

Importance of Cauchy Distribution in the Field of Statistics



Name: Anik Chakraborty

Roll No.: 443

Registration No.: A01-1112-0838-19

Paper Code: HSTD6043D

Department: Statistics

Session: 2019 - '22

Supervisor: Prof. Ayan Chandra

St. Xavier's College (Autonomous), Kolkata

Declaration

I affirm that I have identified all my sources and that no part of my dissertation paper uses unacknowledged materials.

Student's Signature: *Anik Chakraborty*

Date: April 12, 2022

Contents

Introduction	1
1 Research Idea	2
1.1 Background	2
1.2 Importance of the Study	3
1.3 Objectives of the Study	3
1.4 Scheme of Work	3
2 Literature Review	3
3 Cauchy Distribution	4
3.1 Description	4
3.1.1 Genesis	4
3.1.2 Background for developing Cauchy distribution	5
3.2 Probability Laws	5
3.2.1 Probability Density Function	5
3.2.2 Cumulative Distribution Function	5
3.2.3 Visualizing PDF and CDF	6
3.3 Interesting Properties	7
3.3.1 Natrure of the Curve	7
3.3.2 Non-existence of Moments	8
3.3.3 Quantile Measures	9
3.3.4 Mode	10

3.3.5	Points of Inflection	10
3.4	Sampling Properties	11
4	Cauchy and Normal Distribution	14
4.1	Why do we need Cauchy apart from Normal?	14
4.2	Comparison of Cauchy & Normal Distribution	14
4.3	A Simulation Study	15
5	Estimation under Cauchy Population	16
5.1	Unbiasedness	16
5.2	Consistency	18
5.3	Cramer-Rao Inequality	19
5.4	Maximum Likelihood Estimation	20
6	Conclusion	21
7	Appendix	22
7.1	Additive Property of Cauchy Distribution	22
7.2	Regularity Conditions	23
7.3	Expectation and Symmetry	23
7.4	R Codes	24
Acknowledgement		28
References		29

Plots

1	Nature of the curves of the form $\frac{1}{x^2+a^2}$	4
2	PDF of Cauchy distribution	6
3	CDF of Cauchy distribution	6
4	Shape of the $C(0, 1)$ PDF	7
5	Truncated and Standard Cauchy	9
6	Points of Inflection of Standard Cauchy	11
7	X and $\frac{1}{X}$ are Identically Distributed	12
8	$X+Y$ and $2X$ are Identically Distributed	13
9	Densities of Cauchy and Normal	14
10	Comparison of Sample Mean and Sample Median	18

Codes

1	Cauchy and Normal	24
2	Convergence in Cauchy	25

Introduction

Cauchy distribution is an interesting and widely practiced concept in Statistical field. The notion of this distribution comes from the field of Physics and many mathematicians were also involved in developing this distribution. Thus, it is not that the distribution has come solely from Statistical background.

In many applications of Physics Cauchy distribution is widely used, in Economical development and modeling there is a huge impact of this distribution. Uses of Cauchy distribution being multidisciplinary make it so popular. It has been used in many applications such as mechanical and electrical theory, physical anthropology, measurement problems, risk and financial analysis. It was also used to model the points of impact of a fixed straight line of particles emitted from a point source.

Some interesting characteristics make this distribution unique and also make people attracted towards this theoretical distribution. Here, the properties have been studied using analytical and graphical approach. Simulation is done as and when needed. Thus, we have tried to study the Cauchy distribution in a descriptive way.

1 Research Idea

1.1 Background

Cauchy distribution is a well known theoretical distribution which also has a major importance in the statistical study aspects. Let us look what is actually meant by a theoretical distribution and the purpose of developing the notions of probability distributions.

What is a Theoretical Distribution?

In order to simplify the mathematical treatment of population distributions, we take the PMF or PDF to be a sufficiently simple form. Distributions defined in this way are called theoretical distributions because, they are ideal distributions that are hardly expected to reflect in toto the true nature of the population distribution. They are meant simply to give a fairly close approximation to the actual distribution of the variable in the population.

Why do we need to study these?

In practice, we never observe probabilities - what we observe are nothing but the frequency distributions. Once we have, a frequency distribution at our disposal, we can readily construct the corresponding relative frequency distribution. These relative frequency distributions are good approximations of the corresponding probabilities, provided that the total frequency is large enough.

We have on the other hand, some theoretically developed probability distributions, known as **theoretical distributions** at our disposal. Once, we have a relative frequency distribution, we try to ‘match’ it with these theoretical distributions. If it matches well with one of these theoretical distributions then the properties of that distribution may be applied to our relative frequency distribution as well. In fact, in this situation the data we have may be regarded as a sample drawn from that theoretical distribution.

1.2 Importance of the Study

The importance of the study lies in the fact that, Cauchy is a heavy-weight distribution in Statistical and also in many other fields. To study its characteristics will help us improve and make a clear idea about further scopes of this distribution.

1.3 Objectives of the Study

The objective of the study is to grow more attention towards the Cauchy distribution which indirectly helps the Statisticians in inferential aspects and to aware people that normality assumptions may not hold good always.

1.4 Scheme of Work

We first studied the different properties of this well known distribution that paves the path for further studies such as comparison with Normal, different estimation procedures under Cauchy population.

2 Literature Review

In the history of Statistics, there has been a lot of great works involving the Cauchy distribution. Notable mathematicians, physicists, statisticians and people from other fields tried to study the nature of this popular probability distribution.

Daniel Bloch in his article ‘A Note on the Estimation of the Location Parameter of the Cauchy Distribution’ described different procedures for estimating the location parameter of the Cauchy distribution. VD Barnett proposed the different estimators using the order statistics for Cauchy’s location parameter in his article ‘Order Statistics Estimators of the Location of the Cauchy Distribution’.

3 Cauchy Distribution

3.1 Description

3.1.1 Genesis

The Cauchy distribution, named after the French mathematician Augustin-Louis Cauchy, is a continuous probability distribution specially known for its **thicker tails** than the normal curve. The density function: $\frac{1}{\pi(1+x^2)}$, commonly known as Cauchy density or more evidently the curves proportional to $\frac{1}{x^2+a^2}$ have been studied in the mathematical world for over three centuries.

In this aspect the Cauchy density can be thought of as a special case of ‘Witch of Agnesi’ derived from the name of Italian mathematician Maria Agnesi. She had discussed several properties of Cauchy curve and referred to it as ‘Witch’. Several notable persons including Pierre De Fermat, Sir Isaac Newton, Guido Grandi, Gottfried Leibniz also studie the nature and behaviour of this curve.

