Research Article

The Concepts of Pseudo Compound Poisson and Partition Representations in Discrete Probability

Werner Hürlimann

Swiss Mathematical Society, University of Fribourg, 1700 Fribourg, Switzerland

Correspondence should be addressed to Werner Hürlimann; whurlimann@bluewin.ch

Received 8 June 2015; Accepted 24 August 2015

Academic Editor: Shein-chung Chow

Copyright © 2015 Werner Hürlimann. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The mathematical/statistical concepts of pseudo compound Poisson and partition representations in discrete probability are reviewed and clarified. A combinatorial interpretation of the convolution of geometric distributions in terms of a variant of Newton's identities is obtained. The practical use of the twofold convolution leads to an improved goodness-of-fit for a data set from automobile insurance that was up to now not fitted satisfactorily.

1. Introduction

Consider the class of discrete arithmetic random variables X with probability generating function (pgf) $F(s) = \sum_{n \geq 0} f(n) \cdot s^n$ and nonvanishing zero probability $f(0) > 0$. Suppose that the pgf can be written as $F(s) = f(0) \cdot \exp\{G(s)\}$ for some generating function $G(s) = \sum_{k \geq 1} g(k) \cdot s^k$. The pseudo compound Poisson representation concerns the duality between $F(s)$ and $G(s)$, which is best expressed in terms of the identity $F'(s) = G'(s) \cdot F(s)$ or equivalently the recurrence relations

$$nf(n) = \sum_{k=1}^n kg(k) f(n-k), \quad n \geq 1. \quad (1)$$

As a consequence, it has been shown that X is infinitely divisible if and only if one has $g(k) \geq 0, \forall k \geq 1$ (e.g., [1]; [2, page 83]). Around the same time Feller [3, page 290] shows that X is infinitely divisible if and only if it is a compound Poisson random variable with parameter $\lambda = -\ln f(0)$ and severity probabilities $q(k) = g(k)/\lambda$. If X is not infinitely divisible, then (1) still holds for some $g(k), k \geq 1$, with at least one negative value, and this property motivates the naming *pseudo compound Poisson representation*.

Section 1 summarizes its main properties. The power series identity $F(s) = f(0) \cdot \exp\{G(s)\}$ leads also to the more complex and less known expression

$$f(n) = f(0) \cdot \sum_{x \in S_n} \prod_{k=1}^n \frac{g(k)^{x_k}}{x_k!}, \quad n \geq 1, \quad (2)$$

where $S_n = \{x = (x_1, \dots, x_n) \mid \sum_{k=1}^n kx_k = n\}$ is the set of free partitions of weight n , which is called *partition representation*. Although (2) has been applied in Hürlimann [4, Theorem 3.2], to derive an existence criterion for the construction of confidence bounds for discrete sampling distributions, expression (2) is not stated correctly and a proof of it is missing. Section 3 fills in these gaps and provides a combinatorial interpretation of the partition representation.

As a new illustration, Section 4 considers the convolution of m geometric random variables, whose pseudo compound Poisson representation is specified by the k th power sum polynomial $p(k) = kg(k) = \sum_{j=1}^m \theta_j^k, k \geq 1$. The partition representation (2) identifies with the $f(0)$ multiple of the n th complete symmetric function $h(n)$ in the variables θ_j . The recursion (1) is equivalent to a variant of Newton's identities, which motivates us to call this convolution *Newton distribution*. Applied to estimation theory, we show that this distribution satisfies Gauss's principle (the maximum likelihood estimator of the mean is the sample mean) and construct a parameter vector orthogonal to the mean. In the two-parameter case, we derive the maximum likelihood equations and illustrate their use in Section 5 at a specific data set from automobile insurance. Through regrouping of classes, we show that the Newton distribution beats in

goodness-of-fit both the negative binomial and the Poisson inverse Gaussian and at the same time improves the unsatisfactory fit obtained in a previous case study.

