AISM (2006) 58: 555–571
DOI 10.1007/s10463-005-0029-1

Tomasz J. Kozubowski · Seidu Inusah

A skew Laplace distribution on integers

Received: 17 November 2004 / Revised: 24 February 2005 / Published online: 17 June 2006

Abstract We propose a discrete version of the skew Laplace distribution. In contrast with the discrete normal distribution, here closed form expressions are available for the probability density function, the distribution function, the characteristic function, the mean, and the variance. We show that this distribution on integers shares many properties of the skew Laplace distribution on the real line, including unimodality, infinite divisibility, closure properties with respect to geometric compounding, and a maximum entropy property. We also discuss statistical issues of estimation under this model.

Keywords Discrete Laplace distribution · Discrete normal distribution · Double exponential distribution · Exponential distribution · Geometric distribution · Geometric infinite divisibility · Infinite divisibility · Laplace distribution · Maximum entropy property · Maximum likelihood estimation

1 Introduction

Any continuous distribution on $\mathbb{R}$ with p.d.f. f admits a discrete counterpart supported on the set of integers $\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$. This discrete variable has p.m.f. of the form

$$\mathbb{P}(Y = k) = \frac{f(k)}{\sum_{j=-\infty}^{\infty} f(j)}, \quad k \in \mathbb{Z}. \tag{1}$$

T.J. Kozubowski (✉)
Department of Mathematics and Statistics, University of Nevada, Reno, NV 89557, USA
E-mail: tkozubow@unr.edu

S. Inusah
Department of Biostatistics, University of Alabama, Birmingham, AL 35294, USA

Discrete normal (DN) distribution was mentioned in Lisman and van Zuylen (1972) in connection with maximum entropy property, and studied by Kemp (1997) (see also Dasgupta 1993; Szablowski 2001). In this paper we study a distribution on $\mathbb{Z}$ defined via Eq. (1), where

$$f(x) = \frac{1}{\sigma}\frac{\kappa}{1+\kappa^2}\begin{cases} e^{-\frac{\kappa}{\sigma}x} & \text{if } x \geq 0 \\ e^{\frac{1}{\kappa\sigma}x} & \text{if } x < 0 \end{cases} \tag{2}$$

is the p.d.f. of the skew Laplace distribution with a scale parameter $\sigma > 0$ and the skewness parameter κ (see Kotz et al. 2001). In the symmetric case ($\kappa = 1$) this leads to a discrete analog of the classical Laplace distribution, studied in detail by Inusah and Kozubowski (2006). We call this distribution (skew) discrete Laplace (DL), as it shares many properties of the continuous Laplace law:

- A skew Laplace random variable has the same distribution as the difference of two exponential variables. A DL variable has the same representation with two geometric variables.
- Both distributions maximize the entropy within the respective classes of distributions with a given mean and the first absolute moment, which are continuous and supported on $\mathbb{R}$ or discrete and supported on $\mathbb{Z}$.
- Both distributions are unimodal, with explicit forms of the densities, distribution functions, characteristic functions, and moments.
- Both distributions are infinitely divisible, geometric infinitely divisible, and stable with respect to geometric compounding.

The simplicity of this model and its connections with the geometric distribution (which has many applications) and the skew Laplace distribution (which is becoming prominent in recent years, e.g., Kotz et al. 2001) lead to immediate applications of DL distributions. In particular, as discussed in Inusah and Kozubowski (2006), these distributions are of primary importance in analysis of uncertainty in hydroclimatic episodes such as droughts, floods, and El Niño (whose durations are often modeled by the geometric distribution). In this connection, the skew DL distribution is useful in answering the question whether the durations of positive and negative episodes have the same (geometric) distributions.

Our paper is organized as follows. After definitions and basic properties contained in Sect. 2, we present various representations of DL variables in Sect. 3. Then in Sect. 4 we present the main properties of DL laws, which include infinite divisibility, stability with respect to geometric convolutions, and their maximum entropy property. In Sect. 5 we consider statistical issues of parameter estimation. Proofs and technical results are collected in Sect. 6.

2 Definition and basic properties

When the Laplace density Eq. (2) is inserted into Eq. (1), the p.m.f. of the resulting discrete distribution takes on an explicit form in terms of the parameters $p = e^{-\kappa/\sigma}$ and $q = e^{-1/\kappa\sigma}$, leading to the following definition.

Definition 2.1 A random variable Y has the discrete Laplace distribution with parameters $p \in (0, 1)$ and $q \in (0, 1)$, denoted by $DL(p, q)$, if

$$f(k|p,q) = \mathbb{P}(Y = k) = \frac{(1-p)(1-q)}{1-pq} \begin{cases} p^k, & k = 0, 1, 2, 3, \ldots, \\ q^{|k|}, & k = 0, -1, -2, -3, \ldots. \end{cases} \quad (3)$$

The explicit expressions for the cumulative distribution function (c.d.f.) and the characteristic function (ch.f.) corresponding to the DL distribution follow easily from the geometric series formula (see Inusah, 2003 for details).

Proposition 2.1 *Let $Y \sim DL(p, q)$. Then the c.d.f. of Y is given by*

$$F(x|p,q) = \mathbb{P}(Y \le x) = \begin{cases} \frac{(1-p)q^{-[x]}}{1-pq} & \text{if } x < 0 \\ 1 - \frac{(1-q)p^{[x]+1}}{1-pq} & \text{if } x \ge 0, \end{cases} \quad (4)$$

where $[\cdot]$ is the greatest integer function, while the ch.f. of Y is

$$\varphi(t|p,q) = \mathbb{E}e^{itY} = \frac{(1-p)(1-q)}{(1-e^{it}p)(1-e^{-it}q)}, \quad t \in \mathbb{R}. \quad (5)$$

2.1 Special cases

When $p = q$, we obtain the symmetric DL distribution studied by Inusah and Kozubowski (2006). When either p or q converges to zero, we obtain the following two "one-sided" special cases: $Y \sim \mathrm{DL}(p, 0)$ with $p \in (0, 1)$ is a geometric distribution with the p.m.f.

$$f(k|p,0) = \mathbb{P}(Y = k) = (1-p)p^k, \quad k = 0, 1, 2, 3, \ldots, \quad (6)$$

while $Y \sim \mathrm{DL}(0, q)$ with $q \in (0, 1)$ is a geometric distribution on non-positive integers with the p.m.f.

$$f(k|0,q) = \mathbb{P}(Y = k) = (1-q)q^{-k}, \quad k = 0, -1, -2, -3, \ldots. \quad (7)$$

When p and q are both zero, the distribution is degenerate at zero. In contrast, we do not have limiting distributions when either $p \to 1^-$ or $q \to 1^-$.

