Annals of Actuarial Science (2023), **17**, pp. 145–169
doi:10.1017/S1748499522000148

ORIGINAL RESEARCH PAPER

Panjer class revisited: one formula for the distributions of the Panjer (a,b,n) class

Michael Fackler

Actuary (DAV), Munich, Germany
E-mail: michael_fackler@web.de

(Received 07 May 2022; revised 10 August 2022; accepted 26 August 2022; first published online 26 September 2022)

Abstract

The loss count distributions whose probabilities ultimately satisfy Panjer's recursion were classified between 1981 and 2002; they split into six types, which look quite diverse. Yet, the distributions are closely related – we show that their probabilities emerge out of one formula: the binomial series. We propose a parameter change that leads to a unified, practical and intuitive, representation of the Panjer distributions and their parameter space. We determine the subsets of the parameter space where the probabilities are continuous functions of the parameters. Finally, we give an inventory of parameterisations used for Panjer distributions.

Keywords: Panjer class; Unified representation; Discrete loss distribution; Binomial series

1. Introduction

1.1 Motivation

The loss count distributions of the Panjer class are popular in insurance and beyond, in particular the classical models Poisson, binomial and negative binomial (NB). Apart from being realistic and intuitive, they are easy to handle computationally as their probabilities satisfy a recursive formula. For more flexibility, it was successfully tried to extend the class of distributions, while preserving most of their convenient properties. The above three models are the only nontrivial ones fulfilling Panjer's recursion thoroughly (Sundt & Jewell, 1981), but two models satisfy it for all probabilities but the lowest step (p_0 to p_1), namely the logarithmic distribution and the so-called Engen or extended truncated negative binomial distribution (ETNB) (Willmot, 1988). Finally, Hess *et al.* (2002) classified the discrete loss distributions fulfilling the recursion for all but a finite number of initial probabilities, by extending the NB distribution further and adding an extension of the logarithmic distribution. Overall, the resulting class embraces six types of distributions which, however, look quite diverse. So, the picture is complete, but a bit heterogeneous.

1.2 Research context

In the actuarial literature, extensions of the Panjer class beyond the three classical models have been investigated since the 1980s, when Panjer's famous recursive algorithm to calculate the aggregate loss distribution in the collective model emerged (Panjer, 1981). It was successfully tried to generalise the recursive formula for the probabilities in a number of ways, which leads to a huge variety of distributions, see, for example, Sundt & Jewell (1981), Willmot and Panjer (1987), Schröter (1990), Gerhold *et al.* (2010), the survey paper Sundt (2002), and the textbooks

Panjer & Willmot (1992), Klugman *et al.* (2008). However, as for example, Albrecher *et al.* (2017) note in a compact review of the Panjer class, some of the findings are older and date back at least to Johnson & Kotz (1969).

We focus on the particular extension of the Panjer class that keeps the original recursive formula for the probabilities but does not require it for some initial probabilities. We call this the general Panjer class; it combines a *large flexibility* for the initial probabilities with a distribution tail geometry *somewhat extending* that of the three classical distributions. The papers being most relevant for this class were already mentioned: Willmot (1988) and Hess *et al.* (2002) extended the classical three Panjer models to the final total of six distribution types, and Gerhold *et al.* (2010) developed a generalised and numerically superior recursive Panjer algorithm.

1.3 Objective

In this paper, we complement the theory by showing how the probabilities of the general Panjer class emerge out of one formula, namely the binomial series. This enhances and extends unified views on parts of the Panjer class as given by Panjer & Willmot (1992) and Fackler (2011). Beyond being instructive, the resulting representation of the general Panjer class of distributions can ease implementation and use of the models in practice, providing all-in-one formulae for probabilities and moments, expressed in terms of the parameters of Panjer's recursion. A slight transform of the traditional recursion parameters simplifies both the formulae (slightly) and the geometry of the parameter space (greatly).

1.4 Outline

In section 2, we define the general Panjer class of distributions, reconciling diverse definitions appearing in the literature, and discuss useful parameterisations. In section 3, we rearrange the binomial series such that it yields the probabilities of half of the general Panjer class, based on a common parameter space. In section 4, we study some limiting cases, which yield the other half of the class, and determine where the probabilities are continuous functions of their parameters, and where not. Section 5 wraps up the classification; section 6 comments on parameter inference. The appendix provides some technical details and an inventory of parameterisations being used for members of the Panjer class.

2. Preliminaries

For the sake of precision, we must go through some technicalities; however, intuition will be provided on the way. Let us first collect and reconcile definitions of the Panjer class(es) found in the literature, adding an alternative parameter s.

2.1 Definitions

Definition 2.1. For a positive integer k and real $a, b, s = a + b$, *Panjer's recursion* is as follows:

$$p_k = \left(a + \frac{b}{k}\right) p_{k-1} = \left\{a\left(1 - \frac{1}{k}\right) + s\frac{1}{k}\right\} p_{k-1} \tag{1}$$

For any $n \in \mathbb{N}_0$, the *wide Panjer* (a, b, n) *class* is the class of nontrivial counting distributions whose probabilities p_k satisfy Panjer's recursion for some a, b and all integers $k > n$.

The *proper Panjer* (a, b, n) *class* is the subclass thereof where in addition $p_k = 0$ for all integers $k < n$.

The *narrow Panjer* (a, b, n) *class* is the subclass thereof whose distributions cannot be written as a left-truncation of some other distribution of some (a, b, m) class.

Table 1. Narrow Panjer class.

Name	$p_k = P(N = k)$	Parameter space	Support
Binomial	$\binom{n_0}{k} p^k (1-p)^{n_0-k}$	$p \in (0, 1)$	$\{0, 1, ..., n_0\}$
Poisson	$e^{-\lambda} \frac{\lambda^k}{k!}$		$\{0, 1, ...\}$
Negative binomial	$\binom{\alpha+k-1}{k} (1-a)^{\alpha} a^k$	$a \in (0, 1)$	$\{0, 1, ...\}$
Logarithmic	$\frac{1}{-\ln(1-a)} \frac{a^k}{k}$	$a \in (0, 1)$	$\{1, 2, ...\}$
Extended negative binomial	$\frac{\binom{\alpha+k-1}{k} a^k}{(1-a)^{-\alpha} - \sum_{j=0}^{m_0-1} \binom{\alpha+j-1}{j} a^j}$	$a \in (0, 1], \alpha \in (-m_0, 1-m_0)$	$\{m_0, m_0+1, ...\}$
Extended logarithmic	$\frac{a^k / \binom{k}{m_0}}{\sum_{j=m_0}^{\infty} a^j / \binom{j}{m_0}}$	$a \in (0, 1], m_0 > 1$	$\{m_0, m_0+1, ...\}$

Nontrivial means not (almost surely) constant, such that at least two p_k are positive. *Left-truncation* (of discrete loss distributions) means setting some initial probabilities to 0 and rescaling the remaining probabilities accordingly, which yields again a distribution. Truncation to the left of p_m formally means left-truncation at $m-1$; following Hess *et al.* (2002) we call this *m-truncation*:

$$p_k(m) := \begin{cases} 0, & k < m \\ \frac{p_k}{1-\sum_{j=0}^{m-1} p_j}, & k \geq m \end{cases} \tag{2}$$

Instead of the (a, b, n) class one also speaks of the *Panjer distributions of order n*. The *general Panjer class* is the union of the Panjer classes of any order, and likewise has a *wide*, a *proper* and a *narrow* variant.

For the general concept of truncation see Klugman *et al.* (2008), who define the *wide* Panjer distributions (of order $n = 0, 1$). The wide classes are totally ordered: for $n < m$ the (a, b, n) class is a subset of the (a, b, m) class.

Panjer & Willmot (1992) (for $n = 0, 1$) and Hess *et al.* (2002) define the (a, b, n) class in the *proper* sense. Proper Panjer classes of different order are disjoint. If we m-truncate a wide (a, b, m) model, we get the (unique) proper model having the same parameters a, b, m. The original wide model is a mixture of this corresponding proper model with a loss count model having support in $\{0, ..., m-1\}$. In other words, in a wide (a, b, m) model the initial probabilities $p_0, ..., p_{m-1}$ can be chosen (thus are parameters) with the only restriction that their sum must be less than 1. The recursion determines the tail starting from p_m. When Panjer class distributions are analysed, it is sufficient (and common) to look at the proper models; the wider ones are embraced by choosing the initial probabilities as described here.

The distributions of the general *narrow* Panjer class are specified by Hess *et al.* (2002), who call them *basic* claim number distributions and distinguish six types: Poisson, binomial, NB, logarithmic, extended NB and extended logarithmic, see Table 1 with integers n_0, m_0 and further parameters being positive real unless specified otherwise.

If for $n \leq m$ one starts with a narrow (a, b, n) model and m-truncates it, one gets a proper (a, b, m) distribution. All proper distributions can be generated this way, so it is ultimately sufficient to classify narrow Panjer models and generate proper and wide models having the same parameters a, b by changing the first probabilities and rescaling the tail.

Finally, note that the three variants of the Panjer $(a, b, 0)$ class coincide.

https://doi.org/10.1017/S1748499522000148 Published online by Cambridge University Press

2.2 Practical use

Wide Panjer models are useful in situations where one considers the tails of Panjer distributions as realistic but needs much more flexibility for the initial probabilities. In particular, flexibility about p_0 often helps fitting real insurance data, for a discussion see section 6.7 of Klugman *et al.* (2008), who speak of *zero-modification* and *zero-truncation* (which latter is notably 1-truncation in our terminology). We will take a closer look at this combination of “free” initial probabilities and subsequent Panjer tail in section 6.

Proper and narrow models can be adequate for situations where the lowest loss counts are impossible, for example, when modelling the number of risks contributing to an accumulation loss. Reinsurance treaties protecting such losses usually have a *n risks warranty*, that is, losses are only covered when at least n risks are affected. $n=2$ is most common, but in life or personal accident accumulation covers one finds, for example, $n=6$.

Panjer’s classical recursive algorithm to calculate the aggregate loss distribution in the collective model can be extended to all proper Panjer loss count distributions (Hess *et al.*, 2002), so they all share this important ease of calculation that made the classical Panjer distributions so popular in insurance. Gerhold *et al.* (2010) found that Panjer’s algorithm can be numerically unstable with the two “Extended” distribution types and developed (see their section 5) a slower but stable and far more general algorithm, which embraces further loss count distributions, including in particular the wide Panjer class (see their Corollary 4.7).

However, for the latter class one can often alternatively work with the classical algorithm. To see this, let us call the loss count N, the loss severity X and the aggregate loss S. In a wide model of order m, the initial probabilities $p_0, \dots, p_{m-1}$ sum up to $q<1$ and we can write the cdf of S as

$$F_S(x) = \sum_{k=0}^{\infty} p_k F_X^{k*}(x) = \sum_{k=0}^{m-1} p_k F_X^{k*}(x) + \left(1-q\right) F_{\bar{S}}(x)$$

where $\bar{S}$ is the aggregate loss of the collective model with the same severity X but a thinned loss count $\bar{N}$, namely the m-truncation of N with probabilities according to Formula 2. So, F_S can be calculated from the convolutions of F_X up to order $m-1$, plus $F_{\bar{S}}$ resulting from the Panjer algorithm for proper models.

2.3 Consolidation

We will see below that in a narrow $\left(a, b, n\right)$ model the integer n is uniquely determined by a and b and is thus not a parameter. It will further turn out that many (but not all) formulae become simpler if we use the parameter $s=a+b$ instead of b. Both parameterisations are equivalent and easy to convert into each other, so one can switch between them according to convenience.

Lemma 2.2. *For any real a,b (or equivalently a, s) the sequence*

$$r_k := \frac{1}{k!}\prod_{i=1}^{k}\left(b+ai\right) = \frac{1}{k!}\prod_{i=0}^{k-1}\left(s+ai\right), \quad k \in \mathbb{N}_0 \tag{3}$$

(where $r_0 = 1$) satisfies Panjer’s recursion.

