Mathematics

https://www.aimspress.com/journal/Math

AIMS Mathematics, 10(12): 30942–30967.
DOI: 10.3934/math.20251358
Received: 08 October 2025
Revised: 14 December 2025
Accepted: 22 December 2025
Published: 31 December 2025

Research article

On moment convergence of sample quantiles with application to parameter estimations for Cauchy distribution

Yuankang Wang*

School of Mathematics and Statistics, Hunan Normal University, Changsha, Hunan, China

* **Correspondence:** Email: 202520102296@hunnu.edu.cn.

Abstract: Under a general condition on continuous probability distributions of some populations, this study establishes that the distributional convergence of sample quantiles implies moment convergence. As an application, we propose a quick and robust consistent estimator for the shape parameter of the Cauchy distribution under large-sample conditions. This estimator achieves over 99% efficiency relative to the maximum likelihood estimator.

Keywords: asymptotically consistent estimator; Cauchy distribution; convergence in distribution; convergence of moments
Mathematics Subject Classification: 62F10

1. Introduction

The convergence in distribution of a sequence of random variables does not necessarily imply the convergence of moments. For example, consider a sequence of independent and identically distributed random variables $\{X_n, n \geq 1\}$ with common mean μ and variance σ^2. The Central Limit Theorem (CLT) states that the standardized partial sums

$$\left\{ \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma}, n \geq 1 \right\}$$

converge in distribution to the standard normal distribution $\mathrm{N}(0, 1)$ as n tends to infinity. However, the third moment sequence $\left\{ E\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \right)^3, n \geq 1 \right\}$ does not necessarily converge to 0, the third moment of $\mathrm{N}(0, 1)$. In fact, if the third moment of X_i does not exist, then for any given positive integer n, the standardized partial sum $\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma}$ lacks moments of order $m \geq 3$. On the other hand, all moments of order $m \geq 3$ exist for $\mathrm{N}(0, 1)$. Therefore, the convergence in distribution of $\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma}$ to $\mathrm{N}(0, 1)$ does not guarantee the corresponding moment convergence. The convergence of multidimensional random vectors sequence also exhibits similar behavior.

In the application area of parameter estimation, we mainly focus on two types of CLTs: the 'convergence in distribution version' and the 'convergence in moments version'. As direct applications, these two versions serve 'parameter interval estimation' and 'point estimation analysis', respectively.

The property of moment convergence in statistical estimators is crucial for parameter point estimation analysis. When the low-order moments of a population do not exist, we refer to it as an 'ultra-heavy-tailed distribution'. Under such circumstances, traditional moment-based estimation methods generally fail for parameter estimation of the distribution. Furthermore, in the context of ultra-heavy-tailed populations, small-sample methods often yield too few data points to satisfy convergence requirements, making it difficult to obtain accurate and reliable parameter estimates for ultra-heavy-tailed distributions from small samples.

However, in large-sample scenarios, although the 'maximum likelihood estimation' (MLE) is theoretically optimal with desirable asymptotic properties (such as consistency and efficiency), MLE for ultra-heavy-tailed populations frequently lacks closed-form solutions. Moreover, numerical computation of MLE is often computationally intensive and may fail to converge due to the irregularity of the likelihood function. Thus, many researchers have focused on developing alternative methods to MLE to address the limitations of traditional approaches. For instance, Chernozhukov and Hansen [1] incorporated quantile conditions into the 'Generalized Method of Moments framework (Quantile-based GMM)'. Since quantiles are less sensitive to extreme values and exhibit robustness in ultra-heavy-tailed distributions, 'sample quantile methods' have emerged as an important alternative for parameter estimation. It is noteworthy that the asymptotic theory of sample quantiles (order statistics) has also been extensively studied for mixture distributions, which frequently arise in practical scenarios such as modeling wind speed or river discharge [2].

This paper analyzes the moment convergence properties of sample quantiles and references the following three theorems from the literature:

Theorem 1.1. *(see [3]) For a population X distributed according to a continuous probability density function (pdf) $f(x)$ with respect to Lebesgue measure, let p and r be two numbers satisfying* $0 < p \leq r < 1$ *and x_p and x_r be respectively the p-quantile and r-quantile of X satisfying $f(x_p)f(x_r) > 0$. Let $(X_1, \ldots, X_n)$ be a random sample derived from X and $X_{i:n}$ be the i-th sample order statistic. If there are constants $\omega > 0$ and $v \in (-\infty, \infty)$ such that the cumulative distribution function (cdf) $F(x)$ of $\omega X + v$ has an inverse function $G(x)$ which possesses a continuous third-order derivative function $G'''(x)$ in the interval* $(0, 1)$ *satisfying*

$$|G'''(x)| \leq Kx^{-A}(1-x)^{-A},$$

for some given constants $K > 0, A \geq 0$ and all $x \in (0, 1)$, then:
(1) we have, as $n \to \infty$,

$$\lim_{n\to\infty} E\left(\frac{f\left(x_p\right)\left(X_{i:n}-x_p\right)}{\sqrt{p(1-p)/n}}\right)^2 = \lim_{n\to\infty} E\left(\frac{f\left(x_r\right)\left(X_{j:n}-x_r\right)}{\sqrt{r(1-r)/n}}\right)^2 = 1$$

provided $i/n = p + o\left(n^{-1/2}\right)$ and $j/n = r + o\left(n^{-1/2}\right)$;
(2) the correlation coefficient $\operatorname{corr}\left(X_{i:n}, X_{j:n}\right)$ *between $X_{i:n}$ and $X_{j:n}$ satisfies*

$$\lim_{n\to\infty} \operatorname{corr}\left(X_{i:n}, X_{j:n}\right) = \sqrt{\frac{p(1-r)}{r(1-p)}},$$

provided $i/n = p + o(1)$ and $j/n = r + o(1)$ as $n \to \infty$.

Theorem 1.2. *(see [4]) Let the pdf $f(x)$ (with respect to Lebesgue measure) of a population X have a bounded derivative of order m. Given a sample $(X_1, \dots, X_n)$ and $p \in (0,1)$, let x_p be the p-quantile of X satisfying $f(x_p) > 0$. Assume the following two conditions hold:*

(i) The cdf $F(x)$ of X has an inverse function $G(x)$, and there exists constants $B > 0$ and $Q \geq 0$ such that

$$|G(x)| \leq B \cdot x^{-Q}(1-x)^{-Q}$$

holds for all $x \in (0,1)$.

(ii) $i/n = p + O(n^{-1})$ and $a_{i:n} = x_p + O(n^{-1})$.

Then, as $n \to +\infty$, we must have

$$E\left(\frac{f(x_p)(X_{i:n} - a_{i:n})}{\sqrt{p(1-p)/n}}\right)^m = EZ^m + O(n^{-1/2}), \tag{1.1}$$

where Z is a standard normal random variable.

Theorem 1.3. *(see [5] and [6]) Assume $(X_1, ..., X_n)$ to be a sample from a population with a cdf $F(x)$ having positive derivatives $f(x_{p_i})$ at x_{p_i}s where $0 < p_1 < ... < p_m < 1$. Then, as $n \to \infty$, the joint distribution of*

$$\sqrt{n}[(Q_{n,p_1}, Q_{n,p_2}, ..., Q_{n,p_m}) - (x_{p_1}, x_{p_2}, ..., x_{p_m})]$$

converges in distribution to the multi-normal distribution $N_m(0, D)$ where D is the $m \times m$ symmetric matrix with (i, j)th element being $p_i(1-p_j)/(f(x_{p_i})f(x_{p_j})), 1 \leq i \leq j \leq m$. Here $Q_{n,p}$ denotes the p-quantile of the sample $(X_1, ..., X_n)$.

Note: As a direct conclusion of the above three Theorems 1.1–1.3, we see that under conditions of Theorem 1.1, after respect standardizations, two distinct sample quantiles will jointly converge in distribution to a bivariate normal distribution. This convergence also holds in terms of some (suitable order of) moments. Therefore, for any linear function of these two sample quantiles, the moments will be equivalent to the corresponding moments of the associated normal distribution.

However, some scholars (see [7]) have implicitly assumed that when the population expectation does not exist, the expectation of sample order statistics also does not exist. Yet, the conclusion of Theorem 1.2 demonstrates that the expectation of sample p-quantiles generally exists given sufficiently large sample size, even for ultra-heavy-tailed distributions like the Cauchy distribution. This provides new perspectives and approaches for parameter estimation under ultra-heavy-tailed distributions. Particularly, when traditional methods like moment estimation and maximum likelihood estimation face challenges, using sample quantiles for parameter estimation becomes a viable alternative.

Interestingly, some scholars have gone to the other extreme – they directly apply distributional convergence of random sequences to moment convergence. For instance, in large samples, they approximate the moments of standardized sample quantiles directly with the corresponding moments of the standard normal distribution without providing sufficient conditions or theoretical justification. As was pointed out in [3], this stems from some scholars' non-rigorous use of Taylor expansion formulas, neglecting that even when a function has derivatives of all orders, its Taylor series expansion may not converge to the original function itself. They loosely approximate functions with partial sums of Taylor series without rigorous analysis of whether the error terms are infinitesimal (see [8] for an example).

To address these issues, reference [3] presents rigorous arguments showing that as long as the cdf satisfies what we will define as the universal condition (which almost all continuous populations satisfy), the standardized sample quantiles will generally converge in moments to the corresponding moments of the standard normal distribution. As an application, [3] developed a quick unbiased estimator with efficiency exceeding 98% for the location parameter of Cauchy distributions. Here, this paper explores the estimation of the scale parameter for Cauchy distributions. Before proceeding, we give the universal definition for continuous populations as follows:

Definition 1.1. A population X is said to possess universality (or to be universal) if its cdf $F(x)$ has a third-order differentiable inverse function $G(x)$, and there exist constants $K > 0$ and $A \geq 0$ such that the inequation

$$|G'''(x)| \leq K \cdot x^{-A}(1-x)^{-A} \tag{1.2}$$

holds for all $x \in (0, 1)$.

As verified in [4], commonly encountered continuous populations, including the Cauchy distribution with no finite expectation, generally satisfy the above universality condition. For further discussion, here we present the following three theorems:

Theorem 1.4. *If a population X satisfies universality, then for any given constants $\omega \neq 0$ and v, the population $\omega X + v$ also satisfies universality.*

Theorem 1.5. *If the pdf of a population X that satisfies universality is an even function, then the population $|X|$ also possesses universality.*

Theorem 1.6. *If the pdf $f(y)$ of a population Y that satisfies the universality condition has bounded derivatives up to some given order $m \geq 2$, and if y_p denotes the p-quantile of Y, assuming $f(y_p)f(y_r) > 0$ for $0 < p \leq r < 1$, then for the sample $(Y_1, Y_2, \ldots, Y_n)$ with p-quantile $Q_{n,p}$, as the sample size $n \to +\infty$,*

$$EQ_{n,p} = y_p + O(n^{-1}), \operatorname{Var}\left(Q_{n,p}\right) \sim \frac{p(1-p)}{nf\left(y_p\right)^2}, \operatorname{Var}\left(Q_{n,r}\right) \sim \frac{r(1-r)}{nf\left(y_r\right)^2}, \tag{1.3}$$

and the asymptotic correlation coefficient is given by

$$\operatorname{corr}\left(Q_{n,p}, Q_{n,r}\right) \to \sqrt{\frac{p(1-r)}{r(1-p)}}, \quad n \to +\infty. \tag{1.4}$$

The Cauchy distribution, as a fundamental model with quintessential heavy-tailed characteristics, has significant applications across numerous scientific disciplines. In physics (often referred to as the Lorentz distribution), it describes resonance line shapes and provides the mathematical foundation for bidirectional reflectance distribution function (BRDF) modeling, enhancing the accuracy of optical property simulations for materials [9]. In financial econometrics, its variants (such as skewed truncated Cauchy distributions) effectively capture extreme fluctuations and complex statistical features in exchange rate data [10]. Additionally, in communications and information theory, the Cauchy distribution is employed to characterize impulsive noise and analyze certain channel capacity problems [11, 12].

