DOI: 10.37863/umzh.v75i3.6797

UDC 519.21

Y. Akaoka (Department of Mathematics, Faculty of Science, Shinshu University, Nagano, Japan),
K. Okamura[1] (Department of Mathematics, Faculty of Science, Shizuoka University, Japan),
Y. Otobe (Department of Mathematics, Faculty of Science, Shinshu University, Nagano, Japan)

CONFIDENCE DISC AND SQUARE FOR CAUCHY DISTRIBUTIONS[2]

ДОВІРЧИЙ ДИСК І КВАДРАТ ДЛЯ РОЗПОДІЛІВ КОШІ

We construct a confidence region of parameters for a sample of size N from the Cauchy distributed random variables. Although Cauchy distribution has two parameters, a location parameter $\mu \in \mathbb{R}$ and a scale parameter $\sigma > 0$, we infer them simultaneously by regarding them as a single complex parameter $\gamma := \mu + i\sigma$. The region should be a domain in the complex plane. We give a simple and concrete formula to give the region as a disc and as a square.

Побудовано довірчу область параметрів для вибірки розміром N для випадкових величин з розподілом Коші. Хоча розподіл Коші має два параметри, параметр розташування $\mu \in \mathbb{R}$ і параметр масштабу $\sigma > 0$, введено їх одночасно як єдиний комплексний параметр $\gamma := \mu + i\sigma$. Область, що розглядається, повинна бути областю в комплексній площині. Наведено просту і конкретну формулу, за допомогою якої цю область можна подати як диск або квадрат.

1. Introduction. The Cauchy distribution is one of typical bell-shaped and stable laws that has a longer and flatter tail, which makes mathematical handling difficult. But for that reason, typical samples from the Cauchy distribution concentrate on its centre except for a few “outliers” so that there might be cases that to model in Cauchy distribution rather than Gaussian is preferable. From this point of view, in preceding articles [2, 3], we have investigated estimators for the parameters of Cauchy distributions from observations. They are point estimations. We are, however, in the present paper, concerned with an interval estimation, that is, to construct confidence regions. Confidence regions are not only practically often used but also relate to statistical hypothesis testing (see, e.g., [15, § 7.1.2]), so it goes without saying its statistical importance.

On the other hand, to our knowledge, no concrete regions for Cauchy distributions have been known, while few studies examined them. One of them is a paper by Haas, Bain and Antle [6] which is based on [7]. They discussed mainly the maximal likelihood estimators for Cauchy distributions and proposed confidence intervals for the location parameter μ; but unfortunately, the explicit formula for the maximal likelihood estimator is known only for the case of sample sizes of 3 and 4, and furthermore, there is no algebraic closed-form formula for the case of sample size of 5 (see [5, 13]). If the explicit formula is not known, then the Newton – Raphson method has often been used. Hinkley [8, (2.12)] gives confidence regions for the pair of the location and scale by using the maximal likelihood estimator. It is related to the likelihood ratio test, however, it is complicated, and we are hard to imagine the concrete shape of the region. Vrbik also obtained confidence regions [17] based on the maximal likelihood estimator. See also [4, 9, 10] for related inference methods for Cauchy distributions.

We also would like to point out that, though it is certainly true those numerical simulations using a computer have been becoming an effective method for statistically testing or even constructing

[1] Corresponding author, e-mail: okamura.kazuki@shizuoka.ac.jp.
[2] A part of this paper consists of Y. Akaoka’s master’s thesis.

confidence intervals, it is not trivial to apply them for Cauchy distributions. To see it, we see, for example, a simple arithmetic mean of the Cauchy random variables still obeys the Cauchy distribution which is independent of the number of the sizes of samples: we emphasize here that to choose a simple and nice estimator for Cauchy distributions is nontrivial. Let us also mention that few unbiased estimators were known so far except, for instance, using MLEs [11] or trimmed average [14] which actually eliminates the outliers from the observations.

In this situation, we would like to propose in the present paper to take a geometric mean (see Remark 2.2 for the terminology). There are several reasons why it is nicer than other quantities, but since this problem belongs to the theory of point estimation, we would not like to go in this direction further here. We point out here only that the geometric mean has enough high integrability, and is unbiased (see [3] for some additional properties of the geometric mean). Also since the geometric mean is an exponential of the arithmetic mean of the logarithms, it is harmless to compute the quantity. We, however, note that the geometric mean $\prod_{j=1}^{N} X_j^{1/N}$ is a product of $1/N$th power of (probably negative) observations. So we are naturally led to treating all the quantities as complex numbers. In particular, X^p, $0 < |p| < 1$, and $\log X$ are all complex random variables. Therefore, we will gather some basic quantities, mainly expectations of such random variables, in Section 2.

It is well-known that, if X obeys a Cauchy distribution, $\mathbb{E}[|X|^p]$ exists so that $\mathbb{E}[X^p]$ also exists for $|p| < 1$. The explicit formula of $\mathbb{E}[|X|^p]$ is given by [18, (3.0.3)]. However, the derivation of [18, (3.0.3)] is very complicated, so we will give another way of calculations of them along with $\mathbb{E}[X^p]$ in Section 2. They are rather simple and direct. We also would like to expect that, through these manipulations, it becomes clear that considering in the complex plane is not a technical restriction but an essential tool for the Cauchy distributions. In particular, we hope it will be natural to regard the location parameter μ and the scale parameter $\sigma > 0$ of the Cauchy distribution as a single complex parameter $\gamma = \mu + i\sigma$ but two distinct real parameters. Note that this idea was executed by McCullagh [12]. Thus, our statistical inference for the parameters of Cauchy distributions will be actually a problem for a single complex parameter, that is, we guess naturally the location parameter and the scale parameter at once.

Then we will indeed show that a central limit theorem holds for the geometric mean (Lemma 3.3), and then modify it to fit it to derive our confidence region (Lemma 3.4) in the complex plane. It is a standard procedure. Finally, based on the central limit theorem we obtain the confidence region for the Cauchy parameters; as a disc in the complex plane (Theorem 3.1). We will conclude the paper by showing several numerical examples for our confidence disc in Section 4.

We wish to express our grateful thanks to an anonymous referee for careful reading of the manuscript and helpful comments which improve the paper. Especially, Corollaries 3.1 and 3.2 are due to his (or her) suggestions.

2. Auxiliary quantities for Cauchy distribution. In this section, we will gather some basic quantities which will be used in the following sections. In the sequel of the paper, we will denote by $C(\mu,\sigma)$ a Cauchy distribution whose location parameter is $\mu \in \mathbb{R}$ and scale parameter is $\sigma > 0$. We also denote by $X \sim C(\mu,\sigma)$ a real valued random variable $X : \Omega \to \mathbb{R}$ when the push forward measure $P_X \equiv PX^{-1}$ on $\mathbb{R}$ is $C(\mu,\sigma)$, where $(\Omega, \mathbb{F}, P)$ is a probability space that X is defined on. The expectation of X with respect to P will be denoted by $\mathbb{E}[X]$.

