

Article

The Large Arcsine Exponential Dispersion Model—Properties and Applications to Count Data and Insurance Risk

Shaul K. Bar-Lev ^{1,*} and Ad Ridder ² ¹ Faculty of Industrial Engineering and Technology Management, HIT—Holon Institute of Technology, Holon 5810201, Israel² School of Business and Economics, Vrije University, 1081 HV Amsterdam, The Netherlands

* Correspondence: shaulb@hit.ac.il

Abstract: The large arcsine exponential dispersion model (LAEDM) is a class of three-parameter distributions on the non-negative integers. These distributions show the specific characteristics of being leptokurtic, zero-inflated, overdispersed, and skewed to the right. Therefore, these distributions are well suited to fit count data with these properties. Furthermore, recent studies in actuarial sciences argue for the consideration of such distributions in the computation of risk factors. In this paper, we provide a thorough analysis of the LAEDM by deriving (a) the mean value parameterization of the LAEDM; (b) exact expressions for its probability mass function at $n = 0, 1, \dots$; (c) a simple bound for these probabilities that is sharp for large n ; (d) a simulation algorithm for sampling from LAEDM. We have implemented the LAEDM for statistical modeling of various real count data sets. We assess its fitting performance by comparing it with the performances of traditional counting models. We use a simulation algorithm for computing tail probabilities of the aggregated claim size in an insurance risk model.



Citation: Bar-Lev, S.K.; Ridder, A. The Large Arcsine Exponential Dispersion Model—Properties and Applications to Count Data and Insurance Risk. *Mathematics* **2022**, *10*, 3715. <https://doi.org/10.3390/math10193715>

Academic Editor: Manuel Alberto M. Ferreira

Received: 26 August 2022

Accepted: 6 October 2022

Published: 10 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: natural exponential family; exponential dispersion model; variance function; count data; Monte Carlo simulation

MSC: 62E99; 62P99

1. Introduction

In this paper, we consider a class of parameterized probability distributions on the non-negative integers, which are given as members of an exponential dispersion model with specific variance functions. To clarify these concepts, we begin with the classic technique in statistics of representing a family of probability measures (or distributions) on the real line as a natural exponential family (NEF) [1,2] and its subsequent generalization to an exponential dispersion model (EDM) [3]. A key feature of an NEF is that it expresses the variances of the distributions as functions of their means. This leads not only to the mean value parameterization of the NEF [4], but also to its unique pair of mean and variance, which is called its variance function and denoted by $V(m)$ [5].

Next, an interesting question arises as to what functions of $V(m)$ do natural exponential families of probability distributions have as their variance function. In his seminal work, Morris [5] characterized all six NEFs with quadratic variance functions. Letac and Mora [2] characterized all six NEFs with a cubic variance structure. Out of the latter six families, two are absolutely continuous (relative to the Lebesgue measure) and four are discrete, supported by the set of non-negative integers. These four are the Abel, Takács, strict arcsine, and large arcsine families. The probabilities of the first three families can be well expressed both in terms of their Laplace transform and in terms of its mean value representation. This was utilized by Bar-Lev and Ridder [6], who implemented the Abel, Takács, and strict arcsine families for fitting Swedish car insurance claim data. These families were used to model the counting variable N in the random sum $S_N = \sum_{j=1}^N Y_j$ representing the

aggregated claim distribution in an insurance risk model, in which N is the number of claims over a time period and Y_i s are the claim sizes. Bar-Lev and Ridder [6] computed the risk measure $\mathbb{P}(S_N > x)$ for large x and demonstrated superiority over the commonly used counting distributions (e.g., Poisson and negative binomial).

The fourth discrete natural exponential family of probability distributions with a cubic variance function, namely the large arcsine family, was presented by Letac and Mora [2] in their characterization of cubic variance functions of NEFs on $\mathbb{R}$. This was also obtained by Fosam and Shanbhag [7] in their characterization of the constant regression of cubic polynomial statistics on the sample mean. However, beyond these two characterization results, the large arcsine family has not been studied in the literature with respect to its probabilistic and statistical features and thus has not been analyzed as a candidate for modeling any real count data sets. The main reason for this seems to be due to the fact that its Laplace transform does not have an explicit form.

Our contribution in this paper aims to fill this gap. We present probabilistic and statistical aspects of the large arcsine NEF and its associated exponential dispersion model, which we denote by LAEDM. The kernel (or generating measure) of the probability distributions of the LAEDM has a rather cumbersome analytic expression with different forms for even and odd values. Moreover, the respective Laplace transform of its generating measure cannot be expressed explicitly and thus cannot serve as a normalizing constant for computing the relevant LAEDM probabilities. Fortunately, an expressible alternative to this normalizing constant is available in terms of the mean and is obtained by utilizing the mean value parameterization of NEFs.

Our method for determining the LAEDM distributions is based on the mean value parameterization of NEFs and uses the Lagrange formula of Letac and Mora [2]. This leads to exact expressions for the LAEDM probabilities which grant it the ability to run the simulation experiments that are needed for applications such as risk model computations. However, in many other cases, such as the ABM and LM distributions of Bar-Lev and Ridder [8,9], the probabilities can be only computed numerically without being able to use Monte Carlo simulations. Recently, Jørgensen and Kokonendji [10] introduced two- and three-parameter discrete dispersion models by combining convolution with a factorial tilting operation. The Poisson–Tweedie mixture model appears from this approach and is implemented in our data fitting experiments in Section 6.

As opposed to the Abel, Takács, and the strict arcsine families which depend on two parameters, the LAEDM depends on three parameters, a fact that allows more flexibility in statistical modeling. In this paper, we thoroughly study various features of the LAEDM. We rewrite its kernel and derive its probability mass function in terms of its mean value representation. As the latter is rather intricate (particularly, for large positive mass probability n), we study its asymptotic behavior as well as tail probabilities. We provide expressions for the respective moments and show that all members of LAEDM are skewed to right and leptokurtic. In addition, we present a scheme of Monte Carlo simulation for the LAEDM probabilities. Such simulations are needed for many statistical, insurance risk-related, and operational problems. Most of these problems cannot be solved analytically, a fact which requires Monte Carlo computations.

We also consider the effectiveness of LAEDM as a candidate for modeling real count data sets, while comparing it to other frequently used probability models. For this, we utilize various metrics of goodness-of-fit tests such as chi-squared tests, Akaike information criterion, root-mean-square error, and Kullback–Leibler divergence to demonstrate its superior fit.

The paper is organized as follows. For the sake of readability, we summarize in Section 2 some preliminaries of the concept's natural exponential family, exponential dispersion model, mean value parameterization, and variance function. Section 3 presents the main features and properties of the LAEDM, notably, (a) its mean value parameterization; (b) exact expression of the probability mass function $f(n)$; (c) a simple bound for these probabilities that is sharp for large n ; (d) moments, central moments, and coefficients of skewness and kurtosis. In particular, we show that all LAEDM distributions are skewed to the right and

are leptokurtic. Section 4 develops a Monte Carlo algorithm for sampling from the LAEDM distributions. Applications of the LAEDM are given in Sections 5 and 6. In Section 5, we implement the LAEDM in an collective risk model and analyze its statistical performance for a data set of car insurances, and in Section 6, the LAEDM is implemented in the statistical modeling of various real count data sets. Section 7 is devoted to some concluding remarks.

2. Preliminaries

We summarize the concepts of the natural exponential family (NEF), the exponential dispersion model (EDM), mean value parameterization, and variance function.

2.1. Natural Exponential Family

Let μ be a positive non-Dirac Radon measure on $\mathbb{R}$, S the support of μ , and C the convex-hull of S . The Laplace transform of μ is defined by

$$L(\theta) = \int_{\mathbb{R}} e^{x\theta} \mu(dx),$$

where the effective domain D of μ is $D = \{\theta \in \mathbb{R} : L(\theta) < \infty\}$. We assume that $\Theta = \text{int}D \neq \emptyset$, so that Θ is a nonempty open interval. Thus the cumulant transform $k(\theta) = \log L(\theta)$ of μ is well defined by Θ . The family $\mathcal{F}$ of probability measures (or distributions) defined by

$$\mathcal{F} = \{P_{\theta}(dx) = e^{x\theta - k(\theta)} \mu(dx), \quad \theta \in \Theta\}$$

is the NEF generated by μ . Therefore, μ is called a generating kernel for $\mathcal{F}$. It is easy to see that μ is not unique in its generation of $\mathcal{F}$; any exponential shift $\mu^*(dx) = e^{a+bx} \mu(dx)$ generates the same family of distributions.

2.2. Variance Function

The cumulant transform $\lambda_{\theta}(s)$ of a distribution $P_{\theta}(dx) \in \mathcal{F}$ is defined by $\log \int_{\mathbb{R}} e^{sx} P_{\theta}(dx)$, which equals

$$\lambda_{\theta}(s) = \log \int_{\mathbb{R}} e^{sx} e^{\theta x - k(\theta)} \mu(dx) = k(s + \theta) - k(\theta). \quad (1)$$

The first moment and the j -th central moments $j = 2, 3, \dots$ of a distribution $P_{\theta}(dx) \in \mathcal{F}$ can be computed by the derivatives $(d/ds)^j \lambda_{\theta}(s)|_{s=0}$. From (1) we see that these moments are equal to $k^{(j)}(\theta) = (d/d\theta)^j k(\theta)$, $j = 1, 2, \dots$. In particular, $k'(\theta)$ and $k''(\theta)$, and $\theta \in \Theta$, are the respective mean and variance.

The open interval $\mathcal{M} = k'(\Theta)$ is called the mean domain of $\mathcal{F}$. Note that it depends on the generating kernel μ only through its exponential shifts. The map $\theta \mapsto k'(\theta)$ is one to one; thus, its inverse function $(k')^{-1} : \mathcal{M} \rightarrow \Theta$ is well defined. Then, the variance of a distribution $P_{\theta}(dx) \in \mathcal{F}$ can be expressed as a function of the mean m ,

$$V(m) = k''((k')^{-1}(m)). \quad (2)$$

The pair $(V, \mathcal{M})$ is called the variance function of the NEF $\mathcal{F}$. It uniquely determines $\mathcal{F}$ within the class of NEFs (see Letac and Mora [2], Morris [5]).

2.3. Mean Parameterization

The inverse function $(k')^{-1}$, as a function on the mean domain $\mathcal{M}$, is denoted by ψ . Differentiating the ψ function, we obtain the following from (2):

$$\psi'(m) = \frac{1}{k''((k')^{-1}(m))} = \frac{1}{V(m)}.$$

Next, we define the function ϕ on the mean domain by $\phi(m) = k((k')^{-1}(m)) = k(\psi(m))$. Differentiating, we obtain

$$\phi'(m) = k'(\psi(m))\psi'(m) = \frac{m}{V(m)}.$$

Note that in this way, the ψ and ϕ functions can be interpreted as being primitives of the functions $1/V(m)$ and $m/V(m)$, respectively, on the mean domain $\mathcal{M}$.

