The characteristic function of the discrete Cauchy distribution

In Memory of T. Cacoullos

Nickos Papadatos
Department of Mathematics, National and Kapodistrian University of Athens,
Panepistemiopolis, 157 84 Athens, Greece

March 25, 2022

Abstract

A new family of integer-valued Cauchy-type distributions is introduced, the *Cauchy-Cacoullos family.* The characteristic function is evaluated, showing some interesting distributional properties, similar to the ordinary (continuous) Cauchy scale family. The results are extendable to discrete Student-type distributions with odd degrees of freedom.

Keywords: Fourier series; discrete Student distribution; Cauchy-Cacoullos family.

1 Introduction and summary

Some years ago, Cacoullos (Personal Communication), considering discretization of well-known continuous distributions, introduced a (standard) discrete Cauchy random variable (r.v.) X with probability mass function (p.m.f.)

$$\mathbb{P}(X = k) = \frac{1/\pi_0}{1 + k^2}, \quad k \in \mathrm{Z}, \tag{1}$$

by the obvious substitution $k \in \mathrm{Z}$ for $x \in \mathrm{R}$ in the standard Cauchy density

$$f(x) = \frac{1/\pi}{1 + x^2}, \quad x \in \mathrm{R}. \tag{2}$$

Cacoullos immediately raised two natural questions:

(A) While it is expected to be very close to π, what is the exact value of the normalizing constant π_0 in (1)?

(B) While the characteristic function (ch.f.) of (2) is $\phi(t) = e^{-|t|}$, what is the corresponding one, say ϕ_1, of (1)?

We provide explicit answers in section 2. It is well-known that the (continuous) Cauchy distribution appears naturally in statistics and probability. At this point it should be noted though the standard Cauchy r.v. is customarily defined as the ratio of two independent standard normal r.v.'s, or as the tangent of a randomly chosen angle in $[0, 2\pi)$, it has recently been shown ([1], [6], [7]) that the ratio representation still holds if (X, Y) follows any bivariate spherically symmetric distribution.

In [2], Cacoullos showed that if $X = (X_1, \ldots, X_p)'$ ($p \geq 3$) is spherically symmetrically distributed around zero then all polar angle tangent vectors follow a multivariate Cauchy; note that, e.g., Feller (1966) defines the symmetric bivariate and trivariate Cauchy distributions directly through their densities – not as tangent vectors.

In contrast to (2) and its location-scale extension, for which several applications are known both in probability and statistics, for (1) we have been able to find few results related to stochastic processes – see, e.g., [14], p. 383. However, the asymptotic distribution of the sample means for (1), Theorem 4, may serve as a starting point for applications; so appears to be the Cauchy-Cacoullos family defined by (4). These considerations are, however, beyond the scope of the present note.

In section 3 we introduce a novel family of integer-valued distributions, the *Cauchy-Cacoullos family*, sharing similar properties – see Definition 1 and Remark 2. In particular, any distribution in this family has a simple characteristic function that can be written down explicitly, Theorem 2, and the same is valid for the discrete Student-type distributions of Remark 2. Basic inference properties for this family are included in Theorem 3, while some distributional properties are discussed in some detail in Section 4; see Theorems 4–6. We hope that the proposed simple formulae will enlarge the applicability of discrete Cauchy distribution in the future.

2 The characteristic function

Since $\phi_1(t) = \mathbb{E}e^{itX} = \mathbb{E}\cos(tX) + i\mathbb{E}\sin(tX)$ (i denotes the imaginary unit) and X is symmetrically distributed around the origin (hence, $\mathbb{E}\sin(tX) = 0$), both questions, (A), (B), will be answered if we manage to calculate in a closed form the function $g : \mathrm{R} \to \mathrm{R}$, defined by the Fourier series

$$g(t) := \sum_{n=0}^{\infty} \frac{\cos(nt)}{1+n^2}, \qquad t \in \mathrm{R}. \tag{3}$$

Therefore, the problem is to identify which function g is represented as a series of cosines with Fourier coefficients as in (3). Clearly, g is periodic with period 2π. Thus, it suffices to restrict our attention to t-values in the interval $-\pi \le t \le \pi$. On the other hand, since a cosine Fourier series corresponds to an even function, we may further restrict the t-values into the interval $0 \le t \le \pi$.

The key lemma is:

Lemma 1 *For $-2\pi \le t \le 2\pi$,*

$$g(t) = \frac{1}{2} + \frac{\pi\cosh(\pi - |t|)}{2\sinh(\pi)}.$$

We omit the proof because we shall show a more general result in Section 3, below.

Corollary 1 *The normalizing constant π_0 is given by*

$$\pi_0 = 2g(0) - 1 = \frac{\pi\cosh(\pi)}{\sinh(\pi)} = \pi\left(1 + \frac{2}{e^{2\pi} - 1}\right) \simeq 3.15334809493716\ldots\ .$$

The formula for the ch.f., and is an immediate consequence of Lemma 1 and (3):

Theorem 1 *The ch.f. of X is given by $\phi_1(t) = \cosh(\pi - |t|)/\cosh(\pi)$, $-2\pi \le t \le 2\pi$, and it is periodic with period 2π* (see Fig. 1).

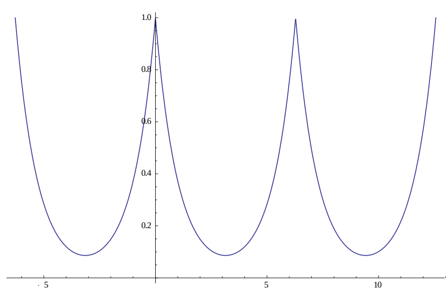

Figure 1: The characteristic function $\phi_1(t)$ in the interval $-2\pi \le t \le 4\pi$.

