Hello! This document details my simulation of the St. Petersburg Paradox and the resulting insights I gained.

The Paradox:

At its core, the St. Petersburg Paradox is a question of probability and how probabilities influence our decision making.

The Paradox starts with you receiving $2 no matter what. Now if you flip a coin and get tails, those $2 double to $4. If you flip the coin and get tails again, your payout of $4 doubles to $8. The game continues in this manner until you don't get tails (otherwise known as getting heads).

So how much should you pay to play this game?

Well, statistics tells us that the expected value of this game is infinite.

$$100\% * \$2 + 50\% * \$4 + 25\% * \$8 + \ldots = \infty$$

Therefore, we should pay any price to play this game. But this probably doesn't sound right to you. And it shouldn't. If you actually play this game a few times in real life, you'll realize that you shouldn't pay anything over $20 to play this game. But why? Didn't statistics say that the expected value is infinite? Is math wrong?

Let's run some simulations to find out.

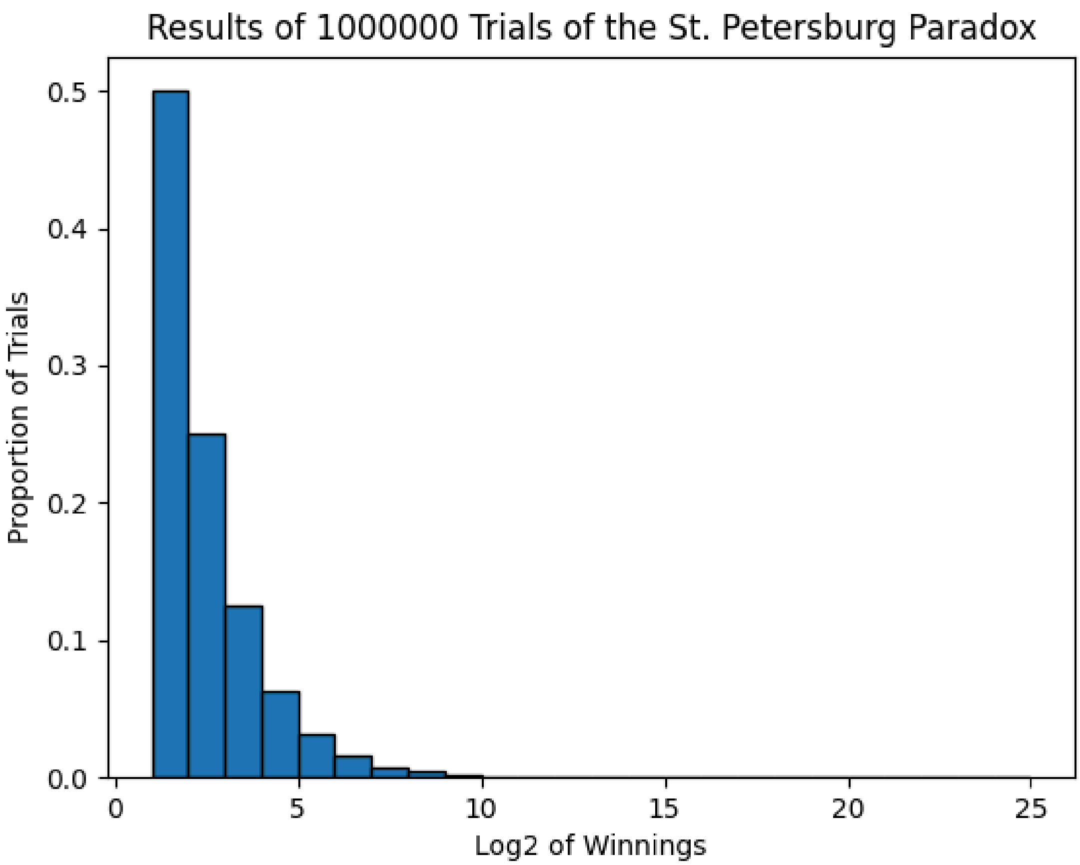
Discovery:

I was introduced to the St. Petersburg Paradox by Dr. Hogue in one of his classes in January of 2024. He actually offered to play the game using real money with those of us who were willing. Basically, Dr. Hogue encouraged student gambling. Don't tell anyone though!

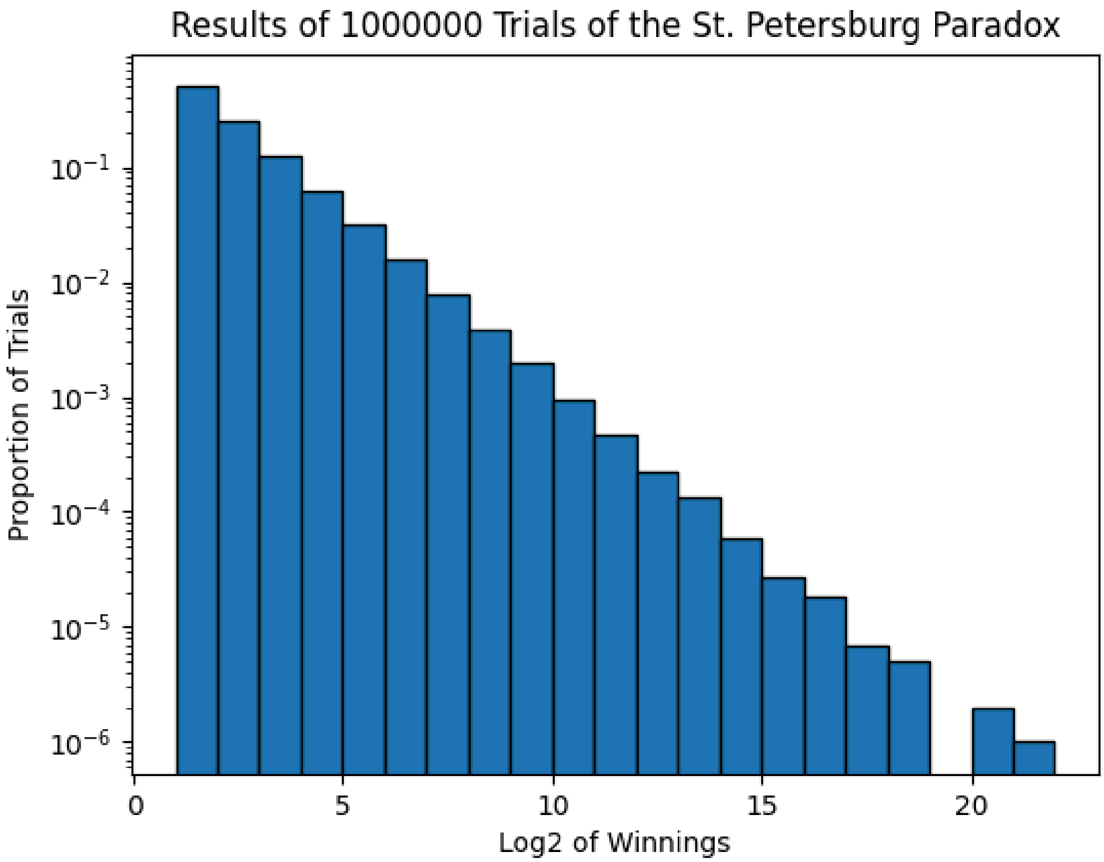
Simulation:

Coding the Paradox itself in Python was actually quite easy (see my code and its comments). The harder part was displaying the results of the trials that I ran of the Paradox in a way where I could make some sense out of them. I ended up plotting the results in a histogram with a logarithmic y-axis.

Results:



This is what a non-log graph looks like. As we can see, it gets hard to tell what happens beyond 10 (as log base two of the winnings). This is why we should graph the results using a logarithmic scale. Let's see the difference on the next page.



