2013

THREE HUNDRED YEARS OF THE ST. PETERSBURG PARADOX

Keguo Huang
Michigan Technological University

Follow this and additional works at: https://digitalcommons.mtu.edu/etds
Part of the Mathematics Commons

THREE HUNDRED YEARS OF THE ST. PETERSBURG PARADOX

By

Keguo Huang

A REPORT

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In Mathematical Sciences

MICHIGAN TECHNOLOGICAL UNIVERSITY

2013

This report has been approved in partial fulfillment of the requirements for the Degree of MASTER OF SCIENCE in Mathematical Sciences.

Department of Mathematical Sciences

Report Advisor: *Iosif Pinelis*

Committee Member: *Jianping Dong*

Committee Member: *Laura E. Brown*

Department Chair: *Mark S. Gockenbach*

Contents

List of Figures

List of Tables

Acknowledgments

First of all, I would like to give my sincere gratitude to my advisor Prof. Iosif Pinelis for the continuous support of my research, for his motivation and enthusiasm. He is the best teacher I've ever had: his guidance helped me in all parts of my research.

Also, I would like to thank my committee members: Prof. Jianping Dong and Prof. Laura Brown, for their encouragement and insightful comments.

Last but not the least, my sincere thanks go to my wife Yi Jiang and my parents Shoukuan Huang and Caimei Chen, for supporting me throughout my life.

Abstract

The St. Petersburg Paradox was first presented by Nicholas Bernoulli in 1713. It is related to a gambling game whose mathematical expected payoff is infinite, but no reasonable person would pay more than $25 to play it. In the history, a number of ideas in different areas have been developed to solve this paradox, and this report will mainly focus on mathematical perspective of this paradox. Different ideas and papers will be reviewed, including both classical ones of 18th and 19th century and some latest developments. Each model will be evaluated by simulation using Mathematica.

Chapter 1

Introduction

The St. Petersburg Paradox was first presented by Nicolas Bernoulli, a prominent Swiss mathematician from the well-known Bernoulli family, first appearing in a letter to another distinguished French mathematician P. R. de Montmort on Sep. 9th, 1713. The first academic article about this paradox was published in 1738 in *Commentaries of the Imperial Academy of Science of Saint Petersburg*, by N. Bernoulli's cousin Daniel Bernoulli. The name of the paradox was coined by D'Alembert in 1768. The paradox relates to a designed game, which will be called "St. Petersburg game" in this text, described by D. Bernoulli [1] in his paper as follows:

Peter tosses a coin and continues to do so until it should land "heads" when it comes to the ground. He agrees to give Paul one ducat if he gets "heads" on the very first throw, two *ducats if he gets it on the second, four if on the third, eight if on the fourth, and so on, so that with each additional throw the number of ducats he must pay is doubled. Suppose we seek to determine the value of Paul's expectation.*

What is a reasonable stake Paul should place to play this game? On the one hand, a reasonable person would not want to spend much on this game, since as the desired payoff raised the corresponding probability decreases very fast. For example, the probability to win a payoff of 16 dollars is only $1/32 \approx 3.125\%$. On the other hand, by the classical probability theory, the "fair price" to play a game shall be the mathematical expectation of payoff of a game, and in this game it turns out to be

$$2^0 \cdot \frac{1}{2^1} + 2^1 \cdot \frac{1}{2^2} + 2^2 \cdot \frac{1}{2^3} + \ldots = \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \ldots = \infty. \tag{1.1}$$

D. Bernoulli [1] states the discrepancy as follows:

. . . Although the standard calculation shows that the value of Paul's expectation is infinitely great, it has . . . to be admitted that any fairly reasonable man would sell his chance, with great pleasure, for twenty ducats.

In the history, a number of ideas in different areas have been developed to solve this paradox. For example, D. Bernoulli's paper is considered to be the root of modern marginal utility theory [10], which stimulating many papers in the area of economics. But our report will mainly focus on mathematical perspective of this paradox. Let us denote the payoff of one game by a random variable X. Then the probability distribution of X is given by

$$P(X = 2^i) = \frac{1}{2^{i+1}}, \text{ for } \text{ for } i = 0, 1, 2, 3, \ldots. \tag{1.2}$$

Also, we define the total and average payoff of n independent games as follows:

$$S_n := \sum_{i=1}^{n} X_i, \text{ and } \overline{X}_n := \frac{S_n}{n}. \tag{1.3}$$

where $X_1, X_2, \ldots$ are independent copies of X.

Chapter 2

Before 1945

2.1 Nicolas Bernoulli and Gabriel Cramer

This paradox exists primarily because the mathematical expectation is infinite. In other words, the series computing the expectation, $\sum_{i=0}^{\infty} 1/2^{i+1} \cdot 2^i$, diverges. We can generalize the setting of this game, with $p(i)$ representing the probability of the outcome that the “head" first lands at i-th toss and $f(i)$ representing the payoff when that occurs. Now the mathematical expectation can be written as

$$\sum_{i=0}^{\infty} p(i) \cdot f(i). \tag{2.1}$$

Then it is straightforward that two methods can be generated to make this series convergent, to modify either $p(i)$ or $f(i)$. In the history, N. Bernoulli himself did the former and another Swiss mathematician Gabriel Cramer did the latter both for the first time. During 18th and 19th century, most of works related to our topic were written in non-English languages. We found this part of history of this chapter mainly from a historical article of Gerard Jorland in 1987 [9] and another article of Jacques Dutka in 1988 [5].

In another letter to de Montmort on Feb. 20th in 1714, N. Bernoulli proposed a resolution that the event of very small probability should not be taken into consideration, although the payoff of these event are large. Applying his solution to our model, $p(i)$ is enforced to be zero for large i and consequently the series becomes finite. Jacob Bernoulli, N. Bernoulli's uncle, had stated a similar concept called *moral certitude* in his famous book *Ars Conjectandi*. G. L. L. Buffon had a good illustration in 1846, cited from Jorland's paper [9], *"As for moral certitude, he (Buffon) had found in a bill of mortality that the odds of death overnight for a 56 year-old man were* 1 *to* $10,189$*, from which he inferred that the mean man of that age does not fear death merely because of his moral certitude that a probability lower than* $1/10,000$ *is as nought."* Then Buffon suggested evaluating probabilities less than $1/2^{13} = 1/8192$ as 0, leading the expectation or the stake to 6.5, which seems reasonable.

