St. Petersburg paradox: adoption of decisions on the basis of data mining and development of software in the sphere of business analytics

Paradoja de san Petersburgo: adopción de decisiones sobre la base de minería de datos y desarrollo de software en el ámbito del análisis empresarial

Mikhail Samuilovich Gasparian*, Irina Anatolievna Kiseleva, Valery Alexandrovich Titov, Natalia Alekseevna Sadovnikova

Plekhanov Russian University of Economics. Moscow, Russian Federation.
*Corresponding author E-mail: gasparian.m.s@mail.ru

(recibido/received: 11-mayo-2021; aceptado/accepted: 25-julio-2021)

ABSTRACT

This article is devoted to analysis of models of St. Petersburg paradox, as well as development of software in the sphere of business analysis. This work is based on mathematical models using theories of probability and games as well as expert survey method. It is demonstrated that the St. Petersburg paradox is a mathematical problem of probability theory with artificial conditions. The influence of this problem on economical theory is exemplified by such provisions as the principle of diminishing marginal utility, the use of expected utility as criterion of decision adoption in uncertain environment, as well as foundations of microeconomics of insurance and risk management, theory of games and some approaches to financial simulation. Adoption of decisions on the basis of the St. Petersburg paradox is analyzed. Review of main decisions of the St. Petersburg paradox and their influence for economic theory has confirmed that the St. Petersburg paradox as a mathematical problem can be used as mathematical model upon implementation of financial simulation. Comparative analysis of available BI solutions has confirmed that most of them propose all major functions, and significant differences can be revealed in penetration of expanded functions.

Keywords: St. Petersburg paradox; Adoption of decisions; Risk; Risk assessment; Data mining.

RESUMEN

Este artículo está dedicado al análisis de modelos de la paradoja de San Petersburgo, así como al desarrollo de software en el ámbito del análisis empresarial. Este trabajo se basa en modelos matemáticos que utilizan teorías de probabilidad y juegos, así como el método de encuesta de expertos. Está demostrado que la paradoja de San Petersburgo es un problema matemático de la teoría de la probabilidad con condiciones artificiales. La influencia de este problema en la teoría económica está ejemplificada por disposiciones tales como el principio de utilidad marginal decreciente, el uso de la utilidad esperada como criterio de adopción de decisiones en un entorno incierto, así como los fundamentos de la microeconomía de los seguros y la gestión de riesgos, la teoría de juegos, y algunos enfoques de la simulación financiera. Se analiza la adopción de decisiones sobre la base de la paradoja de San Petersburgo. La revisión de las
principales decisiones de la paradoja de San Petersburgo y su influencia para la teoría económica ha confirmado que la paradoja de San Petersburgo como problema matemático se puede utilizar como modelo matemático al implementar la simulación financiera. El análisis comparativo de las soluciones de BI disponibles ha confirmado que la mayoría de ellas proponen todas las funciones principales, y se pueden revelar diferencias significativas en la penetración de funciones ampliadas.

Palabras claves: Paradoja de San Petersburgo; Adopción de decisiones; Riesgo; Evaluación de riesgos; Procesamiento de datos.

1. INTRODUCTION

Most people are persuaded that insistence and constant striving are the main elements to achieve success. However, the probability theory proves the opposite. This method does not work for gambling and games involving elements of chance. A person should make a choice in favor of this or that outcome, this decision should necessarily be right (Lobanova et al., 2019). What makes the basis of the adopted decisions and what are the criteria for comparison of various options?

Adoption of decisions includes such cognition processes as preliminary comprehension of a target and ways of its implementation, as well as searching for various options (Zueva et al., 2019). Adoption of decision involves knowledge, interests, and worldview of the individual (Orazbayev et al., 2019). Decision is a social notion adopted by one or several persons. A person's character is determined by actions; thus, decision serves as the basis of human self-identification (Dyatlov et al., 2018). The first stage of decision adoption is the occurrence of unfavorable problematic situation, decision selection is the last stage, that is, an action capable to transform this situation (Diev, 2013).

There are numerous paradoxes in probability theory, they occur due to inaccurate axiomatics and various gaps in human knowledge (Nelyubina, 2015). Despite the fact that numerous authors had not realized the economic content of this problem for a long time, they proposed several approaches, which up till now are of key importance for microeconomy theory (Clark, 2002; Kudryavtsev, 2013; Kraitchik, 1942; Samuelson, 1977; Weirich, 1984).

In 1738 Daniel Bernoulli (1738) published his Specimen Theoriae Novae de Mensura Sortis, where the phenomenon of St. Petersburg paradox was described. However, the initial reference to this phenomenon was mentioned in 1713 in the letter from Nicolaus II Bernoulli (Daniel's brother) to Pierre Remond de Montmort, a French mathematician.

