ON THE INFINITE DIVISIBILITY AND NON INFINITE DIVISIBILITY OF CERTAIN CLASSES OF RANDOM VARIABLES

by

George M. Stukes

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Applied Mathematics

Charlotte

2023

Approved by:

_____
Dr. Stanislav Molchanov

_____
Dr. Isaac Sonin

_____
Dr. Michael Grabchak

_____
Dr. Donald Jacobs

ABSTRACT

GEORGE M. STUKES. On the infinite divisibility and non infinite divisibility of certain classes of random variables. (Under the direction of DR. STANISLAV MOLCHANOV)

In this dissertation we present new results on the classification of limit distributions of random geometric processes. In particular, that develop on the work of Penrose and Wade [1], who were the first to document the phenomenon of infinite divisibility in the case of a particular (uniform) distribution. In this dissertation we put forth not only new results, but a new method of obtaining results through analyzing the sequence of moments produced by random variables. Additionally we have new results in cycle decomposition of the related Dickman-Goncharov distribution. We present a novel proof of the distribution of the three highest order cycles in a random partition.

DEDICATION

I dedicate my dissertation to my wife Rebecca. She is a constant of source loving support, patience, and encouragement. My work would not have been possible without her.

ACKNOWLEDGEMENTS

It is hard to exactly say how grateful I am to my advisor, Dr. Stanislav Molchanov. He possesses so many great qualities of a mathematician, mentor, and human. His deep insight into mathematics is awe-inspiring. His passion is clear and bright. Most important to me however, is that his patience and support never wavered. I will miss our discussions together.

I would also like to thank my committee members Dr. Don Jacobs, Dr. Issac Sonin and Dr. Michael Grabchak. Their advice and encouragement helped me put together all my ideas in one place.

There are so many people in my life who helped me through this journey that it would be impossible to list all of them here. My sons Jozy and George were always a source of joy during difficult moments. My mom was always the first person to offer help in whatever way she could and often didn't even wait for me to ask. I am truly grateful for your support. Lastly, I am indebted to my wife Rebecca for so many things it is impossible to list them here. Thank you.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

The subject of this dissertation involves several different, but related distributions. The details of these distributions will be laid out in subsequent sections. Our early results are focused on the distributions the lengths of cycles that result from decompositions from random permutations of S_n. These cycle lengths were studied extensively by Goncharov [2] and later Vershik [3]. More recently, Molchanov and Panov [4] have produced results that this dissertation expands on.

The bulk of the dissertation is focused on random geometric progressions in the spirit of Penrose and Wade [1] and Vervaat [5]. The authors discovered that under particular conditions, the random variable S,

$$S = Y_1 + Y_1Y_2 + Y_1Y_2Y_3 + Y_1Y_2Y_3Y_4 + \cdots \ , \tag{1.1}$$

is infinitely divisible, where the Y_i are independent, identically distributed random variables. Later work by Grabchak and Molchanov [6] provided more results. This dissertation presents a new method to determine infinite divisibility of such random geometric progressions as well as classifies many examples.

1.1 Motivation for Joint Distribution of Maximal Cycles

Here we will discuss some of the motivation and background behind our first problem, which involves joint distributions of cycle lengths in the decomposition of a random permutation on S_n.

We begin by considering

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & \cdots & n \\ i_1 & i_2 & i_3 & i_4 & \cdots & i_n \end{pmatrix} , \tag{1.2}$$

which is a random permutation on the first n natural numbers. We define 'random' here as every particular element of the group S_n can be selected with uniform probability. Naturally, then, we can say for any permutation $\pi \in S_n$,

$$\mathbb{P}(\pi) = \frac{1}{n!} .$$

The notation in (1.2) indicates that the permutation π maps the element 1 to another element i_1 and 2 to another element i_2 and so on. This permutation can be decomposed into a number of cycles in the following way: $1 \to i_1 \to i_j \to \cdots \to i_r \to 1$, which completes one cycle. The next cycle begins with the smallest element remaining that was not included in the first cycle.

Example 1.1. *Consider the permutation of the first 5 natural numbers*

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 2 & 5 & 4 \end{pmatrix}$$

Here the permutation produces the cycles $1 \to 3 \to 2 \to 1$ and $4 \to 5 \to 4$. We can write the decomposed permutation as $(1, 3, 2)(4, 5)$. Note here that the first cycle has length 3 and the second cycle has length 2.

In our problem, we consider the distribution of the lengths of the maximal cycles in random permutations. We denote the length of the ith cycle as $|c_i|$. Where the ith cycle begins with the smallest element not included in the first $i - 1$ cycles.

In a straightforward manner, we can determine

$$\begin{aligned}
&\mathbb{P}\{|c_1| = 1\} = \frac{1}{n} \quad \text{(the only such cycle is } 1 \to 1) \\
&\mathbb{P}\{|c_1| = 2\} = \frac{(n-1)}{n}\frac{1}{n-1} = \frac{1}{n} \quad \text{(the cycles here are } 1 \to i_1 \to 1, i_1 \neq 1) \ . \\
&\mathbb{P}\{|c_1| = 3\} = \frac{(n-1)}{n}\frac{(n-2)}{(n-1)}\frac{1}{(n-1)} = \frac{1}{n}
\end{aligned}$$

Iterating this process, we arrive at

$$\mathbb{P}\{|c_1| = k\} = \frac{1}{n} \ .$$

If the length of the first cycle is less than n (which it is with probability $\frac{n-1}{n}$), then it is natural to consider the length of the second cycle, conditioned on the length of the first cycle. Understanding that if the first cycle is length k_1 the second cycle has length $n - k_1$, we arrive at the equation

$$\mathbb{P}\left\{|c_2| = k_2 \;\middle|\; |c_1| = k_1\right\} = \frac{1}{n - k_1} \ .$$

In the same way, we can compute the conditional probability for all higher cycles

$$\mathbb{P}\left\{|c_i| = k_i \;\middle|\; |c_{i-1}| = k_{i-1} \bigcap |c_{i-1}| = k_{i-2} \bigcap \cdots \bigcap |c_1| = k_1\right\} = \frac{1}{s_i} \ ,$$

where $1 \leq c_i \leq s_k$ and $s_k = n - k_1 - k_2 - \cdots - k_{i-1}$.

In this dissertation, we are interested in the limiting distribution ($n \to \infty$) of normalized lengths $\frac{|c_i|}{n}$.

We create random variables $Y_1, Y_2, ..., Y_k$ and set

$$\frac{|c_1|}{n} \overset{law}{\to} Y_1, \ \frac{|c_2|}{n} \overset{law}{\to} Y_2, \ \dots, \ \frac{|c_k|}{n} \overset{law}{\to} Y_k \qquad n \to \infty \ .$$

One can see that the distribution of the variables will follow

$Y_1 = X_1 \sim Unif[0,1]$

$Y_2 = X_2(1 - X_1) \quad , X_2 \sim Unif[0,1],$ where X_2, X_1 are independent.

$Y_3 = X_3(1 - X_1)(1 - X_2) \quad , X_3 \sim Unif[0,1],$ where X_3, X_2, X_1 are independent.

$\vdots$

We are primarily interested in the order statistics of the distribution of $\{Y_1, Y_2, Y_3, ...\}$. In particular we develop a novel method to derive a formula for the joint distribution of $(Y_{(1)}, Y_{(2)}, Y_{(3)})$, where Y_i indicates the ith largest value. The distribution of the longest cycle appears in Feller [7] and further cycles were extensively studied by Vershik in [3] and [8]. It is in this context that the Dickman-Goncharov distribution arises. More details and a derivation will be offered in section 2.2.

1.2 Motivation for Random Geometric Progressions

In this part of the dissertation we are going to discuss two concepts that come together in our work, infinite divisibility and random geometric progressions.

1.2.1 Infinite Divisibility

The notion of infinite divisibility began in the early twentieth century through pioneering work by Di Finetti [9], Lévy [10], Kolmogorov [11], and Khintchine [12]. The notion of divisibility of a random variable, X, concerns whether you can represent

X in the following way

$$X \overset{law}{=} Y_1 + Y_2 + Y_3 + \cdots + Y_n$$

$$Y_1, Y_2, ..., Y_n \text{ independent and identically distributed.}$$

If such a representation is possible, X is said to be *n-divisible*. It is interesting to note that $(n+1)$ divisibility does not imply n divisibility. This will make our notion of infinite divisibility slightly more complex than simply a limit of the sum of i.i.d. random variables.

Theorem 1.1. *We call a random variable X infinitely divisible if for every $n \in \mathbb{N}$, you can represent it in the following way*

$$X \overset{law}{=} X_{n,1} + X_{n,2} + \cdots + X_{n,n} \ ,$$

where $X_{n,1}, X_{n,2}, X_{n,n}$ are i.i.d. and there exists some X_n such that $X_{n,j} \overset{law}{=} X_n$ for all j.

Since the infinite divisibility of a random variable is based on it's distribution, we commonly use the term infinitely divisible to describe a random variable, distribution function, or density function (for absolutely continuous distributions).

We can also generalize the definition of infinite divisibility in (1.1) through the following theorem. To do so, we will utilize the following definition.

Here we define a *triangular array*. A triangular array is a double sequence of random variables $X_{i,j}\, i = 1, 2, ..., j; j = 1, 2, 3, ...$ and the variables in the nth row $(X_{1,n}, ..., X_{n,n}$ are mutually independent.

