A Rich Learning Statistical Lesson in the Cauchy

Mohammad Fraiwan Al-Saleh*

m-saleh@yu.edu.jo

Arwa Salem Maabreh

Abstract

Cauchy distribution is a good example of a distribution that has many strange properties and is usually used as a counter example to show that some statistical properties are not always true. In this paper, some hidden properties of the Cauchy distribution are highlighted and studied in details. The focus is to closely look at some very interesting properties of the Cauchy distribution that are rarely mentioned in classrooms. Mentioning these properties in classrooms will have a positive effect on students' attitude toward learning. One of the main properties of the Cauchy distribution is that it has undefined or infinite moments; the odd moments are undefined while the even moments are infinite. For this distribution, the only way to summarize the data is by ordering them (order Statistics); it is the minimal sufficient statistic. The relation of Cauchy to other distributions such as the Normal, and Uniform are highlighted. The above properties several other properties are discussed in details in this paper. We believe that the content of this manuscript will be a real contribution to educational statistics.

Keywords: Cauchy distribution, undefined moments, truncated Cauchy distribution, Jenson's inequality, overlapping coefficient.

* Department of Statistics, Yarmouk University-Jordan.

Received: 5/ 9/2024 . Accepted: 5 / 1/2025 .

درس إحصائي غني بالتعلم في توزيع كوشي

محمد فريوان الصالح*

أروي سالم معابره

ملخص

يعد توزيع كوشي مثالا جيدا لتوزيع يحتوي على العديد من الخصائص الغريبة وعادة ما يستخدم كمثال لإظهار أن بعض الخصائص الإحصائية ليست صحيحة دائما. في هذا البحث، تم تسليط الضوء على بعض الخصائص الخفية لتوزيع كوشي ودراستها بالتفصيل وإلقاء نظرة فاحصة على بعض الخصائص المثيرة للاهتمام لهذا لتوزيع والتي نادرا ما يتم ذكرها في الفصول الدراسية. سيكون لذكر هذه الخصائص في الفصول الدراسية تأثير إيجابي على موقف الطلاب تجاه التعلم. تتمثل إحدى الخصائص الرئيسية لتوزيع كوشي في أنه يحتوي على معالم غير محددة أو لا نهائية. المعالم الفردية غير محددة بينما المعالم الزوجية لا نهائية. بالنسبة لهذا التوزيع، فإن الطريقة الوحيدة لتلخيص البيانات هي ترتيبها (إحصائيات الترتيب) وهي الحد الأدنى من الإحصائية الكافية.

الكلمات المفتاحية: توزيع كوشي، معالم غير معرفة، توزيع كوشي المقطوع، متباينة جونسون، معامل التداخل.

* قسم الإحصاء، جامعة اليرموك، الأردن.

تاريخ تقديم البحث: 2024/9/5. تاريخ قبول البحث: 2025/1/5م .

Introduction:

The Cauchy distribution is a continuous heavy-tailed probability distribution that was named after Augustin Cauchy in 1853. According to (Stigler, 1974) in his historical review of this distribution, many scientists, including Newton, Leibniz, and Maria Agnesi, investigated the Cauchy after it first appeared in Pierre de Fermat's writings in the mid-seventeenth century, and the first explicit analysis of the properties of the Cauchy distribution was published by Poisson in 1824.

There are several applications of the Cauchy distribution; for example, in quantum mechanics, it models the distribution of energy in an unstable state of an atom, where the random variable represents the energy width of the state that decays exponentially with time (Roe, 2012). The Cauchy distribution can be used to model the ratio of two normal random variables (Johnson et al. 1994) and it can be used to model financial returns, which is essentially the ratio between the stock price at time n+1 and time n (Nolan, 2014).

The Cauchy distribution is commonly used as a counterexample for many statistical properties that are rarely untrue; for example, the mean is undefined; the average of a random sample is not a sufficient statistic. Some other interesting properties are highlighted in this paper. These properties, when mentioned by teachers during their classes in mathematics and statistics courses, can be used to motivate students.

Section 2 contains a close look at the probability density function(pdf) and cumulative distribution function(cdf) of the Cauchy distribution. The moments of this distribution and the consequences of having an undefined mean are the content of Section 3. Specifically, the distribution of the sample average is discussed in Section 4. The Correlation between $X - Y$ and $X + Y$ is discussed in Section 5. Section 6 covers some Cauchy-related distributions. The overlapping between Cauchy and Normal distribution is the content of Section 7. The truncated Cauchy distribution is explained in Section 8. More details about using the properties of this distribution in classrooms are given in section 9. The conclusions are the content of Section 10.