Numerous historians and economists believe that the St. Petersburg paradox is the origin of the theory of expected utility (Vatnik, 2008; Weber, 1998; Cowen and High, 1988).

The essence of the St. Petersburg paradox is as follows: a casino offers a game of chance for a single player in which a fair coin is tossed at each stage. The initial stake begins at 2 dollars and is doubled every time heads appears. The first time tails appears, the game ends and the player wins whatever is in the pot. Thus, the player wins 2 dollars if tails appears on the first toss, 4 dollars – if heads appears on the first toss and tails on the second, 8 dollars – if heads appears on the first two tosses and tails on the third, etc.

Thus, what would be a fair price N to pay the casino for profitable entering the game? Profitable: it means that in the case of high number of games, the player wins more than the initial payment.

According to the probability theory, the chance of win is 50% or 0.5 at each coin toss.

The mathematical expectance of win at the first toss is $0.5 \times 1 = 0.5$. At the second toss: $(0.5 \times 0.5) \times 2$. 

1371
The total expected value is the sum of expectances at each game stage, it will equal to $0.5 + 0.5 + 0.5 + ...$ Obviously, the sum of this series is infinitely high amount.

Therefore, the expected monetary prize is infinite amount:

\[ E[x] = \sum_{k=1}^{\infty} x_k \cdot p_k = \sum_{k=1}^{\infty} 2^{k-1} \cdot 2^{-k} = \sum_{k=1}^{\infty} \frac{1}{2} = \infty \] (1)

However, most people refuse to take part in such a game. Why does this happen? The desire to pay infinite money for gambling naturally seemed unreasonable. As a consequence, the problem was referred to as paradox.

The hypothesis of research is as follows:

1. the St. Petersburg paradox can be used as a mathematical model upon implementation of financial simulation;

2. most of BI tools propose all major functions, and significant differences can be revealed in penetration of expanded functions.

### 2. METHODS

In order to prove the formulated hypothesis, the following methods were used:

a) theoretical:

methods of development of mathematical models using probability theory and theory of games;

analysis of the research and analytical references devoted to the problem;

b) empirical: expert survey. The following main questions were proposed to the experts: to select assessment criteria of software for business analytics, to propose 10 BI tools (solutions) for comparison and to compare them according to the selected criteria.

The expert survey was performed online in the Russia language. The experts (32 persons), heads and officers of analytical companies with the working experience in the sphere of analytics of at least 5 years participated in the survey. All participants were warned about the survey target and about future publication of the survey results in generalized form.

### 3. RESULTS

3.1. Review of main solutions to the St. Petersburg paradox and their importance for economic theory

In order to explain the St. Petersburg paradox, D. Bernoulli assumed that people were definitely guided by something else. He assumed that the determination of the value of an item should not be based on the price, but rather on the utility it yielded. Later on, this principle became known as expected utility of gain. Thus, D. Bernoulli replaced the importance of gain with its utility, profit. Then he proposed to apply logarithmic function. After all conversions he obtained the following:
\[
\ln x = \sum_{k=1}^{\infty} \ln(x_k)p_k = \sum_{k=1}^{\infty} \ln(2^{k-1})2^{-k} = \sum_{k=1}^{\infty} \frac{k-1}{2^k} \ln 2 = \ln 2 \sum_{k=1}^{\infty} \frac{k-1}{2^k} = \ln 2
\]

D. Bernoulli concluded that the equivalent sum was finite and equaled to: \( x = 2 < \infty \).

Thus, there exists a certain limit, after which a player does not agree to play.

American economists J. von Neumann and O. Morgenstern, who are commonly considered as the founders of the theory of expected utility, supported the ideas by D. Bernoulli. They continued the analysis and determined that under the conditions of incomplete information, the single reasonable way for an individual to solve this problem would be the selection with maximum expected utility. The player decides whether he participates in the game or not. The selection is made with consideration of not an expected gain but of its real utility. The expected utility of each variant is determined as follows:

\[
U = \sum_{i=1}^{n} p_i x_i; \text{herewith: } \sum_{i=1}^{n} p_i = 1,
\]

where \( U \) is the expected utility, \( p_i \) is the probability of outcome, \( x_i \) is the utility of outcome.

Then, the individual compares the expected utilities of variants and makes selection trying to maximize the expected utility.

