ProbStat Forum, Volume 10, January 2017, Pages 01–15
ISSN 0974-3235

ProbStat Forum is an e-journal.
For details please visit www.probstat.org.in

A new parametric mixture class of distributions on Z

Christophe Chesneau[a], Maher Kachour[b]

[a]*Laboratoire de Mathematiques Nicolas Oresme (LMNO)*
Universite de Caen Basse-Normandie
Campus II, Science 3, 14032 Caen, France.
[b]*Ecole superieure de commerce IDRAC*
47, rue Sergent Michel Berthet
69258 Lyon Cedex 09, France.

Abstract. In this paper we introduce a new class of discrete distributions with support belonging to $\mathbb{Z}$. It is a special case of well known mixtures. 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.

1. Introduction

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

$$p(x) = \sum_{i=1}^{k} \alpha_i p_i(x),$$

where p_i is a probability distribution, $\alpha_i \geq 0$ and $\sum_{i=1}^{k} \alpha_i = 1$. 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 well known finite mixture distribution approach, a new class of discrete distributions with support belonging to $\mathbb{Z}$, denoted by Parametric Mixture(p)$-\mathbb{Z}$ class.

Definition 1.1. *Let $p \in (0,1)$. We say that a random variable Z has a distribution belonging to the Parametric Mixture(p)$-$ class if and only if there exist two discrete non-negative random variables X and Y and an event A independent of X and Y with $\mathbb{P}(A) = p$ such that the following equality in distribution is satisfied :*

$$Z = I_A X - I_{\overline{A}} Y,$$

2010 *Mathematics Subject Classification.* 62M10, 62M20.
Keywords. Mixtures, Discrete distributions, Mixing distribution, Rademacher distribution.
Received: 14 October 2016; Revised: 07 December 2016, Accepted: 08 January 2017

Email addresses: `christophe.chesneau@unicaen.fr` (Christophe Chesneau), `maher.kachour@idraclyon.com` (Maher Kachour)

where I_A denotes the indicator random variable and $\overline{A}$ the contrary event of A.

Such distribution is 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).

The contents of this paper are organized as follows. Various mathematical properties of random variable belongs to the Parametric Mixture(p)$-$ are derived in Section 2. These include, the probability mass function (pmf), the cumulative distribution function, the failure functions, the probability generating function and moments, the mean absolute deviation, the Shannon entropy, the mode(s), and the log-likelihood function. Section 3 is devoted to the proof of our main result. In section 4, we investigate numerically the parameters estimation in a special case where the distribution of X and that of Y are fixed. The practical usefulness of our class is illustrated in section 5 via an application on real data.

2. Mathematical properties

Theorem 2.1. *Let $p \in (0,1)$ and Z be a random variable belonging to the Parametric Mixture(p)$-$ class. Then the following results hold :*

1) *The pmf of Z is given by*

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

2) *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 independent of T.

3) *The random variable $-Z$ belongs to the Parametric Mixture(q)$-$ class, where $q = 1-p$.*

4) *The pmf of $|Z|$ has the following representation :*

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

5) *Let $F_U(x) = \mathbb{P}(U \le x)$ denote the cumulative distribution function of a random variable U and $\lfloor\cdot\rfloor$ denote the floor function. Then we have*

$$F_Z(x) = \begin{cases} (1-p) + pF_X(\lfloor x \rfloor) & \text{if } x \ge 0, \\ \\ (1-p)(1-F_Y(-\lfloor x \rfloor - 1)) & \text{if } x < 0. \end{cases}$$

6) *Let $R_U(x) = \mathbb{P}(U > x)$ denote the survival function of a random variable U. Then we have*

$$R_Z(x) = \begin{cases} pR_X(\lfloor x \rfloor) & \text{if } x \ge 0, \\ \\ 1-(1-p)R_Y(-\lfloor x \rfloor - 1) & \text{if } x < 0. \end{cases}$$

