Demand Forecast Using Markov Chain

Lethícia de Almeida Garcia Silva[1], João Vitor Machado Mariquito [2], Tiago Bittencourt Nazaré[3]

[1] *Student of Industrial Engineering, Faculdades Integradas de Cataguases-FIC/UNIS*
[2] *Student of Industrial Engineering, Faculdades Integradas de Cataguases-FIC/UNIS*
[3] *Master of Engineering Systems Management, Universidade Católica de Petrópolis*
Cataguases/MG, Brazil

Abstract — *Given the need for companies to become increasingly competitive, it is important that they are aware of their customers' profiles and able to serve them to the best of their ability. Thus, forecasting the need to purchase items to be purchased by customers is of paramount importance. The present study is justified in applying the Markov Chain tool to improve the activities of the generalist wholesale company, in order to verify the demand forecast. In order to achieve this goal, Excel software, Markov Chain tool applied by computer simulation using Scilab and evaluated two distinct customer cases. After application, it was possible to identify the demand forecasts for the studied customers and suggest to the company the application of the action in each case.*

Keywords: *Markov chain; Demand Forecasting; Scilab; Computer Simulation; Operational Research.*

I. INTRODUCTION

Technological advancement in the globalized marketplace has led companies to become more competitive and better observe consumer behavior by making them available in the right quantity, at the right place and at the right time. For this purpose, a forecast demand simulation model becomes indispensable. In order for new companies to gain their market share, management effectiveness makes a difference, so planning and controlling resources to meet customer desires, aiming for greater profits and minimal losses, is extremely important.

This work is justified in using the operational research tool, Markov Chain, to improve the company's activities. According to the Brazilian Association of Production Engineering (ABEPRO, 2008), operational research aims to assist in the decision making of real problems through mathematical models, seeking to introduce through decision making processes objectivity and rationality, not allowing subjective data are ignored.

This article aims to analyze the sales history over a period of four months, to identify which products are most profitable for the company in the studied area, to verify how often these products are purchased by the studied customers and based on this information. forecasting the purchase demand of these customers through computer simulation using Markov Chain.

II. LITERATURE REVISION

A. Demand Forecasting

According to Miyata et al. (2010) demand refers to the return of the public as offered by the company. Because of this, good planning requires the use of demand forecasting as a tool for better resource allocation, thus being efficient in delivering the quantity required by the customer at the lowest possible cost.

According to Lustosa et al. (2008) Demand is the propensity of customers to consume goods or services offered by a company or organization. Such demand is influenced by several factors ranging from macroeconomic conditions to operational issues.

According to Lima and Oliveira (2016) forecasting is a science, since it uses historical information and interactions of variables, and through mathematical and statistical methods results in data to be analyzed. These analyzes are forecast checks for the given time. In situations where the results of the analysis are outside pre-established limits, it will be necessary that they be redone or reviewed in order to understand the results obtained. It is well known that predictions are not exact results, they are made so as to have an idea of what might happen in the future if the pattern of situations already occurred is repeated or cause and effect relationships are maintained.

B. Simulation Based Forecasting

The demand forecast can be estimated through a simulation. Taha (2008) defines simulation-based forecasting as a means of reproducing actual behavior through estimates of performance measures, that is, a virtual representation of the random conduct of a situation in order to estimate its performance measures.

Moreira (2010) says that simulation involves the construction of a pattern that comes close to reality. This pattern will be analyzed several times in order to manipulate, control, and understand the possibility of occurrence of a given result. In general, simulation often requires the use of computers and software.

According to Andrade (2009) the use of simulation can be used as an initial test for the development of new decision principles. To formulate the problem, it is necessary to collect data that is sufficient, qualitatively guaranteed and meaningful for decision making. After establishing a model, it is necessary to know the efficiency regarding the objectives of the simulation. The tests performed need to involve the

data obtained in order to be able to verify its occurrence.

C. Markov Chain

In 1907 the Russian mathematician Andrei Andreyevich Markov studied a major type of process, in which only the result obtained from a given current experience can affect the result to be obtained in the next experience, so previous experiences do not influence future ones. experiences. This so-called property is titled Markov Property featuring a Markov Chain (GOLMAKANI et al., 2014).