Proof. For $k>0$, we have

$$\left(a+\frac{b}{k}\right) r_{k-1} = \frac{b+ak}{k}\frac{1}{\left(k-1\right)!}\prod_{i=1}^{k-1}\left(b+ai\right) = \frac{1}{k!}\prod_{i=1}^{k}\left(b+ai\right) = r_k$$

□

https://doi.org/10.1017/S1748499522000148 Published online by Cambridge University Press

Note that we did not require the r_k to be probabilities, not even up to a factor. They may partly equal 0 or have varying sign. Yet, Panjer's recursion holds for all $k > 0$. So, if the r_k have ultimately the same sign, one can possibly use the (rescaled) tail of the sequence as loss count probabilities, which would yield a proper Panjer distribution. This is indeed the key idea of this paper and will lead to a common representation of the key figures of the proper Panjer distributions.

Definition 2.3. We call (r_k) as in Formula 3 the *Panjer sequence*. Where we want to emphasise that r_k is a function of a and b, we write $r_k\left(a, b\right)$; with alternative parameters we write analogously $r_k\left(a, s\right)$, etc.

Before using the Panjer sequence in a general setting, we relate it to the classical Panjer distributions, whose probabilities and recursion parameters can be written in all-in-one formulae (Fackler, 2011):

$$p_k = \left(1 + \frac{\lambda}{\alpha}\right)^{-\alpha} \frac{\lambda^k}{k!} \prod_{i=0}^{k-1} \frac{\alpha + i}{\alpha + \lambda}, \qquad a = \frac{\lambda}{\alpha + \lambda}, \quad s = \frac{\alpha\lambda}{\alpha + \lambda} \tag{4}$$

Here $\lambda > 0$ is the expectation. $\alpha > 0$ yields the NB distribution, while $\alpha \in -\mathbb{N}$ satisfying $-\alpha > \lambda$ yields the binomial distribution. The limits for $\alpha \to \pm\infty$ are well defined and both yield the Poisson distribution. To avoid infinite parameters, one can equivalently work with λ and $c = \frac{1}{\alpha}$; however, here the parameter space is not so easy to write down (see Appendix B for this and alternative parameterisations). It gets even more intricate if we rewrite the probabilities in terms of a and s, but this is the variant that we need for the general treatment of the Panjer class. We have $\alpha = \frac{s}{a}$ and $\lambda = \frac{s}{1-a}$. Some algebra yields

$$p_k = (1-a)^{\frac{s}{a}} \frac{1}{k!} \prod_{i=0}^{k-1} (s + ai) = (1-a)^{\frac{s}{a}} r_k\left(a, s\right) \tag{5}$$

Thus, for appropriate a, s (which we will specify below), the Panjer sequence generates the probability function of the Panjer $\left(a, b, 0\right)$ class. Poisson corresponds to the (well-defined) limit for $a \to 0$, where $s = \lambda$.

Let us introduce a third parameterisation, which only embraces binomial and NB but will be useful below. Using parameters $a \neq 0$ and $\alpha = \frac{s}{a}$, it is a hybrid of the two preceding variants:

$$p_k = (1-a)^{\alpha} r_k\left(a, \alpha\right), \qquad r_k\left(a, \alpha\right) = \frac{a^k}{k!} \prod_{i=0}^{k-1} (\alpha + i) = a^k \binom{\alpha + k - 1}{k} \tag{6}$$

3. Rearranging the Binomial Series

It is well known that for certain real x, γ the Taylor series

$$(1+x)^{\gamma} = \sum_{k=0}^{\infty} \binom{\gamma}{k} x^k = \sum_{k=0}^{\infty} \frac{x^k}{k!} \prod_{i=0}^{k-1} (\gamma - i) \tag{7}$$

converges (absolutely), in particular in the following situations:

- $\gamma \in \mathbb{N}$,
- $|x| < 1$,
- $|x| = 1$ and $\gamma > 0$.

Formula 7 is called the *binomial series* or *generalised binomial formula*.

https://doi.org/10.1017/S1748499522000148 Published online by Cambridge University Press

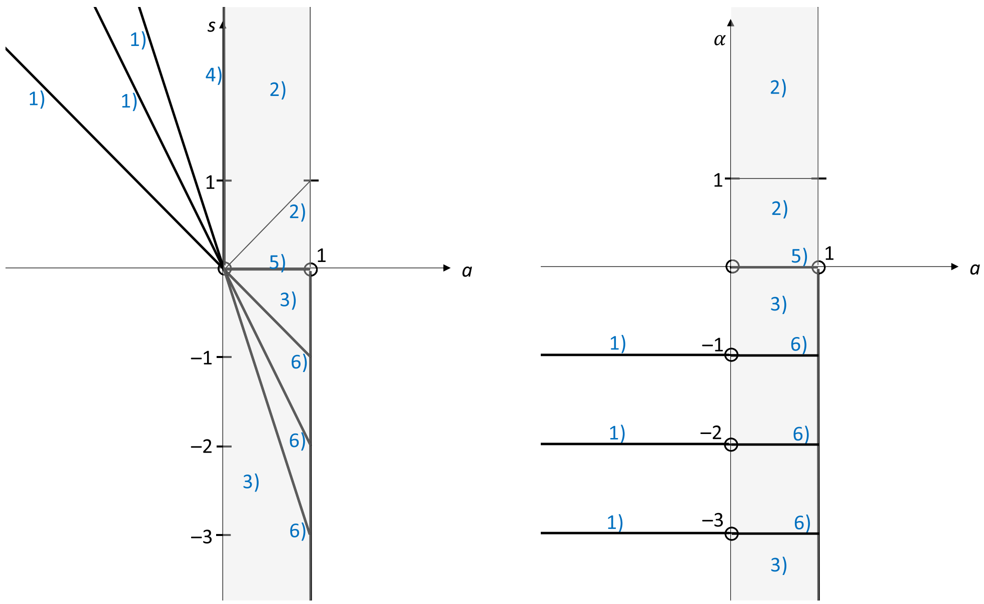

Figure 1. Admissible parameters a, s and a, α.

Now choose real $a \neq 0$ and s; set $x = -a$, $\gamma = -\frac{s}{a}$. If the latter two figures are in the appropriate range, we can rewrite the binomial series in terms of a and s as

$$(1-a)^{-\frac{s}{a}} = \sum_{k=0}^{\infty} \frac{(-a)^k}{k!} \prod_{i=0}^{k-1} \left(-\frac{s}{a} - i\right) = \sum_{k=0}^{\infty} \frac{1}{k!} \prod_{i=0}^{k-1} (s + ai) = \sum_{k=0}^{\infty} r_k(a, s) \tag{8}$$

and have found the (well-defined and finite) sum of the Panjer sequence. In particular, this applies to the following cases, which are special cases of the above three situations, but disjoint and grouped in a different way:

Definition 3.1. We call real *a, s strongly admissible* if

$$(a, s) \in \mathcal{S} := \{(a, -an_0) \mid a \in (-\infty, 0), n_0 \in \mathbb{N}\} \cup (0,1) \times (0, \infty) \cup (0,1] \times ((-\infty, 0) \backslash -\mathbb{N}) \subset \mathbb{R}^2$$

that is, if they lie in one of the following three disjoint *areas* (see also Figure 1, left side):

Area 1) $a < 0$, $s > 0$, $-\frac{s}{a} =: n_0 \in \mathbb{N}$;
2) $0 < a < 1$, $s > 0$;
3) $0 < a \leq 1$, $s < 0$, $-\frac{s}{a} \notin \mathbb{N}$

The *delay* of $(a, s) \in \mathcal{S}$ is the integer

$$m_0(a, s) := \begin{cases} 0, & s > 0 \\ 1 + \left[-\frac{s}{a}\right], & s < 0 \end{cases}$$

where $[x]$ is the largest integer not exceeding x, such that for $s < 0$ we have $0 \leq m_0 - 1 < -\frac{s}{a} < m_0$.

Finally, for $n \in \mathbb{N}_0$ we define the subsets

$$\mathcal{S}_n := \{(a, s) \in \mathcal{S} \mid m_0(a, s) \leq n\}, \quad \mathcal{S}_{-1} := \emptyset, \quad \mathcal{S}_{(n)} := \{(a, s) \in \mathcal{S} \mid m_0(a, s) = n\} = \mathcal{S}_n \backslash \mathcal{S}_{n-1}$$

As we will see, these subsets structure a large parameter space. For orientation, in Figure 1, left side, the $\mathcal{S}_{(n)}$, $n \geq 1$, are the triangles below the a-axis; they include their respective right border and exclude the other borders. The latter borders are where the step function $m_0(a, s)$ changes

https://doi.org/10.1017/S1748499522000148 Published online by Cambridge University Press

value. $\mathcal{S}_0 = \mathcal{S}_{(0)}$ is the union of Areas 1 and 2 above the a-axis, while $\mathcal{S}_n$, $n \geq 1$, unites these areas with the first n triangles below the a-axis.

Lemma 3.2. *For strongly admissible a, s, the RHS of the formula with corresponding delay $m_0 = m_0(a, s)$,*

$$(1-a)^{-\frac{s}{a}} - \sum_{j=0}^{m_0-1} r_j(a, s) = \sum_{k=m_0}^{\infty} r_k(a, s) \tag{9}$$

converges and has either positive or negative summands, except that in Area 1 we ultimately, namely for all $k > n_0$, have $r_k(a, s) = 0$.

Proof. Formula 9 is the rearranged (and convergent) binomial series as given in Formula 8. The summands r_k, $k \geq 0$, are, up to a positive factor, products of the factors $(s + ai)$, where $i = 0, ..., k-1$.

In Area 1, these factors are initially positive, but $(s + an_0) = 0$. Thus, $r_0, ..., r_{n_0}$ are positive, while the further r_k equal 0.

In Area 2, both a and s are positive, such that all $(s + ai)$ and all r_k are positive.

In Area 3, we have $s + a(m_0 - 1) < 0 < s + am_0$. So, the $(s + ai)$ are negative for $i < m_0$ and positive for $i \geq m_0$. Thus, all r_k with $k \geq m_0$ have the same sign, either positive or negative. □

Definition 3.3. For any $k \in \mathbb{N}_0$ and strongly admissible *a, s* with delay $m_0 = m_0(a, s)$, we set

$$p_k(a, s) := \begin{cases} 0, & k < m_0 \\ \frac{r_k(a,s)}{(1-a)^{-s/a} - \sum_{j=0}^{m_0-1} r_j(a,s)}, & k \geq m_0 \end{cases} \tag{10}$$

For any $m \in \mathbb{N}_0$, where $m < n_0$ in Area 1, we set $M(a, s; m) := \max(m, m_0(a, s))$ and

$$p_k(a, s; m) := \begin{cases} 0, & k < M \\ \frac{r_k(a,s)}{(1-a)^{-s/a} - \sum_{j=0}^{M-1} r_j(a,s)}, & k \geq M \end{cases} \tag{11}$$

Proposition 3.4. *The $(p_k(a, s))$ and the $(p_k(a, s; m))$ of Definition 3.3 constitute nontrivial loss count distribution models having (for fixed m) parameter space $\mathcal{S}$. The latter model is the m-truncation of the former. The delay $m_0(a, s)$ indicates the first non-zero probability of $(p_k(a, s))$, while $M(a, s; m)$ is the corresponding index for $(p_k(a, s; m))$. Both distributions satisfy Panjer's recursion with parameters a, s for all but the initial zero probabilities, that is, for $k > m_0$ and $k > M$, respectively.*

Proof. According to the preceding lemma, all $p_k(a, s)$, $k \geq m_0$, have the same sign (apart from possibly ultimately equaling 0) and sum up to 1, so they constitute a discrete loss distribution, which is nontrivial as at least two $p_k(a, s)$ are positive. Obviously, $(p_k(a, s; m))$ is the m-truncation of $(p_k(a, s))$ and in particular a distribution. The constraint $m < n_0$ for Area 1 ensures that it is not the trivial distribution being concentrated at $n_0 \in \mathbb{N}$. Both distributions inherit Panjer's recursion from the $r_k(a, s)$. □

Corollary 3.5. *For $m \leq m_0(a, s)$, that is, $M = m_0$, m-truncations coincide with the original model. For $m = M > m_0$, $(p_k(a, s; m))$ is different from $(p_k(a, s))$ and thus a proper Panjer distribution that is not narrow. Thus, $(p_k(a, s))$ is a narrow Panjer distribution of order m_0, while $(p_k(a, s; m))$, for $m > m_0$, is a proper (but not narrow) distribution of order m.*

The parameters of narrow distributions $\left(p_k(a,s)\right)$ with delay $m_0 \in \mathbb{N}_0$ lie in $\mathcal{S}_{(m_0)}$, while the parameters a, s of proper distributions $\left(p_k(a,s;m)\right)$ with $m \ge m_0(a,s)$ lie in $\mathcal{S}_m$.

