



A new parametric mixture class of distributions on \mathbb{Z}

Christophe Chesneau, Maher Kachour

► To cite this version:

Christophe Chesneau, Maher Kachour. A new parametric mixture class of distributions on \mathbb{Z} . 2016. hal-01381020v1

HAL Id: hal-01381020

<https://hal.science/hal-01381020v1>

Preprint submitted on 13 Oct 2016 (v1), last revised 24 Nov 2016 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A NEW PARAMETRIC MIXTURE CLASS OF DISTRIBUTIONS ON \mathbb{Z}

Christophe Chesneau

Laboratoire de Mathématiques Nicolas Oresme (LMNO)

Université de Caen Basse-Normandie

Campus II, Science 3, 14032 Caen, France.

christophe.chesneau@unicaen.fr

Maher Kachour

Ecole supérieure de commerce IDRAC

47, rue Sergent Michel Berthet

69258 Lyon Cedex 09, France

maher.kachour@idraclyon.com

Abstract. In this paper we introduce, using the finite mixture distribution approach, a new class of discrete distributions with support belongs to \mathbb{Z} , denoted by Parametric Mixture- \mathbb{Z} class. Various mathematical properties of the new class are derived. Estimation procedure, under additional parametric assumption, is also assessed by a numerical study. Finally, a real data example is considered.

Mathematics Subject Classification. 62M10, 62M20.

Keywords. Mixtures, Discrete distributions, Mixing distribution, Rademacher distribution.

1 Introduction

Let k be a fixed non-negative integer. A parametric family of k mixture densities has a probability density function defined as follow

$$p_{\Phi}(x) = \sum_{i=1}^k \alpha_i p_{\phi_i}(x),$$

where, for all $1 \leq i \leq k$, p_{ϕ_i} is a probability distribution with ϕ_i as parameter, $\alpha_i \geq 0$, $\sum_{i=1}^k \alpha_i = 1$, and $\Phi = (\alpha_1, \dots, \alpha_k, \phi_1, \dots, \phi_k)$. A k mixture densities can be interpreted as distribution associated to a population which is composed of k subpopulations, where, for all $1 \leq i \leq k$, α_i represents the proportion of the i^{th} subpopulation and p_{ϕ_i} its associated distribution. Thus, mixture distributions allow for heterogeneity in the sample. For more background information on the finite mixture distributions we refer to McLachlan and Batsford (1988) and Johnson et al. (2005). The finite discrete mixture distributions occur in many practical situations. For example, in actuarial statistics they are used for modeling the number of claims incurred during a given period for an insurance portfolio. The finite discrete mixture distributions topic have been largely investigated in literature, in particular, the

mixture of standard discrete distribution, such as Binomial, Geometric, and Poisson. For a literature review, we refer to Rider (1961), Blischke (1962), Everitt (1981), Harris (1983), Pritchard et al. (2000), Karlis and Xekalaki (2005) and Titterington (2005).

The aim of this paper is to introduce, based on the finite mixture distribution approach, a new class of discrete distributions with support belonging to \mathbb{Z} , denoted by Parametric Mixture- \mathbb{Z} class. Explicitly, let $p \in (0, 1)$. Consider X and Y , two discrete non-negative random variables. We said that a random variable Z has a distribution belonging to the Parametric Mixture- \mathbb{Z} class if and only if its probability mass function (pmf) has the following representation :

$$\mathbb{P}(Z = k) = p\mathbb{P}(X = k) + (1 - p)\mathbb{P}(Y = -k), \quad k \in \mathbb{Z}. \quad (1.1)$$

The above pmf can naturally be interpreted as distribution associated with a statistical population which is mixture of two subpopulations : positive values (with p as mixing proportion and associated distribution that of the random variable X) and negative values (with $1 - p$ as mixing proportion and associated distribution that of the random variable Y).

Remark 1 *Note that the Parametric Mixture(p) - \mathbb{Z} class can be considered as a generalization of the Rademacher(p) - \mathbb{N} class originally introduced by Chesneau and Kachour (2012). Indeed, if X and Y have the same distribution than a random variable T , then Z can be rewritten as the following form*

$$Z = RT,$$

where R is a Rademacher random variable with parameter p and independent of T . In particular, if T is a Poisson random variable with parameter $\lambda > 0$, then Z follows the Extended Poisson (E-Po) distribution studied by Bakouch et al. (2016).

Remark 2 *If Z belongs to Parametric Mixture(p) - \mathbb{Z} class, then $-Z$ belongs to Parametric Mixture(q) - \mathbb{Z} class, where $q = 1 - p$.*

Remark 3 *If Z belongs to Parametric Mixture(p) - \mathbb{Z} class, then probability mass function (pmf) of $|Z|$ has the following representation*

$$\mathbb{P}(|Z| = k) = p\mathbb{P}(X = k) + (1 - p)\mathbb{P}(Y = k), \quad k \in \mathbb{N}.$$

Thus, $|Z|$ has a 2-mixture discrete distribution.

The contents of this paper are organized as follows. Various mathematical properties of random variable belongs to the Parametric Mixture(p) - \mathbb{Z} are derived in Section 2. These include, the failure functions (Section 2.1), the probability generating function and moments (Section 2.2), the mean absolute deviation (Section 2.3), the Shannon entropy (Section 2.4), the mode (Section 2.5), and the log-likelihood function (Section 2.6). In section 3, we investigate numerically the parameters estimation in a special case where the distribution of X and that of Y are fixed. Finally, Section 4 illustrates practical usefulness of our class via an application on real data.

2 Mathematical properties

Throughout this section we consider that Z belongs to the Parametric Mixture- \mathbb{Z} class, such that its probability mass function is defined by (1.1).