According to Horn (2016) Markov chains are stochastic processes that have discrete states. In general, the parameter used is time, which can be discrete or continuous. Markov chains are characterized by the fact that their future state depends only on their current state and is not altered by past states. The Markov Chain, being one of the oldest discrete state model description systems, has become the most direct way of representing such states. Markov Chains are Markovian stochastic processes. A Markovian process is considered a Markov Chain, when random variables Xt constitute a particular dependency relationship with time, such variables being defined in a discrete state space.

For Pinheiro (2013) a stochastic process is defined as a collection of random variables (X (t)) attached by a parameter t belonging to a set T. Such a set is often considered to be the set of nonnegative integers and X (t) is the representation of a measurable characteristic of interest in time t.

D. Computational Simulation in Operational Research

According to Pedgen (1990) simulation is the means of designing a computational model of a real system and conducting experiments with this model in order to understand behavior and evaluate strategies for operation. Thus, computer simulation encompasses the entire study process, from model construction to experiment, allowing an analysis of a system without interference on it.

Simulation has been increasingly used as a tool that enables analysts from various industries, such as administrators, biologists, engineers, to verify problems and develop solutions in order to achieve the expected objectives (FREITAS, 2008).

Modeling and simulation are based on operational research, assisting the decision making process. Operational research uses simulation to solve problems, building models that enable a better understanding of the analyzed environment, identifying problems, developing strategies and improvements for the system in question (MORABITO; PUREZA, 2011). According to Law (2014) simulation is the most applied technique in the operational research area, as it has several applications as a tool.

III METHODOLOGY

For the development of this scientific research we used a literature review on the themes "demand forecasting", "simulation" and "Markov Chain", serving as the research base the Google Scholar tool. The target company of the case study was from the generalist wholesale branch, being located in a city in the interior of the state of Minas Gerais, having been founded in July 2017. This company operates in the distribution of food items, beverages, personal hygiene items, cleaning products, disposables and packaging in the region of the forest zone.

For the development of the research, it was necessary to perform a documentary analysis of the sales flow from June to September 2019. The research focused on customers of the food service business, which is being developed by the company. Currently the company has 10 clients in this branch. The products analyzed were the top 10 sales to these customers during the last four months, during which time the company operates in this field. The choice of the quantity of products was based on the software to be used in the simulation, since it works with square matrices. As a source of the information used, we made use of the sales management report provided by the ERP software used by this company. After identifying this information, Excel 2010 software was used to analyze which customers bought which products, so that the frequency of product purchases per customer was identified. The information obtained was grouped in a 10x10 square matrix that was inserted in the Scilab 5.5.2 software, responsible for the computational simulations of the Markov Chain. Scilab analyzed the frequencies identified through Excel and then computed simulations of the expected demand for the second order of these customers.

IV. RESULTS AND DISCUSSION

Through the ERP used by the company, we extracted the product sales information from June/2019 to September/2019, and this information was entered into Excel to calculate the item purchase frequency per customer, using the formula “CONTALUES” considering the number of orders placed per customer. As shown below in the case of product 1, where it is observed that it was purchased by customers 2, 4 and 5, with the calculated purchase frequency being 78.26%, 87.50% and 75.00% respectively.

Product 1																							
Client	REQUEST																						Frequency
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
Client 1																								0,00%
Client 2	x			x	x	x	x	x	x		x	x		x	x	x	x	x	x	x	x		x	78,26%
Client 3																								0,00%
Client 4	x	x	x	x	x	x	x																	87,50%
Client 5	x		x	x		x	x	x																75,00%
Client 6																								0,00%
Client 7																								0,00%
Client 8																								0,00%
Client 9																								0,00%
Client 10																								0,00%

Fig. 1 -- Frequency of purchase - Product 1

Subsequently, the information obtained from the frequency of purchase of the analyzed products was grouped in a matrix, shown in figure 2:

	Product 1	Product 2	Product 3	Product 4	Product 5	Product 6	Product 7	Product 8	Product 9	Product 10
Client 1	0,00%	0,00%	61,90%	38,10%	0,00%	14,29%	14,29%	0,00%	0,00%	0,00%
Client 2	78,26%	17,39%	0,00%	0,00%	0,00%	21,74%	0,00%	17,39%	0,00%	0,00%
Client 3	0,00%	0,00%	0,00%	71,43%	14,29%	0,00%	0,00%	0,00%	0,00%	14,29%
Client 4	87,50%	62,50%	0,00%	0,00%	62,50%	0,00%	0,00%	87,50%	0,00%	0,00%
Client 5	75,00%	75,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	50,00%	87,50%
Client 6	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	44,44%	0,00%	33,33%	0,00%
Client 7	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Client 8	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Client 9	0,00%	0,00%	0,00%	0,00%	0,00%	50,00%	0,00%	12,50%	0,00%	0,00%
Client 10	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	45,45%	0,00%

Fig. 2 -- Purchase Frequencies

Later this information was used in Scilab to assemble the matrix to be used as a basis for the development of the Markov Chain study, as can be seen in figure 3:

```
Scilab 5.5.2 Console

Execução de iniciação:
  carregando o ambiente inicial

-->a=[0.00 0.00 0.62 0.38 0.00 0.14 0.14 0.00 0.00 0.00; 0.78 0.17 0.00 0.00 0.00
 a  =

    0.     0.     0.62   0.38   0.     0.14   0.14   0.     0.     0.
    0.78   0.17   0.     0.     0.     0.22   0.     0.17   0.     0.
    0.     0.     0.     0.71   0.14   0.     0.     0.     0.     0.14
    0.88   0.63   0.     0.     0.63   0.     0.     0.88   0.     0.
    0.75   0.75   0.     0.     0.     0.     0.     0.     0.5    0.88
    0.     0.     0.     0.     0.     0.     0.44   0.     0.33   0.
    0.     0.     0.     0.     0.     0.     0.     0.     0.     0.
    0.     0.     0.     0.     0.     0.     0.     0.     0.     0.
    0.     0.     0.     0.     0.     0.5    0.     0.13   0.     0.
    0.     0.     0.     0.     0.     0.     0.     0.     0.45   0.
```

Fig. 3 -- Markov Chain Matrix - Initial State

For the analysis of this information, a graph representing the initial state data concentration was generated, presented in figure 4:

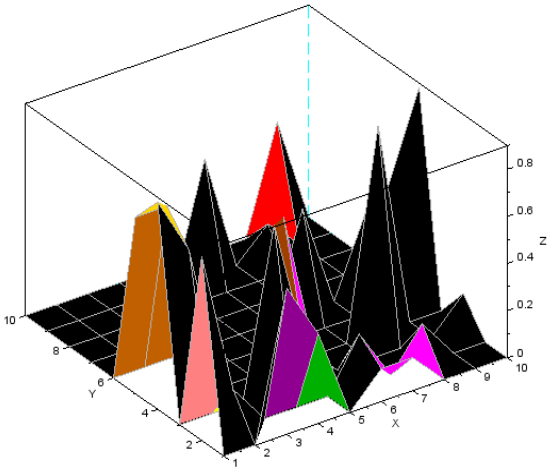

Fig. 4 -- Data Concentration - Initial State

In figure 4, the X axis represents the studied products, the Y axis represents the analyzed customers and the Z axis represents the data concentration considering items ordered by customer. Based on the data presented in figures 3 and 4, it is possible to identify the recurrence of the purchase of each item by the studied customers, and it can be observed that the studied customers have different purchasing profiles, some buy more products and a larger quantity of products. different items, others buy certain items with high frequency but less variation of items, others make more diverse purchases but less often purchase such items.