In our triangular arrays, we are interested in rows where individual components $X_{i,j}$ do not exert significant influence over the sum of the elements in the row. To

this end, we impose the additional constraint that for each $\epsilon > 0$,

$$\mathbb{P}\left\{|X_{i,j}| > \epsilon\right\} < \epsilon \qquad (i = 1, 2, ..., j) \tag{1.3}$$

for sufficiently large n. These arrays are sometimes called *null arrays*.

Theorem 1.2. *Let*

$$S_n = \sum_{i=i}^{n} X_{i,n} \tag{1.4}$$

be the sum of rows of a triangular array, $X_{i,j}$ with property (1.3)*. If there exists an S such that $S_n \overset{law}{\to} S$, then S is infinitely divisible.*

This remarkable fact generalizes (1.1) and drops the requirement that the components $X_{i,j}$ have some common distribution. For more details, the reader is directed to [13].

It is straightforward to show that a infinitely divisible distribution F can be written for every $n \in \mathbb{N}$ as the n-fold convolution of some other distribution F_n with itself. Similarly, a characteristic function ϕ will be infinitely divisible iff for every $n \in \mathbb{N}$ it is the nth power of some characteristic function ϕ_n:

$$F = F_n^{\star n} \ \ \forall n \in \mathbb{N} \qquad \phi(z) = \Big(\phi_n(z)\Big)^n \ \ \forall n \in \mathbb{N}\,.$$

Next we show the some of the most common types of infinitely divisible distributions.

Example 1.2. *Let X have a Normal (μ, σ^2) distribution. Then X is infinitely divisible and for every $n \in \mathbb{N}$ $X_n \sim N(\frac{\mu}{n}, \frac{\sigma^2}{n})$.*

Example 1.3. *Let X have a Poisson (λ) distribution. Then X is infinitely divisible*

with $X_n \sim Poisson(\frac{\lambda}{n})$ *for every* $n \in \mathbb{N}$.

Example 1.4. *Let* X *have a exponential distribution with parameter* k. *Then* X *is infinitely divisible with* $X_n \sim Exp(\frac{k}{n})$ *for every* $n \in \mathbb{N}$.

Example 1.5 (Penrose Wade)**.** *Let* $\{Y_i\}_i$ *be a sequence of random variables with distribution* $Y_i \sim Unif[0,1]$. *Then the sum,* S

$$S = Y_1 + Y_1Y_2 + Y_1Y_2Y_3 + \cdots$$

is infinitely divisible.

Alternative characterizations are discussed extensively in [14]. These characterizations will be utilized in chapter 3.

1.2.2 Random Geometric Progressions

In a paper about Minimal Directed Spanning Trees, Penrose and Wade [1] discovered that it is possible to generate an infinitely divisible distribution as the limit of a random geometric progression

$$S = Y_1 + Y_1Y_2 + Y_1Y_2Y_3 + Y_1Y_2Y_3Y_4 \cdots \ .$$

In the paper the authors demonstrated that if each of the Y_i are $U^{\frac{1}{\Theta}}(\Theta > 0)$, and $U \sim Unif[0,1]$ that the resulting geometric progression is infinitely divisible. We will recreate this result later in the paper. What is interesting to note is the specific parameters around the distribution of the Y_i variables. The main thrust of this dissertation is to make more clear what distribution of Y_i might result in an infinitely divisible distribution.

It is straightforward to show the condition $U^{\frac{1}{\Theta}}(\Theta > 0)$, and $U \sim Unif[0,1]$ corresponds to a random variable with a distribution of $Beta[\Theta, 1]$. In our search for

other potentially infinitely divisible progressions, we will naturally look at distributions that are very similar to the uniform and beta distributions. Details and proofs are presented in chapter 3.

1.3 Important Distributions

There are three main distributions that will be discussed in this dissertation. The first of these was discovered in the 1930's by Karl Dickman when researching natural numbers free of large prime factors [15]. The density of the Dickman distribution satisfies

$$\mathbb{P}\left\{p_1(\xi) \leq n^{1/a}\right\} \overset{n\to\infty}{\longrightarrow} \mathcal{D}(a) = \int_a^\infty d(u)\, du \ , \tag{1.5}$$

where ξ is a random integer with uniform distribution on $\{1, ..., n\}$ and $p_1(\xi)$ is it's smallest prime factor. It is well established that this Dickman distribution is infinitely divisible. The Dickman function $\mathcal{D}(a)$ satisfies the delay differential equation

$$a\mathcal{D}'(a) + \mathcal{D}(a-1) = 0 \qquad \mathcal{D}(a) = 1 \quad \forall a \in [0,1] \tag{1.6}$$

A similar distribution that will be studied is the Penrose-Wade distribution. Consider a random variable $\mathfrak{D}$ defined as

$$\mathfrak{D} = U_1 + U_1U_2 + U_1U_2U_3 + \cdots$$
$$\mathfrak{D} \overset{law}{=} U(1+\mathfrak{D})$$

$\mathfrak{D}$ has a density that satisfies the Dickman equation (1.6) with the initial condition $p(x) = e^{-\gamma}\ \forall\ x \in [0,1]$. It can also be written

$$p(x) = e^{-\gamma}\mathcal{D}(x)$$

where $\mathcal{D}$ again is the Dickman function from (1.6).

Our third and final distribution was discovered by Vasily Leonidovich Goncharov in the 1940s while investigating asymptotics of the type discussed in section 1.1. Let π be a random permutation of n elements and consider the cycle decomposition of π

$$\pi = c_1(\pi)c_2(\pi)\cdots c_j(\pi) \tag{1.7}$$

If we take $c_{(1)}(\pi)$ to be the longest cycle in the decomposition, we have for any $a \in [0,1]$

$$\mathbb{P}\left\{\frac{|c_{(1)}(\pi)|}{n} \leq a\right\} \overset{n\to\infty}{\longrightarrow} G(a) = \int_0^a g(u)\,du \tag{1.8}$$

Where $G(a)$ is the distribution function for the Goncharov law. Notably if a random variable X has the Goncharov distribution, then $Y = 1/X$ then Y has the distribution function $F(y) = 1 - \mathcal{D}(y)$ where $\mathcal{D}$ is the Dickman function from (1.6). This remarkable connection between the two distributions was not discovered until the 1980s by Vershik in [8].

CHAPTER 2: JOINT DISTRIBUTION OF SEVERAL HIGHER ORDER CYCLES

2.1 Preliminaries

Here we will derive two important identities for the Goncharov distribution. These identities appear multiple times in literature that predates this paper. The reader is directed to [4]. Consider a random variable $\mathcal{G}$ that has the distribution

$$\mathcal{G} \stackrel{law}{=} \max\{1-U, \tilde{\mathcal{G}}U\}$$

Where U is uniformly distributed on $[0,1]$, $\mathcal{G}$ and $\tilde{\mathcal{G}}$ are independent and have the same distribution. Note that this distribution is analogous to the situation in (1.8). Because $\mathcal{G}$ is concentrated on $[0,1]$ we can write the following equation with $\Phi(x)$ representing the distribution of $\mathcal{G}$

$$\mathbb{P}\{\mathcal{G} \leq x\} = \Phi(x) = \int_0^1 \mathbb{P}\{\max\{t, (1-t)\mathcal{G}\} \leq x\}dt$$
$$\Phi(x) = \int_0^x \Phi\Big(\frac{x}{1-t}\Big)dt \ .$$

Here we will substitute $z = \frac{x}{1-t}$ and we obtain

$$\frac{\Phi(x)}{x} = \int_x^{\frac{x}{1-x}} \frac{\Phi(z)}{z^2}dz$$

If we let $\phi(x) = \Phi'(x)$ be the density of $\mathcal{G}$ we can take differentiate both sides and we arrive at the first Goncharov identity:

$$x\phi(x) = \Phi\left(\frac{x}{1-x}\right) \ . \tag{2.1}$$

To arrive at the second Goncharov identity, we differentiate both sides of (2.1).

$$\begin{aligned} x\phi(x) + \phi'(x) &= \phi\big(\frac{x}{1-x}\big)\frac{1}{(1-x)^2} \\ \phi'(x) &= \phi\big(\frac{x}{1-x}\big)\frac{1}{(1-x)^2} - x\phi(x)\ . \end{aligned} \tag{2.2}$$

Now we can see that the identity below

$$\phi(x)(1-x) = \int_x^{\frac{x}{1-x}} \frac{\phi(z)}{z} dz \tag{2.3}$$

is trivially equal at $x = 0$. We differentiate both sides of (2.3) to arrive at

$$\phi(x) - x\phi'(x) = \frac{\phi\big(\frac{x}{1-x}\big)}{x} - \frac{\phi(x)}{x}$$

Substituting (2.2) for $\phi'(x)$ we get

$$\phi(x)(1-x) = \frac{\phi\big(\frac{x}{1-x}\big)}{(1-x)^2} + \frac{\phi\big(\frac{x}{1-x}\big)}{x} - \frac{\phi(x)}{x}$$

and we now have the verified the identity in (2.3)

$$\phi(x)(1-x) = \int_x^{\frac{x}{1-x}} \frac{\phi(z)}{z} dz$$

The case $k = 2$ was established in [4], and is listed here:

$$r_2(a_1, a_2) = \frac{1}{a_1(1-a_1)}\phi\left(\frac{a_2}{1-a_1}\right) \tag{2.4}$$

where ϕ is the Dickman-Goncharov distribution.

