An Explanation of The St. Petersburg Paradox and How to Interpret the Result

Yansı Evre Yüksel

Boğaziçi University

Department of Mathematics

The St. Petersburg Paradox was put forward by Nicolas Bernoulli during the 18th century. The problem suggested a particular game that leads to an infinite expected value while the game only appeared to be worth a small amount to the participants. A number of solutions were proposed for this problem however most of these solutions have accepted the result as false. In this paper, I will be discussing why the result diverges to infinity, why it could be more valid than we think and our mistaken view of expected values.

You are being offered a game of chance. A coin will be flipped until it comes up heads and the prize you win will depend on the number of flips the game has lasted. If the game lasts *n* turns then you acquire 2^n dollars. So if you get heads on the first flip, you win two dollars; if you get heads on the third flip, you earn eight dollars and so on. How much would you be willing to pay for this game? We generally make use of expected values from probability theory to figure out how much a bet is worth. When we apply expected value to this game we acquire a very interesting result. The probability of getting heads on the first flip is ½ and the prize is 2; for the second flip its ¼ to 4 and so on. The expected value of this game is

$$\left(\tfrac{1}{2} \times 2\right) + \left(\tfrac{1}{4} \times 4\right) + \left(\tfrac{1}{8} \times 8\right) + \ldots = 1 + 1 + 1 + 1 + \ldots = \infty$$

A rational gambler would prefer to enter a game for a price less than the expected value but the infinite result indicates that a rational gambler should be willing to enter this game for any price while most of us wouldn't be willing to pay more than $25.

Let us consider the case where the same game is proposed but with a reward other than $\$2^n$. Say the reward is constant, for example $3. Obviously, no matter how long the game lasts you will win $3. If we consider a reward of $*n* then clearly, the expected value converges to $2. These results seem reasonable. What if the reward was $\$(1.3)^n$?

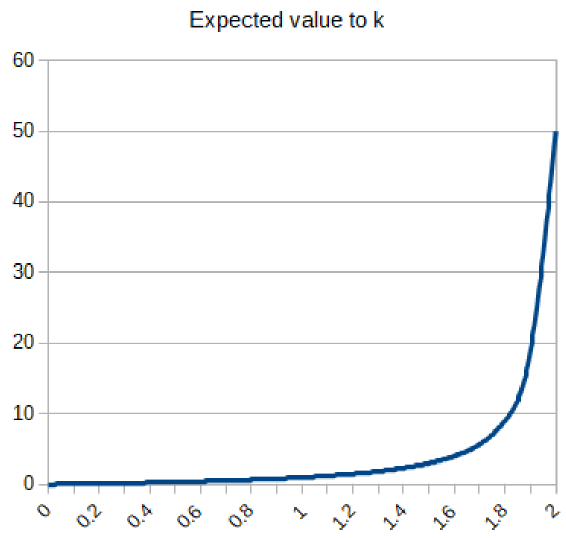
$$\mathbf{E.V} = \sum_{n=1} \left(\tfrac{1}{2}\right)^n \times (1.3)^n = 1.86 \,.$$

It's certainly interesting that this gives such a small value when 2^n diverges to infinity.

Here are some results for different values:

- **E.V($1.5^n) = $\sum_{n=1}$ (½)^n×(1.5)^n = 3** [The rewards in order are: 1.5; 2.25; 3.375; 5.06; ...]
- **E.V($1.8^n) = $\sum_{n=1}$ (½)^n×(1.8)^n = 9**
- **E.V($1.93^n) = $\sum_{n=1}$ (½)^n×(1.93)^n = 27.6**
- **E.V($1.98^n) = $\sum_{n=1}$ (½)^n×(1.97)^n = 99** [The rewards in order are: 1.98; 3.9; 7.7; 15.3; …]