However, this solution has two drawbacks. One is that the choice of threshold is subjective thus can be very arbitrary. Also, as we will see from the result from simulation study later, although the probability is small, high payoff still have a big influence on the outcomes especially when we increase the number of games. Nowadays, high-speed computers allow us to simulate games within seconds. Figure 2.1 is a result of simulated St. Petersburg games using Mathematica. For each point on the graph, the x-coordinate represents the number of games Paul chooses to play, and y-coordinate is the average payoff $\overline{X}_n$, and we simulate the situation from one game to ten thousands games. As it showed, mostly the average payoff lands between 5 and 10. But there are still many times, Paul obtains considerable average payoffs. It happens mainly because Paul is lucky enough to obtain one (or two) very high payoff(s) on some game, even though the probability of occurrence is extremely small. For example, all points in circle A on the graph are under the situation that Paul obtains one payoff of $2^{19} = 524,288$ coins in some game, which totally dominates and leads the average payoff to a high lever. Also all points in circle B indicates occurrence of one payoff of $2^{18} = 262,144$ coins in some game. Apparently, these situations can not be ignored, in fact they act as a strong motivation for gamblers to play this game.

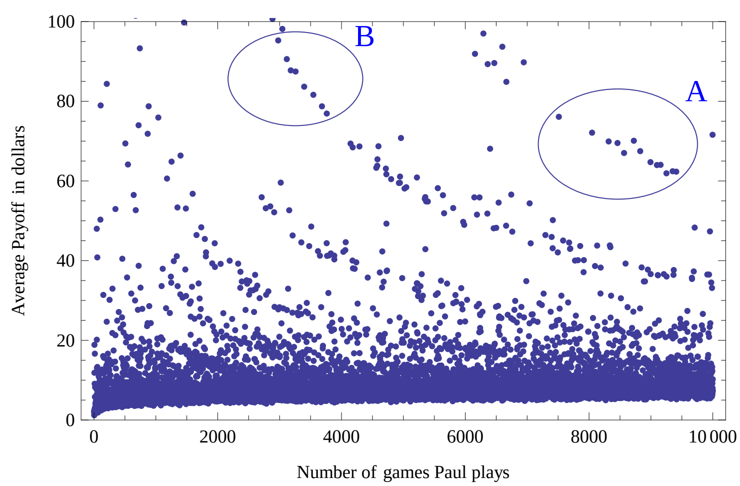

Figure 2.1: Average payoffs for different number of games

Later, N. Bernoulli told the paradox to G. Cramer, and the latter states his solution in a letter to the former in May 21st 1728. To solve the paradox, Cramer applied a concept, which would be recognized as utility: *"Mathematicians value money in proportion its quantity, whereas reasonable people value it in proportion to its use."* He continued to argue that any payoff beyond certain large quantity, which he proposed 2^{24}, will just bring the same pleasure as 2^{24} does. Applying this solution to our model (2.1), it is enforced that $f(i) = \min\{2^i, 2^{24}\}$, which becomes bounded and leads to a finite expectation and thus a finite stake

$$\sum_{i=0}^{24} \frac{1}{2^{i+1}} \cdot 2^i + \sum_{i=25}^{\infty} \frac{1}{2^{i+1}} \cdot 2^{24} = 12 + 1 = 13. \tag{2.2}$$

Similarly, we can evaluate this solution by simulation, and we will make the payoff be 2^{24} if it exceeds 2^{24}. First, we make Figure 2.2 in the similar way as we did in Figure 2.1. As we can see, mostly the points has a payoff below 10, thus below the stake 13. In fact, 88.09% of points are below 13. This result undermines the Cramer's solution.

Also, to be more precisely, we can consider average payoff of a group of 100 games. After playing $10,000$ groups three times, proportion of groups whose average payoffs exceeds the stake 13 are correspondingly $0.0956, 0.095, 0.1008$. This result suggests if Paul plays 100 consecutive games, he has about 10% of chance getting a average payoff higher than the stake 13.

This result can be implied by Chebyshev's Inequality. To model Cramer's solution, we can

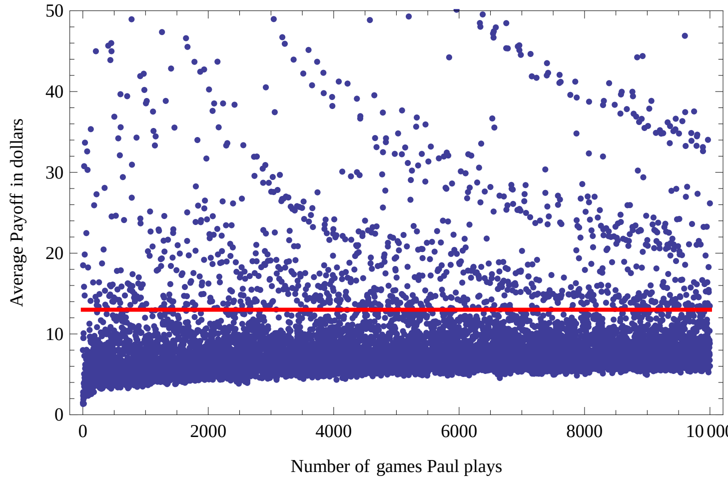

Figure 2.2: Average payoffs for different number of games under Cramer's bound

define $Y := \min\{X, 2^{24}\}$ as the payoff of "Cramer's criterion". Then it can be obtained that Y has a large variance: $\mathbf{Var}Y \approx 3 \times 2^{23} \approx 2.5 \times 10^7$, then by Chebyshev's Inequality,

$$\mathbf{P}(|\overline{Y}_n - EY| \geq \varepsilon) \leq \frac{\mathbf{Var}\,Y}{n\varepsilon^2}. \tag{2.3}$$

In our case $n = 100$, then to make $\frac{\mathbf{Var}Y}{n\varepsilon^2} < 1$, we still need $\varepsilon > \sqrt{2.5 \times 10^5}$. So $\overline{Y}_n$ can appear very far away from EY, suggesting our result earlier is reasonable.

