Importance of Cauchy Distribution in the Field of Statistics

Descriptive statistics

Name: Mst. Saima Jahan
Registration No.: 19236134588
Department: Statistics
Session: 2019 – 20
Self Practice Project

Declaration

I affirm that I have identified all my sources and that no part of my dissertation paper uses unacknowledged materials and it is a practice project and can be done by others too.

Contents

Introduction

1 Research Idea

1.1 Background
1.2 Importance of the
1.3 Objectives of the Study
1.4 Scheme of Work

2 Literature Review

3 Cauchy Distribution

3.1 Description
3.1.1 Genesis
3.1.2 Background for developing Cauchy distribution

3.2 Probability Laws

3.2.1 Probability Density Function
3.2.2 Cumulative Distribution Function
3.2.3 Visualizing PDF and CDF

3.3 Interesting Properties

3.3.1 Natrure of the Curve
3.3.2 Non-existence of Moments
3.3.3 Quantile Measures
3.3.4 Mode
3.3.5 Points of Inflection

3.4 Sampling Properties

4 Cauchy and Normal Distribution

4.1 Why do we need Cauchy apart from Normal?
4.2 Comparison of Cauchy & Normal Distribution
4.3 A Simulation Study

5 Estimation under Cauchy Population 16
5.1 Unbiasedness
5.2 Consistency
5.3 Cramer-Rao Inequality
5.4 Maximum Likelihood Estimation .

6 Conclusion

7 Appendix 22

7.1 Additive Property of Cauchy Distribution
7.2 Regularity Conditions
7.3 Expectation and Symmetry
7.4 R Codes

References

Plots

1 Nature of the curves of the form
x2+a2
2 PDF of Cauchy distribution
3 CDF of Cauchy distribution
4 Shape of the C(0,1) PDF
5 Truncated and Standard Cauchy
6 Points of Inflection of Standard Cauchy7 X and X1 are Identically Distributed
8 X+Y and 2X are Identically Distributed
9 Densities of Cauchy and Normal
10 Comparison of Sample Mean and Sample Median

Codes

1 Cauchy and Normal
2 Convergence in Cauchy

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 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 truenature 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: 1/π(1+x2), commonly known as Cauchy density or more evidently the curves proportional to 1/x2+a2 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 behavior of this curve.

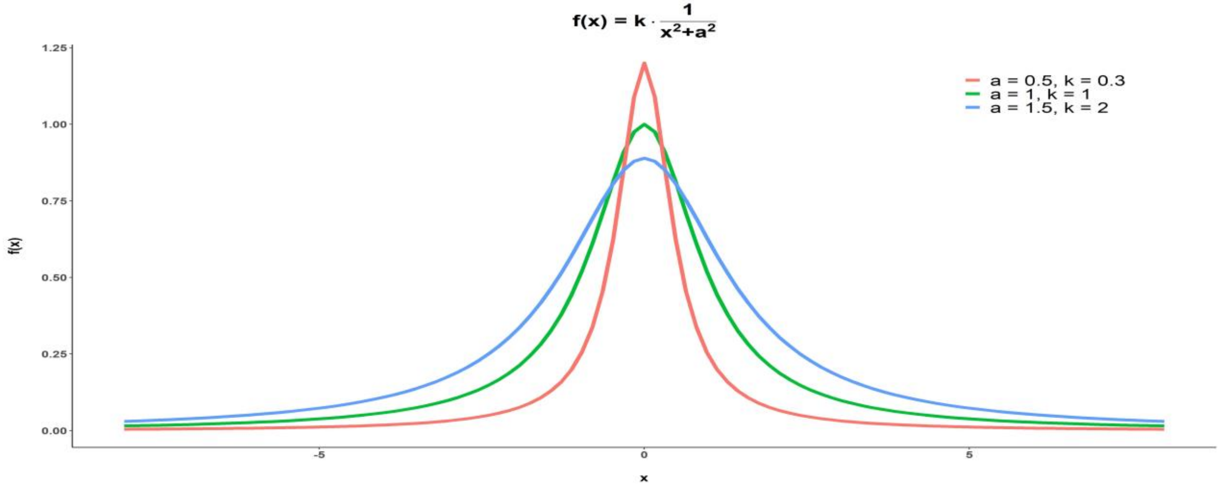

Figure 1: Nature of the curves of the form 1/x2+a2

3.1.2 Background for developing Cauchy distribution

The idea of Cauchy distribution was initially developed by French mathematician and physicist Simeon Denis Poisson who observed that the distribution with density 1/π(1+x2) 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 Bienayme’s article showing Legendre’s least-squares fails to provide “the most probable results” in case of an 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(μ,σ) .

