

# On Binomial Thinning and Mixing

Offer Kella<sup>\*†</sup> and Andreas Löpker<sup>‡</sup>

February 26, 2022

## Abstract

In this paper we consider the notions of binomial thinning, binomial mixing, their generalizations, certain interplay between them, associated limit theorems and provide various examples.

*Keywords:* Binomial thinning, binomial mixture, alternate probability generating function.

*AMS Subject Classification (MSC2020):* Primary 60E05, Secondary: 60E10.

## 1. Introduction

Given a non-negative integer-valued random variable  $Z$  and a probability  $\alpha$ , a (binomially)  $\alpha$ -thinned random variable denoted  $\alpha \circ Z$  is a random variable that has, given  $Z = n$ , a binomial distribution with parameters  $n$  and  $\alpha$ . By definition, a 1-thinned random variable has the same distribution as the original random variable and a 0-thinned random variable is distributed like the constant zero.

While, for every  $\alpha \in [0, 1]$  and any arbitrary non-negative integer-valued  $Z$ , the  $\alpha$ -thinning operation is well defined, there are non-negative integer-valued random variables that cannot be written as an  $\alpha$ -thinned ‘version’ of some non-negative integer-valued  $Z$  for any  $\alpha < 1$ . It turns out that for any  $X$  there is a minimal  $\alpha \in [0, 1]$  (possibly one or zero) allowing such a representation. This observation gives rise to a nested hierarchy of subsets of all distributions concentrated on the non-negative integers, parametrized by  $\alpha$ . We investigate conditions for the membership to these classes, specialize results to distributions with finite support and provide examples.

In addition we also focus on the dual concept of binomial mixtures denoted  $W \circ n$ , where initially the support of the distribution of  $W$  is contained in  $[0, 1]$ .  $W \circ n$  is the notation for a random variable of which conditional distribution given  $W$  is binomial with

---

<sup>\*</sup>Department of Statistics and Data Science, the Hebrew University of Jerusalem, Jerusalem 9190501, Israel (offer.kella@huji.ac.il).

<sup>†</sup>Supported in part by grant No. 1647/17 from the Israel Science Foundation and the Vigevani Chair in Statistics.

<sup>‡</sup>Department of Computer Science and Mathematics, HTW Dresden, University of Applied Sciences, Friedrich-List-Platz 1, D-01069 Dresden, Germany (lopker@htw-dresden.de).

$n$  and  $W$  (to be made more precise later). One question that we explore in this context is whether one can perform the same type of binomial mixing operation when the support is not contained in  $[0, 1]$ . We provide a nontrivial example where  $W$  has the Gamma distribution and thus has an unbounded support. Another question that arises is regarding the relation between the minimal  $\alpha$  mentioned in the previous paragraph with respect to the  $\alpha$ -thinning operation associated with  $W \circ n$ , when well defined, and the minimal  $\alpha$  such that the distribution of  $(W/\alpha) \circ n$  is well defined. It turns out that these two values are equal.

For both thinning and mixing concepts we provide some limiting results where the limits turn out to have Poisson and mixed Poisson distributions, respectively.

Binomial thinning has become a useful tool when studying count data time series ([1], [16], [6]). The binomial thinning operation is also performed in a Galton-Watson branching process with Bernoulli offspring, where each individual survives with probability  $\alpha$  and has no descendants. One can also think of applications in queueing, e.g. in modified M/G/1 systems with impatient customers ([5], [3]).

A somewhat related study to ours has been conducted in [18], where the emphasis is on subclasses of non-negative integer-valued random variables of the form  $\{\alpha \circ Z : \alpha \in (0, 1], Z \in \mathbb{M}_1^*\}$  (in the notations of the present paper). Structural aspects of binomial thinning are also studied in [12] (the operation is called ‘dilation’ there).

Regarding the mixing operation a prominent mixed distribution is the beta-binomial distribution, finding applications in Bayesian statistics. A reference for this and other distributions appearing in this paper is [11].

The paper is organized as follows. In Section 2 we discuss the basic  $\alpha$ -thinning operation and prove various preliminary results. In Section 3 we give a more in depth study of various concepts introduced in Section 2. In Section 4 and 5 we discuss the case where the support of the random variables is bounded and provide various examples. In particular we give a necessary and sufficient characterization as to when an  $X$  does not have an  $\alpha$ -thinned distribution for any  $\alpha < 1$ . Finally, in Section 6 we discuss the concept of binomial mixing and provide two nontrivial examples.

## 2. Thinning preliminaries

In what follows, for a random variable  $X$ , *support of  $X$*  abbreviates *support of the distribution of  $X$*  and *a.s.* abbreviates almost surely (with respect to the probability discussed). Unless stated otherwise, random variables are assumed to be a.s. finite. The notation ‘ $\sim$ ’ abbreviates ‘distributed as’, ‘ $\Rightarrow$ ’ denotes convergence in distribution. We write  $\text{Gamma}(\alpha, \beta)$ ,  $\text{Exp}(\lambda)$ ,  $\text{Erlang}(m, \lambda)$ ,  $\text{Uniform}[a, b]$ ,  $\text{Uniform}(i_1, \dots, i_m)$ ,  $\text{Poisson}(\lambda)$ ,  $\text{Bernoulli}(p)$ ,  $\text{Binomial}(n, p)$  for the gamma/exponential/continuous uniform/discrete uniform distributions with the usual notations for the parameters. Finally,  $1_A$  denotes the indicator of a set  $A$ .

Let  $\mathbb{M}_1$  denote the family of all discrete probability distributions with support contained in  $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$ . When we write  $X \in \mathbb{M}_1$  we mean that the distribution of  $X$  is in  $\mathbb{M}_1$ , or equivalently, that  $\mathbb{P}(X \in \mathbb{Z}_+) = 1$ .

Given  $X \in \mathbb{M}_1$  and  $\alpha \in [0, 1]$  the binomial  $\alpha$ -thinning operation is defined as

$$\alpha \circ X \sim \sum_{i=1}^X B_i, \quad (1)$$

where an empty sum (when  $X = 0$ ) is defined to be zero and  $(B_i)_{i \in \mathbb{Z}_+}$  are i.i.d. Bernoulli r.v., independent of  $X$ , with  $\mathbb{P}(B_i = 1) = \alpha$  (see [15]). In particular, note that  $0 \circ X = 0$  a.s. and  $1 \circ X \sim X$ .

It is simple to show that for independent  $X, Y \in \mathbb{M}_1$ , independent  $\alpha \circ X, \alpha \circ Y$  and  $\alpha, \beta \in (0, 1]$  the following relations hold.

- (i)  $\alpha \circ (X + Y) \sim \alpha \circ X + \alpha \circ Y$ .
- (ii)  $\alpha \circ (\beta \circ X) \sim \beta \circ (\alpha \circ X) \sim (\alpha\beta) \circ X$ .
- (iii)  $\mathbb{E}(\alpha \circ X) = \alpha \mathbb{E}X$  (finite or infinite).
- (iv)  $\text{Var}(\alpha \circ X) = \alpha^2 \text{Var}(X) + \alpha(1 - \alpha)\mathbb{E}X$ , whenever  $\mathbb{E}X^2 < \infty$ .

The set of all  $\alpha$ -thinned distributions is denoted by  $\mathbb{M}_\alpha$  and when we write  $X \in \mathbb{M}_\alpha$  we mean that  $X \sim \alpha \circ Z$  for some  $Z \in \mathbb{M}_1$ .

For  $X \in \mathbb{M}_1$  we denote the probability generating function (p.g.f.) of  $X$  by

$$P_X(s) = \begin{cases} \mathbb{E}s^X & s \in [-1, 0) \cup (0, 1] \\ P(X = 0) & s = 0, \end{cases}$$

which is continuous on  $[-1, 1]$  and infinitely differentiable on  $(-1, 1)$ . Often it is more convenient to work with the function  $G_X(s) = P_X(1 - s)$ , the alternate probability generating function (a.p.g.f.) of  $X$ , which is continuous on  $[0, 2]$  and is infinitely differentiable on  $(0, 2)$ . We recall that  $n!P(X = n) = P_X^{(n)}(0) = (-1)^n G_X^{(n)}(1)$ . One reason to use  $G_X(s)$  rather than  $P_X(s)$  is the useful relation

$$G_{\alpha \circ X}(s) = G_X(\alpha s), \quad (2)$$

for  $s \in [0, 1)$ , which is straightforward to show. In fact, this equality holds on  $[0, 2]$ . The reason is that upon changing the order of summation we have that

$$\sum_{i=0}^{\infty} |1 - s|^i \sum_{k=i}^{\infty} \binom{k}{i} \alpha^i (1 - \alpha)^{k-i} P(X = i) = \mathbb{E}(1 - \alpha + |1 - s|\alpha)^X \quad (3)$$

Thus, the sum converges absolutely whenever  $1 - \alpha + |1 - s|\alpha \leq 1$  which holds if and only if  $s \in [0, 2]$ . Therefore, it is allowed to change the order of summation when  $(1 - s)^i$  replaces  $|1 - s|^i$  which gives  $G_X(\alpha s)$  on the right hand side.

A real valued function  $f$  is called *absolutely monotone* (resp., *completely monotone*) on an open interval  $(a, b)$  ( $a$  or  $b$  could be infinite) if it is smooth (=infinitely differentiable) and satisfies  $f^{(n)}(t) \geq 0$  (resp.,  $(-1)^n f^{(n)}(t) \geq 0$ ) for all  $t \in (a, b)$ .

We recall that a function  $f : (0, 1] \rightarrow \mathbb{R}$  is a p.g.f. of some, nonnegative, integer valued random variable if and only if it is absolutely monotone on  $(0, 1)$  and satisfies  $f(1-) = f(1) = 1$  (e.g., Th. 3a, p. 146 of [17]). Therefore,  $G : [0, 1) \rightarrow \mathbb{R}$  is an a.p.g.f. if and only if it is completely monotone on  $(0, 1)$  and satisfies  $G(0+) = G(0) = 1$ .

We note that although  $P_X(s)$  (as a function on  $[-1, 1]$ ) is absolutely monotone on  $(0, 1)$  (and hence, also at 0), it is not necessarily so on  $(-1, 0)$ . Similarly,  $G_X(s)$  (as a

function on  $[0, 2]$  need not be completely monotone on  $(1, 2)$ . For a trivial example, take  $P(X = 1) = 1$  and observe that  $P_X(s) = s$  is negative on  $(-1, 0)$  and  $G_X(s) = 1 - s$  is negative on  $(1, 2)$ .

When  $Y$  is an a.s. nonnegative random variable we denote by  $L_Y(s) = \mathbb{E}e^{-sY}$  the Laplace-Stieltjes transform (LST) of  $Y$ . We recall that a function  $h : [0, \infty) \rightarrow \mathbb{R}$  is an LST of some a.s. nonnegative random variable if and only if it is completely monotone on  $(0, \infty)$ , is continuous (from the right) at zero and satisfies  $h(0) = 1$ .

For ease of reference,  $X$  has a *mixed Poisson* distribution if for some nonnegative random variable  $Y$ , the conditional distribution of  $X$  given  $Y$  is  $\text{Poisson}(Y)$ , where  $\text{Poisson}(0)$  is the distribution of the constant zero. Equivalently, when  $X \sim N(Y)$ , where  $\{N(t) | t \geq 0\}$  is a Poisson process with rate 1 and  $Y$  is an independent nonnegative random variable. We note that when  $X$  has a mixed Poisson distribution, then  $P(X = k) = \mathbb{E}e^{-Y}Y^k/k!$ ,  $\mathbb{E}X = \mathbb{E}Y$  and  $\text{Var}(X) = \text{Var}(Y) + \mathbb{E}Y$  whenever  $\mathbb{E}Y^2 < \infty$ . It is easy to verify that, since  $G_X(s)$  converges absolutely on  $[0, 2]$ , then  $G_X(s) = L_Y(s)$  for  $s \in [0, 2]$ . Also, if  $G_X(s) = L_Y(s)$  on some  $S \subset (0, 2)$  containing an accumulation point, then since  $G_X(\cdot)$  and  $L_Y(\cdot)$  are analytic on  $(0, 2)$  and continuous on  $[0, 2]$ , it follows that  $G_X(s) = L_Y(s)$  on  $[0, 2]$ , thus necessarily  $X \sim N(Y)$ .

**Remark 1.** We note that, although  $L_Y(s)$  is finite on  $[0, \infty)$  and is equal to  $G_{N(Y)}(s)$  on  $[0, 2]$ , it is possible that  $G_{N(Y)}(s)$  is infinite or undefined for  $s \notin [0, 2]$ . One such example is when  $Y$  has the density  $y^{-2}1_{[1, \infty)}(y)$ . It is easy to verify that in this case, letting  $S_{n-1} \sim \text{Gamma}(n-1, 1)$  and recalling  $N(1) \sim \text{Poisson}(1)$ , then for  $n \geq 2$  we have

$$n(n-1)\mathbb{E}e^{-Y}\frac{Y^n}{n!} = \int_1^\infty e^{-y}\frac{y^{n-2}}{(n-2)!}dy = \mathbb{P}(S_{n-1} > 1) = \mathbb{P}(N(1) \leq n-2). \quad (4)$$

Thus, when  $s \notin [0, 2]$  we have that  $|1-s| > 1$ , which implies that

$$\lim_{n \rightarrow \infty} |1-s|^n \mathbb{P}(N(Y) = n) = \lim_{n \rightarrow \infty} \frac{|1-s|^n}{n(n-1)} \mathbb{P}(N(1) \leq n-2) = \infty \quad (5)$$

and thus  $\sum_{i=0}^\infty (1-s)^i \mathbb{P}(N(Y) = i)$  is either infinite (for  $s < 0$ ) or undefined (for  $s > 2$ ).  $\diamond$

An immediate consequence of the discussion above is the following.