However, its heavy-tailed nature (where even the mean and variance do not exist) poses fundamental challenges for parameter estimation. For the Cauchy distribution, moment-based estimation methods are entirely ineffective. Although maximum likelihood estimation (MLE) possesses asymptotic optimality [5], its numerical implementation faces well-documented difficulties: the likelihood equation may have multiple solutions, causing numerical algorithms to be sensitive to initial values and prone to converging to local optima or failing to converge altogether [13, 14]. Even when convergence is achieved, MLE may lack robustness against outliers in finite samples.

To address these challenges, researchers have developed a range of alternative estimation methods, primarily focused on improving robustness and computational feasibility: Quantile-based (L-estimation) methods: These construct estimators using linear combinations of sample quantiles. Such methods are inherently robust and, for the Cauchy distribution, can achieve asymptotic relative efficiency (ARE) between 82% and 98%, while remaining computationally efficient [15–17]. Robust M-estimation: For example, Huber's M-estimation provides resistance to outliers while maintaining a certain level of ARE (e.g., 90%) [11, 18]. Bayesian methods: Posterior inference via Markov chain Monte Carlo (MCMC) techniques can effectively handle multimodal likelihood functions, though at higher computational costs [19]. One-step corrected estimation: Starting from a robust initial estimate (e.g., the median), a single Newton-Raphson iteration is performed to balance high asymptotic efficiency with computational simplicity. Its theoretical properties (e.g., Bahadur efficiency) have been extensively studied [20].

Despite the variety of methods, a core optimization problem remains unresolved when estimating the scale parameter of the Cauchy distribution: how to optimally select and weight multiple sample quantiles to construct an estimator with minimal asymptotic variance (i.e., highest efficiency). Existing approaches often fail to achieve the theoretical upper bound of efficiency. This is precisely the motivation for our study. Theorem 1.6 shows that, even under universal conditions, the sample quantiles from a super-heavy-tailed population still possess the property of moment convergence. This provides a crucial theoretical foundation for rigorously analyzing and optimizing the asymptotic variance of linear combinations of quantiles. It thus enables us to theoretically derive a near-optimal weighted quantile estimator for the scale parameter of the Cauchy distribution.

As a direct application of our theoretical framework, Section 3 of this paper focuses on estimating the scale parameter of the Cauchy distribution. Through rigorous theoretical analysis and extensive simulations, we demonstrate that the proposed method achieves both high computational efficiency and estimation efficacy exceeding 99% (as measured by asymptotic relative efficiency).

1.1. Structure of the paper

The remainder of this paper is organized as follows:

- Section 2 presents the detailed proofs of Theorems 1.4–1.6, establishing the universality condition and moment convergence properties for sample quantiles.
- Section 3 applies these results to the Cauchy distribution:
  - Subsection 3.1 introduces the Cauchy model and its statistical properties.

  - Subsection 3.2 derives a robust estimator for the scale parameter c (Theorem 1.4) under known location parameter μ, leveraging quantile-based asymptotics.

- Section 4 validates the method through numerical simulations, efficiency comparisons (e.g., 99% relative efficiency to MLE) and discusses extensions.
- Section 5 concludes with a summary of key contributions and potential applications.

2. Main proofs

2.1. *Proof of Theorem 1.4*

Proof. (i) Now that the population X satisfies universality, we assume condition (1.2) holds. Note that for any given constant c, the quantile function of the population $X + c$ is $G_{X+c}(x) = G_X(x) + c$, and thus

$$\left|G'''_{X+c}(x)\right| \leq K \cdot x^{-A}(1-x)^{-A}.$$

This implies that $X + c$ also satisfies universality.

(ii) Suppose the population X satisfies universality and the inequality (1.2) holds.

- Case $\omega > 0$: The quantile function of ωX is $G_{\omega X}(x) = \omega G_X(x)$, so

$$\left|G'''_{\omega X}(x)\right| = \omega \left|G'''_X(x)\right| \leq (\omega K) \cdot x^{-A}(1-x)^{-A}.$$

Thus, ωX satisfies universality.

- Case $\omega < 0$: The quantile function of ωX is $G_{\omega X}(x) = \omega G_X(1-x)$, and hence

$$\left|G'''_{\omega X}(x)\right| \leq (-\omega K) \cdot x^{-A}(1-x)^{-A}.$$

Therefore, ωX also satisfies universality.

From (ii), we conclude that if X satisfies universality, then ωX must also satisfy universality. Combining (i) and (ii), we see that linear transformations of the population preserve universality. That is, if X satisfies universality, then for any constants $\omega \neq 0$ and $v \in \mathbb{R}$, the population $\omega X + v$ also satisfies universality. □

2.2. *Proof of Theorem 1.5*

Proof. For convenience, let G_Y denote the quantile function corresponding to the cdf F_Y of a random variable Y. Under the universal condition for X, we assume there exist constants $K > 0$ and $A \geq 0$ such that

$$\left|G'''_X(p)\right| \leq K \cdot p^{-A}(1-p)^{-A}.$$

By the symmetry of the population distribution $F(-x) = 1 - F(x)$, the distribution function of $|X|$ is given by

$$F_{|X|}(x) = P(|X| \leq x) = 2F(x) - 1, \quad x \geq 0.$$

Its quantile function $G_{|X|}(u)$ satisfies $F_{|X|}\left(G_{|X|}(u)\right) = u$, which implies $2F\left(G_{|X|}(u)\right) - 1 = u$. Thus, we obtain

$$G_{|X|}(u) = G_X\left(\frac{1+u}{2}\right), \ \textit{namely}\ G_{|X|}(x) = G_X\left(\frac{1+x}{2}\right).$$

Taking the third derivative of $G_{|X|}(x) = G_X\left(\frac{1+x}{2}\right)$, we have

$$G'''_{|X|}(x) = \frac{1}{8} G'''_X\left(\frac{1+x}{2}\right),$$

and hence

$$\left|G'''_{|X|}(x)\right| = \frac{1}{8}\left|G'''_X\left(\frac{1+x}{2}\right)\right| \leq \frac{1}{8} \cdot K\left(\frac{1+x}{2}\right)^{-A}\left(1-\frac{1+x}{2}\right)^{-A} = \frac{2^{2A}K}{8} \cdot (1+x)^{-A}(1-x)^{-A}.$$

Noting that for $x \in (0,1)$, $(1+x)^{-A} \leq 2^{-A}$ (since $1+x \geq 1$), we derive

$$\left|G'''_{|X|}(x)\right| \leq \frac{2^{2A}K}{8} \cdot 2^{-A}(1-x)^{-A} = \frac{K \cdot 2^A}{8}(1-x)^{-A} \leq \frac{K \cdot 2^A}{8} x^{-A}(1-x)^{-A}.$$

Therefore, $|X|$ also satisfies the universal condition. □

2.3. *Proof of Theorem 1.6*

Proof. First, the conclusion that as $n \to +\infty$, the limit conclusion

$$\operatorname{corr}\left(Q_{n,\mathrm{p}}, Q_{n,\mathrm{r}}\right) \to \sqrt{\frac{p(1-r)}{r(1-p)}}$$

follows directly from Theorem 1.1. Next, by the population universality, we first prove that condition (i) of Theorem 1.2 holds. For this purpose, we will prove the proposition: For a differentiable function $G(x)$, if there exist constants $M > 0$ and $D \geq 0$ such that the inequation

$$|G'(x)| \leq M \cdot x^{-D}(1-x)^{-D}$$

holds for $\forall x \in (0,1)$, then there exists a constant $L > 0$ such that for $\forall x \in (0,1)$,

$$|G(x)| \leq L \cdot x^{-(D+1)}(1-x)^{-(D+1)}. \tag{2.1}$$

In fact, since $G(x) - G\left(\frac{1}{2}\right) = \int_{\frac{1}{2}}^{x} G'(t)dt$ holds for $\forall x \in (0,1)$, combined with $|G'(x)| \leq M \cdot x^{-D}(1-x)^{-D}$, we have:

1) For $x \in (0, 1/2]$,

$$\begin{aligned} |G(x)| \leq |G(1/2)| + \left|\int_{1/2}^{x} G'(t)dt\right| &= |G(1/2)| + \left|\int_{x}^{1/2} G'(t)dt\right| \\ &\leq |G(1/2)| + \int_{x}^{1/2} |G'(t)|\, dt \\ &\leq |G(1/2)| + \int_{x}^{1/2} M \cdot t^{-D}(1-t)^{-D}dt. \end{aligned}$$

Since $t \leq 1/2$, we have $1 - t \geq 1/2$, thus $(1-t)^{-D} \leq 2^D$. Therefore,

$$|G(x)| \quad \leq \quad |G(1/2)| + \mathrm{M}\int_{x}^{1/2} t^{-D}(1-t)^{-D}dt \leq |G(1/2)| + 2^D\mathrm{M}\int_{x}^{1/2} t^{-D}dt$$

$$= \begin{cases} \left|G\left(\frac{1}{2}\right)\right| + 2^{\mathrm{D}} M \frac{\left(\frac{1}{2}\right)^{1-D} - x^{1-D}}{1-D}, & D \neq 1 \\ \left|G\left(\frac{1}{2}\right)\right| + 2M(-\ln 2 - \ln x), & D = 1 \end{cases}.$$

Clearly, as long as $D \geq 0$, the conclusion $\lim_{x \to 0^+} G(x)[x(1-x)]^{D+1} = 0$ holds.
2) For $x \in [1/2, 1)$, similarly we obtain

$$|G(x)| \leq |G(1/2)| + M \int_{1/2}^{x} t^{-D}(1-t)^{-D} dt.$$

Since $1/2 \leq t \leq 1$ implies $t^{-D} \leq 2^D$, we have

$$\begin{aligned} |G(x)| &\leq |G(1/2)| + 2^D M \int_{1/2}^{x} (1-t)^{-D} dt = |G(1/2)| - 2^D M \int_{1/2}^{1-x} v^{-D} dv \\ &= \begin{cases} \left|G\left(\frac{1}{2}\right)\right| + 2^D M(-\ln(1-x) - \ln 2), D = 1 \\ \left|G\left(\frac{1}{2}\right)\right| + 2^D M \frac{\left(\frac{1}{2}\right)^{1-D} - (1-x)^{1-D}}{1-D}, \quad D \neq 1 \end{cases}. \end{aligned}$$

Clearly, when $D \geq 0$, the conclusion $\lim_{x \to 1-} G(x)[x(1-x)]^{D+1} = 0$ must hold.

Combining both cases 1) and 2), we evidently obtain the conclusion

$$\lim_{x \to 0^+} G(x)[x(1-x)]^{D+1} = \lim_{x \to 1^-} G(x)[x(1-x)]^{D+1} = 0.$$

Note that the function $G(x)[x(1-x)]^{D+1}$ is also continuous on $(0, 1)$. Therefore, there exists a sufficiently large positive constant L such that the inequality $\left|G(x)[x(1-x)]^{D+1}\right| \leq L$ holds for $\forall x \in (0, 1)$. Equivalently the inequation

$$|G(x)| \leq Lx^{-(D+1)}(1-x)^{-(D+1)}$$

holds for $\forall x \in (0, 1)$ and thus we complete the proof of the proposition (2.1).