7) *Let* $k_U(x) = \mathbb{P}(U = x \mid U \geq x)$ *denote the failure rate function of a random variable* U*. Then we have*

$$k_Z(x) = \begin{cases} \dfrac{\mathbb{P}\left(X = x\right)}{R_X(\lfloor x \rfloor - 1)} & \text{if } x \in \{1, 2, \ldots\}, \\[2ex] \dfrac{p\mathbb{P}\left(X = 0\right) + (1-p)\mathbb{P}\left(Y = 0\right)}{p + (1-p)\mathbb{P}(Y = 0)} & \text{if } x = 0, \\[2ex] \dfrac{(1-p)\mathbb{P}\left(Y = -x\right)}{1 - (1-p)R_Y(-\lfloor x \rfloor)} & \text{if } x \in \{\ldots, -2, -1\}. \end{cases}$$

8) *Let* $\tau_U(x) = \mathbb{P}(U = x \mid U \leq x)$ *denote the reverse failure rate function of a random variable* U*. Then we have*

$$\tau_Z(x) = \begin{cases} \dfrac{p\mathbb{P}\left(X = x\right)}{(1-p) + pF_X(\lfloor x \rfloor)} & \text{if } x \in \{1, 2, \ldots\}, \\[2ex] \dfrac{p\mathbb{P}\left(X = 0\right) + (1-p)\mathbb{P}\left(Y = 0\right)}{(1-p) + p\mathbb{P}(X = 0)} & \text{if } x = 0, \\[2ex] \dfrac{(1-p)\mathbb{P}\left(Y = -x\right)}{(1-p)(1 - F_Y(-\lfloor x \rfloor - 1))} & \text{if } x \in \{\ldots, -2, -1\}. \end{cases}$$

9) *Let* $G_U(s) = \mathbb{E}(s^U)$ *denote the probability generating function of a random variable* U*. Then the pmf of* Z *has the following representation :*

$$\mathbb{P}(Z = k) = \begin{cases} \dfrac{G^{(k)}_{ZI_{\{Z \geq 0\}}}}{k!}(0) & \text{if } k > 0, \\[2ex] \dfrac{\left(G_{ZI_{\{Z \leq 0\}}}\left(\dfrac{1}{s}\right)\right)^{(-k)}}{(-k)!} \Big|_{s \to 0} & \text{if } k < 0, \\[2ex] G_{ZI_{\{Z \geq 0\}}}(0) & \text{if } k = 0. \end{cases}$$

10) *Let* $M_U(t) = \mathbb{E}(e^{tU})$ *denote the moment-generating function of* U*. Then we have*

$$M_Z(t) = pM_X(t) + (1-p)M_Y(-t), \qquad t \in \mathbb{R}.$$

11) *For any non-negative integer* n*, we have*

$$\mathbb{E}\left(Z^n\right) = p\mathbb{E}\left(X^n\right) + (-1)^n(1-p)\mathbb{E}\left(Y^n\right).$$

12) *The variance of* Z *is given by*

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

13) *Let $CV(Z) = \sigma_Z/\mathbb{E}(Z)$, with $\sigma_Z = \sqrt{\mathbb{V}(Z)}$, denote the coefficient of variation of Z. Then we have*

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

14) *Let $Skewness(Z) = \mathbb{E}(((Z-\mathbb{E}(Z))/\sigma_Z)^3)$ denote the skewness of Z. Then we have*

$$Skewness(Z) = \frac{\mathbb{E}\left(Z^3\right) - 3\mathbb{E}\left(Z^2\right)\mathbb{E}\left(Z\right) + 2\left(\mathbb{E}\left(Z\right)\right)^3}{\sigma_Z^3}.$$

15) *Let $Kurtosis(Z) = \mathbb{E}(((Z-\mathbb{E}(Z))/\sigma_Z)^4)$ denote the kurtosis of Z. Then we have*