3 The Cauchy-Cacoullos family of discrete distributions

If we multiply a continuous Cauchy r.v. by a constant $\lambda > 0$ we stay in the same family of distributions – the Cauchy scale family. More precisely, if X is standard Cauchy, the density of λX is given by

$$f(x) = \frac{1}{\pi} \frac{\lambda}{\lambda^2 + x^2}, \quad x \in \mathrm{R}, \quad \lambda > 0.$$

However, this is no longer true for a discrete Cauchy X, since the support of λX is not the set of integers. Motivated from this observation, we define a family of discrete integer-valued distributions as follows:

Definition 1 The discrete Cauchy-Cacoullos family ($\mathcal{CC}$, for short) contains the p.m.f.'s

$$f_\lambda(k) = \frac{\tanh(\lambda\pi)}{\pi} \frac{\lambda}{\lambda^2 + k^2}, \qquad k \in \mathrm{Z}, \quad \lambda > 0. \tag{4}$$

For completeness of the presentation, it is convenient to include the limiting case $\lambda = 0$, which corresponds to a degenerate r.v. at zero.

Although this family has several interesting properties, similar to the Cauchy, it does not seem to have been studied elsewhere. Clearly, for $\lambda = 1$ we get (1). At a first glance, it is not entirely obvious to verify that the normalizing constant is as in (4). This is a by-product of the following result.

Lemma 2 *For $-\pi \le t \le \pi$ and $\lambda > 0$,*

$$\cosh(\lambda t) = \frac{\lambda \sinh(\lambda\pi)}{\pi} \left\{ \frac{1}{\lambda^2} + 2 \sum_{n=1}^{\infty} \frac{(-1)^n \cos(nt)}{\lambda^2 + n^2} \right\}.$$

Proof: We express the even function $h(t) = \cosh(\lambda t)$ in a cosine Fourier series to get $h(t) \sim \sum_{n=0}^{\infty} \alpha_n \cos(nt)$. Simple calculations show that

$$\alpha_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \cosh(\lambda u) du = \frac{\sinh(\lambda\pi)}{\pi\lambda}$$

and

$$\alpha_n = \frac{1}{\pi} \int_{-\pi}^{\pi} \cosh(\lambda u) \cos(nu) du = (-1)^n \frac{2\lambda \sinh(\lambda\pi)}{\pi(\lambda^2 + n^2)}, \quad n = 1, 2, \ldots\ .$$

Since h is differentiable in $[-\pi, \pi]$ with $h(-\pi) = h(\pi)$, the lemma is proved (and the series converges uniformly to h). Q.E.D.

If we set $\lambda = 1$ and $t \to t - \pi$ in Lemma 2 we obtain Lemma 1 with g as in (3).

Corollary 2 *We have*

$$\sum_{k=-\infty}^{\infty} \frac{1}{\lambda^2 + k^2} = \frac{\pi}{\lambda \tanh(\lambda\pi)},$$

and hence, (4) *defines a p.m.f. for any* $\lambda > 0$.

Proof: Substitute $t = \pi$ in Lemma 2. Q.E.D.

As for the case $\lambda = 1$, we can obtain the ch.f. of $X_\lambda \sim f_\lambda$ in a closed form.

Theorem 2 *The ch.f. of* X_λ *with p.m.f.* $f_\lambda \in \mathcal{CC}$ *is given by*

$$\phi_\lambda(t) = \frac{\cosh\big(\lambda(t-\pi)\big)}{\cosh(\lambda\pi)}, \quad 0 \le t \le 2\pi,$$

and it is periodic with period 2π. *More precisely,*

$$\phi_\lambda(t) = \frac{\cosh\Big(\lambda\Big(t - 2\pi\lfloor\frac{t}{2\pi}\rfloor - \pi\Big)\Big)}{\cosh(\lambda\pi)}, \quad -\infty < t < \infty,$$

where $\lfloor x \rfloor$ *denotes the integer part of* x.

Proof: As is well-known, all integer-valued r.v.'s have periodic ch.f.'s, with period 2π. The particular r.v. is symmetrically distributed around zero, and thus, its ch.f. is real and even, so that $\phi_\lambda(t) = \mathbb{E}\cos(tX_\lambda)$. To calculate this, we may restrict our attention in the interval $0 \le t \le 2\pi$. Then, since $-\pi \le t - \pi \le \pi$ and $\cos(nt) = (-1)^n \cos(n(t-\pi))$,

$$\begin{aligned} \phi_\lambda(t) &= \frac{\lambda\tanh(\lambda\pi)}{\pi}\left\{\frac{1}{\lambda^2} + 2\sum_{n=1}^{\infty}\frac{(-1)^n\cos(n(t-\pi))}{\lambda^2+n^2}\right\} \\ &= \frac{\lambda\tanh(\lambda\pi)}{\pi}\,\frac{\pi\cosh(\lambda(t-\pi))}{\lambda\sinh(\lambda\pi)}, \end{aligned}$$

where the second equality follows from Lemma 2. Q.E.D.

Statistical inference for the parameter λ is facilitated from the fact that the p.m.f.'s and the ch.f.'s in $\mathcal{CC}$ have tractable forms.

