Notes About Cauchy Distribution[1]

The standard Cauchy distribution and the corresponding random variable ζ can be characterized in one of several equivalent ways:

(1) PDF (probability density function)

$$f_\zeta(x) = \frac{1}{\pi}\frac{1}{1+x^2},\ x \in (-\infty, +\infty).$$

(2) CDF (cumulative distribution function)

$$F_\zeta(x) = \frac{1}{2} + \frac{1}{\pi}\arctan(x),\ x \in (-\infty, +\infty).$$

(3) Characteristic functions

$$\varphi_\zeta(t) = e^{-|t|},\ t \in (-\infty, +\infty).$$

This is best established by first computing the characteristic function of the double-exponential (Laplace) distribution, with pdf $\propto e^{-|x|}$, and then using self-duality of the Fourier transform.

(4) $\zeta \stackrel{d}{=} X/Y$, where X and Y are iid standard Gaussian random variables and *the equality is in distribution*. This can be easily verified directly by using the formula for the pdf of the quotient of two independent random variables

$$f_{X/Y}(x) = \int_{-\infty}^{+\infty} f_X(xy)f_Y(y)|y|dy.$$

It is also a particular case of a more general fact: if the joint pdf of the random vector (X, Y) is rotation-invariant, then X/Y has the standard Cauchy distribution.

(5) $\zeta \stackrel{d}{=} \tan(V)$, where the random variable V is uniform on $(-\pi/2, \pi/2)$. This is a particular case of the *inverse* probability integral transform, also known under many other names, such as the `inverse transform sampling` or `inverse transformation method`: if U is uniform on $(0, 1)$ and ξ is a random variable with cdf F_ξ, then $\xi \stackrel{d}{=} F_\xi^{\leftarrow}(U)$, where $F_\xi^{\leftarrow}(t) = \min\{x : F_\xi(x) \geq t\}$ is the `left inverse` of F_ξ. Of course, if F_ξ is strictly increasing and continuous, then $F_\xi^{\leftarrow}(t) = F_\xi^{-1}(t)$, the usual inverse function.

Examples.

(1) $t_1 \stackrel{d}{=} \zeta$, where t_n is Student's t distribution with n degrees of freedom.

(2) $\zeta \stackrel{d}{=} -\zeta \stackrel{d}{=} 1/\zeta$: standard Cauchy distribution is symmetric and self-reciprocal.

(3) If (X, Y) is a Gaussian vector with mean zero and *correlation coefficient* $\rho \in [-1, 1]$, then

$$\mathbb{P}(XY > 0) = \frac{1}{2} + \frac{1}{\pi}\arctan\left(\frac{\rho}{\sqrt{1-\rho^2}}\right) = \frac{1}{2} + \frac{1}{\pi}\arcsin(\rho).$$

This follows after putting variances equal to one and writing $X \stackrel{d}{=} Z_1, Y \stackrel{d}{=} \rho Z_1 + \sqrt{1-\rho^2}Z_2$ for iid standard normal Z_1, Z_2.

(4) `Laser/lighthouse example.` If a laser is at $(0, 1)$ on the plane and is pointed at random at the x axis (that is, the angle with the y-axis is uniform on $(-\pi/2, \pi/2)$), then ζ is the distribution of the point on the x axis hit by the light. Instead of laser, one can think of a lighthouse with rotating light.

(5) If $\zeta_1, \ldots, \zeta_n$ are iid standard Cauchy and $S_n = \zeta_1 + \ldots + \zeta_n$, then

$$\frac{S_n}{n} \stackrel{d}{=} \zeta.$$

This is best seen using the characteristic functions. In other words, the standard Cauchy distribution is `stable of order one`.

(6) $\mathbb{E}|\zeta|^r < \infty$ if and only if $|r| < 1$, and $\mathbb{E}e^{t\zeta} = +\infty$ for every $t \neq 0$. In particular, ζ has no moment generating function and the random variable $W = e^\zeta$ has no moments: $\mathbb{E}W^p = +\infty$ for every $p \neq 0$.

