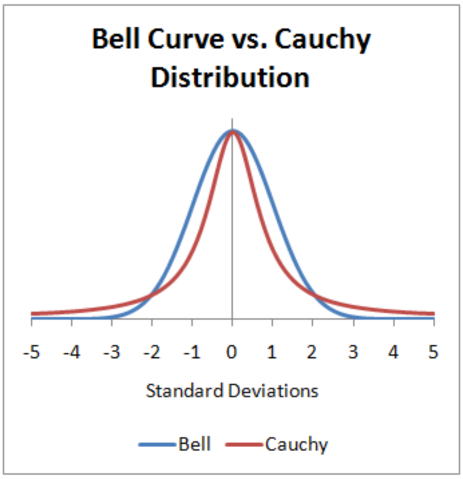
Why does the Cauchy distribution have no mean?

From the distribution density function we could identify a mean (=0) for Cauchy distribution just like the graph below shows. But why do we say Cauchy distribution has no mean?



mathematical-statistics pdf

edited Sep 11 '12 at 21:29 gung ♦ 95.8k 28 221 467

asked Sep 10 '12 at 15:28 Flying pig 1,719 5 23 26

2 I recommend the reference Cabeza G., U. A.. (2013). La Media de la Distribución de Cauchy. In the blog Apoyo en Matemáticas about the mean of Cauchy distribution. – user26162 May 27 '13 at 12:57

Se my answer here: stats.stackexchange.com/questions/94402/... – kjetil b halvorsen Dec 7 '15 at 14:50

10 Answers

You can mechanically check that the expected value does not exist, but this should be physically intuitive, at least if you accept Huygens' principle and the Law of Large Numbers. The conclusion of the Law of Large Numbers fails for a Cauchy distribution, so it can't have a mean. If you average n independent Cauchy random variables, the result does not converge to 0 as $n \to \infty$ with probability 1. It stays a Cauchy distribution of the same size. This is important in optics.

The Cauchy distribution is the normalized intensity of light on a line from a point source. Huygens' principle says that you can determine the intensity by assuming that the light is re-emitted from any line between the source and the target. So, the intensity of light on a line 2 meters away can be determined by assuming that the light first hits a line 1 meter away, and is re-emitted at any forward angle. The intensity of light on a line n meters away can be expressed as the n-fold convolution of the distribution of light on a line 1 meter away. That is, the sum of n independent Cauchy distributions is a Cauchy distribution scaled by a factor of n.

If the Cauchy distribution had a mean, then the 25th percentile of the n-fold convolution divided by n would have to converge to 0 by the Law of Large Numbers. Instead it stays constant. If you mark the 25th percentile on a (transparent) line 1 meter away, 2 meters away, etc. then these points form a straight line, at 45 degrees. They don't bend toward 0.

This tells you about the Cauchy distribution in particular, but you should know the integral test because there are other distributions with no mean which don't have a clear physical interpretation.

edited Sep 10 '12 at 16:37 answered Sep 10 '12 at 16:20

Douglas Zare
7,709 1 23 40

30 +1 Now *there* is an illuminating answer :-) (sorry). By the way, the principle is named for Christiaan Huygens, not Huygen. Huygens was the first to appreciate new developments in probability published in the 1650's by Pascal (based on his letters with Fermat): it was Huygens' account of these ideas (1657), including that of expectation, which originally got probability theory on a mathematical footing and paved the way for the seminal (posthumous) treatise of Jakob Bernoulli (*Ars Conjectandi*, 1713). – whuber ♦ Sep 10 '12 at 16:26

1 The amplitudes are propagated, not the intensities. – Doru Constantin Jul 22 '15 at 14:10

Answer added in response to @whuber's comment on Michael Chernicks's answer (and re-written completely to remove the error pointed out by whuber.)

