Decision theory meets the Witch of Agnesi

J. McKenzie Alexander
Department of Philosophy, Logic and Scientific Method
London School of Economics and Political Science

26 April 2010

In the course of history, many individuals have the dubious honor of being remembered primarily for an eponym of which they would disapprove. How many are aware that Joseph-Ignace Guillotin actually opposed the death penalty? Another noteable case is that of Maria Agnesi, an Italian woman of privileged, but not noble, birth who excelled at mathematics and philosophy during the 18th century. In her treatise of 1748, *Instituzioni Analitche*, she provided a comprehensive summary of the current state of knowledge concerning both integral calculus and differential equations. Later in life she was elected to the Balogna Academy of Sciences and, in 1762, was consulted by the University of Turin for an opinion on the work of an up-and-coming French mathematician named Joseph-Louis Lagrange.

Yet what Maria Agnesi is best remembered for now is a geometric curve to which she lent her name. That curve, given by the equation $y = \frac{a^3}{a^2+x^2}$, for some positive constant a, was known to Italian mathematicians as *la versiera di Agnesi*. In 1801, John Colson, the Cambridge Lucasian Professor of Mathematics, misread the curve's name during translation as ***l'avversiera** di Agnesi*. An unfortunate mistake, for in Italian "l'avversiera" means "wife of the devil". Ever since, that curve has been known to English geometers as the "Witch of Agnesi". I suspect Maria, who spent the last days of her life as a nun in Milan, would have been displeased.

The name of Maria's graceful curve seems doubly misfortunate to contemporary eyes, for it has no apparent bewitching properties. To those acquainted with the delights of Cantor's paradise, it is a rather dull specimen. It is absolutely continuous, smooth, symmetric, and it converges to 0 in the limit as x approaches both ∞ and $-\infty$. Yet I will argue that the Witch of Agnesi deserves her name when she encounters contemporary decision theory. As we will see, this 250-year old curve presents us with a problem for which we still have no answer.

1. The witch casts her spell

Consider the following proposition: at the point $(1,1)$, I draw a line to the x-axis at an angle θ selected at random from $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. Let x denote the point where the line intersects the horizontal axis. If $x > 0$, you win \$$x$, but if $x < 0$, you have to pay \$$x$. (See figure 1 for an illustration.) How much should you be willing to pay to participate in this gamble?

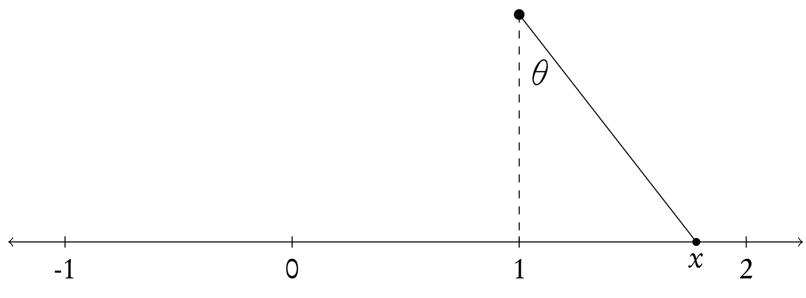


Figure 1: Illustration of the proposed gamble

One natural way to think about how to put a price on this gamble is to determine the probability distribution induced over the real line by the above construction. Basic algebra, and a little calculus, yields:

$$\theta = \arctan\left(\frac{x-1}{1}\right)$$

and so

$$\frac{d\theta}{dx} = \frac{1}{1+(x-1)^2}$$

The value $x-1$ appears, rather than x, because we have to adjust for the fact that we are drawing the line through the point $(1,1)$ instead of $(0,1)$. Because the angle θ is drawn from the open interval $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, dividing by π gives the probability distribution we need: $\frac{1}{\pi} \cdot \frac{1}{1+(x-1)^2}$.

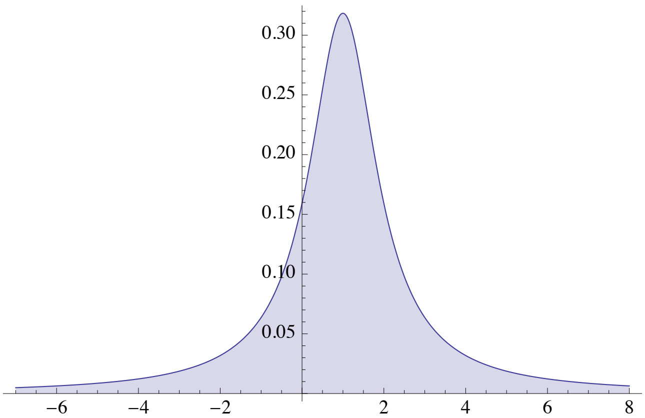


Figure 2: The probability distribution

Figure 2 provides a plot the probability distribution. There are two things worth noting. First, although the distribution appears similar to the normal distribution, in that both are "bell-shaped", it has considerably heavier tails. Second, the distribution has the form of the Witch of Agnesi, with the only change being that it is shifted one unit to the right and multiplied by the renormalisation constant of $\frac{1}{\pi}$. When viewed as a probability distribution, the Witch is better known as the "Cauchy distribution" even though knowledge of the geometric curve defining it predated Cauchy by a considerable margin. (To tell the truth, knowledge of the curve predated Agnesi by a considerable margin as well, for it was studied by Fermat in 1630.) In physics, the Cauchy distribution appears as a common approximation to the solution of the resonance equation for the damped, driven harmonic oscillator.