2.2 The case k=3

Here we continue the work of Molchanov and Panov [4] and also Vershik [3]. We seek to prove a joint distribution for the lengths of the three highest order cycles in a permutation selected at uniformly at random from S_n. We utilize the relationship

$$\left\{Y_{(1)}, Y_{(2)}, Y_{(3)}\right\} \stackrel{law}{=} \left\{3 \text{ largest of: } X_1, (1-X_1)\tilde{Y}_1, (1-X_1)\tilde{Y}_2, (1-X_1)\tilde{Y}_3\right\}$$

where X_1 and $\tilde{Y}_1$ are independent, uniformly distributed random variables on $[0,1]$. As the argument is laid out in [4], we designate events:

$$\begin{aligned}
&\Omega_1 : \left\{3 \text{ largest elements - } (1-X_1)\tilde{Y}_1, (1-X_1)\tilde{Y}_2, (1-X_1)\tilde{Y}_3\right\} && X_1 < (1-X_1)\tilde{Y}_3\\
&\Omega_2 : \left\{3 \text{ largest elements - } (1-X_1)\tilde{Y}_1, (1-X_1)\tilde{Y}_2, X_1\right\} && (1-X_1)\tilde{Y}_3 < X_1 < (1-X_1)\tilde{Y}_2\\
&\Omega_3 : \left\{3 \text{ largest elements - } (1-X_1)\tilde{Y}_1, X_1, (1-X_1)\tilde{Y}_2\right\} && (1-X_1)\tilde{Y}_2 < X_1 < (1-X_1)\tilde{Y}_1\\
&\Omega_4 : \left\{3 \text{ largest elements - } X_1, (1-X_1)\tilde{Y}_1, (1-X_1)\tilde{Y}_2\right\} && (1-X_1)\tilde{Y}_1 < X_1
\end{aligned}$$

From [4] we denote the joint distribution for $\left\{Y_{(1)}, Y_{(2)}, Y_{(3)}\right\}$ as:

$$r(a_1,a_2,a_3)da_1da_2da_3 = \mathbb{P}\left\{Y_{(1)} \in (a_1, a_1+da_1), Y_{(2)} \in (a_2, a_2+da_2), Y_{(3)} \in (a_3, a_3+da_3)\right\} \tag{2.5}$$

and additionally we designate the right hand side of (2.5) as Ω_0.

Theorem 2.1. *The density for* $\left\{Y_{(1)}, Y_{(2)}, Y_{(3)}\right\}$, $a_1 > a_2 > a_3$ *is*

$$r(a_1,a_2,a_3) = \frac{1}{a_1a_2(1-a_1-a_2)}\phi\left(\frac{a_3}{1-a_1-a_2}\right) \tag{2.6}$$

Proof. We first show that (2.6) is a valid density function on $[0,1]$.

$$\int_0^1\int_0^1\int_0^1 r(a_1,a_2,a_3)da_1da_2da_3 = \int_0^1\int_0^1\int_0^1 \frac{1}{a_1a_2(1-a_1-a_2)}\phi\left(\frac{a_3}{1-a_1-a_2}\right)da_1da_2da_3$$

$$= \int_0^1 \frac{1}{a_1}\int_0^{a_1}\frac{1}{a_2}\int_0^{a_2}\frac{1}{(1-a_1-a_2)}\phi\left(\frac{a_3}{1-a_1-a_2}\right)da_1da_2da_3$$

let $t = \dfrac{a_3}{1-a_1-a_2}$

$$= \int_0^1 \frac{1}{a_1}da_1\int_0^{a_1}\frac{1}{a_2}da_2\int_0^{\frac{a_2}{1-a_1-a_2}}\phi(t)\,dt$$

let $\gamma = \dfrac{a_2}{1-a_1}$

$$= \int_0^1 \frac{1}{a_1}da_1\int_0^{a_1}\frac{1}{a_2}da_2\int_0^{\frac{\gamma}{1-\gamma}}\phi(t)\,dt$$

$$= \int_0^1 \frac{1}{a_1}da_1\int_0^{a_1}\frac{1}{a_2}\Phi\left(\frac{\gamma}{1-\gamma}\right)da_2$$

and by (2.1)

$$= \int_0^1 \frac{1}{a_1}da_1\int_0^{a_1}\frac{1}{1-a_1}\phi\left(\frac{a_2}{1-a_1}\right)da_2$$

$$= \int_0^1 \phi(a_1)\,da_1 = 1$$

We will construct the formula for $r(a_1,a_2,a_3)$ using the cases outlined above. We proceed with the individual cases:

1. $\Omega_0 \cap \Omega_1$

$$= \left\{(1-X_1)\tilde{Y}_1 \in (a_1, a_1+da_1), (1-X_1)\tilde{Y}_2 \in (a_2, a_2+da_2), (1-X_1)\tilde{Y}_3 \in (a_3, a_3+da_3)\right\}$$

$$X_1 < (1-X_1)\tilde{Y}_3$$

hence,

$$r_{(1)}(a_1,a_2,a_3)=\int_0^{a_3}\frac{dx}{(1-x)^3}r\left(\frac{a_1}{1-x},\frac{a_2}{1-x},\frac{a_3}{1-x}\right)$$
$$=\frac{1}{a_1a_2}\int_0^{a_3}\frac{1}{1-a_1-a_2-x}\phi\left(\frac{a_3}{1-a_1-a_2-x}\right)dx$$

let $t=\frac{a_3}{1-a_1-a_2-x}$

$$=\frac{1}{a_1a_2}\int_{\frac{a_3}{1-a_1-a_2}}^{\frac{a_3}{1-a_1-a_2-a_3}}\frac{1}{t}\phi\left(t\right)dt$$

let $\beta=\frac{a_3}{1-a_1-a_2}$, note also that $\frac{\beta}{1-\beta}=\frac{a_3}{1-a_1-a_2-a_3}$,

$$=\frac{1}{a_1a_2}\int_{\beta}^{\frac{\beta}{1-\beta}}\frac{\phi\left(t\right)}{t}dt$$

by equation (2.3),

$$=\frac{1}{a_1a_2}\left(1-\beta\right)\phi(\beta)$$
$$r_{(1)}(a_1,a_2,a_3)=\frac{1-a_1-a_2-a_3}{a_1a_2(1-a_1-a_2)}\phi\left(\frac{a_3}{1-a_1-a_2}\right)\tag{2.7}$$

2. $\Omega_0\cap\Omega_2$

$$=\left\{(1-X_1)\tilde{Y}_1\in(a_1,a_1+da_1),(1-X_1)\tilde{Y}_2\in(a_2,a_2+da_2),X_1\in(a_3,a_3+da_3)\right\},X_1>(1-X$$

So we have:

$$r_{(2)}(a_1,a_2,a_3)=\frac{1}{(1-a_3)^2}\int_0^{\frac{a_3}{1-a_3}}r\left(\frac{a_1}{1-a_3},\frac{a_2}{1-a_3},x\right)dx$$

By equation (2.6),

$$=\frac{1}{(1-a_3)^2}\int_0^{\frac{a_3}{(1-a_3)}}\frac{(1-a_3)^3}{a_1a_2(1-a_1-a_2-a_3)}\phi\left(\frac{x(1-a_3)}{1-a_1-a_2-a_3}\right)$$

let $t = \frac{x(1-a_3)}{1-a_1-a_2-a_3}$ and we have:

$$= \frac{1}{a_1 a_2} \int_0^{\frac{a_3}{1-a_1-a_2-a_3}} \phi(t)dt$$

and now let $\gamma = \frac{a_3}{1-a_1-a_2}$,

$$= \frac{1}{a_1 a_2} \int_0^{\frac{\gamma}{1-\gamma}} \phi(t)dt$$

by equation (2.3),

$$= \frac{1}{a_1 a_2} \Phi\left(\frac{\gamma}{1-\gamma}\right)$$

and by (2.1) we can say

$$= \frac{1}{a_1 a_2} \gamma \phi(\gamma)$$

hence

$$r_{(2)}(a_1, a_2, a_3) = \frac{a_3}{a_1 a_2 (1-a_1-a_2)} \phi\left(\frac{a_3}{1-a_1-a_2}\right) \tag{2.8}$$

3. $\Omega_0 \cap \Omega_3$

$$= \left\{(1-X_1)\tilde{Y}_1 \in (a_1, a_1+da_1), X_1 \in (a_2, a_2+da_2), (1-X_1)\tilde{Y}_2 \in (a_3, a_3+da_3)\right\}$$

$$r_{(3)}(a_1, a_2, a_3) = \frac{1}{(1-a_2)^2} r_2\left(\frac{a_2}{1-a_1}, \frac{a_3}{1-a_1}\right)$$

and by (2.4)

$$r_{(3)}(a_1, a_2, a_3) = \frac{1}{a_2(1-a_1-a_2)} \phi\left(\frac{a_3}{1-a_1-a_2}\right) \tag{2.9}$$

4. $\Omega_4 \cap \Omega_0$

$$= \left\{ X_1 \in (a_1, a_1 + da_1), (1 - X_1)\tilde{Y}_1 \in (a_2, a_2 + da_2), (1 - X_1)\tilde{Y}_2 \in (a_3, a_3 + da_3) \right\}$$

$$r_{(4)} = \frac{1}{(1 - a_1)^2} r_2 \left(\frac{a_2}{1 - a_1}, \frac{a_3}{1 - a_1} \right)$$

and by (2.4)

$$r_{(4)} = \frac{1}{a_1(1 - a_1 - a_2)} \phi \left(\frac{a_3}{1 - a_1 - a_2} \right) \tag{2.10}$$