2.1 Failure functions

Let $F_Z(x) = \mathbb{P}(Z \leq x)$ denote the cumulative distribution function of Z and $\lfloor \cdot \rfloor$ the floor function. Let $F_X(x) = \mathbb{P}(X \leq x)$ (resp. $F_Y(x) = \mathbb{P}(Y \leq x)$) denote the cumulative distribution function of X (resp. Y).

For all $x < 0$, we have

$$\begin{aligned} F_Z(x) &= \sum_{k=-\infty}^{\lfloor x \rfloor} \mathbb{P}(Z = k) \\ &= (1-p) \sum_{k=-\lfloor x \rfloor}^{\infty} \mathbb{P}(Y = k) \\ &= (1-p) \mathbb{P}(Y \geq -\lfloor x \rfloor) \\ &= (1-p)(1 - F_Y(-\lfloor x \rfloor - 1)). \end{aligned}$$

Suppose now that $0 \leq x < 1$, then

$$\begin{aligned} F_Z(x) &= F_Z(-1) + \mathbb{P}(Z = 0) \\ &= (1-p)(1 - \mathbb{P}(Y = 0)) + p\mathbb{P}(X = 0) + (1-p)\mathbb{P}(Y = 0) \\ &= (1-p) + p\mathbb{P}(X = 0). \end{aligned}$$

Finally, for $x \geq 1$, we have

$$\begin{aligned} F_Z(x) &= \sum_{k=-\infty}^{\lfloor x \rfloor} \mathbb{P}(Z = k) \\ &= F_Z(0) + \sum_{k=1}^{\lfloor x \rfloor} \mathbb{P}(Z = k) \\ &= F_Z(0) + p \sum_{k=1}^{\lfloor x \rfloor} \mathbb{P}(X = k) \\ &= (1-p) + p\mathbb{P}(X = 0) + p \sum_{k=1}^{\lfloor x \rfloor} \mathbb{P}(X = k) \\ &= (1-p) + pF_X(\lfloor x \rfloor). \end{aligned}$$

Let R_Z , k_Z , and τ_Z denote, respectively, the survival function, failure rate function,

and the reverse failure function of Z . Then, we have

$$\begin{aligned}
 R_Z(x) &= \mathbb{P}(Z > x) = 1 - F_Z(x) \\
 &= \begin{cases} 1 - (1 - p)R_Y(-\lfloor x \rfloor - 1), & \text{if } x < 0, \\ p(1 + \mathbb{P}(X = 0)) - 2p(1 - p)\mathbb{P}(Y = 0), & \text{if } 0 \leq x < 1, \\ pR_X(x) + 2(1 - p)\mathbb{P}(Y = 0), & \text{if } x \geq 1, \end{cases}
 \end{aligned}$$

$$\begin{aligned}
 k_Z(x) &= \mathbb{P}(Z = x \mid Z \geq x) = \frac{\mathbb{P}(Z = x)}{R_Z(x - 1)} \\
 &= \begin{cases} \frac{p\mathbb{P}(X = x)}{pR_X(x - 1) + 2(1 - p)\mathbb{P}(Y = 0)}, & \text{if } x \geq 2, \\ \frac{p\mathbb{P}(X = 1)}{p(1 + \mathbb{P}(X = 0)) - 2p(1 - p)\mathbb{P}(Y = 0)}, & \text{if } x = 1, \\ \frac{p\mathbb{P}(X = 0) + (1 - p)\mathbb{P}(Y = 0)}{p(1 + \mathbb{P}(X = 0)) - 2p(1 - p)\mathbb{P}(Y = 0)}, & \text{if } x = 0, \\ \frac{(1 - p)\mathbb{P}(Y = -x)}{1 - (1 - p)R_Y(-\lfloor x \rfloor - 1)}, & \text{if } x \leq -1, \end{cases}
 \end{aligned}$$

and

$$\begin{aligned}
 \tau_Z(x) &= \mathbb{P}(Z = x \mid Z \leq x) = \frac{\mathbb{P}(Z = x)}{F_Z(x)} \\
 &= \begin{cases} k_Y(-x), & \text{if } x < 0, \\ \frac{p\mathbb{P}(X = 0) + (1 - p)\mathbb{P}(Y = 0)}{(1 - p)(1 - 2\mathbb{P}(Y = 0)) + p\mathbb{P}(X = 0)}, & \text{if } 0 \leq x < 1, \\ \frac{p\mathbb{P}(X = x)}{(1 - p)(1 - 2\mathbb{P}(Y = 0)) + pF_X(x)}, & \text{if } x \geq 1, \end{cases}
 \end{aligned}$$

where R_Y (resp. R_X) is the the survival function of Y (resp. X) and k_Y its failure rate function.

2.2 Probability generating function and Moments

Let G_X (resp. G_Y) be the pgf of X (resp. Y). Define

$$\begin{aligned} G_{Z,+}(s) &= \sum_{k=0}^{\infty} s^k \mathbb{P}(Z = k) \\ &= \mathbb{P}(Z = 0) + \sum_{k=1}^{\infty} s^k \mathbb{P}(Z = k) \\ &= pG_X(s) + (1-p)\mathbb{P}(Y = 0), \end{aligned}$$

and

$$\begin{aligned} G_{Z,-}\left(\frac{1}{s}\right) &= \sum_{k=-\infty}^0 \left(\frac{1}{s}\right)^k \mathbb{P}(Z = k) \\ &= \mathbb{P}(Z = 0) + \sum_{k=-\infty}^{-1} \left(\frac{1}{s}\right)^k \mathbb{P}(Z = k) \\ &= (1-p)G_Y(s) + p\mathbb{P}(X = 0). \end{aligned}$$