Theorem 3 *Consider a random sample* $X_1, \ldots, X_n \sim f_\lambda \in \mathcal{CC}$ *with* $\lambda > 0$ *unknown.*
(i) *The minimal sufficient statistic is* $T = (Y_1, \ldots, Y_n)$, *with* $Y_1 \le Y_2 \le \cdots \le Y_n$ *being the order statistics of* $|X_1|, \ldots, |X_n|$.
(ii) *The Fisher Information* (of a single observation) *is*

$$I(\lambda) = \frac{1}{2\lambda^2} + \frac{\pi}{\lambda}w(\lambda) \quad \text{where} \quad w(\lambda) = \frac{\lambda\pi}{\cosh(\lambda\pi)^2} - \frac{1}{\sinh(2\lambda\pi)}. \tag{5}$$

(iii) *The MLE* $\widehat{\lambda}_n$ *of* λ *is unique; it is given as the unique solution in* $[0,\infty)$ *of the equation*

$$\frac{\pi\lambda}{\sinh(2\pi\lambda)}+\frac{1}{n}\sum_{i=1}^{n}\frac{X_i^2}{\lambda^2+X_i^2}=\frac{1}{2}. \tag{6}$$

(iv) *The MLE is consistent and asymptotically efficient,*

$$\sqrt{n}\left(\widehat{\lambda}_n-\lambda\right)\stackrel{d}{\to}N(0,1/I(\lambda)),$$

where $\stackrel{d}{\to}$ *denotes weak convergence.*

Proof: Let $\mathbf{x}=(x_1,\ldots,x_n)$ and $\mathbf{y}=(y_1,\ldots,y_n)$ be two vectors in Z^n. Then, the likelihood ratio is given by

$$\frac{L(\mathbf{x};\lambda)}{L(\mathbf{y};\lambda)}=\prod_{i=1}^{n}\frac{\lambda^2+y_i^2}{\lambda^2+x_i^2},$$

and it has the same form as in the continuous Cauchy scale-family. Obviously, this ratio is independent of $\lambda>0$ if and only if the ordered squared values of $\mathbf{x}$ and $\mathbf{y}$ are identical, and this verifies (i). Now, a straightforward computation yields the score function

$$S(k;\lambda):=\frac{\partial}{\partial\lambda}\log f_\lambda(k)=\frac{1}{\lambda}+\frac{2\pi}{\sinh(2\pi\lambda)}-\frac{2\lambda}{\lambda^2+k^2}.$$

Let $X_\lambda\sim f_\lambda$. Using Remarks 1, 2 below, it is seen that $\mathbb{E}S(X_\lambda;\lambda)=0$ and $\mathbb{E}S(X_\lambda;\lambda)^2=I(\lambda)$ with $I(\lambda)$ as in (5). Note that $I(\lambda)=-\mathbb{E}\frac{\partial^2}{\partial\lambda^2}\log f_\lambda(X_\lambda)$, since the regularity conditions are obviously fulfilled; both formulae require computation of the series $\sum_n(\lambda^2+n^2)^{-s}$, $s=1,2,3$. Moreover, one can easily verify that the log-likelihood is given by

$$\frac{\partial}{\partial\lambda}\log L(\mathbf{x};\lambda)=\frac{2n}{\lambda}\left(\frac{\pi\lambda}{\sinh(2\pi\lambda)}-\frac{1}{2}+\frac{1}{n}\sum_{i=1}^{n}\frac{x_i^2}{\lambda^2+x_i^2}\right). \tag{7}$$

For fixed $\mathbf{x}\in\mathrm{Z}^n$, the positive function $u(\lambda):=\pi\lambda/\sinh(2\pi\lambda)+n^{-1}\sum_{i=1}^n x_i^2/(\lambda^2+x_i^2)$ decreases to zero as $\lambda\to\infty$ and has a limit $u(0+)\geq 1/2$ (it equals to $1/2$ iff $\mathbf{x}=\mathbf{0}$). Since u is strictly decreasing and continuous, the likelihood is first increasing and then decreasing, reaching its global maximum at λ_0, where $u(\lambda_0)=1/2$. This shows that the MLE is the unique solution of (6), it equals to 0 iff $\mathbf{X}=\mathbf{0}$, and it is otherwise positive. Finally, in order to prove (iv), fix $\lambda=\lambda_0$ and $c\in(0,\lambda_0)$, and assume that λ varies in the interval $(\lambda_0-c,\lambda_0+c)$. Then, $\frac{\partial^3}{\partial\lambda^3}\log f_\lambda(k)=A(\lambda)+B(\lambda,k)$ where

$$A(\lambda)=4\pi^3\,\frac{3+\cosh(4\lambda\pi)}{\sinh(2\lambda\pi)^3}+\frac{2}{\lambda^3},\quad B(\lambda,k)=4\lambda\frac{3k^2-\lambda^2}{(\lambda^2+k^2)^3}.$$

The function A is decreasing and positive, so that $|A(\lambda)| < A(\lambda_0 - c)$. Moreover,

$$|B(\lambda, k)| < 4\lambda \frac{3k^2 + 3\lambda^2}{(\lambda^2 + k^2)^3} < \frac{12(\lambda_0 + c)}{((\lambda_0 - c)^2 + k^2)^2} \leq \frac{12(\lambda_0 + c)}{(\lambda_0 - c)^4}.$$

It follows that we can find a finite constant $M = M(\lambda_0, c)$ such that $|\frac{\partial^3}{\partial \lambda^3} \log f_\lambda(k)| < M$ uniformly in $k \in \mathrm{Z}$, $\lambda \in (\lambda_0 - c, \lambda_0 + c)$, and the result follows by applying Theorem 3.10 in [11]. Q.E.D.

Unfortunately, the MLE does not admit a closed form and, hence, numerical procedures should be employed. On the other hand, we can construct closed-form consistent estimators, due to the fact that the ch.f. admits a simple form. For example, $\phi_\lambda(\pi) = 1/\cosh(\lambda\pi) = \beta$, say, equals to the difference $\mathbb{P}(X_\lambda \text{ even}) - \mathbb{P}(X_\lambda \text{ odd})$. This can be consistently and unbiasedly estimated by $\widehat{\beta}_n = n^{-1}\sum_{i=1}^n (-1)^{X_i}$, and a trivial application of the CLT leads to $\sqrt{n}(\widehat{\beta}_n - \beta) \xrightarrow{d} N(0, 1 - \beta^2)$, while the SLLN shows that $\widehat{\beta}_n$ is eventually positive w.p. 1. Applying the delta-method (see [16]) with $g(\beta) = \pi^{-1}\Big(\log(1 + \sqrt{1 - \beta^2}) - \log(\beta)\Big)$, so that $g(\beta) = \lambda$, we obtain

$$\sqrt{n}\left(g(\widehat{\beta}_n) - \lambda\right) \xrightarrow{d} N\Big(0, \cosh(\pi\lambda)^2/\pi^2\Big).$$

However, compared to the MLE, the closed-form estimator $g(\widehat{\beta}_n)$ is by far less efficient. Thus, it is natural to seek for closed-form highly efficient estimators, and this may be possible as in the continuous case. In the continuous case it is shown that the asymptotic relative efficiency of the geometric mean of the absolute values of the observations is $8/\pi^2 \simeq 81\%$, and in [10] a more efficient closed-form estimate is proposed. Also, highly efficient estimators that are based on the ch.f. may be obtained by adapting the methodology of [9] to the present discrete case. However, such results are beyond the scope of the present note. Note that the Fisher information in the continuous Cauchy scale family equals to $1/(2\lambda^2)$ (compare to (5)), and the likelihood equation is as in (7), with the absence of the term $\pi\lambda/\sinh(2\pi\lambda)$.

Remark 1 The series in Corollary 2 is of some interest in itself, because of the computation of the sum $\sum_{n=1}^\infty (\lambda^2 + n^2)^{-1}$ in a closed form. Then, e.g., taking limits as $\lambda \searrow 0$, we arrive at the famous Euler sum, $\sum_{n=1}^\infty n^{-2} = \pi^2/6$. Moreover, differentiating term by term with respect to λ we can evaluate the series

$$\sum_{n=1}^\infty \frac{1}{(\lambda^2 + n^2)^2}.$$

From this, taking limits as $\lambda \searrow 0$, we arrive at the sum for $\zeta(4)$, that is, $\sum_{n=1}^\infty n^{-4} = \pi^4/90$; clearly, this process can be continued to evaluate all $\zeta(2s)$ values, as well as the series $\sum_{n=1}^\infty (\lambda^2 + n^2)^{-s}$, $s = 1, 2, \ldots$.

Remark 2 Differentiating m times with respect to λ^2 the series in Lemma 2, it is possible to introduce and investigate discrete Student-type families with $\nu = 2m+1$ degrees of freedom, that is, p.m.f.'s of the form

$$f_{\nu;\lambda}(k) = \frac{c_{\nu;\lambda}}{(\lambda^2 + k^2)^{(\nu+1)/2}}, \quad k \in \mathrm{Z}, \quad \nu = 1, 3, 5, \ldots, \quad \lambda > 0, \tag{8}$$

admitting closed-form ch.f.'s $\phi_{\nu;\lambda}(t)$ and explicit normalizing constants $c_{\nu;\lambda}$. However, the situation becomes quite complicated for even values of ν.

4 Some distributional properties of the $\mathcal{CC}$ family

We observe that the ch.f. $\phi_\lambda(t)$ is not differentiable at the points $t = 2k\pi$, $k \in \mathrm{Z}$ (c.f. Fig. 1). It is known that a random variable Y_1 satisfies a weak law of large numbers, that is,

$$\overline{Y}_n := \frac{Y_1 + \cdots + Y_n}{n} \to \text{ some constant } c, \text{ in probability,}$$

if and only if its ch.f., ϕ_{Y_1}, is differentiable at $t = 0$; then, $\phi'_{Y_1}(0) = ic$ where i is the imaginary unit (the problem was treated by A. Zygmund and E.J.G. Pitman, and it is closely connected to Khintchine's weak law of large numbers; see Feller 1966, p. 528 and van der Vaart 1998, p. 15). Hence, the distributions of the $\mathcal{CC}$ family do not satisfy the weak law of large numbers, since their ch.f.'s are not differentiable at $t = 0$. Therefore, it is of some interest to study the asymptotic behavior of the sample means from a $\mathcal{CC}$ random variable with p.m.f. as in (4). Recall the well-known continuous counterpart, which says that $\overline{X}_n$ is the same Cauchy for all n (Cauchy r.v.'s are stable).

We have the following result.

Theorem 4 *If $X_1, X_2, \ldots$ are independent identically distributed random variables with p.m.f. as in* (4) *then*

$$\overline{X}_n \xrightarrow{d} \lambda \tanh(\lambda\pi)\ Z,$$

where Z is standard (continuous) *Cauchy with density* (2).

Proof: Fix $t \geq 0$. Theorem 2 shows that the ch.f. of $\overline{X}_n$ is given by

$$\phi_\lambda(t/n)^n = \left(\frac{\cosh\big(\lambda(\pi - t/n)\big)}{\cosh(\lambda\pi)} \right)^n, \ n \geq \frac{t}{2\pi}.$$

Using this, it is easy to verily (e.g., by taking logarithms) that $\phi_\lambda(t/n)^n \to e^{-ct}$, $t \geq 0$, where $c = \lambda \tanh(\lambda\pi)$. Finally, from the fact that ϕ_λ is even, it follows that $\phi_\lambda(t/n)^n \to e^{-c|t|}$ for all $t \in \mathrm{R}$, which is the ch.f. of cZ, and the result follows from the continuity theorem of characteristic functions. Q.E.D.

Unlike the usual Cauchy scale family, the $\mathcal{CC}$ family is not convolution closed; however, it is "almost" closed. More precisely, the following result holds.

Theorem 5 *For independent r.v.'s X, Y in $\mathcal{CC}$ with $X \sim f_{\lambda_1}$ and $Y \sim f_{\lambda_2}$, the ch.f. of $X+Y$ is given by*

$$\phi_{X+Y}(t) = \frac{\alpha(\lambda_1+\lambda_2)}{2\alpha(\lambda_1)\alpha(\lambda_2)}\phi_{\lambda_1+\lambda_2}(t) + \frac{\alpha(|\lambda_2-\lambda_1|)}{2\alpha(\lambda_1)\alpha(\lambda_2)}\phi_{|\lambda_2-\lambda_1|}(t), \quad t \in R,$$

where $\phi_0(t) \equiv 1$ is the ch.f. of the degenerate r.v. X_0 with $\mathbb{P}(X_0 = 0) = 1$, and $\alpha(\lambda) := \cosh(\lambda\pi)$, $\lambda \geq 0$. Consequently, $X+Y$ is a mixture of two r.v.'s that are members of $\mathcal{CC}$ family,

$$\mathbb{P}(X+Y=k) = \frac{\alpha(\lambda_1+\lambda_2)}{2\alpha(\lambda_1)\alpha(\lambda_2)}f_{\lambda_1+\lambda_2}(k) + \frac{\alpha(|\lambda_2-\lambda_1|)}{2\alpha(\lambda_1)\alpha(\lambda_2)}f_{|\lambda_2-\lambda_1|}(k), \quad k \in Z.$$

Proof: Set

$$p = \frac{\alpha(\lambda_1+\lambda_2)}{2\alpha(\lambda_1)\alpha(\lambda_2)}, \qquad q = \frac{\alpha(|\lambda_2-\lambda_1|)}{2\alpha(\lambda_1)\alpha(\lambda_2)}.$$

Obviously, $p > 0$ and $q > 0$. Also, using the formula

$$\cosh(x)\cosh(y) = \frac{1}{2}\cosh(x+y) + \frac{1}{2}\cosh(y-x) \tag{9}$$

it is easily seen that $p+q=1$. Restricting our attention to the interval $0 \leq t \leq 2\pi$, we have

$$\phi_{X+Y}(t) = \phi_{\lambda_1}(t)\phi_{\lambda_2}(t) = \frac{\cosh(\lambda_1(t-\pi))\cosh(\lambda_2(t-\pi))}{\alpha(\lambda_1)\alpha(\lambda_2)}$$

and a final application of (9) to the numerator, taking into account Theorem 2, completes the proof. Q.E.D.

Remark 3 If X, Y are i.i.d. from f_λ then, since $\alpha(0) = 1$ and $f_0(k) = I(k=0)$, we get

$$\mathbb{P}(X+Y=k) = \begin{cases} \dfrac{1}{2\cosh(\lambda\pi)^2} + \dfrac{\tanh(\lambda\pi)}{2\lambda\pi}, & k = 0, \\[2ex] \dfrac{\tanh(\lambda\pi)}{\pi}\,\dfrac{2\lambda}{(2\lambda)^2+k^2}, & k \in \mathrm{Z}^*. \end{cases}$$

This formula quantifies the fact that the p.m.f. of $X+Y$ lies outside $\mathcal{CC}$, but it is close, in some sense, to $f_{2\lambda}$; in fact, the ratio $f_{X+Y}(k)/f_{2\lambda}(k)$ does not vary with $k \in \mathrm{Z}^*$.

A ch.f. ϕ (or the corresponding r.v. X) is called infinitely divisible (i.d.) if for each n, we can find a ch.f. ϕ_n such that $\phi_n^n = \phi$; equivalently, if $X_{1,n} + \cdots + X_{n,n}$ has the same distribution as X, where $X_{1,n}, \ldots, X_{n,n}$ are i.i.d. with ch.f. ϕ_n. Properties of this kind are included in what is called "arithmetic of probability laws" ([12], [13]), and a vast bibliography exists, see, e.g., [5], [3], [8], [12], [13], [15], and references therein.

Since the notion of i.d. is related to limit theorems of sums of independent r.v.'s, it would be useful to know whether the $\mathcal{CC}$ family is i.d. This is indeed the case, and it follows immediately from a result of Polya, because the ch.f. ϕ_λ is even, log-convex in $[0, 2\pi]$ and 2π periodic, see [8], [13]. In fact, ϕ_λ^α is a ch.f. for all $\lambda \geq 0$ and $\alpha \geq 0$.