The mathematical and statistical results mentioned above which will be discussed in detail in this paper are not new, but they are new ideas for mathematics and statistics teachers to use in classrooms. For some other published work on educational statistics, see (Al-Saleh et. al., 2010) and

(Al-Saleh, 2008), (Al-Saleh and Adam, 2024). For current work on Cauchy distribution, see (Maabreh and Al-Saleh, 2023)

The Cauchy pdf and cdf

The pdf of the Cauchy distribution with location parameter θ and scale parameter λ, denoted $C(\theta,\lambda)$, is:

$$f(x;\theta,\lambda)=\frac{1}{\pi\lambda}\frac{1}{1+\left(\frac{x-\theta}{\lambda}\right)^2}, \qquad |x|<\infty, |\theta|<\infty, 0<\lambda<\infty.$$

Its cdf is: $$F(x;\theta,\lambda)=0.5+\frac{1}{\pi}tan^{-1}\left(\frac{x-\theta}{\lambda}\right).$$

The standard Cauchy distribution has a location parameter of 0 and a scale parameter of 1; $X \sim C(0,1)$ with pdf: $f(x)=\frac{1}{\pi}\frac{1}{\left(1+x^2\right)}, \quad |x|<\infty$.

The graphs of the Cauchy pdf with the same location (taken to be zero) and 3 different shape parameters is given in Figure 1(It is obtained using Scientific Work Place Package(SWP)). Clearly, it is symmetric around zero which is the mode and the median of this distribution and its height is $1/\lambda\pi$. For $C(\theta,\lambda)$, θ is the mode and the median of this distribution and its height is $1/\lambda\pi$. The standard normal and Cauchy distribution are given in Figure 2.

Figure (1) Pdf of Cauchy with θ=0 and $\lambda=1,2,3$

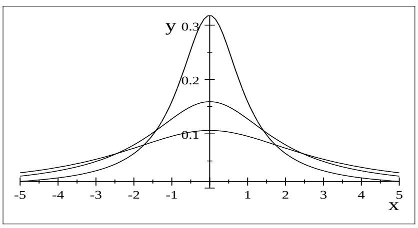
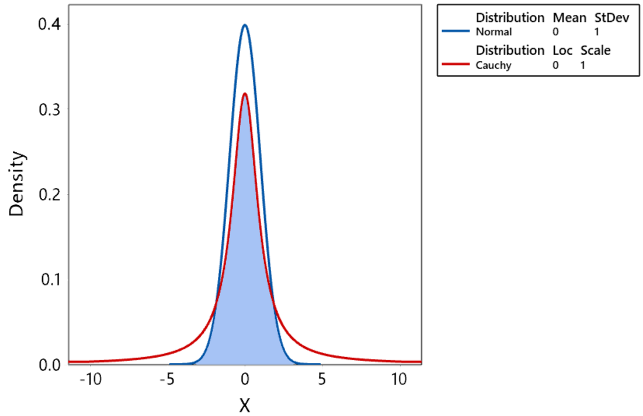


Figure (2) N(0,1) and C(0,1)



Moments of the Cauchy Distribution:

If X is $C(\theta,\lambda)$ then $(X-\theta)/\lambda$ is $C(0,1)$. Thus, without loss of generality, we will concentrate on $C(0,1)$.

$$\mathrm{E}(X)=\frac{1}{\pi}\int_{-\infty}^{\infty} x\frac{1}{1+x^2}dx = \frac{1}{2\pi}\ \ln(1+\mathrm{x}^2)\,|_{-\infty}^{\infty}=\infty-\infty.$$

$$\mathrm{E}(X^2)= \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{x^2}{1+x^2}dx =(1/\pi)\left((1/2)\pi+x-\tan^{-1}x\right)|_{-\infty}^{\infty}=\infty.$$

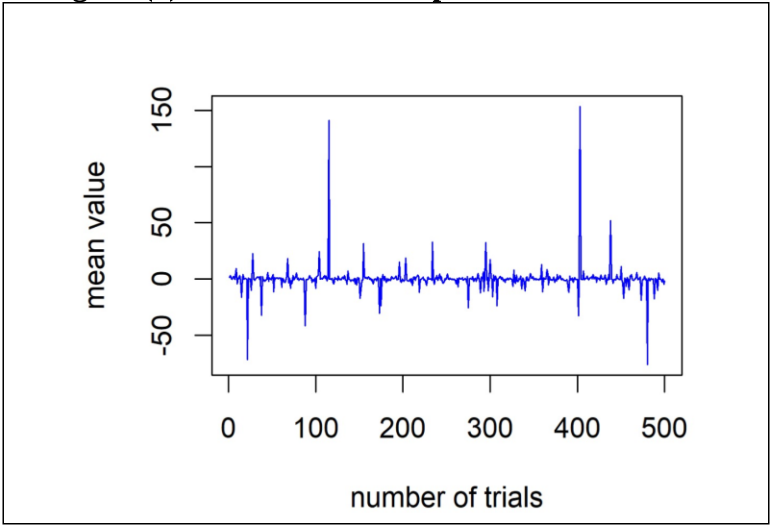
$$Var(X)=\sigma^2=E(X^2)-\left(E(X)\right)^2.$$

Thus, the Cauchy distribution has undefined(indeterminate)mean and variance. One way to visualize this is to look at a chart of the running average of random samples drawn from a Cauchy distribution. If we generate many random samples each of size n from $C(0,1)$, then the sample mean will not converge to a specific value. This behaviour is caused by the heavy tail, where the tail of the pdf goes to zero very slowly.