**Theorem 1.** Let  $X \in \mathbb{M}_1$  and  $\alpha \in (0, 1)$ .

- (i) There exists  $f : [0, 1/\alpha) \rightarrow \mathbb{R}$  completely monotone on  $(0, 1/\alpha)$  with  $f(0+) = f(0)$  for which  $f(s) = G_X(s)$  on  $S \subset (0, \min(\alpha^{-1}, 2))$  containing an accumulation point if and only if  $X \in \mathbb{M}_\alpha$ . In this case there exists  $Z \in \mathbb{M}_1$  such that  $G_X(s) = G_Z(\alpha s)$  on  $[0, 2]$ .
- (ii) There exists  $f : [0, \infty) \rightarrow \mathbb{R}$  completely monotone on  $(0, \infty)$  with  $f(0) = f(0+)$  for which  $f(s) = G_X(s)$  on  $S \subset (0, 2)$  containing an accumulation point if and only if  $X$  has a mixed Poisson distribution. Equivalently, there exists some nonnegative  $Y$  such that  $G_X(s) = L_Y(s)$  on  $[0, 2]$ .

*Proof.* In (i) we note that since  $G_X(\cdot)$  and  $f(\cdot)$  are analytic on  $(0, \alpha^{-1} \wedge 2)$  and continuous on  $[0, \alpha^{-1} \wedge 2]$  then  $G_X(s) = f(s)$  on  $[0, \alpha^{-1} \wedge 2]$ . Also,  $f_\alpha(t) = f(t\alpha^{-1})$  is completely monotone on  $(0, 1)$ , continuous at 0, with  $f_\alpha(0) = 1$  and thus there exists  $Z \in \mathbb{M}_1$  for which  $G_Z(s) = f(s\alpha^{-1})$  for  $s \in [0, 1)$ . Therefore  $G_Z(\alpha s) = f(s)$  for  $s \in [0, \alpha^{-1})$  and

in particular for  $s \in [0, 1)$  we have  $G_Z(\alpha s) = f(s) = G_X(s)$  which is equivalent to  $X \sim \alpha \circ Z$ , hence  $X \in \mathbb{M}_\alpha$ . The converse is obvious. The fact that  $G_X(s) = G_Z(\alpha s)$  for  $s \in [0, 2]$  was discussed above. (ii) is also immediate from the discussion above.  $\square$

**Corollary 2.** *The following are equivalent.*

- (i)  $X \in \mathbb{M}_\alpha$  for some  $\alpha \in (0, 1)$ .
- (ii)  $G_X(s)$  is completely monotone on  $(1, c)$  for some  $c \in (1, 2]$ .
- (iii)  $P_X(s)$  is absolutely monotone on  $(-\epsilon, 0)$  for some  $\epsilon \in (0, 1)$ .

We note that replacing  $(0, 1)$  in (i) by  $[0, 1)$  is of no consequence since if (i) holds for  $\alpha = 0$  then it also holds for any  $\alpha \in (0, 1)$  and if  $\alpha \in (0, 1)$  then also  $\alpha \in [0, 1)$ .

*Proof.* The equivalence of (ii) and (iii) is obvious, thus it remains to show the equivalence of (i) and (ii). If  $X \in \mathbb{M}_\alpha$  then we take  $Z \in \mathbb{M}_1$  such that  $X \sim \alpha \circ Z$ . Then  $G_X(s) = G_Z(\alpha s)$  for  $s \in [0, 2]$ . Since  $G_Z(\alpha s)$  is completely monotone on  $(0, 1/\alpha)$  then  $G_X(s)$  must be completely monotone on  $(0, c)$  (hence on  $(1, c)$ ) where  $c = \min(2, 1/\alpha) \in (1, 2]$ . For the converse, if  $G_X(s)$  is completely monotone on  $(1, c)$  (hence, on  $(0, c)$ ) then we take  $\alpha = 1/c$  and apply part (i) of Theorem 1.  $\square$

**Corollary 3.**  *$X \sim N(Y)$  for some nonnegative  $Y$  if and only if there exists  $f : [0, \infty) \rightarrow \mathbb{R}$  completely monotone on  $(0, \infty)$  with  $f(0) = f(0+)$  which agrees with  $P_X(-s)$  on some  $S \subset [0, 1)$  containing an accumulation point.*

*In this case there exists an nonnegative random variable  $\tilde{Y}$  and  $p_0 \in (0, 1]$  such that  $P_X(-s) = p_0 L_{\tilde{Y}}(s)$  for  $s \in (0, 1]$ . If  $p_0 = 1$  then  $Y$  is a.s. zero and otherwise  $Y$  satisfies  $\mathbb{P}(Y \in dy) = p_0 e^y \mathbb{P}(\tilde{Y} \in dy)$ .*

We will demonstrate the use of Corollary 3 in Example 7.

*Proof.* If  $X \sim N(Y)$  then  $P_X(-s) = G_X(1+s) = L_Y(1+s)$  on  $S = [0, 1)$  where clearly  $f(s) = L_Y(1+s)$  is completely monotone on  $(0, \infty)$  with  $f(0) = f(0+)$ .

For the converse, if  $P_X(-s) = f(s)$  on  $S$ , then  $G_X(s) = f(s-1)$  on  $S+1 = \{s+1 \mid s \in S\} \subset [1, 2) \subset (0, 2)$ . Thus, if we define  $g(s) = G_X(s)1_{[0,1)}(s) + f(s-1)1_{[1,\infty)}(s)$  then  $g$  is completely monotone on  $(0, 1) \cup (1, \infty)$  and since  $g$  is smooth on  $(0, 2)$  and in particular at  $s = 1$  it follows that  $g$  is completely monotone on  $(0, \infty)$  so that by (ii) of Theorem 1 it follows that  $X \sim N(Y)$  for some nonnegative  $Y$ .

Now, clearly  $p_0 = P_X(0) = G_X(1) = L_Y(1) > 0$ . Hence  $\mathbb{P}(\tilde{Y} \in dy) = p_0^{-1} e^{-y} \mathbb{P}(Y \in dy)$  defines a probability distribution satisfying  $p_0 L_{\tilde{Y}}(s) = L_Y(1+s)$ . This implies that  $P_X(-s) = p_0 L_{\tilde{Y}}(s)$  and that  $\mathbb{P}(Y \in dy) = p_0 e^y \mathbb{P}(\tilde{Y} \in dy)$ . Finally  $p_0 = 1$  if and only if  $L_Y(1) = 1$  which is equivalent to  $\mathbb{P}(Y = 0) = 1$ .  $\square$

### 3. The sets $\mathbb{M}_\alpha$ and $\mathbb{M}_\alpha^\star$

We recall that, for a given  $\alpha \in [0, 1]$ ,  $\mathbb{M}_\alpha$  denotes the set of all  $\alpha$ -thinned distributions of  $\mathbb{M}_1$ . Clearly, for each  $\alpha \in [0, 1]$ ,  $\mathbb{M}_\alpha$  is nonempty since  $\mathbb{M}_1$  is nonempty.

For  $\alpha \in [0, 1)$  define  $\mathbb{M}_{\alpha+} = \bigcap_{\gamma \in (\alpha, 1]} \mathbb{M}_\gamma$ . Note that  $\mathbb{M}_0$  contains only the distribution of the constant zero. Recall (Theorem 1) that  $X \in \mathbb{M}_\alpha$  if and only if  $G_X(s) = G_Z(\alpha s)$  for

some  $Z \in \mathbb{M}_1$  and all  $s \in [0, 2]$ . Finally, denote  $\mathbb{M}_0^* = \mathbb{M}_{0+}$  and for  $\alpha \in (0, 1]$  let

$$\mathbb{M}_\alpha^* \equiv \mathbb{M}_\alpha \setminus \bigcup_{\gamma \in [0, \alpha)} \mathbb{M}_\gamma. \quad (6)$$

By definition,  $\{\mathbb{M}_\alpha^* \mid \alpha \in [0, 1]\}$  are pairwise disjoint with  $\bigcup_{\alpha \in [0, 1]} \mathbb{M}_\alpha^* = \mathbb{M}_1$ .

**Theorem 4.** *For every  $\alpha \in (0, 1]$ ,  $M_{\alpha+} = M_\alpha$ . Moreover,  $\mathbb{M}_{0+} \neq M_0$  and for  $0 < \alpha < \beta \leq 1$ ,*

$$\mathbb{M}_0 \subset \mathbb{M}_{0+} \subset \mathbb{M}_\alpha \subset \mathbb{M}_\beta. \quad (7)$$

*Proof.*  $\mathbb{M}_{0+}$  contains any Poisson (as well as mixed Poisson) distribution, therefore it is non-empty and is different from  $\mathbb{M}_0$  which contains only the distribution of the constant 0, which is also in  $\mathbb{M}_{0+}$  and thus  $\mathbb{M}_0 \subset \mathbb{M}_{0+}$ . Let  $X \sim \alpha \circ 1$  (a Bernoulli distribution with  $\mathbb{P}(X = 1) = \alpha$ ). Assume that for some  $\beta \in (0, 1]$  and some  $Z \in \mathbb{M}_1$  we have that  $X \sim \beta \circ Z$ . Then clearly also  $Z$  has a Bernoulli distribution (possibly with  $P(Z = 1) = 1$ ) and  $\alpha = P(X = 1) = \beta P(Z = 1) \leq \beta$ . Therefore,  $X \notin \mathbb{M}_\beta$  for any  $\beta < \alpha$  and thus  $X \in \mathbb{M}_\alpha^*$ . Hence,  $\mathbb{M}_\alpha^*$  is non-empty. We also see that if  $X \sim \alpha \circ Z$ , then for  $\beta > \alpha$  we have that  $X \sim \beta \circ (\alpha/\beta) \circ Z$  and thus  $\mathbb{M}_\alpha \subset \mathbb{M}_\beta$ . The fact that they are contained in  $\mathbb{M}_1$  and contain  $\mathbb{M}_0$  is obvious. Finally, assume that for every  $\beta > \alpha$ ,  $X \sim \beta \circ Z_\beta$  for some  $Z_\beta \in \mathbb{M}_1$ . Then, since for  $\alpha < \beta_1 < \beta_2 < 1$  we have that  $X \sim \beta_1 \circ Z_{\beta_1} \sim \beta_2 \circ (\beta_1/\beta_2) \circ Z_{\beta_1}$ , we have that  $Z_{\beta_2} \sim (\beta_1/\beta_2) \circ Z_{\beta_1}$ , where the right side is stochastically smaller than  $Z_{\beta_1}$ . Therefore  $Z_\beta$  converges in distribution to some  $Z_\alpha$  as  $\beta \downarrow \alpha$  where  $Z_\alpha$  may possibly have a defective distribution (infinite with positive probability). Since  $G_{Z_\beta}(s) = G_X(s/\beta)$  for  $s \in [0, \beta]$  (hence for  $s \in [0, \alpha]$ ), and  $G_X(s/\beta)$  converges to  $G_X(s/\alpha)$  for every  $s \in [0, \alpha]$ , it follows that  $G_{Z_\alpha}(s) = G_X(s/\alpha)$  for all  $s \in [0, \alpha]$  and hence  $G_X(s) = G_{Z_\alpha}(\alpha s)$  with  $1 = G_X(0) = G_{Z_\alpha}(0)$  which implies that the distribution of  $Z_\alpha$  is proper and hence  $X \in \mathbb{M}_\alpha$ . This implies that  $\mathbb{M}_{\alpha+} = \mathbb{M}_\alpha$ .  $\square$

For every  $X \in \mathbb{M}_1$  the value  $\rho(X) = \inf\{\alpha \in [0, 1] : X \in \mathbb{M}_\alpha\}$  is well defined and

$$X \in \mathbb{M}_\alpha^* \iff \rho(X) = \alpha. \quad (8)$$

The mapping  $\rho(\cdot) : \mathbb{M}_1 \rightarrow \mathbb{R}^+$  shares some properties with a norm as follows.

**Theorem 5.** *Let  $X, Y \in \mathbb{M}_1$  be independent and  $\alpha \in [0, 1]$ . Then,*

- (i)  $\rho(X + Y) \leq \max(\rho(X), \rho(Y)) (\leq \rho(X) + \rho(Y))$ ,
- (ii)  $\rho(\alpha \circ X) = \alpha \rho(X)$ ,
- (iii) *When  $X$  is a.s. bounded then  $\rho(X) = 0$  if and only if  $X$  is a.s. 0.*

*Proof.* (i) For every  $\alpha, \beta \in [0, 1]$  such that  $X \in \mathbb{M}_\alpha \subset \mathbb{M}_{\max(\alpha, \beta)}$  and  $Y \in \mathbb{M}_\beta \subset \mathbb{M}_{\max(\alpha, \beta)}$  we have  $X + Y \in \mathbb{M}_{\max(\alpha, \beta)}$  and thus  $\rho(X + Y) \leq \max(\alpha, \beta)$ . Setting  $\alpha = \rho(X)$  when  $\rho(X) > 0$  or letting  $\alpha \downarrow 0$  when  $\rho(X) = 0$  and similarly for  $Y$  completes the proof.

(ii) Obviously  $0 \circ X \sim 0$  and thus  $\rho(0 \circ X) = 0 = 0\rho(X)$ , thus it suffices to prove this for  $\alpha \in (0, 1]$ .

We first note that if  $X \in \mathbb{M}_\beta$  then  $\alpha \circ X \in \mathbb{M}_{\alpha\beta}$  and thus  $\rho(\alpha \circ X) \leq \alpha\beta$ . By taking  $\alpha = \rho(X)$  when  $\rho(X) > 0$  and letting  $\beta \downarrow 0$  when  $\rho(X) = 0$ , this implies that  $\rho(\alpha \circ X) \leq \alpha\rho(X)$ .

When  $\rho(X) = 0$  we have  $0 \leq \rho(\alpha \circ X) \leq \alpha\rho(X) = 0$  and thus  $\rho(\alpha \circ X) = \alpha\rho(X) = 0$ .

When  $\rho(X) \in (0, 1]$  then  $X \sim \rho(X) \circ Z$  for some  $Z \in \mathbb{M}_1$ . If there was a  $\beta$  with  $\rho(\alpha \circ X) < \beta < \alpha\rho(X)$  then there was a  $W \in \mathbb{M}_1$  such that  $\alpha \circ X \sim \beta \circ W$ . Since  $\beta/\alpha < \rho(X) \leq 1$  we can write  $\alpha \circ X = \alpha \circ (\beta/\alpha) \circ W$  (recalling  $\alpha > 0$ ), implying that  $X \sim (\beta/\alpha) \circ W$  and hence  $\beta/\alpha \geq \rho(X)$ , contradicting  $\beta < \alpha\rho(X)$ .

