SIMULATION MODELING OF INVESTMENT PROJECTS IN THE SERVICE SECTOR

Viktor Dzis¹, Olena Dyachynska²

Abstract. Scientific article is devoted to researching the possibility of developing a simulation model of an investment project for a travel agency. In recent years, the tourism industry in the world is experiencing rapid development. The development of the tourism industry requires new investments. The investor must assess the degree of risks, the likelihood of profit or loss. It is noted that the most important advantage of simulation modeling is that it makes it possible to study economic systems at the design stage. Due to this, simulation models can be used as a universal tool in making appropriate decisions under conditions of uncertainty and taking into account those factors that are difficult to predict and take into account, which is why simulation is so often used in the development of investment projects. When assessing the risks of investment projects, the collection of information requires significant financial expenses, is quite time-consuming, and sometimes impossible. The subject of the study is a set of the theoretical, methodological and organizational problems related to the management decisions regarding the investment of the projects in the field of tourism. The methodology is based on the stochastic simulation modeling of economic processes. The aim of the study is to develop a simulation model for an investment project in the tourism sector, which will allow the investor to appreciate the degree of the risk and the likelihood of return on the investment. The article reveals the features of computer simulation in the MathCAD system. It is proved that it is convenient to develop simulation models of investment projects in the MathCAD system. It has a powerful mathematical support, remains one of the systems in which the description of the solution of mathematical problems is given with the use of conventional mathematical formulas and symbols and does not require special training in programming. The simulation model of the investment project of the travel agency was developed in the MathCAD system. Statistical processing of the results of the experiments with the model has been carried out. The necessity of the detailed study of critical intervals of the histogram at the transition from loss-making to profitable NPV values has been substantiated. Based on the results of simulation experiments with the model, the errors of the model were estimated, the law of NPV random variables distribution was established. The developed simulation model allows the investor to estimate the risk factor, the probability of profit or loss from the investment, to assess the possible uncertainty of the results of their own decision to invest in the project. Conclusion: the simulation model of the investment project built and investigated in MathCAD system allows investor to estimate the risk factor, probability of profit or loss for choosing the optimal pricing policy and optimization of economic strategy of tourist agency.

Key words: simulation model, MathCAD, investments, tourism industry, risk factor, probability, histogram, critical intervals, errors.

JEL Classification: C53, C63

1. Introduction

Statement of the Problem. Simulation modeling is a method of research in which the system under study is replaced by a mathematical model that describes the real system with sufficient accuracy, and experiments are conducted with it in order to obtain information about this system. Experiments with a model are called simulation. Simulation is a numerical method of conducting computer experiments with mathematical models that describe the behavior of
complex systems over long periods of time. Simulation changes the parameters of the model without resorting to experiments on the real object. Thus, simulation consists of developing a robust mathematical model of the system and performing numerous experiments on that model.

The information about the simulation model has a logical and mathematical nature and is presented as a set of algorithms describing the process of system functioning. Thus, the simulation model is more its software implementation on the computer, and simulation modeling is reduced to conducting experiments (simulations) with the model by repeatedly running the program with some set of data.

When you investigate a system, you can act on its inputs and observe its outputs. Of course, you can change all the input variables during the simulation. Thus, the simulation model can function in the same way as the real system under study. During simulation, it is necessary to ensure the software implementation of algorithmic descriptions, to organize, plan and perform computational experiments with mathematical models on the computer. Simulation requires a lot of computation and multiple runs of the program, so it is done with the help of computers. Simulation models are quite convenient to develop in the application package: MATLAB (Simulink application) (Zhitenko, 2011), Scilab (Xcos application) (Dubovoy, 2018), AnyLogic (Taranenko, 2015), MathCAD (Gavrilenko, 2007). MS EXCEL is often used to build simulation models of economic and financial systems (Gavrilenko, 2007). However, many people prefer to write their simulation programs in high-level languages such as Python, C++ a few runs of a program with some data set.

Simulation statistical modeling plays an important role in modeling. Simulation statistical models are used to simulate probabilistic systems and processes under uncertainty. Stochastic modeling is an extremely important area of simulation modeling and is used to study the dynamics of nondeterministic systems, i.e., systems whose exact characteristics are unknown a priori. Economic models are a typical example of such a system.

Simulation statistical modeling is a numerical method of conducting computational experiments on a computer with mathematical models that simulate the behavior of real objects, processes and systems over a certain period of time. Most economic processes and systems are probabilistic, so statistical simulation models are used to study them.