The value of the integral for the expected value of a Cauchy random variable is said to be undefined because the value can be "made" to be anything one likes. The integral

$$\int_{-\infty}^{\infty} \frac{x}{\pi(1+x^2)}\,\mathrm{d}x$$

(interpreted in the sense of a Riemann integral) is what is commonly called an improper integral and its value must be computed as a limiting value:

$$\int_{-\infty}^{\infty} \frac{x}{\pi(1+x^2)}\,\mathrm{d}x = \lim_{T_1\to-\infty}\lim_{T_2\to+\infty}\int_{T_1}^{T_2} \frac{x}{\pi(1+x^2)}\,\mathrm{d}x$$

or

$$\int_{-\infty}^{\infty} \frac{x}{\pi(1+x^2)}\,\mathrm{d}x = \lim_{T_2\to+\infty}\lim_{T_1\to-\infty}\int_{T_1}^{T_2} \frac{x}{\pi(1+x^2)}\,\mathrm{d}x$$

and or course, both evaluations should give the same finite value. If not, the integral is said to be undefined. This immediately shows why the mean of the Cauchy random variable is said to be undefined: the limiting value in the inner limit diverges.

The Cauchy principal value is obtained as a single limit:

$$\lim_{T\to\infty}\int_{-T}^{T} \frac{x}{\pi(1+x^2)}\,\mathrm{d}x$$

instead of the double limit above. The principal value of the expectation integral is easily seen to be 0 since the limitand has value 0 for all T. But this cannot be used to say that the mean of a Cauchy random variable is 0. That is, the mean is defined as the value of the integral in the usual sense and not in the principal value sense.

For $\alpha > 0$, consider instead the integral

$$\begin{aligned}\int_{-T}^{\alpha T} \frac{x}{\pi(1+x^2)}\,\mathrm{d}x &= \int_{-T}^{T} \frac{x}{\pi(1+x^2)}\,\mathrm{d}x + \int_{T}^{\alpha T} \frac{x}{\pi(1+x^2)}\,\mathrm{d}x\\ &= 0 + \left.\frac{\ln(1+x^2)}{2\pi}\right|_{T}^{\alpha T}\\ &= \frac{1}{2\pi}\ln\left(\frac{1+\alpha^2T^2}{1+T^2}\right)\\ &= \frac{1}{2\pi}\ln\left(\frac{\alpha^2+T^{-2}}{1+T^{-2}}\right)\end{aligned}$$

which approaches a limiting value of $\dfrac{\ln(\alpha)}{\pi}$ as $T\to\infty$. When $\alpha=1$, we get the principal value 0 discussed above. Thus, we cannot assign an unambiguous meaning to the expression

$$\int_{-\infty}^{\infty} \frac{x}{\pi(1+x^2)}\,\mathrm{d}x$$

without specifying how the two infinities were approached, and to ignore this point leads to all sorts of complications and incorrect results because things are not always what they seem when the milk of principal value masquerades as the cream of value. This is why the mean of the Cauchy random variable is said to be undefined rather than have value 0, the principal value of the integral.

If one is using the measure-theoretic approach to probability and the expected value integral is defined in the sense of a Lebesgue integral, then the issue is simpler. $\int g$ exists only when $\int |g|$ is finite, and so $E[X]$ is undefined for a Cauchy random variable X since $E[|X|]$ is not finite.

edited Sep 10 '12 at 19:16 answered Sep 10 '12 at 16:24

Dilip Sarwate

26.1k 1 43 123

8 The evaluation of the middle integral is incorrect: it's zero, not a logarithm. The problem actually lies with evaluating the two limits implicit in the infinite integrals. – whuber ♦ Sep 10 '12 at 16:28

@whuber Thanks for pointing out the error. I have completely re-written my answer and your comment no longer applies. – Dilip Sarwate Sep 10 '12 at 18:53

+1 I love the improved exposition: thank you! – whuber ♦ Sep 10 '12 at 19:05

I don't understand why the expectation of the ratio doesn't exist. If X and Y are jointly normally distributed with mean different than zero, then the mean of $Z = \frac{X}{Y}$ is given by $\int\int \frac{x}{y} p(x,y)\,dxdy$, what am I missing? – Royi Apr 13 '15 at 18:53

1 @Drazick Look into whether your integral *exists* at all. In general, if the density of X is continuous in a neighborhood of 0, E[X^{-1}]$ does not exist. – Dilip Sarwate Apr 14 '15 at 1:36