(iii) Clearly, if  $P(X = 0) = 1$  then  $\rho(X) = 0$ . Conversely, when  $\rho(X) = 0$  then either  $X \sim 0 \circ Z \sim 0$  for some  $Z \in \mathbb{M}_1$  and we are done or for every  $\alpha \in (0, 1]$  there exists some  $Z_\alpha \in \mathbb{M}_1$  such that  $X = \alpha \circ Z_\alpha$ . Let  $n$  be such that  $P(X \leq n) = 1$ . Then  $P(Z_\alpha \leq n) = 1$  for all  $\alpha \in (0, 1]$ . For every  $1 \leq i \leq n$  we have that  $P(X = i) = \sum_{k=i}^n \binom{k}{i} \alpha^k (1-\alpha)^{i-k} P(Z_\alpha = k)$  which converges to zero as  $\alpha \downarrow 0$  and thus  $P(X = 0) = 1$ .  $\square$

**Remark 2.** It is tempting to conjecture that the left inequality in (i) of Theorem 5 is actually an equality. However, this turns out to be false in general. In an attempt to salvage one's dignity, another guess could be that it should always hold that  $\rho(X + Y) \geq \min(\rho(X), \rho(Y))$ . This is false as well. In Corollary 16 to follow we will give an example where  $\rho(X + Y)$  is strictly smaller than both  $\rho(X)$  and  $\rho(Y)$ .  $\diamond$

**Remark 3.** It is interesting to identify a connection between Corollary 2 and  $\rho(X)$ . First note that  $\rho(X) = 1$  if and only if  $X \in \mathbb{M}_1^\star$  which holds if and only if  $G_X(s)$  is not completely monotone on  $(1, c)$  for any  $c \in (1, 2]$ . From the proof of Corollary 2 it may be concluded that, when  $\rho(X) \in [1/2, 1)$ ,  $X \in \mathbb{M}_{\rho(X)}$  if and only if  $G_X(s)$  is completely monotone on  $(0, 1/\rho(X))$ . In this case  $X \sim \rho(X) \circ Z$  where  $Z \in \mathbb{M}_1^\star$  so that  $G_Z(s)$  is not completely monotone on  $(1, c)$  for any  $c \in (1, 2]$  and thus  $G_X(s) = G_Z(\rho(X)s)$  is not completely monotone on  $(1/\rho(X), d)$  for any  $d \in (1/\rho(X), 2]$ .

Assume that for some  $n \geq 0$  we have that  $\rho(X) \in [2^{-(n+1)}, 2^{-n})$ . Let  $Z \in \mathbb{M}_1^\star$  be such that  $X \sim \rho(X) \circ Z \sim 2^{-n} \circ Z_n$  where  $Z_n \sim (2^n \rho(X)) \circ Z$ . Then  $G_{Z_n}(s)$  is completely monotone on  $(0, 1/(2^n \rho(X)))$  and is not completely monotone on  $(1/(2^n \rho(X)), d]$  for any  $d \in (1/(2^n \rho(X)), 2]$ .

This implies the following procedure. If  $G_X(s)$  is completely monotone on  $(1, 2)$ , then take  $Z_1 \in \mathbb{M}_1$  such that  $X \sim (1/2) \circ Z_1$ . If  $G_{Z_1}(s)$  is completely monotone on  $(1, 2)$  then take  $Z_2 \in \mathbb{M}_1$  such that  $Z_1 \sim (1/2) \circ Z_2$ . We continue like this until the first  $n$  for which  $G_{Z_n}(s)$  is not completely monotone on  $(1, 2)$ . If it is not completely monotone on  $(1, c)$  for any  $c \in (1, 2]$  then  $\rho(X) = 2^{-n}$ . Otherwise we take the largest  $c \in (1, 2)$  for which  $G_{Z_n}(s)$  is completely monotone on  $(1, c)$  (necessarily strictly between 1 and 2) and then  $\rho(X) = 1/(2^n c)$ . Note that if we take  $Z \in \mathbb{M}_1^\star$  such that  $Z_n \sim (1/c) \circ Z$  then we have that

$$X \sim (1/2) \circ Z_1 \sim \dots \sim (1/2^n) \circ Z_n \sim (1/2^n) \circ (1/c) \circ Z \sim \rho(X) \circ Z. \quad (9)$$

How do we identify that  $\rho(X) = 0$ ? This will happen if for any  $n$ ,  $G_{Z_n}(s)$  is completely monotone on  $(1, 2)$ , so that in this case this process never not ends. Since  $G_{Z_n}(s/2^n)$  is completely monotone on  $(0, 2^n)$  this will eventually result in a function which is completely monotone on  $(0, \infty)$  and agrees with  $G_X(s)$  on  $[0, 1)$  (hence, as discussed earlier, also on  $[0, 2]$ ), as expected (see Theorem 6 below).  $\diamond$

If  $0 < \rho(X) \leq \alpha \leq 1$  then it makes sense to let  $\alpha^{-1} \circ X$  denote the random variable  $Z$  on the right hand side of  $X = \alpha \circ Z$ , so that  $X \sim \alpha \circ (\alpha^{-1} \circ X)$  for every  $\alpha \in [\rho(X), 1]$ . On the other hand, from  $\alpha \circ X \sim \alpha \circ Z$  for  $Z \sim X$  it follows that  $\alpha^{-1} \circ (\alpha \circ X) \sim X$  for all  $\alpha \in (0, 1]$ .

When  $\rho(X) = 0$ , then for  $\alpha, s \in (0, 1]$  we have

$$L_{\alpha(\alpha^{-1} \circ X)}(s) = G_{(\alpha^{-1} \circ X)}(1 - e^{-\alpha s}) = G_X\left(\frac{1 - e^{-\alpha s}}{\alpha s} s\right)$$

which converges to  $G_X(s)$  as  $\alpha \downarrow 0$ . Since  $G_X(s) \rightarrow 1$  as  $s \downarrow 0$ , then by the continuity theorem for LST's (which also holds when restricting to  $[0, 1]$ ) we have that  $\alpha(\alpha^{-1} \circ X)$  converges in distribution to some nonnegative  $Y$  satisfying  $G_X(s) = L_Y(s)$  on  $[0, 1]$  (hence, also on  $[0, 2]$ ). Therefore,  $X \sim N(Y)$  as explained in the paragraph preceding Theorem 1. Also, if  $X \sim N(Y)$  then since  $\alpha \circ N(Y/\alpha) \sim N(Y)$  for every  $\alpha \in (0, 1]$ , this implies that  $\rho(X) = 0$ . Therefore we have the following known result that can be traced back to [14], Satz 4.2 (see also [9], Theorem 2.3 and [12]).

**Theorem 6.**  $X \in \mathbb{M}_0^\star$  if and only if  $X$  has a mixed Poisson distribution.

**Remark 4.** Since  $N(Y)$  has unbounded support whenever  $Y$  is not identically zero, any random variable with bounded support, other than the constant zero, cannot be in  $\mathbb{M}_0^\star$ .  $\diamond$

The following theorem generalizes Proposition 2.1 in [12].

**Theorem 7.** [Poisson approximation] Let  $S_n = \sum_{k=1}^n X_{kn}$ , where  $X_{1n}, X_{2n}, \dots, X_{nn}$  are i.i.d. non-negative integer-valued random variables with  $X_{1n} \Rightarrow X$  and  $\mathbb{E}X_{1n} \rightarrow \mathbb{E}X$  as  $n \rightarrow \infty$ , where  $\mathbb{E}X < \infty$ . Also assume that  $a_n \in [0, 1]$  with  $na_n \rightarrow \lambda \in (0, \infty)$  as  $n \rightarrow \infty$ . Then

$$a_n \circ S_n \Rightarrow N(\lambda \mathbb{E}X), \quad n \rightarrow \infty.$$

*Proof.* From (2) it follows that for every  $s \in [0, 1]$ ,

$$G_{a_n \circ S_n}(s) = G_{S_n}(a_n s) = G_{X_{1n}}^n(a_n s) = \exp\left(\frac{\log G_{X_{1n}}(a_n s)}{a_n s} \cdot na_n s\right).$$

Thus, showing that the right hand side converges to  $e^{-\lambda \mathbb{E}X s} = L_{\lambda \mathbb{E}X}(s) = G_{N(\lambda \mathbb{E}X)}(s)$ , will complete the proof. Since  $na_n \rightarrow \lambda$ , it remains to show that

$$\frac{\log G_{X_{1n}}(a_n s)}{a_n s} \rightarrow -\mathbb{E}X$$

as  $n \rightarrow \infty$ . By Jensen's inequality it follows that

$$\frac{\log G_{X_{1n}}(a_n s)}{a_n s} = \frac{\log \mathbb{E}(1 - a_n s)^{X_{1n}}}{a_n s} \geq \mathbb{E}X_{1n} \frac{\log(1 - a_n s)}{a_n s} \quad (10)$$

where the right hand side converges to  $-\mathbb{E}X$  as  $n \rightarrow \infty$ . Now, since  $\log x \leq x - 1$  for  $x > 0$  we have that

$$\frac{\log G_{X_{1n}}(a_n s)}{a_n s} \leq \frac{\mathbb{E}(1 - a_n s)^{X_{1n}} - 1}{a_n s}. \quad (11)$$

Note that for every  $u \in (0, 1)$  and every  $x \in [0, \infty)$  we have that, for some  $v \in (0, u)$ ,

$$\frac{(1 - u)^x - 1}{u} = -x(1 - v)^{x-1} \leq -x(1 - u)^{(x-1)^+}. \quad (12)$$

Therefore, for every  $\epsilon \in (0, 1)$  and  $n \geq s/\epsilon$  we have that

$$\frac{\mathbb{E}(1 - a_n s)^{X_{1n}} - 1}{a_n s} \leq -\mathbb{E}X_{1n}(1 - a_n s)^{(X_{1n}-1)^+} \leq -\mathbb{E}X_{1n}(1 - \epsilon)^{(X_{1n}-1)^+}. \quad (13)$$



As  $x(1 - \epsilon)^{(x-1)^+}$  is bounded and continuous (in  $x$ ) on  $[0, \infty)$ , it follows that the right side converges, as  $n \rightarrow \infty$ , to  $-\mathbb{E}X(1 - \epsilon)^{(X-1)^+}$ , which in turn (monotone convergence) converges to  $-\mathbb{E}X$  as  $\epsilon \downarrow 0$ . Therefore we have shown that

$$-\mathbb{E}X \leq \liminf \frac{\log G_{X_{1n}}(a_ns)}{a_ns} \leq \limsup \frac{\log G_{X_{1n}}(a_ns)}{a_ns} \leq -\mathbb{E}X \quad (14)$$

and we are done.  $\square$

**Remark 5.** As a sanity check, note that for  $X_{1n} = 1$  we have that  $a_n \circ S_n \sim \text{Binomial}(n, a_n)$  which implies the well known fact that  $\text{Binomial}(n, a_n) \Rightarrow \text{Poisson}(\lambda)$  whenever  $na_n \rightarrow \lambda$ . Also note that  $n^{-1} \circ S_n \rightarrow N(\mathbb{E}X)$  which is a distributional version of a law of large numbers in the current context. Finally note that if either the supports of  $X_{1n}$  are contained in a (common) bounded set or  $X_{1n}$  is stochastically increasing in  $n$ , then the condition  $\mathbb{E}X_{1n} \rightarrow \mathbb{E}X$  is redundant.  $\diamond$

For  $X \in \mathbb{M}_1$ , denote  $\text{supp}(X) = \{i \mid i \in \mathbb{Z}_+, \mathbb{P}(X = i) > 0\}$  (support of the distribution of  $X$ ). We will say that  $\text{supp}(X)$  *contains holes* if for some  $n \geq 1$ ,  $P(X = n) > 0 = P(X = n - 1)$ . Note that a special case is when  $P(X = n) > 0$  for some  $n \geq 1$  and  $P(X = 0) = 0$ . Also note that trivially the support of the constant zero does not contain holes, but that of any other positive integer constant does.

**Lemma 8.** *For every  $\alpha \in (0, 1)$ ,  $\text{supp}(\alpha \circ X)$  does not contain holes and*

$$\sup\{i \mid i \in \mathbb{Z}_+, P(X = n) > 0\} = \sup\{i \mid i \in \mathbb{Z}_+, P(\alpha \circ X = n) > 0\}. \quad (15)$$

*Proof.* Immediate from the fact that for each  $n \geq 1$  with  $P(X = n) > 0$  we have, for every  $0 \leq k \leq n$ , that  $P(\alpha \circ X = k) \geq \binom{n}{k} \alpha^k (1 - \alpha)^{n-k} P(X = n) > 0$ .  $\square$

**Theorem 9.**  $\rho(X) = \alpha \in (0, 1)$  (equivalently,  $X \in \mathbb{M}_\alpha^\star$ ) if and only if there is some  $Z \in \mathbb{M}_1^\star$  such that  $X \sim \alpha \circ Z$ . Moreover, any  $X \in \mathbb{M}_1$  such that  $\text{supp}(X)$  contains holes is in  $\mathbb{M}_1^\star$ .