Cramer also improved his utility model by taking square root of quantity as the utility, which is already very close to Danial Bernoulli's later resolution where a logarithmic function is used. This model makes more sense, since it models the fact that we always obtain more pleasure if we have more payoff, no matter how large that is. Two billion will surely bring more happiness than one billion, but just not as twice as much. Under this new assumption, the mathematical expectation of utility, or equivalently utility mean, becomes

$$\sum_{i=1}^{\infty} \frac{\sqrt{2^{i-1}}}{2^i} = \frac{1}{2} \sum_{i=1}^{\infty} \frac{1}{\sqrt{2^{i-1}}} = \frac{1}{2-\sqrt{2}}. \tag{2.4}$$

Then the stake Paul places would also has the same utility, which would be the squared utility mean $(1/(2-\sqrt{2}))^2 \approx 2.914$, which is more reasonable than the previous stake 13.

2.2 Danial Bernoulli

Although D. Bernoulli's paper [1] was published in 1738 as stated in the introduction, he finished and submitted his paper in 1731, about three years after Cramer's resolution. In fact, he did not know Cramer's idea until his publication. In this paper, he discussed in detail about the concept of utility: the value of wealth should not depend on the amounts but the utility. For example, it is usually true that one thousand dollars have less value to a rich person than to a beggar, although the amount is the same. Thus, the utility (y) can be considered as a function of wealth (x). He made an assumption that, rephrased using today's college calculus terminology, the rate of change of utility with respect to wealth is inversely proportional to the initial wealth, a.e.,

$$\frac{dy}{dx} = \frac{k}{x}, \tag{2.5}$$

where k is a positive constant. After integration, we have the solution

$$y = k\ln x - k\ln a, \tag{2.6}$$

where a is the initial wealth, which yields zero utility, i.e. $y(a) = 0$. If x is the wealth after a St. Petersburg game, then $x - a$ is the payoff in wealth, and the expected utility gained from the payoff can be calculated as a weighted mean as follows

$$\sum_{i=0}^{\infty} \frac{1}{2^{i+1}} \cdot (k\ln(a+2^i) - k\ln a) = k\sum_{i=0}^{\infty} \frac{1}{2^{i+1}} \ln(a+2^i) - k\ln a \tag{2.7}$$

Then if we denote the suggested stake as s, then its utility would be the same as the expected gain utility, more precisely,

$$k\ln(a+s) - k\ln a = k\sum_{i=0}^{\infty} \frac{1}{2^{i+1}} \ln(a+2^i) - k\ln a \tag{2.8}$$

which yields

$$s = \prod_{i=0}^{\infty} (a+2^i)^{\frac{1}{2^{i+1}}} - a. \tag{2.9}$$

Essentially, D. Bernoulli's resolution contends that Paul's stake should depends on his initial wealth. He made a comment *"This result sheds light on a statement which is universally accepted in practice: it may be reasonable for some individuals to invest in a doubtful enterprise and yet be unreasonable for others to do so."* Setting different values of initial wealth a, we can compute the corresponding stakes. See Table 2.1

Table 2.1
Estimated stake price for different initial wealth

Initial Wealth a	0	10	100	1,000	10^4	10^5	10^6
Suggested Stake s	2.0	3.0	4.4	6.0	7.6	9.3	10.9

In spite of the absurd first column where the suggested stake for a zero initial wealth is 2.0, the other six columns seem reasonable. Now we can simulate games to test if D. Bernoulli's resolution is reasonable.

First consider the case with an initial wealth 10^6 dollars and a stake 10.9 coins. Suppose Paul has an initial wealth of 10^6, and let us assume Paul chooses to play $10,000$ games. Each time he pays a stake of 10.9 and receives the payoff. We will design a algorithm to simulate this process. In order to achieve a better estimation, let us consider 200 Pauls with the same initial wealth playing these games. Of course, some of them will gain wealth whereas others will lose coins. We collect the results of these 200 Pauls, and display them in a histogram. See Figure 2.3.

We see about $3/4$ of Pauls (146 out of 200) lose some amount of money, but for those Pauls who earned, quite a few of them have earn a big amount of money, especially those three that cannot be displayed in the graph, about one million, two million, four million. In fact, the mean value of coins earned/lost is 31074.4. In total, this is a reasonable game to play for Paul if he has one million coins.

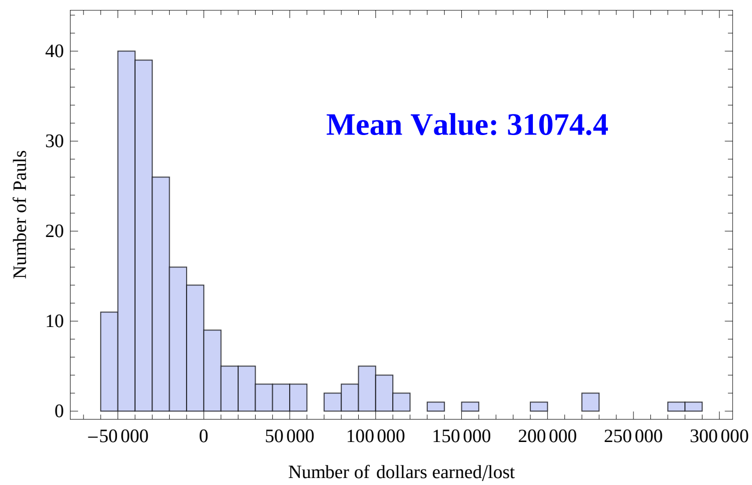

Figure 2.3: Histogram of gains/losses for 200 Pauls with initial wealth 10^6 playing $10,000$ games. [a]

[a] There are three gain $1.0049 \times 10^6, 2.0796 \times 10^6, 4.14824 \times 10^6$, not displaying in the graph, since they are far apart from others.