As a result of the non-existent of the mean, the law of large numbers (LLN) and the Central Limit Theorem (CLT) both fail for the Cauchy distribution. In simple word, the LLN states that the average of a random sample from a distribution converges in probability to the mean of the distribution. The CLT states that the distribution of the average of a random sample from a distribution with finite mean and variance converges to normal distribution.

500 random samples of size 1000 each were chosen from $C(0,1)$. The mean for each sample was obtained. Figure 3 is the graph of the means (the x-axis is the sample number and the y-axis is the sample mean). Certainly, the sample mean is unstable (does not converge); it fluctuates from about -50 to 150.

Figure (3) Means of 500 samples, each of size 1000



The following numerical example clarifies the meaning of the undefined mean and variance, the defined median, and the heavy tail property of the Cauchy distribution. Using the statistical package, MINITAB, five random samples of size 20,000 each were generated from C (0,1). The summary of the data is given in the following table:

Table (1)Sample mean, median, standard deviation, minimum, and maximum for 5 samples of size 20,000 each from C (0,1)

Sample #	Mean	Median	St. Dev.	Min.	Max.
1	4.6	0.0	568.5	-2749.9	77004.1
2	0.3	0.0	131.2	-6890	13355.1
3	-7.0	0.0	1124	-128829.0	62013
4	-5.8	0.0	556.1	-67706.5	1666.4
5	-0.5	0.0	138.3	-9181.9	12746.4

The mean and the standard deviation fluctuate, which is an indication of the undefined mean and standard deviation of the distribution, while the median is around zero. The range ($Min. - Max.$) is very large, which is a strong indication of the heavy tail of this distribution.

Definition: For a given random variable X, we have two associated non-negative random variables [8]:

$$X^{+} = \max\{X, 0\} = \begin{cases} X & if\ X \geq 0 \\ 0 & if\ X < 0 \end{cases} \quad \&$$

$$X^{-} = \max\{-X, 0\} = \begin{cases} 0 & if\ X \geq 0 \\ -X & if\ X < 0 \end{cases} .$$

Clearly, $X = X^{+} - X^{-}$, and $X^{+}X^{-} = 0$. Therefore, the expected value of a random variable X is $\mathrm{E}(X) = \mathrm{E}(X^{+}) - \mathrm{E}(X^{-})$. $\mathrm{E}(X)$ is defined(exists) if $\mathrm{E}(X^{+})$ or $\mathrm{E}(X^{-})$ is finite. If both of them are infinite, then $\mathrm{E}(X)$ is undefined.

For the Cauchy distribution,

$$\mathrm{E}(X^{+}) = \mathrm{E}(X^{-}) = \infty. \text{ and } \mathrm{E}(X^{2}) = \mathrm{E}\left((X^{+})^{2}\right) + \mathrm{E}\left((X^{-})^{2}\right) = \infty .$$

To study $\mathrm{E}(X^{r})$, for any positive integer r, we need to use the following inequality:

Jensen’s inequality: If X is a random variable and $g(x)$ is a real continuous and convex function ($g''(x) > 0$) then $g\left(E(X)\right) \leq E\left(g(X)\right)$, [9].

$g(x) = x^{2}$ is convex because, $g''(x) = 2$. Thus, using Jenson’s inequality:

$$\left(\mathrm{E}(X^{+})\right)^{2} \leq \mathrm{E}\left((X^{+})^{2}\right) \Rightarrow \mathrm{E}\left((X^{+})^{2}\right) = \infty,$$
$$\left(\mathrm{E}(X^{-})\right)^{2} \leq \mathrm{E}\left((X^{-})^{2}\right) \Rightarrow \mathrm{E}\left((X^{-})^{2}\right) = \infty. .$$

Therefore, $\mathrm{E}(X^{2}) = \infty + \infty = \infty.$

$$\mathrm{E}(X^{3}) = \mathrm{E}\left((X^{+} - X^{-})^{3}\right) = \mathrm{E}\left((X^{+})^{3} - 3X^{+2}X^{-} + 3X^{+}X^{-2} - (X^{-})^{3}\right)$$

$$= \mathrm{E}\left((X^{+})^{3} - (X^{-})^{3}\right) = \mathrm{E}\left((X^{+})^{3}\right) - \mathrm{E}\left((X^{-})^{3}\right).$$

$$\left(\mathrm{E}(X^{+})\right)^{3} \le \mathrm{E}\left((X^{+})^{3}\right) \to \mathrm{E}\ \left((X^{+})^{3}\right) = \infty, \left(\mathrm{E}(X^{-})\right)^{3} \le \mathrm{E}\left((X^{-})^{3}\right) \to \mathrm{E}\ \left((X^{-})^{3}\right) = \infty.$$

.