Given the nature of the distribution, it looks as though the gamble it represents is one worth taking. Suppose that the distribution were centered at 0, rather than 1. In that case, the symmetry of the distribution around 0 would mean that the chance of winning an amount in the range $(n-\varepsilon, n+\varepsilon)$, for some $n, \varepsilon > 0$, exactly equaled the chance of losing an amount in the range $(-(n-\varepsilon), -(n+\varepsilon))$. In other words, the symmetry of the distribution around 0 generates a fair bet. Shifting the distribution one unit to the right, making it symmetric around 1, only serves to tilt the gamble in our favour: the chance of winning an amount in the range $(n-\varepsilon, n+\varepsilon)$ now exceeds the chance of losing an amount in the range $(-(n-\varepsilon), -(n+\varepsilon))$. Or, to put the point in more flowery language: if the Witch were sitting at 0, a rational agent would be indifferent towards playing, but when the Witch sits at 1, she looks positively enticing. Although we do not yet know how much we should be willing to *pay* to play this game, it seems certain that we should be willing to pay *something*. The question becomes, how much?

The natural way to figure out what we should be willing to pay is to calculate the expected value of the gamble. Determining the expected value of the gamble requires calculating the value of the following improper integral:

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

Yet because the integrand has the following antiderivative

$$\int \frac{x}{1+(x-1)^2}\,dx = \arctan(x-1) + \frac{1}{2}\log(1+(x-1)^2)$$

it turns out that the integral yielding the expected value has the indeterminate form $\infty - \infty$, and does not exist.[1] Hence, curiously, the gamble has no expected value.

Gambles which are difficult to value are not unfamiliar to decision theory, but normally for different reasons. Perhaps the most familiar type of hard-to-value gamble concerns choice under uncertainty, where the outcomes are known but one does not have a probability distribution over them. In *that* case, it comes as no surprise that one cannot easily say what one would be willing to pay for the gamble. Another type of hard-to-value gamble concerns cases where the outcomes themselves are unknown. Suppose I hold in my hand a brown paper bag. What should you be willing to pay for it? In this case, the bag could literally contain *anything*, good or bad.[2] But our gamble here is hard to value for very different reasons. By construction, you *have* all the relevant probabilistic information, so this is not a decision problem under uncertainty. And you *know* what all of the possible outcomes are, for that is just a matter of either winning or losing a certain amount of money. It just happens to be the case that, nevertheless, no expected value exists.

A number of gambles lack expected values. Perhaps the most famous example is the one giving rise to the St. Petersburg paradox, which inspired Daniel Bernoulli in 1738 to develop expected utility theory in the first place. A more recent example is the Pasadena game, invented by Nover and Hájek in 2004.[3] In the Pasadena game, you flip a fair coin until it lands heads for the first time; if it lands heads on the nth toss, the outcome is \$$(-1)^{n-1}\cdot\frac{2^n}{n}$, where a positive value indicates the amount of money you receive and a negative value indicates the amount of money you have to pay. The interesting thing about the Pasadena game is that payoffs were carefully constructed so as to yield the terms of the alternating harmonic series. Consequently, it

---

[1] The problematic portion of the antiderivative is $\log(1+(x-1)^2)$, since that diverges as $x\to\infty$ and as $x\to-\infty$.

[2] Even an elephant, since the bag could contain a deed transferring ownership. And in a world run by absolute libertarians, where the legal system acts only to enforce agreements but not to judge which agreements are acceptable, the bag could contain arbitrarily bad outcomes, too: a document requiring the bearer to sacrifice his or her kidney, or worse.

[3] Harris Nover and Alan Hájek, ``Vexing Expectations,'' *Mind*, vol. 113, no. 450, pp. 237--249 (2004).

lacks an expected value because the possible payoffs have no natural order in which they are to be summed, and the conditionally convergent nature of the alternating harmonic series means that it can be made to converge to any value whatsoever (or even to diverge) by suitably rearranging the terms.

Yet it would be a mistake to think that simply because the Pasadena game lacks an expected value no principled way exists for assigning it one. In a recent paper, Kenny Easwaran[4] argued that the Pasadena game may be assigned the intuitively compelling value of $\log(2)$ using his method of "weak expectations".[5] The method is compelling because not only does it assign a plausible value to the Pasadena game, it does so in a way which respects dominance reasoning applied to variations of the game. Consider, for example, the Altadena game, defined in the same way as the Pasadena game except that the payoff for each outcome has $+1$ added to it. Dominance reasoning suggests that, regardless of how we value the Pasadena game, we ought to value the Altadena game as "one better".[6] And, indeed, Easwaran's method of weak expectations assigns a value of $\log(2) + 1$ to the Altadena game.