While the above answers are valid explanations of why the Cauchy distribution has no expectation, I find the fact that the ratio X_1/X_2 of two independent normal $\mathcal{N}(0,1)$ variates is Cauchy just as illuminating: indeed, we have

$$\mathbb{E}\left[\frac{|X_1|}{|X_2|}\right] = \mathbb{E}\left[|X_1|\right] \times \mathbb{E}\left[\frac{1}{|X_2|}\right]$$

and the second expectation is $+\infty$.

edited Oct 21 '12 at 20:31

answered Oct 21 '12 at 19:45

Xi'an

42.4k 5 78 270

Is $\left|\frac{X_1}{X_2}\right|$ a 'folded' Cauchy random variable when I know that $\frac{X_1}{X_2}$ is standard Cauchy? How can one find the distribution of $\left|\frac{X_1}{X_2}\right|$? – StubbornAtom Dec 16 '17 at 11:54

Yes, this is the absolute value of a Cauchy variate, which has thus the density $f(x)+f(-x)$ over the positive real numbers. – Xi'an Dec 16 '17 at 12:40

Oh that was too easy. Thanks. – StubbornAtom Dec 16 '17 at 13:49

The Cauchy has no mean because the point you select (0) is not a mean. It is a *median* and a *mode*. The mean for an absolutely continuous distribution is defined as $\int xf(x)dx$ where f is the density function and the integral is taken over the domain of f (which is $-\infty$ to ∞ in the case of the Cauchy). For the Cauchy density, this integral is simply not finite (the half from $-\infty$ to 0 is $-\infty$ and the half from 0 to ∞ is ∞).

edited Oct 2 '16 at 1:15

Firebug

5,823 2 23 63

answered Sep 10 '12 at 15:36

Michael Chernick

32.3k 6 49 107

8 I'm not criticizing you, @Dilip: I'm augmenting your observation. What is very interesting is that the existence of a zero principal value might tempt us to define the mean of the Cauchy distribution (or the mean of any RV) as the principal value of the integral. This probes much more deeply into the nature of this question, which is glossed over by declaring that the integral is either infinite or undefined: namely, *why doesn't the principal value work*? Why would it not be legitimate to use that as a mean? – whuber ♦ Sep 10 '12 at 16:03

5 @whuber It is also interesting that if you truncate the integral at -a and +a for any a>0 you get 0. So taking the limit as a approaches ∞ of the symmetric integral gives 0. Another reason to ask why isn't 0 the mean. – Michael Chernick Sep 10 '12 at 16:08

10 @whuber: I take your last question in your penultimate remark to be rhetorical; at any rate we want absolute convergence and "the" reason in my mind is that we want things to behave like areas. In particular, we need to be able to chop things (functions) into pieces and rearrange them at will without disturbing the answer we obtain. We cannot do this chopping and rearranging for a linear function wrt a Cauchy distribution, so we must insist that its mean does not exist. – cardinal ♦ Sep 10 '12 at 16:24

9 That, @cardinal, is a good answer! I wasn't merely being rhetorical, because the question itself asks "why do we say [the] Cauchy distribution has no mean?" Asserting the expectation is undefined may satisfy the uncurious, but the possibility that a reasonable alternative definition of the integral may exist--and yields an intuitively correct answer!--ought to trouble people. Your answer is close to what I had in mind, but it's still incomplete. I think a satisfactory answer would identify important theorems of statistical theory that fail when we work with conditionally convergent integrals. – whuber ♦ Sep 10 '12 at 16:39

7 @Dilip I thought so too, but upon reflection find this to be a little more challenging than you seem to suggest. For instance, there's no problem with the Central Limit Theorem: requiring a variance automatically guarantees an expectation, of course. And a lot of theorems are proven using Chebyshev's Inequality, where once more we're guaranteed a mean. So I really am curious: what are the big theorems used in the practice of statistics where we really have to be cognizant of the problems with conditionally convergent, but not convergent, expectations? – whuber ♦ Sep 10 '12 at 19:04

The Cauchy distribution is best thought of as the uniform distribution on a unit circle, so it would be surprising if averaging made sense. Suppose f were some kind of "averaging function". That is, suppose that, for each finite subset X of the unit circle, $f(X)$ was a point of the unit circle. Clearly, f has to be "unnatural". More precisely f cannot be equivariant with respect to rotations. To obtain the Cauchy distribution in its more usual, but less revealing, form, project the unit circle onto the x-axis from (0,1), and use this projection to transfer the uniform distribution on the circle to the x-axis.

