Annals of the Institute of Statistical Mathematics (2024) 76:679–709
https://doi.org/10.1007/s10463-024-00903-y

A delineation of new classes of exponential dispersion models supported on the set of nonnegative integers

Shaul K. Bar-Lev[1] · Gérard Letac[2] · Ad Ridder[3]

Received: 5 July 2023 / Revised: 20 February 2024 / Accepted: 26 February 2024 /
Published online: 22 April 2024

Abstract

The aim of this paper is to delineate a set of new classes of natural exponential families and their associated exponential dispersion models whose probability distributions are supported on the set of nonnegative integers with positive mass on 0 and 1. The new classes are obtained by considering a specific form of their variance functions. We show that the distributions of all these classes are supported on nonnegative integers, that they are infinitely divisible, and that they are skewed to the right, leptokurtic, over-dispersed, and zero-inflated (relative to the Poisson class). Accordingly, these new classes significantly enrich the set of probability models for modeling zero-inflated and over-dispersed count data. Furthermore, we elaborate on numerical techniques how to compute the distributions of our classes, and apply these to an actual data experiment.

Keywords Discrete distribution · Exponential dispersion model · Infinitely divisible · Lagrange inversion formula · Natural exponential family · Numerical computations · Variance function

✉ Ad Ridder
ad.ridder@vu.nl

Shaul K. Bar-Lev
shaulb@hit.ac.il

Gérard Letac
gerard.letac@math.univ-toulouse.fr

[1] Faculty of Industrial Engineering and Technology Management, Holon Institute of Technology, 52 Golomb Street, POB 305, Holon 5810201, Israel

[2] Institut de Mathématique de Toulouse, Université Paul Sabatier, 118 route de Narbonne, F-31062 Toulouse Cedex 9, France

[3] School of Business and Economics, Vrije University, de Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

1 Introduction

A classic tool in statistical modeling and analysis is the concept of natural exponential families (NEFs). These are classes of probability distributions that are convenient in practice for executing statistical analysis and numerical computations. In its generality, a NEF is defined to be the set of probability measures (Barndorff-Nielsen 1978),

$$\mathcal{F} = \mathcal{F}(\mu) = \left\{ P(\theta, \mu)(\mathrm{d}x) = e^{\theta x - k_\mu(\theta)}\, \mu(\mathrm{d}x) \,:\, \theta \in \Theta_\mu \right\}, \tag{1}$$

where

- μ is a positive Radon measure on $\mathbb{R}$ with closed convex support; it is called the generating kernel or measure of the NEF;
- $\theta \in \mathbb{R}$ is a canonical parameter;
- $D_\mu = \left\{ \theta \in \mathbb{R} \,:\, L_\mu(\theta) = \int_{\mathbb{R}} e^{\theta x}\, \mu(\mathrm{d}x) < \infty \right\}$ is the domain of the Laplace transform of the kernel;
- its interior $\Theta_\mu = \operatorname{int} D_\mu$ is nonempty;
- $k_\mu(\theta) = \log L_\mu(\theta)$ is the cumulant transform of measure μ.

Taking derivatives of the cumulant transform gives that $k'_\mu(\theta)$ and $k''_\mu(\theta)$ are the respective mean and variance corresponding to probability measure $P(\theta, \mu)$. The image $\mathcal{M} = \mathcal{M}_{\mathcal{F}} = k'(\Theta_\mu)$ is an open interval, and is called the mean domain of $\mathcal{F}$. As k_μ is strictly convex on Θ_μ, the inverse mapping of its derivative $k'_\mu \,:\, \theta \mapsto m$, $\Theta_\mu \to \mathcal{M}$, is well defined. Hence, the canonical parameter θ becomes a function of the mean m. This is denoted by $\theta = \psi(m)$. In this way, the variance corresponding to $P(\theta, \mu)$ becomes also a function of the mean, $V(m) = k''_\mu(\psi_\mu(m))$ for $m \in \mathcal{M}$.

The pair $(V, \mathcal{M})$ is called a variance function (VF). We cite an important property (Morris, 1982).

> *A VF uniquely determines the NEF within the class of NEFs and is not dependent on the specific measure μ.*

A natural generalization is the concept of exponential dispersion models (EDMs). An EDM associated with a NEF $\mathcal{F}(\mu)$ is a collection $\bigcup_{p\in\Lambda} \mathcal{F}(\mu_p)$ of NEFs where the generating kernels are defined implicitly by (Jørgensen, 1987),

$$\Lambda = \left\{ p \in \mathbb{R}^+ \,:\, L^p_\mu \text{ is a Laplace transform of some measure } \mu_p \right\}.$$

The parameter p is called a dispersion parameter. For any dispersion value $p \in \Lambda$, the associated NEF is the set of probability distributions of the form

$$\mathcal{F}_p = \mathcal{F}(\mu_p) = \left\{ P(\theta, \mu_p)(\mathrm{d}x) = e^{\theta x - p k_\mu(\theta)}\, \mu_p(\mathrm{d}x) \,:\, \theta \in \Theta_{\mu_p} = \Theta_\mu \right\}. \tag{2}$$

The contribution of our paper is that we delineate a set of new classes of EDMs whose probability distributions are concentrated on the set of nonnegative integers $\mathbb{N}_0$ with positive mass on 0 and 1. We will characterize these classes by their variance functions. Particularly, we show that any generating function $f(m) = 1 + \sum_{n=1}^{\infty} a_n m^n$ of coefficients $a_n \geq 0$ with positive radius of convergence can be used to generate a VF of the form $V(m) = mf(m)$, whose associated NEF (and EDM) is valued in $\mathbb{N}_0$. In this way we can construct an enormous number of classes of EDMs supported on $\mathbb{N}_0$. The vast majority of these classes have not been considered before, mainly because being not known. Consequently, these new classes enrich significantly the set of competitive models to those in frequent and widespread use (as the Poisson, negative binomial, Poisson-inverse Gaussian, Poisson-Tweedie and the like). Furthermore, we show that the distributions of all these classes are infinitely divisible, and we show that these distributions are skewed to the right, leptokurtic, over-dispersed, and zero-inflated (relative to the Poisson NEF).

Next to these structural properties, the second contribution of our study is that we present two methods for retrieving the generating measure μ of a NEF $\mathcal{F}$ supported on $\mathbb{N}_0$ by solely utilizing the given form of its respective VF $(V, \mathcal{M})$.

– Method I utilizes the techniques presented in Proposition 4.4 of Letac and Mora (1990) in which the mean value parametrization of NEFs plays a crucial role for the numerical computations of the related probability mass functions. Method I was instrumental in Letac and Mora (1990) for describing all cubic VFs of NEFs supported on $\mathbb{N}_0$.
– Method II uses directly the Lagrange formula and enables to give information about the generating function of the kernel μ (which is a discrete measure on $\mathbb{N}_0$).

Recently, Method I has been used in Bar-Lev and Ridder (2021) to generate two classes of EDMs, namely, the ABM and LM classes. These were applied in Bar-Lev and Ridder (2023) for modeling various real count datasets, and in Awad et al. (2022) for mortality projections. Method II is more limited in its applicability than Method I because of the level of analytical complexity involved. On the other hand, it has its importance because it presents analytical expressions for μ_n.

The paper is organized as follows. Section 2 presents various basic preliminaries on NEFs and EDMs, the mean value parameterization of NEFs and the Lagrange formula. Section 3 introduces our main tool (Proposition 2) for constructing new classes of VFs associated with NEFs/EDMs supported on $\mathbb{N}_0$. We also prove that the resulting NEFs/EDMs distributions are over-dispersed, zero-inflated, skewed to the right and leptokurtic. In fact we prove that the latter properties hold for the larger set of infinitely divisible NEFs supported on $\mathbb{N}_0$. Section 4 elaborates the practical steps for implementing Methods I and II, and in Sect. 5 we apply our methods to a range of variance functions. Numerical experiments are reported in Sect. 6. Finally, in Sect. 7 we summarize our findings.

2 Preliminaries

In this section, we summarize some necessary details of the mean value parametrization of NEFs and EDMs from Bar-Lev and Kokonendji (2017)), and some relevant facts related to the Lagrange formula from Letac and Mora (1990).

2.1 The mean value parameterization

NEFs.

For a given VF $(V, \mathcal{M})$ of a NEF $\mathcal{F}$ generated by a measure μ, choose any two primitives $\psi(m)$ and $\varphi(m)$ of $1/V(m)$ and $m/V(m)$, respectively, i.e.,

$$\begin{aligned}\psi(m) &= \int \frac{1}{V(m)}\, dm;\\ \varphi(m) &= \int \frac{m}{V(m)}\, dm.\end{aligned} \tag{3}$$

The following observation is obvious, but crucial (Letac and Mora, 1990). The proof is based on the fact that the measure μ is not unique for the NEF $\mathcal{F}$. Any exponential shift of μ, i.e., $\mu^*(\mathrm{d}x) = e^{a+bx}\mu(\mathrm{d}x)$ for some real a, b, generates the same NEF, i.e., $\mathcal{F}(\mu) = \mathcal{F}(\mu^*)$.

Corollary 1 *For any pair of primitives as defined in* (3), *there exists a positive Radon measure* $\mu^* \in \mathcal{M}$ *such that*

$$\varphi(m) = \log \int_{\mathbb{R}} e^{\psi(m)x}\, \mu^*(\mathrm{d}x), \quad m \in \mathcal{M}.$$

This leads to a reparameterization of a NEF.

Definition 1 The mean value parameterization of a NEF $\mathcal{F}$ is

$$\mathcal{F} = \left\{ P(m, \mu^*)(\mathrm{d}x) = e^{x\psi(m) - \varphi(m)}\, \mu^*(\mathrm{d}x)\, :\, m \in \mathcal{M} \right\}. \tag{4}$$

As there are infinitely many such primitives, $\psi(m)$ and $\varphi(m)$ should be chosen such that (4) is a family of probability distributions. We shall come back to this point in Sect. 4.

EDMs.