3.2.1 Probability Density Function

The Probability Density Function (PDF) of C(μ,σ) is given by,
$f\ (x) = f(x) = f(x) = (1/\pi) \cdot \sigma\ /\ [\sigma^2 + (x - \mu)^2],\ x \in \mathbb{R}$

3.2.2 Cumulative Distribution Function

The Cumulative Distribution Function (CDF) of C(μ,σ) is given by,
$F(x) = 1/2 + (1/\pi)\ \tan^{-1}[(x - \mu)/\sigma],\ x \in \mathbb{R}$

Here, $\mu \in$ Ris 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 canthink of a standardized Cauchy distribution by putting, $\mu = 0$ and $\sigma = 1$.

3.2.3 Visualizing PDF and CDF

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

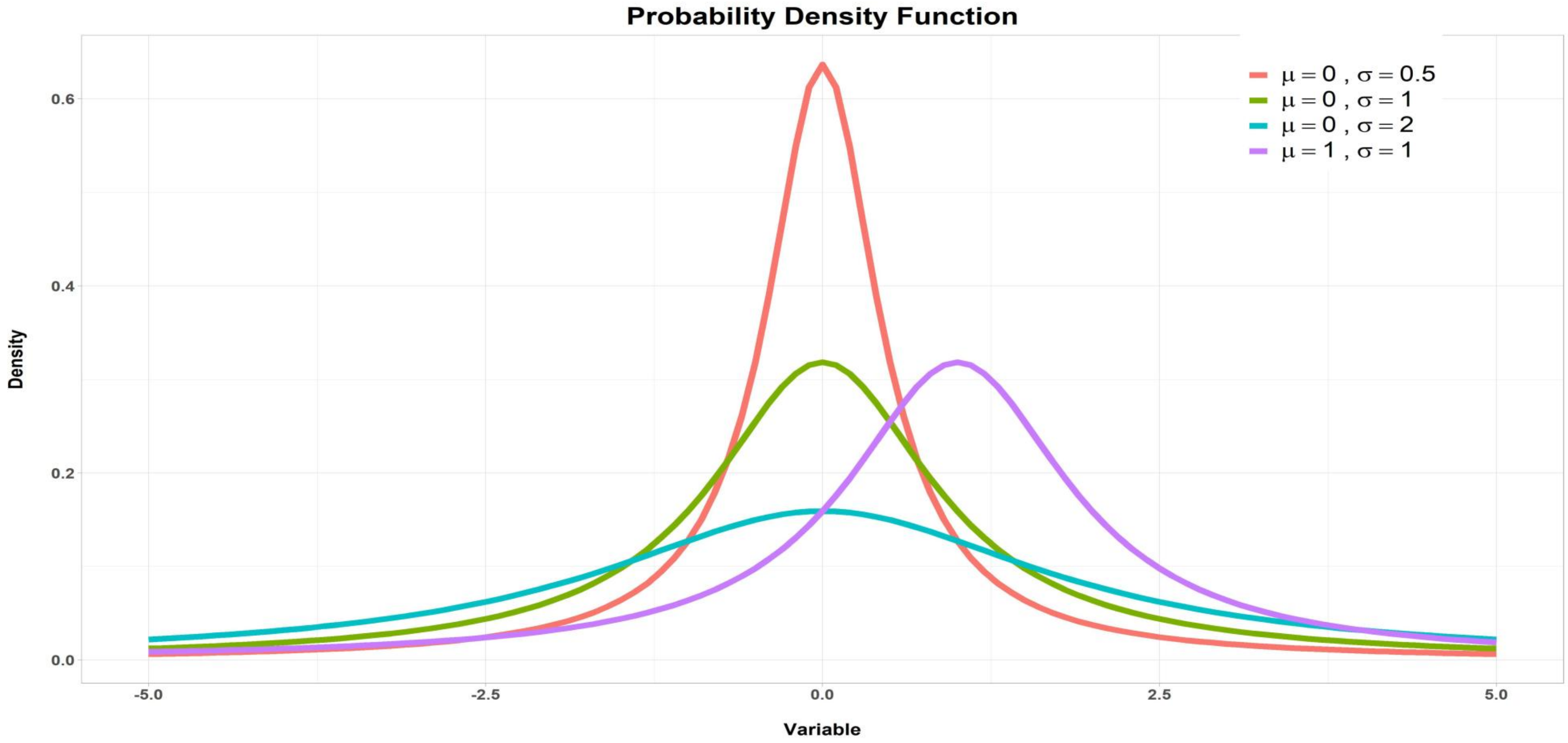

Figure 2: PDF of Cauchy distribution

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

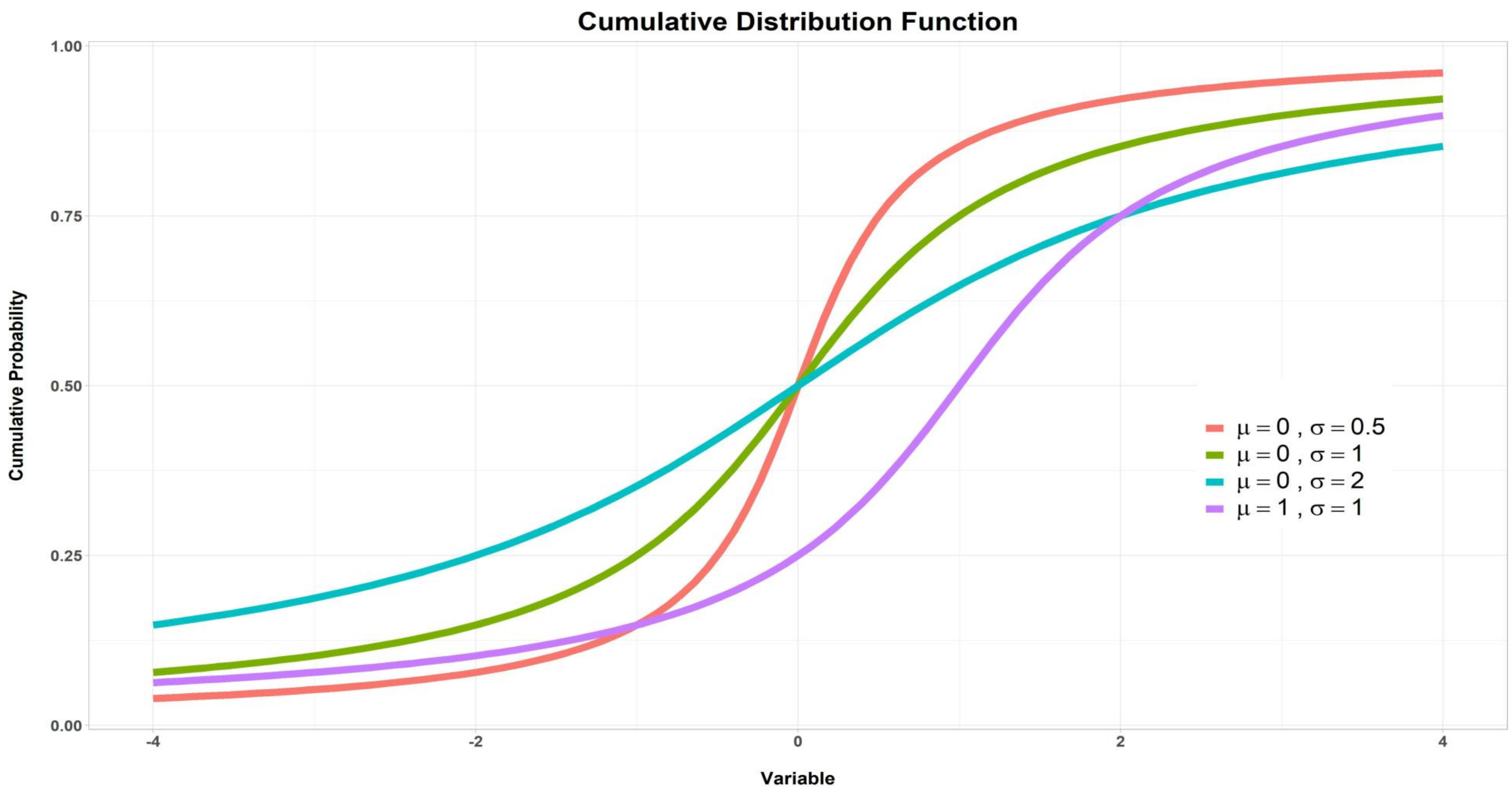

Figure 3: CDF of Cauchy distribution

3.3 Interesting Properties

3.3.1 Natrure of the Curve

The values of f (x) = σ / {π[σ² + (x − μ)²]}, x ∈ ℝ} are plotted against the values of x, we get a graph as shown below -