Now, let there be given a variance function $(V, \mathcal{M})$ of an NEF $\mathcal{F}$, without having specified the generating kernel μ . We choose two primitives ψ and ϕ of $1/V(m)$ and $m/V(m)$, respectively, and then there is a positive Radon measure μ on the real line such that [2],

$$\phi(m) = \log \int_{\mathbb{R}} e^{\psi(m)x} \mu(dx), \quad m \in \mathcal{M}.$$

This leads to expressing the NEF $\mathcal{F}$ in terms of the mean as

$$\mathcal{F} = \{P_m(dx) = e^{x\psi(m) - \phi(m)} \mu(dx), \quad m \in \mathcal{M}\}. \quad (3)$$

The representation (3) is called the mean value parametrization of $\mathcal{F}$ [2,4].

The generating kernel is not unique, as any exponential shift generates the same family. This corresponds to the fact that the functions $\psi(m)$ and $\phi(m)$ are not unique as primitives. Indeed, the set of such primitives is infinitely uncountable. We shall detail how to choose appropriate primitives in subsequent work. Parameterization (3) is important for two reasons. One is related to the fact that the parameter m is as meaningful as the mean, and therefore is much more significant than the canonical parameter θ which is just the argument of the corresponding Laplace transform. The second reason is related to situations in which the corresponding Laplace transforms are not explicitly expressed, while primitives ψ and ϕ are easy. Numerous examples of the latter situation are presented in Bar-Lev and Kokonendji [4], Awad et al. [11].

Additionally, our study relates to the latter situation. We consider an NEF given by the variance function $(V, \mathcal{M})$ in which the mean domain $\mathcal{M} = \mathbb{R}^+$, and variance

$$V(m) = m \left(1 + 2m + \frac{1+a^2}{a^2} m^2 \right), \quad m \in \mathcal{M}, \quad (4)$$

where $a > 0$ is a positive parameter. In Section 3, we will derive primitives ϕ and ψ and an associated kernel μ for the NEF represented by the mean value parametrization (3). This NEF is called the large arcsine family in Letac and Mora [2].

2.4. Exponential Dispersion Model

Let $\mathcal{F}$ be an NEF generated by a kernel μ with a variance function $(V, \mathcal{M})$, a Laplace transform $L(\theta)$, and a cumulant transform $k(\theta)$. Consider the set

$$\Lambda = \{p \in \mathbb{R}^+ : L^p \text{ is a Laplace transform of a kernel } \mu_p\}.$$

Then, Λ is nonempty due to convolution. This is called the Jørgensen set (or the dispersion parameter space). It has been shown that $\Lambda = \mathbb{R}^+$ if μ is infinitely divisible (then also all distributions in the family $\mathcal{F}$ are infinitely divisible). For any $p \in \Lambda$, the NEF generated by μ_p is the set of probability measures of the form

$$\mathcal{F}_p = \{P_{\theta,p}(dx) = e^{x\theta - pk(\theta)} \mu_p(dx), \quad \theta \in \Theta\}. \quad (5)$$

The set of NEFs

$$\cup_{p \in \Lambda} \mathcal{F}_p$$

is the EDM associated with μ [12]. The parameter p is called the dispersion parameter. In particular if $\Lambda = \mathbb{R}^+$ (i.e., μ is infinitely divisible) then EDMs are used to describe the distribution

of the error component in generalized linear models (see Nelder and Wedderburn [13], and Jørgensen [3,12] for numerous applications). Note that an EDM is an uncountable set and that the NEF with which we began to construct the EDM is just a special case with a unit dispersion parameter.

Next, let us develop the mean value parameterization of an EDM. The NEF $\mathcal{F}_p$ of (5) has a variance function of $(V_p, \mathcal{M}_p)$ with $\mathcal{M}_p = p\mathcal{M}$. Clearly,

$$\text{if } \mathcal{M} = \mathbb{R}^+ \text{ then } \mathcal{M}_p = p\mathcal{M} = \mathbb{R}^+. \quad (6)$$

The variance satisfies (where $k_p(\theta) = p k(\theta)$),

$$V_p(m) = k_p''((k_p')^{-1}(m)) = p k''((k')^{-1}(m/p)) = pV(m/p). \quad (7)$$

Then, it is easy to see that we can choose primitives $\psi_p(m)$ of $1/V_p(m)$, and $\phi_p(m)$ of $m/V_p(m)$ that satisfy

$$\psi_p(m) = \psi(m/p), \quad \phi_p(m) = p\phi(m/p). \quad (8)$$

Then, we obtain the mean value parameterization corresponding to (5),

$$\mathcal{F}_p = \{P_{m,p}(dx) = e^{x\psi(m/p) - p\phi(m/p)} \mu_p(dx), \quad m \in \mathcal{M}_p\}. \quad (9)$$

Specifically, in our study of the large arcsine family with variance given in (4) for the NEF ($p = 1$), the variance for the EDM (any $p > 0$) becomes

$$V_p(m) = m \left(1 + 2\frac{m}{p} + \frac{1+a^2}{a^2} \frac{m^2}{p^2} \right). \quad (10)$$

This will be the variance function of our study. We denote the associated exponential dispersion model by LAEDM (large arcsine exponential dispersion model).

2.5. Literature Review on Discrete EDMs and Related Distributions

Exponential dispersion models (EDMs) are considered to be powerful tools for statistical analysis because of their modeling flexibility, convolution properties, their usage for generalized linear models [3,12], and their mean value parameterization feature [2,4,5]. The latter enables the modeling of the variances of the distributions as functions of the mean. Jørgensen [12] showed that all polynomial variance functions with non-negative coefficients and a zero constant coefficient correspond to infinitely divisible NEFs. A discrete EDM is an EDM for which its distributions have a discrete domain (not necessarily the integers). The LAEDM is a special case of a discrete exponential dispersion model where its domain is the non-negative integers (denoted $\mathbb{N}_0$) and it has a polynomial variance function of degree 3 which satisfies the infinite divisibility property. As a consequence, its variance function (10) fulfills (6).

In the introductory section, we already reflected on the three other discrete EDMs on $\mathbb{N}_0$ with cubic variance functions identified by Letac and Mora [2] and implemented in Bar-Lev and Ridder [6]. Other simple discrete EDMs on $\mathbb{N}_0$ are the classic Poisson with variance function $V(m) = m$ (degree 1 polynomial), the binomial with $V(m) = m(1 - m/N)$ (degree 2), and the negative binomial with $V(m) = m(1 + m/r)$ (degree 2) identified by Morris [5]. For ease, we mention here the variance function of the default NEF. The general variance function for dispersion p has the form $V_p(m) = pV(m/p)$ (see (7)).

In Bar-Lev and Ridder [8,9], we analyzed discrete EDMs in $\mathbb{N}_0$ with polynomial variance functions in the form $V(m) = m(1 + m)^r$ for any $r = 0, 1, \dots$ and discrete EDMs on $\mathbb{N}_0$ with rational variance functions of the form $V(m) = m/(1 - m)^r$, $r = 1, 2, \dots$. We showed that these models give excellent fitting performances of data showing zero-inflation, overdispersion, and a large amount of skewness and kurtosis.

Kokonendji et al. [14] investigated two discrete EDMs, the first is a class of Poisson mixture with positive Tweedie mixing distributions, hence called a Poisson–Tweedie EDM, which is concentrated on $\mathbb{N}_0$. Its variance function has the form $V(m) = m + m^\gamma \exp((2 - \gamma)\psi(m))$, where $\gamma \geq 1$, and $\psi(m)$ the inverse of the derivative of the cumulant function (see Section 2.3). The second discrete EDM in [14] has the variance function $V(m) = m + m^\gamma$, $\gamma > 1$, and is called Hinde–Demétrio class. If $\gamma = 1, 2, \dots$, it is concentrated on $\mathbb{N}_0$. The probability mass functions of the Poisson–Tweedie EDM and the Hinde–Demétrio EDM are generally not easy, except for in cases with specific parameter values. Kokonendji et al. [14] considered two data sets of car insurance claims [15] and fit Poisson–Tweedie EDM with $\gamma = 2$ (which is negative binomial model) and the Hinde–Demétrio EDM with $\gamma = 3$, which is the strict arcsine model (see also Kokonendji and Khoudar [16]). The generating measure of the Poisson–Tweedie EDM is the Poisson–Tweedie mixture distribution whose probability mass function is more easily computable, and has three parameters that allow for maximum likelihood estimation. Therefore, it has been applied in a wide range of modeling data, such as crash and traffic accident data [17,18], species abundance data [19], and longitudinal RNA-sequencing data [20]. In Section 6, we implement the Poisson–Tweedie mixture distribution as one of the models for fitting count data.

Two of the three parameters of the Poisson–Tweedie mixture model are the dispersion p and the power γ of the Tweedie EDM [12]. Then, the variance of the Poisson–Tweedie mixture distribution is expressed by its mean as $\text{Var} = m + pm^\gamma$. Recently, Abid et al. [21] extended this model to the Poisson–exponential Tweedie model by the relationship $\text{Var} = m + m^2 + pm^\gamma$, and applied this model for fitting overdispersed count data sets. Although they are closely related, note however that these distributions are not members of an EDM.

3. The Large Arcsine Exponential Dispersion Model

This section is dedicated to our main study. We first establish the mean value parametrization of the LAEDM and the choice of appropriate primitives $\psi_p(m)$ and $\phi_p(m)$ for any dispersion parameter p . From these, we argue that the probability distributions $P_{m,p}(dx)$ (see (9)) of the LAEDM are concentrated on the non-negative integers. Therefore, we denote these as the probability mass functions $f_{m,p}(n)$, $n = 0, 1, \dots$. We then present a generating kernel $\mu_p = \{\mu_p(n), n = 0, 1, \dots\}$ from which we obtain exact expressions for the probability mass functions $f_{m,p}(n)$. As these expressions are rather cumbersome, we derive simple bounds of $f_{m,p}(n)$ for any n . These bounds are sharp, as $n \rightarrow \infty$. Furthermore, in the next section, the bounds serve for developing a sampling algorithm based on the accept–reject method.

We end this section by presenting general expressions for the central moments of LAEDM and show that all LAEDM members are skewed to the right and are leptokurtic, i.e., in terms of shape, a leptokurtic distribution has fatter tails.