2. Characterization through Pseudo Compound Poisson Representation

Recall the pseudo compound Poisson representation in discrete probability theory from Hürlimann [5, 6]. Let X be a discrete arithmetic random variable defined on the natural numbers with probabilities $f(n) = \Pr(X = n)$, $n \geq 0$, such that $f(0) > 0$. Besides the probability generating function (pgf) $F(s) = \sum_{n \geq 0} f(n) \cdot s^n$ one considers the *cumulant* pgf defined by

$$G(s) = \ln \{F(s)\} + \lambda = \sum_{k \geq 1} g(k) \cdot s^k, \quad (3)$$

$$\lambda = -\ln f(0) = G(1).$$

Its name is motivated by the series expansion of the cumulant generating function (cgf)

$$C(s) = \ln \{F(e^s)\} = G(e^s) - \lambda = \sum_{k \geq 1} (e^{ks} - 1) g(k). \quad (4)$$

Proposition 1 (pseudo compound Poisson representation). *Let X be a discrete arithmetic random variable with pgf $F(s)$ such that $f(0) > 0$ and set $\lambda = -\ln f(0)$. Then the probabilities $f(n)$ satisfy Panjer's recursion*

$$f(0) = e^{-\lambda},$$

$$nf(n) = \sum_{k=1}^n kg(k) f(n-k), \quad n \geq 1, \quad (5)$$

and the following pseudo compound Poisson representation $X \sim \text{PCP}(\lambda, q)$ holds:

$$f(n) = \sum_{k \geq 0} \frac{\lambda^k}{k!} e^{-\lambda} q^{*k}(n), \quad (6)$$

$$q(k) = \frac{1}{\lambda} g(k).$$

Proof. The recursion (5) is an immediate consequence of (3). Indeed, the derivative of the equation $F(s) = f(0) \cdot \exp\{G(s)\}$ satisfies the relation $F'(s) = G'(s) \cdot F(s)$, which is equivalent to identities (5). For further details consult Hürlimann [5, Corollary 2], Hürlimann [6, Theorem 1]. $\square$

Remarks 1. Clearly, if $g(k) \geq 0$, $\forall k \geq 1$, the distribution is compound Poisson with parameter λ and severity probabilities $q(k) \geq 0$, $\forall k \geq 1$, a fact denoted by $X \sim \text{CP}(\lambda, q)$. In this situation, the recursion (5) is named Panjer recursion after Panjer [7]. With Feller [3, Section XII.2], one knows that X is compound Poisson if and only if it is infinitely divisible (see also [6] for a more general characterization). Otherwise, one says that X has a pseudo compound Poisson distribution. In the terminology of Sundt [8], it belongs to

the class $R_{\infty}[0, b]$ with $b(k) = kg(k)$ (see also [9]). The theoretical and practical usefulness of pseudo compound Poisson distributions have been demonstrated by the author in numerous publications. Sometimes, the sequence $\{g(n)\}$ defined by (5) is called the *De Pril transform* of $\{f(n)\}$ after De Pril [10] (e.g., [11]). The interest of recurrence relation (5) extends beyond discrete probability to the general context of integer sequences. Let Id be the identity map such that $\text{Id}(n) = n$ for all nonnegative integers. Then, the convolution equation $\text{Id} \cdot f = g * f$ for integer sequences $\{f(n)\}$ and $\{g(n)\}$ occurs in many areas of mathematics. An important problem is the relationship between the asymptotic behaviors of the two sequences $\{f(n)\}$ and $\{g(n)\}$ (e.g., [12] and its references).

The converse of Proposition 1, although used in applications (e.g., [6, 13–15]), has been less studied. In case of negative $g(k)$'s a necessary condition on the cumulant pgf such that (5) defines a true probability distribution has been first identified by Lévy.

