 doi:10.20944/preprints202208.0481.v1

Random normalization and thinning for discrete random variables

Lev B. Klebanov[*], Michal Šumbera[†]

Abstract

Different variants of thinning for discrete random variables are studied. The thinning procedure allows to introduce an analog of scale parameter for positive integer-valued random variables. Sufficient and necessary conditions for the existence of such a scale are given.

Key words: Random normalization; thinning operators; Bernstein Theorem; problem of moments; Sibuya distribution.

1 Introduction

In the preprint different variants of thinning for discrete random variables are studied. For some cases, the thinning procedure allows to introduce an analog of scale parameter for positive integer-valued random variables. Sufficient and necessary conditions for the existence of such scale are given. These conditions lead to different variants and modifications of the classical problem of moments.

2 Random normalization for objects connected to positive integer-valued random variables

Let $X_1, X_2, \dots$ be a sequence of independent identically distributed (i.i.d.) non-negative integer-valued random variables. In many branches of Prob-

[*]Department of Probability and Mathematical Statistics, Charles University, Prague, Czech Republic. e-mail: levbkl@gmail.com

[†]Nuclear Physics Institute, Czech Academy of Sciences, 25068 Řež, Czech Republic. e-mail: sumbera@ujf.cas.cz

ability and Statistics it is important to know a limit $n \to \infty$ behavior of $\max_{1\le j\le n} X_j$, $\min_{1\le j\le n} X_j$ or $\sum_{j=1}^{n} X_j$, see e.g. [10]. To obtain non-degenerate limit distribution of such objects one needs to make a normalization of them. Most popular (but not unique) type of normalization in the classical limit theory is a multiplication by a (non-random) variable depending on n only. However, this type of normalization destroys integer structure of the objects under consideration. Therefore, one needs another type of the normalization. One of alternative ways to make a normalization of the sums of i.i.d. random variables is to use the multiplication of the corresponding variable X_j by non-negative integer random variable ε_j ($j = 1, \dots, n$). In this Section we will consider this approach in more details. However, there are also different methods of normalization. Some of them we consider in other Sections.

2.1 Random normalization of sums of random variables

For the case of normalization of sums of random variables $\sum_{j=1}^{n} X_j$ by multiplication by independent random variable ε_j there are two simple possibilities.

1. Let $\{\varepsilon_j(n),\ j = 1, \dots, n\}$ be a sequence of i.i.d. random variables having Bernoulli distribution with the probability of success $\mathbb{P}\{\varepsilon_j(n) = 1\} = \lambda/n$. We suppose that $\varepsilon_j(n)$ are independent of the sequence $\{X_j,\ j = 1, \dots, n\}$. In terms of $f(t)$, the characteristic function of the random variable X_j, and $1 - \lambda/n(1 - \exp(it))$, the characteristic function of the random variable $\varepsilon_j(n)$, the characteristic function of normalized variable $\widetilde{X}_j = \varepsilon_j(n) \cdot X_j$ is
$$1 - \lambda(1 - f(t))/n.$$
The characteristic function of the normalized sum $\sum_{j=1}^{n} \widetilde{X}_j$ equals
$$\bigl(1 - \lambda(1 - f(t))/n\bigr)^n \xrightarrow[n\to\infty]{} \exp\{-\lambda(1 - f(t))\}.$$

2. Let $\{\varepsilon_j(n),\ j = 1, \dots, n\}$ be a sequence of i.i.d. random variables with geometric distribution $\mathbb{P}\{\varepsilon_j(n) = k\} = (1 - \lambda/n)(\lambda/n)^k$, $k = 0, 1, \dots$. We suppose that $\varepsilon_j(n)$ is independent of the sequence $\{X_j\}$. Characteristic function of $\widetilde{X}_j = \varepsilon_j(n) \cdot X_j$ is
$$\tilde{f}_n(t) = (1 - \lambda/n) \sum_{k=0}^{\infty} f(kt) \lambda^k / n^k,$$

 doi:10.20944/preprints202208.0481.v1

where $f(t)$ is characteristic function of X_j. It is not difficult to verify that the characteristic function of the normalized sum $\sum_{j=1}^{n} \widetilde{X}_j$ has the same $n \to \infty$ limit as for the case of Bernoulli distribution

$$(\tilde{f}_n(t))^n \xrightarrow[n\to\infty]{} \exp\{-\lambda(1 - f(t))\}.$$

Of course, it is possible to use some other distributions of the "normalizing" variables ε. However, it is more interesting to consider possible normalization of extremums. The normalization by multiplying seems not to be natural in that situation.

2.2 Random normalization for extrema of random variables

Consider now a possibility to make a normalization for the minimums of a sequence of i.i.d. random variables. Let $\{X_j,\ j = 1, \dots, n\}$ be a sequence of i.i.d. random variables with probability distribution function $F(x) = \mathbb{P}(X < x)$. Suppose that $\{\varepsilon_j(n)\},\ j = 1, \dots, n$ is a sequence of i.i.d. Bernoulli-distributed random variables with $\mathbb{P}\{\varepsilon_j(n) = 1\} = 1 - \lambda/n$ which is independent on the sequence of $\{X_j\}$. The probability distribution function of maximum-normalized random variable $\widetilde{X}_j = \max(\varepsilon_j(n), X_j)$ then reads

$$\mathbb{P}(\widetilde{X}_j < x) = G_n(x) = \begin{cases} 0, \ \text{for } x < 0; \\ \frac{\lambda}{n}F(x), \ \text{for } 0 \le x < 1; \\ F(x), \ \text{for } x \ge 1. \end{cases}$$

It is clear that the distribution function of $\min_{1\le j\le n} \widetilde{X}_j$ is

$$1 - (1 - G_n(x))^n \xrightarrow[n\to\infty]{} \begin{cases} 0, \ \text{for } x < 0; \\ 1 - \exp\{-\lambda F(x)\}, \ \text{for } 0 \le x < 1; \\ 1, \ \text{for } x \ge 1. \end{cases}$$

Much more strange is the following method of normalization. Let $\{X_j,\ j = 1, \dots, n\}$ be a sequence of i.i.d. random variables with probability distribution function $F(x)$. Then the probability distribution function of the maximum or minimum of the above sequence is $\widetilde{X}(p) = \mathbb{P}\{\max_{1\le k\le n} X_k < x\} = (F(x))^n$ and $\widetilde{Y}(p) = \mathbb{P}\{\min_{1\le k\le n} X_k < x\} = 1 - (1 - F(x))^n$, respectively.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 August 2022 doi:10.20944/preprints202208.0481.v1

Suppose that $\{\nu_\gamma,\ \gamma \in (0,1)\}$ is a family of integer-valued random variables with Sibuya distribution having probability generating function (p.g.f.)