The essence of statistical simulation:
- numerical values of stochastic parameters for each experiment are taken as random numbers from confidence ranges;
- confidence range – the range limited by the optimistic and pessimistic value of the parameter;
- random numbers are generated using a random number generator (often using a uniform distribution);
- the obtained results are analyzed as statistical data.

The most important advantage of simulation modeling is that it makes it possible to study economic systems at the design stage. Because of this, simulation models can be used as a universal tool in making appropriate decisions under conditions of uncertainty and taking into account those factors that are difficult to predict, so simulation is often used in the development of investment projects. When assessing the risk of investment projects, the collection of information requires significant financial expenses, is quite labor-intensive, and sometimes impossible. To adequately assess risk, it is necessary to have a sufficiently large set of data. In cases where it is difficult or impossible to obtain physical data, they are replaced by values obtained in the course of a simulation experiment. The results of the simulation allow the investor to assess the uncertainty of the decision to invest in the project, to determine the feasibility of this decision and the degree of risk. Thus, building simulation models and improving the reliability of forecasting in the service sector is relevant.

2. Analysis of recent research and publications

The work of many scientists is devoted to the development and study of simulation models of economic systems.

Gavrilenko V. V., Shumeiko O. A. in their work using MS EXCEL and MATHCAD revealed the essence of simulation modelling of the investment risks (Gavrilenko, Shumeiko, 2007). The problems related to the use of new information technologies in the management analysis, namely simulation modelling were considered by Kulinch M. B. (Kulinch, 2009).

Yanekova J., Fabiarova J., Kadarova J. presented the methodology of choosing the optimal investment option using the Monte Carlo modelling and OptQuest optimization (Janekova, 2021).

The method of construction and research of simulation models of technical systems (Dubovoy, 2018; Dubovoy, 2017; Yeliseyev, 2021; Zhitenko, 2011), is developed in detail. In the field of medicine, it is fundamental work on simulation model of the spread and reduction of the impact of the COVID-19 virus (Christine, 2020).

3. Investment project simulation model

Consider an investment project for the development of a travel agency. In the model we put the conditional parameters (Table 1):

Evaluate the effectiveness of the investment project and the risk of its implementation by simulation modeling.
Investing in any project involves risk. The results of an investment project are ambiguous. This is due to the fact that the criteria for assessing the effectiveness of investments depend on cash flows, which in turn may be influenced by random factors in future periods: to increase or decrease. Risk and uncertainty factors should be taken into account when assessing the effectiveness of the project, as under different possible conditions of the project the results will be different.

The main criteria of the investment project:
- cash flow;
- the net present value of the project is Net Present Value (NPV).

Cash-flow (CF) is used in the analysis of securities, development of investment projects and reporting of enterprises. Net cash flow from operating activities is called operational cash flow, net cash flow from investing activities is called investment cash flow. The difference between inflows and outflows in financial activities (in the narrow sense) – Cash-flow from financial activities.

Net Cash-flow (cash flow) of the enterprise for the reporting period is calculated as the arithmetic sum of Cash-flow from all activities. It will be equal to the growth of cash and cash equivalents for the specified period (Gavrilenko, 2007; Janekova, 2021).

\[ CF = (Q \cdot (P - V) - PV - AM) \cdot (1 - PD) + AM, \]  

where \( Q \) – sales, \( P \) – the cost of one ticket, \( PV \) – variable costs per tourist, \( AM \) – depreciation, \( PD \) – taxes.

The Net Present Value (NPV) of a project. This is the best known and most commonly used criterion. NPV is the discounted value of a project (the present value of the income or benefits of the investment). (Gavrilenko, 2007; Janekova, 2021). The net present value of a project is the difference between the amount of cash flow discounted at a reasonable rate of return and the amount of investment. To calculate the NPV of a project, you must determine the discount rate, use it to discount the cost and benefit streams, and add up the discounted benefits and costs (minus sign costs). NPV calculations are cumbersome, so they are done on a computer using built-in functions:
- MS EXCEL function of PS \((r, \tau, CF)\);
- MathCAD function pv \((r, \tau, CF)\),

where \( r \) – is the discount rate (rate) for the period, \( \tau \) – is the total number of payment periods (in our case 3 periods, the number of periods is equal to the number of years of implementation), \( CF \) – is the payment made in each period (enter as a negative value).

\[ NPV = PV(r, \tau, CF) - IN. \]  

For a preliminary analysis of NPV value distribution series, absolute value histograms are most often used and the absolute value of frequencies is plotted along the ordinate axis rather than the frequency density. Histograms for modeling investment projects should be constructed so that individual intervals contain only positive (profitable) or negative (unprofitable) NPV values. If one of the histogram intervals contains both negative and positive NPV values, the results of the simulation experiment can be significantly distorted. It is also necessary to examine in detail the critical intervals of the histogram at the transition from loss-making to profitable NPV values. They may contain a significant number of frequencies with insignificant gains or losses for the project. Errors of average NPV values in these intervals can reach quite large values of 100-500%, so average values in these intervals are not very reliable. In some cases, the results of simulation modeling it is possible to establish a density function of the distribution of the random variable under study, then the probability of the event is estimated by the classical relations of mathematical statistics and probability theory.