The mean value parameterization of EDMs are defined in a manner similar to NEFs. Indeed, given a VF $(V, \mathcal{M})$ of a NEF then the mean parameter space and variance corresponding to (2) are

$$\mathcal{M}_p = pM; \quad V_p(m) = pV(m/p),$$

whereas the respective primitives of $1/V_p(m)$ and $m/V_p(m)$ are given by, respectively,

$$\psi_p(m) = \int 1/V_p(m)\,dm;$$

$$\varphi_p(m) = \int m/V_p(m)\,dm.$$

Then, it is readily seen that

$$\psi_p(m) = \psi(m/p); \quad \varphi_p(m) = p\varphi(m/p), \tag{5}$$

and thus the mean value parameterization of (2) has the form

$$\mathcal{F}_p = \left\{ P(m, \mu_p^*)(\mathrm{d}x) = e^{x\psi(m/p) - p\varphi(m/p)}\ \mu_p^*(\mathrm{d}x)\ :\ m \in p\mathcal{M}, p > 0 \right\}. \tag{6}$$

All results presented in the sequel concerning NEFs hold also for their respective EDMs by just replacing $\psi(m)$ and $\varphi(m)$, respectively, by $\psi_p(m)$ and $\varphi_p(m)$ as given in (5).

2.2 The Lagrange formula

The Lagrange inversion formula is a powerful way of considering an equation of the type $z = f(m)$ (z as a function of m), and invert (or solve) it for m (under analytic condition on f). A special case is when $f(m)$ is of the form $m/g(m)$ for some analytic function $g(m)$. Commonly, then the formula is called the Lagrange–Bürmann formula, which has several presentations in the literature, essentially split between analysis and combinatorics.

Proposition 1 *Let $z = m/g(m)$ be an analytic (complex) function defined on an open disc centered at* 0, *such that $g(0) \neq 0$. Then, the inverse function has the form*

$$m = h(z) = \sum_{n=1}^{\infty} \frac{z^n}{n!} \left[\left(\frac{\mathrm{d}}{\mathrm{d}m} \right)^{n-1} (g(m))^n \right]_{m=0}. \tag{7}$$

More generally, let $z = u/g(u)$ as above, and suppose that $m = G(u)$ for some analytic function, then the general Lagrange formula is

$$m = G(h(z)) = \sum_{n=1}^{\infty} \frac{z^n}{n!} \left[\left(\frac{\mathrm{d}}{\mathrm{d}u} \right)^{n-1} G'(u)(g(u))^n \right]_{u=0}. \tag{8}$$

The Lagrange formula has been exploited in Letac and Mora (1990) for deriving explicit expressions for the measure μ of four NEFs supported on $\mathbb{N}_0$ and having cubic VFs. In Section 5 we shall give other applications.

3 EDMs on Nonnegative Integers

Suppose that a given VF $(V, \mathcal{M})$ satisfies

- The mean domain $\mathcal{M} = (0, b)$ for some $0 < b \leq \infty$.
- There exists a real analytic function φ' on $\mathcal{M}$ such that $\varphi'(m) = m/V(m)$ on $\mathcal{M}$, and such that

$$\lim_{m \to 0} \varphi'(m) = 1. \tag{9}$$

Proposition 4.4 of Letac and Mora (1990) shows that the corresponding NEF $\mathcal{F}$ is concentrated on $\mathbb{N}_0$. Moreover, this proposition provides a way to compute a generating kernel μ. Let φ be a primitive of φ', choose a primitive $\psi(m)$ of $1/V(m)$, satisfying

$$\lim_{m \to 0} m\, e^{-\psi(m)} = 1, \tag{10}$$

and define $G(m) = m\, e^{-\psi(m)}$. Then, the formulas for μ_n's are given by

$$\begin{cases} \mu_0 = e^{\varphi(m)}\big|_{m=0}; \\ \mu_n = \frac{1}{n!}\left[\left(\frac{\mathrm{d}}{\mathrm{d}m}\right)^{n-1} e^{\varphi(m)} \times \varphi'(m) \times (G(m))^n)\right]\big|_{m=0}, \quad n \geq 1. \end{cases} \tag{11}$$

Note that the necessary condition (9) leaves infinitely many choices for the primitive φ, and hence for a generating measure μ. By demanding

$$\lim_{m \to 0} \varphi(m) = 0, \tag{12}$$

we get by (11), $\mu_0 = 1$, and also $\mu_1 = 1$.

The above formulas for the μ_n's are rather cumbersome for numerical calculations. They form the basis of our Method I, whose implementation steps will be summarized in Sect. 4.1. A first successful attempt of applying Proposition 4.4 of Letac and Mora (1990) was made in Awad et al. (2022) for obtaining NEFs (and their EDM generalizations) that are defined by variance function of the form

$$V(m) = m(1+m)^r, \quad r = 0, 1, \dots; \quad \mathcal{M} = (0, \infty). \tag{13}$$

Clearly, it includes the Poisson ($r = 0$), negative binomial ($r = 1$) and Abel ($r = 2$, also known as generalized Poisson) EDMs as special cases. Moreover, the corresponding primitives $\psi(m)$ and $\varphi(m)$ are explicitly expressible. Because the authors implemented these EDMs in a Bayesian model, there was no need to compute the μ_n's explicitly, as these quantities are cancelled out when deriving appropriate posterior distributions. The class is named ABM after the authors names (Awad—Bar-Lev—Makov).

A second application of Proposition 4.4 of Letac and Mora (1990) was made by Bar-Lev and Ridder (2021, 2023) who again considered EDMs with variance functions (13), but also NEFs (and their EDM generalizations) defined by variance functions

$$V(m) = \frac{m}{(1-m)^r}, \quad r = 0, 1, \dots; \quad \mathcal{M} = (0, 1).$$

From the exact expressions of the primitives $\psi(m)$ and $\varphi(m)$, they used the formulas (11) for computing numerically the μ_n's. The interest was in fitting various real count data. The class is named LM after the authors names (Letac–Mora).

3.1 An instrumental proposition

In this section we introduce an instrumental proposition which provides a general tool for constructing VFs which satisfy the premises of Proposition 4.4 in Letac and Mora (1990), and thus their corresponding NEFs and EDMs are supported on $\mathbb{N}_0$. We shall use this proposition throughout this paper.

Proposition 2 *Assume that $f(m)$ admits a power series*

$$f(m) = 1 + \sum_{n=1}^{\infty} a_n m^n \tag{14}$$

with radius of convergence $R > 0$ and $a_n \geq 0$ for all n and let

$$V(m) = mf(m), \quad m \in (0, R).$$

Then $(V, \mathcal{M}) = (V, (0, R))$ is a VF of an infinitely divisible NEF $\mathcal{F} = \mathcal{F}(\mu)$ such that $\mu = \sum_{n=0}^{\infty} \mu_n \delta_n$ is supported on $\mathbb{N}_0$. Furthermore, μ_0 and μ_1 are positive.

Proof We deduce the proof by a result of Bar-Lev (1987) and its reformulation in Letac and Mora (1990, Theorem 3.2). Indeed, V is positive on (0, R). So we have to show that

$$\int_0^{m_0} \frac{\mathrm{d}m}{V(m)} = \infty,$$

for an arbitrary $m_0 \in (0, R)$. Clearly, $f(m)$ is positive in an interval $(-\alpha, R)$, $\alpha > 0$, and $1/f(m)$ is expandable in the power series

$$\frac{1}{f(m)} = 1 + b_1 m + b_2 m^2 + \cdots$$

where $b_i, i = 1, 2, \ldots$ are real coefficients and $1/f$ has a positive radius of convergence R_1. Hence,

$$\frac{1}{V(m)} = \frac{1}{m} + b_1 + b_2 m + \cdots = \frac{1}{m} - g(m)$$

where

$$g(m) = \frac{f(m) - 1}{mf(m)}$$

is a continuous function in $[0, R_1)$. Hence,

$$\int_0^{m_0} \frac{\mathrm{d}m}{V(m)} = \int_0^{m_0} \frac{\mathrm{d}m}{m} - \int_0^{m_0} g(m)\mathrm{d}m = \infty.$$

In order to use Proposition 4.4 of Letac and Mora (1990), we need to prove that the masses on 0 and 1 are positive. To this end we write: if μ is any measure which generates $\mathcal{F}$,

$$\mathrm{d}\theta = \psi'_\mu(m)\mathrm{d}m = \frac{\mathrm{d}m}{m} - g(m)\mathrm{d}m.$$

Let

$$A(m) = \int_0^m g(t)\mathrm{d}t$$

be the primitive of g on $[0, R_1)$ with $A(0) = 0$. Adding a first condition on μ we get

$$\theta = \psi_\mu(m) = \log m - A(m) \Rightarrow z = e^\theta = me^{-A(m)} \Rightarrow ze^{A(m)} = m.$$

The fact that $A(m) \to_{m\to 0} 0$ insures $m \sim_{m\to 0} z$. As a consequence $m = k'_\mu(\theta) = z + O(z^2) = e^\theta + O(z^2)$ since $A(0) = 0$ and—adding a second condition on μ—we get $k_\mu(\theta) = e^\theta + O(z^2) = z + O(z^2)$. Therefore,

$$L_\mu(\theta) = e^{k_\mu(\theta)} = e^{z+O(z^2)} = 1 + z + \sum_{n=2}^{\infty} \mu_n z^n.$$

With the two choices of constants we made we get $\mu_0 = \mu_1 = 1 > 0$. Hence by Bar-Lev (1987) it follows that $(V, \mathcal{M}) = (V, (0, R))$ is a VF of infinitely divisible NEF which satisfies the conditions of Proposition 4.4 of Letac and Mora (1990). In particular $\mathcal{F}$ is supported on $\mathbb{N}_0$. □

Proposition 2 inspires the following definition.

Definition 2

(i). $\mathcal{D}$ is the set of NEFs (or VFs) which fulfills the premises of Proposition 2.
(ii). $\mathcal{ID}$ is the set of infinitely divisible distributions on $\mathbb{N}_0$ with positive mass on 0 and 1.
(iii). $\mathcal{CP}$ is the set of infinitely divisible distributions on $\mathbb{N}_0$ with positive mass on 0.

Trivially $\mathcal{ID} \subset \mathcal{CP}$, and in Proposition 3 below, we shall see that $\mathcal{D}$ is strictly contained in $\mathcal{ID}$, thus various properties of the elements of $\mathcal{CP}$ are inherited by those of $\mathcal{D}$. Note that $\mathcal{D}$ is an infinite set of classes of NEFs/EDMs as any function of the form (14) can be used to construct such a class. Thus, the class $\mathcal{D}$ significantly enriches the set of count probability models aiming at count data modeling as well as at other applications.