[1] Sergey Lototsky, USC; updated July 31, 2023

The shifted and scaled Cauchy distribution/random variable $\zeta_{\mu,\sigma}$ is characterized by the `location parameter` $\mu \in \mathbb{R}$ (which coincides with the mode and the median) and the `scale parameter` $\sigma > 0$:

$$\zeta_{\mu,\sigma} \overset{d}{=} \mu + \sigma\zeta; \qquad \zeta \overset{d}{=} \frac{\zeta_{\mu,\sigma} - \mu}{\sigma}.$$

In the laser/lighthouse example, this will be the distribution of the point on the x axis if the light is at the point (μ, σ). Note that μ is NOT the expected value and σ is NOT the standard deviation: for Cauchy distribution, neither one is defined.

The corresponding pdf and cdf are

$$f_{\mu,\sigma}(x) = \frac{1}{\pi\sigma}\,\frac{1}{1 + \left(\frac{x-\mu}{\sigma}\right)^2} = \frac{1}{\pi}\,\frac{\sigma}{\sigma^2 + (x-\mu)^2}, \quad F_{\mu,\sigma}(x) = \frac{1}{2} + \frac{1}{\pi}\arctan\left(\frac{x-\mu}{\sigma}\right), \ x \in (-\infty, +\infty).$$

The characteristic function is

$$\varphi_{\mu,\sigma}(t) = e^{\mathfrak{i}t\mu - |t|\sigma},$$

where $\mathfrak{i} = \sqrt{-1}$ is the imaginary unit.

Complex parametrization. The location and scale parameters of the Cauchy distribution can be combined into a single complex number $z(\zeta_{\mu,\sigma}) = \sigma + \mathfrak{i}\mu$. Recall that, for a complex number $x + \mathfrak{i}y$, $x, y \in \mathbb{R}$, we have $\overline{x + \mathfrak{i}y} = x - \mathfrak{i}y$ (complex conjugation) and $|x + \mathfrak{i}y|^2 = (x + \mathfrak{i}y)(\overline{x + \mathfrak{i}y}) = x^2 + y^2$.

Then,[2] by direct computation, if Q is a Cauchy random variable with arbitrary location and scale parameters, then

$$z(-Q) = \overline{z(Q)}, \quad z(1/Q) = \frac{1}{\overline{z(Q)}} = \frac{z(Q)}{|z(Q)|^2},$$

and if R is a Cauchy random variable independent of Q, then $z(Q+R) = z(Q) + z(R)$.

Further Notes.

1. The study of the pdf of the standard Cauchy distribution, without the $1/\pi$ normalization, goes back to at least 1659 (by Fermat). The graph of the function has a number of remarkable properties and is known as the `witch`[3] `of Agnesi`, after the Italian mathematician `Maria Gaetana Agnesi` (1718–1799) included it as an example in her 1748 calculus textbook. Meanwhile, several people, including Liebnitz, integrated the corresponding Taylor expansion term-by-term to deduce the equality

$$\frac{\pi}{4} = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}.$$

2. In 1824, the French mathematician Siméon Denis Poisson (1781–1840) noticed that the sample mean from a Cauchy population has the same distribution as the population itself and therefore is NOT a good estimator of the location parameter, but apparently nobody took the result seriously, including Poisson himself. Later, another French mathematician Augustin-Louis Cauchy (1789–1857) made a similar observation[4] in connection with the least squares method. This time, more people paid attention. Now the distribution is named after Cauchy and is often used to test *robustness* of an estimation algorithm.

3. The laser/lighthouse example has major ramifications in physics, both classical and quantum, in particular, in connection with resonance phenomena. Accordingly, the Cauchy distribution can go under other names, such as `non-relativistic Breit-Wigner distribution` or `Lorentz distribution`.

---

[2] An immediate exercise: what happens if we consider $\mu + \mathfrak{i}\sigma$ instead?

[3] Reasons for this name are unclear; most probably due to ambiguities of the original Italian name used by Agensi herself.

[4] Cauchy, A.L. (1853), *Sur les résultats moyens d'observations de même nature, et sur les résultats les plus probables*, Comptes Rendus de l'Académie des Sciences, Paris, 37, 198–206.