Now, since the population universality condition is satisfied, the quantile function $G(x)$ is three times differentiable, and there exist constants $K > 0$ and $A \geq 0$ such that the inequation

$$|G'''(x)| \leq K \cdot x^{-A}(1-x)^{-A}$$

holds for all $x \in (0, 1)$. By repeatedly applying the proven proposition, we conclude that there exist constants $C_1 > 0, C_2 > 0$, and $C_3 > 0$ such that for all $x \in (0, 1)$, the following inequalities

$$|G''(x)| \leq C_1 \cdot x^{-(A+1)}(1-x)^{-(A+1)}, |G'(x)| \leq C_2 \cdot x^{-(A+2)}(1-x)^{-(A+2)}$$

and

$$|G(x)| \leq C_3 \cdot x^{-(A+3)}(1-x)^{-(A+3)}$$

hold. Here the last inequality demonstrates that condition (i) of Theorem 1.2 is satisfied. It is evident that the other conditions of Theorem 1.2 are also satisfied. Therefore, according to the Eq (1.1) in Theorem 1.2, we have

$$E\left(\frac{f\left(y_p\right)(Y_{i:n} - a_{i:n})}{\sqrt{p(1-p)/n}}\right)^m = EZ^m + O\left(n^{-1/2}\right), \tag{2.2}$$

where $a_{i:n} = y_p + O\left(n^{-1}\right)$, and $Y_{i:n}$ is the i-th order statistic of the sample $(Y_1, Y_2, \ldots, Y_n)$.

As conditions also hold for $1 \le m \le 2$, by taking $m = 1$ in Eq (2.2) we have the conclusion

$$E\left(\frac{f\left(y_p\right)(Y_{i:n} - a_{i:n})}{\sqrt{p(1-p)/n}}\right) = O\left(n^{-1/2}\right),$$

which implies $E\left(Y_{i:n} - a_{i:n}\right) = O\left(n^{-1}\right)$, namely,

$$EY_{i:n} = a_{i:n} + O\left(n^{-1}\right) = y_p + O\left(n^{-1}\right). \tag{2.3}$$

Obviously, we can also have

$$EY_{i+1:n} = a_{i:n} + O\left(n^{-1}\right) = y_p + O\left(n^{-1}\right),$$

and accordingly

$$EQ_{n,p} = a_{i:n} + O\left(n^{-1}\right) = y_p + O\left(n^{-1}\right). \tag{2.4}$$

Moreover, the result (2.3) indicates that $EY_{i:n}$ can replace $a_{i:n}$ in Eq (2.2) and thus

$$E\left(\frac{f\left(y_p\right)(Y_{i:n} - EY_{i:n})}{\sqrt{p(1-p)/n}}\right)^m = EZ^m + O\left(n^{-1/2}\right).$$

In particular, setting $m = 2$ in (2.2) again yields

$$E\left(\frac{f\left(y_p\right)(Y_{i:n} - EY_{i:n})}{\sqrt{p(1-p)/n}}\right)^2 = 1 + O\left(n^{-1/2}\right).$$

That indicates

$$\mathrm{Var}\left(Y_{i:n}\right) = \frac{p(1-p)}{nf\left(y_p\right)^2} + O\left(n^{-3/2}\right).$$

Clearly, replacing i with $i + 1$ gives

$$\mathrm{Var}\left(Y_{i+1:n}\right) = \frac{p(1-p)}{nf\left(y_p\right)^2} + O\left(n^{-3/2}\right).$$

Furthermore, according to Theorem 1.1 and by taking $p = r$ and $j = i + 1$, we obtain

$$\lim_{n\to\infty} \mathrm{corr}\left(Y_{i:n}, Y_{i+1:n}\right) = \sqrt{\frac{p(1-p)}{p(1-p)}} = 1.$$

Accordingly, by denoting $\mathrm{corr}\left(Y_{i:n}, Y_{j:n}\right) = 1 + o(1)$, we see that

$$\mathrm{Var}\left(\frac{Y_{i:n} + Y_{i+1:n}}{2}\right) = \frac{1}{4}\left[\mathrm{Var}\left(Y_{i:n}\right) + \mathrm{Var}\left(Y_{i+1:n}\right) + 2\,\mathrm{corr}\left(Y_{i:n}, Y_{i+1:n}\right)\sqrt{\mathrm{Var}\left(Y_{i:n}\right)\mathrm{Var}\left(Y_{i+1:n}\right)}\right]$$

$$= \frac{1}{2}\left\{\frac{p(1-p)}{nf\left(y_p\right)^2} + O\left(n^{-3/2}\right) + (1+o(1))\left[\frac{p(1-p)}{nf\left(y_p\right)^2} + O\left(n^{-3/2}\right)\right]\right\}$$
$$= \frac{p(1-p)}{nf\left(y_p\right)^2} + o\left(n^{-1}\right) \sim \frac{p(1-p)}{nf\left(y_p\right)^2}.$$

In summary, regardless of whether the sample $(Y_1, Y_2, \ldots, Y_n)$ uses $Q_{n,p} = Y_{i:n}$ or $Q_{n,p} = \frac{Y_{i:n}+Y_{i+1:n}}{2}$ as its p-quantile estimator, we have

$$\mathrm{Var}\left(Q_{n,p}\right) \sim \frac{p(1-p)}{nf\left(y_p\right)^2}.$$

Obviously we also have

$$\mathrm{Var}\left(Q_{n,r}\right) \sim \frac{r(1-r)}{nf\left(y_r\right)^2}.$$

□

3. Parameter estimation for Cauchy distributions

3.1. Cauchy distribution and some relative application models

Assume a radioactive substance is placed at the unit point $(0, 1)$ on the y-axis of the xoy plane coordinate system. At each moment, it may emit a particle outward randomly, with the emission angle uniformly distributed over the interval $[0, 2\pi]$. It can be proven (see, e.g., [13] Johnson et al., 1994) that the position coordinate X where the emitted particle intersects the x-axis follows a Cauchy distribution; Similarly, the Cauchy distribution can describe the light intensity distribution along a line beneath a point light source; If the radioactive substance is analogized to represent the economic influence of a major city, the distribution of its impact on surrounding areas can also be characterized by the Cauchy distribution. Currently, the Cauchy distribution finds important applications in various fields such as physics, signal processing, financial engineering, and biology. The Cauchy distribution serves as a crucial complement to the normal distribution.

By writing

$$X \sim f(x;\mu,c) = \frac{c}{\pi\left[c^2 + (x-\mu)^2\right]}, x \in \mathbb{R}, \quad or\ X \sim \mathrm{Cauchy}(\mu, c),$$

we denote a Cauchy distribution with μ being the location parameter and $c > 0$ being the scale parameter, particularly, the standard Cauchy distribution has a pdf $f(x; 0, 1) = \frac{1}{\pi(1+x^2)}$.

In addition to the nonexistence of mean, the Cauchy distribution possesses the following properties:

(i) **Scale invariance:**

$$X \sim \mathrm{Cauchy}(\mu, c) \implies aX + b \sim \mathrm{Cauchy}(a\mu + b, |a|c), \quad a \neq 0.$$

(ii) **Closure property:** If $X_1 \sim \mathrm{Cauchy}(\mu_1, c_1)$ and $X_2 \sim \mathrm{Cauchy}(\mu_2, c_2)$ are independent, then:

$$X_1 + X_2 \sim \mathrm{Cauchy}(\mu_1 + \mu_2, c_1 + c_2).$$

Consequently, when a population follows a Cauchy distribution, the corresponding sample mean converges in distribution to the population distribution itself rather than to any constant.

3.2. *Scale estimation for Cauchy distribution* Cauchy(μ, **c**) *when* μ *is known*

Since [3] has already established estimation methods for the location parameter of Cauchy distributions, this paper focuses on developing efficient estimators for the scale parameter c.

Given that $X \sim Cauchy(\mu, c)$ implies $X - \mu \sim Cauchy(0, c)$, we can, without loss of generality, restrict our analysis to the case where $\mu = 0$. Thus, we consider $X \sim Cauchy(0, c)$, where $c > 0$ is the unknown scale parameter and have the following conclusion:

Theorem 3.1. *For a Cauchy-distributed population* $X \sim \text{Cauchy}(0, c)$*, where* $c > 0$ *is the unknown scale parameter, we construct a new sample* $(Y_1, \ldots, Y_n) = (|X_1|, \ldots, |X_n|)$ *from the original sample* $(X_1, \ldots, X_n)$*. Let* $Q_{n,p}$ *be the p-quantile of the new sample* ($Y_1, \ldots, Y_n$)*. Then, for any given* $p \in (0, 1)$*, the estimator*

$$\hat{c}_{n,p} := \frac{Q_{n,p}}{\tan\left(\frac{p\pi}{2}\right)}$$

is an asymptotically unbiased estimator of c, with an equivalent variance expression

$$\text{Var}\left(\hat{c}_{n,p}\right) \sim \frac{p(1-p)c^2\pi^2}{n\sin^2(\pi p)}. \tag{3.1}$$

If $0 < p \le r < 1$*, then the covariance* $\text{cov}\left(\hat{c}_{n,p}, \hat{c}_{n,r}\right)$ *has an equivalent expression*

$$\text{cov}\left(\hat{c}_{n,p}, \hat{c}_{n,r}\right) \sim \frac{c^2\pi^2 p(1-r)}{n\sin(p\pi)\sin(r\pi)}. \tag{3.2}$$

Proof. For the new population $Y = |X|$, the p-quantile $y_p > 0$ satisfies

$$\int_0^{y_p} \frac{2c}{\pi\left(x^2+c^2\right)}dx = p \quad \Rightarrow \quad y_p = c \cdot \tan\left(\frac{p\pi}{2}\right).$$

Clearly, the conditions of Theorems 1.4–1.6 are all satisfied. Therefore, by Theorem 1.6, the sample p-quantile $Q_{n,p}$ of $(Y_1, Y_2, \ldots, Y_n)$ satisfies

$$\text{E}Q_{n,p} = y_p + O\left(n^{-1}\right) = c \cdot \tan\left(\frac{p\pi}{2}\right) + O\left(n^{-1}\right).$$

Consequently,

$$\text{E}\hat{c}_{n,p} = \text{E}\frac{Q_{n,p}}{\tan\left(\frac{p\pi}{2}\right)} = c + O\left(n^{-1}\right),$$

which shows that $\hat{c}_{n,p}$ is an asymptotically unbiased estimator of c. Furthermore, Theorem 1.6 gives

$$\begin{aligned}\text{Var}\left(Q_{n,p}\right) \sim \frac{p(1-p)}{nf\left(y_p\right)^2} &= \frac{p(1-p)}{n\left\{\frac{2c}{\pi\left[\left(c\cdot\tan\left(\frac{p\pi}{2}\right)\right)^2+c^2\right]}\right\}^2} = \frac{p(1-p)}{n\left\{\frac{2}{\pi c\left(\sec\left(\frac{p\pi}{2}\right)\right)^2}\right\}^2} \\ &= \frac{p(1-p)}{n\left(\frac{2}{\pi c}\right)^2\left(\cos\left(\frac{p\pi}{2}\right)\right)^4} = \frac{\pi^2c^2}{4n}\frac{p(1-p)}{\left(\cos\left(\frac{p\pi}{2}\right)\right)^4}.\end{aligned}$$