The errors of the simulation model depend on the type of mathematical dependence of the model, the number of variables, and the number of model runs. The relative error of the number of values falling within one histogram interval is estimated by the ratio (Dubovoy, 2018):

<table>
<thead>
<tr>
<th>Model parameters</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable costs per hotel voucher (V), UAH</td>
<td>3000</td>
<td>4600</td>
</tr>
<tr>
<td>Number of sold hotel vouchers (Q), units</td>
<td>3000</td>
<td>6000</td>
</tr>
<tr>
<td>The cost of one hotel voucher (P), UAH</td>
<td>3600</td>
<td>5500</td>
</tr>
<tr>
<td>Initial investment (INV), UAH</td>
<td>1000000</td>
<td></td>
</tr>
<tr>
<td>Fixed costs (PV), UAH</td>
<td>500000</td>
<td></td>
</tr>
<tr>
<td>Depreciation (AM), UAH</td>
<td>200000</td>
<td></td>
</tr>
<tr>
<td>Tax (POD), %</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>Project implementation period (( \tau )), years</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Discount rate (( r )), %</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

Source: developed by the authors
\( \varepsilon = \frac{1 - P_i}{N} \cdot 100\% \), \( \sigma \) is the probability of the value falling into a separate interval of the histogram expressed in fractions of units \( P_i = \frac{n_i}{N} \), \( n_i \) – frequency; \( N \) – sample size (number of runs of the program on the computer).

The relative error of the arithmetic mean \( \varepsilon \) and the absolute root-mean-square error \( \sigma \) of the mean NPV value on a separate histogram interval are estimated by coefficients:

\[
\varepsilon = \frac{1}{n \sum_{i=1}^{n} \frac{N_{P_i} - NPV_i}{NPV}} \cdot 100\% \),
\]

\[
\sigma = \sqrt{\frac{1}{n(n-1) \sum_{i=1}^{n} (N_{P_i} - NPV)^2}} ,
\]

where \( n \) – frequency; \( NPV \) – the mean value of NPV in a single interval.

The model error can be estimated only on the basis of simulation results. For this purpose, preliminary simulation experiments are performed with the model. The number of runs of the program is chosen so that the average value of the frequency error in the interval does not exceed a predetermined level, for example 2.5%. Note that the number of histogram intervals increases with the number of program runs \( N \).

4. Construction of a simulation model

We will build the simulation model in the MathCAD system (Listing 1), because it has powerful mathematical software, remains one of the systems in which the description of solving mathematical problems is given using conventional mathematical formulas and symbols and does not require special training in programming (Dzis, Diachynska, 2020; Dzis, Levchuk, Diachynska, 2020; Dzis, 2019).

Vectors of random values of the variable parameters of the model (the cost of servicing one tourist, the number of tourists, the cost of the tour) will be set according to the uniform distribution law.

Dependencies for NPV estimation (formulas 1, 2) are non-linear functions, which nullifies the results of correlation analysis, so for a detailed analysis of the results of the simulation experiment we build summary statistical tables (Tables 2, 3).

Testing the results of simulation experiments according to the Pearson agreement criterion \( \chi^2 \)
Simulation model "Investment in the tourism business"

\[ \text{ORIGIN} := 1 \]

\[ \text{INV} := 1000000 \quad \text{Investments, UAH} \]
\[ \text{PV} := 500000 \quad \text{Fixed costs, UAH} \]

\[ \text{AM} := 200000 \quad \text{Depreciation, UAH} \]
\[ r := 0.1 \quad \text{Discount rate} \]

\[ \text{POD} := 0.6 \quad \text{Taxes} \]
\[ \tau := 3 \quad \text{Period, years} \]
\[ N := 2000 \quad \text{Number of runs of program} \]

\[ V_{\text{min}} := 3000 \quad V_{\text{max}} := 4600 \quad \text{Service costs for 1 tourist, UAH} \]

\[ Z_{\text{min}} := 3000 \quad Z_{\text{max}} := 6000 \quad \text{Number of tourists} \]

\[ P_{\text{min}} := 3600 \quad P_{\text{max}} := 5500 \quad \text{The cost of one hotel voucher, UAH} \]

\[ V := \text{runif}(N, V_{\text{min}}, V_{\text{max}}) \quad \text{Random values vector of maintenance costs for 1 tourist (uniform distribution law), UAH} \]

\[ Z := \text{runif}(N, Z_{\text{min}}, Z_{\text{max}}) \quad \text{Random values vector of the number of tourists (uniform distribution law), UAH} \]

\[ P := \text{runif}(N, P_{\text{min}}, P_{\text{max}}) \quad \text{Random values vector of the cost of vouchers (uniform distribution law), UAH} \]