Proposition 2. *Given is a finite generating function $G(s) = \sum_{k=1}^N g(k) \cdot s^k$ with $\lambda = G(1) > 0$. Assume that $g(1), g(N-1), g(N) \neq 0$. In order that (5) defines a discrete probability distribution, it is necessary that the condition (NC) $g(1), g(N-1), g(N) > 0$, is fulfilled.*

Proof. The condition (NC) is found in Lukacs [16, page 252] and Johnson et al. [17, page 356]. Proofs are found in Lévy [18, 19] and Cuppens [20, Section 8.4 and Appendix B]. $\square$

Remarks 2. Precise sufficient conditions (SC) in Proposition 2 are not known. In principle, any of the remaining $g(k)$'s could be negative provided they are sufficiently small in absolute value; that is, $|g(k)| < \varepsilon$, $k \neq 1, N-1, N$, for some $\varepsilon > 0$ (cf. [18, page 263]; [13, Remark 1 to Theorem 3.1]). Lévy [18, page 263] points out that $N = 4$ is the simplest case admitting a negative $g(2) < 0$. Van Harn [21, page 84] obtains four inequality constraints on $g(2) < 0$ in terms of the other $g(k)$'s. Lukacs [16, page 251] shows that if $g(1) = g(3) = g(4) = 1$, $0 < |g(2)| \leq 1/4$, then (5) defines a discrete probability distribution (cf. [13, Remark 3 to Theorem 3.1]).

3. Partition Representation and Combinatorial Interpretation

Besides the pseudo compound Poisson representation the power series identity $F(s) = f(0) \cdot \exp\{G(s)\}$ has another important immediate consequence.

Proposition 3 (partition representation). *Let X be a discrete arithmetic random variable with pgf $F(s) = \sum_{n \geq 0} f(n) \cdot s^n$ such that $f(0) > 0$, and let $G(s) = \ln\{F(s)\} - \ln f(0) = \sum_{k \geq 1} g(k) \cdot s^k$ be its cumulant pgf. Then, the following partition representation holds:*

$$f(n) = f(0) \cdot \sum_{x \in S_n} \prod_{k=1}^n \frac{g(k)^{x_k}}{x_k!}, \quad n \geq 1, \quad (7)$$

where $S_n = \{x = (x_1, \dots, x_n) \mid \sum_{k=1}^n kx_k = n\}$ is the set of free partitions of weight n .

Proof. This result follows from the calculation

$$\begin{aligned} F(s) &= f(0) \cdot \exp \{G(s)\} = f(0) \cdot \exp \left\{ \sum_{k \geq 1} g(k) \cdot s^k \right\} \\ &= f(0) \cdot \prod_{k \geq 1} \exp \{g(k) \cdot s^k\} \\ &= f(0) \cdot \prod_{k \geq 1} \sum_{x_k} \frac{\{g(k) \cdot s^k\}^{x_k}}{x_k!} \\ &= f(0) \cdot \sum_{n \geq 0} \left(\sum_{x \in S_n} \prod_{k=1}^n \frac{g(k)^{x_k}}{x_k!} \right) \cdot s^n, \end{aligned} \quad (8)$$

which implies identities (7). $\square$

Remarks 3. Although variants existed before (e.g., [22, Section 2]; [23, Equation (3.4)]), the general partition representation (7) for pseudo compound Poisson arithmetic distributions has been first stated in Corollary 3.1 of Hürlimann [4]. Unfortunately, the given formula contains a misprint: the multiplicity factorial $x_k!$ has been wrongly replaced by $k!$. Note that this mistake remains without influence on the validity of Theorem 3.2 derived from this representation. However, the claimed but missing proof by induction is best replaced by the present proof. The simple power series manipulation has already been used by Macdonald [24] in his proof on page 25 of a similar but more special result in the theory of symmetric functions (see later Section 4.3). Section I.1 of the mentioned textbook is recommended reading for definitions and properties around partitions (in particular weight, length, and multiplicity of a partition).