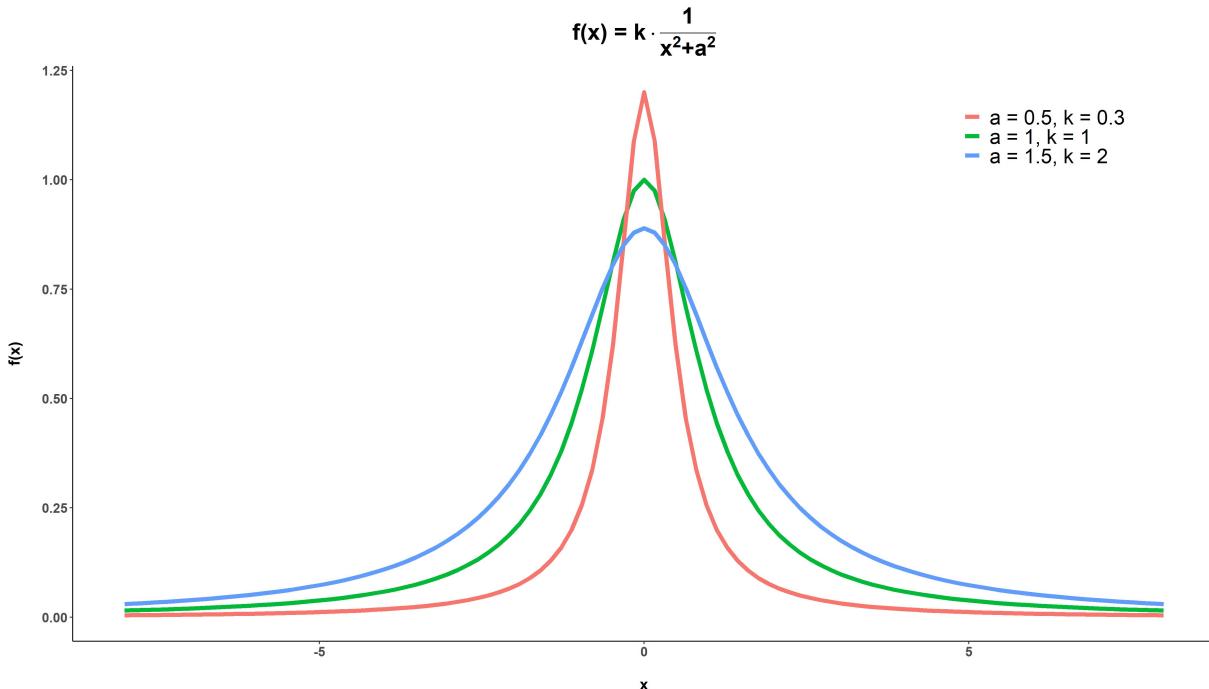


Figure 1: Nature of the curves of the form $\frac{1}{x^2+a^2}$

3.1.2 Background for developing Cauchy distribution

The idea of Cauchy distribution was initially developed by French mathematician and physicist Siméon Denis Poisson who observed that the distribution with density $\frac{1}{\pi(1+x^2)}$ has some peculiar properties. French mathematician Legendre introduced the famous and widely used ‘Principle of Least Squares’ with the assumption of “normally distributed errors”. Poisson showed Laplace’s large sample justification of Legendre’s Least Square Theory is not valid by presenting the characteristic function of the Cauchy distribution.

The distribution became associated with Augustin-Louis Cauchy as he responded to the criticism of Bienaym 's article showing Legendre’s least-squares fails to provide “the most probable results” in case of a interpolation method suggested by Cauchy as it does in case of normality of errors.

3.2 Probability Laws

Cauchy distribution is a theoretical distribution of an absolutely continuous random variable. It is mostly used for those random variables which are likely to produce extreme values. The Cauchy distribution with parameters ‘ μ ’ and ‘ σ ’ is denoted by $C(\mu, \sigma)$.

3.2.1 Probability Density Function

The Probability Density Function (PDF) of $C(\mu, \sigma)$ is given by,

$$f(x) = \frac{1}{\pi} \frac{\sigma}{\{\sigma^2 + (x - \mu)^2\}}, \quad x \in \mathbb{R}$$

3.2.2 Cumulative Distribution Function

The Cumulative Distribution Function (CDF) of $C(\mu, \sigma)$ is given by,

$$F(x) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1} \left(\frac{x - \mu}{\sigma} \right), \quad x \in \mathbb{R}$$

Here, $\mu \in \mathbb{R}$ is the location parameter

$\sigma > 0$ is the scale parameter.

NOTE: Here, μ and σ are not the expectation and variance of the distribution. Though, we can think of a standardized Cauchy distribution by putting, $\mu = 0$ and $\sigma = 1$.

3.2.3 Visualizing PDF and CDF

- The PDF of $C(\mu, \sigma)$ for different choices of μ and σ is plotted below:

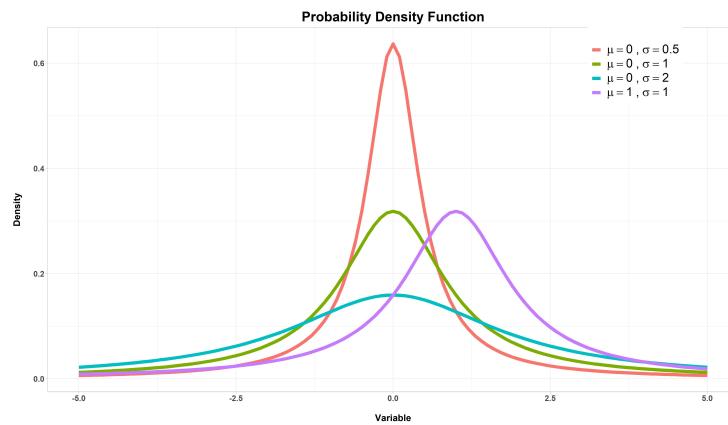


Figure 2: PDF of Cauchy distribution

- The CDF of $C(\mu, \sigma)$ for different choices of μ and σ is plotted below:

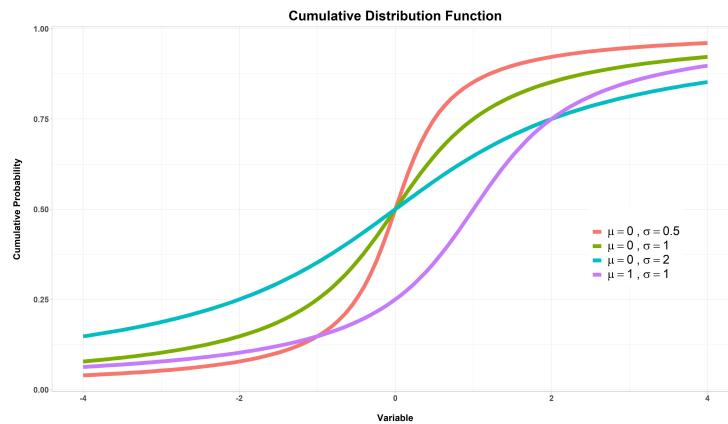


Figure 3: CDF of Cauchy distribution

3.3 Interesting Properties

3.3.1 Nature of the Curve

The values of $f(x) = \frac{1}{\pi} \frac{\sigma}{\{\sigma^2 + (x - \mu)^2\}}$ are plotted against the values of x , we get a graph as shown below -

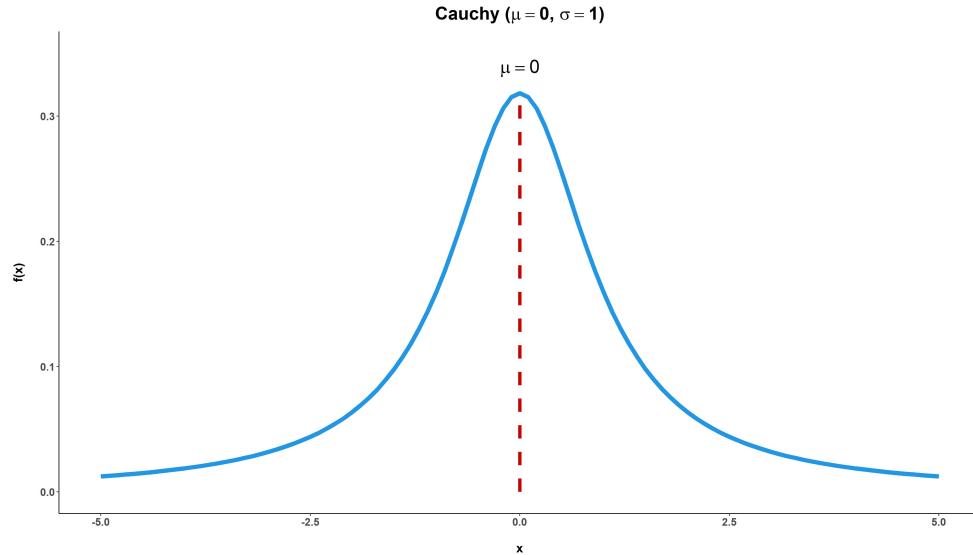


Figure 4: Shape of the $C(0, 1)$ PDF

From ‘Figure 4’, we observe that the graph is bell-shaped, symmetric about μ and the tails of the distribution are much thicker.