For noninteger real number $p \in \mathbb{R}$, we almost surely define X^p by $\exp\{p \log X\}$ as usual when $P(X = 0) = 0$. Here the logarithmic function log is defined on a Riemann surface $\{(z, \theta);\ z \in \mathbb{C}^*,\ \theta = \arg z\} \subset \mathbb{C}^* \times \mathbb{R}$, $\mathbb{C}^* := \mathbb{C} \setminus \{0\}$, that is, $\log\colon (z, \theta) \mapsto \log|z| + i\theta$ using the usual $\log\colon R_+ \to \mathbb{R}$; log is a single-valued holomorphic function. We may simply write $\log z$ as a function on $\mathbb{C}^*$ for $\log(z, \arg z)$ if there is no confusion. In this case, we regard the logarithmic function has branches, that is, $\log z = \log|z| + (\arg z + 2k\pi)i$, $0 \le \arg z < 2\pi$, $k \in \mathbb{Z}$, is a multivalued function on $\mathbb{C}^*$.

***Remark* 2.1.** Since we have an option to take a branch of log, the values of 1^p and $(-1)^p$ may differ, for instance if $p = 1/3$,

branch	k		
	$\ldots, -3, 0, 3, \ldots$	$\ldots, -2, 1, 4, \ldots$	$\ldots, -1, 2, 5, \ldots$
$1^{1/3}$	$e^{0\pi i} = 1$	$e^{2/3\pi i} = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$	$e^{4/3\pi i} = -\frac{1}{2} - \frac{\sqrt{3}}{2}i$
$(-1)^{1/3}$	$e^{1/3\pi i} = \frac{1}{2} + \frac{\sqrt{3}}{2}i$	$e^{\pi i} = -1$	$e^{5/3\pi i} = \frac{1}{2} - \frac{\sqrt{3}}{2}i$

Note that, even if $x > 0$, x^p may not be a real number but contains an imaginary part. It does not affect any result while we keep arbitrariness to choose a branch.

Let us emphasise that both $1^{1/3} = 1$ and $(-1)^{1/3} = -1$ never hold together at the same time. If we prefer $1^{1/3} = 1$ then $(-1)^{1/3}$ is no more real, and if we prefer $(-1)^{1/3} = -1$ then $1^{1/3}$ becomes complex. Otherwise, neither $1^{1/3}$ nor $(-1)^{1/3}$ is real.

Hence, it is natural to handle complex valued random variables. For such a complex valued random variable Z, the expectation is defined as usual: $\mathbb{E}[Z] := E[\mathrm{Re}(Z)] + i\mathbb{E}[\mathrm{Im}(Z)]$. We can also define a variance as a nonnegative real number: $\mathrm{Var}(Z) := \mathbb{E}[|Z - \mathbb{E}[Z]|^2] = \mathbb{E}[|Z|^2] - |\mathbb{E}[Z]|^2$. Moreover, we may define pseudovariance $\mathrm{PV}(Z) := \mathbb{E}[(Z - \mathbb{E}[Z])^2] = \mathbb{E}[Z^2] - (\mathbb{E}[Z])^2$. If a complex random variable Z satisfies (i) $\mathbb{E}[Z] = 0$, (ii) $\mathrm{Var}(Z) < \infty$, and (iii) $\mathbb{E}[Z^2] = 0$, it is called proper. A proper random variable has a vanishing pseudovariance. We set $\gamma := \mu + i\sigma$, and we also write $C(\gamma)$ in place of $C(\mu, \sigma)$.

Proposition 2.1. *Let* $f\colon \mathbb{C} \setminus \{0\} \to \mathbb{C}$ *be holomorphic and sublinear growth, that is,* $\lim_{R\to\infty} \frac{1}{R} \sup_{|z|=R} |f(z)| = 0$ *and* $\lim_{\varepsilon\to 0} \varepsilon \sup_{|z|=\varepsilon} |f(z)| = 0$. *For* $X \sim C(\gamma)$, *we see that* $\mathbb{E}[f(X)] = f(\gamma)$. *In particular, the following hold*:

(1) $\mathbb{E}[X^p] = \gamma^p$ *for* $|p| < 1$;

(2) $\mathbb{E}[(\log X)^p] = (\log \gamma)^p$ *for all* $p \ge 0$.

Proof. Note that the probability density function for Cauchy distribution is

$$p_\gamma(x) = \frac{\sigma}{\pi} \frac{1}{(x-\mu)^2 + \sigma^2} = \frac{1}{2\pi i} \left(\frac{1}{x - \gamma} - \frac{1}{x - \bar{\gamma}} \right). \tag{2.1}$$

Since $\sigma > 0$, $1/(z - \bar{\gamma})$ is holomorphic on the upper half plane. Therefore, the assertion follows immediately from Cauchy's integral formula.

***Corollary* 2.1.** *If* $X_1, X_2, \ldots, X_N \sim C(\gamma)$, $N \ge 2$, *are independent, then*:

(1) *their geometric mean is an unbiased estimator for* $\gamma = \mu + i\sigma$, *that is*, $\mathbb{E}\left[\prod_{j=1}^N X_j^{1/N}\right] = \gamma$;

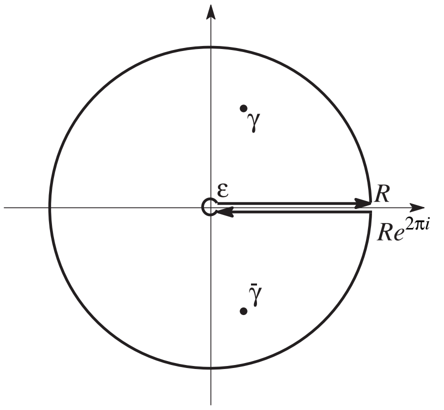


Fig. 1. Integral curve for $\int_0^\infty$.

(2) $\exp\left\{\mathbb{E}\left[\frac{1}{N}\sum_{j=1}^{N}\log X_j\right]\right\} = \gamma.$

***Remark* 2.2.** In this paper, we call $\prod_{j=1}^{N} X_j^{1/N}$ the *geometric mean*. Note that $\prod_{j=1}^{N} X_j^{1/N}$ is not equal to $\left(\prod_{j=1}^{N} X_j\right)^{1/N}$ which may be usually referred to as the geometric mean.

