Article

Entrance fees and a Bayesian approach to the St. Petersburg Paradox

Diego Marcondes ‡, Cláudia Peixoto ‡*, Kdson Souza ‡ and Sergio Wechsler ‡

Institute of Mathematics and Statistics, University of São Paulo, Butantã, São Paulo 03178-200, Brazil; diegormmarcondes@gmail.com; kdson19@gmail.com; sw@ime.usp.br

* Correspondence: claudiap@ime.usp.br; Tel.: +55 11 3091-6142

‡ These authors contributed equally to this work.

Abstract: In his best-selling book *An Introduction to Probability Theory and its Applications*, W. Feller established a way of ending the St. Petersburg Paradox by the introduction of an entrance fee, and provided it for the case in which the game is played with a fair coin. A natural generalization of his method is to establish the entrance fee for the case in which the probability of head is θ ($0 < \theta < 1/2$). The deduction of those fees is the main result of Section 2. We then propose a Bayesian approach to the problem. When the probability of head is θ ($1/2 < \theta < 1$) the expected gain of the St. Petersburg Game is finite, therefore there is no paradox. However, if one takes θ as a random variable assuming values in $(1/2, 1)$ the paradox may hold, what is counter-intuitive. On Section 3 we determine a necessary and sufficient condition for the absence of paradox on the Bayesian approach and on Section 4 we establish the entrance fee for the case in which θ is uniformly distributed in $(1/2, 1)$, for in this case there is paradox.

Keywords: St. Petersburg Paradox; entrance fees; Bayesian analysis

MSC: 60F05

---

1. Introduction

The St. Petersburg Paradox was first discussed in letters between Pierre Rémond de Montmort and Nicholas Bernoulli, dated from 1713, and was supposedly invented by the latter [13]. Since then, the paradox has been one of the most famous examples in probability theory and has been generalized by economists, philosophers and mathematicians, being widely applied in Decision Theory and Theory of Games.

The Paradox arises from a very simple coin tossing game. A player tosses a fair coin until it falls head. If this occurs at the *ith* toss, the player receives 2^i money unities. However, the paradox appears when one engage in determining the fair quantity that should be paid by a player to play a trial of this game. Indeed, in the eighteenth century, when the paradox was first studied, the probability theory main goal was to answer questions that involved gambling, especially the problem of defining fair quantities that should be paid to play a certain game or if a game was interrupted before its end. Those quantities were then named *moral value* or *moral price* and are what we call *expected utility* nowadays. It is intuitive to take the expected gain at a trial of the game or the expected gain of a player if the game were not to end now, respectively, as *moral values*, and that is what the first probabilists established as those fair quantities. Nevertheless, when the expected gain was not finite, they would not know what to do, and that is exactly what happens at the St. Petersburg Game.

In fact, let X denote the gain of a player at a trial of the game. Then, the expected gain at a trial of the game is given by

$$E(X) = \sum_{i=1}^{\infty} 2^i \frac{1}{2^i} = \infty,$$

for the random variable that represents the toss in which the first head appears has a Geometric distribution with parameter $1/2$.

Besides the fact that the gain is indeed a random variable with infinite expected value, the St. Petersburg Paradox drew the attention of many mathematicians and philosophers throughout the years because it treats a game that nobody in its right state of mind would want to pay large quantities to play, and that is what intrigued great mathematicians as Laplace, Gabriel Cramer and the Bernoulli family. Indeed, at that time, not finite expected values were quite disturbing, for not even the simplest Laws of Large Numbers had been proved and the difference between a random variable and its expected value was not clear yet. In this scenario, the St. Petersburg Paradox emerged and took its place in the probability theory.

The magnitude of the paradox may be exemplified by the comparison between the odds of wining in the lottery and recovering the value paid to play the game, as showed in [9], in the following way. Suppose a lottery that pays 50 million money unities to the gambler that hit the 6 numbers drawn out of sixty. For simplification, suppose that there is no possibility of a draw and that the winner always gets the whole 50 million. Now, presume a gambler can bet on 16 numbers, i.e., bet on $\binom{16}{6} = 8008$ different sequences of 6 numbers, paying 10.000 money unities, and let p be his probability of wining. Then, it is straightforward that

$$p = \frac{\binom{16}{6}}{\binom{60}{6}} = 1,59 \times 10^{-4}.$$

On the other hand, if one pays 10.000 money unities to play the St. Petersburg Game, he will get his money back (and eventually profit) if the first head appears on the 14^{th} toss or later. Therefore, letting q be the probability of the gambler recovering its money, it is easy to see that

$$q = \mathbb{P}\{X \geq 14\} = 1 - \mathbb{P}\{X < 14\} = 1 - \sum_{i=1}^{13} \frac{1}{2^i} = 1,22 \times 10^{-4}$$

and $p > q$. Then, winning the lottery can be more likely than recovering the money at a St. Petersburg Game. Of course, it is true only for a 10.000 money unities (or more) lottery bet. Thus, if one is interested in investing some money, the lottery may be a better investment than the St. Petersburg Game. Although, I would rather take my chances at Wall Street for p and q are way too small.

Nevertheless, the paradox gets even more interesting when we compare the lottery on the scenario above and the St. Petersburg Game about their expected gain. On the one hand, as seen before, the expected gain at the St. Petersburg Game is infinite. On the other hand, the expected gain at the lottery on the scenario above is

$$E(\text{gain at the lottery}) = 50p \times 10^6 - 10^4 \approx -2050,$$

and there it is the St. Petersburg Paradox brought to a twenty-first century context in which even the layman can understand its most intrinsic problem: how can one expect to win an infinite amount of money, but have at the same time less probability of winning anything at all than someone that expects to loose money? Now imagine the impact of this result on the eighteenth century mathematicians, who did not know the modern probability theory as invented by Kolmogorov [6] and that facilitates in a great deal the understanding of the St. Petersburg Paradox.