2.2 Moments

The moments of $Y \sim \mathrm{DL}(p, q)$ can be easily obtained using Eq. (3) and the combinatorial identity

$$\sum_{k=1}^{\infty} k^n p^k = \sum_{k=1}^{n} S(n,k) \frac{k!p^k}{(1-p)^{k+1}}$$

(see, e.g., formula (7.46), p. 337, of Graham et al. 1989), where

$$S(n,k) = \frac{1}{k!} \sum_{i=0}^{k-1} (-1)^i \binom{k}{i} (k-i)^n \quad (8)$$

is the Stirling number of the second kind (the number of ways of partitioning a set of n elements into k nonempty subsets, see, e.g., Graham et al. 1989).

Proposition 2.2 *If $Y \sim DL(p,q)$, then for any integer $n \geq 1$ we have*

$$\mathbb{E}|Y|^n = \frac{1}{1-pq} \sum_{k=1}^{n} k!S(n,k) \left((1-q) \left(\frac{p}{1-p} \right)^k + (1-p) \left(\frac{q}{1-q} \right)^k \right), \quad (9)$$

$$\mathbb{E}Y^n = \frac{1}{1-pq} \sum_{k=1}^{n} k!S(n,k) \left((1-q) \left(\frac{p}{1-p} \right)^k + (-1)^n (1-p) \left(\frac{q}{1-q} \right)^k \right), \quad (10)$$

with $S(n,k)$ as above.

In particular, we have

$$\mathbb{E}Y = \frac{1}{1-p} - \frac{1}{1-q} = \frac{p}{1-p} - \frac{q}{1-q}, \quad \mathbb{E}|Y| = \frac{q(1-p)^2 + p(1-q)^2}{(1-qp)(1-q)(1-p)}, \quad (11)$$

and

$$\mathbb{V}\text{ar}Y = \frac{1}{(1-p)^2(1-q)^2} \left(\frac{q(1-p)^3(1+q) + p(1-q)^3(1+p)}{1-pq} - (p-q)^2 \right). \quad (12)$$

3 Representations

There are various representations of $Y \sim \text{DL}(p,q)$ r.v. related to the geometric distribution with parameter $u \in (0,1)$, whose p.m.f. is

$$g(k) = (1-u)^{k-1}u, \quad k = 1, 2, 3, \ldots. \quad (13)$$

As shown in Inusah and Kozubowski (2006), in the symmetric case ($p = q$) we have

$$Y \stackrel{d}{=} X_1 - X_2, \quad (14)$$

where the X_i's are i.i.d. geometric (Eq. 13) with $u = 1 - p$. Similar representation holds for the skew case as well.

Proposition 3.1 *If $Y \sim DL(p,q)$ then (Eq. 14) holds, where the X_i's are independent geometric variables (Eq. 13) with $u = 1 - p$ and $u = 1 - q$, respectively.*

Remark 3.1 The skew Laplace distribution Eq. (2) arises as the difference of two exponential random variables (see, e.g., Kotz et al. 2001). Since the geometric distribution is a discrete analog of the exponential distribution, it is natural to name the distribution of the difference of two geometric variables a "discrete Laplace".

Remark 3.2 A discrete normal random variable Y admits the representation (14) with i.i.d. Heine variables (see Kemp 1997). If the X_i's have Poisson distributions, then the distribution of their difference is rather complicated, see Irwin (1937).

Remark 3.3 Note that the representation (14) holds if the X_i's are replaced by the shifted variables $X_i + m$. In particular, if they have a shifted geometric distribution with p.m.f.

$$P(X = k) = (1 - u)^k u, \quad k = 0, 1, 2, \dots . \tag{15}$$

We now consider representations of the form

$$Y \stackrel{d}{=} IX, \tag{16}$$

where I takes on the values 0 and ± 1, X has a discrete distribution on non-negative integers, and the two variables are independent (if Y has the skew Laplace distribution (2) then a similar representation holds with an exponentially distributed X, see, e.g., Kotz et al. 2001). If $p = q$, then Eq. (16) holds if and only if $\alpha \geq 2p/(1 + p)$, where $0 < \alpha \leq 1$ is such that $\mathbb{P}(I = 0) = 1 - \alpha$ and $\mathbb{P}(I = 1) = \mathbb{P}(I = -1) = \alpha/2$, in which case

$$\mathbb{P}(X = 0) = 1 - \frac{2p}{\alpha(1 + p)} \quad \text{and } \mathbb{P}(X = k) = \frac{2}{\alpha}\frac{1 - p}{1 + p}p^k, \qquad k \geq 1, \tag{17}$$

see Inusah and Kozubowski (2006). In particular, if $\alpha = 2p/(1 + p)$, then X has the geometric distribution 13 with $u = 1 - p$ (in which case I takes on the value of zero with a positive probability). It turns out that the asymmetric case is quite different, as Eq. (16) does not hold with any X.

Proposition 3.2 *Let $\mathbb{P}(I = 1) = \alpha$, $\mathbb{P}(I = -1) = \beta$, and $\mathbb{P}(I = 0) = 1 - \alpha - \beta$, where α and β are positive real numbers satisfying $\alpha + \beta \leq 1$. Then, if $Y \sim DL(p, q)$ with $p \neq q$ then the relation (16) cannot hold with any discrete r.v. X, which is independent of I and supported on non-negative integers.*

Next, we seek representations of the form

$$Y \stackrel{d}{=} IX - (1 - I)W, \tag{18}$$

where I is a Bernoulli random variable with $\mathbb{P}(I = 1) = 1 - P(I = 0) = \alpha$ while X and W have discrete distributions supported on non-negative integers, and all three variables are mutually independent. Recall that if Y has the skew Laplace distribution (2) then Eq. (18) holds with the above I and independent, exponentially distributed X and W, see, e.g., Kotz et al. 2001). The following result holds in the above notation.