Given the success of weak expectations in valuing games which had previously proven hard to value, can we use weak expectations to value the Witch of Agnesi? Unfortunately, the answer is no. In what follows, I will first show that mathematical considerations prevent the method of weak expectations from applying. I will then consider alternative methods of valuing the Witch and argue that none of them is adequate because there is little reason to take their outcomes as providing a rational value. Finding rational grounds for valuing the Witch thus remains an open problem.

2. Why weak expections do not suffice.

In his proof that a weak expectation exists for the Pasadena game, Easwaran invokes the following form of the Weak Law of Large Numbers:

> Let $Y_1, Y_2, \ldots$ be [independent, identically distributed random variables] with
> $$y\Pr(|Y_i| > y) \to 0 \text{ as } y \to \infty.$$
> Let $S_n = Y_1 + \cdots + Y_n$ and let μ_n be the expectation of the random variable that is 0 whenever $|Y_1| > n$ and equal to Y_1 otherwise. Then $\frac{S_n}{n} - \mu_n \to 0$ in probability.

When this holds, and μ_n converges to ω as $n \to \infty$, the limit ω is said to be a *weak expectation*. From this we see that whether a weak expectation exists depends upon the convergence behaviour of the sum of n independent, identically distributed (i.i.d.) random variables. Let us consider this for the gamble under consideration.

To begin, take the simple case where X and Y are two i.i.d. random variables, each with a Cauchy distribution centered at 1. As we have seen, the probability density functions for X and Y are as follows:

---

4 "Strong and Weak Expectations" *Mind*, vol. 117, no. 467, pp. 633–641 (2008).

5 Some believe $\log(2)$ to be the intuitively correct value of the Pasadena game because that is the value to which the alternating harmonic series converges.

6 Although see Fine, "Evaluating the Pasadena, Altadena, and St. Petersburg Gambles" *Mind*, vol. 117, no. 467, pp. 613–632 (2008), in which it is argued that one may assign values to the Pasadena and Altadena games which contradict this dominance ordering without contradicting the axioms of utility theory.

$$p_X(x) = p_Y(x) = \frac{1}{\pi(1 + (x-1)^2)}.$$

Now consider the random variable $Z = X + Y$. Probability theory tells us that the density function of Z is given by the convolution of the densities of X and Y:

$$p_Z(z) = \frac{1}{\pi^2}\int_{-\infty}^{+\infty} \frac{1}{1 + (z-y-1)^2} \cdot \frac{1}{1+(y-1)^2}\, dy = \frac{2}{\pi(8 - 4z + z^2)}.$$

What, then, can be said the *mean* of X and Y? If we let $M = \frac{X+Y}{2} = \frac{1}{2}Z$ denote the mean, probability theory again says[7] that the density of M is related to the density of Z according to $p_M(x) = 2p_Z(2x)$. Hence we find that the mean of two i.i.d. random variables with a Cauchy distribution centered around 1 is

$$p_M(x) = 2 \cdot \frac{2}{\pi(8 - 4(2x) + (2x)^2)} = \frac{1}{\pi(1 + (x-1)^2}.$$

That is, the *mean* of *two* random variables has the same distribution as an *individual* random variable.

For the Cauchy distribution, this fact holds in general. It can be shown[8] that, if $X_1, X_2, \ldots, X_n$ are i.i.d. random variables with a Cauchy distribution, then the mean $S_n = (X_1 + \cdots + X_n)/n$ is also Cauchy distributed. This has several important implications for our problem. For one, notice that it implies the problem of valuing *repeated* plays of the gamble, where you are interested in your average payoff, is equivalent to the problem of valuing an *individual* play of the gamble. That is very odd.

This matters because, normally, repeating a gamble a large number of times is advantageous in the sense that the mean of the repeated plays of the gamble approximates, increasingly closely, the expected value of the gamble. (This is one of the messages of the central limit theorem.) Suppose we were considering a gamble defined by the uniform probability over $[0,1]$, where you just win the amount selected. The expected value of this gamble is, of course, $\frac{1}{2}$. Now suppose you were thinking about repeating the gamble 10, 100, or 1000 times. What does the distribution of your average winnings to look like? Figure 3(a) contrasts these three distributions. As the central limit theorem says, repeating the gamble makes your average earning approximately normally distributed around the expected value of the gamble and, *more importantly*, also significantly reduces the variance in your average earning. (The distribution is more sharply peaked around the expected value.) So you can use the sample mean as a measure of the expected value of the gamble. This is just basic statistics.