A first attempt to solve the paradox was made independently by Daniel Bernoulli on a 1738 paper translated to [2] and Gabriel Cramer on a 1728 letter to Nicholas Bernoulli available on [13]. In Bernoulli's approach, *the value of an item is not to be based on its price, but rather on the utility it yields*. In fact, the price of an item depends solely on the item itself and is equal for every buyer, albeit the utility is dependent on the particular circumstances of the purchase and the person making it. In this context, we have his famous assertion that *a gain of one thousand ducats is more significant to a pauper than to a rich man though both gain the same amount*.

In his approach, the *gain* at a St. Petersburg Game is interchanged by the *utility* it yields, i.e., a *linear* utility function that represents the monetary gain of the game is interchanged by a *logarithm*

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 February 2017 doi:10.20944/preprints201702.0064.v1

Peer-reviewed version available at Philosophies 2017, 2, , 11; doi:10.3390/philosophies2020011

utility function. If a logarithm utility function is used instead, the paradox disappear, for now the expected incremental utility is finite and, therefore, may be used to determine a fair quantity to be paid to play a trial of the game. The logarithm utility function takes into account the total wealth of the person willing to play the game, what makes the utility function dependent on whom is playing the game, and not only on the game itself, in the following way.

Let w be the total wealth of a player, i.e., all the money he has available to gamble at a given moment. We define the utility of a monetary quantity m as $U(m) = \log m$, in which the logarithm is the natural one. Therefore, letting c be the value that should be paid to play a trial of the game, the expected incremental utility, i.e., how much utility the player expects to gain, is given by

$$\Delta E(U) = \sum_{i=1}^{\infty} \frac{1}{2^i} \Big[U(w + 2^i - c) - U(w)\Big] = \Big[\sum_{i=1}^{\infty} \frac{1}{2^i} \log(w + 2^i - c)\Big] - \log(w) < \infty,$$

and we have an implicit relation between the total wealth of a player and how much he must pay to expect to gain a finite utility $\Delta E(U)$. Table 1 shows the values of c for which $\Delta E(U) = 0$, i.e., the maximum monetary unities someone must be willing to pay to play the game, for different values of w.

Table 1. The maximum value c someone with total wealth w must be willing to pay to play a trial of the St. Petersburg Game according to the logarithm utility function, for different values of w.

Total Wealth (w)	**Maximum Values** (c)	**Total Wealth** (w)	**Maximum Values** (c)
2	3,35	500	9,99
4	4	1.000	10,96
20	5,77	10.000	14,24
50	6,90	100.000	17,56
100	7,80	500.000	19,88
200	8,73	1.000.000	20,88

Note that, if $w \leq 4$, one must be willing to give all his money and eventually take a loan to play the game. In fact, the game is advantageous for players that have little wealth and, as richer the player is, less will be the percentage of his wealth that he must be willing to pay to play the game.

Although the logarithm utility function solves the St. Petersburg Paradox, the paradox immediately reappears if the gain when the first head occurs at the *ith* trial is interchanged from 2^i to $2^{M(i)}$, in which $M(i)$ is another function of i and not the identity one. For example, if $M(i) = 2^i$ the paradox holds even with the logarithm utility function. This fact was outlined by [8] that created the so-called Super St. Petersburg Paradoxes, in which $M(i)$ may take different forms.

There are many others generalizations and solutions to the St. Petersburg Paradox, as shown in [10], [11] and [12], for example. However, the main goal of this paper is to present and generalize W. Feller's solution by the introduction of an entrance fee, as will be defined on the next section. We generalize his method for the game in which the coin used is not fair, i.e., its probability of head is $\theta \neq 1/2$, and for the case in which the coin is a random one, i.e., its probability θ of head is a random variable defined on $(0, 1)$.

2. Entrance Fee

A way of ending the paradox, i.e., making the game fair in the sense discussed by [5], is to define an entrance fee e_n that should be paid by a player so he can play n trials of the St. Petersburg Game. This solution differs from the solutions given above on the fact that it does not take into account the utility function of a player, although it is rather theoretical, for it determines the cost of the game based on the convergence in probability of a convenient random variable. Furthermore, Feller's solution is given for the scenario in which the player pays to play n trials of the game, i.e., the coin is tossed until n heads appear and a trial of the game ends when a head appears. In Feller's definition, the entrance fee e_n will be fair if

$$\lim_{n\to\infty} \mathbb{P}\left\{\left|\frac{S_n}{e_n} - 1\right| > \delta\right\} = 0$$

in which X_k is the gain at the *kth* trial of the game, $S_n = \sum_{k=1}^{n} X_k$ and $\delta > 0$. If the coin is fair, it was proved by [4] that $e_n = n \log_2 n$.

Note that, of course, $\lim\limits_{n\to\infty} e_n = \infty$. However, e_n gives the *rate* in which the accumulated gain at the trials of the game increase, so that paying the entrance fee e_n may be considered fair, for it will be as closely to the accumulated gain in n trials of the game as desired with probability 1, as the number of trials diverges. Although classic and well-known in probability theory, Feller's solution may still be generalized to the case in which the coin used at the game is not fair, in the following way.

As a generalization of the St. Petersburg Game, one could toss a coin with probability θ $(0 < \theta < 1)$ of head. If $0 < \theta \leq 1/2$ the paradox still holds, for

$$E(X) = \sum_{i=1}^{\infty} 2^i \theta (1-\theta)^{i-1} \geq \sum_{i=1}^{\infty} 2^i \frac{\theta}{2^{i-1}} = \infty. \tag{1}$$

However, if $1/2 < \theta < 1$,

$$E(X) = \sum_{i=1}^{\infty} 2^i \theta (1-\theta)^{i-1} = \frac{2\theta}{2\theta - 1} < \infty, \tag{2}$$

so there is no paradox. Applying the same method used by [4], it can be shown that for a coin with probability θ $(0 < \theta < \frac{1}{2})$ of head the entrance fee is given by

$$e_n = \frac{2\theta}{1-2\theta} n \left(n^{-1+\log_\alpha 2} \log_\alpha n - 1\right); \alpha = \frac{1}{1-\theta}.$$

Then, we have the following theorem that remained an open problem for the last decades and is originally solved in this paper. Proofs for all results are presented on the Appendix.

