

Dediscretization of Distributions Arising in Macroevolution Models

Jaakko Astola and Eduard Danielian

Abstract: The standard birth-death process with intensities of moderate growth generates stationary skewed distributions suitable for modelling frequency distributions of events arising in large-scale biomolecular systems.

We study a large class of such distributions that can be used to model, for instance, frequency distributions of the number of expressed genes in the transcriptome, the number of protein domain occurrences in the proteomes, etc.

In the present paper a new *dediscretization* approach is suggested, discussed and applied to the chosen class. This approach conserves the *qualitative* properties of the original class of distributions.

The advantages of the approach consist in following:

1. It simplifies the form of distributions;
2. It allows simple mathematical analysis of the properties of the original class by applying the tools mathematical analysis continuous functions
3. It allows to find out the *optimal* form of stationary distributions, i.e. suggests new classes of distributions for biomolecular applications.

The deviations of the dediscretized continuous distribution functions from the original distribution functions is estimated.

Several typical examples are considered which illustrate the possibilities of the dediscretization approach.

The reverse procedure to *dediscretization*, i.e. the procedure of *discretization*, back to discrete distributions is described.

Keywords: Dediscretization, distribution, distribution function, regular variation, skewness, convexity, biomolecular systems.

Manuscript received on April 17, 2007.

J. Astola is with TICSP, Tampere University of Technology, Tampere, Finland (e-mail: Jaakko.Astola@tut.fi). E. Danielian is with Dept. of Mathematics, Yerevan State University, Yerevan, Armenia.

1 Introduction

The mechanism of the dynamic of a large-scale biomolecular system often is explained with the help of standard birth-death process with various specific constraints on its intensities (coefficients). The stationary distributions of the process, which *always* have a skew to the right, and may be used as frequency distributions of different events taking place in large-scale biomolecular systems. Recently, a huge class of such distributions with *moderate* growth of the intensities of the process has been obtained [1]. Moreover, in [2] this class was verified to satisfy will known *empirical facts* in macroevolution of biomolecular systems. As a result of this, under some very natural assumptions on the process' intensities we extracted a *subclass*, of *regularly varying* distributions. The last class includes *all* well known and widely used distributions designed to model the empirical *frequency* distribution of the number of different type of events arising in large-scale biomolecular systems, in particular, of the number of expressed genes in the transcriptome and the number of protein domain occurrences in the proteomes (see, for information, [3,4]).

1.1 The Description of the Class

In order to define the above mentioned class denote by Λ the class of regularly varying with exponent $\alpha \in [1, +\infty)$ (see, [5]) increasing sequences $\{\delta_n\}$ with $\delta_0 = 1$ satisfying following conditions: $\{\delta_n\}$ is

- (a) downward convex, i.e. for $n = 1, 2, \dots$

$$\delta_{n-1} + \delta_{n+1} > 2\delta_n,$$

- (b) log-upward convex, i.e. for $n = 1, 2, \dots$

$$\delta_n^2 > \delta_{n-1} \cdot \delta_{n+1}.$$

The assumptions (a) and (b) on convexity are necessary to have an analyzable class that satisfies empirical facts in [6].

A *Special* class Λ_0 includes increasing sequences $\{\delta_n\}$ of the type

$$\delta_n = 1 + \frac{n}{A}(1 + o(1)), \quad n \rightarrow +\infty, \quad A \in \mathbb{R}^+ = (0, +\infty). \quad (1.1)$$

In other words, (1.1) means that

$$\delta_n = n \cdot L(n) + 1, \quad n = 0, 1, 2, \dots,$$

where $\{L(n)\}$ is a *slowly varying* sequence satisfying condition

$$\lim_{n \rightarrow +\infty} L(n) = A^{-1}.$$

The case of sequences of type (1.1) with $o(1) = 0$ in (1.1) or $L(n) = A^{-1}$, $n = 0, 1, 2, \dots$, is called the *linear* case.

Now, if $\{\delta_n\} \in \Lambda$, then the following limit exists (see, [2])

$$\lim_{n \rightarrow +\infty} \frac{n}{\delta_n} = 0, \quad (1.2)$$

and then under (1.2) two possibilities exist. Either

$$\sum_{n \geq 1} \frac{1}{\delta_n} = +\infty \quad (\text{then, denote } \delta_n = \delta_n^-, n = 1, 2, \dots), \quad (1.3)$$

or

$$\sum_{n \geq 1} \frac{1}{\delta_n} < +\infty \quad (\text{then, denote } \delta_n = \delta_n^+, n = 1, 2, \dots). \quad (1.4)$$

Let the classes Λ_+ and Λ_- be formed by sequences $\{\delta_n^+\}$ and $\{\delta_n^-\}$, respectively. Then,

$$\Lambda = \Lambda_+ \cup \Lambda_- \quad \text{and} \quad \Lambda_+ \cap \Lambda_- = \emptyset.$$

For the *linear* case we have (1.3) but the condition (1.2) does not hold.

By introducing a sequence $\{\varepsilon_n\} \in \Lambda \cup \Lambda_0$ *asymptotically equivalent* to $\{\delta_n\}$, i.e.

$$\lim_{n \rightarrow +\infty} \frac{\varepsilon_n}{\delta_n} = 1, \quad (1.5)$$

we are able to write down the distributions in the *center* of our attention in the present paper.

Any $\{\delta_n^-\}$ from Λ_- , or from Λ_0 together with $\{\varepsilon_n\}$ (see, (1.5)) generates a *family* of distributions $\{p_n^-\}$ of the type

$$\begin{cases} p_0^- = (1 + (1-b) \sum_{n \geq 1} \frac{1}{\varepsilon_n} \prod_{m=1}^{n-1} (1 - \frac{b}{\delta_m^-}))^{-1}, & 0 < b < 1, \\ p_K^- = \frac{p_0^- \cdot (1-b)^{K-1}}{\varepsilon_K} \prod_{m=1}^{K-1} (1 - \frac{b}{\delta_m^-}), & K = 1, 2, \dots \end{cases} \quad (1.6)$$

where $\prod_{m=1}^0 = 1$. In this case, the probability p_0^- has a simple expression [6],

$$p_0^- = \frac{b}{1-b}D, \quad (1.7)$$

where

$$D = \sum_{k \geq 1} \frac{\varepsilon_k}{\delta_k^-} p_k^-, \quad \in R_+,$$

and, in particular, when $\{\varepsilon_n\} = \{\delta_n^-\}$, we have $D = 1$ and

$$p_0 = \frac{b}{1-b}. \quad (1.8)$$

Any $\{\delta_n^+\} \in \Lambda_+$ together with $\{\varepsilon_n\} \in \Lambda_+$ (see, (1.5)) generates a family of distributions of the type

$$\begin{cases} p_0^+ = (1 + (1+b) \sum_{n \geq 1} \frac{1}{\varepsilon_n} \prod_{m=1}^{n-1} (1 + \frac{b}{\delta_m^+}))^{-1}, & -1 < b < +\infty, \\ p_K^+ = \frac{p_0^- \cdot (1+b)^{K-1}}{\varepsilon_K} \prod_{m=1}^{K-1} (1 + \frac{b}{\delta_m^+}), & K = 1, 2, \dots \end{cases} \quad (1.9)$$

Distributions $\{p_n^+\}$ and $\{p_n^-\}$ vary regularly with some exponent $(-\rho)$, where $\rho \in [1, +\infty)$.

Remark 1. Let us consider the standard birth-death process with intensities λ_n and μ_{n+1} , $n = 1, 2, \dots$ (see, [1]).

The distributions $\{p_n^-\}$ and $\{p_n^+\}$ may be interpreted as stationary distributions of the standard birth-death process with intensities

$$\lambda_0 = 1 - b, \quad \lambda_n = \varepsilon_n \cdot (1 + \frac{b}{\delta_n^-}), \quad \mu_n = \varepsilon_n, \quad n = 1, 2, \dots,$$

or

$$\lambda_0 = 1 + b, \quad \lambda_n = \varepsilon_n \cdot (1 - \frac{b}{\delta_n^+}), \quad \mu_n = \varepsilon_n, \quad n = 1, 2, \dots,$$

respectively.