However, as figures 3(b) and 3(c) show, this does not happen with the Cauchy distribution. Figure 0 shows the empirical distribution (determined by simulation) of 100,000 draws from the symmetric Cauchy distribution centred at 1. Figure 0 shows the distribution of 100,000 sample means, where each sample consisted of 1,000 draws from the symmetric Cauchy distribution centered at 1. As noted earlier, increasing the sample size has no benefit whatsoever,

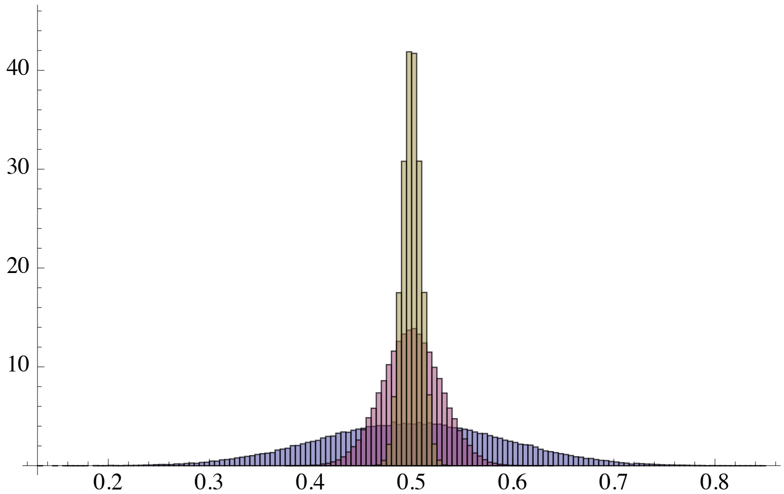
---

7 In general, if U and V are two continuous random variables with probability density functions $p_U(x)$ and $p_V(x)$, and it is the case that $U = aV$, then $p_U(x) = \frac{1}{a} p_V\left(\frac{x}{a}\right)$.

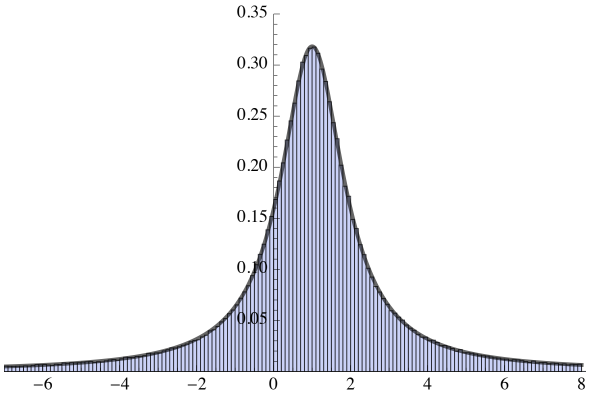
8 Allan Gut, *Probability: A Graduate Course* (New York: Springer, 2005).

because the distribution of sample means (of any size whatsoever) is *absolutely identical* to the distribution of an individual random variable. In short, attempting to figure out how much you should be willing to pay for the gamble by thinking about the distribution of your average earnings if you repeated the gamble n times gets you nowhere.

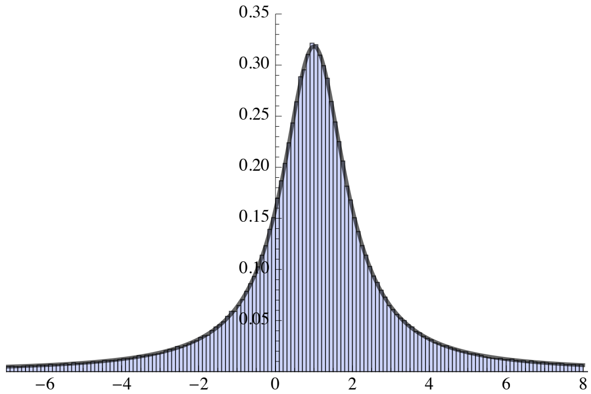
[Insert Figure 3 somewhere around here.]



(a) Distributions of S_{10}, S_{100}, and S_{1000}, based on the uniform distribution over $[0,1]$.



(b) The empirical distribution of 500,000 occurrences of a random variable drawn from the symmetric Cauchy distribution centered at 1.



(c) The empirical distribution of 500,000 occurrences of S_{1000}, based on the symmetric Cauchy distribution centered at 1.

Figure 3: An illustration of the failure of the central limit theorem for the Cauchy distribution.

Finally, notice the strong and weak laws of large numbers fail to hold as well. The strong law fails because it says the sample mean converges to the expected value of the distribution; but, as we have seen, the Cauchy distribution has no expected value![9] What is crucial for us, though, is that the failure of the weak law of large numbers to hold means no weak expectation exists for this particular game, either.

The reason why no weak expectation exists can be seen relatively easily, for it is connected to the fact that the distribution of S_n is an ordinary Cauchy distribution, for all n. In order for a weak expectation ω to exist, it needs to be the case that the sequence of random variables $\langle S_n \rangle$ converges in probability to ω. But the sequence $\langle S_n \rangle$ cannot converge in probability to ω because each S_n, as noted above, is distributed according to the original

[9] *Ibid.*, pp. 272–273.