Proof of Symmetry

A continuous random variable X with PDF, $f(x)$ is said to have a symmetric distribution about the line ‘ $x = a$ ’ if,

$$f(a - x) = f(a + x), \forall x \in \mathbb{R}$$

Here, $f(\mu - x) = \frac{1}{\pi} \frac{\sigma}{\{\sigma^2 + (\mu - x)^2\}} = f(\mu + x), \forall x \in \mathbb{R}$

Hence, $C(\mu, \sigma)$ is symmetric about the line ‘ $x = \mu$ ’ i.e. about its location parameter.

3.3.2 Non-existence of Moments

To check the existence of r^{th} order moment of $C(\mu, \sigma)$, let us first find -

$$\begin{aligned}
E\left(\left|\frac{X-\mu}{\sigma}\right|^r\right) &= \int_{-\infty}^{\infty} \left|\frac{x-\mu}{\sigma}\right|^r \cdot \frac{1}{\pi \{\sigma^2 + (x-\mu)^2\}} dx \\
&= \frac{1}{\pi} \int_{-\infty}^{\infty} |z|^r \cdot \frac{1}{(1+z^2)} dz && \left[\text{substituting } \frac{x-\mu}{\sigma} = z \right] \\
&= \frac{2}{\pi} \int_0^{\infty} \frac{z^r}{1+z^2} dz && [\because \text{integrand is an even function of } z] \\
&= \frac{1}{\pi} \int_0^{\infty} \frac{u^{\frac{r+1}{2}-1}}{(1+u)^{\frac{r+1}{2}+\frac{1-r}{2}}} du && [\text{substituting } z^2 = u]
\end{aligned}$$

which converges, if and only if, $r+1 > 0$ and $1-r > 0$, i.e. $-1 < r < 1$

Hence, $E\left(\left|\frac{X-\mu}{\sigma}\right|^r\right)$ does not exist for $r \geq 1$.

As such, for Cauchy distribution the moments of order $r \geq 1$ do not exist. Consequently, the MGF of Cauchy distribution does not exist.

Remark

Let us consider $C(0, 1)$ here. For some $0 < k < \infty$,

$$\begin{aligned}
E(|X|^r) &= \int_{-\infty}^{\infty} \frac{|x|^r}{\pi(1+x^2)} dx \\
&= \int_{|x|<k} \frac{|x|^r}{\pi(1+x^2)} dx + \int_{|x|\geq k} \frac{|x|^r}{\pi(1+x^2)} dx \\
&= I_1 + I_2 && \dots\dots (*)
\end{aligned}$$

Note that, as the tails of Cauchy are thicker the propensity of producing extreme values ($|x| \geq k$) is high in this case i.e. the problem of ‘non-existence of moments’ is associated with the integral I_2 .

- **Truncated Cauchy Distribution**

We can get rid of the above problem if we consider the truncated symmetric Cauchy distribution i.e. if we take the support of the distribution to be $-k < x < k$ for some $0 < k < \infty$. From (*) we can readily observe that, $E(|X|^r)$ exists and converges to a finite value as I_2 vanishes.

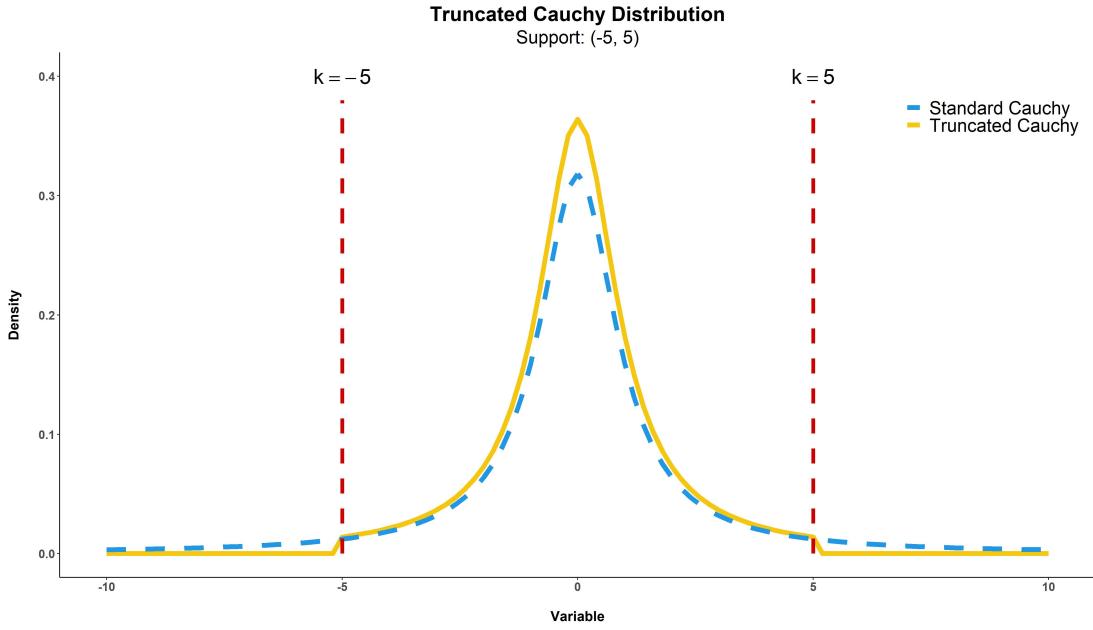


Figure 5: Truncated and Standard Cauchy

Hence, all the moments exist for a truncated Cauchy distribution (truncated over a finite interval) and for this reason Cauchy and truncated Cauchy is interesting in the statistical field.

3.3.3 Quantile Measures

Let, ξ_p be the p^{th} quantile of the $C(\mu, \sigma)$ distribution. Then,

$$\begin{aligned}
 F(\xi_p) &= p \\
 \implies \frac{1}{2} + \frac{1}{\pi} \tan^{-1} \left(\frac{\xi_p - \mu}{\sigma} \right) &= p \\
 \implies \xi_p &= \mu + \sigma \tan \left[\pi \left(p - \frac{1}{2} \right) \right] \quad \dots\dots (**)
 \end{aligned}$$

From (**) -

For $p = \frac{1}{2}$, median $= \xi_{\frac{1}{2}} = \mu$ and for $p = \frac{1}{4}$, $Q_1 = \xi_{\frac{1}{4}} = \mu - \sigma$

From symmetry, $Q_3 = \mu + \sigma$

The quartile deviation is $\frac{Q_3 - Q_1}{2} = \sigma$

3.3.4 Mode

From ‘Figure 4’, we can see that the Cauchy distribution is unimodal and the mode is at $x = \mu$.

The mathematical justification is as follows:

$$f(x) = \frac{1}{\pi} \frac{\sigma}{\{\sigma^2 + (x - \mu)^2\}}$$

$$f'(x) = -\frac{2(x - \mu)}{\sigma^2 + (x - \mu)^2} \cdot f(x)$$

Now, $f'(x) \geq 0$ according as $x \leq \mu$ since $f(x) > 0, \forall x$

Hence, the mode of the distribution is at $x = \mu$ and the modal value is $f(\mu) = \frac{1}{\pi\sigma}$.