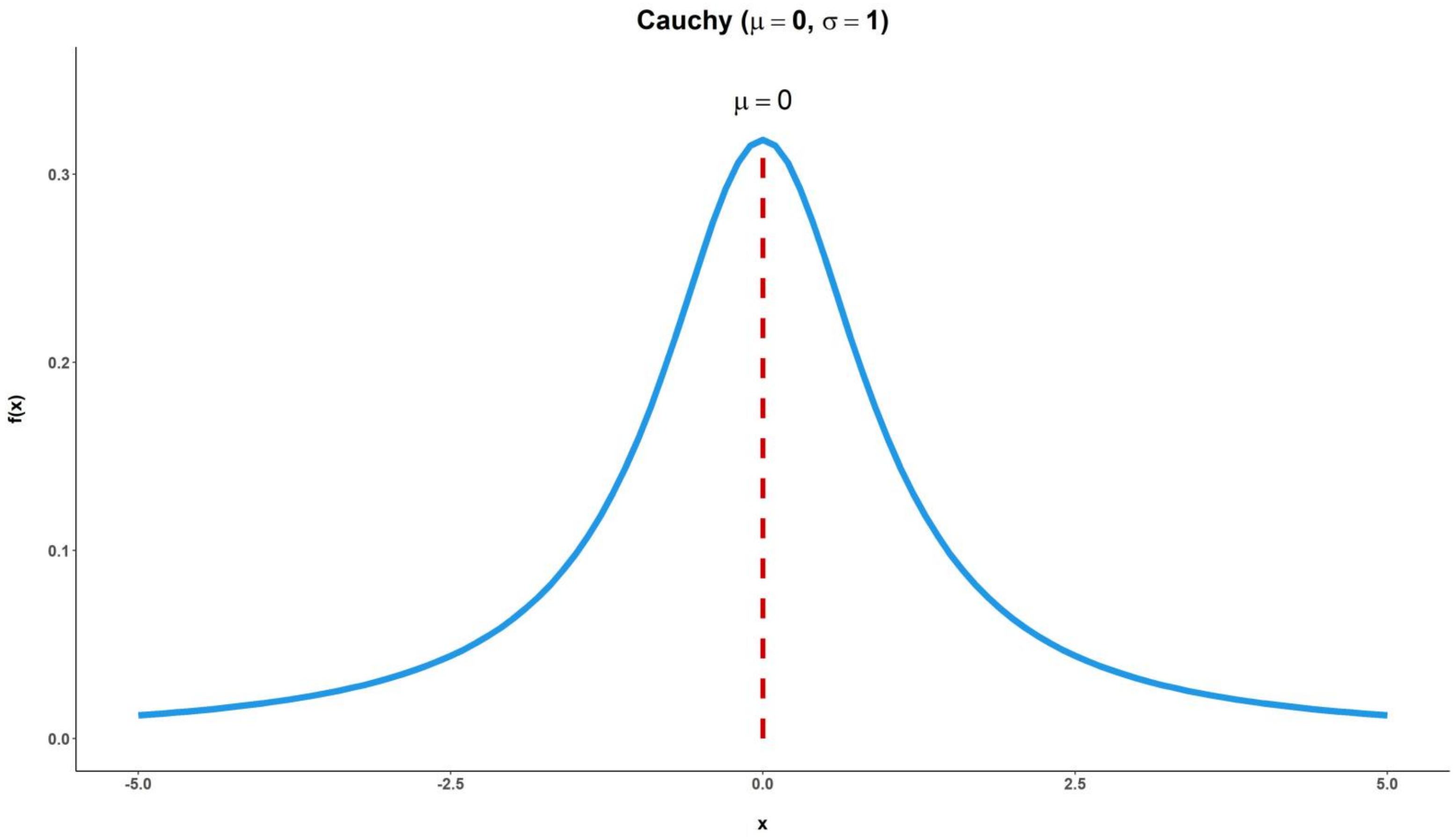

Figure 4: Shape of the C(0,1) PDF

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,

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

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

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 rth order moment of $C(\mu,\sigma)$, let us first find –

$E\,|(X - \mu)/\sigma|^r$

$= \int_{-\infty}^{\infty} |(x - \mu)/\sigma|^r \cdot [\sigma / \{\pi(\sigma^2 + (x - \mu)^2)\}]\, dx$

Let $z = (x - \mu)/\sigma$, so $dx = \sigma\, dz$. Then,

$= (1/\pi) \int_{-\infty}^{\infty} |z|^r / (1 + z^2)\, dz$

Since the integrand is an even function of z,

$= (2/\pi) \int_0^{\infty} z^r / (1 + z^2)\, dz$

Now substitute $u = z^2$, so $du = 2z\, dz$. Hence,

$= (1/\pi) \int_0^{\infty} u^{(r-1)/2} / (1 + u)\, du$

which converges, if and only if, $r+1>0$ and $1-r>0$, i.e. $-1<r<1$
Hence, $E[|(X-\mu)/\sigma|r]$ 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 $0 < k < \infty$,

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

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 I2.

- **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 I2 vanishes.