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

16) *Let $ID\left(Z\right) = \mathbb{V}(Z)/\mathbb{E}(Z)$ denote the index of dispersion of Z. Then we have*

$$ID\left(Z\right) = \frac{p\mathbb{E}(X^2) + (1-p)\mathbb{E}(Y^2) - p^2\left(\mathbb{E}(X)\right)^2 - (1-p)^2\left(\mathbb{E}(Y)\right)^2 + 2p(1-p)\mathbb{E}(X)\mathbb{E}(Y)}{p\mathbb{E}(X) - (1-p)\mathbb{E}(Y)}.$$

17) *Let $MAD(Z) = \mathbb{E}\left(|Z-m|\right)$ denote the mean absolute deviation of Z. Then we have*

$$MAD(Z) = 2mF_Z(\lfloor m \rfloor) - m - \mathbb{E}\left(Z\right) + 2\sum_{k=\lfloor m \rfloor+1}^{\infty} k\mathbb{P}\left(Z=k\right),$$

with

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

18) *Let P_U denote the probability mass function of a random variable U and $H(U)$ denote the Shannon entropy of U. Then we have*

$$\begin{aligned} H(Z) &= (1-p)P_Y(0)\left(-\log\left(P_Z(0)\right) + \log\left(1-p\right) + \log\left(P_Y(0)\right)\right) \\ &+ pP_X(0)\left(-\log\left(P_Z(0)\right) + \log\left(p\right) + \log\left(P_X(0)\right)\right) \\ &+ pH(X) + (1-p)H(Y) - p\log\left(p\right) - (1-p)\log\left(1-p\right). \end{aligned}$$

19) *Set $\mathcal{E}_X = \{mode(s) \text{ of } X\}$ and $\mathcal{F}_Y = \{-k_Y;\ k_Y \text{ is a mode of } Y\}$. Then the mode(s) of Z is a subset of $\mathcal{E}_X \cup \{0\} \cup \mathcal{F}_Y$.*

20) *Let* $l_U(k_1, k_2, \ldots, k_n) = \log\left(\prod_{k\in\{k_1,k_2,\ldots,k_n\}} \mathbb{P}(U = k)\right)$ *denote the log-likelihood function associated to a random variable* U*. For any* $B = \{b_1, \ldots, b_v\}$*,* $\sharp(B) = v$ *denotes its cardinal and* $|B| = \{|b_1|, \ldots, |b_v|\}$*, where* $|b_i|$ *equals to the absolute value of* b_i*. We set* $I_0 = \{k_i;\ 1 \le i \le n, k_i = 0\}$*,* $I_+ = \{k_i;\ 1 \le i \le n, k_i > 0\}$ *and* $I_- = \{k_i;\ 1 \le i \le n, k_i < 0\}$*. Then, for all* $(k_1, k_2, \ldots, k_n) \in \mathbb{Z}^n$*, we have*

$$\begin{aligned} & l_Z(k_1, k_2, \ldots, k_n) \\ &= \sharp(I_0)\log(\mathbb{P}(Z = 0)) + \log(p)\, l_X(I_+) + \log(1-p)\, l_Y(|I_-|). \end{aligned}$$

Remark 2.2. *Owing to* **2)***, note that the Parametric Mixture(p)− class can be considered as a generalization of the Rademacher(p)− class originally introduced by Chesneau and Kachour (2012). 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 2.3. *It follows from* **4)** *that* $|Z|$ *has a* $2-$*mixture discrete distribution.*

Remark 2.4. *Using* **16)***, 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* $\mathbb{E}(X) > (1-p)/p\mathbb{E}(Y)$ *(resp.* $\mathbb{E}(X) < (1-p)/p\mathbb{E}(Y)$*).*

Remark 2.5. *Let* $\varphi_U(t) = \mathbb{E}(e^{itU})$ *denote the characteristic function of a random variable* U*. Proceeding as in* **10)***, we show that* $\varphi_Z(t) = \mathbb{E}\left(e^{itZ}\right) = p\varphi_X(t) + (1-p)\varphi_Y(-t)$*,* $t \in \mathbb{R}$*.*

3. Proofs of Theorem 2.1

1) For any $k \in \mathbb{Z}$, since A is independent of X and Y, we have

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

2) If X and Y have the same distribution than a random variable T, then we have the equality in distribution : $Z = I_A X - I_{\overline{A}} Y = RT$, where $R = I_A - I_{\overline{A}} \in \{-1, 1\}$ is a Rademacher random variable with parameter p independent of T.