3.3.5 Points of Inflection

Definition: The point on a curve where it changes from concavity to convexity or vice-versa is called a point of inflection.

In other words, at the points of inflection, the rate of change of slope of that curve changes. Now,

$$f''(x) = \frac{3(x - \mu)^2 - \sigma^2}{[\sigma^2 + (x - \mu)^2]^2} \cdot f(x) \begin{cases} > 0, & \text{if } x < \mu - \frac{\sigma}{\sqrt{3}} \\ < 0, & \text{if } \mu - \frac{\sigma}{\sqrt{3}} \leq x \leq \mu + \frac{\sigma}{\sqrt{3}} \\ > 0, & \text{if } x > \mu + \frac{\sigma}{\sqrt{3}} \end{cases}$$

Hence, the points of inflection of the Cauchy distribution are $\left(\mu - \frac{\sigma}{\sqrt{3}}, \frac{3}{4\pi\sigma}\right)$ and $\left(\mu + \frac{\sigma}{\sqrt{3}}, \frac{3}{4\pi\sigma}\right)$.

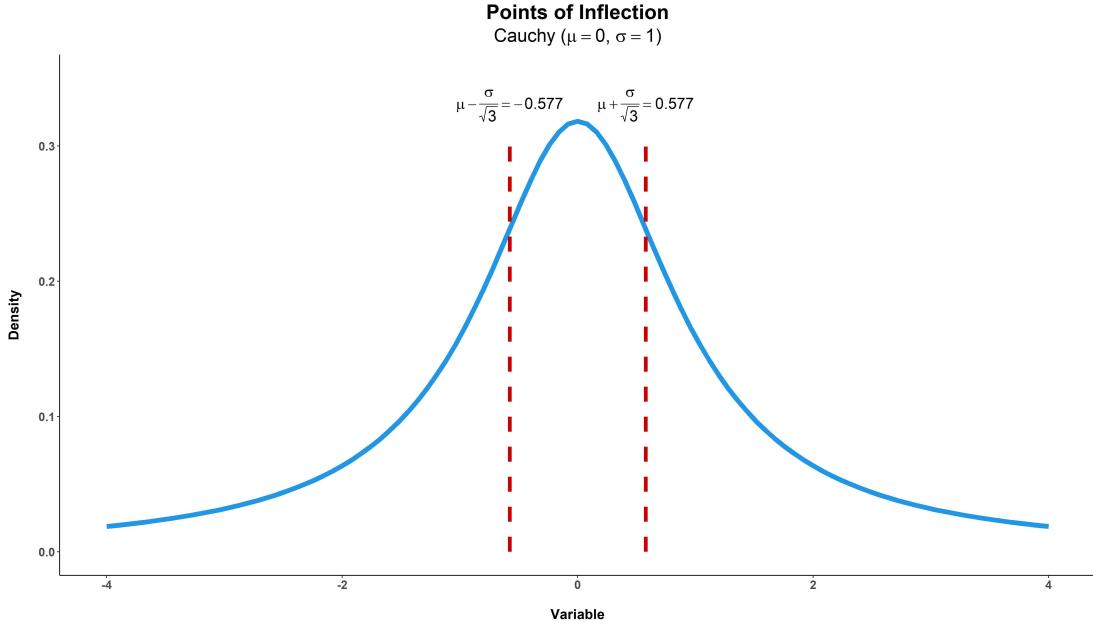


Figure 6: Points of Inflection of Standard Cauchy

For $C(0, 1)$ the points of inflection are $(-0.577, 0.239)$ and $(0.577, 0.239)$ i.e. at these points the curve changes from concavity to convexity and vice-versa.

3.4 Sampling Properties

- $X \sim C(0, 1) \implies \frac{1}{X} \sim C(0, 1)$

Proof. Let, $g(\cdot)$ be the PDF of $Y = \frac{1}{X}$.

Jacobian of the transformation is $\left| \frac{\partial x}{\partial y} \right| = \frac{1}{y^2}$ and $y \in \mathbb{R}$

Then, the PDF of Y is given by,

$$\begin{aligned}
 g(y) &= f\left(\frac{1}{y}\right) \cdot \left| \frac{\partial x}{\partial y} \right|, \quad y \in \mathbb{R} & [f(\cdot) \text{ is the PDF of } X] \\
 &= \frac{1}{\pi \left(1 + \frac{1}{y^2}\right)} \cdot \frac{1}{y^2}, \quad y \in \mathbb{R} \\
 &= \frac{1}{\pi(1+y^2)}, \quad y \in \mathbb{R}
 \end{aligned}$$

Hence, if X is a random variable having a $C(0, 1)$ distribution then $\frac{1}{X}$ also have the identical distribution. \square

The above property is visualized below:

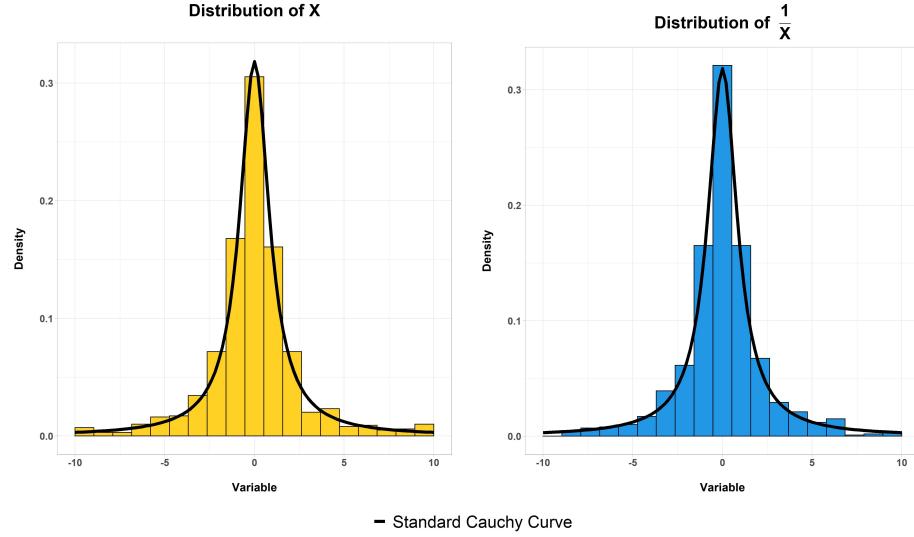


Figure 7: X and $\frac{1}{X}$ are Identically Distributed

- $X, Y \stackrel{iid}{\sim} C(0, 1) \implies X + Y \stackrel{D}{=} 2X \sim C(0, 2)$

Proof. Let, $F(\cdot)$ be the common CDF of X, Y .

Now, using Additive property of Cauchy distribution (proof given in Appendix 7.1),

$$X + Y \sim C(0, 2) \quad \dots (i)$$

Again, let $G(\cdot)$ be the CDF of $U = 2X$. Then, for $u \in \mathbb{R}$

$$\begin{aligned} G(u) &= P(U \leq u) \\ &= P(2X \leq u) \\ &= F\left(\frac{u}{2}\right) \\ &= \frac{1}{2} + \frac{1}{\pi} \tan^{-1}\left(\frac{u}{2}\right) \end{aligned}$$

Hence, $2X \sim C(0,2)$... (ii)

From (i) and (ii), $X + Y \stackrel{D}{=} 2X \sim C(0,2)$

□

This property does not hold in general for any distribution. As such, this is a unique and an interesting property of Cauchy distribution. Also, we can justify from this property that Cauchy distribution has no finite variance.