Proof. The assertions follow immediately from the construction of the $p_k(a,s)$ and $p_k(a,s;m)$, and from the definition of the parameter sets $\mathcal{S}_n$ and $\mathcal{S}_{(n)}$. □

Example 3.6. For $s>0$ (and $m_0=0$), Formula 10 defines the combined *binomial* (Area 1)/*NB* (Area 2) distribution as parameterised in Formula 5. If $m>0$, Formula 11 yields respective m-truncations of the binomial and NB distributions that are proper, but not narrow, $\left(a,b,m\right)$ models.

Let us now rewrite Formula 11, which for $m=m_0$ embraces Formula 10, in terms of the parameters a and $\alpha = \frac{s}{a}$ used in Formula 6. As $a \neq 0$ for all strongly admissible a, s, this is possible and an equivalent parameterisation.

$$p_k(a,\alpha;m) := \frac{r_k(a,\alpha)}{(1-a)^{-\alpha} - \sum_{j=0}^{M-1} r_k(a,\alpha)} = \frac{a^k\binom{\alpha+k-1}{k}}{(1-a)^{-\alpha} - \sum_{j=0}^{M-1} a^j\binom{\alpha+j-1}{j}}, \quad k \ge M \qquad (12)$$

Area 3 in this parameterisation means $0<a\le 1$, $\alpha<0$, $\alpha \notin -\mathbb{N}$, which is a union of squares, see Figure 1, right side. In particular, we have $0 \le m_0 - 1 < -\alpha < m_0$.

Example 3.7. If $m=m_0$, Formula 12 with parameters (a,α) in Area 3 parameterises the *extended negative binomial* (ENB) distribution as defined by Hess *et al.* (2002), a further of their six types of narrow $\left(a,b,n\right)$ models. $M=m>m_0$ yields m-truncations of ENB that are proper, but not narrow, $\left(a,b,m\right)$ models.

The special case $m=m_0=1$, $-1<\alpha<0$ (first square below the a-axis) has been known for long and was termed ETNB distribution (Willmot 1988). However, like Hess *et al.* (2002), we prefer the name ENB and confirm their reasoning from our perspective: ETNB has $p_0=0$, but it is no 1-truncation of any Panjer distribution – it is a narrow $\left(a,b,1\right)$ model. While one can indeed interpret ETNB as an extension of 1-truncated NB (which we will generalise below), an alternative perspective fits better to the unified view we are developing here: via Formula 10, ETNB and the rest of ENB ($\alpha<-1$, that is, the further squares below the a-axis, or equivalently triangles in terms of a, s) are an extension of the $\left(a,b,0\right)$ class; focusing on $a>0$ (Areas 2, 3), one can see ENB as an extension of the *non-truncated* NB model.

4. Embracing Limiting Cases

Definition 3.3 yielded three of the six narrow $\left(a,b,n\right)$ distributions, plus the corresponding proper distributions emerging via m-truncation. Now we show that it yields the other three types too, as limiting cases at (part of) the border of the parameter space. To this end, let us complement Definition 3.1.

Definition 4.1. We call real *a, s admissible* if

$$(a,s) \in \mathcal{A} := \{(a,-an_0) \mid a \in (-\infty,0), n_0 \in \mathbb{N}\} \cup \{0\}\times(0,\infty) \cup (0,1)\times\mathbb{R} \cup \{1\}\times(-\infty,0) \subset \mathbb{R}^2$$

that is, if they lie in one of the following six disjoint *areas* (see also Figure 1, left side):

$$\begin{array}{llll}
(1) & a<0, & s>0, & -\frac{s}{a} =: n_0 \in \mathbb{N}; \\
(2) & 0<a<1, & s>0; & \\
(3) & 0<a\le 1, & s<0, & -\frac{s}{a} \notin \mathbb{N}; \\
(4) & a=0, & s>0; & \\
(5) & 0<a<1, & s=0; & \\
(6) & 0<a\le 1, & s<0, & -\frac{s}{a} \in \mathbb{N}
\end{array}$$

https://doi.org/10.1017/S1748499522000148 Published online by Cambridge University Press

The *delay* of $(a, s) \in \mathcal{A}$ is the integer

$$m_0(a, s) := \begin{cases} 0, & s > 0 \\ 1 + \left[-\frac{s}{a}\right], & s \leq 0 \end{cases} \tag{13}$$

such that for $s \leq 0$ we have $0 \leq m_0 - 1 \leq -\frac{s}{a} < m_0$.

Finally, for $n \in \mathbb{N}_0$ we define the subsets

$$\mathcal{A}_n := \{(a, s) \in \mathcal{A} \mid m_0(a, s) \leq n\}, \quad \mathcal{A}_{-1} := \emptyset, \quad \mathcal{A}_{(n)} := \{(a, s) \in \mathcal{A} \mid m_0(a, s) = n\} = \mathcal{A}_n \backslash \mathcal{A}_{n-1}$$

One sees at a glance that strongly admissible parameters are admissible and that the two definitions of the delay are consistent – as are the definitions of the respective subsets of $\mathcal{A}$ and $\mathcal{S}$:

Lemma 4.2. *We have* $\mathcal{A}_0 = \mathcal{A}_{(0)} = \{(a, s) \in \mathcal{A} \mid s > 0\}$ *and, for* $n \in \mathbb{N}$,

$$\mathcal{A}_n = \{(a, s) \in \mathcal{A} \mid s > 0 \vee s > -na\}, \qquad \mathcal{A}_{(n)} = \{(a, s) \in \mathcal{A} \mid -na < s \leq (1-n)\, a\}$$

For the subsets of $\mathcal{S}$ *the analogous formulae hold and, for* $n \in \mathbb{N}_0$, $\mathcal{S}_n = \mathcal{S} \cap \mathcal{A}_n$, $\mathcal{S}_{(n)} = \mathcal{S} \cap \mathcal{A}_{(n)}$.

Proof. Straightforward algebra. □

Formulae 10 and 11, which generate narrow/proper Panjer distributions, hold for Areas 1, 2, 3. They extend in a straightforward way to Area 4, which is adjacent (in topological sense) to Areas 1, 2.

Proposition 4.3. *The* $p_k(a, s)$ *and* $p_k(a, s; m)$ *from Definition 3.3 are continuous on* $\mathcal{S}$ *and can be continuously extended to Area 4.*

Proof. As all $p_k(a, s; m)$ are algebraic functions of the $p_k(a, s)$, $k \in \mathbb{N}_0$, it is sufficient to show continuity on $\mathcal{S}$ for the latter, which is clear from Formulae 10 and 3.

For $s > 0$, we can rearrange Formula 9 as

$$(1-a)^{-\frac{s}{a}} - \sum_{j=0}^{M-1} r_j(a, s) = \sum_{k=M}^{\infty} r_k(a, s) \tag{14}$$

where all summands but the first are continuous on $\mathbb{R}^2$. Now recall from the binomial series that for $-1 < a < 1$ and $\gamma \in \mathbb{R}$, $(1-a)^{\gamma}$ is well defined, finite and continuous in both variables. Thus, $(1-a)^{-\frac{s}{a}}$ is well defined and continuous for $(a, s) \in ((-1, 1)\backslash\{0\}) \times (0, \infty)$. The continuous extension to $\{0\} \times (0, \infty)$, that is, Area 4, is straightforward via $\lim_{a \to 0}(1-a)^{-\frac{s}{a}} = e^s$. So, both sides of Equation (14) are continuous functions on $\mathcal{S}$ plus Area 4, and further non-zero, such that $r_k(a, s)$ divided by the LHS yields $p_k(a, s; m)$ being a continuous function on that domain. □

Example 4.4. With $r_k(a, s) = \frac{s^k}{k!}$ on Area 4, Definition 3.3 here yields the *Poisson* distribution and its m-truncations, which are limiting cases of (m-truncated) binomial and NB distributions having the same parameter s.

The continuity of the $p_k(a, s)$ and $p_k(a, s; m)$ on the "upper" $\mathcal{A}_0$ (Areas 1, 2, 4) is remarkable as $\mathcal{A}_0$ has a weird geometry, see Figure 1, left side. Instead, continuity on Area 3 does not mean too much as this area is separated (in topological sense) from $\mathcal{A}_0$ and splits into small separated pieces: the triangular $\mathcal{S}_{(n)}$, $n > 0$. However, things become interesting once we involve the remaining areas: Area 6 connects the pieces of Area 3, while Area 5 connects Area 3 and Area 2, see again Figure 1, left side.

Our aim is to define the $p_k(a,s)$ and $p_k(a,s;m)$ on Areas 5 and 6 in a reasonable way, which would mean having them on the whole parameter set $\mathcal{A}$. With them being defined on $\mathcal{S}$ in a discrete manner via the step function $m_0(a,s)$, we cannot expect them to be continuous on $\mathcal{A}$. Yet, it will turn out that we get surprisingly close.

We can in the following treat Areas 5 and 6 largely in parallel, but the setting is a bit intricate. It is more convenient to work with the parameters a, α, which for $a \neq 0$ ($\mathcal{A}$ without Area 4) are equivalent to a, s and can be converted via smooth functions ($\alpha = \frac{s}{a}$, $s = \alpha a$), such that properties like continuity are not affected by this parameter change. For orientation see the right side of Figure 1 emphasising that the geometry of $\mathcal{A}$ is much simpler (utterly rectangular) in terms of a, α – as is the delay formula:

Definition 4.5. We call real a, α *(strongly) admissible* if $a \neq 0$, $s = \alpha a$ are (strongly) admissible. We call the set of such parameters $\mathcal{A}^*$ ($\mathcal{S}^*$) and write, equivalently to earlier formulae,

$$m_0(a,\alpha) := \begin{cases} 0, & \alpha > 0 \\ 1 + [-\alpha], & \alpha \leq 0 \end{cases}$$

$$p_k(a,\alpha;m) := \begin{cases} 0, & k < M = \max(m, m_0(a,\alpha)) \\ \dfrac{r_k(a,\alpha)}{(1-a)^{-\alpha} - \sum_{j=0}^{M-1} r_j(a,\alpha)}, & k \geq M \end{cases}$$

which for $m = m_0$ (and $M = m_0$) embraces $p_k(a,\alpha)$. We write finally, for $n \in \mathbb{N}_0$,

$$\mathcal{A}_n^* := \left\{(a,\alpha) \in \mathcal{A}^* \mid m_0(a,\alpha) \leq n\right\}, \quad \mathcal{A}_{-1}^* := \emptyset, \qquad \mathcal{A}_{(n)}^* := \left\{(a,\alpha) \in \mathcal{A}^* \mid m_0(a,\alpha) = n\right\} = \mathcal{A}_n^* \backslash \mathcal{A}_{n-1}^*$$

and analogously for the subsets of $\mathcal{S}^*$.

Lemma 4.6. *For $n \in \mathbb{N}$ we have*

$$\mathcal{A}_0^* = \mathcal{A}_{(0)}^* = \left\{(a,\alpha) \in \mathcal{A}^* \mid s > 0\right\}, \ \mathcal{A}_n^* \left\{(a,\alpha) \in \mathcal{A}^* \mid s > 0 \vee \alpha > -n\right\},$$
$$\mathcal{A}_{(n)}^* \left\{(a,\alpha) \in \mathcal{A}^* \mid -n < \alpha \leq 1-n\right\}$$

For the subsets of $\mathcal{S}^*$ the analogous formulae hold and, for $n \in \mathbb{N}_0$, $\mathcal{S}_n^* = \mathcal{S}^* \cap \mathcal{A}_n^*$, $\mathcal{S}_{(n)}^* = \mathcal{S}^* \cap \mathcal{A}_{(n)}^*$.