*Proof.* Assume that  $\rho(X) = \alpha$  (equivalently,  $X \in \mathbb{M}_\alpha^\star$ ). Then there is a  $Z \in \mathbb{M}_1$  such that  $X \sim \alpha \circ Z$ . Assume to the contrary that  $Z \in \mathbb{M}_\beta$  for some  $\beta \in (0, 1)$ . Then  $Z \sim \beta \circ W$  for some  $W \in \mathbb{M}_1$  and thus  $X \sim \alpha \circ (\beta \circ W) \sim (\alpha\beta) \circ W$  which contradicts the assumption that  $\rho(X) = \alpha$ . Thus, necessarily  $Z \in \mathbb{M}_1^\star$ . Now, let  $Z \in \mathbb{M}_1^\star$  and take  $X \sim \alpha \circ Z$ . Assume to the contrary that  $X \in \mathbb{M}_\beta$  for some  $\beta < \alpha$ . Then, for some  $W \in \mathbb{M}_1$ ,  $X \sim \beta \circ W$ . Since  $\alpha \circ (\beta/\alpha) \circ W \sim \beta \circ W \sim \alpha \circ Z$ , then  $Z \sim (\beta/\alpha) \circ W$ . This contradicts the assumption that  $Z \in \mathbb{M}_1^\star$ . Finally, by Lemma 8, any distribution in  $\mathbb{M}_1$  of which support contains holes cannot be in  $\mathbb{M}_\alpha$  for any  $\alpha \in (0, 1)$  and is thus in  $\mathbb{M}_1^\star$ .  $\square$

The next theorem shows that any mixed Poisson distribution not concentrated at zero is the distribution of a product of some random variable in  $\mathbb{M}_\alpha^\star$  (necessarily with unbounded support) with some independent Bernoulli random variable.

**Theorem 10.** *Let  $\alpha \in (0, 1]$  and let  $Y$  be a non-negative random variable with  $P(Y = 0) < 1$ . Then there exist  $X \in \mathbb{M}_\alpha^\star$  and an independent Bernoulli  $B$ , such that  $BX \sim N(Y)$ .*

*Proof.* Let  $P(Z = 0) = 0$  and for  $n \geq 1$

$$P(Z = n) = \frac{1}{1 - L_Y(1/\alpha)} \mathbb{E} \left( e^{-Y/\alpha} \frac{(Y/\alpha)^n}{n!} \right). \quad (16)$$

From Theorem 9,  $Z \in \mathbb{M}_1^*$  and thus  $X \sim \alpha \circ Z \in \mathbb{M}_\alpha^*$ . Let  $B \sim \text{Bernoulli}(1 - L_Y(1/\alpha))$  be independent of  $Z$ . It is easy to check that  $BZ \sim N(Y/\alpha)$  and that

$$B(\alpha \circ Z) \sim \alpha \circ BZ \sim \alpha \circ N(Y/\alpha) \sim N(Y) \quad (17)$$

and the proof is complete.  $\square$

For a general  $X \in \mathbb{M}_1$  let  $X_0 \sim X|(X \geq 1)$ . Obviously, as  $\alpha \rightarrow 0$  the thinned variable  $\alpha \circ X$  tends to zero, but  $(\alpha \circ X)_0$  may still converge weakly to a nonconstant random variable.

**Theorem 11.** *Let  $X \in \mathbb{M}_1$  with  $\mathbb{P}(X = 0) < 1$ . If  $X$  has a finite mean then*

$$(\alpha \circ X)_0 \Rightarrow 1, \quad \alpha \rightarrow 0.$$

*If instead  $\mathbb{P}(X \geq k) = k^{-\gamma}L(k)$  with  $\gamma \in (0, 1)$  and  $L$  slowly varying at infinity then*

$$(\alpha \circ X)_0 \Rightarrow \text{Sibuya}(\gamma), \quad \alpha \rightarrow 0.$$

**Remark 6.** The Sibuya distribution is described in Example 6. Recall that a function  $L$  is slowly varying if  $L(cx)/L(x) \rightarrow 1$  as  $x \rightarrow \infty$  for every  $c > 0$ . The theorem shows that if we apply binomial thinning repeatedly to an  $X$  with finite mean then  $\alpha^n \circ X$ , given that it is still positive, will most probably be one if  $n$  is large. This is no longer true in the infinite mean case.  $\diamond$

*Proof.* The a.p.g.f. of  $X_0$  is given by  $G_{X_0}(s) = \frac{G_X(s) - G_X(1)}{1 - G_X(1)}$ . Provided that  $\mathbb{E}X < \infty$  we have

$$G_{(\alpha \circ X)_0}(s) = 1 - \frac{1 - G_X(\alpha s)}{1 - G_X(\alpha)} = 1 - \frac{\alpha s \mathbb{E}X + o(\alpha)}{\alpha \mathbb{E}X + o(\alpha)} \rightarrow 1 - s,$$

which is the a.p.g.f. of the constant one. Now assume that  $a_k := \mathbb{P}(X \geq k) = k^{-\gamma}L(k)$  and let  $A(s) = \sum_{k=0}^{\infty} (1-s)^k a_k$ . By Karamata's theorem for power series ([2], Corollary 1.7.3)  $a_n = k^{-\gamma}L(k)$  implies that

$$A(s) \sim s^{\gamma-1}L(1/s)\Gamma(1-\gamma)$$

as  $s \rightarrow 0$ . Now

$$sA(s) = s \sum_{k=0}^{\infty} (1-s)^k \sum_{i=k}^{\infty} p_i = s \sum_{i=0}^{\infty} p_i \sum_{k=0}^i (1-s)^k = 1 - (1-s)G_X(s).$$

i.e. as  $s \rightarrow 0$

$$1 - G_X(s) = 1 - \frac{1 - sA(s)}{1 - s} \sim \Gamma(1-\gamma)s^\gamma L(1/s).$$

This implies that

$$1 - G_{(\alpha \circ X)_0}(s) = \frac{1 - G_X(\alpha s)}{1 - G_X(\alpha)} \sim \frac{(\alpha s)^\gamma L(1/(\alpha s))}{\alpha^\gamma L(1/\alpha)} \rightarrow s^\gamma$$

as  $\alpha \rightarrow 0$  by the properties of slowly varying functions.  $G_{(\alpha \circ X)_0}(s) = 1 - s^\gamma$  is the a.p.g.f. of a Sibuya distribution.  $\square$

Recall the following. For a sequence  $\{a_k | k \geq 0\}$  define  $r = \limsup_{n \rightarrow \infty} |a_k|^{1/k}$ , then the radius of convergence of  $S(x) = \sum_{k=0}^{\infty} a_k x^k$  is  $R = 0$  when  $r = \infty$ ,  $R = \infty$  when  $r = 0$

and  $R = 1/r$  when  $r \in (0, \infty)$ . When  $|x| < R$ ,  $S(x)$  converges absolutely and is infinitely differentiable with  $S^{(n)}(x) = n! \sum_{k=n}^{\infty} \binom{k}{n} x^{k-n} a_k$  which also has a radius of convergence  $R$  and thus converges absolutely for  $|x| < R$ .

**Theorem 12.** Let  $X \in \mathbb{M}_1$  and let  $R_X$  denote the radius of convergence of  $P_X(s) = G_X(1-s) = \sum_{k=0}^{\infty} s^k p_k$ , where  $p_k = \mathbb{P}(X = k)$ . Then  $R_X \geq 1$  (possibly infinite) and the following holds:

- (i) If  $(1 + R_X)^{-1} < \alpha < 1$  (0 on the left when  $R_X = \infty$ ) then  $G_X(s/\alpha)$  is well defined for  $s \in [0, 1]$  and

$$p_j^*(\alpha) \equiv \sum_{k=j}^{\infty} \binom{k}{j} \alpha^{-k} (\alpha - 1)^{k-j} p_k \quad (18)$$

converges absolutely with  $\sum_{j=0}^{\infty} p_j^*(\alpha) = 1$ . Moreover, if  $X \in \mathbb{M}_\alpha$  with  $X = \alpha \circ Z$ ,  $Z \in \mathbb{M}_1$  then necessarily  $p_j^*(\alpha) \geq 0$  for all  $j \geq 0$  and  $\mathbb{P}(Z = j) = p_j^*(\alpha)$ .

- (ii) If  $R_X > 1$ ,  $2(1 + R_X)^{-1} < \alpha < 1$  and  $p_j^*(\alpha) \geq 0$  for all  $j \in \mathbb{Z}_+$  then  $X \in \mathbb{M}_\alpha$  with  $X = \alpha \circ Z$ ,  $Z \in \mathbb{M}_1$  and  $\mathbb{P}(Z = j) = p_j^*(\alpha)$ .

**Remark 7.** Note that since  $R_X \geq 1$ , then  $(1 + R_X)^{-1} \leq 1/2$  and if  $R_X > 1$  then  $2(1 + R_X)^{-1} < 1$ . In particular, the first part of Theorem 12 is valid for any  $\alpha \in (1/2, 1)$ . Also note that the first part also implies that when  $(1 + R_X)^{-1} < \alpha < 1$  and there exists  $j \geq 0$  such that  $p_j^*(\alpha) < 0$ , then there is no  $Z \in \mathbb{M}_1$  such that  $X \sim \alpha \circ Z$ . In particular, if for any  $\alpha \in (1 - \epsilon, 1)$ , where  $0 < \epsilon \leq 1/2$ , there is some  $j$  for which  $p_j^*(\alpha) < 0$  then necessarily  $X \in \mathbb{M}_1^*$ .  $\diamond$

*Proof.* (i)  $G_X(s/\alpha)$  is well defined for  $\alpha(1 - R_X) < s < \alpha(1 + R_X)$ . Since  $\alpha > (1 + R_X)^{-1}$  it follows that  $[0, 1] \subset [\alpha(1 - R_X), \alpha(1 + R_X)]$ . Since  $(1 - \alpha)/\alpha < R_X$ , the series (18) converges absolutely. If  $X = \alpha \circ Z$  then

$$G_Z(\alpha s) = G_X(s), \quad |1 - s| < R_X. \quad (19)$$

Since  $\alpha > (1 + R_X)^{-1}$  it follows that (19) holds in particular for  $s = 1/\alpha$  and

$$\mathbb{P}(Z = j) = \frac{(-1)^j}{j!} G_Z^{(j)}(1) = \frac{(-1/\alpha)^j}{j!} G_X^{(j)}(1/\alpha) = p_j^*(\alpha). \quad \square$$

- (ii) Since  $\alpha > (1 + R_X)^{-1}$  the function

$$\begin{aligned} f(s) &\equiv G_X(s/\alpha) = \sum_{k=0}^{\infty} (1 - s/\alpha)^k p_k \\ &= \sum_{k=0}^{\infty} \sum_{j=0}^k \left( \frac{1 - \alpha}{\alpha} \right)^k p_k \binom{k}{j} \left( \frac{1 - s}{1 - \alpha} \right)^j (-1)^{k-j} \end{aligned} \quad (20)$$

is well defined for  $s \in [0, 1]$  and  $f(0) = 1$ . The sum of the absolute values is

$$\begin{aligned} \sum_{k=0}^{\infty} \left( \frac{1 - \alpha}{\alpha} \right)^k p_k \sum_{j=0}^k \binom{k}{j} \left( \frac{1 - s}{1 - \alpha} \right)^j &= \sum_{k=0}^{\infty} \left( \frac{1 - \alpha}{\alpha} \right)^k \left( 1 + \frac{1 - s}{1 - \alpha} \right)^k p_k \\ &= \sum_{k=0}^{\infty} \left( \frac{2 - \alpha - s}{\alpha} \right)^k p_k \end{aligned}$$

which converges provided that  $|2 - \alpha - s|/\alpha < R_X$ . That is, when

$$-\alpha(R_X - 1) - 1 < 1 - s < \alpha(R_X + 1) - 1$$

where the left hand side is negative and, since  $\alpha \geq 2/(1 + R_X)$ , the right hand side is  $\geq 1$ . Therefore for all  $s \in [0, 1]$  the sum  $f(s)$  is absolutely convergent, so that we can interchange the order of summation in (??) and obtain, after some rearrangements,

$$f(s) = G_X(s/\alpha) = \sum_{j=0}^{\infty} (1-s)^j p_j^*(\alpha).$$

If  $p_j^*(\alpha) \geq 0$  for  $j \in \mathbb{Z}_+$  then, since  $\sum_{j=0}^{\infty} p_j^*(\alpha) = 1$ ,  $f(s)$  is the a.p.g.f. of a random variable  $Z \in \mathbb{M}_1$  with  $\mathbb{P}(Z = j) = p_j^*(\alpha)$  and  $X = \alpha \circ Z$ .

We have seen that if the support of a distribution in  $\mathbb{M}_1$  contains a hole then  $X \in \mathbb{M}_1^*$ . In Theorem 9 in the next section we will show that for distributions with bounded support the converse is also true. The following example demonstrates that this converse fails when the support is unbounded. That is, there are distributions in  $\mathbb{M}_1^*$  with unbounded supports that do not contain holes.

**Example 1.** For  $0 < d < c < 1$  let  $p_k = P(X = k)$  be defined as follows.

$$p_k = \frac{1}{A} \cdot \begin{cases} c^k & k \text{ even,} \\ d^k & k \text{ odd} \end{cases}$$

where  $A = (1 - c^2)^{-1} + d(1 - d^2)^{-1}$ . The support of this distribution is unbounded and does not contain holes. Clearly,

$$2AP_X(s) = 2 \left( \frac{1}{1 - (cs)^2} + \frac{ds}{1 - (ds)^2} \right) = \frac{1}{1 - cs} + \frac{1}{1 + cs} + \frac{1}{1 - ds} - \frac{1}{1 + ds} \quad (21)$$

so that for odd  $n$  we have

$$\frac{2A}{n!} P_X^{(n)}(s) = \frac{c^{-1}}{(c^{-1} - s)^{n+1}} - \frac{c^{-1}}{(c^{-1} + s)^{n+1}} + \frac{d^{-1}}{(d^{-1} - s)^{n+1}} + \frac{d^{-1}}{(d^{-1} + s)^{n+1}}. \quad (22)$$

Since  $d < c$  then for  $s \in (-1, 0)$  we have  $c^{-1} + s < \max(c^{-1} - s, d^{-1} - s, d^{-1} + s)$  from which it follows that

$$(c^{-1} + s)^{n+1} \cdot \frac{2A}{n!} P_X^{(n)}(s) \rightarrow -c^{-1} < 0 \quad (23)$$

when  $n$  is odd and  $n \rightarrow \infty$ . Therefore  $P_X(s)$  is not absolutely monotone on  $(-\epsilon, 0)$  for any  $\epsilon \in (0, 1)$ . From Corollary 2, this is equivalent to  $X \in \mathbb{M}_1^*$ .