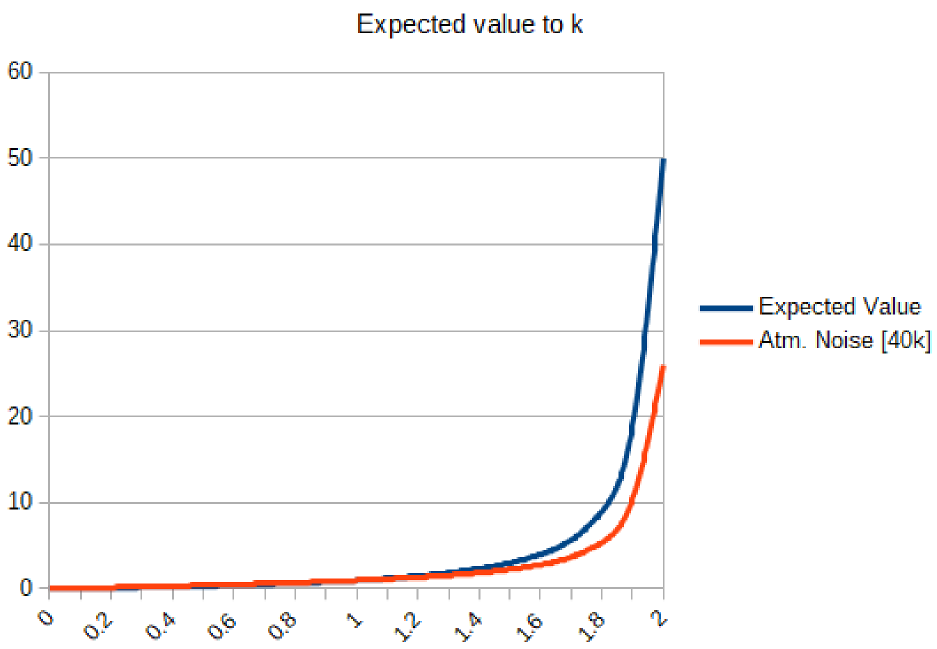
These results are very intriguing. It seems that up to somewhere around 1.9, expected value gives a reasonable result but after that, it becomes much harder to believe. Especially when you consider the rewards for the $1.5^n and the $1.98^n games. The probability of these game ending in any of the first 4 turns is 93% and the maximum difference between the rewards is $10 and yet the expected value tells us that 99 dollars is a fair price to pay for the $1.98^n game. This tells us something important about the nature of expected value calculation. Even though we have accepted that it gives us the fair value and we label the $2^n game a paradox, it certainly isn't fair for the $1.98^n game which is supposedly a proper result, as the result is similarly absurd.



The graph above shows us the results of E.V(k) for the games with $k^n rewards. The expected value rises exponentially as k approaches 2. To see how accurate any of these values are we need to compare it to experimental values. Since it is quite difficult to gather enough data in real life, a random number generator would be a more efficient method. For this, I have decided to use a random integer sequence generator that is based on atmospheric noise. The atmospheric noise makes the generation quite unpredictable, making it a good alternative to using pseudorandom number generators.

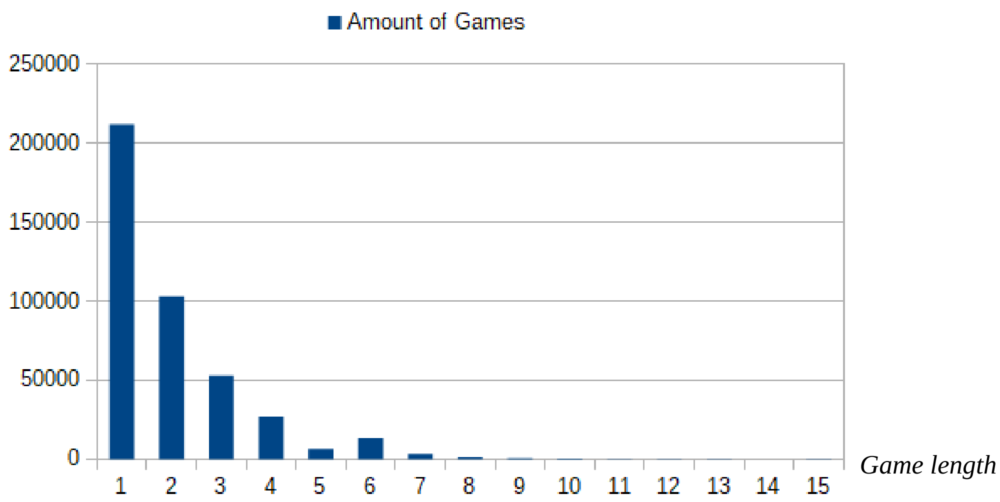
(I was unable to graph the line as going up to infinity. I appointed the value of 50 to 2 to ensure that the rest of the graph is readable.)