Therefore,

$$\mathrm{Var}\left(\hat{c}_{n,p}\right) = \frac{\mathrm{Var}\left(Q_{n,p}\right)}{\left(\tan\left(\frac{p\pi}{2}\right)\right)^2} \sim \frac{\pi^2 c^2}{4n}\frac{p(1-p)}{\left(\cos\left(\frac{p\pi}{2}\right)\right)^4\left(\tan\left(\frac{p\pi}{2}\right)\right)^2} = \frac{\pi^2 c^2}{n}\frac{p(1-p)}{(\sin(p\pi))^2}.$$

Finally, for $0 < p \leq r < 1$, we can readily obtain

$$\begin{aligned}\mathrm{cov}\left(\hat{c}_{n,p}, \hat{c}_{n,r}\right) &= corr(\hat{c}_{n,p}, \hat{c}_{n,r})\sqrt{Var(\hat{c}_{n,p})Var(\hat{c}_{n,r})} = corr(Q_{n,p}, Q_{n,r})\sqrt{Var(\hat{c}_{n,p})Var(\hat{c}_{n,r})} \\ &\sim \sqrt{\frac{p(1-r)}{r(1-p)}}\sqrt{\frac{\pi^2c^2}{n}\frac{p(1-p)}{(\sin(p\pi))^2}\frac{\pi^2c^2}{n}\frac{r(1-r)}{(\sin(r\pi))^2}} = \frac{c^2\pi^2 p(1-r)}{n\sin(p\pi)\sin(r\pi)},\end{aligned}$$

and thus we complete the proof of

$$\mathrm{cov}\left(\hat{c}_{n,p}, \hat{c}_{n,r}\right) \sim \frac{c^2\pi^2 p(1-r)}{n\sin(p\pi)\sin(r\pi)}.$$

□

Corollary 3.1. *Under the conditions of Theorem 1.4, the estimator* $\hat{c}_{n,p} := \frac{Q_{n,p}}{\tan(\frac{p\pi}{2})}$ *is optimal when* $p = 0.5$. *That is, if we use the linear function of only one quantile of the new sample* $(Y_1, \ldots, Y_n)$ *to estimate c, then* $\hat{c}_{n,0.5}$ *is asymptotically optimal.*

In this case of $X \sim Cauchy(0, c)$, *we denote the asymptotically optimal unbiased estimator as*

$$\widehat{E1} := \hat{c}_{n,0.5} = \left.\frac{Q_{n,p}}{\tan\left(\frac{p\pi}{2}\right)}\right|_{p=0.5} = Q_{n,0.5} = \mathrm{median}(Y_1, \ldots, Y_n) = \mathrm{median}(|X_1|, \ldots, |X_n|),$$

where $\widehat{E1}$ *has an asymptotic relative efficiency*

$$\frac{\mathrm{CRLB}}{\mathrm{Var}(\widehat{\mathrm{E1}})} = \frac{2c^2}{n}/\frac{c^2\pi^2}{4n} = \frac{8}{\pi^2} \approx 0.8106.$$

Here, CRLB represents the asymptotic variance of the maximum likelihood estimator, or equivalently the lower bound of the Cramér-Rao inequation.

Remark 3.1. In Corollary 3.1, although the distribution of Y is asymmetric, the asymptotic variance $\frac{p(1-p)c^2\pi^2}{n\sin^2(\pi p)}$ of the asymptotically unbiased estimator $\hat{c}_{n,p}$ constructed from its corresponding samples reaches its minimum at $p = 0.5$ and is symmetric about $p = 0.5$ (see Appendix 1), as is displayed in the following Figure 1. This inspires us that when constructing an asymptotically unbiased estimator for c using a linear combination of several sample statistics $\hat{c}_{n,p}$ from $(Y_1, \ldots, Y_n)$, if the statistic $\hat{c}_{n,p}$ is selected, the corresponding $\hat{c}_{n,1-p}$ should also be selected with equal weights. Additionally, if an odd number of $\hat{c}_{n,p}$s is given, it must include $\hat{c}_{n,0.5}$.

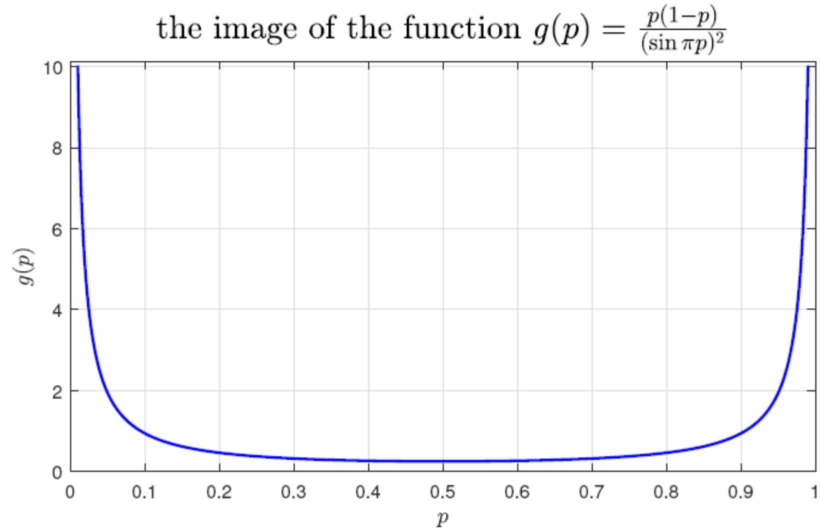

Figure 1. Symmetry & minimum of $\frac{p(1-p)}{(sin(\pi p))^2}$ at $p = 0.5$.

Now we firstly try to find an optimal estimator which we denoted as $\widehat{\text{E2}}$ for c by using a linear combination of two sample quantiles from the new sample. The following obtained result of $\widehat{\text{E2}}$ verifies the above analysis. $\widehat{\text{E2}}$ should be the best estimator among the estimators formed as

$$\text{E2} = \hat{c}_{n,p}u + \hat{c}_{n,r}(1-u), 0 < p \leq r < 1.$$

Noting that

$$\begin{aligned}
\text{Var}\left(u\hat{c}_{n,p} + (1-u)\hat{c}_{n,r}\right) &= u^2 \,\text{Var}\left(\hat{c}_{n,p}\right) + (1-u)^2 \,\text{Var}\left(\hat{c}_{n,r}\right) + 2u(1-u)\,\text{cov}\left(\hat{c}_{n,p}, \hat{c}_{n,r}\right) \\
&\sim u^2 \frac{p(1-p)c^2\pi^2}{n\sin^2(\pi p)} + (1-u)^2 \frac{r(1-r)c^2\pi^2}{n\sin^2(\pi r)} + 2u(1-u)\frac{c^2\pi^2 p(1-r)}{n\sin(p\pi)\sin(r\pi)} \\
&= \frac{c^2\pi^2}{n}\left[u^2\frac{p(1-p)}{\sin^2(\pi p)} + (1-u)^2\frac{r(1-r)}{\sin^2(\pi r)} + \frac{2u(1-u)p(1-r)}{\sin(\pi p)\sin(\pi r)}\right] := \frac{c^2\pi^2}{n}h(u,p,r),
\end{aligned}$$

to determine the value of u that minimizes the expression $h(u, p, r)$ above, we first take the partial derivative of $h(u, p, r)$ with respect to u and set it to zero. Solving for u and substituting it back into the expression $h(u, p, r)$, we obtain:

$$\frac{p(1-r)(r-p)}{(r-r^2)\sin(\pi p)^2 - 2p\sin(\pi r)(1-r)\sin(\pi p) + p\sin(\pi r)^2(1-p)} := f(p,r).$$

Subsequently, by visualizing the objective function $f(p, r)$ around the initial point $(0.25, 0.75)$, we numerically determined the minimum point $(p, r) = (0.371010, 0.628990)$ using Maple. The corresponding optimal u value is 0.5 , with the minimized expression $h(u, p, r)$ yielding a value of 0.219642 (see Appendix 2).

As we can see, the calculated results are in accordance with the previous analysis in Remark 3.1. What is more, when using two sample quantiles to estimate c, we can also see that the asymptotically optimal unbiased estimator

$$\widehat{\text{E2}} = 0.5\hat{c}_{n,0.371010} + 0.5\hat{c}_{n,0.6290} = 0.5 \cdot \frac{Q_{n,0.3710}}{\tan\left(\frac{0.3710\pi}{2}\right)} + 0.5 \cdot \frac{Q_{n,0.6290}}{\tan\left(\frac{0.6290\pi}{2}\right)}$$

$$= \quad 0.7586Q_{n,0.3710} + 0.3296Q_{n,0.6290},$$

has an asymptotic variance of $\frac{0.219642c^2\pi^2}{n}$, corresponding to an asymptotic efficiency of

$$\frac{\text{CRLB}}{\text{Var}(\widehat{\text{E2}})} = \frac{\frac{2c^2}{n}}{\text{Var}(\widehat{\text{E2}})} = \frac{2c^2}{n} / \left(\frac{0.219642c^2\pi^2}{n}\right) = 0.92260. \tag{3.3}$$

We secondly try to find an optimal estimator which we denoted as $\widehat{\text{E7}}$ for c by using a linear combination of 7 sample quantiles from the new sample. Namely, $\widehat{\text{E7}}$ attains the minimum asymptotic variance among estimators of the form

$$\begin{aligned}
\text{E7} &= u\left(\hat{c}_{n,p} + \hat{c}_{n,1-p}\right) + v\left(\hat{c}_{n,q} + \hat{c}_{n,1-q}\right) + w\left(\hat{c}_{n,r} + \hat{c}_{n,1-r}\right) + (1 - 2u - 2v - 2w)\hat{c}_{n,0.5} \\
&= u\left(\frac{Q_{n,p}}{\tan\left(\frac{p\pi}{2}\right)} + \frac{Q_{n,1-p}}{\tan\left(\frac{(1-p)\pi}{2}\right)}\right) + v\left(\frac{Q_{n,q}}{\tan\left(\frac{q\pi}{2}\right)} + \frac{Q_{n,1-q}}{\tan\left(\frac{(1-q)\pi}{2}\right)}\right) \\
&\quad + w\left(\frac{Q_{n,r}}{\tan\left(\frac{r\pi}{2}\right)} + \frac{Q_{n,1-r}}{\tan\left(\frac{(1-r)\pi}{2}\right)}\right) + (1 - 2u - 2v - 2w)\frac{Q_{n,0.5}}{\tan\left(\frac{\pi}{4}\right)},
\end{aligned}$$

where $0 < p \leq q \leq r \leq 0.5$.