$$\mathcal{S}(z) = \sum_{n=1}^{\infty} z^n \mathbb{P}\{\nu_\gamma = n\} = 1 - (1-z)^\gamma \tag{2.2.1}$$

which is independent on the sequence of X_j. Define the following normalizations $\widetilde{X}(\gamma)$ and $\widetilde{Y}(\gamma)$ of the random variable X:

$$\widetilde{X}(\gamma) = \max_{1\le k\le \nu_\gamma} X_k\,,\ \widetilde{Y}(\gamma) = \min_{1\le k\le \nu_\gamma} X_k$$

It is clear that

$$\mathbb{P}\{\widetilde{X}(\gamma) < x\} = \sum_{n=0}^{\infty} \mathbb{P}\{\max_{1\le k\le \nu_\gamma} X_k < x\}\mathbb{P}\{\nu_\gamma = n\} = \mathcal{S}(F(x)) = 1 - \big(1 - F(x)\big)^\gamma$$

and

$$\mathbb{P}\{\widetilde{Y}(\gamma) < x\} = \sum_{n=0}^{\infty} \mathbb{P}\{\min_{1\le k\le \nu_\gamma} X_k < x\}\mathbb{P}\{\nu_\gamma = n\} = 1 - \mathcal{S}(1 - F(x)) = (F(x))^\gamma$$

In this situation the role of normalized $\min_{1\le j\le n} X_j$ plays $\min_{1\le j\le n} \widetilde{X}_j(\gamma)$. Simple calculations show that the distribution function of this normalized minimum equals to

$$1 - \big(1 - F(x)\big)^{n\gamma}.$$

In the case of $\gamma = 1/n$ the distribution of the normalized minimum coincides with the initial distribution $F(x)$.

The same is true for the normalized maximum with the normalization by mean of random minima. Hence we have obtained new characterization of the Sibuya distribution.

Theorem 2.2.1. *Let $X_1, X_2, \dots$ be a sequence of i.i.d. (continuous or discrete) random variables with the probability distribution function $F(x)$ and $\widetilde{X}(p) = \max_{1\le k\le N} X_k$ ($\widetilde{Y}(p) = \min_{1\le k\le N} X_k$) be the corresponding maximum (minimum) taken over its first N terms. If N is Sibuya-distributed random variable with parameter $\gamma = 1/n$ then the distribution of normalized minima $\min_{1\le j\le n} \widetilde{X}_j(p)$ (maxima $\max_{1\le j\le n} \widetilde{Y}_j(p)$) coincides with $F(x)$.*

It would be nice to write it as unique characterization of Sibuya distribution e.g. as

 doi:10.20944/preprints202208.0481.v1

Theorem 2.2.2. *Let $X_1, X_2, \ldots$ be a sequence of i.i.d. (continuous or discrete) random variables with the probability distribution function $F(x)$ and $\widetilde{X}(p) = \max_{1\le k\le N} X_k$ ($\widetilde{Y}(p) = \min_{1\le k\le N} X_k$) be the corresponding maximum (minimum) taken over its first N terms. If N is a random variable and the distribution of normalized minima $\min_{1\le j\le n} \widetilde{X}_j(p)$ (maxima $\max_{1\le j\le n} \widetilde{Y}_j(p)$) coincides with $F(x)$ then $\mathbb{P}(N=n)$ is Sibuya distribution with parameter $\gamma = 1/n$.*

Now we see that each probability distribution function is min-stable (max-stable) with corresponding random normalization. The normalization works for arbitrary types of the random variables (including integer-valued).

Example 2.2.1. *Consider particles with random velocity V and distribution function $F(v) = \mathbb{P}(V < v)$. In ensemble of n particles the fastest one has the velocity distribution $\mathbb{P}(V_{max} < v) = (F(v))^n$ and the slowest one $\mathbb{P}(V_{min} < v) = 1-(1-F(v))^n$. Now let's assume that the number of particles is not fixed but fluctuates according to some distribution $\mathbb{P}(N=n)$ with the the p.g.f. $\mathcal{P}(z)$. In this case the velocity distribution of the fastest particle is*

$$\mathbb{P}(V_{max} < v) = \sum_{n=1}^{\infty} (F(v))^n \mathbb{P}(N=n) = \mathcal{P}(F(v))$$

Similarly $\mathbb{P}(V_{min} < v) = 1-\mathcal{P}(1-F(v))$ describes the velocity distribution of the slowest particle. It is now obvious that for the p.g.f. of Sibuya distribution (2.2.1) we obtain $\mathbb{P}(V_{max} < v) = 1-(1-F(v))^\gamma$ and $\mathbb{P}(V_{min} < v) = (F(v))^\gamma$.

3 Bernoulli thinning and a scale for positive integer random variables

3.1 General considerations

Returning back to the case of the sums of random variables let's consider a family of Bernoulli p.g.f.'s $\{Q_a(z) = 1-a+az,\ a \in (0,1)\}$. The superposition $\mathcal{P} \circ Q_a$, where $\mathcal{P}(z)$ is a p.g.f. corresponds to replacement of a random variable X with p.g.f. $\mathcal{P}$ by another random variable $\widetilde{X}(a) = \sum_{k=1}^{X} \varepsilon_k$ as it was mentioned in the Introduction. Dependence on the parameter $a \in (0,1)$ of the $\widetilde{X}(a)$ distribution is similar to that of a continuous random variate

 doi:10.20944/preprints202208.0481.v1

with scale parameter a. However, it is true for $0 < a < 1$ only. In other words, a positive integer random variable has an analogue of the scale for small (less than 1) values.

It is worth noting that for the negative binomial distribution (NBD) the superposition $\mathcal{P} \circ Q_a$ remains p.g.f. also for $a > 1$:

$$\mathcal{P}_{NBD}(z) = \left[1 + \frac{\langle n \rangle}{k}(1-z)\right]^{-k} \to \mathcal{P}_{NBD} \circ Q_a = \left[1 + a\frac{\langle n \rangle}{k}(1-z)\right]^{-k} \quad (3.1.1)$$

Using the fact that the NBD p.g.f. can be expressed as a Laplace integral of the Gamma distribution probability density with parameters k and $b = k/\langle n \rangle$

$$\left[1 + \frac{1}{b}(1-z)\right]^{-k} = \int_0^\infty e^{-(1-z)x} d\mathcal{A}(x)\,, \; d\mathcal{A}(x) = \frac{b^k}{\Gamma(k)} x^{k-1} e^{-bx} dx \quad (3.1.2)$$

we can express the superposition $\mathcal{P}_{NBD} \circ Q_a$ as:

$$\left[1 + \frac{a}{b}(1-z)\right]^{-k} = \int_0^\infty e^{-a(1-z)x} d\mathcal{A}(x) = \int_0^\infty e^{-(1-z)y} d\mathcal{A}(y/a)\,. \quad (3.1.3)$$

Starting from this insight we can ask ourselves: Is it possible to define corresponding variant of scale for large values of the parameter a, namely, for $a > 1$? The following result gives one of the possible answers.