Also, consider the case with an initial wealth 10 coins and a stake 3.0 coins. Similarly, suppose $2,000$ Paul has an initial wealth of 10, and each time he pays a stake of 3.0 and receives the payoff. But it is a little different in this case, for one group of Pauls may run out of money very quickly if they gets bad luck. Also there will be another group of Pauls whose wealth may keep growing. Let us assume that Paul will at most play $1,000$ games under any case. Figure 2.4 and Figure 2.5 are histogram of the first group and the second group of Pauls.

It's been shown in Figure 2.4 that $1,383$ out of $2,000$ Pauls run out of money before $1,000$ games end, and in fact most of them end it within the first 20 games. From Figure 2.5, most of Pauls earn less than $5,000$ coins, but some Pauls get very luck to have a very big payoff with one reach one million coins started from 10 coins. In total, it also suggests that 3.0 is a reasonable price of a Paul with an initial wealth 10.

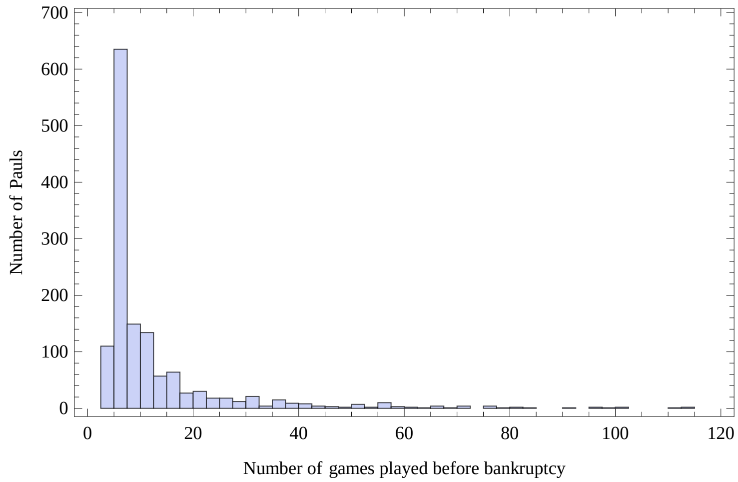

Figure 2.4: Histogram of number of games played for 1383 Pauls who go bankrupt

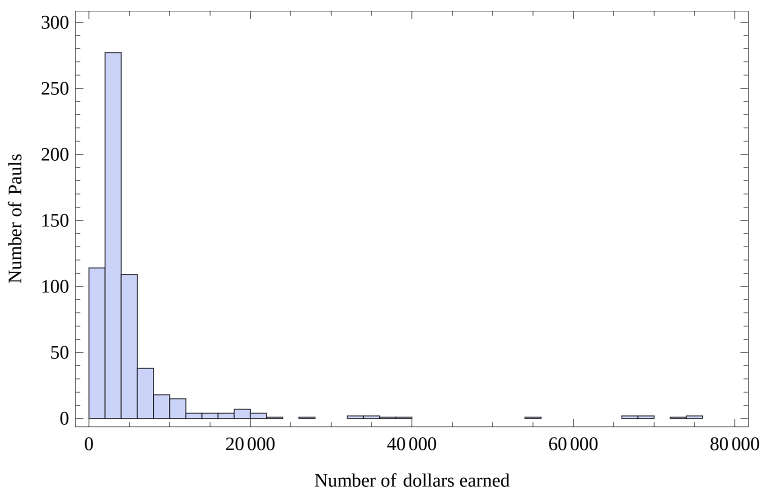

Figure 2.5: Histogram of number of coins earned for 617 Pauls[a]

[a]There are seven gain 132866, 134883, 137413, 267984, 269753, 526220, 1068322, not displaying in the graph, since they are far apart from others.

The most striking result from this simulation is that even though the suggested stake of fist case 10.9 is more than three times greater than that of second case 3.0 (even higher than the initial wealth of the second case 10), they are still reasonable according to the experimental results. This also supports D. Bernoulli's resolution.

2.3 G. L. L. Buffon

In the history, Buffon is the first one who actually did the games, he found a child play this game 2048 times to examine his result. Instead of the payoff of a St. Petersburg Paradox, Buffon considers the total number of tosses when it lands a head for the first time. In a paper of 1777, he published his resolution. In a St. Petersburg game, total number of tosses when the first head appears is denoted by T, e.g. a head appears at T-th toss for the first time. Then $\mathbf{P}(T=k) = 1/2^k$, for $k = 1,2,\cdots$. Now suppose Peter and Paul agree that they will play a total of $N = 2^s$ games, then we will have a sequence of $T_1, T_2, \ldots, T_N$ as independent copies of T. Thus, the event $\{T = k\}$ is expected to happen in

$$\mathbf{E}\sum_{j=1}^{N} I\{T_j = k\} = N \cdot \mathbf{P}(T = k) = \frac{N}{2} = 2^{s-k} \tag{2.10}$$

games. More precisely, it can be achieved that

$$
\begin{aligned}
&\{T=1\} \text{ is expected to happen in } 2^{s-1} \text{ games,}\\
&\{T=2\} \text{ is expected to happen in } 2^{s-2} \text{ games,}\\
&\qquad\qquad\vdots\\
&\{T=s\} \text{ is expected to happen in 1 games,.}
\end{aligned}
\tag{2.11}
$$

If we add all games above, there will be $2^s - 1$ games and one game will be missing. Buffon claims that the number of tosses in last incomplete game cannot be well estimated, and he believes that discarding that game will not cause any significant error. Using result above, we can naturally calculate the expected total payoff $\sum_{i=1}^{s} 2^{s-i} \cdot 2^{i-1} = N \cdot s/2$, or equivalently, expected average payoff $s/2$. Then $s/2$ will also be Paul's reasonable stake. Here, the significance of Buffon's resolution is that the average payoff of games depends on the number of games (N) Paul will play. His result is displayed in Table 2.2

In 1889, Sydney Lupton [11] generalized Buffon's result. If Paul will play N St. Petersburg game, then Paul's stake should be of the same order of $1/2 \cdot \log_2 N$.