3) For any $k \in \mathbb{Z}$, we have

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

Hence $-Z$ belongs to the Parametric Mixture(q)− class.

4) For any $k \in \mathbb{N}$, we have

$$\begin{aligned} \mathbb{P}(|Z| = k) &= \mathbb{P}(\{|Z| = k\} \cap A) + \mathbb{P}(\{|Z| = k\} \cap \overline{A}) \\ &= \mathbb{P}(\{X = k\} \cap A) + \mathbb{P}(\{Y = k\} \cap \overline{A}) \\ &= p\mathbb{P}(X = k) + (1-p)\mathbb{P}(Y = k). \end{aligned}$$

5) For $x < 0$, we have

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

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

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

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

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

Hence

$$F_Z(x) = \begin{cases} (1-p) + pF_X(\lfloor x \rfloor) & \text{if } x \geq 0, \\ \\ (1-p)(1 - F_Y(-\lfloor x \rfloor - 1)) & \text{if } x < 0. \end{cases}$$

6) 7) 8) These points are immediate consequences of the equalities :

$$R_Z(x) = \mathbb{P}(Z > x) = 1 - F_Z(x), \quad k_Z(x) = \mathbb{P}(Z = x \mid Z \geq x) = \frac{\mathbb{P}(Z = x)}{R_Z(x-1)}$$

and

$$\tau_Z(x) = \mathbb{P}(Z = x \mid Z \leq x) = \frac{\mathbb{P}(Z = x)}{F_Z(x)}$$

and the points **1)** and **5)**.

9) We have

$$\begin{aligned} G_{ZI_{\{Z \geq 0\}}}(s) &= \mathbb{E}(s^{ZI_{\{Z \geq 0\}}}) = \mathbb{E}(s^{ZI_{\{Z \geq 0\}}} I_A) + \mathbb{E}(s^{ZI_{\{Z \geq 0\}}} I_{\overline{A}}) \\ &= \mathbb{E}(s^{XI_{\{X \geq 0\}}} I_A) + \mathbb{E}(s^{-YI_{\{-Y \geq 0\}}} I_{\overline{A}}) \\ &= pG_X(s) + (1-p)\mathbb{E}(s^{-YI_{\{Y = 0\}}}) = pG_X(s) + (1-p)\mathbb{P}(Y = 0) \end{aligned}$$

and

$$
\begin{aligned}
& G_{ZI_{\{Z\leq 0\}}}\left(\frac{1}{s}\right) = \mathbb{E}\left(\left(\frac{1}{s}\right)^{ZI_{\{Z\leq 0\}}}\right) \\
&= \mathbb{E}\left(\left(\frac{1}{s}\right)^{ZI_{\{Z\leq 0\}}} I_A\right) + \mathbb{E}\left(\left(\frac{1}{s}\right)^{ZI_{\{Z\leq 0\}}} I_{\overline{A}}\right) \\
&= \mathbb{E}\left(\left(\frac{1}{s}\right)^{XI_{\{X\leq 0\}}} I_A\right) + \mathbb{E}\left(\left(\frac{1}{s}\right)^{-YI_{\{-Y\leq 0\}}} I_{\overline{A}}\right) \\
&= p\mathbb{E}\left(\left(\frac{1}{s}\right)^{XI_{\{X=0\}}}\right) + (1-p)G_Y(s) = p\mathbb{P}(X=0) + (1-p)G_Y(s).
\end{aligned}
$$