***Corollary* 2.2.** *If* $X \sim C(\gamma)$, *then*:

(1) *pseudovariances* $\mathrm{PV}(X^p) = 0,\ 0 < p < 1/2,$ *and* $\mathrm{PV}(\log X) = 0$;

(2) $X^p - \gamma^p,\ 0 < p < 1/2,$ *and* $\log X - \log\gamma$ *are proper complex random variables.*

Proposition 2.2. *If* $X \sim C(\gamma)$ *and* $|p| < 1,$ *then*

(1) $\mathbb{E}[X^p, X > 0] = \dfrac{\gamma^p - \bar{\gamma}^p}{1 - e^{2p\pi i}} = |\gamma|^p \dfrac{\sin p(\pi - \arg\gamma)}{\sin p\pi}$;

(2) $\mathbb{E}[|X|^p] = \dfrac{\gamma^p + (-\bar{\gamma})^p}{1 + e^{p\pi i}} = |\gamma|^p \dfrac{\cos(p(\arg\gamma - \pi/2))}{\cos(p\pi/2)}$;

(3) $\mathbb{E}[\log|X|] = \log|\gamma|$;

(4) $\mathbb{E}[(\log|X|)^2] = (\log|\gamma|)^2 + \arg\gamma(\pi - \arg\gamma)$;

(5) $\mathrm{Var}(\log|X|) = \arg\gamma(\pi - \arg\gamma)$;

(6) $\mathrm{Var}(\log X) = 2\,\mathrm{Var}(\log|X|) = 2\arg\gamma(\pi - \arg\gamma)$.

Proof. We will use the probability density of the Cauchy distribution in a form given by (2.1).

First, we will compute an integration along a curve like Fig. 1. Using Cauchy's integral formula and letting $R \to \infty$ and $\varepsilon \to 0$, we clearly get

$$\int_0^\infty x^p p_\gamma(x)\,dx + \int_\infty^0 e^{2p\pi i} x^p p_\gamma(x)\,dx = \gamma^p - \bar{\gamma}^p.$$

Note that two poles γ and $\bar{\gamma}$ stay on the same leaf but each segment in the left-hand side integrals stays on different leaves. Thus, we have (1):

$$\int_0^\infty x^p p_\gamma(x)\,dx = \frac{\gamma^p - \bar{\gamma}^p}{1 - e^{2p\pi i}} = |\gamma|^p \frac{\sin p(\arg\gamma - \pi)}{\sin p\pi}.$$

Then, to compute (2), we first note that

$$\mathbb{E}[|X|^p] = \int_{-\infty}^{\infty} |x|^p p_\gamma(x)\, dx = \int_0^{\infty} x^p (p_\gamma(x) + p_\gamma(-x))\, dx.$$

Similarly to (1), we get, by Cauchy's integral formula again,

$$(1 - e^{2p\pi i})\mathbb{E}[|X|^p] = \gamma^p - \bar{\gamma}^p + (-\bar{\gamma})^p - (-\gamma)^p.$$

Noting that with setting $\theta := \arg\gamma$,

$$\gamma^p + (-\bar{\gamma})^p = |\gamma|\big(e^{ip\theta} + e^{ip(\pi-\theta)}\big) \qquad \text{and} \qquad \bar{\gamma}^p + (-\gamma)^p = e^{ip\pi}\big(\gamma^p + (-\bar{\gamma})^p\big),$$

we finally obtain (2):

$$\mathbb{E}[|X|^p] = |\gamma| \frac{e^{ip\theta} + e^{ip(\pi-\theta)}}{1 + e^{p\pi i}}.$$

Now, we note that $g(p) := \mathbb{E}[|X|^p] = \mathbb{E}[e^{p\log|X|}]$ is the moment generating function of $\log|X|$. We also note that $z \mapsto \alpha^z$, $\alpha \in \mathbb{C}$ and $z \mapsto 1 + e^{z\pi i}$ are holomorphic. Therefore, expanding them as

$$\alpha^p = 1 + (\log\alpha)p + \frac{1}{2!}(\log\alpha)^2 p^2 + \frac{1}{3!}(\log\alpha)^3 p^3 + \ldots,$$

$$\frac{1}{1 + e^{p\pi i}} = \frac{1}{2} - \frac{1}{4}\pi i p - \frac{1}{48}\pi^3 i p^3 + \ldots,$$

allows us the following expansion:

$$g(p) = \frac{\gamma^p + (-\bar{\gamma})^p}{1 + e^{p\pi i}} = 1 + \frac{1}{2}(\log\gamma + \log(-\bar{\gamma}) - \pi i)p$$
$$+ \frac{1}{4}[((\log\gamma)^2 + (\log(-\bar{\gamma}))^2) - (\log\gamma + \log(-\bar{\gamma}))\pi i]p^2 + \ldots,$$

so that $\log\gamma + \log(-\bar{\gamma}) = \log(-|\gamma|^2) = 2\log|\gamma| + \pi i$ and $g'(0) = \log|\gamma| = \mathbb{E}[\log|X|]$ provides (3). To show (4) is routine; letting $\gamma = |\gamma|e^{i\theta}$, $\log\gamma = \log|\gamma| + i\theta$ and $\log(-\gamma) = \log|\gamma| + i(\theta - \pi)$; $\log|\gamma| \in \mathbb{R}$. Since $g''(0) = (\log|\gamma|)^2 - \theta(\theta - \pi) = \mathbb{E}[(\log|X|)^2]$, (5) also follows.

To show (6), we first note that $\mathrm{Var}(\log X) = \mathbb{E}[|\log X - \log\gamma|^2] = \mathbb{E}[|\log X|^2] - |\mathbb{E}[\log X]|^2 = \mathbb{E}[|\log X|^2] - |\log\gamma|^2$ from the definition. Now we may fix a primary branch, that is, $\log X = \log|X| + 1_{\{X<0\}}\pi i$. Hence, $|\log X|^2 = (\log|X|)^2 + 1_{\{X<0\}}\pi^2$. Thus, $\mathbb{E}[|\log X|^2] = (\log|\gamma|)^2 + \arg\gamma(\pi - \arg\gamma) + \pi\arg\gamma$. Since $|\log\gamma|^2 = (\log|\gamma|)^2 + (\arg\gamma)^2$, we have the conclusion.

Proposition 2.2 is proved.

For the geometric mean, we obtain explicit values of the covariances of the real and imaginary parts.