Cauchy distribution. That means the probability of S_n being greater than r, for any real r, is given by the constant $\frac{1}{\pi}\int_r^{+\infty} \frac{1}{1+(x-1)^2}\, dx$, which does not depend upon n at all! Hence there can be no convergence in probability, and so we cannot use the method of weak expectations to assign a value to the gamble.

To summarise: when the Witch of Agnesi meets decision theory, we find ourselves in the curious position of having a gamble which clearly seems to be beneficial, but yet we are unable to determine what a rational agent should be willing to pay for the opportunity to play. Neither strong nor weak expectations do the trick. What other alternatives are there?

3. Four proposed solutions.

Bound the utility function. One natural response to the problem would be to argue that it arises simply because we allow for unbounded payoffs, both positive and negative. Unbounded payoffs have been known to be problematic ever since they were used to generate the St. Petersburg paradox. Similarly, unbounded payoffs, both positive *and* negative, played a crucial part in generating the problem of assigning a value to the Pasadena game. One could argue that a most important mistake was made in the very formulation of the gamble under discussion by explicitly assuming that unbounded payoffs, either in terms of money or utility, were possible. (In other words, the best way to avoid the problem is to not let the Witch into your house in the first place.)

Although I have some sympathy with the spirit in which this "solution" is offered, it largely misses the point. The best response to this argument, I think, can be found in Nover and Hájek's response to the same objection made with respect to the Pasadena game.[10] They employ an elegant divide-and-conquer approach, beginning with the question of what the complaint about unbounded payoffs (or utility) really means. If it means that people's utility functions are bounded *as a matter of conceptual necessity* (their italics), the claim seems ad hoc and unlikely. If it means that people's utility functions are bounded *as a matter of contingent fact*, it fails to appreciate that we are after an understanding of what it means to *rationally assign* a value to a gamble. Whether something is rational or not should not depend upon arbitrary or accidental facts about human nature, like whether or not real utility functions are bounded.

Use the limit of approximations to the expected value. The suggestion here is that although

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

diverges, one can consider the limit of ever-greater approximations.[11] In fact, it does turn out that

$$\lim_{n\to\infty} \int_{-(n-1)}^{n+1} \frac{x}{\pi(1+(x-1)^2)}\, dx = 1.$$

Note that the limits of integration run from $-(n-1)$ to $n+1$, rather than from $-n$ to n, as one might expect, because the point of symmetry is shifted one unit to the right.

This suggestion appears promising, but one crucial problem is that, because the expectation is not absolutely convergent, the limit one gets depends on the particular way one

10. "Vexing Expectations," pg. 248.

11. That is, use the Cauchy principal value of the improper integral.

constructs the approximation. Suppose you are particularly pessimistic and you thus decide to include twice as much of the negative tail as the positive tail. That is, you approximate the expectation by

$$\int_{-2(n-1)}^{n+1} \frac{x}{\pi(1+(x-1)^2)}\, dx.$$

If you take the limit as $n \to \infty$, the limit value turns out to be $1 - \frac{\log(4)}{2\pi}$. If you are particularly optimistic and include twice as much of the positive tail as the negative tail, you will find the limit to be $1 + \frac{\log(4)}{2\pi}$. This serves to undercut the plausibility of using the approximations to the expected value as the actual value of the gamble.

Use the median. Although the gamble does not have a *expected value*, the underlying probability distribution does have a *median*, which equals 1. Using this as the value of the gamble gains some support from the fact that whenever a symmetric distribution has both a mean and a median, the value of those two coincide. Therefore one might argue that since we have a symmetric distribution which lacks a mean but has a median, we might as well propose to use the median of the distribution as the value of the gamble, since that *would* be its expected value *if* it had one.

The first problem with this suggestion is that is utterly *ad hoc*. What is the rational justification for using the median as the price, other than the fact that it exists? We need a more substantive reason than convenience for adopting the median as the value of the gamble.

The second problem with this solution is that it relies heavily on the symmetry of the underlying probability distribution for its motivation. Even if we are willing to treat the median as a proxy for the nonexistent expected value, we must recognise that this solution is, at best, a stop-gap. It will not generalise to asymmetric distributions.

Consider the underlying mechanism. Another proposal points to the fact that, for the gamble under consideration, we have information about the causal process which generates the probability distribution. Recall that the amount we win or lose derives from an angle θ selected at random from $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. The expected value of θ is, of course, simply 0, and the corresponding amount won when $\theta = 0$ is \$1. Intuitively, this valuation seems correct. If the distribution, when centered at 0, yields a fair bet, we should be willing to pay nothing; and if the price is \$0 when centered at 0, the price should be \$1 when centered at \$1.