To understand why the mean doesn't exist, think of x as a function on the unit circle. It's quite easy to find an infinite number of disjoint arcs on the unit circle, such that, if one of the arcs has length d, then x > 1/4d on that arc. So each of these disjoint arcs contributes more than 1/4 to the mean, and the total contribution from these arcs is infinite. We can do the same thing again, but with x < -1/4d, with a total contribution minus infinity. These intervals could be displayed with a diagram, but can one make diagrams for Cross Validated?

answered Sep 11 '12 at 21:17
David Epstein
141 2

1 Welcome to the site, @DavidEpstein. You can make images w/ your preferred software & upload them into your answer by clicking the little picture icon (to launch the wizard) above the answer field. Unfortunately however, you need >=10 rep to do so. I'm sure you'll have that soon enough; in the interim, if you can post the image anywhere else on the internet & post a link to it in your answer, a higher rep user can fetch it & post it for you. – gung ♦ Sep 11 '12 at 21:27

2 I wasn't aware of Cauchy being interpreted as a uniform on a circle but it certainly makes sense. A topological argument shows that there can be no continuous function on a circle that has the properties of an averaging function. – johnny Oct 22 '12 at 6:29

The mean or expected value of some random variable X is a Lebesgue integral defined over some probability measure P:

$$EX = \int XdP$$

The nonexistence of the mean of Cauchy random variable just means that the integral of Cauchy r.v. does not exist. This is because the tails of Cauchy distribution are heavy tails (compare to the tails of normal distribution). However, nonexistence of expected value does not forbid the existence of other functions of a Cauchy random variable.

edited Sep 12 '12 at 16:01
Michael Chernick
32.3k 6 49 107

answered Sep 10 '12 at 15:43
Tomas
793 6 15

5 The tails are "heavy" in the sense that they do not decay fast enough in either direction to cause the integral to converge. This concept has nothing to do with normal distributions (or any reference distribution). – whuber ♦ Sep 10 '12 at 15:51

4 Yes, thanks for this correction. I have not intended to imply any rigour connection between heavy tails and normal distribution. However, I think that comparing normal distribution (with light tails) and heavy-tailed distribution visually makes (not always) a bit easier to grasp the concept of the "heavy" tails. – Tomas Sep 10 '12 at 16:06

Here is more of a visual explanation. (For those of us that are math challenged.). Take a cauchy distributed random number generator and try averaging the resulting values. Here is a good page on a function for this. https://math.stackexchange.com/questions/484395/how-to-generate-a-cauchy-random-variable You will find that the "spikiness" of the random values cause it to get larger as you go instead of smaller. Hence it has no mean.

edited Apr 13 '17 at 12:19
Community ♦
1

answered Feb 4 '14 at 3:05
Paul
51 1 1

Just to add to the excellent answers, I will make some comments about why the nonconvergence of the integral is relevant for statistical practice. As others have mentioned, if we allowed the principal value to be a "mean" then the slln are not anymore valid! Apart from this, think about the implications of the fact that , in practice, all models are approximations. Specifically, the Cauchy distribution is a model for an unbounded random variable. In practice, random variables are bounded, but the bounds are often vague and uncertain. Using unbounded models is way to alleviate that, it makes unnecessary the introduction of unsure (and often unnatural) bounds into the models. But for this to make sense, important aspects of the problem should not be affected. That means that, if we were to introduce bounds, that should not alter in important ways the model. But when the integral is nonconvergent that does not happen! The model is unstable, in the sense that the expectation of the RV would depend on the largely arbitrary bounds. (In applications, there is not necessarily any reason to make the bounds symmetric!)