Thus, one can remark that

$$\mathbb{P}(Z = k) = \begin{cases} \frac{G_{Z,+}^{(k)}(0)}{k!}, & \text{if } k > 0, \\ \frac{G_{Z,-}^{(|k|)}(0)}{|k|!}, & \text{if } k < 0, \\ G_{Z,+}(0), & \text{if } k = 0. \end{cases}$$

Now, for any integer $n \geq 1$, using the moment-generating function properties, we know that

$$\mathbb{E}(Z^n) = M_Z^{(n)}(0),$$

where

$$\begin{aligned} M_Z^{(n)}(t) &= \mathbb{E}(e^{tZ}) = \sum_{k=-\infty}^{+\infty} e^{tk} \mathbb{P}(Z = k) \\ &= \sum_{k=-\infty}^{-1} e^{tk} \mathbb{P}(Z = k) + \mathbb{P}(Z = 0) + \sum_{k=1}^{+\infty} e^{tk} \mathbb{P}(Z = k) \\ &= (1-p)G_Y\left(\frac{1}{e^t}\right) + pG_X(e^t) \\ &= (1-p)G_Y\left(\frac{1}{e^t}\right) + pM_X(t). \end{aligned}$$

Moreover, we deduce that

$$G_Y^{(n)}\left(\frac{1}{e^t}\right) = \begin{cases} -\frac{1}{e^t}M_Y(t), & \text{if } n \text{ is odd,} \\ \frac{1}{e^t}M_Y(t), & \text{if } n \text{ is even.} \end{cases}$$

Therefore, we have

$$\mathbb{E}(Z^n) = \begin{cases} p\mathbb{E}(X^n) - (1-p)\mathbb{E}(Y^n), & \text{if } n \text{ is odd,} \\ p\mathbb{E}(X^n) + (1-p)\mathbb{E}(Y^n), & \text{if } n \text{ is even} \end{cases} \quad (2.1)$$

Thus, the first four moments of Z are

- $\mathbb{E}(Z) = p\mathbb{E}(X) - (1-p)\mathbb{E}(Y)$,
- $\mathbb{E}(Z^2) = p\mathbb{E}(X^2) + (1-p)\mathbb{E}(Y^2)$,
- $\mathbb{E}(Z^3) = p\mathbb{E}(X^3) - (1-p)\mathbb{E}(Y^3)$,
- $\mathbb{E}(Z^4) = p\mathbb{E}(X^4) + (1-p)\mathbb{E}(Y^4)$.

The corresponding variance, coefficient of variation, skewness and kurtosis are

$$\mathbb{V}(Z) = p\mathbb{E}(X^2) + (1-p)\mathbb{E}(Y^2) - p^2(\mathbb{E}(X))^2 - (1-p)^2(\mathbb{E}(Y))^2 + 2p(1-p)\mathbb{E}(X)\mathbb{E}(Y),$$

$$\text{CV}(Z) = \frac{\sigma_Z}{p\mathbb{E}(X) - (1-p)\mathbb{E}(Y)},$$

$$\text{Skewness}(Z) = \frac{\mathbb{E}(Z^3) - 3\mathbb{E}(Z^2)\mathbb{E}(Z) + 2(\mathbb{E}(Z))^3}{\sigma_Z^3},$$

and

$$\text{Kurtosis}(Z) = \frac{\mathbb{E}(Z^4) - 4\mathbb{E}(Z^3)\mathbb{E}(Z) + 6\mathbb{E}(Z^2)(\mathbb{E}(Z))^2 - 3(\mathbb{E}(Z))^4}{\sigma_Z^4} - 3,$$

where $\sigma_Z = \sqrt{\mathbb{V}(Z)}$.

The index of dispersion is defined as the ratio of the variance to the mean, i.e.,

$$\begin{aligned} ID(Z) &= \frac{\mathbb{V}(Z)}{\mathbb{E}(Z)} \\ &= \frac{p\mathbb{E}(X^2) + (1-p)\mathbb{E}(Y^2) - p^2(\mathbb{E}(X))^2 - (1-p)^2(\mathbb{E}(Y))^2 + 2p(1-p)\mathbb{E}(X)\mathbb{E}(Y)}{p\mathbb{E}(X) - (1-p)\mathbb{E}(Y)}. \end{aligned}$$

Obviously, if $p = 1$ (resp. $p = 0$), then Z has the same index of dispersion of X (resp. Y). Moreover, $ID(Z) > 0$ (resp. $ID(Z) < 0$) if and only if $\mathbb{E}(X) > \frac{1-p}{p}\mathbb{E}(Y)$ (resp. $\mathbb{E}(X) < \frac{1-p}{p}\mathbb{E}(Y)$).

The characteristic function of Z is given by

$$\varphi_Z(t) = \mathbb{E}(e^{itZ}) = p\varphi_X(t) + (1-p)\varphi_Y(-t), \quad (2.2)$$

where φ_X (resp. φ_Y) is the characteristic function of X (resp. Y).