[Justification

If possible, let $Var(X) = Var(Y) = \sigma^2 < \infty$ [Here, $\sigma^2 > 0$ as Cauchy is not degenerated]

Again, we have $X + Y \stackrel{D}{=} 2X$.

Then, $X + Y$ and $2X$ will have the same variance which gives us $4\sigma^2 = 2\sigma^2$.

Now, this is only possible if: $\sigma^2 = +\infty$ or $\sigma^2 = 0$

Hence, $\sigma^2 = +\infty$ ($\because \sigma^2 > 0$).

This justifies that the fact that the variance of Cauchy is not finite.]

The above property is visualized below:

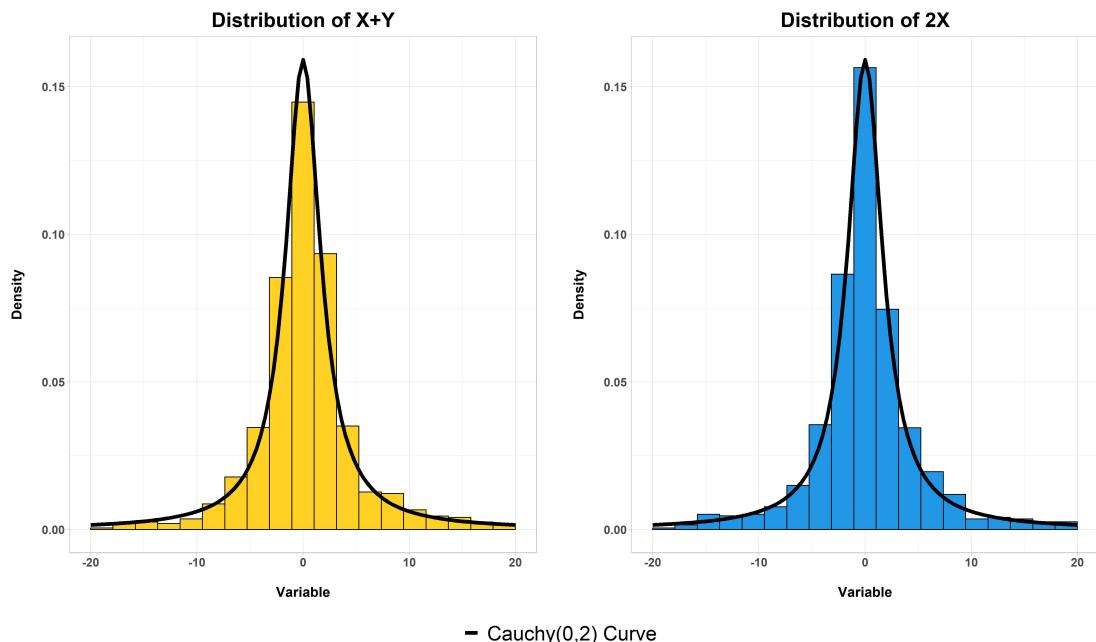


Figure 8: **X+Y and 2X are Identically Distributed**

4 Cauchy and Normal Distribution

4.1 Why do we need Cauchy apart from Normal?

In Parametric Inference, we usually use Normal as our Parent distribution and derive test rules, confidence intervals based on the assumption that normality holds everywhere. But in practice, this is not the case always as in most of the real life problems ‘Normality Assumption’ will lead us to erroneous conclusions.

Cauchy distribution is such a distribution which may not be used frequently to deal with the real world problems but can make us understand the fact that assumption of normal laws cannot yield satisfactory results always, in the domain of Parametric Inference only. That is why, the Cauchy distribution is so much important in inferential statistics apart from the Normal distribution.

4.2 Comparison of Cauchy & Normal Distribution

Let us consider two random variables X and Y where, $X \sim N(0, 1)$ and $Y \sim C(0, 1)$. Their PDFs are respectively given by, $f_X(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$ and $f_Y(t) = \frac{1}{\pi(1+t^2)}$, $t \in \mathbb{R}$. Then,

$$\begin{aligned} \text{(a)} \quad & \max_{t \in \mathbb{R}} f_X(t) = \frac{1}{\sqrt{2\pi}} = f_X(0) \text{ and } \max_{t \in \mathbb{R}} f_Y(t) = \frac{1}{\pi} = f_Y(0) \\ & \implies f_X(0) > f_Y(0) \end{aligned}$$

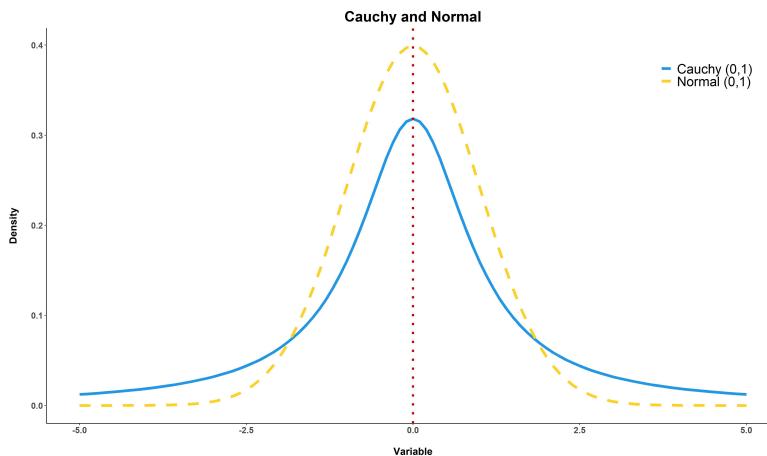


Figure 9: Densities of Cauchy and Normal

(b) Both Cauchy and Normal distribution are symmetric about ‘0’.

But $E(X) = \text{median} = \text{mode} = 0$ for X whereas $\text{median} = \text{mode} = 0$ but $E(X)$ does not exist for Y .

(c) $f_X(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$ decreases to ‘0’ more rapidly than $f_Y(t) = \frac{1}{\pi(1+t^2)}$, as $|t| \rightarrow \infty$.

(d) Here, from ‘Figure 9’ observe that, $P(|X| > k) < P(|Y| > k)$, $\forall k > 0$.

As Cauchy has thick tails compared to Normal distribution and this is the reason of non-existence of moments of Cauchy distribution (discussed in subsection 3.3.2).

4.3 A Simulation Study

Objective

Here, for comparison, we ran a simulation to check how the empirical level behaves if we draw samples from Cauchy and consider a testing problem on the basis of normality assumption.

Results and Comments

The results are given in the following table -

Table 1: Empirical Levels

Sample Size	Cauchy	Normal
5	0.529	0.037
10	0.661	0.054
25	0.745	0.055
50	0.833	0.046
100	0.878	0.055
200	0.908	0.051
500	0.949	0.042
1000	0.960	0.050

From ‘Table 1’, it can be observed that, under normality assumptions, the Cauchy distribution is behaving badly.

5 Estimation under Cauchy Population

5.1 Unbiasedness

We know that, ‘sample mean is an unbiased estimator of the population mean’ but in case of Cauchy distribution the population mean itself does not exist. So, we cannot use the fact here. Instead of this, for Cauchy distribution sample median can be used as a suitable estimate of the location parameter.

Let us consider a random sample of size $n = 2k + 1, k \in \mathbb{Z}$ from $C(\mu, \sigma)$ distribution. The population distribution is symmetric about μ (proof in subsection 3.3.1).

Hence, $f(\mu - x) = f(\mu + x)$ and $F(\mu - x) = 1 - F(\mu + x), \forall x \in \mathbb{R}$.