For this reason, it is better to say the integral is divergent than saying it is "infinite", the last being close to imply some definite value when no exists!

edited Dec 7 '15 at 14:53

answered Sep 11 '12 at 22:35
kjetil b halvorsen
20.8k 9 61 147

I wanted to be a bit picky for a second. The graphic at the top is wrong. The x-axis is in standard deviations, something that does not exist for the Cauchy distribution. I am being picky because I use the Cauchy distribution every single day of my life in my work. There is a practical case where the confusion could cause an empirical error. Student's t-distribution with 1 degree of freedom is the standard Cauchy. It will usually list various sigmas required for significance. These sigmas are NOT standard deviations, they are probable errors and mu is the mode.

If you wanted to do the above graphic correctly, either the x-axis is raw data, or if you wanted them to have equivalent sized errors, then you would give them equal probable errors. One probable error is .67 standard deviations in size on the normal distribution. In both cases it is the semi-interquartile range.

Now as to an answer to your question, everything that everyone wrote above is correct and it is the mathematical reason for this. However, I suspect you are a student and new to the topic and so the counter-intuitive mathematical solutions to the visually obvious may not ring true.

I have two nearly identical real world samples, drawn from a Cauchy distribution, both have the same mode and the same probable error. One has a mean of 1.27 and one has a mean of 1.33. The one with a mean of 1.27 has a standard deviation of 400, the one with the mean of 1.33 has a standard deviation of 5.15. The probable error for both is .32 and the mode is 1. This means that for symmetric data, the mean is not in the central 50%. It only takes ONE additional observation to push the mean and/or the variance outside significance for any test. The reason is that the mean and the variance are not parameters and the sample mean and the sample variance are themselves random numbers.

The simplest answer is that the parameters of the Cauchy distribution do not include a mean and therefore no variance about a mean.

It is likely that in your past pedagogy the importance of the mean was in that it is usually a sufficient statistic. In long run frequency based statistics the Cauchy distribution has no sufficient statistic. It is true that the sample median, for a Cauchy distribution with support over the entire reals, is a sufficient statistic, but that is because it inherits it from being an order statistic. It is sort of coincidentally sufficient, lacking an easy way to think about it. Now in Bayesian statistics there is a sufficient statistic for the parameters of the Cauchy distribution and if you use a uniform prior then it is also unbiased. I bring this up because if you have to use them on a daily basis, you have learned about every way there is to perform estimations on them. This is due to the fact that inference between long run frequency based statistics and Bayesian statistics runs in opposite directions.

There are no valid order statistics that can be used as estimators for truncated Cauchy distributions, which are what you are likely to run into in the real world, and so there is no sufficient statistic in frequency based methods for most but not all real world applications.

What I suggest is to step away from the mean, mentally, as being something real. It is a tool, like a hammer, that is broadly useful and can usually be used. Sometimes that tool won't work.

A mathematical note on the normal and the Cauchy distributions. When the data is received as a time series, then the normal distribution only happens when errors converge to zero as t goes to infinity. When data is received as a time series, then the Cauchy distribution happens when the errors diverge to infinity. One is due to a convergent series, the other due to a divergent series. Cauchy distributions never arrive at a specific point at the limit, they swing back and forth across a fixed point so that fifty percent of the time they are on one side and fifty percent of the time on the other. There is no median reversion.

edited Sep 11 '12 at 21:42 answered Sep 11 '12 at 21:20
DE Harris
19 2

7 There is some confusion in this response! For instance, it says " Now in Bayesian statistics there is a sufficient statistic for the parameters of the Cauchy distribution and if you use a uniform prior then it is also unbiased. ". It is difficult to make any sense of this! First, the Frequentist and Bayesian concepts of sufficiency are very close (and I believe can differ only in some strange, infinite-dim sample spaces, so for the real line are the same). There is no sufficient statistic for the Cauchy model, of fixed dimension!, simply (the complete data obviously are sufficient). – kjetil b halvorsen Sep 12 '12 at 3:39

To put it simply, the area under the curve approaches infinity as you zoom out. If you sample a finite region, you can find a mean for that region. However, there is no mean for infinity.

answered Nov 10 '12 at 17:38
Paul
1

6 The area under the PDF equals 1, by definition, so you must mean something else by "the curve." What is it? – whuber ♦ Nov 10 '12 at 23:29