For the sake of convenient utilization of Maple, we define a function

$$\text{fcov(p, r)} := \frac{c^2\pi^2 p(1 - r)}{n\sin(p\pi)\sin(r\pi)}, \quad 0 < \text{p} \leq \text{r} < 1.$$

According to the conclusion (3.2), we are clear that as $n \to \infty$,

$$cov(\hat{c}_{n,p}, \hat{c}_{n,r}) \sim \text{fcov(p, r)}, \quad 0 < \text{p} \leq \text{r} < 1;$$

Moreover, for $0 < \text{p} \leq \text{r} \leq 0.5$, since

$$\begin{aligned}
cov(\hat{c}_{n,p} + \hat{c}_{n,1-p}, \hat{c}_{n,r} + \hat{c}_{n,1-r}) &\sim \text{fcov}(p, r) + \text{fcov}(p, 1 - r) + \text{fcov}(r, 1 - p) + \text{fcov}(1 - r, 1 - p) \\
&= \frac{c^2\pi^2[p(1 - r) + pr + rp + (1 - r)p]}{n\sin(p\pi)\sin(r\pi)} \\
&= \frac{2c^2\pi^2 p}{n\sin(p\pi)\sin(r\pi)},
\end{aligned} \tag{3.4}$$

we define another function

$$\text{mfcov(p, r)} := \frac{2c^2\pi^2 p}{n\sin(p\pi)\sin(r\pi)}, 0 < p \leq r \leq 0.5.$$

Now noting that the variance of the estimator $E7$ can be expressed as

$$\begin{aligned}
\text{Var}(E7) =& u^2 \text{Var}\left(\hat{c}_{n,p} + \hat{c}_{n,1-p}\right) + v^2 \text{Var}\left(\hat{c}_{n,q} + \hat{c}_{n,1-q}\right) + w^2 \text{Var}\left(\hat{c}_{n,r} + \hat{c}_{n,1-r}\right) \\
&+ (1 - 2u - 2v - 2w)^2 \text{Var}\left(\hat{c}_{n,0.5}\right) + 2uv\,\text{cov}\left(\hat{c}_{n,p} + \hat{c}_{n,1-p}, \hat{c}_{n,q} + \hat{c}_{n,1-q}\right) \\
&+ 2uw\,\text{cov}\left(\hat{c}_{n,p} + \hat{c}_{n,1-p}, \hat{c}_{n,r} + \hat{c}_{n,1-r}\right) + 2u(1 - 2u - 2v - 2w)\,\text{cov}\left(\hat{c}_{n,p} + \hat{c}_{n,1-p}, \hat{c}_{n,0.5}\right) \\
&+ 2vw\,\text{cov}\left(\hat{c}_{n,q} + \hat{c}_{n,1-q}, \hat{c}_{n,r} + \hat{c}_{n,1-r}\right) + 2v(1 - 2u - 2v - 2w)\,\text{cov}\left(\hat{c}_{n,q} + \hat{c}_{n,1-q}, \hat{c}_{n,0.5}\right) \\
&+ 2w(1 - 2u - 2v - 2w)\,\text{cov}\left(\hat{c}_{n,r} + \hat{c}_{n,1-r}, \hat{c}_{n,0.5}\right),
\end{aligned}$$

we define the third following function fun7 for convenient use later in Maple software

$$\begin{aligned}\mathrm{Var}(E7) \quad \sim \quad & u^2 \cdot \mathrm{mfcov}(p,p) + v^2 \cdot \mathrm{mfcov}(q,q) + w^2 \cdot \mathrm{mfcov}(r,r) \\ & +(1-2u-2v-2w)^2 \cdot \mathrm{fcov}(1/2,1/2) + 2uv \cdot \mathrm{mfcov}(p,q) + 2uw \cdot \mathrm{mfcov}(p,r) \\ & +u(1-2u-2v-2w) \cdot \mathrm{mfcov}(p,1/2) + 2vw \cdot \mathrm{mfcov}(q,r) \\ & +v(1-2u-2v-2w) \cdot \mathrm{mfcov}(q,1/2) + w(1-2u-2v-2w) \cdot \mathrm{mfcov}(r,1/2) \\ := \quad & \frac{c^2\pi^2}{n}\mathrm{fun7}. \end{aligned} \tag{3.5}$$

Through Maple optimization to minimize the asymptotic variance, we obtain

$$u = 0.08864913209, v = 0.1436902755, w = 0.1750267626,$$

$$p = 0.194668, q = 0.309483, r = 0.407633$$

and thereby derive

$$\begin{aligned}\widehat{\mathrm{E7}} =& 0.2808152418 Q_{n,0.195} + 0.02798519251 Q_{n,0.805} + 0.2719174618 Q_{n,0.309} \\ & + 0.07593074418 Q_{n,0.691} + 0.2349278975 Q_{n,0.408} + 0.1303990201 Q_{n,0.592} + 0.1852676596 Q_{n,0.5}.\end{aligned}$$

The minimum function value of fun7 can be obtained to be 0.20465990432 according to Maple software (see Appendix 3) and accordingly the asymptotic efficiency of $\widehat{\mathrm{E7}}$ can be worked out as

$$\frac{\mathrm{CRLB}}{\mathrm{Var}(\widehat{\mathrm{E7}})} = \frac{\frac{2c^2}{n}}{\mathrm{Var}(\widehat{\mathrm{E7}})} = \frac{\frac{2c^2}{n}}{0.20465990432 \times \frac{c^2\pi^2}{n}} = \frac{2}{0.20465990432\pi^2} = 0.9901420012.$$

As is implied by the above highly efficiency rate 0.9901420012, the estimator $\widehat{E7}$ has an asymptotic variance that quite closes to the variance of the classical MLE. The estimator $\widehat{E7}$ proposed in this paper can be regarded as a quick estimator for the shape parameter of the Cauchy distribution. It clearly possesses asymptotic consistency, and its efficiency exceeding 99% demonstrates its potential as a substitute for MLE.

For Cauchy distribution, the reference [21] provides an asymptotically unbiased estimator Q for the shape parameter c with

$$\mathrm{Q} = d \times \left\{ \left| X_i - X_j \right| ; i < j \right\}_{(k)} = 1.2071 \times \left\{ \left| X_i - X_j \right| ; i < j \right\}_{(k)},$$

here k represents the combination number $\binom{[\frac{n}{2}]+1}{2}$ where $\left[\frac{n}{2}\right]$ denotes the largest integer not exceeding n/2. The subscript k in the expression $\left\{ \left| X_i - X_j \right| : i < j \right\}_{(k)}$ stands for the k-th smallest order statistic in the set $\left\{ \left| X_i - X_j \right| : i < j \right\}$. The text mentions that the estimator Q achieves an asymptotic efficiency of $e(Q) = 98\%$ compared to the MLE (without rigorous proof or detailed algorithm). However, the construction of this estimator is highly complex, and its computational implementation is not significantly simpler than that of MLE.

Finally, as summarized in [6], the joint asymptotic distribution of different sample quantiles follows a multivariate normal distribution. Combined with the conclusions of Theorems 1.4 and 1.5 in references [3] and [4], it is evident that the estimator $\widehat{\mathrm{E7}}$ is asymptotically normally distributed, with moment equivalence to its corresponding normal distribution moment.

4. Simulation study

Regardless of whether the scale parameter is known or unknown, reference [3] has already provided estimation methods for the location parameter of the Cauchy distribution. Here, we only conduct simulation studies for the estimation of the scale parameter, considering both cases where the location parameter is known and unknown.

4.1. Estimating the scale parameter for Cauchy distribution with known location

Assume the true population follows $X \sim \text{Cauchy}(\mu, c) = \text{Cauchy}(4, 2)$. Using the transformed sample $(Y_1, \ldots, Y_n) = (|X_1 - 4|, \ldots, |X_n - 4|)$ where $n = 10000$, each time we respectively apply $\widehat{\text{E1}}$, $\widehat{\text{E2}}$ and $\widehat{\text{E7}}$ to estimate c and define that an experiment. We totally do 20 the same experiments independently with the same population and form the following Figure 2 indicating the effectiveness comparison of these estimators (see Appendix 3).

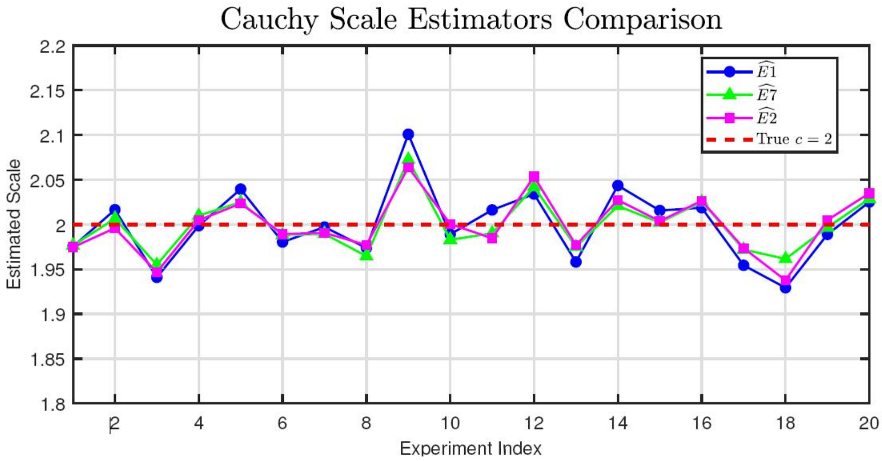

Figure 2. Comparing estimators $\widehat{\text{E1}}, \widehat{\text{E2}}$ and $\widehat{\text{E7}}(n = 10000)$ in estimating $c = 2$.

As is displayed in Figure 2, estimator $\widehat{E7}$ generally performs better than either $\widehat{E1}$ or $\widehat{E2}$.

4.2. Estimating the scale parameter for Cauchy distribution with unknown location

Let $m_{n,p}$ denote the p-quantile of the sample ($X_1, X_2, \ldots, X_n$), and define $R_n(p) := \frac{m_{n,p}+m_{n,1-p}}{2}$. Reference [4] demonstrates that the statistic

$$\begin{aligned} E_{5,n} = &- 0.0192R_n(0.0632) - 0.0747R_n(0.1347) + 0.2953R_n(0.3577) \\ &+ 0.3799R_n(0.4199) + 0.4187R_n(0.4739) \end{aligned}$$

is an unbiased estimator for the location parameter of a Cauchy-distributed population X, regardless of whether the scale parameter is known. The variance of $E_{5,n}$ has an asymptotic equivalence $\text{Var}(E_{5,n}) \sim 2.0314/n$. Compared to the MLE, the asymptotic relative efficiency of $E_{5,n}$ is $2/2.0314 = 0.9845$. In large samples, $E_{5,n}$ exhibits high stability, with values tightly concentrated around the true location parameter μ.

Generally, for a Cauchy distribution with unknown location μ and scale c, we can firstly estimate the unknown μ by the observation of $E_{5,n}$, then we consider the population distribution as known location situation, and then estimate the scale parameter c as the previous procedure of obtaining estimator $\widehat{E7}$, we can finally get estimator for c which we denoted as $\widehat{E7*}$.

For example, with $n = 10000$, the asymptotic normality of $E_{5,n}$ implies

$$P\left(\left|E_{5,n} - \mu\right| \geq 2\sqrt{\mathrm{Var}\left(E_{5,n}\right)}\right) = 2(1 - \Phi(2)) = 0.044,$$

meaning that

$$\mathrm{P}\left(\left|E_{5,n} - \mu\right| \geq 2\sqrt{2.0314/10000}\right) = 0.044 \rightarrow \mathrm{P}\left(\left|E_{5,n} - \mu\right| \geq 0.029\right) = 0.044.$$

This implies that when using $E_{5,10000}$ to estimate μ, it is highly improbable for the absolute error to exceed 0.029 . Therefore, we approximate μ as being equal to $E_{5,10000}$. Under this approximation, the population $\mathrm{X} - E_{5,n} \sim \mathrm{Cauchy}(0, \mathrm{c}), \mathrm{c} > 0$, and the corresponding sample is transformed into

$$(\mathrm{Y}_1, \ldots, \mathrm{Y}_\mathrm{n}) = \left(\left|\mathrm{X}_1 - E_{5,n}\right|, \ldots, \left|\mathrm{X}_\mathrm{n} - E_{5,n}\right|\right).$$

Following the aforementioned approach, we utilize the statistic $\widehat{E7^*}$ constructed from $(\mathrm{Y}_1, \ldots, \mathrm{Y}_\mathrm{n})$ to estimate the scale parameter c:

$$\begin{aligned}\widehat{E7^*} =& 0.2808152418Q^*_{n,0.195} + 0.02798519251Q^*_{n,0.805} + 0.2719174618Q^*_{n,0.309} \\ &+ 0.07593074418Q^*_{n,0.691} + 0.2349278975Q^*_{n,0.408} \\ &+ 0.1303990201Q^*_{n,0.592} + 0.1852676596Q^*_{n,0.5},\end{aligned}$$

where $Q^*_{n,p}$ represents the p-quantile of the new sample $\left(Y^*_1, \ldots, Y^*_n\right) := \left(\left|X_1 - E_{5,n}\right|, \ldots, \left|X_n - E_{5,n}\right|\right)$.