This concept was described in details by J. von Neumann and O. Morgenstern (19853) in Theory of Games and Economic Behavior. It follows from the axioms that there always exists such utility function, the mathematical expectance of which corresponds to the preferences of subject. Indeed, this is really true, an individual always knows what is more preferred for him for the moment: acquisition of any wealth or purchase of a lottery ticket.

And what is the player's attitude to risk? Nowadays risk is an inherent part of life of any person. In everyday life risk is a possibility of a certain unfavorable event resulting in various material or moral loss. The analysis demonstrates that at present there are numerous definitions of risk, disclosing its essence in terms of various sciences. In economic sciences risk is also interpreted differently. For instance, in economic theory risk is generally considered as a certain negative product which can be an object of free purchase/sale (von Neumann and Morgenstern, 1953; Shumpeter, 2001). Risks are redistributed among the participants in business activity by means of various financial tools. Herewith, some participants insure themselves against risk by diversification and hedging of their portfolios, others purchase the risk, trying to provide higher profitability.

In financial theory risk is more frequently considered as uncertainty in forecast of transaction outcome, its possible deviation from expected or planned result. Various definitions of financial risk can be found in (Cowen and High, 1988; Al'gin, 1991; Alexander, 1996).

It is generally accepted that existence of risk is related with uncertainty which is homogeneous both in form and in content. At common level these notions are often identified. Analysis demonstrates that such identification can be met in financial publications (Chekanskii and Frolova, 2016; Buffon, 1829).

Different people tend to treat risk differently, too. In economic theory it is common to identify people who are neutral to risk, risk lovers, and people who are against any types of risk.

There are situations, when mathematical expectation upon execution of risky activity could equal in

Let us assume such situation: there is a debtor and the person, who gave this debt, that is, a creditor. The debtor, instead of returning 100 dollars to the creditor at once, offers to toss coin. If the creditor wins, then he will receive twice as much, that is, 200 dollars instead of 100 dollars. Therefore, pure gain of the creditor will be 100 dollars. However, there is a chance equaling to 50% that the creditor loses. In this case the debtor does not repay his debt. Thus, the creditor loses his 100 dollars. The mathematical expectance is: (0.5*100$) + (0.5*(-100$)) = 0. It can be concluded that for the creditor it makes no difference, either to toss coin with the debtor or just to demand his money, since upon nonrisky situation he will return his debt, or to take risk and to receive twice as much.

However, in such situation one person can take the risk hoping to receive more, and for another person it would be preferable not to take any risky actions. In order to explain, how and on what the choice of economic agents depends, it is required to consider the concept of expected utility in addition to the analysis.

It has been demonstrated in practice that in most cases people are not prone to make decisions involving high or at least medium levels of risk. In addition to human mentality, this could be attributed to purely economic reason, namely: the action of law of decreasing ultimate utility (Barguzin.ru, 2013).

Let us assume that an individual has 100 dollars. He is referred to the neutral to risk persons. The individual decides to play roulette and bets 50 dollars on red. If he wins, he will have 150 dollars: 50 dollars remaining in his pocket, since he does not bet them, plus 50$*2: the gain itself. Therefore, the player increases by 50 dollars his initial wealth of 100 dollars. In the case of loss, he will have only 50 dollars, that is, he decreases his initial wealth by 50 dollars. The mathematical expectance of gain in monetary form will be: (0.5*50$) + (0.5*(-50$)) = 0.

However, the ultimate utility decreases, as can be seen in the plot of total utility, thus, in conventional units of utility, the expected utility will be negative: (0.5*(-2)) + (0.5*1) = -1.

![Figure 1. Curve of total utility for an individual who antipathizes risk](image)

In other words, if the player loses, then the loss in conventional units of utility will be higher than his acquisition in the case of gain. Therefore, in terms of utility, the situation seems to be different than in monetary form. And we conclude that in this situation the player cannot maximize the utility of gain and, most likely, he decides not to play.

Therefore, it is obvious that the player would be glad to receive more than he has, though, the loss of his wealth would be significant for him, that is, those 100 dollars possessed by him. In economic theory this
The phenomenon is known as the endowment effect. The endowment effect is comprised of the fact that a person estimates higher his possession rather than that which he can get but does not yet possesses.

Returning to the St. Petersburg paradox, we can state that individuals, refusing to toss coin despite infinite high mathematical expectancy, are governed, according to the hypothesis of D. Bernoulli, mainly by the expected utility of gain. And the ultimate utility of profit decreases with its each increase. Upon decreasing ultimate utility of monetary gain, people will ask for increasing payments in order to compensate their risk in the case of loss.

The Austrian economist Karl Menger in his works proposed another solution to the St. Petersburg paradox (Menger, 1934, 1959, 2005; Shubik, 1967). He considered this phenomenon in terms of economic behavior. Menger formulated certain restrictions for solution to this problem.