The PDF of $(k+1)^{\text{th}}$ order statistics is given by,

$$f_{X_{(k+1)}}(x) = \frac{(2k+1)!}{k! k!} \{F(x)\}^k f(x) \{1 - F(x)\}^k, \quad x \in \mathbb{R}$$

Note that, $X_{(k+1)}$ is the sample median of the sample of size $(2k+1)$. Now,

$$\begin{aligned} f_{X_{(k+1)}}(\mu - x) &= \frac{(2k+1)!}{k! k!} \{F(\mu - x)\}^k f(\mu - x) \{1 - F(\mu - x)\}^k \\ &= \frac{(2k+1)!}{k! k!} \{1 - F(\mu + x)\}^k f(\mu + x) \{F(\mu + x)\}^k \\ &= f_{X_{(k+1)}}(\mu + x) \end{aligned}$$

Hence, the distribution of the sample median of odd sample size is symmetric about $x = \mu$. Also, the quantile measures exist for all the distributions and hence $E[X_{(k+1)}]$ exists and is finite.

We have for any symmetric distribution, mean of that distribution is the point of symmetry, if the expectation is finite. (proof in appendix 7.3)

Hence, here $E[X_{(k+1)}] = \mu$

Thus, if the sample size is odd then the sample median is an unbiased estimator of the location parameter for Cauchy distribution.

Sample Mid-range

Here, X_i 's have distributions symmetric about μ . Hence, for $i = 1, 2, \dots, n$,

$$\begin{aligned} P(\mu - X_i \leq y) &= P(X_i \geq \mu - y) \\ &= P(X_i \leq \mu + y) && [\text{Due to symmetry}] \\ &= P(X_i - \mu \leq y) \end{aligned}$$

Thus, $\mu - X_i \stackrel{D}{=} X_i - \mu$, $i = 1, 2, \dots, n$.

Hence, $\max_{i=1(1)n} \{\mu - X_i\} \stackrel{D}{=} \max_{i=1(1)n} \{X_i - \mu\}$ i.e. $\mu - X_{(1)} \stackrel{D}{=} X_{(n)} - \mu$

Hence, $E[\mu - X_{(1)}] = E[X_{(n)} - \mu]$

Thus, $E\left[\frac{X_{(1)} + X_{(n)}}{2}\right] = \mu$.

Hence, for Cauchy distribution sample mid-range can also be considered as an unbiased estimator of the location parameter.

Some Other Estimators

'Quick Estimators' can be considered as an estimator of the location parameter of the Cauchy distribution. A quick estimator is the weighted average (can be thought as a linear combination) of some order statistics.

Another good estimator of the location parameter of Cauchy is the trimmed mean. It used the central 24% of the sample order statistics and can be considered as a better estimator than the sample median.

In the above cases, our target is to minimize the Asymptotic Relative Efficiency (A.R.E.) of the concerned estimator with respect to the best estimator.

5.2 Consistency

As the Expectation of Cauchy does not exist, we cannot use Khintchine's Weak Law of Large Numbers to say that 'Sample Mean is a consistent estimator of the Population Mean'. In case of Cauchy, sample median, as a robust estimator becomes consistent for the location parameter of the Cauchy distribution.

Let, us visualize this graphically:

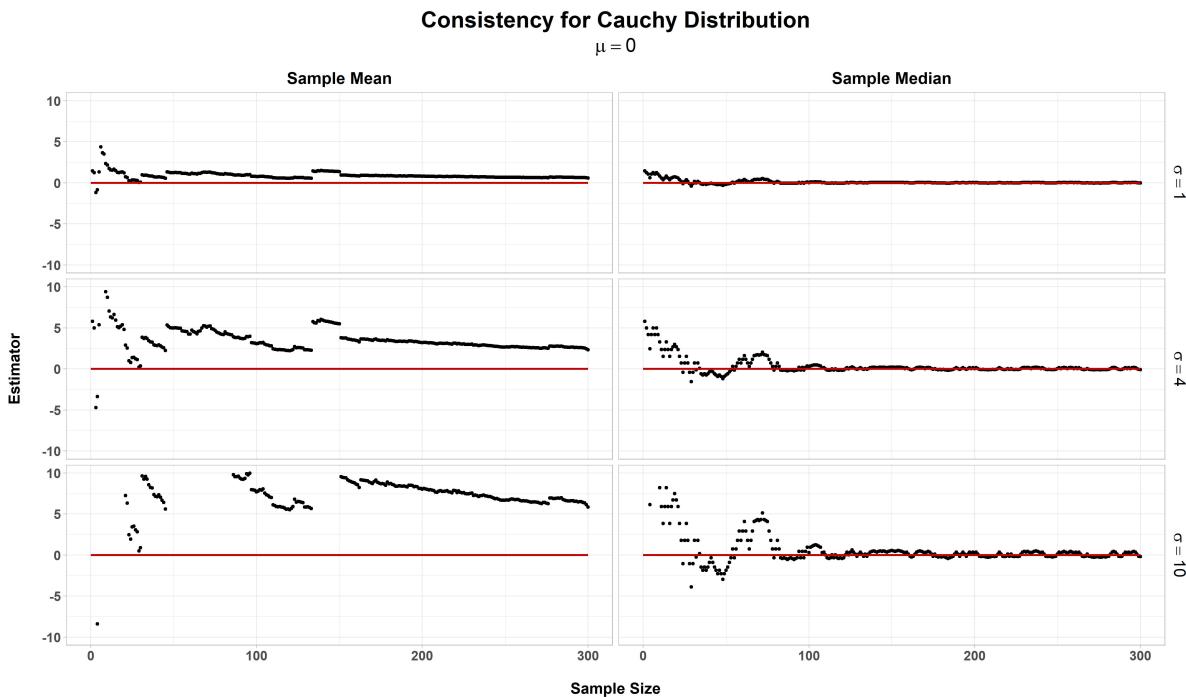


Figure 10: Comparison of Sample Mean and Sample Median

From 'Figure 10' we observe that, sample median converges to μ in probability, more rapidly than sample mean which indicates that Sample median is better in terms of consistency than sample mean in case of Cauchy.

Also, note that as we increase the scale parameter, more sample size is required to converge in probability for sample median.

5.3 Cramer-Rao Inequality

Consider, the family $\{C(\theta, 1) : \theta \in \mathbb{R}\}$ of Cauchy distributions which satisfies all the regularity conditions required for Cramer-Rao Inequality. (The conditions are given in appendix 7.2).

$$\text{Here, } f(x; \theta) = \frac{1}{\pi\{1 + (x - \theta)^2\}}, \quad x \in \mathbb{R}, \theta \in \mathbb{R}$$

The support $S = \mathbb{R}$ is independent of θ and the parameter space $\Omega = \mathbb{R}$ is an open interval. It also satisfies the other regularity conditions.

Hence, the Cramer-Rao Inequality exists for $\{C(\theta, 1) : \theta \in \mathbb{R}\}$.

\therefore The Cramer-Rao Inequality for the variance of an unbiased estimator T of θ is given by,

$$Var(T) \geq \frac{1}{I_n(\theta)} = \frac{1}{nI_1(\theta)}$$

$$\text{Now, } \frac{\partial}{\partial \theta} \ln f(x; \theta) = \frac{2(x - \theta)}{1 + (x - \theta)^2}$$

$$\begin{aligned} I_1(\theta) &= E \left[\frac{\partial}{\partial \theta} \ln f(x; \theta) \right]^2 \\ &= \int_{-\infty}^{\infty} \frac{4(x - \theta)^2}{\{1 + (x - \theta)^2\}^2} \cdot \frac{dx}{\pi\{1 + (x - \theta)^2\}} \\ &= \frac{4}{\pi} \int_{-\infty}^{\infty} \frac{u^2}{(1+u)^3} du && [u = x - \theta] \\ &= \frac{8}{\pi} \int_0^{\infty} \frac{u^2}{(1+u)^3} du && [u = x - \theta] \\ &= \frac{4}{\pi} \int_0^{\infty} \frac{t^{\frac{3}{2}-1}}{(1+t)^{\frac{3}{2}+\frac{3}{2}}} dt && [t = u^2] \\ &= \frac{4}{\pi} \cdot B\left(\frac{3}{2}, \frac{3}{2}\right) \\ &= \frac{1}{2} \end{aligned}$$

Hence, the Cramer-Rao Inequality reduces to,

$$Var(T) \geq \frac{2}{n}$$

As $\{C(\theta, 1) : \theta \in \mathbb{R}\}$ does not belong to One Parameter Exponential Family, the equality does not hold in Cramer-Rao Inequality i.e. there is no MVUE for any parametric function $\psi(\theta)$.