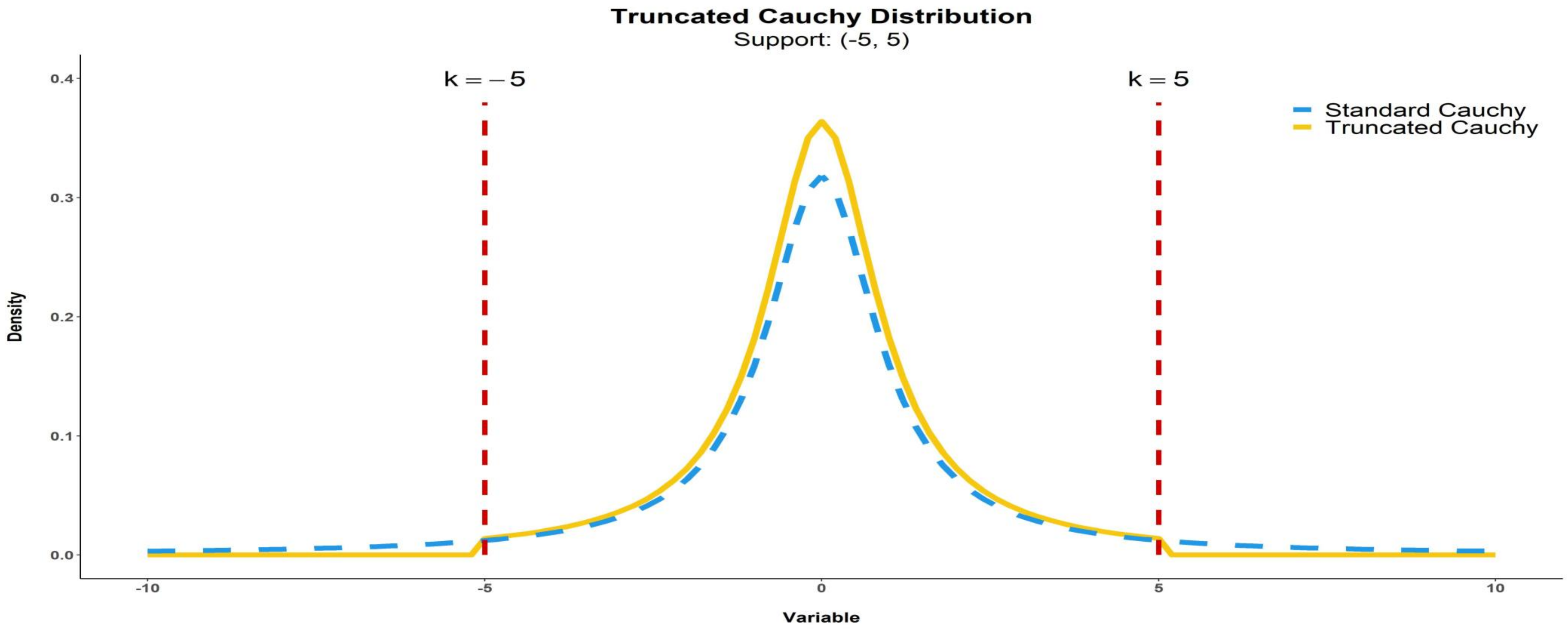

Figure 5: Truncated and Standard Cauchy

3.3.3 Quantile Measures

Let ξ_p be the *p*-th quantile of the **$C(\mu, \sigma)$** distribution. Then,

$F(\xi_p) = p$

$\Rightarrow 1/2 + (1/\pi) \tan^{-1}[(\xi_p - \mu)/\sigma] = p$

$\Rightarrow \tan^{-1}[(\xi_p - \mu)/\sigma] = \pi(p - 1/2)$

$\Rightarrow (\xi_p - \mu)/\sigma = \tan[\pi(p - 1/2)]$

$\Rightarrow \xi_p = \mu + \sigma \tan[\pi(p - 1/2)]$

From **(****)**:

For **$p = 1/2$**, the **Median** = $\xi_{1/2} = \mu$

For **$p = 1/4$**, the first quartile is **$Q_1 = \xi_{1/4} = \mu - \sigma$**

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

Hence, the **quartile deviation** is **$(Q_3 - Q_1) / 2 = \sigma$**

3.3.4 Mode

From ‘Figure 4’, we can see that the Cauchy distribution is unimodal and the mode is at x = μ. The mathematical justification is as follows:

$f(x) = \sigma / \{\pi[\sigma^2 + (x - \mu)^2]\}$

$f'(x) = -\, 2(x - \mu) / [\sigma^2 + (x - \mu)^2] \cdot f(x)$

Now, since **$f(x) > 0$** for all *x*, **$f'(x) \gtreqless 0$ according as $x \lesseqgtr \mu$**

Hence, the function increases for $x < \mu$ and decreases for $x > \mu$. Therefore, the **mode of the distribution** is at **$x = \mu$** and the **modal value** is **$f(\mu) = 1 / (\pi\sigma)$**

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,

The second derivative of the pdf is
$f''(x) = [3(x - \mu)^2 - \sigma^2] / [\sigma^2 + (x - \mu)^2]^2 \cdot f(x)$
Since $f(x) > 0$ for all *x*, the sign of $f''(x)$ depends on
$3(x - \mu)^2 - \sigma^2$.
Hence,
$f''(x)$
> 0, if $x < \mu - \sigma/\sqrt{3}$
< 0, if $\mu - \sigma/\sqrt{3} \leq x \leq \mu + \sigma/\sqrt{3}$
> 0, if $x > \mu + \sigma/\sqrt{3}$