Table 2.2
Theoretical and Experimental results of Buffon

		Distribution of games	
Rank of toss ending the game	Value of the game	Empirical	Binomial
1	1	1,061	1,024
2	2	494	512
3	4	232	256
4	8	137	128
5	16	56	64
6	32	29	32
7	64	25	16
8	128	8	16
9	256	6	8
10	512		4
11	1,1024		2
Total value of games		10,057	11,264

Chapter 3

After 1945

For the sake of convenience, most of researchers modified the original game by multiplying the payoff by 2. To be consistent with their results we will do that as well. So from now on, the distribution of payoff in (1.2) would be redefined as follows,

$$\mathbf{P}(X = 2^i) = \frac{1}{2^i}, \qquad \text{for } i = 1, 2, 3, \ldots. \tag{3.1}$$

3.1 William Feller

Feller's paper in 1945 [6] really brings the study of St. Petersburg Paradox into a level of maturity. In this paper Feller applied a generalized Weak Law of Large Number (WLLN) of his another paper in 1937 [8] into this problem. The normal WLLN can be stated as follows: suppose $X_1, X_2, \ldots$ are independent and identically distributed random variables with common expectation $EX_1 = \xi < \infty$, then

$$\forall \varepsilon > 0, \mathbf{P}(|\overline{X}_n - \xi| < \varepsilon) \to 1, \text{ as } n \to \infty, \tag{3.2}$$

where $\overline{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i$. The limit above can also be denoted as $\overline{X}_n \xrightarrow[n\to\infty]{P} \xi$, meaning $\overline{X}_n$ converges to ξ in probability. However, this WLLN cannot be used in the St. Petersburg games, since the expectation of payoff is infinite as we stated in the introduction. In the paper of 1937, he generalized a WLLN for an iid random sample whose expectation is infinite. After applying this result, he achieved a WLLN for St. Petersburg paradox: if $X_1, X_2, \ldots$ denotes payoffs of our St. Petersburg games, then

$$\mathbf{P}\left(\left|\frac{\overline{X}_n}{\log_2 n} - 1\right| < \varepsilon\right) \to 1, \text{ as } n \to \infty, \tag{3.3}$$

In the sense of the convergence in probability, Feller suggest the fair stake for playing n games should be of the order with $n \log_2 n$, to which Buffon's result is close. A detail discussion can be found in Chapter X of his book in 1968 [7].

There are some improvements of Feller's result. In 1961, Y. S. Chow and Herbert Robbins [2] proved that the corresponding Strong Law of Large Number (SLLN) of St. Petersburg Games does not hold. The SSLN asserts that: $X_1, X_2, \ldots$ are independent and identically distributed random variables with common expectation $EX_1 = \xi < \infty$,

$$\mathbf{P}(\lim_{n\to\infty} X_n = \xi) = 1. \tag{3.4}$$

The limit above can also be denoted as $\overline{X}_n \xrightarrow[n\to\infty]{a.s.} \xi$, meaning $\overline{X}_n$ converges to ξ almost surely. (3.4) is equivalent to

$$\forall \varepsilon > 0, \mathbf{P}(\sup_{m\geq n} |\overline{X}_m - \xi| < \varepsilon) \to 1, \text{ as } n \to \infty, \tag{3.5}$$

Comparing (3.2) and (3.5), it is clear that almost sure convergence implies convergence in probability, but not vice versa. Y. S. Chow and Herbert Robbins show that (3.4) does not hold for St. Petersburg games. In fact, Applying Borel-Cantelli lemma, they achieve that

$$P\left(\lim_{n\to\infty} \frac{\overline{X}_n}{\log_2 n} = 1\right) = 0. \tag{3.6}$$

So St. Petersburg Paradox provides a good example of a sequence of random variables converges in probability but not almost surely.

Another improvement was made by Anders Martin-Löf [12] in 1985, who obtained a limit distribution for Paul's average gain. Suppose Paul plays a number of $N = 2^n$ games in total with a total stakes $N \log_2 N$ suggested by Feller and the total gain is denoted by S_N, then the average gain will be

$$\frac{S_N - N\log_2 N}{N} = \frac{S_N}{N} - n. \tag{3.7}$$

Martin-Löf proves that this average gain in (3.7) has a limit distribution $G(x)$ whose characteristic function can be expressed explicitly. He actually gives a heuristic argument that the limit distribution $G(x)$ which is the same as that of

$$S := \sum_{k=-\infty}^{0} (Z_k - 2^{-k})2^k + \sum_{k=1}^{\infty} Z_k 2^k, \tag{3.8}$$

where $Z_1, Z_2, \dots$ independently and identically follow the Poisson distribution with parameter 2^{-k}. Later he proves that

$$\mathbf{P}(\frac{S_N}{N} - n > 2^m + x) \approx 2^{-m}(2 - G(x)) \tag{3.9}$$

when n is not too small and $m \geq 5$. This approximation will help Peter to determine the stake in order to achieve a low probability for Paul to gain from him. For example, setting $x = 0$, $G(x)$ can be calculate as 1.7925. Then if Peter wants a low probability $10^{-3} \approx 2^{-10}$, he takes $m = 11$, then the stake would be $n + 2048$.