Proof. Straightforward translation from the (a,s) world. □

Definition 4.7. For all $(a,-n)$ in Areas 5 and 6 of $\mathcal{A}^*$ (where $a > 0$, $n \in \mathbb{N}_0$, and $m_0 = n+1$), we set

$$p_k(a, \alpha = -n) := p_k(a,\alpha)\big|_{\alpha=-n} := \lim_{\alpha \nearrow -n} p_k(a,\alpha) \tag{15}$$

and set, for $m \in \mathbb{N}_0$, the m-truncations $p_k(a, \alpha = -n; m)$ thereof as usual, according to Formula 2.

Proposition 4.8. *The $p_k(a, \alpha = -n)$ and the $p_k(a, \alpha = -n; m)$ of the preceding definition constitute well-defined discrete loss distributions on Areas 5 and 6. Together with the earlier analogous definitions on $\mathcal{S}^*$, they define $\left(p_k(a,\alpha)\right)$ and its m-truncations on $\mathcal{A}^*$.*

$p_k(a,\alpha)$ is continuous on $\mathcal{S}^*$, while on $\mathcal{A}^*$ it is continuous in a (for fixed α) and left-continuous in α with discontinuities at Areas 5 and 6 (for $k > n = -\alpha$). Instead, the m-truncation $p_k(a,\alpha;m)$ is continuous on $\mathcal{A}_m^*$.

https://doi.org/10.1017/S1748499522000148 Published online by Cambridge University Press

For $(a, \alpha) \in (0,1) \times \{0\}$ *(Area 5) or* $(a, \alpha) \in (0,1] \times -\mathbb{N}$ *(Area 6), we have with* $m_0(a, \alpha) = 1 - \alpha$ *that* $p_k(a, \alpha) = 0$ *for* $k < m_0$*, while for* $k \geq m_0$

$$\begin{aligned} p_k(a, \alpha = -n) &= \frac{a^k / \binom{k}{m_0}}{\sum_{j=m_0}^{\infty} a^j / \binom{j}{m_0}} \\ &= \frac{a^k / \binom{k}{m_0}}{(-1)^{m_0} m_0 \left\{ (\ln(1-a))(1-a)^{m_0-1} - \sum_{j=1}^{m_0-1} \left((-a)^j \binom{m_0-1}{j} \sum_{i=m_0-j}^{m_0-1} \frac{1}{i} \right) \right\}} \end{aligned} \tag{16}$$

with m_0 *notably equaling* $1 + n$.

For $m \in \mathbb{N}_0$ *and* $M = \max(m, m_0)$*, we have analogously* $p_k(a, \alpha) = 0$ *for* $k < M$*, while for* $k \geq M$

$$\begin{aligned} p_k(a, \alpha = -n; m) &= \frac{a^k / \binom{k}{m_0}}{\sum_{j=M}^{\infty} a^j / \binom{j}{m_0}} \\ &= \frac{a^k / \binom{k}{m_0}}{(-1)^{m_0} m_0 \left\{ (\ln(1-a))(1-a)^{m_0-1} - \sum_{j=1}^{m_0-1} \left((-a)^j \binom{m_0-1}{j} \sum_{i=m_0-j}^{m_0-1} \frac{1}{i} \right) \right\} + \sum_{j=m_0}^{M-1} a^j / \binom{j}{m_0}} \end{aligned} \tag{17}$$

Proof. The proof and more details are set out in Appendix A. □

Note that Formulae 16 and 17 contain the parameter α only indirectly, via m_0, which will below make conversion to the parameters a, s very easy. Both formulae represent the respective probabilities in two ways, a compact and a complex one. The former contains an infinite sum, which may complicate its use in practice, so it is good to have both variants.

One sees quickly that the parameter spaces can be allocated just as in Corollary 3.5.

Corollary 4.9. *The parameters of narrow distributions* $(p_k(a, \alpha))$ *with delay* $m_0 \in \mathbb{N}_0$ *lie in* $\mathcal{A}^*_{(m_0)}$*, where* $-m_0 < \alpha \leq 1 - m_0$*. The parameters* a, α *of proper distributions* $(p_k(a, \alpha; m))$ *with* $m \geq m_0(a, \alpha)$ *lie in* $\mathcal{A}^*_m$*, where* $\alpha > -m$.

Example 4.10. The compact part of Formula 16 (with the infinite sum in the denominator) is given, for $\alpha = 1 - m_0 \in -\mathbb{N}, 0 < a \leq 1$, by Hess *et al.* (2002), who term the resulting narrow distribution type *extended logarithmic* distribution. It corresponds to Area 6 of the parameter space and Formula 17 yields the corresponding m-truncations and proper models.

The given name becomes clear if we at last look at Area 5, where $\alpha = 1 - m_0 = 0, 0 < a < 1$. Here Formula 16 simplifies to

$$p_k(a, \alpha = 0) = \frac{a^k / k}{-\ln(1-a)}, \qquad k \geq 1$$

which is the *logarithmic* distribution, one of the two well-known narrow Panjer models of order 1 and the sixth narrow model type as specified by Hess *et al.* (2002). Again, Formula 17 yields the corresponding m-truncations and proper distributions.

For ease of orientation, we restate the parameter space of the narrow Panjer class in formulae and in Figure 2, in terms of both (a, s) and (a, α). We add the *geometric* distribution, which is the special NB model having $\frac{s}{a} = \alpha = 1$.

https://doi.org/10.1017/S1748499522000148 Published online by Cambridge University Press

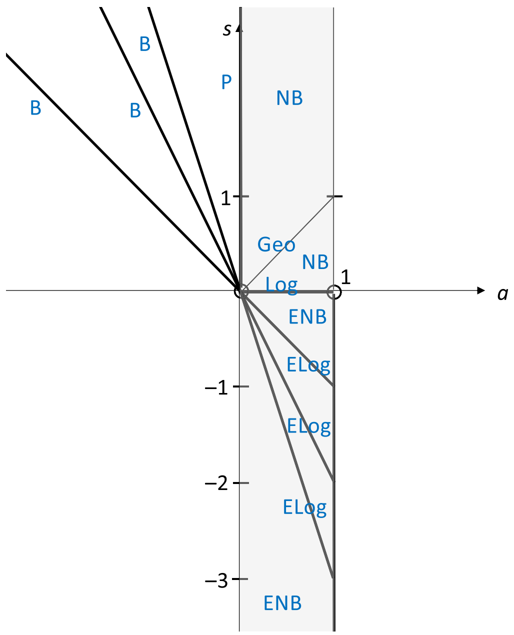

Figure 2. Narrow Panjer class.

(1)	B	$a < 0,$	$s > 0,$	$\alpha \in -\mathbb{N};$
(2)	NB	$0 < a < 1,$	$s > 0,$	$\alpha > 0;$
	Geo	$0 < a < 1,$	$s = a,$	$\alpha = 1;$
(3)	ENB	$0 < a \le 1,$	$s < 0,$	$\alpha \in (-\infty, 0) \backslash -\mathbb{N};$
	ETNB	$0 < a \le 1,$	$-1 < s < 0,$	$\alpha \in (-1, 0)\,;$
(4)	P	$a = 0,$	$s > 0;$	
(5)	Log	$0 < a < 1,$	$s = 0,$	$\alpha = 0;$
(6)	ELog	$0 < a \le 1,$	$s < 0,$	$\alpha \in -\mathbb{N}$

The left side of Figure 2 is a redesign of Figure 7.2.1 of Panjer & Willmot (1992), which represents the narrow $(a, b, 0)$ and $(a, b, 1)$ classes in terms of (a, b). We have added ELog and the ENB parts beyond ETNB (below the $s = -a$ line), which constitute the narrow models of higher order. To the right we show the same parameter space apart from Poisson ($a = 0$), translated to (a, α).

As noted earlier, the overall ENB parameter space splits into separate pieces, with ELog models in between. The delay m_0 is constant on each ENB piece and the (E)Log model bordering above. The logarithmic model separates ENB from NB, which has a topologically connected parameter space, but the intermediate geometric distribution can be seen as separating the very skewed NB models between Log and Geo from the "higher" NB models being closer to the limiting Poisson model.

As for the subsets of the parameter space, in both charts of Figure 2, $\mathcal{A}_0^{(*)} = \mathcal{A}_{(0)}^{(*)}$ corresponds to the part of the parameter space above (and excluding) the a-axis, plus the area left of the α-axis. $\mathcal{A}_{(1)}^{(*)}$ is the log line segment plus the first ENB piece (ETNB) immediately below, which on the left side is triangular and on the right side square. $\mathcal{A}_{(2)}^{(*)}$ is the second ENB triangle/square including its upper border, which is the first part of ELog; and so on.

The continuity of $p_k\ (a, \alpha; m)$ on $\mathcal{A}_m^*$ gives another insight if we for the moment leave the (topologically separated) binomial area to the left of the α-axis aside. The lowest part of $\mathcal{A}_m^*$ (to the right

https://doi.org/10.1017/S1748499522000148 Published online by Cambridge University Press

of the α-axis) is $\mathcal{A}^*_{(m)}$, the parameter space of the narrow Panjer model of order m having probabilities $p_k(a,\alpha) = p_k(a,\alpha;m)$. Instead, the highest part of $\mathcal{A}^*_m$ is Area 2, where the $p_k(a,\alpha;m)$ represent the *m-truncated NB* distribution. By this perspective, $p_k(a,\alpha;m)$ can be seen as representing a continuous extension of m-truncated NB from Area 2 to $\mathcal{A}^*_m$, or say: from $\alpha > 0$ down to $\alpha > -n$. In particular for $m = 1$, this means extending 1-truncated NB to the first ENB piece. This justifies the traditional name ETNB; however, it shall be noted that this extension also embraces the logarithmic model in between, which allows for a continuous passage from 1-truncated NB to ETNB.

Remark 4.11. While the NB model is often interpreted as being opposed to binomial ($\alpha = \pm1, \pm2, \pm3, \ldots$, starting with geometric vs Bernoulli), Figure 2 reveals another opposition, on the right side: binomial vs extended logarithmic. For each $\alpha = -1, -2, -3, \ldots$ there is a pair of models, which are indeed based on the *same binomial series* with exponent $-\alpha \in \mathbb{N}$:

to the left of the α-axis $(1-a)^{-\alpha}$, $a < 0$, generating a binomial model of "size" $n_0 = -\alpha$, where the r_k (and p_k) are initially positive, but for $k \geq n_0 + 1 = 1 - \alpha$ equal 0, which leads to a model being concentrated on $\{0, 1, \ldots, -\alpha\}$,

to the right of the α-axis $(1-a)^{-\alpha}$, $0 < a \leq 1$, generating an ELog model with positive probabilities starting at m_0, where likewise the r_k, for $k \geq m_0 = 1 - \alpha$, equal 0, but can be "rescaled" to $p_k > 0$ via L'Hôpital's rule.

If one, for the sake of symmetry with ELog, restricts binomial to $-1 \leq a < 0$, one gets the binomial models having probability $0 < p \leq \frac{1}{2}$. The conversion of parameters for this and other Panjer distributions can be deduced most easily from Table 3 in Appendix B.

5. Wrap Up

Let us collect and complement the main results, formulating them in terms of the parameters a, s, which work for the whole Panjer class including Area 4.

Theorem 5.1. *The general narrow Panjer class Pan (a,s) and the general proper Panjer class Pan $(a,s;m)$ of distribution models can be described by the parameters*

$$(a,s) \in \mathcal{A} = \{(a, -an_0) \mid a \in (-\infty, 0), n_0 \in \mathbb{N}\} \cup \{0\} \times (0,\infty) \cup (0,1) \times \mathbb{R} \cup \{1\} \times (-\infty, 0) \subset \mathbb{R}^2$$

and (for the latter)

$$\mathbb{N}_0 \ni m \geq m_0(a,s) = \begin{cases} 0, & s > 0 \\ 1 + \left[-\frac{s}{a}\right], & s \leq 0 \end{cases}$$

with the additional constraint $m < n_0$ for $(a,s) \in \{(a, -an_0) \mid a \in (-\infty, 0), n_0 \in \mathbb{N}\}$. Technically, one can drop the constraint $m \geq m_0$ but needs it if one wants uniqueness of parameters.