3.1. The Mean Value Parameterization of the LAEDM

From now on and for convenience, we will use the abbreviation LAEDM, both to indicate the variance function (10) and to its associated EDM. In this section, we will derive the mean value parameterization (9) of the LAEDM. To obtain the mean value parameterization, we start with computing the primitives $\psi(m)$ of $1/V(m)$, and $\phi(m)$ of $m/V(m)$ of the NEF whose variance is given for the unit dispersion in (4).

Lemma 1. Consider the large arcsine NEF with variance function given in (4). Then,

$$\begin{aligned}\int \frac{1}{V(m)} dm &= \log m - \frac{1}{2} \log((1+a^2)m^2 + 2a^2m + a^2) \\ &\quad - a \arctan \frac{(1+a^2)m + a^2}{a} + c, \\ \int \frac{m}{V(m)} dm &= a \arctan \frac{(1+a^2)m + a^2}{a} + d,\end{aligned}$$

where $c, d \in \mathbb{R}$ are integration constants.

Proof. Use partial fraction for

$$\begin{aligned}\frac{1}{V(m)} &= \frac{1}{m(1+2m+\frac{1+a^2}{a^2}m^2)} = \frac{a^2}{m(a^2+2a^2m+(1+a^2)m^2)} \\ &= \frac{1}{m} - \frac{1}{2} \log \frac{2a^2+2(1+a^2)m}{a^2+2a^2m+(1+a^2)m^2} - \frac{a^2}{a^2+2a^2m+(1+a^2)m^2}.\end{aligned}$$

The first two terms have primitives $\log m$ and $\frac{1}{2} \log(1+2m+(1+a^2)m^2)$, respectively. For the third term, apply square completion of the denominator to obtain after rewriting

$$\frac{1+a^2}{1+\left(\frac{(1+a^2)m+a^2}{a}\right)^2},$$

with primitive $a \arctan \frac{(1+a^2)m+a^2}{a}$.

Similarly,

$$\begin{aligned}\frac{m}{V(m)} &= \frac{1}{1+2m+\frac{1+a^2}{a^2}m^2} = \frac{a^2}{a^2+2a^2m+(1+a^2)m^2} \\ &= \frac{1+a^2}{1+\left(\frac{(1+a^2)m+a^2}{a}\right)^2},\end{aligned}$$

with primitive $a \arctan \frac{(1+a^2)m+a^2}{a}$. $\square$

Primitives of the variance function (10) for any dispersion parameter $p > 0$ follow from (8),

$$\begin{cases} \psi_p^{(c)}(m) &= \psi(m/p) = \log m - \frac{1}{2} \log((1+a^2)m^2 + 2a^2mp + a^2p^2) \\ &\quad - a \arctan \frac{(1+a^2)m+a^2p}{ap} + c, \\ \phi_p^{(d)}(m) &= p\phi(m/p) = ap \arctan \frac{(1+a^2)m+a^2p}{ap} + d, \end{cases} \quad (11)$$

where we have added the superscripts c and d to indicate the integration constants that are free parameters at this moment.

Corollary 1. The LAEDM is a class of probability distributions on the non-negative integers.

Proof. Proposition 4.4 of Letac and Mora [2] provides the necessary and sufficient conditions under which a given variance function is associated with an NEF concentrated on the non-negative integers. We check these conditions for any NEF of an LAEDM given by the variance function (10). The first two conditions ($\mathcal{M} = (0, b)$ for some $0 < b \leq \infty$ and $(\phi_p^{(d)})' = m/V_p(m)$ is a real analytic on $\mathcal{M}$) are clearly satisfied. The third condition is

$$\lim_{m \rightarrow 0} (\phi_p^{(d)})'(m) = 1. \quad (12)$$

Because $(\phi_p^{(d)})'(m) = m/V_p(m) = (1 + 2m/p + (1 + a^2)m^2/(ap)^2)^{-1}$, the condition (12) is immediate. $\square$

Now, the question becomes how to choose, appropriately, the integration constants c and d . The most convenient method is to impose [2,8]

$$\lim_{m \rightarrow 0} \phi_p^{(d)}(m) = 0, \quad \text{and} \quad \lim_{m \rightarrow 0} me^{-\psi_p^{(c)}(m)} = 1. \quad (13)$$

We denote the resulting primitives just by $\psi_p(m)$ and $\phi_p(m)$. Then, under these conditions, a generating kernel μ_p can be represented (and computed) by [2,8],

$$\mu_p(n) = \frac{1}{n!} \left(\frac{d}{dm} \right)^{n-1} e^{\phi_p(m)} \phi_p'(m) (me^{-\psi_p(m)})^n \Big|_{m=0}. \quad (14)$$

Proposition 1. *The mean value parameterization of LAEDM that satisfies the conditions (13) is given by primitives*

$$\begin{cases} \psi_p(m) &= \log am p - \frac{1}{2} \log((1 + a^2)m^2 + 2a^2mp + a^2p^2) \\ &\quad - a \left(\arctan \frac{(1+a^2)m+a^2p}{ap} - \arctan a \right), \\ \phi_p(m) &= ap \left(\arctan \frac{(1+a^2)m+a^2p}{ap} - \arctan a \right). \end{cases} \quad (15)$$

Proof. It suffices to determine the integration constants c and d in (11) such that (13) holds. Concerning the $\phi_p^{(d)}$ primitive in (11), we immediately obtain $d = -ap \arctan a$. To compute c , let $\xi_p^{(c)}(m) = \psi_p^{(c)}(m) - \log m$, then

$$me^{-\psi_p^{(c)}(m)} = e^{\log m - \psi_p^{(c)}(m)} = e^{-\xi_p^{(c)}(m)},$$

where (see (11))

$$\xi_p^{(c)}(m) = -\frac{1}{2} \log((1 + a^2)m^2 + 2a^2mp + a^2p^2) - a \arctan \frac{(1 + a^2)m + a^2p}{ap} + c.$$

Thus ,

$$\lim_{m \rightarrow 0} me^{-\psi_p^{(c)}(m)} = 1 \Leftrightarrow \lim_{m \rightarrow 0} \xi_p^{(c)}(m) = 0 \Leftrightarrow c = \log ap + a \arctan a.$$

Now, we substitute the found constants c and d in (11) to obtain (15). $\square$

The final component of the mean value parameterization (9) is the generating kernel $\{\mu_p(n), n = 0, 1, \dots\}$.

Proposition 2. *Consider the discrete measure on the non-negative integers given by*

$$\begin{aligned} \mu_p(2n) &= \frac{p}{p+2n} \frac{1}{(ap)^{2n}} \frac{1}{(2n)!} \prod_{k=0}^{n-1} (a^2(2n+p)^2 + 4k^2), \quad n = 0, 1, \dots, \\ \mu_p(2n+1) &= \frac{ap}{(ap)^{2n+1}} \frac{1}{(2n+1)!} \prod_{k=0}^{n-1} (a^2(2n+1+p)^2 + (2k+1)^2), \quad n = 0, 1, \dots, \end{aligned} \quad (16)$$

(with empty products equal one). This measure generates the mean value parameterization of LAEDM, and serves as a kernel.

Proof. When the primitives ψ_p and ϕ_p satisfy the conditions (13), a kernel can be computed by (14) for $n = 1, 2, \dots$, with $\mu_p(0) = 0$. This has been elaborated in Letac and Mora [2] by applying the Lagrange formula, to become

$$\mu_p(n) = \frac{p}{p+n} \frac{1}{(ap)^n} \frac{1}{n!} \pi_n(a(n+p)), \quad n = 0, 1, \dots,$$

where the polynomials $(\pi_n)_{n=0}^\infty$ are defined by

$$\pi_{2n}(x) = \prod_{k=0}^{n-1} (x^2 + 4k^2), \quad \pi_{2n+1}(x) = x \prod_{k=0}^{n-1} (x^2 + (2k+1)^2).$$

By computation, we can obtain the expressions in (16). $\square$

3.2. Computation of the Probability Mass Functions

Let $\{f_{m,p}(n), n = 0, 1, \dots\}$ be the probability mass function as member of LAEDM with a specific mean $m > 0$ and dispersion $p > 0$. It is represented by the mean value parameterization

$$f_{m,p}(n) = \mu_p(n) e^{n\psi_p(m) - \phi_p(m)}, \quad n = 0, 1, \dots, \quad (17)$$

where the kernel μ_p and the primitives ψ_p and ϕ_p are computed in Section 3.1. We present here their expressions that result after straightforward computations. For completeness, we give the calculus in Appendix A.

Lemma 2. The LAEDM probability mass functions are given by

$$\begin{aligned} f_{m,p}(2n) &= e^{-apB} \frac{p}{p+2n} \frac{1}{C^n} \frac{1}{(2n)!} \prod_{k=0}^{n-1} (a^2(2n+p)^2 + 4k^2), \\ f_{m,p}(2n+1) &= ap e^{-apB} \frac{1}{\sqrt{C}} \frac{1}{C^n} \frac{1}{(2n+1)!} \prod_{k=0}^{n-1} (a^2(2n+1+p)^2 + (2k+1)^2), \end{aligned} \quad (18)$$

where

$$\begin{cases} B = \arctan \frac{(1+a^2)m+a^2p}{ap} - \arctan a, \\ C = (1+a^2)(1+D)e^{2aB}, \\ D = \frac{a^2}{1+a^2} \frac{p}{m} \left(2 + \frac{p}{m}\right). \end{cases} \quad (19)$$

These exact expressions are rather cumbersome, and not very helpful for recognizing structural properties. However, in Appendix B we shall prove the following theorem expressing simple bounds and asymptotics. For these, define

$$\begin{cases} E = \frac{1}{2} \log \left(1 + \frac{1}{a^2}\right) - 1 + a \arctan \frac{1}{a}, \\ \gamma = e^{-apB} \frac{pe^p}{\sqrt{2\pi}}, \\ \rho = \frac{ae^{1+E}}{\sqrt{C}}. \end{cases} \quad (20)$$

All these parameters depend only on the variance function parameters m, p, a .

Proposition 3.

(i) For all $n = 1, 2, \dots$,

$$f_{m,p}(n) \leq \gamma n^{-3/2} \rho^n.$$

(ii) The bound is asymptotically sharp.

$$f_{m,p}(n) = \gamma n^{-3/2} \rho^n + o(1), \quad \text{as } n \rightarrow \infty.$$

(iii) For all $m, p, a > 0, 0 < \rho < 1$.

Hence, we see that the tail of an LAEDM distribution decays to zero faster than the geometric distribution with parameter ρ . Notwithstanding, we call it the geometric parameter of LAEDM.