Let $\alpha = \frac{1}{1-\theta}$. If $0 < \theta < 1/2$ and $e_n = \frac{2\theta}{1-2\theta} n \left(n^{-1+\log_\alpha 2} \log_\alpha n - 1\right)$ then

$$\lim_{n\to\infty} \mathbb{P}\left\{\left|\frac{S_n}{e_n} - 1\right| > \delta\right\} = 0.$$

Figure 1 shows the Entrance Fees for selected n and θ. It is important to note that the limit of e_n, as defined in Theorem 2, as θ increases to $1/2$, does not equal the entrance fee $n \log_2 n$ for the fair coin game. In fact,

$$\lim_{\theta\to 1/2} e_n = \lim_{\theta\to 1/2} \frac{2\theta}{1-2\theta} n \left(\log_2 n - 1\right) = \infty$$

for all values of n. Therefore, the Entrance Fee have the interesting phase transition property on $\theta = 1/2$. Furthermore, as smaller the value of θ, bigger will be entrance fee for any fixed n. This behaviour is expected for as smaller the value of θ, longer it will take for the first head to show up and more money the player expect to gain.

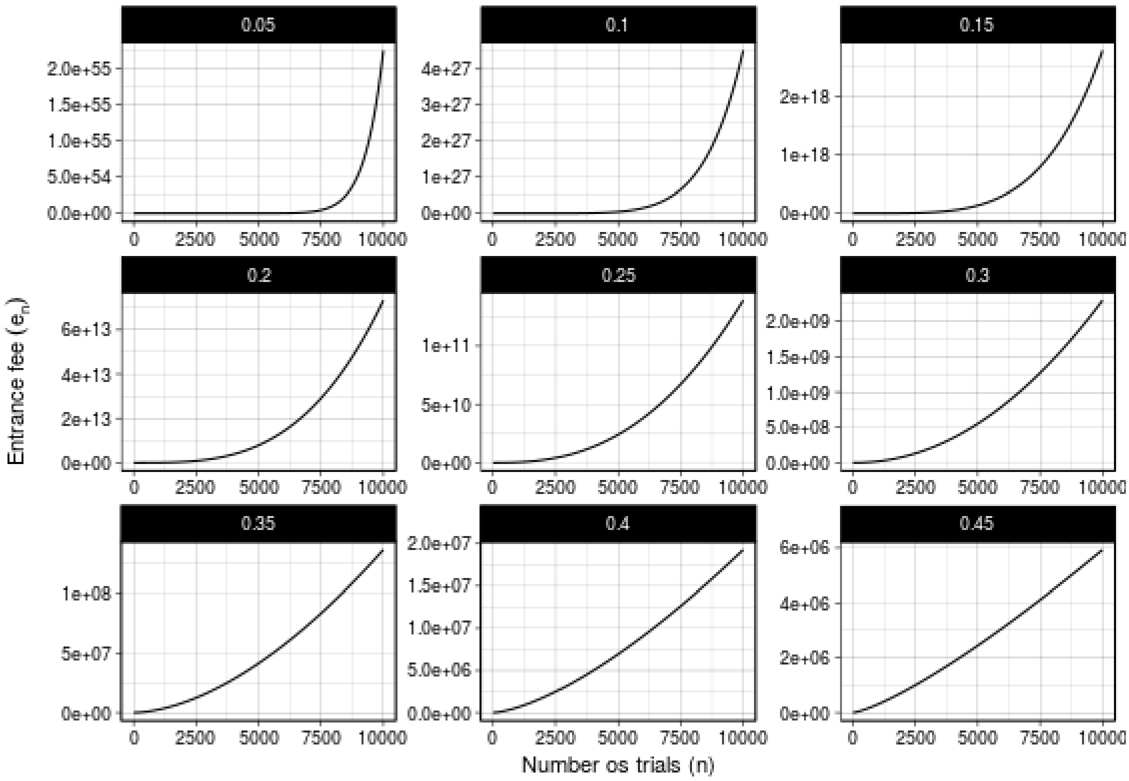

Figure 1. Entrance fees for different values of n and for $\theta \in \{0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45\}$.

3. Bayesian Approach

Persi Diaconis once said that there is much further work to be done by incorporating uncertainty and prior knowledge into basic probability calculations and developed himself a Bayesian version of three classical problems on [3]. Following his advice, we will develop a Bayesian approach to the St. Petersburg Paradox.

The Bayesian paradigm started with Reverend Thomas Bayes, whose work led to what is known today as Baye's theorem, and was first published posthumously on 1763 by Richard Price [1]. Although his work inspired great probabilists and statisticians of the following centuries, it was not until the second half of the twentieth century that the Bayesian paradigm was popularized within the statistical society, for the development of computer technology made it possible to apply Bayesian Analysis on a variety of theoretical and practical problems.

In the centuries following Baye's work, another branch of statistics, the frequentist one, was developed and exhaustively studied by probabilists, mathematicians and, of course, statisticians. The main characteristic of the frequentist statistics is that it treats the probability of an event as the limit of the ratio between the number of times the event occur and the number of times the experiment, in which the event is a possible outcome, is repeated, the limit being taken as the latter diverges. The sharper reader may immediately conclude that this interpretation of probability has an intrinsic problem: not every experiment can be *replicated*. Consider, for example, the probability of raining tomorrow. By a frequentist perspective, this probability is the limit of the ratio between the number of times that *rain* tomorrow and the number of times that tomorrow *happens*. However, it is naive to think of probability in this way, for tomorrow will happen only once. Therefore, there is a need for a more general interpretation of probability, and that is where the Bayesian approach comes into play.

On the Bayesian approach, the concept of probability may be interpreted as measure of knowledge: how much it is known about the odds of an event to occur. Of course, the concept of probability on Bayesian statistics may be considered by a frequentist statistician as subjective, for

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 February 2017 doi:10.20944/preprints201702.0064.v1

Peer-reviewed version available at *Philosophies* **2017**, *2*, , 11; doi:10.3390/philosophies2020011

each person may have a different knowledge about the likelihood of an event and, therefore, different people may have distinct conclusions analysing the same data. The debate between the Bayesian and the frequentist approach has been going on for the last decades, and we are not going to get into its merits. However, for a succinct and introductory presentation of the Bayesian approach see [7].

On the context of the St. Petersburg Paradox, the main difference between the frequentist and Bayesian approach relies upon the way in which the parameter θ, i.e., the probability of head of the game's coin, is treated. On the frequentist approach, θ is a *number* on the interval $(0,1)$, either known or unknown. If it is known, as it was on Sections 1 and 2, there is nothing to be done from a statistical point of view. However, if it is unknown, it may be estimated and hypothesis about it may be tested. On the other hand, on the Bayesian approach, the parameter θ is a random variable, i.e, has a probability distribution $F(a) = \mathbb{P}\{\theta \leq a\}$, that is called the *priori* distribution of θ. The St. Petersburg Game played with a coin with probability θ_0 may also be interpreted from a Bayesian point of view, as it is enough to take

$$F(a) = \begin{cases} 1, & \text{if } a \geq \theta_0 \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

However, in this section, we will determine conditions on F for which the paradox does not hold, i.e., $E(X) < \infty$. For this purpose, we have first to define the probability distribution of X (the gain at a trial of the Petersburg game) from a Bayesian point of view. Applying basics properties of the probability measure we have that

$$\mathbb{P}\{X = x\} = \int_0^1 \mathbb{P}\{X = x|\theta\} dF(\theta) = \int_0^1 \theta(1-\theta)^{x-1} dF(\theta), \tag{4}$$

for $X|\theta$ has a Geometric distribution with parameter θ. Therefore, the expected gain $E(X)$ at a trial of the game is given, applying (4), by

$$E(X) = \sum_{x=1}^{\infty} x \int_0^1 \mathbb{P}\{X = x|\theta\} dF(\theta) = \int_0^1 \sum_{x=1}^{\infty} x\mathbb{P}\{X = x|\theta\} dF(\theta) = \int_0^1 E(X|\theta) dF(\theta)$$

If F is given by (3), it was proved on (1) and (2) that if $\theta_0 \leq 1/2$ the paradox holds, whilst if $\theta_0 > 1/2$ there is no paradox. This case has the same conclusion from a frequentist point of view, therefore is not of interest. The Bayesian approach gets more appealing when F takes other forms, distinct from (3). As a motivation for the Bayesian approach, we present the case in which θ is uniformly distributed in $(1/2, 1)$, that raises some interesting questions about the Bayesian approach.

From (2), we know that if $1/2 < \theta < 1$, the expected gain of the game is finite, therefore there is no paradox. However, suppose one incorporates prior knowledge to θ and takes it as a random variable assuming values in $(1/2, 1)$ with probability 1 and distribution F. It would be expected that the paradox would not hold for any distribution F with such properties. However, it does not happen.

Suppose θ uniformly distributed in $(1/2, 1)$. Then,

$$E(X) = \int_{1/2}^{1} \sum_{x=1}^{\infty} 2^x \theta(1-\theta)^{x-1} 2d\theta = \sum_{x=1}^{\infty} 2^{x+1} \int_0^{1/2} (1-p)p^{x-1} dp = \infty. \tag{5}$$

The result (5) raises a couple of interesting questions about the Bayesian approach to the St. Petersburg Paradox. Firstly, one may ask for which prior distributions F the paradox holds. Theorems 3 and 3 answer this question, giving a necessary and sufficient condition for the absence of paradox on the Bayesian approach. Then, one could inquire what is the entrance fee for the case in which θ is a random variable uniformly distributed in $(1/2, 1)$. On Section 4, we establish the entrance fee for this case.

We now treat the case in which the probability θ of head in each trial of the game is a random variable, defined on $(\Omega, \mathbb{F}, \mathbb{P})$, taking values on the interval $(1/2, 1)$ with probability 1 and distribution F. The case in which $F(1/2) > 0$ is not of interest, for the paradox holds for any F with such property. Therefore, in this scenario, the expected gain at a trial of the game is given by

$$E(X) = \int_{1/2}^{1} E(X|\theta)dF(\theta) = \int_{1/2}^{1/2+\epsilon} E(X|\theta)dF(\theta) + \int_{1/2+\epsilon}^{1} E(X|\theta)dF(\theta)$$

for any fixed $0 < \epsilon < 1/2$. From now on, consider ϵ fixed in $(0, 1/2)$, unless said otherwise. Now,

$$\int_{1/2+\epsilon}^{1} E(X|\theta)dF(\theta) \leq \int_{1/2+\epsilon}^{1} E(X|\theta = 1/2+\epsilon)dF(\theta) \leq \frac{1+2\epsilon}{2\epsilon} < \infty$$

so that the paradox will exist if, and only if, $\int_{1/2}^{1/2+\epsilon} E(X|\theta)dF(\theta) = \infty$. The lemma bellow gives a sufficient condition for the absence of paradox that will be used to prove the main result of this section.

Suppose that F is absolutely continuous in $(1/2, 1/2+\epsilon)$ and $F'(\theta) = f(\theta), \theta \in (1/2, 1/2+\epsilon)$. If $f(\theta) \leq c(2\theta - 1), \forall \theta \in (1/2, 1/2+\epsilon)$ and for some $c \in \mathbb{R}_{+}$, there is no paradox.

The theorems below characterize the prior distributions for which the paradox does not hold. As the probability measure may be decomposed in a singular, a discrete and a continuous measure, we may treat first the singular and discrete case, and then the continuous one. If F is not absolutely continuous, i.e, is singular and/or discrete, in $(1/2, 1/2+\epsilon)$, the paradox does not hold. Now, if F is absolutely continuous in $(1/2, 1/2+\epsilon)$, the paradox holds if, and only if, the limit of the probability density of θ at $1/2$ is greater than zero.

If there exists a countable set $C \subset (1/2, 1/2+\epsilon)$ such that $\mathbb{P}\{\theta \in C, \theta \in (1/2, 1/2+\epsilon)\} = \mathbb{P}\{\theta \in C\}$ then $\int_{1/2}^{1/2+\epsilon} E(X|\theta)dF(\theta) < \infty$.

Suppose that F is absolutely continuous in $(1/2, 1/2+\epsilon)$ and $F'(\theta) = f(\theta), \theta \in (1/2, 1/2+\epsilon)$. There is no paradox if, and only if, $\lim\limits_{\theta \to 1/2} f(\theta) = 0$.

In summary, there will be no paradox if, and only if, there is no probability mass on $1/2$ nor around it. If F is not absolutely continuous on a neighbourhood of $1/2$ the paradox will clearly not hold, for all the probability mass will be at least a distance $\epsilon' > 0$ away from $1/2$. However, if F is absolutely continuous on a neighbourhood of $1/2$, there will be no paradox if, and only if, the probability density tends to zero as θ tends to $1/2$. Otherwise, there would be a probability mass in *every* neighbourhood ϵ' of $1/2$.

4. Uniformly Distributed θ

The method of [4] will now be applied to find the entrance fee for the game in which the probability θ of head in each trial is a random variable with Uniform distribution in $(1/2, 1)$, i.e. $f(\theta) = 2\mathbb{1}_{\{\theta \in (1/2,1)\}}$. As $\lim\limits_{\theta \to 1/2} f(\theta) = 2$, by Theorem 3, in this scenario there is a paradox. It will be shown that the entrance fee in this case is $e_n = n \log_e(\log_2 n)$.

If the probability of head θ in each trial of the game is a random variable uniformly distributed in $(1/2, 1)$ and $e_n = n \log_e(\log_2 n)$ then

$$\lim_{n \to \infty} \mathbb{P}\left\{\left|\frac{S_n}{e_n} - 1\right| > \delta\right\} = 0.$$

Figure 2 shows the entrance fee for different values of n.

Peer-reviewed version available at *Philosophies* **2017**, *2*, , 11; doi:10.3390/philosophies2020011

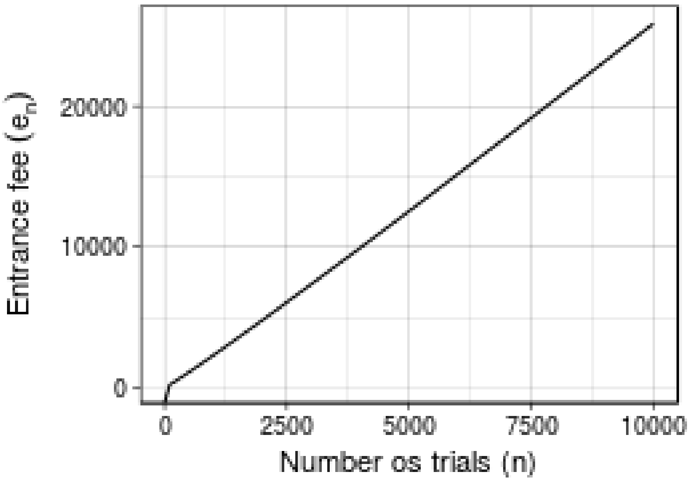

Figure 2. Entrance fees for different values of n for θ uniformly distributed in $(1/2, 1)$

Note that the entrance fee in this case is smaller then the entrance for the case in which $\theta = 1/2$ with probability 1. It is expected, for the probability mass that was first concentrated on $1/2$ is now spread equally over the interval $(1/2, 1)$, an interval in which there is no paradox. In fact, what is really interesting in this Bayesian approach is the fact that, if θ is uniformly distributed in $(1/2, 1)$, the probability of a θ, for which the paradox holds, be chosen is zero, although the expected gain is infinite and, in overall, there is paradox.

A suggestion for further studies would be to establish the entrance fee for random θ with other distributions, e.g. truncated Pareto like distributions with scale parameter equals to $1/2$ and Beta distributions.

"The authors declare no conflict of interest."

Appendix

Proof of Theorem 1. Let $\{(U_k, V_k) : k = 1, 2, \dots\}$ be a sequence of random variables such that

$$\begin{cases} (U_k, V_k) = (X_k, 0), & \text{if } X_k \leq (n \log_\alpha(\log_\alpha n))^{\log_\alpha 2} \\ (U_k, V_k) = (0, X_k), & \text{if } X_k > (n \log_\alpha(\log_\alpha n))^{\log_\alpha 2}. \end{cases}$$

Therefore,

$$\mathbb{P}\left\{\left|\frac{S_n}{e_n} - 1\right| > \delta\right\} \leq \mathbb{P}\left\{\left|\sum_{k=1}^{n} U_k - e_n\right| > \delta e_n\right\} + \mathbb{P}\left\{\sum_{k=1}^{n} V_k \neq 0\right\}. \tag{6}$$

Let r be the greatest integer such that $\alpha^r \leq n \log_\alpha(\log_\alpha n) < \alpha^{r+1}$. Then,

$$r = \log_\alpha \alpha^r \leq \log_\alpha(n \log_\alpha(\log_\alpha n)) < \log_\alpha \alpha^{r+1} = r + 1$$
$$\implies 2^r \leq 2^{\log_\alpha(n \log_\alpha(\log_\alpha n))} \leq 2^{r+1}.$$

Now, defining $\beta = n \log_\alpha(\log_\alpha n)$,

$$2^{\log_\alpha(n \log_\alpha(\log_\alpha n))} = 2^{\log_\alpha \beta} = \beta^{\log_\beta 2^{\log_\alpha \beta}} = \beta^{\log_\alpha 2} = (n \log_\alpha(\log_\alpha n))^{\log_\alpha 2}$$
$$\implies 2^r \leq (n \log_\alpha(\log_\alpha n))^{\log_\alpha 2} \leq 2^{r+1}.$$

Peer-reviewed version available at Philosophies 2017, 2, , 11; doi:10.3390/philosophies2020011

Applying the inequalities above to the second term of (6) we get

$$\begin{aligned}\mathbb{P}\left\{\sum_{k=1}^{n} V_k \neq 0\right\} \leq & n\mathbb{P}\{X_1 > (n\log_\alpha(\log_\alpha n))^{\log_\alpha 2}\} \leq n\mathbb{P}\{X_1 > 2^r\} \leq \\ \leq & n(1-\theta)^r \leq \frac{n(1-\theta)^{-1}}{\alpha^{r+1}} \leq \frac{n(1-\theta)^{-1}}{n\log_\alpha(\log_\alpha n)} \xrightarrow{n\to\infty} 0.\end{aligned}$$

Note that

$$\begin{aligned} E(U_k) &= \sum_{i=1}^{r} 2^i\theta(1-\theta)^{i-1} = \frac{2\theta}{1-2\theta}\left\{[2(1-\theta)]^r - 1\right\} \\ E(U_k^2) &= \sum_{i=1}^{r} 2^{2i}\theta(1-\theta)^{i-1} = \frac{4\theta}{3-4\theta}\left\{[4(1-\theta)]^r - 1\right\} \\ Var\left(\sum_{k=1}^{n} U_k - nE(U_k)\right) &\leq nE(U_k^2) \leq 2n[4(1-\theta)]^r = 2n\alpha^{\log_\alpha[4(1-\theta)]^r} \leq \\ &\leq 2n[\alpha^r]^{-1+2\log_\alpha 2} \leq 2n(n\log_\alpha(\log_\alpha n))^{-1+2\log_\alpha 2}.\end{aligned}$$

Taking

$$e_n = \frac{2\theta}{1-2\theta} n(n^{-1+\log_\alpha 2}\log_\alpha n - 1) \geq \frac{\theta}{1-2\theta} n\left(n^{-1+\log_\alpha 2}\log_\alpha n\right)$$

and applying the Chebyshev's inequality to the first term of (6) we get

$$\begin{aligned}\mathbb{P}\left\{\left|\sum_{k=1}^{n} U_k - nE(U_k)\right| > \delta e_n\right\} &\leq \frac{2n^{2\log_\alpha 2}(\log_\alpha(\log_\alpha n))^{-1+2\log_\alpha 2}}{\delta^2\left(\frac{\theta}{1-2\theta}\right)^2 n^{2\log_\alpha 2}(\log_\alpha n)^2} \leq \\ &\leq \frac{2(\log_\alpha(\log_\alpha n))^{-1+2\log_\alpha 2}}{\delta^2\left(\frac{\theta}{1-2\theta}\right)^2 (\log_\alpha n)^2} \xrightarrow{n\to\infty} 0.\end{aligned}$$

Now, it is enough to prove that $nE(U_k)$ may be approximated by e_n for great values of n. By definition,

$$\begin{aligned} &\alpha^r \leq n\log_\alpha(\log_\alpha n) < \alpha^{r+1} \implies \log_\alpha n \leq r \leq \log_\alpha n + \log_\alpha(\log_\alpha(\log_\alpha n)) \\ &\implies \lim_{n\to\infty} \frac{\log_\alpha n}{\log_\alpha n + \log_{2(1-\theta)}(\log_\alpha n)} \leq \lim_{n\to\infty} \frac{r}{\log_\alpha n + \log_{2(1-\theta)}(\log_\alpha n)} \\ &\qquad \leq \lim_{n\to\infty} \frac{\log_\alpha n + \log_\alpha(\log_\alpha(\log_\alpha n))}{\log_\alpha n + \log_{2(1-\theta)}(\log_\alpha n)} \\ &\implies 1 \leq \lim_{n\to\infty} \frac{r}{\log_\alpha n + \log_{2(1-\theta)}(\log_\alpha n)} \leq 1 \implies r \approx \log_\alpha n + \log_{2(1-\theta)}(\log_\alpha n) \\ &\implies r \approx \log_\alpha n + (\log_{2(1-\theta)}\alpha)(\log_\alpha(\log_\alpha n)) \implies r \approx \log_\alpha[n(\log_\alpha n)^{\log_{2(1-\theta)}\alpha}] \\ &\qquad \implies \alpha^r \approx n(\log_\alpha n)^{\log_{2(1-\theta)}\alpha}.\end{aligned}$$

Hence

$$\begin{aligned} &[2(1-\theta)]^r = \alpha^{\log_\alpha[2(1-\theta)]^r} = (\alpha^r)^{\log_\alpha[2(1-\theta)]} \approx [n(\log_\alpha n)^{\log_{2(1-\theta)}\alpha}]^{\log_\alpha[2(1-\theta)]} \\ &\qquad \implies [2(1-\theta)]^r \approx n^{-1+\log_\alpha 2}\log_\alpha n \\ &\implies nE(U_k) = \frac{2\theta}{1-2\theta} n\left\{[2(1-\theta)]^r - 1\right\} \approx \frac{2\theta}{1-2\theta} n\left\{n^{-1+\log_\alpha 2}\log_\alpha n - 1\right\} = e_n\end{aligned}$$

for great values of n. Therefore, for $e_n = \frac{2\theta}{1-2\theta} n \left\{ n^{-1+\log_\alpha 2} \log_\alpha n - 1 \right\}$,

$$\mathbb{P}\left\{ \left| \sum_{k=1}^{n} U_k - n\frac{2\theta}{1-2\theta}\left[n^{-1+\log_\alpha 2}\log_\alpha n - 1 \right] \right| > \delta e_n \right\} \xrightarrow{n\to\infty} 0$$
$$\implies \mathbb{P}\left\{ \left| \sum_{k=1}^{n} U_k - e_n \right| > \delta e_n \right\} \xrightarrow{n\to\infty} 0 \implies \lim_{n\to\infty} \mathbb{P}\left\{ \left| \frac{S_n}{e_n} - 1 \right| > \delta \right\} = 0.$$

□

Proof of Lemma 1:

$$\lim_{\theta\to 1/2} E(X|\theta)f(\theta) = \lim_{\theta\to 1/2} \frac{2\theta}{2\theta-1} f(\theta) \leq \lim_{\theta\to 1/2} \frac{2\theta}{2\theta-1} c(2\theta-1) = c < \infty$$
$$\implies \int_{1/2}^{1/2+\epsilon} E(X|\theta)dF(\theta) < \infty.$$

□

Proof of Theorem 2: Let C be a fixed countable set such that $C \subset (1/2, 1/2+\epsilon)$ and $\mathbb{P}\{\theta \in C, \theta \in (1/2, 1/2+\epsilon)\} = \mathbb{P}\{\theta \in C\}$. If $\mathbb{P}\{\theta = y\} = 0, \forall y \in C$, there is no paradox, for $\int_{1/2}^{1/2+\epsilon} E(X|\theta)dF(\theta) = 0$. Now suppose that exists a $y \in C$ such that $\mathbb{P}\{\theta = y\} > 0$ and let $\gamma = \min\{y \in C : \mathbb{P}\{\theta = y\} > 0\}$. Then,

$$\int_{1/2}^{1/2+\epsilon} E(X|\theta)dF(\theta) \leq \frac{2\gamma}{2\gamma - 1} < \infty.$$

□

Proof of Theorem 3: ($\implies$) Consider the extension g of f on $(1/2-\epsilon, 1/2+\epsilon)$ given by

$$g(\theta) = f(\theta)\mathbb{1}_{\{\theta\in[1/2,1/2+\epsilon)\}} - f(1-\theta)\mathbb{1}_{\{\theta\in(1/2-\epsilon,1/2)\}}$$

in which $\mathbb{1}$ is the indicator function. Note that

$$\int_{1/2}^{1/2+\epsilon} E(X|\theta)dF(\theta) = \int_{1/2}^{1/2+\epsilon} E(X|\theta)g(\theta)d\theta.$$

As g is odd around $1/2$ and continuous for a ϵ small enough, $g'(1/2)$ exists and $g''(1/2) = 0$. Therefore, for a ϵ small enough, g' and g'' exist for $\theta \in I = (1/2, 1/2+\epsilon)$. Consider ϵ such that g' and g'' exist on I and note that $0 \leq g'(\theta) < \infty, \theta \in I$, for f is a probability density.

On the one hand, if $g''(\theta) > 0$ for $\theta \in I$, g is concave upward and

$$g(\theta) \leq \frac{g(1/2+\epsilon)}{2\epsilon}(2\theta-1), \forall\theta \in I.$$

Taking $c = \frac{g(1/2+\epsilon)}{2\epsilon}$, the result follows from Lemma 3.

On the other hand, if $g''(\theta) \leq 0$ for $\theta \in I$, by the Taylor expansion of g around $1/2$, for $1/2 \leq \bar{\theta} \leq \theta$ and $\forall\theta \in I$,

$$g(\theta) = g'(1/2)(\theta - 1/2) + \frac{g''(\bar{\theta})}{2}(\theta-1/2)^2 \leq \frac{g'(1/2)}{2}(2\theta-1).$$

Taking $c = \frac{g'(1/2)}{2}$ the result follows from Lemma 3.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 February 2017 doi:10.20944/preprints201702.0064.v1

Peer-reviewed version available at Philosophies 2017, 2, , 11; doi:10.3390/philosophies2020011

($\Longleftarrow$) Suppose that there is no paradox, but $\lim\limits_{\theta\to 1/2} f(\theta) = \xi > 0$. Then,

$$\infty > E(X) \geq \int_{1/2}^{1/+\epsilon} E(X|\theta)dF(\theta) \geq \lim_{\theta\to 1/2} \frac{2\theta\xi\epsilon'}{2\theta-1} = \infty, 0 < \epsilon' < \epsilon.$$

□

Proof of Theorem 4: Let $\{(U_k, V_k) : k = 1, 2, \dots\}$ be a sequence of random variables such that

$$\begin{cases} (U_k, V_k) = (X_k, 0), & \text{if } X_k \leq n\log_e(\log_2 n) \\ (U_k, V_k) = (0, X_k), & \text{if } X_k > n\log_e(\log_2 n). \end{cases}$$

Therefore,

$$\mathbb{P}\left\{\left|\frac{S_n}{e_n} - 1\right| > \delta\right\} \leq \mathbb{P}\left\{\left|\sum_{k=1}^{n} U_k - e_n\right| > \delta e_n\right\} + \mathbb{P}\left\{\sum_{k=1}^{n} V_k \neq 0\right\}. \tag{7}$$

Defining r as the greatest integer such that $2^r \leq n\log_2(\log_2 n) < 2^{r+1}$ implies that

$$\log_2 n < r \leq \log_2 n + \log_2(\log_2(\log_2 n)).$$

Applying the inequalities above to the second term of (7) we get

$$\begin{aligned} \mathbb{P}\left\{\sum_{k=1}^{n} V_k \neq 0\right\} &\leq n\mathbb{P}\{X_1 > n\log_2(\log_2 n)\} \leq n\mathbb{P}\{X_1 > 2^r\} \leq n(1-\theta)^r \leq \\ &\leq n(1/2)^r = \frac{2n}{2^{r+1}} \leq \frac{2n}{n\log_2(\log_2 n)} = \frac{2}{\log_2(\log_2 n)} \xrightarrow{n\to\infty} 0. \end{aligned}$$