**Cash-flow and NPV project evaluation**

\[ i := 1 \ldots N \]

\[ \text{CF}_i := \left[ Z_i \left( P_i - V_i \right) - \text{PV} - \text{AM} \right] (1 - \text{POD}) + \text{AM} \quad \text{Cash-flow \text{i}-th run of the program, UAH} \]

\[ \text{NPV}_i = \text{pv}(r, \tau, -\text{CF}_i) - \text{INV} \quad \text{NPV \text{i}-th run of the program, UAH} \]

**NPV value (UAH)**
\[ S_1 = \sum_{i=1}^{N} \text{if} \{NPV_i < 0, NPV_i, 0 \} = -1061601508 \]  \(\leftarrow\) The total amount of losses for all runs of the program, UAH

\[ S_2 = \sum_{i=1}^{N} \text{if} \{NPV_i \geq 0, NPV_i, 0 \} = 5498714565 \]  \(\leftarrow\) The total amount of profits for all program runs, UAH

\[ \text{KRZK} = \frac{|S_1|}{|S_1| + S_2} = 0.16 \]  \(\leftarrow\) Risk factor

\[ \hat{z}(u) := \begin{cases} n & \text{if } n \leq 0 \\ z_1 & \text{if } 1 \leq n \\ z_2 & \text{otherwise} \end{cases} \]  \(\leftarrow\) Function for calculating the number of profitable \(N_p\) and unprofitable \(N_z\) consequences of program runs

\[ \text{NPV}_{\text{sort}} := \text{sort}(NPV) \]  \(\leftarrow\) Arrange the \(NPV\) matrix as the elements grow.

\[ j_1 := 1 \ldots N_z \quad \text{NPV}_{j_1} := \text{NPV}_{j_1} \]  \(\leftarrow\) Formation of matrices of values of unprofitable and profitable consequences of model runs.

\[ j_2 := N_z + 1 \ldots N \quad \text{NPV}_{j_2} := \text{NPV}_{j_2} \]

\[ W_p := \frac{N_p}{N} = 73.05 \]  \(\leftarrow\) Probability of return on investment, %

\[ W_z := \frac{N_z}{N} = 26.95 \]  \(\leftarrow\) Probability of loss on investment, %

\[ \text{NPV}_{\text{min}} := \min(NPV) = -6848980 \quad \text{NPV}_{\text{max}} := \max(NPV) = 12388568 \]

\[ a := \text{mean}(NPV) = 2218556.5 \]  \(\leftarrow\) The average value of \(NPV\), UAH

\[ \sigma = \text{Stddev}(NPV) = 3451187.1 \]  \(\leftarrow\) \(NPV\) standard deviation

\[ \text{Var}(NPV) = 11910592397194.8 \]  \(\leftarrow\) \(NPV\) dispersion

\[ \text{As} := \text{skew}(NPV) = 0.175 \]  \(\leftarrow\) \(NPV\) asymmetry

\[ \text{Ek} := \text{kurt}(NPV) = -0.21 \]  \(\leftarrow\) Excess \(NPV\)

\[ k := \text{round}(1 + \log(N, 2)) = 12 \]  \(\leftarrow\) The number of intervals of the histogram (Sturgess formula)

\[ Y_r := \left( \text{NPV}_{\text{min}}, 10^{-6}, -4.5, -3, -1.5, 0, 1.5, 3, 4.5, 6, 7.5, 9, 10.5, \text{NPV}_{\text{max}}, 10^{-6} \right) 10^6 \]

\[ Y := Y_r^T \]  \(\leftarrow\) Boundaries of histogram intervals
(Continuation of the listing 1)

\[ HIST = \text{histogram}(Y, \text{NPVS}) \quad \leftarrow \text{Matrix for histogram construction} \]

\[ \text{NP} = HIST^{(1)} \quad \leftarrow \text{The average value of NPV in the interval, UAH} \]

\[ i := 1 \ldots \text{rows}(HIST) \]

\[ n := HIST^{(2)} \quad \leftarrow \text{Frequency values} \]

\[ n^T = (24 \quad 58 \quad 158 \quad 249 \quad 309 \quad 350 \quad 297 \quad 217 \quad 161 \quad 65 \quad 40 \quad 22) \]

\[ h_i := \frac{n_i}{Y_{i+1} - Y_i} \quad \leftarrow \text{Frequency density in the interval} \]

\[ f(\text{NPVS}) := \text{dnorm(\text{NPVS}, \mu, \sigma)} \quad \leftarrow \text{The density distribution function is random NPV values} \]