◇

#### 4. Distributions with bounded support

We now consider random variables  $X$  with support in the finite set  $I_n \equiv \{0, 1, 2, \dots, n\}$ . Let us write  $\mathbb{M}_\alpha(n) = \{X \in \mathbb{M}_\alpha : \text{supp}(X) \subseteq I_n\}$  and  $\mathbb{M}_\alpha^*(n) = \{X \in \mathbb{M}_\alpha^* : \text{supp}(X) \subseteq I_n\}$

for the respective subsets of  $\mathbb{M}_\alpha$  and  $\mathbb{M}_\alpha^\star$ . Let as before for  $\alpha \in (0, 1]$

$$p_j^\star(\alpha) = \sum_{k=j}^n \binom{k}{j} \alpha^{-k} (\alpha - 1)^{k-j} p_k, \quad (24)$$

where  $p_k = \mathbb{P}(X = k)$ . We note that here  $G_X(s)$ , being a polynomial, is well defined and finite for all  $s \in \mathbb{R}$ .

**Theorem 13.** *For  $n \geq 1$  let  $X \in \mathbb{M}_1(n)$  with  $p_n > 0$ . Then  $\rho(X) \geq p_n^{1/n} > 0$  and for  $\alpha \in (0, 1]$ ,  $X \in \mathbb{M}_\alpha(n)$  if and only if  $p_j^\star(\alpha) \geq 0$  for all  $j \in I_{n-1}$ , in which case  $X \sim \alpha \circ Z$  with  $P(Z = j) = p_j^\star(\alpha)$  for  $j \in I_n$ .*

*Proof.* First observe that for every  $\alpha \in (0, 1]$  we have that  $p_n^\star(\alpha) = p_n/\alpha^n$  and that  $p_n^\star(\rho(X)) \leq 1$ . This implies that  $p_n^\star(\alpha) > 0$  for all  $\alpha \in (0, 1]$  and that  $\rho(X) \geq p_n^{1/n}$ . Now, for every  $\alpha \in (0, 1]$  and  $s \in \mathbb{R}$ , we have,

$$\begin{aligned} G_X(s/\alpha) &= \sum_{k=0}^n (1 - s/\alpha)^k p_k = \sum_{k=0}^n (1 - s)^k p_k \sum_{j=0}^k \binom{k}{j} \alpha^{-k} (1 - s)^{j-k} (\alpha - 1)^{k-j} \\ &= \sum_{j=0}^n (1 - s)^j \sum_{k=j}^n \binom{k}{j} \alpha^{-k} (\alpha - 1)^{k-j} p_k = \sum_{j=0}^n (1 - s)^j p_j^\star(\alpha), \end{aligned} \quad (25)$$

where, setting  $s = 0$ , gives  $\sum_{j=0}^n p_j^\star(\alpha) = G_X(0) = 1$ . Therefore,  $G_X(s/\alpha)$  is the a.p.g.f. of some  $Z \in \mathbb{M}_1(n)$  (i.e.,  $X \sim \alpha \circ Z$ , necessarily with  $P(Z = j) = p_j^\star(\alpha)$  for  $j \in I_n$ ).

□

**Remark 8.** We mention that since  $\mathbb{M}_{\alpha_1}(n) \subset \mathbb{M}_{\alpha_2}(n)$  for  $0 < \alpha_1 < \alpha_2 \leq 1$ , then it follows that  $p_i^\star(\alpha) \geq 0$  for all  $i \in I_n$  and  $\alpha \in [\rho(X), 1]$ . For all  $\alpha \in (0, \rho(X))$  there is at least one  $i \in I_{n-1}$  for which  $p_i^\star(\alpha) < 0$ . This will be illustrated in Remark 11 following Example 3. ◇

**Theorem 14.** *Let  $X \in \mathbb{M}_1$  with bounded support. Then  $X \in \mathbb{M}_1^\star$  if and only if its support contains a hole.*

*Proof.* We have already seen (Lemma 8) that if the support of  $X$  has a hole then it must be in  $\mathbb{M}_1^\star$ . For the converse, assume that the support of  $X$  has no holes. Therefore, either  $P(X = 0) = 1$  and then  $X \in \mathbb{M}_0^\star$  (hence, not in  $\mathbb{M}_1^\star$ ), or we let  $n = \max\{i \mid P(X = i) > 0\}$ . Since  $G_X(s)$  is a polynomial of degree  $n$ , then for all  $k > n$  we have that  $G^{(k)}(s) = 0$  for all real  $s$ . For  $k \leq n$  we have that  $(-1)^k G^{(k)}(1) = P(X = k) > 0$  so that by the continuity of  $G^{(k)}(s)$  there must be some  $c_k > 0$  such that  $(-1)^k G^{(k)}(s) > 0$  for all  $s \in (1, c_k)$ . Taking  $c = \min\{c_k \mid k = 0, \dots, n\} > 1$  gives that  $G_X$  is completely monotone on  $(0, c)$  and by Corollary 2 we have that  $X \notin \mathbb{M}_1^\star$ . □

The following implies that  $\rho(\cdot)$  is continuous on  $\mathbb{M}_1(n) \setminus \mathbb{M}_1(n-1)$  (but not on  $\mathbb{M}_1(n)$ ). Although it is possible that this may be concluded from the continuity property of (possibly complex) roots of polynomials of the form  $x^n + \sum_{i=0}^{n-1} c_i x^i$  in  $c_0, \dots, c_{n-1}$  (and the continuity of the maximum function), we prefer to give a simple self contained derivation demonstrating how it is in fact a by product of Theorem 14.

**Theorem 15.** Assume that  $X_k \in \mathbb{M}_1(n)$  for  $k \geq 1$  and that  $X_k \Rightarrow X$  with  $P(X = n) > 0$ . Then  $\rho(X_k) \rightarrow \rho(X) > 0$  as  $k \rightarrow \infty$ .

*Proof.* First assume that  $\rho(X_k) = 1$  for all  $k \geq 1$ . Since  $P(X_k = n) \rightarrow P(X = n) > 0$  as  $n \rightarrow \infty$  then there exists some  $K$  such that  $P(X_k = n) > 0$  for all  $k \geq K$ . By Theorem 14 for each  $k \geq K$  the support of  $X_k$  contains a hole (which cannot be  $n$ ) and thus there exists some  $i \in \{0, \dots, n-1\}$  and a subsequence  $X_{k_m}$  such that  $P(X_{k_m} = i) = 0$ . Therefore  $P(X = i) = 0$  and since  $P(X = n) > 0$  then the support of  $X$  contains a hole and thus  $\rho(X) = 1$ .

In general, let  $\alpha_k = \rho(X_k)$  and let  $Z_k \in \mathbb{M}_1^*(n)$  be such that  $X_k \sim \alpha_k \circ Z_k$ . It suffices to show that if  $\alpha_k$  converges then it necessarily converges to  $\rho(X)$ . The reason is that this would imply that any convergent subsequence of  $\alpha_k$  necessarily converges to  $\rho(X)$  which in turn implies that  $\alpha_k \rightarrow \rho(X)$ . Therefore, we assume that  $\alpha_k \rightarrow \alpha$  as  $k \rightarrow \infty$ . Take  $Z_k \in \mathbb{M}_1^*(n)$  such that  $X_k \sim \alpha_k \circ Z_k$ . There exists a subsequence  $Z_{k_m}$  that converges in distribution to some  $Z$ . Now,

$$\alpha^n P(Z = n) = \lim_{m \rightarrow \infty} \alpha_{k_m}^n P(Z_{k_m} = n) = \lim_{m \rightarrow \infty} P(X_{k_m} = n) = P(X = n) > 0 \quad (26)$$

and thus  $\alpha > 0$  and  $P(Z = n) > 0$ . By the first part of the proof we have that  $Z \in \mathbb{M}_1^*(n)$  and since  $X_{k_m} \sim \alpha_{k_m} \circ Z_{k_m} \Rightarrow \alpha \circ Z$  we necessarily have that  $X \sim \alpha \circ Z$  and thus  $\rho(X) = \alpha > 0$ .  $\square$

**Remark 9.** We note that without the condition  $P(X = n) > 0$ , Theorem 15 would no longer be always valid. For example, for  $n \geq 2$  take  $p_0, \dots, p_{n-2} > 0$  with  $\sum_{i=0}^{n-2} p_i = 1$  and assume that  $P(X_k = i) = (1 - k^{-1})p_i$  for  $0 \leq i \leq n-2$ ,  $P(X_k = n-1) = 0$  and  $P(X_k = n) = k^{-1}$ . For each  $k \geq 1$ ,  $\rho(X_k) = 1$  since the support of  $X_k$  contains a hole. However, as  $k \rightarrow \infty$  we have that  $X_k \Rightarrow X$  where the support of  $X$  is  $I_{n-2}$  and contains no holes and thus  $\rho(X) < 1$ . An extreme case is when  $n = 2$  and then  $\rho(X_k) = 1$  for  $k \geq 1$  but  $\rho(X) = 0$ .  $\diamond$

**Corollary 16.** There exist independent  $X, Y \in M_1$  such that  $\rho(X + Y)$  is strictly smaller than both  $\rho(X)$  and  $\rho(Y)$ .

*Proof.* Suppose that  $X, Y$  are independent and both have support  $S = \{0, 1, 3, 4\}$ . Since  $2 \notin S$  then  $S$  contains a hole and thus  $\rho(X) = \rho(Y) = 1$ . However the support of  $X + Y$  is  $S + S = \{i \mid 0 \leq i \leq 8\}$  which has no holes and thus  $\rho(X + Y) < 1$ .  $\square$

**Remark 10.** If one is not happy with the example given in the proof of Corollary 16 and is looking for an example with  $\rho(X) \neq \rho(Y)$ . Then one may replace  $X$  by  $\alpha \circ X$ . Since  $\alpha \circ X + Y \Rightarrow X + Y$  as  $\alpha \uparrow 1$  and  $P(X + Y = 8) > 0$ , then Theorem 15 implies that  $\rho(\alpha \circ X + Y) \rightarrow \rho(X + Y) < 1$  as  $\alpha \uparrow 1$ . Therefore there exists an  $\alpha \in (0, 1)$  such that

$$\rho(\alpha \circ X + Y) < \alpha = \rho(\alpha \circ X) < 1 = \rho(Y). \quad (27)$$

For an example with

$$\rho(X + Y) < \min(\rho(X), \rho(Y)) < \max(\rho(X), \rho(Y)) < 1, \quad (28)$$

simply replace  $X$  and  $Y$  by  $(\alpha\beta) \circ X$  and  $\beta \circ Y$  where  $\beta \in (0, 1)$  and  $\alpha$  is what we chose for (27).  $\diamond$

## 5. Examples

**Example 2.** We have already seen that a Bernoulli random variable  $B$  with  $\mathbb{P}(B = 1) = \alpha$  is a member of  $\mathbb{M}_\alpha^\star$ . Since the constant  $n$  is in  $\mathbb{M}_1^\star$  a binomial random variable  $\alpha \circ n \sim \text{Binomial}(n, \alpha)$  is also a member of  $\mathbb{M}_\alpha^\star$ .  $\diamond$

**Example 3.** Assume that the support of  $X$  is  $\{0, 1, 2\}$  and that  $p_i = P(X = i) > 0$  for  $i = 0, 1, 2$ . Then it is easy to check that

$$\begin{aligned} p_0^\star(\alpha) &= p_0 - \left(\frac{1}{\alpha} - 1\right)p_1 + \left(\frac{1}{\alpha} - 1\right)^2 p_2 \\ p_1^\star(\alpha) &= \frac{2p_2}{\alpha} \left(1 + \frac{p_1}{2p_2} - \frac{1}{\alpha}\right) \\ p_2^\star(\alpha) &= \frac{p_2}{\alpha^2} \geq 0 \quad \forall \alpha \in (0, 1]. \end{aligned} \tag{29}$$

The quadratic function  $P_X(-x) = p_2x^2 - p_1x + p_0$  is nonnegative when  $\Delta = p_1^2 - 4p_0p_2 \leq 0$  which implies that  $p_0^\star(\alpha) \geq 0$  for  $\alpha \in (0, 1]$ . Therefore, in this case the equation for  $p_1^\star(\alpha)$  implies that  $\rho(X) = \frac{1}{1 + \frac{p_1}{2p_2}}$ .

When  $\Delta > 0$  the two (positive) roots of  $P_X(-x)$  are  $\frac{p_1 \pm \sqrt{\Delta}}{2p_2}$ . From this it follows that  $p_0^\star(\alpha) \geq 0$  when either

$$\frac{1}{1 + \frac{p_1 - \sqrt{\Delta}}{2p_2}} \leq \alpha \leq 1 \tag{30}$$

or when

$$0 < \alpha \leq \frac{1}{1 + \frac{p_1 + \sqrt{\Delta}}{2p_2}}. \tag{31}$$

Evidently, (31) is irrelevant since the right hand side is strictly less than  $(1 + \frac{p_1}{2p_2})^{-1}$  and for any such  $\alpha$ ,  $p_1^\star(\alpha) < 0$ . Any  $\alpha$  that satisfies (30) is also larger than  $(1 + \frac{p_1}{2p_2})^{-1}$  and for such  $\alpha$  we also have that  $p_1^\star(\alpha) \geq 0$ . Summarizing, with  $\Delta^+ = \max(\Delta, 0)$ , we have that

$$\rho(X) = \frac{1}{1 + \frac{p_1 - \sqrt{\Delta^+}}{2p_2}} = \begin{cases} \frac{1}{1 + \frac{p_1}{2p_2}} & \Delta < 0 \\ \frac{1}{1 + \frac{p_1}{2p_2}} = \frac{1}{1 + \frac{2p_0}{p_1}} & \Delta = 0 \\ \frac{1}{1 + \frac{2p_0}{p_1 + \sqrt{\Delta}}} & \Delta > 0. \end{cases} \tag{32}$$

Therefore, if  $X \sim \rho(X) \circ Z$ , noting that  $p_1^\star(\rho(X)) = 0$  when  $\Delta \leq 0$  and that  $p_0^\star(\rho(X)) = 0$  when  $\Delta \geq 0$  (recalling that  $p_2^\star(\alpha) > 0$  for all  $\alpha \in (0, 1]$  and in particular for  $\alpha = \rho(X)$ ), we have

$$\text{supp}(Z) = \begin{cases} \{2\} & \Delta = 0 \\ \{1, 2\} & \Delta > 0 \\ \{0, 2\} & \Delta < 0. \end{cases} \tag{33}$$

Not surprisingly, these are also all of the possible supports which contain at least one hole.