3.2 Properties of the set $\mathcal{CP}$ inherited by $\mathcal{D}$

Let us adopt the following notation: if $u = (u_n)_{n=0}^{\infty}$, we write

$$f_u(z) = \sum_{n=0}^{\infty} u_n z^n. \tag{15}$$

Suppose that $\mu = (\mu_n)_{n=0}^{\infty}$ is a probability measure on $\mathbb{N}_0$. Recall that μ is said to be infinitely divisible if and only if for all $t > 0$ there exists a probability measure $\mu(t) = (\mu_n(t))_{n=0}^{\infty}$ such that for all $|z| \le 1$ one has

$$\left(f_\mu(z)\right)^t = f_{\mu(t)}(z).$$

From Feller (1968, page 271) we cite the following result describing the class $\mathcal{CP}$:

Lemma 1 *If $\mu_0 > 0$, then μ is infinitely divisible if and only if there exists a probability measure $\nu = (\nu_n)_{n=1}^{\infty}$ on $\mathbb{N}$ and a number $\lambda > 0$ such that f_μ has the form*

$$f_\mu(z) = e^{\lambda(f_\nu(z)-1)}. \tag{16}$$

Note that the latter representation of f_μ is equivalent to saying that all NEFs supported on $\mathbb{N}_0$ considered here are of a compound Poisson type in the following sense. Suppose that $Y_1, Y_2, \ldots$ are i.i.d. with distribution ν and consider an independent Poisson random variable N with mean λ. Then the distribution of $Y_1 + \cdots + Y_N$ is μ by a classical calculation using conditioning by N. This explains the notation $\mathcal{CP}$.

We shall use Lemma 1 in the following proposition recalling some important properties of $\mathcal{CP}$ (and thus of $\mathcal{D}$).

Proposition 3 *Consider the sets $\mathcal{D}, \mathcal{ID}$, and $\mathcal{CP}$ as defined in Definition 2, then,*

(a) *The inclusions $\mathcal{D} \subset \mathcal{ID} \subset \mathcal{CP}$ are strict.*
(b) *The elements of $\mathcal{CP}$ are zero-inflated relative to the Poisson NEF, which means $\mu_0 \ge \exp\left(-\sum_0^{\infty} n\mu_n\right)$ with equality if and only if μ is Poisson.*
(c) *The elements of $\mathcal{CP}$ are overdispersed relative to the Poisson NEF, which means $V_{\mathcal{F}(\mu)}(m) - m \ge 0$ with equality if and only if μ is Poisson.*
(d) *The elements of $\mathcal{CP}$ are skewed to the right, which means the third centered moment of μ is positive.*
(e) *The elements of $\mathcal{CP}$ are leptokurtic, which means that μ has a positive kurtosis.*

Proof (a). For seeing that $\mathcal{D} \ne \mathcal{ID}$, we consider a variance function of the form $V_{\mathcal{F}}(m) = m + \sum_{n=2}^{\infty} c_n m^n$ that belongs to $\mathcal{ID}$ but has at least one negative coefficient. For instance, let $f_\nu(z) = \frac{1}{2}(z + z^2)$, and $f_\mu(z) = e^{f_\nu(z)-1}$. By expanding $f_\mu(z)$ we get $\mu_0 > 0$ and $\mu_1 > 0$; hence, it is an element of $\mathcal{ID}$. However, standard calculations show that

$$V_{\mathcal{F}(\mu)}(m) = \frac{1}{8}(1 + 16m - \sqrt{1 + 16m})$$

which has $c_{2n+1} < 0$ for $n \geq 1$. Thus, $V_{\mathcal{F}}(m) \notin \mathcal{D}$. The origin of this example is the Hermite distribution, obtained as the compound Poisson (λ) sum of binomial ($n = 2, p$) variables, which is infinitely divisible. Specifically, $\lambda = 9/8, p = 2/3$, see Section 9.4 in Johnson et al. (2005).

For seeing that $\mathcal{ID} \neq \mathcal{CP}$, consider the probability ν such that $f_\nu(z) = \frac{1}{2}(z^2 + z^3)$, and define the probability μ by $f_\mu(z) = e^{f_\nu(z)-1}$. As μ satisfies the conditions of Lemma 1, it is an element of $\mathcal{CP}$. However, by expanding $f_\mu(z)$ we get $\mu_1 = 0$, which means that $\mu \notin \mathcal{ID}$.

(b)–(e). For the proofs of (b), (c), (d), and (e), we consider μ infinitely divisible, concentrated on $\mathbb{N}_0$ and such that $\mu_0 > 0$. Then, the statements follow from similar properties of compound Poisson distributions as set out in Section 9.3 in Johnson et al. (2005). □

A crucial observation of part (a) of Proposition 3 is that the NEFs (and EDMs) satisfying the conditions of Proposition 2 do not cover all NEFs on $\mathbb{N}_0$ which are infinitely divisible. Specifically this holds for compound Poisson distributions with generating functions $f_\mu(z) = e^{\lambda(f_\nu(z)-1)}$. These distributions are also called multiple Poisson distributions, they are infinitely divisible, and they come in a large variety of combinations, see Chapter 9 in Johnson et al. (2005). Each combination may or may not have a VF of the form in Proposition 2. We have seen that compounding Poisson with the Binomial $n = 2$ leads to an infinite divisible NEF not lying in $\mathcal{D}$. Also interesting is to consider the Poisson(λ)-Poisson(ϕ) compounding distribution, i.e., the distribution of $S_N = \sum_{j=1}^{N} Y_j$ where $N \sim$ Poisson(λ) and $Y \sim$ Poisson(ϕ). This distribution is also called the Neyman type A distribution (Johnson et al. 2005, Section 9.6). Hence, its generating function is

$$f_\mu(z) = e^{\lambda\left(e^{\phi(z-1)}-1\right)}.$$

We use μ as kernel of the NEF

$$\left\{P(\theta, \mu)(n) = \mu_n\, e^{\theta n - k_\mu(\theta)}\right\}$$

on $\mathbb{N}_0$, where the cumulant is obtained by $k_\mu(\theta) = \log f_\mu(e^\theta)$. Computing first and second derivatives, we obtain its variance function, cf. Vinogradov (2013, Theorem 5.1)

$$V(m) = m\left(1 + W\left(\frac{me^\phi}{\lambda}\right)\right), \quad m > 0,$$

in which $W(x)$ is the Lambert W function defined by $W(x)e^{W(x)} = x$. Whereas Vinogradov (2013) examined the NEF generated by the Neyman type A distribution in relation to the Lambert function, Mselmi (2022) introduced many more NEFs with variance functions that depend on the Lambert function, and therefore called Lambert classes. In Mselmi (2022), these NEFs are generated by subordinated Gamma or Lévy processes with a Poisson subordinator. Expansion of $W(x)$ in a Taylor series is done by the Lagrange inversion formula (7), and shows that the coefficients are alternating positive and negative (Vinogradov 2013, Equation (1.3))

$$a_n = \lim_{x \to 0} \frac{1}{n!} \left(\frac{\mathrm{d}}{\mathrm{d}x} \right)^{n-1} e^{-nx} = \frac{(-n)^{n-1}}{n!}.$$

We conclude that the NEF generated by the Poisson-Poisson compounded distribution is not an element of our set $\mathcal{D}$. Further investigation of other Lambert classes goes beyond this paper's scope.

4 Practical implementation steps

Since the set $\mathcal{D}$ is infinite, we shall select a number of classes that lead to relatively explicit computation. For this, Proposition 2 delineates functions $f(m)$ of the form (14) with radius of convergence $R > 0$ and $a_n \geq 0$ for all n. Then we use the fact that $V(m) = mf(m)$ is a VF of an infinitely divisible NEF or EDM supported on $\mathbb{N}_0$. This allows us in general to compute numerically μ_n by (11).

The choice of appropriate primitives of $\int \frac{\mathrm{d}m}{V(m)}$ and $\int \frac{m\mathrm{d}m}{V(m)}$ that fulfills conditions (9), (10), and (12) is crucial for implementing Proposition 2. Indeed, if $\tilde{\psi}(m)$ and $\tilde{\varphi}(m)$ are some specific primitives, then the sets of all corresponding primitives are given by $\{\tilde{\psi}(m) + c,\ c \in \mathbb{R}\}$, and $\{\tilde{\varphi}(m) + d,\ \mathrm{d} \in \mathbb{R}\}$, in which case c and d should be chosen to comply with (9)–(12) for selecting the specific $\psi(m)$ and $\varphi(m)$. However, we found that the following is a simpler way for such a choice. This is introduced in the following proposition.

Proposition 4 *Consider the VF $V(m) = mf(m)$ of Proposition 2 associated with a NEF $\mathcal{F} = \mathcal{F}(\mu)$, where $\mu = \sum_{n=0}^{\infty} \mu_n \delta_n$. Let*

$$\begin{aligned} g(m) &= \frac{f(m) - 1}{mf(m)}; \\ A(m) &= \int_0^m g(t)\mathrm{d}t. \end{aligned} \tag{17}$$

Furthermore, let ψ and φ satisfying the requirements of Proposition 4.4 *in Letac and Mora* (1990). *Then*

(*a*). *As μ_0 and μ_1 are positive, one can choose $\mu_0 = \mu_1 = 1$.*
(*b*). *The choice $\mu_0 = \mu_1 = 1$ is equivalent to requiring that*

$$\begin{aligned} \psi(m) &= \log m - A(m); \\ \varphi(m) &= \int_0^m \frac{\mathrm{d}t}{f(t)}. \end{aligned} \tag{18}$$

Proof

(a) By Proposition 2, μ_0 and μ_1 are positive. Recall that if μ is a generating measure, then $e^{ax+b}\mu(\mathrm{d}x)$ is also a generating measure. We can choose a and b such that $e^b\mu_0 = 1$, and $e^{a+b}\mu_1 = 1$.