For the development of demand forecasting, the Markov Chain methodology was applied through Scilab, considering the second request, a ^ 2, as shown in Figure 5:

```
Scilab 5.5.2 Console
  carregando o ambiente inicial

-->a=[0.00 0.00 0.62 0.38 0.00 0.14 0.14 0.00 0.00 0.00; 0.78 0.17 0.00 0.00 0.00 0.22 0.00 0.17 0.0
a  =

    0.     0.     0.62   0.38   0.     0.14   0.14   0.     0.     0.
    0.78   0.17   0.     0.     0.     0.22   0.     0.17   0.     0.
    0.     0.     0.     0.71   0.14   0.     0.     0.     0.     0.14
    0.88   0.63   0.     0.     0.63   0.     0.     0.88   0.     0.
    0.75   0.75   0.     0.     0.     0.     0.     0.     0.5    0.88
    0.     0.     0.     0.     0.     0.     0.44   0.     0.33   0.
    0.     0.     0.     0.     0.     0.     0.     0.     0.     0.
    0.     0.     0.     0.     0.     0.     0.     0.     0.     0.
    0.     0.     0.     0.     0.     0.5    0.     0.13   0.     0.
    0.     0.     0.     0.     0.     0.     0.     0.     0.45   0.

-->a^2
ans  =

    0.3344   0.2394   0.     0.4402   0.3262   0.     0.0616   0.3344   0.0462   0.0868
    0.1326   0.0289   0.4836   0.2964   0.     0.1466   0.206    0.0289   0.0726   0.
    0.7298   0.5523   0.     0.     0.4473   0.     0.     0.6248   0.133    0.1232
    0.9639   0.5796   0.5456   0.3344   0.     0.2618   0.1232   0.1071   0.315    0.5544
    0.585    0.1275   0.465    0.285    0.     0.52     0.105    0.1925   0.396    0.
    0.     0.     0.     0.     0.     0.     0.165    0.     0.0429   0.     0.
    0.     0.     0.     0.     0.     0.     0.     0.     0.     0.     0.
    0.     0.     0.     0.     0.     0.     0.     0.     0.     0.     0.
    0.     0.     0.     0.     0.     0.     0.     0.22   0.     0.165    0.
    0.     0.     0.     0.     0.     0.225    0.     0.0585   0.     0.
```

Fig. 5 -- Markov Chain Matrix - Second Order

Figure 6 presents the data concentration graph of the matrix a ^ 2:

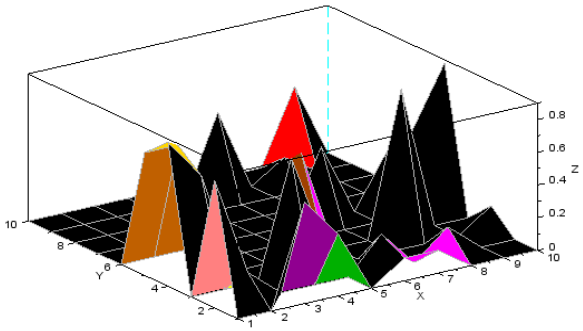

Fig. 6-- Data Concentration - Second Order

Based on the data presented in figures 5 and 6, it is possible to analyze the expected demand for future customer purchases. Two customer profiles were analyzed: one that presented a high variety of items purchased with high frequency - customer 4, which presented a high variety of items purchased with low frequency - customer 2.

Through the results obtained it is noted: that customer 4 has a demand forecast with high variety of items and high frequency of these items, and it is expected that he will buy 9 of the 10 studied items; Client 2 presents a demand forecast with a high variety of items and a low frequency of purchase of these items, and it is expected that he will buy 8 of the 10 studied items.

From these results it is possible to plan what will be the sales strategy that the company will use for each customer. Considering the customer profile 4, the company will not need to offer a differentiated sales proposal, as it will buy 90% of the studied items and, with such frequency, a high frequency.

According to customer profile 2, it is noted that despite the forecast that he will buy 80% of the items studied, such purchases have a low frequency. Therefore, it is suggested that the company offer a personalized sales proposal with the objective of maintaining the regularity of purchases and increasing the frequency with which the customer will purchase such items.

From the results obtained, it is suggested that the company apply the suggested sales actions, but follow the customers' orders in order to verify if they resulted in the expected objectives so that they can be improved.