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

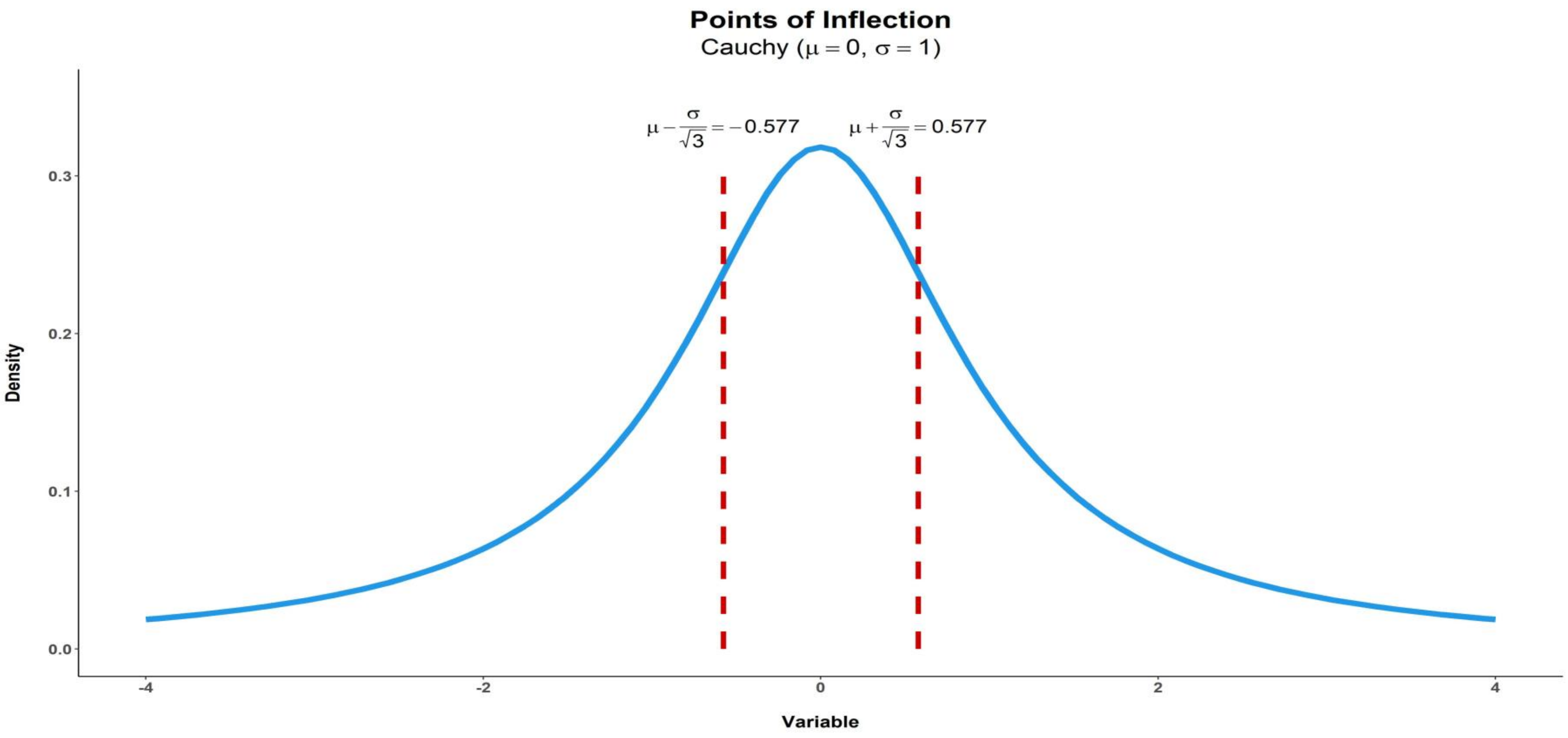

Figure 6: Points of Inflection of Standard Cauchy

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

• $X \sim C(0,1) \Longrightarrow X1 \sim C(0,1)$

Proof

Let $g(\cdot)$ be the PDF of $Y = 1/X$.
The Jacobian of the transformation is
$|dx/dy| = 1 / y^2$, with $y \in \mathbb{R}$.
Then, the PDF of Y is given by
$g(y) = f(1/y) \cdot |dx/dy|$, $y \in \mathbb{R}$,
where $f(\cdot)$ is the PDF of X.
Therefore,
$g(y) = (1 / \pi) \cdot 1 / (1 + 1/y^2) \cdot 1 / y^2$, $y \in \mathbb{R}$
$= 1 / \{\pi(1 + y^2)\}$, $y \in \mathbb{R}$

Hence, if X is a random variable having a C(0,1) distribution then X1 also have the identical distribution.