Subsequently, we perform simulation studies to compare the estimation performance between two scenarios with the actual value of the scale parameter $c = 2$.

As is indicated by the simulated results (Figure 3), estimator $\widehat{\mathrm{E7}}$ and $\widehat{E7*}$ are close to each other in estimating the scale parameter $c = 2$ under the large sample size $n = 10000$.

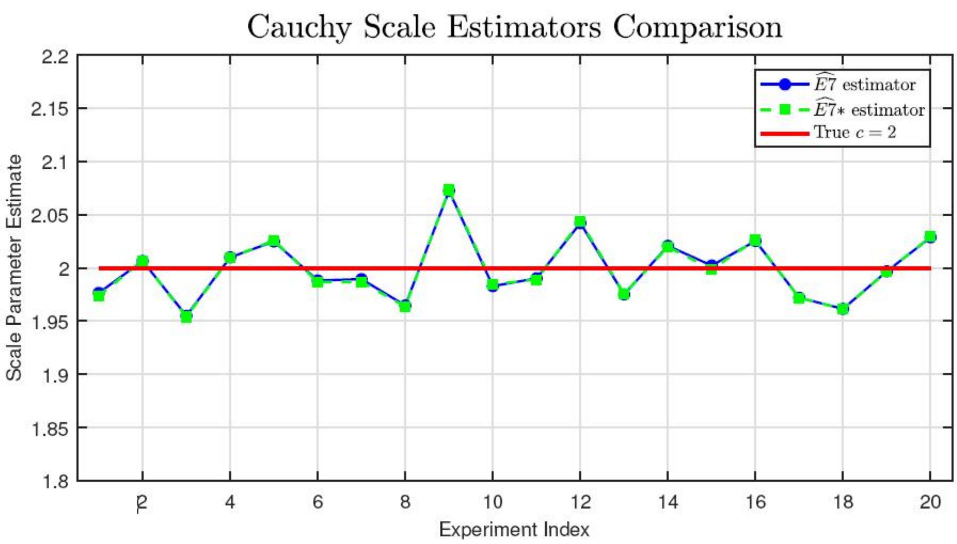

Figure 3. Comparing estimators $\widehat{\mathrm{E7}}$ *and* $\widehat{\mathrm{E7}*}$($n = 10000$) in estimating $c = 2$.

To examine the estimation performance with a moderately reduced sample size ($n = 600$), we present two subplots Figure 4(a) and 4(b). Subplot (a) compares the results of 20 estimations using $\widehat{E2}$ and $\widehat{E2*}$ (where $\widehat{E2*}$ is obtained by treating the sample median as the true population location parameter and then estimating the scale parameter using two quantiles from a new sample). Subplot (b) similarly compares 20 estimation results of $\widehat{E7}$ and $\widehat{E7*}$, as described earlier.

The results (see Figure 4 as well as Table 1) show that as the sample size n decreases, some discrepancies arise between $\widehat{E7}$ and $\widehat{E7*}$, while the differences between $\widehat{E2}$ and $\widehat{E2*}$ are more pronounced. The superiority of $\widehat{E7}$ over $\widehat{E2}$ is evident, as is the advantage of $\widehat{E7*}$ over $\widehat{E2*}$.

Table 1. Squared errors of 4 estimates in 20 experiments.

	$(\widehat{E2} - c)^2$	$(\widehat{E7} - c)^2$	$(\widehat{E2*} - c)^2$	$(\widehat{E7*} - c)^2$
1	0.03174569	0.03351053	0.04486434	0.03308145
2	0.00227712	0.00475764	0.00040686	0.00715247
3	0.01151640	0.00746694	0.01646004	0.00667703
4	0.00101425	0.00677288	0.00308368	0.00501002
5	0.00004881	0.00255630	0.00104718	0.00235408
6	0.00060215	0.00227445	0.00283797	0.00294717
7	0.00223157	0.00130517	0.00187146	0.00151906
8	0.00577856	0.00270196	0.00690803	0.00679727
9	0.01387076	0.01503183	0.01235822	0.01255513
10	0.00137906	0.00046466	0.00049879	0.00012351
11	0.01774646	0.01161823	0.01594170	0.01203941
12	0.00207789	0.00130980	0.00438729	0.00048977
13	0.01405944	0.00289688	0.01366790	0.00375828
14	0.03302616	0.04562693	0.03122447	0.04538348
15	0.01132997	0.00589189	0.01053595	0.00635019
16	0.00919226	0.00578929	0.00461350	0.00615215
17	0.00083791	0.00294791	0.00017231	0.00274508
18	0.05214600	0.02398681	0.04773047	0.02531572
19	0.00168733	0.00224212	0.00065236	0.00214939
20	0.00642312	0.00542582	0.00185151	0.00631507
Average	0.01094955	0.00922890	0.01105570	0.00944579

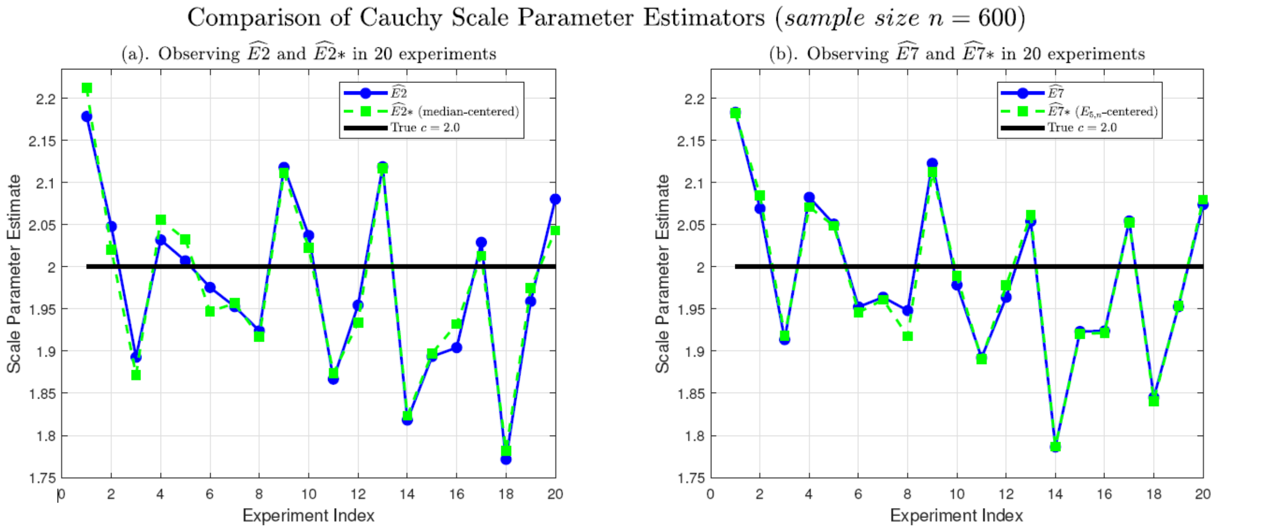

Figure 4. Comparing estimators $\widehat{\text{E2}}, \widehat{\text{E2}*}, \widehat{\text{E7}}$ *and* $\widehat{\text{E7}*}(n = 600)$ in estimating $c = 2$.

5. Concluding remarks

This study has established a theoretical foundation for the moment convergence of sample quantiles under a general universality condition, a property that holds for a wide range of continuous populations, including the Cauchy distribution. Leveraging this theory, we have developed a novel, computationally efficient estimator for the scale parameter of the Cauchy distribution by employing an optimal linear combination of seven sample quantiles from the transformed variable $Y = |X - \mu|$. Extensive simulation experiments confirm that the proposed estimator $\widehat{E7}$ achieves an asymptotic relative efficiency exceeding 99% relative to the maximum likelihood estimator (MLE), while circumventing the computational complexities and convergence issues associated with iterative MLE algorithms.

The methodology presented offers a robust and practical alternative for parameter estimation in ultra-heavy-tailed distributions, effectively balancing high statistical efficiency with computational feasibility. Future research could explore the extension of this quantile-based framework to other heavy-tailed distributions, such as the Pareto or stable families. Furthermore, investigating adaptive methods for optimal quantile selection in finite-sample scenarios or for multivariate heavy-tailed distributions presents a promising and challenging avenue for further work.

Use of Generative-AI tools declaration

The author declares he have not used Artificial Intelligence (AI) tools in the creation of this article.

Conflict of interest

The author declares no conflict of interest.

References

1. V. Chernozhukov, C. Hansen, An IV model of quantile treatment effects, *Econometrica*, **73** (2005), 245–261.
2. H. M. Barakat, M. A. Abd Elgawad, M. A. Alawady, Asymptotic order statistics of mixtures of distributions, *Rocky Mountain J. Math.*, **50** (2020), 429–443. https://doi.org/10.1216/rmj.2020.50.429
3. J. Wang, F. Wang, S. Hu, On asymptotic correlation coefficient for some order statistics, *AIMS Mathematics*, **8** (2023), 6763–6776. http://dx.doi.org/10.3934/math.2023344
4. J. Wang, C. Deng, J. Li, On moment convergence for some order statistics, *AIMS Mathematics*, **7** (2022), 17061–17079. https://doi.org/10.3934/math.2022938
5. J. Shao, Mathematical statistics, In: *Springer texts in statistics*, New York: Springer, 2003. https://doi.org/10.1007/b97553
6. A. Dasgupta, Asymptotic theory of statistics and probability, In: *Springer texts in statistics*, New York: Springer, 2008. https://doi.org/10.1007/978-0-387-75971-5
7. G. Stoops, D. Barr, Moments of certain Cauchy order statistics, *Amer. Statist.*, **25** (1971), 51–54. https://doi.org/10.1080/00031305.1971.10477307
8. H. A. David, H. N. Nagaraja, *Order statistics*, John Wiley & Sons, Inc., 2003. https://doi.org/10.1002/0471722162
9. F. Guo, J. Zhu, L. Huang, H. Li, J. Deng, X. Zhang, et al., A modified BRDF model based on Cauchy-Lorentz distribution theory for metal and coating materials, *Photonics*, **10** (2023), 773. https://doi.org/10.3390/photonics10070773
10. S. Shah, P. J. Hazarika, S. Chakraborty, The Balakrishnan Alpha skew truncated Cauchy distribution with applications in modelling currency exchange rate, *Sains Malays.*, **49** (2020), 2565–2571. http://dx.doi.org/10.17576/jsm-2020-4910-22
11. C. Haydaroğlu, B. Gümüş, Fault detection in distribution network with the Cauchy-M estimate—RVFLN method, *Energies*, **16** (2022), 252. https://doi.org/10.3390/en16010252
12. S. Pang, W. Zhang, Information rates of channels with additive white Cauchy noise, *IEEE Commun. Lett.*, **29** (2025), 115–119. https://doi.org/10.1109/LCOMM.2024.3502753
13. N. L. Johnson, S. Kotz, N. Balakrishnan, *Continuous univariate distributions*, 2 Eds., John Wiley & Sons, 1994.
14. 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.*, **74** (2022), 895–923. https://doi.org/10.1007/s10463-021-00818-y
15. J. A. Reeds, Asymptotic number of roots of Cauchy location likelihood equations, *Ann. Statist.*, **13** (1985), 775–784. https://doi.org/10.1214/aos/1176349554
16. R. Koenker, G. Bassett, Regression quantiles, *Econometrica*, **46** (1978), 33–50. https://doi.org/10.2307/1913643

17. G. Shevlyakov, P. Smirnov, Robust estimation of the correlation coefficient: An attempt of survey, *Aust. J. Stat.*, **40** (2011), 147–156.