Using a sequence of 40 000 numbers made up of 1 and 2, I have appointed one to heads and another to tails. In order to avoid any biases of the generator as much as possible, I calculated the results the other way as well and averaged the result.



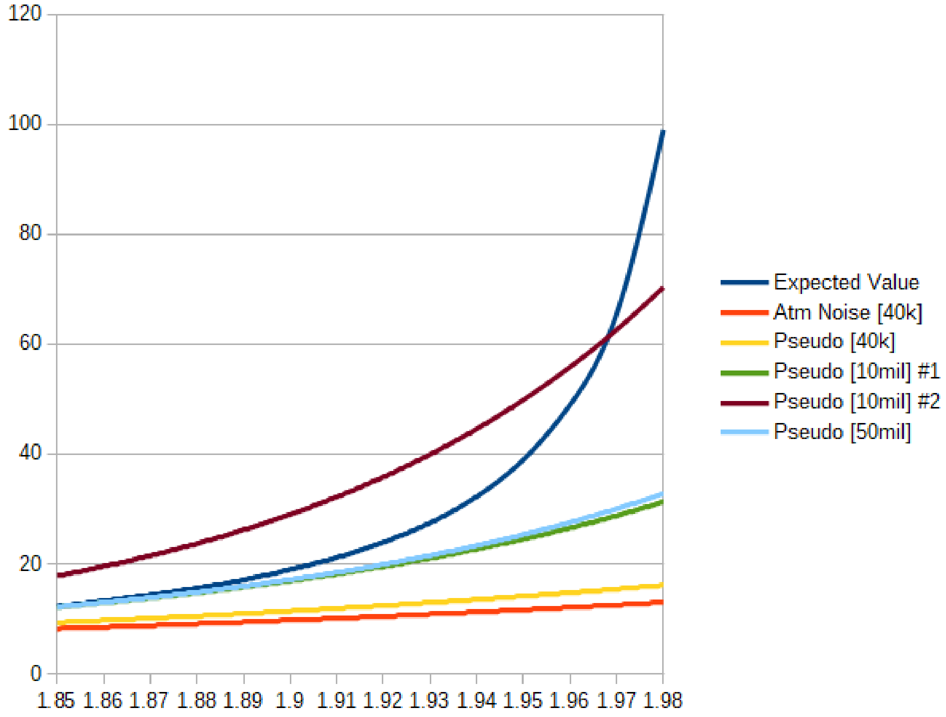
We can see that there is a clear distinction between the two graphs as k approaches 2. Atmospheric Noise games certainly don't get close to infinity at 2 but it is rather low, similar to what our intuition tells us. $25 is a reasonable price for the original game. The expected value doesn't seem to be unrelated to the problem considering the shape of both lines. While the values are off, there is indeed a connection between these two functions.

The set of generated flips includes around 20 000 games of coin tossing. In all of these games, the one that lasted longest was 15 flips. The reward being 2^{15} = $32 768. The highest we could have possibly won in this series of 20 000 games is infinitesimally low compared to what the expected value says is a fair price (any amount). If a fair price for this game was $40 000, we wouldn't have been able to make a profit in any of these 20 000 games. But what about the games where the expected value does give an answer? For the $$1.98^n$ game, the fair price was $99. Since we're using the same set of games, our longest game is 15 flips. 1.98^{15} = $28 182. The price does seem fair this time, however let's look at the entire set.



An overwhelming amount of games have lasted less than 7 turns. The maximum reward possible for these is 1.98^6 = $60.25. It seems we're back where we started. The simulation shows that the average reward for the $$1.98^n$ game is $18. Meaning that you would only make a profit in the case of 5 flips. This seems much more reasonable than $99.

We can all agree that we need to generate even more games to produce more accurate results. One would expect 40 000 games to be enough to reach a correct average. However, in our experiment even just a few occurrences of long games is enough to have a big impact on the average. I was unable to produce a larger amount of games using atmospheric noise and had to opt for a pseudorandom algorithm.