Now it remains to combine equations (2.7) - (2.10):

$$\begin{aligned} r(a_1, a_2, a_3) &= r_{(1)} + r_{(2)} + r_{(3)} + r_{(4)} \\ &= \left(\frac{1 - a_1 - a_2 - a_3}{a_1 a_2 (1 - a_1 - a_2)} + \frac{a_3}{a_1 a_2 (1 - a_1 - a_2)} + \frac{1}{a_2 (1 - a_1 - a_2)} \right. \\ &\quad \left. + \frac{1}{a_1 (1 - a_1 - a_2)} \right) \phi \left(\frac{a_3}{1 - a_1 - a_2} \right) \\ r(a_1, a_2, a_3) &= \frac{1}{a_1 a_2 (1 - a_1 - a_2)} \phi \left(\frac{a_3}{1 - a_1 - a_2} \right) \end{aligned}$$

which concludes the proof. □

CHAPTER 3: RANDOM GEOMETRIC PROGRESSIONS

In this chapter we consider the infinite divisibility of particular classes of random variables that arise from a random geometric progression of the type

$$S = Y_1 + Y_1Y_2 + Y_1Y_2Y_3 + Y_1Y_2Y_3Y_4 + \cdots \ , \tag{3.1}$$

where the Y_i are independent, identically distributed random variables. As shown by Penrose and Wade [1], when the Y_i are distributed uniformly on the interval $[0, 1]$, the resulting limit is an infinitely divisible distribution.

We begin by showing some relationships that will be used throughout the chapter. The random variable S has the following distribution equality

$$S \stackrel{law}{=} Y_1(1 + \tilde{S}) \ , \tag{3.2}$$

where S and $\tilde{S}$ are independent and equal in distribution. and from this we can say

$$\mathbb{E}(S) = \mathbb{E}(Y_1)\mathbb{E}(1 + S). \tag{3.3}$$

Using this relation, we can calculate all moments of S

$$\begin{aligned} M_n = \mathbb{E}(S^n) &= \mathbb{E}(Y_1^n)\mathbb{E}((1+S)^n) \\ &= \mathbb{E}(Y_1^n)\sum_{j=0}^{n-1}\binom{n}{j}M_j + \mathbb{E}(Y_1^n)M_n \\ M_n &= \frac{\mu_n}{1-\mu_n}\sum_{j=0}^{n-1}\binom{n}{j}M_j \qquad (3.4) \\ &\text{with } M_0 = 1 \end{aligned}$$

where μ_n denotes the nth moment of Y_i. This relationship seems to have deep connections with the tail probabilities of infinitely divisible distributions shown in [14], [16], and [17].

As will be seen later in the paper, cumulants play an important role in infinitely divisible distributions. The cumulant generating function of a distribution is defined as

$$K(t) = \log \mathbb{E}(e^{tX}) .$$

We can calculate individual cumulants κ_i by the Maclaurin series

$$K(t) = \sum_{i=1}^{\infty} \kappa_i \frac{t^n}{n!} = \kappa_1 \frac{t}{1!} + \kappa_2 \frac{t^2}{2!} + \cdots$$

Similarly to moments and the moment generating function we can calculate

$$\kappa_n = K^{(n)}(0) .$$

In particular, we should note that $\kappa_1 = \mu_1 = \mathbb{E}(X)$ and that $\kappa_2 = \sigma^2 = Var(X)$.

Cumulants are so-named because of the property

$$\kappa_n(X_1 + X_2 + \cdots + X_j) = \kappa_n(X_1) + \kappa_n(X_2) + \cdots + \kappa_n(X_j)$$

We will make extensive use of the relationship between moments and cumulants from [18],

$$\kappa_n = M_n - \sum_{j=1}^{n-1} \binom{n-1}{j} \kappa_{n-j} M_j \ ; \tag{3.5}$$

Where κ_n denotes the nth cumulant of the distribution.

Before we get into specific distributions, we will demonstrate an important relationship between the moments μ_n of Y_i and the moments of the resulting random geometric progression M_n using (3.4) .

$$M_n = \frac{\mu_n}{1-\mu_n} \sum_{j=0}^{n-1} \binom{n}{j} M_j, \quad M_0 = 1 \tag{3.6}$$

and here we calculate the first moment, M_1

$$M_1 = \frac{\mu_1}{1-\mu_1} M_0$$
$$M_1 = \frac{\mu_1}{1-\mu_1}$$

and to simplify notation we will introduce the term

$$a_n = \frac{\mu_n}{1-\mu_n} \ .$$

So we can rewrite

$$M_1 = a_1 \ .$$

Continuing with (3.4) recursively, we can solve for M_2

$$M_2 = \frac{\mu_2}{1-\mu_2} \sum_{j=0}^{1} \binom{2}{j} M_j$$
$$M_2 = a_2 + \binom{2}{1} a_1 a_2 \ .$$

Again we can use (3.4), we can find M_3

$$M_3 = a_3 \sum_{j=0}^{2} \binom{3}{j} M_j$$
$$M_3 = a_3 \left(\binom{3}{0} M_0 + \binom{3}{1} M_1 + \binom{3}{2} M_2 \right)$$
$$M_3 = a_3 + \binom{3}{1} a_3 a_1 + \binom{3}{2} a_3 a_2 + \binom{3}{2}\binom{2}{1} a_3 a_2 a_1$$
$$M_3 = a_3 \sum_{j=1}^{2} \binom{3}{j} a_3 a_j + \sum_{j=2}^{3} \sum_{i=1}^{j} \binom{3}{j}\binom{j}{i} a_3 a_j a_i \ .$$

From there we can generalize the pattern to finding any M_n without the need for recursion. We arrive at the formula

$$M_n = a_n + \sum_{j=1}^{n} \binom{n}{j} a_n a_j + \sum_{j=2}^{n} \sum_{i=1}^{j} \binom{n}{j}\binom{j}{i} a_n a_j a_1$$
$$+ \cdots + \sum_{i_k=k}^{n} \sum_{i_{k-1}}^{i_k} \cdots \sum_{i_1}^{i_2} \binom{n}{i_k}\binom{i_k}{i_{k-1}} \cdots \binom{i_2}{i_1} a_{i_1} a_{i_2} \cdots a_{i_k} a_n \tag{3.7}$$

which gives moments M_n of the random geometric progression based only on the moments μ_n of Y_i. We will use equation (3.7) extensively to calculate moments of various infinite geometric progressions.

Consider, as an example, the infinite progression studied in the paper by Vervaat [5] and Penrose and Wade [1], where

$$S = Y_1 + Y_1Y_2 + Y_1Y_2Y_3 + \cdots$$

and

$$Y_i \sim Unif[0, 1] \ .$$

We know this density to possess infinite divisibility, and we will use our new method based on equation (3.7) to demonstrate. We begin by calculating

$$\mu_n = \frac{1}{n+1} \ ,$$

which gives

$$a_n = \frac{1}{n} \ .$$

Here we will deviate slightly from the method slightly to show an important property of S defined as by Penrose Wade. Recall the cumulants to moments equation (3.5):

$$\kappa_n = M_n - \sum_{j=1}^{n-1} \binom{n-1}{j} \kappa_{n-j} M_j \ .$$

To simplify the notation, consider the situation when $n = 3$

$$\begin{aligned}\kappa_3 &= M_3 - \sum_{j=1}^{2} \binom{2}{j} \kappa_{3-j} M_j \\ &= a_3 + \binom{3}{1} a_3 a_1 + \binom{3}{2} a_3 a_2 + \binom{3}{2}\binom{2}{1} a_3 a_2 a_1 - \binom{2}{1} \kappa_2 a_1 - \kappa_1 \big(a_2 + 2a_1 a_2\big).\end{aligned}$$

Using the recursion for κ_n, we can insert the values for κ_1 and κ_2

$$\begin{aligned}\kappa_3 &= a_3 + \binom{3}{1} a_3 a_1 + \binom{3}{2} a_3 a_2 + \binom{3}{2}\binom{2}{1} a_3 a_2 a_1 \\ &\quad + \binom{2}{1} \big(a_2 + 2a_1 a_2 - a_1 a_1\big) a_1 + a_1 \big(a_2 + 2a_1 a_2\big). \\ &= \frac{1}{3} + \binom{3}{1} \frac{1}{3}\frac{1}{2} + \binom{3}{2}\binom{2}{1} \frac{1}{3}\frac{1}{2} - \binom{2}{1}\frac{1}{2} - \binom{2}{1} - \binom{2}{1} - \frac{1}{2} + 1 \\ \kappa_3 &= \frac{1}{3}\ .\end{aligned}$$

This demonstrates the unique quality of a geometric progression based on the Y_i having uniform density on $[0, 1]$, which can be extended to any cumulant. Because the nth moment for Y_i is $\frac{1}{n+1}$, produces coefficients a_n exactly equal to $\frac{1}{n}$ for every n.

When combined with the cumulant formula (3.5), the moments of S cancel exactly the binomial coefficients that occur in the recursion formulas for cumulants and moments. As we will show in the paper, if those moments are changed slightly, the resulting variable $\tilde{S}$ loses the property of infinite divisibility.

Now we will show the method novel to this paper. Using the equation (3.7) we will demonstrate that the S defined as above has moments that are not decreasing sufficiently quickly to produce negative moments. Recall that if a random variable distributed on $\mathbb{R}^+$ has negative cumulants, it is not infinitely divisible. For details the reader is referred to [14] chap. III, Theorem 7.1.