The partition representation has a nice combinatorial interpretation. For convenience, set $h(k) = kg(k)$, $k \geq 1$. Then, the expression

$$\frac{f(n)}{f(0)} = \sum_{x \in S_n} \prod_{k=1}^n \frac{h(k)^{x_k}}{x_k! k^{x_k}}, \quad n \geq 1, \quad (9)$$

derived from (7) identifies with the cycle index polynomial (or cycle indicator) of the symmetric group of order n in the variables $h(k)$ (e.g., [24, Example 9.(a), page 29]). Indeed, the coefficient in (9) of an arbitrary monomial $h(1)^{x_1} \cdot h(2)^{x_2} \cdot \dots \cdot h(n)^{x_n}$ is equal to the fraction of all permutations that have x_1 fixed points, x_2 cycles of length 2, $\dots$, and x_n cycles of length n . We note that this combinatorial interpretation is crucial to the novel example studied in the next two sections.

4. The Convolution of Geometric Random Variables: A Newton Type Distribution

To illustrate further the discrete probability concepts of Sections 2 and 3, consider the power sum like specification of the cumulant pgf in (3); namely,

$$g(k) = g(k; \theta) = k^{-1} \cdot \sum_{j=1}^m \theta_j^k, \quad k \geq 1, \quad (10)$$

where $\theta = (\theta_1, \theta_2, \dots, \theta_m)$, $0 < \theta_j < 1$, $\theta_i \neq \theta_j$, $i, j = 1, \dots, m$, are m unknown parameters. A calculation shows that the quantity $\lambda = \sum_{k \geq 1} g(k) = -\sum_{j=1}^m \ln(1 - \theta_j) > 0$ is finite. The condition (C1) of Proposition 2 is fulfilled and therefore (10) belongs to the random variable X of some compound Poisson distribution, to be determined. Clearly, one has

$$G(s) = \sum_{k \geq 1} g(k) \cdot s^k = -\sum_{j=1}^m \ln(1 - \theta_j s), \quad (11)$$

$$f(0) = e^{-\lambda} = \prod_{j=1}^m (1 - \theta_j),$$

from which one obtains

$$F(s) = f(0) \cdot \exp \{G(s)\} = \prod_{j=1}^m \frac{1 - \theta_j}{1 - \theta_j s}, \quad (12)$$

which is the pgf of the *convolution of m geometric random variables*; hence

$$X =_d \bigoplus_{j=1}^m \text{Geom}(1 - \theta_j). \quad (13)$$

Consider the partition representation of Proposition 3. One observes that $p(k) = kg(k) = \sum_{j=1}^m \theta_j^k$ is nothing else other than the k th *power sum polynomial* in the variables θ_j , $j = 1, \dots, m$. Identity (9) is herewith equal to

$$\frac{f(n)}{f(0)} = \sum_{x \in S_n} \prod_{k=1}^n \frac{p(k)^{x_k}}{x_k! k^{x_k}}, \quad n \geq 1, \quad (14)$$

and shows with Macdonald [24, Equation (2.14'), page 25] that the n th *probability* of X , namely, $f(n) = f(0) \cdot h(n)$, is the $f(0)$ multiple of the n th *complete symmetric function* $h(n)$ in the variables θ_j , $j = 1, \dots, m$. In this combinatorial context the Panjer recursion (5) is equivalent to the identities

$$nh(n) = \sum_{k=1}^n p(k) h(n-k), \quad n \geq 1, \quad h(0) = 1, \quad (15)$$

which is nothing else other than a variant of *Newton's identities* (see [24, Equations (2.11) and (2.11'), page 23]). Therefore, the distribution (14) of the convolution of geometric random variables derived from the recursion (15) could also be called *Newton distribution*. The distribution corresponding to the random variable (13) is abbreviated by $X \sim N_m(\theta_1, \dots, \theta_m)$.