First of all, we can see what's going on even beyond 20 (as log base 2 of the winnings). Secondly, we can see that the proportion of trials decreases as the winnings increase.

Analysis:

Looking at the graph, we can clearly see that there are outcomes where you win more than \$2^20, or roughly \$1 million. If we ran even more trials (than 1 million), there probably would be outcomes where you win even more than \$1 billion. The thing is, the probability of these scenarios occurring is extremely, extremely, extremely small, such as 1 in a million or even less.

Therefore, this game is much like a lottery, and even more like a scratch card. Most of the time, you either lose money or win little. But statistically, there is a really, really, really small chance that you hit the jackpot and win a huge amount of money, like a million dollars. And what's more, there's an infinitesimally small chance that you win an infinite amount of money.

So if you had an infinite amount of money, you could pay any price to play this game an infinite amount of times because eventually you would win an infinite amount of money. How much that matters to someone who already has an infinite amount of money, I'm not sure. But the point is that the expected value of this game is indeed infinite; it's just that you would have to play the game an infinite amount of times to realize it.

And as a dealer without an infinite amount of money, you can't let anyone play this game too many times because, as aforementioned, someone will eventually get extremely lucky and run your wallet dry.

So how much money should a player pay to play this game? And how much money should a dealer charge a player to play this game?

Well, as a player, you obviously want to pay as little as possible. I wouldn't pay anything over $8 to play the game, since in order to make a profit, you'd have to get tails 3 times in a row (to win $16), which you have just over a 10% chance of doing. Also consider that you'll probably be playing the game many times, where you'll most likely either lose money or just break even. So even when you make a profit on a single round, you may be just getting back money that you lost in the rounds prior.

And therefore as a dealer, I’d want to charge a player more than $8 to play. More importantly, I’d make sure that they play for only a few rounds.

I think the moral of the story here is either not to gamble, or if you can, gamble a lot. Don’t do anything half-heartedly.

Reflection:

Was your answer accurate? How do you know this?

I’m very confident in the validity of my simulation because I was purely following the description of the Paradox and therefore did not make any assumptions.

If your answer was not accurate, where do you think you went off track in your simulation?

I don’t think anything about my simulation was inaccurate. My analysis of the results may be flawed, however. I will need to consult with my Statistics teacher.

UPDATE

I presented my analysis to my Statistics teacher (Ms. Neul), the other Statistics teacher (Mr. Kopeikin), and Dr. Hogue, who all agreed with what I said. So I think it's safe to say that my analysis is in the clear.

What kind of biases did you bring to this project? What did you do to reduce bias as much as possible? What biases were unavoidable?

Again, because I was merely following the description of the Paradox in my simulation, I don't think there were any biases present. However, my analysis may have been influenced by the sources that I used in my research of the Paradox. To promote original thinking and prevent plagiarism (of words and ideas), I didn't look at my sources at all while writing my analysis.

How would you expect your simulation might vary if you altered your question (different sample group, different location, etc.)?

I'm particularly curious how the results of the Paradox would change if your winnings were doubled 33% of the time instead of 50%.

100% * $2 + 33% * $4 + 11% * 8 + … = $2 + $1.33 + $0.88 + … = ?

It seems as though the expected value will converge at some number below $10.

What did you learn about simulations/the world through this process?

I learned to trust math. The math never lied in this Paradox. The expected value of the game is indeed infinite. But because it seems so unlikely to us that we start to doubt our math. Instead, we should find an explanation for why we don't think our math is correct.

What would you adjust/do differently if you had more time or the opportunity to tackle this simulation (or something similar) again?

I think I'd look into different variations of this Paradox, such as the one I suggested previously about your winnings doubling 33% of the time. I'd also like to think about the real-world applications of this Paradox. I suppose there aren't many because this Paradox deals with infinites when there Earth's resources are not infinite.