In the sets of games with 40 000 flips, the pseudorandom generator and atmospheric noise generator are fairly close so we can rule out the biases of the generators for the rest. There are two game sets of 10 million flips and very interestingly, there's a massive difference between the two of them. The second set is even higher than the expected value before 1.97. Consider the set Pseudo [10mil] #1. Using a bigger set of data has definitely changed the result from Pseudo [40k]. Turns out 40k certainly wasn't enough to produce a good enough result.

We can clearly see that the more flips we generate, the closer we get to the expected value but even when the number of flips is as high as 10 million, we can have the averages deviate such a marginal amount. So far, all of this evidence points to the fact that if infinite games were played then the average would indeed get closer to the expected value. In this case, *what exactly is the problem?*

Imagine that you are going to play the $\$2^n$ game 5 million times and each time you pay $25. In 4 999 999 of these games you only go as far as 1 flip and in one game you get 27 flips. The reward you get from that game is so large that you profit by $19 217 726. What if you were to play this game *forever*? When we're considering an infinite amount of games, the possibility of getting any amount of flips at least once is infinitely large. No matter how much you pay, as long as you continue playing the game you are *virtually guaranteed* to profit a huge amount. In such a case you can stop playing and if you decide to keep going you are virtually guaranteed again to win an even bigger profit as long as you keep going. Meaning it is a priceless bet that is always in your favor.

The expected value calculations have been giving us the correct result all this time and yet we have fixated on the fact that the results are incorrect by our intuitions. They *are* incorrect by our intuitions! It's unfair to expect any human to play the game forever. If someone can't play forever, why would he pay an enormous amount of money?

Banker	Bankroll	Expected value of lottery
Friendly Game	$100	$7.56
Millionaire	$1,000,000	$20.91
Billionaire	$1,000,000,000	$30.86
Bill Gates (2015)	$79,200,000,000	$37.15

Graph taken from Wikipedia

One of the solutions offered for the paradox was setting a limit on the reward, making it impossible for the game to last forever. The graph above shows us the expected values for multiple examples. Clearly, we produce more believable results.

We can compare these results to our generated games.

- **Average reward of $\$2^n$ with \$100 limit for Pseudo [10mil] #1 = \$7.05**
- **Average reward of $\$2^n$ with \$1,000,000 limit for Pseudo [10mil] #1 = \$19.92**
- **Average reward of $\$2^n$ with \$1,000,000,000 limit for Pseudo [10mil] #1 = \$29.79**

We have been using expected values incorrectly in two ways. We were trying to apply it to a situation that we simply can't comprehend, leading us into thinking the results must be wrong and we were trying to figure out how much a single game is worth paying for when the results don't make sense for one time games. Someone that could play indefinitely would not accept to play this game for too big an amount if he could only play once but he would be very likely to profit if he could play forever. This tells us a lot about how to utilize expected value calculations.

Consider a case of dice. The expected value of a dice roll is 3.5. Why would anyone in their right minds expect a dice to roll 3.5? However, it is reasonable to expect the average to come up 3.5 after a large number of rolls. We have shown that the expected value produced for this game have reasons to be more valid than it appears, because different uses of the expected value shows that it cannot be applied to single games. In the context in which it can be applied, the infinite result does not seem as absurd. If we had run simulations with bigger data sets we might have produced results even closer to the expected value. As you can see, the main reason we call St. Petersburg Paradox a paradox is because it makes no sense for us with finite point of views while in actuality, someone who isn't subject to our finite cognition might have found it very believable that the game is too good to be offered.

Sources

The St. Petersburg Paradox. (2008). In *Stanford Encyclopedia of Philosophy*. Retrieved from https://stanford.library.sydney.edu.au/archives/sum2008/entries/paradox-stpetersburg/

St. Petersburg Paradox (n.d). Retrieved January 6, 2019 from Wikipedia: https://en.wikipedia.org/wiki/St._Petersburg_paradox

MLA: Haahr, Mads. “True Random Integer Generator.” *RANDOM.ORG: True Random Number Service.* Randomness and Integrity Services Ltd., 5 Jan. 2019 Web. 5 Jan. 2019.

Urbaniak, G. C., & Plous, S. (2013). Research Randomizer (Version 4.0) [Computer software]. Retrieved on January 5, 2019, from http://www.randomizer.org/

Löve2D [computer software]