2.3 Mean Absolute Deviation

The Mean Absolute Deviation (MAD) of the random variable Z with respect to any measure of central tendency $m = m(Z)$ is given by

$$\begin{aligned} MAD(Z) &= \mathbb{E}(|Z - m|) = \sum_{k=-\infty}^{\infty} |k - m| \mathbb{P}(Z = k) \\ &= \sum_{k=-\infty}^{\lfloor m \rfloor} (m - k) \mathbb{P}(Z = k) + \sum_{k=\lfloor m \rfloor+1}^{\infty} (k - m) \mathbb{P}(Z = k) \\ &= 2mF_Z(\lfloor m \rfloor) - m - \mathbb{E}(Z) + 2 \sum_{k=\lfloor m \rfloor+1}^{\infty} k \mathbb{P}(Z = k), \end{aligned}$$

where F_Z denote the cumulative distribution function of Z , $\lfloor \cdot \rfloor$ the floor function and

$$\sum_{k=\lfloor m \rfloor+1}^{\infty} k \mathbb{P}(Z = k) = \begin{cases} p \left(\mathbb{E}(X) - \sum_{k=1}^{\lfloor m \rfloor} k \mathbb{P}(X = k) \right), & \text{if } \lfloor m \rfloor \geq -1, \\ p \mathbb{E}(X) - (1-p) \sum_{k=1}^{-\lfloor m \rfloor-1} k \mathbb{P}(Y = k), & \text{if } \lfloor m \rfloor < -1. \end{cases}$$

In particular, if $m = \mathbb{E}(Z)$ (resp. $m = \text{Median}(Z)$), the MAD is then called the mean absolute deviation about mean (resp. median).

2.4 Shannon entropy

Let P_Z denotes the probability mass function of Z . The Shannon entropy is given by

$$\begin{aligned}
H(Z) &= \mathbb{E}(-\log(P_Z)) = - \sum_{k=-\infty}^{\infty} P_Z(k) \log(P_Z(k)) \\
&= -P_Z(0) \log(P_Z(0)) - \sum_{k=-\infty}^{-1} P_Z(k) \log(P_Z(k)) - \sum_{k=1}^{\infty} P_Z(k) \log(P_Z(k)) \\
&= -P_Z(0) \log(P_Z(0)) - (1-p) \sum_{k=1}^{\infty} P_Y(k) \log((1-p)P_Y(k)) - p \sum_{k=1}^{\infty} P_X(k) \log(pP_X(k)) \\
&= (1-p)P_Y(0) (-\log(P_Z(0)) + \log(1-p) + \log(P_Y(0))) \\
&\quad + pP_X(0) (-\log(P_Z(0)) + \log(p) + \log(P_X(0))) \\
&\quad + pH(X) + (1-p)H(Y) - p \log(p) - (1-p) \log(1-p),
\end{aligned}$$

where $H(X)$ (resp. $H(Y)$) is the Shannon entropy of X (resp. Y) and P_X (resp. P_Y) denotes its probability mass function.

2.5 Mode

Let $k_X \geq 0$ (resp. $k_Y \geq 0$) denotes the mode of X (resp. Y). By using (1.1), we remark that :

- If $k_X = k_Y = 0$, then for all $k \in \mathbb{Z}$, we have

$$\begin{aligned}
\mathbb{P}(Z = k) &= p\mathbb{P}(X = k) + (1-p)\mathbb{P}(Y = -k) \\
&\leq p\mathbb{P}(X = 0) + (1-p)\mathbb{P}(Y = 0) = \mathbb{P}(Z = 0).
\end{aligned}$$

- Let k denotes any positive integer ($\neq 0$). We distinguish two cases :

– If $k_X > 0$, then we have

$$\mathbb{P}(Z = k) = p\mathbb{P}(X = k) \leq p\mathbb{P}(X = k_X) = \mathbb{P}(Z = k_X).$$

– If $k_X = 0$, then we have

$$\begin{aligned}
\mathbb{P}(Z = k) &= p\mathbb{P}(X = k) \\
&\leq p\mathbb{P}(X = 0) = \mathbb{P}(Z = 0) - (1-p)\mathbb{P}(Y = 0) \\
&\leq \mathbb{P}(Z = 0).
\end{aligned}$$

- Let k denotes any negative integer ($\neq 0$). We distinguish two cases :

– If $k_Y > 0$, then we have

$$\mathbb{P}(Z = k) = (1-p)\mathbb{P}(Y = -k) \leq (1-p)\mathbb{P}(Y = k_Y) = \mathbb{P}(Z = -k_Y).$$

– If $k_Y = 0$, then we have

$$\begin{aligned}\mathbb{P}(Z = k) &= (1 - p)\mathbb{P}(Y = -k) \\ &\leq (1 - p)\mathbb{P}(Y = 0) = \mathbb{P}(Z = 0) - p\mathbb{P}(X = 0) \\ &\leq \mathbb{P}(Z = 0).\end{aligned}$$

One can deduce that the mode(s) is a subset of $\{-k_Y, 0, k_X\}$. For example, if $\mathbb{P}(Z = 0) \geq \max\{\mathbb{P}(Z = k_X), \mathbb{P}(Z = -k_Y)\}$, then 0 is the unique mode of Z . On the other hand, if $\mathbb{P}(Z = k_X) = \mathbb{P}(Z = -k_Y) \geq \mathbb{P}(Z = 0)$, the Z has a bimodal distribution, with k_X and $-k_Y$ as modes. Obviously, if $\mathbb{P}(Z = k_X) = \mathbb{P}(Z = -k_Y) = \mathbb{P}(Z = 0)$, then the distribution of Z is trimodal.