3.3. Moments, Central Moments, Skewness and Kurtosis

Consider the NEF $\mathcal{F}_p$ of an LAEDM with an arbitrary dispersion parameter $p > 0$ (see (5)). In Section 2.4, we saw that $\mathcal{F}_p$ has the cumulant $k_p(\theta)$ as function of the natural parameter, variance function $V_p(m)$ as function of the mean parameter, and $\psi_p(m)$ as a primitive of $1/V_p(m)$. Then, we can define the derivatives of the cumulant in the mean parameterization by

$$k_p^{(j)}(m) = \left(\frac{d}{d\theta}\right)^j k_p(\theta) \Big|_{\theta=\psi_p(m)}.$$

Recursively, it can easily be shown that

$$k_p^{(j+1)}(m) = V_p(m) (k_p^{(j)})'(m), \quad j = 1, 2, \dots; \quad m \in \mathcal{M}_p. \quad (21)$$

Now, consider the random variable $X_{m,p}$ associated with the probability mass function (17). By definition, its mean is given by $k_p^{(1)}(m) = m$. Its higher central moments are

$$C_p^{(j)}(m) = \mathbb{E}[(X_{m,p} - m)^j], \quad j = 2, 3, \dots$$

Using the recursion (21) for the cumulant derivatives, we can easily obtain that

$$\begin{aligned} C_p^{(2)}(m) &= k_p^{(2)}(m) = V_p(m), \\ C_p^{(3)}(m) &= k_p^{(3)}(m) = V(m)V'(m), \\ &\dots \\ C_p^{(r+2)}(m) &= k_p^{(r+2)}(m) + \sum_{j=2}^r \binom{r+1}{j} C_p^{(j)}(m) k_p^{(r-j+2)}(m), \quad r \geq 2. \end{aligned}$$

The LAEDM shares the following properties.

Proposition 4. *The LAEDM distributions are*

- (a) *Overdispersed (relative to the Poisson NEF).*
- (b) *Zero inflated (relative to the Poisson NEF).*
- (c) *Skewed to the right and leptokurtic.*

Proof.

- (a) This is simple as $V_p(m) > m$.
- (b) Let $X_{m,p} \sim$ be an LAEDM and $Y_m \sim$ be a Poisson, both with mean m . We need to show that

$$\mathbb{P}(X_{m,p} = 0) > \mathbb{P}(Y_m = 0), \quad \text{for all } m, p.$$

Calculate

$$\mathbb{P}(X_{m,p} = 0) = f_{m,p}(0) = \mu_p(0)e^{-\phi_p(m)} = e^{-\phi_p(m)},$$

where by (15),

$$\phi_p(m) = ap \left(\arctan \frac{(1+a^2)m + a^2p}{ap} - \arctan a \right).$$

Because $\mathbb{P}(Y_m = 0) = e^{-m}$, it suffices to show that

$$m > ap \left(\arctan \frac{(1+a^2)m + a^2p}{ap} - \arctan a \right) \quad (22)$$

for all m, a and p . Define

$$h(m) = m - ap \left(\arctan \frac{(1+a^2)m + a^2p}{ap} - \arctan a \right),$$

then

$$h'(m) = m \frac{m + a^2m + 2a^2p}{a^2m^2 + 2a^2mp + a^2p^2 + m^2} > 0,$$

for all $m > 0, a > 0, p > 0$. As $h(0) = 0$, (22) follows.

- (c) Recall that a distribution is skewed to the right if its skewness coefficient $\gamma_1 > 0$. It is called leptokurtic if it has a positive excess kurtosis $\gamma_2 > 0$ (i.e., in terms of shape it has fatter tails). For the LAEDM variance function, we have by (21) and denoting $k_j \equiv k_p^{(j)}(m)$, $C_j \equiv C_p^{(r)}(j)$, that $k_j > 0$ for all $j \geq 1$, implying

$$\begin{aligned} \gamma_1 &= \frac{C_3}{C_2^{3/2}} = \frac{k_3}{k_2^{3/2}} > 0, \\ \gamma_2 &= \frac{C_4}{C_2^2} = \frac{k_4 + 3C_2k_2}{k_2^2} = \frac{k_4 + 3k_2^2}{k_2^2} > 0. \end{aligned}$$

□

4. Monte Carlo Simulation Algorithm for Sampling from LAEDM Distributions

From the upper bound in Proposition 3, we are able to construct a Monte Carlo simulation algorithm for sampling from any LAEDM distribution. The method is based on the accept-reject method [22].

Let X be the random variable on the non-negative integers with probability mass function $f_{m,p}(n)$, $n = 0, 1, \dots$, as given in Lemma 2, and with the upper bound displayed in Proposition 3. We shall consider two random variables, Y and Z , both on the positive integers, as candidates for majorizing $X|X \geq 1$ in the accept-reject sampling method. We define the probability in zero,

$$q = f_{m,p}(0) = e^{-apB}.$$

1. Consider the random variable Y on $\{1, 2, \dots\}$ with probability mass function (or density)

$$g(n) = \frac{\rho^{n-1}}{n^{3/2}} - \frac{\rho^n}{(n+1)^{3/2}}, \quad n = 1, 2, \dots \quad (23)$$

Now, we majorize the density of X given $X \geq 1$:

$$\begin{aligned} f_{m,p}(n|n \geq 1) &= \frac{f_{m,p}(n)}{1-q} \leq \frac{q}{1-q} \frac{pe^p}{\sqrt{2\pi}} \frac{\rho^n}{n^{3/2}} \\ &= \frac{q}{1-q} \frac{pe^p}{\sqrt{2\pi}} \frac{\rho}{1-\rho} \frac{(1-\rho)\rho^{n-1}}{n^{3/2}} \\ &= \frac{q}{1-q} \frac{pe^p}{\sqrt{2\pi}} \frac{\rho}{1-\rho} \left(\frac{\rho^{n-1}}{n^{3/2}} - \frac{\rho^n}{(n+1)^{3/2}} \right) \\ &\leq \frac{q}{1-q} \frac{pe^p}{\sqrt{2\pi}} \frac{\rho}{1-\rho} \left(\frac{\rho^{n-1}}{n^{3/2}} - \frac{\rho^n}{(n+1)^{3/2}} \right). \end{aligned}$$

In this way, we obtain the inequality $f_{m,p}(n|n \geq 1) \leq C_y g(n)$ for all $n = 1, 2, \dots$, with the majorizing constant

$$C_y = \frac{q}{1-q} \frac{pe^p}{\sqrt{2\pi}} \frac{\rho}{1-\rho}. \quad (24)$$

2. Consider the random variable Z on $\{1, 2, \dots\}$ with density

$$h(n) = \frac{1}{\sqrt{n+1}} \left(\sqrt{1 + \frac{1}{n}} - 1 \right), \quad n = 1, 2, \dots \quad (25)$$

This easily shows that

$$\frac{1}{n^{3/2}} \leq \frac{\sqrt{2}}{\sqrt{2}-1} h(n).$$

Because $\rho < 1$, we find

$$\begin{aligned} f_{m,p}(n|n \geq 1) &= \frac{f_{m,p}(n)}{1-q} \leq \frac{q}{1-q} \frac{pe^p}{\sqrt{2\pi}} \frac{\rho^n}{n^{3/2}} \\ &\leq \frac{q}{1-q} \frac{pe^p}{\sqrt{2\pi}} \frac{\sqrt{2}}{\sqrt{2}-1} h(n). \end{aligned}$$

So now, we obtain the inequality $f_{m,p}(n|n \geq 1) \leq C_z h(n)$ for all $n = 1, 2, \dots$, with the majorizing constant

$$C_z = \frac{q}{1-q} \frac{pe^p}{\sqrt{2\pi}} \frac{\sqrt{2}}{\sqrt{2}-1}. \quad (26)$$

Consequently, sampling from an LAEDM distribution can be executed by an accept–reject algorithm using the dominating Y or Z defined above. In fact, we choose the one with the smallest dominating factor C (highest chance of acceptance), thus

$$\begin{aligned} \frac{\rho}{1-\rho} &< \frac{\sqrt{2}}{\sqrt{2}-1} \Rightarrow \text{use } Y; \\ \frac{\rho}{1-\rho} &> \frac{\sqrt{2}}{\sqrt{2}-1} \Rightarrow \text{use } Z. \end{aligned}$$

Hence, given the parameters p, m, a of the LAEDM, we compute the probability in zero q and the majorizing constants C_y, C_z as in (24) and (26). Then, Algorithm 1 summarizes the accept–reject method in the case of sampling using the Z majorant. The case of using the Y majorant is similar.

Algorithm 1 Sampling X from LAEDM Using Z

```

1: Generate  $U \sim \text{uniform}(0, 1)$ 
2: if  $U < q$  then
3:   return 0
4: else
5:   repeat
6:     Generate  $X \sim h(\cdot)$ 
7:     Compute  $P = \frac{f(X)|_{X \geq 1}}{C_z h(X)}$ 
8:     Generate  $U \sim \text{uniform}(0, 1)$ 
9:   until  $U < P$ 
10:  return  $X$ 
11: end if

```

In the following sections, we sketch the process of generating samples of the majorant distributions of Y and Z . As a performance test, we sampled $N = 10000$ samples of the LAEDM distribution with parameters $a = 2.3, p = 1.5$, and $m = 3.7$. The geometric

parameter ρ satisfies $\rho/(1-\rho) > \sqrt{2}/(\sqrt{2}-1)$, and thus the Z majorant was used. The acceptance ratio in Algorithm 1 for these LAEDM parameters was $1/C_Z = 0.31$. The simulated samples were tested (chi-square) against expected numbers. The p -value was 0.46, which verifies that the algorithm was implemented correctly. Figure 1 shows the simulated and expected number of data.

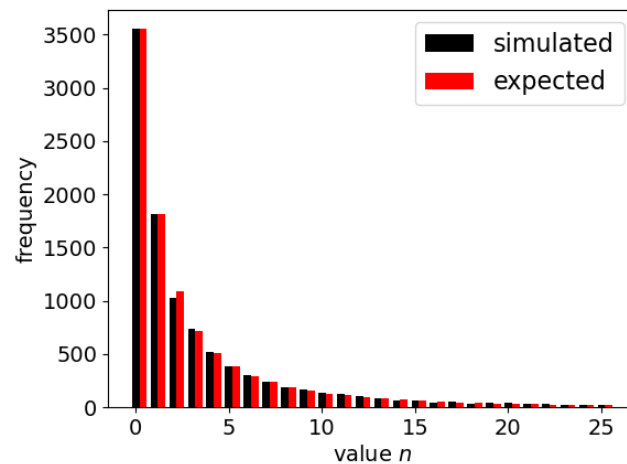


Figure 1. Histogram of the simulated data and their expected numbers.