3.2 Hugo Steinhaus

In 1949, Steinhaus [13] suggested his resolution in a short paper of sixty six lines. Observing that payoffs of 2 are expected to happen in $1/2$ of games, payoffs of 4 are expected to happen in $1/4$ of games, payoffs of 8 are expected to happen in $1/8$ of games, etc, he built the a sequence for stakes of the game,

$$2,4,2,8,2,4,2,16,2,4,2,8,2,4,2,32,2,4,2,8,2,4,2,16,2,4,2,\cdots \tag{3.10}$$

where 2 appears every other position, 4 appears every four positions, 8 appears every eight positions, and in general 2^k appears every k positions. Then in order to play the game, Paul needs to pay 2 in the first game, 4 in the second game, pay 2 in the third game, etc. We can call this sequence “Steinhaus Sequence", which will be referred as $\{a_n\}$ in this section. Also we define

$$s_n := \sum_{i=1}^{n} a_n, \text{ and } \overline{a}_n := \frac{s_n}{n} \tag{3.11}$$

We can compare Steinhaus’ average stake $\{\overline{a}_n\}$ with Feller’s average stake $\{\log_2 n\}$. Figure 3.1 shows these two sequences as n ranges from 1 to 1000, with blue dots indicating $\{\overline{a}_n\}$ and red dots as $\{\log_2 n\}$. It can be noticed that Steinhaus’ game is more expensive to play than Feller’s. More precisely, large differences will happen when n is chosen to be a power of 2, since big stakes will be added into the sequence $\{a_n\}$. Also, small differences will happen when n is one unit less than a power of 2. In fact, when $n = 2^k - 1, k = 1, 2, \cdots,$

$$s_n = 2 \cdot 2^{k-1} + 4 \cdot 2^{k-2} + 8 \cdot 2^{k-3} + \cdots + 2^{k-1} \cdot 2 + 2^k \cdot 1 = k2^k = kn, \tag{3.12}$$

then $\overline{a}_n = k = \log_2(n+1) \approx \log_2 n$ when n is large. So one shortcoming of Steinhaus' resolution is if Paul knows the stake sequence in advance, he can stop playing games as the number of games gets closer to a power of 2 to avoid a large stake. Table 3.1 shows the average stake price of 2^k and $2^k - 1$ St. Petersburg games for some k. The striking result of the table is the big differences of average stake prices between two groups even though they only differ from one game.

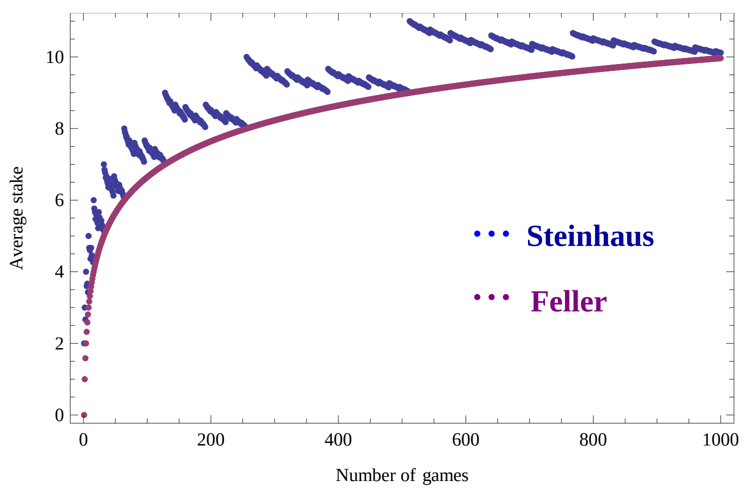

Figure 3.1: Comparison of Stake prices of Feller and Steinhaus.

Sándor Csörgö and Gordon Simons [4] presented a deeper study based on Steinhaus' work in a paper of 1993.

Table 3.1
Required average stake prices for 2^k and 2^k-1 suggested by Steinhaus

Value of k	7	8	9	10	11	12	13
Average Stake for 2^k	9	10	11	12	13	14	15
Average Stake for 2^k-1	7.06	8.03	9.02	10.01	11.01	12.00	13.00

First, they make efforts to asymptotic property of the sequence $\{s_n\}$ compared with $\{\log_2 n\}$, and it turns out that

$$\frac{\overline{a}_n}{\log_2 n} \xrightarrow[n\to\infty]{} 1$$

More precisely, they explicitly find a function, denoted as δ_n here, such that

$$\frac{\overline{a}_n}{\log_2 n} = 1 + \delta_n,$$

where $\delta_n > 0$ and $\delta_n \xrightarrow[n\to\infty]{} 0$.

Moreover, based on Steinhaus sequence, they design “Steinhaus games". They first build a “empirical distribution function" of Steinhaus sequence.

$$\tilde{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} I\{a_i \leq x\}, \quad x \in \mathbf{R}$$

Then, the Steinhaus game is defined as a game whose payoff, denotes as X_n' for each n, follows the distribution function above. Namely, the Steinhaus game will depend on the first n elements of the Steinhaus sequences. For example, for $n=7$,

$$X_7' = \begin{cases} 2, & \text{with probability } 4/7 \\ 4, & \text{with probability } 2/7 \\ 8, & \text{with probability } 1/7 \end{cases}$$

They also make a comparison between the Steinhaus and St. Petersburg game. Which one would Paul prefer? They show that the payoff that Paul achieves will be stochastically larger in St. Petersburg game, i.e.,

$$1 - F(x) = \mathbf{P}(X > x) \geq \mathbf{P}(X_n' > x) = 1 - \tilde{F}_n(x), \quad x \in \mathbf{R}$$

where X is the payoff of St. Petersburg game defined in (3.1), and $F(x)$ is the corresponding distribution function.

3.3 Sándor Csörgö and Gordon Simons

Apart from the extension of Steinhaus' resolution, Csörgö and Simons have published several papers on this paradox. In one of them [3] , they presented the so-called "Two-Paul Paradox", described as follows,

Suppose Peter agrees to play exactly one St. Petersburg game with each of two players, Paul$_1$ and Paul$_2$. Question: Are Paul$_1$ and Paul$_2$ better off (i) accepting their individual *winnings, X_1 and X_2, say, or (ii) agreeing, before they play, to divide their total winnings in half, so that each receives $(X_1 + X_2)/2$?*

Contrary to the common sense, these two strategies do make a difference, and in fact, two Pauls are better off with the strategy (ii). More precisely, they showed that

$$\mathbf{P}(X_1 + X_2 \geq x) \geq \mathbf{P}(2X_1 \geq x), \text{ for all } x > 2 \tag{3.13}$$

where X_1 and X_2 are independent copies X defined in (3.1). To improve the result, we explicitly find out the expression of two probability in (3.13). See Appendix A.1. If we denote $p_{21} := \mathbf{P}(X_1 + X_2 \geq x)$ and $p22 := \mathbf{P}(2X_1 \geq x)$. Figure 3.3 shows the relative difference between them, $(p_{21} - p_{22})/p_{22}$, for each integer x from 1 to 1000.