Proposition 3.3 *If $Y \sim DL(p, q)$ then the relation (18) holds with I, X, W as above if and only if*

$$\frac{p(1 - q)}{1 - pq} \leq \alpha \leq \frac{1 - q}{1 - pq}, \tag{19}$$

in which case

$$\mathbb{P}(X = 0) = 1 - \frac{p(1 - q)}{\alpha(1 - pq)}, \quad \mathbb{P}(W = 0) = 1 - \frac{q(1 - p)}{(1 - \alpha)(1 - pq)}, \tag{20}$$

$$\mathbb{P}(X = k) = \frac{(1 - p)(1 - q)}{\alpha(1 - pq)}p^k, \qquad \mathbb{P}(W = k) = \frac{(1 - p)(1 - q)}{(1 - \alpha)(1 - pq)}q^k, \quad k \geq 1. \tag{21}$$

Remark 3.4 In order for X or W to have a geometric distribution, α must take one of the boundary values in Eq. (19). In this case, if $\alpha = p(1-q)/(1-pq)$, then X has the (shifted) geometric distribution (15) with $u = 1-p$ and W has the geometric distribution (13) with $u = 1-q$. Similarly, if $\alpha = (1-q)/(1-pq)$, then X has the geometric distribution (13) with $u = 1-p$ and W has the (shifted) geometric distribution (15) with $u = 1-q$.

Our last representation shows that the DL distribution can be obtained by a certain "discretization" of the skew Laplace distribution. The result below generalizes the symmetric case $p = q$ discussed in Inusah and Kozubowski (2006).

Proposition 3.4 *Let X have a skew Laplace distribution given by the p.d.f. (2), and let $p = \mathrm{e}^{-\frac{\kappa}{\sigma}}$ and $q = \mathrm{e}^{-\frac{1}{\kappa\sigma}}$. Then,*

$$\varepsilon_1 = \frac{\log\left(\frac{q-pq}{1-pq}\frac{\log pq}{\log p}\right)}{\log q} \in (0,1), \quad \varepsilon_2 = \frac{\log\left(\frac{p-pq}{1-pq}\frac{\log pq}{\log q}\right)}{\log p} \in (0,1), \tag{22}$$

and the variable

$$Y = \begin{cases} n & \text{if} n - \varepsilon_1 < X \le n + 1 - \varepsilon_1, \ n \le -1, \\ 0 & \text{if} -\varepsilon_1 < X < \varepsilon_2 \\ n & \text{if} n - 1 + \varepsilon_2 \le X < n + \varepsilon_2, \ n \ge 1, \end{cases} \tag{23}$$

where $n \in \mathbb{Z}$, has the $DL(p,q)$ distribution.

4 Further properties

In this section, we show that DL laws are infinitely divisible, establish their stability properties with respect to geometric compounding, and discuss their maximum entropy property.

4.1 Infinite divisibility

Since the geometric distribution (15) is infinitely divisible, in view of the representation (14), it is clear that $Y \sim \mathrm{DL}(p,q)$ is infinitely divisible. Moreover, by well-known factorization properties of the geometric law (see, e.g., Feller 1957), $Y \sim \mathrm{DL}(p,q)$ admits a representation involving negative binomial $\mathrm{NB}(r,u)$ variables with the p.m.f.

$$f(k) = \binom{r+k-1}{k} u^r (1-u)^k, \quad k = 0,1,2,\cdots, \quad r > 0, \quad u \in (0,1). \tag{24}$$

Namely, we have

$$Y \stackrel{d}{=} X_{n1} + X_{n2} + \cdots + X_{nn}, \quad n = 1,2,3,\ldots, \tag{25}$$

where $X_{ni} \stackrel{d}{=} W_{n1} - W_{n2}$ are i.i.d. and W_{n1}, W_{n2} are independent, NB$(1/n, 1-p)$ and NB$(1/n, 1-q)$ variables, respectively. A canonical representation of the DL(p, q) ch.f. follows from the compound Poisson representations of the geometric variables X_1 and X_2 appearing in Eq. (14),

$$X_1 \stackrel{d}{=} \sum_{i=1}^{Q_p} Z_i, \quad X_2 \stackrel{d}{=} \sum_{j=1}^{Q_q} T_j, \tag{26}$$

see, e.g., Feller (1957). Here, the i.i.d. variables Z_i and T_j have logarithmic distributions (see, e.g., Johnson et al. 1993) with the p.m.f.'s

$$\mathbb{P}(Z_1 = k) = \frac{1}{\lambda_p}\frac{p^k}{k}, \quad \mathbb{P}(T_1 = k) = \frac{1}{\lambda_q}\frac{q^k}{k}, \quad k = 1, 2, 3, \ldots, \tag{27}$$

respectively, where $\lambda_p = -\log(1-p)$ and $\lambda_q = -\log(1-q)$. The variables Q_p and Q_q have Poisson distributions with means λ_p and λ_q, respectively, and the variables appearing in each of the representations in Eq. (26) are independent. When we write Eq. (26) and (14) in terms of the ch.f.'s, then (after some algebra) we obtain the following canonical representation of the DL(p, q) ch.f.,

$$\varphi(t|p,q) = \exp\{\lambda(\phi(t) - 1)\} = \exp\left\{\lambda \int_{-\infty}^{\infty} \left(\mathrm{e}^{\mathrm{i}tx} - 1\right) \mathrm{d}G(x)\right\}, \quad t \in \mathbb{R},$$

where $\lambda = \lambda_p + \lambda_q$ and ϕ, G are the ch.f. and the c.d.f. corresponding to a skew *double logarithmic* distribution with the p.m.f.

$$f(k) = \frac{1}{\lambda}\begin{cases} \frac{p^k}{k}, & k = 1, 2, 3, \ldots, \\ \frac{q^{|k|}}{|k|}, & k = -1, -2, -3, \ldots. \end{cases} \tag{28}$$

We summarize this discussion in the result below.

Proposition 4.1 *Let $Y \sim DL(p, q)$ and let Q have a Poisson distribution with mean $\lambda = -\log(1-p) - \log(1-q)$. Then Y is infinitely divisible and admits the representation*

$$Y \stackrel{d}{=} \sum_{j=1}^{Q} V_j, \tag{29}$$

where the V_j's have the double-logarithmic distribution (28), and are independent of Q.