4.1. Sampling Y

The cumulative distribution function associated with density (23) is

$$G(n) = 1 - \mathbb{P}(Y \geq n+1) = 1 - \frac{\rho^n}{(n+1)^{3/2}}, \quad n = 1, 2, \dots$$

Hence, the inverse transform method applies,

$$Y = \inf\{n = 1, 2, \dots : G(n) \geq U\},$$

where $U \sim \text{uniform}(0, 1)$. Solving for $u \in (0, 1)$:

$$G(n) \geq u \Leftrightarrow \frac{\rho^n}{(n+1)^{3/2}} \leq 1 - u.$$

Thus, let $x > 0$ be the solution to

$$\frac{\rho^x}{(x+1)^{3/2}} = 1 - u,$$

then $Y = \lceil x \rceil$. Now,

$$\begin{aligned} \frac{\rho^x}{(x+1)^{3/2}} &= 1 - u \\ \Leftrightarrow \left(\frac{\rho^{x+1}}{\rho(x+1)^{3/2}} \right)^{2/3} &= (1 - u)^{2/3} \\ \Leftrightarrow (x+1)\rho^{-(2/3)(x+1)} &= (\rho(1 - u))^{-2/3} \\ \Leftrightarrow (x+1)e^{(-(2/3)\log\rho)(x+1)} &= (\rho(1 - u))^{-2/3} \\ \Leftrightarrow (-(2/3)\log\rho)(x+1)e^{(-(2/3)\log\rho)(x+1)} &= (-(2/3)\log\rho)(\rho(1 - u))^{-2/3} \end{aligned}$$

This reads as $w e^w = z$ (for positive z) whose solution is the Lambert W function. Hence,

$$\left(-\frac{2}{3} \log \rho\right)(x+1) = W\left(\left(-\frac{2}{3} \log \rho\right)(\rho(1-u))^{-2/3}\right).$$

Finally, solving for x , we obtain the random sample

$$Y = \left\lceil \frac{3}{-2 \log \rho} W\left(\left(-\frac{2}{3} \log \rho\right)(\rho(1-U))^{-2/3}\right) - 1 \right\rceil$$

4.2. Sampling Z

Recall the density (25) of Z ,

$$h(n) = \frac{1}{\sqrt{n+1}} \left(\sqrt{1 + \frac{1}{n}} - 1 \right), \quad n = 1, 2, \dots$$

It is well known (for instance, see page 550 in Devroye [22]) that such a density is associated with the random variable $\lfloor U^{-2} \rfloor$, where U is the uniform $(0, 1)$ variate. Sampling is easy.

5. Applications of the LAEDM to Risk Measures

Risk measures are statistical indicators that are used by investors, financial institutions, and financial regulators for assessing investment risk. Their main purpose is to determine an amount of capital to keep in reserve in order to cope with risk. Here we consider risk measures for a collective risk model $S_N = \sum_{j=1}^N Y_j$, where

- $Y_1, Y_2, \dots$ are i.i.d. positive random variables representing the individual claims (or losses) at, for instance, an insurance (or financial) company.
- $N \in \mathbb{N}_0 = \{0, 1, \dots\}$ is a random variable designating the total number of claims (or losses) occurring during a certain time period. Commonly it is called the frequency of the claims.
- N and the Y_j s are independent.

Risk measures could be

- Catastrophic risk, or the loss probability [23]

$$\mathbb{P}(S_N > x) \tag{27}$$

for large levels of x . Such a case is very familiar to actuaries since many insurance policies include a deductible and reinsurance contracts which involve some level of retention from the insurer [24].

- The value at risk (VaR) at confidence level $q \in (0, 1)$:

$$\text{VaR}_q = \inf\{x : \mathbb{P}(S_N \leq x) \geq q\}.$$

- The tail conditional expectation at level $q \in (0, 1)$ [25,26]:

$$\text{TCE}_q = \mathbb{E}[S_N | S_N > \text{VaR}_q].$$

This is also known to be the expected shortfall, or the conditional value-at-risk. It is interpreted as the expected worst possible loss, given that this loss exceeds the value at risk. Typically in practice, q is taken to be larger than 0.9.

For a case study, we chose the data from Bar-Lev and Ridder [6] to compute the loss probability (27) in a car insurance company. Typically for automobile insurance, only a small percentage of the policyholders will file claims in any given year. Hence, data sets for insurance risk modeling are highly zero-inflated, and moreover, they involve overdispersed distributions (see Lee [27] and references therein). The data that we considered were taken from a Swedish car insurance company and are publicly available [28,29]. These data sat-

isfy the aforementioned properties (see Table 1 below). From Bar-Lev and Ridder [6], we obtained the frequency distributions of Abel, strict arcsine, and Takács EDMs. These are EDMs on the non-negative integers with cubic variance functions, and their distributions are zero-inflated and overdispersed. We took the claim distributions from gamma and inverse Gaussian distributions and distributions from the natural exponential family (NEF) generated by positive α -stable random variables. The reason for taking these claim distributions is that each forms a natural exponential family also containing their convolutions. Our findings showed that implementing Abel, strict arcsine, or Takács for the frequency with any of the mentioned claim size distributions is significantly superior than the use of the classical Poisson or negative binomial frequency distributions, according to various goodness-of-fit metrics.

Table 1. Statistics of insurance data.

Variable	Zeros p_0	Average m	Variance V
Frequency N	0.06349	70.60	52,181.5
Aggregate S_N		329.2	1,153,532.3
Claim Y		4.663	265.3

As a follow-up to our previous study, we now investigate whether the fourth EDM with cubic variance function, the large arcsine EDM, might be an alternative for modeling the frequency distribution. The data set is the same as the one in Bar-Lev and Ridder [6]: 630 observations (n_i, s_i) where n_i is the i -th frequency (sample of N) and s_i the i -th observation of the aggregated claim (sample of S_N). The statistics of the data are shown in Table 1. Note that the individual claim data (y_{ij}) are not observed, but that their sample average can be computed:

$$\hat{m}_Y = \frac{\sum_i s_i}{\sum_i n_i}.$$

For the sample variance of the individual claims, we use the well-known identity for the variance of the aggregated sum S_N :

$$\text{Var}(S_N) = (\mathbb{E}[N])(\text{Var}(Y)) + (\text{Var}(N))(\mathbb{E}[Y])^2.$$

From Table 1 we see that the claim frequency is overdispersed ($V/m = 52181.5/70.60 \gg 1$) and zero-inflated ($p_0 = 0.06349 \gg 2.189 \times 10^{-31} = e^{-m}$). The LAEDM parameters (a, p, m) are estimated by matching the fraction of zeros p_0 , the mean m , and the variance $V(m)$. The distribution of the claim size Y is chosen from gamma, inverse Gaussian, and NEF positive α -stable distributions. These are two-parameter distributions, estimated by two moment matching.

With these parameters, we have fit the frequency distribution and the claim distribution. Then, we ran simulations of aggregated claim sizes in these models and executed the chi-square test for goodness of it (hypothesizing that the samples came from the same distribution). Furthermore, we show the histograms in Figure 2 and Q–Q plots of the data S_N versus simulated S_N in these three cases in Figure 3.

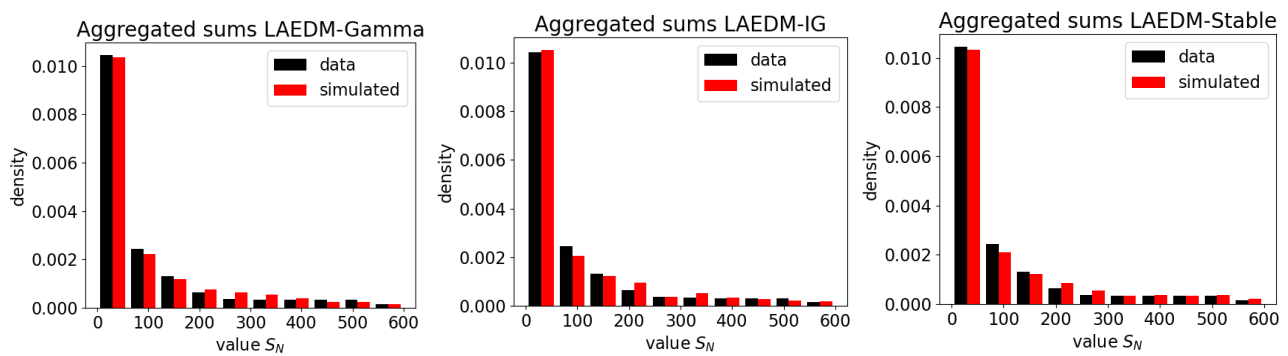


Figure 2. Histograms of the aggregated claim observed data and 2000 simulated samples (normed to form pdf's), using LAEDM frequency.

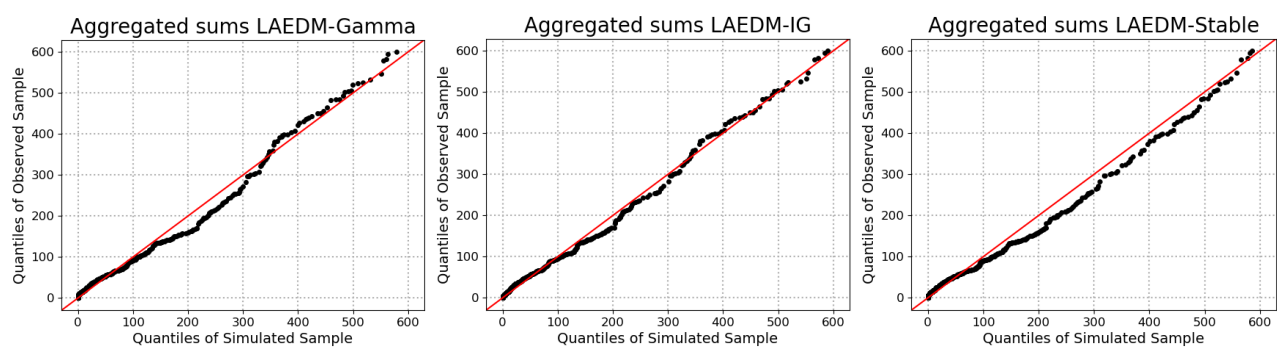


Figure 3. Q–Q plots of the aggregated claim observed data and 2000 simulated samples, using LAEDM frequency.

Table 2 summarizes the test results in terms of p -values.

Table 2. p -values of the fitted models.

Frequency Model	Claim Model	p -Value
LAEDM	Gamma	0.3709
LAEDM	Inverse Gaussian	0.3646
LAEDM	NEF Stable	0.6042

As a comparison with classic modeling, we have fitted the negative binomial (NB) frequency distribution. When combining this with with gamma, inverse Gaussian, and NEF-stable claim distributions, all three combinations gave p -values of the order 10^{-5} of the goodness of fit for the aggregated sum samples. Figure 4 shows the three Q–Q plots. Clearly, the NB distribution does not catch the skewness, and is outperformed by the LAEDM.