2.6 The log-likelihood function

Let $A = \{a_1, \dots, a_v\}$ and $\sharp(A) = v$ denotes its cardinal. Moreover, we denote $|A| = \{|a_1|, \dots, |a_v|\}$, where $|a_i|$ equals to the absolute value of a_i , for all $1 \leq i \leq v$. The log-likelihood function associated to Z is defined by

$$\begin{aligned}l_Z(k_1, k_2, \dots, k_n) &= \log(L_Z(k_1, k_2, \dots, k_n)) = \log\left(\prod_{k=k_1}^{k_n} \mathbb{P}(Z = k)\right) \\ &= \sum_{k=k_1}^{k_n} \log(\mathbb{P}(Z = k)).\end{aligned}$$

Let the following sets

$$\begin{aligned}I_0 &= \{k_i, 1 \leq i \leq n, k_i = 0\}, \\ I_+ &= \{k_i, 1 \leq i \leq n, k_i > 0\},\end{aligned}$$

and

$$I_- = \{k_i, 1 \leq i \leq n, k_i < 0\}.$$

Note that, $0 \leq \sharp(I_0) \leq n$, $0 \leq \sharp(I_+) \leq n$, $0 \leq \sharp(I_-) \leq n$, and

$$\sharp(I_0) + \sharp(I_+) + \sharp(I_-) = n.$$

Thus, based on (1.1), one can deduce

$$\begin{aligned}l_Z(k_1, k_2, \dots, k_n) &= \sum_{k \in I_0} \log(\mathbb{P}(Z = 0)) + \sum_{k \in I_+} \log(p\mathbb{P}(X = k)) + \sum_{k \in I_-} \log((1 - p)\mathbb{P}(Y = -k)) \\ &= \sharp(I_0) \log(\mathbb{P}(Z = 0)) + \log(p) \sum_{k \in I_+} \log(\mathbb{P}(X = k)) \\ &\quad + \log(1 - p) \sum_{k \in I_-} \log(\mathbb{P}(Y = |k|)) \\ &= \sharp(I_0) \log(\mathbb{P}(Z = 0)) + \log(p) l_X(I_+) + \log(1 - p) l_Y(|I_-|),\end{aligned}\tag{2.3}$$

where l_X (resp. l_Y) is the log-likelihood function associated to X (resp. Y).

3 Special case

In this section, we consider that X is a Poisson random variable with $\lambda > 0$ as parameter and Y is a Geometric random variable with $0 < a < 1$ as parameter. Let now Z be a random variable which belongs to the Parametric Mixture- \mathbb{Z} class, with X and Y as considered below, and $p \in (0, 1)$. In other words, based on (1.1), the pmf of Z is defined by

$$\mathbb{P}(Z = k) = \begin{cases} pe^{-\lambda} \frac{\lambda^k}{k!}, & \text{if } k > 0, \\ pe^{-\lambda} + (1-p)a, & \text{if } k = 0, \\ (1-p)a(1-a)^{-k}, & \text{if } k < 0. \end{cases} \quad (3.1)$$

Next, we denote the above distribution PG-M (λ, a, p) .

Remark 4 In Figure 1, we have plotted the distribution for Z , defined by (3.1), for a combination of parameters. It is interesting the wide range of shapes of the distributions derived.

Therefore, using some properties introduced in Section 2, one can see that

- $\mathbb{E}(Z) = p\lambda - (1-p)\frac{1-a}{a}$,
- $\mathbb{V}(Z) = \lambda^2 p(1-p) + \lambda p \left(1 + \frac{2(1-p)(1-a)}{a} \right) + \frac{(1-p)(1-a)(2-a - (1-p)(1-a))}{a^2}$,
- $\varphi_Z(t) = pe^{-\lambda(1-e^{it})} + (1-p)\frac{a}{1 - (1-a)e^{-it}}$,
- Mode(s) is a subset of $\{0, \lfloor \lambda \rfloor\}$, where $\lfloor \cdot \rfloor$ is the floor function.
- The log-likelihood function is given by

$$\begin{aligned} l_Z(k_1, k_2, \dots, k_n) &= \#(I_0) \log(pe^{-\lambda} + (1-p)a) \\ &+ \log(p) \left(-\lambda \#(I_+) + \left(\sum_{I_+} k_i \right) \log(\lambda) - \log \left(\prod_{I_+} k_i \right) \right) \\ &+ \log(1-p) \left(\#(I_-) \log(a) + \left(\sum_{I_-} |k_i| \right) \log(1-a) \right). \end{aligned} \quad (3.2)$$

Remark 5 Let $K = (K_1, \dots, K_n)$ be a n -sample of Z and $\theta = (\lambda, a, p)$ be the vector of parameters associated to the distribution (3.1). We suppose that $\theta \in \Theta$, where Θ is a compact subset of $\mathbb{R} \times [0, 1] \times [0, 1]$. Therefore, the Maximum Likelihood Estimator of θ , denoted by $\hat{\theta}_n = (\hat{\lambda}_n, \hat{a}_n, \hat{p}_n)$, is defined by

$$\hat{\theta}_n = \operatorname{argmax}_{\theta \in \Theta} l_Z(K_1, \dots, K_n),$$

where l_Z is the log-likelihood function defined by (3.2). Here, we omit the standard theoretical details concerning asymptotic properties of $\hat{\theta}_n$. For example, standard errors can be derived from the Hessian as usual (see, e.g. Lehmann and Casella (1998)). Next, in Section 3.1, standard numerical methods will be used to derive the estimates and to illustrate these asymptotic properties. We have used simple code in R which worked without problems.