4.2 Stability with respect to geometric summation

Let $X_1, X_2, \ldots$ be i.i.d. DL(p, q) variables, and let N_u be a geometric variable with p.m.f. Eq. (13), and independent of the X_i's. The quantity

$$Y \stackrel{d}{=} \sum_{i=1}^{N_u} X_i \tag{30}$$

is the geometric convolution (compound) of the X_i's. It is well known that the class of Laplace distributions is closed with respect to taking geometric convolutions, see, e.g., Kotz et al. (2001). Our next result, which extends the symmetric case of Inusah and Kozubowski (2006), shows that this property is also shared by the class of skew DL laws.

Proposition 4.2 *Let $X_1, X_2, \ldots$ be i.i.d. $DL(p,q)$ variables, and let N_u be a geometric variable (13), independent of the X_i's. Then the variable (30) has the $DL(s,r)$ distribution with*

$$s = \frac{2p}{p+q+(1-p)(1-q)u+\sqrt{[p+q+(1-p)(1-q)u]^2-4pq}}, \quad r = \frac{sq}{p}. \tag{31}$$

Next, we study the question whether for each $u \in (0, 1)$ a DL(s, r) r.v. Y admits the representation (30) with some i.i.d. variables X_i, in which case Y is said to be *geometric infinitely divisible* (see Klebanov et al. 1984). Our next results, which an extension of the symmetric case establish by Inusah and Kozubowski (2006), shows that this is indeed the case, and the variables X_i are DL distributed themselves. This shows that skew DL distributions are geometric infinitely divisible, which is in contrast with the geometric distribution itself (see, e.g., Kozubowski and Panorska, 2005).

Proposition 4.3 *If $Y \sim DL(s,r)$ and N_u is geometric with p.m.f. (13), then the representation (30) holds. The X_i's are i.i.d. and $DL(p,q)$ distributed, where*

$$p = \frac{2su}{(r+s)u+(1-s)(1-r)+\sqrt{[(r+s)u+(1-s)(1-r)]^2-4sru^2}}, \quad q = \frac{rp}{s}. \tag{32}$$

4.3 Maximum entropy property

The entropy of a one-dimensional random variable X with density (or probability function) f is defined as

$$H(X) = \mathbb{E}(-\log f(X)). \tag{33}$$

The principle of maximum entropy states that, of all distributions that satisfy certain constraints, one should select the one with the largest entropy, as this distribution does not incorporate any extraneous information other than that specified by the relevant constrains (see Jaynes, 1957). This concept has been successfully applied in a many fields, including statistical mechanics, statistics, stock market analysis, queuing theory, image analysis, reliability estimation (see, e.g., Kapur, 1993). It is well known that the skew Laplace distribution (2) maximizes the entropy among all continuous distributions on $\mathbb{R}$ with specified mean and the first absolute moment, see Kotz et al. (2002). Below, we show that the DL$(p,\ q)$ distribution maximizes the entropy under the same conditions among all *discrete distributions on integers*.

Proposition 4.4 *Consider the class $\mathcal{C}$ of all discrete distributions on integers with non vanishing densities such that*

$$\mathbb{E}X = c_1 \in \mathbb{R} \quad \text{and} \quad \mathbb{E}|X| = c_2 > 0 \quad \text{for} \quad X \in \mathcal{C}. \tag{34}$$

Then the entropy Eq. (33) is maximized by $Y \sim DL(p, q)$, where

$$q = \frac{(c_2 - c_1)(1 + c_1)}{1 + (c_2 - c_1)c_1 + \sqrt{1 + (c_2 - c_1)(c_2 + c_1)}}, \quad p = \frac{q + c_1(1-q)}{1 + c_1(1-q)} \tag{35}$$

when $c_1 \geq 0$ and

$$p = \frac{(c_2 + c_1)(1 - c_1)}{1 - (c_2 + c_1)c_1 + \sqrt{1 + (c_2 + c_1)(c_2 - c_1)}}, \quad q = \frac{p - c_1(1-p)}{1 - c_1(1-p)} \tag{36}$$

when $c_1 \leq 0$. Moreover,

$$\begin{aligned}\max_{X \in \mathcal{C}} H(X) = H(Y) = &-\log \frac{(1-p)(1-q)}{1-pq} \\ &- \frac{(1-p)(1-q)}{1-pq}\left(\frac{p \log p}{(1-p)^2} + \frac{q \log q}{(1-q)^2}\right).\end{aligned} \tag{37}$$

Remark 4.1 Note that if $c_1 = 0$, the DL distribution Y above is symmetric with

$$p = q = \frac{c_2}{1 + \sqrt{1 + c_2^2}} \in (0, 1) \tag{38}$$

while the maximum entropy is

$$\max_{X \in \mathcal{C}} H(X) = H(Y) = -\log \frac{1-p}{1+p} - \frac{2p \log p}{1 - p^2}. \tag{39}$$

The same solution is obtained even if we drop the condition $\mathbb{E}X = c_1$ from Eq. (34) altogether (see Kozubowski and Inusah 2006).

5 Estimation

In this section we derive maximum likelihood and method of moments estimators of p and q, and establish their asymptotic properties. We begin with the Fisher information matrix

$$I(p, q) = \left[-\mathbb{E}\left(\frac{\partial^2}{\partial \gamma_i \partial \gamma_j} \log f(Y|\gamma_1, \gamma_2)\right)\right]_{i,j=1,2}, \tag{40}$$

where Y has the DL(p, q) distribution with the vector-parameter $\gamma = (\gamma_1, \gamma_2)' = (p, q)'$ and density f given by Eq. (3). Routine albeit lengthy calculations produce

$$I(p, q) = \frac{1}{(1-pq)^2}\begin{bmatrix} \frac{(1-q)(1-qp^2)}{p(1-p)^2} & -1 \\ -1 & \frac{(1-p)(1-pq^2)}{q(1-q)^2} \end{bmatrix}. \tag{41}$$

Let $X_1, \dots, X_n$ be a random sample from a DL(p, q) distribution with density Eq. (3), and let $x_1, \dots, x_n$ be its particular realization. Then the log-likelihood function is

$$\log\, L(p,q) = n\left(\log(1-p)+\log(1-q) - \log(1-pq)+\bar{x}_n^+ \log p+\bar{x}_n^- \log q\right), \tag{42}$$

where

$$\bar{x}_n^+ = \frac{1}{n}\sum_{i=1}^n x_i^+, \quad \bar{x}_n^- = \frac{1}{n}\sum_{i=1}^n x_i^-,$$

and x^+ and x^- are the positive and the negative parts of x, respectively: $x^+ = x$ if $x \geq 0$ and zero otherwise, $x^- = (-x)^+$.