---

**Histogram of absolute frequencies for NPV**

![Histogram of absolute frequencies for NPV](image)

**Histogram and distribution density function for NPV**

![Histogram and distribution density function for NPV](image)
Three Seas Economic Journal

Vol. 3 No. 1, 2022

(Continuation of the listing 1)

\[ n_5 = 309 \quad j_5 := 1..n_5 \quad PP_j := NFV_j \]

\[ k_5 := \text{trunc}(1 + \log(n_5, 2)) = 9 \]

\[ \text{HIST}_5 = \text{histogram}(k_5, PP) \quad \text{NP}_5 := \text{HIST}_5^{(1)} \]

\[ m_5 := \text{HIST}_5^{(2)} \]

The histogram of absolute frequencies for NPV (5th interval)

\[ W_1 = 100 \left[ \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right]_{1000000}^{2000000} = 11.3 \]

\[ Q(X_1, X_2, Y) = \begin{cases} 
    n \leftarrow 0 \\
    \text{for } x \in Y \\
    n \leftarrow n + 1 \text{ if } x \geq X_1 \land x \leq X_2 \\
    n
\end{cases} \]

\[ n_x := Q(1000000, 2000000, \text{NPV}) = 236 \]

\[ W_2 = \frac{n_x}{N} \cdot 100 = 11.8 \]

Source: developed by the authors based on [1; 2; 4; 6; 7]
shows that the random values of NPV are distributed according to the normal distribution law (Listing 1, histogram of the distribution of absolute NPV frequencies, Table 2, 3):

\[ f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}, \tag{8} \]

where \( a = 2218556.5 \) is the sample mean NPV, \( \sigma = 3451187.1 \) is the standard deviation of NPV.

The function for the distribution of the NPV random variable density (formula 8), which allows us to estimate the probability of profit or loss from the investment, is obtained. For example, the probability that the profit will be in the range from \( X_1 = 1000000 \) UAH up to \( X_2 = 2000000 \) UAH is estimated by the formula:

\[ W_1 = 100\% \times \int_{x_1}^{x_2} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}} \, dx. \tag{9} \]

If the type of distribution function of random variables NPV can not be established, the probability \( W \) of profit or loss is estimated by the usual relations of probability theory

\[ W_2 = \frac{n}{N} 100\%. \tag{10} \]

The probabilities of return on investment estimated by two methods (formulas 9, 10) are consistent with each other within the errors of the model \( W_1 = 11.3\% \), \( W_2 = 11.8 \% \) (Listing 1).

5. Conclusions and suggestions

The developed simulation model is based on a scenario approach, which in turn allows you to consider and analyze different options for the development of a travel agency.

Cumbersome mathematical formulas are embedded in the simulation model. The model requires a significant number of simulation experiments on the computer, so it is advisable to build it in MathCAD, which has all the necessary mathematical tools and powerful graphical capabilities to visualize the model (Dzis, Levchuk, Diachynska, 2020).

To generate vectors of random values of model variables (cost per tourist, number of tourists, cost of the tour) it is recommended to use the uniform law of distribution.

When constructing a histogram of the distribution of random NPV values, you should assume that the individual intervals of the histogram should contain only positive (profitable) or negative (unprofitable) NPV values.

Particular attention should be paid to model accuracy. The accuracy of the model depends on the number of model runs on the computer and the type of mathematical model underlying it. It is recommended to choose such a number of program runs that the average value of frequency error in the interval does not exceed a specified level, for example, 2.5%.

In some cases, it is possible to establish the law of probability distribution of random NPV values, which allows the use of reliable statistical methods for analyzing the NPV distribution.

On the basis of the simulation experiments the law of probability density distribution of random NPV values was established.

The developed simulation model allows the investor to estimate the risk factor, the probability of profit or loss, to assess the possible uncertainty of the results of his own decision to invest in the project.

The model, based on simulation experiments, allows investors and entrepreneurs to choose the optimal pricing policy and optimize the economic strategy of the travel agency.

The scientific novelty of our results lies in the improvement of existing and introduction of new approaches to simulation modeling of simulation processes in the predictive activity of tourist infrastructure.

References:


