\mathrm{Re}(z_i) \in (0, 1]$ and $\mathrm{Im}(z_i) = 0$ for all $i = 1, \ldots, k$.

The probability generating function of SDS distribution is defined for $\left|z + \frac{1}{z}\right| \leq 2$. Therefore

$$\begin{aligned} \mathcal{A} =& \left\{\mathbf{z} \in \mathbb{C}^k : \left|z_i + \frac{1}{z_i}\right| \leq 2, \ \left|z_i \bar{z}_j + \frac{1}{z_i \bar{z}_j}\right| \leq 2, \ i, j = 1, \ldots, k\right\} \\ =& \{\mathbf{z} \in \mathbb{C}^k : |z_i| = 1, \ i = 1, \ldots, k\}. \end{aligned}$$

The case of DS distribution leads also to $\mathcal{A} = \{\mathbf{z} \in \mathbb{C}^k : |z_i| = 1, \ i = 1, \ldots, k\}$.

Speed of convergence of $\widehat{\mathbf{I}}(\theta)$ The goal is to have a quick estimation method with high asymptotic efficiency; however, these two properties are as usually in contradiction. With the optimal choice of the linear space $\mathcal{L}_k = \{1, z_1^x, \ldots, z_k^x\}$ we can achieve, given a fixed size of the linear space k, the highest possible efficiency. The speed of convergence of the optimal approximated Fisher information matrix for different values of the parameter γ for PDS distribution is displayed in Figure 9.1 and for SDS distribution in Figure 9.2. We can see that for PDS distribution the size $k = 5$ is enough, for SDS distribution depending on the unknown value of γ the size k to achieve high asymptotic efficiency might be significantly higher.

For the solution of the optimization problem (9.4) one need to know the unknown parameter θ. The AML estimation is thus done sequentially in four steps as described in Algorithm 9.4.

Algorithm 9.4.

Step 1: Choose $k \in \mathbb{N}$ and $\mathbf{z} = (z_1, \ldots, z_k)^T$ with z_i uniformly and independently distributed over the set $\mathcal{A}$.

Step 2: Find initial estimate $\widehat{\theta}^{*(0)}$ by solving (9.3) with randomly chosen $\mathbf{z}$ from Step 1.

Step 3: Use the initial estimate $\widehat{\theta}^{*(0)}$ to find the optimal value of the vector $\mathbf{z}$ by maximizing $\left|\widehat{\mathbf{I}}\left(\widehat{\theta}^{*(0)}\right)\right|$.

Step 4: Find final AML estimate $\widehat{\theta}^{*}$ by solving (9.3) with optimally chosen $\mathbf{z}$ from Step 3.

This algorithm where the values of $\mathbf{z}$ are chosen optimally instead of randomly have a significant effect on the quality of the estimator in terms of the efficiency. We will show on a simulation study how the random and optimal choice affect the resulting estimator in the next Section.

9.3 Simulation study

In the current Section we do two simulation studies. First we show the asymptotic behaviour of the AML estimator on simulated samples from SDS distribution and prove that the optimal choice of $\mathbf{z}$ significantly improves the results of the estimation. Secondly we will compare the results of the AML estimation with the results of the $k-L$ method described in (Feuerverger and McDunnough, 1981) on simulated samples of PDS and SDS distributions. The simulation algorithms can be derived from the stochastic representations of discrete stable distributions given in Part I of the thesis (Corollary 5.4 and Corollary 5.24).

9.3.1 Asymptotic behaviour of the AML estimate of SDS distribution

Here we look at the asymptotic behaviour of the AML estimator as a function of k. We repeatedly (1000 times) simulate a sample of size 1000 from SDS(0.8, 1) distribution and

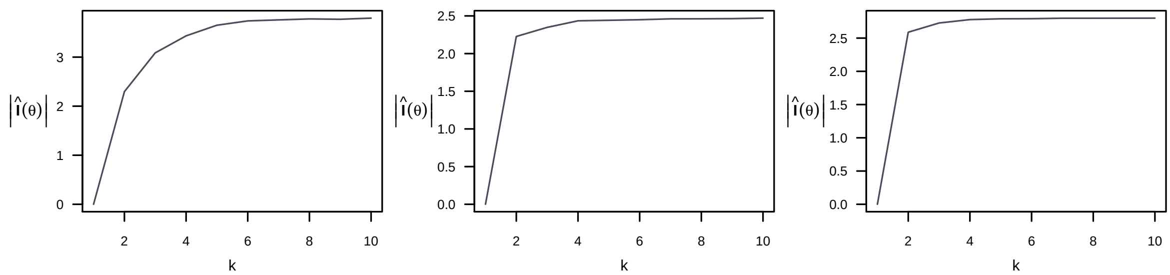

Figure 9.1: Speed of convergence of $|\widehat{\mathbf{I}}(\theta)|$ for PDS$(\gamma, 1)$ with $\gamma = 0.4$ (left), $\gamma = 0.6$ (middle) and $\gamma = 0.8$ (right).

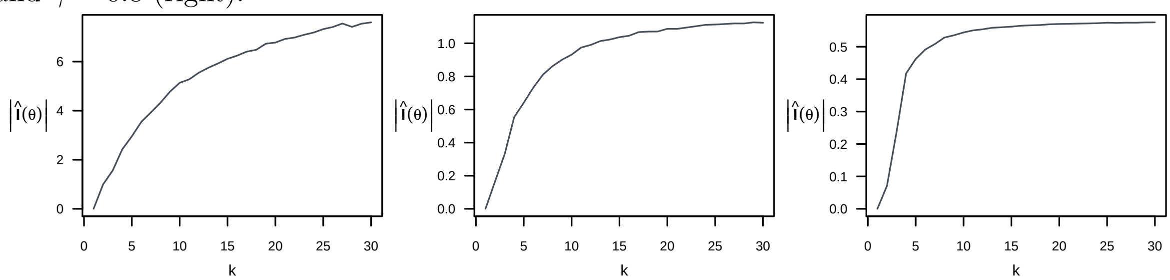

Figure 9.2: Speed of convergence of $|\widehat{\mathbf{I}}(\theta)|$ for SDS$(\gamma, 1)$ with $\gamma = 0.3$ (left), $\gamma = 0.6$ (middle) and $\gamma = 0.9$ (right).

for every $k \in \{3, \ldots, 25\}$ we estimate the parameters using the Algorithm 9.4. The mean square errors of the estimators of γ and λ are displayed in Figure 9.3. The behaviour of the determinant of the approximated Fisher information matrix is displayed in Figure 9.4. The optimal choice of $\mathbf{z}$ leads to best possible estimates as the determinant of the approximated Fisher information matrix approaches very closely the theoretical value (computed with the real values of parameters and optimally chosen vector $\mathbf{z}$). The random choice of $\mathbf{z}$ cannot compete with the optimal one in sense of efficiency. In Figure 9.5 we see the asymptotic behaviour of the estimates of γ and λ respectively as a function of k.

Remark 9.5. The algorithm from the previous Section is slightly modified in the simulation study. To achieve more precise estimates we added one step at the end of the algorithm. In this additional step we estimate parameter θ_1 in the presence of a nuisance parameter $(\theta_2, \ldots, \theta_d)$. This method was proposed by Klebanov and Melamed (1978) and the idea is in modifying the likelihood function (in our case the informant $\widehat{\mathbf{J}}(x, \theta)$) as

$$\tilde{J}_1(x, \theta) = \widehat{J}_1 - \widehat{\mathrm{E}}_\theta[\widehat{J}_1 | \widehat{J}_2, \ldots, \widehat{J}_d],$$

where $\widehat{\mathrm{E}}_\theta$ is the mathematical expectation in the wide sense, i.e. we solve linear regression problem of $\widehat{J}_1$ on $\widehat{J}_2, \ldots, \widehat{J}_d$. We first estimate all parameters together using the algorithm and then we estimate parameter λ with a nuisance parameter γ. Using this method the MSE of the estimate of parameter λ significantly decreased.

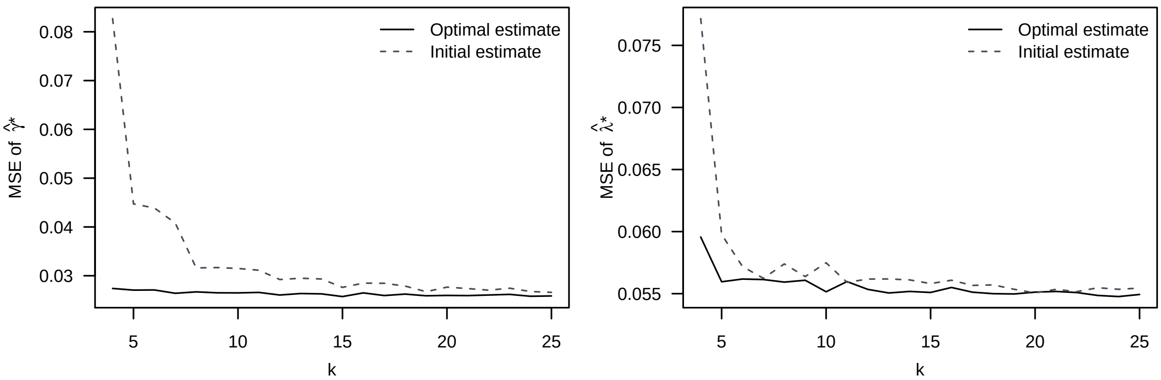

Figure 9.3: Mean square error of AML estimator of γ (left) and λ (right) as a function of k. Parameters estimated from simulated sample of size 1000 from SDS$(0.8, 1)$, with 1000 repetitions. The dashed line correspond to the initial estimate $\hat{\theta}^{*(0)}$ and the solid line to the optimal estimate $\hat{\theta}^*$.

9.3.2 Comparison of results of the AML method with the $k - L$ method

The $k-L$ procedure introduced by Feuerverger and McDunnough (1981) uses the asymptotic distribution of the empirical characteristic function at k fixed points $t_1, \ldots, t_k \in \mathbb{R}$. Let us denote f_n the empirical characteristic function, i.e.

$$f_n(t) = \frac{1}{n} \sum_{j=1}^{n} e^{\mathrm{i}tx_j},$$

where $x_1, \ldots, x_n$ is the observed sample. The characteristic function can be obtained from probability generating function as $f_\theta(t) = \mathcal{P}(e^{\mathrm{i}t})$. We use the following notation

$$V_n = (\mathrm{Re}(f_n(t_1)), \ldots, \mathrm{Re}(f_n(t_k)), \mathrm{Im}(f_n(t_1)), \ldots, \mathrm{Im}(f_n(t_k)))^T,$$
$$V_\theta = (\mathrm{Re}(f_\theta(t_1)), \ldots, \mathrm{Re}(f_\theta(t_k)), \mathrm{Im}(f_\theta(t_1)), \ldots, \mathrm{Im}(f_\theta(t_k)))^T,$$

and $\Sigma = \mathrm{cov}(V_n)$. The ECF estimate is given as the solution of the minimization problem

$$\min_{\theta \in \Theta} (V_n - V_\theta)^T \Sigma^{-1} (V_n - V_\theta).$$

Feuerverger and McDunnough (1981) show that this estimator is asymptotically normal and asymptotically efficient.

In our simulation study we compare results of the AML method and the $k - L$ method. We simulate samples of size 2000 from SDS$(0.8, 1)$ and PDS$(0.5, 4)$. We simulate 100 samples and we compare the mean square errors of the AML and $k - L$ estimates. We use $k = 10$ for

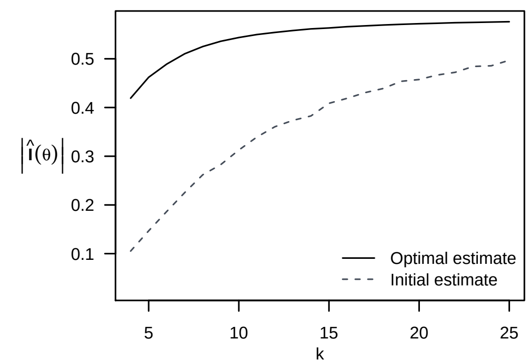

Figure 9.4: Asymptotic behaviour of $|\widehat{\mathbf{I}}(\theta)|$ as a function of k. Parameters estimated from simulated sample of size 1000 from SDS$(0.8, 1)$, with 100 repetitions. The dashed line correspond to the initial estimate $\hat{\theta}^{*(0)}$ and the solid line to the optimal estimate $\hat{\theta}^*$.

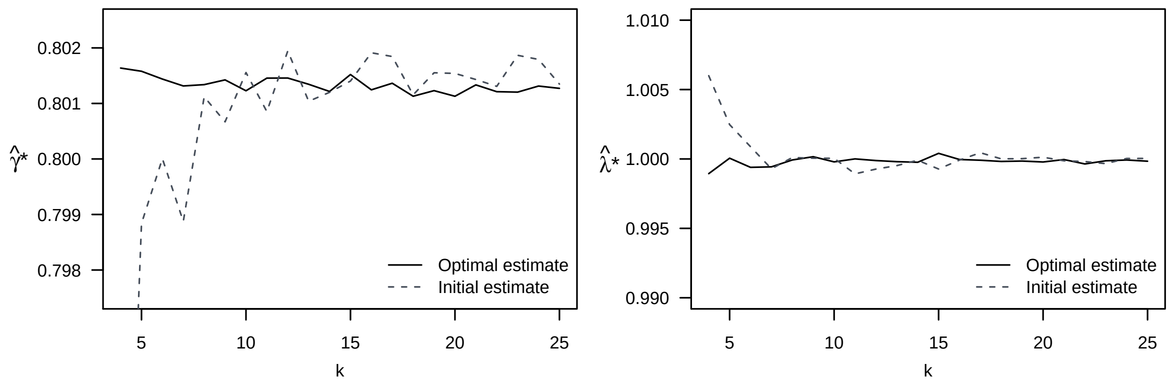

Figure 9.5: Asymptotic behaviour of AML estimators of γ (left) and λ (right) as a function of k. Parameters estimated from simulated sample of size 1000 from SDS$(0.8, 1)$, with 1000 repetitions. The dashed line correspond to the initial estimate $\hat{\theta}^{*(0)}$ and the solid line to the optimal estimate $\hat{\theta}^*$.

the AML method in case of SDS distribution and $k = 5$ for PDS distribution. We use $k = 10$ points in the $k - L$ method. The results are given in Table 9.1.

	γ	λ
AML	0.801	0.999
	(0.018)	(0.036)
$k - L$	0.806	1.001
	(0.033)	(0.038)

	γ	λ
AML	0.500	4.002
	(0.006)	(0.080)
$k - L$	0.496	3.978
	(0.028)	(0.212)

Table 9.1: Estimated parameters of SDS(0.8, 1) (left) and PDS(0.5, 4) (right) distributions from simulated samples of size 2000 with 100 repetitions. The mean square errors of the estimators are given in parentheses.

10. Discrete stable GARCH models and application to option pricing

Majority of financial models is based on the assumption that asset prices are continuous random variables. These models include both continuous and discrete time models. The continuous time modelling approach started with geometric Brownian motion by Black and Scholes (1973) and Merton (1973), followed by stochastic volatility models of Hull and White (1987) and Heston (1993), where both the asset price and volatility are stochastic processes. Models including jump part were introduced first by Merton (1976) who considered jump diffusion process governing the asset price. His work was followed by many others and an extensive overview of jump diffusion models can be found in Cont and Tankov (2004). Finally Eberlein and Jacod (1997) proposed pure jump processes as a model of asset prices. There is an apparent trend in loosening the assumption of continuity of asset prices toward their discreteness. Discrete time sampling of asset prices lead to consideration of discrete time models instead of continuous time; Duan (1995) introduced a discrete time model of asset prices, where the continuously compounded returns follow the generalized ARCH (GARCH) process of Bollerslev (1986). Similar model was studied by Heston and Nandi (2000).

The aim of many new models is the ability to describe the behaviour of financial markets more accurately. Asset price behaviour can be summarized into several stylized facts: stochastic volatility, volatility clustering, appearance of heavy tails and leverage effect where the negative returns have bigger effect on the volatility than the positive returns. In financial literature, heavy tails usually denote “tails heavier than Gaussian”. The discreteness is another property of asset prices that should not be ignored – stock and futures prices are quoted on exchanges in discrete manner, the smallest possible price change being called a tick; foreign exchange (FX) rates are also discrete with the smallest price change called a pip; interest rates are usually rounded from three to five decimal places. Already in the 80’s several authors noted that ignoring the discrete character of prices leads to biased moment estimators and suggested new estimators for variance and higher moments. Gottlieb and Kalay (1985) considered geometric Brownian model of asset prices and introduced an approximate correction factor for variance and market beta computed from the discrete observed prices; Ball (1988) assumed that the true asset price follow Brownian motion and suggested estimators based on Sheppard’s corrections (see (Craig, 1936)) of the discrete observed prices; Cho and Frees (1988) introduced yet another unbiased estimator of variance based on stopping times of geometric Brownian motion leading the asset prices. Since then many authors considered the discreteness as microstructure noise to otherwise continuous price process (Aït-Sahalia (2002), Aït-Sahalia et al. (2005), Andersen et al. (2011) among others). The microstructure noise is very strongly observed on low frequencies where the continuity assumption collide with the discreteness of prices that is very apparent on the short time scale. To summarize we may consider price discreteness as another stylized fact of the financial markets and study models where the discrete character of the market would not be ignored. Amilon (2003) considered discrete prices and introduced a GARCH model with hidden fair price that is estimated from the discrete market prices. Barndorff-Nielsen et al. (2012) introduced an integer-valued Lévy process for modelling low latency market data.

The aim of this paper is to introduce a new GARCH model taking into account the discrete character of the market by modelling the innovations by a discrete distribution. The GARCH model of asset prices introduced by Duan (1995) has as key assumption the normality of innovations, or conditional normality of asset returns. It was noted already by Mandelbrot (1963) that the asset returns exhibit leptokurtic behaviour and tails heavier than Gaussian and thus the Gaussian distribution does not provide a very good model. Stable distributions (or Paretian distributions) provide a modelling tool to describe the heavy tails in market data, however, the infinite variance of stable distributions and the fact that financial returns have heavier tails on a short time scale and almost Gaussian on a long scale brings into question the appropriateness of the stable model of returns. Grabchak and Samorodnitsky (2010) studied the paradox and showed, using the pre-limit theorem by Klebanov et al. (1999), that the sum of a large number of independent and identically distributed random variables behave as a stable random variable even though the tails of the random variables are not heavy. This motivated the introduction of other models that depart from the stable distribution but whose tails are not heavy. There are several ways how to obtain a distribution with exponential tails and stable behaviour in the "centre" of the distribution. Menn and Rachev (2005) used α-stable innovations with truncated tails. Menn and Rachev (2009) applied smooth truncation of tails of stable distribution, where the centre of the distribution correspond to stable distribution and the tails follow Gaussian distribution. Another way to come from heavy tails to a model with similar body and non-heavy tails is by tempering. In this approach the polynomially decaying tails are tempered with exponential rate. Stable distributions with exponentially tempered tails have been considered in the literature under different names – truncated Lévy flights (Koponen (1995)), CGMY model (Carr et al. (2002)) and finally tempered stable distributions (Rosiński (2007)). Tempered stable innovations and other generalizations are broadly used in GARCH option pricing framework, see, for example, Kim et al. (2009), Kim et al. (2010), Rachev et al. (2011), Mercuri (2008). To include the leverage effect, Menn and Rachev (2009) considered the nonlinear asymmetric GARCH (or shortly NGARCH) model, introduced by Engle and Ng (1993). This approach allows for asymmetric volatility smile in option prices and therefore is a better tool for modelling stock prices, that exhibit leverage effect.

The contribution of this paper is twofold. We introduce a new NGARCH model with discrete distribution of innovations, that takes into account the stylized facts of the market such as stochastic volatility, volatility clustering, leverage effect and discrete character of market prices. On the other hand we consider discrete distribution of innovations allowing for heavier tails by introducing two new discrete probability distributions – tempered discrete stable distribution and approximated discrete stable distribution, both approaches were considered already in Chapter 3. By tempering the tails of the distribution or by truncating large jumps of discrete stable distribution, we obtain distributions that behave as discrete stable in the middle but have exponential or Gaussian tails. We study the quality of the models on empirical data of the S&P 500 index compared to the normal and classical tempered stable (CTS) NGARCH models studied by Rachev et al. (2011). We study the option-pricing performance of our models compared to the normal NGARCH model and the CTS-NGARCH model. The Chapter is concluded by the study of option traders' market sentiment incorporated in the option prices. We analyse the market sentiment by computing the so called information ratio that measures the relative change from a benchmark market (Black-Scholes or Gaussian NGARCH) towards a heavier tailed market. By comparing the relatively calm recent period in March 2014 with a volatile period at the beginning of the financial crisis in September 2008, we show that the market is still relatively heavy-tailed (in the sense that the tails of the returns distribution are heavier than Gaussian) with a possibility of big losses; however, this risk decreased significantly since 2008.

This Chapter contains results from Slámová et al. (2014).

10.1 The model

In this Section we introduce the discrete stable NGARCH model. We will consider the discrete stable distribution in the limit sense with binomial thinning and modified geometric thinning operators that was introduced and studied in Part I of the thesis. The characteristic functions of the corresponding distributions $f_1(t)$ and $f_2(t)$ respectively are given by the following formulas:

$$f_1(u) = \exp\left\{-\lambda_1(1-e^{\mathrm{i}u})^\alpha - \lambda_2(1-e^{-\mathrm{i}u})^\alpha\right\}, \quad \alpha \in (0,1], \lambda_1, \lambda_2 > 0, \tag{10.1}$$

$$f_2(u) = \exp\left\{-\lambda_1\left(\frac{1-e^{\mathrm{i}u}}{1-\kappa e^{\mathrm{i}u}}\right)^\alpha - \lambda_2\left(\frac{1-e^{-\mathrm{i}u}}{1-\kappa e^{-\mathrm{i}u}}\right)^\alpha\right\}, \quad \begin{array}{l}\alpha \in (0,1], \lambda_1, \lambda_2 > 0,\\ \kappa \in [0,1).\end{array} \tag{10.2}$$

We can directly see that $f_2(u)$ reduces to $f_1(u)$ with $\kappa = 0$. Discrete stable distribution can be seen as compound Poisson distribution with intensity of jumps $\lambda_1+\lambda_2$ and distribution of jumps given by the following characteristic functions:

$$h_1(u) = 1 - \frac{\lambda_1}{\lambda_1+\lambda_2}(1-e^{\mathrm{i}u})^\alpha - \frac{\lambda_2}{\lambda_1+\lambda_2}(1-e^{-\mathrm{i}u})^\alpha,$$
$$h_1(u) = 1 - \frac{\lambda_1}{\lambda_1+\lambda_2}\left(\frac{1-e^{\mathrm{i}u}}{1-\kappa e^{\mathrm{i}u}}\right)^\alpha - \frac{\lambda_2}{\lambda_1+\lambda_2}\left(\frac{1-e^{-\mathrm{i}u}}{1-\kappa e^{-\mathrm{i}u}}\right)^\alpha.$$

10.1.1 Tempered discrete stable distributions

The tempered discrete stable distribution is constructed by exponential tempering of the tails of discrete stable distribution (10.1). This was done already in Chapter 3, here we provide more details on how the tempered discrete stable distributions are obtained.

Discrete stable distribution is a compound Poisson distribution with Lévy-Khintchine representation

$$\log f_1(u) = \int_{-\infty}^{\infty}\left(e^{\mathrm{i}ux}-1\right)\nu(\mathrm{d}x),$$

where the Lévy measure ν takes the following form

$$\nu(\mathrm{d}x) = (\lambda_1+\lambda_2)\sum_{k=-\infty}^{\infty} p_k\delta_k(\mathrm{d}x),$$

where

$$p_k = \begin{cases} \frac{\lambda_1}{\lambda_1+\lambda_2}(-1)^{k+1}\binom{\alpha}{k}, & k>0,\\ \frac{\lambda_2}{\lambda_1+\lambda_2}(-1)^{k+1}\binom{\alpha}{|k|}, & k<0,\\ 0, & k=0,\end{cases}$$

and δ_k is the Dirac measure, i.e. $\delta_x(A) = 1$ if $x \in A$ and 0 otherwise.

We temper the heavy tails of the discrete stable distribution using the tempering function

$$q(x) = e^{-\theta_1 x}\mathbf{1}_{x>0} + e^{-\theta_2|x|}\mathbf{1}_{x<0}, \quad \theta_1, \theta_2 > 0.$$

The tempered discrete stable distribution is then obtained by multiplying the corresponding Lévy measure by this tempering function. The parameters θ_1 and θ_2 give the tail decay rates.

Definition 10.1. Let $\alpha \in (0, 1]$, $\lambda_1, \lambda_2, \theta_1, \theta_2 > 0$. An infinitely divisible distribution is called *tempered discrete stable* (denoted TDS) with parameters $\alpha, \lambda_1, \lambda_2, \theta_1, \theta_2$, if its Lévy triplet (σ^2, ν, γ) is given by $\gamma = 0$, $\sigma = 0$ and

$$\nu(\mathrm{d}x) = (\lambda_1 + \lambda_2) \sum_{k=-\infty}^{\infty} p_k \delta_k(\mathrm{d}x),$$

where

$$p_k = \begin{cases} \frac{\lambda_1}{\lambda_1+\lambda_2} e^{-\theta_1 k}(-1)^{k+1}\binom{\alpha}{k}, & k > 0, \\ \frac{\lambda_2}{\lambda_1+\lambda_2} e^{-\theta_2 |k|}(-1)^{k+1}\binom{\alpha}{|k|}, & k < 0, \\ 0, & k = 0. \end{cases}$$

The characteristic function of $X \sim \mathrm{TDS}(\alpha, \lambda_1, \lambda_2, \theta_1, \theta_2)$ is given by

$$f(u) = \exp\left\{-\lambda_1 \left(1 - e^{\mathrm{i}u} e^{-\theta_1}\right)^\alpha - \lambda_2 \left(1 - e^{-\mathrm{i}u} e^{-\theta_2}\right)^\alpha + \lambda_1 \left(1 - e^{-\theta_1}\right)^\alpha + \lambda_2 \left(1 - e^{-\theta_2}\right)^\alpha\right\}.$$

This function can be extended via analytical continuation into the strip $-\theta_2 < \mathrm{Im}(z) < \theta_1$. All moments of TDS distribution exists. The first two central moments are given by the following formulas:

$$\begin{aligned} m_1 = \mathrm{E}X &= \alpha \left[\lambda_1 e^{-\theta_1} \left(1 - e^{-\theta_1}\right)^{\alpha-1} - \lambda_2 e^{-\theta_2} \left(1 - e^{-\theta_2}\right)^{\alpha-1}\right], \\ m_2 = \mathrm{E}X^2 &= \alpha \left[\lambda_1 e^{-\theta_1} \left(1 - e^{-\theta_1}\right)^{\alpha-1} + \lambda_2 e^{-\theta_2} \left(1 - e^{-\theta_2}\right)^{\alpha-1}\right] \\ &\quad - \alpha(\alpha - 1) \left[\lambda_1 e^{-2\theta_1} \left(1 - e^{-\theta_1}\right)^{\alpha-2} + \lambda_2 e^{-2\theta_2} \left(1 - e^{-\theta_2}\right)^{\alpha-2}\right]. \end{aligned}$$

If we substitute

$$\begin{aligned} \lambda_2 &= \lambda_1 \frac{e^{-\theta_1}\left(1 - e^{-\theta_1}\right)^{\alpha-1}}{e^{-\theta_2}\left(1 - e^{-\theta_2}\right)^{\alpha-1}}, \\ \lambda_1 &= \left\{\alpha e^{-\theta_1} \left(1 - e^{-\theta_1}\right)^{\alpha-1} \left[2 - (\alpha - 1)\left(\frac{1}{e^{\theta_1} - 1} + \frac{1}{e^{\theta_2} - 1}\right)\right]\right\}^{-1}, \end{aligned}$$

we obtain a tempered discrete stable distribution with zero mean and variance equal to 1. Such distribution will be called standard tempered discrete stable distribution with parameters $\alpha, \theta_1, \theta_2$ and denoted $\mathrm{stdTDS}(\alpha, \theta_1, \theta_2)$. The log-Laplace transform of the random variable $X \sim \mathrm{TDS}(\alpha, \lambda_1, \lambda_2, \theta_1, \theta_2)$, is defined for $u \in [-\theta_2, \theta_1]$ and is given as

$$L(u) = \log \mathrm{E}[e^{uX}] = -\lambda_1 \left(1 - e^{u} e^{-\theta_1}\right)^\alpha - \lambda_2 \left(1 - e^{-u} e^{-\theta_2}\right)^\alpha + \lambda_1 \left(1 - e^{-\theta_1}\right)^\alpha + \lambda_2 \left(1 - e^{-\theta_2}\right)^\alpha.$$

10.1.2 Approximate discrete stable distributions

Another approach to obtain a distribution with finite variance is truncating its tails. In Chapter 3 we introduced approximate symmetric discrete stable distribution. Here we use a similar approach and we truncate the large jumps of the jump distribution of discrete stable distribution (10.2). Let $m_1, m_2 \in \mathbb{N}$. Define the following characteristic function

$$f(u) = \exp\Bigg\{-\lambda_1 \sum_{n=0}^{m_1} (-1)^n \binom{\alpha}{n} (1-\kappa)^n \frac{e^{\mathrm{i}un}}{(1-\kappa e^{\mathrm{i}u})^n} - \lambda_2 \sum_{n=0}^{m_2} (-1)^n \binom{\alpha}{n} (1-\kappa)^n \frac{e^{-\mathrm{i}un}}{(1-\kappa e^{-\mathrm{i}u})^n} + \lambda_1 \sum_{n=0}^{m_1} (-1)^n \binom{\alpha}{n} + \lambda_2 \sum_{n=0}^{m_2} (-1)^n \binom{\alpha}{n}\Bigg\}. \tag{10.3}$$

It is obvious that (10.3) converges pointwise with $m_1, m_2 \to \infty$ to the characteristic function (10.2) of discrete stable distribution.

Definition 10.2. Let $\alpha \in (0,1]$, $\kappa \in [0,1)$ and $m_1, m_2 \in \mathbb{N}$. An infinitely divisible distribution is called *approximate discrete stable* (denoted ADS) with parameters $\alpha, \lambda_1, \lambda_2, \kappa, m_1, m_2$, if its characteristic function is given by (10.3).

The characteristic function f is analytic and the log-Laplace transform $L(u) = \log \mathrm{E}[e^{uX}]$ exists for all $u \in \mathbb{R} \setminus \{\log(\kappa), \log(1/\kappa)\}$ and is given as

$$L(u) = -\lambda_1 \sum_{n=1}^{m_1} (-1)^n \binom{\alpha}{n} (1-\kappa)^n \frac{e^{un}}{(1-\kappa e^{u})^n} - \lambda_2 \sum_{n=1}^{m_2} (-1)^n \binom{\alpha}{n} (1-\kappa)^n \frac{e^{-un}}{(1-\kappa e^{-u})^n} + \lambda_1 \sum_{n=1}^{m_1} (-1)^n \binom{\alpha}{n} + \lambda_2 \sum_{n=1}^{m_2} (-1)^n \binom{\alpha}{n}.$$

All moments of ADS distribution exists. The first two central moments are given by the following formulas:

$$m_1 = \mathrm{E}X = -\frac{\lambda_1}{1-\kappa} \sum_{n=1}^{m_1} (-1)^n \binom{\alpha}{n} n + \frac{\lambda_2}{1-\kappa} \sum_{n=1}^{m_2} (-1)^n \binom{\alpha}{n} n,$$

$$m_2 = \mathrm{E}X^2 = m_1^2 + \frac{\lambda_1}{(1-\kappa)^2} \sum_{n=1}^{m_1} (-1)^n \binom{\alpha}{n} n(n+\kappa) + \frac{\lambda_2}{(1-\kappa)^2} \sum_{n=1}^{m_2} (-1)^n \binom{\alpha}{n} n(n+\kappa).$$

We can choose λ_1 and λ_2 in such a way that $m_1 = 0$ and $m_2 = 1$. We then obtain approximate discrete stable distribution with zero mean and variance equal to 1. Such distribution will be called standard approximate discrete stable distribution with parameters α, κ, m_1, m_2 and denoted stdADS$(\alpha, \kappa, m_1, m_2)$.

10.1.3 Discrete stable NGARCH option pricing model

In this Subsection we introduce a new NGARCH model with discrete stable innovations. In particular we will study two models with stdTDS and stdADS distributed innovations. We will denote these models TDS-NGARCH and ADS-NGARCH model respectively and together they will be called discrete stable NGARCH models, or DS-NGARCH. Let us assume that the asset price process $(S_t, t \in \mathbb{N})$ is governed by the following dynamic under the real-world probability measure $\mathbb{P}$:

$$\log\left(\frac{S_t}{S_{t-1}}\right) = r_t + \lambda\sigma_t - L(\sigma_t) + \sigma_t\varepsilon_t, \quad t \in \mathbb{N}, \tag{10.4}$$

where S_t is the asset price at time t, r_t is the time-t risk-free interest rate for the period $[t-1, t]$, λ is the market price of risk. The processes r_t is assumed to be predictable. The innovations ε_t are independent and identically distributed random variables coming from distribution F, $L(u)$ is the log-Laplace transform of the innovations ε_t, i.e. $L(u) = \log \mathrm{E}[\exp\{uX\}]$ with $X \sim F$. L is defined either on a closed interval $[-a, b]$ with $a, b > 0$ or on the whole real line $\mathbb{R}$. The conditional variance σ_t^2 follows a NGARCH(1,1) process defined as

$$\sigma_t^2 = (\alpha_0 + \alpha_1\sigma_{t-1}^2(\varepsilon_{t-1} - \gamma)^2 + \beta_1\sigma_{t-1}^2) \wedge \rho, \quad t \in \mathbb{N}, \tag{10.5}$$

with $\varepsilon_0 = 0$, $\alpha_0 > 0$, $\alpha_1, \beta_1 \geq 0$ and $\alpha_1(1+\gamma^2) + \beta_1 < 1$ in order to guarantee the existence of a strong stationary solution with finite unconditional mean, and where $0 < \rho \leq b^2$ if L is not defined on the whole real line and $\rho = \infty$ otherwise.

The model reduces to classical GARCH model of Duan (1995) when $\gamma = 0$ and the distribution of innovations is Gaussian. The parameter γ allows for asymmetry in the volatility when negative returns have higher impact on the volatility than positive returns. The model moreover reduces to discrete-time Black-Scholes model when the volatility is constant, i.e. when $\alpha_1 = \beta_1 = 0$.

The innovations $(\varepsilon_t, t \in \mathbb{N})$ are assumed to be i.i.d. random variables with one of the following distribution

- $\varepsilon_t \sim \mathrm{stdTDS}(\alpha, \theta_1, \theta_2)$ in case of the TDS-NGARCH model,
- $\varepsilon_t \sim \mathrm{stdADS}(\alpha, \kappa, m_1, m_2)$ in case of the ADS-NGARCH model.

For comparative purposes we will also consider Gaussian and classical tempered stable (CTS) distribution of innovations.

- $\varepsilon_t \sim \mathrm{Normal}(0, 1)$ in case of the Normal-NGARCH model,
- $\varepsilon_t \sim \mathrm{stdCTS}(\alpha, \theta_1, \theta_2)$ in case of the CTS-NGARCH model.

The model is then a generalized version of models considered in Duan (1995) and Kim et al. (2010) respectively (with $\gamma = 0$).

In case of most securities the smallest price change is 0.01 or 0.001, i.e. the tick size is 0.01 or 0.001. We will therefore assume that the innovations take discrete values on a lattice with size $\Delta = 0.001$. It is easy to modify the definitions of TDS and ADS distributions to this case by considering characteristic functions of the form $f(\Delta u)$.

For pricing of derivatives we need the price process (10.4), (10.5) under a risk-neutral measure $\mathbb{Q}$. As explained in Rachev et al. (2011) the market given by (10.4) is incomplete and therefore there exist more than one equivalent martingale measure. We can however impose an especially simple form of the risk-neutral measure $\mathbb{Q}$ (as was done in Duan (1995)) under which the price process has the following dynamics

(10.6) $$\log\left(\frac{S_t}{S_{t-1}}\right) = r_t - L(\sigma_t) + \sigma_t\varepsilon_t, \quad t \in \mathbb{N},$$

where ε_t have the same distribution under the market measure and risk-neutral measure and where

(10.7) $$\sigma_t^2 = (\alpha_0 + \alpha_1\sigma_{t-1}^2(\varepsilon_{t-1} - \lambda - \gamma)^2 + \beta_1\sigma_{t-1}^2) \wedge \rho, \quad t \in \mathbb{N}.$$

10.2 Empirical analysis

In this Section, we study the performance of our discrete stable NGARCH models on the real data compared to the normal NGARCH and classical tempered stable NGARCH models.

As it is not possible to obtain closed form formula for the option prices

$$C(T, K) = e^{-rT}\mathrm{E}_{\mathbb{Q}}\left[(S_T - K)^+\right]$$

based on the model (10.6), (10.7), we will use the Monte Carlo simulation approach, i.e. we simulate M sample paths of the price process S_t, $(S_t^{(j)}, 1 \leq t \leq T)$ for $j = 1, \ldots, M$. The option price is then estimated by

$$\tilde{C}(T, K) = e^{-rT}\frac{1}{M}\sum_{j=1}^{M}(S_T^{(j)} - K)^+.$$

In the first part we describe the data used for our analysis. In the second Subsection, we investigate the statistical properties of our models under the objective probability measure $\mathbb{P}$. We will study the goodness of fit of the discrete stable NGARCH models on historical data, using two-step maximum likelihood method. We do a backtest to see the predictive power of the models under consideration. In the third Subsection we price the options under the risk-neutral measure $\mathbb{Q}$ using NGARCH parameters and parameters of the innovations distribution estimated by the maximum likelihood method. We will use the option prices to estimate the market price of risk and the spot volatility as these values are time-varying. In the last Subsection, we calibrate the risk-neutral parameters of the considered NGARCH models using market prices of options. This is done by minimizing the square error between the market and model prices. In this procedure one does not consider the market measure $\mathbb{P}$ but only the risk neutral measure $\mathbb{Q}$. Model estimated in this way can be then used for pricing more complicated options.

10.2.1 Data description

In our empirical analysis we will consider the S&P 500 index and call options on this index. We use adjusted closing prices of the S&P 500 index from March 25, 2004 to March 25, 2014, obtained from Bloomberg. We use IRX index for our daily interest rate r_t. IRX is an index created by Chicago Board Options Exchange and is computed from actual prices of 13-week Treasury Bills.

As for the options prices we will consider prices of liquid call options. We use two sets of data to study performance under different market conditions. The first set is from volatile period shortly before the outburst of the 2008 financial crisis – September 10, 2008. Second set is from much calmer recent period of March 25, 2014. All data were obtained from Bloomberg. The options have maturities ranging from 6 days to 200 days, with moneyness (ratio of the strike price and the current asset price) between 0.9 and 1.1; in total of 69 option prices for the 2008 data set and 69 option prices for the 2014 data set.

10.2.2 Market parameters estimation

In this Subsection we review results of market estimation – we estimate parameters of different NGARCH models from the historical market prices. We will use quasi maximum likelihood estimation where we approximate the non-normal NGARCH model by normal NGARCH. The reason for this approximation is the appearance of the log-Laplace transform L of the innovations distribution in (10.4). This inhibit the direct use of the maximum likelihood estimation, where the likelihood function depends on the unknown parameters of the innovations distribution. This problem is overcome once we use the Gaussian distribution of innovations: the log-Laplace transform L then reduces to $-1/2x^2$. So in order to estimate the NGARCH parameters $(\alpha_0, \alpha_1, \beta_1, \lambda, \gamma)$ we use normal approximation – that is we assume that the innovations have Gaussian distribution. In this setting the NGARCH parameters can be easily estimated using the MLE method. The innovations estimated from the Normal NGARCH model are then used to estimate the parameters of the innovations distribution. We use the empirical characteristic function method (or $k - L$ method) to estimate the distribution of innovations.

The estimated parameters for the index S&P 500 are in Table 10.1, the goodness of fit statistics are in Table 10.2, the QQ-plots are in Figure 10.1. To assess the goodness of fit we use Kolmogorov-Smirnov (KS) statistic,

$$KS = \sup_x |F_n(x) - F(x)|,$$

and Anderson-Darling (AD) statistic,

$$AD = \int_{-\infty}^{\infty} \frac{(F_n(x) - F(x))^2}{F(x)(1 - F(x))} \mathrm{d}F(x),$$

where F_n is the empirical cumulative distribution function of the innovations. The AD statistic focus on the tails of the distribution whereas the KS statistics more on its middle. We also give the p-values of the Kolmogorov-Smirnov and Anderson-Darling tests. It is well known that the p-values of the KS and AD tests are exact only if the distribution is continuous. In the case of discrete distributions we therefore use a randomized tests of Lemeshko et al. (2007). To compute the distribution of the AD test statistics we use the procedure of Marsaglia and Marsaglia (2004).

The CTS, TDS and ADS distributions offer all very good fit for the innovations distribution compared to the normal distribution. The best fit on the tails and in the middle of the distribution is achieved with the ADS distribution.

Backtest of models. To further assess the quality of our models, we perform a backtest, a procedure that assess the predictive power of the model. We base the backtest on the paper of Campbell (2005). We divide our sample into 2 parts, the first part serves for market estimation, the second part, of length N, serves for backtesting. For every data point in the

Table 10.1: SPX Market Parameters estimated on the time series from March 25, 2004 to March 25, 2014.

Model	Parameters				
	α_0	α_1	β_1	λ	γ
NGARCH	1.00e-10	0.0990	0.8645	0.6802	0.0064
	α	θ_+	θ_-		
stdCTS	-0.5313	2.6662	2.2678		
	α	θ_1	θ_2		
stdTDS	-0.6526	0.0035	0.0022		
	α	κ	m_1	m_2	
stdADS	0.4311	0.9519	55	87	

Table 10.2: Goodness of fit statistics for different distributional assumptions on the innovations of SPX returns.

Model	KS	(p-value)	AD	(p-value)
stdNormal	0.0541	7.84e-7	11.0625	4.54e-6
stdCTS	0.0117	0.8757	0.6597	0.5934
stdTDS	0.0097	0.9673	0.2086	0.9879
stdADS	0.0078	0.9966	0.1019	1.0000

second part of the data we compute Value at Risk at level α based on our model. We then compare this value with the return on the next day. We compute the number of exceedances in our given sample, i.e. we count the number of days when the loss is bigger than the Value at Risk from previous day. If our model is good, the number of exceedances should be equal to α. Campbell (2005) suggests a method where the number of exceedances on different levels is computed. We divide the interval [0,1] into several subintervals $I_1, I_2, \ldots, I_k$, i.e. [0,0.005), [0.005,0.01), [0.01, 0.05), [0.05,0.1) and [0.1, 1]. Under the null hypothesis that our model is correct, the number of exceedances falling into one of the intervals I_i, denoted N_i, should be equal to $N(u_i - l_i)$, where l_i and u_i are the lower and upper bounds of the interval I_i, respectively. We then perform the Pearson's chi-squared test, i.e. we compute a test statistic

$$Q = \sum_{i=1}^{k} \frac{(N_i - N(u_i - l_i))^2}{N(u_i - l_i)}.$$

Under the null hypothesis the statistics Q should have chi-squared distribution with $k-1$ degrees of freedom.

We perform the backtest on our 10 year data, we use first 8 years for market estimation and the last 2 years for backtesting, i.e. $N = 500$. The results of the backtest are summarized in the Table 10.3. We see that the discrete stable NGARCH models perform very well in the backtest, and are slightly more successful in predicting large losses than the tempered stable NGARCH model. The predictive power of the normal NGARCH model is bad as can be expected, as it fails to predict the heavier tails of returns.

Table 10.3: Results of the backtest on the time series from March 25, 2012 to March 25, 2014, with model estimated on time series from March 25, 2004 to March 24, 2012.

	Number of exceedances at level α				Pearson's Chi-squared test	
Model	0.005	0.01	0.05	0.1	Q	p-value
Normal-NGARCH	13	18	34	51	49.9606	3.6e-10
CTS-NGARCH	3	10	31	54	8.4441	0.0766
TDS-NGARCH	3	7	30	52	1.8190	0.7690
ADS-NGARCH	3	7	28	48	2.0590	0.7249
Theoretical	2.5	5	25	50		
Upper bound	0	1	16	37		
Lower bound	5	9	34	63		

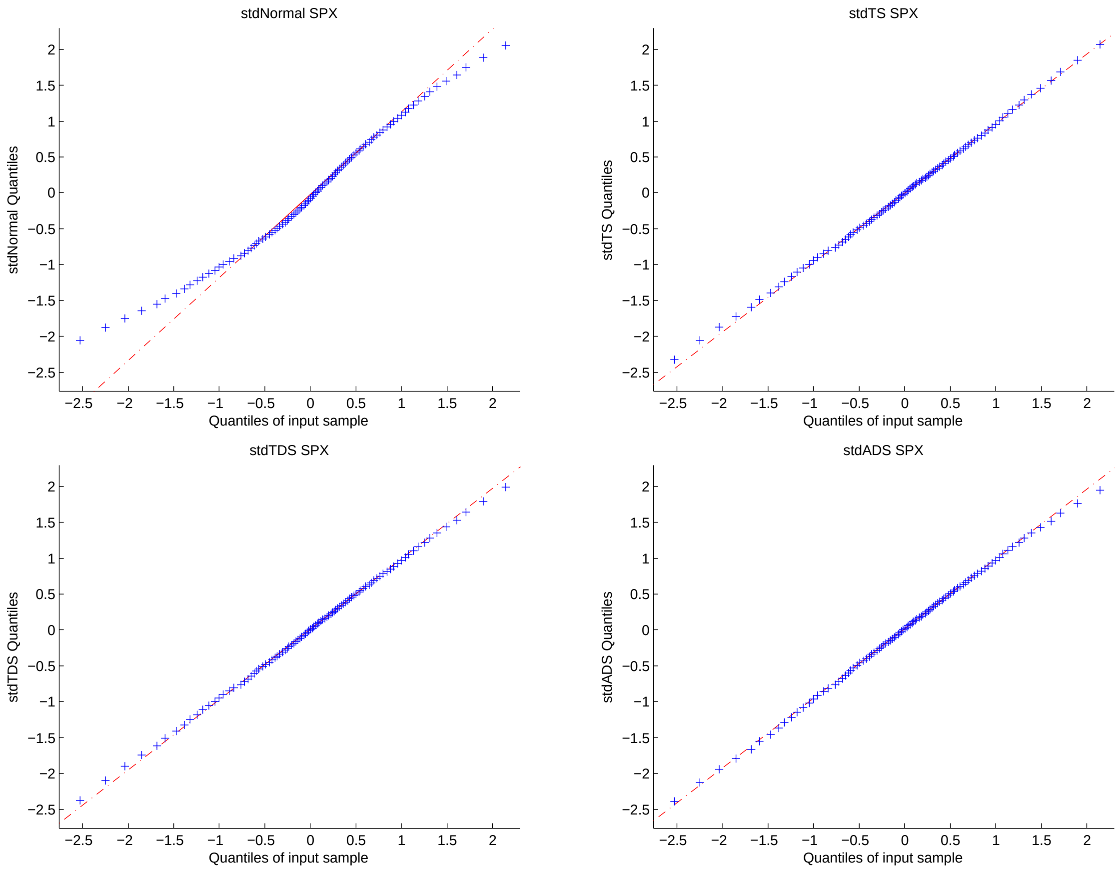

Figure 10.1: QQ-Plots for S&P 500 innovations

10.2.3 Risk-neutral option pricing

To price the options we need a risk-neutral probability measure $\mathbb{Q}$. As was already mentioned, we can choose a simple form of the measure $\mathbb{Q}$ under which the NGARCH asset price model takes form (10.6), (10.7). We therefore use the market parameters estimated from the historical prices under the measure $\mathbb{P}$ and use them in the risk-neutral model (10.6), (10.7). To price the options we use the Monte Carlo simulation approach. The problem that appears when using Monte Carlo simulations is that the generated sample path of the discounted price process is not a martingale any more. We thus use the empirical martingale simulation of Duan and Simonato (1998) to overcome this problem. In this procedure the simulated asset price is corrected in every step to impose the martingale property and to reduce the variance of the simulated sample.

We simulate the price process under the measure $\mathbb{Q}$ with parameters $\alpha_0, \alpha_1, \beta_1, \gamma$ of the NGARCH process and parameters of the innovations distributions. The market price of risk λ and the spot variance σ_0^2 are estimated from the market prices of options by fitting the model prices of options to the market prices. This procedure was considered in Menn and Rachev (2009) as MLE/fitted approach. The reason for this method is that the parameters of the NGARCH process and of the innovations distribution are considered stable over time, however the market price of risk is time varying and very difficult to estimate from historical data. The spot variance influences the option prices to a big extent, and it is reasonable to assume that the current level of volatility is better contained in the current prices of options than in the 10 years history of index values.

Denote by N the number of options and $C_i(T_i, K_i)$, $i = 1, \ldots, N$ the price of an option with maturity T_i and strike K_i observed on the market. We estimate the parameter $\theta = (\lambda, \sigma_0^2)$ by calibrating the model under the measure $\mathbb{Q}$ to the market prices of options. We calibrate the model in the following way. Let us denote $\tilde{C}_i^\theta(T_i, K_i)$ the model price of the i-th option. We estimate the parameter θ by fitting the model prices to the observed market prices by solving the following problem

$$\min_\theta \sum_{i=1}^N (C_i(T_i, K_i) - \tilde{C}_i^\theta(T_i, K_i))^2.$$

We compare the option pricing performance of our discrete stable NGARCH models together with the CTS-NGARCH model and the Normal-NGARCH model. For comparative purposes we include also results of the corresponding GARCH models and the classical Black-Scholes model. To compare the models we use the error estimators defined in Schoutens (2003), namely absolute percentage error (APE), average absolute error (AAE), root mean square error (RMSE) and average relative percentage error (ARPE). Their definitions are as follows (we denote $\bar{C}$ the average market option price):

$$\text{APE} = \frac{1}{\bar{C}} \sum_{i=1}^N \frac{|C_i - \tilde{C}_i|}{N} \qquad \text{AAE} = \sum_{i=1}^N \frac{|C_i - \tilde{C}_i|}{N}$$

$$\text{RMSE} = \sqrt{\sum_{i=1}^N \frac{(C_i - \tilde{C}_i)^2}{N}} \qquad \text{ARPE} = \frac{1}{N} \sum_{i=1}^N \frac{|C_i - \tilde{C}_i|}{C_i}$$

The results of risk-neutral pricing of European call options on S&P 500 from September 10, 2008, and for the March 25, 2014 data set are presented in Table 10.4. As expected, we can see that the performance of the Black-Scholes model with constant volatility is significantly worse then any other model that captures the volatility clustering using the GARCH model

Table 10.4: Risk-neutral in-sample pricing (MLE/fitted) of S&P 500 European Call Options on September 10, 2008 (left) and on March 25, 2014 (right).

Model	APE	AAE	RMSE	ARPE	Model	APE	AAE	RMSE	ARPE
Black-Scholes	0.6539	19.927	23.838	0.808	Black-Scholes	0.1500	11.297	14.668	0.315
Normal-GARCH	0.0797	2.4301	3.0370	0.1462	Normal-GARCH	0.0763	5.7487	8.6069	0.2986
CTS-GARCH	0.0697	2.1253	2.6541	0.1326	CTS-GARCH	0.0705	5.3119	7.8806	0.2760
TDS-GARCH	0.0688	2.0956	2.6170	0.1303	TDS-GARCH	0.0689	5.1888	7.6646	0.2633
ADS-GARCH	0.0603	1.8362	2.2969	0.1153	ADS-GARCH	0.0683	5.1453	7.5834	0.2580
Normal-NGARCH	0.0487	1.4837	1.8355	0.1008	Normal-NGARCH	0.0217	1.6345	2.3147	0.0335
CTS-NGARCH	0.0375	1.1440	1.4323	0.0820	CTS-NGARCH	0.0272	2.0468	2.8188	0.0434
TDS-NGARCH	0.0369	1.1232	1.4049	0.0803	TDS-NGARCH	0.0310	2.3311	3.1747	0.0522
ADS-NGARCH	0.0313	0.9538	1.1858	0.0686	ADS-NGARCH	0.0306	2.3051	3.1454	0.0521

for volatility. The incorporation of the leverage effect in the NGARCH model improves the results of the GARCH model. The non-Gaussian NGARCH models overperform the Normal-NGARCH model in the volatile period in 2008, however in 2014, the Normal-NGARCH model performs the best. The GARCH effect itself already captures the heavier tails in the unconditional distribution of the asset prices, and by adding non-Gaussian distribution of innovations we are able to capture the non-Gaussian tails into further extent. The period of September 2008, shortly before the outburst of the financial crisis, was very volatile and the returns exhibited heavier tails than in the relatively calm period of March 2014. Therefore in 2008 the GARCH model itself was not enough and by adding non-Gaussian innovations we improved the model. In 2014 it appears that the tail behaviour is captured by the GARCH effect of volatility and therefore non-Gaussian NGARCH models do not improve the results.

10.2.4 Market calibration

In this Subsection we calibrate our model using the market option prices. Our goal is to obtain risk-neutral parameters of our model by extracting the available information from the current option prices.

We estimate the risk-neutral parameters of the model (10.6), (10.7) by calibrating the model to the market prices of options that reflect the risk-neutral probability space. We estimate the set of parameters θ of the innovations distribution and we use the parameters of the NGARCH model estimated from the historical times series from this Section. The set of parameters θ is estimated by fitting the model prices to the observed market prices, i.e. we solve the following problem

$$\min_{\theta} \sum_{i=1}^{N} (C_i(T_i, K_i) - \tilde{C}_i^{\theta}(T_i, K_i))^2.$$

Once we estimate the parameter θ of the innovations distribution, we use the Monte Carlo simulations to price the options as in previous Subsection.

The results of the market calibration in terms of pricing errors are given in Table 10.5 for the data from September 2008 and March 2014. We see that in both cases the non-Gaussian GARCH and NGARCH models perform better than the corresponding Normal-GARCH and NGARCH models.

Table 10.5: Market calibration of S&P 500 European Call Option Prices on September 10, 2008 (left) and on March 25, 2014 (right).

Model	APE	AAE	RMSE	ARPE	Model	APE	AAE	RMSE	ARPE
Normal-GARCH	0.1283	3.9094	4.7835	0.2109	Normal-GARCH	0.0397	2.9878	3.8099	0.1340
CTS-GARCH	0.1108	3.3771	4.1085	0.1481	CTS-GARCH	0.0334	2.5160	3.3420	0.0855
TDS-GARCH	0.1125	3.4285	4.1717	0.1508	TDS-GARCH	0.0323	2.4334	3.2843	0.0773
ADS-GARCH	0.1138	3.4679	4.2041	0.1601	ADS-GARCH	0.0353	2.6605	3.4399	0.1033
Normal-NGARCH	0.0142	0.4334	0.7114	0.0218	Normal-NGARCH	0.0257	1.9330	2.6512	0.0519
CTS-NGARCH	0.015	0.4559	0.6571	0.0288	CTS-NGARCH	0.0202	1.5222	2.1576	0.0455
TDS-NGARCH	0.0152	0.464	0.6664	0.0289	TDS-NGARCH	0.0208	1.5695	2.2187	0.0399
ADS-NGARCH	0.0144	0.4394	0.6691	0.0251	ADS-NGARCH	0.0225	1.6961	2.4410	0.0406

10.3 Comparison of options markets sentiment in 2008 and 2014

Options markets can serve as an interesting indicator of the market sentiment – the prices are formed by traders who translate into the prices their prediction of market behaviour. If traders see the market as very volatile with a possibility of big losses, the option prices will react accordingly. The question is how to quantify this sentiment and how to extract it from the current market option prices.