(b) Assume that $\mu_0 = \mu_1 = 1$. Then

$$k_\mu(\theta) = \log\left(1 + e^\theta + \sum_{n=2}^{\infty} \mu_n e^{n\theta}\right)$$
$$\Rightarrow \quad m = k'_\mu(\theta) = \frac{e^\theta + \sum_{n=2}^{\infty} n\mu_n e^{n\theta}}{1 + e^\theta + \sum_{n=2}^{\infty} \mu_n e^{n\theta}}.$$

Therefore

$$m \to 0 \quad \Leftrightarrow \quad \theta \to -\infty,\ e^\theta \sim_{\theta\to-\infty} k'_\mu(\theta).$$

We get

$$e^{\psi_\mu(k'_\mu(\theta))} \sim_{\theta\to-\infty} k'_\mu(\theta);\ e^{\psi_\mu(m)} \sim_{m\to 0} m;\ e^{\psi_\mu(m)-\log m} \to_{m\to 0} 1.$$

And finally $\psi_\mu(m) - \log m \to_{m\to 0} 0$. Now,

$$\frac{\mathrm{d}}{\mathrm{d}m}\big(\psi_\mu(m) - \log m\big) = \frac{1}{V(m)} - \frac{1}{m} = -g(m).$$

Thus $\psi_\mu(m) - \log m$ is the primitive of $-g$ vanishing at 0, namely, $-A(m) = \int_0^m -g(t)\mathrm{d}t$. Since $\varphi' = 1/f$ and since for satisfying to the requirements of Proposition 4.4 of Letac and Mora (1990) it is required that $1 = \mu_0 = e^{\varphi(0)}$, we get $\varphi(0) = 0$ and therefore $\varphi(m) = \int_0^m \mathrm{d}t/f(t)$.

Conversely if $\psi(m) = \log m - A(m)$ and $\varphi(m) = \int_0^m \mathrm{d}t/f(t)$, then $\varphi(0) = 0$. From Letac and Mora (1990, (4.16)) we have $\mu_0 = e^{\varphi(0)}$; therefore, $\mu_0 = 1$. We now prove that $\mu_1 = 1$. Imitating the direct part we write

$$k_\mu(\theta) = \log\left(1 + \mu_1 e^\theta + \sum_{n=2}^{\infty} \mu_n e^{n\theta}\right)$$
$$\Rightarrow \quad m = k'_\mu(\theta) = \frac{\mu_1 e^\theta + \sum_{n=2}^{\infty} n\mu_n e^{n\theta}}{1 + e^\theta + \sum_{n=2}^{\infty} \mu_n e^{n\theta}}.$$

Therefore,

$$m \to 0 \quad \Leftrightarrow \quad \theta \to -\infty;\ \mu_1 e^\theta \sim_{\theta\to-\infty} k'_\mu(\theta),$$

and

$$\mu_1 e^{\psi_\mu(k'_\mu(\theta))} \sim_{\theta\to-\infty} k'_\mu(\theta);\ \mu_1 e^{\psi_\mu(m)} \sim_{m\to 0} m;\ \mu_1 e^{-A(m)} \to_{m\to 0} 1.$$

□

We shall divide the search for such $f(m)$'s falling into various categories of elementary functions as polynomial functions, rational functions, trigonometric functions, exponential functions and their inverse functions, including the inverse trigonometric functions and logarithms. However, for a practical reasoning it is more convenient to confine the search to those $f(m)$'s for which at least one of $\psi(m)$ and $\varphi(m)$ is expressible as elementary function. Otherwise the numerical computations of μ_n in (11) might turn out to be rather cumbersome.

4.1 Method I

The steps for implementing Method I, and some accompanying comments are basically elaborated in Bar-Lev and Ridder (2021, Appendix A) and Bar-Lev and Ridder (2023, Section 3). To summarize:

Step 1 Start with a VF $(V, \mathcal{M}) \in \mathcal{D}$ associated with a NEF $\mathcal{F} = \mathcal{F}(\mu)$ of probability distributions generated by kernel μ, and use its mean value parameterization (4).
Step 2 Next, compute the functions $f(m) = V(m)/m$, and $g(m)$, $A(m)$ as in (17).
Step 3 Then, compute the primitives ψ and φ by (18).
Step 4 Let, for $n = 2, 3, \ldots$,

$$H_n(m) = \varphi(m) + \log \varphi'(m) + nA(m),$$

compute its derivatives $H_n^{(j)}(m)$, $j = 1, \ldots, n-1$, and evaluate $H_n^{(j)}(0)/j!$.
Step 5 Next, compute the kernel μ by $\mu_0 = \mu_1 = 1$, and for $n = 2, 3, \ldots$ by applying (11) and the general chain rule,

$$\begin{aligned}
\mu_n &= \frac{1}{n!}\left(\left(\frac{d}{dm}\right)^{n-1} e^{\varphi(m)} \times \varphi'(m) \times e^{n(\log m - \psi(m))}\right)\Big|_{m=0} \\
&= \frac{1}{n!}\left(\left(\frac{d}{dm}\right)^{n-1} e^{H_n(m)}\right)\Big|_{m=0} \\
&= \frac{1}{n!} \sum \frac{(n-1)!}{k_1!k_2!\cdots k_{n-1}!}\left(\left(\frac{d}{dx}\right)^{k} e^{x}\right)_{x=H_n(0)} \prod_{j=1}^{n-1}\left(\left(\frac{H_n^{(j)}(m)}{j!}\right)\Big|_{m=0}\right)^{k_j} \\
&= \frac{1}{n} \sum \frac{1}{k_1!k_2!\cdots k_{n-1}!} e^{H_n(0)} \prod_{j=1}^{n-1}\left(\frac{H_n^{(j)}(0)}{j!}\right)^{k_j} \\
&= \frac{1}{n} e^{H_n(0)} \sum \prod_{j=1}^{n-1} \prod_{t=1}^{k_j} \frac{H_n^{(j)}(0)/j!}{t},
\end{aligned}$$

where the sum is over all nonnegative integer solutions of the Diophantine equation $\sum_{j=1}^{n-1} jk_j = n-1$, and where $k = \sum_{j=1}^{n-1} k_j$.
Step 6 Finally, the probabilities of the NEF are computed by

$$P_n(m,\mu) = \mu_n e^{n\psi(m)-\varphi(m)}, \quad n = 0, 1, \dots .$$

Remark 1 In a similar manner, the probabilities of an associated EDM with dispersion parameter p are computed. Just replace $V(m)$ by $V_p(m) = pV(m/p)$ in Step 1; $f(m)$ by $f_p(m) = f(m/p)$, $g(m)$ by $g_p(m) = g(m/p)/p$, and $A(m)$ by $A_p(m) = A(m/p)$ in Step 2; $\psi(m)$ by $\psi_p(m) = \log m - A_p(m)$, and $\varphi(m)$ by $\varphi_p(m) = p\varphi(m/p)$ in Step 3; and substitute these in $H_{n,p}(m) = \varphi_p(m) + \log \varphi_p'(m) + nA_p(m)$. Step 6 delivers the kernel $\mu_{n,p}$, which is substituted in

$$\mu_{n,p} e^{n\psi_p(m)-\varphi_p(m)}, \quad n = 0, 1, \dots .$$

4.2 Method II

In this section, we use the fact that NEFs belonging to $\mathcal{D}$ are valued in $\mathbb{N}_0$, and that their elements are infinitely divisible. Given a NEF of the type of Proposition 2, Method II tries to find from the knowledge of $V_{\mathcal{F}}(m) = mf(m)$ a probability measure ν on $\mathbb{N}$ and $\lambda > 0$ such that $\mathcal{F} = \mathcal{F}(\mu)$ is generated by a probability measure μ with the form (16), namely $f_\mu = e^{\lambda(f_\nu - 1)}$.

The implementation of Method II starts with the same first 3 steps of Method I. It goes on as follows.

Step 4 Using the natural parameter θ, the distributions of the NEF are given by (1), i.e.,

$$P_n(\theta,\mu) = \mu_n e^{\theta n - \kappa_\mu(\theta)}, \ n = 0, 1, \dots .$$

Recall from (15) the generating function $f_\mu(z) = \sum_{n=0}^{\infty} \mu_n z^n$ of a kernel μ, and the cumulant function

$$\kappa_\mu(\theta) = \log \sum_{n=0}^{\infty} \mu_n e^{\theta n} = \log f_\mu(e^\theta),$$

whose derivative $\kappa_\mu'(\theta) = m$, the mean of $P(\theta,\mu)$. Hence, we get the relations

$$\theta = \theta(m) = (\kappa_\mu')^{-1}(m) = \psi(m), \ \text{ and } \ \mathrm{d}\theta = \frac{\mathrm{d}m}{V(m)}. \tag{19}$$

In this way we get $z = z(\theta) = e^\theta$ as function of θ, and $z = z(m) = e^{\theta(m)}$ as function of m.

Step 5 Consider its inverse, $m = m(z)$, i.e., solve $z = e^{\psi(m)}$ for m. Defining the function $\ell(z) = \log f_\mu(z)$, we get a differential equation for this function, by

$$m(z) = \frac{\mathrm{d}}{\mathrm{d}\theta}\kappa_\mu(\theta) = \frac{\mathrm{d}}{\mathrm{d}\theta}\ell(e^\theta) = \frac{\mathrm{d}}{\mathrm{d}z}\ell(z) \times \frac{\mathrm{d}}{\mathrm{d}\theta}z(\theta) = \ell'(z)\, z, \tag{20}$$

with boundary condition $\ell(0) = 0$. Suppose that this equation is solvable, then we would obtain $\ell(z)$ explicitly.

Step 6 If Step 5 is not feasible, we consider $z = z(m)$ as function of m as in Step 4,

$$z = e^{\psi(m)} = e^{\log m - A(m)} = \frac{m}{e^{A(m)}}.$$

Apply the Lagrange–Bürmann inversion formula (7),

$$m = \sum_{n=1}^{\infty} \frac{z^n}{n!} \underbrace{\left(\left(\frac{d}{dm} \right)^{n-1} e^{nA(m)} \right) \Big|_{m=0}}_{a_n}$$

Then (20) gives

$$z\ell'(z) = \sum_{n=1}^{\infty} a_n \frac{z^n}{n!} \quad \Rightarrow \quad \ell(z) = \sum_{n=1}^{\infty} a_n \frac{z^n}{n \times n!}. \tag{21}$$

Step 7 Given $\ell(z)$ either by Step 5 or Step 6, we get an explicit expression for the generating function of kernel μ,

$$f_\mu(z) = e^{\ell(z)} = \sum_{n=0}^{\infty} \mu_n z^n.$$

Let $g_n = \mu_n e^{n\psi(m)}$, then an explicit expression is available for

$$G(z) = f_\mu\left(e^{\psi(m)} z\right) = \sum_{n=0}^{\infty} \mu_n \left(e^{\psi(m)} z\right)^n = \sum_{n=0}^{\infty} g_n z^n.$$

Step 8 Numerical inversion of the generating function $G(z)$ yields the coefficients $g_n, n = 0, 1, \ldots$, and from these the NEF probabilities,

$$P_n(m, \mu) = \mu_n e^{n\psi(m) - \varphi(m)} = g_n\, e^{-\varphi(m)}, \quad n = 0, 1, \ldots.$$