For $n \in \mathbb{N}_0$, the narrow Panjer class of order n corresponds to the parameter subspace

$$\mathcal{A}_{(n)} = \{(a,s) \in \mathcal{A} \mid -na < s \leq (1-n)\,a\}; \qquad \mathcal{A}_{(0)} = \{(a,s) \in \mathcal{A} \mid s > 0\}$$

while the proper (a,b,n) class corresponds to parameters in

$$\mathcal{A}_n \times \{n\}, \qquad \mathcal{A}_n = \left\{(a,s) \in \mathcal{A} \,\middle|\, s > -n\,a^+\right\}$$

$\mathcal{A}$ splits into six areas corresponding to six types of distributions:

(1)	*B*	$a < 0$,	$s > 0$,	$-\frac{s}{a} = n_0 \in \mathbb{N}$;
(2)	*NB*	$0 < a < 1$,	$s > 0$;	
(3)	*ENB*	$0 < a \leq 1$,	$s < 0$,	$-\frac{s}{a} \notin \mathbb{N}$;
(4)	*P*	$a = 0$,	$s > 0$;	
(5)	*Log*	$0 < a < 1$,	$s = 0$;	
(6)	*ELog*	$0 < a \leq 1$,	$s < 0$,	$-\frac{s}{a} = n \in \mathbb{N}$;

https://doi.org/10.1017/S1748499522000148 Published online by Cambridge University Press

The probabilities of the general narrow/proper Panjer distribution are given, with $p_k(a,s) = p_k(a,s;m_0)$, by the following formula:

$$p_k(a,s;m) = \begin{cases} 0, & k < M \\ \dfrac{r_k(a,s)}{(1-a)^{-s/a} - \sum_{j=0}^{M-1} r_j(a,s)}, & k \geq M \end{cases} \tag{18}$$

which emerges from the binomial series and where

$$r_k(a,s) = \frac{1}{k!}\prod_{i=0}^{k-1}(s+ai)\,,\ k \in \mathbb{N}_0; \qquad M(a,s;m) = \max(m, m_0(a,s))$$

Formula 18 holds in a strict sense on the parameter subspace (types B, NB, ENB)

$$(a,s) \in \mathcal{S} = \{(a,-an_0) \mid a \in (-\infty,0)\,,\, n_0 \in \mathbb{N}\} \cup (0,1)\times(0,\infty) \cup (0,1]\times((-\infty,0)\backslash -\mathbb{N}) \subset \mathcal{A}$$

and extends to $\mathcal{A}\backslash\mathcal{S}$ as follows:

type P: for $(a,s) \in \{0\}\times(0,\infty)$ by simply replacing $(1-a)^{-s/a}$ by e^s, which formally means setting

$$p_k(a=0,s;m) := p_k(a,s;m)\big|_{a=0} := \lim_{a\to 0} p_k(a,s;m)$$

types Log, ELog: for $0 < a \leq 1, -\frac{s}{a} = n \in \mathbb{N}_0$, but without $(a,s) = (1,0)$, by setting

$$p_k(a,s=-na;m) := p_k(a,s;m)\big|_{s=-na} := \lim_{s\nearrow -na} p_k(a,s;m)$$

Here $m_0 = n+1$ and for $k \geq M$ we have

$$p_k(a,s=-na;m) = \frac{a^k/\binom{k}{m_0}}{\sum_{j=M}^{\infty} a^j/\binom{j}{m_0}}$$

$$= \frac{a^k/\binom{k}{m_0}}{(-1)^{m_0} m_0 \left\{(\ln(1-a))(1-a)^{m_0-1} - \sum_{j=1}^{m_0-1}\left((-a)^j \binom{m_0-1}{j} \sum_{i=m_0-j}^{m_0-1} \frac{1}{i}\right)\right\} + \sum_{j=m_0}^{M-1} a^j/\binom{j}{m_0}} \tag{19}$$

With these limiting cases, $p_k(a,s)$ is continuous on $\mathcal{A}$ apart from Areas 5 and 6 ($s = -na$, $n \in \mathbb{N}_0$), where we have discontinuity for $k > n$. Instead, for fixed $m \geq m_0$, $p_k(a,s;m)$ is a continuous function on $\mathcal{A}_m$.

The general formula for the probability generating function of the proper general Panjer class (which embraces the narrow class for $m = m_0$) is

$$P_{(a,s;m)}(z) = \frac{(1-az)^{-s/a} - \sum_{j=0}^{M-1} r_j(az,s)}{(1-a)^{-s/a} - \sum_{j=0}^{M-1} r_j(a,s)} \tag{20}$$

which likewise holds in a strict sense for B, NB, ENB, and extends

to P via $a \to 0$, which simply means replacing $(1-a)^{-s/a}$ by e^s and $(1-az)^{-s/a}$ by e^{sz};

https://doi.org/10.1017/S1748499522000148 Published online by Cambridge University Press

to (E)Log via $s \nearrow -na$. Here, we have

$$P_{(a,s=-na;\,m)}\left(z\right) := P_{(a,s;\,m)}\left(z\right)\Big|_{s=-na} = \frac{\sum_{j=M}^{\infty}\left(az\right)^{j}/\binom{j}{m_0}}{\sum_{j=M}^{\infty}a^{j}/\binom{j}{m_0}}$$

$$= \frac{\left(-1\right)^{m_0} m_0 \left\{\left(\ln\left(1-az\right)\right)\left(1-az\right)^{m_0-1} - \sum_{j=1}^{m_0-1}\left(\left(-az\right)^{j}\binom{m_0-1}{j}\sum_{i=m_0-j}^{m_0-1}\frac{1}{i}\right)\right\} + \sum_{j=m_0}^{M-1}\left(az\right)^{j}/\binom{j}{m_0}}{\left(-1\right)^{m_0} m_0 \left\{\left(\ln\left(1-a\right)\right)\left(1-a\right)^{m_0-1} - \sum_{j=1}^{m_0-1}\left(\left(-a\right)^{j}\binom{m_0-1}{j}\sum_{i=m_0-j}^{m_0-1}\frac{1}{i}\right)\right\} + \sum_{j=m_0}^{M-1}a^{j}/\binom{j}{m_0}} \tag{21}$$

With the notation $x_{(n)} = \binom{x}{n}n! = x\left(x-1\right)\ldots\left(x-n+1\right)$, $x \in \mathbb{R}$, $n \in \mathbb{N}_0$ (with $x_{(0)} = 1$), we have, for $a < 1$, the following recursion for the factorial moments of $N \sim Pan\left(a, s; m\right)$:

$$E\left(N_{(n)}\right) = \frac{1}{1-a}\left\{\left(\left(n-1\right)a+s\right)E\left(N_{(n-1)}\right) + M_{(n)}p_M\right\}, \qquad n \in \mathbb{N} \tag{22}$$

The respective first and second moments, written compactly in terms of a, s, M, and the first non-zero probability p_M, are

$$E\left(N\right) = E\left(N_{(1)}\right) = \frac{s + Mp_M}{1-a} \tag{23}$$

$$E\left(N_{(2)}\right) = \frac{s\left(a+s\right) + \left\{\left(1-a\right)M + 2a + s - 1\right\}Mp_M}{\left(1-a\right)^2} \tag{24}$$

$$Var\left(N\right) = \frac{s + \left\{\left(1-a\right)M + a - s\right\}Mp_M - M^2p_M^2}{\left(1-a\right)^2} \tag{25}$$

Geometrically, all narrow and proper Panjer distributions have unimodal probability functions with an overall maximum probability and no further local maxima. For $s \leq 0$ (Log, ENB, ELog) the mode is the first non-zero probability.

Proof. Recall that Hess *et al.* (2002) defined the six types of narrow and proper Panjer distributions (using diverse parameterisations), proving in particular that no more types exist. Our system describes the same distributions, rewritten uniformly in terms of the parameters a, s, m, and thus also covers the whole Panjer class. The optional constraint $m \geq m_0$ (which implies $M = m$) for the proper models ensures unique parameters: for all $m \leq m_0$ one has $p_k\left(a, s; m\right) = p_k\left(a, s; m_0\right) = p_k\left(a, s\right)$.

The correspondence between the classes of order n and the subsets $\mathcal{A}_n$, $\mathcal{A}_{(n)}$ of $\mathcal{A}$ results quickly from Corollaries 3.5 and 4.9.

The formulae for the probabilities and the assertions on continuity are earlier results of this paper (in particular from Proposition 4.8), translated from the (a, α) parameterisation where necessary. The latter does not cover Area 4, but this is topologically separated from Areas 5 and 6 and does not affect their continuity properties.

The general pgf formula is given in Table 1 of Hess *et al.* (2002) for the (E)NB distributions, in terms of a, α (which they call ϑ, β). Converting to a, s, embracing B and P and generalising to the respective m-truncations is straightforward. The compact variant of the ELog formula is given there too; generalising to m-truncations is straightforward. The second (complex) pgf representation results from Formula 27 in Appendix A; both variants embrace Log ($n = 0$, $0 < a < 1$, $m_0 = 1$).

The recursion for the factorial moments results from the corresponding differential equation for the pgf, as given in Theorem 2.1(b) of Hess *et al.* (2002), after translation to our notation. The expectation follows immediately, the second moment after some algebra.

https://doi.org/10.1017/S1748499522000148 Published online by Cambridge University Press

As for modes, if a probability function, on some interval, is positive and first rises then decreases, constituting a (possibly local) maximum, this implies that some $\frac{p_k}{p_{k-1}}$ are greater than 1, being followed by values smaller than 1. For the non-zero probabilities of proper Panjer distributions we have $\frac{p_k}{p_{k-1}} = a + \frac{b}{k}$, which is a monotonic function in k for all admissible a, b and can shift from being > 1 to < 1 at most once. So, proper Panjer distributions cannot have a local maximum probability other than the global one, thus are unimodal.

As is well known, for $s > 0$ the mode of the narrow distributions (B, NB, P) may be positive and may consist of two subsequent probabilities, which is, however, still interpreted as unimodal. This property is inherited by their left-truncations, the corresponding proper Panjer distributions.

For a proper Panjer distribution with $s \le 0$ we have $a > 0$, $b = s - a < 0$, such that overall for the non-zero probabilities we have $\frac{p_k}{p_{k-1}} = a + \frac{b}{k}$ with a factor $0 < a + \frac{b}{k} < a \le 1$. Thus, the probabilities strictly decrease from the first non-zero probability onwards. □

Note that the recursion for the moments does not embrace the case $a = 1$, where at least the higher moments are infinite (Hess *et al.*, 2002). We have included this border case in our theory for the sake of completeness; it is, however, not of much practical interest.

Example 5.2. For the Panjer $(a, b, 0)$ class, the formulae for expectation and variance are simple and well known:

$$\mathrm{E}\,(N) = \frac{s}{1-a}, \qquad \mathrm{Var}\,(N) = \frac{s}{(1-a)^2}$$

The corresponding formulae for the proper $(a, b, 1)$ class are

$$\mathrm{E}\,(N) = \frac{s+p_1}{1-a}, \qquad \mathrm{Var}\,(N) = \frac{(s+p_1)\,(1-p_1)}{(1-a)^2}$$

They are very compact; however, some complexity is hidden in p_1. Note that they hold for both the narrow distributions of order 1 (Log and ETNB) and the 1-truncations of Poisson, binomial and NB. Instead, the better-known alternative formulae (see Appendix B.3.1 of Klugman *et al.* (2008))

$$\mathrm{E}\,(N) = \frac{s}{(1-a)\,(1-p_0^*)}, \qquad \mathrm{Var}\,(N) = \frac{s\,(1-(1+s)\,p_0^*)}{(1-a)^2\,(1-p_0^*)^2}$$

are more complex and hold only for the 1-truncations of the models of order 0, from which they take the no-loss probability p_0^*.