Two problems exist. First, this proposal departs from the traditional decision-theoretic point of view which holds that all of the information required to value the gamble is given by the probability distribution and the values of the outcomes. This proposal identifies a value but does not say why that value, and only that value, is rational.[12] Second, in this case we happened to know the underlying causal mechanism. If we had been ignorant of the mechanism, the problem would remain.

4. Stable distributions.

In the 1920s, the French mathematician Paul Lévy discovered a fascinating family of probability distributions while studying sums of independent and identifically distributed random

[12] Note that weak expectations in the Pasadena game suffer from the same problem: although the Pasadena game has a unique weak expectation, Fine has proved that one can assign any value whatsoever to the Pasadena game without contradicting the axioms of utility theory.

variables.[13] The defining property of these distributions is their stability, in the following sense: if X_1 and X_2 are two i.i.d. random variables of some type X, then X is said to be *stable* if for any positive constants a and b, there exist additional constants c and d, with $c > 0$, such that $aX_1 + bX_2$ is equal in distribution to $cX + d$.[14] When the constant $d = 0$, the distribution X is said to be *strictly stable*.

The most well-known example of stable distributions is the normal distribution. If X_1 and X_2 are two independent normally distributed random variables, then the sum $X_1 + X_2$ is also normally distributed, and the mean (and variance) of the sum is simply the sum of the means (and variances) of X_1 and X_2. This fact, often given as an exercise to students of probability theory, shows that the normal distribution is a stable distribution.

As we have seen previously, the Cauchy distribution is also stable.[15] It turns out that these two examples very nearly exhaust the set of stable distributions whose probability density functions can be written down in a closed-form (that is, using the traditional elementary functions of mathematical analysis).[16] Although one cannot write down the probability density functions for most stable distributions, they can be characterised by specifying numerical values for four parameters. These parameters control key properties of the distribution: the characteristic exponent α influences its shape, the skewness β determines its "tilt", the scale parameter γ determines how spread out the distribution is over the real line, and a location parameter δ shifts the distribution to the left or right.[17] Given this, a stable distribution is often

---

[13] See Nolan, *Stable Distributions: Models for Heavy Tailed Data* (Boston: Birkhaüser, 2010).

[14] If X_1 and X_2 are two random variables defined on a subset S of the real line, then X_1 and X_2 are *equal in distribution* if $\Pr(X_1 \leq x) = \Pr(X_2 \leq x)$ for all $x \in S$.

[15] In fact, when X_1 and X_2 are independent copies of the Cauchy distribution X, we have shown that $X_1 + X_2$ is equal in distribution to X, and so the Cauchy distribution is strictly stable.

[16] As we have seen, the probability function of the Cauchy distribution, centred at 0, is $\frac{1}{\pi(1+x^2)}$. The probability function of the normal distribution, with mean μ and standard deviation σ, is $\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$. The one remaining stable distribution capable of being expressed using elementary functions is the Lévy distribution, defined by $\frac{\sqrt{\sigma}}{\sqrt{2\pi}} \cdot \frac{e^{-\frac{\sigma}{2(x-\mu)}}}{(x-\mu)^{3/2}}$ for $x > \mu$ and 0 elsewhere. (The parameter μ specifies the location of the distribution and σ is a dispersion parameter.) The Lévy distribution is a one-tailed distribution arising in both spectroscopy and financial modeling.

[17] The characteristic function of a random variable X is the expected value of e^{-itX} (which is a function of t). Unlike the ordinary expected value, which may not exist, the expected value of e^{-itX} always exists because $|e^{-itX}| \leq 1$. Given this, it often proves useful in probability theory to work with the characteristic function of a random variable rather than its probability function. If $\varphi_X(t)$ is the characteristic function of the random variable X and is integrable, then one can obtain the probability density function of $p_X(x)$ via the following equality:

$$p_X(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-itx}\varphi_X(t)\,dt.$$

A general form of the characteristic function for stable distributions is as follows:

$$\phi(t;\alpha,\beta,\gamma,\delta) = \begin{cases} e^{it\delta-\gamma|t|^{\alpha}\left(1-i\beta\frac{t}{|t|}\tan\left(\frac{\pi\alpha}{2}\right)\right)}, & \text{if } \alpha \neq 1 \\ e^{it\delta-\gamma|t|\left(1+\frac{2i\beta\frac{t}{|t|}\log(|t|)}{\pi}\right)}, & \text{if } \alpha = 1 \end{cases}$$

where $\alpha \in (0,2]$, $\beta \in [-1,1]$, $\gamma > 0$, and $\delta \in \mathbb{R}$. There are actually several different ways of parameterising the characteristic function of stable distributions. The form listed above is one frequently used in economics. For a discussion of alternate parameterizations, see Robert H. Rimmer and John P. Nolan, "Stable Distributions in *Mathematica*," *The Mathematica Journal*, vol. 9, no. 4 (2005).

denoted by $\mathcal{S}(\alpha, \beta, \gamma, \delta)$.