As is well known, the notion of self-decomposability, as well as that of stability, do not apply to discrete r.v.'s. Recall that X is stable if, for each n, we can find constants $\alpha_n > 0$ and $\beta_n \in \mathrm{R}$ such that X and $(X_1 + \cdots + X_n)/\alpha_n - \beta_n$ have the same distribution, where $X_1, \ldots, X_n$ are i.i.d. copies X. Obviously, the class of stable distributions is a proper subset of i.d. distributions. Due to a fundamental result of Lévy, stable distributions are very important because their class contains exactly all possible limits of (properly) normalized sums of i.i.d. r.v.'s. Every stable distribution has a ch.f. that can be expressed in a closed form, and the corresponding r.v. is absolutely continuous. The subclass of symmetric stable ch.f.'s, after a location-scale transformation, can be written as $\mathcal{S} = \{\phi_\alpha(t) = e^{-|t|^\alpha},\ 0 < \alpha \leq 2\}$. Only the densities that correspond to $\alpha = 1/2$ (Lévy), $\alpha = 1$ (Cauchy) and $\alpha = 2$ (Normal), have known explicit forms.

It is natural to ask whether the $\mathcal{CC}$ family contains discrete stable distributions, in the sense of [15]. However, the definitions in [15] are designed for non-negative integer-valued r.v.s, and are based on probability generating functions; it is not obvious how to extend these results to the $\mathcal{CC}$ case. The following definition provides a different approach that seems to be natural for our case.

Definition 2 Let Λ be a set of indices, consider a parametric family $\mathcal{F} = \{\phi_\lambda,\ \lambda \in \Lambda\}$ of discrete, integer-valued, ch.f.'s, and let $\mathcal{F}'$ be the corresponding family of random variables. Then, $\mathcal{F}$ is called discrete stable (DSF) if for each $\phi_\lambda \in \mathcal{F}$, we can find a sequence of indices $\{\lambda_n\}_{n=1}^\infty \subset \Lambda$ such that $\phi_{\lambda_n}^n \to \phi_\lambda$. Equivalently, if every random variable in $\mathcal{F}'$ is the weak limit of sums of i.i.d. r.v.'s from $\mathcal{F}'$.

The usual Poisson family is DSF, as well as the Negative Binomial. In order for such a model to be useful in practice, the family $\mathcal{F}$ should not contain "too many" ch.f.'s. Also, it is plausible to consider those DSF's that satisfy some kind of discrete attraction, in the sense that (non-normalized) sums of several i.i.d. discrete r.v.'s converge weakly to one of the members of the DSF. It is clear that the Compound Poisson that is produced by a fixed discrete ch.f. ψ, namely, $\mathcal{F} = \{\phi_\lambda(t) = e^{\lambda(\psi(t)-1)},\ \lambda \geq 0\}$, is such a useful DSF model. On the other hand, the complete Compound Poisson model (allowing any ψ in the exponent) seems to be too wide. Regarding the $\mathcal{CC}$ family we have the following result.

Theorem 6 *The $\mathcal{CC}$ family is not DSF. To be more specific, suppose $\{\phi_{\lambda_n}\}_{n=1}^\infty \subset \mathcal{CC}$ where $\lambda_n \geq 0$ is an arbitrary sequence, and ϕ_{λ_n} is as in Theorem* 2*. Then,* (i) *and* (ii) *below are equivalent.*

(i) *There is a point* $t_0 \in (0, 2\pi)$ *such that* $\lim_n \phi_{\lambda_n}(t_0)^n = \delta > 0$.

(ii) *It holds* $\lambda_n = \theta/\sqrt{n} + o(1/\sqrt{n})$, *where* $\theta = (-2\log\delta)^{1/2}(t_0(2\pi - t_0))^{-1/2} \geq 0$.

If (i) *or* (ii) *is satisfied then* $\phi_{\lambda_n}(t)^n \to \psi(t) := \exp(-\theta^2 t(2\pi - t)/2)$ *uniformly in* t, $0 \leq t \leq 2\pi$, *and the limiting ch.f.* ψ (extended to be 2π-periodic) *is an infinitely divisible ch.f.*

Before proving Theorem 6, we provide some remarks. The limiting ch.f. ψ is a Compound Poisson one. Indeed, the exponent can be written as $\lambda(\psi_1(t) - 1)$, where $\psi_1(t) = 1 - \theta^2 t(\pi - t/2)/\lambda$ and, e.g., $\lambda \geq \pi^2\theta^2/2$ (we shall see below that the minimum value of λ for which ψ_1 is a ch.f. is $\lambda_0 = \pi^2\theta^2/3$). Then, it follows that the even, 2π-periodic function ψ_1 is nonnegative, decreasing and convex in $[0, \pi]$, and so, by Polya's sufficiency criterion (see [8]) it is a ch.f. of an integer-valued r.v. Clearly, the parametric family produced by all possible limits from $\mathcal{CC}$, namely, $\mathcal{F} = \{\psi_\lambda(t) = e^{-\lambda t(2\pi - t)},\ \lambda \geq 0,\ t \in [0, 2\pi]\}$, forms a DSF according to Definition 2. By applying the inversion formula for ch.f.'s of integer-valued r.v.'s, namely,

$$\mathbb{P}(X = k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ikt}\phi_X(t)dt, \quad k \in \mathbb{Z},$$

it is recognized that the p.m.f.'s in $\mathcal{F}$ do not admit closed forms. Indeed, if $Y_\lambda \sim \psi_\lambda$ then the preceding formula reduces to

$$\mathbb{P}(Y_\lambda = k) = \frac{1}{\pi}\int_0^{\pi} \cos(kt)e^{-\lambda t(2\pi - t)}dt, \quad k \in \mathbb{Z},$$

and this integral cannot be computed in terms of elementary functions (unless $\lambda = 0$). Moreover, if we make use of the preceding formula with ψ_1 instead of ψ, we can easily obtain the p.m.f. of the r.v. W with ch.f. ψ_1. Setting for convenience $c = \theta^2/\lambda$ one finds $\mathbb{P}(W = 0) = 1 - c\pi^2/3$ (so that $c \leq 3/\pi^2$ and, hence, $\lambda \geq \theta^2\pi^2/3$) and $\mathbb{P}(W = k) = c/k^2$, $k \in \mathbb{Z}^*$. According to Theorem 6, these remarks provide a detailed description of the class of the limiting distributions of sums of i.i.d. r.v.'s from $\mathcal{CC}$.