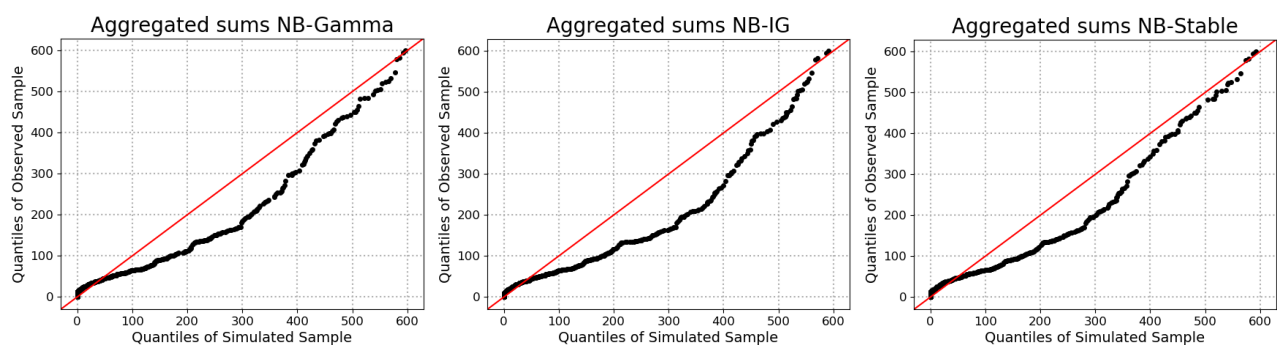


Figure 4. Q–Q plots of the aggregated claim observed data and 2000 simulated samples, using NB frequency.

6. Statistical Modeling

In this section, we will investigate the usage of LAEDM distributions for fitting real-world count data. We consider only data that show the four properties of Proposition 4, i.e., those that are zero-inflated, overdispersed, skewed to the right, and leptokurtic. Many data sets fulfilling these conditions are available in areas such as actuarial sciences, labor economics, health economics, and behavioral science. We choose five data sets with values given in Table 3. Their origins and the model fitting analyses are given below in Section 6.1.

Table 3. Five data sets.

	0	1	2	3	4	5	6	7	8	9	10	11
Set 1	6984	2452	433	100	26	5						
Set 2	29,087	2952	464	108	40	9	5	2	3	1	1	
Set 3	65	14	10	6	4	2	2	2	1	1	1	2
Set 4	3541	599	176	48	20	12	5	1	4			
Set 5	256	54	14	10	1	2						

Their statistics are listed in Table 4, showing the required properties for the zero-inflation ($p_0 > e^{-m}$), the dispersion $\delta = V(m)/m > 1$, the skewness $\gamma_1 > 0$, and the excess kurtosis $\gamma_2 > 0$ (see Section 3.3).

Table 4. The statistics of the data sets.

	p_0	m	e^{-m}	δ	γ_1	γ_2
Set 1	0.6984	0.3747	0.6875	1.127	2.054	5.455
Set 2	0.8903	0.1376	0.8714	1.485	5.119	45.25
Set 3	0.5909	1.391	0.2488	4.394	2.230	4.807
Set 4	0.8037	0.2960	0.7438	1.882	3.963	22.80
Set 5	0.7596	0.3739	0.6881	1.742	2.748	8.786

6.1. Model Fit Analysis

Given a data set, we estimate the mean parameter m of the LAEDM distribution by the data average, and the other two parameters (a and dispersion p) by the maximum likelihood method. Then, we compute the performance of using the LAEDM- (a, p, m) distribution as a model for the data. In addition to the usual measures, Akaike information criterion (AIC) and the χ^2 value (chi sq) with degrees of freedom (df) [30], we also computed the difference between the observed data frequencies and the expected frequencies under the fitted LAEDM distribution by their root mean squared error (RMSE). Finally, we also computed the difference of the empirical (data) distribution and the fitted LAEDM distribution by their Kullback–Leibler divergence (KL).

For assessing the quality of these measures, we also computed these values when using other models for fitting. Because the LAEDM distributions have three parameters, we consider other three-parameter distributions: (a) distributions from the ABM and LM exponential dispersion models. We introduced these models in Bar-Lev and Ridder [8], and analyzed their fits for insurance and crash data in Bar-Lev and Ridder [9]. (b) Poisson–Tweedie distributions (PTD), introduced in Kokonendji et al. [14] and analyzed for fitting purposes in Saha et al. [18]. These distributions are based on Poisson mixing of distributions from the Tweedie exponential dispersion model, and in this way, they are flexible in the sense that they include classical two-parameter distributions such as negative binomial, Poisson-inverse Gaussian, and the geometric-Poisson distributions. (c) Zero-inflated negative binomial distribution (ZINB), which is a traditional model. To illustrate that two-parameter distributions typically perform worse, we included the Poisson-inverse Gaussian (PIG) distribution which has been used for modeling insurance data in Wilmot [15].

data set 1 Automobile claim data from the Central African Republic, 1984 [14,16], presented in Table 5. Surprisingly, all models have difficulty fitting these data except the LAEDM, which, in fact, gave an excellent fit. The statistics of these data show a minor zero-inflation ($p_0 = 0.6984 > 0.6875 = e^{-m}$) and a minor overdispersion ($\delta = 1.127$). We will consider data sets with larger zero-inflation and overdispersion for which other models perform just as well as LAEDM.

Table 5. Performance for fitting models of data set 1.

Model	AIC	chi sq	df	RMSE	KL
LAEDM	15,943.9	0.2596	1	3.004	0.0002498
PTD	16,028.1	140.6	1	72.75	0.004459
ABM	15,965.5	25.65	1	53.15	0.001332
LM	15,963.1	22.45	1	52.50	0.001208
ZINB	15,967.6	28.22	1	53.00	0.001432
PIG	15,961.2	22.70	2	52.64	0.001214

data set 2 Single-vehicle roadway departure fatal crashes on rural two-lane horizontal curves in Texas 2003–2008 [31], presented in Table 6. The best fit to these data was the Poisson–Tweedie distribution, while our LAEDM performs well and much better than the other models. Again, zero-inflation and overdispersion is minor, though larger than in data set 1.

Table 6. Performance for fitting models of data set 2.

Model	AIC	chi sq	df	RMSE	KL
LAEDM	27,062.8	6.183	3	6.177	0.0001844
PTD	27,058.0	2.868	3	4.764	0.0001123
ABM	27,075.7	20.10	3	25.99	0.0003818
LM	27,059.6	4.509	3	10.65	0.0001361
ZINB	27,105.2	58.17	3	38.70	0.0008345
PIG	27,062.6	8.294	4	14.56	0.0002120

data set 3 The counts of cysts of kidneys using steroids in 2010 [32], presented in Table 7. These data show a large zero-inflation ($p_0 = 0.5909 > 0.2488 = e^{-m}$, and large overdispersion (4.394). Here, we see that the traditional zero-inflated negative binomial distribution performs just as well as the Poisson–Tweedie model (and, actually better than the distribution proposed in El-Morshedy et al. [32]). These two are clearly the best. The optimal a parameter of the LAEDM is rather large, making it the same distribution as ABM with a power parameter of two.

Table 7. Performance for fitting models of data set 3.

Model	AIC	chi sq	df	RMSE	KL
LAEDM	343.5	1.900	2	1.713	0.04392
PTD	340.0	0.1371	2	0.6587	0.02827
ABM	343.5	1.900	2	1.713	0.04392
LM	347.6	4.454	2	2.628	0.06257
ZINB	340.1	0.1217	2	0.6482	0.02857
PIG	345.7	4.755	3	2.763	0.06333

data set 4 The length of stays after admission in a USA hospital among the elderly population, aged 65 years or more in 1997 [33], presented in Table 8. The data show moderate zero-inflation and overdispersion, but a large excess kurtosis ($\gamma_2 = 22.80$). All optimal distributions from the exponential dispersion models (LAEDM, PTD, ABM, LM) perform about the same (and about the same as proposed in Bhati and Bakouch [33]). The traditional ZINB and PIG perform worse.

Table 8. Performance for fitting models of data set 4.

Model	AIC	chi sq	df	RMSE	KL
LAEDM	6021.1	3.533	4	5.428	0.001268
PTD	6021.1	3.504	4	5.885	0.001264
ABM	6020.8	3.105	4	5.434	0.001235
LM	6020.9	3.091	4	5.932	0.001240
ZINB	6025.2	8.303	4	7.386	0.001735
PIG	6020.5	4.715	5	9.994	0.001421

data set 5 Falls of older people in a randomized controlled study in Sydney 2007 [34], presented in Table 9. Again, these data have moderate zero-inflation and overdispersion, now also with moderate excess kurtosis. All six models of our comparison analysis perform about the same, including the two-parameter PIG.

Table 9. Performance for fitting models of data set 5.

Model	AIC	chi sq	df	RMSE	KL
LAEDM	542.1	3.383	1	2.094	0.008977
PTD	541.8	2.979	1	2.303	0.008497
ABM	542.1	3.382	1	2.093	0.008977
LM	542.7	3.941	1	2.190	0.009763
ZINB	541.8	3.023	1	2.256	0.008541
PIG	541.1	4.472	2	2.366	0.01043

7. Conclusions

In this study, we gave a comprehensive analysis of a three-parameter distribution (LAEDM) on the non-negative integers as an alternative for classical Poisson, negative binomial, and their zero-inflated (or Hurdle) variations. The main feature of LAEDM is that it is introduced by the variance function of an exponential dispersion model. This leads to an ability to easily derive statistical properties such as zero-inflation, overdispersion, right skewness, and leptokurtic. With these properties, the LAEDM distributions are well suited to fit insurance, accident, or crash data, as these show typically the mentioned statistics. Indeed, we showed that for small sets of count data from insurance and accidents, the fitted LAEDM performs very well and much better than traditional distributional models when zero-inflation and overdispersion is present (but not too much). We also compared the model fitting of the LAEDM distributions with other distributions from exponential dispersion models. The results of this comparison showed that LAEDM performs as good as the other distributions in most cases. An advantage of LAEDM distributions is that they are easy to sample from. This is useful for larger stochastic problems such as insurance risk measures in an insurance risk model, as we showed. In conclusion, we believe that the LAEDM distributions form a useful tool for modeling and analyzing data that show zero-inflation, overdispersion, right skewness, and leptokurtic properties. Future work concerns the use of the LAEDM distributions for analyzing data in the context of a generalized linear model, in which, for instance, the mean is linked to covariates.