Now, we apply our concepts to study some estimation properties of the Newton distribution. In a first step, we construct a parameter vector $\delta = (\delta_2, \delta_3, \dots, \delta_m)$ orthogonal to the mean

$$\mu = \mu(\theta) = \sum_{k \geq 1} kg(k) = \sum_{j=1}^m \frac{\theta_j}{1 - \theta_j}. \quad (16)$$

In general, given a random variable X we suppose that the mean μ is functionally independent of a certain parameter vector $\delta = \delta(\theta)$, that is, $\partial\mu/\partial\delta_k = 0$, $k = 2, \dots, m$, and denote the log-likelihood of X by $\ell(x; \mu, \delta)$. The mean μ is *orthogonal* to the parameter vector δ , denoted by $\mu \perp \delta$, if one has $E[(\partial^2 \ell / \partial \mu \partial \delta_k)] = E[(\partial \ell / \partial \mu)(\partial \ell / \partial \delta_k)] = 0$, $k = 2, \dots, m$. Further, let $\bar{C}$ be the class of all discrete arithmetic distributions for which the maximum likelihood estimator of the mean is the sample mean, that is, such that $\hat{\mu} = \bar{X}$ (Gauss's principle). The subclass $\bar{C}^\perp$ of $\bar{C}$, called mean orthogonal class, consists of all those distributions, which besides Gauss's principle satisfy the *mean orthogonal property* $\mu \perp \delta$. It is known that the class $\bar{C}^\perp$ is closed under convolution (see [15, Theorem 2.2]) and can be characterized as follows. Suppose there exist a parameter α and a one-to-one coordinate transformation mapping (α, δ) to (μ, δ) , and set $\lambda = -\ln \Pr(X = 0)$. Then $X \in \bar{C}^\perp$ with $\mu = \mu(\alpha, \delta) \perp \delta$ is equivalent to the following condition (see [15, Lemma 3.1, Equation (3.8)]):

$$\mu \cdot \frac{\partial g(k)}{\partial \alpha} = \frac{\partial \lambda}{\partial \alpha} \cdot kg(k), \quad k \geq 1, \quad (17)$$

where $g(k)$, $k \geq 1$, is determined by the pseudo compound Poisson representation of X . Applied to the Newton distribution, one sees after some calculation that (17) with $\alpha = \theta_1$, $\theta_{j,1} = \partial\theta_j/\partial\theta_1$, $j = 2, \dots, m$, is fulfilled provided the partial differential equation

$$\begin{aligned} & \sum_{j=2}^m (\theta_1 \theta_{j,1} - \theta_j) \cdot \left(\frac{\theta_1^{k-1}}{1 - \theta_j} - \frac{\theta_j^{k-1}}{1 - \theta_1} \right) \\ & + \sum_{2 \leq i < j \leq m} (\theta_j \theta_{i,1} - \theta_i \theta_{j,1}) \cdot \left(\frac{\theta_j^{k-1}}{1 - \theta_i} - \frac{\theta_i^{k-1}}{1 - \theta_j} \right) \\ & = 0, \quad k \geq 1, \end{aligned} \quad (18)$$

can be solved. It is easy to see that this will be fulfilled provided one has

$$\begin{aligned} \theta_1 \theta_{j,1} &= \theta_j, \quad j = 2, \dots, m, \\ \theta_j \theta_{i,1} &= \theta_i \theta_{j,1}, \quad i \leq i < j \leq m. \end{aligned} \quad (19)$$

A solution to these partial differential equations is

$$\theta_j = \delta_j \theta_1, \quad \delta_j \in (0, \theta_1^{-1}), \quad \delta_j \neq 1, \quad j = 2, \dots, m. \quad (20)$$

It follows that the mean μ is orthogonal to the parameter vector $\delta = \delta(\theta) = (\theta_2/\theta_1, \dots, \theta_m/\theta_1)$.