3.1 Numerical study

In this section, we introduce a simulation study to present the assessment of the performance of the estimators for this special case. Hence, for this study, we make the following steps :

1. we generate ten thousand samples of size n from the distribution of Z , defined by (3.1),
2. we compute, based on (3.2), the Maximum Likelihood estimators for the ten thousand samples, say $(\hat{\lambda}_{n,i}, \hat{a}_{n,i}, \hat{p}_{n,i})$ for $i = 1, \dots, 10000$,
3. we compute the biases and mean squared errors given by :

$$\text{Bias}_1(n) = \frac{\sum_{i=1}^{10000} (\hat{\lambda}_{n,i} - \lambda_0)}{10000},$$

$$\text{Bias}_2(n) = \frac{\sum_{i=1}^{10000} (\hat{a}_{n,i} - a_0)}{10000},$$

$$\text{Bias}_3(n) = \frac{\sum_{i=1}^{10000} (\hat{p}_{n,i} - p_0)}{10000},$$

$$\text{MSE}_1(n) = \frac{\sum_{i=1}^{10000} (\hat{\lambda}_{n,i} - \lambda_0)^2}{10000},$$

$$\text{MSE}_2(n) = \frac{\sum_{i=1}^{10000} (\hat{a}_{n,i} - a_0)^2}{10000},$$

$$\text{MSE}_3(n) = \frac{\sum_{i=1}^{10000} (\hat{p}_{n,i} - p_0)^2}{10000},$$

4. we repeated these steps for $n = 50, 100, 200, 500, 1000, 2000, 10000$ with $\lambda_0 = 1$, $a_0 = 0.3$ and $p_0 = 0.4$.

Figure 2 shows how biases and mean squared errors vary with respect to n (broken line in figures corresponds to a zero horizontal line) Thus, one can remark that the biases and the mean squared errors for each estimator decrease to zero as $n \rightarrow \infty$. On the other hand, fitting to a Gaussian distribution is illustrated in Figure 3, for the ML estimators.

4 Real data application

The candidates who have passed the entrance test to IDRAC Business School (Lyon, France) are required to answer a questionnaire concerning their expectations in several areas, in particular, the academic one. Thus, the participants note (based on a seven-point Likert-scale) how the IDRAC teaching reputation influenced their choice of the school (note that, for scale used, 1 represents a very low influence and 7 is considered as very high influence). The results of this questionnaire were collected during July 2015 (240 of new students participated to the survey). At the end of the first year in the school, the same students are subjected to another questionnaire. This time, they have to indicate the level to which this expectation where met after attending the School. Once again a seven-point Likert-scale is used where 1 is equivalent to "expectation not met", and 7 is equivalent to "expectation fully met". The results of this second questionnaire were collected during June 2016. For our real data application, we will take the difference between post perception of the teaching reputation (measured one year after the entrance test) and the prior expectation (measured at the entrance test). Note that, by definition, values of this new variable vary from -6 to 6 . Indeed, a positive value can be interpreted as a increasing perception, a negative value indicate a degression, and a null value represents a stability. Data are depicted in Figure 4. The results of the runs test show that the data set is a random sample (p -value = 0.5534). Moreover, based on empirical ACF and PACF (see Figure 5), one can see that there is no cut-off at any lag. Therefore, observations can be considered as no correlated. Now, in order to fit the data, we propose PG-M (λ, a, p) distribution, defined by (3.1). First, we estimated the distribution parameters (MLE) by using (3.2). It follows that $\hat{\lambda}_n = 1.1232694$, $\hat{a}_n = 0.7192641$, and $\hat{p}_n = 0.6454687$. The Pearson Chi-square test is performed to test the fitting. The null hypothesis is that the sample comes from PG-M distribution and the alternative hypothesis is that sample does not come from this distribution. The results of Pearson Chi-square test are given in the following table.

Modalities	Observed	Expected	$(O - E)^2/E$
≤ -2	10	8.090517	0.4506663
-1	20	17.181147	0.4624796
0	115	111.580226	0.1048112
1	60	56.590125	0.2054642
2	24	30.782978	1.4946179
≥ 3	11	15.775006	1.4453676
Total	240	240	4.163407

$\sum \frac{(O - E)^2}{E} = 4.163407 < \chi_{2,0.95}^2 = 5.9915$ indicating that the PG-M distribution fits the data well. we seek to compare the quality adjustment for the observed data of Skellam (S) distribution Skellam (1946), Extended Poisson (E-Po) Bakouch et al. (2016), Extended Binomial (EB) distribution Alzaid and Omair (2012), discrete analogue of the Laplace (DAL) distribution Inusah and Kozubowski (2006), and our PG-M distribution. In fact, using the data, we estimate parameter values associated to each distribution, we simulated 1000 series of length 240 from each distribution and we kept the expected relative frequencies for each series. The reported frequencies are the mean over the 1000 series. Results are represented in the following table.

Relative frequency of students number and fitted distributions.

Modalities	Observed frequency	S frequency	E-Po frequency	EB frequency	DAL frequency	PG-M frequency
≤ -2	4.166667	2.916667	1.25	2.3333	1.6666	3.750
-1	8.333333	14.583333	12.08333	7.251	6.2554	10.91667
0	47.916667	45.000	45.416667	42.456	44.4522	48.33333
1	25.000	23.333333	26.250	30.1	31.133	20.750
2	10.000	12.083333	11.666	15.13	14.234	11.250
≥ 3	4.583333	2.083333	2.7297	16.19	2.2588	5.000
Total	100	100	100	100	100	100

One can see that the PG-M distribution is the more appropriate to fit the data, compared to the other distributions.

