

**APLICAÇÃO DE CADEIAS DE MARKOV NA ANÁLISE DA VARIAÇÃO DO  
NÚMERO DE FURTOS NA CIDADE DE CAXIAS DO SUL**

**Felipe Marcon**

Universidade de Caxias do Sul – UCS  
R. Atílio Andreazza, 3520A - Sagrada Família, Caxias do Sul-RS  
fmarcon4@ucs.br

**Franck Schaffer**

Universidade de Caxias do Sul – UCS  
Avenida Abramo Randon, 777 anexo E, Caxias do Sul-RS  
frschaffer@ucs.br

**Silvio Ismael Schvantes**

Universidade de Caxias do Sul – UCS  
R. Willibaldo Ludwig, 1500, Caxias do Sul-RS  
sischvantes@ucs.br

**Leandro Luis Corso**

Universidade de Caxias do Sul – UCS  
R. Francisco Getúlio Vargas, 1130, Petrópolis, Caxias do Sul-RS  
llcorso@ucs.br

**Michele Otobelli Bertéli**

Universidade de Caxias do Sul – UCS  
R. Francisco Getúlio Vargas, 1130, Petrópolis, Caxias do Sul-RS  
mobertel@ucs.br

**RESUMO**

Este estudo tem como principal objetivo desenvolver uma aplicação utilizando o método de Cadeias de Markov, realizado no ano de 2019, com ênfase na previsão do número de furtos analisados durante os anos de 2016 a março de 2019. Com a intenção de projetar prognósticos com fundamentação matemática, para que seja possível prever o número de furtos, foi construído a matriz de transição por meio das Cadeias de Markov onde foram definidos intervalos da variação de ocorrências de furtos, e em seguida, analisado os cenários por meio das equações de estado estável que indica a probabilidade de encontrar a variação de ocorrências. Outra análise foi o tempo esperado de recorrência destas probabilidades. O estudo mostrou que o intervalo com maior probabilidade de ocorrer é da variação da faixa entre -12,5% a -1,0% (diminuição de ocorrências) e com tempo de recorrência de 3,53 meses.

**Palavra-chave:** Cadeias de Markov; Prognósticos; Estado estável; Furtos.

**ABSTRACT**

The main objective of this study is to develop an application using the Markov Chains method, conducted in 2019, with an emphasis on predicting the number of thefts analyzed from 2016 to March 2019. With the intention of projecting predictions based on In order to predict the number of thefts, the transition matrix was constructed using the Markov Chains, where intervals of variation of thefts were defined, and then the scenarios were

analyzed using the steady state equations. Indicates the probability of finding the variation of occurrences. Another analysis was the expected recurrence time of these probabilities. The study showed that the interval most likely to occur ranges from -12.5% to -1.0% (decrease in occurrences) and with a recurrence time of 3.53 months.

**Keywords:** Markov Chains; Predictions; Steady state; Thefts.

#### Como Citar:

MARCON, Felipe; SCHAFFER, Franck R.; SCHVANTES, Silvio I.; CORSO, Leandro L.; BERTELI, Michele O.. Aplicação de cadeias de Markov na análise da variação do número de furtos na cidade de Caxias do Sul. *In: SIMPÓSIO DE PESQUISA OPERACIONAL E LOGÍSTICA DA MARINHA*, 19., 2019, Rio de Janeiro, RJ. **Anais** [...]. Rio de Janeiro: Centro de Análises de Sistemas Navais, 2019.

## 1. INTRODUÇÃO

Atualmente, a crise econômica que o Brasil tem passado traz consigo uma série de fatores que desafiam o bem-estar da população (Mascaro (2016)). A situação da criminalidade e insegurança podem ser considerados como alguns dos principais. Segundo relatório do Programa das Nações Unidas (PNUD), o Brasil está entre os países com as maiores taxas de roubos do mundo. Ao mesmo tempo, segundo Jaitan et al. (2017), o Brasil é um dos países com um dos maiores custos nominais da América Latina. Entretanto, sabe-se que aplicações por meio de modelos matemáticos podem trazer excelentes benefícios ao apoio a segurança pública. Pode-se citar a Pesquisa Operacional (PO) como uma área que possui modelos para esta finalidade.