6. Parameter Inference

With the multitude of situations where Panjer distributions could be applied, from abundant empirical data (wide models) to cases where low loss counts are impossible (proper models), a full exploration of parameter inference is beyond the scope of this paper; this would probably yield enough content for a separate paper (available data for examples permitting). Yet, some ideas for potentially successful procedures can be derived from generalising a graphical approach explained in section 6.5 of Klugman *et al.* (2008) for the classical distributions of order 0.

If the probabilities of a loss count distribution ultimately fulfil Panjer's recursion, for large enough k, the empirical loss counts n_k fulfil approximately, apart from random fluctuation, the equivalent formulae

$$\frac{n_k}{n_{k-1}} \approx a + \frac{b}{k}, \qquad k\frac{n_k}{n_{k-1}} \approx ka + b$$

https://doi.org/10.1017/S1748499522000148 Published online by Cambridge University Press

Table 2. Loss count example.

k	n_k	$k\frac{n_k}{n_{k-1}}$
0	7,840	
1	1,317	0.17
2	239	0.36
3	42	0.53
4	14	1.33
5	4	1.43
6	4	6.00
7	1	1.75
8+	0	
Total	9,461	

For high k randomness and low likelihood will let most n_k equal zero, for the lowest k the recursion may not hold; however, if we observe an intermediate interval of values k with non-zero n_k, we can plot the $k\frac{n_k}{n_{k-1}}$ against k and check whether they approximate an affine linear function $ka + b = \left(k - 1\right) a + s$. If so, the distribution is a candidate for a general-Panjer fit and we can get initial (rough) estimates for a, b, s and m_0, the latter being an approximate lower bound for m.

The assessment of m is crucial: if we can narrow down m to one or a few integers, we can do the final estimation of the further parameters for fixed m, which greatly reduces complexity of the Maximum Likelihood (or other) estimation algorithm.

- If we want to fit a proper model, m is often determined by the setting of the problem. For example, in the case of the six risks warranty (Cat reinsurance) mentioned earlier, it is clear that m most probably equals 6.
- Wide models have many parameters and need a lot of data. In such situations the plot should make clear where the Panjer recursion starts to hold. The initial probabilities before the Panjer tail are parameters and simply estimated by the respective sample frequencies.

Having fixed m, a look at the initial estimates for a and s will in many cases narrow down the parameter space to consider: to a rather small subset of $\mathcal{A}_m$. The most benign cases are clear indications for either (possibly truncated) binomial or NB. ENB is slightly more complicated with the ELog stripes in between; however, as the $p_k\,(a, s; m)$ are continuous on $\mathcal{A}_m$, complexity is lower than the above ELog formulae indicate. Bordering cases require some more work, just as in the classical case: when one is not sure whether binomial or NB fits better, one has to calculate both, plus the intermediate Poisson.

In such cases inference may technically work better when done separately per distribution type, possibly using parameters other than the unified ones. For example, for NB the parameterisation NB2 as given in Appendix B, which uses shape parameter α and expectation λ, is preferable as the respective MLE estimators are asymptotically independent, see section 9.8 of Panjer & Willmot (1992).

Example 6.1. For illustration we revisit Example 6.2 of Klugman *et al.* (2008) treating an MTPL loss count data set, see Table 2. The $k\frac{n_k}{n_{k-1}}$ show a slow but clear upward trend, albeit one must acknowledge that the last points are based on very scarce data. However, the data seem to indicate that if modelled with a Panjer distribution, the resulting parameter a should be positive.

So, among the models of order 0, the first choice would be NB. Here one would expect to get a rather small $a \in (0, 1)$ and an even smaller $s > 0$, which would lead to a likewise small $\alpha = \frac{s}{a} > 0$ and thus to a rather skewed NB distribution.

https://doi.org/10.1017/S1748499522000148 Published online by Cambridge University Press

Alternatively, one can try wide $(a, b, 1)$ models, which are called zero-modified by Klugman *et al*. (2008) and have a free parameter p_0 followed by a Panjer tail. Willmot (1988) derived (in his section 6) common MLE formulae for zero-modified NB/ETNB, embracing the intermediate Log as limiting case. This means inference based on the parameter space $\mathcal{A}_1$, or more exactly $\mathcal{A}_1^*$ as the parameters a, α were used. The MLE estimators were applied to a number of MTPL data sets, including (as Data Set 5) the one discussed here. The resulting MLE estimates are the sample frequency $\widehat{p_0} = \frac{7840}{9461} = 0.829$ and $\widehat{\alpha} = -0.103$, $\widehat{a} = 0.380$, which means $\widehat{s} = \widehat{a}\widehat{\alpha} = -0.391$. The resulting fit turns out to be statistically reasonable; the negative, but quite small, $\widehat{\alpha}$ means that the estimated tail model is ETNB and quite similar to Log ($\alpha = 0$).

However, noting that the slope of the function $k\frac{n_k}{n_{k-1}}$ rises after the first steps, one could alternatively try wide models of a bit higher order. After the second step the function $k\frac{n_k}{n_{k-1}}$ looks rather steep; an affine linear approximation in this area would lead to a larger a than with the above models, possibly greater than 0.5, and to a negative s maybe close to -1. The resulting α would be in or around the interval $[-2, -1]$. So, plausible values for m_0 are 1, 2 and possibly 3. Having already tried $m = 0, 1$, the next (and due to data scarcity probably last reasonable) option is $m = 2$. In a wide model of order 2, p_0 and p_1 are estimated by the empirical frequencies. The further probabilities constitute the Panjer tail; the respective parameter space for a, s is $\mathcal{A}_2$; tail inference is based on the observed n_k with $k \geq 2$. Here we have 304 such observations, which is not abundant but could suffice for reliable inference of the two tail parameters.

7. Conclusion

We have seen that the representation of the general Panjer class of distributions in terms of a, s yields (mostly) simpler formulae and more intuitive insight than the traditional representation in terms of a, b. Other parameterisations, for example, a, α, can yield still simpler formulae and additional insight, but they usually do not work for the whole class, see the overview in Appendix B and there in particular Table 3. It seems that some of the most common parameterisations, while being very practical for a specific model type, obscure the links to other parts of the Panjer class.

Beyond being instructive, the unified view of the general Panjer class of distributions via the parameters a, s, m can ease implementation and use of the models in practice, providing (mostly compact) all-in one formulae for their probabilities and moments.

References

Albrecher, H., Beirlant, J. & Teugels, J.L. (2017). *Reinsurance: Actuarial and Statistical Aspects*. Wiley.

Fackler, M. (2009). Panjer class united – one formula for the Poisson, Binomial, and Negative Binomial distribution. In *ASTIN* Colloquium *2009*.

Fackler, M. (2011). Panjer Class United: one formula for the probabilities of the Poisson, Binomial, and Negative Binomial distribution. *Anales del Instituto de Actuarios Españoles*, **17**, 1–12.

Fackler, M. (2017). *Experience rating of (re)insurance premiums under uncertainty about past inflation*. PhD thesis, Universität Oldenburg, urn:nbn:de:gbv:715-oops-33087.

Gerhold, S., Schmock, U. & Warnung, R. (2010). A generalization of Panjer's recursion and numerically stable risk aggregation. *Finance and Stochastics*, **14**(1), 81–128.

Hess, K.T., Liewald, A. & Schmidt, K.D. (2002). An extension of Panjer's recursion. *ASTIN Bulletin*, **32**(2), 283–297.

Johnson, N.L. & Kotz, S. (1969). *Distributions in Statistics: Discrete Distributions*. Houghton Mifflin.

Johnson, N.L., Kotz, S. & Kemp, A.W. (2005). *Univariate Discrete Distributions*. Wiley.

Klugman, S.A., Panjer, H.H. & Willmot, G.E. (2008). *Loss Models: From Data to Decisions*. Wiley.

Panjer, H.H. (1981). Recursive evaluation of a family of compound distributions. *ASTIN Bulletin*, **12**(1), 22–26.

Panjer, H.H. & Willmot, G.E. (1992). *Insurance Risk Models*. Society of Actuaries.

Schröter, K.J. (1990). On a family of counting distributions and recursions for related compound distributions. *Scandinavian Actuarial Journal*, **1990**(2–3), 161–175.

Sundt, B. (2002). Recursive evaluation of aggregate claims distributions. *Insurance: Mathematics and Economics*, **30**(3), 297–322.
Sundt, B. & **Jewell, W.S.** (1981). Further results on recursive evaluation of compound distributions. *ASTIN Bulletin*, **12**(1), 27–39.
Willmot, G.E. (1988). Sundt and Jewell's family of discrete distributions. *ASTIN Bulletin*, **18**(1), 17–29.
Willmot, G.E. & **Panjer, H.H.** (1987). Difference equation approaches in evaluation of compound distributions. *Insurance: Mathematics and Economics*, **6**(1), 43–56.

A Technical Details for the (E)Log Models

Lemma A.1. *For $0 < a < 1$ and real α we have*

$$\frac{d}{d\alpha}(1-a)^{-\alpha} = (-\ln(1-a))(1-a)^{-\alpha}$$

This extends to the case $a = 1, \alpha < 0$ if we interpret the RHS as the limit for $a \nearrow 1$.

Proof. The formula follows immediately from $(1-a)^{-\alpha} = \exp(-\alpha \ln(1-a))$. For $a = 1, -\alpha > 0$ both sides of the formula yield 0, which for the RHS can be seen quickly via L'Hôpital's rule. □

Lemma A.2. *For $0 < a \leq 1$, integers $0 \leq j < n$, and real $\alpha \notin \{1-j, \ldots, -1, 0\}$ we have*

$$\frac{d}{d\alpha}\prod_{i=0}^{j-1}(\alpha+i) = \left(\prod_{i=0}^{j-1}(\alpha+i)\right)\sum_{i=0}^{j-1}\frac{1}{\alpha+i}$$

such that

$$\left.\frac{d}{d\alpha}\right|_{\alpha=1-n} r_j(a,\alpha) = -(-a)^j\binom{n-1}{j}\sum_{i=n-j}^{n-1}\frac{1}{i}$$

In particular for $j = 0$, the derivatives equal 0.

Proof. The case $j = 0$ is clear: the indexed product and sums are empty; the former equals 1 and has derivative 0, the latter equal 0. For $j > 0$ the first formula is clear. As for the second,

$$\left.\frac{d}{d\alpha}\right|_{\alpha=1-n} r_j(a,\alpha) = \frac{a^j}{j!}\left(\prod_{i=0}^{j-1}(1-n+i)\right)\sum_{i=0}^{j-1}\frac{1}{1-n+i} = -\frac{(-a)^j}{j!}\left(\prod_{i=n-j}^{n-1} i\right)\sum_{i=n-j}^{n-1}\frac{1}{i}.$$

□

Lemma A.3. *For $0 < a \leq 1$ and integers $k \geq n \geq 1$, $\frac{r_k(a,\alpha)}{\alpha+n-1}$ is well defined at $\alpha = 1-n$ and*

$$\left.\frac{r_k(a,\alpha)}{\alpha+n-1}\right|_{\alpha=1-n} = \frac{(-1)^{n+1}}{n}\frac{a^k}{\binom{k}{n}}$$

Proof. We have

$$\frac{r_k(a,\alpha)}{\alpha+n-1} = \frac{a^k}{k!}\left(\prod_{i=0}^{n-2}(\alpha+i)\right)\prod_{i=n}^{k-1}(\alpha+i)$$

https://doi.org/10.1017/S1748499522000148 Published online by Cambridge University Press

which for $\alpha = 1 - n$ equals

$$\frac{a^k}{k!}\left(\prod_{i=0}^{n-2}(1-n+i)\right)\prod_{i=n}^{k-1}(1-n+i) = \frac{a^k}{k!}(-1)^{n-1}(n-1)!\,(k-n)! = \frac{(-1)^{n+1}}{n}a^k\frac{n!\,(k-n)!}{k!}$$

□

Proposition A.4. *For any integers $m \geq 1$, $k \geq 0$, the probability $p_k(a, \alpha; m)$ can be continuously extended from $\mathcal{S}^*$:*

to $(a, \alpha) \in (0, 1] \times \{1 - m\}$ (in Area 6) for $m > 1$,

to $(a, \alpha) \in (0, 1) \times \{0\}$ (being Area 5) for $m = 1$.