Thus, $\mathrm{E}(X^{3}) = \infty - \infty = \textit{undefined}$. In general, for any positive integer r,

$$\mathrm{E}(X^{r}) = \mathrm{E}\left((X^{+} - X^{-})^{r} \right) = \mathrm{E}\left(\sum_{k=0}^{r} (-1)^{k} \binom{r}{k} (X^{-})^{k} (X^{+})^{r-k} \right)$$

$$= \mathrm{E}\left(\binom{r}{0} (X^{+})^{r} - \binom{r}{1} (X^{-})(X^{+})^{r-1} + \ldots + (-1)^{r} \binom{r}{r} (X^{-})^{r} \right)$$ Th

$$= \mathrm{E}\left((X^{+})^{r} + (-1)^{r} (X^{-})^{r}\right), \qquad \text{since} \quad X^{+}X^{-} = 0$$

erefore,

$$\mathrm{E}(X^{r}) = \begin{cases} \mathrm{E}\left((X^{+})^{r}\right) + \mathrm{E}\left((X^{-})^{r}\right), & \text{if } r \text{ is even} \\ \mathrm{E}\left((X^{+})^{r}\right) - \mathrm{E}\left((X^{-})^{r}\right), & \text{if } r \text{ is odd} \end{cases} = \begin{cases} \infty & \text{if } r \text{ is even} \\ \textit{undefined} & \text{if } r \text{ is odd}. \end{cases}$$ T

he above implies that any central moment (that depends on the mean) is undefined, so the variance is also undefined. As a result, the Cauchy distribution does not have a moments generating function. However, it does have a characteristic function:

If X is $C(\theta, \lambda)$ then $Y = (X - \theta)/\lambda$ is $C(0,1)$. The characteristic function of X is

$$C_{X}(t) = \mathrm{E}(e^{itX}) = \mathrm{E}(e^{it(\theta + \lambda Y)}) = e^{it\theta}\mathrm{E}(e^{it\lambda Y}) = e^{it\theta}e^{-\lambda|t|} = e^{it\theta - \lambda|t|}.$$

The characteristic function has no derivative at t=0, indicating that there is no finite moment for this distribution (Florescu & Tudor,2013).

In summary, the Cauchy distribution has infinite or undefined moments, has no moments generating function, the average of a random sample failing to converge as the sample size increases, and θ is the median and mode but not the mean.

The Distribution of the Sample Mean $\bar{X}$

Let $X_1, X_2, \ldots, X_n$ be $iid\ C(\theta, \lambda)$. The characteristic function of $\overline{X}$ is

$$C_{\bar{X}}(t) = \mathrm{E}\,(e^{it\bar{X}}) = \mathrm{E}\ (e^{it\sum_{i=1}^{n} X_i/n}) = \mathrm{E}\left(\prod_{i=1}^{n} e^{i\frac{tX_i}{n}}\right) = \mathrm{E}\left(e^{it\frac{X}{n}}\right)^n = \mathrm{E}\,(e^{itX}) = e^{it\theta-\lambda|t|}.$$

Thus, $\overline{X}$ has $C(\theta,\lambda)$ distribution; hence $\bar{X}$ provides no more information about the parameters than any of the individual observations. This seems odd; applying the factorization theorem, when λ is known, assumes $\lambda = 1$, the likelihood function is

$$L(\theta \mid x_1, x_2, \ldots, x_n) = \frac{1}{\pi^n}\prod_{i=1}^{n}\frac{1}{1+\left(x_i-\theta\right)^2} = \frac{1}{\pi^n}\prod_{i=1}^{n}\frac{1}{1+\left(x_{(i)}-\theta\right)^2}, \quad \left(x_1, x_2, \ldots, x_n\right) \in \mathbb{R}^n.$$

Thus, $\overline{X}$ is not a sufficient statistic and $(X_{(1)}, X_{(2)}, \ldots, X_{(n)})$ is a minimal sufficient statistic.

Based on the above, there is no sufficient way to summarize the data in a random sample from this distribution. The only thing we can do is to put the sample values in an ascending order, $(X_{(1)}, X_{(2)}, \ldots, X_{(n)})$. The information in $\overline{X}$ about the parameters is the same as the information in a sample of size 1.

The Correlation Between $X - Y$ and $X + Y$

For any identically distributed random variables $X \,\&\, Y$, with a finite first and second moment, it is commonly known that the correlation coefficient between $X - Y$ and $X + Y$ is zero, since

$$C\mathrm{ov}(X - Y, X + Y) = E(X^2 - Y^2) = E(X^2) - E(Y^2) = 0.$$

However, if X and Y are iid $C(\theta, 1)$ then the correlation coefficient between $X - Y$ and $X + Y$ is undefined because $E(X^2) = E(Y^2) = \infty$.

This may be demonstrated by repeatedly taking two independent samples from the Cauchy distribution and calculating the correlation between the values of $X - Y$ and $X + Y$ for each two random samples. As seen in Figure 3 below, the correlation coefficient is unstable; fluctuating from -1 to +1. The fluctuation in these values is an indication of the undefined correlation coefficient. However, when the same procedure is used with the Normal distribution, the correlation coefficient remains approximately constant at zero.

This is another uncommonly known property of the Cauchy distribution. The correlation coefficient, which takes a value between -1 and 1 and measures the linear relation between $X-Y$ and $X+Y$, is undefined in the case of the Cauchy distribution.