References

- Alzaid, A. A. and Omair, M. A. (2012). An Extended Binomial Distribution with Applications. *Communications in Statistics: Theory and Methods*, **41**, 3511-3527.
- Bakouch C., Kachour, M. and Nadarajah, S. (2014). An extended Poisson distribution, *Communications in Statistics - Theory and Methods*, **45**, 22, 6746-6764.
- Blischke, W. R. (1962). Moment estimators for the parameters of a mixture of two binomial distributions, *The Annals of Mathematical Statistics*, **33**, 2, 444-454.
- Chesneau C. and Kachour, M. (2012). A parametric study for the first-order signed integer-valued autoregressive process, *Journal of Statistical Theory and Practice*, **6**, 4, 760-782.
- Everitt, B. S. (1981). Mixture Distributions. John Wiley & Sons, Inc.
- Inusah, S. and Kozubowski, T. J. (2006). A discrete analogue of the Laplace distribution, *Journal of Statistical Planning and Inference*, **136**, 1090-1102.
- Harris, C. M. (1983). On finite mixtures of geometric and negative binomial distributions, *Communications in Statistics-Theory and Methods*, **12**, 9, 987-1007.
- Johnson, N. L., Kemp, A. W., and Kotz, S. (2005). Univariate discrete distributions (Vol. 444). John Wiley & Sons.
- Karlis, D. and Xekalaki, E. (2005). Mixed poisson distributions. *International Statistical Review*, 73(1), pp. 35-58.
- Lehmann, E.L. and Casella, G. (1998). *Theory of Point Estimation* (2nd ed.), Springer.
- McLachlan G.J. and Batsford K.E. (1988). Mixture Models: Inference and Applications to Clustering. Decker, New York.
- Pritchard J.K., Stephens M., and Donnelly P. (2000). Inference of population structure using multilocus genotype data, *Genetics*, 155: 945-959.
- Rider, P. R. (1961). Estimating the parameters of mixed Poisson, binomial, and Weibull distributions by the method of moments, *Bulletin de l'Institut International de Statistiques*, 38.
- Skellam, J. G. (1946). The frequency distribution of the difference between two Poisson variates belonging to different populations, *Journal of the Royal Statistical Society, A*, **109**, 296.
- Titterton, D. M. (2005). Statistical analysis of finite mixture distributions (*Doctoral dissertation, Institute of Philosophy*).

Figures

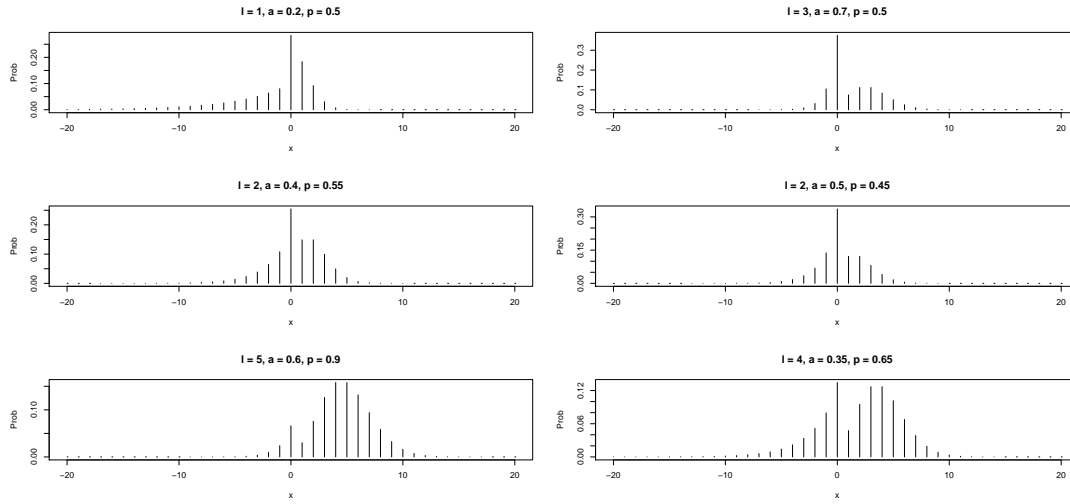


Figure 1: Plot of the marginal distribution, defined by (3.1), for several parameter combinations.

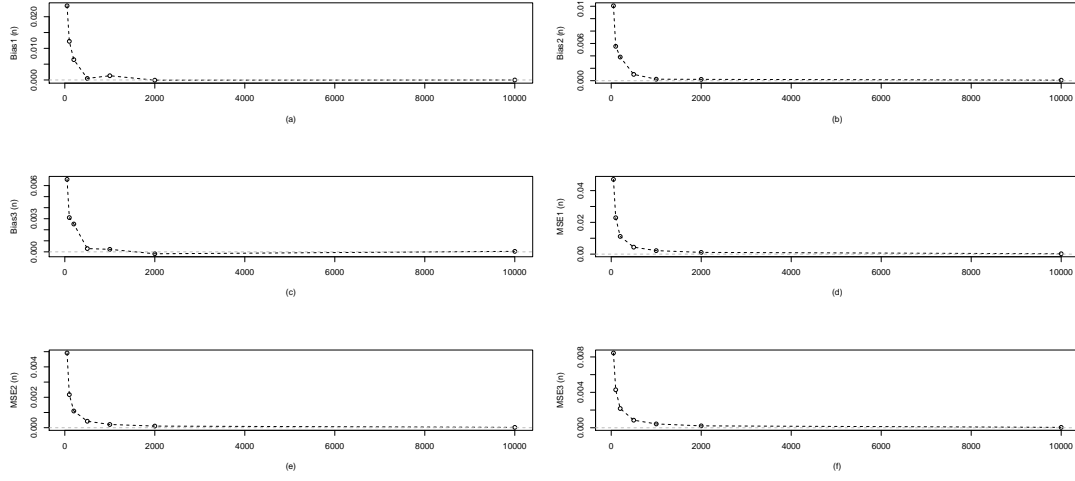


Figure 2: From top to bottom and from left to right : Bias and Mean squared errors of MLE, versus $n = 50, 100, 200, 500, 1000, 2000, 10000$. (a) Bias of $\hat{\lambda}_n$ (b) Bias of \hat{a}_n (c) Bias of \hat{p}_n (d) MSE of $\hat{\lambda}_n$ (e) MSE of \hat{a}_n (f) MSE of \hat{p}_n .

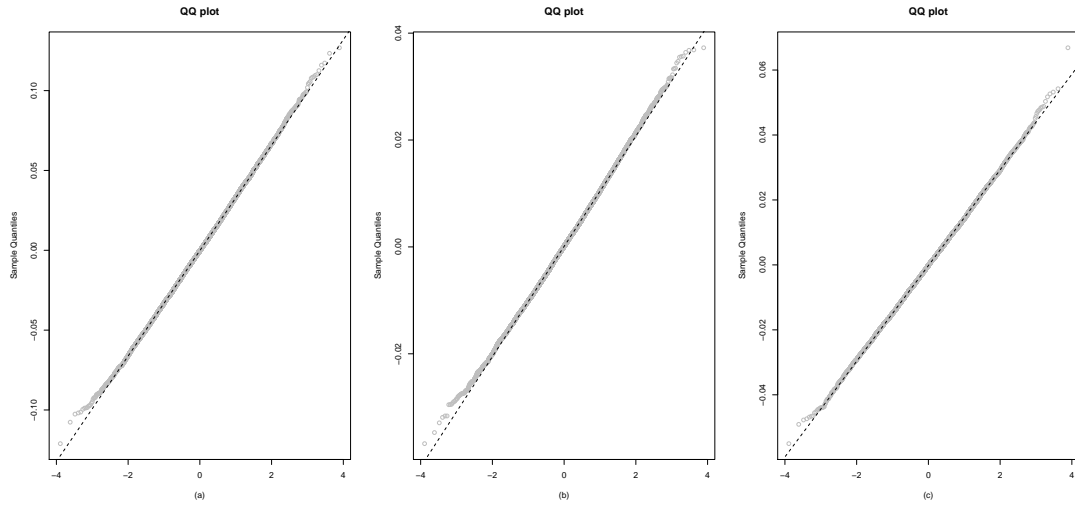


Figure 3: Normal Q-Q plots for the errors (ML estimators), when Z is a mixture of Poisson distribution and Geometric distribution and $n = 1000$. (a) Normal Q-Q plots of $(\hat{\lambda}_n - \lambda)$ (b) Normal Q-Q plots of $(\hat{a}_n - a)$ (c) Normal Q-Q plots of $(\hat{p}_n - p)$.

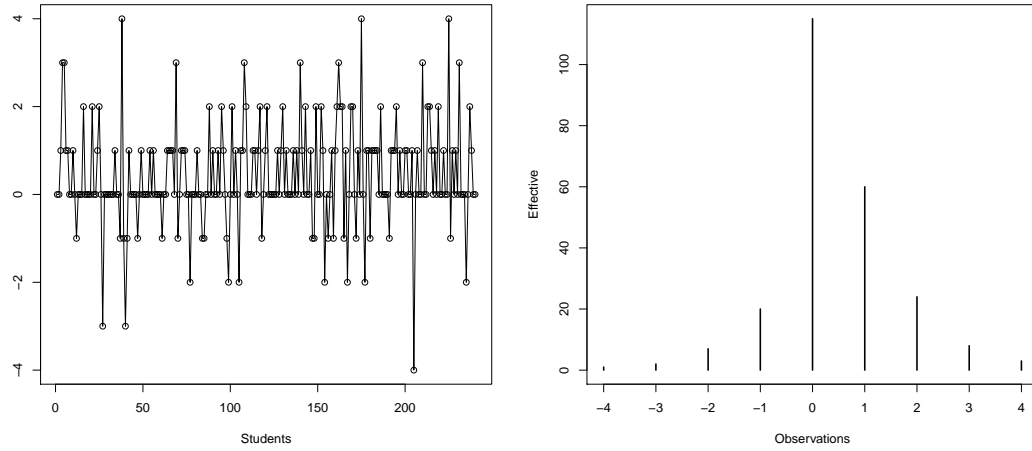


Figure 4: Plot of the real data : the difference between post perception of the teaching reputation (measured during July 2016) and the prior expectation (measured during June 2015).

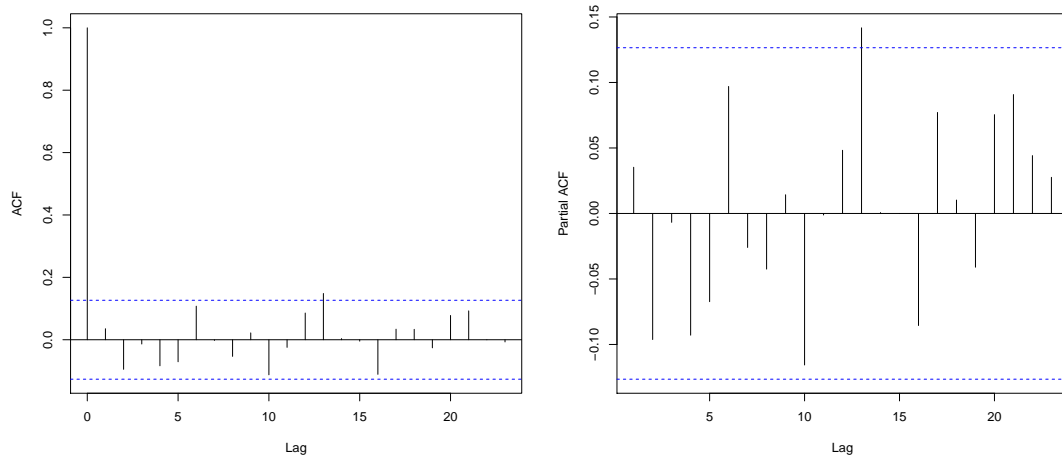


Figure 5: ACF and PACF of the data.