In this Section we introduce the Information ratio, giving an information of how much the market diverges from a benchmark market. The benchmark can be the classical Black-Scholes market with constant volatility where the asset prices follow geometric Brownian motion; or a Gaussian market with volatility clustering ruled by Normal-NGARCH model. The information ratio can be then used as an indicator of future big losses on the options market – if the ratio is close to zero or negative, the traders sentiment is such that the markets behave as the benchmark market. If on the other hand the ratio is large, it can serve as an indicator of a strongly non-Gaussian market where the risk of big losses is higher than in the benchmark market. To measure the "heavy-tails" of the market we use Value at Risk, a measure that indicates the largest possible loss in a specified time frame at a given confidence level. A 99% Value at Risk is a value of a loss that will not be exceeded with a 99 % probability. This value, as it measures a tail event, is able to capture to what extent is the market heavy-tailed: a big Value at Risk signalizes a significant probability of large loss, or that the tails of the returns distribution are fat (not necessarily heavy-tailed with infinite variance).

We will use the calibrated models from previous Section because they contain the information about current option prices and therefore the option traders' market sentiment. We will compute 10 day Value at Risk from options at 99% confidence level. We do this in the following way. For every model we simulate $M = 1000$ sample paths of the asset price S_t for 10 days. For every sample path, $i = 1, \dots, M$, we compute the price of every option using the Monte Carlo simulation of the previous Section, conditioned that S_0 is the last value of every simulated sample path. We use N options ($N = 69$ options for the 2008 data set and $N = 69$ for the 2014 data set) with maturities ranging from $T = 6$ to $T = 200$ days. We then compute the returns of the options, by computing $\log(C_{10}^{(i)}/C_0)$, $i = 1, \dots, M$, where $C_{10}^{(i)}$ is the future price of the option given that the new S_0 corresponds to the i-th simulated sample path and C_0 is the current market option price. We compute the Value at Risk of

Table 10.6: Information ratios and information spread with respect to the Black-Scholes benchmark.

Model	IR		IS
	2008	2014	
Normal-GARCH	1.4774	1.4768	−0.04%
CTS-GARCH	1.4374	1.4456	0.57%
TDS-GARCH	1.4386	1.4392	0.04%
ADS-GARCH	1.4424	1.4910	3.37%
Normal-NGARCH	1.5544	1.4719	−5.31%
CTS-NGARCH	1.5561	1.4736	−5.30%
TDS-NGARCH	1.5585	1.4771	−5.22%
ADS-NGARCH	1.5922	1.5302	−3.89%

every option $j = 1, \dots, N$ based on these returns, for every model. For every option in our data set we compute the spread between the Value at Risk of the option in our model and a benchmark model,

$$\mathrm{S}_j^{\mathrm{model}} = \mathrm{VaR}_j^{\mathrm{model}}(99\%) - \mathrm{VaR}_j^{\mathrm{benchmark}}(99\%), \quad j = 1, \dots, N.$$

This spread express the change from the benchmark market to the model market. We define *information ratio* as normalized version of S over all options, i.e.

$$\mathrm{IR}^{\mathrm{model}} = \frac{\sqrt{N}\sum_{j=1}^{N}\mathrm{S}_j^{\mathrm{model}}}{\sqrt{\sum_{j=1}^{N}\left(\mathrm{S}_j^{\mathrm{model}} - \overline{\mathrm{S}^{\mathrm{model}}}\right)^2}}.$$

We can measure the change of the market by measuring the relative change of the information ratios between two periods of time T_1 and T_2, we will call such measure *information spread*,

$$\mathrm{IS}^{\mathrm{model}} = \frac{\mathrm{IR}^{\mathrm{model}}(T_2) - \mathrm{IR}^{\mathrm{model}}(T_1)}{\mathrm{IR}^{\mathrm{model}}(T_1)}.$$

The information ratios for the benchmark Black-Scholes market are reported in Table 10.6. The information ratios for the benchmark Black-Scholes market are positive for all models and both time periods under consideration, signalizing that the market sentiment is towards non-Gaussian models with volatility clustering. The information spread is rather insignificant, however we can see a slight decrease for all the NGARCH models, meaning the market in 2014 is closer to Black-Scholes market than in 2008.

The results with respect to the Normal-NGARCH benchmark are presented in Table 10.7. Here the results are more interesting. We can see that the information ratios are positive for all models and for both periods, however the ratios take larger values for the 2008 period, the information ratio decreased by around 40-50 % between the 2008 and 2014 periods. This signalizes that the market in 2008 was more heavy-tailed than in 2014. The market in 2014 still presents heavier tails than the Normal-NGARCH model where heavy tails originate only from volatility clustering.

Table 10.7: Information ratios and information spread with respect to the Normal-NGARCH benchmark.

Model	IR		IS
	2008	2014	
CTS-NGARCH	3.1925	1.3046	−59.14%
TDS-NGARCH	4.2306	2.0814	−50.80%
ADS-NGARCH	4.5323	2.6660	−41.18%

10.4 Conclusion

In this Chapter we introduced a new discrete stable NGARCH option pricing model, which captures the market stylized facts as volatility clustering, heavy tails and leverage effect. We made an attempt to include the discreteness of market prices as another stylized fact by considering the innovations of the NGARCH model to be discretely distributed. This model is not ideal as it does not preserve the discreteness in asset prices. Even despite this drawback we showed that the two discrete stable NGARCH models have a very good modelling and predictive performance even compared to the classical tempered stable NGARCH model – the fit of the innovations distribution with historical prices and the predictive power of the models in terms of the number of exceedances of Value at Risk are better for the discrete stable NGARCH models than for the CTS-NGARCH model. Finally we applied our models on option pricing. We showed that the pricing performance of non-Gaussian NGARCH models is better than the Gaussian NGARCH model. Finally we have showed that the option traders' market sentiment changed significantly since the 2008 financial crisis. Whereas the 2008 option markets were very volatile with a possibility of big losses, the analysis of the market sentiment showed that the risk of big losses decreased, as the market in 2014 approached the Gaussian NGARCH market quite significantly.

11. Conclusion

In this thesis we dealt with generalizations of the strict stability property and we focused our interest mainly on ways, how to define stability for discrete random variables. By introducing discrete stable distributions with different types of thinning operators, we provided tools for both researches and practitioners that deal with discrete distributions both with heavy and exponential tails. The discrete stable distributions studied in Chapter 5 can be considered as discrete analogies of classical stable distributions – they share many properties with them and with the lattice of the distribution going to zero, they converge to stable distributions. The discrete stability on the set of all integers is however defined only in the limit sense as the definition in the algebraic sense turned out to be infeasible.

When we generalized the random normalization procedure on continuous random variables and introduced casual stable distribution, we found out that many probability distributions, both with heavy and exponential tails, are in fact stable in this new sense. It also turned out that it is possible to find discrete distributions on $\mathbb{Z}$ that are discrete stable in the algebraic sense. The newly introduced symmetric geometric distribution (a discrete analogy of Laplace distribution) and symmetric negative binomial distribution are casual stable and thus discrete stable in the algebraic sense. It is essential to note, that there are many examples of casual (and discrete) stable distributions with exponentially decreasing tails, which contradicts to the popular opinion that stability (or self-similarity) property is connected to heavy tails. My studies of application of discrete stable distributions to option pricing show that the presence of exponential tails and central body similar to that of discrete stable distribution is essential for such kind of applications too.

Bibliography

Klebanov, L. B. and Slámová, L. (2013). Integer valued stable random variables. *Statistics & Probability Letters*, 83(6):1513–1519.

Klebanov, L. B. and Slámová, L. (2014). Discrete stable and casual stable random variables. arXiv preprint `arXiv:1406.3748 [math.PR]`.

Slámová, L., Kim, Y. S., and Klebanov, L. B. (2014). Discrete stable GARCH models and option pricing. Submitted.

Slámová, L. and Klebanov, L. B. (2012). Modelling financial returns with discrete stable distributions. In *Proceedings of 30th International Conference Mathematical Methods in Economics*, pages 805–810. Silesian University in Opava, School of Business Administration in Karviná.

Slámová, L. and Klebanov, L. B. (2014a). On discrete approximations of stable distributions. arXiv preprint `arXiv:1403.3671 [math.PR]`.

Slámová, L. and Klebanov, L. B. (2014b). Approximated maximum likelihood estimation of parameters of discrete stable family. *Kybernetika*. To appear.

Aït-Sahalia, Y. (2002). Telling from discrete data whether the underlying continuous–time model is a diffusion. *The Journal of Finance*, 57(5):2075–2112.

Aït-Sahalia, Y., Mykland, P. A., and Zhang, L. (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *Review of Financial studies*, 18(2):351–416.

Amilon, H. (2003). GARCH estimation and discrete stock prices: an application to low-priced Australian stocks. *Economics Letters*, 81(2):215–222.

Andersen, T. G., Bollerslev, T., and Meddahi, N. (2011). Realized volatility forecasting and market microstructure noise. *Journal of Econometrics*, 160(1):220–234.

Anderson, D. N. (1992). A multivariate Linnik distribution. *Statistics & Probability Letters*, 14(4):333–336.

Anderson, D. N. and Arnold, B. C. (1993). Linnik distributions and processes. *Journal of Applied Probability*, 30(2):330–340.

Ball, C. A. (1988). Estimation bias induced by discrete security prices. *The Journal of Finance*, 43(4):841–865.

Barabesi, L. and Pratelli, L. (2014). Discussion of “On simulation and properties of the stable law” by L. Devroye and L. James. *Statistical Methods & Applications*, pages 1–7.

Barndorff-Nielsen, O. E., Pollard, D. G., and Shephard, N. (2012). Integer-valued Lévy processes and low latency financial econometrics. *Quantitative Finance*, 12(4):587–605.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *The Journal of Political Economy*, pages 637–654.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327.

Bunge, J. (1996). Composition semigroups and random stability. *The Annals of Probability*, 24(3):1476–1489.

Campbell, S. D. (2005). A review of backtesting and backtesting procedures. Working paper. Federal Reserve Board, Washington.

Carr, P., Geman, H., Madan, D. B., and Yor, M. (2002). The fine structure of asset returns: An empirical investigation. *The Journal of Business*, 75(2):305–333.

Cho, D. C. and Frees, E. W. (1988). Estimating the volatility of discrete stock prices. *The Journal of Finance*, 43(2):451–466.

Christoph, G. and Schreiber, K. (1998). Discrete stable random variables. *Statistics & Probability Letters*, 37(3):243–247.

Christoph, G. and Schreiber, K. (2000). Scaled Sibuya distribution and discrete self-decomposability. *Statistics & Probability Letters*, 48(2):181–187.

Christoph, G. and Wolf, W. (1992). *Convergence theorems with a stable limit law*, volume Mathematical Research, vol. 70. Akademie Verlag.

Cont, R. and Tankov, P. (2004). *Financial Modelling with jump processes.* Chapman & Hall/CRC: London.