Theorem 3.1.1. *A p.g.f. $\mathcal{P}$ possess a Bernoulli analogue of scale parameter, i.e. $\mathcal{P} \circ Q_a$ is a p.g.f. for any $a > 0$ if and only if*

$$\mathcal{P}(z) = \varphi(1-z), \quad (3.1.4)$$

where $\varphi(s)$ is Laplace transform of a probability distribution function on positive semi-axis.

Proof. 1. Suppose that

$$\varphi(s) = \int_0^\infty e^{-sx} d\mathcal{A}(x)$$

is Laplace transform of a probability distribution function on positive semi-axis. Let us define $\mathcal{P}$ using (3.1.4) and prove it is a p.g.f. Really,

$$\mathcal{P}(z) = \varphi(1-z) = \int_0^\infty e^{-x} e^{zx} d\mathcal{A}(x) = \sum_{k=0}^\infty \int_0^\infty e^{-x} x^k d\mathcal{A}(x) \cdot \frac{z^k}{k!}.$$

 doi:10.20944/preprints202208.0481.v1

It is clear the power series has positive coefficients and converges in the unit circle on the complex plain. Because $\mathcal{A}(x)$ is a probability distribution function the sum of coefficients equals to 1.

2. If $\mathcal{P}(z) = \varphi(1-z)$ then $\mathcal{P}(1-a+az)$ is a p.g.f. for any $a > 0$. Really, $\mathcal{P}(1-a+az) = \varphi(a(1-z))$ and we can apply 1. to $\varphi(as)$.

3. Suppose that $\mathcal{P}(1-a+az)$ is a p.g.f. for any $a > 0$ and prove it has representation (3.1.4). Define $\varphi(s) = \mathcal{P}(1-s)$. It is necessary to proof that $\varphi(s)$ is Laplace transform of a distribution function or, equivalently, that it is absolutely monotone function. In other words we have to proof that

$$\varphi^{(k)}(s) = (-1)^k A_k(s), \quad k = 0, 1, \dots, \tag{3.1.5}$$

where the functions $A_k(s)$ are non-negative for $s > 0$. For any $z \in (0,1)$ and any $a > 0$ we have $\varphi(a(1-z)) = \mathcal{P}(1-a+az)$. Therefore,

$$\frac{d^k}{dz^k}\varphi(a(1-z)) = (-1)^k p^k \varphi^{(k)}(a(1-z)) = (-1)^k p^k \mathcal{P}^{(k)}(1-a+az). \tag{3.1.6}$$

Because $\mathcal{P}(1-a+az)$ is probability generating function for any $a > 0$ the terms $p^k\mathcal{P}^{(k)}(1-a+az)$ are non-negative for all $0 < z < 1$ and all positive a. It proves absolutely monotones of $\varphi(s)$. The result follows from classical S.N. Bernstein theorem (see, for example [7]).

4. Let us give another (more simpler) proof of the statement 3.Because $\mathcal{P}(1-a+az)$ is probability generating function for any $a > 0$ the function $\mathcal{P}(1-n(1-e^{-s/n}))$ is Laplace transform of a distribution. We have $\lim_{n\to\infty} n(1-e^{-s/n}) = s$. Therefore $\mathcal{P}(1-s)$ is also Laplace transform of a distribution. $\square$

Denote by $\mathfrak{B}$ the class p.g.f.s of non-negative discrete random variable which possess Bernoulli analogue of scale parameter for any $a > 0$ and can therefore be represented by Eq. (3.1.4). In the following we will sometimes also call this class of p.g.f.s as the $\mathfrak{B}$-scalable

Remark 3.1.1. *Let us note that for $\mathcal{P} \in \mathfrak{B}$ and $a > 0$ the one-parameter transformations*

$$T_a\mathcal{P}(z) = T_a\varphi(1-z) = \varphi(a(1-z)) = \mathcal{P}_a(z)\,. \tag{3.1.7}$$

are automorphisms of $\mathfrak{B}$. With $T_aT_b\mathcal{P}(z) = \mathcal{P}_{ab}(z) = T_bT_a\mathcal{P}(z)$ and $T_a^{-1}\mathcal{P}(z) = T_{1/a}\mathcal{P}(z)$ it is obvious that transformations T_a form the Abelian multiplicative group.

 doi:10.20944/preprints202208.0481.v1

Corollary 3.1.1. *Consider the function* $\mathcal{R}(b,z) = \mathcal{P}(bz)/\mathcal{P}(b)$ *where* $\mathcal{P}(z) \in \mathfrak{B}$ *and* $b > 0$*. Obviously for* $0 < b < 1$ *we have* $\mathcal{R}(b,z) = \varphi_b(1-z)$ *where*

$$\varphi_b(s) = \int_0^\infty e^{-sx} d\mathcal{A}_b(x)\,, \quad d\mathcal{A}_b(x) = e^{-(1-b)x/b} \frac{d\mathcal{A}(x/b)}{\mathcal{P}(b)}\,. \tag{3.1.8}$$

For $b > 1$ *the function* $\mathcal{R}(b,z)$ *is p.g.f. if and only if*

$$\mathcal{P}(b) = b \int_0^\infty e^{-(1-b)x} d\mathcal{A}(x) < \infty\,. \tag{3.1.9}$$

Remark 3.1.2. *It is worth mentioning that the p.g.f.* $\mathcal{R}(b,z) = \mathcal{P}(bz)/\mathcal{P}(b)$ *has the form which is characteristic for a power series distribution [9]. It can be shown that in this case the cummulants* $\kappa_r, r = 1,2,\dots$ *satisfy the simple recurrence* $\kappa_{r+1} = bd\kappa_r/db$ *[14]. Thus knowing the function* $\kappa_1(b)$ *it is possible to obtain all higher cumulants and moments.*

3.2 Bernoulli thinning of the $\mathfrak{B}$-scalable p.g.f.s

Theorem 3.2.1. *[12] Let* $\mathcal{P}(z) \in \mathfrak{B}$*. Then the probabilities* $p_n = \mathcal{P}^{(n)}(0)/n!$ *satisfy the one-step recurrence*

$$(n+1)p_{n+1} = p_n g(n)\,. \tag{3.2.10}$$

Proof. From Eq.(3.1.4) we obtain

$$g(n) = \frac{(n+1)p_{n+1}}{p_n} = \frac{\mathcal{P}^{(n+1)}(0)}{\mathcal{P}^{(n)}(0)} = \frac{\int_0^\infty e^{-x} x^{n+1} d\mathcal{A}(x)}{\int_0^\infty e^{-x} x^n d\mathcal{A}(x)}\,. \tag{3.2.11}$$

□

For $T_a\mathcal{P}(z) = \mathcal{P}_a(z)$ Eq. (3.2.10) reads

$$g_a(n) = \frac{\mathcal{P}_a^{(n+1)}(0)}{\mathcal{P}_a^{(n)}(0)} = \frac{\int_0^\infty e^{-x} x^{n+1} d\mathcal{A}(x/a)}{\int_0^\infty e^{-x} x^n d\mathcal{A}(x/a)}\,. \tag{3.2.12}$$

Let us note that in general the functional dependence of $g_a(n)$ on variable n need not be the same of $g(n)$.

 doi:10.20944/preprints202208.0481.v1

Example 3.2.1. *For the NBD Eqs.(3.1.2), (3.2.10) and (3.2.12) yield*

$$g(n) = \frac{\langle n\rangle(k+n)}{\langle n\rangle + k}\,, \quad g_a(n) = \frac{a\langle n\rangle(k+n)}{a\langle n\rangle + k} = g(n)\cdot\frac{a(\langle n\rangle + k)}{a\langle n\rangle + k}\,.$$

In this case the functional dependence of $g(n)$ and $g_a(n)$ on n is the same.