The classic Witch of Agnesi, centred at 0, is the stable distribution $\mathcal{S}(1,0,1,0)$. The translated Witch, centred at 1, that underlies the gamble we are trying to value is the stable distribution $\mathcal{S}(1,0,1,1)$. Notice that both of these have the key skewness parameter β equal to zero. If we set β to a nonzero value, say $\beta = \frac{1}{2}$, we obtain the *skewed* Cauchy distribution illustrated in figure 4.

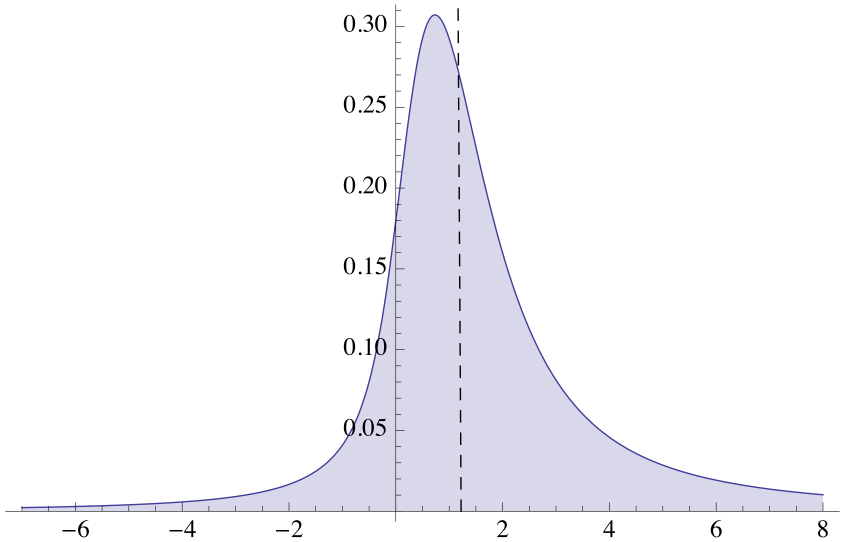


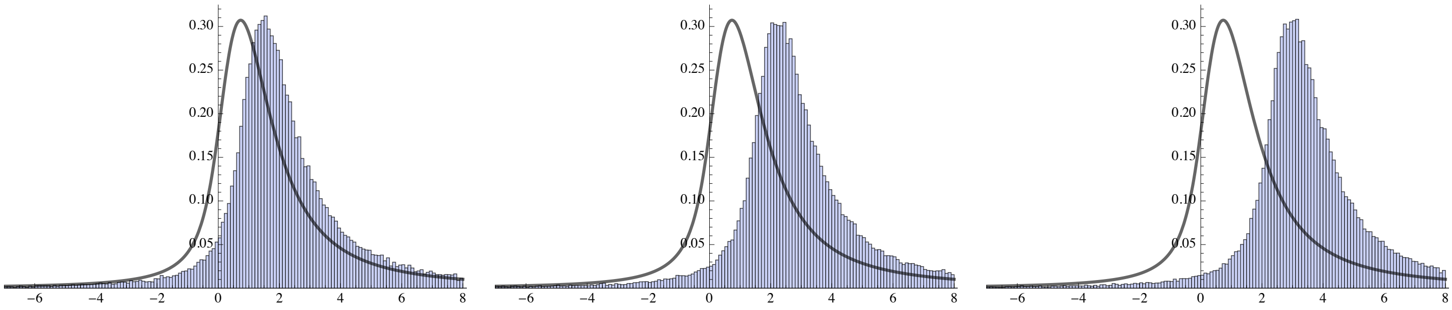
Figure 4: A "skewed Cauchy" distribution, corresponding to the stable distribution $\mathcal{S}\left(1, \frac{1}{2}, 1,1\right)$. The vertical dashed line denotes the location of the distribution's median.

Now consider the following gamble: a real number r is selected at random according to the probability distribution shown in figure 4. If $r > 0$, you win that amount of money, and if $r < 0$, you lose that amount of money. (If you dislike talk of playing the game for money because there is not that much money in the world, think of the payoff in terms of utility, instead.) Unlike the gamble defined earlier using the Witch of Agnesi, one cannot really tell a similar story motivating this gamble because the probability function of the $\mathcal{S}\left(1, \frac{1}{2}, 1,1\right)$ distribution cannot be expressed in a closed-form.

First question: should you find the gamble attractive? Although the probability distribution is no longer symmetric, obviously you should find the gamble attractive. The right tail is significantly heavier than the left tail, which means that the probability of winning an amount in the interval $(n - \varepsilon, n + \varepsilon)$ exceeds the probability of losing an amount in the interval $(-(n - \varepsilon), -(n + \varepsilon))$. This dominance reasoning concerning the amount you are likely to win indicates that the gamble, again, is in your favour.

Second question: what should you, or any other rational agent, be willing to *pay* to participate in this gamble? Herein lies the problem. Earlier we saw that the symmetric Cauchy distribution did not have an expected value. This was not an isolated phenomena, for one can show that stable distributions, when the characteristic exponent α is less than or equal to 1, in general do not have expected values.[18] Hence the skewed gamble does not have an expected value, and we find ourselves in the same position as before: attempting to attach a value to a choiceworthy gamble where the standard techniques of decision theory are silent.