When  $\Delta = 0$ , since  $p_0^*(\rho(X)) = p_1^*(\rho(X)) = 0$ , we necessarily have that  $p_2^*(\rho(X)) = 1$ . Therefore,

- When  $\Delta = 0$ ,  $X \sim \text{Binomial}(2, \rho(X))$  (with  $P(Z = 2) = 1$ ).
- When  $\Delta > 0$ ,  $Z - 1 \sim \text{Bernoulli}(p_2^*(\rho(X)))$ .
- When  $\Delta < 0$ ,  $Z/2 \sim \text{Bernoulli}(p_2^*(\rho(X)))$ .

When  $p_i = 1/3$  for  $i = 0, 1, 2$  we have that  $\Delta = 1/3 - 4(1/3)^2 < 0$  and thus  $\Delta^+ = 0$  so that

$$\rho(X) = \frac{1}{1 + \frac{p_1}{2p_2}} = \frac{1}{1 + \frac{1/3}{2/3}} = \frac{1}{1 + \frac{1}{2}} = \frac{2}{3} \quad (34)$$

which is consistent with Example 9 appearing later. In this case  $Z/2 \sim \text{Bernoulli}(3/4)$  since

$$p_2^*(\rho(X)) = \frac{p_2}{\rho(X)^2} = \frac{1/3}{(2/3)^2} = \frac{3}{4} \quad (35)$$

Equivalently,  $\frac{2}{3} \circ [2(\frac{3}{4} \circ 1)]$  has the uniform distribution on  $\{0, 1, 2\}$ .  $\diamond$

**Remark 11.** To illustrate Remark 8 assume that  $\Delta > 0$ , set

$$\alpha_1 = \left(1 + \frac{p_1 + \sqrt{\Delta}}{2p_2}\right)^{-1} \quad \alpha_2 = \left(1 + \frac{p_1}{2p_2}\right)^{-1} \quad \alpha_3 = \rho(X) = \left(1 + \frac{p_1 - \sqrt{\Delta}}{2p_2}\right)^{-1}. \quad (36)$$

and recall that  $p_2^*(\alpha) > 0$  for all  $\alpha \in (0, 1]$ . Now,

- (i) For  $\alpha \in [\alpha_3, 1]$ ,  $p_0^*(\alpha), p_1^*(\alpha) \geq 0$ .
- (ii) For  $\alpha \in [\alpha_2, \alpha_3)$ ,  $p_0^*(\alpha) < 0 \leq p_1^*(\alpha)$ .
- (iii) For  $\alpha \in (\alpha_1, \alpha_2)$ ,  $p_0^*(\alpha), p_1^*(\alpha) < 0$ .
- (iv) For  $\alpha \in (0, \alpha_1]$ ,  $p_1^*(\alpha) < 0 \leq p_0^*(\alpha)$ .

As observed in Remark 8, we have that for  $\alpha \in [\rho(X), 1]$ ,  $p_i^*(\alpha) \geq 0$  for  $i \in I_2$  and for  $\alpha \in (0, \rho(X))$  there is at least one  $i \in I_1$  such that  $p_i^*(\alpha) < 0$ .  $\diamond$

**Remark 12.** Denoting the right hand side of (32) by  $r(p_0, p_1, p_2)$  and defining  $r(1, 0, 0) = 0$  it is easy to check that  $r(\cdot)$ , defined on the closed simplex

$$S = \{(p_0, p_1, p_2) \mid p_0, p_1, p_2 \geq 0, p_0 + p_1 + p_2 = 1\}, \quad (37)$$

gives the correct formula for  $\rho(X)$  also for the cases where at least one of the  $p_i$ 's is zero. In particular, when  $p_2 = 0$  and  $0 < p_1 < 1$ ,  $X \sim \text{Bernoulli}(p_1)$  with  $\rho(X) = p_1$  (see Example 2). For  $(0, p_1, p_2)$  with any choice of  $p_1, p_2$  or  $(p_0, 0, p_2)$  with  $p_2 > 0$  the support contains a hole so that  $\rho(X) = 1$ . For  $(1, 0, 0)$ ,  $\rho(X)$  is clearly zero. Also, as mentioned in Theorem 15,  $r$  is continuous on  $\{p \mid p \in S, p_2 > 0\}$  as well as on  $\{p \mid p \in S, p_1 > 0, p_2 = 0\}$ .  $\diamond$

**Example 4.** Poisson and more generally mixed Poisson random variables are in  $\mathbb{M}_0^*$  (see Theorem 6). If  $X$  has a Poisson distribution and  $X = N(Y)$  then  $Y$  is a.s. a constant.  $\diamond$

**Example 5.**  $X$  has a generalized negative binomial distribution with probability function (see, e.g., [10, 11]) if, for some  $a > 0$ ,

$$\mathbb{P}(X = k) = \frac{\Gamma(a + k)}{\Gamma(a)k!} (1 - p)^k p^a, \quad (38)$$



for  $k \geq 0$ . In this case

$$G_X(s) = \left( \frac{\frac{p}{1-p}}{\frac{p}{1-p} + s} \right)^\alpha. \quad (39)$$

for  $s \in (-\frac{p}{1-p}, 2 + \frac{p}{1-p}) \supset [0, 2]$  and does not exist elsewhere. As a function on  $[0, \infty)$  the right side is clearly completely monotone on  $(0, \infty)$  and continuous at zero. It is in fact the LST of  $Y \sim \text{Gamma}(\alpha, \frac{p}{1-p})$ . Hence ((ii) of Theorem 1) the generalized negative binomial distribution is in  $\mathbb{M}_0^\star$ . Of course, the geometric and the negative binomial distributions, where in both we count only failures, are special cases.  $\diamond$

**Example 6.** In a sequence of independent experiments with decreasing success probabilities  $\gamma/k$ ,  $k = 1, 2, \dots$ ,  $\gamma \in (0, 1)$ , the number of trials  $X$  until the first success is observed has a so called *Sibuya distribution* ([7]), with probabilities given by

$$\mathbb{P}(X = k) = \frac{\gamma}{k} \prod_{j=1}^{k-1} \left( 1 - \frac{\gamma}{j} \right) = (-1)^{k+1} \binom{\gamma}{k}, \quad k = 1, 2, 3, \dots$$

We write  $X \sim \text{Sibuya}(\gamma)$  in this case. The associated a.p.g.f. is  $G_X(s) = 1 - s^\gamma$ . It follows from Hardy and Littlewood's Tauberian theorem that  $\mathbb{P}(X \geq k) \sim k^{-\gamma}/\Gamma(1 - \gamma)$  as  $k \rightarrow \infty$ , in particular  $\mathbb{E}(X^c) = \infty$  for  $c \geq \gamma$ , implying that  $X$  has an infinite mean. Also, since the support of  $X$  has a gap at zero,  $X \in \mathbb{M}_1^\star$ .

If  $X \sim \text{Sibuya}(\gamma)$  then  $\alpha \circ X$  has a *scaled Sibuya distribution* ([4]) with scale parameter  $\alpha$ , i.e.,  $G_{\alpha \circ X}(s) = 1 - (\alpha s)^\gamma$  and

$$\mathbb{P}(\alpha \circ X = k) = \begin{cases} 1 - \alpha^\gamma & k = 0 \\ (-1)^{k+1} \alpha^\gamma \binom{\gamma}{k} & k = 1, 2, 3, \dots \end{cases}$$

Obviously,  $\alpha \circ X \in \mathbb{M}_\alpha^\star$ .  $\diamond$

**Example 7.** With the notations from Example 6, consider  $X - 1$ , the number of failures until the first success, when  $X \sim \text{Sibuya}(\gamma)$ . It is known (e.g., [7]) That  $X - 1 \sim N(Y)$  where  $Y \sim \frac{Z_1 Z_{1-\gamma}}{Z_\gamma}$  and  $Z_1, Z_{1-\gamma}, Z_\gamma$  and  $N(\cdot)$  are independent with  $Z_\alpha \sim \text{Gamma}(\alpha, 1)$  for  $\alpha = 1, 1 - \gamma, \gamma$  and hence  $X - 1 \in \mathbb{M}_0^\star$ .

We would like demonstrate the use of Corollary 3 to infer that  $X - 1 \in \mathbb{M}_0^\star$ , thereby giving an alternate (necessarily equivalent) representation of  $Y$ .

If  $U \sim \text{Uniform}[0, 1]$  and  $V \sim \text{Gamma}(1 - \gamma, 1)$  are independent then, for  $s > 0$ ,

$$L_{UV}(s) = \mathbb{E}L_V(Us) = \mathbb{E} \left( \frac{1}{1 + sU} \right)^{1-\gamma} = \int_0^1 (1 + su)^{\gamma-1} du = \frac{(1 + s)^\gamma - 1}{\gamma s}, \quad (40)$$

so that, with  $p_0 = \mathbb{P}(X - 1 = 0) = \mathbb{P}(X = 1) = \gamma$  and  $\tilde{Y} = UV$ ,

$$P_{X-1}(-s) = \frac{(1 + s)^\gamma - 1}{s} = p_0 \mathbb{E}e^{-s\tilde{Y}} \quad (41)$$

for every  $s \in (0, 1]$ , where the right side is well defined on  $[0, \infty)$  and completely monotone on  $(0, \infty)$ . Thus, if we take  $Y$  with  $\mathbb{P}(Y \in dy) = \gamma e^y \mathbb{P}(\tilde{Y} \in dy)$ , then by Corollary 3,  $X - 1 \sim N(Y)$  where  $Y$  and  $N$  are independent.

This necessarily implies that

$$\frac{L_{UV}(s-1)}{L_{UV}(-1)} = L_{\frac{Z_1 Z_{1-\gamma}}{Z_\gamma}}(s). \quad (42)$$

where we observe that  $UV \sim e^{-Z_1} Z_{1-\gamma}$ .

Although this is not the main point of this example, for the sake of completeness, it is straightforward to show that the density of  $\tilde{Y}$  is given by

$$\frac{\Gamma(-\gamma, y)}{\Gamma(1-\gamma)} \quad (43)$$

where  $\Gamma(-\gamma, y) = \int_y^\infty e^{-t} t^{-\gamma-1} dt$  is the upper incomplete Gamma function. Since  $\Gamma(-\gamma, y) = e^{-y} U(1+\gamma, 1+\gamma, y)$  where  $U$  is the confluent hypergeometric function of the second kind, then the density of  $Y$  can be written as

$$\frac{\gamma U(1+\gamma, 1+\gamma, y)}{\Gamma(1-\gamma)} = \frac{U(1+\gamma, 1+\gamma, y)}{-\Gamma(-\gamma)}. \quad (44)$$

A related expression may be found in Proposition 6 of [13].  $\diamond$

**Example 8 (Mixtures).** Let  $(\Theta, \mathcal{G}, Q)$  be some probability space and assume that, for each  $\theta \in \Theta$ ,  $p(\theta, \cdot)$  is a probability mass function on  $\mathbb{Z}_+$  and that for each  $n \in \mathbb{Z}_+$ ,  $p(\cdot, n)$  is  $\mathcal{G}$ -measurable. That is,  $P(\theta, A) = \sum_{n \in A} p(\theta, n)$ , where  $\theta \in \Theta$  and  $A \subset \mathbb{Z}_+$ , is a (normal) Markov kernel. Let  $K_\theta$  be such that  $P(K_\theta = n) = p(\theta, n)$ , let  $K_\Theta$  be such that  $P(K_\Theta = n) = \int_\Theta p(\theta, n) Q(d\theta)$  and for  $\alpha \in (0, 1)$  let  $(\alpha \circ K)_\Theta$  satisfy

$$P((\alpha \circ K)_\Theta = n) = \int_\Theta P(\alpha \circ K_\theta = n) Q(d\theta). \quad (45)$$

Then it is straightforward to check that  $(\alpha \circ K)_\Theta \sim \alpha \circ K_\Theta$ . Similarly if  $\alpha > \kappa \equiv \sup_{\theta \in \Theta} \rho(K(\theta))$  then

$$\alpha^{-1} \circ K_\Theta \sim (\alpha^{-1} \circ K)_\Theta,$$

where  $(\alpha^{-1} \circ K)_\Theta$  obeys (45) with  $\alpha^{-1}$  replacing  $\alpha$ , so that  $K_\Theta \in \mathbb{M}_\alpha$  and  $\rho(K_\Theta) \geq \kappa$ .

The mixed Poisson distribution is, of course, a special case of this setup. For this case  $\alpha \circ K_\Theta$  takes the form  $\alpha \circ N(Y)$ , while for  $(\alpha \circ K)_\Theta$  we first take  $\alpha \circ N(y) \sim N(\alpha y)$  and then uncondition with respect to  $y$  to obtain  $N(\alpha Y)$ . This results in  $\alpha \circ N(Y) \sim N(\alpha Y)$  which we have already seen earlier. For this case for all  $y \geq 0$  we have that  $\rho(N(y)) = \kappa = \rho(N(Y)) = 0$ .  $\diamond$

## 6. Binomial mixtures

There is a second mixing operation that can be performed on binomial random variables.

We first note that if  $X, U_1, U_2, \dots$ , are independent random variables with  $X \in \mathbb{M}_1$  and  $U_i \sim \text{Uniform}[0, 1]$ , then, for  $\alpha \in [0, 1]$  we have that

$$\alpha \circ X \sim \sum_{i=1}^X 1_{\{U_i \leq \alpha\}}$$

Let us replace  $X$  by  $n \in \mathbb{Z}_+$  and  $\alpha$  by some independent  $W$  with support in  $[0, 1]$ . Denote

a random variable having such a distribution by  $W \circ n$ , that is,

$$W \circ n \sim \sum_{i=1}^n 1_{\{U_i \leq W\}}$$

and let us call this distribution a *binomial mixture*. If we take a Poisson process  $N$  with rate 1, independent of  $W$ , then this is also the conditional distribution of  $N(W)$  given  $N(1) = n$ .