18. J. Zhang, A highly efficient L-estimator for the location parameter of the Cauchy distribution, *Comput. Stat.*, **25** (2010), 97–105. https://doi.org/10.1007/s00180-009-0163-y

19. S. Fegyverneki, A simple robust estimation for parameters of Cauchy distribution, *Miskolc Math. Notes*, **14** (2013), 887–892. https://doi.org/10.18514/MMN.2013.830

20. L. Zhang, R. Smith, Robust M-estimation for heavy-tailed models, *J. Nonparametr. Stat.*, **30** (2022), 567–589.

21. P. J. Rousseeuw, C. Croux, Alternatives to the median absolute deviation, *J. Amer. Statist. Assoc.*, **88** (1993), 1273–1283. https://doi.org/10.1080/01621459.1993.10476408

A. Appendix 1

Proposition A.1. *The function* $g(x) = \frac{x(1-x)}{\sin^2(\pi x)}$ *defined on* $(0, 1)$ *is symmetric about* $x = 1/2$ *and attains its global minimum there.*

Proof. The symmetry $g(1-x) = g(x)$ is immediate from the definition, as substituting $x \mapsto 1-x$ leaves the expression unchanged.

To prove that $x = 1/2$ is the global minimizer, we show that $g(x) \geq g(1/2) = 1/4$ for all $x \in (0, 1)$. This inequality is equivalent to

$$4x(1 - x) - \sin^2(\pi x) \geq 0.$$

Define $h(x) = 4x(1 - x) - \sin^2(\pi x)$. Due to the symmetry of $g(x)$, it suffices to prove $h(x) \geq 0$ for $x \in [0, 1/2]$, with equality only at the endpoints. We have $h(0) = 0$ and $h(1/2) = 0$.

We analyze $h'(x) = 4(1 - 2x) - \pi \sin(2\pi x)$. At $x = 0$, $h'(0) = 4 > 0$. At $x = 1/2$, $h'(1/2) = 0$. The second derivative $h''(x) = -8 - 2\pi^2 \cos(2\pi x)$ has exactly one zero in $(0, 1/2)$ because $\cos(2\pi x)$ is strictly decreasing there. Hence, by Rolle's theorem, $h'(x)$ can have at most two zeros in $(0, 1/2]$. Now that $h'(1/2) = 0$, we see that $h'(x)$ can have at most one zero in $(0, 1/2)$.

Since $h'(1/2) = 0$ and $h''(1/2) > 0$, $h'(x)$ is increasing at $x = 1/2$, so $h'(x) < 0$ immediately left of $1/2$. Given $h'(0) > 0$ and $h'(x) < 0$ near $1/2^-$, by the Intermediate Value Theorem $h'(x)$ has at least one zero in $(0, 1/2)$. Thus $h'(x)$ has exactly one zero $x_1 \in (0, 1/2)$ and one zero at $x = 1/2$.

Therefore, $h'(x) > 0$ on $(0, x_1)$, $h'(x) < 0$ on $(x_1, 1/2)$. This implies $h(x)$ increases on $[0, x_1]$ and decreases on $[x_1, 1/2]$. Since $h(0) = h(1/2) = 0$, it follows that $h(x) > 0$ for $x \in (0, 1/2)$. □

B. Appendix 2

Maple DAids in obtaining the best linear combination of two new sample quantiles for estimating the scale parameter c:

restart :

$$h := (u, p, r) \to u^2 \cdot \left(\frac{p(1-p)}{\sin^2(\pi p)}\right) + (1-u)^2 \cdot \left(\frac{r(1-r)}{\sin^2(\pi r)}\right) \quad + 2u(1-u) \cdot \frac{p(1-r)}{\sin(\pi p)\sin(\pi r)},$$

$$u := \text{solve}\left(\frac{\partial}{\partial u}(h(u,p,r)) = 0, u\right) : \text{simplify}(h(u,p,r))$$

$$-\frac{p(-1+r)(p-r)}{(r^2-r)\sin^2(\pi p) - 2p\sin(\pi r)(-1+r)\sin(\pi p) + \sin^2(\pi r)p(-1+p)},$$

$$f := (p,r) \to -\frac{p(-1+r)(-r+p)}{(r^2-r)\sin(\pi p)^2 - 2\sin(\pi r)p(-1+r)\sin(\pi p) + p\sin(\pi r)^2(-1+p)} :$$

with (Optimization):

Find the numerical minimum near (0.25, 0.75) with constraints $0 < p < r < 1$

result := Minimize ($f(p,r)$, $\{p \leqslant 0.2, \quad r \leqslant p + 0.01, \quad r \leq 0.8\}$, initialpoint $= \{p = 0.25, \quad r = 0.75\}$) :

Extract the minimum point and value

min _p := rhs(result[2][1]); min _r := rhs(result[2][2]); min_value := result[1];
p := min _p; r := min _r; eval(u) #the last step to calculate weight$u$,
min _p = 0.371009648203599,
min _r = 0.628990351796402,
min_value = 0.219641801446212925,
u = 0.500000000214519.

C. Appendix 3

Maple Aids in obtaining the best linear combination of 7 new sample quantiles for estimating the scale parameter c

$$\text{restart: } fcov := (p,r) \to \frac{c^2 \cdot \pi^2 \cdot p \cdot (1-r)}{n \cdot \sin(p\pi) \cdot \sin(r\pi)} : mfcoV := (p,r) \to \frac{2c^2 \cdot \pi^2 \cdot p}{n \cdot \sin(p\pi) \cdot \sin(r\pi)} :$$

$$\text{VarE}\,7 := \text{simplify}\Big(u^2 \cdot mf\,\text{cov}(p,p) + v^2 \cdot mf\,\text{cov}(q,q) + w^2 \cdot mf\,\text{cov}(r,r) + (1 - 2u - 2v - 2w)^2 \cdot f\,\text{cov}\left(\frac{1}{2}, \frac{1}{2}\right) + 2uv \cdot mf\,\text{cov}(p,q) + 2uw \cdot mf\,\text{cov}(p,r) + u \cdot (1 - 2u - 2v \cdot (1 - 2u - 2v - 2w) \cdot mf\,\text{cov}\left(r, \frac{1}{2}\right)\Big) : \text{ fun } 7 := \text{simplify}\left(\frac{n}{c^2 \cdot \pi^2} \cdot \text{VarE}\,7\right) :$$

$$u := -\frac{1}{2\left(2p\csc(p\pi)^2 - 4p\csc(p\pi) + 1\right)}(4p\csc(p\pi)\csc(q\pi)v + 4p\csc(p\pi)\csc(r\pi)w - 4qv\csc(q\pi) - 4rw\csc(r\pi) - 4p\csc(p\pi)v^{-}4p\csc(p\pi)w + 2p\csc(p\pi) + 2v + 2w^{-}1),$$

$$v := \text{solve}\left(\frac{\partial}{\partial v}(\text{ fun } 7) = 0, v\right); \text{ fun } 7 := \text{simplify}(\text{ fun } 7) :$$

$$v := -\left(4p^2\csc(p\pi)^2\csc(q\pi)\csc(r\pi)w{-}4p\csc(p\pi)^2\csc(q\pi)q\csc(r\pi)w\right.$$
$$-4p^2\csc(p\pi)^2\csc(q\pi)w + 4p\csc(p\pi)^2\csc(q\pi)qw - 4\csc(r\pi)\csc(p\pi)^2p^2w$$
$$+4\csc(r\pi)\csc(p\pi)^2prw + 4p\csc(p\pi)\csc(q\pi)q\csc(r\pi)w - 4p\csc(p\pi)\csc(q\pi)\csc(r\pi)wr$$
$$+2p^2\csc(p\pi)^2\csc(q\pi) - 2p\csc(p\pi)^2\csc(q\pi)q + 4\csc(p\pi)^2p^2w - 4p\csc(p\pi)\csc(q\pi)qw$$
$$-4\csc(r\pi)\csc(p\pi)prw + 4\csc(q\pi)q\csc(r\pi)wr - 2\csc(p\pi)^2p^2 - 2\csc(p\pi)^2pw$$
$$+2p\csc(p\pi)\csc(q\pi)q + 2p\csc(p\pi)\csc(q\pi)w + 2p\csc(p\pi)\csc(r\pi)w$$
$$\left.-2\csc(q\pi)q\csc(r\pi)w + p\csc(p\pi)^2 - p\csc(p\pi)\csc(q\pi)\right)/\left(2\left(2p^2\csc(p\pi)^2\csc(q\pi)^2\right.\right.$$
$$-2p\csc(p\pi)^2\csc(q\pi)^2q - 4p^2\csc(p\pi)^2\csc(q\pi) + 4p\csc(p\pi)^2\csc(q\pi)q + 2\csc(p\pi)^2p^2$$
$$\left.\left.-4p\csc(p\pi)\csc(q\pi)q + 2\csc(q\pi)^2q^2 - p\csc(p\pi)^2 + 2p\csc(p\pi)\csc(q\pi) - \csc(q\pi)^2q\right)\right),$$

$$w := \text{solve}\left(\frac{\partial}{\partial w}(\text{ fun7 }) = 0, w\right);\ \text{fun7 } := \text{simplify(fun7) :}$$

$$w := \left(\sin(r\pi)p\left(2\sin(r\pi)pq - 2\sin(r\pi)q^2 - 2p\sin(q\pi)r + 2q\sin(q\pi)r - \sin(r\pi)p + \sin(r\pi)q\right.\right.$$
$$\left.\left.+p\sin(q\pi) - q\sin(q\pi) - 2pq + 2pr + 2q^2 - 2qr\right)\right)/\left(2\left(2\sin(r\pi)^2p^2q - 2\sin(r\pi)^2pq^2\right.\right.$$
$$+\sin(r\pi)^2pq - 2pq^2 + 2p^2q + 2rpq + 2pq\sin(p\pi)\sin(q\pi) + q\sin(p\pi)^2r + p\sin(q\pi)^2r$$
$$+2q^2\sin(p\pi)^2r - 2q\sin(p\pi)^2r^2 + 2p^2\sin(q\pi)^2r{-}2p\sin(q\pi)^2r^2 + 4p\sin(p\pi)\sin(q\pi)r^2$$
$$-2p\sin(p\pi)\sin(q\pi)r + 4\sin(r\pi)pq\sin(q\pi)r + 2\sin(r\pi)p^2\sin(q\pi) - 4\sin(r\pi)p^2q + 4\sin(r\pi)p^2r$$
$$+4\sin(r\pi)pq^2 - 4p\sin(p\pi)\sin(q\pi)qr - 2\sin(r\pi)pq\sin(q\pi) - 4\sin(r\pi)pqr{-}4\sin(r\pi)p^2\sin(q\pi)r$$
$$\left.\left.-2rp^2 - \sin(r\pi)^2p^2 - q^2\sin(p\pi)^2 - p^2\sin(q\pi)^2\right)\right),$$