Our question is what are the conditions for factorization of $g_a(n)$ as a product of functions dependent on n and a separately. From Eqs. 3.2.11 and 3.2.12 we obtain

$$\frac{g_a(n)}{g(n)} = \frac{\int_0^\infty e^{-x}x^{n+1}d\mathcal{A}(x/a)}{\int_0^\infty e^{-x}x^{n}d\mathcal{A}(x/a)} \Big/ \frac{\int_0^\infty e^{-x}x^{n+1}d\mathcal{A}(x)}{\int_0^\infty e^{-x}x^{n}d\mathcal{A}(x)}. \tag{3.2.13}$$

Obviously, it is equivalent to writing

$$\frac{\int_0^\infty e^{-x}x^{n+1}d\mathcal{A}(x/a)}{\int_0^\infty e^{-x}x^{n}d\mathcal{A}(x/a)} = h_n\cdot k(a)\,. \tag{3.2.14}$$

With $\psi(a) = \int_0^\infty e^{-x}d\mathcal{A}(x/a) = \int_0^\infty e^{-ax}d\mathcal{A}(x)$ relation (3.2.14) can be written as

$$\frac{\psi^{(n+1)}(a)}{\psi^{(n)}(a)} = -h_n\cdot\frac{k(a)}{a}\,. \tag{3.2.15}$$

Integrating both sides of Eq.(3.2.15) with respect to a from zero to u we obtain

$$\log(\psi^{(n)}(u)) = -h_n\cdot K(u) + log(c_n)\ ,\ K(u) = \int_0^u \frac{k(a)}{a}\,da$$

where $c_n = \psi^{(n)}(0)$ does not depend on u. Therefore,

$$\psi^{(n)}(u)) = c_n\exp\{-h_n\cdot K(u)\} \tag{3.2.16}$$

and from (3.2.15)

$$c_{n+1}\exp\{-h_{n+1}\cdot K(u)\} = -h_n c_n\exp\{-h_n\cdot K(u)\}\cdot K'(u). \tag{3.2.17}$$

For $n = 1$ the equation (3.2.17) reduces to

$$K'(u)\exp\{(h_2 - h_1)\cdot K(u)\} = -D_1, \tag{3.2.18}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 August 2022 doi:10.20944/preprints202208.0481.v1

where $D_1 = h_1 c_1 / c_2$ is a constant. After integration of this relation with respect to u we obtain

$$K(u) = \frac{1}{h_2 - h_1} \log\Big(D_2 - D_1(h_2 - h_1)u\Big),$$

where D_2 is a constant. This expression of $K(u)$ allows us to find $\psi'(u)$, and consequently, $\psi^n(u)$ for all $n \geq 1$. Namely,

$$\psi'(u) = -\int_0^\infty x e^{-ux} d\mathcal{A}(x) = c_1 \cdot (D_2 - D_1(h_2 - h_1)u)^{-\gamma}, \tag{3.2.19}$$

where $\gamma = h_1/(h_2 - h_1)$. The function $\psi'(u)$ is the first derivative of the Laplace transform of a probability measure. Therefore, $c_1 < 0$, $D_1 < 0$, $h_2 > h_1$ and $\gamma > 0$. It is easy to see that for this function the relation (3.2.15) holds. Comparing equations (3.2.19) and (3.1.2) we find that $\mathcal{P}(z) = \mathcal{P}_{NBD}(z)$ with parameters $k = \gamma - 1$ and $b = -D_1(h_2 - h_1)/D_2$. So, we proved the following result.

Theorem 3.2.2. *Let $\mathcal{P}(z)$ be a p.g.f. of the form*

$$\mathcal{P}(z) = \int_0^\infty \exp\{-(1-z)x\} d\mathcal{A}(x),$$

where $\mathcal{A}(x)$ is a non-degenerate probability distribution function on $[0, \infty)$. The relation $g_a(n) = g(n) \cdot k(a)$ holds if and only if $\mathcal{P}(z)$ is p.g.f. of negative binomial distribution.

3.3 $\mathfrak{B}$-scalability and moments

Let us give now different condition for the existence of Bernoulli analogue of scale parameter.

Theorem 3.3.1. *Suppose that X is non-negative integer random variable with the p.g.f. $\mathcal{P}(z)$ having the probabilities of its values as*

$$p_k = \mathbb{P}\{X = k\}.$$

Suppose that Stieltjes problem of moments

$$p_k = \frac{1}{k!} \int_0^\infty x^k e^{-x} d\mathcal{A}(x), \quad k = 0, 1, \ldots \tag{3.3.20}$$

has a solution $\mathcal{A}$. Then $\mathcal{P} \in \mathfrak{B}$.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 August 2022 doi:10.20944/preprints202208.0481.v1

Proof. Let $\mathcal{A}$ be a solution of the problem (3.3.20). Because $\sum_{k=0}^{\infty} p_k$ equals to 1 then $\mathcal{A}$ is a probability distribution function. Define

$$\varphi(s) = \int_0^{\infty} e^{-sx} d\mathcal{A}(x).$$

Now it is easy to calculate that $\mathcal{P}(z) = \varphi(1-z)$ and the rest follows from Theorem 3.1.1. □

Corollary 3.3.1. *The relations (3.3.20) show that for non-trivial p.g.f. $\mathcal{P} \in \mathfrak{B}$ the following statements are true:*

1. $p_k > 0$ *for all* $k = 0, 1, \ldots$. *Particularly, p.g.f. of a random variable taking finite set of values is not an element of* $\mathfrak{B}$;

2. $$\Delta_n = \det \begin{pmatrix} p_o & p_1 & \ldots & n! \cdot p_n \\ p_1 & 2 \cdot p_2 & \ldots & (n+1)! \cdot p_{n+1} \\ \ldots & \ldots & \ldots & \ldots \\ n! \cdot p_n & (n+1)! \cdot p_{n+1} & \ldots & (2n)! \cdot p_{2n} \end{pmatrix} \geq 0 \tag{3.3.21}$$

   for all $n = 0, 1, \ldots$. *It follows from considering Hamburger problem of moments for the function* $e^{-s} d\mathcal{A}(x)$ *(see [4]). It is necessary condition for the solution of corresponding Stieltjes problem of moments. For example, if* $2p_o p_2 < p_1^2$ *then* $\mathcal{P} \notin \mathfrak{B}$. *This relation holds for Binomial distribution with the parameter* $0 < p < 1$. *Really, for Binomial distribution*

   $$2p_o p_2 - p_1^2 = -np^2(1-p)^{2n-2} < 0.$$

3. *It is clear that* $x^k e^{-x}/k! \leq 1/e$ *for all* $x > 0$ *and all positive integers* k. *Therefore,* $p_k \leq 1/e$ *for* $k = 0, 1, 2 \ldots$. *Therefore, if we have a random sample from the population with probability distribution function* $\mathcal{A}(x)$ *we can easily estimate probabilities* p_k *of corresponding discrete distribution.*

It is clear that a random variable with Poisson distribution possesses Bernoulli analogue of scale parameter.