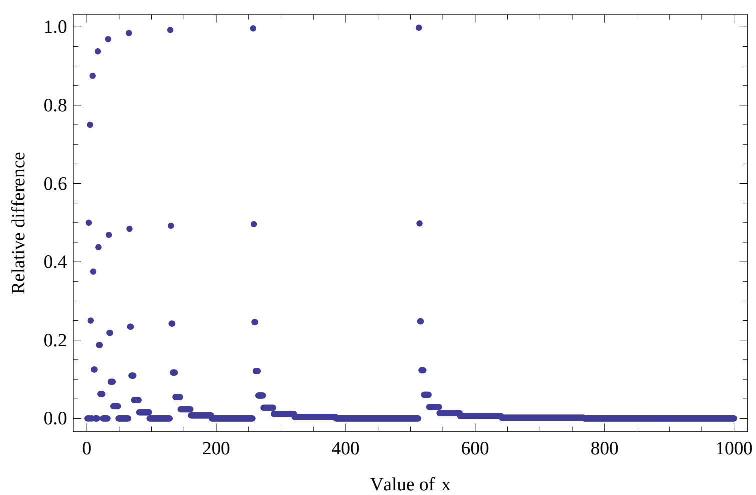

Figure 3.2: Two Paul Paradox, relative difference

After a little algebra A.2, we can obtain the relative difference for four-Paul problem in Figure 3.3.

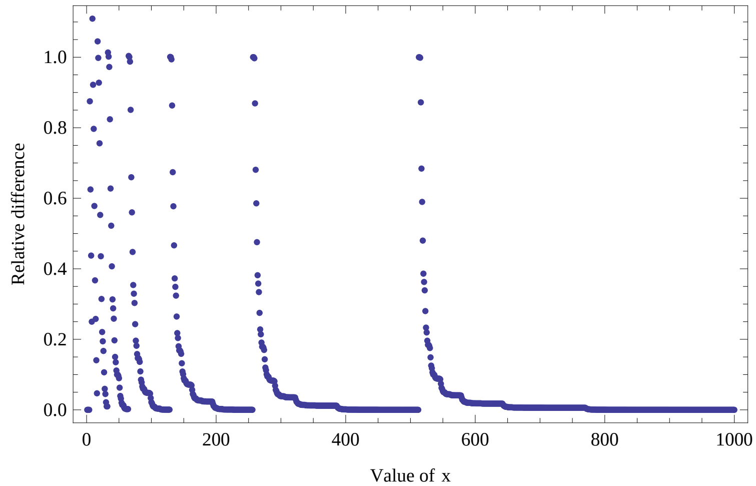

Figure 3.3: Four Paul Paradox, relative difference

In their paper, Csörgö and Simons prove that 2^k-Paul problem holds, i.e.,

$$\mathbf{P}(\sum_{i=1}^{2^k} X_i \geq x) \geq \mathbf{P}(2^k X_1 \geq x) \text{ for all } x > 2^k, \tag{3.14}$$

where $k \in \mathbf{N}$, but Three-Paul problem does not hold.

References

[1] Daniel Bernoulli. Exposition of a new theory on the measurement of risk. *Econometrica: Journal of the Econometric Society*, pages 23–36, 1954.

[2] Y.S. Chow and Herbert Robbins. On sums of independent random variables with infinite moments and "fair" games. *Proceedings of the National Academy of Sciences of the United States of America*, 47(3):330, 1961.

[3] S Csörgo and G Simons. The two-paul paradox and the comparison of infinite expectations. *Limit theorems in probability and statistics*, 1:427–455, 2002.

[4] Sándor Csörgö and Gordon Simons. On steinhaus' resolution of the st. petersburg paradox. *Probability and Mathematical Statistics*, 14.

[5] Jacques Dutka. On the st. petersburg paradox. *Archive for History of Exact Sciences*, 39(1):13–39, 1988.

[6] William Feller. Note on the law of large numbers and" fair" games. *The Annals of Mathematical Statistics*, 16(3):301–304, 1945.

[7] Willliam Feller. *An introduction to probability theory and its applications*, volume I. John Wiley & Sons, 3rd edition, 1968.

[8] Willy Feller. Ueber das gesetz der grossen zahlen. *Acta Litt. Scient. Regiae Univ. Hungaricae Francisco-Iosephinae*, pages 1936–37, 1937.

[9] Gérard Jorland. The saint petersburg paradox 1713–1937. *The probabilistic revolution*, 1:157–190, 1987.

[10] Emil Kauder. Genesis of the marginal utility theory: from aristotle to the end of the eighteenth century. *The Economic Journal*, 63(251):638–650, 1953.

[11] Sydney Lupton. The st. petersburg problem. *Nature*, 41:165–166, 1889.

[12] Anders Martin-Löf. A limit theorem which clarifies the'petersburg paradox'. *Journal of Applied Probability*, pages 634–643, 1985.

[13] Hugo Steinhaus. The so-called petersburg paradox. *Colloq. Math*, 2:56–58, 1949.

Appendix A

Proofs

A.1 Two-Paul problem

Given X, X_1, X_2 are independent and identically distributed random variables, where the common distribution is

$$\mathbf{P}(X = 2^j) = \frac{1}{2^{j+1}}, \quad j = 0, 1, \ldots$$

Show that for all $x > 0$,

$$\mathbf{P}(X_1 + X_2 \geq x) \geq \mathbf{P}(2X_1 \geq x) \tag{A.1}$$

Proof: The right side of (A.1) is easy to find. In fact,

$$\mathbf{P}(2X_1 \geq x) = \mathbf{P}(X_1 \geq x/2) = \sum_{j=\lceil \log_2(x/2) \rceil}^{\infty} \frac{1}{2^{j+1}} = \frac{1}{2^{\lceil \log_2(x/2) \rceil}} = \left(\frac{1}{2}\right)^{j_x}, \tag{A.2}$$

where $\lceil x \rceil$ represents the ceiling function of x, which returns the smallest integer not less than x. Also, j_x in the formula is defined as a function of x, $j_x := \lceil \log_2 x \rceil - 1$.