Consider first the case $\bar{x}_n^+ = \bar{x}_n^- = 0$ (all sample values are zero). Then it is easy to see that the log-likelihood function is maximized by $p = q = 0$ (corresponding to a distribution degenerate at zero). Next, when $\bar{x}_n^+ > 0$ but $\bar{x}_n^- = 0$ (all sample values are non-negative), the log-likelihood function is decreasing in $q \in [0, 1)$ for each fixed $p \in (0, 1)$ (as can be verified by taking the derivative). Therefore, for each $p, q \in (0, 1)$ we have

$$\log\, L(p,q) \leq \log\, L(p,0) = n\left(\log(1-p) + \bar{x}_n^+ \log\, p\right).$$

By taking the derivative of the right-hand-side above it is easy to verify that the maximum likelihood estimators (MLE's) of p and q are unique values given by

$$\hat{p}_n = \frac{\bar{X}_n^+}{1+\bar{X}_n^+}, \quad \hat{q}_n = 0, \tag{43}$$

corresponding to a geometric distribution with the p.m.f. (6). Similar considerations show that in case $\bar{x}_n^+ = 0$ but $\bar{x}_n^- > 0$ (all sample values are non-positive), the MLE's of p and q are unique values given by

$$\hat{p}_n = 0, \quad \hat{q}_n = \frac{\bar{X}_n^-}{1+\bar{X}_n^-}, \tag{44}$$

corresponding to a geometric distribution on non-positive integers Eq. (7).

Finally, consider the case $\bar{x}_n^+ > 0$ and $\bar{x}_n^- > 0$ (at least one positive and one negative sample value). Since the log-likelihood function (42) is continuous on the closed unit square (taking negative infinity on its boundary) and differentiable on the open unit square, it is clear that it takes the maximum value in its interior, at the point (p, q) where the partial derivatives are zero. This leads to the system of equations

$$\frac{p}{1-pq} + \frac{\bar{x}_n^-}{q} = \frac{1}{1-q}, \quad \frac{q}{1-pq} + \frac{\bar{x}_n^+}{p} = \frac{1}{1-p}. \tag{45}$$

It turns out that this system of equations is equivalent to Eq. (60) in Proposition 6.2, where $c_1 = \bar{x}_n^+ - \bar{x}_n^- = \bar{x}_n$ (the sample mean) and $c_2 = \bar{x}_n^+ + \bar{x}_n^- = \bar{|x|}_n$ (the sample first absolute moment). Indeed, the first of the equations in Eq. (60) is obtained by multiplying both sides of the two equations in Eq. (45) by q and p, respectively, and subtracting the corresponding sides. The second equation arises by solving the two equations in Eq. (45) for $\bar{x}_n^+$ and $\bar{x}_n^-$ and adding the corresponding sides of the resulting expressions. Thus, by Proposition 6.2, the MLE's of p and q are the unique values of p and q given in Proposition 4.4 with the above c_1 and c_2. Note that we chose the solution (35) or (36) according to whether $\bar{x}_n \geq 0$ or $\bar{x}_n \leq 0$, respectively. Further, these solutions include the special cases $\bar{x}_n^+=0$ or $\bar{x}_n^-=0$ discussed above, so that the MLE's of p and q are always given by the unique solution of the system of equations (45), or equivalently, the system Eq. (60) with the above c_1 and c_2. The following result summarizes this discussion.

Proposition 5.1 *Let $X_1, \ldots, X_n$ be i.i.d. variables from the $DL(p, q)$ distribution. Then the MLE's of p and q are unique values given by Eqs. (35), (36) with $c_1 = \bar{X}_n^+ - \bar{X}_n^-$ and $c_2 = \bar{X}_n^+ + \bar{X}_n^-$, that is*

$$\hat{q}_n = \frac{2\bar{X}_n^-(1+\bar{X}_n)}{1+2\bar{X}_n^-\bar{X}_n+\sqrt{1+4\bar{X}_n^-\bar{X}_n^+}}, \quad \hat{p}_n = \frac{\hat{q}_n+\bar{X}_n(1-\hat{q}_n)}{1+\bar{X}_n(1-\hat{q}_n)} \tag{46}$$

when $\bar{X}_n \geq 0$ and

$$\hat{p}_n = \frac{2\bar{X}_n^+(1-\bar{X}_n)}{1-2\bar{X}_n^+\bar{X}_n+\sqrt{1+4\bar{X}_n^-\bar{X}_n^+}}, \quad \hat{q}_n = \frac{\hat{p}_n-\bar{X}_n(1-\hat{p}_n)}{1-\bar{X}_n(1-\hat{p}_n)} \tag{47}$$

when $\bar{X}_n \leq 0$.

Remark 5.1 Equating $\mathbb{E}Y$ and $\mathbb{E}|Y|$ given by Eq. (11) with the corresponding sample moments results in equations Eq. (60) with the same values of c_1 and c_2 as above. Thus, the method of moments estimators of p and q coincide with the MLE's. In addition, note that in view of Proposition 4.4 the same estimators are obtained by using the principle of maximum entropy without assuming any particular distribution, under the condition that the mean and the first absolute moment are given (and approximated by their sample counterparts).

Our final result describes asymptotic properties of the maximum likelihood (and the method of moments) estimators of DL parameters.

Proposition 5.2 *The MLE's of p and q given in Proposition 5.1 are*

(i) Consistent;
(ii) Asymptotically bivariate normal, with the asymptotic covariance matrix

$$\Sigma_{\mathrm{MLE}} = \frac{pq(1-p)(1-q)}{1+pq}\begin{bmatrix} \frac{(1-p)(1-pq^2)}{q(1-q)^2} & 1 \\ 1 & \frac{(1-q)(1-qp^2)}{p(1-p)^2} \end{bmatrix}; \tag{48}$$

(iii) Asymptotically efficient, that is, the above asymptotic covariance matrix coincides with the inverse of the Fisher information matrix (41).