Remark 2 A particular feature of Method II is the fact we do not have to bother to the extension of the method to EDM. Indeed, the corresponding $\ell(z)$ when we replace μ by μ_p is simply $p\ell(z)$. Thus, defining $g_{p,n} = \mu_{p,n} e^{n\psi_p(m)}$, an explicit expression is available for

$$G_p(z) = e^{\ell_p(e^{\psi_p(m)} z)} = f_{\mu_p}\left(e^{\psi_p(m)} z\right) = \sum_{n=0}^{\infty} \mu_{p,n} \left(e^{\psi_p(m)} z\right)^n = \sum_{n=0}^{\infty} g_{p,n} z^n.$$

Numerical inversion of the generating function $G_p(z)$ yields the coefficients $g_{p,n}, n = 0, 1, \ldots$, and from these the EDM probabilities,

$$P_{p,n}(m, \mu) = \mu_{p,n} e^{n\psi_p(m) - \varphi_p(m)} = g_{p,n}\, e^{-\varphi_p(m)}, \quad n = 0, 1, \ldots.$$

Remark 3 For numerical inversion of a generating function $G(z)$ we applied the Fast Fourier Transform (FFT) method. The FFT assumes that the series converges for $|z| = 1$. Suppose that the radius of convergence $R < 1$. Then, choose any $0 < r < R$, and define

$$G^*(z) = G(rz) = \sum_{n=0}^{\infty} g_n (rz)^n = \sum_{n=0}^{\infty} g_n^* z^n,$$

which converges for $|z| = 1$. Numerical inversion gives the coefficients g_n^*, and hence $g_n = g_n^*/r^n$. Thus, the NEF probabilities are

$$P_n(m, \mu) = g_n\, e^{-\varphi(m)} = g_n^*\, e^{-n \log r - \varphi(m)} \quad n = 0, 1, \ldots.$$

5 Examples

In this section, we present several examples of classes in $\mathcal{D}$. Each class is given by the VF $(V_\omega, \mathcal{M})$ of the NEF, where ω represents extra parameter(s). The associated primitives $\psi(m)$ and $\varphi(m)$ in the examples of Sects. 5.1, 5.2, and 5.3 are both elementary functions, whereas in Sects. 5.4 and 5.5 one is not. For the associated EDM we add dispersion parameter p, thus VF $(V_{\omega,p}, \mathcal{M}_p)$. Moreover, the examples have the following distinctions.

– In Sect. 5.1 we exploit Method II, where the mean m can be solved explicitly in Step 5.
– Section 5.2 uses the Lagrange formula for applying Method II.
– The example in Sect. 5.3 is worked out by computing Method I.
– Section 5.4 is worked out by Method I, and Sect. 5.5 by Method II.

We prefer to be able to compute Method II, because when Method II is analytically solvable, and numerically implementable, it is faster, and more stable than an implementation of Method I.

5.1 Polynomial VF

In their review paper, Bar-Lev and Kokonendji (2017) presented the following class of VFs as an example for which both primitives $\psi(m)$ and $\phi(m)$ have explicit forms. This class has extra parameters $a > 0$ and $r \in \mathbb{N}_0$, and is given by VF

$$V_{a,r}(m) = m(1+m)(1+m/a)^r, \quad \mathcal{M} = (0, \infty).$$

Then, by setting $\begin{pmatrix} r \\ r+1 \end{pmatrix} = 0$,

$$f_{a,r}(m) = (1+m)(1+m/a)^r = 1 + \sum_{i=1}^{r+1} \frac{1}{a^{i-1}} \left(\binom{r}{i-1} + \binom{r}{i} \frac{1}{a} \right) m^i,$$

which shows that $V_{a,r} \in \mathcal{D}$.

We work out the special (and simple) case of $a = 2$ and $r = 1$, thus

$$V(m) = m(1+m)(1+m/2), \quad m > 0.$$

The functions in Step 2 and Step 3 are derived easily.

$$\begin{aligned}
f(m) &= \frac{V(m)}{m} = (1+m)(1+m/2);\\
g(m) &= \frac{f(m)-1}{V(m)} = \frac{(1+m)(1+m/2)-1}{m(1+m)(1+m/2)} = \frac{3/2+m/2}{(1+m)(1+m/2)};\\
A(m) &= \int_0^m g(t)\,\mathrm{d}t = 2\log(1+m) - \log(1+m/2);\\
\psi(m) &= \log m - A(m) = \log m - 2\log(1+m) + \log(1+m/2);\\
\varphi(m) &= \int_0^m \frac{\mathrm{d}t}{f(t)} = 2\log(1+m) - \log(1+m/2).
\end{aligned}$$

We show the computations for Method II. In Step 5 we derive m as function of z by solving $z = e^{\psi(m)}$.

$$\begin{aligned}
z = e^{\psi(m)} = \frac{m(1+m/2)}{(1+m)^2} \quad &\Leftrightarrow \quad m + \frac{1}{2}m^2 = z + 2zm + zm^2\\
&\Leftrightarrow \quad (z - \frac{1}{2})m^2 + (2z-1)m + z = 0,
\end{aligned}$$

with solution

$$m = \frac{1}{\sqrt{1-2z}} - 1,$$

for $0 < z < 1/2$. Thus, the differential Eq. (20) is

$$z\ell'(z) = m = \frac{1}{\sqrt{1-2z}} - 1,$$

such that $\ell(0) = 0$. Solution is (see Appendix A),

$$\ell(z) = \log \frac{1-\sqrt{1-2z}}{z(1+\sqrt{1-2z})} + \log 2.$$

Note that $\ell(z) < \infty$ for $|z| < 1/2$. Since we compute in Step 7 the series $G(z)$ for $\ell(e^{\psi(m)}z)$, we need

$$e^{\psi(m)}|z| < \frac{1}{2} \quad \Leftrightarrow \quad |z| < \frac{1}{2}\, e^{-\psi(m)} = \frac{(1+m)^2}{2m(1+m/2)} = 1 + \frac{1}{2m(1+m/2)} = R.$$

Thus, $R > 1$, and the FFT method is applied in the usual manner.

5.2 Sibuya VF

This class has an extra parameters $a \in (0,1)$, such that $a = 1/(r+1)$ for some $r = 1, 2, \ldots$. The VF is

$$V_a(m) = \frac{1-(1-m)^a}{a}, \quad \mathcal{M} = (0,1).$$

Recall

$$(1-m)^a = \sum_{i=0}^{\infty} (-1)^i \binom{a}{i} m^i, \quad \text{where } \binom{a}{i} = \frac{a(a-1)\cdots(a-i+1)}{i!}.$$

Thus,

$$(1-m)^a = 1 - am + \frac{1}{2}a(a-1)m^2 + \sum_{i=3}^{\infty} (-1)^i \binom{a}{i} m^i.$$

Then,

$$f_a(m) = V_a(m)/m = 1 + \frac{1}{2}(1-a)m + \frac{1}{a}\sum_{i=2}^{\infty} (-1)^i \binom{a}{i+1} m^i,$$

which shows that $V_a \in \mathcal{D}$.

We work out the special (and simple) case of $a = 1/2$ ($r = 1$), thus

$$V(m) = 2(1 - \sqrt{1-m}), \quad 0 < m < 1.$$

The functions in Step 2 and Step 3 are as follows, see Appendix B for the calculus.

$$\begin{aligned}
f(m) &= \frac{V(m)}{m} = \frac{2(1-\sqrt{1-m})}{m};\\
g(m) &= \frac{f(m)-1}{V(m)} = \frac{2(1-\sqrt{1-m})-m}{2m(1-\sqrt{1-m})};\\
A(m) &= \int_0^m g(t)\,\mathrm{d}t = 1 - \sqrt{1-m} - \log f(m);\\
\psi(m) &= \log m - A(m);\\
\varphi(m) &= \int_0^m \frac{\mathrm{d}t}{f(t)} = \frac{m}{2} + \frac{1-(1-m)^{3/2}}{3}.
\end{aligned}$$

We apply Method II. We cannot directly solve $z = e^{\psi(m)}$ for m. However, by substituting $u = 1 - \sqrt{1-m}$, we get

$$m = G(u) = 1 - (1-u)^2.$$

Using the relation (19) for $\mathrm{d}\theta$, and the chain rule, leads to $\theta = \theta(u) = \log u - u$. Then $z = z(u)$ as a function of u is given by,

$$z = e^{\theta} = \frac{u}{e^u}.$$

There is no explicit solution to this equation; therefore, we invert it by the Lagrange formula to give u as function of z. Because $m = G(u)$, the generalized Lagrange formula (8) gives m as function of z,

$$m = \sum_{n=1}^{\infty} \frac{z^n}{n!} \underbrace{\left[\left(\frac{\mathrm{d}}{\mathrm{d}u} \right)^{n-1} 2(1-u)e^{nu} \right]_{u=0}}_{a_n}.$$

Using Leibniz formula of differentiating for $n = 2, 3, \ldots$,

$$a_n = 2 \left[\sum_{k=0}^{n-1} \binom{n-1}{k} \left(\frac{\mathrm{d}}{\mathrm{d}u} \right)^{n-1-k} (1-u) \times \left(\frac{\mathrm{d}}{\mathrm{d}u} \right)^{k} e^{nu} \right]_{u=0} = 2n^{n-2},$$

and $a_1 = 2$. Hence, Step 6 gives

$$\ell(z) = \sum_{n=1}^{\infty} a_n \frac{z^n}{n \times n!} = 2z + 2 \sum_{n=2}^{\infty} \frac{z^n}{n!} n^{n-3}.$$

The radius of convergence of $\ell(z)$ is given by

$$\lim_{n\to\infty} \frac{n^{n-3}}{n!} \Big/ \frac{(n+1)^{n-2}}{(n+1)!} = \lim_{n\to\infty} \left(1 - \frac{1}{n+1}\right)^{n-3} = 1/e.$$

For the series $\ell(e^{\theta(u)} z)$ in Step 7, we need

$$e^{\theta(u)} |z| < \frac{1}{e} \quad \Leftrightarrow \quad |z| < \frac{1}{e} e^{-\theta(u)} = \frac{1}{u} e^{u-1}.$$

Easy to see that $e^{u-1}/u > 1$ on $(0, 1)$. Thus $R > 1$, and the FFT method is applied in the usual manner.