Here, $\{C(\theta, 1) : \theta \in \mathbb{R}\}$ is a regular family in the sense that Cramer-Rao Inequality holds for this family but the Cramer-Rao Lower Bound is not an attainable lower bound.

5.4 Maximum Likelihood Estimation

Suppose, X_1, X_2, \dots, X_n be a random sample of size n from $C(\theta, 1)$ distribution. To find the Maximum Likelihood Estimator of the parameter θ on the basis of the above sample:

The PDF of X is given by,

$$f(x) = \frac{1}{\pi \{1 + (x - \theta)^2\}}, \quad x \in \mathbb{R}, \theta \in \mathbb{R}$$

Given $(X_1, X_2, \dots, X_n) = (x_1, x_2, \dots, x_n)$ the likelihood function of θ is given by,

$$L(\theta | x) = \frac{1}{\pi^n \prod_{i=1}^n \{1 + (x_i - \theta)^2\}}, \quad \theta \in \mathbb{R}$$

The log-likelihood function is given by,

$$\ln L(\theta) = -n \ln \pi + \sum_{i=1}^n \ln \{1 + (x_i - \theta)^2\}$$

Let, $\hat{\theta}$ be the MLE of θ . Hence, we have,

$$\left. \frac{\partial}{\partial \theta} \ln L(\theta) \right|_{\theta=\hat{\theta}} = \sum_{i=1}^n \left\{ \frac{2(x_i - \theta)}{1 + (x_i - \theta)^2} \right\} \Big|_{\theta=\hat{\theta}} = 0 \quad \dots (I)$$

Now, we will find $\hat{\theta}$ iteratively as a root of the equation (I).

Let, $\theta_0 = \text{median}(x)$ is an initial solution of (I). By Taylor's expansion we have,

$$\frac{\partial}{\partial \theta} \ln L(\theta) \Big|_{\theta=\theta_0} + (\hat{\theta} - \theta_0) \frac{\partial^2}{\partial \theta^2} \ln L(\theta) \Big|_{\theta=\theta_0} \simeq 0 \quad [\text{we ignore the higher order terms}]$$

Let, the first approximation of $\hat{\theta}$ be $\hat{\theta}^{(1)}$. Then,

$$\hat{\theta}^{(1)} \simeq \theta_0 + \frac{\frac{\partial}{\partial \theta} \ln L(\theta) \Big|_{\theta=\theta_0}}{I(\theta_0)} \quad [\text{where, } I(\theta) = \frac{n}{2}]$$

Again, $\hat{\theta}^{(2)} \simeq \hat{\theta}^{(1)} + \frac{\frac{\partial}{\partial \theta} \ln L(\theta) \Big|_{\theta=\hat{\theta}^{(1)}}}{I(\hat{\theta}^{(1)})}$

The above procedure is repeated until the desired level of convergence is attained.

6 Conclusion

Cauchy distribution is a huge topic. Here, we considered some of its interesting properties along with different estimation procedures briefly in analytical and graphical approach. While doing the project, we observe that, estimation of the location parameter of the Cauchy distribution is not an easy task, be it unbiasedness or maximum likelihood estimation as we cannot use any moment estimator.

In one hand, the thick tails of this distribution make it an ‘outlier producing distribution’ while on the other hand as the moment estimates cannot be found we can indulge ourselves in quantile measures only and study their properties thoroughly which is a good aspect indeed.

Thus, there are many future aspects of this project as we can dive deep into the estimation for Cauchy distribution. There are further scopes of doing research based projects on the impact of Cauchy distribution as one can try to estimate the scale parameter, multivariate Cauchy parameters, testing procedures and confidence intervals etc.

7 Appendix

7.1 Additive Property of Cauchy Distribution

Proof. Let, X_1, X_2 be a random sample from $C(0, 1)$ distribution. Their joint PDF is given by,

$$f(x_1, x_2) = \frac{1}{\pi^2(1+x_1^2)(1+x_2^2)}, \quad x_1, x_2 \in \mathbb{R}$$

Let, $U = X_1 + X_2$, $V = X_2$

Then, $X_1 = U - V$, $X_2 = V$ and the jacobian of the transformation is 1.

The joint PDF of (U, V) is given by,

$$g(u, v) = \frac{1}{\pi^2\{1+(u-v)^2\}(1+v^2)}, \quad u, v \in \mathbb{R}$$

. : For $u \in \mathbb{R}$ the marginal PDF of U is,

$$\begin{aligned} h(u) &= \int_{-\infty}^{\infty} \frac{1}{\pi^2\{1+(u-v)^2\}(1+v^2)} dv \\ &= \int_{-\infty}^{\infty} \frac{1}{\pi^2 u^2 (u^2 + 4)} \left\{ \frac{2uv}{1+v^2} + \frac{u^2}{1+v^2} + \frac{2u(u-v)}{1+(u-v)^2} + \frac{u^2}{1+(u-v)^2} \right\} dv \\ &= \frac{1}{\pi^2 u^2 (u^2 + 4)} \left[u \ln \left\{ \frac{1+v^2}{1+(u-v)^2} \right\} + u^2 \left\{ \tan^{-1} v - \tan^{-1}(u-v) \right\} \right]_{-\infty}^{\infty} \\ &= \frac{1}{\pi^2 u^2 (u^2 + 4)} \cdot u^2 \cdot 2\pi \\ &= \frac{2}{\pi(u^2 + 4)} \end{aligned}$$

Thus, $U = X_1 + X_2 \sim C(0, 2)$

Hence, the proof. □

7.2 Regularity Conditions

Let X has a distribution from $\{f(x; \theta) : \theta \in \Omega\}$ satisfying the following regularity conditions -

- (i) The parameter space, Ω is an open interval in \mathbb{R} i.e. $\Omega = \{\theta : a < \theta < b\}$
- (ii) The support, $S = \{x : f(x; \theta) > 0\}$ is independent of θ
- (iii) For each $x \in S$, $\frac{\partial}{\partial \theta} [\ln f(x; \theta)]$ exists and finite
- (iv) The identity “ $\sum_{x \in S} f(x; \theta) = 1$ ” or “ $\int_S f(x; \theta) dx = 1$ ” can be differentiated under the summation or integral sign.
- (v) $T \in U_\psi = \left\{ T(X) : E_\theta [T(X)] = \psi(\theta), \text{Var}_\theta [T(X)] < \infty, \forall \theta \in \Omega \right\}$ is an UE of $\psi(\theta)$ such that the derivative of $\psi(\theta) = E[T(X)]$ with respect to θ can be evaluated by differentiating under the summation or integral sign.