V. CONCLUSIONS

Given the competitive market and the need for competitiveness of companies, developing sales strategies based on the customer profile is a very important factor for them to establish themselves in the market, so it is important to know the customer and their needs. Therefore, this work aimed to understand the customer and their purchase profile and perform the computer simulation of the expected demand.

The Scilab computational simulation software was used to carry out the present research, through which the Markov Chain operational research tool was applied, thus developing the expected purchase demand of the customers studied in the second request following the research.

Through the application of the tools used we studied the demand forecast of two customers with different profiles, customers 4 and 2, and the best sales strategy to be applied according to the data obtained in the simulation. It is concluded that customer 4 buys various items and with high frequency, so a personalized sales proposal is not necessary. Customer 2, on the other hand, buys several items with low frequency, thus requiring a specific sales proposal. From the results obtained, it is suggested that the company apply the suggested sales actions. It is recommended that it tracks customer orders reviewed to verify that sales practices have resulted in expected objectives so that they can be improved.

REFERENCES

[1] ABEPRO. Áreas e Subáreas de Engenharia de Produção. Disponível em <http://www.abepro.org.br/interna.asp?m=424&s=1&c=362>>. Acesso em 20 mar. 2019.

[2] ANDRADE, E. L. Introdução à pesquisa operacional: métodos e técnicas para análise de decisão. Rio de Janeiro: LTC, 2009

[3] FREITAS FILHO, Paulo José. Introdução À Modelagem E Simulação De Sistemas Com Aplicações Arena. 2 ed. Visual Books. 2008.

[4] GOLMAKANI, Ali et al. Cadeias de Markov. Texto apresentado durante a VII Bienal da Sociedade Brasileira de Matemática realizada em Maceió no ano de 2014. Disponível em: <http://www.im.ufal.br/evento/bsbm/download/minicurso/Cadeias.pdf>. Acesso em 20 mar. 2019.

[5] HORN, M. M. Modelagem e Simulação de Uma Solução de Integração Para Identificação de Gargalos de Desempenho Baseadas em Formalismo Matemático. Dissertação de Mestrado. Universidade Regional do Noroeste do Estado do Rio Grande do Sul. Ijuí, 2016.

[6] LAW, A. M. Simulation Modeling and Analysis. 5. ed. Boston: McGraw-Hill Science, 2014

[7] LIMA, Vladimir de; OLIVEIRA, Pedro. Previsão de Demanda: O Básico Que Você Precisa Saber. Edição: 1ª. Editora Baraúna, 2016.

[8] LUSTOSA, L; MESQUITA, M.A.; QUELHAS, O.; OLIVEIRA, R. J. Planejamento e Controle da Produção. Rio de Janeiro. Editora Elsevier, 2008.

[9] MIYATA, H. H.; BARRETO, A. Yoshida, M. V.; ARAÚJO, A. O.. Previsão de Demanda por Simulação de Monte Carlo em Uma Empresa Especializada em Produtos Odontológicos. IV Encontro de Engenharia de Produção Agroindustrial, Campo Mourão - Paraná, 2010.

[10] MORABITO, R. N..; PUREZA, V. 2011. In: MIGUEL, P. A. C. (Org.) Metodologia da Pesquisa Científica em Engenharia de Produção e Gestão de Operações. 2ª ed. Rio de Janeiro: Elsevier (ABEPRO). cap. 5.

[11] MOREIRA, D. A. Pesquisa Operacional: Curso Introdutório. 2ª edição. São Paulo: Cengage Learning, 2010.

[12] TAHA, H. A. Pesquisa Operacional: Uma visão geral. 8ª edição. São Paulo: Person Prentice Hall, 2008

[13] PEGDEN, C. D.;SHANNON, R. E.; SADOWSKI, R. P. Introduction to simulation using SIMAN. McGraw-Hill, NY. 2 ed., 1990

[14] Pinheiro 2013: G. Pinheiro. Teoria de filas e sistemas de comunicação. Apostila de Aula. Departamento de Engenharia Eletrônica e Telecomunicações, Universidade Estadual do Rio de Janeiro–UERJ, 2013.