The above property is visualized below:

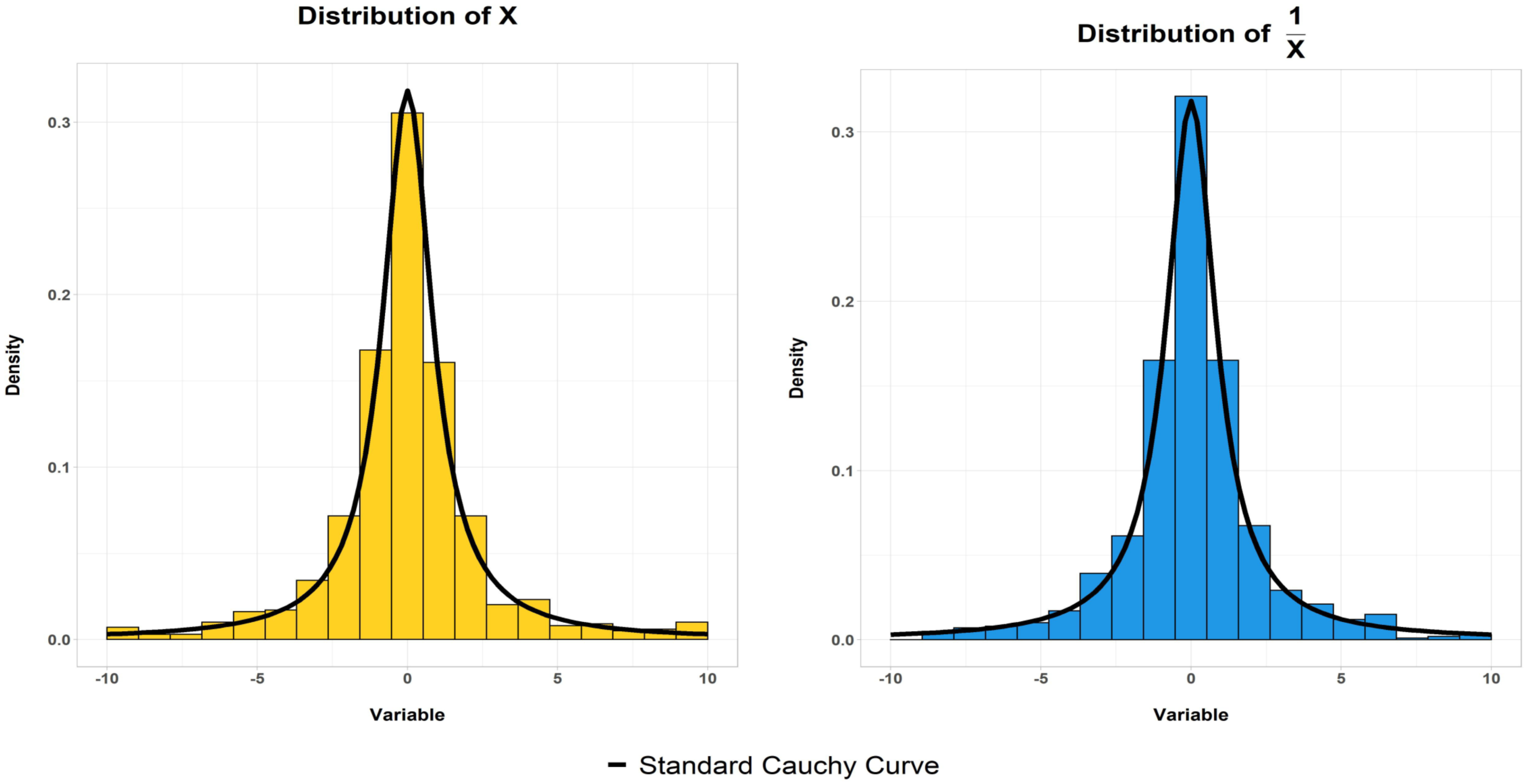

Figure 7: X and X1 are Identically Distributed