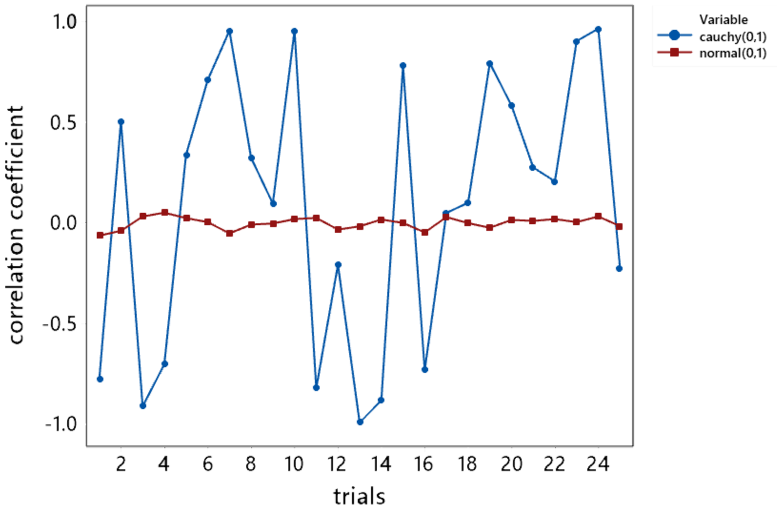


Figure (4) Correlation coefficient of X-Y and X+Y of random variables have standard Normal and Cauchy distributions.

Some Distributions Related to the Cauchy Distribution

The Cauchy distribution is related to other important distributions; some relationships are given below.

If U *is Uniform* $(-\frac{\pi}{2},\frac{\pi}{2})$, then $Y=\tan(U)$ *is* $C(0,1)$.

To show this result, let $Y=\tan(U)$, then the cdf of Y is

$$G_Y(y)=p(Y\le y)=p(\tan(U)\le y)=p(U\le \tan^{-1}(y))=F_U(\tan^{-1}(y)).$$

Therefore, the pdf of Y is

$$g_Y(y) = \frac{d}{dy}\left[F_U(\tan^{-1}(y))\right] = f_U(\tan^{-1}(y)) \times \frac{1}{1+y^2} = \frac{1}{\pi}\frac{1}{1+y^2}.$$

Thus, $Y = \tan(U) \sim C(0,1)$.

Using this result, random numbers from a Cauchy distribution can be simulated by taking random values from $\mathit{Uniform}(-\pi/2, -\pi/2)$, and taking the tan of the values. Note that if W is U(0,1) then $\pi W - \frac{\pi}{2}\ \mathit{Uniform}(-\pi/2, \pi/2)$. "Generating from U(0,1) " can be performed on many hand calculators.

$C(0,1)$ **is a t-distribution with one degree of freedom.**

This can be seen by looking at the pdf of the t-distribution with v degrees of freedom:

$$f(t) = \frac{\Gamma(\frac{v+1}{2})}{\Gamma(\frac{v}{2})\sqrt{\pi v}}\left(1+\frac{t^2}{v}\right)^{-\frac{v+1}{2}}, \qquad t \in \mathbb{R}, v > 0.$$

If $v = 1$, then $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. Thus, $f(t) = \frac{1}{\pi}\frac{1}{1+t^2}$, which is the pdf of $C(0,1)$.

The ratio of two independent standard normal variables is $C(0,1)$.

To show this result, let X, Y *be iid* $N(0,1)$, then their joint pdf is

$$f_{X,Y}(x,y) = \frac{1}{2\pi}e^{-\frac{1}{2}(x^2+y^2)} \quad x, y \in \mathbb{R}.$$

Let $W = \frac{X}{Y}$ (so that $X = WY$); the pdf of W is (using the transformation method):

$$f_W(w) = \int_{-\infty}^{\infty} |y|\ f_{X,Y}(wy, y)\, dy = 2\int_0^{\infty} y\ \frac{1}{2\pi}e^{-\frac{1}{2}(w^2y^2+y^2)}\, dy = \frac{1}{\pi}\int_0^{\infty} y\ e^{-\frac{1}{2}(w^2+1)y^2}\, dy$$

.

Thus,

$$f_W(w) = \frac{1}{\pi}\frac{1}{2(1+w^2)}\frac{e^{-\frac{1}{2}(w^2+1)y^2}}{-(1/2)}\Bigg|_0^{\infty} = \frac{1}{\pi}\frac{1}{(1+w^2)}(-0+1) = \frac{1}{\pi}\frac{1}{(1+w^2)}.$$

Thus, $\frac{X}{Y} \sim C(0,1)$ and $W = \frac{X}{|Y|}$ is also $C(0,1)$.

So, from a very well-known and very popular distribution (the normal distribution), we generate a very strange one (the Cauchy distribution).

Overlapping Between Cauchy and Normal Distribution

(Weitzman, 1970) introduced the concept of the overlapping coefficient; this coefficient measures the similarity, agreement, or closeness of two probability distributions. If $f_1(x)$ and $f_2(x)$ are two probability density functions of two continuous random variables, then the overlapping coefficient (Δ) of Weitzman is:

$$\Delta = \int_{-\infty}^{\infty} \min(f_1(x), f_2(x))dx .$$

It was simplified by [12] as follows:

$$\Delta = 1 - \frac{1}{2}\int_{-\infty}^{\infty} | f_1(x) - f_2(x) | dx.$$

It was called the “Indifference Zone”. It is simply the common area of the two distributions. The Cauchy distribution is symmetric about its location parameter and is similar in appearance to the normal distribution but with a heavy tail [3]. The overlapping between the Cauchy distribution and the Normal distribution is shown in Figure 4. The indifference zone is the shaded region; Δ is the area of this region.

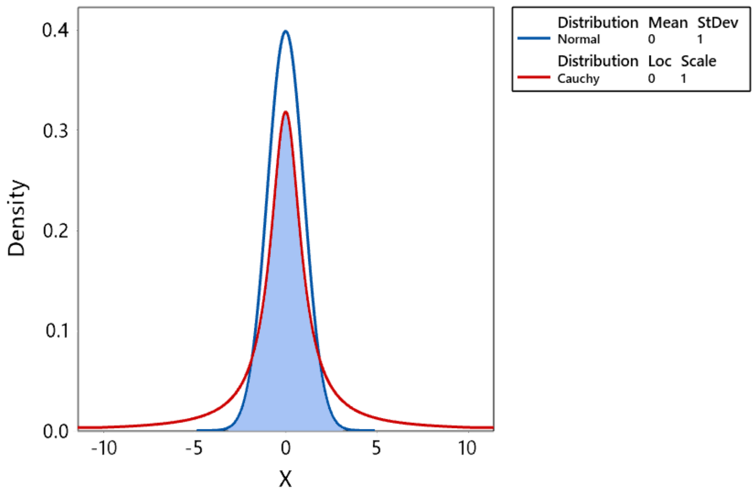


Figure (5) Overlapping between C(0,1) and N(0,1).

To find Δ, the intersection points first must be found:

If $f_1(x) = f_2(x)$ then

$$\frac{1}{\pi}\frac{1}{(1+x^2)} = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} \Rightarrow (1+x^2)e^{-\frac{x^2}{2}} = \sqrt{\frac{2}{\pi}}.$$

Solving this equation numerically, we obtained $x \approx \pm 1.85123$; let $x_1 = 1.85123$, and $x_2 = -1.85123$. Thus,

$$\Delta = \int_{-\infty}^{-1.85} \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx + \int_{-1.85}^{1.85} \frac{1}{\pi}\frac{1}{(1+x^2)}\,dx + \int_{1.85}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\,dx$$

$$= p(Z \le -1.85) + \left[\frac{1}{\pi}\tan^{-1}x\Big|_{-1.85}^{1.85}\right] + p(Z \ge 1.85)$$

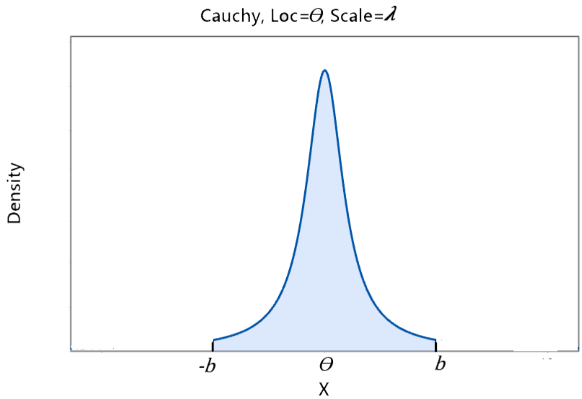
$$= \Phi(-1.85) + 0.6845 + [1 - \Phi(1.85)] = 0.0322 + 0.68452 + [1 - 0.9678] = 0.7489.$$

Truncated Cauchy Distribution:

The absence of finite moments limits the applications of the Cauchy distribution. To circumvent this restriction, a truncated version was proposed by (Johnson and Kotz, 1970). Consider the standard case, where $\theta = 0$ & $\lambda = 1$. A continuous random variable X is distributed as truncated

Cauchy, $TC(0,1,a,b)$ on the interval $[a,b]$ if the pdf of X is: $g(x)=\frac{1}{D}\frac{1}{\left(1+x^2\right)}$ $a\le x\le b$, where $D=\tan^{-1}(b)-\tan^{-1}(a)$.If $a=-b$, then $D=\tan^{-1}(b)-\tan^{-1}(-b)=2\tan^{-1}(b)$. The pdf is: $g(x)=\frac{1}{D}\frac{1}{\left(1+x^2\right)}$ $-b\le x\le b$, $\lambda>0$. This is the standard case of truncated Cauchy, see Figure 6 below.

Figure (6) Truncated Cauchy distribution on the interval [-b,b] where $\theta\in[-b,b]$.



Now, $\theta=0\ \&\ \lambda=1$ then

$$\mathrm{E}(X)=\int_{-b}^{b}x\ g(x)\,dx=\frac{1}{D}\int_{-b}^{b}\frac{x}{\left(1+x^2\right)}dx=0$$

$$\mathrm{E}(X^2)=\frac{1}{D}\int_{-b}^{b}\frac{x^2}{\left(1+x^2\right)}dx=\frac{2}{D}(\mathrm{b}-\tan^{-1}(b))=\frac{2b}{D}-1.$$

(Rohatgi, 1976) derived the first two moments in the standard case of truncated Cauchy. Moreover, an explicit expression of $\mathrm{E}(X^n)$, $n\ge 1$ has been provided by Stigler, 1974).