Next, we derive the maximum likelihood equations for the two-parameter Newton distribution $N_2(\alpha, \beta)$ obtained from the convolution $X = \text{Geom}(1 - \alpha) \oplus \text{Geom}(1 - \beta)$ of two geometric distributions with parameters $0 < \alpha, \beta < 1, \beta \neq \alpha$. The probabilities in the above orthogonal parameterization $(\alpha, \delta) = (\alpha, \beta/\alpha)$ are given by

$$f(n) = (1 - \alpha)(1 - \delta\alpha)\alpha^n \left(\frac{1 - \delta^{n+1}}{1 - \delta} \right), \quad n \geq 0. \quad (21)$$

For each $i \geq 0$ let N_i be the number of observations of the random variable X equal to i , and let $N = \sum_{i=0}^k N_i$ be the total number of observations, where $N_i = 0$ for all $i > k$. From (21) one obtains the scaled log-likelihood function

$$\begin{aligned} \frac{1}{N} \ell(X; \alpha, \delta) &= \ln(1 - \alpha) + \ln(1 - \delta\alpha) \\ &+ \frac{1}{N} \sum_{i=0}^k i N_i \ln(\alpha) \\ &+ \frac{1}{N} \sum_{i=0}^k (i + 1) N_i \ln \left(\frac{1 - \delta^{i+1}}{1 - \delta} \right). \end{aligned} \quad (22)$$

The maximum likelihood equations $N^{-1} \ell_\alpha(X; \alpha, \delta) = N^{-1} \ell_\delta(X; \alpha, \delta) = 0$ are given by

$$\begin{aligned} \mu &= \frac{\alpha}{1 - \alpha} + \frac{\delta\alpha}{1 - \delta\alpha} = \bar{X}, \quad \bar{X} = \frac{1}{N} \sum_{i=0}^k i N_i, \\ \frac{\delta\alpha}{1 - \delta\alpha} &= \frac{\delta}{1 - \delta} - \frac{1}{N} \sum_{i=0}^k (i + 1) N_i \frac{\delta^{i+1}}{1 - \delta^{i+1}}. \end{aligned} \quad (23)$$

This system of two nonlinear equations in the two variables (α, δ) can be solved with substitution setting $x = \delta\alpha/(1 - \delta\alpha)$. Insert $\alpha = x/\delta(1 + x)$ in the first equation to see that

$$(1 - \delta)x^2 - [1 + \delta + (1 - \delta)\mu]x + \delta\mu = 0. \quad (24)$$

The only feasible solution of this quadratic equation takes the minus sign and is given by

$$\begin{aligned} x = x(\delta, \mu) &= \frac{1}{2(1 - \delta)} \left\{ 1 + \delta + (1 - \delta)\mu \right. \\ &\quad \left. - \sqrt{(1 + \delta)^2 + (1 - \delta)^2 \mu(\mu + 2)} \right\}. \end{aligned} \quad (25)$$

Using that $\mu = \bar{X}$ and inserting (25) into the second equation of (23) one sees that the latter depends besides the observations only upon δ and is determined by

$$\frac{\delta}{1 - \delta} - x(\delta, \bar{X}) - \frac{1}{N} \sum_{i=0}^k (i + 1) N_i \frac{\delta^{i+1}}{1 - \delta^{i+1}} = 0. \quad (26)$$

The preceding simple derivation of maximum likelihood properties is possible in virtue of the pseudo compound Poisson representation (mainly (17)). Let us also illustrate the mathematical benefit of the partition representation. For this

we rely on Theorem 3.2 in Hürlimann [4], which guarantees the existence of well-defined confidence bounds for the mean of a count distribution provided the following regularity assumption is fulfilled.

(RA) There exist a positive integer m and a constant $c(\mu, \delta)$ such that $E[X^m] < \infty$ and $\sum_{j=1}^k x_j (\partial \ln g(j) / \partial \mu) \leq k^m c(\mu, \delta)$ for all partitions $x \in S_k$ such that $\sum_{j=1}^k j x_j = k$, $k \geq 1$.