Theorem 3.3.2. *Let* $\mathcal{P}(z)$ *be a p.g.f. from* $\mathfrak{B}$ *possessing finite second derivative* λ *at* $z = 1$. *Then*

$$\mathcal{P}(z) \geq \exp\{\lambda(z-1)\}. \tag{3.3.22}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 August 2022 doi:10.20944/preprints202208.0481.v1

If the equality in (3.3.22) is attend at one $z_o < 1$ then the equality holds for all $|z| \leq 1$.

Proof. Because $\mathcal{P} \in \mathfrak{B}$ then $\mathcal{P}(z) = \varphi(1-z)$ where $\varphi(s) = \int_0^\infty \exp\{-sx\} d\mathcal{A}(x)$ for some probability distribution function $\mathcal{A}$.

Let us show that

$$\varphi(s)\varphi''(s) - \big(\varphi'(s)\big)^2 \geq 0. \tag{3.3.23}$$

Really,

$$\varphi(s)\varphi''(s) - \big(\varphi'(s)\big)^2 = \int_0^\infty e^{-su} d\mathcal{A}(u) \int_0^\infty u^2 \cdot e^{-su} d\mathcal{A}(u) - \Big(\int_0^\infty u \cdot e^{-su} d\mathcal{A}(u)\Big)^2.$$

But the Cauchy-Bunyakovski inequality gives

$$\int_0^\infty u \cdot e^{-su} d\mathcal{A}(u) \leq \Big(\int_0^\infty e^{-su} d\mathcal{A}(u) \cdot \int_0^\infty u^2 \cdot e^{-su} d\mathcal{A}(u)\Big)^{1/2}.$$

Therefore (3.3.23) is true.

Inequality (3.3.23) is equivalent to

$$\frac{d^2 \log(\varphi(s))}{ds^2} \geq 0,$$

which means $\log(\varphi(s))$ is a convex function. Therefore $\log\big(\varphi(s)\big) \geq -\lambda s$ so that we have (3.3.22).

Moreover, if for any $z_o < 1$ there is equality in (3.3.22) then, in view of convexity, there must be equality on the interval $[z_o, 1]$. From the analytic character of $\mathcal{P}$ the equality holds for all $|z| \leq 1$. □

Let us note that the Theorem 3.3.2 can be obtained from one result by S.N. Bernstein [8].

Corollary 3.3.2. *Suppose that X is a random variable with p.g.f. $\mathcal{P}(z) \in \mathfrak{B}$ and second factorial moment $\lambda = \mathbb{E}(X(X-1))$. Then for $0 < \lambda < \infty$ and $1/2 < r < 1$*

$$\mathbb{E}X^r \leq \frac{1}{-\Gamma(-r)} \int_0^1 \frac{1 - \exp(\lambda(z-1))}{z(-\log(z))^{1/r}} dz \tag{3.3.24}$$

with equality if and only if X has Poisson distribution with the parameter λ.

Proof. The result follows from Theorem 3.3.2 and the expression for fractional moments from [12]. Note that $\mathbb{E}(X) > 1$ implies $\lambda > 0$. □

It is clear that the p.g.f. from $\mathfrak{B}$ cannot have zeros for $z \geq 0$ because $\varphi(s)$ is Laplace transform of a probability distribution function. Therefore, any p.g.f. taking zero value at point $z = 0$ is not an element of $\mathfrak{B}$.

Let us give some conditions for a p.g.f. to be an element of the class $\mathfrak{B}$. Suppose that $\mathcal{P}(z)$ is a p.g.f. with finite moments of all orders. Its factorial moments may be calculated as $\mathcal{P}^{(n)}(1)$, $n = 0, 1, 2, \ldots$. If $\mathcal{P} \in \mathfrak{B}$ then $\mathcal{P}(z) = \varphi(1-z)$, where $\varphi(s) = \int_0^\infty \exp\{-sx\} d\mathcal{A}(x)$. From here we see that

$$\mathcal{P}^{(n)}(1) = \int_0^\infty x^n d\mathcal{A}(x), \quad n = 0, 1, \ldots. \tag{3.3.25}$$

Hence the necessary condition for $\mathcal{P} \in \mathfrak{B}$ is that the sequence $\{\mathcal{P}^n(1),\ n = 0, 1, \ldots\}$ represents a solution of Stieltjes problem of moments. To have corresponding solution of this problem it is necessary and sufficient that the forms

$$\sum_{i,k=0}^{n} \mathcal{P}^{(i+k)}(1) x_i x_k, \quad \sum_{i,k=0}^{n} \mathcal{P}^{(i+k+1)}(1) x_i x_k \tag{3.3.26}$$

are positive for all $n \in \mathbb{N}$ (see, for example [4]). Of course, if the condition (3.3.26) is true and the function reconstructs by the moments in the unique way then the (3.3.25) holds, and we see that $\mathcal{P} \in \mathfrak{B}$. It means the following result holds.

Theorem 3.3.3. *Let $\{\mu_n,\ n = 0, 1, \ldots\}$, ($\mu_0 = 1$) be a sequence of positive numbers. Suppose that the forms*

$$\sum_{i,k=0}^{n} \mu_{i+k} x_i x_k, \quad \sum_{i,k=0}^{n} \mu_{i+k+1} x_i x_k \tag{3.3.27}$$

are positive for all $n \in \mathbb{N}$. Then there exists a p.g.f. $\mathcal{P}(z)$ such that: 1) $\mathcal{P} \in \mathfrak{B}$; 2) μ_n is the n-th factorial moment of $\mathcal{P}(z)$ for all $n \in \mathbb{N}$.