Now, if $b=10^{100}$, then this truncated distribution, $TC(0,1,-10^{100},10^{100})$ has all moments, and also the central limit theorem, as well as the weak and strong law of large numbers, can be applied to an iid random variables. For all practical purposes, this truncated distribution is not very different from the Cauchy distribution [15].

Using the Properties of Cauchy Distribution in the Classroom

The uses of the above properties and other related properties of the Cauchy distribution in the classroom are highlighted in this section. They can be a very good motivation for students to become more interested in statistical topics. The properties are classified into 5 parts.

(1) Properties Based on the pdf and cdf of the Cauchy distribution

- For N(0,1), the mean= the median=the mode=0, while for C(0,1), 0 is only the median and the mode. The maximum value of the pdf is $1/\sqrt{2\pi}$ for N(0,1) and $1/\pi$ for C(0,1), both occur at $x=0$.
- It is well known that π is the ratio of the circumference of any circle to its diameter; it is also the area of any circle with radius 1. The following are some other interesting facts about π:

If X is $N(0,1)$ then $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = 1 \Rightarrow \pi = \frac{1}{2}\left(\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} dx \right)^2$.

If X is $C(0,1)$ then $\int_{-\infty}^{\infty} \frac{1}{\pi} \frac{1}{1+x^2} dx = 1 \Rightarrow \pi = \int_{-\infty}^{\infty} \frac{1}{1+x^2} dx$.

- The cdf of C(0,1) is $F(x) = \int_{-\infty}^{x} \frac{1}{\pi} \frac{1}{1+u^2} du = \frac{1}{2} + \frac{1}{\pi} \tan^{-1}(x)$. Thus, $\frac{d}{dx} \tan^{-1}(x) = \frac{1}{1+x^2}$.
- It is well known that any continuous CDF, regarded as a random variable, has the $U(0,1)$ distribution. Thus, for this distribution, $Y = F(X) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1}(X)$ is $U(0,1)$.
- Based on that we have: $\tan^{-1}(X)$ is $U(-\frac{\pi}{2}, \frac{\pi}{2})$; $E(\tan^{-1}(X)) = 0$ and $Var(\tan^{-1}(X)) = \frac{\pi^2}{12}$. Thus, X has an undefined mean and variance but a function of it, $\tan^{-1}(X)$, has a finite mean and variance.

- $C(0,1)$ is a t-distribution with one degree of freedom.
- The minimal sufficient statistic for a random sample $X_1, X_2, ..., X_n$ from $C(\theta,1)$ is the order statistic $\left(X_{(1)}, X_{(2)}, ..., X_{(n)}\right)$. There is no sufficient statistic of lower dimension. Thus, the only way to summarize the data from Cauchy is to put the values in ascending order.

(2) Cauchy Distribution as a Counter Example

- The even moments of the Cauchy distribution exit but are infinite and the odd moments don't exist(undefined);

$$E(X^r) = \begin{cases} \infty & \text{if } r \text{ is even} \\ \text{undefined (indeterminate)} & \text{if } r \text{ is odd.} \end{cases}$$

- This is a good counter example of a random variable with undefined or infinite moments and of an undefined integral.
- The Central Limit Theorem simply states that under some conditions, the sample mean converges in distribution to normal. For the Cauchy distribution, the sample mean is Cauchy. The law of large number can't be applied here; the sample mean doesn't converge in probability.
- If X &Y are iid with finite first and second moment, then it is well-known that the correlation coefficient between $X-Y$ and $X+Y$ is zero, but if X &Y are iid C(0,1) then the correlation coefficient between $X-Y$ and $X+Y$ is undefined because $E(X^2) = E(Y^2) = \infty$.
- The numerical examples mentioned in section 2, are a suitable way to clarify the meaning of undefined mean for students.

(3) Generation of Cauchy Distribution from very Well-known Distribution

- If X is $U(-\frac{\pi}{2}, \frac{\pi}{2})$ then $Y = \tan(U) \sim C(0,1)$.
- If X_1, X_2 are iid $N(0,1)$ then $Y = \frac{X_1}{X_2} \sim C(0,1)$.

Thus, this distribution with very strange properties can be generated from a very well-known distribution. If X_1, X_2 have a finite mean, their ratio may have an undefined mean.

(4) Similarity Between the Cauchy and Normal distribution

- The overlapping, (Indifference Zone), of the two densities of C(0,1) and N(0,1) is the common region between the two regions under the graphs of the two densities.
- The overlapping coefficient Δ is the area of the Indifference Zone.
- The two densities are equal at $x \approx \pm 1.85123$ and $\Delta = 0.7489$. Thus, the two distributions have about 75% overlapping (common area).
- The overlapping coefficient (Δ) shows a very large similarity between the two distributions. This feature motivates the researchers to use truncated Cauchy distribution (75% of it) to overcome the problems caused by the non-existing moments of this distribution.