Loosely speaking, this means that the conditional distribution of  $X$  given  $W = p$  is  $\text{Binomial}(n, p)$ . Binomial mixtures appear for example in connection with exchangeable experiments: If  $Y_1, Y_2, Y_3, \dots$  is an infinite sequence of exchangeable Bernoulli random variables then de Finetti's theorem ([8], VII.4) implies that  $S_n = Y_1 + Y_2 + \dots + Y_n$  is a binomial mixture. In fact, for this case the same  $W$  can be used for every  $n$ . It is to be noted that the requirement that the sequence  $(Y_n)_{n \in \mathbb{Z}_+}$  is infinite is crucial in this setup.

The a.p.g.f. of  $W \circ n$  is given by

$$G_{W \circ n}(s) = \mathbb{E}(1 - sW)^n. \quad (46)$$

It turns out that under certain conditions on  $W$  the r.h.s. of (46) defines a proper a.p.g.f. even if  $\text{supp}(W) \cap (1, \infty)$  is nonempty. In fact, we have the following.

**Theorem 17.** *Let  $W$  be a nonnegative random variable and set  $n \geq 1$ . Then (46) is a proper a.p.g.f. if and only if  $\mathbb{E}W^n < \infty$  and  $\mathbb{E}W^k(1 - W)^{n-k} \geq 0$  for every  $k \in K_n$  where*

$$K_n = \{k \mid k \in I_n, n - k \text{ is odd}\}. \quad (47)$$

**Remark 13.** Observe that for  $n = 1$  the conditions of Theorem 17 are met if and only if  $\mathbb{E}W \leq 1$ . For  $n = 2$  they are met if and only if  $\mathbb{E}W^2 \leq \mathbb{E}W < \infty$  and from  $(EW)^2 \leq EW^2$  this also implies that  $EW \leq 1$ . Thus  $\mathbb{B}(2) \subset \mathbb{B}(1)$  (also see Corollary 18 below). In both cases, the distribution of  $W$  may have unbounded support. e.g.,  $\exp(1)$  for  $n = 1$  and  $\exp(2)$  for  $n = 2$  (also see Example 10 below).  $\diamond$

*Proof.* If (46) is a proper a.p.g.f. then in particular it is nonnegative and nonincreasing in  $s$  and thus  $\mathbb{E}(1 - W)^n \in [0, 1]$ . Since  $\mathbb{E}(1 - W)^n 1_{\{W \leq 1\}} \in [0, 1]$ , then a separate consideration for  $n$  odd and even implies that  $\mathbb{E}(W - 1)^n 1_{\{W > 1\}} \in [0, 1]$ . Hence,  $\mathbb{E}|W - 1|^n < \infty$ , which is equivalent to  $\mathbb{E}W^n < \infty$ . Obviously, since  $W$  is nonnegative,  $\mathbb{E}W^k(1 - W)^{n-k} \geq 0$  for  $k \in K_n$  if and only if it holds for  $k \in I_n$ .

It therefore remains to proof the equivalence under the condition that  $\mathbb{E}W^n < \infty$ . Since  $\sum_{k=0}^n (1 - s)^k p_k$  is an a.p.g.f. if and only if  $p_k \geq 0$  for  $k \in I_n$  and  $\sum_{k=0}^n p_k = 1$  and since

$$\mathbb{E}(1 - sW)^n = \sum_{k=0}^n (1 - s)^k \binom{n}{k} \mathbb{E}W^k (1 - W)^{n-k}, \quad (48)$$

(in particular for  $s = 0$ ) the equivalence is immediate.  $\square$

From here on we will agree to write  $W \in \mathbb{B}(n)$  whenever  $W$  satisfies the conditions of Theorem 17. In this case the notation  $W \circ n$  denotes a random variable having a distribution with a.p.g.f. satisfying (46).

**Corollary 18.** *For every  $n \geq 1$ ,  $\mathbb{B}(n) \subset \mathbb{B}(n - 1)$ .*

*Proof.* Let  $W \in \mathbb{B}(n)$ . It is easy to verify that for  $k \in I_{n-1}$

$$\mathbb{E}W^k(1-W)^{n-1-k} = \mathbb{E}W^k(1-W)^{n-k} + \mathbb{E}W^{k+1}(1-W)^{n-k-1} \quad (49)$$

where, by Theorem 17, both terms on the right are non-negative. Therefore the expression on the left is nonnegative for all  $k \in I_{n-1}$  and in particular for  $k \in K_{n-1}$ . Applying Theorem 17 again we have that  $W \in \mathbb{B}(n-1)$ .  $\square$

**Remark 14.** It is natural to wonder whether there exists a random variable  $W$  with  $\text{supp}(W) \cap (1, \infty)$  nonempty such that  $W \in \mathbb{B}(n)$  for all  $n \geq 1$ . The answer is that this is impossible. The reason is that due to Theorem 17 this would be equivalent to  $\mathbb{E}W^k(1-W)^\ell \geq 0$  for all  $k, \ell$ . From monotone convergence it would then follow that for odd  $\ell$

$$\mathbb{E}W^k(1-W)^\ell 1_{\{W>1\}} \rightarrow -\infty$$

as  $k \rightarrow \infty$ , while  $\mathbb{E}W^k(1-W)^\ell 1_{\{W \leq 1\}} \leq 1$  for all  $k$ . Thus, for sufficiently large  $k$  we would necessarily have that  $\mathbb{E}W^k(1-W)^\ell < 0$ .  $\diamond$

**Remark 15.** Is it true that for every  $n \geq 1$  there exists  $W \in \mathbb{B}(n)$  with  $\text{supp}(W) \cap (1, \infty)$  nonempty? Is it true that  $\mathbb{B}(n)$  contains distributions with infinite support? Example 10, to appear later, implies that the answer to both questions is yes.  $\diamond$

We would like to emphasize two facts. The first is that for  $W \circ n$  an interpretation as a binomially mixed random variable is available only if  $\text{supp}(W) \subset [0, 1]$ . The second is that for any  $W \in \mathbb{B}(n)$  the first  $n$  moments of  $W$  identify the distribution of  $W \circ n$  and thus, in contrast to the case of binomial thinning, the distribution  $W$  is not uniquely determined from the distribution of  $W \circ n$ .

As for the  $\circ$  operation there are a number of immediate rules:

- (i)  $(\alpha W) \circ n \sim \alpha \circ (W \circ n)$ ,  $\alpha \in [0, 1]$ .
- (ii)  $\mathbb{P}(W \circ n = k) = \binom{n}{k} \mathbb{E}W^k(1-W)^{n-k}$ .
- (iii)  $\mathbb{E}\binom{W \circ n}{j} = \binom{n}{j} \mathbb{E}W^j$  in particular  $\mathbb{E}(W \circ n) = n\mathbb{E}W$ . Recall  $\binom{k}{j} = 0$  when  $k < j$ .

For a non-negative random variable  $W$  with  $\mathbb{E}W^n < \infty$  we let

$$\sigma_n(W) = \inf\{\alpha \in (0, \infty) : W/\alpha \in \mathbb{B}(n)\}.$$

In contrast to the often difficult evaluation of  $\rho(X)$ , it turns out that the value  $\sigma_n(W)$  is a root of a certain polynomial. Recall (47).

**Theorem 19.** When  $\mathbb{P}(W = 0) = 1$ ,  $\sigma_n(W) = 0$ . Otherwise, for any  $W$  with  $\mathbb{E}W^n < \infty$ , for each  $k \in K_n$  there is a unique positive root  $\alpha_k$  of the polynomial  $g_k(\alpha) = \mathbb{E}W^k(\alpha - W)^{n-k}$  and

$$\sigma_n(W) = \max\{\alpha_k \mid k \in K_n\} \quad (50)$$

and  $\rho(W/\sigma_n(W)) \circ n = 1$ . Consequently, for any  $W \in \mathbb{B}(n)$  we have that  $\rho(W \circ n) = \sigma_n(W)$ .

*Proof.* We first note that when  $W = 0$  a.s. then  $(W/\alpha) \circ n = 0$  a.s. for all  $\alpha \in (0, \infty)$  (so that  $W/\alpha \in \mathbb{B}(n)$ ) and thus  $\sigma_n(W) = 0$ . As was discussed earlier, for this case we also clearly have that  $\rho(W \circ n) = 0$ . Therefore, from here on it suffices to assume that  $\mathbb{P}(W = 0) < 1$ .

From Theorem 17 we have, for  $W \in \mathbb{B}(n)$  (hence, necessarily  $\mathbb{E}W^n < \infty$ ) and  $\alpha \in (0, \infty)$ , that  $W/\alpha \in \mathbb{B}(n)$  if and only if  $g_k(\alpha) \geq 0$  for  $k \in K_n$ , in which case

$$\mathbb{P}((W/\alpha) \circ n = k) = \binom{n}{k} \alpha^{-n} g_k(\alpha). \quad (51)$$

For  $k \in K_n$ ,  $g_k$  is continuous, strictly increasing with  $g_k(0) = -\mathbb{E}W^k < 0$  and  $\lim_{\alpha \rightarrow \infty} g_k(\alpha) = \infty$ . Therefore there exists a unique positive  $\alpha_k$  such that  $g_k(\alpha_k) = 0$ . Consider

$$k(W) = \arg \max\{\alpha_k \mid k \in K_n\}. \quad (52)$$

It is clear from Theorem 17 that for every  $\alpha \in [\alpha_{K(W)}, 1]$  we have that  $W/\alpha \in \mathbb{B}(n)$  and that for every  $\alpha \in [0, \alpha(W))$  we have that there exists some  $k \in K_n$  such that  $g_k(\alpha) < 0$  so that by Theorem 17,  $W \notin \mathbb{B}(n)$ . Therefore  $\sigma_n(W) = \alpha_{k(W)}$  and (50) is satisfied.

Since  $g_{k(W)}(\alpha_{k(W)}) = 0$ ,  $g_k(\alpha_{k(W)}) \geq 0$  for  $k \in K_n$  (since  $\alpha_{K(W)} \geq \alpha_k$ ) and  $g_n(\alpha_{K(W)}) = \mathbb{E}W^n > 0$ , it follows from (51) that the support of the distribution of  $(W/\sigma_n(W)) \circ n$  contains  $n$  and has a hole in  $k(W)$ . Thus, Theorem 14 implies that  $(W/\sigma_n(W)) \circ n \in \mathbb{M}_1^*(n)$ . Since whenever  $W \in \mathbb{B}(n)$  we have that  $\sigma_n(W) \leq 1$ , then  $W \circ n \sim \sigma_n(W) \circ ((W/\sigma_n(W)) \circ n)$  and it follows from Theorem 9 that  $W \circ n \in \mathbb{M}_{\sigma_n(W)}^*(n)$  and thus  $\rho(W \circ n) = \sigma_n(W)$ .  $\square$

**Remark 16.** In order to efficiently find the solutions of  $EW^k(z - W)^{n-k} = 0$  for  $k \in K_n$ , let us write it as  $a(z) = b(z)$ , where

$$a(z) = \mathbb{E}W^i(z - W)^{n-i} 1_{\{W \leq z\}}, \quad b(z) = \mathbb{E}W^i(W - z)^{n-i} 1_{\{W > z\}}. \quad (53)$$

The function  $a(x)$  is non-negative and increasing with  $a(0) = 0$ , while  $b(x)$  is non-negative, decreasing with  $b(0) > 0$ . Hence, in order to solve  $a(x) = b(x)$  a binary search algorithm might be employed: start with an initial value  $z_0 \in (0, 1)$  and update according to

$$z_{n+1} = \begin{cases} z_{n-1} + 2^{-n-1} & \text{if } a(z_{n-1}) > b(z_{n-1}) \\ z_{n-1} - 2^{-n-1} & \text{if } a(z_{n-1}) < b(z_{n-1}). \end{cases}$$

$\diamond$

**Remark 17.** If, for some  $\alpha \in (0, 1]$ ,  $\text{supp}(W) \subset [0, \alpha]$  then clearly  $W/\alpha \in \mathbb{B}(n)$ . Therefore,  $\rho(W \circ n) = \sigma_n(W) \leq \alpha$  and thus  $W \circ n \in \mathbb{M}_\alpha$ .  $\diamond$

The natural place of the following example is at the end of Section 4. It appears here since Theorem 19 allows a smoother derivation.