Craig, C. C. (1936). Sheppard's corrections for a discrete variable. *The Annals of Mathematical Statistics*, 7(2):55–61.

Devroye, L. (1990). A note on Linnik's distribution. *Statistics & Probability Letters*, 9(4):305–306.

Devroye, L. (1993). A triptych of discrete distributions connected to stable law. *Statistics & Probability Letters*, 18:349–351.

Doray, L. G., Jiang, S. M., and Luong, A. (2009). Some simple method of estimation for the parameters of the discrete stable distribution with the probability generating function. *Communications in Statistics – Simulation and Computation*, 38(9):2004–2017.

Duan, J.-C. (1995). The GARCH option pricing model. *Mathematical Finance*, 5(1):13–32.

Duan, J.-C. and Simonato, J.-G. (1998). Empirical martingale simulation for asset prices. *Management Science*, 44(9):1218–1233.

Eberlein, E. and Jacod, J. (1997). On the range of options prices. *Finance and Stochastics*, 1(2):131–140.

Engle, R. F. and Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *The Journal of Finance*, 48(5):1749–1778.

Erdélyi, A. et al. (1953a). *Higher transcendental functions*, volume 2. McGraw-Hill, New York.

Erdélyi, A. et al. (1953b). *Higher transcendental functions*, volume 1. McGraw-Hill, New York.

Erdoğan, B. and Ostrovskii, I. (1998). On mixture representation of the Linnik density. *Journal of the Australian Mathematical Society (Series A)*, 64(03):317–326.

Faa di Bruno, C. F. (1857). Note sur une nouvelle formule de calcul différentiel. *Quarterly J. Pure Appl. Math*, 1:359–360.

Feller, W. (1968). *An introduction to probability theory and its applications*, volume 1. John Wiley & Sons.

Feller, W. (1970). *An introduction to probability theory and its applications*, volume 2. John Wiley & Sons.

Feuerverger, A. and McDunnough, P. (1981). On the efficiency of empirical characteristic function procedure. *Journal of the Royal Statistical Society B*, 43(1):20–27.

Gerlein, O. V. and Kagan, A. M. (1979). Hilbert space methods in classical problems of mathematical statistics. *Journal of Soviet Mathematics*, 12(2):184–213.

Gnedenko, B. V. (1983). On limit theorems for a random number of random variables. In *Probability Theory and Mathematical Statistics*, pages 167–176. Springer.

Gnedenko, B. V. and Kolmogorov, A. N. (1949). *Limit distributions for sums of independent random variables (in Russian)*. Gostekhizdat, Moscow.

Gnedenko, B. V. and Kolmogorov, A. N. (1954). *Limit distributions for sums of independent random variables*. Addison-Wesley Publishing Company.

Gnedenko, B. V. and Korolev, V. Y. (1996). *Random summation: limit theorems and applications*. CRC Press.

Gottlieb, G. and Kalay, A. (1985). Implications of the discreteness of observed stock prices. *The Journal of Finance*, 40(1):135–153.

Grabchak, M. and Samorodnitsky, G. (2010). Do financial returns have finite or infinite variance? A paradox and an explanation. *Quantitative Finance*, 10(8):883–893.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–343.

Heston, S. L. and Nandi, S. (2000). A closed-form GARCH option valuation model. *Review of Financial Studies*, 13(3):585–625.

Hull, J. and White, A. (1987). The pricing of options on assets with stochastic volatilities. *The Journal of Finance*, 42(2):281–300.

Ibragimov, I. A. and Linnik, Y. V. (1971). *Independent and stationary sequences of random variables*. Wolters-Noordhoff, Groningen.

Johnson, N. L., Kemp, A. W., and Kotz, S. (2005). *Univariate discrete distributions*, volume 444. John Wiley & Sons.

Kagan, A. M. (1976). Fisher information contained in a finite-dimensional linear space, and a correctly posed version of the method of moments (in Russian). *Problemy Peredachi Informatsii*, 12(2):20–42.

Kakosyan, A. V., Klebanov, L. B., and Melamed, I. A. (1984). *Characterization of distributions by the method of intensively monotone operators*. Springer.

Khintchine, A. Y. (1938). *Limit laws of sums of independent random variables (in Russian)*. ONTI, Moscow.

Kim, Y. S., Rachev, S. T., Bianchi, M. L., and Fabozzi, F. J. (2010). Tempered stable and tempered infinitely divisible GARCH models. *Journal of Banking & Finance*, 34(9):2096–2109.

Kim, Y. S., Rachev, S. T., Chung, D. M., and Bianchi, M. L. (2009). The modified tempered stable distribution, GARCH models and option pricing. *Probability and Mathematical Statistics*, 29:91–117.

Klebanov, L., Melamed, J., and Manija, C. (1987). υ_p-strictly stable laws and estimation of their parameters. In Kalashnikov, V. V., Penkov, B., and Zolotarev, V. M., editors, *Stability Problems for Stochastic Models*, Lecture Notes in Mathematics, pages 23–31. Springer.

Klebanov, L. B. (2003). *Heavy tailed distributions*. Matfyzpress.

Klebanov, L. B., Kakosyan, A. V., Rachev, S. T., and Temnov, G. (2012). On a class of distributions stable under random summation. *Journal of Applied Probability*, 49(2):303–318.

Klebanov, L. B., Maniya, G. M., and Melamed, I. A. (1985). A problem of Zolotarev and analogs of infinitely divisible and stable distributions in a scheme for summing a random number of random variables. *Theory of Probability & Its Applications*, 29(4):791–794.

Klebanov, L. B. and Melamed, I. A. (1978). Several notes on Fisher information in presence of nuisance parameters. *Statistics: A Journal of Theoretical and Applied Statistics*, 9(1):85–90.

Klebanov, L. B. and Rachev, S. T. (1996). Sums of random number of random variables and their approximations with ν-accompanying infinitely divisible laws. *Serdica*, 22:471–498.

Klebanov, L. B., Rachev, S. T., and Szekely, G. J. (1999). Pre-limit theorems and their applications. *Acta Applicandae Mathematicae*, 58:159–174.

Koponen, I. (1995). Analytic approach to the problem of convergence of truncated Lévy flights towards the gaussian stochastic process. *Physical Review E*, 52(1):1197.

Kotz, S. and Ostrovskii, I. (1996). A mixture representation of the Linnik distribution. *Statistics & Probability Letters*, 26(1):61–64.

Kozubowski, T. J. and Rachev, S. T. (1999). Univariate geometric stable laws. *Journal of Computational Analysis and Applications*, 1(2):177–217.

Lachout, P. (2004). *Teorie pravděpodobnosti*. Karolinum.

Lemeshko, B. Y., Chimitova, E. V., and Kolesnikov, S. S. (2007). Nonparametric goodness-of-fit tests for discrete, grouped or censored data. In *Proceedings of ASMDA 2007 International Conference*, page 1–7.

Lévy, P. (1925). *Calcul des probabilités*. Gauthier-Villars Paris.

Lévy, P. (1937). *Théorie de l'addition des variables aléatoires*, volume 1. Gauthier-Villars Paris.

Lin, G. D. (1998). On the Mittag–Leffler distributions. *Journal of Statistical Planning and Inference*, 74(1):1–9.

Linnik, Y. V. (1953). Linear forms and statistical criteria, I, II. *Selected Transl. Math. Statist. and Prob*, 3:41–90.

Linnik, Y. V. (1964). *Decomposition of probability distributions*. Oliver & Boyd.

Lukacs, E. (1970). *Characteristic functions*. Griffin London, second edition.

Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business*, 36:294–419.

Marcheselli, M., Baccini, A., and Barabesi, L. (2008). Parameter estimation for the discrete stable family. *Communications in Statistics – Theory and Methods*, 37(6):815–830.

Marsaglia, G. and Marsaglia, J. (2004). Evaluating the Anderson-Darling distribution. *Journal of Statistical Software*, 9(2):1–5.

Menn, C. and Rachev, S. T. (2005). A GARCH option pricing model with α-stable innovations. *European Journal of Operational Research*, 163(1):201–209.

Menn, C. and Rachev, S. T. (2009). Smoothly truncated stable distributions, GARCH models, and option pricing. *Mathematical Methods of Operations Research*, 69(3):411–438.

Mercuri, L. (2008). Option pricing in a GARCH model with tempered stable innovations. *Finance research letters*, 5(3):172–182.

Merton, R. C. (1973). Theory of rational option pricing. *The Bell Journal of Economics and Management Science*.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1):125–144.

Peterson, G. J., Pressé, S., and Dill, K. A. (2010). Nonuniversal power law scaling in the probability distribution of scientific citations. *Proceedings of the National Academy of Sciences*, 107(37):16023–16027.

Petrov, V. V. (1975). *Sums of independent random variables*. Springer-Verlag.

Phillips, M. (1978). Sums of random variables having the modified geometric distribution with application to two-person games. *Advances in Applied Probability*, pages 647–665.

Pillai, R. (1990). On Mittag-Leffler functions and related distributions. *Annals of the Institute of Statistical Mathematics*, 42(1):157–161.

Pillai, R. N. and Jayakumar, K. (1995). Discrete Mittag-Leffler distributions. *Statistics & Probability Letters*, 23:271–274.

Price, D. J. d. S. (1965). Network of scientific papers. *Science*, 149:510?515.

Price, D. J. d. S. (1976). A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science*, 27(5):292–306.

Rachev, S. T., Kim, Y. S., Bianchi, M. L., and Fabozzi, F. J. (2011). *Financial models with Lévy processes and volatility clustering*, volume 187. John Wiley & Sons.

Rivlin, T. J. (1974). *The Chebyshev polynomials*. John Wiley & Sons.

Robbins, H. (1948). Mixture of distributions. *The Annals of Mathematical Statistics*, 19(3):360–369.

Rosiński, J. (2007). Tempering stable processes. *Stochastic Processes and their Applications*, 117(6):677–707.

Samorodnitsky, G. and Taqqu, M. S. (1994). *Stable non-gaussian random processes: Stochastic models with infinite variance*. Chapman & Hall.

Schoutens, W. (2003). *Lévy processes in finance: Pricing Financial Derivatives*. John Wiley & Sons.

Simkin, M. V. and Roychowdhury, V. P. (2007). A mathematical theory of citing. *Journal of the American Society for Information Science and Technology*, 58(11):1661–1673.

Skellam, J. G. (1946). The frequency distribution of the difference between two Poisson variates belonging to different populations. *Journal of the Royal Statistical Society*, 109(3):p. 296.

Steutel, F. W. and van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *Annals of Probabability*, 7(5):893–899.

Uchaikin, V. V. and Zolotarev, V. M. (1999). *Chance and Stability: Stable Distributions and their applications*. Walter de Gruyter.

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Addison–Wesley Press.

Zolotarev, V. M. (1986). *One-dimensional stable distributions*, volume 65. American Mathematical Soc.

