The Cauchy distribution

Abdelwahid Benslimane

wahid.benslimane@gmail.com

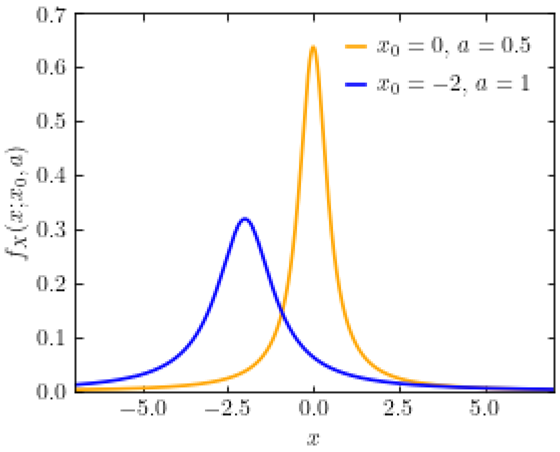
The Cauchy distribution is mainly used in physics, however, it is also involved in many machine learning techniques. For example, in estimation of distribution algorithms (EDA), which are stochastic optimization methods, replacing the univariate Gaussian distribution with a univariate Cauchy distribution as a search operator may be wise in some cases as, due to its heavy tails, it allows for larger jumps in the search space, and thus mitigates the problem of premature convergence of the Gaussian distribution when the population is far from the optimum.

For further details, you can read the following publication: https://researchcommons.waikato.ac.nz/bitstream/handle/10289/10506/cec2016.pdf?sequence=6. If you don't know what an EDA is, you can consult the wilipedia page dedicated to it: https://en.wikipedia.org/wiki/Estimation_of_distribution_algorithm.

The Cauchy distribution admits 2 parameters, a scale parameter a and a position parameter x_0 which allows to translate it horizontally. In general, the distribution is centered in 0. A random variable X follows the Cauchy distribution $\mathcal{C}(a, x_0)$ if it is absolutely continuous and has as its density function :

$$f_X(x; x_0, a) = \frac{a}{\pi} \frac{1}{(x-x_0)^2+a^2}$$

Below, the representative curves of the density function of the Cauchy distribution for different values of x_0 and a.



The integrals that define the expectation and variance for this distribution do not converge, so that a Cauchy random variable has neither expected value (1^{st} order moment), nor variance (2^{nd} order moment), nor any moment, and that is what we will examine here.

A random variable X has an expected value only if it is integrable, or in other words, if the mathematical expectation of its absolute value is finite.

In the following, we will consider the Cauchy distribution centered in 0 because it is in this form that it is used in the vast majority of situations. Moreover the calculation of E(|X|) is in this case very simple, although it is not much more complicated to proceed to the same calculation for $x_0 \neq 0$.

$E(|X|) = \int_{-\infty}^{+\infty} \frac{a}{\pi} \frac{|x|}{x^2+a^2} \, \mathrm{d}x$

It is the integral of an even function, thus we can write :

$E(|X|) = 2 \int_0^{+\infty} \frac{a}{\pi} \frac{x}{x^2+a^2} \, \mathrm{d}x = \frac{a}{\pi} \int_0^{+\infty} \frac{2x}{x^2+a^2} \, \mathrm{d}x$

Let be the function $U(x) = x^2 + a^2$, then its derivative $U'(x) = 2x$, and therefore we can write:

$E(|X|) = \frac{a}{\pi} \int_0^{+\infty} \frac{U'(x)}{U(x)} \, \mathrm{d}x$

As $\frac{U'}{U}$ is the derivative of $\ln(U)$, it turns out that:

$E(|X|) = \frac{a}{\pi} [\ln(x^2 + a^2)]_0^{+\infty}$

As $\ln(x^2 + a^2) \to +\infty$ when $x \to +\infty$, the integral $\int_{-\infty}^{+\infty} \frac{a}{\pi} \frac{|x|}{x^2+a^2} \, \mathrm{d}x$ diverges and X has no expected value.

X does not belong to the space of integrable random variables: $X \notin \mathcal{L}^1$.

$\mathcal{L}^2$ is the space of square integrable random variables (those which admit a variance). As $\mathcal{L}^2 \subset \mathcal{L}^1, X \notin \mathcal{L}^2 => X$ has no variance nor any moment.