Now, for the Newton distribution with two parameters such that $g(j) = (1 + \delta^j) \alpha^j / j$, $j \geq 1$, one has

$$\begin{aligned} \frac{\partial \ln g(j)}{\partial \mu} &= \frac{\partial \ln g(j)}{\partial \alpha} \cdot \frac{\partial \alpha(\mu, \delta)}{\partial \mu} \\ &= \frac{j}{\alpha(\mu, \delta)} \cdot \frac{\partial \alpha(\mu, \delta)}{\partial \mu}, \end{aligned} \quad (27)$$

and thus $\sum_{j=1}^k x_j (\partial \ln g(j) / \partial \mu) = (k / \alpha(\mu, \delta)) \cdot (\partial \alpha(\mu, \delta) / \partial \mu)$. The regularity assumption with $m = 1$ is fulfilled.

5. A Numerical Example from Automobile Insurance

The choice of an appropriate claim number model in automobile insurance is a prominent actuarial research topic whose last 50-year-old story begins with Bichsel [25], who used the *negative binomial* (NB) distribution to construct a bonus-malus system. Among the other first good choices, the *Poisson inverse Gaussian* (PIG) has often been advocated (see Example 3.4 in [4] and discussions). Due to the penalty in the SBC (Schwartz Bayesian criterion) score the best overall fit is in any case obtained for a two-parameter distribution (see [4, end of Section 4]). In this context, we ask whether the Newton distribution might compete with the NB and PIG. For a specific example and in a very precise sense (regrouping of classes), it beats both of them and at the same time improves the unsatisfactory fitting of data set 4 in Table 4.5 of Hürlimann [4].

Since the method of moments is inappropriate, as shown by Gossiaux and Lemaire [26], we use the maximum likelihood estimation (MLE) method, which is best in view of its asymptotic properties. The goodness-of-fit is established on the basis of both the SBC score and the p value of the chi-square statistic with appropriate regrouping of the classes. The improved results are found in Tables 1 and 2. Although the minimum SBC score is attained for the PIG the chi-square and p value (with the last two regrouped classes) are best for the Newton distribution. The decrease in goodness-of-fit of the PIG is due to the loose in fit by regrouping the last two classes and the deficiency in fitting the class with two observed insurance claims.

Let us conclude the present work. Our aim was to give a systematic brief overview of the concepts of pseudo compound Poisson and partition representations in discrete probability and present a new application. Through the detailed study of a single new example, we have exemplified some typical features and further potential of our methodology.

TABLE 1: Data set and MLE fit with the last two regrouped classes.

# of claims	Observations	Discrete arithmetic distribution		
		Newton	NB	PIG
0	96978	96981.10	96979.96	96978.53
1	9240	9231.66	9235.87	9240.40
2	704	706.52	702.06	697.63
3	43	54.71	56.12	57.44
4	9	—	—	—

TABLE 2: Goodness-of-fit (chi-square and p value with the last two regrouped classes).

Distribution	SBC score	Chi-square value	p value in %
Newton	36123.40	0.15132	69.73
NB	36123.14	0.30936	57.81
PIG	36123.06	0.57386	44.87