Firstly, these are the restrictions of actual world, which do not allow to achieve infinitely high results. It should be mentioned that there exists certain restriction of capital involved in the game. As a consequence, the game cannot last infinitely long. The second restriction is the restriction of the game itself. Therefore, the number of coin tosses is finite. Menger believed that infinite mathematical expectancy of gain was not paradoxical in its essence. In other words, he considers the mathematical expectancy in wider sense than fair price, that is, as an estimation form. According to Menger, the real sense of the paradox is that there exists inconsistence between the mathematical model and the observed behavior. Therefore, the analysis of solutions to the St. Petersburg paradox presented in his work includes their interpretation in terms of their adequacy. He supported implementation of the St. Petersburg paradox and interrelated concepts into economic theory.

D. Bernoulli also considered such notion as decreasing ultimate utility. The level of gain utility depends on initial state, financial status of the player. Thus, a person of high financial status has minimum gain utility in this game. Whereas a person with low level of profit hopes to win a lot of money. Obviously, a poor man and a rich man will treat differently the win of 100 rubles.

Let us imagine such situation. A person, possessing certain property, is offered to play the game with the following rules. There are 100 envelopes in a hat, in one of them there is the card with 1 billion rubles. However, when the person selects another envelope, then he loses absolutely all his property. Let us assume that the player's property is 1 million rubles. In this case, the mathematical expectancy is very high. However, in this case it is obvious that most people would refuse to play this game. Because a person estimates higher the current property than that he could possess.

In addition, it is possible to consider the process of investments in terms of the St. Petersburg paradox. The American economist D. Durand (1957) analyzed this issue in his work Growth stocks and the St. Petersburg paradox in 1957.

Very often it happens that incomes of companies increase faster than the overall income of economy during total economic growth. In this case, the shares of such companies are referred to as the growth shares. This problem was especially urgent in late 1990s–early 2000s. Up to 2000 the growth dynamic of share prices was positive. Such growth caused continuous discussions. Sceptics declared that the growth of inflation expectancy resulted from the growth of share prices. Nevertheless, some people purchased sufficiently high amounts of shares of IT companies. Investors believed that the growth rate of share price would last infinitely. However, this is impossible, and the share price gets higher only if the number of people who intend to buy gets higher. Therefore, the investors, expecting for huge profit, were purchasing shares, thus increasing their price. In late 2000s the share price sharply dropped, resulting in high losses.
for share holders.

3.2. Software in the sphere of business analytics: comparative analysis based on expert survey

At the first stage of the expert survey, the following comparison criteria of BI tools (solutions) in the sphere of business analytics were highlighted, presenting a list of expanded functions in BI tools, which users could expect/required upon searching for software in the sphere of business analytics:

Master Data/Data Quality Management (MD/MDQ);

Self-Service BI: process, where final users developed and deployed their reports and analyses in the frames of approved and maintained architecture and tools portfolio;

Cloud BI: availability of cloud solution;

Mobile BI: opportunity to apply mobile devices;

Real-Time Analytics: decision adoption in real time;

Agile BI Development: agile and scaled architecture including rapid iterative development and commercialization of data warehouse;

Data Warehouse Modernization: capability to design architecture of data warehouses and to use other engineering options, such as memory processing, cloud warehouse, or data warehouse automation;

Data Preparation for Business Users: data cleaning, structuring, and beneficiation for research and expanded analytics;

Data Storytelling: data visualization, infographics, dashboards, data presentations, etc.;

Using External/Open Data: using data from social networks, customers, market, meteorological, geographical, and demographical data, as well as from existing analytical results;

IoT Analytics: data analysis of Internet of things in real time.

Big Data Analytics presents tools for analysis of large data arrays, variety and speed, acquired from internal and external sources, including texts, sensors, geolocation, etc.

At the second stage of the expert survey, the following BI tools in the sphere of business analytics, being still available at present in the market, were highlighted (see Table 1). In addition to the solutions proposed by leading suppliers (for instance, Microsoft, IBM, SAP), the experts included smaller suppliers, who proposed some interesting functions (for instance, Avlino, Sisense). Some them, such as Domo, Sisense, Tableau, Power BI, and Qlik, according to the experts' opinion, should be included into the list of best business analytics software. In general, this list includes simple user-friendly solutions supporting overall spectrum of analytical capabilities of working process, which do not require for significant participation of IT specialists.
Table 1. BI tools selected for comparison