Let us note the following facts. Let $\{s_k, k = 0, 1, \ldots\}$ be a sequence of real numbers. This sequence is positive if and only if all determinants

$$D_n = \det \begin{pmatrix} s_0 & s_1 & \ldots & s_n \\ s_1 & s_2 & \ldots & s_{n+1} \\ \ldots & \ldots & \ldots & \ldots \\ s_n & s_{n+1} & \ldots & s_{2n} \end{pmatrix} \quad (n = 0, 1, \ldots) \tag{3.3.28}$$

are non-negative.

 doi:10.20944/preprints202208.0481.v1

Suppose that $\mathcal{P}(z)$ is a p.g.f. determining by the sequence $\{\mathcal{P}^{(n)}(1)\}$, $n \in \mathbb{N}$ in unique way (for example, analytic in a neighborhood of the point $z = 1$). Connected to the forms (3.3.3) are the determinants

$$D_n^{(1)} = \det \begin{pmatrix} 1 & \mathcal{P}'(1) & \dots & \mathcal{P}^{(n)}(1) \\ \mathcal{P}'(1) & \mathcal{P}''(1) & \dots & \mathcal{P}^{(n+1)}(1) \\ \dots & \dots & \dots & \dots \\ \mathcal{P}^{(n)}(1) & \mathcal{P}^{(n+1)}(1) & \dots & \mathcal{P}^{(2n)}(1) \end{pmatrix} \tag{3.3.29}$$

and

$$D_n^{(2)} = \det \begin{pmatrix} \mathcal{P}'(1) & \mathcal{P}''(1) & \dots & \mathcal{P}^{(n+1)}(1) \\ \mathcal{P}''(1) & \mathcal{P}^{(3)}(1) & \dots & \mathcal{P}^{(n+2)}(1) \\ \dots & \dots & \dots & \dots \\ \mathcal{P}^{(n+1)}(1) & \mathcal{P}^{(n+2)}(1) & \dots & \mathcal{P}^{(2n+1)}(1) \end{pmatrix}. \tag{3.3.30}$$

If all determinants (3.3.29) and (3.3.29) are non-negative then the p.g.f. $\mathcal{P}(z)$ belongs to the class $\mathfrak{B}$. Additionally, if for some n_o at least one of the determinants is zero then $\mathcal{P}(z)$ is a mixture of no more than n_o Poisson distributions. Let us mention that the results on positive sequences are taken from [4] and [5].

Now we can give a characteristic property of Poisson distribution.

Theorem 3.3.4. *Let $\mathcal{P}(z) \in \mathfrak{B}$ be a p.g.f. with finite second moment. Denote μ_j $(j = 1, 2)$ its factorial moments. $\mathcal{P}(z)$ is Poisson p.g.f. if and only if $\mu_2/\mu_1^2 = 1$.*

Proof. Because $\mathcal{P} \in \mathfrak{B}$ then

$$\mathcal{P}(z) = \int_0^\infty e^{-(1-z)x} d\mathcal{A}(x) \quad \text{and} \quad \mathcal{P}^{(n)}(1) = \int_0^\infty x^n d\mathcal{A}(x). \tag{3.3.31}$$

The condition $\mu_2/\mu_1^2 = 1$ is equivalent to $D_1^{(1)} = 0$. Therefore the distribution function $\mathcal{A}$ has only one point of growth. The statement follows from (3.3.31). □

What can one say if the condition $D_1^{(1)} = 0$ change by $D_1^{(1)} \le \varepsilon^2$, where $\varepsilon > 0$ is small enough? The answer is given by the following result.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 August 2022 doi:10.20944/preprints202208.0481.v1

Theorem 3.3.5. *Let $\mathcal{P}(z) \in \mathfrak{B}$ be a p.g.f. with finite second moment. Denote μ_j $(j = 1, 2)$ its factorial moments. Suppose that $0 \le \mu_2 - \mu_1^2 \le \varepsilon^2$ for a fixed $\varepsilon \ge 0$. Then*

$$\sup_{|z|\le 1}\Big|\mathcal{P}(z) - \exp\{\mu_1(z-1)\}\Big| \le \varepsilon^2. \tag{3.3.32}$$

The relation (3.3.32) means that $\mathcal{P}(z)$ is close to Poisson p.g.f. for real $z \in [-1, 1]$.

Proof. Because $\mathcal{P} \in \mathfrak{B}$ we have

$$\mathcal{P}(z) = \int_0^\infty \exp\{(z-1)x\} d\mathcal{A}(x). \tag{3.3.33}$$

As it was shown before, the factorial moments of $\mathcal{P}$ are the usual moments of $\mathcal{A}$. Therefore, $\mu_2 - \mu_1^2 = \sigma^2$ is the variance of $\mathcal{A}$ and we know that $\sigma \le \varepsilon$. It is enough to estimate the difference

$$\mathcal{P}(z) - \exp\{\mu_1(z-1)\} = \int_0^\infty \Big(\exp\{(z-1)x\} - \exp\{\mu_1(z-1)\}\Big) d\mathcal{A}(x).$$

However, according to Taylor formula

$$\exp\{(z-1)x\} - \exp\{\mu_1(z-1)\} = (z-1)(x-\mu_1)\exp\{\mu_1(z-1)\}+$$
$$+\frac{(z-1)^2}{2}(x-\mu_1)^2\exp\{\tilde{x}(z-1)\},$$

where $\tilde{x}$ is a point, depending on $x \ge 0$ and z and lying between x and μ_1. For $-1 \le z \le 1$ we have

$$\Big|\frac{(z-1)^2}{2}(x-\mu_1)^2\exp\{\tilde{x}(z-1)\}\Big| \le (x-\mu_1)^2.$$

Therefore,

$$\Big|\mathcal{P}(z) - \exp\{\mu_1(z-1)\}\Big| = \Big|\int_0^\infty \Big(\exp\{(z-1)x\} - \exp\{\mu_1(z-1)\}\Big) d\mathcal{A}(x)\Big| \le$$

$$\le \Big|\int_0^\infty \Big((z-1)(x-\mu_1)\exp\{\mu_1(z-1)\} + \frac{(z-1)^2}{2}(x-\mu_1)^2\exp\{\tilde{x}(z-1)\}\Big) d\mathcal{A}(x)\Big|$$

$$\le \int_0^\infty (x-\mu_1)^2 d\mathcal{A}(x) = \mu_2 - \mu_1^2 \le \varepsilon^2.$$

$\square$

 doi:10.20944/preprints202208.0481.v1

Let us note that the condition $\mathcal{P} \in \mathfrak{B}$ in Theorem 3.3.4 is essential. Really, the mixture distribution with equal probabilities of Geometric distribution with parameter p_1 and Bernoulli distribution with parameter $p_2 = 1/(1+p_1)$ gives $\mu_2 - \mu_1^2 = 0$ for factorial moments μ_1 and μ_2. It is obvious that Theorem 3.3.5 is not true for this case as well.

4 Other thinning operators

The following definition had been given in [6].

Definition 4.0.1. *Let X be a random variable with p.g.f. $\xi(z)$, and let $\mathcal{Q} = \{Q_a(z),\ a \in (0,1)\}$ be a family of p.g.f.s. We say the family $\mathcal{Q}$ is thinning with respect to $\xi(z)$ if*

$$\xi(Q_a(z)) = (1-a) + a\xi(z) \tag{4.0.1}$$

for all $|z| \le 1$ and $a \in [0,1]$.

Remark 4.0.1. *Consider group $\mathbb{G}$ of all non-negative strictly monotone functions $g(z)$ equipped with the binary operation $g_1 \circ g_2 = g_1(g_2(z))$ and unit element $g^{-1} \circ g = z$. Obviously, if $\xi(z)$ is the p.g.f. then $\xi(z) \in \mathbb{G}$ as well as its inverse function $\xi^{-1}(z) \in \mathbb{G}$. Moreover, since $\xi^{-1}(1) = \xi(1) = 1$ the group elements with $g(1) = 1$ form the subgroup $\mathbb{G}_1 \subset \mathbb{G}$. Consequently, Eq. (4.0.1) is equivalent to the similarity transformation*

$$Q_a(z) = \xi^{-1} \circ \mathcal{B}_a \circ \xi\,, \ \mathcal{B}_a(z) = 1 - a + az\,. \tag{4.0.2}$$

between two conjugate elements of the group $\mathbb{G}_1$ – Bernoulli p.g.f. $\mathcal{B}_a(z)$ and function $Q_a(z)$. Let us recall that $Q_a(z) = \sum_n a_n z^n$ it is the p.g.f. if $\forall n\,,\ a_n \ge 0$.

Following [6] note that if the family $\mathcal{Q}$ is thinning with respect to ξ then

$$\mathcal{R}(z) = \exp\{\lambda\big(\xi(z) - 1\big)\}, \quad (\lambda > 0) \tag{4.0.3}$$

is discrete stable distribution with the thinning operator $\mathcal{Q}$ (see the definitions in [3] and [2]). So, each example of the thinning family gives at the same moment an example of the corresponding discrete stable distribution.

One of essential properties of the thinning families is their commutativity under superposition.

Proposition 4.0.1. *[6] Let $\mathcal{Q} = \{Q_a(z)\ a \in (0,1)\}$ be a family of the thinning operators with respect to p.g.f. $\xi(z)$. Then*

$$Q_a \circ Q_b = Q_b \circ Q_a \tag{4.0.4}$$

for all $a, b \in (0,1)$, where $\circ$ is the superposition sign.

It is also clear that $Q_a \circ Q_b = Q_{ab}$ for all $a, b \in [0,1]$.

Many examples of commuting families of p.g.f.s are given in [1, 6]. Below we give some additional examples of such type in connection to the thinning families.

Example 4.0.1. *Starting from the p.g.f. of Sibuya distribution $\xi(z) = 1 - (1-z)^\gamma$ where $0 \le \gamma < 1$. With $\xi^{-1}(u) = 1 - (1-u)^{1/\gamma}$ Eq.(4.0.2) yields the family*

$$Q_a(z) = 1 - a^{1/\gamma}(1-z)\,,$$

which is the thinning with respect to $\xi(z)$. The p.g.f.

$$\mathcal{R}(z) = e^{\lambda(1-z)^\gamma}$$

corresponds to ordinary discrete stable distribution satisfying

$$\mathcal{R}^n(Q_{1/n}(z)) = \mathcal{R}(z)$$

for all $n \in \mathbb{N}$.

Example 4.0.2. *Suppose that a p.g.f. is defined as $\xi(z) = pz/(1-(1-p)z)$ for a fixed $0 < p < 1$. Then $\xi^{-1}(u) = u/(p+(1-p)u)$. Define a family of p.g.f.s*

$$Q_a(z) = \frac{1-a-(1-a-p)z}{1-a(1-p)-(1-p)(1-a)z},\ \ 0 < a \le 1-p.$$

It is clear that

$$\xi(Q_a(z)) = 1 - a + a\xi(z), \quad \text{and} \quad 1 - \xi(Q_a(z)) = a(1-\xi(z)).$$

Now we see the family $\{Q_a(z),\ a \in (0, 1-p)\}$ is the thinning with respect to $\xi(z)$. The p.g.f.

$$\mathcal{R}(z) = \exp\left\{\frac{\lambda(z-1)}{1-(1-p)z}\right\}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 August 2022 doi:10.20944/preprints202208.0481.v1

is an analogue of discrete stable distribution in the sense

$$\mathcal{R}^n(Q_{1/n}(z)) = \mathcal{R}(z)$$

for all $n \in \mathbb{N}$.

Consider the p.g.f. $\xi(z) = (1-q)z/(1-qz)$ of random variable $Y = X+1$ where X has the geometric distribution. From $\xi(z)$ and its inverse function $\xi^{-1}(u) = u/(1-q+qu)$ we obtain the family of p.g.f.s

$$Q_a(z) = \xi^{-1} \circ \mathcal{B}_a \circ \xi = \frac{1-a-(q-a)z}{1-aq-q(1-a)z}, \; 0 < a \le q$$

The p.g.f.

$$\mathcal{R}(z) = \exp\Big\{\frac{\lambda(z-1)}{1-qz}\Big\}$$

is an analogue of discrete stable distribution in the sense

$$\mathcal{R}^n(Q_{1/n}(z)) = \mathcal{R}(z)$$

for all $n \in \mathbb{N}$.

Suppose that $\varphi(s)$ is Laplace transform of a probability distribution function, ξ and Q_a are defined in Example 4.0.2. The function $\varphi(1-\xi(z))$ is a p.g.f. in view of Theorem 3.1.1. Using Eq. (4.0.1), we have $\varphi(1-\xi(Q_a(z))) = \varphi(a(1-\xi(z)))$. Although, Q_a is not p.g.f. for sufficiently large values of $a > 0$ the function $\varphi(1-\xi(Q_a(z)))$ remains to be p.g.f. for all positive a. Therefore, we have special example of a random variable possessing an analogue of scale parameter. It is not Bernoulli analogue, but a new type of scale parameter.

This example of the scale parameter definition is rather general one.

Definition 4.0.2. *Let $\{Q_a(z), \; a \in (0,\varepsilon)\}$ for $\varepsilon > 0$ be a family of p.g.f.s thinning for $\xi(z)$. Suppose that $\varphi(s)$ is a Laplace transform of a probability distribution on $\mathbb{R}_+$. Then $\mathcal{P}(z) = \varphi(1-\xi(z))$ possesses Q_a-type of scale parameter $a > 0$.*

Example 4.0.3. *Let $\xi(z) = \xi(z, b, p, m)$ be a p.g.f. of the form*

$$\xi(z) = 1 - b\Big(\frac{1-z^m}{1-\kappa z^m}\Big)^p,$$

 doi:10.20944/preprints202208.0481.v1

where $0 < p \leq 1$, $m \in \mathbb{N}$, $0 < b \leq 1$ *and* $0 \leq \kappa < 1$. *For* $b = m = p = 1$ *we come to Example 4.0.2, for* $b = m = \kappa = 1$ *to Example 4.0.1. Therefore, consider the case* $m > 1$. *With* $\xi^{-1}(u) = ((1-u)^{1/p} - b^{1/p})/((1-u)^{1/p}\kappa - b^{1/p})$ *we obtain*

$$Q_a(z) = \xi^{-1} \circ \mathcal{B}_a \circ \xi = \Big(\frac{1 - a^{1/p} + (a^{1/p} - \kappa)z^m}{1 - \kappa a^{1/p} - \kappa(1 - a^{1/p})z^m}\Big)^{1/m},$$

where $0 < a < \kappa^p < 1$. *It is not difficult to verify that both* ξ *and* Q_a *are p.g.f.s. Simple calculations shows that*

$$\xi(Q_a(z)) = 1 - ab\Big(\frac{1 - z^m}{1 - \kappa z^m}\Big)^p = 1 - a + a\xi(z).$$

For $\varphi(1-\xi(z))$ *we have* Q_a*-type of scale parameter* $a > 0$. *The discrete stable version is*

$$\mathcal{R}(z) = \exp\{\lambda(\xi(z) - 1)\}$$

for which we have

$$\mathcal{R}^n(Q_{1/n}(z)) = \mathcal{R}(z).$$

The next Example has essentially different form.

Example 4.0.4. *Let*

$$\xi(z) = 1 - b\frac{\log(1 + B - Bz)}{\log(1 + B)},$$

where $B > 0$ *and* $b \in (0, 1)$ *are parameters. It is clear that*

$$\xi(z) = 1 - b + b\sum_{k=1}^{\infty} \frac{(1 - 1/(1 + B))^k}{k \log(1 + B)} z^k$$

is p.g.f. Let

$$Q_a(z) = \frac{(B + 1)}{B} - \frac{(B + 1)^a}{B}\Big(1 - \frac{B}{B + 1}z\Big)^a, \quad a \in (0, 1).$$

It is easy to verify $Q_a(z)$ *is a p.g.f. Also we have*

$$\xi(Q_a(z)) = 1 - a + a\xi(z),$$

so that the family $\{Q_a(z),\ a \in (0, 1)\}$ *is thinning with respect to* $\xi(z)$.

Define

$$\mathcal{R}(z) = \exp\{\lambda(\xi(z) - 1)\}.$$

Clearly

$$\mathcal{R}^n(Q_{1/n}(z)) = \mathcal{R}(z).$$

In other words, $\mathcal{R}(z)$ is a discrete stable distribution. It is not difficult to see that $\mathcal{R}(z)$ is p.g.f. of negative binomial distribution.

Let us give a construction of p.g.f. allowing generalized scale parameter. Suppose that

$$\varphi(s) = \int_0^\infty \exp\{-sx\} d\mathcal{A}(x)$$

and a family $\{Q_a(z),\ a \in (0,1)\}$ is thinning with respect to $\xi(z)$. Define the following p.g.f

$$\mathcal{P}(z) = \varphi(1 - \xi(z)). \tag{4.0.5}$$

For arbitrary $a \in (0,1)$ let us consider

$$\mathcal{P}_{1/a}(z) = \varphi((1 - \xi(z))/a),$$

which is again a p.g.f. We have

$$\mathcal{P}_{1/a}(Q_a(z)) = \varphi((1 - \xi(Q_a(z)))/a) = \varphi((1 - (1 - a) - a\xi(z))/a)$$

$$= \varphi(1 - \xi(z)) = \mathcal{P}(z).$$

This relation allows us to introduce generalized scale parameter.

5 Moments of integer-valued heavy-tailed random variables

Consider integer-valued non-negative random variable X with the p.g.f. $Q(z)$. If $\mathbb{E}X < \infty$ then for $\forall r$ such that $0 < r < a < 1$ also $\mathbb{E}X^r < \infty$. In the opposite case the random variable X is called heavy-tailed and its r-th moments $0 < r < a < 1$ do exist provided [12]

$$\int_0^1 \frac{1 - Q(z)}{(-\log z)^{1+r} z} dz < \infty\,. \tag{5.0.1}$$

 doi:10.20944/preprints202208.0481.v1

Consider now a special case of such heavy-tailed integer-valued random variable whose the p.g.f. can be $\forall n \in \mathbb{N}$ expressed as $Q = \xi_1 \circ \xi_2 \circ \ldots \circ \xi_n$ where $\xi_i(z) = \xi(z), i = \overline{1,n}$ are the p.g.f.s of independent equally distributed non-negative integer-valued random variables $X_i, i = \overline{1,n}$. Obviously $X \stackrel{\mathrm{d}}{=} X_1 \cdot X_2 \cdot \ldots \cdot X_n$, where $\stackrel{\mathrm{d}}{=}$ means equality in distribution, and for $r = m/n, m < n, m \in \mathbb{N}, n \in \mathbb{N}$ the r-th absolute moment $\mathbb{E}X^r = \mathbb{E}(X_1 \cdot X_2 \cdot \ldots \cdot X_n)^r$ can be interpret as the mean of the m-th power of the geometric average $\sqrt[n]{X_1 \cdot X_2 \cdot \ldots \cdot X_n}$ of random variables $X_i, i = \overline{1,n}$.

References

[1] Klebanov L.B., Kakosyan A.V., Rachev S.T., Temnov G. (2012) On a class of distributions stable under random summation. – J. Appl. Prob., **49**, 303–318.

[2] Klebanov L.B., Slamova L. (2014) Discrete stable and casual stable random variables. – arXiv:1406.3748v1 [math.PR], 1–8.

[3] Steutel, F. W. and van Harn, K. (1979). Discrete analogues of selfdecomposability and stability.–Annals of Probabability, **7** (5):893899.

[4] Akhiezer, N. I., The Classical Moment Problem and Some Related Questions in Analysis, Oliver & Boyd, Edinburg and London 1965.

[5] Krein M.G., Nudelman A. A., The Markov Moment Problem and Extremal Problems, Translations of Mathematical Monographs, Vol. 50, American Mathematical Society (December 31, 1977).

[6] Klebanov L.B. (2021) On normalizations of integer-valued random variables (in Russian), Zapiski Nauchnih Seminarov POMI, v. 505, 138–146.

[7] Widder D.V. (1946) The Laplace Transform, Princeton University Press, London.

[8] Bernstein S.N. ...

[9] Norman L. Johnson, Adrienne W. Kemp, Samuel Kotz, Univariate Discrete Distributions, John Wiley and Sons, 2005.

 doi:10.20944/preprints202208.0481.v1

[10] W. Feller, An Introduction to Probability Theory and Its Applications, Vol. 1, John Wiley and Sons, Inc., New York 1971

[11] Wiliam J. Anderson, Continuous-time Markov chains, Springer series in statistic, 1991

[12] Klebanov L.B., Šumbera M., On Sibuya-like distributions in branching and birth-and-death processes, arXiv:2204.08871 [math.PR], 1–18.

[13] G. Christoph and K. Schreiber, Scaled Sibuya distribution and discrete self-decomposability, Statistics and Probability Letters 48(2000)181.

[14] Khatri, C. G., On certain properties of power-series distributions, Biometrika, 46, 486–490.