Now we define $M := \max\{X_1, X_2\}$, and consider the event on left side of (A.1) as a union of three exclusive events: (i) $\{M < 2^{j_x}, X_1 + X_2 \geq x\}$, (ii) $\{M > 2^{j_x}, X_1 + X_2 \geq x\}$, (iIi) $\{M = 2^{j_x}, X_1 + X_2 \geq x\}$.

For (i), if $M < 2^{j_x}$, then $M \leq 2^{j_x - 1}$ and hence $X_1 + X_2 \leq 2M \leq 2^{j_x} < x$. Thus, $X_1 + X_2 \geq x$ implies $M \geq 2^{j_x}$. So the probability of the event in (i) is 0.

For (ii), if $M \geq 2^{j_x+1}$, then $X_1 + X_2 \geq M + 1 \geq 2^{j_x+1} \geq x$. Then

$$\begin{aligned}
\mathbf{P}(M > 2^{j_x}, X_1 + X_2 \geq x) &= \mathbf{P}(X \geq 2^{j_x+1}) \\
&= 1 - \mathbf{P}(M < 2^{j_x+1}) \\
&= 1 - \mathbf{P}(X_1 \leq 2^{j_x}) \cdot \mathbf{P}(X_1 \leq 2^{j_x}) \\
&= 1 - \left(1 - \frac{1}{2^{j_x+1}}\right)\left(1 - \frac{1}{2^{j_x+1}}\right) \\
&= \left(\frac{1}{2}\right)^{j_x} - \left(\frac{1}{2}\right)^{2j_x+2}
\end{aligned}$$

For (iii), one has

$$\begin{aligned} X_1 \geq x - 2^{j_x} &\Leftrightarrow \log_2 X_1 \geq \lceil \log_2(x - 2^{j_x}) \rceil =: k_x \\ &\Leftrightarrow X_1 \geq 2^{k_x}. \end{aligned}$$

Then

$$\begin{aligned} &\{M = 2^{j_x}, X_1 + X_2 \geq x\} \\ &= \{X_1 = 2^{j_x}, x - 2^{j_x} \leq X_2 \leq 2^{j_x}\} \cup \{X_2 = 2^{j_x}, x - 2^{j_x} \leq X_1 < 2^{j_x}\} \\ &= \{X_1 = 2^{j_x}, 2^{k_x} \leq X_2 \leq 2^{j_x}\} \cup \{X_2 = 2^{j_x}, 2^{k_x} \leq X_1 \leq 2^{j_x - 1}\}. \end{aligned}$$

Since X_1 and X_2 are independent, one has

$$\begin{aligned} &\mathbf{P}(M = 2^{j_x}, X_1 + X_2 \geq x) \\ &= \mathbf{P}(X_1 = 2^{j_x}) \cdot \mathbf{P}(2^{k_x} \leq X_2 \leq 2^{j_x}) + \mathbf{P}(X_2 = 2^{j_x}) \cdot \mathbf{P}(2^{k_x} \leq X_1 \leq 2^{j_x - 1}) \\ &= \frac{1}{2^{j_x+1}} \left(\frac{1}{2^{k_x}} - \frac{1}{2^{j_x+1}} \right) + \frac{1}{2^{j_x+1}} \left(\frac{1}{2^{k_x}} - \frac{1}{2^{j_x}} \right) \\ &= \left(\frac{1}{2}\right)^{j_x+k_x} - \left(\frac{1}{2}\right)^{2j_x+2} - \left(\frac{1}{2}\right)^{2j_x+1} \end{aligned}$$

Combining results from (i), (ii) and (iii), one has

$$\mathbf{P}(X_1 + X_2 \geq x) = \left(\frac{1}{2}\right)^{j_x} + \left(\frac{1}{2}\right)\left[\left(\frac{1}{2}\right)^{k_x} - \left(\frac{1}{2}\right)^{j_x}\right] \tag{A.3}$$

Comparing (A.2) and (A.3), with the fact that $k_x \leq j_x$ for all $x > 0$, one can obtain the wanted (A.1).

This completes the proof.

$\square$

The significance of the difference can be seen in Figure 3.2.

A.2 Four-Paul problem

For the corresponding four Paul problem, one needs to Show that for all $x > 0$,

$$\mathbf{P}(X_1 + X_2 + X_3 + X_4 \geq x) \geq \mathbf{P}(4X_1 \geq x) \tag{A.4}$$

The right side of (A.4) is easy to find. In fact,

$$\mathbf{P}(4X_1 \geq x) = \mathbf{P}(X_1 \geq x/4) = \sum_{j=\lceil \log_2(x/4) \rceil}^{\infty} \frac{1}{2^{j+1}} = \frac{1}{2^{\lceil \log_2(x/4) \rceil}} = \left(\frac{1}{2}\right)^{j_x - 1}, \tag{A.5}$$

where $j_x := \lceil \log_2 x \rceil - 1$. The right hand side of (A.4) can be found by

$$\mathbf{P}(X_1+X_2+X_3+X_4 \geq x) = \sum_{y=2}^{x-2} \mathbf{P}(X_1+X_2=y)\mathbf{P}(X_3+X_4 \geq x-y) + P(X_1+X_2 \geq x-1). \tag{A.6}$$

In the right hand right of this expression, the last two probability can be found by (A.3). The first probability can also be found by using the following expression and (A.3)

$$\mathbf{P}(X_1+X_2=y) = \mathbf{P}(X_1+X_2 \geq y) - \mathbf{P}(X_1+X_2 \geq x+1) \tag{A.7}$$

The expression of (A.6) can not be simplified, but we can use Mathematica to calculate their values for each x. The significance of the difference can be seen in Figure 3.3.