<table>
<thead>
<tr>
<th>No.</th>
<th>BI solution</th>
<th>Web page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Tableau</td>
<td><a href="https://www.tableau.com/">https://www.tableau.com/</a></td>
</tr>
<tr>
<td>2</td>
<td>MicroStrategy</td>
<td><a href="https://www.microstrategy.com/us">https://www.microstrategy.com/us</a></td>
</tr>
<tr>
<td>3</td>
<td>Sisense</td>
<td><a href="https://www.sisense.com/">https://www.sisense.com/</a></td>
</tr>
<tr>
<td>4</td>
<td>Domo</td>
<td><a href="https://www.domo.com/">https://www.domo.com/</a></td>
</tr>
<tr>
<td>5</td>
<td>Power BI</td>
<td><a href="https://powerbi.microsoft.com/en-us">https://powerbi.microsoft.com/en-us</a></td>
</tr>
<tr>
<td>6</td>
<td>Qlik</td>
<td><a href="https://wwwqlik.com/us">https://wwwqlik.com/us</a></td>
</tr>
<tr>
<td>7</td>
<td>SAP Crystal Cloud</td>
<td><a href="https://www.sap.com/products/crystal-bi.html">https://www.sap.com/products/crystal-bi.html</a></td>
</tr>
<tr>
<td>8</td>
<td>IBM Cognos Analytics</td>
<td><a href="https://www.ibm.com/products/cognos-analytics">https://www.ibm.com/products/cognos-analytics</a></td>
</tr>
<tr>
<td>9</td>
<td>Avlino</td>
<td><a href="https://avlino.com/">https://avlino.com/</a></td>
</tr>
</tbody>
</table>

The third stage of the expert survey was devoted to comparative analysis of the selected BI tools in terms of their functional capabilities in the form of tool/function matrix presented in Table 2.

Table 2. BI tools and their capabilities

<table>
<thead>
<tr>
<th>BI tools</th>
<th>Main data /Management of data quality</th>
<th>BI self servicing</th>
<th>Cloud BI</th>
<th>Mobile BI</th>
<th>Real-time analytics</th>
<th>Agile architecture</th>
<th>Modernization of data warehouse</th>
<th>Data preparation for business</th>
<th>Data storytelling</th>
<th>Use of external/open data</th>
<th>Analytics of Internet of things</th>
<th>Big data analytics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tableau</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>MicroStrategy</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Sisense</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Domo</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Power BI</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Qlik</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>SAP Crystal Cloud</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>IBM Cognos Analytics</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Avlino</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

The analysis demonstrated that overwhelming majority of the considered tools proposed data analysis, dashboards, specialized request tools, special reports, key indicators of efficiency and performances, which referred to the main group of functions. In addition, they proposed data quality management, data detection/visualization, BI self service, deployment of cloud BI, mobile BI, real time analysis and big data analysis.

On the basis of the obtained results, it is possible to conclude that Domo, Sisense, Tableau, Power BI, and Qlik are among the best variants of data mining.
4. CONCLUSION

One of the most important results in solution to the St. Petersburg paradox is introduction of the function of expected utility, on the basis of which the theory of decreasing utility was created.

The influence of this problem on economic theory can be traced, for instance, for the principle of decreased ultimate utility, the use of expected utility as a criterion of decision adoption in uncertain environment, as well as for the foundations of microeconomy of insurance and risk management.

It particular, attempts have been made to present the paradox as a predecessor of the average–risk concept actively applied in the portfolio theory, as well as a precursor of econophysics: one of the popular alternatives to modern financial theory.

Therefore, the first part of the hypothesis has been confirmed that the mathematical problem of St. Petersburg paradox can be used as mathematical model upon implementation of financial simulation.

In order to get the most for a company of its data, it is required to determine a BI solution, which is the most suitable for the company's demands. It is necessary to determine, which functions should be solved by the solution, whether it is easily adapted and user friendly, which support can be provided by it, how good it solves the problem of security, and, finally, how much it costs.

Therefore, the comparative analysis of the available BI solutions has confirmed the second part of the hypothesis that most of the BI solutions propose all main functions, and significant differences can be revealed in penetration of expanded functions.

Further studies can be aimed at expansion of the selected criteria for analysis of software tools of business analytics and the number of analyzed BI solutions.

ACKNOWLEDGMENTS

This research was performed in the framework of the state task in the field of scientific activity of the Ministry of Science and Higher Education of the Russian Federation, project "Development of the methodology and a software platform for the construction of digital twins, intellectual analysis and forecast of complex economic systems", grant no. FSSW-2020-0008.

REFERENCES


Chekanskii, A. N. and Frolova, N. A. (2016). Mikroekonomika, promezhutochnyi uroven'