Therefore

$$
\begin{aligned}
\mathbb{P}(Z=k) &= \begin{cases} p\mathbb{P}(X=k) & \text{if } k>0, \\ (1-p)\mathbb{P}(Y=-k) & \text{if } k<0, \\ \mathbb{P}(Z=0) & \text{if } k=0, \end{cases} = \begin{cases} p\dfrac{G_X^{(k)}}{k!}(0) & \text{if } k>0, \\ (1-p)\dfrac{G_Y^{(-k)}}{(-k)!}(0) & \text{if } k<0, \\ \mathbb{P}(Z=0) & \text{if } k=0, \end{cases} \\
&= \begin{cases} \dfrac{G_{ZI_{\{Z\geq 0\}}}^{(k)}}{k!}(0) & \text{if } k>0, \\ \dfrac{\left(G_{ZI_{\{Z\leq 0\}}}\left(\frac{1}{s}\right)\right)^{(-k)}}{(-k)!}\Big|_{s\to 0} & \text{if } k<0, \\ G_{ZI_{\{Z\geq 0\}}}(0) & \text{if } k=0. \end{cases}
\end{aligned}
$$

10) For any $t \in \mathbb{R}$, we have

$$
\begin{aligned}
M_Z(t) &= \mathbb{E}(e^{tZ}) = \mathbb{E}(e^{tZ}I_A) + \mathbb{E}(e^{tZ}I_{\overline{A}}) = \mathbb{E}(e^{tX}I_A) + \mathbb{E}(e^{-tY}I_{\overline{A}}) \\
&= pM_X(t) + (1-p)M_Y(-t).
\end{aligned}
$$

11) It follows from **10)** that

$$M_Z^{(n)}(t) = pM_X^{(n)}(t) + (-1)^n(1-p)M_Y^{(n)}(-t).$$

So

$$\mathbb{E}\left(Z^n\right) = M_Z^{(n)}(0) = p\mathbb{E}\left(X^n\right) + (-1)^n(1-p)\mathbb{E}\left(Y^n\right).$$

12) Owing to **11)**, we have $\mathbb{E}(Z) = p\mathbb{E}(X) - (1-p)\mathbb{E}(Y)$ and

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

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

13) 14) 15) 16) These points are immediate consequences of the definitions and the point **12)**.

17) We have

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

with, by **1)**,

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

18) Using **1)**, we have

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

19) Let $k_X \geq 0$ (resp. $k_Y \geq 0$) be a mode of X (resp. Y).

- If $k > 0$ and

– 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

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

- If $k < 0$ and

– 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

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

Therefore, if we set $\mathcal{E}_X = \{\text{modes of } X\}$ and $\mathcal{F}_Y = \{-k_Y \text{ such that } k_Y \text{ is a mode of } Y\}$, one can deduce that the mode(s) iof Z s a subset of $\mathcal{E}_X \cup \{0\} \cup \mathcal{F}_Y$.

20) Using **1)**, we obtain

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

4. Numerical study

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(p)− class, with X and Y as considered below, and $p \in (0, 1)$. In other words, based on (1), the pmf of Z is defined by

$$\mathbb{P}(Z = k) = \begin{cases} pe^{-\lambda}\dfrac{\lambda^k}{k!}, & \text{if k} > 0,\\ pe^{-\lambda} + (1-p)a, & \text{if k} = 0,\\ (1-p)a(1-a)^{-k}, & \text{if k} < 0.\end{cases} \quad (2)$$

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

Remark 4.1. *In Figure 1, we have plotted the distribution for Z, defined by (2), for a combination of parameters. It is interesting the wide range of shapes of the distributions derived.*

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

- $\mathbb{E}(Z) = p\lambda - (1-p)\dfrac{1-a}{a}$,
- $\mathbb{V}(Z) =$
  $\lambda^2 p(1-p) + \lambda p\left(1 + \dfrac{2(1-p)(1-a)}{a}\right) + \dfrac{(1-p)(1-a)(2-a-(1-p)(1-a))}{a^2}$,
- $M_Z(t) = pe^{-\lambda(1-e^t)} + (1-p)\dfrac{a}{1-(1-a)e^{-t}}$,
- The set of the mode(s) of Z is a subset of $\{0, \lfloor\lambda\rfloor\}$ if λ is not an integer, or a subset of $\{0, \lambda-1, \lambda\}$ if λ is an integer,
- The log-likelihood function is given by