The resulting $\lim_{\alpha \to 1-m} p_k(a, \alpha; m)$ is at the same time the (well-defined) limit of $p_k(a, \alpha)$ for $\alpha \nearrow 1 - m$; it equals 0 if $k < m$, else

$$\lim_{\alpha \to 1-m} p_k(a, \alpha; m) = \lim_{\alpha \nearrow 1-m} p_k(a, \alpha) =$$

$$= \frac{a^k/\binom{k}{m}}{(-1)^m m\left\{(\ln(1-a))(1-a)^{m-1} - \sum_{j=1}^{m-1}\left((-a)^j\binom{m-1}{j}\sum_{i=m-j}^{m-1}\frac{1}{i}\right)\right\}}$$

$$= \frac{a^k/\binom{k}{m}}{\sum_{j=m}^{\infty} a^j/\binom{j}{m}}, \qquad k \geq m \geq 1 \tag{26}$$

These limits are positive for $k \geq m$, while their sum (over k) equals 1, such that overall one has a loss count distribution.

Proof. For $(a, \alpha) \in \mathcal{S}^*_{(m)}$, where $-m < \alpha < 1 - m \leq 0$, we have $m_0(a, \alpha) = m$, such that $p_k(a, \alpha)$ and $p_k(a, \alpha; m)$ coincide. Thus, if $\lim_{\alpha \to 1-m} p_k(a, \alpha; m)$ exists, so does $\lim_{\alpha \nearrow 1-m} p_k(a, \alpha)$ and is equal.

Now consider $(a, \alpha) \in \mathcal{S}^*$ with α being in the proximity of $1 - m \leq 0$ (which implies in particular $\alpha < 1$). Then we have $-m_0 < \alpha < 1 - m_0 \leq 1$. Thus, $m_0 = m$ if $\alpha < 1 - m$ and $m_0 = m - 1$ if $\alpha > 1 - m$, such that in both cases $M = m$. Consider further the binomial series

$$(1-a)^{-\alpha} = \sum_{j=0}^{\infty} r_j(a, \alpha), \qquad r_j(a, \alpha) = \frac{a^j}{j!}\prod_{i=0}^{j-1}(\alpha + i)$$

and recall that in Area 3, where $\alpha < 0$, this embraces the border case $a = 1$, where $(1 - a)^{-\alpha} = 0$. For all $j \geq m$, $r_j(a, \alpha)$ contains the factor $(\alpha + m - 1)$, such that $\lim_{\alpha \to 1-m} r_j(a, \alpha) = 0$. Thus, we must also have

$$\lim_{\alpha \to 1-m}\left\{(1-a)^{-\alpha} - \sum_{j=0}^{m-1} r_j(a, \alpha)\right\} = 0$$

Now we calculate $\lim_{\alpha \to 1-m} p_k(a, \alpha; m)$, which is trivial for $k < m$ as $p_k(a, \alpha; m) \equiv 0$ by definition (for strongly admissible parameters). For $k \geq m$ consider a sequence $(a_l, \alpha_l) \in \mathcal{S}^*$ converging to a point $(a, 1 - m)$ on the given interval. With L'Hôpital's rule and the three preceding lemmas we get

https://doi.org/10.1017/S1748499522000148 Published online by Cambridge University Press

$$\lim_{l\to\infty} p_k(a_l,\alpha_l;m) = \lim_{l\to\infty} \frac{r_k(a_l,\alpha_l)}{(1-a_l)^{-\alpha_l} - \sum_{j=0}^{m-1} r_j(a_l,\alpha_l)}$$
$$= \lim_{l\to\infty} \frac{\frac{r_k(a_l,\alpha_l)}{\alpha_l+m-1}}{\frac{(1-a_l)^{-\alpha_l} - \sum_{j=0}^{m-1} r_j(a_l,\alpha_l)}{\alpha_l+m_0-1}}$$
$$= \frac{\frac{(-1)^{m+1}}{m} a^k / \binom{k}{m}}{(-\ln(1-a))(1-a)^{m-1} - \sum_{j=1}^{m-1}\left(-(-a)^j\binom{m-1}{j}\sum_{i=m-j}^{m-1}\frac{1}{i}\right)}$$

This limit of positive figures must be nonnegative, obviously it does not equal 0 and is thus positive. The sum of the limits over k equals 1 because it is the limit of analogous series equaling 1. The final representation of Formula 26 is thus clear as the infinite sum in the denominator converges. □

Corollary A.5. *For* $0 < a \le 1$ *and integer* $m > 1$ *the following identity holds:*

$$(-1)^m m\left\{(\ln(1-a))(1-a)^{m-1} - \sum_{j=1}^{m-1}\left((-a)^j\binom{m-1}{j}\sum_{i=m-j}^{m-1}\frac{1}{i}\right)\right\} = \sum_{j=m}^{\infty}\frac{a^j}{\binom{j}{m}} \tag{27}$$

For $m=1$ and $0<a<1$ the formula holds as well and is the well-known logarithmic series.

Remark A.6. Gerhold et al. (2010) prove (in their Lemma 2.1) a formula of similar complexity that is apparently (but not obviously) equivalent. One can generalise their resulting Formula 2.7 to obtain an alternative to our Formula 21 representing the pgf of the (possibly m-truncated) extended logarithmic distribution.

Corollary A.7. *For any integers* $m\ge 1$, $k\ge 0$, *the probability* $p_k(a,\alpha;m)$ *can be continuously extended to* $(a,\alpha)\in((0,1]\times\{0,\dots,1-m\})\backslash(1,0)$, *which means that* $p_k(a,\alpha;m)$ *is (made) continuous on* $\mathcal{A}_m^*$. *The limits equal 0 if* $k<m$, *else we have, for* $k\ge m>0$, $n=0,\dots,m-1$,

$$\lim_{\alpha\to -n} p_k(a,\alpha;m) = \frac{a^k/\binom{k}{n+1}}{\sum_{j=n+1}^{\infty} a^j/\binom{j}{n+1}}$$
$$= \frac{a^k/\binom{k}{n+1}}{(-1)^{n+1}(n+1)\left\{(\ln(1-a))(1-a)^n - \sum_{j=1}^{n}\left((-a)^j\binom{n}{j}\sum_{i=n+1-j}^{n}\frac{1}{i}\right)\right\} - \sum_{j=n+1}^{m-1} a^j/\binom{j}{n+1}} \tag{28}$$

The resulting $p_k(a,\alpha=-n;m)$, $k\ge m$, *are positive and sum up to 1, such that one has a loss count distribution.*

Proof. For integers $m>n\ge 0$, $\left(p_k(a,\alpha;m)\right)$ is an m-truncation of $\left(p_k(a,\alpha;n+1)\right)$, which can be continuously extended at $\alpha=-n$ as shown in Proposition A.4, leading to the loss count distribution $\left(p_k(a,\alpha=-n;n+1)\right)$. For $k\ge m$ we have

$$p_k(a,\alpha=-n;n+1) = \frac{a^k/\binom{k}{n+1}}{(-1)^{n+1}(n+1)\left\{(\ln(1-a))(1-a)^n - \sum_{j=1}^{n}\left((-a)^j\binom{n}{j}\sum_{i=n+1-j}^{n}\frac{1}{i}\right)\right\}}$$

and calculating $p_k(a,\alpha=-n;m)$ out of the $p_j(a,\alpha=-n;n+1)$, one easily gets the asserted formula, including the alternative representation with the infinite sum.

https://doi.org/10.1017/S1748499522000148 Published online by Cambridge University Press

Collecting the continuous extensions for $n = 0, 1, ..., m-1$, one finds the gaps in the original domain $\mathcal{S}^*$ of $p_k(a, \alpha; m)$ closed for $\alpha = 0, -1, ..., 1-m$, such that $p_k(a, \alpha; m)$ is made continuous for the part of $\mathcal{A}^*$ having $a > 0, \alpha > -m$, which together with (the topologically separated) Area 1 constitutes $\mathcal{A}_m^*$. □

Having shown the continuity of the $p_k(a, \alpha; m)$ on large parts of $\mathcal{A}^*$ (in particular for large m), we finally show that $p_k(a, \alpha)$, which in Areas 5 and 6 is defined as $\lim_{\alpha \nearrow -n} p_k(a, \alpha)$, has right-discontinuities in α.

Proposition A.8. *$p_k(a, \alpha)$ is continuous on $\mathcal{S}^*$ and (for fixed α) continuous in a on the whole parameter space $\mathcal{A}^*$. However, at points $(a, -n)$ in Areas 5 and 6 (with $a > 0$, $n \in \mathbb{N}_0$), for $k \geq m_0 = n+1$, we have*

$$\lim_{\alpha \nearrow -n} p_k(a, \alpha) = p_k(a, \alpha = -n) > 0 = \lim_{\alpha \searrow -n} p_k(a, \alpha)$$

Proof. We know from Proposition 4.3 that $p_k(a, \alpha)$ is continuous on $\mathcal{S}^*$ (Areas 1, 2, 3) and from Formula 26 it is clear that, for fixed α, $p_k(a, \alpha)$ is a continuous function in a on Areas 5 and 6 too. As a function in α we have left-continuity by definition:

$$p_k(a, \alpha = -n) = \lim_{\alpha \nearrow -n} p_k(a, \alpha) > 0$$

To assess $\lim_{\alpha \searrow -n} p_k(a, \alpha)$, let α be slightly above $-n$, which implies $m_0(a, \alpha) = n$ and

$$p_k(a, \alpha) = \frac{r_k(a, \alpha)}{(1-a)^{-\alpha} - \sum_{j=0}^{n-1} r_j(a, \alpha)}$$

As $k \geq n+1$, the numerator contains the factor $(\alpha + n)$ and tends to 0 for $\alpha \searrow -n$. Instead, the denominator cannot tend to 0 because, as shown for Proposition A.4, $(1-a)^{-\alpha} - \sum_{j=0}^{n} r_j(a, \alpha)$ tends to 0 and its last summand $r_n(a, \alpha)$, which makes the difference, tends to

$$\frac{a^n}{n!} \prod_{i=0}^{n-1} (i-n) = (-a)^n \neq 0$$

Thus, $\lim_{\alpha \searrow -n} p_k(a, \alpha) = 0$. □

B Overview of the $(a, b, 0)$ Class

We unite the Panjer $(a, b, 0)$ class of distributions in one table, assembling parameterisations that are common, useful and/or instructive, not least the ones that can describe some distributions of higher order. Note that generally most negative binomial parameterisations can be formally extended to binomial, and some to Poisson, but not all of them yield formulae being easy to interpret or at least compact.

The large Table 3 below extends and enhances similar tables given in Fackler (2011) and (the better-known non-refereed version) Fackler (2009), where the parameterisations are discussed and the first all-in-one model is introduced.