To use our method, we start with the coefficient corresponding to $\mu_n = \frac{1}{n+1}$

$$a_n = \frac{1}{n}$$

and combine it with the equation (3.7) to get

$$M_n = \frac{1}{n} + \sum_{j=1}^{n} \frac{\binom{n}{j}}{jn} + \sum_{j=2}^{n} \sum_{i=1}^{j-1} \frac{\binom{n}{j}\binom{j}{i}}{jni} + \cdots$$

Here we split the summands of each of the series:

$$\begin{aligned} T_1 &= \frac{1}{n} \\ T_2 &= \frac{\binom{n}{j}}{jn} \\ T_3 &= \frac{\binom{n}{j}\binom{j}{i}}{jni} \\ &\vdots \\ T_k &= \frac{\binom{n}{i_k} \cdots \binom{i_2}{i_1}}{i_k \cdots i_2 i_1 n} \\ &\vdots \end{aligned}$$

Now we perform a Taylor series expansion of each of the T_i terms about an arbitrary

point k:

$$
\begin{aligned}
T_1 &= \frac{1}{k} - \frac{1}{k^2}(n-k) + O\left((n-k)^2\right) \\
T_2 &= \frac{\Gamma(k+1)}{j^2 k \Gamma(j)\Gamma(-j+k+1)} \\
&\quad + \frac{\Gamma(k+1)\big(k(\psi^{(0)}(k+1) - \psi^{(0)}(k+1-j)) - 1\big)}{j^2 k^2 \Gamma(j)\Gamma(-j+k+1)}(n-k) + O\left((n-k)^2\right) \\
T_3 &= \frac{\Gamma(k+1)\binom{j}{i}}{ij^2 k \Gamma(j)\Gamma(-j+k+1)} \\
&+ \frac{\Gamma(k+1)\binom{j}{i}(-k\psi^{(0)}(-j+k+1) + k\psi^{(0)}(k+1) - 1)}{ij^2 k^2 \Gamma(j)\Gamma(-j+k+1)}(n-k) + O\left((n-k)^2\right) \\
&\vdots
\end{aligned}
$$

Where $\phi^{(0)}(z)$ is the digamma function, defined as

$$\phi^{(0)} = \frac{\Gamma'(z)}{\Gamma(z)} \ . \tag{3.8}$$

We look only at the first order terms in the expansion of T_n to get a sense of how quickly moments are changing. Recall for terms T_2 and higher, we must sum over all $j < k$. Examining T_2, we can produce the second coefficient in the series expansion

$$T_2 = \frac{\Gamma(k+1)\big(k(\psi^{(0)}(k+1) - \psi^{(0)}(k+1-j)) - 1\big)}{j^2 k^2 \Gamma(j)\Gamma(-j+k+1)}$$

We should note here that T_2 increases with k, which we should expect. The important part of the dynamics here is that for any k, the second term in the Taylor expansion is positive, and finite. From here we can say that

$$\frac{dM}{dn}\Big|_{n=k} > 0 \ . \tag{3.9}$$

This produces a sequence of moments that are increasing with n at a rate that will not produce negative cumulants from (3.5).

Here we will show that if a random variable Y_i has moments that produce coefficients a_n which are grow infinitesimally larger than $\frac{1}{n}$, then the random variable based on the infinite progression of Y_i is not infinitely divisible.

Theorem 3.1. *Let Y_i be a sequence of i.i.d. random variables with the moment of each Y_i defined as follows*

$$\mu_n = \frac{1}{n^{1+\epsilon}+1} \ .$$

Then the random variable S

$$S = Y_1 + Y_1Y_2 + Y_1Y_2Y_3 + Y_1Y_2Y_3Y_4 + \cdots$$

is not infinitely divisible.

Proof. Again we construct the moments of the geometric progression based on $\mu_n = \frac{1}{n^{1+\epsilon}+1}$,

$$\begin{aligned} a_n &= \frac{\mu_n}{1-\mu_n} \\ &= \frac{\frac{1}{n^{1+\epsilon}+1}}{1-\frac{1}{n^{1+\epsilon}+1}} \\ &= \frac{1}{n^{1+\epsilon}} \end{aligned}$$

Now we can find a formula for M_n, the nth moment of the progression via the formula

$$\begin{aligned} M_n &= a_n + \sum_{j=1}^{n} \binom{n}{j} a_n a_j + \sum_{j=2}^{n} \sum_{i=1}^{j} \binom{n}{j}\binom{j}{i} a_n a_j a_1 + \cdots \\ &= \frac{1}{n^{1+\epsilon}} + \sum_{j=1}^{n-1} \binom{n}{j} \frac{1}{(nj)^{1+\epsilon}} + \sum_{j=2}^{n-1} \sum_{i=1}^{j-1} \binom{n}{j}\binom{j}{i} \frac{1}{(inj)^{1+\epsilon}} + \cdots \end{aligned}$$

And here we perform our Taylor series expansion about the point $\epsilon = 0$ for $T_1 = \frac{1}{n^{1+\epsilon}}$

$$T_1 = \frac{1}{n} - \frac{\log(n)}{n}\epsilon + O\left(\epsilon^2\right) \ .$$

Similarly we find expansions for T_2 and T_3

$$\begin{aligned} T_2 &= \frac{\binom{n}{j}}{jn} - \frac{\binom{n}{j}\log(jn)}{jn}\epsilon + O\left(\epsilon^2\right) \\ T_3 &= \frac{\binom{j}{i}\binom{n}{j}}{nij} - \frac{\binom{j}{i}\binom{n}{j}\log(nij)}{nij}\epsilon + O\left(\epsilon^2\right) \\ &\vdots \end{aligned}$$

This provides with the first order coefficient in the Taylor Series expansion for M_n about $\epsilon = 0$

$$\frac{dM_n}{d\epsilon}\Big|_{\epsilon=0} = \sum_{n=0}^{\infty}\left(\sum_{j=1}^{n-1} \frac{-\log(n)}{n} - \sum_{j=2}^{n-1}\sum_{i=1}^{j-1} \frac{\binom{j}{i}\binom{n}{j}\log(nij)}{nij} - \cdots\right)$$

Further T_i have similarly negative first-order terms, which brings us to our conclusion. Since the first order terms are negative and increasing without bound around the point $\epsilon = 0$, since the M_n are strictly positive, we can conclude that the sequence M_n converges pointwise to 0 for any $\epsilon > 0$. Hence there exists an N such that provides a sufficiently small M_N to produce a negative K_N. It follows that S defined in this

way is not infinitely divisible. □

3.1 Uniform Type Distributions

We start with the case where each Y_i follows the distribution

$$Y_i \sim \begin{cases} 0 & q \\ Unif[0,1] & p \ . \end{cases}$$

As we will see soon, this distribution has strong connections to distributions that are uniform on sub-intervals of $[0, 1]$.

Theorem 3.2. *The random variable S,*

$$S = Y_1 + Y_1Y_2 + Y_1Y_2Y_3 + Y_1Y_2Y_3Y_4 + \cdots$$

with

$$Y_i \sim \begin{cases} 0 & q \\ Unif[0,1] & p \ . \end{cases}$$

is not infinitely divisible.

Proof. To begin, we first calculate moments, μ_n of Y_i:

$$\mu_n = \mathbb{E}(Y_i) = q \cdot 0 + p \cdot \left(\frac{1}{1+n}\right)$$
$$\mu_n = \frac{p^n}{1+n} \ .$$

To simplify notation, we will create the coefficient

$$\begin{aligned} a_n &= \frac{\mu_n}{1-\mu_n} \\ a_n &= \frac{p^n}{(n+1)-p^n} \ . \end{aligned} \tag{3.10}$$

We now calculate the moments of $S = Y_1 + Y_1Y_2 + Y_1Y_2Y_3 + \cdots$ as a function of p, the probability that Y_i is non-zero. Using the formula (3.4) gives the following recursion relation.

$$\begin{aligned} M_n(p) &= \frac{\mu_n}{1-\mu_n} \sum_{j=0}^{n-1} \binom{n}{j} M_j \\ &= a_n \sum_{j=0}^{n-1} \binom{n}{j} M_j(p) \\ &= \frac{p^n}{(n+1)-p^n} \sum_{j=0}^{n-1} \binom{n}{j} M_j(p) \end{aligned}$$

Solving for the recursion yields

$$\begin{aligned} M_n(p) =& \frac{p^n}{(n+1)-p^n} + \sum_{j=1}^{n-1} \frac{\binom{n}{j} p^{n+j}}{((n+1)-p^n)((j+1)-p^j)} \\ &+ \sum_{j=2}^{n-1} \sum_{i=1}^{j-1} \frac{\binom{n}{j}\binom{j}{i} p^{n+i+j}}{((n+1)-p^n)((j+1)-p^j)((i+1)-p^i)} + \\ &+ \sum_{a_k=k}^{n} \sum_{a_{k-1}=k-1}^{a_k} \cdots \sum_{a_1=1}^{a_2} \frac{\binom{n}{a_k}\binom{a_k}{a_{k-1}} \cdots \binom{a_2}{a_1} p^{\sum a_i + k}}{((n+1)-p^n)((a_k+1)-p^a_{k-1}) \cdots ((a_1+1)-p^{a_1})} \end{aligned} \tag{3.11}$$

We should take note of several important facts about $M_n(p)$: It is analytic on the interval $(0,1)$, it is non-negative on [0,1] and it's only real root on the interval $[0,1]$ is $M_n(0) = 0$ for all n. With these facts in hand, we proceed by performing a Taylor

expansion about the point $p = 1$ for each of the terms in (3.11).