**Example 9** (Discrete uniform distribution). If  $X \sim \text{Uniform}(I_n)$  then  $\rho(X) = \frac{n}{n+1}$ .

To see this, let  $W \sim \text{Uniform}[0, 1]$  and set  $X \sim W \circ n$ . Then, for  $k \in I_n$ ,

$$\mathbb{P}(X = k) = \binom{n}{k} \int_0^1 t^k (1-t)^{n-k} dt = \frac{1}{n+1}, \quad (54)$$

so that indeed  $X \sim \text{Uniform}(I_n)$ . By Theorem 19  $\rho(X) = \sigma_n(W)$  is the maximum of the unique roots  $\alpha_k$  of  $g_k(\alpha) = \int_0^1 t^k (\alpha - t)^{n-k} dt$  in  $(0, 1]$  for  $k \in K_n$ . When  $k \geq 2$  (and

$n - k$  is odd), integration by parts leads to

$$\begin{aligned} g_k(\alpha) &= \frac{k}{n-k+1} g_{k-1}(\alpha) - \frac{(1-\alpha)^{n-k+1}}{n-k+1} \\ &= \frac{k}{n-k+1} \frac{k-1}{n-k+2} g_{k-2}(\alpha) - \frac{(1-\alpha)^{n-k+1}}{n-k+1} + \frac{k}{n-k+1} \frac{(1-\alpha)^{n-k+2}}{n-k+2}. \end{aligned}$$

Since  $g_k(\alpha_k) = 0$  and  $g_k(1) > 0$  we have that  $\alpha_k < 1$ . Therefore,

$$g_{k-2}(\alpha_k) = \frac{1}{k-1} (1-\alpha_k)^{n-k+1} \left( \alpha_k + \frac{n+2}{k} \right) > 0.$$

so that necessarily  $\alpha_k > \alpha_{k-2}$ . Noting that  $n - (n-1) = 1$  is odd, it follows that  $\alpha_{n-1}$  is the desired maximal root which determines  $\rho(X) = \sigma_n(W)$ . From

$$g_{n-1}(\alpha) = \int_0^1 t^{n-1} (\alpha - t) dt = \frac{\alpha}{n} - \frac{1}{n+1},$$

it follows that  $\alpha_{n-1} = n/(n+1)$  and thus  $\rho(X) = n/(n+1)$  or equivalently  $X \in \mathbb{M}_{n/(n+1)}^\star$ .  $\diamond$

**Remark 18.** When  $X_n \sim \text{Uniform}(I_n)$  and  $X_n \sim \frac{n}{n+1} \circ Z_n$ , it is straightforward to show that (24) with  $\alpha = \frac{n}{n+1}$  becomes

$$\mathbb{P}(Z_n = j) = p_j^\star \left( \frac{n}{n+1} \right) = (n+1)^{j-1} \sum_{i=j}^n \binom{i}{j} \frac{1}{n^i} (-1)^{i-j}. \quad (55)$$

In particular,

$$P(Z_n = 0) = \frac{n^{n+1} - (-1)^{n+1}}{(n+1)^2 n^n}, \quad P(Z_n = n-1) = 0, \quad P(Z_n = n) = \frac{(n+1)^{n-1}}{n^n}, \quad (56)$$

noting that the distribution of  $Z_n \in \mathbb{M}_1^\star$  has a hole in  $n-1$ .

Clearly,  $\mathbb{P}(Z_n = n)/\mathbb{P}(X_n = n)$  is asymptotic to  $e$  while  $\mathbb{P}(Z_n = 0)/\mathbb{P}(X_n = 0)$  is asymptotic to one.  $\diamond$

The following theorem may be compared with Theorem 7, where we recall that  $N(W)$  has a mixed Poisson distribution with conditional mean  $W$ .

**Theorem 20.** Suppose  $W$  is a nonnegative random variable,  $(a_n)_{n \in \mathbb{N}}$  is a positive sequence such that  $a_n/n \rightarrow 1$  as  $n \rightarrow \infty$ ,  $W/a_n \in \mathbb{B}(n)$  for each  $n \geq 1$  and  $\mathbb{E}e^{\epsilon W} < \infty$  for some  $\epsilon > 0$ . Then  $(W/a_n) \circ n \Rightarrow N(W)$ .

*Proof.* Since  $a_n/n \rightarrow 1$ , we have by bounded convergence that, for  $s > 0$ ,

$$\mathbb{E} \left( 1 - s \frac{W}{a_n} \right)^n 1_{\{W \leq a_n/s\}} = \mathbb{E} \left( \left( 1 - s \frac{W}{a_n} \right)^{a_n} \right)^{n/a_n} 1_{\{W \leq a_n/s\}} \quad (57)$$

converges to  $\mathbb{E}e^{-sW}$ , which is also equal to  $G_{N(W)}(s)$  on  $[0, 1]$ . Now,

$$\mathbb{E} \left( s \frac{W}{a_n} - 1 \right)^n 1_{\{W > a_n/s\}} = e^{-\epsilon a_n/s} \left( \frac{s}{a_n} \right)^n \mathbb{E} \left( W - \frac{a_n}{s} \right)^n e^{-\epsilon(W - a_n/s)} e^{\epsilon W} 1_{\{W > a_n/s\}}. \quad (58)$$

Since  $x^n e^{-\epsilon x}$  achieves its maximum at  $x = n/\epsilon$ , this implies that the right side of (58) is

bounded above by

$$\left(\frac{s}{a_n}\right)^n \left(\frac{n}{\epsilon}\right)^n e^{-n-\epsilon a_n/s} \mathbb{E} e^{\epsilon W} = \left(\frac{ns}{\epsilon a_n} e^{-1-\frac{\epsilon a_n}{ns}}\right)^n \mathbb{E} e^{\epsilon W} = f\left(\frac{ns}{\epsilon a_n}\right)^n \mathbb{E} e^{\epsilon W}, \quad (59)$$

where  $f(x) \equiv x e^{-1-x^{-1}}$  is continuous and strictly increasing on  $(0, \infty)$ . Let  $s^* \approx 3.59112$  be the unique positive solution of  $f(x) = 1$ . For  $s \in (0, \epsilon s^*)$  let  $N_s$  be such that for  $n \geq N_s$

$$\frac{n}{a_n} \leq \frac{1 + \epsilon s^*/s}{2}. \quad (60)$$

Then,

$$f\left(\frac{ns}{\epsilon a_n}\right) \leq f\left(\frac{s^* + s/\epsilon}{2}\right) < f(s^*) = 1 \quad (61)$$

and thus  $f(\frac{ns}{\epsilon a_n})^n \rightarrow 0$  as  $n \rightarrow \infty$ . Since  $\mathbb{E} e^{\epsilon W} < \infty$ , it follows that the left hand side of (58) vanishes as  $n \rightarrow \infty$  on  $[0, \epsilon s^*)$ . All of the above implies that  $G_{(W/a_n) \circ n}(s) \rightarrow G_{N(W)}(s)$  on some positive right neighborhood of 0 from which the result follows.  $\square$

The next lemma will be needed for Example 10 to follow.

**Lemma 21.** *Let  $W \sim \exp(1)$ . For every odd integer  $\ell \geq 1$  and real  $\gamma \geq 0$  the root of*

$$\mathbb{E} W^\gamma (z - W)^\ell \quad (62)$$

*is positive and strictly smaller than  $\gamma + \ell/2 + 1$ . For  $\ell = 1$  the root is precisely  $z = \gamma + 1$ .*

*Proof.* For the first part it suffices to show that for  $z = \gamma + \ell/2 + 1$  we have that  $\mathbb{E} W^\gamma (z - W)^\ell > 0$ . Now,

$$\mathbb{E} W^\gamma (z - W)^\ell = \int_0^z w^\gamma (z - w)^\ell e^{-w} dw - \int_z^\infty w^\gamma (w - z)^\ell e^{-w} dw. \quad (63)$$

Dividing the right hand side by  $z^{\gamma+\ell+1}$ , making the change of variable  $x = w/z$  in the first integral and  $x = z/w$  in the second, gives

$$\int_0^1 (1-x)^\ell \left( x^\gamma e^{-zx} - \frac{1}{x^{\gamma+\ell+2}} e^{-z/x} \right) dx. \quad (64)$$

A sufficient condition for (64) to be strictly positive is that the integrand is strictly positive on  $(0, 1)$  which is equivalent to

$$x^{2\gamma+\ell+2} e^{z(1/x-x)} > 1. \quad (65)$$

Noting that  $(2\gamma + \ell + 2)/z = 2$  (since  $z = \gamma + (\ell/2) + 1$ ) this is equivalent to

$$f(x) = 2 \log x + x^{-1} - x > 0. \quad (66)$$

Since

$$f'(x) = \frac{2}{x} - \frac{1}{x^2} - 1 = -\left(\frac{1}{x} - 1\right)^2, \quad (67)$$

it follows that  $f$  is strictly decreasing on  $(0, 1]$  with  $f(1) = 0$  and thus  $f(x) > 0$  for all  $x \in (0, 1]$ . To complete the derivation we observe that

$$\mathbb{E} W^\gamma (z - W) = z\Gamma(\gamma + 1) - \Gamma(\gamma + 2) = (z - \gamma - 1)\Gamma(\gamma + 1) \quad (68)$$

so that the root is clearly  $z = \gamma + 1$ .  $\square$



The following example demonstrates the applicability of Theorems 19 and 20. As promised in Remark 15, it also implies that for every  $n \geq 1$  there are distributions in  $\mathbb{B}(n)$  with unbounded support.

**Example 10.** [The gamma distribution]  $W \sim \text{Gamma}(a, b)$  where  $a, b > 0$ . Then  $\sigma_n(W) = \frac{a+n-1}{b}$  for every  $n \geq 1$ . Moreover,  $(W/\sigma_n(W)) \circ n \Rightarrow X$  where  $X$  has a generalized negative binomial distribution with  $a$  and  $p = 1/2$  (see (38)).

To see this, we first note that since  $bW \sim \text{Gamma}(a, 1)$  and  $\sigma_n(bW) = b\sigma_n(W)$ , it suffices to consider the case  $b = 1$ . We observe that if  $W \sim \text{Gamma}(a, 1)$  and  $X \sim \text{Exp}(1)$  then

$$\mathbb{E}W^k(z - W)^{n-k} = \frac{1}{\Gamma(a)} \mathbb{E}X^{k+a-1}(z - X)^{n-k}. \quad (69)$$

If  $n \geq 1$  and  $k = n - 1$  then we have from Lemma 21, with  $\gamma = k + a - 1 = a + n - 2$  and  $\ell = n - k = 1$ , that the root of  $\mathbb{E}W^{n-1}(z - W)$  is  $z = \gamma + 1 = a + n - 1$ . If  $\gamma = k + a - 1$  is such that  $\ell = n - k$  is odd and greater than or equal to 3 then from (69) and Lemma 21 it follows that the root of  $\mathbb{E}W^k(z - W)^{n-k}$  is strictly smaller than  $\gamma + \ell/2 + 1 = k + a - 1 + \frac{n-k}{2} + 1 = a + \frac{k+n}{2}$ . Since  $n - k \geq 3$ , it follows that  $k \leq n - 3$  and thus  $a + \frac{k+n}{2} \leq a + n - \frac{3}{2} < a + n - 1$ . Therefore, if  $\alpha_k$  are the roots of  $\mathbb{E}W^k(z - W)^{n-k}$  for  $k \in K_n$ , then all but one are strictly less than  $a + n - 1$  and one is equal to  $a + n - 1$ . From Theorem 19 it therefore follows that  $\sigma_n(W) = a + n - 1$ .

Now, we recall that  $\mathbb{E}e^{-sW} = G_{N(W)}(s)$  for  $s \in [0, 1]$ . Also, from Example 5, when  $W \sim \text{Gamma}(a, p/(1-p))$ , then  $N(W)$  has a generalized negative binomial distribution with  $a$  and  $p$ . Since  $p/(1-p) = 1$  then necessarily  $p = 1/2$ .

Finally, for every  $\epsilon \in (0, 1)$  we have that  $\mathbb{E}e^{\epsilon W} = (\frac{1}{1-\epsilon})^a < \infty$  and, clearly,  $\sigma_n(W)/n = 1 + \frac{a-1}{n} \rightarrow 1$  as  $n \rightarrow \infty$ . Thus the convergence result follows from Theorem 20.  $\diamond$

**Remark 19.** Special cases of Example 10 are:

- $W \sim \text{Exp}(\lambda)$  with  $\sigma_n(W) = n/\lambda$ .
- $W \sim \text{Erlang}(m, \lambda)$  (independent sum of  $m$   $\text{Exp}(\lambda)$  distributed random variables) with  $\sigma_n(W) = (m + n - 1)/\lambda$ .
- $W \sim \chi_{(m)}^2$  (independent sum of squares of  $m$  standard normally distributed random variables) with  $\sigma_n(W) = m + 2(n - 1)$ .

$\diamond$

## References

- [1] M. Al-Osh and E.-E. A. Aly. First order autoregressive time series with negative binomial and geometric marginals. *Communications in Statistics-Theory and Methods*, 21(9):2483–2492, 1992.
- [2] N. H. Bingham, C. M. Goldie, J. L. Teugels, and J. Teugels. *Regular variation*. Number 27. Cambridge university press, 1989.
- [3] O. Boxma, A. L pker, and M. Mandjes. On two classes of reflected autoregressive processes. *Journal of Applied Probability*, 57(2):657–678, 2020.



- [4] G. Christoph and K. Schreiber. Scaled sibuya distribution and discrete self-decomposability. *Statistics & probability letters*, 48(2):181–187, 2000.
- [5] J. Cohen. *The single server queue, 2nd ed.* North-Holland Publ. Cy., 1982.
- [6] R. A. Davis, S. H. Holan, R. Lund, and N. Ravishanker. *Handbook of discrete-valued time series*. CRC Press, 2016.
- [7] L. Devroye. A triptych of discrete distributions related to the stable law. *Statistics & Probability Letters*, 18(5):349–351, 1993.
- [8] W. Feller. *An introduction to probability theory and its applications*, volume 2. Wiley, 1966.
- [9] J. Grandell. *Mixed poisson processes*, volume 77. CRC Press, 1997.
- [10] M. Greenwood and G. U. Yule. An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. *Journal of the Royal statistical society*, 83(2):255–279, 1920.
- [11] N. L. Johnson, A. W. Kemp, and S. Kotz. *Univariate discrete distributions*, volume 444. John Wiley & Sons, 2005.
- [12] B. Jørgensen and C. C. Kokonendji. Discrete dispersion models and their tweedie asymptotics. *AStA Advances in Statistical Analysis*, 100(1):43–78, 2016.
- [13] T. J. Kozubowski and K. Podgórski. A generalized Sibuya distribution. *Annals of the Institute of Statistical Mathematics*, 70:855–887, 2018.
- [14] J. Mecke. Eine charakteristische Eigenschaft der doppelt stochastischen Poisson-schen Prozesse. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 11(1):74–81, 1968.
- [15] F. Steutel and K. van Harn. Discrete analogues of self-decomposability and stability. *Ann. Probab.*, 7:893–899, 1979. ISSN 0091-1798.
- [16] C. H. Weiß. Thinning operations for modeling time series of counts – a survey. *AStA Advances in Statistical Analysis*, 92(3):319–341, 2008.
- [17] D. V. Widder. *The Laplace transform*. Princeton Math. Series, 1946.
- [18] C. Wiuf and M. P. Stumpf. Binomial subsampling. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 462(2068):1181–1195, 2006.