For each parameterisation, we show the following quantities:

- probability function
- probability generating function
- probability of no loss
- expectation
- variance
- squared coefficient of variation: $\mathrm{Var}(N) / \mathrm{E}^2(N)$

https://doi.org/10.1017/S1748499522000148 Published online by Cambridge University Press

Table 3. Panjer $(a, b, 0)$ class.

	pf	pgf	p_0	E (N)	Var (N)	CV$^2(N)$	D (N)	OD (N)	Ct (N)	s	a	b	
P	$\frac{\lambda^k}{k!} e^{-\lambda}$	$e^{\lambda(z-1)}$	$e^{-\lambda}$	λ	λ	$\frac{1}{\lambda}$	1	0	0	λ	0	λ	P
B1	$\binom{n}{k} p^k (1-p)^{n-k}$	$(1-p+pz)^n$	$(1-p)^n$	np	$np\,(1-p)$	$\frac{1-p}{np}$	$1-p$	$-p$	$-\frac{1}{n}$	$\frac{np}{1-p}$	$\frac{-p}{1-p}$	$\frac{(n+1)p}{1-p}$	B1
B2	$\binom{n}{k} \frac{\lambda^k (n-\lambda)^{n-k}}{n^n}$	$\left(1+\frac{\lambda}{n}(z-1)\right)^n$	$\left(1-\frac{\lambda}{n}\right)^n$	λ	$\lambda\left(1-\frac{\lambda}{n}\right)$	$\frac{1}{\lambda}-\frac{1}{n}$	$1-\frac{\lambda}{n}$	$-\frac{\lambda}{n}$	$-\frac{1}{n}$	$\frac{n\lambda}{n-\lambda}$	$\frac{-\lambda}{n-\lambda}$	$\frac{(n+1)\lambda}{n-\lambda}$	B2
NB1	$\binom{\alpha+k-1}{k} p^\alpha (1-p)^k$	$\left(\frac{1-(1-p)z}{p}\right)^{-\alpha}$	p^α	$\frac{\alpha(1-p)}{p}$	$\frac{\alpha(1-p)}{p^2}$	$\frac{1}{\alpha(1-p)}$	$\frac{1}{p}$	$\frac{1-p}{p}$	$\frac{1}{\alpha}$	$\alpha\,(1-p)$	$1-p$	$(\alpha-1)(1-p)$	NB1
NB2	$\binom{\alpha+k-1}{k}\left(\frac{\alpha}{\alpha+\lambda}\right)^\alpha\left(\frac{\lambda}{\alpha+\lambda}\right)^k$	$\left(1-\frac{\lambda}{\alpha}(z-1)\right)^{-\alpha}$	$\left(\frac{\alpha}{\alpha+\lambda}\right)^\alpha$	λ	$\lambda\left(1+\frac{\lambda}{\alpha}\right)$	$\frac{1}{\lambda}+\frac{1}{\alpha}$	$1+\frac{\lambda}{\alpha}$	$\frac{\lambda}{\alpha}$	$\frac{1}{\alpha}$	$\frac{\alpha\lambda}{\alpha+\lambda}$	$\frac{\lambda}{\alpha+\lambda}$	$\frac{(\alpha-1)\lambda}{\alpha+\lambda}$	NB2
NB3	$\binom{\alpha+k-1}{k}\frac{\vartheta^\alpha}{(1+\vartheta)^{\alpha+k}}$	$\left(1-\frac{z-1}{\vartheta}\right)^{-\alpha}$	$\left(\frac{\vartheta}{1+\vartheta}\right)^\alpha$	$\frac{\alpha}{\vartheta}$	$\frac{\alpha}{\vartheta}\left(1+\frac{1}{\vartheta}\right)$	$\frac{1+\vartheta}{\alpha}$	$1+\frac{1}{\vartheta}$	$\frac{1}{\vartheta}$	$\frac{1}{\alpha}$	$\frac{\alpha}{1+\vartheta}$	$\frac{1}{1+\vartheta}$	$\frac{\alpha-1}{1+\vartheta}$	NB3
NB4	$\frac{(1+\beta)^{-k-\lambda/\beta}}{k!}\prod_{i=0}^{k-1}(\lambda+\beta i)$	$(1-\beta\,(z-1))^{-\frac{\lambda}{\beta}}$	$(1+\beta)^{-\frac{\lambda}{\beta}}$	λ	$\lambda\,(1+\beta)$	$\frac{1+\beta}{\lambda}$	$1+\beta$	β	$\frac{\beta}{\lambda}$	$\frac{\lambda}{1+\beta}$	$\frac{\beta}{1+\beta}$	$\frac{\lambda-\beta}{1+\beta}$	NB4
BNB1	$\binom{\alpha+k-1}{k}\frac{\beta^k}{(1+\beta)^{\alpha+k}}$	$(1-\beta\,(z-1))^{-\alpha}$	$(1+\beta)^{-\alpha}$	$\alpha\beta$	$\alpha\beta\,(1+\beta)$	$\frac{1+\beta}{\alpha\beta}$	$1+\beta$	β	$\frac{1}{\alpha}$	$\frac{\alpha\beta}{1+\beta}$	$\frac{\beta}{1+\beta}$	$\frac{(\alpha-1)\beta}{1+\beta}$	BNB1
BNB2	$(1-a)^\alpha a^k \binom{\alpha+k-1}{k}$	$\left(\frac{1-a}{1-az}\right)^\alpha$	$(1-a)^\alpha$	$\frac{\alpha a}{1-a}$	$\frac{\alpha a}{(1-a)^2}$	$\frac{1}{\alpha a}$	$\frac{1}{1-a}$	$\frac{a}{1-a}$	$\frac{1}{\alpha}$	αa	a	$(\alpha-1)\,a$	BNB2
Pan1a	$\left(1+\frac{\lambda}{\alpha}\right)^{-\alpha}\frac{\lambda^k}{k!}\prod_{i=0}^{k-1}\frac{\alpha+i}{\alpha+\lambda}$	$\left(1-\frac{\lambda}{\alpha}(z-1)\right)^{-\frac{1}{c}}$	$\left(\frac{\alpha}{\alpha+\lambda}\right)^\alpha$	λ	$\lambda\left(1+\frac{\lambda}{\alpha}\right)$	$\frac{1}{\lambda}+\frac{1}{\alpha}$	$1+\frac{\lambda}{\alpha}$	$\frac{\lambda}{\alpha}$	$\frac{1}{\alpha}$	$\frac{\alpha\lambda}{\alpha+\lambda}$	$\frac{\lambda}{\alpha+\lambda}$	$\frac{(\alpha-1)\lambda}{\alpha+\lambda}$	Pan1a
Pan1b	$(1+c\lambda)^{-\frac{1}{c}}\frac{\lambda^k}{k!}\prod_{i=0}^{k-1}\frac{1+ci}{1+c\lambda}$	$(1-c\lambda\,(z-1))^{-\frac{1}{c}}$	$(1+c\lambda)^{-\frac{1}{c}}$	λ	$\lambda\,(1+c\lambda)$	$\frac{1}{\lambda}+c$	$1+c\lambda$	$c\lambda$	c	$\frac{\lambda}{1+c\lambda}$	$\frac{c\lambda}{1+c\lambda}$	$\frac{(1-c)\lambda}{1+c\lambda}$	Pan1b
Pan2a	$(1-a)^{1+\frac{b}{a}}\frac{1}{k!}\prod_{i=1}^{k}(b+ai)$	$\left(\frac{1-a}{1-az}\right)^{1+\frac{b}{a}}$	$(1-a)^{1+\frac{b}{a}}$	$\frac{a+b}{1-a}$	$\frac{a+b}{(1-a)^2}$	$\frac{1}{a+b}$	$\frac{1}{1-a}$	$\frac{a}{1-a}$	$\frac{a}{a+b}$	$a+b$	a	b	Pan2a
Pan2b	$(1-a)^{\frac{s}{a}}\frac{1}{k!}\prod_{i=0}^{k-1}(s+ai)$	$\left(\frac{1-a}{1-az}\right)^{\frac{s}{a}}$	$(1-a)^{\frac{s}{a}}$	$\frac{s}{1-a}$	$\frac{s}{(1-a)^2}$	$\frac{1}{s}$	$\frac{1}{1-a}$	$\frac{a}{1-a}$	$\frac{a}{s}$	s	a	$s-a$	Pan2b
Pan3	$\frac{(1-a)^{\lambda\frac{1-a}{a}}}{k!}\prod_{i=0}^{k-1}(\lambda(1-a)+ai)$	$\left(\frac{1-a}{1-az}\right)^{\lambda\frac{1-a}{a}}$	$(1-a)^{\lambda\frac{1-a}{a}}$	λ	$\frac{\lambda}{1-a}$	$\frac{1}{\lambda(1-a)}$	$\frac{1}{1-a}$	$\frac{a}{1-a}$	$\frac{a}{\lambda(1-a)}$	$\lambda\,(1-a)$	a	$\lambda\,(1-a)-a$	Pan3
	pf	pgf	p_0	E (N)	Var (N)	CV$^2(N)$	D (N)	OD (N)	Ct (N)	s	a	b	

- dispersion: $\mathrm{D}(N) = \mathrm{Var}(N)/\mathrm{E}(N)$
- overdispersion: $\mathrm{OD}(N) = \mathrm{D}(N) - 1$
- contagion: $\mathrm{Ct}(N) = \mathrm{OD}(N)/\mathrm{E}(N)$
  For a general treatment of this most useful quantity, see Chapter 5 of Fackler (2017).
- Panjer's recursion parameters s, a, b

The presented models and parameters are the following:

- Poisson (1 variant)
  - *P*: expectation λ
- Binomial (2 variants): number of trials n
  - *B1*: with probability of success p
  - *B2*: with expectation λ
- Negative binomial (4 variants): shape parameter α
  - *NB1*: with probability p. Note that this probability is not comparable to the one used in B1.
  - *NB2*: with expectation λ
  - *NB3*: Poisson-Gamma model inheriting both parameters α and ϑ from the Gamma distribution with density $x^{\alpha-1}e^{-\vartheta x}\vartheta^{\alpha}/\Gamma(\alpha)$
  - *NB4*: without α; combines expectation λ and overdispersion β
- Negative binomial/binomial (2 variants): shape parameter (or negative number of trials) α
  - *BNB1*: with overdispersion β. For binomial, the parameters are the negatives of the B1 parameters, while for negative binomial they have a Poisson-Gamma interpretation with β being the inverse of ϑ from NB3 (which latter in the literature is often also denoted by β, such that the parameterisations are easily mixed up).
    - ∗ Used for negative binomial and the $(a, b, 1)$ class in Panjer & Willmot (1992), Klugman *et al.* (2008); extension to binomial is straightforward.
  - *BNB2*: with a from Panjer's recursion. For negative binomial a can be interpreted as the probability $1 - p$. Works for the general Panjer class but the Poisson type, which is the limit for $a \to 0$ with constant $\alpha a = s$.
    - ∗ Used in this paper to derive (E)Log from (E)NB.
    - ∗ Used by Willmot (1988) for (ET)NB, by Hess *et al.* (2002) for (E)NB.
    - ∗ Used by Schröter (1990) for a generalisation of Panjer's recursion.
- Panjer $(a, b, 0)$ class united (5 variants)
  - *Pan1a/b*: expectation λ with shape parameter α or contagion $c = 1/\alpha$
    - ∗ Discussed by Fackler (2011).
  - *Pan2a/b*: Panjer's recursion parameters a with b or s. Work for the general Panjer class.
    - ∗ Pan2b was used to formulate most theory developed in this paper (apart from a substantial detour to BNB2).
  - *Pan3*: combines the above λ and a. Used for negative binomial in some old literature (Johnson *et al.* 2005), but extension is straightforward.

The respective parameter spaces are as follows; some can be enlarged to embrace Panjer distributions of higher order.

- $p \in (0, 1)$
- $n \in \mathbb{N}$; in B2 $n > \lambda$
- $\lambda, \vartheta, \beta, s \in (0, \infty)$; or $\beta \in (-1, 0)$ in BNB1
- $\alpha \in (0, \infty)$; or $-\alpha \in \mathbb{N}$ in BNB2; or $\lambda < -\alpha \in \mathbb{N}$ or $\alpha = \pm\infty$ in Pan1a
- $c \in [0, \infty)$ or $\lambda < -\frac{1}{c} \in \mathbb{N}$
- $a \in (-\infty, 1)$; in BNB2 $a \neq 0$; in Pan2b $-\frac{s}{a} \in \mathbb{N}$ for $a < 0$; in Pan3 $\lambda \frac{a-1}{a} \in \mathbb{N}$ for $a < 0$
- $(a, s) \in \mathcal{A}_0$; $\big(a, b\big)$ as shown in Figure 6.6.1 of Panjer & Willmot (1992):

$$(a, s) \in \{(a, -an_0) \mid a \in (-\infty, 0)\,, n_0 \in \mathbb{N}\} \cup [0, 1)\times(0, \infty)$$

$$\big(a, b\big) \in \{(a, -a\,(n_0 + 1)) \mid a \in (-\infty, 0)\,, n_0 \in \mathbb{N}\} \cup \Big\{\big(a, b\big) \in [0, 1)\times\mathbb{R} \;\Big|\; a + b > 0\Big\}$$

Cite this article: Fackler M (2023). Panjer class revisited: one formula for the distributions of the Panjer (a,b,n) class, *Annals of Actuarial Science*, **17**, 145–169. https://doi.org/10.1017/S1748499522000148