A. Probabilities of discrete stable distributions

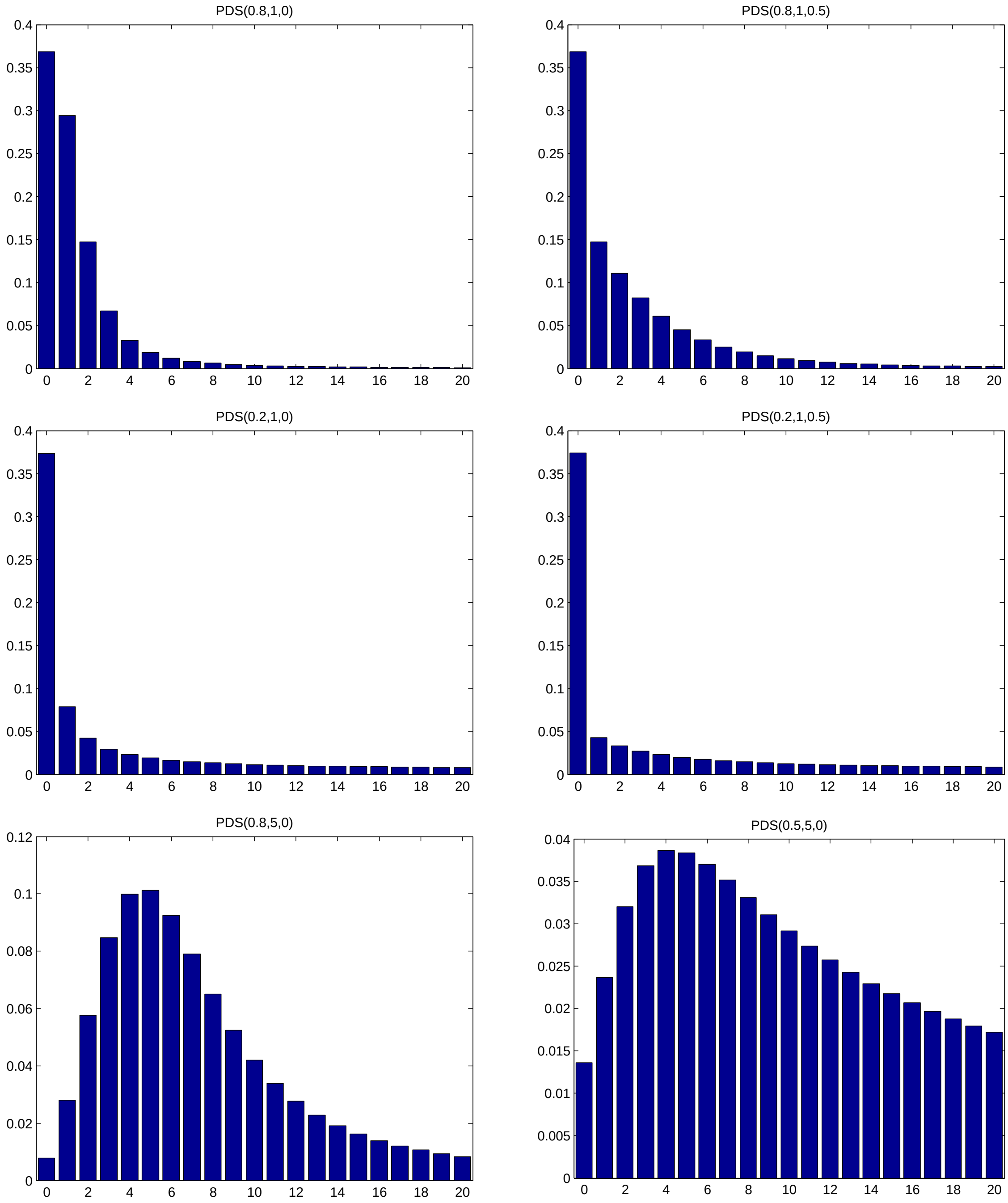

Figure A.1: Probabilities of positive discrete stable PDS(γ, λ, κ) random variables for different values of parameters.

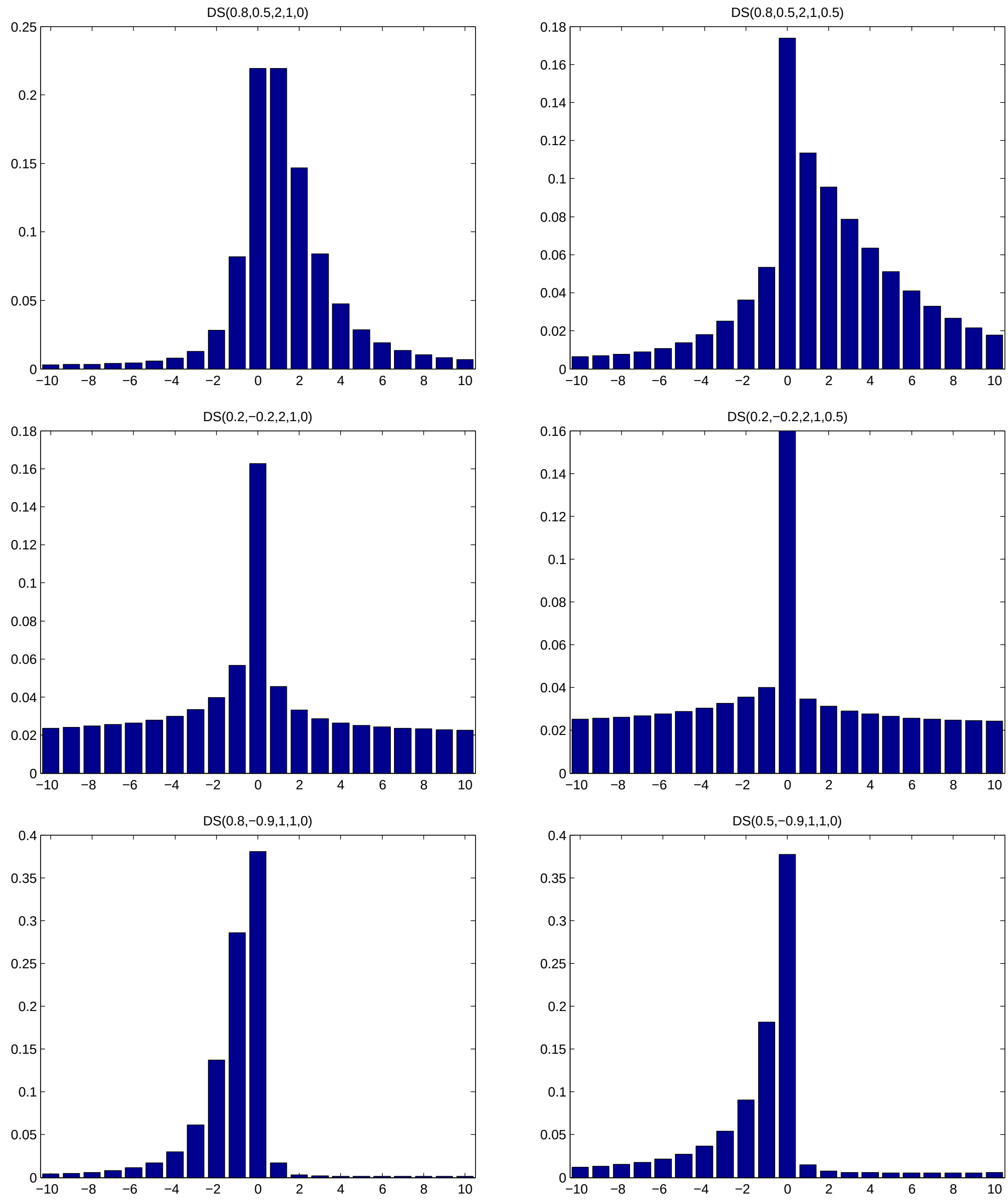

Figure A.2: Probabilities of discrete stable $\mathrm{DS}(\gamma, \beta, \lambda, 1, \kappa)$ random variables for different values of parameters.

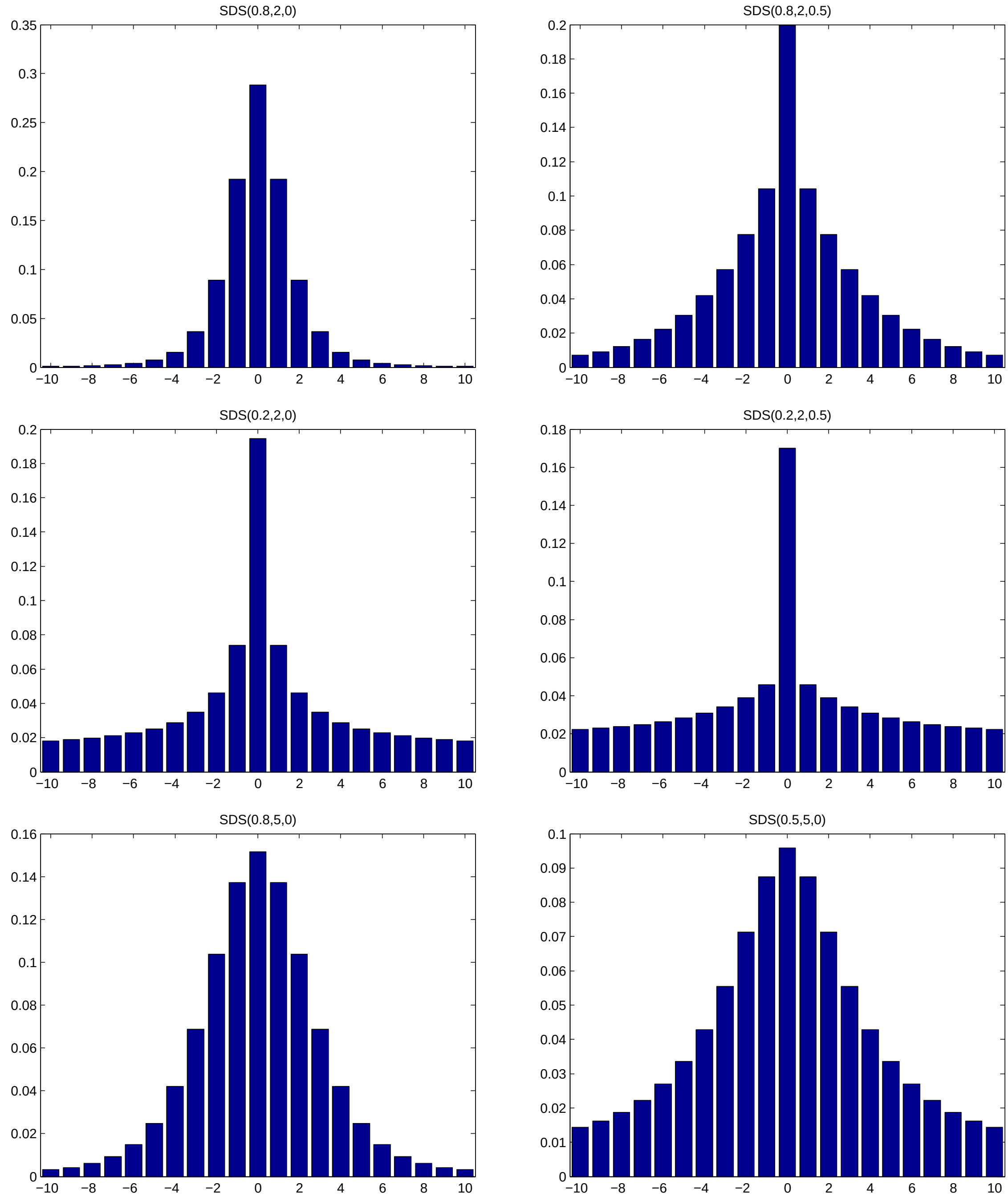

Figure A.3: Probabilities of symmetric discrete stable SDS(γ, λ, κ) random variables for different values of parameters.