$$divisions := 10 : min_val := 0. : max_val := 0.5 : step := (max_val - min_val)/divisions :$$

$$grid_points := [seq(i * step + min_val, i = 0..divisions)] : min_value := infinity : min_point := [0,0,0] :$$

```
for p_val in grid_points do
for q_val in grid_points do
for r_val in grid_points do
if 0 < p_val and p_val < q_val and q_val < r_val and r_val < 0.5 then
    current_value := eval(fun7, [p = p_val, q = q_val, r = r_val]):
    if type(current_value, numeric) then
        if current_value < min_value then
            min_value := current_value;
            min_point := [p_val, q_val, r_val]:
        end if;
    end if;
end if;
end do; end do; end do;

with(Optimization):
epsilon := 0.00001:
constraints := {epsilon <= p, r <= 0.5 - epsilon,
```

```
    p + epsilon <= q, q + epsilon <= r}:
fun7_num := unapply(fun7, [p, q, r]):
verification := NLPSolve(fun7_num(p, q, r), constraints,
    initialpoint = [p = min_point[1], q = min_point[2], r = min_point[3]],
    method = sqp):
if type(verification, list) and 2 < nops(verification) then
    opt_value := verification[1]:
    opt_point := verification[2]:
    p_opt := eval(p, opt_point):
    q_opt := eval(q, opt_point):
    r_opt := eval(r, opt_point):
    printf("\nOptimization verification results:\n"):
    printf("Minimum point: [p = %.6f, q = %.6f, r = %.6f]\n",
        p_opt, q_opt, r_opt):
    printf("Minimum function value: %.15f \n", opt_value):
else
    printf("\nOptimization verification failed"):
end if;
```

opt_value := 0.204659904317932712

opt_point := [p = 0.194667583229499, q = 0.309482928923083, r = 0.407633277770656]

p_opt := 0.194667583229499

q_opt := 0.309482928923083

r_opt := 0.407633277770656

$Optimizationverificationresults$:

$Minimumpoint$: [p = 0.194668, q = 0.309483, r = 0.407633]

$Minimumfunctionvalue$: 0.204659904317933 restart : $p := 0.194668 : q := 0.309483 : r := 0.407633 :$

$$
\begin{aligned}
w :=\ & \Big(\sin(r\pi)p\Big(8\sin(p\pi)\sin(q\pi)pqr - 4p^2r + 2\sin(r\pi)p^2 + 4pqr - \sin(p\pi)^2\sin(r\pi)q - 2\sin(q\pi)p^2\\
& + \sin(p\pi)^2\sin(q\pi)q + 2\sin(p\pi)^2\sin(r\pi)q^2 + 2\sin(p\pi)^2qr + \sin(p\pi)^2\sin(r\pi)p - \sin(p\pi)^2\sin(q\pi)p\\
& - 2\sin(p\pi)^2pr + 2\cos(p\pi)^2\sin(r\pi)pq + 2\sin(p\pi)^2\sin(q\pi)pr + 8\sin(p\pi)\sin(r\pi)p^2q\\
& - 8\sin(p\pi)\sin(r\pi)pq^2 + 4\sin(p\pi)\sin(r\pi)pq - 8\sin(p\pi)pqr - 4\sin(q\pi)pqr - 4\sin(p\pi)\sin(r\pi)p^2\\
& + 8\sin(p\pi)p^2r - 4\sin(r\pi)p^2q + 4\sin(r\pi)pq^2 + 4\sin(q\pi)p^2r - 4\sin(r\pi)pq + 2\sin(q\pi)pq\\
& - 2\sin(p\pi)^2\sin(q\pi)qr + 4\sin(q\pi)\sin(p\pi)p^2 + 2\sin(p\pi)^2pq - 8\sin(p\pi)p^2q + 8\sin(p\pi)pq^2\\
& +4p^2q - 4pq^2 - 2\sin(p\pi)^2q^2 - 8\sin(p\pi)\sin(q\pi)p^2r - 4\sin(p\pi)\sin(q\pi)pq\Big)\Big) \,/\, \Big(2\Big(4\sin(q\pi)^2p^3\\
& - 2\sin(p\pi)^2\sin(r\pi)\sin(q\pi)p^2 + 4\sin(p\pi)^2\sin(r\pi)p^2q - 4\sin(p\pi)^2\sin(r\pi)p^2r - 4\sin(p\pi)^2\sin(r\pi)pq^2\\
& + 8\sin(p\pi)\sin(r\pi)\sin(q\pi)p^3 - 16\sin(p\pi)\sin(r\pi)p^3q + 16\sin(p\pi)\sin(r\pi)p^3r\\
& + 16\sin(p\pi)\sin(r\pi)p^2q^2 + 8\sin(q\pi)\sin(r\pi)p^3r + 4\sin(q\pi)\sin(r\pi)p^2q + 8\sin(r\pi)p^2qr\\
& + 4\cos(r\pi)^2p^3q - 4\sin(p\pi)^2\sin(r\pi)\sin(q\pi)pqr + 16\sin(p\pi)\sin(r\pi)\sin(q\pi)p^2qr - 4p^2qr\\
& - 4\sin(q\pi)\sin(r\pi)p^3 + 8\sin(r\pi)p^3q - 8\sin(r\pi)p^3r - 8\sin(r\pi)p^2q^2 + 4\sin(r\pi)^2\cos(p\pi)^2p^2q\\
& - 2\sin(q\pi)\sin(p\pi)^3pq + 8\sin(q\pi)\sin(p\pi)^2p^2q - 4\sin(q\pi)\sin(p\pi)p^2q - 4\sin(p\pi)^3pq^2\\
& + 6\sin(p\pi)^2pq^2 + \sin(p\pi)^4q^2 - \sin(r\pi)^2\sin(p\pi)^2pq + 4\sin(r\pi)^2\sin(p\pi)p^2q + 2\sin(r\pi)^2\cos(q\pi)^2p^3\\
& + \sin(r\pi)^2\sin(p\pi)^2p^2 - 4\sin(r\pi)^2\sin(p\pi)p^3 + 8\sin(r\pi)^2p^2q^2 - 6\sin(r\pi)^2p^2q\\
& + 16\sin(p\pi)p^3q - 16\sin(p\pi)p^2q^2 - 8p^3q + 2\cos(q\pi)^2\sin(p\pi)^2p^2r - 8\cos(q\pi)^2\sin(p\pi)p^3r\\
& - 4p\sin(p\pi)^3r^2\sin(q\pi) + 8p\sin(p\pi)^3rq^2 - 8p\sin(p\pi)^3r^2q + 2\sin(p\pi)^2\sin(q\pi)^2pr^2\\
& + 16\sin(p\pi)^2\sin(q\pi)p^2r^2 - 8\sin(p\pi)\sin(q\pi)^2p^2r^2 + 2p\sin(p\pi)^3r\sin(q\pi) + 4p\sin(p\pi)^3rq\\
& - \sin(p\pi)^2\sin(q\pi)^2pr - 4\sin(p\pi)^2pq^2r + 4\sin(p\pi)^2pqr^2 + 4\sin(p\pi)\sin(q\pi)^2p^2r\\
& - 8\sin(p\pi)\sin(q\pi)p^2r^2 + 4\sin(p\pi)\sin(q\pi)p^2r + 4\cos(r\pi)^2\cos(p\pi)^2p^2q^2 + 4\cos(r\pi)^2\sin(p\pi)^2p^2q^2\\
& - 2\cos(r\pi)^2\sin(p\pi)^2p^2q - 2\cos(r\pi)^2\sin(p\pi)^2pq^2 - 8\sin(q\pi)\sin(p\pi)^2p^2r - 4\sin(p\pi)^2pqr\\
& + 8\sin(p\pi)p^2qr + 4p\sin(p\pi)^3r\sin(q\pi)q - 16\sin(p\pi)^2\sin(q\pi)p^2qr + 8\sin(p\pi)\sin(q\pi)p^2qr\\
& - 8\sin(p\pi)\cos(r\pi)^2p^3q + 8\sin(p\pi)\cos(r\pi)^2p^2q^2 - 2\cos(r\pi)^2\sin(q\pi)^2p^3 - 2q^2\sin(p\pi)^4r\\
& + 2q\sin(p\pi)^4r^2 + 4\cos^2(q\pi)^2p^3r - q\sin(p\pi)^4r + \sin(p\pi)^2\sin(q\pi)^2p^2 - 4\sin(p\pi)\sin(q\pi)^2p^3\\
& + 4\sin(q\pi)^2p^2r^2 - 2\sin(q\pi)^2p^2r + 4\sin(p\pi)^2\sin(r\pi)\sin(q\pi)p^2r - 16\sin(p\pi)\sin(r\pi)\sin(q\pi)p^3r\\
& + 2\sin(p\pi)^2\sin(r\pi)\sin(q\pi)pq + 4\sin(p\pi)^2\sin(r\pi)pqr - 8\sin(p\pi)\sin(r\pi)\sin(q\pi)p^2q\\
& - 16\sin(p\pi)\sin(r\pi)p^2qr - 8\sin(q\pi)\sin(r\pi)p^2qr\Big)\Big),
\end{aligned}
$$

$$w := 0.1750266863.$$

$$
\begin{aligned}
v := \Big(&4\csc(q\pi)w\csc(r\pi)q\sin(p\pi)^2 r - 2\csc(q\pi)w\csc(r\pi)q\sin(p\pi)^2 + 4\csc(q\pi)w\csc(r\pi)p\sin(p\pi)q \\
&- 4\csc(q\pi)prw\sin(p\pi)\csc(r\pi) + 4\csc(q\pi)p^2 w\csc(r\pi) - 4\csc(q\pi)w\csc(r\pi)pq \\
&- 4prw\sin(p\pi)\csc(r\pi) - 4\csc(q\pi)wp\sin(p\pi)q + 2\sin(p\pi)pw\csc(r\pi) - 4p^2 w\csc(r\pi) + 4prw\csc(r\pi) \\
&+ 2\csc(q\pi)p\sin(p\pi)q + 2\csc(q\pi)\sin(p\pi)pw^{-}4\csc(q\pi)p^2 w + 4\csc(q\pi)wpq - \csc(q\pi)p\sin(p\pi) \\
&+ 2\csc(q\pi)p^2 - 2\csc(q\pi)pq + 4p^2 w^{-}2p^2 - 2pw + p\Big) \Big/ \Big(2\Big(-2\csc(q\pi)^2\sin(p\pi)^2 q^2 \\
&+ \csc(q\pi)^2 q\sin(p\pi)^2 + 2p^2\cot(q\pi)^2 - 4\csc(q\pi)^2 p^2 + 2\csc(q\pi)^2 pq + 4\csc(q\pi)p\sin(p\pi)q \\
&- 2\csc(q\pi)p\sin(p\pi) + 4\csc(q\pi)p^2 - 4\csc(q\pi)pq + p)\Big),
\end{aligned}
$$

$$v := 0.1436903088,$$

$$
\begin{aligned}
u := -\frac{1}{2\left(2p\csc(p\pi)^2 - 4p\csc(p\pi) + 1\right)}&(4p\csc(p\pi)\csc(r\pi)w + 4p\csc(p\pi)\csc(q\pi)v \\
&-4rw\csc(r\pi) - 4qv\csc(q\pi) - 4p\csc(p\pi)v - 4p\csc(p\pi)w + 2p\csc(p\pi) + 2v + 2w - 1),
\end{aligned}
$$

$$u := 0.08864913264,$$

$$1 - 2u - 2v - 2w,$$

$$0.1852677445.$$

AIMS Press