$$
\begin{aligned}
l_Z(k_1, k_2, \ldots, k_n) &= \sharp(I_0)\log\left(pe^{-\lambda} + (1-p)a\right) \\
&+ \log(p)\left(-\lambda\sharp(I_+) + \left(\sum_{I_+} k_i\right)\log(\lambda) - \log\left(\prod_{I_+} k_i\right)\right) \\
&+ \log(1-p)\left(\sharp(I_-)\log(a) + \left(\sum_{I_-} |k_i|\right)\log(1-a)\right).
\end{aligned} \tag{3}
$$

Remark 4.2. *Let $K = (K_1, \ldots, K_n)$ be a n-sample of Z and $\theta = (\lambda, a, p)$ be the vector of parameters associated to the distribution (2). We suppose that $\theta \in \Theta$, where Θ is a compact subset of $(0,\infty) \times [0,1] \times [0,1]$. Therefore, the Maximum Likelihood Estimator of θ, denoted by $\hat{\theta}_n = \left(\hat{\lambda}_n, \hat{a}_n, \hat{p}_n\right)$, is defined by*

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

where l_Z is the log-likelihood function defined by (3). 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)).

We now present a simulation study to present the assessment of the performance of the estimators for this special case. We have used simple code in `R` which worked without problems. For this study, we make the following steps :

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

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

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

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

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

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

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

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 $n \to \infty$. On the other hand, fitting to a Gaussian distribution is illustrated in Figure 3, for the ML estimators.

5. 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 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 (2). First, we estimated the distribution parameters (MLE) by using (3). 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.

Pearson chi-square test for PG-M distribution.

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

$\sum(O-E)^2/E = 4.163407 < \chi^2_{2,0.95} = 5.9915$ 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.

Acknowledgement

The authors would like to thanks the referee's for carefully reading the paper and for helpful suggestions which greatly improved the paper.

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), 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.

Titterington, D. M. (2005). Statistical analysis of finite mixture distributions (*Doctoral dissertation, Institute of Philosophy*).

Figures

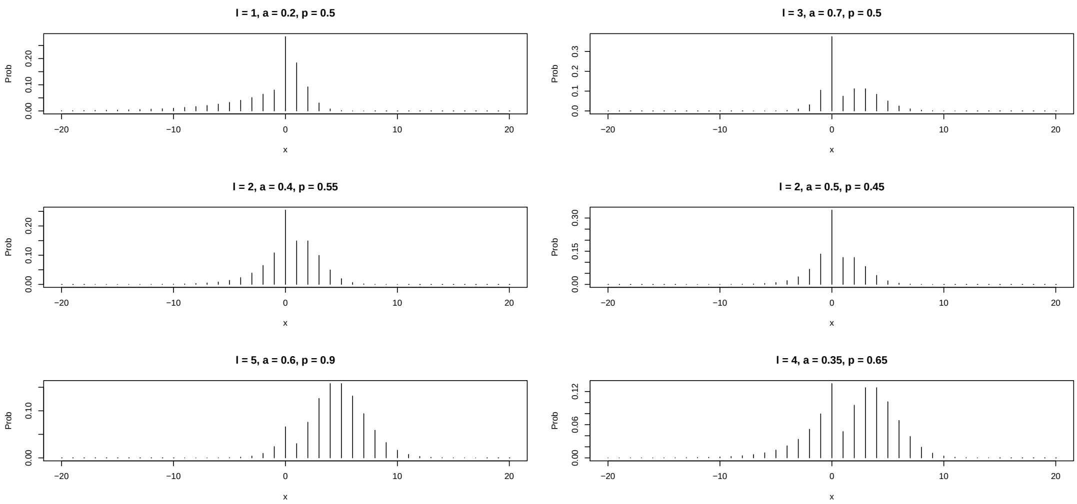

Figure 1: Plot of the marginal distribution, defined by (2), for several parameter combinations.

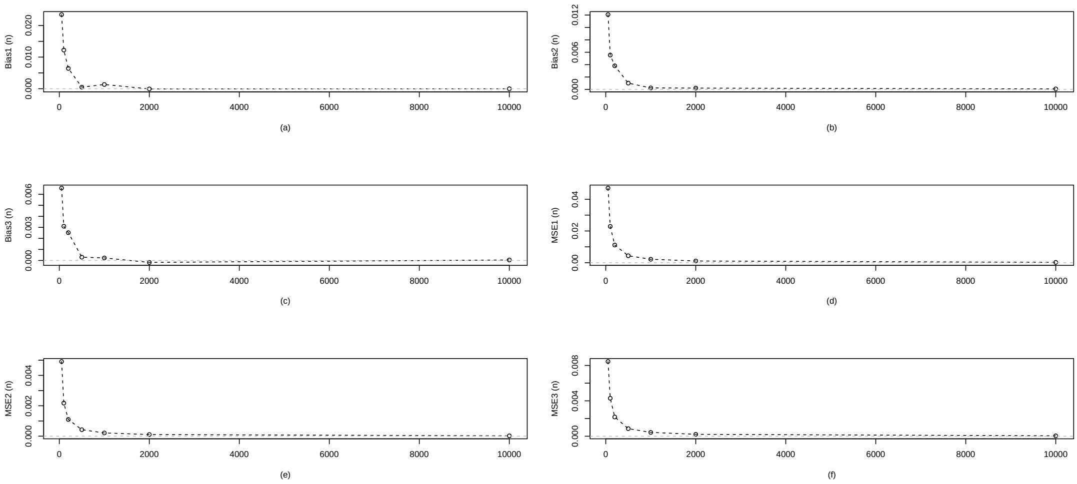

Figure 2: From top to bottom and from left to right : Bias and Mean squared errors of MLE, versus $n = 50, 100, 200, 500, 1000, 2000, 10000$. (a) Bias of $\hat{\lambda}_n$ (b) Bias of $\hat{a}_n$ (c) Bias of $\hat{p}_n$ (d) MSE of $\hat{\lambda}_n$ (e) MSE of $\hat{a}_n$ (f) MSE of $\hat{p}_n$.

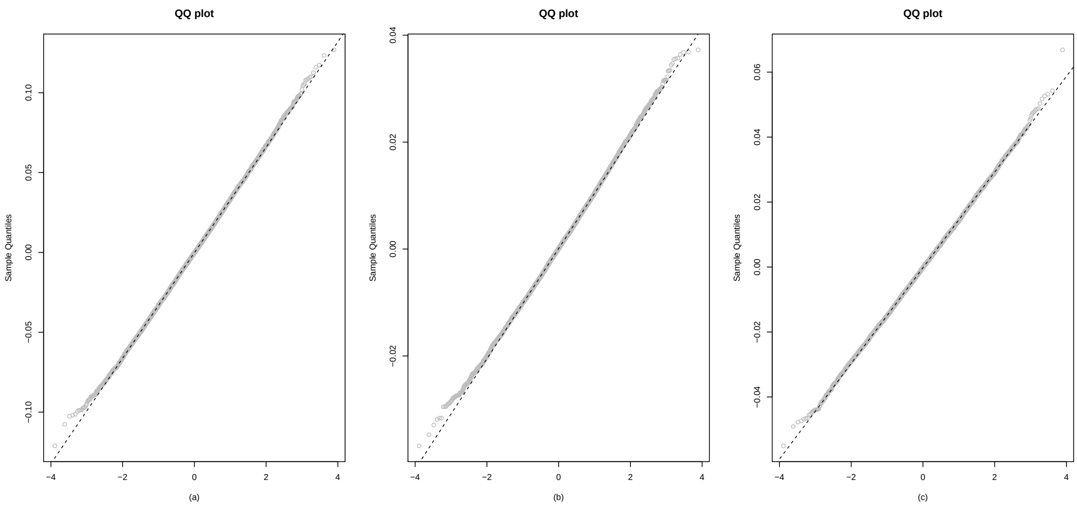

Figure 3: Normal Q-Q plots for the errors (ML estimators), when Z is a mixture of Poisson distribution and Geometric distribution and $n = 1000$. (a) Normal Q-Q plots of $(\hat{\lambda}_n - \lambda)$ (b) Normal Q-Q plots of $(\hat{a}_n - a)$ (c) Normal Q-Q plots of $(\hat{p}_n - p)$.

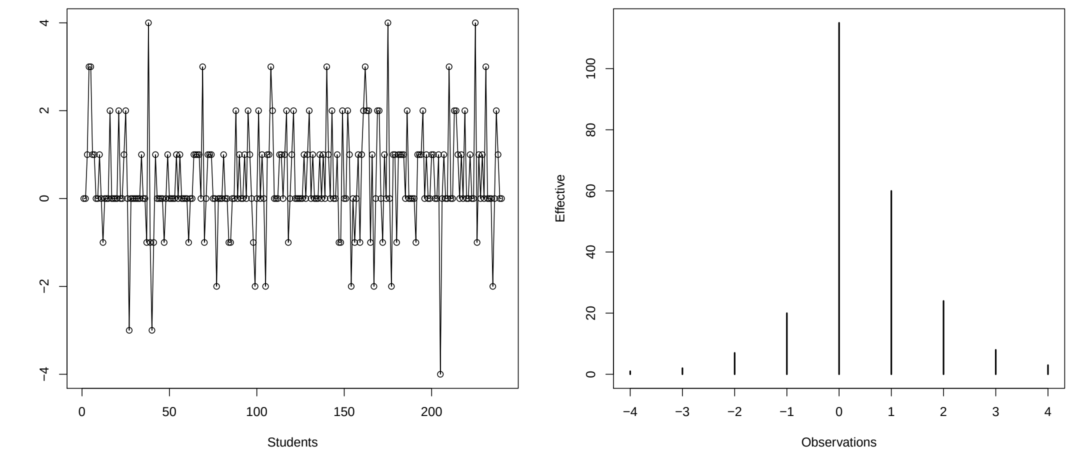

Figure 4: Plot of the real data : the difference between post perception of the teaching reputation (measured during July 2016) and the prior expectation (measured during June 2015).

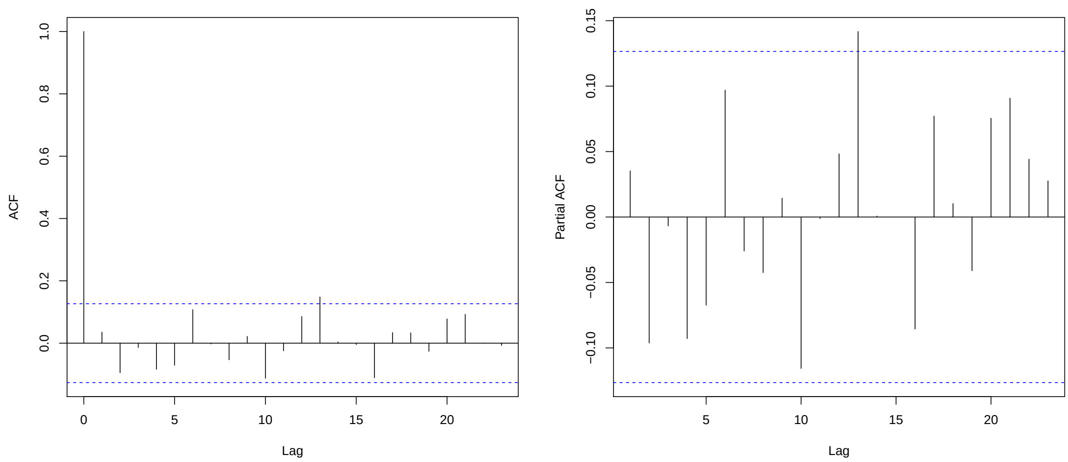

Figure 5: ACF and PACF of the data.