$$\begin{aligned}
T_1 &= \frac{p^n}{(n+1-p^n)} \\
&= \left(\frac{1}{n} + \left(1 + \frac{1}{n}\right)(p-1) + O\left((p-1)^2\right)\right) \\
T_2 &= \sum_{j=1}^{n-1} \frac{\binom{n}{j} p^{n+j}}{((n+1)-p^n)((j+1)-p^j)} \\
&= \sum_{j=1}^{n-1} \left(\frac{\binom{n}{j}}{j} + \frac{\binom{n}{j}(j+n+2)}{j}(p-1) + O\left((p-1)^2\right)\right) \\
T_3 &= \sum_{j=2}^{n-1}\sum_{i=1}^{j-1} \frac{\binom{n}{j}\binom{j}{i} p^{n+i+j}}{((n+1)-p^n)((j+1)-p^j)((i+1)-p^i)} \\
&= \sum_{j=2}^{n-1}\sum_{i=1}^{j-1} \frac{\binom{n}{j}\binom{j}{i}}{ijn} + \frac{\binom{n}{j}\binom{j}{i}(i+j+n+3)}{ijn}(p-1) + O\left((p-1)^2\right)
\end{aligned}$$

Recombining the terms yields

$$\begin{aligned}
M_n(p) =& T_1 + T_2 + T_3 + \cdots \\
&= \left(\frac{1}{n} + \sum_{j=1}^{n-1} \frac{\binom{n}{j}}{j} + \sum_{j=2}^{n-1}\sum_{i=1}^{j-1} \frac{\binom{n}{j}\binom{j}{i}}{ijn}\right) \\
&\quad + \left(1 + \frac{1}{n}\frac{\binom{n}{j}(j+n+2)}{j} + \sum_{j=2}^{n-1}\sum_{i=1}^{j-1} \frac{\binom{n}{j}\binom{j}{i}(i+j+n+3)}{ijn}\right)(p-1) + O\left((p-1)^2\right)
\end{aligned}$$

Further terms T_n can be calculated in the same way–although it is very tedious. The other terms produce the same important properties. Our Taylor series expansion provides a key insight into the behavior of the sequence of moments $M_n(p)$, namely that the first term of the is consistently positive. This indicates a positive derivative, increasing with n, for every M_n. Combining this with properties discussed above yields pointwise convergence in the limit, $M_n(p) \to 0$ as $n \to \infty$. Now for any $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that $M_N(p) < \epsilon,\ \forall p \in [0,1]$. $\square$

When we consider the cumulants-to-moments formula,

$$\kappa_n = M_n - \sum_{j=1}^{n-1} \binom{n-1}{j} \kappa_{n-j} M_j \ ,$$

we can now guarantee the existence of an N such that M_N is small enough to produce a corresponding negative κ_N. It follows that S is not infinitely divisible for any p.

Here we consider the scenario where Y_i is distributed uniformly on particular subintervals of [0,1]. We begin by considering

$$Y_i \sim Unif[0, 1-\epsilon] \ .$$

It is important to note here that since it has finite support, the non-infinite divisibility of this random variable is already known. We proceed here to demonstrate the method novel to this paper.

Theorem 3.3. *The random variable S,*

$$S = Y_1 + Y_1Y_2 + Y_1Y_2Y_3 + Y_1Y_2Y_3Y_4 + \cdots$$

with $Y_i \sim Unif[0, 1-\epsilon]$ is not infinitely divisible.

Proof. Recall that the moments of Y_i follow the relation

$$\mu_n = \frac{(1-\epsilon)^n}{1+n} \ .$$

To simplify notation, we will create the coefficient

$$\begin{aligned} a_n &= \frac{\mu_n}{1-\mu_n} \\ a_n &= \frac{(1-\epsilon)^n}{(n+1)-(1-\epsilon)^n} \ . \end{aligned} \tag{3.12}$$

This demonstrates a remarkable similarity between S as defined here and S from the previous section (in which $Y_i \sim Unif[0,1]$ with probability p and 0 with probability q). Setting $(1-\epsilon) = p$, the two distributions have identical moments M_n for all n. It immediately follows that our S with $Y_i \sim Unif[0, 1-\epsilon]$ is also not infinitely divisible for any $\epsilon > 0$. □

3.2 Beta Distributed Random Variables

In this chapter we will consider the infinite divisibility of the random variable S, $S = Y_1 + Y_1Y_2 + Y_1Y_2Y_3 + Y_1Y_2Y_3Y_4 + \cdots$ where each Y_i follows a Beta distribution with $\beta > 1$, $Y_i \sim Beta[\alpha, \beta]$. For simplicity of calculation, we will consider $Y_i \sim Beta[1, \beta]$ WLOG. Note that Penrose and Wade showed that S is infinitely divisible for when $Y_i \sim Beta[\alpha, 1]$ in [1].

Theorem 3.4. *The random variable S,*

$$S = Y_1 + Y_1Y_2 + Y_1Y_2Y_3 + Y_1Y_2Y_3Y_4 + \cdots \tag{3.13}$$

with $Y_i \sim Beta[1, \beta]$, $\beta > 1$ *is not infinitely divisible.*

Proof. The moments for each of the Y_i are defined by the relation

$$\mu_n(\beta) = \prod_{j=0}^{n-1} \frac{1+j}{1+\beta+j} \ .$$

Which produces the coefficient for the sequence M_n

$$a_n = \frac{\mu_n}{1-\mu_n}$$

$$a_n = \frac{\beta! n!}{(\beta+n)!} \ . \tag{3.14}$$

As in previous sections, we apply (3.4) to determine the sequence of moments of S

$$M_n(\beta) = a_n \sum_{j=0}^{n-1} \binom{n}{j} M_j(\beta)$$

Now we have all moments in terms of β:

$$\begin{aligned} M_n(\beta) = \frac{\beta! n!}{(\beta+n)!} + \sum_{j=1}^{n} \frac{\binom{n}{j}(\beta!)^2 n! j!}{(\beta+n)!(\beta+j)!} + \sum_{j=2}^{n} \sum_{i=1}^{j-1} \frac{\binom{n}{j}\binom{j}{i} n! j! i! (\beta!)^3}{(\beta+n)!(\beta+j)!(\beta+i)!} \\ + \sum_{a_k=k}^{n} \sum_{a_{k-1}=k-1}^{a_{k-1}} \cdots \sum_{a_1=1}^{a_2-1} \frac{\binom{n}{a_k}\binom{a_k}{a_{k-1}} \cdots \binom{a_2}{a_1} n! a_k! \cdots a_1! (\beta!)^n}{(\beta+n)!(\beta+a_k)! \cdots (\beta+a_1)!} \end{aligned} \tag{3.15}$$

Following the method established in the previous sections, we examine the first three terms in the sum,

$$\begin{aligned} T_1 &= \frac{\beta! n!}{(\beta+n)!} \\ T_2 &= \frac{\binom{n}{j}(\beta!)^2}{(\beta+n)!(\beta+j)!} \\ T_3 &= \frac{\binom{n}{j}\binom{j}{i} n! j! i! (\beta!)^3}{(\beta+n)!(\beta+j)!(\beta+i)!} \end{aligned}$$

And for each of the terms, we examine the Taylor series around $\beta = 1$:

$$
\begin{aligned}
T_1 &= \frac{1}{(n+1)} + \frac{n!(\gamma - 1 + \psi^{(0)}(n+2))}{n+1!}(\beta-1) + O\left((\beta-1)^2\right) \\
T_2 &= \frac{\binom{n}{j} n! j!}{(j+1)(n+1)!} \\
&\quad - \frac{\binom{n}{j} n! j! (j\psi^{(0)}(n+2) + 2\gamma j - 2j + \psi^{(0)}(n+2) + 2\gamma - 1)}{(j+1)!(n+1)!}(\beta-1) \\
&\quad + O\left((\beta-1)^2\right) \\
T_3 &= \frac{\binom{n}{j}\binom{j}{i}}{(i+2)!(j+2)!(n+2)!} \\
&\quad - \frac{\binom{n}{j}\binom{j}{i}(\psi^{(0)}(i+2) + \psi^{(0)}(j+2) + \psi^{(0)}(n+2) + 3\gamma - 3)}{(i+2)!(j+2)!(n+2)!}(\beta-1) \\
&\quad + O\left((\beta-1)^2\right)
\end{aligned}
$$

Now we will try to classify the behavior of the first order terms in the series expansion of $M_n(\beta)$.

$$
\begin{aligned}
\frac{dM_n(\beta)}{d\beta}\Big|_{\beta=1} &= \sum_{n=0}^{\infty} \frac{n!(\gamma - 1 + \psi^{(0)}(n+2))}{n+1!} \\
&\quad - \sum_{n=2}^{\infty}\sum_{j=1}^{n} \frac{\binom{n}{j} n! j! (j\psi^{(0)}(n+2) + 2\gamma j - 2j + \psi^{(0)}(n+2) + 2\gamma - 1)}{(j+1)!(n+1)!} \\
&\quad - \sum_{n=3}^{\infty}\sum_{j=2}^{n-1}\sum_{i=1}^{j-1} \frac{\binom{n}{j}\binom{j}{i}(\psi^{(0)}(i+2) + \psi^{(0)}(j+2) + \psi^{(0)}(n+2) + 3\gamma - 3)}{(i+2)!(j+2)!(n+2)!}
\end{aligned}
$$

We can see here again that the linear terms in the expansion around $\beta = 1$ are negative and decreasing without bound and our sequence M_n converges pointwise to 0 for any $\beta > 1$ and the distribution defined this way is not infinitely divisible. □

3.3 Bernoulli type variables

Here we continue the work of determining whether infinitesimally changed distributions preserve infinite divisibility under the random geometric progression.