6 Proofs

Proof of Proposition 3.2 If the relation (16) holds with some discrete r.v. X supported on non-negative integers, then $\mathbb{P}(Y = k) = \alpha\mathbb{P}(X = k)$ for $k \geq 1$ and $\mathbb{P}(Y = k) = \beta\mathbb{P}(X = -k)$ for $k \leq -1$. In view of Eq. (3), we obtain

$$\frac{1}{\alpha}\frac{(1-p)(1-q)}{1-pq}p^k = \mathbb{P}(X = k) = \frac{1}{\beta}\frac{(1-p)(1-q)}{1-pq}q^k, \quad k \geq 1. \tag{49}$$

But this implies $\alpha/\beta = (p/q)^k$, $k \geq 1$, which is clearly impossible (unless $p = q$).

Proof of Proposition 3.3 Let $p_k = \mathbb{P}(X = k)$ and $q_k = \mathbb{P}(W = k)$, where $k = 0, 1, 2, \ldots$. Suppose that Eq. (18) holds. Then α as well as the p_k's and the q_k's must satisfy the equations

$$\mathbb{P}(Y = 0) = \frac{(1-p)(1-q)}{1-pq} = \alpha p_0 + (1-\alpha)q_0 \tag{50}$$

and

$$\mathbb{P}(Y = k) = \frac{(1-p)(1-q)}{1-pq}p^k = \alpha p_k, \quad \mathbb{P}(Y = -k) = \frac{(1-p)(1-q)}{1-pq}q^k = (1-\alpha)q_k \ (k \geq 1). \tag{51}$$

Relations (51) lead to the probabilities (21). In turn, summing these probabilities from one to infinity and subtracting from 1 leads to the expressions for p_0 and q_0 given in Eq. (20). It can be verified by direct substitution that the probabilities given by Eq. (20) satisfy Eq. (50). The requirement that the probabilities in Eqs. (20) and (21) be not greater than one and nonnegative, respectively, leads to the condition on α given in Eq. (19). The other implication is obtained by reversing these steps.

Proof of Proposition 3.4 First, let us note that $\varepsilon_1, \varepsilon_2 \in (0, 1)$ whenever $p, q \in (0, 1)$. This is equivalent to the following double inequality:

$$p < \frac{p - pq}{1 - pq}\frac{\log pq}{\log q} < 1, \quad p, q \in (0, 1). \tag{52}$$

After simple algebra, we find that the first inequality in Eq. (52) is equivalent to $g(p) > q$, $p, q \in (0, 1)$, where for each $q \in (0, 1)$,

$$g(p) = \frac{\log p}{\log q}(1-q) + pq. \tag{53}$$

It is easy to see that the function g is decreasing on (0,1], so that $g(p) > g(1) = p$ for $p \in (0, 1)$, as desired. Similarly, the second inequality in Eq. (52) follows by noting that for each fixed $q \in (0, 1)$ the function

$$h(p) = \frac{p - pq}{1 - pq}\frac{\log pq}{\log q} \tag{54}$$

is increasing on (0, 1] (which can be verified by differentiation), so that $h(p) < h(1) = 1$ for $p \in (0, 1)$. Verification that the p.m.f. of Y defined by Eq. (23) is given by Eq. (3) is straightforward.

Proof of Proposition 4.2 If X_i's are i.i.d. with DL(p, q) distribution and characteristic function $\varphi(\cdot|p, q)$ given by Eq. (5), and N_u is a geometric variable with mean $1/u$ and p.m.f. (13), then the characteristic function of the geometric compound on the right-hand-side of Eq. (30) is

$$\begin{aligned}\varphi_Y(t) &= \frac{\varphi(t|p, q)u}{1-(1-u)\varphi(t|p, q)}\\ &= \frac{(1-p)(1-q)u}{(1-e^{it}p)(1-e^{-it}q)-(1-u)(1-p)(1-q)}.\end{aligned} \tag{55}$$

We show below that this function coincides with the characteristic function of the DL(s, r) distribution with s, r given by Eq. (31). Indeed, setting Eq. (55) equal to $\varphi(t|s, r)$, which is the characteristic function of the DL(s, r) distribution (see Eq. (5) with $p = s$ and $q = r$), produces the equation

$$\begin{aligned}&(1-p)(1-q)u(1-e^{it}s)(1-e^{-it}r)\\ &\quad = (1-s)(1-r)[(1-e^{it}p)(1-e^{-it}q)-(1-u)(1-p)(1-q)],\end{aligned} \tag{56}$$

which holds for each $t \in \mathbb{R}$. Some algebra shows that this is so whenever the following two equations hold simultaneously:

$$r(1-p)(1-q)u=q(1-s)(1-r), \quad s(1-p)(1-q)u=p(1-s)(1-r). \tag{57}$$

Dividing the corresponding sides of the equations above leads to $r = sq/p$. Substituting this term into the first one of these equations results in the following quadratic equation in s:

$$h(s) = qs^2 - s[p + q + (1-p)(1-q)] + p = 0. \tag{58}$$

Since $h(0) = p > 0$ and $h(1) = -u(1-p)(1-q) < 0$, it is clear that Eq. (58) admits a unique root in the interval $(0, 1)$. Applying the quadratic formula leads to the expression for s given in Eq. (31), and the result follows.

Proof of Proposition 4.3 The proof is analogous to that of Proposition 4.2. Here, we need to show that Eq. (56) can be solved for $p, q \in (0, 1)$ for each $u, s, r \in (0, 1)$. This follows if the Eq. (57) hold simultaneously. Substituting $q = rp/s$ into the first one of these equations results in the following quadratic equation in p:

$$g(p) = rup^2 - p[(r+s)u + (1-s)(1-r)] + su = 0. \tag{59}$$

Since $g(0) = su > 0$ and $g(1) = -(1-s)(1-r) < 0$, we conclude that Eq. (59) admits a unique solution in the interval $(0, 1)$, and the quadratic formula produces Eq. (32).