1.2 Particular Cases

The first particular case has been introduced and thoroughly investigated in [6]. The case is the following: for $n = 1, 2, \dots$

$$\varepsilon_n = \delta_n^- \text{ in (1.6) and } \varepsilon_n = \delta_n^+ \text{ in (1.9)}. \quad (1.10)$$

The second particular case is given by equalities: for $n = 1, 2, \dots$

$$\frac{1}{\varepsilon_n} = -\frac{1}{b} \ln\left(1 - \frac{b}{\delta_n^-}\right) \text{ in (1.6),} \tag{1.11}$$

and

$$\frac{1}{\varepsilon_n} = -\frac{1}{b} \ln\left(1 + \frac{b}{\delta_n^+}\right) \text{ in (1.9).} \tag{1.12}$$

The third particular case is given by equalities: for $n = 1, 2, \dots$

$$\begin{cases} p_0^- = \left(1 + (1-b) \sum_{n \geq 1} \frac{1}{\delta_n^-} \exp\left\{-b \cdot \sum_{m=1}^{n-1} \frac{1}{\delta_m^-}\right\}\right)^{-1} & 0 < b < 1. \\ p_k^- = \frac{p_0^- \cdot (1-b)}{\delta_k^-} \exp\left\{-b \cdot \sum_{m=1}^{k-1} \frac{1}{\delta_m^-}\right\}, & k = 1, 2, \dots, \end{cases} \tag{1.13}$$

$$\begin{cases} p_0^+ = \left(1 + (1+b) \sum_{n \geq 1} \frac{1}{\delta_n^+} \exp\left\{b \cdot \sum_{m=1}^{n-1} \frac{1}{\delta_m^+}\right\}\right)^{-1}, & -1 < b < +\infty. \\ p_k^+ = \frac{p_0^+ \cdot (1+b)}{\delta_k^+} \exp\left\{b \cdot \sum_{m=1}^{k-1} \frac{1}{\delta_m^+}\right\}, & k = 1, 2, \dots, \end{cases} \tag{1.14}$$

To see the reasons behing these forms note that in order to obtain (1.13) and (1.14), we represent

$$\prod_{m=1}^{n-1} \left(1 - \frac{b}{\delta_m^-}\right) = \exp\left\{\sum_{m=1}^{n-1} \ln\left(1 - \frac{b}{\delta_m^-}\right)\right\}$$

(1.15)

and

$$\prod_{m=1}^{n-1} \left(1 + \frac{b}{\delta_m^+}\right) = \exp\left\{\sum_{m=1}^{n-1} \ln\left(1 + \frac{b}{\delta_m^+}\right)\right\}$$

in (1.6) and (1.9), respectively.

The reasons for the case (1.11)-(1.12) are clear.

Next, using the expansion

$$\ln(1-x) = -\sum_{n \geq 1} \frac{x^n}{n}, \quad x \in R^1 = (-1, +1), \tag{1.16}$$

the third case is obtained as follows. We take the first terms at the right-hand-sides of equalities in (1.15) (see, (1.16)) and substitute them into the right-hand-side of (1.6) and (1.9).

By the L'Hopital rule

$$\lim_{x \rightarrow 0} \frac{\pm \ln(1 \pm x)}{x} = 1, \tag{1.17}$$

so, in the second case the condition (1.5) is fulfilled.

Similarly, the third particular case may be obtained from the second one if we replace the sequences

$$\left\{ -\frac{1}{b} \ln\left(1 - \frac{b}{\delta_n^-}\right) \right\} \text{ and } \left\{ -\frac{1}{b} \ln\left(1 + \frac{b}{\delta_n^+}\right) \right\}$$

by the asymptotically equivalent sequences

$$\frac{1}{\delta_n^-} \text{ and } \frac{1}{\delta_n^+},$$

respectively. Then, due to (1.17), the condition (1.5) holds.

In the next *Section* a new *dediscretization* approach is discussed and applied to the above classes of distributions.

2 The Dediscretization Approach

The method boils down to replacing the sums in (1.6) and (1.9) (we represent $\prod a_n$ by $\exp \sum_n \ln a_n$) by integrals and it will not change the *qualitative* behavior of distribution.

This operation simplifies many formulas and allows to derive *new* distributions with the *same* qualitative properties as before for biomolecular applications.

We call this approach a *Dediscretization Method*.

2.1 Dediscretization

Given some class of distributions $\{p_n\}$.

What does, in general, the dediscretization mean for $\{p_n\}$?

By our understanding elaborated below, this is some procedure on a class of distributions leading to a *concrete* construction of corresponding and in some sense "close" class of "smooth enough" (for instance, infinite differentiable) *distribution functions*.

The constructed class of distribution functions has to satisfy *definite* restrictions.

Necessarily, it must conserve the main *qualitative* properties of distributions of the original class such as: monotonicity; convexity; moments' existence; regular variation with the same exponent, etc.

What is the reason for dediscretization?

Sometimes the dediscretization leads to more simple expressions for distribution functions from the obtained class in comparison with the original class.

The next advantage presented *always* consists in following. Infinite differentiability, monotony, convexity, etc., allow to use *deeply* developed and well-understood *tools* of mathematical analysis.

Finally, it is more convenient to deal with continuous functions and suggest interpolations and approximations for them.

The further *discretization* (the reverse procedure) does not necessarily produce the original discrete distribution but a "close" one that may again be useful for applications.

2.2 The First Problem

In *Dediscretization* the following *problem* arises. We must replace the regularly varying sequences $\{\delta_n\}$ and $\{\varepsilon_n\}$ by regularly varying functions $\delta(t)$ and $\varepsilon(t)$, respectively. In other words, due to representations

$$\delta_n^\pm = 1 + n^\alpha L^\pm(n), \quad n = 0, 1, 2, \dots, \quad (2.1)$$

and

$$\varepsilon_n = 1 + n^\alpha L(n), \quad n = 0, 1, 2, \dots, \quad (2.2)$$

where

$$\lim_{n \rightarrow +\infty} (L^\pm(n)/L(n)) = 1,$$

$\alpha \in [1, +\infty)$, $\{L^\pm(n)\}$, and $\{L(n)\}$ are slowly varying sequences, we must replace the sequences $\{L^\pm(n)\}$ and $\{L(n)\}$ by slowly varying functions $\hat{L}^\pm(t)$ and $\hat{L}(t)$, respectively, in order to get, at least, continuous analogs of (2.1) and (2.2):

$$\delta^\pm(t) = 1 + t^\alpha \hat{L}^\pm(t) \quad \text{and} \quad \varepsilon(t) = 1 + t^\alpha \hat{L}(t), \quad t \in R^+. \quad (2.3)$$

In *typical* existing cases:
either

$$\lim_{n \rightarrow +\infty} L^+(n) = c \in R^+, \quad (2.4)$$

or

$$\lim_{n \rightarrow +\infty} L^\pm(n) = +\infty, \quad (2.5)$$

where

$$\{L^\pm(n)\} \text{ increases and is upward convex (see, [6]).} \quad (2.6)$$

The "breadth" of the class Λ is "so large" and the frequency distributions in biomolecular systems possess so many "smoothness" properties that the "restricting" Λ by additional assumptions (2.4)-(2.6) is quite natural and reasonable.

The *problem* is the same for $\{\delta_n^\pm\}$ and $\{\varepsilon_n\}$, so, let us consider it for $\{\varepsilon_n\}$.

The function $\hat{L}(t)$ has to be built in a *constructive* way and to satisfy following restrictions (if possible):

1. $\hat{L}(n) = L(n)$, $n = 1, 2, \dots$;
2. $\lim_{n \rightarrow +\infty} (\hat{L}(n)/L(n)) = 1$;
3. $\hat{L}(t)$ is infinite differentiable;
4. $\hat{L}(t)$ increases (decreases) if $\{L(n)\}$ increases (decreases);
5. $\hat{L}(t)$ is convex if $\{L(n)\}$ is convex

etc.

It is enough to solve the *problem* for the case $\lim_{n \rightarrow +\infty} L(n) = +\infty$. Indeed, if $\lim_{n \rightarrow +\infty} L(n) = c \in R^+$, then we take a sequence $\{L(n) \cdot \ln(n+1)\}$, which is slowly varying and satisfies condition $\lim_{n \rightarrow +\infty} L(n) \ln(n+1) = +\infty$. It reduces this case to the previous one.

If $\lim_{n \rightarrow +\infty} L(n) = 0$, then we take the sequence $\{1/L(n)\}$.

2.3 The solution of the problem

The function \hat{L} may be constructed in various ways. If the form of $\{L(n)\}$ is known and given by an elementary formula, replacing the discrete argument n by the continuous argument $t \in R^+$ we obtain a function $\hat{L}(t)$ defined on R^+ , which very often has *all* properties we need.

Let us give *examples*. Denote

$$e_{(0)} = 1, \quad e_{(1)} = e = \lim_{n \rightarrow +\infty} \left(1 + \frac{1}{n}\right)^n, \quad e_{(K+1)} = \exp(e_{(K)}), \quad K = 1, 2, \dots$$

For any given $K = 1, 2, \dots$ introduce the sequences $\{L_K^{(1)}(n)\}$ and $\{L_K^{(2)}(n)\}$ defined by following elementary formulas and satisfying restrictions (2.5)-(2.6):

$$L_K^{(1)}(n) = \prod_{j=1}^K \underbrace{\ln \ln \dots \ln n}_j$$

and

$$L_K^{(2)}(n) = L_K^{(1)}(n) \cdot \underbrace{\ln \ln \cdots \ln n}_K \text{ for integers } n > e_{(K)}.$$

These sequences when suitably defined for $n = 0, 1, \dots, [e_{(K)}]$ vary slowly. It can be seen by induction on K using the following two obvious facts:

1. If $\{L_1\}$ and $\{L_2\}$ vary slowly, then $\{L_1 \cdot L_2\}$ vary slowly;
 2. If $\{L(n)\}$ vary slowly and $\lim_{n \rightarrow +\infty} L(n) = +\infty$, then $\{\ln L(n)\}$ vary slowly.
- Indeed,

$$\lim_{n \rightarrow +\infty} \frac{\ln L(sn)}{\ln L(n)} = \lim_{n \rightarrow +\infty} \left\{ 1 + \frac{1}{\ln L(n)} \ln \frac{L(sn)}{L(n)} \right\} = 1 \text{ for } s = 2, 3, \dots.$$

Then, for $K = 1, 2, \dots$; $n > e_{(K)}$; $i = 1, 2$ the positive numbers $\delta_{n,K}^{(i)} = 1 + nL_K^{(i)}(n)$ (compare to (2.1)) form for any K sequences $\{\delta_{n,K}^{(i)}\}$ of type (1.3) if $i = 1$ and of type (1.4) if $i = 2$.

Therefore we get the form (2.3) by putting

$$\hat{L}_K^{(1)}(t) = \prod_{j=1}^K \underbrace{\ln \ln \cdots \ln t}_j$$

and

$$\hat{L}_K^{(2)}(t) = \hat{L}_K^{(1)}(t) \underbrace{\ln \ln \cdots \ln t}_K \text{ for } K = 1, 2, \dots$$

and $t \in [e_{(K)}, +\infty)$.

These functions when suitably defined for $t \in (0, e_{(K)})$ vary slowly and increase. Moreover, they are *upward convex*. Indeed, the multiplication of increasing, upward convex functions leads to increasing, upward convex function.

It is easily verify that the restrictions 1-5. for these functions are fulfilled.

In general, we have the following *interpolation* theorem

Theorem 1. *There is a slowly varying function $\hat{L}(t)$ satisfying restrictions 1-5.*

Proof. If the form of $\{L(n)\}$ is not given by elementary formula which itself leads to desired interpolation, then we proceed as follows. For the sequence $\{L(n)\}$ draw a piecewise linear curve passing through points $(0, L(0)), (1, L(1)), (2, L(2)), \dots$ on the plane. This "broken" line, say $L_0(t)$, satisfies conditions $L_0(n) = L(n), n = 0, 1, 2, \dots$.

We say that $\underline{L_0(t)}$ is a linear continuous analog of the sequence $\{L(n)\}$.

For any $t \in R^+$ and $n = 1, 2, \dots$, due to monotonicity of L_0 we have

$$\begin{aligned} \frac{L(nK)}{L(K+1)} &= \frac{L_0(n \cdot [t])}{L_0([t]+1)} \leq \frac{L_0(nt)}{L_0(t)} \leq \frac{L_0(n[t]+1)}{L_0([t])} \\ &= \frac{L(nK+n)}{L(K)} = \frac{L(nK+n)}{L(K)} = \frac{L(nK')}{L(K'+1)}, \quad K = K' - 1 = [t], \end{aligned}$$

where $[t]$ denotes the entire part of the positive number t .

Taking into account that for $n = 2, 3, \dots$

$$\lim_{K \rightarrow +\infty} \frac{L(Kn)}{L(K)} = 1 \text{ and, as a result of this, } \lim_{K \rightarrow +\infty} \frac{L(K+1)}{L(K)} = 1,$$

we conclude that

$$\lim_{t \rightarrow +\infty} \frac{L_0(xt)}{L_0(t)} = 1 \tag{2.7}$$

for $x = n$ with $n = 1, 2, \dots$. Putting $t = (t'/m)$ with $m = 1, 2, \dots$ we get (2.7) for $x = m^{-1}$ with $m = 1, 2, \dots$. Combining these two cases we conclude that (2.7) holds for $x = \frac{m}{n}$ with $m = 1, 2, \dots$ and $n = 1, 2, \dots$, i.e. for *all* positive *rational* numbers x . The set of such numbers is *everywhere dense* in R^+ . Therefore, by *Lemma 1*, p.275 [7], we see that the *continuous analog* $L_0(t)$ of $\{L(n)\}$ *varies slowly*. By generalization of Adamovic Interpolation Theorem on slowly varying function with condition $\lim_{t \rightarrow +\infty} L_0(t) = +\infty$, there is a *constructive* method of building of an increasing, convex, infinite differentiable, slowly varying function $\hat{L}(t)$ satisfying restrictions:

$$\hat{L}(n) = L_0(n), \quad n = 1, 2, \dots \text{ and } \lim_{t \rightarrow +\infty} (\hat{L}(t)/L_0(t)) = 1 \text{ (see, [8])}$$

For this function, as we see, conditions 1-5 are fulfilled.

2.4 The procedure

Denote

$$f_b^\pm(x) = \ln\left(1 \pm \frac{b}{\delta^\pm(x)}\right), \quad x \in R^+,$$

where $\delta^\pm(x)$ is the above described *interpolation* of a given sequence $\{\delta_n^\pm\} \in \Lambda_\pm$, which together with *interpolation* $\varepsilon(x)$ of a sequence $\{\varepsilon_n\}$ generate one-parametric *family* of distributions $\{p_n^\pm\} = \{p_n^\pm(b)\}$ of types (1.6) and (1.9) with $0 < b < 1$ and condition (1.3) for sign "-" and with $-1 < b < +\infty$ and condition (1.4) for sign "+".

Definition 1. We say that the function

$$\hat{F}_{\pm}(x, b) \stackrel{\text{def.}}{=} \frac{\int_{0-}^x \frac{1}{\varepsilon(t)} \exp \left\{ \int_{0-}^t f_b^{\pm}(u) du \right\} dt}{\int_{0-}^{+\infty} \frac{1}{\varepsilon(t)} \exp \left\{ \int_{0-}^t f_{\varepsilon}^{\pm}(u) du \right\} dt} \quad (2.8)$$

defined on $[0, +\infty)$ is a dediscretization of distribution $\{p_n^{\pm}(b)\}$ generated by $(\{\varepsilon_n\}, \{\delta_n^{\pm}\})$.

The reasons behind the forms $\hat{F}_+(x, b)$ and $\hat{F}_-(x, b)$ are elaborated below.

For a given sequence $\{\delta_n^- \in \Lambda_-\}$ denote $f_n^- = f_n^-(b) = \ln(1 - \frac{b}{\delta_n^-})$, $n = 0, 1, 2, \dots$, $0 < b < 1$. The distribution function $F_-(x) = F_-(x, b)$, $x \in [0, +\infty)$, which corresponds to distribution $\{p_n^-(b)\}$ generated by $(\{\varepsilon_n\}, \{\delta_n^-\})$, takes the form

$$F_-(x) = F_-(+0) \cdot \sum_{n=0}^{[x]} \exp \left\{ \sum_{m=0}^{n-1} f_m^- \right\}, \quad (2.9)$$

where $0 < b < 1$ and $\sum_{m=0}^{-1} = 0$, $\varepsilon_0 = 1$. Indeed, putting $\prod_{m=1}^0 = 1$, $\sum_{m=1}^0 = 0$, due to (1.6), for $x \in R$ and $0 < b < 1$ we have

$$\begin{aligned} F_-(x) &= p_0^- \cdot \left\{ 1 + (1-b) \sum_{n=1}^{[x]} \frac{1}{\varepsilon_n} \sum_{m=1}^{n-1} \left(1 - \frac{b}{\delta_m^-}\right) \right\} \\ &= F_-(0) \left\{ 1 + (1-b) \sum_{n=1}^{[x]} \frac{1}{\varepsilon_n} \exp \left(\sum_{m=1}^{n-1} f_m^- \right) \right\} \\ &= F_-(+0) \left\{ \frac{1}{\varepsilon_0} \exp \left(\sum_{m=0}^{-1} f_m^- \right) + \sum_{n=1}^{[x]} \frac{1}{\varepsilon_n} \exp(f_0^-) \exp \left(\sum_{m=1}^{n-1} f_m^- \right) \right\} \\ &= F_-(+0) \sum_{n=0}^{[x]} \frac{1}{\varepsilon_n} \exp \left(\sum_{m=0}^{n-1} f_m^- \right). \end{aligned}$$

By (2.9), for $x \in R^+$ a function $g_b^-(x) \stackrel{\text{def.}}{=} (F_-(x)/F_-(+0))$, $0 < b < 1$, $g_b^-(-0) = 0$, represents a finite, positive, discrete measure, and takes the form

$$g_b^-(x) = F_-(+0) \sum_{n=0}^{[x]} \frac{1}{\varepsilon_n} \exp \left\{ \sum_{m=0}^{n-1} f_m^- \right\}. \quad (2.10)$$

Similarly, the *distribution function* $F_+(x) = F_+(x, b), x \in [0, +\infty)$, which corresponds to distribution $\{p_n^+(b)\}$ generated by $(\{\varepsilon_n\}, \{\delta_n^+\})$, takes the form

$$F_+(x) = F_+(+0) \sum_{n=0}^{[x]} \frac{1}{\varepsilon_n} \exp \left\{ \sum_{m=0}^{n-1} f_m^+ \right\} \quad (2.11)$$

(F_+ and F_- have jumps p_0^+ and p_0^- at zero), where $f_n^+ = f_n^+(b) = \ln(1 + \frac{b}{\delta_n^+})$, $n = 0, 1, 2, \dots, -1 < b < +\infty$, and

$$g_b^+(x) = (F_+(x)/F_+(+0)) = \sum_{n=0}^{[x]} \frac{1}{\varepsilon_n} \exp \left\{ \sum_{m=0}^{n-1} f_m^+ \right\}, \quad g_b^+(-0) = 0, \quad (2.12)$$

is a *finite, positive, discrete* measure.

The form (2.10) and (2.12) of g_b^- and g_b^+ is *suitable* for *dediscretization* procedure. Namely, measures

$$\hat{g}^\pm(x) = \int_{0-}^x \frac{1}{\varepsilon(t)} \exp \left\{ \int_{0-}^t f_b^\pm(u) du \right\} dt$$

with $0 < b < 1$ for sign ”-” and $-1 < b < +\infty$ for sign ”+”, *absolutely continuous* on R^+ , after *normalization* give \hat{F}_+ and \hat{F}_- .

2.5 Justification

The justification of *dediscretization* procedure is based on

Theorem 2. $\hat{F}_+(x, b)$ and $\hat{F}_-(x, b)$ are *distribution functions*.

Proof. Due to (2.8), we have to prove

$$0 < g_b^\pm(+\infty) < +\infty, \quad (2.13)$$

where the first inequality is obvious.

The cases: (a) with sign ”+” and (b) with sign ”-” are considered separately.

By (1.5) and limit equality $\lim_{n \rightarrow +\infty} (\delta_{n+1}^\pm / \delta_n^\pm) = 1$ for $\varepsilon \in (0, 1)$ starting from some index n_0 simultaneously

$$\frac{1 - \varepsilon}{\delta_n^\pm} < \frac{1}{\varepsilon_n} < \frac{1 + \varepsilon}{\delta_n^\pm} \quad \text{and} \quad \frac{1 - \varepsilon}{\delta_{n+1}^\pm} < \frac{1}{\delta_n^\pm} < \frac{1 + \varepsilon}{\delta_{n+1}^\pm} \quad \text{for } n = n_0, n_0 + 1, \dots \quad (2.14)$$

(a) Here there is a simple subcase $b = 0$: $\hat{F}_+(x, 0) = (\int_0^x \frac{dt}{\varepsilon(t)}) / (\int_0^{+\infty} \frac{dt}{\varepsilon(t)})$, $x \in [0, +\infty)$.

By (1.4) and (2.14),

$$0 < \int_{n_0}^{+\infty} \frac{dt}{\varepsilon(t)} = \sum_{n \geq n_0} \int_n^{n+1} \frac{dt}{\varepsilon(t)} < \sum_{n \geq n_0} \frac{1}{\varepsilon_n} < (1 + \varepsilon) \cdot \sum_{n \geq n_0} \frac{1}{\delta_n^+} < +\infty.$$

Thus, $g_0^+(+\infty) < +\infty$.

If $-1 < b < 0$, then $g_b^+(+\infty) < g_0^+(+\infty) < +\infty$.

Let $0 < b < +\infty$. In accordance with the inequalities (2.14) and

$$\frac{b}{\delta^+(t)} \leq \ln\left(1 + \frac{b}{\delta^+(t)}\right) = f_b^+(t), \quad t \in [0, +\infty),$$

we have

$$\begin{aligned} \hat{g}_b^+(+\infty) - \hat{g}_b^+(n_0) &= \sum_{n \geq n_0} \int_n^{n+1} \frac{1}{\varepsilon^+(t)} \exp\left\{\int_{0-}^t f_b^+(u) du\right\} dt \\ &< \sum_{n \geq n_0} \frac{1 + \varepsilon}{\delta_n^+} \int_n^{n+1} \exp\left\{\int_{0-}^t f_b^+(u) du\right\} dt \\ &< (1 + \varepsilon)^2 \cdot \sum_{n \geq n_0} \frac{1}{\delta_{n+1}^+} \int_n^{n+1} \exp\left\{\int_{0-}^t f_b^+(u) du\right\} dt \\ &< (1 + \varepsilon)^2 \cdot \sum_{n \geq n_0} \int_n^{n+1} \frac{1}{\delta^+(t)} \exp\left\{\int_{0-}^t f_b^+(u) du\right\} dt \\ &< \frac{(1 + \varepsilon)^2}{b} \int_n^{+\infty} f_b^+(t) \exp\left\{\int_{0-}^t f_b^+(u) du\right\} dt \\ &= \frac{(1 + \varepsilon)^2}{b} (\exp(\int_{0-}^{+\infty} f_b^+(u) du) - \exp(\int_{0-}^{n_0} f_b^+(u) du)). \end{aligned} \tag{2.15}$$

By L'Hopital rule

$$\lim_{x \rightarrow +\infty} \frac{\int_x^{+\infty} f_b^+(u) du}{a \cdot \int_x^{+\infty} \frac{du}{\delta^+(u)}} = \lim_{x \rightarrow +\infty} \frac{f_b^+(x)}{\frac{a}{\delta^+(x)}} = 1.$$

But, by (1.4),

$$0 \leq \int_{0-}^{+\infty} \frac{du}{\delta^+(u)} = \sum_{n \geq 0} \int_n^{n+1} \frac{du}{\delta^+(u)} < 1 + \sum_{n \geq 1} \frac{1}{\delta_n^+} < +\infty.$$

Therefore,

$$\int_{0-}^{+\infty} f_b^+(u) du < +\infty, \tag{2.16}$$

and, due to (2.15), (2.13) holds in this case.

(b) In accordance with the inequalities (2.14) and

$$f_b^-(t) \leq -\frac{b}{\delta^-(t)} + \frac{1}{2}\left(\frac{b}{\delta^-(t)}\right)^2, \quad 0 < b < 1, \quad t \in [0, +\infty),$$

we have similarly to (2.15)

$$\begin{aligned} \hat{g}_b^-(+\infty) - \hat{g}_b^-(n_0) &< (1 + \varepsilon)^2 \cdot \sum_{n \geq n_0} \int_n^{n+1} \frac{1}{\delta^-(t)} \exp \left\{ \int_{0^-}^t f_b^-(u) du \right\} dt \\ &< (1 + \varepsilon)^2 \cdot c(b) \cdot \int_{n_0}^{+\infty} \frac{1}{\delta^-(t)} \exp \left(-b \cdot \int_{0^-}^t \frac{du}{\delta^-(u)} \right) dt \quad (2.17) \\ &= \frac{(1 + \varepsilon)^2 \cdot c(b)}{b} \left(-\exp \left(-\int_{0^-}^{+\infty} \frac{du}{\delta^-(u)} \right) + \exp \left(-\int_{0^-}^{n_0} \frac{du}{\delta^-(u)} \right) \right), \end{aligned}$$

where

$$c(b) = \exp \left(\frac{b^2}{2} \int_{0^-}^{+\infty} \frac{du}{(\delta^-(u))^2} \right). \quad (2.18)$$

Taking into account that for any $\{\delta_n^-\} \in \Lambda_-$, due to $\delta_0^- = 1$ and $\{\delta_n^-\}$ is downward convex with $\lim_{n \rightarrow +\infty} (n/\delta_n^-) = 0$, we have

$$\frac{1}{\delta_n^-} < \frac{1}{n}, \quad n = 1, 2, \dots \quad (2.19)$$

Now, with the help of (1.3) and (2.19) we obtain

$$\begin{aligned} \int_{0^-}^{+\infty} \frac{du}{\delta^-(u)} &= \sum_{n \geq 0} \int_n^{n+1} \frac{du}{\delta^-(u)} > \sum_{n \geq 1} \frac{1}{\delta_n^-} = +\infty, \\ 0 \leq \int_{0^-}^{+\infty} \frac{du}{(\delta^-(u))^2} &= \sum_{n \geq 0} \int_n^{n+1} \frac{du}{(\delta^-(u))^2} < \sum_{n \geq 1} \frac{1}{(\delta_n^-)^2} < \sum_{n \geq 1} \frac{1}{n^2} < +\infty. \end{aligned} \quad (2.20)$$

Therefore, from (2.17)-(2.18) we conclude that (2.13) holds in this case too. *Theorem 2* is proved.

Now, we may say that $\hat{F}_\pm(x, b)$ is strictly increasing. In *Section 4* the properties of distribution functions \hat{F}_+ and \hat{F}_- are investigated. In the next *Section* some typical examples are considered.

3 Typical Examples

The structure of stochastic birth-death process is comparatively complicated from the point of view of computations stationary distributions. This structure automatically transforms also on *dediscretization*. Thus, if we change nothing besides in some sense replace the discrete argument by a continuous one, then we can not expect to get any *essential* simplification in the form *in general*. However, there are possibilities to apply the *tools* of mathematical analysis (not only for proofs of statements, but also in order to get simplifications in particular cases).

In the present *Section* we consider some *typical* examples of distributions $\{p_n\}$ of forms (1.6) and (1.9), respectively.

3.1 On distributions $\{p_n\}$

It is possible to write the distributions $\{p_n\}$ in one unified symmetric form, which depends on two parameters p and q satisfying condition

$$\begin{cases} 0 < p < q < +\infty & \text{if } \sum_{n \geq 1} \frac{1}{\zeta_n} = +\infty, \\ 0 < p < +\infty, 0 < q < +\infty & \text{if } \sum_{n \geq 1} \frac{1}{\zeta_n} < +\infty. \end{cases} \quad (3.1)$$

The form uses sequences $\{\zeta_n\}$ and $\{\lambda_n\}$ instead of $\{\delta_n\}$ and $\{\varepsilon_n\}$, respectively, where

$$\frac{1}{q} \zeta_n = \delta_n - 1, \quad \frac{1}{q} \lambda_n = \varepsilon_n - 1, \quad n = 0, 1, 2, \dots \quad (3.2)$$

Distributions $\{p_n\}$ take the form

$$\begin{cases} p_0 = \left(1 + p \cdot \sum_{n \geq 1} \frac{1}{q + \lambda_n} \prod_{m=1}^{n-1} \frac{p + \zeta_m}{q + \zeta_m} \right)^{-1}, \\ p_K = \frac{p \cdot p_0}{q + \lambda_K} \cdot \prod_{m=1}^{K-1} \frac{p + \zeta_m}{q + \zeta_m}, K = 1, 2, \dots \end{cases} \quad (3.3)$$

The corresponding distribution function is

$$F_{p,q}(x) = F_{p,q}(+0) \cdot \sum_{n=0}^{[x]} \frac{1}{q + \lambda_n} \exp \left\{ \sum_{m=0}^{n-1} \ln \frac{p + \zeta_m}{q + \zeta_m} \right\}, \quad x \in [0, +\infty);$$

the *dediscretization* of distribution function $F_{p,q}(x)$ is

$$\hat{F}_{p,q}(x) = c(p, q) \int_{0-}^x \frac{1}{q + \lambda(t)} \exp \left\{ \int_{0-}^t \ln \left(\frac{p + \zeta(u)}{q + \zeta(u)} \right) du \right\} dt \quad (3.4)$$

with *normalization* factor

$$c(p, q) = \left(\int_{0-}^{+\infty} \frac{1}{q + \lambda(t)} \exp \left\{ \int_{0-}^t \ln \left(\frac{p + \zeta(u)}{q + \zeta(u)} \right) du \right\} dt \right)^{-1}.$$

Here

$$\frac{1}{q} \zeta(t) = \delta(t) - 1, \quad \frac{1}{q} \lambda(t) = \varepsilon(t) - 1, \quad t \in [0, +\infty),$$

where $\delta(t)$ and $\varepsilon(t)$ are already obtained interpolations of $\{\delta_n\}$ and $\{\varepsilon_n\}$, respectively.

The form (3.3) with restrictions (3.1) of distributions $\{p_n\}$ is chosen because exactly this form for *etalon linear case* $\lambda_n = \zeta_n = n$ implies the *traditional* form of a *family of Waring Distributions* (WD)

$$\begin{cases} p_0 = \left(1 + p \cdot \sum_{n \geq 1} \frac{1}{q+n} \prod_{m=1}^{n-1} \frac{p+m}{q+m} \right)^{-1}, & 0 < p < q < +\infty, \\ p_K = \frac{p \cdot p_0}{q+K} \prod_{m=1}^K \frac{p+m}{q+m}, & K = 1, 2, \dots \end{cases} \quad (3.5)$$

3.2 Typical examples

Denote

$$f(t) = \ln \left(\frac{p + \zeta(t)}{q + \zeta(t)} \right), \quad t \in [0, +\infty).$$

Then, for $t \in [0, +\infty)$ we have

$$\begin{aligned} \int_{0-}^t f(x) dx &= t \cdot f(t) - \int_{0-}^t x df(x) \\ &= t \cdot \ln \left(\frac{p + \zeta(t)}{q + \zeta(t)} \right) - \int_{0-}^t x d \{ \ln(p + \zeta(t)) - \ln(q + \zeta(t)) \} \\ &= t \cdot \ln(p + \zeta(t)) - t \ln(q + \zeta(t)) - \int_{0-}^t \frac{x \cdot \zeta'(x) dx}{p + \zeta(x)} + \int_{0-}^t \frac{x \cdot \zeta'(x) dx}{q + \zeta(x)}. \end{aligned} \quad (3.6)$$

If there is a simple relationship between $\zeta(x)$ and $\zeta'(x)$, then integrals in (3.6) can be evaluated, and in (3.4) we may get rid of the inner integral. The two typical examples considered below are based on this *idea*.

1. Let us consider *linear* $\{\zeta_n\}$: $\zeta_n = n, n = 0, 1, 2, \dots$. For this case we get two-parametric *family* of WD (3.5). By (3.6), taking $\zeta(t) = t, t \in [0, +\infty)$, we obtain

$$\begin{aligned} \int_{0-}^t f(x) dx &= t \cdot \ln(p+t) - t \ln(q+t) - \int_{0-}^t \frac{x dx}{p+x} - \int_{0-}^t \frac{x dx}{q+x} \\ &= (t+p) \ln(t+p) - (q+t) \ln(q+t) - p \ln p + q \ln q \\ &= \ln\left(\frac{(p+t)^{p+t}}{(q+t)^{q+t}} \cdot \frac{q^q}{p^p}\right). \end{aligned} \quad (3.7)$$

Substituting the last expression into (3.4) we have

$$\begin{aligned} \hat{F}_{p,q}(x) &= c(p,q) \cdot \int_{0-}^x \frac{1}{q+\lambda(t)} \cdot \exp\left\{\ln\left(\frac{(p+t)^{p+t}}{(q+t)^{q+t}} \cdot \frac{q^q}{p^p}\right)\right\} dt \\ &= \hat{c}(p,q) \cdot \int_{0-}^x \frac{(p+t)^{p+t}}{(q+t)^{q+t}} \cdot \frac{1}{q+\lambda(t)} dt, \quad x \in [0, +\infty), \end{aligned} \quad (3.8)$$

where the *normalization* factor takes the form

$$\hat{c}(p,q) = \left(\int_{0-}^{+\infty} \frac{(p+t)^{p+t}}{(q+t)^{q+t}} \cdot \frac{1}{q+\lambda(t)} dt\right)^{-1}.$$

Note that the limit exists (the asymptotical equivalency of $\lambda(t)$ and t)

$$\lim_{t \rightarrow +\infty} \frac{\lambda(t)}{t} = 1. \quad (3.9)$$

Let us present two *corollaries* of the obtained result (3.8).

First of all, from (3.8) we are able to figure out the tail's asymptotic

$$\begin{aligned} 1 - \hat{F}_{p,q}(x) &= \hat{c}(p,q) \cdot \int_x^{\infty} \frac{(p+t)^{p+t}}{(q+t)^{q+t}} \frac{dt}{q+\lambda(t)} \\ &= \hat{c}(p,q) \cdot \int_x^{\infty} \frac{(p+t)^p}{(q+t)^q} \frac{(1+(p/t))^t}{(1+(q/t))^t} \frac{dt}{q+\lambda(t)} \\ &\approx \hat{c}(p,q) e^{-(q-p)} \cdot \int_x^{\infty} \frac{(p+t)^p}{(q+t)^q} \frac{dt}{q+\lambda(t)} \\ &\approx \hat{c}(p,q) e^{-(q-p)} \cdot \int_x^{\infty} t^{-(q+1-p)} dt \\ &= \hat{c}(p,q) e^{-(q-p)} \frac{1}{q-p} x^{-(q-p)}, \quad x \rightarrow +\infty. \end{aligned}$$

It means that the tail $1 - \hat{F}_{p,q}(x)$, $x \in [0, +\infty)$ ($\hat{F}_{p,q}(-x) = 0$) of distribution function $\hat{F}_{p,q}$ varies regularly with exponent $-(q-p)$ and in representation

$$1 - \hat{F}_{p,q}(x) = x^{-(q-p)} \cdot L(x), \quad x \in R^+,$$

of the tail the *slowly varying* function $L(x)$ satisfies condition

$$\lim_{x \rightarrow +\infty} L(x) = \hat{c}(p, q) e^{-q-p} \cdot \frac{1}{q-p}.$$

Secondly, the evaluations carry to the end for *asymptotically linear* $\lambda(t)$ (see, (3.9)) of a *special* form

$$\lambda(t) = \frac{p-q}{\ln\left(\frac{p+t}{q+t}\right)} - q, \quad t \in [0, +\infty). \quad (3.10)$$

Let us show that the form (3.10) indeed presents an asymptotically linear $\lambda(t)$. Due to L'Hopital rule,

$$\lim_{t \rightarrow +\infty} \frac{\ln\left(1 - \frac{q-p}{q+t}\right)}{-\frac{q-p}{q+t}} = 1.$$

Therefore, from (3.10) we have

$$\lambda(t) = \frac{p-q}{\ln\left(1 - \frac{q-p}{q+t}\right)} - q \approx \frac{p-q}{-\frac{q-p}{q+t}} - q = t, \quad t \rightarrow +\infty.$$

Thus, (3.10) represents asymptotically linear (etalon) $\lambda(t)$.

Substituting (3.10) into (3.4) with $\zeta(t) = t, t \in [0, +\infty)$, we obtain

$$\begin{aligned} \hat{F}_{p,q}(x) &= \frac{c(p,q)}{q-p} \int_{0-}^x \left| \ln\left(\frac{p+t}{q+t}\right) \right| \exp \left\{ - \int_{0-}^t \left| \ln\left(\frac{p+u}{q+u}\right) \right| du \right\} dt = \\ &= \frac{c(p,q)}{q-p} \left(1 - \exp \left\{ - \int_{0-}^x \left| \ln\left(\frac{p+u}{q+u}\right) \right| du \right\} \right), \end{aligned} \quad (3.11)$$

where we took into account that

$$\ln\left(\frac{p+u}{q+u}\right) = - \left| \ln\left(\frac{p+u}{q+u}\right) \right| \text{ for } 0 < p < q < +\infty \text{ and } u \in R^+.$$

Evaluations may be carried out with the help of (3.7). By (3.7),

$$\int_{0-}^{+\infty} \left| \ln\left(\frac{p+u}{q+u}\right) \right| du = +\infty, \text{ so, in (3.11) we have } \frac{c(p,q)}{q-p} = 1.$$

It means that (3.11) may be rewritten in the form (see, (3.7))

$$\hat{F}_{p,q}(x) = 1 - \exp \left\{ \ln \left(\frac{(p+t)^{p+t} q^q}{(q+t)^{q+t} p^p} \right) \right\} = 1 - \frac{q^q}{p^p} \cdot \frac{(p+x)^{p+x}}{(q+x)^{q+x}}. \quad (3.12)$$

The expression at the right-hand-side of (3.12) gives for the tail's asymptotic the final result. In this case the slowly varying function in representation of $1 - \hat{F}_{p,q}(x)$ as $x \rightarrow +\infty$ tends to the following constant

$$\lim_{x \rightarrow +\infty} \frac{q^q}{p^p} x^{(q-p)} \cdot \frac{(p+x)^{p+x}}{(q+x)^{q+x}} = \frac{q^q}{p^p} \lim_{x \rightarrow +\infty} \frac{(1 + (p/x))^{x+p}}{(1 + (q/x))^{x+q}} = \frac{q^q}{p^p} \cdot e^{-(q-p)}.$$

The applied *idea*: to choose an asymptotically equivalent to $\zeta(t)$ function $\lambda(t)$ which allows to evaluate arising integrals is very *fruitful* in order to find the most simple stationary distributions generated by the birth-death process with conserving the *qualitative* properties of distributions.

2. For *powerform* of $\{\zeta_n\}$, i.e. $\zeta_n = n^\alpha, n = 0, 1, 2, \dots, 1 < \alpha < +\infty$, the first *idea* does not work. By (3.6), taking $\zeta(t) = t^\alpha, t \in [0, +\infty)$, we proceed

$$\begin{aligned} 0 < \int_{0-}^t f(x) dx &= t \cdot \ln(p+t^\alpha) - t \cdot \ln(q+t^\alpha) - \alpha \int_{0-}^t \frac{x^\alpha dx}{p+x^\alpha} + \alpha \int_{0-}^t \frac{x^\alpha dx}{q+x^\alpha} \\ &= \ln \left(\left(\frac{p+t^\alpha}{q+t^\alpha} \right)^t \right) + \alpha p \int_{0-}^t \frac{dx}{p+x^\alpha} - \alpha q \int_{0-}^t \frac{dx}{q+x^\alpha} < +\infty. \end{aligned} \quad (3.13)$$

The last inequality is clear because the integrals at the right-hand-side of (3.13) converge as $t \rightarrow +\infty$ and are finite for finite $t \in R^+$.

Substituting the last expression into (3.4) we find out a distribution function

$$\hat{F}_{p,q}(x) = \hat{c}(p,q) \int_{0-}^x \left(\frac{p+t^\alpha}{q+t^\alpha} \right)^t \frac{1}{q+\lambda(t)} \exp \left\{ \alpha p \int_{0-}^t \frac{dx}{p+x^\alpha} - \alpha q \int_{0-}^t \frac{dx}{q+x^\alpha} \right\} dt, \quad (3.14)$$

$$0 \leq x < +\infty,$$

with the corresponding *normalization* factor $\hat{c}(p,q)$.

Note that the limit exists (the asymptotical equivalency of $\lambda(t)$ and t^α)

$$\lim_{t \rightarrow +\infty} \frac{\lambda(t)}{t^\alpha} = 1. \quad (3.15)$$

Let us present two *corollaries* of formula (3.14).

From (3.14) we may figure out the tail's asymptotic. Denote

$$I_\alpha(z) = \alpha \cdot z \cdot \int_{0-}^{\infty} \frac{dx}{z+x^\alpha}, \quad 1 < \alpha < +\infty, \quad 0 < z < +\infty.$$

We have

$$\begin{aligned} 1 - \hat{F}_{p,q}(x) &= \hat{c}(p,q) \int_x^{+\infty} \left(\frac{p+t^\alpha}{q+t^\alpha}\right)^t \frac{1}{q+\lambda(t)} \exp \left\{ \alpha p \int_{0-}^t \frac{dx}{p+x^\alpha} - \alpha q \int_{0-}^t \frac{dx}{q+x^\alpha} \right\} dt \\ &\approx \hat{c}(p,q) e^{I_\alpha(p) - I_\alpha(q)} \int_x^{\infty} \frac{(1+(p/t^\alpha))^t}{(1+(q/t^\alpha))^t} \frac{1}{q+\lambda(t)} dt \\ &\approx \hat{c}(p,q) e^{I_\alpha(p) - I_\alpha(q)} \int_x^{\infty} \frac{dt}{q+\lambda(t)} \approx \hat{c}(p,q) e^{I_\alpha(p) - I_\alpha(q)} \int_x^{\infty} \frac{dt}{q+t^\alpha} \\ &= \hat{c}(p,q) e^{I_\alpha(p) - I_\alpha(q)} \frac{1}{\alpha-1} (x+q)^{-(\alpha-1)}, \\ &\approx \hat{c}(p,q) e^{I_\alpha(x) - I_\alpha(q)} \frac{1}{\alpha-1} \frac{1}{x^{\alpha-1}}, \quad x \rightarrow +\infty. \end{aligned}$$

It means that the tail $1 - \hat{F}_{p,q}(x), x \in R^+$ of distribution function $\hat{F}_{p,q}$ varies regularly with exponent $-(\alpha-1)$ and in representation

$$1 - \hat{F}_{p,q}(x) = x^{-(\alpha-1)} L_1(x), \quad x \in R^+$$

of the tail the slowly varying function $L_1(x)$ satisfies condition

$$\lim_{x \rightarrow +\infty} L_1(x) = \hat{c}(p,q) e^{I_\alpha(p) - I_\alpha(q)} \frac{1}{\alpha-1}.$$

The second *idea* being used for etalon linear case works in the present case.

Indeed, let

$$\lambda(t) = \frac{p-q}{\ln\left(\frac{p+t^\alpha}{q+t^\alpha}\right)} - q, \quad t \in [0, +\infty). \quad (3.16)$$

Due to L'Hopital rule,

$$\lim_{t \rightarrow +\infty} \frac{\ln\left(1 - \frac{p-q}{q+t^\alpha}\right)}{-\frac{q-p}{q+t^\alpha}} = 1.$$

Therefore, similarly to the etalon linear case $\lambda(t) \approx t^\alpha, t \rightarrow +\infty$, so, (3.15) holds.

Substituting (3.16) into (3.4) with $\zeta(t) = t^\alpha, 1 < \alpha < +\infty, t \in [0, +\infty)$, we obtain

$$\begin{aligned}\hat{F}_{p,q}(x) &= \frac{c(p,q)}{p-q} \int_{0-}^x \ln\left(\frac{p+t^\alpha}{q+t^\alpha}\right) \exp\left\{\int_{0-}^x \ln\left(\frac{p+u^\alpha}{q+u^\alpha}\right) du\right\} dt \\ &= \frac{c(p,q)}{p-q} (\exp\left\{\int_{0-}^x \ln\left(\frac{p+u^\alpha}{q+u^\alpha}\right) du\right\} - 1), \quad x \in [0, +\infty).\end{aligned}\quad (3.17)$$

Note that $\frac{c(p,q)}{p-q}$ is *positive* if $0 < q < p < +\infty$ and is *negative* if $0 < p < q < +\infty$. By (3.13), the integral

$$\begin{aligned}\int_{0-}^x \ln\left(\frac{p+t^\alpha}{q+t^\alpha}\right) dt &= x \cdot \left\{ \ln\left(1 + \frac{p}{x^\alpha}\right) - \ln\left(1 + \frac{q}{x^\alpha}\right) \right\} \\ &\quad + \alpha p \cdot \int_{0-}^x \frac{dt}{p+t^\alpha} - \alpha q \cdot \int_{0-}^x \frac{dt}{q+t^\alpha}\end{aligned}\quad (3.18)$$

is *positive* if $0 < q < p < +\infty$ and is *negative* if $0 < p < q < +\infty$. Since

$$\lim_{x \rightarrow +\infty} x \cdot \left\{ \ln\left(1 + \frac{p}{x^\alpha}\right) - \ln\left(1 + \frac{q}{x^\alpha}\right) \right\} = 0,$$

therefore,

$$\int_{0-}^{\infty} \ln\left(\frac{p+t^\alpha}{q+t^\alpha}\right) dt = I_\alpha(p) - I_\alpha(q).$$

It means that

$$\frac{c(p,q)}{p-q} (\exp\left\{\int_{0-}^{+\infty} \ln\left(\frac{p+u^\alpha}{q+u^\alpha}\right) du\right\} - 1) = \hat{F}_{p,q}(+\infty) = 1,$$

and, as a result of this,

$$\frac{c(p,q)}{p-q} = (\exp\{I_\alpha(p) - I_\alpha(q)\} - 1)^{-1}.$$

Finally, note that in case $p = q$, even in general situation, from (3.4) we come to the following formula

$$\hat{F}_{p,q}(x) = \frac{\int_{0-}^x \frac{dt}{q + \lambda(t)}}{\int_{0-}^{\infty} \frac{dt}{q + \lambda(t)}}, \quad x \in [0, +\infty).$$

4 On the Class of Distribution Functions

In *Section 2* with the help of *Dediscretization Approach* we constructed a wide class of infinite differentiable on R^+ distribution functions defined on $[0, +\infty)$ with jumps at zero. Considering this class as *a priori given* (we forget how it was obtained) with *definite* restrictions, its *discretization* (procedure reverse to dediscretization procedure and more simple) includes *all* distributions generated by birth-death process with regularly varying intensities. By the way, now, we do not need *Theorem 1*.

In the present *Section* this class is defined, discussed and investigated.

4.1 The description of the class

The description of the class. Denote by Ω the class of regularly varying with exponent $\alpha \in [1, +\infty)$, increasing, infinite differentiable on R^+ functions $\delta(t)$ with $\delta(0) = 1$, satisfying following conditions: $\delta(t)$ is

- (a) downward convex, i.e. $\frac{d\delta(t)}{dt} > 0$ and $\frac{d^2\delta(t)}{dt^2} > 0$ for $t \in R^+$,
- (b) log-upward convex, i.e. $\frac{d \ln \delta(t)}{dt} > 0$ and $\frac{d^2 \ln \delta(t)}{dt^2} < 0$ for $t \in R^+$.

A *Special* class Ω_0 includes increasing functions $\delta(t), t \in [0, +\infty)$, of the type

$$\delta(t) = 1 + \frac{t}{A}(1 + o(1)), \quad t \rightarrow +\infty, \quad A \in R^+. \quad (4.1)$$

We single out the *linear* case: $o(1) \equiv 0$.

Now, if $\delta(t) \in \Omega$, then we *assume* that the limit exists

$$\lim_{t \rightarrow +\infty} \frac{t}{\delta(t)} = 0, \quad (4.2)$$

and in accordance with (4.2) two situations arise: either

$$\int_{0+}^{\infty} \frac{dt}{\delta(t)} = +\infty \quad (\text{then, denote } \delta(t) = \delta^-(t), \quad t \in R^+), \quad (4.3)$$

or

$$0 < c_1 = \int_{0+}^{\infty} \frac{dt}{\delta(t)} < +\infty \quad (\text{then, denote } \delta(t) = \delta^+(t), \quad t \in R^+). \quad (4.4)$$

In both situations (4.3) and (4.4), due to (4.2), we have

$$0 < c_n = \int_{0+}^n \frac{dt}{\delta(t)} < +\infty, \quad n = 2, 3, \dots$$

Denote $\Omega_+ = \{\delta^+(t)\}$, $\Omega_- = \{\delta^-(t)\}$.

Then, $\Omega = \Omega_+ \cup \Omega_-$ and $\Omega_+ \cap \Omega_- = \emptyset$.

For the *linear* case we are in situation (4.3), but the condition (4.2) does not hold.

Introducing *asymptotically equivalent* to $\delta(t)$ function $\varepsilon(t) \in \Omega$, or $\varepsilon(t) \in \Omega_?$ we are able to describe the required class of distribution functions.

Any pairs $(\delta^-(t), \varepsilon(t))$ and $(\delta^+(t), \varepsilon(t))$, $t \in R^+$, generates a *family* of distribution functions

$$\hat{F}_-(x, b) \text{ with } 0 < b < 1 \text{ and } \hat{F}_+(x, b) \text{ with } -1 < b < +\infty$$

given by formula (2.8), where

$$f_b^\pm(t) = \ln\left(1 \pm \frac{b}{\delta^\pm(t)}\right), \quad t \in [0, +\infty).$$

4.2 Particular cases

1. $\varepsilon(t) = \delta^-(t)$ and $\varepsilon(t) = \delta^+(t)$.

A thorough investigation for the class of stationary distributions generated by standard birth-death process with special restrictions on the process' intensities has been done in [6]. The dedcretization of this class leads exactly to the present particular case.

Let us write down the expansions

$$f_b^-(x) = -|f_b^-(x)| = \sum_{n \geq 1} (-1)^n \cdot \frac{1}{n} \left(\frac{b}{\delta^-(x)}\right)^n, \quad x \in R^+, \quad -1 < b < +\infty, \quad (4.5)$$

and

$$f_b^+(x) = |f_b^+(x)| = \sum_{n \geq 1} \frac{1}{n} \left(\frac{b}{\delta^+(x)}\right)^n, \quad x \in R^+, \quad 0 < b < 1. \quad (4.6)$$

The *partial* sums of the right-hand-side expressions in (4.5) and (4.6) denote by

$$[f_b^-(x)]_N = \sum_{n=1}^N (-1)^n \cdot \frac{1}{n} \left(\frac{b}{\delta^-(x)}\right)^n$$

and

$$[f_b^+(x)]_N = \sum_{n=1}^N \frac{1}{n} \left(\frac{b}{\delta^+(x)}\right)^n \quad \text{for } N = 1, 2, \dots$$

Then, we have following inequalities

$$\begin{aligned} [f_b^\pm(x)]_1 &< [f_b^\pm(x)]_3 < \cdots < [f_b^\pm(x)]_{2K-1} < \cdots < f_b^\pm(x) < \cdots \\ &\cdots < [f_b^\pm(x)]_{2K} < \cdots < [f_b^\pm(x)]_4 < [f_b^\pm(x)]_2 \end{aligned} \quad (4.7)$$

for sign "+" with $0 < b < 1$ and for sign "-" with $-1 < b < 0$.

Similarly, for sign "+" with $0 < b < +\infty$ the following inequalities hold

$$[f_b^+(x)]_1 < [f_b^+(x)]_2 < \cdots < [f_b^+(x)]_K < \cdots < f_b^+(x). \quad (4.8)$$

Easily seen that

$$[f_b^\pm(x)]_1 = \pm \frac{b}{\delta^\pm(x)}, \quad [f_b^\pm(x)]_2 = \pm \frac{b}{\delta^\pm(x)} + \frac{1}{2} \left(\frac{b}{\delta^\pm(x)} \right)^2. \quad (4.9)$$

Lemma 1. The functions $[f_b^\pm(x)]_N$ for any $N = 1, 2, \dots$ and $f_b^\pm(x)$ are asymptotically equivalent.

Proof. Due to (4.5)-(4.9) and $\lim_{x \rightarrow +\infty} \delta^\pm(x) = +\infty$, it is enough to prove that

$$\lim_{x \rightarrow +\infty} \frac{f_b^\pm(x)}{\pm(b/\delta^\pm(x))} = 1. \quad (4.10)$$

But (4.10) follows from the L'Hopital rule.

In accordance with *Lemma 1* the following sequence of examples, particular cases, arises. Instead of $\delta^\pm(t)$ we take $([f_b^\pm]_{N_1}/b)$ for *some* $N_1 = 1, 2, \dots$, and put $\varepsilon(t) = ([f_b^\pm]_{N_2}/b)$ for *some* $N_2 = 1, 2, \dots$. The case $b = 0$ is excluded.

The following subsequence is of interest. We take instead of $\delta^\pm(t)$ the function $([f_b^\pm]_N/b)$ and put $\varepsilon(t) = ([f_b^\pm]_N/b)$ for $N = 1, 2, \dots$.

Then, for the corresponding distribution functions we obtain the following expressions

$$\begin{aligned} \hat{F}_\pm(x, b) &= c_\pm(b) \cdot \int_{0-}^x \frac{b}{[f_b^\pm]_N} \exp \left\{ \int_{0-}^t [f_b^\pm(u)]_N du \right\} dt = \\ &= c_\pm(b) \cdot \left\{ \exp \left\{ \int_{0-}^x [f_b^\pm(u)]_N du \right\} - 1 \right\} \hat{F}_\pm^{(N)}(x, b). \end{aligned}$$

Note that the conditions (4.3) and (4.4) are equivalent to the following ones

$$\int_{0+}^{\infty} [f_b(t)]_N dt = +\infty \quad \text{and} \quad \int_{0-}^{\infty} [f_b(t)]_N dt < +\infty$$

for any (some) $N = 1, 2, \dots$, respectively.

The constant $c_+(b)$ is evaluated. Indeed, for sign ”-” with $0 < b < 1$ and for sign ”+” with $-1 < b < 0$ the function $[f_b^\pm(x)]_N$ is *negative*. In the first case $\int_{0-}^{\infty} [f_b^-(u)]_N du = -\alpha$. Therefore,

$$\hat{F}_-^{(N)}(x, b) = 1 - \exp \left\{ - \int_{0-}^x |[f_b^-(u)]_N| du \right\}, \text{ and } c_-(b) = 1.$$

In the second case

$$\hat{F}_+^{(N)}(x, b) = c_-(b) \cdot \left\{ 1 - \exp \left\{ - \int_{0-}^x |[f_b^+(u)]_N| du \right\} \right\},$$

where

$$c_-(b) = \left(1 - \exp \left\{ - \int_0^{\infty} |[f_b^+(u)]_N| du \right\} \right)^{-1}.$$

For the sign ”+” with $0 < b < +\infty$ we get *positive* $[f_b^+(x)]_N$ and because of this

$$\hat{F}_+^{(N)}(x, b) = c_+(b) \cdot \left(\exp \left\{ - \int_{0-}^x |[f_b^+(u)]_N| du \right\} - 1 \right),$$

where

$$c_+(b) = \left(\exp \left\{ + \int_{0+}^{\infty} |[f_b^+(u)]_N| du \right\} \right)^{-1}.$$

The consideration above shows that the particular case 1. is an *upper* bound for distribution function \hat{F}_- when $\varepsilon(t)$ is chosen among introduced particular cases. The same is true for $(\hat{F}_+(x, b)/c_+(b))$ in situations with $c_{\pm}(b) \neq 1$. Anyway, we have to indicate two particular cases.

$$2. \frac{1}{\varepsilon(t)} = -\frac{1}{b} \ln \left(1 - \frac{b}{\delta^-(t)} \right), 0 < b < 1 \text{ and } \frac{1}{\varepsilon(t)} = \frac{1}{b} \ln \left(1 + \frac{b}{\delta^+(t)} \right),$$

$$-1 < b < +\infty, b \neq 0.$$

$$3. \text{ We take } \delta^\pm(t) \text{ instead of } (\pm b / \ln(1 \pm \frac{b}{\delta^\pm(t)})) \text{ and put } \varepsilon(t) = \delta^\pm(t).$$

4.3 The result

Now we have

Theorem 3.

1. $\hat{F}_-(x, b)$ and $\hat{F}_+(x, b)$ with $-1 < b \leq 0$ are upward convex on R^+ ;

2. For $\hat{F}_+(x, b)$ with $0 < b < +\infty$ there is a point $x_0 \in R^+$ such that \hat{F}_+ is downward convex in $(0, x_0)$ and upward convex in $(x_0, +\infty)$;
3. $1 - \hat{F}_\pm(x, b)$ varies regularly with exponent $-(\alpha - 1 + bA)$.

Proof. $\hat{F}_\pm(x, b)$ has a density on R^+ .

$$\hat{\phi}_\pm(x, b) = \frac{d\hat{F}_\pm(x, b)}{dx} = \frac{\tilde{c}_\pm(b)}{\varepsilon(x)} \exp \left\{ \int_{0-}^x f_b^\pm(u) du \right\}, \quad (4.11)$$

where

$$\tilde{c}_\pm(b) = \left(\int_{0-}^{\infty} \frac{1}{\varepsilon(t)} \exp \left\{ \int_{0-}^t f_b^\pm(u) du \right\} dt \right)^{-1}.$$

For $x \in R^+$, due to (4.11), we have for $x \in R^+$

$$\frac{d}{dx} \hat{\phi}_\pm(x, b) = \frac{\tilde{c}_\pm(b)}{\varepsilon(x)} \exp \left\{ \int_{0-}^x f_b^\pm(u) du \right\} \cdot \left\{ f_b^\pm(x) - \frac{1}{\varepsilon(x)} \frac{d}{dx} \varepsilon(x) \right\}. \quad (4.12)$$

For sign "-" with $0 < b < 1$, and for the sign "+" with $-1 < b < 0$ we have $f_b^\pm(x) < 0$, and, therefore, because of (4.12), $(d\hat{\phi}_\pm(x, b)/dx) < 0$, which proves the statement 1. with $b \neq 0$. The case $b = 0$ is obvious.

For sign "+" with $0 < b < +\infty$ the limit exists $\lim_{x \rightarrow +\infty} (f_b^+(x)/(\frac{a}{\delta^+(x)})) = 1$ and $\frac{d}{dx} \varepsilon(x) > 0$, $\lim_{x \rightarrow +\infty} \frac{d}{dx} \varepsilon(x) = +\infty$. Due to asymptotical equivalency of $\varepsilon(x)$ and $\delta^+(x)$, and to continuity of terms at the right-hand-side of (4.12), we conclude that there is $x_0 \in R^+$ such that $\frac{d}{dx} \hat{\phi}_+(x, b) > 0$ for $x \in (0, x_0)$ and $\frac{d}{dx} \hat{\phi}_+(x, b) < 0$ for $x \in (x_0, +\infty)$. It proves the statement 2.

For $1 < s < +\infty$ let us evaluate the limit

$$\begin{aligned} \lim_{x \rightarrow +\infty} \frac{1 - \hat{F}_\pm(sx, b)}{1 - \hat{F}_\pm(x, b)} &= \lim_{x \rightarrow +\infty} \frac{\int_{sx}^{\infty} \frac{1}{\varepsilon(t)} \exp \left\{ \int_{0-}^t f_b^\pm(u) du \right\} dt}{\int_x^{\infty} \frac{1}{\varepsilon(t)} \exp \left\{ \int_{0-}^t f_b^\pm(u) du \right\} dt} \\ &= \left(\lim_{x \rightarrow +\infty} \frac{\varepsilon(sx)}{\varepsilon(x)} \right)^{-1} \cdot s \cdot \exp \left\{ \lim_{x \rightarrow +\infty} \int_x^{sx} f_b^\pm(u) du \right\} \quad \text{if exists,} \end{aligned}$$

where the L'Hopital rule is applied.

Let us show that if $A = 0$, then $\lim_{x \rightarrow +\infty} \int_x^{sx} f_b^\pm(u) du = 0$. Indeed, if (4.4) holds then $\lim_{x \rightarrow +\infty} \int_x^{sx} \frac{du}{\delta^+(u)} = 0$. If (4.3) holds and $A = 0$, then for $\varepsilon \in (0, 1)$ starting from some x_0 we have $\frac{1}{\delta^-(x)} < \varepsilon \cdot \frac{1}{x}$ for $x \in (x_0, +\infty)$, and $\int_x^{sx} \frac{du}{\delta^-(u)} < \varepsilon \cdot \ln s$. So, letting $\varepsilon \downarrow 0$ we obtain

$$\lim_{x \rightarrow +\infty} \int_x^{sx} \frac{du}{\delta^-(u)} = 0.$$

But, due to the L'Hopital rule $\int_x^{sx} f_b^\pm(u)du$ and $\pm b \cdot \int_x^{sx} \frac{du}{\delta^\pm(u)}$ are equivalent as $x \rightarrow +\infty$. Thus, the statement for these cases is established.

Since $\lim_{x \rightarrow +\infty} (\varepsilon(sx)/\varepsilon(x)) = s^\alpha$, therefore from (4.13) we get statement 3. for $A = 0$. (If $0 < s < 1$ we overturn the ratio at the left-hand-side in (4.13)).

The remaining case is $0 < A < +\infty$. Now, the problem consists of evaluation of the limit

$$\lim_{x \rightarrow +\infty} \int_x^{sx} f_b^-(u)du, \quad x_0 \leq x < +\infty. \quad (4.13)$$

Since $f_b^-(x)$ and $(-b/\delta^-(x))$, $(-b/\delta^-(x))$ and $(-bA/x)$ are asymptotically equivalent, therefore for $\varepsilon \in (0, 1)$ starting from some $x_0 > 0$ the inequalities hold

$$\frac{bA}{x}(1 - \varepsilon) < |f_b^-(x)| < \frac{bA}{x}(1 + \varepsilon), \quad x_0 \leq x < +\infty.$$

Thus, for $1 < s < +\infty$ we obtain

$$b \cdot A(\ln s)(1 - \varepsilon) \leq \lim_{x \rightarrow +\infty} \int_x^{sx} |f_b^-(u)|du \leq \overline{\lim}_{x \rightarrow +\infty} \int_x^{sx} |f_b^-(u)|du \leq b \cdot A(\ln s)(1 + \varepsilon).$$

Letting $\varepsilon \downarrow 0$ we come to the relationship

$$\lim_{x \rightarrow +\infty} \int_x^{sx} f_b^-(u)du = -b \cdot A \ln s, \quad 1 < s < +\infty. \quad (4.14)$$

Finally, (4.13) and (4.15) imply the statement 3. for $0 < A < +\infty$.

Theorem 3 is proved.

5 Conclusion

The standard birth-death process with intensities of moderate growth generates stationary skewed distributions suitable for modelling frequency distributions of events arising in large-scale biomolecular systems.

We studied a large class of distributions that can be used to model, for instance, frequency distributions of the number of expressed genes in the transcriptome, the number of protein domain occurrences in the proteomes, etc.

In particular, a new *dediscretization* approach was suggested, discussed and applied to the chosen class. This approach conserves the *qualitative* properties of the original class of distributions. The obtained distributions are often simpler

in form and also often easily analyzed by the tools of mathematical analysis that have been developed for continuous functions. We also studied several typical examples which illustrate the possibilities of the dediscretization approach. The reverse procedure of *dediscretization*, i.e. the procedure of *discretization* was also studied.

References

- [1] E. Danielian and J. Astola, "On the steady-state of birth-death process with coefficients' moderate growth," *Facta Univ., Ser.: Elec. and Energ.*, vol. 17, no. 3, pp. 405–419, Dec. 2004. [Online]. Available: factae.elfak.ni.ac.yu/fu2k43/Astola.pdf
- [2] J. Astola and E. Danielian, "On regularly varying distributions generated by birth-death process," *Facta Univ., Ser.: Elec. and Energ.*, vol. 19, no. 3, pp. 109–131, Apr. 2004. [Online]. Available: factae.elfak.ni.ac.yu/fu2k61/astola.pdf
- [3] V. A. Kuznetsov, "Family of skewed distributions associated with the gene expression and prometome evolution," *Signal Processing*, vol. 83, no. 4, pp. 889–910, 2003.
- [4] —, "Distributions associated with stochastic processes of gene expression in a single eukaryotic cell," *EURASIP Journal on Applied Signal processing*, no. 4, pp. 258–296, 2001.
- [5] E. Seneta, *Regularly Varying Functions: Lecture Notes in Mathematics*. Springer-Verlag, 1976.
- [6] J. Astola and E. Danielian, "Regularly varying skewed distributions generated by birth-death process," *Tampere, TICSP Series #*, 2004.
- [7] W. Feller, *An Introduction to Probability Theory and Its Applications*, 2nd ed. John Wiley and Sons, 1971, vol. II.
- [8] E. Danielian, "Regularly varying functions," *Tampere, TICSP Series 16*, 2001.