5.3 Rational VF

The rational VF is defined by

$$V_c(m) = \frac{m(1 + 2cm - m^2)}{1 - m^2}, \quad \mathcal{M} = (0, 1),$$

with parameter $c > 0$. Then

$$f_c(m) = V_c(m)/m = \frac{1 + 2cm - m^2}{1 - m^2} = 1 + \frac{2cm}{1 - m^2},$$

which shows that $V_c \in \mathcal{D}$. We use the following identities. Let $c = \sinh a$, $a > 0$, then

$$\frac{2\cosh a}{1 + 2x\sinh a - x^2} = \frac{2\cosh a}{(e^a - x)(e^{-a} + x)} = \frac{1}{e^a - x} + \frac{1}{e^{-a} + x},$$

and

$$\frac{2x\cosh a}{1 + 2x\sinh a - x^2} = \frac{2x\cosh a}{(e^a - x)(e^{-a} + x)} = \frac{e^a}{e^a - x} - \frac{e^{-a}}{e^{-a} + x}.$$

The functions in Step 2 and Step 3 are

$$\begin{aligned}
g_c(m) &= \frac{f_c(m) - 1}{mf_c(m)} = \frac{2c}{1 + 2cm - m^2} = \frac{2\sinh a}{(e^a - m)(e^{-a} + m)} \\
&= \tanh a\Big(\frac{1}{e^a - m} + \frac{1}{e^{-a} + m}\Big); \\
A_c(m) &= \int_0^m g_c(t)\,dt = \tanh a(2a - \log(e^a - m) + \log(e^{-a} + m)); \\
\psi_c(m) &= \log m - A_c(m); \\
\varphi_c(m) &= \int_0^m \frac{dt}{f_c(t)} = m - 2a\frac{\sinh^2 a}{\cosh a} + \tanh a(e^a\log(e^a - m) + e^{-a}\log(e^{-a} + m)).
\end{aligned}$$

The function $H_n(m)$, $n = 2, 3, \ldots$ in Step 4 of Method I is

$$H_n(m) = m + \log(1 - m) + \log(1 + m) + \alpha_n \log(e^a - m) + \beta_n \log(e^{-a} + m) + \gamma_n,$$

where

$$\begin{aligned}
\alpha_n &= (e^a - n)\tanh a - 1; \\
\beta_n &= (e^{-a} + n)\tanh a - 1; \\
\gamma_n &= -2a\frac{\sinh^2 a}{\cosh a} + 2na\tanh a.
\end{aligned}$$

Derivatives of $H_n(m)$ are computed by considering separately its terms, which gives

$$\frac{H_n^{(j)}(0)}{j!} = -\frac{1}{j} + (-1)^{j-1}\frac{1}{j} - \alpha_n\frac{e^{-ja}}{j} + \beta_n(-1)^{j-1}\frac{e^{ja}}{j}.$$

These derivatives are implemented in Step 5 for the kernel μ_c, from which the NEF probabilities follow in Step 6.

Method II is numerically more demanding, and left out.

5.4 Exponential VF 1

Our first exponential VF is defined by

$$V(m) = m\,e^m, \quad \mathcal{M} = (0, \infty).$$

Then

$$f(m) = V(m)/m = e^m = 1 + \sum_{n=1}^{\infty} \frac{1}{n!}\, m^n,$$

which shows that $V \in \mathcal{D}$. The functions in Step 2 and Step 3 are

$$\begin{aligned}
g(m) &= \frac{f(m) - 1}{mf(m)} = \frac{e^m - 1}{m\,e^m} = \frac{1 - e^{-m}}{m} = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{m^{n-1}}{n!};\\
A(m) &= \int_0^m g(t)\,\mathrm{d}t = \int_0^m \sum_{n=1}^{\infty} (-1)^{n-1} \frac{t^{n-1}}{n!}\,\mathrm{d}t = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n}\, \frac{m^n}{n!};\\
\psi(m) &= \log m - A(m);\\
\varphi(m) &= \int_0^m \frac{1}{f(t)}\,\mathrm{d}t = \int_0^m e^{-t}\,\mathrm{d}t = 1 - e^{-m}.
\end{aligned}$$

This VF has no elementary primitive $\psi(m)$. However, consider the function $H_n(m)$ is Step 4,

$$\begin{aligned}
&1 - e^{-m} - m + n \sum_{i=1}^{\infty} (-1)^{i-1} \frac{1}{i}\, \frac{m^i}{i!}\\
&= \sum_{i=1}^{\infty} (-1)^{i-1} \frac{m^i}{i!} - m + n \sum_{i=1}^{\infty} (-1)^{i-1} \frac{1}{i}\, \frac{m^i}{i!}\\
&= -m + \sum_{i=1}^{\infty} (-1)^{i-1} \left(1 + \frac{n}{i}\right) \frac{m^i}{i!}.
\end{aligned}$$

We see immediately that $H_n(0) = 0$. Derivatives of $H_n(m)$ are computed by considering that for $j = 1, 2, \ldots$ and $i = 1, 2, \ldots,$

$$\left(\frac{\mathrm{d}}{\mathrm{d}m}\right)^j m^i = \begin{cases} \frac{i!}{(i-j)!} m^{i-j} & \text{for } j \le i;\\ 0 & \text{for } j > i. \end{cases}$$

Thus, for $j = 2, 3, \ldots,$ we get

$$\frac{H_n^{(j)}(0)}{j!} = (-1)^{j-1} \left(1 + \frac{n}{j}\right) \frac{1}{j!} = (-1)^{j-1} \frac{1}{j!} + n(-1)^{j-1} \frac{1}{j(j!)}.$$

For the first derivative $H_n^{(1)}(0)$, add -1. These derivatives are implemented in Step 5 for the kernel μ. In Step 6, we need to compute the infinite series of the function

$A(m)$ (in the primitive $\psi(m)$). The series is approximated by computing the first $k-1$ terms,

$$A_k(m) = \sum_{i=1}^{k-1} (-1)^{i-1} \frac{1}{i} \frac{m^i}{i!}.$$

Then, for $k > me$,

$$|A(m) - A_k(m)| \le \sum_{i=k}^{\infty} \frac{1}{i} \frac{m^i}{i!} \approx \sum_{i=k}^{\infty} \frac{1}{i} \frac{m^i}{\sqrt{2\pi i}(i/e)^i}$$
$$\le \sum_{i=k}^{\infty} \left(\frac{me}{k}\right)^i = \left(\frac{me}{k}\right)^k \sum_{i=k}^{\infty} \left(\frac{me}{k}\right)^{i-k} = \frac{(me/k)^k}{1-(me/k)}.$$

For a small $\epsilon > 0$, find k such that

$$\frac{(me/k)^k}{1-(me/k)} < \epsilon \quad \Leftrightarrow \quad \left(\frac{me}{k}\right)^k < \epsilon\left(1 - \frac{me}{k}\right) \quad \Leftrightarrow \quad \left(1 - \frac{me}{k}\right)k^k > \frac{(me)^k}{\epsilon}.$$

Let $k_0 = \lceil me \rceil$, thus $me/k_0 < 1$, and

$$1 - \frac{me}{k} > 1 - \frac{me}{k_0}.$$

Thus, find $k > k_0$ such that

$$\left(\frac{me}{k}\right)^k < \epsilon\left(1 - \frac{me}{k_0}\right) \quad \Leftrightarrow \quad k \log(me/k) < \log \epsilon\left(1 - \frac{me}{k_0}\right).$$

In this way, the probabilities in Step 6 of Method I are computed. Method II is cumbersome, and not implemented.

5.5 Exponential VF 2

Our second exponential VF has the form

$$V(m) = e^m - 1 \quad \mathcal{M} = (0, \infty),$$

which has appeared in Letac (2022). Then

$$f(m) = V(m)/m = \frac{e^m - 1}{m} = \sum_{n=1}^{\infty} \frac{1}{n!} m^{n-1} = 1 + \sum_{n=1}^{\infty} \frac{1}{(n+1)!} m^n, \tag{22}$$

which shows that $V \in \mathcal{D}$. The functions in Step 2 and Step 3 are

$$
\begin{aligned}
g(m) &= \frac{f(m)-1}{mf(m)} = \frac{1}{m} - \frac{1}{e^m - 1};\\
A(m) &= \int_0^m g(t)\,\mathrm{d}t = \log m - \int_0^m \frac{e^t + 1 - e^t}{e^t - 1}\,\mathrm{d}t\\
&= \log m - \int_0^m \frac{e^t}{e^t-1}\,\mathrm{d}t + \int_0^m 1\,\mathrm{d}t = \log m - \log(e^m - 1) + m;\\
\psi(m) &= \log m - A(m) = \log(e^m - 1) - m;\\
\varphi(m) &= \int_0^m \frac{1}{f(t)}\,\mathrm{d}t = \int_0^m \frac{t}{e^t - 1}\,\mathrm{d}t = \int_0^m \frac{te^{-t}}{1-e^{-t}}\,\mathrm{d}t\\
&\overset{u=1-e^{-t}}{=} -\int_0^{1-e^{-m}} \frac{\log(1-u)}{u}\,\mathrm{d}u = \mathrm{Li}_2(1-e^{-m}),
\end{aligned}
$$

where $\mathrm{Li}_2(z)$ is the dilogarithm function given by

$$
\mathrm{Li}_2(z) = -\int_0^z \frac{\log(1-u)}{u}\,\mathrm{d}u \quad z \in \mathbb{C}. \tag{23}
$$

In case that $|z| < 1$, we get by expanding $\log(1-u)$,

$$
\mathrm{Li}_2(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^2}. \tag{24}
$$

A useful identity is (Abramowitz and Stegun 1965, Eq. 27.7.3).

$$
\mathrm{Li}_2(z) + \mathrm{Li}_2(1-z) = \frac{\pi^2}{6} - (\log z)(\log(1-z)).
$$

Applying this identity to $z = 1 - e^{-m}$, we get

$$
\varphi(m) = \frac{\pi^2}{6} - \mathrm{Li}_2(e^{-m}) + m\log(1-e^{-m}).
$$

We shall work out the computations for Method II. We have

$$
z = e^{\psi(m)} = e^{\log(e^m-1)-m} = 1 - e^{-m} \quad \Leftrightarrow \quad m = -\log(1-z),
$$

for $|z| < 1$. The differential equation in Step 5 is

$$
z\ell'(z) = -\log(1-z) \quad \Leftrightarrow \quad \ell'(z) = -\frac{\log(1-z)}{z}.
$$

The solution is the dilogarithm function $\ell(z) = \mathrm{Li}_2(z)$, see (23). The series $G(z)$ in Step 7 is defined for

$$
e^{\psi(m)}|z| < 1 \quad \Leftrightarrow \quad |z| < e^{-\psi(m)} = \frac{1}{1-e^{-m}} = R,
$$

where clearly $R > 1$. The FFT applies to $G(z)$.