To prove Proposition 4.4, we need two auxiliary results.

Proposition 6.1 *If* $\sum a_i$ *and* $\sum b_i$ *are convergent series of positive numbers such that* $\sum a_i \geq \sum b_i$ *then* $\sum a_i \log \frac{b_i}{a_i} \leq 0$, *with equality if and only if* $a_i = b_i$.

This result is taken from Rao (1965), Result 1e.6 p. 47.

Proposition 6.2 *The system of equations*

$$\frac{p}{1-p} - \frac{q}{1-q} = c_1 \in \mathbb{R}, \quad \frac{p(1-q)^2 + q(1-p)^2}{(1-p)(1-q)(1-pq)} = c_2 \geq 0, \tag{60}$$

where $|c_1| \leq c_2$ *and* $p, q \in [0, 1)$, *admits a unique solution given by Eq. (35) when* $c_1 \geq 0$ *and (36) when* $c_1 \leq 0$.

Proof Assume first that $c_1 \geq 0$. Using the first equation in Eq. (60) to express p in terms of q we obtain the second equation in Eq. (35). Substituting this expression into the second equation in Eq. (60), after tedious algebra we obtain the following equation for q:

$$g(q) = (c_1 - 1)(c_1 - c_2)q^2 - 2(c_1^2 - c_1c_2 - 1)q + (c_1 + 1)(c_1 - c_2) = 0. \tag{61}$$

If $c_1 = c_2$, this equation produces $q=0$, which coincides with the expression for q in Eq. (35). Otherwise, if $c_1=1$, Eq. (61) reduces to a linear equation with solution $q = (c_2 - 1)/c_2$, which again coincides with the expression for q in Eq. (35). Finally, if $c_1 \neq c_2$ and $c_1 \neq 1$, the function g is quadratic in q with $g(0) = (c_1 + 1)(c_1 - c_2) < 0$ and $g(1) = 2 > 0$, so there is a unique solution of Eq. (61) in $(0, 1)$. An application of the quadratic formula shows that this solution is given by the expression for q in Eq. (35). Assume now that $c_1 \leq 0$. Then, since the first equation in (60) can be written as $q/(1-q) - p/(1-p) = -c_1 \geq 0$ while the second equation is symmetric in p and q, it follows that the solution is the same as that in case $c_1 \geq 0$, where c_1 is replaced by $-c_1$ and p is interchanged with q. This leads to Eq. (36), and the result follows. □

Proof of Proposition 4.4 Let Y have a DL(p, q) distribution with the p.m.f. p_k given by Eq. (3), satisfying the constraint Eq. (34). Let the variable X have any distribution on the integers with non-vanishing p.m.f. $q_k = \mathbb{P}(X = k)$, $k \in \mathbb{Z}$, also satisfying Eq. (34). As before, denote by X^+ and X^- the positive and the negative parts of X, respectively. Then by Eq. (34) we have

$$\mathbb{E}X^+ = \sum_{k=0}^{\infty} kq_k = \frac{c_1 + c_2}{2}, \quad \mathbb{E}X^- = -\sum_{k=-\infty}^{0} kq_k = \frac{c_2 - c_1}{2}. \tag{62}$$

Next, by Proposition 6.1 with $a_i = q_i$ and $b_i = p_i$ (so that $\sum a_i = \sum q_i = 1 \geq 1 = \sum p_i = \sum b_i$), it follows that

$$H(X) = -\sum_{k=-\infty}^{\infty} q_k \log q_k \leq -\sum_{k=-\infty}^{\infty} q_k \log p_k = -\log \frac{(1-p)(1-q)}{1-pq} - Q, \tag{63}$$

where

$$Q = \sum_{k=-\infty}^{0} |k|q_k \log q + \sum_{0}^{\infty} kq_k \log p = \log q\mathbb{E}X^- + \log p\mathbb{E}X^+. \tag{64}$$

Thus, in view of Eq. (62), we have the following upper bound for the entropy $H(X)$ of $X \in \mathcal{C}$:

$$H(X) \leq -\log \frac{(1-p)(1-q)}{1-pq} - \left(\log q \frac{c_2 - c_1}{2} + \log p \frac{c_1 + c_2}{2}\right). \quad (65)$$

A straightforward calculation shows that this bound is actually attained by Y. If we now use Eq. (34) along with the expressions for $\mathbb{E}Y$ and $\mathbb{E}|Y|$ given by Eq. (11), Proposition 6.2 produces the values of p and q given in Eqs. (35) and (36). Substituting c_1 and c_2 into the right-hand-side of Eq. (65) yields the maximum entropy Eq. (37).

To prove Proposition 5.2 we need the following lemma, which can be established by a straightforward albeit lengthy algebra.

Lemma 6.1 *Let $X \sim DL(p,q)$, and define $\mathbf{W} = [X^+, X^-]'$. Then the mean vector and the covariance matrix of $\mathbf{W}$ are*

$$\mathbb{E}\mathbf{W} = \left[\frac{p(1-q)}{(1-p)(1-pq)}, \frac{q(1-p)}{(1-q)(1-pq)}\right]' \quad (66)$$

and

$$\Sigma_{\mathbf{W}} = \frac{1}{(1-pq)^2}\begin{bmatrix} \frac{p(1-q)(1-qp^2)}{(1-p)^2} & -pq \\ -pq & \frac{q(1-p)(1-pq^2)}{(1-q)^2} \end{bmatrix}. \quad (67)$$

Proof of Proposition 5.2 We start with Part (i). Note that according to Eq. (46), when $\bar{X}_n \geq 0$ the MLE's are of the form $\hat{p}_n = F_1(\bar{X}_n^+, \bar{X}_n^-)$ and $\hat{q}_n = F_2(\bar{X}_n^+, \bar{X}_n^-)$, where

$$F_1(x,y) = \frac{2y + (x-y)(1+\sqrt{1+4xy})}{(1+\sqrt{1+4xy})(1+x-y)},$$
$$F_2(x,y) = \frac{2y(1+x-y)}{1+2y(x-y)+\sqrt{1+4xy}}. \quad (68)$$