7.3 Expectation and Symmetry

Proof. Let, the random variable X has a continuous distribution symmetric about the line $x = a$ with PDF $f(\cdot)$. Then, we have, $f(a-x) = f(a+x)$, $\forall x \in \mathbb{R}$. Now,

$$\begin{aligned}
 E(X-a) &= \int_{-\infty}^{\infty} (x-a)f(x)dx \\
 &= \int_{-\infty}^a (x-a)f(x)dx + \int_a^{\infty} (x-a)f(x)dx \\
 &= \int_0^{\infty} -uf(a-u)du + \int_0^{\infty} vf(a+v)dv \\
 &= -\int_0^{\infty} uf(a-u)du + \int_0^{\infty} vf(a-v)dv \\
 &= \begin{cases} 0, & \text{if } E(X) \text{ exists then } \int_0^{\infty} xf(a-x)dx \text{ is finite.} \\ -\infty + \infty, & \text{if } E(X) \text{ does not exist then } \int_0^{\infty} xf(a-x)dx = +\infty \end{cases}
 \end{aligned}$$

Hence, Expectation, if exists is the point of symmetry for a symmetric distribution. \square

7.4 R Codes

Code 1: Cauchy and Normal

```
1 #####  
2 # DISSERTATION: Simulation_01  
3 #####  
4 # Cauchy and Normal  
5 #-----  
6 emp_lvl = function(sizes, mu0 = 0, los = 0.05)  
7 {  
8     # Specifications of the argument  
9     # mu0: Hypothesized value of the location parameter  
10    # los: level of significance  
11  
12    # Testing -> H0: mu = mu0 against H1: mu != mu0  
13    # Cauchy(mu,1) vs Normal(mu,1)  
14  
15    R = 1000                      # simulation number  
16    n = sizes                     # sample sizes  
17    sigma = 1                      # Scale parameter  
18  
19    # Storing Empirical Levels  
20    emp.level = data.frame(Sample_Size = 0, Cauchy = 0, Normal = 0)  
21  
22    for (i in seq_along(n))  
23    {  
24        # Drawing Samples  
25        x.c = replicate(R, mean(rcauchy(n[i], mu0, sigma)))    #  
26        Cauchy  
27        x.n = replicate(R, mean(rnorm(n[i], mu0, sigma)))      #  
28        Normal  
29  
30        # Test Statistic (Normal theory based)  
31        ts.c = sqrt(n[i])*(x.c - mu0)/sigma    # Cauchy  
32        ts.n = sqrt(n[i])*(x.n - mu0)/sigma    # Normal  
33  
34        # Critical Point  
35        crit = qnorm(los/2, lower.tail = F)  
36  
37        # Empirical Levels  
38        emp.c = mean(abs(ts.c) > crit)  
39        emp.n = mean(abs(ts.n) > crit)
```

```

39         emp.level[i,] = c(n[i], emp.c, emp.n)
40     }
41
42     return(emp.level)
43 }
44
45 # Choosing Sample Sizes
46 n = c(5, 10, 25, 50, 100, 200, 500, 1000)
47
48 # Simulation Results
49 set.seed(2022)           # Initializing the seed at '2022'
50 output = emp_lvl(n); output # Output of the Program
51
52 =====

```

Code 2: Convergence in Cauchy

```

1 =====
2 # DISSERTATION: Simulation_02
3 =====
4 setwd("E:/St. Xavier's College/DISSERTATION")
5 getwd()
6
7 library(ggplot2)
8 library(dplyr)
9
10 # Theme ----
11 theme_new = theme(plot.title = element_text(size = 22,
12                                              face = "bold",
13                                              hjust = 0.5),
14     legend.title = element_text(size = 15,
15                                              face = "bold"),
16     plot.subtitle = element_text(size = 19,
17                                              hjust = 0.5),
18     plot.caption = element_text(size = 14),
19     plot.margin = margin(1, 1.5, 0.7, 0.7, "cm"),
20     legend.text = element_text(size = 13),
21     axis.title = element_text(size = 15,
22                                              face = "bold"),
23     axis.text = element_text(size = 13,
24                                              face = "bold"),
25     strip.background = element_blank(),
26     strip.placement = "outside",

```

```

27             strip.text = element_text(size = 16,
28                                         color = 1,
29                                         face = "bold"))
30
31 theme_set(theme_light() + theme_new)
32
33 # -----
34
35 # Sample Mean vs Sample Median
36 # -----
37 cauchy = function(size_max, mu)
38 {
39     n = 1:size_max
40     sigma = c(1, 4, 10)
41     k = length(sigma)
42     s = 100
43
44     # storage of probability
45     data = vector(mode = "list", length = k)
46     names(data) = seq_along(sigma)
47
48     for (i in 1:k)
49     {
50         mean_med = function(size)
51         {
52             set.seed(s)
53             y = rcauchy(size, mu, sigma[i])
54
55             m = mean(y)
56             med = median(y)
57
58             return(c(m, med))
59         }
60
61         est = matrix(0, size_max, 2)
62         colnames(est) = c("Sample Mean", "Sample Median")
63
64         for (j in n)
65             est[j,] = mean_med(j)
66
67         data[[i]] = est
68     }
69

```

```

70   data = list2DF(data) %>%
71     mutate(size = rep(n, 2),
72           estimator = rep(colnames(est),
73                             each = size_max) %>%
74             factor()) %>%
75   tidyr::pivot_longer(-c(estimator, size), "sigma") %>%
76   mutate(sigma = factor(sigma))
77
78 # levels(data$sigma) = paste("Scale =", sigma)
79 sgm_names = c('1' = "sigma==1",
80               '2' = "sigma==4",
81               '3' = "sigma==10")
82
83 data %>%
84   ggplot(aes(size, value)) +
85   geom_point(size = 1.2) +
86   labs(title = "Consistency for Cauchy Distribution",
87         subtitle = bquote(mu == .(mu)),
88         x = "\nSample Size",
89         y = "Estimator\n") +
90   geom_segment(aes(0, mu, xend = size_max, yend = mu),
91                 col = "red3", lwd = 1) +
92   ylim(mu-10, mu+10) +
93   facet_grid(sigma ~ estimator,
94             labeller = labeller(sigma = as_labeller(sgm_
95                               names,
96                               label_
97                               parsed
98                               ))))
99
100 cauchy(300, 0)
101
102 #-----
```

Acknowledgement

I would like to express my special gratitude to Father Principal Rev. Dr. Dominic Savio, Sj. and the Department of Statistics who gave the golden opportunity to do this wonderful project on ‘Importance of Cauchy Distribution in the Field of Statistics’. I would also like to thank my supervisor and dissertation guide Professor Ayan Chandra for his guidance throughout the duration of completion of my project. I am also very much grateful to all my respected professors of the St. Xavier’s College Statistics faculty, who have inculcated in me, a strong research mindset, curiosity and intrinsic capability to pursue this subject. Lastly, I would like to thank St. Xavier’s College for the opportunity to prepare a dissertation project paper on a topic of my choice, as well as for imbuing in me a drive to polish my research mindset.

References

- [1] VD Barnett. Order statistics estimators of the location of the cauchy distribution. *Journal of the American Statistical Association*, 61(316):1205–1218, 1966.
- [2] Daniel Bloch. A note on the estimation of the location parameter of the cauchy distribution. *Journal of the American Statistical Association*, 61(315):852–855, 1966.
- [3] Norman L. Johnson, Samuel Kotz, and N. Balakrishnan. *Continuous Univariate Distributions (Vol. I)*. Wiley Series in Probability and Mathematical Statistics. Wiley-Interscience Publication, 1994.
- [4] Glenn Stoops and Donald Barr. Moments of certain cauchy order statistics. *The American Statistician*, 25(5):51–51, 1971.
- [5] Wikipedia. Cauchy distribution. https://en.wikipedia.org/wiki/Cauchy_distribution.