Author Contributions: Conceptualization, S.K.B.-L.; methodology, S.K.B.-L. and A.R.; software, A.R.; validation, S.K.B.-L. and A.R.; formal analysis, S.K.B.-L. and A.R.; investigation, S.K.B.-L. and A.R.; resources, S.K.B.-L. and A.R.; data curation, S.K.B.-L. and A.R.; writing—original draft preparation, S.K.B.-L.; writing—review and editing, A.R.; funding acquisition, A.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by STAR (Stochastics—Theoretical and Applied Research) and NWO (Netherlands Organization for Scientific Research) grant number 040.11.711.

Data Availability Statement: The automobile insurance data that we used in Section 5 are publicly available at <http://www.statsci.org/data/>.

Acknowledgments: The authors thank the reviewers for their comments and suggestions to improve our paper.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Proof of Lemma 2

We give here the computations for obtaining the LAEDM probability mass function. Basically, it can be summarized as computing the mean value parameterization for the even and odd terms.

The even terms

$$f_{m,p}(2n) = \mu_p(2n) e^{2n\psi_p(m) - \phi_p(m)},$$

where $\mu_p(2n)$ is given in (16), the primitives $\psi_p(m)$ and $\phi_p(m)$ are given in (15), and the parameters B, C, D are given in (19).

Start with

$$\begin{aligned} e^{2n\psi_p(m)} &= \exp \left[2n \left(\log amp - \frac{1}{2} \log ((1+a^2)m^2 + 2a^2mp + a^2p^2) - aB \right) \right] \\ &= (ap)^{2n} \frac{m^{2n}}{((1+a^2)m^2 + 2a^2mp + a^2p^2)^n} e^{-2naB}. \end{aligned}$$

Go on with

$$\begin{aligned} \frac{m^{2n}}{((1+a^2)m^2 + 2a^2mp + a^2p^2)^n} &= \left(\frac{m^2}{(1+a^2)m^2 + 2a^2mp + a^2p^2} \right)^n \\ &= \frac{1}{\left((1+a^2) + \frac{2a^2p}{m} + \frac{a^2p^2}{m^2} \right)^n} = (1+a^2)^{-n} \left(1 + \frac{2p}{m} \frac{a^2}{1+a^2} + \frac{p^2}{m^2} \frac{a^2}{1+a^2} \right)^{-n} \\ &= (1+a^2)^{-n} \left(1 + \frac{a^2}{1+a^2} \frac{p}{m} \left(2 + \frac{p}{m} \right) \right)^{-n} = (1+a^2)^{-n} (1+D)^{-n}. \end{aligned}$$

Next, note that

$$e^{-\phi_p(m)} = e^{-apB}.$$

Hence, collecting all ingredients,

$$\begin{aligned} f_{m,p}(2n) &= (\mu_p(2n)) \times (e^{2n\psi_p(m)}) \times (e^{-\phi_p(m)}) \\ &= \left(\frac{p}{p+2n} \frac{1}{(ap)^{2n}} \frac{1}{(2n)!} \prod_{k=0}^{n-1} (a^2(2n+p)^2 + 4k^2) \right) \\ &\quad \times \left((ap)^{2n} (1+a^2)^{-n} (1+D)^{-n} e^{-2naB} \right) \times (e^{-apB}) \\ &= e^{-apB} \frac{p}{p+2n} \frac{1}{C^n} \frac{1}{(2n)!} \prod_{k=0}^{n-1} (a^2(2n+p)^2 + 4k^2). \end{aligned}$$

The odd terms

$$\begin{aligned} &e^{(2n+1)\psi_p(m)} \\ &= \exp \left[(2n+1) \left(\log amp - \frac{1}{2} \log ((1+a^2)m^2 + 2a^2mp + a^2p^2) - aB \right) \right] \\ &= (ap)^{2n+1} \frac{m^{2n+1}}{((1+a^2)m^2 + 2a^2mp + a^2p^2)^{(2n+1)/2}} e^{-(2n+1)aB}, \end{aligned}$$

with

$$\frac{m^{2n+1}}{((1+a^2)m^2 + 2a^2mp + a^2p^2)^{(2n+1)/2}} \\ = ((1+a^2)(1+D))^{-n} ((1+a^2)(1+D))^{-1/2}.$$

Thus,

$$f_{m,p}(2n+1) = (\mu_p(2n+1)) \times (e^{(2n+1)\psi_p(m)}) \times (e^{-\phi_p(m)}) \\ = \left(\frac{ap}{(ap)^{2n+1}} \frac{1}{(2n+1)!} \prod_{k=0}^{n-1} (a^2(2n+1+p)^2 + (2k+1)^2) \right) \\ \times \left((ap)^{2n+1} ((1+a^2)(1+D))^{-n} ((1+a^2)(1+D))^{-1/2} e^{-(2n+1)apB} \right) \times (e^{-apB}) \\ = ap e^{-apB} \frac{1}{\sqrt{C}} \frac{1}{C^n} \frac{1}{(2n+1)!} \prod_{k=0}^{n-1} (a^2(2n+1+p)^2 + (2k+1)^2).$$

Appendix B. Proof of Proposition 3

(i) *The even terms*

The following inequalities are elementary analysis.

- By Stirling approximation, for any $n \geq 1$,

$$n! \geq \sqrt{2\pi n} \left(\frac{n}{e}\right)^n.$$

- Differential calculus gives for any $n \geq 1$,

$$\left(\frac{2n+p}{2n}\right)^{2n} = \left(1 + \frac{p}{2n}\right)^{2n} \leq e^p.$$

- Because all parameters are positive,

$$\frac{2}{a(2n+p)} \leq \frac{1}{na}.$$

- By Riemann integration,

$$\int_0^{1/a} \log(1+x^2) dx \geq \sum_{k=0}^{n-1} \frac{1}{na} \log\left(1 + \left(\frac{1}{na}k\right)^2\right) \\ \geq \sum_{k=0}^{n-1} \frac{2}{2na} \log\left(1 + \left(\frac{2}{a(2n+p)}k\right)^2\right),$$

where

$$\int_0^{1/a} \log(1+x^2) dx = \left[x \log(1+x^2) - 2(x - \arctan x) \right]_0^{1/a} \\ = \frac{1}{a} \log\left(1 + \frac{1}{a^2}\right) - \frac{2}{a} + 2 \arctan \frac{1}{a} = \frac{2}{a} E..$$

- Using the previous inequalities,

$$\sum_{k=0}^{n-1} \log\left(1 + \left(\frac{2}{a(2n+p)}k\right)^2\right) \leq \frac{2na}{2} \int_0^{1/a} \log(1+x^2) dx = 2nE.$$

Now, collect all bounds in the expression of the probability $f_{m,p}(2n)$ given in (18), as follows.

$$\begin{aligned}
 f_{m,p}(2n) &= e^{-apB} \frac{p}{p+2n} \frac{1}{C^n} \frac{1}{(2n)!} \prod_{k=0}^{n-1} (a^2(2n+p)^2 + (2k)^2) \\
 &= e^{-apB} \frac{p}{p+2n} \frac{1}{C^n} \frac{1}{(2n)!} a^{2n} (2n+p)^{2n} \prod_{k=0}^{n-1} \left(1 + \left(\frac{2k}{a(2n+p)}\right)^2\right) \\
 &\leq e^{-apB} \frac{p}{p+2n} \frac{1}{C^n} \frac{1}{\sqrt{2\pi(2n)}} \left(\frac{e}{2n}\right)^{2n} a^{2n} (2n+p)^{2n} e^{2nE} \\
 &= e^{-apB} \frac{p}{\sqrt{2\pi}} \frac{2n}{p+2n} \frac{1}{2n\sqrt{2n}} \left(\frac{2n+p}{2n}\right)^{2n} \left(\frac{ae}{\sqrt{C}} e^E\right)^{2n} \\
 &\leq e^{-apB} \frac{pe^p}{\sqrt{2\pi}} \frac{1}{2n\sqrt{2n}} \rho^{2n} = \gamma(2n)^{-3/2} \rho^{2n}.
 \end{aligned}$$

The odd terms

The analysis is similar.

$$\begin{aligned}
 f_{m,p}(2n+1) &= ap e^{-apB} \frac{1}{\sqrt{C}} \frac{1}{C^n} \frac{1}{(2n+1)!} \prod_{k=0}^{n-1} (a^2(2n+1+p)^2 + (2k+1)^2) \\
 &= ap e^{-apB} \frac{1}{\sqrt{C}} \frac{1}{C^n} \frac{1}{(2n+1)!} a^{2n} (2n+1+p)^{2n} \prod_{k=0}^{n-1} \left(1 + \left(\frac{2k+1}{a(2n+1+p)}\right)^2\right) \\
 &\leq p e^{-apB} \frac{1}{\sqrt{C}} \frac{1}{C^n} \frac{1}{\sqrt{2\pi(2n+1)}} \left(\frac{e}{2n+1}\right)^{2n+1} a^{2n+1} (2n+1+p)^{2n} e^{(2n+1)E} \\
 &= e^{-apB} \frac{p}{\sqrt{2\pi}} \frac{1}{(2n+1)\sqrt{2n+1}} \left(\frac{2n+1+p}{2n+1}\right)^{2n} \frac{(ae)^{2n+1} e^{(2n+1)E}}{C^n \sqrt{C}} \\
 &\leq e^{-apB} \frac{pe^p}{\sqrt{2\pi}} \frac{1}{(2n+1)\sqrt{2n+1}} \rho^{2n+1} = \gamma(2n+1)^{-3/2} \rho^{2n+1}.
 \end{aligned}$$

- (ii) All inequalities above become asymptotics as $n \rightarrow \infty$. For instance,

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n, \quad n \rightarrow \infty,$$

and

$$\left(\frac{2n+p}{2n}\right)^{2n} = \left(1 + \frac{p}{2n}\right)^{2n} \sim e^p, \quad n \rightarrow \infty,$$

where $\sim$ indicates that the ratio of the left and right part trends toward 1 as $n \rightarrow \infty$.

- (iii) *The geometric parameter ρ*

Recall

$$\rho = \frac{ae^{1+E}}{\sqrt{C}},$$

where B, C, D , and E are given in (19)–(20).