Further, using this notation we can write the expressions in Eq. (47) corresponding to the case $\bar{X}_n \leq 0$ as $\hat{p}_n = F_2(\bar{X}_n^-, \bar{X}_n^+)$ and $\hat{q}_n = F_1(\bar{X}_n^-, \bar{X}_n^+)$, respectively. This allows us to express the MLE's as $(\hat{p}_n, \hat{q}_n) = G(\frac{1}{n}\sum_{j=1}^n \mathbf{W}_j)$, where the $\mathbf{W}_j$'s are i.i.d. bivariate copies of $\mathbf{W}$ defined in Lemma 6.1 and $G(x,y) = (G_1(x,y), G_2(x,y))$ where

$$G_1(x,y) = \begin{cases} F_1(x,y) & \text{when } x \geq y \\ F_2(y,x) & \text{when } x \leq y \end{cases} \quad \text{and } G_2(x,y) = \begin{cases} F_2(x,y) & \text{when } x \geq y \\ F_1(y,x) & \text{when } x \leq y \end{cases}. \quad (69)$$

Since the $\mathbf{W}_j$'s satisfy the law of large numbers and the function G is continuous, we have

$$G\left(\frac{1}{n}\sum_{j=1}^n \mathbf{W}_j\right) \xrightarrow{d} G(\mathbb{E}\mathbf{W}) = (G_1(\mathbb{E}\mathbf{W}), G_2(\mathbb{E}\mathbf{W})).$$

Substituting $\mathbb{E}\mathbf{W}$ given in Lemma 6.1 into G_1 and G_2 defined above utilizing the expressions for F_1 and F_2 given in Eq. (68) results in (p, q). This proves consistency. We now move to Part (ii). By (bivariate) central limit theorem, we have

$$\sqrt{n}\left(\frac{1}{n}\sum_{j=1}^{n}\mathbf{W}_j - \mathbb{E}\mathbf{W}\right) \xrightarrow{d} N(\mathbf{0},\ \Sigma_{\mathbf{W}}),$$

where $\mathbb{E}\mathbf{W}$ is given by Eq. (66) as before and the right-hand-side is a bivariate normal distribution with mean zero and covariance matrix given by Eq. (67). Then, by standard large sample theory results (see, e.g., Rao 1965), we have

$$\sqrt{n}\left(G\left(\frac{1}{n}\sum_{j=1}^{n}\mathbf{W}_j\right) - G(\mathbb{E}\mathbf{W})\right) \xrightarrow{d} N(\mathbf{0},\ \Omega),$$

where $\Omega = D\Sigma_{\mathbf{W}}D'$ and

$$D = \left[\left.\frac{\partial G_i}{\partial x_j}\right|_{(x_1,x_2)=\mathbb{E}\mathbf{W}}\right]_{i,j=1,2}$$

is the matrix of partial derivatives of the functions G_1 and G_2 defined by Eq. (69). Rather lengthy computations produce

$$D = \begin{bmatrix} \dfrac{1-pq^2}{1+pq}\dfrac{(1-p)^2}{1-q} & \dfrac{p(1-p)(1-q)}{1+pq} \\[2ex] \dfrac{q(1-p)(1-q)}{1+pq} & \dfrac{1-qp^2}{1+pq}\dfrac{(1-q)^2}{1-p} \end{bmatrix}.$$

After straightforward but laborious matrix multiplications we find that $D\Sigma_{\mathbf{W}}D'$ reduces to Eq. (48) given in the statement of Proposition 5.2. Finally, to establish Part (iii), take the inverse of the Fisher information matrix (41) and verify that it coincides with the asymptotic covariance matrix Ω. This concludes the proof.

Acknowledgements We thank the referees for their valuable comments and corrections. Research was partially supported by NSF grant DMS-0139927.

References

Dasgupta, R. (1993). Cauchy equation on discrete domain and some characterization theorems. *Theory of Probability and its Applications*, *38*(3), 520–524.

Feller, W. (1957). *An Introduction to Probability Theory and Its Applications* (2nd Ed). New York: Wiley.

Graham, R.L., Knuth, D.E., Patashnik, O. (1989). *Concrete Mathematics*. Reading: Adison-Wesley Publishing Company.

Inusah, S. (2003). *Discrete Laplace Distributions*. Thesis, Department of Mathematics and Statistics, University of Nevada at Reno.

Inusah, S., Kozubowski, T.J. (2006). A discrete analogue of the Laplace distribution. *Journal of Statistical Planning and Inference*, *136*, 1090–1102.

Irwin, J.O. (1937). The frequency distribution of the difference between two independent variates following the same Poisson distribution. *Journal of the Royal Statistical Society*, *100*(3), 415–416.

Jaynes, E.T. (1957). Information theory and statistical mechanics I. *Physical Review*, *106*, 620–630.

Johnson, N.L., Kotz, S., Kemp, A.W. (1993). *Univariate Discrete Distributions* (2nd Ed). New York: Wiley.

Kapur, J.N. (1993). *Maximum-Entropy Models in Science and Engineering* (revised edition). New York: Wiley.

Kemp, A.W. (1997). Characterizations of a discrete normal distribution. *Journal of Statistical Planning and Inference*, *63*, 223–229.

Klebanov, L.B., Maniya, G.M., Melamed, I.A. (1984). A problem of Zolotarev and analogs of infinitely divisible and stable distributions in a scheme for summing a random number of random variables. *Theory of Probability and its Applications*, *29*, 791–794.

Kotz, S., Kozubowski, T.J., Podgórski, K. (2001). *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*. Boston: Birkhäuser.

Kotz, S., Kozubowski, T.J., Podgórski, K. (2002). Maximum entropy characterization of asymmetric Laplace distribution. *International Mathematical Journal*, *1*, 31–35.

Kozubowski, T.J., Panorska, A.K. (2005). A mixed bivariate distribution with exponential and geometric marginals. *Journal of Statistical Planning and Inference*, *134*, 501–520.

Lisman, J.H.C., van Zuylen, M.C.A. (1972). Note on the generation of most probable frequency distributions. *Statistica Neerlandica*, *26*(1), 19–23.

Rao, C.R. (1965). *Linear Statistical Inference and Its Applications*, New York: Wiley

Szablowski, P.J. (2001). Discrete normal distribution and its relationship with Jacobi Theta functions. *Statistics and Probability Letters*, *52*, 289–299.