Theorem 3.5. *Let X be a random variable with the Bernoulli distribution*

$$Y_i = \begin{cases} 0 & p \\ 1 & q \end{cases} \tag{3.16}$$

the resulting geometric progression

$$S = Y_1 + Y_1Y_2 + Y_1Y_2Y_3 + Y_1Y_2Y_3Y_4 + \cdots \tag{3.17}$$

is infinitely divisible.

Proof. To show this, we let

$$\phi_{\mathcal{S}}(\lambda) = \mathbb{E}\left(e^{-\lambda\mathcal{S}}\right) = e^{-\lambda Y_i(1+\mathcal{S})} \tag{3.18}$$

$$= p + qe^{-\lambda}\mathbb{E}(e^{-\lambda\mathcal{S}})$$

$$\phi_{\mathcal{S}}(\lambda) = \frac{p}{1 - qe^{-\lambda}} \tag{3.19}$$

Here we can make the substitution $t = \ln\frac{1}{\lambda}$ and $f(\lambda) = \frac{1}{\lambda}\ln\left(\frac{1}{1-qe^{-\lambda}}\right)$ and (3.19) can be represented as

$$\phi_{\mathcal{S}}(\lambda) = e^{-t+tf(\lambda)} \ . \tag{3.20}$$

We should note the representation in (3.20) determines that the distribution is infinitely divisible, see [7]. □

Qualitatively, one can inspect the scenario from (3.17) and (3.21) is a representation of the exponential distribution. It is known that the exponential distribution is infinitely divisible.

In the spirit of the paper, we will construct a random variable Y_i that has a related but different distribution and determine whether the resulting geometric progression

is infinitely divisible.

Theorem 3.6. *The random variable S,*

$$S = Y_1 + Y_1Y_2 + Y_1Y_2Y_3 + Y_1Y_2Y_3Y_4 + \cdots$$

with Y_i defined as

$$Y_i = \begin{cases} 0 & p \\ \frac{1}{2} & \epsilon \\ 1 & q - \epsilon \end{cases} \tag{3.21}$$

is not infinitely divisible.

Proof. Here we see that, with probability ϵ each term in the sequence will be reduced by a factor of $\frac{1}{2}$. We begin with the method by computing the sequence of moments μ_n.

$$\begin{aligned} \mu_n &= \epsilon \left(\frac{1}{2}\right)^n + (q-\epsilon)(1)^n \\ &= \left(\left(\frac{1}{2}\right)^n - 1\right)\epsilon + q \end{aligned}$$

As before, we then compute our coefficient a_n

$$\begin{aligned} a_n &= \frac{\mu_n}{1-\mu_n} \\ &= \frac{\left(2^{-n}-1\right)\epsilon + q}{-\left(2^{-n}-1\right)\epsilon - q + 1} \end{aligned}$$

Now we are able to compute the moments M_n of the resulting geometric progression using (3.7).

$$M_n = a_n + \sum_{j=1}^{n} \binom{n}{j} a_n a_j + \sum_{j=2}^{n} \sum_{i=1}^{j} \binom{n}{j} \binom{j}{i} a_n a_j a_1$$
$$+ \cdots + \sum_{i_k=k}^{n} \sum_{i_{k-1}}^{i_k} \cdots \sum_{i_1}^{i_2} \binom{n}{i_k} \binom{i_k}{i_{k-1}} \cdots \binom{i_2}{i_1} a_{i_1} a_{i_2} \cdots a_{i_k} a_n$$

$$\begin{aligned} M_n = & \frac{(2^{-n}-1)\,\epsilon + q}{-(2^{-n}-1)\,\epsilon - q + 1} \\ & + \sum_{j=1}^{n-1} \frac{\binom{n}{j}\,((2^{-j}-1)\,\epsilon + q)\,((2^{-n}-1)\,\epsilon + q)}{(-(2^{-j}-1)\,\epsilon - q + 1)\,(-(2^{-n}-1)\,\epsilon - q + 1)} \\ & + \sum_{j=2}^{n-1} \left(\sum_{i=1}^{j} \frac{\binom{j}{i}\binom{n}{j}\,((2^{-i}-1)\,\epsilon + q)\,((2^{-j}-1)\,\epsilon + q)\,((2^{-n}-1)\,\epsilon + q)}{(-(2^{-i}-1)\,\epsilon - q + 1)\,(-(2^{-j}-1)\,\epsilon - q + 1)\,(-(2^{-n}-1)\,\epsilon - q + 1)} \right) \\ & + \cdots \end{aligned}$$

Here we split the summands into our terms $T_1, T_2, T_3, ...$

$$T_1 = \frac{(2^{-n}-1)\,\epsilon + q}{-(2^{-n}-1)\,\epsilon - q + 1}$$
$$T_2 = \frac{\binom{n}{j}\,((2^{-j}-1)\,\epsilon + q)\,((2^{-n}-1)\,\epsilon + q)}{(-(2^{-j}-1)\,\epsilon - q + 1)\,(-(2^{-n}-1)\,\epsilon - q + 1)}$$
$$T_3 = \frac{\binom{j}{i}\binom{n}{j}\,((2^{-i}-1)\,\epsilon + q)\,((2^{-j}-1)\,\epsilon + q)\,((2^{-n}-1)\,\epsilon + q)}{(-(2^{-i}-1)\,\epsilon - q + 1)\,(-(2^{-j}-1)\,\epsilon - q + 1)\,(-(2^{-n}-1)\,\epsilon - q + 1)}$$

and perform a Taylor expansion about the point $\epsilon = \frac{1}{3}$

$$T_1 = -\frac{9\left(2^n\left(2^n-1\right)\right)\left(\epsilon-\frac{1}{3}\right)}{\left(3\ 2^n q-2^{n+2}+1\right)^2}-\frac{3\ 2^n q-2^n+1}{3\ 2^n q-2^{n+2}+1}+O\left(\left(\epsilon-\frac{1}{3}\right)^2\right)$$

$$T_2 =$$

$$T_3 =$$

$$\vdots$$

The terms T_2 and T_3 are too large to conveniently fit onto the page, they are listed in the appendix in section A.1

When we combine the first order terms, we can determine $\frac{d}{d\epsilon}M_n$ at $\epsilon = 0$ The first order terms of the series expansion are all negative (the terms are listed in the appendix in section A.2) and decreasing without bound for all $n > 1$, hence M_n converges pointwise to 0 as $n \to \infty$ for all $\epsilon > \frac{1}{3}$.

□

CHAPTER 4: CONCLUSIONS AND FURTHER WORK

The method described in the paper is novel, and there are many possible applications and improvements. Notably, the method currently is only able to prove the negative with regard to infinite divisibility. It is only able to determine cases where the adjustment of a parameter causes moments to converge pointwise to zero on some interval of the domain of the function. There are scenarios indicated by (3.5) where the growth of M_n is slow enough to become smaller in magnitude than sum of the subtracted product of cumulants and moments. The method presented in this paper is not able to determine infinite divisibility in such cases. The dynamics of the moments to cumulant equation

$$\kappa_n = M_n - \sum_{j=1}^{n-1} \binom{n-1}{j} \kappa_{n-j} M_j$$

suggest that there may be a method similar to the one presented here.

There is also significant work left to be done to improve the method with respect to its application to discrete random variables. It seems reasonable to think the result in Theorem 3.6 could be expanded to discrete random variables with arbitrarily many divisions of $[0, 1]$.

CHAPTER A: APPENDIX

A.1 Bernoulli Type Variables 1

The terms T_2 and T_3 from section 3.3 are presented on the following page.

$$T_2 = \Big(9\left(-3q2^{j+n+2} + 21q2^{2j+n} + 5\ 2^{j+n+1} - 9\ 2^{2j+n} - 9\ 2^{j+2n} + 2^{2j+2n+3} - 2^j + 2^{2j} - 2^n + 2^{2n}\right.$$
$$\left. -9q^2 2^{2j+n} - 9q^2 2^{j+2n} + 9q^2 2^{2j+2n+1} - 15q2^{2j+2n+1} + (+21)q2^{j+2n}\right)\binom{n}{j}\Big)$$
$$/((1 - 2^(2 + j) + 32^j q)^2 (1 - 2^(2 + n) + 32^n q)^2)$$

$$T_3 = \frac{(2^{-n} - 1)\left(\frac{1}{3}(2^{-i} - 1) + q\right)\left(\frac{1}{3}(2^{-j} - 1) + q\right)}{\left(\frac{1}{3}(1 - 2^{-i}) - q + 1\right)\left(\frac{1}{3}(1 - 2^{-j}) - q + 1\right)\left(\frac{1}{3}(1 - 2^{-n}) - q + 1\right)}$$
$$+ \frac{1}{3}\left(2^{-n} - 1\right) + q\Bigg(\Bigg(-\frac{9\ 2^n(2^n - 1)\left(\frac{1}{3}(2^{-i} - 1) + q\right)\left(\frac{1}{3}(2^{-j} - 1) + q\right)}{\left(\frac{1}{3}(1 - 2^{-i}) - q + 1\right)\left(\frac{1}{3}(1 - 2^{-j}) - q + 1\right)(3\ 2^n q - 2^{n+2} + 1)^2} +$$
$$\frac{\frac{\left(2^{-j}-1\right)\left(\frac{1}{3}\left(2^{-i}-1\right)+q\right)}{\left(\frac{1}{3}(1-2^{-i})-q+1\right)\left(\frac{1}{3}(1-2^{-j})-q+1\right)} + \left(\frac{1}{3}(2^{-j} - 1) + q\right)\left(\frac{\frac{2^{-i}-1}{\frac{1}{3}(1-2^{-i})-q+1} - \frac{9\ 2^i\left(2^i-1\right)\left(\frac{1}{3}\left(2^{-i}-1\right)+q\right)}{(3\ 2^i q - 2^{i+2}+1)^2}}{\frac{1}{3}(1-2^{-j})-q+1} - \frac{9\ 2^j\left(2^j-1\right)\left(\frac{1}{3}\left(2^{-i}-1\right)+q\right)}{\left(\frac{1}{3}(1-2^{-i})-q+1\right)(3\ 2^j q - 2^{j+2}+1)^2}\right)}{\frac{1}{3}(1 - 2^{-n}) - q + 1}\Bigg)\Bigg)\binom{j}{i}\binom{n}{j}\Bigg)$$