Note that

$$\begin{aligned} E(U_k) &= \int_{1/2}^{1} \sum_{i=1}^{r} 2^i\theta(1-\theta)^{i-1}2d\theta = \sum_{i=1}^{r} 2^{i+1}\int_0^{1/2}(1-p)p^{i-1}dp = \sum_{i=1}^{r}\left[\frac{2}{i} - \frac{1}{i+1}\right] \\ &= \frac{r}{r+1} + \sum_{i=1}^{r}\frac{1}{i}. \end{aligned}$$

Now,

$$\sum_{i=1}^{r}\frac{1}{i} = 1 + \sum_{i=2}^{r}\frac{1}{i} \leq 1 + \log_e r \text{ and } \sum_{i=1}^{r}\frac{1}{i} \geq \int_1^{r+1}\frac{dx}{x} = \log_e(r+1)$$

so that

$$\log_e r \leq \log_e(r+1) \leq \sum_{i=1}^{r}\frac{1}{i} \leq E(U_k) \leq \frac{r}{r+1} + 1 + \log_e r \leq 2 + \log_e r.$$

The moment of order two of U_k is given by

$$\begin{aligned} E(U_k^2) &= \int_{1/2}^{1}\sum_{i=1}^{r} 2^{2i}\theta(1-\theta)^{i-1}2d\theta = \sum_{i=1}^{r} 2^{2i+1}\int_0^{1/2}(1-p)p^{i-1}dp = \\ &= \sum_{i=1}^{r}\left[\frac{2^{i+1}}{i} - \frac{2^i}{i+1}\right] \leq \sum_{i=1}^{r}\frac{2^{i+1}}{i} \leq \sum_{i=1}^{r} 2^{i+1} \leq 2^{r+2} = (4)2^r \leq 4n\log_2(\log_2 n). \end{aligned}$$

Finally,

$$Var\left(\sum_{k=1}^{n} U_k - nE(U_k)\right) \leq nVar(U_k) \leq nE(U_k^2) \leq 4n^2 \log_2(\log_2 n).$$

Taking $e_n = n \log_e(\log_2 n)$ and applying the Chebyshev's inequality to the first term of (7) we get

$$\mathbb{P}\left\{\left|\sum_{k=1}^{n} U_k - n\left[\frac{r}{r+1} + \sum_{i=1}^{r}\frac{1}{i}\right]\right| > \delta e_n\right\} \leq \frac{4n^2 \log_2(\log_2 n)}{\delta^2 n^2 (\log_e(\log_2 n))^2} \leq$$
$$\leq \frac{4}{\delta^2 (\log_e 2) \log_e(\log_2 n)} \xrightarrow{n\to\infty} 0.$$

Again, it is enough to show that $nE(U_k)$ may be approximated by e_n for great values of n. From the inequalities established above,

$$\log_e(\log_2 n) \leq \log_e r \leq E(U_k) \leq 2 + \log_e r \leq 2 + \log_e(\log_2 n + \log_2(\log_2(\log_2 n))) \leq$$
$$\leq 2 + \log_e(\log_2 n + \log_2 n) \leq 2 + \log_e 2 + \log_e(\log_2 n) \leq 3 + \log_e(\log_2 n)$$

and

$$1 = \lim_{n\to\infty} \frac{\log_e(\log_2 n)}{\log_e(\log_2 n)} \leq \lim_{n\to\infty} \frac{E(U_k)}{\log_e(\log_2 n)} \leq \lim_{n\to\infty} \frac{3 + \log_e(\log_2 n)}{\log_e(\log_2 n)} = 1.$$

□

References

1. Bayes, T.; Price, R.; Canton, J.. An essay towards solving a problem in the doctrine of chances. Printer to the Royal Society of London, **1763**.
2. Bernoulli, D.. Exposition of a new theory on the measurement of risk. *Econometrica* **1954**, 23-36.
3. Diaconis, P.; Holmes, S.. A Bayesian peek into Feller volume I. *Sankhy* **2002**, *Series A*, 820-841.
4. Feller, W.. Chapter X: Law of Large Numbers. In *An introduction to probability theory and its applications*; Volume I; New York: John Wiley & Sons, 1968; pp. 243-263.
5. Feller, W.. Note on the law of large numbers and fair games. *Ann. Math. Stat.* **1945**, *16*, *3*, 301-304.
6. Kolmogorov, A. N.. Foundations of the theory of probability. In *Grundbegriffe der Wahrscheinlichkeitsrechnung*; Springer, Berlin, 1933.
7. Lee, P. M.. Chapter 1 Preliminaries. In *Bayesian statistics: an introduction*;2nd ed., John Wiley & Sons, 1997: pp. 1-27
8. Menger, K.. Zeit. Nationaloekonomie. English transl. In *Essays in Mathematical Economics in Honor of Oskar Morgenstern*; ed. Princeton University Press, Princeton, NJ, 1967; Chap. 16.
9. Peixoto, C. M.; Wechsler, S.. O Paradoxo de São Petersburgo. *Boletim da Seção Brasileira da ISBA* **2013**, *6*, *2*, 6-7.
10. Peters, O.. The time resolution of the St Petersburg paradox. *Phil. Trans. R. Soc. A* **2011**, *369*, *1956*, 4913-4931.
11. Samuelson, P. A.. St. Petersburg paradoxes: Defanged, dissected, and historically described. *J. Econ. Lit.* **1977**, *15*, *1*, 24-55.
12. Seidl, C.. The St. Petersburg Paradox at 300.*J Risk Uncertain* **2013**, *46*, *3*, 247-264.
13. Van der Waerden, B. L.. In *Die Werke von Jakob Bernoulli*; Basel, 1975.