(5) The Use of Truncated Cauchy Instead of the of the Cauchy

The unsuitable properties of the Cauchy distribution such as the undefined or infinite moments and the inapplicability of the central limit theorem and laws of large numbers can be dealt with if we use truncated Cauchy distribution. Instead of the domain $(-\infty, \infty)$, we can use $(-a, a)$; the distribution is a truncated Cauchy; its pdf is

$$\frac{1}{c}\frac{1}{\pi(x^2+1)} \text{ for } -a \le x \le a, c = \int_{-a}^{a} \frac{1}{\pi\left(x^2+1\right)} dx = \frac{2}{\pi}\tan^{-1}(a).$$

- If a=1000, then $c = \int_{-1000}^{1000} \frac{1}{\pi\left(x^2+1\right)} dx = 0.99936$. Thus, the truncated Cauchy at $(-1000, 1000)$ is almost the same as C(0,1). Furthermore, this truncated distribution has finite moments, the central limit theorem and the laws of large numbers can be used.
- To simulate from Truncated Cauchy, just simulate from C(0,1) and ignore values that are larger than a or smaller than $-a$.

Conclusions:

In this paper, we have examined several interesting properties of the Cauchy distribution. Although all these properties are not new, some of them are not widely known or taught. We looked at the fact that the Cauchy distribution has an undefined mean, which makes the central limit theorem and the law of large numbers inapplicable. We also showed that the correlation between X-Y and X+Y for independent random variables X and Y from the Cauchy distribution is undefined. It is shown that the average of any random sample from the Cauchy distribution does not reveal any extra information compared to any individual value, hence, the order statistics is the minimal sufficient statistic. We also showed that the truncated Cauchy distribution can be utilized in a closed interval to overcome the issue of not having finite moments. Moreover, the relationships between the Cauchy distribution and other distributions are highlighted. We are certain that the contents of this paper will make significant contributions to educational Statistics.

ACKNOWLEDGMENT: The authors are very thankful to the referees for the comments and the editor of the journal for his efforts.

Funding and/or Conflicts of Interests/Competing Interests: The authors declare that there is no conflict of interest regarding the publication of this article. The article is not funded.

References:

Al-Saleh, M. F. Adam, H. Y. (2024). Moving Set Size Ranked Set Sampling. International Journal of Statistics and Probability, 13(4), to appear.

Al-Saleh, M.F. (2007). On the similarity structure of order statistics. Communications in Statistics—Theory and Methods, 36(7), 1433-1439.

Al-Saleh, M. F. (2008). On the "Independence of Trials-Assumption" in Geometric Distribution. International Journal of Mathematical Education in Science and Technology 39, 944 - 948.

Al-Saleh, M. F., Ali, D. & Dahshal, L. (2010). Toward a Reference Curve for the Grades of Each Course. International Journal of Mathematical Education in Science and Technology 41, 547-555.

Arnold, B. C., Balakrishnan, N., & Nagaraja, H. N. (2008). A first course in order

statistics. Society for Industrial and Applied Mathematics.

Bhat, B. R. (2007). Modern probability theory. New Age International.

Chakraborty, S. (2015). Generating discrete analogues of continuous probability

distributions-A survey of methods and constructions. Journal of Statistical

Distributions and Applications, 2(1), 1-30.

Florescu, I., & Tudor, C. A. (2013). Handbook of Probability. John Wiley & Sons.

Hampel, F., & Zurich, E. (1998). Is statistics too difficult? Canadian Journal of

Statistics, 26(3), 497-513.

Jensen, J. L. W. V. (1906). Sur les fonctions convexes et les inégalités entre les valeurs

moyennes. Acta Mathematica, 30(1), 175-193.

Johnson, N.L., Kotz, S. and Balakrishnan, N. (1994). Continuous Univariate

Distributions. Vol. 1, 2nd Edition, John Wiley & Sons Ltd., New York.

Maabreh, A. S. and Al-Saleh, M. F. (2023). Estimation of the Location Parameter of Cauchy Distribution Using Some Variations of the Ranked Set Sampling Technique. Mathematics and Statistics, 11(2), 263-276.

Nadarajah, S. (2011). Making the Cauchy work. Brazilian Journal of Probability and

Statistics, 25(1), 99-120.

Nolan, J. P. (2014). Financial modeling with heavy-tailed stable distributions. Wiley

Interdisciplinary Reviews: Computational Statistics, 6(1), 45-55.

Roe, B. P. (2012). Probability and statistics in experimental physics. Springer Science

& Business Media.

Rohatgi, V.K., 1976, An Introduction to Probability Theory and Mathematical

Statistics. (New York: John Wiley and Sons).

Stigler, S. M. (1974). Studies in the History of Probability and Statistics. XXXIII Cauchy and the witch of Agnesi: An historical note on the Cauchy distribution, Biometrika, 375-380.

Weitzman, M. S. (1970). Measures of overlap of income distributions of white and

Negro families in the United States (Vol. 3). US Bureau of the Census.