A.2 Bernoulli Type Variables 2

The first order terms from the Taylor series in section 3.3 are listed on the following page.

$$\frac{dM_n(\epsilon)}{d\epsilon}\Big|_{\epsilon=\frac{1}{3}} =$$

$$\frac{9\ 2^n\left(2^n-1\right)}{\left(3\ 2^n q-2^{n+2}+1\right)^2}+\sum_{j=1}^{n-1}\frac{1}{\left(3\ 2^j q-2^{j+2}+1\right)^2\left(3\ 2^n q-2^{n+2}+1\right)^2}-9\left(5\ 2^{j+n+1}-9\ 2^{2j+n}-2^j+2^{2j}-2^n+2^{2n}\right.$$

$$\left.-9q^2 2^{2j+n}-9q^2 2^{j+2n}+9q^2 2^{2j+2n+1}-3q2^{j+n+2}+21q2^{2j+n}+21q2^{j+2n}-15q2^{2j+2n+1}-9\ 2^{j+2n}+2^{2j+2n+3}\right)\binom{n}{j}$$

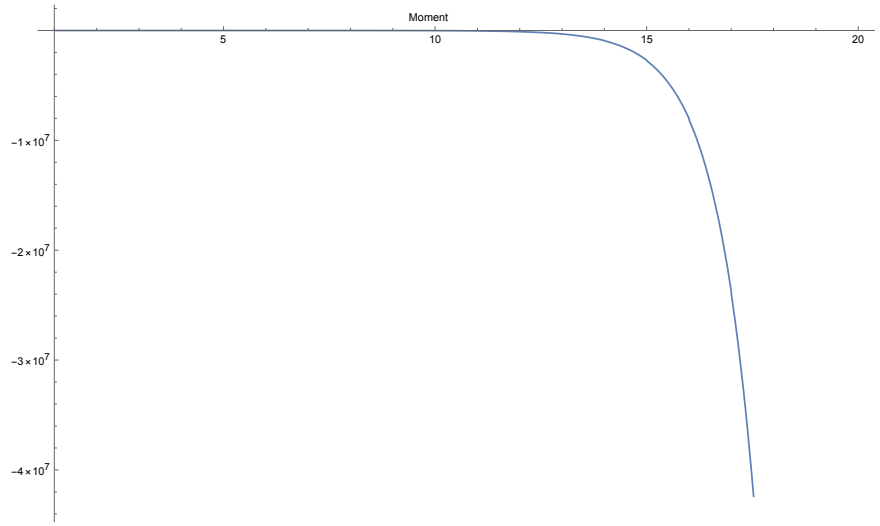
$$+\sum_{j=2}^{n-1}\left(\sum_{i=1}^{j-1}\left(\frac{\left(2^{-n}-1\right)\left(\frac{1}{3}\left(2^{-i}-1\right)+q\right)\left(\frac{1}{3}\left(2^{-j}-1\right)+q\right)}{\left(\frac{1}{3}\left(1-2^{-i}\right)-q+1\right)\left(\frac{1}{3}\left(1-2^{-j}\right)-q+1\right)\left(\frac{1}{3}\left(1-2^{-n}\right)-q+1\right)}+\right.\right.$$

$$\frac{1}{3}\left(2^{-n}-1\right)+q\left(-\frac{9\ 2^n\left(2^n-1\right)\left(\frac{1}{3}\left(2^{-i}-1\right)+q\right)\left(\frac{1}{3}\left(2^{-j}-1\right)+q\right)}{\left(\frac{1}{3}\left(1-2^{-i}\right)-q+1\right)\left(\frac{1}{3}\left(1-2^{-j}\right)-q+1\right)\left(3\ 2^n q-2^{n+2}+1\right)^2}+\right.$$

$$\left.\left.\left.\frac{\left(2^{-j}-1\right)\left(\frac{1}{3}\left(2^{-i}-1\right)+q\right)}{\left(\frac{1}{3}(1-2^{-i})-q+1\right)\left(\frac{1}{3}(1-2^{-j})-q+1\right)}+\left(\frac{1}{3}\left(2^{-j}-1\right)+q\right)\left(\frac{\frac{2^{-i}-1}{\frac{1}{3}(1-2^{-i})-q+1}-\frac{9\ 2^i\left(2^i-1\right)\left(\frac{1}{3}\left(2^{-i}-1\right)+q\right)}{\left(3\ 2^i q-2^{i+2}+1\right)^2}}{\frac{1}{3}(1-2^{-j})-q+1}-\frac{9\ 2^j\left(2^j-1\right)\left(\frac{1}{3}\left(2^{-i}-1\right)+q\right)}{\left(\frac{1}{3}(1-2^{-i})-q+1\right)\left(3\ 2^j q-2^{j+2}+1\right)^2}\right)\right)}{\frac{1}{3}\left(1-2^{-n}\right)-q+1}\right)\right)\binom{j}{i}\binom{n}{j}\right)$$

For clarity, a plot of the function is on the following page.

Figure A.1: The linear terms in the expansion of M_n about $\epsilon = 0$



REFERENCES

[1] M. D. Penrose and A. R. Wade, "Random minimal directed spanning trees and dickman-type distributions.," *Advances in applied probability.*, vol. 36, pp. 691–714, January 2004.

[2] V. Goncharov, "On the field of combinatory analysis," *Amer. Math. Soc. Transl.*, vol. 19, pp. 1–46, 1962.

[3] A. M. Vershik and A. A. Shimdt, "Limit measures arising in the asympyotic theory of symmetric groups. i.," *Theory of Probability & Its Applications*, vol. 22, no. 1, pp. 70–85, 1977.

[4] S. A. Molchanov and V. A. Panov, "The Dickman-Goncharov distribution," *Russian Mathematical Surveys*, vol. 75, p. 1089, dec 2020.

[5] W. Vervaat, "On a stochastic difference equation and a representation of non-negative infinitely divisible random variables," *Advances in Applied Probability*, vol. 11, no. 4, pp. 750–783, 1979.

[6] M. Grabchak, S. A. Molchanov, and V. A. Panov, "Around the infinite divisibility of the dickman distribution and related topics," *Zapiski Nauchnykh Seminarov POMI*, vol. 515, no. 0, pp. 91–120, 2022.

[7] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 1. Wiley, January 1968.

[8] A. M. Vershik, "The asymptotic distribution of factorizations of natural numbers into prime divisors," *Sov. Math., Dokl.*, vol. 34, pp. 57–61, 1987.

[9] B. Di Finetti, "Sulle funzioni a incremento aleatorio," *Rend. Acad. Lincei*, vol. 6, no. 10, pp. 163–168, 325–329, 548–553, 1929.

[10] P. Lévy, "Sur les intégrales dont les éléments sont des variables aléatoires indépendantes," *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, vol. 3, no. 3-4, pp. 337–366, 1934.

[11] A. N. Kolmogorov, "Sulla forma generale di un processo stocastico omogeneo," *Atti Accad. Naz. Lincei*, vol. 15, pp. 805–808, 1932.

[12] A. Khintchine, "Zur theorie der unbeschrankt teilbaren verteilungsgesetze," *Mathematical collections (in Russian)*, vol. 2, no. 1, pp. 79–119, 1937.

[13] W. Feller, *An introduction to probability theory and its applications. Vol. II.* Second edition, New York: John Wiley & Sons Inc., 1971.

[14] F. W. Steutel and K. Van Harn, *Infinite divisibility of probability distributions on the real line.* CRC Press, 2003.

[15] K. Dickman, "On the frequency of numbers containing prime factors of a certain relative magnitude," *Arkiv for Matematik, Astronomi och Fysik*, vol. 22, pp. A–10, Jan. 1930.

[16] S. Ken-Iti, *Lévy processes and infinitely divisible distributions*, vol. 68. Cambridge university press, 1999.

[17] A. Bose, A. Dasgupta, and H. Rubin, "A contemporary review and bibliography of infinitely divisible distributions and processes," *SankhyÄ: The Indian Journal of Statistics, Series A (1961-2002)*, vol. 64, no. 3, pp. 763–819, 2002.

[18] P. J. Smith, "A recursive formulation of the old problem of obtaining moments from cumulants and vice versa," *The American statistician*, vol. 49, no. 2, pp. 217–, 1995.