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

References

- [1] S. K. Katti, "Infinite divisibility of integer-valued random variables," *Annals of Mathematical Statistics*, vol. 38, pp. 1306–1308, 1967.
- [2] F. Steutel, *Preservation of Infinite Divisibility under Mixing and Related Results*, vol. 33 of *Mathematical Centre Tracts*, Mathematical Centre, Amsterdam, The Netherlands, 1970.
- [3] W. Feller, *An Introduction to Probability and Its Applications*, vol. 1, John Wiley & Sons, 3rd edition, 1968.
- [4] W. Hürlimann, "Robust confidence bounds for the mean of some count data models," *Blätter der DGVFM*, vol. 25, no. 4, pp. 795–811, 2002.
- [5] W. Hürlimann, "On maximum likelihood estimation for count data models," *Insurance: Mathematics and Economics*, vol. 9, no. 1, pp. 39–49, 1990, Correction note: *Insurance: Mathematics and Economics*, vol. 10, p. 81, 1991.
- [6] W. Hürlimann, "Pseudo compound poisson distributions in risk theory," *ASTIN Bulletin*, vol. 20, no. 1, pp. 57–79, 1990.
- [7] H. H. Panjer, "Recursive evaluation of a family of compound distributions," *The Astin Bulletin*, vol. 12, no. 1, pp. 22–26, 1981.
- [8] B. Sundt, "On some extensions of Panjer's class of counting distributions," *ASTIN Bulletin*, vol. 22, no. 1, pp. 61–80, 1992.
- [9] B. Sundt and R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications*, vol. 2 of *EAA Lecture Notes*, Springer, 2009.
- [10] N. De Pril, "The aggregate claims distribution in the individual model of risk theory," *ASTIN Bulletin*, vol. 19, no. 1, pp. 9–24, 1989.
- [11] J. Dhaene and B. Sundt, "On approximating distributions by approximating their De Pril transforms," *Scandinavian Actuarial Journal*, vol. 1998, no. 1, pp. 1–23, 1998.
- [12] A. Baltrunas and E. Omey, "Second-order subexponential sequences and the asymptotic behavior of their De Pril transform," *Lithuanian Mathematical Journal*, vol. 41, no. 1, pp. 17–27, 2001.

- [13] R. K. Milne and M. Westcott, "Generalized multivariate Hermite distributions and related point processes," *Annals of the Institute of Statistical Mathematics*, vol. 45, no. 2, pp. 367–381, 1993.
- [14] W. Hürlimann, "Quasi-exact numerical evaluation of collateralized debt obligations prices," *Journal of Numerical Mathematics and Stochastics*, vol. 2, no. 1, pp. 64–75, 2010.
- [15] W. Hürlimann, "A characterization of the compound multiparameter hermite gamma distribution via Gauss's principle," *The Scientific World Journal*, vol. 2013, Article ID 468418, 2013.
- [16] E. Lukacs, *Characteristic Functions*, Griffin, UK, 2nd edition, 1970.
- [17] L. N. Johnson, S. Kotz, and A. W. Kemp, *Univariate Discrete Distributions*, John Wiley & Sons, New York, NY, USA, 2nd edition, 1992.
- [18] P. Lévy, "Sur les exponentielles de polynômes et sur l'arithmétique des produits de lois de Poisson," *Annales Scientifiques de l'École Normale Supérieure Série III*, vol. 54, pp. 231–292, 1937.
- [19] P. Lévy, "L'arithmétique des lois de probabilités et les produits finis de lois de Poisson," *Actualités Scientifiques et Industrielles*, no. 736, pp. 25–59, 1938, (Colloque de Genève III), Hermann, Paris, France.
- [20] R. Cuppens, *Decomposition of Multivariate Probability*, vol. 29 of *Probability and Mathematical Statistics*, Academic Press, New York, NY, USA, 1975.
- [21] K. Van Harn, "Classifying infinitely distributions by functional equations," Mathematical Centre Tract 103, Mathematisch Centrum, Amsterdam, The Netherlands, 1978.
- [22] W. Hürlimann, "Negative claim amounts, bessel functions, linear programming and Miller's algorithm," *Insurance: Mathematics and Economics*, vol. 10, no. 1, pp. 9–20, 1991.
- [23] W. Hürlimann, "Predictive Stop-Loss Premiums," *ASTIN Bulletin*, vol. 23, no. 1, pp. 55–76, 1993.
- [24] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, Oxford Mathematical Monographs, Oxford University Press, New York, NY, USA, 2nd edition, 1995.
- [25] F. Bichsel, "Erfahrungs-Tarifierung in der Motorfahrzeughaftpflicht-Versicherung," *Bulletin of the Swiss Association of Actuaries*, vol. 64, pp. 119–143, 1964.
- [26] A. Gossiaux and J. Lemaire, "Méthodes d'ajustement de distributions de sinistres," *Bulletin of the Association of Swiss Actuaries*, vol. 1981, pp. 87–95, 1981.