The following lemma will be used in the proof of Theorem 6.

Lemma 3 (i) *Let* $\{\beta_n\}_{n=1}^{\infty} \subset (0, 1]$, *assume that* $\beta_n^n \to \beta \in (0, 1]$ *and set* $B = -\log\beta$. *Then,* $\beta_n = 1 - B/n + o(1/n)$.

(ii) *Fix* $x_0 \in [0, 1)$, *and define the function* $f(y) := \cosh(x_0 y)/\cosh(y)$, $y \geq 0$. *Suppose that* $\{\alpha_n\}_{n=1}^{\infty} \subset [0, \infty)$ *and that* $f(\alpha_n)^n \to \delta \in (0, 1]$. *Then,* $\alpha_n = \alpha/\sqrt{n} + o(1/\sqrt{n})$, *where* $\alpha = \sqrt{(-2\log\delta)/(1 - x_0^2)}$.

Proof: (i) Despite the fact that (i) is known, we provide a very quick proof here. The inequality $y \leq -\log(1 - y) \leq y/(1 - y)$ $(0 \leq y < 1)$, applied $y = 1 - \beta_n$, yields $\beta_n(-n\log\beta_n) \leq n(1 - \beta_n) \leq -n\log\beta_n$, and since the upper bound implies that

$\beta_n \to 1$, both bounds converge to B.
(ii) The sequence $n\alpha_n^2$ is bounded. Indeed, assuming the contrary, it follows that for any $M > 0$ (arbitrarily large) we can find a subsequence n_k such that $\alpha_{n_k} > M/\sqrt{n_k}$ for all k. Since it is easily checked that $f'(y) < 0$ for $y > 0$, the positive continuous function f is strictly decreasing, with $f(0) = 1$, $f(\infty) = 0$ (recall that $0 \le x_0 < 1$). Therefore, $f\big(\alpha_{n_k}\big)^{n_k} \le f\big(M/\sqrt{n_k}\big)^{n_k} \to \exp\big(-M^2(1-x_0)^2/2\big)$, as $k \to \infty$. Thus, $\liminf f(\alpha_n)^n \le \exp\big(-M^2(1-x_0)^2/2\big)$, and since $M > 0$ is arbitrary, $\liminf f(\alpha_n)^n \to 0$. This contradicts the hypothesis $f(\alpha_n)^n \to \delta > 0$, and verifies that the sequence $n\alpha_n^2$ is, indeed, bounded. Hence, $\alpha_n \to 0$. By applying a Taylor development to the function f it can be checked that for $y \ge 0$, sufficiently close to zero,

$$1 - \frac{1}{2}(1 - x_0^2)y^2 \le f(y) \le 1 - \frac{1}{2}(1 - x_0^2)y^2 + \frac{1}{24}(1 - x_0^2)(5 - x_0^2)y^4, \quad 0 \le y < \epsilon.$$

Substituting $y = \alpha_n$ (which tends to zero) we obtain the inequality

$$An(1 - f(\alpha_n)) \le n\alpha_n^2 \le An(1 - f(\alpha_n)) + B\alpha_n^2(n\alpha_n^2), \quad n \ge n_0,$$

with $A = 2/(1-x_0^2)$, $B = (5-x_0^2)/12$. Since $f(\alpha_n)^n \to \delta \in (0, 1]$ (and $0 < f(\alpha_n) \le 1$), it follows from part (i) that $n(1 - f(\alpha_n)) \to -\log \delta$, and the preceding inequality shows that $n\alpha_n^2 \to (-\log \delta)A$, completing the proof. Q.E.D.

Proof of Theorem 6: Assume first that (ii) holds, that is, $\lambda_n = \theta/\sqrt{n} + o(1/\sqrt{n})$ for some $\theta \ge 0$. It is straightforward to verify that $\phi_{\lambda_n}(t)^n$ converges pointwise to $\psi(t)$ as given, and from the fact that ψ is continuous at the origin, the convergence is uniform at compacts, and in particular, in $[0, 2\pi]$. Obviously, (i) is satisfied for (any choice of) $t_0 \in (0, 2\pi)$ with $\delta = \psi(t_0) = \exp(-\theta^2 t_0(2\pi - t_0)/2) > 0$.

Assume now that (i) holds, i.e., suppose that for a fixed $t_0 \in (0, 2\pi)$, $\phi_{\lambda_n}(t_0)^n \to \delta > 0$. Due to symmetry ($\phi_{\lambda_n}(t) = \phi_{\lambda_n}(2\pi - t)$), we can further assume that $0 < t_0 \le \pi$. Set $\alpha_n = \pi\lambda_n$, $x_0 = 1 - t_0/\pi \in [0, 1)$, and consider the function $f(y) = \cosh(x_0 y)/\cosh(y)$, $y \ge 0$, as in Lemma 3. Then, $\phi_{\lambda_n}(t_0) = f(\alpha_n)$, and by assumption, $f(\alpha_n)^n \to \delta > 0$ (certainly, $\delta \le 1$). Hence, from Lemma 3(ii) we conclude that $n\alpha_n^2 \to (-2\log\delta)/(1 - x_0^2)$, that is, $n\lambda_n^2 \to (-2\log\delta)/(t_0(2\pi - t_0))$, which verifies (ii). Q.E.D.

It is of some interest to observe that, according to Theorem 6, the limiting ch.f. exists if we can merely show the convergence $\phi_{\lambda_n}(t_0)^n \to \delta > 0$ for a single nontrivial point t_0 (i.e., $t_0 \ne 2k\pi$). Then, $\psi(t)$ is uniquely determined from the pair (t_0, δ), Also, the limiting distribution is degenerate at zero if and only if $\delta = 1$ (which is corresponds to $\theta = 0$ in Theorem 6(ii)).

Another related problem concerns the extended $\mathcal{CC}$ class, defined as the family of ch.f.'s $\mathcal{CC}^+ := \{\phi^\alpha : \ \phi \in \mathcal{CC}, \ \alpha > 0\}$. Since every $\phi \in \mathcal{CC}$ is 2π-periodic, decreases in $[0, \pi]$ and is log-convex in $[0, 2\pi]$, the same is true for all ch.f.'s in $\mathcal{CC}^+$. Hence,

$\mathcal{CC}^+$ is a family of i.d. ch.f.'s. This family is similar to the (continuous) Cauchy scale family. Cramér [3] showed that all stable centered distributions with exponent $\alpha < 2$ are not factor closed. This means that, e.g., the ch.f. of the standard Cauchy, $e^{-|t|}$, can be written as $\phi_1\phi_2$, with ϕ_i $(i = 1, 2)$ lying outside the class of Cauchy ch.f.'s. So, it is fairly expected that the same is true for $\mathcal{CC}^+$. Indeed, it can be proved that this is the case, and, as a concrete example, we provide the following 2π-periodic ϕ_i's:

$$\begin{aligned}
\phi_1(t) &= \left(\frac{\cosh(t-\pi)}{\cosh(\pi)}\right)^{1/2}\left(\frac{1+\pi^4}{1+(t-\pi)^4}\right)^{1/50}, \quad 0 \le t \le 2\pi, \\
\phi_2(t) &= \left(\frac{\cosh(t-\pi)}{\cosh(\pi)}\right)^{1/2}\left(\frac{1+(t-\pi)^4}{1+\pi^4}\right)^{1/50}, \quad 0 \le t \le 2\pi.
\end{aligned}$$

It can be checked that both functions are positive, decreasing in $[0, \pi]$, and convex (ϕ_1 is log-convex) in $[0, 2\pi]$ and hence, their 2π-periodic extensions (which are even functions) are ch.f.'s, see [13]. Obviously, these ch.f.'s lie outside $\mathcal{CC}^+$, and, trivially, their product equals to the standard discrete Cauchy ch.f. of Theorem 1.

Acknowledgements. I would like to thank an anonymous referee who motivated several of the results that are presented in Section 4, regarding infinite divisibility and stability in the discrete case.

References

[1] Arnold, B.C.; Brockett, P.L. (1992). On distributions whose component ratios are Cauchy. *Amer. Statist.*, **46**, 25–26.

[2] Cacoullos, T. (2014). Polar angle tangent vectors follow Cauchy distributions under spherical symmetry. *J. Multivariate Anal.*, **128**, 147–153.

[3] Cramér, H. (1949). On the factorization of certain probability distributions, *Aekiv f or Matematik*, **7**, 61–65.

[4] Feller, W. (1966). *An Introduction to Probability Theory and its Applications*, Vol. II. Wiley, N.Y.

[5] Geluk, J.L; de Haan, L. (2000). Stable probability distributions and their domains of attraction: a direct approach *Probab. Math. Statist.*, **20**, 169–188.

[6] Jones, M.C. (1999). Distributional relationships arising from simple trigonometric formulas. *Amer. Statist.*, **53**, 99–102.

[7] Jones, M.C. (2008). The distribution of the ratio X/Y for all centred elliptically symmetric distributions. *J. Multivariate Anal.*, **99**, 572–573.

[8] Keilson, J.; Steutel, F.W. (1972). Families of infinitely divisible distributions closed under mixing and convolution. *Ann. Math. Statist.*, **43**, 242-250.

[9] Koutrouvelis. I.A. (1982). Estimation of location and scale in Cauchy distributions using the empirical characteristic function. *Biometrika*, **69**, 205–213.

[10] Kravchuk, O.Y; Pollett, P.K. (2012). Hodges-Lehmann scale estimator for Cauchy distribution. *Commun. Statist.–Theory Meth.*, **41**, 3621–3632.

[11] Lehmann, E.L.; Gasella, G. (1998). *Theory of Point Estimation*, 2nd ed. Springer., N.Y.

[12] Lukacs, E. (1961). Recent developments in the theory of characteristic functions. *Proc. Fourth Berkeley Symposium*, **2**, 307–335.

[13] Lukacs, E. (1972). A survey of the theory of characteristic functions. *Adv. Appl. Probab.*, **4**, 1–38.

[14] Renshaw, E. (2011). *Stochastic Population Processes: Analysis, Approximations, Simulations.* Oxford University Press.

[15] Steutel, F.W.; van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *Ann. Probab.*, **7**, 893–899.

[16] van der Vaart, A.W. (1998). *Asymptotic Statistics.* Cambridge University Press, N.Y.