Table 1 Parameter values of the distribution classes

VF	p	Extra	Method	Remark
Polynomial	7.17		II	Section 5.1
Sibuya	3.43	0.5	II	Section 5.2
Rational	5.05	0.33	I	Section 5.3
Exponential 1	5.66		I	Section 5.4
Exponential 2	3.09		II	Section 5.5
Poisson				Traditional
Negative binomial	4.13			Traditional

6 Numerical experiments

To see our distributions in action, we run two experiments. The first experiment is a visualization of distributions coming from the five classes presented in Sect. 5. The second experiment is fitting these distributions to a count dataset. In both experiments we compare with the traditional count models of Poisson and negative binomial. These experiments have illustrative purposes only, and are by no means exhaustive for showing the capabilities of our distributions. More on this issue will be said in the concluding Sect. 7.

6.1 Visualization

We compute the probability mass functions of distributions in all five classes presented in Sect. 5. The parameters are chosen such that all these distributions have mean $\mu = 3.4$, and variance $\sigma^2 = 6.2$. These numbers are chosen arbitrarily. Distributions from EDM classes have always the mean parameter m, and the dispersion parameter p, and thus a two-moment fit is possible. When distributions have an extra parameter, we fit skewness to be 2. As for the computations, when possible we apply Method II, otherwise Method I. The code for computing the distribution of the Sibuya class by Method II (Sect. 5.2) is available at https://github.com/adriaanridder/sibuya

The code for the other distributions follow by similar implementations. For comparison, we consider the Poisson distribution with mean $m = \mu$, and the negative binomial distribution with mean $m = \mu$ and dispersion parameter p. The relation to the usual parameterization π for success probability, and r for the number of failures is $r = p$, and $\pi = m/(m + p)$. Table 1 lists the specifics.

In Fig. 1, we show the probability mass functions of the five distributions versus the Poisson and negative binomial.

Next, in Fig. 2 we show the distributions versus the Poisson by their cumulative probabilities. We give these for the Sibuya and exponential 1 only, the other distributions show similar graphs. In Fig. 3, we show two distributions (rational and exponential 2) versus the Poisson by their quantiles. Again, the other distributions show similar graphs.

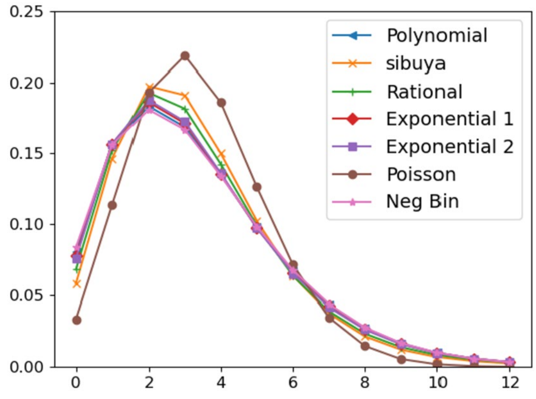

Fig. 1 Probability mass functions of the distributions with mean $\mu = 3.4$, and variance $\sigma^2 = 6.2$

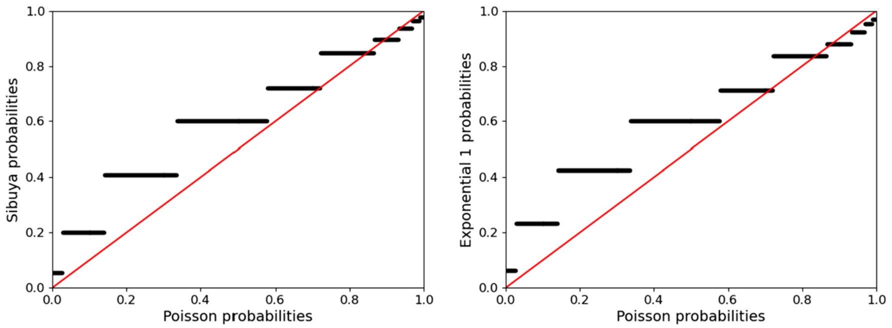

Fig. 2 P-P plots of distributions versus the Poisson

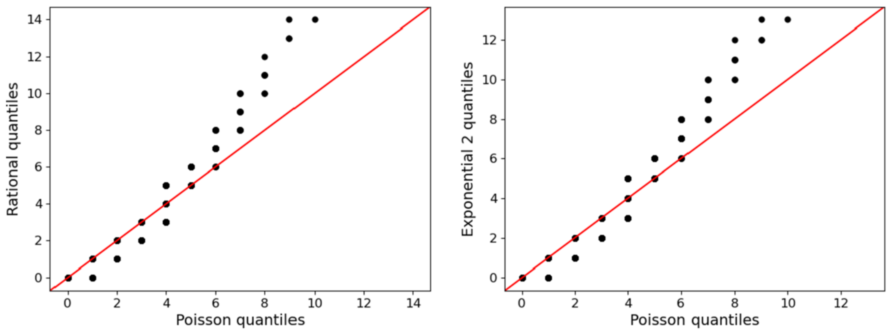

Fig. 3 Q-Q plots of distributions versus the Poisson

The overall picture is that our distributions show more probability mass at small values than the Poisson (the zero-inflated property of Proposition 3), and have fatter tails (the leptokurtic property of Proposition 3).

Table 2 The count data

Value	0	1	2	3	4	5	6	7	8	9	10	11	12
Frequency	1333	404	133	43	25	10	4	4	1	2	2	0	1

Table 3 Moments of the data

Mean	Variance	Skewness	Kurtosis
0.5398	1.1781	3.7316	21.6006

6.2 Data fitting

We consider a dataset of the number of direct job changes from the German Socio-Economic Panel (Winkelmann, 2008), given in Table 2.

For fitting parametric distributions we have applied moment matching, similar to Sect. 6.1, and maximum likelihood estimation (MLE). The latter seems obvious considering that we deal with exponential families which makes MLE rather straightforward. However, in case of our EDMs, the generating measure $\mu = \mu_p$ depends also on the dispersion parameter that we wish to estimate, see their expressions (2) and (6). This makes the numerical optimization of the likelihood function more demanding. Table 3 lists the moments of the data for the moment matching method for estimating the distributional parameters, as given in Table 4. This table gives also the parameters when estimating these by MLE. For the Sibuya class we increased the power parameter to $r = 2$ (thus parameter $a = 1/(r + 1) = 1/3$) for getting a feasible solution. The derivations for this case follow similar to the exposition in Sect. 5.2, though with a bit more effort.

Each fitted distribution is compared with the empirical distribution of the real data, and the fit is assessed by the root mean square errors (RMSE), and the Kullback–Leibler divergence (KLD), see Table 4, using the MLE parameters.

We see indeed that, except for the Sibuya class, our new distributions perform much better than the traditional Poisson and negative binomial distributions, mainly because they capture the zero-inflated and leptokurtic characteristics of

Table 4 Fitted parameter values of the distributions and the resulting performance

	Moment match		MLE		Performance	
VF	p	Extra par	p	Extra par	RMSE	KLD
Polynomial	0.8326		0.9576		0.004197	0.003886
Sibuya	0.5533	0.3333	0.5408	0.3333	0.01105	0.006365
Rational	1.0039	0.7826	0.8086	0.5381	0.001739	0.002387
Exponential 1	0.6915		0.7360		0.002685	0.002629
Exponential 2	0.3855		0.3979		0.001926	0.002263
Poisson					0.040780	0.114849
neg. binomial	0.4564		0.5708		0.005369	0.005505

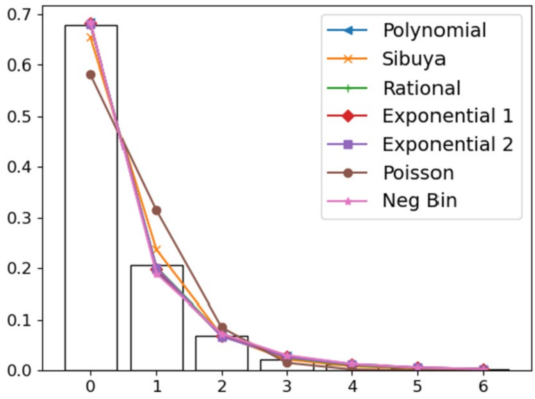

Fig. 4 Probability mass functions of the fitting distributions

the dataset. The rational, the exponential 1, and exponential 2 classes are compatible for RMSE and KLD criteria. We hypothesize that the Sibuya class will perform better by further increasing the power parameter r. The probability mass functions of the fitted distributions are visualized in Fig. 4.

7 Conclusion

Modeling count data by parametric families of counting distributions has always been part of the statistical literature with an enormous number of applications in various fields. Originally, count data models were based on the Poisson and the negative binomial distributions (Bliss and Fisher, 1953). Since then, many models have been developed and analyzed for their properties, see Coly et al. (2016) for the historic development of these models. Moreover, reviews of numerous count data models can be found in Winkelmann (2008), and Bar-Lev and Ridder (2023).

All these models share that their probability functions can be expressed explicitly in a closed form, allowing relatively easy computations and parameter estimation. In our opinion, this is the main reason these distributions have been used for statistical modeling. Researchers refrain from using an analytically complex mechanism that only allows numerical calculations of probabilities. Consequently, they prefer

– Add extra probability mass on zero, such as Hurdle models, or zero-inflated Poisson or negative binomial distributions (Feng, 2021).
– Generalize existing models, such as generalized Poisson or negative binomial distributions (Consul, 1989).
– Discretize versions of continuous distributions, such as discrete Lindley distribution (Gomez-Deniz and Calderin-Ojeda, 2011), and discrete Gamma distribution (Chakraborty and Chakravarty, 2012).
– Mixture distributions, such as Poisson-inverse Gaussian (Willmot, 1987), and Poisson-Tweedie (Kokonendji et al., 2004).