[18] John P. Nolan, "Modeling financial data with stable distributions," in *Handbook of Heavy Tailed Distributions in Finance I*, S. T. Rachev, ed. (Amsterdam: Elsevier, 2003).



(a) Distribution of S_{10}.

(b) Distribution of S_{100}.

(c) Distribution of S_{1000}.

Figure 5: The distribution of sample means, for samples of three different sizes, from the skewed Cauchy distribution.

What, then, of the method of weak expectations? Recall that a weak expectation ω exists if and only if $S_n = \frac{X_1+\cdots+X_n}{n}$ converges in probability to ω as $n \to \infty$. Figure 4 illustrates the distribution of sample means for S_{10}, S_{100}, and S_{1000} when the sample is drawn from the skewed Cauchy distribution. The way to think of these distributions is as follows: Suppose that 100,000 people each play the gamble 10 times (for figure 5(a)), or 100 times (for figure 5(b)) or 1,000 times (for figure 5(c)). Once each person has played the gamble the appropriate number of times, each person calculates their average payoff, and then we plot a histogram of the outcomes, normalising the height of the bars so that it becomes a probability distribution. This illustrates the probability distribution of winning (or loosing) $\$X$ when the game is played 10 or 100 or 1000 times.

Perhaps the most striking thing we find in the distribution of sample means is that, although the *shape* of the distribution remains that of the skewed Cauchy distribution (as it must, because it is a stable distribution), the distribution moves further to the right as the sample size increases. This contrasts with the case of our gamble defined in terms of the symmetric Cauchy distribution, where the distribution of sample means was always identical to the original distribution, regardless of the sample size.

What does this mean? Well, for starters, it means that the argument we gave before for the nonexistence of a weak expectation holds in a modified form. Suppose that ω is proposed as a candidate for being a weak expectation. Because the distribution of S_n moves increasingly far to the right as $n \to \infty$, we see that S_n cannot converge in probability to ω. Why? The probability that $S_n > k$, for any $k > \omega$, actually *converges to 1* as $n \to \infty$!

Furthermore, this means that using S_n as a guide for the price you should pay for a single play of the gamble actually serves to *mislead*. The larger n is, the further to the right the distribution for S_n is translated. If S_n were taken as a guide for the price of an individual play of the gamble, it would lead you to set a price which was too high.

What of the suggestion that the *median* of the underlying distribution be used for the price of the gamble? Although this had some intuitive support in the symmetric case, that support disappears in the asymmetric case. If the distribution $\mathcal{S}\left(1, \frac{1}{2}, 1{,}1\right)$ had a mean (which it does not), it would not equal the median. The median, indicated by the dashed vertical line in figure 3, becomes just another attribute of the distribution and completely unmotivated as the price of the gamble.

Finally, the suggestion that we look at the causal process which generates the distribution cuts no ice here. That information, as so often in life, is unavailable in this case. We are asked to value the gamble presented without knowing how it was generated. And although we have complete information of the relevant probabilities and outcomes, we are at a loss as to what to say.

5. Conclusion

Decision theory provides a straightforward rule for making choices under risk: maximise your expected utility. When we apply this rule to the pricing of gambles, it says that a risk-neutral rational agent would accept the expected value of the gamble as its price. Although this rule normally applies in cases where one has full information of the possible outcomes and attached probabilities, it may still fail to give a recommendation to a rational agent for because no expected value exists.

Sometimes one can find alternative ways of assigning values. The Pasadena game was one such gamble to which decision theory was incapable of assigning a value. However, as Easwaran showed, one could assign a reasonable value to the Pasadena game by considering the convergence behavior of repeated plays. What we have seen is that even when decision theory is augmented by the method of weak expectations, which extends the capability of the theory to assign values to gambles[19], it remains incomplete. The Witch of Agnesi shows that well-defined gambles exist which are incapable of being valued by all currently known methods.

The incompleteness of decision theory in this respect matters because these distributions — the symmetric and skewed Cauchy distribution — are not mere mathematical curiosities. Stable distributions of various kinds are used in finance to model a variety of phenomena, ranging from stock prices to wheat markets. The heavy tails of these distributions (which cause the expected value not to exist in the cases we have seen) makes them suitable for modeling phenomena where outliers still occur surprisingly often. But attaching values to gambles featuring these distributions can be difficult, if not impossible, using current methods. Yet often for reasons of practical importance values must be attached. Of course one *can* assign values to such gambles, either by picking a number out of the air (the median) or by ad hoc methods, such as bounding the utility function. These methods succeed in assigning *a* value, but what we really want to know is what counts as a *rational* value. And why.

[19] Weak expectations augment decision theory because, as Fine showed, it is consistent with the axioms of decision theory to assign any value whatsoever to the Pasadena game.