Clearly, ρ is positive. We prove that $\rho^2 < 1$. Solve

$$\begin{aligned}\rho^2 &= \frac{(ae)^2}{C} e^{2E} = \frac{a^2}{1+a^2} \times \frac{e^2}{1+D} \times e^{2E-2aB} \\ &= \frac{a^2}{1+a^2} \times \frac{e^2}{1 + \frac{a^2}{1+a^2} \frac{p}{m} (2 + \frac{p}{m})} \\ &\quad \times \exp \left[\log \left(1 + \frac{1}{a^2} \right) - 2 + 2a \arctan \frac{1}{a} \right. \\ &\quad \left. - 2a \left(\arctan \frac{(1+a^2)m + a^2p}{ap} - \arctan a \right) \right] \\ &= \frac{a^2}{1+a^2} \times \frac{e^2}{1 + \frac{a^2}{1+a^2} \frac{p(2m+p)}{m^2}} \\ &\quad \times \frac{a^2+1}{a^2} e^{-2} \exp \left[2a \left(\arctan \frac{1}{a} - \arctan \frac{(1+a^2)m + a^2p}{ap} + \arctan a \right) \right] \\ &= \frac{1}{1 + \frac{a^2}{1+a^2} \frac{p(2m+p)}{m^2}} \exp \left[2a \left(\frac{\pi}{2} - \arctan \frac{(1+a^2)m + a^2p}{ap} \right) \right].\end{aligned}$$

Let

$$x = \frac{a^2}{1+a^2} \frac{p}{m}. \quad (\text{A1})$$

We obtain

$$\begin{aligned}\rho^2 &= \frac{1}{1+x(2+p/m)} \exp \left(2a \left(\frac{\pi}{2} - \arctan a(1+1/x) \right) \right) \\ &= \frac{1}{1+x(2+x(1+a^2)/a^2)} \exp \left(2a \left(\frac{\pi}{2} - \arctan a(1+1/x) \right) \right).\end{aligned} \quad (\text{A2})$$

- Fix a and p , then ρ^2 as function of m is increasing. This follows by differentiating:

$$\frac{d}{dm} \rho^2 = \frac{2a^2(1+a^2)x^2}{m(x^2 + a^2(1+x)^2)^2} \exp \left[2a \left(\frac{\pi}{2} - \arctan a(1+1/x) \right) \right] > 0.$$

From (A1), we obtain $\lim_{m \rightarrow \infty} x = 0$, hence,

$$\begin{aligned}\lim_{m \rightarrow \infty} \rho^2 &= \left(\lim_{x \rightarrow 0} \frac{1}{1+x(2+x(1+a^2)/a^2)} \right) \\ &\quad \times \left(\lim_{x \rightarrow 0} \exp \left[2a \left(\frac{\pi}{2} - \arctan a(1+1/x) \right) \right] \right) = 1 \times \exp(0) = 1.\end{aligned}$$

Conclude, for any pair (a, p) , the geometric parameter satisfies $0 < \rho < 1$ for all $m > 0$.

- In the same way, fix a and m , and consider ρ^2 as function of p .

$$\frac{d}{dp} \rho^2 = -\frac{2a^2(1+a^2)x^2}{p(x^2 + a^2(1+x)^2)^2} \exp \left[2a \left(\frac{\pi}{2} - \arctan a(1+1/x) \right) \right] < 0.$$

Thus, ρ^2 is decreasing (as function of p) with (see (A1)–(A2)),

$$\lim_{p \rightarrow 0} x = 0 \Rightarrow \lim_{p \rightarrow 0} \rho^2 = \lim_{x \rightarrow 0} \rho^2 = 1,$$

which gives $0 < \rho < 1$ for all $p > 0$ (and any pair (a, m)).

- Finally, fix m and p and consider ρ^2 as function of a . The derivative with respect to a is a large expression which is negative for all positive a . Thus, ρ^2 is decreasing (as function of a), with (see (A1)–(A2)),

$$\lim_{a \rightarrow 0} x = 0 \Rightarrow \lim_{a \rightarrow 0} \rho^2 = \lim_{x \rightarrow 0} \rho^2 = 1,$$

which gives $0 < \rho < 1$ for all $a > 0$ (and any pair (p, m)).

Conclude that $0 < \rho < 1$ for any triple (a, p, m) of the LAEDM parameters.

References

1. Barndorff-Nielsen, O. *Information and Exponential Families in Statistical Theory*; Wiley: Chichester, UK, 1978.
2. Letac, G.; Mora, M. Natural real exponential families with cubic variance functions. *Ann. Stat.* **1990**, *18*, 1–37. [\[CrossRef\]](#)
3. Jørgensen, B. The Theory of Exponential Dispersion Models. In *Monographs on Statistics and Probability*; Chapman and Hall: London, UK, 1997; Volume 76.
4. Bar-Lev, S.K.; Kokonendji, C.C. On the mean value parameterization of natural exponential families—A Revisited Review. *Math. Methods Stat.* **2017**, *26*, 159–175. [\[CrossRef\]](#)
5. Morris, C.N. Natural exponential families with quadratic variance functions. *Ann. Stat.* **1982**, *10*, 65–80. [\[CrossRef\]](#)
6. Bar-Lev, S.K.; Ridder, A. Monte Carlo methods for insurance risk computation. *Int. J. Stat. Probab.* **2019**, *8*, 54–74. [\[CrossRef\]](#)
7. Fosam, E.B.; Shanbhag, D.N. An extended Laha-Lukacs characterization results based on a regression property. *J. Stat. Plan. Inference* **1997**, *63*, 173–186. [\[CrossRef\]](#)
8. Bar-Lev, S.K.; Ridder, A. New exponential dispersion models for count data—The ABM and LM classes. *ESAIM Probab. Stat.* **2021**, *25*, 31–52. [\[CrossRef\]](#)
9. Bar-Lev, S.K.; Ridder, A. Exponential dispersion models for overdispersed zero-inflated count data. *Commun.-Stat.-Simul. Comput.* **2021**, 1–19. [\[CrossRef\]](#)
10. Jørgensen, B.; Kokonendji, C.C. Discrete dispersion models and their Tweedie asymptotics. *ASTA Adv. Statistical Anal.* **2016**, *100*, 43–78. [\[CrossRef\]](#)
11. Awad, Y.; Bar-Lev, S.K.; Makov, U. A new class of counting distributions embedded in the Lee-Carter model for mortality projections: A Bayesian Approach. *Risks* **2022**, *10*, 111. [\[CrossRef\]](#)
12. Jørgensen, B. Exponential dispersion models (with discussion). *J. R. Stat. Soc. Ser. B* **1987**, *49*, 127–162.
13. Nelder, J.A.; Wedderburn, R.W.M. Generalized linear models. *J. R. Stat. Soc. Ser. A* **1972**, *135*, 370–384. [\[CrossRef\]](#)
14. Kokonendji, C.C.; Dossou-Gbété, S.; Demétrio, C.G.B. Some discrete exponential dispersion models: Poisson-Tweedie and Hinde-Demétrio classes. *Stat. Oper. Res. Trans.* **2004**, *28*, 201–214.
15. Willmot, G. The Poisson-Inverse Gaussian distribution as an alternative to the negative binomial. *Scand. Actuar. J.* **1987**, 3–4, 113–127. [\[CrossRef\]](#)
16. Kokonendji, C.C.; Khoudar, M. On strict arcsine distribution. *Commun. Stat.-Theory Methods* **2004**, *33*, 993–1006. [\[CrossRef\]](#)
17. Debrabant, B.; Halekoh, U.; Bonat, W.H.; Hansen, D.L.; Hjelmberg, J.; Lauritsen, J. Identifying traffic accident black spots with Poisson-Tweedie models. *Accid. Anal. Prev.* **2018**, *111*, 147–154. [\[CrossRef\]](#)
18. Saha, D.; Alluri, P.; Dumbaugh, E.; Gan, A. Application of the Poisson-Tweedie distribution in analyzing crash frequency data. *Accid. Anal. Prev.* **2020**, *137*, 105456. [\[CrossRef\]](#)
19. El-Shaarawi, A.H.; Zhu, R.; Joe, H. Modelling species abundance using th Poisson-Tweedie family. *Environmetrics* **2011**, *22*, 152–164. [\[CrossRef\]](#)
20. Signorelli, M.; Spitali, P.; Tsonaka, R. Poisson-Tweedie mixed-effects model: A flexible approach for the analysis of longitudinal RNA-seq data. *Stat. Model.* **2021**, *21*, 520–545. [\[CrossRef\]](#)
21. Abid, R.; Kokonendji, C.C.; Masmoudi, A. On Poisson-exponential-Tweedie models for ultra-overdispersed count data. *ASTA Adv. Statistical Anal.* **2021**, *105*, 1–23. [\[CrossRef\]](#)
22. Devroye, L. *Non-Uniform Random Variate Generation*; Springer: New York, NY, USA, 1986.
23. Kunreuther, H.; Novemsky, N.; Kahneman, D. Making low probabilities useful. *J. Risk Uncertain.* **2001**, *23*, 103–120. [\[CrossRef\]](#)
24. Kaas, R.; Goovaerts, M.; Dhaene, J.; Denuit, M. *Modern Actuarial Risk Theory*, 2nd ed.; Springer: Heidelberg, Germany, 2008.
25. Furman, E.; Landsman, Z. On some risk-adjusted tail-based premium calculation principles. *J. Actuar. Pract.* **2006**, *13*, 175–190.
26. McNeil, A.J.; Frey, R.; Embrechts, P. *Quantitative Risk Management: Concepts, Techniques and Tools*, Revised ed.; Princeton University Press: Princeton, NJ, USA, 2015.
27. Lee, S. Addressing imbalanced insurance data through zero-inflated Poisson regression boosting. *ASTIN Bull.* **2021**, *51*, 27–55. [\[CrossRef\]](#)
28. Hallin, M.; Ingenbleek, J.-F. The Swedish automobile portfolio in 1977. *Scand. Actuar. J.* **1983**, *1*, 49–64. [\[CrossRef\]](#)
29. Smyth, G.K. Third Party Motor Insurance in Sweden. Australasian Data and Story Library (OzDASL). 2011. Available online: <http://www.statsci.org/data/> (accessed on 17 August 2022).
30. Hilbe, J.M. *Modeling Count Data*; Cambridge University Press: Cambridge, UK, 2014.

-
31. Lord, D.; Geedipally, S.R. The negative binomial-Lindley distribution as a tool for analyzing crash data characterized by a large amount of zeros. *Accid. Anal. Prev.* **2011**, *43*, 1738–1742. [[CrossRef](#)]
 32. El-Morshedy, M.; Eliwa, M.S.; Nagy, H. A new two-parameter exponentiated discrete Lindley distribution: Properties, estimation and applications. *J. Appl. Stat.* **2020**, *47*, 354–375. [[CrossRef](#)]
 33. Bhati, D.; Bakouch, H.S. A new infinitely divisible discrete distribution with applications to count data modeling. *Commun. Stat.-Theory Methods* **2019**, *48*, 1401–1416. [[CrossRef](#)]
 34. Ullah, S.; Finch, C.F.; Day, L. Statistical modelling for fall count data. *Accid. Anal. Prev.* **2010**, *42*, 384–392. [[CrossRef](#)]