The present paper aims to introduce a new set of counting distributions based on explicit variances functions belonging to what we called class $\mathcal{D}$. We have provided numerical comparisons with only a limited number of classical distributions (such as Poisson and negative binomial) to demonstrate the applicability of our results. We have shown that the class $\mathcal{D}$ has a large number of exciting features:

1. It establishes an excellent set of new exponential dispersion models which can be used to describe the random error distribution in generalized linear models (GLMs). Also, as all such EDMs are infinitely divisible, their dispersion space is $\mathbb{R}^+$ as indeed required by GLM methodology.
2. All members of $\mathcal{D}$ are skewed to the right, leptokurtic, over-dispersed, and zero-inflated.
3. All cumulants, moments, and central moments of each EDM in this class are easily obtainable using the following. Let $\kappa_r, r \geq 1$, be the r-th cumulant of a NEF with VF (V, M) with $\kappa_1 = m$ and $\kappa_2 = V(m)$. Then $\kappa_{r+1} = \kappa_{r+1}(m)$ is given by

   $$\kappa_{r+1}(m) = V(m)\kappa_r'(m), r \geq 2,$$

   from which the corresponding moments and central moments can be deduced.
4. Four cubic VFs of discrete EDMs, all of which belong to class $\mathcal{D}$, were characterized by Letac and Mora (1990).
5. Another essential point to note is that there are of course infinitely many kernels (generating measures) that can generate NEFs (and thus EDMs). However, for a large part of them, their Laplace function cannot be expressed explicitly and, therefore, practically, cannot serve as a normalizing constant for computing the relevant NEFs' probabilities. Fortunately, the $\mathcal{D}$ class provides expressible alternatives to these normalizing constants in terms of the mean value parameterization, as elaborated in Section 2.1.

We trust that our presentation of the $\mathcal{D}$ class will play a significant role in (a) model fit for real count datasets; (b) describing the random error distribution in GLM with covariates (as alternatives to the Poisson, negative binomial, or any other distribution in use); (c) use in insurance risk models. These are interesting avenues for future work in which we intend to consider again the VF examples of Sect. 5.

Acknowledgements The authors are grateful to an Associate Editor and two reviewers for their constructive comments and suggestions for improvements.

Appendix A Polynomial VF calculus for section 5.1

Here, we show how to solve

$$z\ell'(z) = \frac{1}{\sqrt{1-2z}} - 1 \quad \Leftrightarrow \quad \ell'(z) = \frac{1}{z\sqrt{1-2z}} - \frac{1}{z},$$

such that $\ell(0) = 0$. The $1/z$ term delivers $\log z$. Let

$$y' = y'(z) = \frac{1}{z\sqrt{1-2z}} \quad \Leftrightarrow \quad \mathrm{d}y = \frac{\mathrm{d}z}{z\sqrt{1-2z}} \quad \Leftrightarrow \quad y = \int \frac{\mathrm{d}z}{z\sqrt{1-2z}}.$$

Substitute

$$u = \sqrt{\Delta} = \sqrt{1-2z} \quad \Leftrightarrow \quad z = \frac{1}{2}(1-u^2),$$

and $\mathrm{d}z = -u\,\mathrm{d}u$. Then,

$$y = -\int \frac{2\,\mathrm{d}u}{1-u^2} = -\int \left(\frac{1}{1-u} + \frac{1}{1+u}\right)\mathrm{d}u = \log(1-u) - \log(1+u) + C.$$

Thus,

$$\ell(z) = \log(1-\sqrt{\Delta}) - \log(1+\sqrt{\Delta}) - \log z + C = \log \frac{1-\sqrt{\Delta}}{z(1+\sqrt{\Delta})} + C,$$

where $C = \log 2$ to ensure $\ell(0) = 0$.

Appendix B Sibuya VF calculus for section 5.2

The VF with parameter $a = 1/(r+1)$, for $r = 0, 1, \ldots$,

$$V_a(m) = \frac{1-(1-m)^a}{a}, \quad \mathcal{M} = (0,1).$$

Then,

$$f_a(m) = V_a(m)/m = \frac{1-(1-m)^a}{am}.$$

from which

$$g_a(m) = \frac{f_a(m)-1}{mf_a(m)} = \frac{(r+1)\big(1-(1-m)^{1/(r+1)}\big) - m}{m(r+1)\big(1-(1-m)^{1/(r+1)}\big)}.$$

A long line of manipulations for

$$
\begin{aligned}
A_a(m) &= \int_0^m g_a(t)\,\mathrm{d}t \\
&\overset{u=(1-t)^{1/(r+1)}}{=} \int_{(1-m)^{1/(r+1)}}^{1} \frac{(r+1)(1-u)-(1-u^{r+1})}{(1-u)(1-u^{r+1})} u^r\,\mathrm{d}u \\
&= \int_{(1-m)^{1/(r+1)}}^{1} \left(\frac{(r+1)u^r}{1-u^{r+1}} - \frac{u^r}{1-u} \right) \mathrm{d}u \\
&= \int_{(1-m)^{1/(r+1)}}^{1} \left(\frac{(r+1)u^r}{1-u^{r+1}} + \frac{1-u^r}{1-u} - \frac{1}{1-u} \right) \mathrm{d}u \\
&= \int_{(1-m)^{1/(r+1)}}^{1} \left(\frac{(r+1)u^r}{1-u^{r+1}} + (1+u+u^2+\cdots+u^{r-1}) - \frac{1}{1-u} \right) \mathrm{d}u \\
&= \left[-\log\left(\frac{1-u^{r+1}}{1-u} \right) + \sum_{i=1}^{r} \frac{u^i}{i} \right]_{(1-m)^{1/(r+1)}}^{1} \\
&= -\log(r+1) + \sum_{i=1}^{r} \frac{1}{i} - \log\frac{1-(1-m)^{1/(r+1)}}{m} - \sum_{i=1}^{n} \frac{1}{i}(1-m)^{i/(n+1)} \\
&= -\log(r+1) + \sum_{i=1}^{r} \frac{1}{i} - \log\frac{f_a(m)}{r+1} - \sum_{i=1}^{r} \frac{1}{i}(1-m)^{i/(r+1)} \\
&= \sum_{i=1}^{r} \frac{1-(1-m)^{i/(r+1)}}{i} - \log f_a(m).
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
\varphi_a(m) &= \int_0^m \frac{1}{f_a(t)}\,\mathrm{d}t = \int_{(1-m)^{1/(r+1)}}^{1} \frac{1-u^{r+1}}{1-u} u^r\,\mathrm{d}u \\
&= \int_{(1-m)^{1/(r+1)}}^{1} (1+u+u^2+\cdots+u^r)u^r\,\mathrm{d}u \\
&= \left[\sum_{i=r+1}^{2r+1} \frac{u^i}{i} \right]_{(1-m)^{1/(r+1)}}^{1} = \sum_{i=r+1}^{2r+1} \frac{1-(1-m)^{i/(r+1)}}{i}.
\end{aligned}
$$

References

Abramowitz, M., Stegun, I. A. (1965). *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*. New York: Dover Publications Inc.

Awad, Y., Bar-Lev, S. K., Makov, U. (2022). A new class of counting distributions embedded in the Lee-Carter model for mortality projections: A Bayesian Approach. *Risks, 10*, 111–127.

Bar-Lev, S. K. (1987). Discussion on paper by B. Jørgensen, Exponential dispersion models. *Journal of the Royal Statistical Society, Series B, 49*, 153–154.

Bar-Lev, S. K., Kokonendji, C. C. (2017). On the mean value parameterization of natural exponential families - a Revisited Review. *Mathematical Methods of Statistics, 26*, 159–175.

Bar-Lev, S.K., Ridder, A. (2021). New exponential dispersion models for count data – The ABM and LM classes. *ESAIM: Probability and Statistics*, 25, 31-52.
Bar-Lev, S. K., Ridder, A. (2023). Exponential dispersion models for overdispersed zero-inflated count data. *Communications in Statistics-Simulation and Computation, 52*, 3286–3304.
Barndorff-Nielsen, O. (1978). *Information and Exponential Families in Statistical Theory*. Chichester: Wiley.
Bliss, C. I., Fisher, R. A. (1953). Fitting the negative binomial distribution to biological data. *Biometrics, 9*, 176–200.
Chakraborty, S., Chakravarty, D. (2012). Discrete Gamma distributions: properties and parameter estimations. *Communications in Statistics - Theory and Methods, 41*, 3301–3324.
Coly, S., Yao, A.-F., Abrial, D., Charras-Garrido, M. (2016). Distributions to model overdispersed count data. *Journal de la Société Française de Statistique, 157*, 39–63.
Consul, P. C. (1989). *Generalized Poisson distributions: properties and applications*. New York: Marcel Dekker.
Feller, W. (1968). *An Introduction to Probability Theory and Its Applications,* (3rd ed., Vol. I). New York: John Wiley.
Feng, C.X. (2021). A comparison of zero-inflated and hurdle models for modeling zero-inflated count data. *Journal of Statistical Distributions and Applications* 8, article 8.
Gomez-Deniz, E., Calderin-Ojeda, E. (2011). The discrete Lindley distribution: properties and applications. *Journal of Statistical Computation and Simulation, 81*, 1405–1416.
Johnson, N. L., Kemp, A. W., Kotz, S. (2005). *Univariate Discrete Distributions* (3rd ed.). Hoboken, New Jersey: Wiley.
Jørgensen, B. (1987). Exponential dispersion models (with discussion). *Journal of the Royal Statistical Society, Series B, 49*, 127–162.
Kokonendji, C. C., Dossou-Gbété, S., Demétrio, C. G. B. (2004). Some discrete exponential dispersion models: Poisson-Tweedie and Hinde-Demétrio classes. *Statistics and Operations Research Transactions, 28*, 201–214.
Letac, G. (2022). Duality for real and multivariate exponential families, *Journal of Multivariate Analysis* 188, article 104811.
Letac, G., Mora, M. (1990). Natural real exponential families with cubic variance functions. *The Annals of Statistics, 18*, 1–37.
Morris, C. N. (1982). Natural exponential families with quadratic variance functions. *The Annals of Statistics, 10*, 65–80.
Mselmi, F. (2022). Generalized linear model for subordinated Lévy processes. *Scandinavian Journal of Statistics, 49*, 772–801.
Rainville, E. D. (1960). *Special Functions*. Bronx, New York: Chelsea Publishing Company.
Vinogradov, V. (2013). Some utilizations of Lambert W function in distribution theory. *Communications in Statistics - Theory and Methods, 42*, 2025–2043.
Willmot, G. (1987). The Poisson-inverse Gaussian distribution as an alternative to the negative binomial. *Scandinavian Actuarial Journal, 1987*, 113–127.
Winkelmann, R. (2008). *Econometric Analysis of Count Data* (5th ed.). Berlin, Heidelberg: Springer.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.