Proposition 2.3. *If* $n \geq 3$, *then*

$$\mathrm{Var}\Big(\mathrm{Re}\big(X_1^{1/n}\ldots X_n^{1/n}\big)\Big) = \mathrm{Var}\Big(\mathrm{Im}\big(X_1^{1/n}\ldots X_n^{1/n}\big)\Big) = \frac{r^2}{2}\left(\left(\frac{\cos((2\arg\gamma - \pi)/n)}{\cos(\pi/n)}\right)^n - 1\right)$$

and

$$\mathrm{Cov}\Big(\mathrm{Re}\big(X_1^{1/n}\ldots X_n^{1/n}\big), \mathrm{Im}\big(X_1^{1/n}\ldots X_n^{1/n}\big)\Big) = 0.$$

Proof. Let $\mu := \mathrm{Re}(\gamma)$ and $\sigma := \mathrm{Im}(\gamma)$. By the unbiasedness, we see that

$$\mathrm{Var}\Big(\mathrm{Re}\big(X_1^{1/n}\dots X_n^{1/n}\big)\Big) = E\Big[\mathrm{Re}\big(X_1^{1/n}\dots X_n^{1/n}\big)^2\Big] - \mu^2$$

and

$$\mathrm{Var}\Big(\mathrm{Im}\big(X_1^{1/n}\dots X_n^{1/n}\big)\Big) = E\Big[\mathrm{Im}\big(X_1^{1/n}\dots X_n^{1/n}\big)^2\Big] - \sigma^2.$$

Since $X_1,\dots,X_n \in \mathbb{R}$,

$$\big(X_1^{1/n}\dots X_n^{1/n}\big)^2 = \big(X_1^{1/n}X_1^{1/n}\big)\dots\big(X_n^{1/n}X_n^{1/n}\big) = X_1^{2/n}\dots X_n^{2/n}.$$

Since

$$\mathrm{Re}\big(X_1^{1/n}\dots X_n^{1/n}\big)^2 - \mathrm{Im}\big(X_1^{1/n}\dots X_n^{1/n}\big)^2 = \mathrm{Re}\big(X_1^{2/n}\dots X_n^{2/n}\big),$$

we see that

$$E\Big[\mathrm{Re}\big(X_1^{1/n}\dots X_n^{1/n}\big)^2\Big] - E\Big[\mathrm{Im}\big(X_1^{1/n}\dots X_n^{1/n}\big)^2\Big] = \mu^2 - \sigma^2.$$

Hence,

$$\mathrm{Var}\Big(\mathrm{Re}\big(X_1^{1/n}\dots X_n^{1/n}\big)\Big) = \mathrm{Var}\Big(\mathrm{Im}\big(X_1^{1/n}\dots X_n^{1/n}\big)\Big).$$

By this and Proposition 2.2 (3), we see that $\mathrm{Var}\Big(\mathrm{Re}\big(X_1^{1/n}\dots X_n^{1/n}\big)\Big)$ and $\mathrm{Var}\Big(\mathrm{Im}\big(X_1^{1/n}\dots X_n^{1/n}\big)\Big)$ are both equal to $\dfrac{r^2}{2}\left(\left(\dfrac{\cos((\pi - 2\arg\gamma)/n)}{\cos(\pi/n)}\right)^n - 1\right)$.

Since

$$2\mathrm{Re}\big(X_1^{1/n}\dots X_n^{1/n}\big)\mathrm{Im}\big(X_1^{1/n}\dots X_n^{1/n}\big) = \mathrm{Im}\big(X_1^{2/n}\dots X_n^{2/n}\big),$$

we see that

$$E\Big[\mathrm{Re}\big(X_1^{1/n}\dots X_n^{1/n}\big)\mathrm{Im}\big(X_1^{1/n}\dots X_n^{1/n}\big)\Big] = \mu\sigma.$$

Proposition 2.3 is proved.

3. Central limit theorem and asymptotic confidence disc and square for geometric mean. To state a central limit theorem for the geometric mean of the Cauchy random variables, let us begin with recalling that a complex random variable $Z_{0,1}$ is standard complex normal if $\mathrm{Re}\, Z_{0,1}$ and $\mathrm{Im}\, Z_{0,1}$ are independent, and both $\mathrm{Re}\, Z_{0,1}$ and $\mathrm{Im}\, Z_{0,1}$ obey $N(0, 1/2)$. Now let us assume that $X_1, X_2, \dots$ are independent and identically distributed random variables with $\mathbb{E}[X_1] = m$ whose real and imaginary parts have variance $v/2$ and covariance 0. Then, it holds that as $n \to \infty$, $Z_N = \Big[\sum_{j=1}^N X_j - nm\Big] \Big/ \big(\sqrt{Nv}\big)$ converge in law to the standard complex normal random variable [16, Example 2.18].

We already know from Corollary 2.2 that, if $X \sim C(\gamma)$, $\log X - \log\gamma$ is proper, that is, the real part and the imaginary part are independent and have a same distribution with mean 0. And we also have computed in Proposition 2.2 (6) that the variance of $\log X$ is $v \equiv 2\arg\gamma(\pi - \arg\gamma)$. Therefore, we get the following central limit theorem for the logarithms of the Cauchy random variables.

Lemma 3.1. *Let $X_1, X_2, \ldots \sim C(\gamma)$ be independent. Then it holds that*

$$\frac{\sqrt{N}}{\sqrt{v}}\left(\frac{1}{N}\sum_{j=1}^{N}\log X_j - \log\gamma\right) \to Z_{0,1}$$

as $N \to \infty$ in law.

To get a central limit theorem for the geometric mean around $\gamma = \mu + i\sigma$, we need the following complex version of the delta-method [15, Theorem 1.12], for which the proof is omitted since it is straightforward.

Lemma 3.2. *Let $X_1, X_2, \ldots$ and Y be complex random variables satisfying*

$$b_n(X_n - m) \to Y$$

as $n \to \infty$ in law, where $m \in \mathbb{C}$ and $\{b_n\}$ is a sequence of positive numbers with $b_n \to \infty$. Let $f : \mathbb{C} \to \mathbb{C}$ be a holomorphic function on a domain containing $m \in \mathbb{C}$, then

$$b_n\big[f(X_n) - f(m)\big] \to f'(m)Y$$

as $n \to \infty$ in law.

Applying the complex delta-method for a complex function $f(z) := e^z$, we get the following version of the central limit theorem.

Lemma 3.3. *Let $X_1, X_2, \ldots \sim C(\gamma)$ be independent. Then it holds that*

$$\frac{\sqrt{N}}{\sqrt{v}\gamma}\left(\prod_{j=1}^{N} X_j^{1/N} - \gamma\right) \to Z_{0,1} \tag{3.1}$$

as $N \to \infty$ in law.

Unfortunately Lemma 3.3 above could not be applied to get a confidence disc for Cauchy distributed random variables. We will derive the following variant of the central limit theorem.

Lemma 3.4. *Let $X_1, X_2, \ldots \sim C(\gamma)$ be independent. Setting*

$$P_N := \prod_{j=1}^{N} X_j^{1/N}$$

and

$$V_N := \frac{1}{N-1}\sum_{j=1}^{N}|\log X_j|^2 - \left|\frac{1}{N}\sum_{j=1}^{N}\log X_j\right|^2, \tag{3.2}$$

which is an unbiased and consistent estimator for $v := \mathrm{Var}(\log X_1) = 2\arg\gamma(\pi - \arg\gamma)$, we have

$$\frac{\sqrt{N}}{\sqrt{V_N}P_N}(P_N - \gamma) \to Z_{0,1}$$

as $N \to \infty$ in law.

Proof. Since V_N converges to $2\arg\gamma(\pi - \arg\gamma)$ and P_N converges to γ as $N \to \infty$ in probability, Lemma 3.3 implies the conclusion.

Now we can state a formula for our asymptotic confidence disc.

Theorem 3.1. *Let $X_1, X_2, \ldots, X_N$ be independent Cauchy random variables with location $\mu \in \mathbb{R}$ and scale $\sigma > 0$. Set P_N and V_N as in Lemma* 3.4. *Then we have*

$$\lim_{N\to\infty} P\left(\gamma \in B\left(P_N, \frac{\sqrt{V_N}}{\sqrt{N}} R_\alpha |P_N|\right)\right) = 1 - \alpha \tag{3.3}$$

for $0 < \alpha \leq 1$, where $R_\alpha = \sqrt{-\log\alpha}$ and $B(p, r)$ denotes a disc in $\mathbb{C}$ with centre p and radius r.

Proof. We first note that $P\big(|Z_{0,1}| \leq \sqrt{-\log\alpha}\,\big) = 1 - \alpha$. Combining it with Lemma 3.4 asserts that

$$\lim_{N\to\infty} P\left(\left|\frac{\sqrt{N}}{\sqrt{V_N} P_N}(P_N - \gamma)\right| \leq R_\alpha\right) = 1 - \alpha,$$

which immediately leads to the conclusion.

Theorem 3.1 is proved.

A similar argument enables us to derive an asymptotic confidence square that is slightly smaller than the disc.

***Corollary* 3.1.** *We denote by $Q(p, r)$ a square in the complex plane whose centre is r and the length of one side is $2r$, namely, $Q(p, r) = [p - r, p + r] \times i[p - r, p + r]$. Let ρ_α be the upper α-quantile of standard real normal distribution, that is, $\displaystyle\int_{-\infty}^{\rho_\alpha} \frac{1}{\sqrt{2\pi}} \exp\{-x^2/2\}\, dx = 1 - \alpha$. Under same settings of Theorem* 3.1 *and $\beta = \dfrac{1}{2}\big(1 - \sqrt{1-\alpha}\,\big)$, we have*

$$\lim_{N\to\infty} P\left(\gamma \in Q\left(P_N, \frac{\sqrt{V_N}}{\sqrt{2N}} |P_N| \rho_\beta\right)\right) = 1 - \alpha.$$

Proof. Recall that $Z_{0,1}$ is rotation invariant in law, that is, for any complex number w with $|w| = 1$, $Z_{0,1}$ and $wZ_{0,1}$ share a same distribution. This fact guarantees that we may take an absolute value in the coefficient of (3.1), that is, we have

$$\frac{\sqrt{N}}{\sqrt{v}|\gamma|}(P_N - \gamma) \to Z_{0,1}$$

as $N \to \infty$ in law. Therefore, a same argument with Lemma 3.4 leads

$$\lim_{N\to\infty} P\left(\gamma \in Q\left(P_N, \frac{\sqrt{V_N}|P_N|}{\sqrt{2N}} \rho_\beta\right)\right) = P\Big(Z_{0,1} \in Q\Big(0, \rho_\beta/\sqrt{2}\Big)\Big).$$

Since $\mathrm{Re}(Z_{0,1})$ and $\mathrm{Im}(Z_{0,1})$ are independent, and obeys $N(0, 1/2)$, the right-hand side equals to $(1 - 2\beta)^2 = 1 - \alpha$.

Corollary 3.1 is proved.

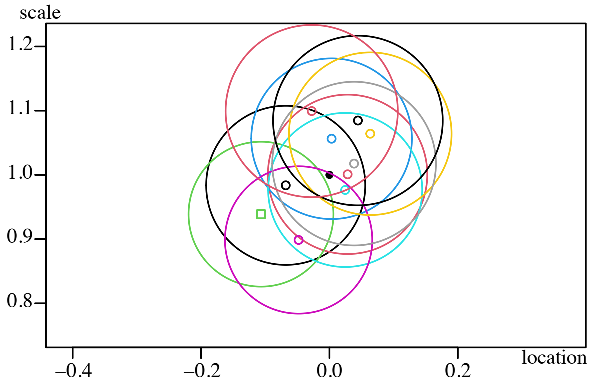
We can apply the idea of Corollary 3.1 to get confidence intervals for μ and σ, respectively.

***Corollary* 3.2.** *Under the same settings of Corollary* 3.1, *we have*

$$\lim_{N\to\infty} P\left(\mathrm{Re}(P_N) - \frac{\sqrt{V_N}|P_N|}{\sqrt{2N}} \rho_{\alpha/2} \leq \mu \leq \mathrm{Re}(P_N) + \frac{\sqrt{V_N}|P_N|}{\sqrt{2N}} \rho_{\alpha/2}\right) = 1 - \alpha,$$

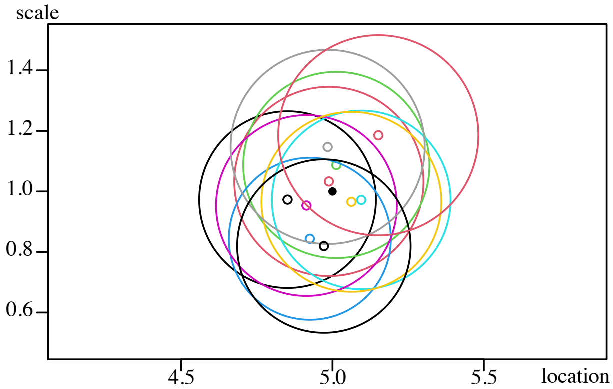
$$\lim_{N\to\infty} P\left(\mathrm{Im}(P_N) - \frac{\sqrt{V_N}|P_N|}{\sqrt{2N}} \rho_{\alpha/2} \leq \sigma \leq \mathrm{Im}(P_N) + \frac{\sqrt{V_N}|P_N|}{\sqrt{2N}} \rho_{\alpha/2}\right) = 1 - \alpha.$$

Proof. Since we have $P\big(Z_{0,1} \in \big[-\rho_{\alpha/2}/\sqrt{2}, \rho_{\alpha/2}/\sqrt{2}\,\big] \times i\mathbb{R}\big) = 1 - \alpha$, the results follow from the same computation with Corollary 3.1.



geom. mean	log-variance	radius
$-0.067+0.968i$	5.271	0.124
$0.028+0.985i$	5.130	0.124
$-0.105+0.925i$	4.762	0.113
$0.003+1.042i$	4.683	0.125
$0.024+0.962i$	5.037	0.120
$-0.047+0.884i$	5.428	0.115
$0.063+1.049i$	4.710	0.127
$0.038+1.002i$	5.233	0.127
$0.044+1.069i$	4.947	0.132
$-0.027+1.083i$	4.966	0.134
$0.124+1.091i$	5.087	0.138

Fig. 2. $\mu=0,\ \sigma=1,\ N=1000,\ \alpha=0.05,$ 10 trials.



geom. mean	log-variance	radius
$4.835+0.970i$	1.162	0.292
$4.969+1.029i$	1.257	0.313
$4.995+1.084i$	1.200	0.308
$4.911+0.841i$	0.958	0.268
$5.078+0.969i$	1.080	0.295
$4.896+0.950i$	1.188	0.299
$5.046+0.962i$	1.110	0.297
$4.964+1.142i$	1.312	0.321
$4.955+0.817i$	1.079	0.287
$5.131+1.181i$	1.306	0.331

Fig. 3. $\mu=5,\ \sigma=1,\ N=1000,\ \alpha=0.05,$ 10 trials.

4. Numerical examples. In this section we will show several numerical examples of Theorem 3.1.

4.1. $\boldsymbol{\mu=0,\ \sigma=1,\ N=1000,\ \alpha=0.05,\ 10}$ ***trials.*** Figure 2 shows 10 confidence discs; each of the samples has of size 1,000. Namely, we generated samples of $X_1^1, X_2^1, \ldots, X_{1000}^1, X_1^2, X_2^2, \ldots, X_{1000}^2, \ldots, X_1^{10}, X_2^{10}, \ldots, X_{1000}^{10} \sim C(0,1)$. Then we computed the geometric mean $P_{1000}^i := \prod_{j=1}^{1000} (X_j^i)^{1/1000}$ by the formula $P_{1000}^i = \exp\left\{\dfrac{1}{1000}\sum_{j=1}^{1000} \log X_j^i\right\}$, and sample variance V_{1000}^i by (3.2). We drew 10 discs $B^1, \ldots, B^{10}$ using these quantities and (3.3). A symbol $\bullet$ indicates the true value $(\mu,\sigma)=(0,1)$. Symbols $\circ$ point the centre of the disc containing the true value, while $\square$ is the point that fails.

4.2. $\boldsymbol{\mu=5,\ \sigma=1,\ N=1000,\ \alpha=0.05,\ 10}$ ***trials.*** Figure 3 also shows the same size of trials and samples, but $\mu=5$ and $\sigma=1$ for larger true values; we fix $\sigma=1$ to compare with Fig. 2 easily. Note that, since $\arg(5+i)=\tan^{-1}\dfrac{1}{5}\sim 0.197,\ \mathrm{Var}(\log X)\sim 1.162$. Note that $\mathrm{Var}(\log X)$ is small if $|\mu|$ is large but the radius of the confidence discs increases as $|\mu|$ increases.

4.3. $\boldsymbol{\mu=0,\ \sigma=1,\ N=30,\ \alpha=0.05,\ 1000}$ ***trials.*** Since the formula (3.3) to compute our confidence disc is based on a central limit theorem (Theorem 3.1). Therefore, we may fail to guess the parameters if N is small. Figure 4 (omitting to draw circles) shows the case $N=30$ with $(\mu,\sigma)=(0,1)$. In this example, computed $\alpha=0.05$, only 925 trials of $1,000$ contain the true value.

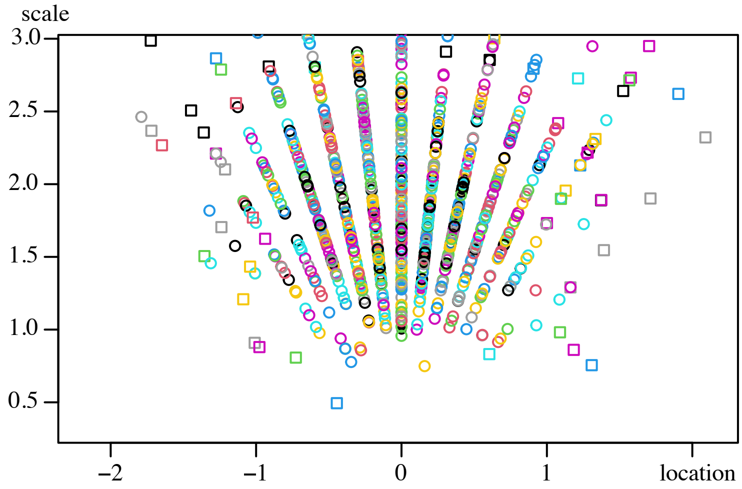
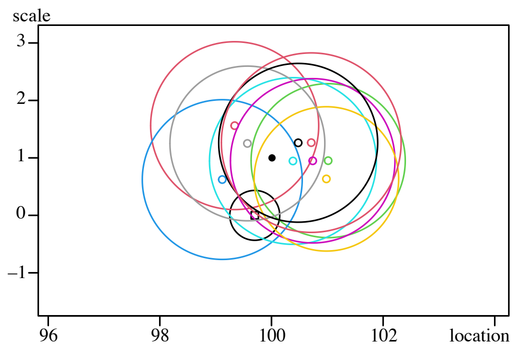


Fig. 4. $\mu = 0,\ \sigma = 1,\ N = 30,\ \alpha = 0.05,$ 1000 trials.

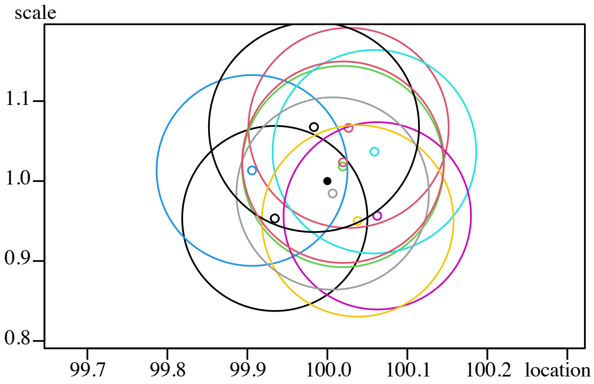


geom. mean	median	log-variance	radius
$99.698 + 0.000i$	99.934	0.006	0.434
$99.330 + 1.560i$	100.026	0.072	1.460
$100.958 + 0.952i$	100.019	0.059	1.340
$99.116 + 0.623i$	99.906	0.066	1.390
$100.343 + 0.946i$	100.059	0.070	1.450
$100.687 + 0.949i$	100.062	0.067	1.430
$100.930 + 0.634i$	100.038	0.052	1.256
$99.554 + 1.251i$	100.007	0.061	1.346
$100.436 + 1.262i$	99.983	0.063	1.381

Fig. 5. $\mu = 100,\ \sigma = 1,\ N = 1000,\ \alpha = 0.05,$ 10 trials.

4.4. Large μ and small σ. When $|\mu|$ is much larger than σ, all data may have the same sign. In such a case, there is a possibility that the geometric mean has the vanishing imaginary part so that we shall guess σ as zero. Figure 5 shows such a case. One practical way to avoid the situation is to subtract a constant from each datum. It is clear that, putting $Y_i := X_i - c$, an estimate using Y_i's plus c is equal to that of X_i's. We also emphasise that subtracting a constant may decrease the absolute value of the geometric mean so that our estimated discs may be smaller, i.e., we may improve the estimation. However, it is impossible to get a suitable constant to make the data consist of both positive and negative values. Therefore, practically speaking, it may be a candidate to subtract the median $\mathrm{med}(X_1, \ldots, X_N)$ from all X_i's. Then the data will consist of same numbers of positive and negative ones. Figure 6 shows the effect of such manipulations. Let us, however, mention that $X_i - \mathrm{med}(X_1, \ldots, X_N)$ are neither Cauchy nor independent though the central limit theorem may still hold. It seems difficult for us to get a suitable random variable to subtract from each datum. Figure 6 is shown just for convenience; we will not provide any detailed mathematically reasonable explanation for this sort of manipulations.

4.5. Outlier and Gaussian estimation. Here we deal with one-dimensional location-scale families and consider the estimation of the location when the scale is not known. It is common to assume the data are Gaussian distributed if the shape of the frequency forms a bell shape. But when the data



geom. mean	log-variance	radius
$99.934 + 0.939i$	4.922	0.116
$100.026 + 1.052i$	4.597	0.125
$100.019 + 1.003i$	5.104	0.126
$99.906 + 0.999i$	4.628	0.119
$100.059 + 1.021i$	5.030	0.127
$100.062 + 0.942i$	5.007	0.117
$100.038 + 0.935i$	5.305	0.120
$100.007 + 0.970i$	4.982	0.120
$99.983 + 1.051i$	5.050	0.131
$100.020 + 1.008i$	5.057 0.126	

Fig. 6. Same data with Fig. 5 with subtracting the median.

Table 1. Outlier's effect for t-distribution

sample	AM	radius	interval	AM	radius	interval
1	-0.049	0.200	$[-0.249, 0.150]$	-0.002	0.223	$[-0.225, 0.221]$
2	0.061	0.165	$[-0.103, 0.226]$	0.116	0.191	$[-0.075, 0.308]$
3	0.075	0.197	$[-0.122, 0.271]$	0.127	0.219	$[-0.093, 0.346]$
4	-0.076	0.185	$[-0.261, 0.109]$	-0.010	0.208	$[-0.219, 0.198]$
5	-0.001	0.219	$[-0.221, 0.218]$	0.029	0.237	$[-0.209, 0.266]$
6	0.042	0.222	$[-0.180, 0.264]$	0.105	0.241	$[-0.136, 0.346]$
7	-0.124	0.199	$[-0.323, 0.074]$	-0.062	0.222	$[-0.285, 0.160]$
8	0.094	0.228	$[-0.134, 0.322]$	0.147	0.248	$[-0.101, 0.395]$
9	-0.091	0.181	$[-0.272, 0.091]$	-0.032	0.207	$[-0.238, 0.175]$
10	0.180	0.203	$[-0.024, 0.383]$	0.243	0.223	$[0.020, 0.465]$

have some outliers, using t-distributions may give an inappropriate confidence interval. But it should be noted that the Cauchy distribution often gives such "outliers" due to its heavy tail of the density so that we may expect that an inference that is applicable to the Cauchy distribution is more robust.

Here we show a numerical example, which consists of 10 samples; each of them contains 100 values from a standard normal distribution. Then we replace each 100th datum with "5" to mimic an outlier. Tables 1 and 2 exhibit these differences. The left parts of those tables are the original data, and the right parts are the modified data containing the outliers.

Table 1 shows how such contamination varies the confidence intervals when we use the t-distribution with the degree of freedom 99, while Table 2 shows those by using Corollary 3.2. Here the AM and GM denote the arithmetic mean and the real part of the geometric mean, respectively.

It can be seen that the confidence intervals computed using Corollary 3.2 are not so affected by outliers.

4.6. Remark about subtraction. As we see in Subsection 4.4, the manipulation of subtracting the median works well when the location is far from the origin. However, subtraction entails a delicate problem.

Table 2. Outlier's effect for Corollary 3.2

sample	GM	radius	interval	GM	radius	interval
1	0.017	0.143	$[-0.126, 0.160]$	0.017	0.148	$[-0.130, 0.165]$
2	0.025	0.110	$[-0.085, 0.135]$	0.038	0.113	$[-0.075, 0.152]$
3	0.032	0.136	$[-0.104, 0.168]$	0.050	0.141	$[-0.091, 0.191]$
4	-0.131	0.138	$[-0.269, 0.006]$	-0.117	0.140	$[-0.257, 0.024]$
5	-0.061	0.163	$[-0.224, 0.102]$	-0.061	0.165	$[-0.227, 0.104]$
6	0.017	0.150	$[-0.133, 0.167]$	0.034	0.153	$[-0.119, 0.187]$
7	0.016	0.137	$[-0.121, 0.153]$	0.032	0.140	$[-0.108, 0.172]$
8	0.099	0.162	$[-0.063, 0.261]$	0.122	0.167	$[-0.046, 0.289]$
9	-0.155	0.123	$[-0.278, -0.032]$	-0.144	0.126	$[-0.270, -0.018]$
10	0.206	0.162	$[0.044, 0.368]$	0.227	0.165	$[0.063, 0.392]$

Let μ and σ be the real and imaginary parts of γ. Consider

$$Z_N := Y_N + \prod_{i=1}^{N} (X_i - Y_N)^{1/N}.$$

There is a sequence of random variable Y_N such that $Y_N \to \mu$ a.s. and Z_N in the above does *not* converge to γ in probability as $N \to \infty$.

This occurs if we let $Y_N := X^{(\lfloor N/2 \rfloor + 1)}$, where $\{X_1, \ldots, X_N\} = \left\{X^{(1)} \leq \ldots \leq X^{(N)}\right\}$ and $\lfloor r \rfloor$ denotes the integer part of $r \in \mathbb{R}$. In that case we see that $Z_N = Y_N$ and, in particular, $\operatorname{Im}(Z_N) = 0$. Y_N is the median of $\{X_1, \ldots, X_N\}$ if N is *odd*. In the above subsection, we deal with the case that N is *even*.

Let I be a closed interval containing the location μ in its interior. Then

$$\sup_{\theta \in I} \left| \theta + \prod_{j=1}^{N} (X_j - \theta)^{1/N} - \gamma \right| \geq \sigma.$$

Since $P(Y_N \in I) \to 1,\ N \to \infty$, we see that

$$P\left(\sup_{\theta \in I} \left| \theta + \prod_{j=1}^{N} (X_j - \theta)^{1/N} - \gamma \right| \geq \sigma \right) \to 1, \quad N \to \infty.$$

We also have the following proposition.

Proposition 4.1. *Let I be a closed interval. Then*

$$\left(\sup_{\theta \in I} \left| \theta + \prod_{j=1}^{N} (X_j - \theta)^{1/N} - \gamma \right| \right)_N$$

is uniformly integrable.

Proof. Let $M := \max\{|\theta - \mu| : \theta \in I\}$. Then, for every $\theta \in I$,

$$\left|\theta + \prod_{j=1}^{N}(X_j - \theta)^{1/N} - \gamma\right| \le |\gamma| + M + \prod_{j=1}^{N}\big(|X_j - \mu| + M\big)^{1/N}.$$

Hence,

$$E\left[\sup_{\theta \in I}\left|\theta + \prod_{j=1}^{N}(X_j - \theta)^{1/N}\right|\right] \le M + E\Big[\big(|X_1 - \mu| + M\big)^{1/N}\Big]^N.$$

Since

$$\lim_{N\to\infty} E\Big[\big(|X_1 - \mu| + M\big)^{1/N}\Big]^N = \exp\big(E\big[\log(|X_1 - \mu| + M)\big]\big),$$

we see that

$$\sup_N E\Big[\big(|X_1 - \mu| + M\big)^{1/N}\Big]^N < +\infty.$$

Proposition 4.1 is proved.

On the other hand, if we let $Y_N := \big(X^{(\lfloor N/2\rfloor+1)} + X^{(\lfloor N/2\rfloor)}\big)/2$, then, for even number N, Y_N is the median of $\{X_1, \ldots, X_N\}$, and by numerical computations, Z_N approximates γ well, and we conjecture that $\lim_{N\to\infty} Z_N = \gamma$, a.s. However, we do not have any proofs of it. We do not see why the case that $Y_N = X^{(\lfloor N/2\rfloor+1)}$ is bad, but, on the other hand, the case that $Y_N = (X^{(\lfloor N/2\rfloor+1)} + X^{(\lfloor N/2\rfloor)})/2$ is good, although $X^{(\lfloor N/2\rfloor+1)}$ and $(X^{(\lfloor N/2\rfloor+1)} + X^{(\lfloor N/2\rfloor)})/2$ are close to each other with high probability if N is large.

One way to resolve this delicate issue is to consider the estimator

$$\theta - \epsilon i + \prod_{j=1}^{N}(X_j - \theta + \epsilon i)^{1/N}$$

for some $\epsilon > 0$. This is also an unbiased estimator of γ. This is easier to handle due to the following proposition.

Proposition 4.2. *Let I be a closed interval. Then*

$$\sup_{\theta \in I}\left|\theta - \epsilon i + \prod_{j=1}^{N}(X_j - \theta + \epsilon i)^{1/N} - \gamma\right|$$

converges to 0 *a.s. and in* L^1 *as* $N \to \infty$.

Proof. The a.s. convergence follows from the uniform law of large numbers for $\big(\log(X_N - \theta + \epsilon i) - \log(\mu - \theta + (\sigma + \epsilon)i)\big)_N$.

In the same manner as in the proof of Proposition 4.1, we can show that

$$\left(\sup_{\theta \in I}\left|\theta - \epsilon i + \prod_{j=1}^{N}(X_j - \theta + \epsilon i)^{1/N} - \gamma\right|\right)_N$$

is uniformly integrable. Thus, we see the L^1 convergence.

Proposition 4.2 is proved.

There is also a problem in this resolution because there are many possibilities of $\epsilon > 0$, and we do not see which $\epsilon > 0$ should be taken. We do not delve into this issue in this paper.

The authors wish to give their gratitude to Dr. Thomas Simon for references. The second and third authors were supported by JSPS KAKENHI 19K14549 and 16K05196, respectively.

References

1. Y. Akaoka, *Parameter estimation using complex valued moments for Cauchy distributions*, Master's thesis, Shinshu Univ. (2020).
2. Y. Akaoka, K. Okamura, Y. Otobe, *Bahadur efficiency of the maximum likelihood estimator and one-step estimator for quasi-arithmetic means of the Cauchy distribution*, Ann. Inst. Statist. Math. (2022).
3. Y. Akaoka, K. Okamura, Y. Otobe, *Limit theorems for quasiarithmetic means of random variables with applications to point estimations for the Cauchy distribution*, Braz. J. Probab. and Stat. (2022).
4. G. V. C. Freue, *The Pitman estimator of the Cauchy location parameter*, J. Statist. Plann. and Inference, **137**, № 6, 1900 – 1913 (2007).
5. T. S. Ferguson, *Maximum likelihood estimates of the parameters of the Cauchy distribution for samples of size 3 and 4*, J. Amer. Statist. Assoc., **73**, № 361, 211 – 213 (1978).
6. G. Haas, L. Bain, C. Antle, *Inferences for the Cauchy distribution based on maximum likelihood estimators*, Biometrika, **57**, № 2, 403 – 408 (1970).
7. G. N. Haas, *Statistical inferences for the Cauchy distribution based on maximum likelihood estimators*, Doct. diss., Univ. Missouri-Rolla (1969).
8. D. V. Hinkley, *Likelihood inference about location and scale parameters*, Biometrika, **65**, № 2, 253 – 261 (1978).
9. O. Y. Kravchuk, P. K. Pollett, *Hodges – Lehmann scale estimator for Cauchy distribution*, Commun. Statist. Theory and Meth., **41**, № 20, 3621 – 3632 (2012).
10. J. F. Lawless, *Conditional confidence interval procedures for the location and scale parameters of the Cauchy and logistic distributions*, Biometrika, **59**, № 2, 377 – 386 (1972).
11. P. McCullagh, *On the distribution of the Cauchy maximum-likelihood estimator*, Proc. Roy. Soc. London. Ser. A, **440**, 475 – 479 (1993).
12. P. McCullagh, *Möbius transformation and Cauchy parameter estimation*, Ann. Statist., **24**, № 2, 787 – 808 (1996).
13. K. Okamura, Y. Otobe, *Characterizations of the maximum likelihood estimator of the Cauchy distribution*, Lobachevskii J. Math. **43**, № 9, 2576 – 2590 (2022).
14. T. J. Rothenberg, F. M. Fisher, C. B. Tilanus, *A note on estimation from a Cauchy sample*, J. Amer. Statist. Assoc., **59**, № 306, 460 – 463 (1964).
15. Jun Shao, *Mathematical statistics*, 2nd ed., Springer New York (2003).
16. A. W. van der Vaart, *Asymptotic statistics*, Cambridge Ser. in Statistical and Probabilistic Mathematics, Cambridge Univ. Press (1998).
17. J. Vrbik, *Accurate confidence regions based on MLEs*, Adv. and Appl. Stat., **32**, № 1, 33 – 56 (2013).
18. V. M. Zolotarev, *One-dimensional stable distributions*, Amer. Math. Soc. (1986).

Received 29.06.21