Segundo Carter e Price (2001), a PO, que teve surgimento durante a 2ª Guerra Mundial, proporcionou diversos avanços em termos de aplicações. Logo após o término da guerra, foram aplicadas em diversas áreas, também fora do âmbito militar. Em outras aplicações, o foco de segurança também foi dado. Por exemplo, Taylor e Huxley (1989) aplicaram na polícia de São Francisco um modelo matemático para apoio na escala, obtendo um valor de economia de U\$ 11 milhões/ano, além de propiciar melhora na eficiência e atendimento a sociedade. Em outro trabalho, Curtin, Hayslett-mccall e Qiu (2007) apresentaram um modelo matemático capaz de reduzir em 19% a distância percorrida pelas patrulhas, fato este que mostra ganho em eficiência e redução de custo. Outras técnicas de PO também podem ser avaliadas para dados que são aleatórios, fazendo com que se possam obter parâmetros otimizados e ao mesmo tempo confiáveis (CORSO; WALLACE (2015)).

Este trabalho tem como principal objetivo desenvolver a aplicação de Cadeias de Markov, com ênfase na previsão do número de furtos, da cidade de Caxias do Sul, nordeste do Rio Grande do Sul. A cidade escolhida para análise é considerada parte do segundo maior polo metalmeccânico do país. Esta aplicação tem a intenção de projetar prognósticos, com fundamentação matemática em processos estocásticos. Desta forma, visa-se mostrar a metodologia como possibilidade de avaliar probabilidades de transição para este tipo de situação, bem como avaliar o estado estável da base de dados definida para este trabalho.

## 2. REFERENCIAL

Segundo (CARTER; PRICE, 2001), as Cadeias de Markov são definidas como um modelo matemático, baseado em princípios de teoria de probabilidade estudados por Andrei Andreyevich Markov, no início do século XX. Esse método considera uma sequência de variáveis aleatórias, onde a probabilidade de transição de uma para a outra pode variar ao longo do tempo. De acordo com Hillier (2013), as cadeias de Markov possuem a propriedade especial de que as probabilidades que envolvem como o processo evolui no futuro dependem apenas do estado atual do processo e, portanto, são independentes de eventos passados. O princípio deste método está na consideração de processos estocásticos. Os processos estocásticos podem ser definidos como um conjunto de variáveis aleatórias  $\{X_t, t \in T\}$ , em um determinado espaço de tempo. Segundo Taha (2008), o processo estabelece que cada variável (x) pertencente ao conjunto de dados, caracteriza o estado do sistema  $X_t$ , em postos discretos do tempo  $t = \{1, 2, 3, \dots\}$ . Estes processos devem seguir o princípio Markoviano, onde se diz que o estado futuro é dependente do presente, e não do futuro, tornando a Equação 1, que é uma definição de geral com passado, presente e futuro, na Equação 2, que mostra a relação do futuro com o presente.

$$P\{X_{t+1}=j \mid X_0=k_0, X_1=k_1, \dots, X_{t-1}=k_{t-1}, X_t=i\} = P\{X_{t+1}=j \mid X_t=i\} \quad (1)$$

$$p_{ij} = P\{X_t = j \mid X_{t-1} = i\} \quad i, j = 1, 2, \dots, n \mid t = 0, 1, 2, \dots, T \quad (2)$$

Para representação das transições entre os estados, considera a seguinte Matriz

$$P = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1N} \\ p_{21} & p_{22} & \dots & p_{2N} \\ \dots & \dots & \dots & \dots \\ p_{N1} & p_{N2} & \dots & p_{NN} \end{bmatrix} \quad (3)$$

Onde  $p_{ij}$  representa a probabilidade de transição do estado  $i$  para o estado  $j$  em um passo de tempo. Dentro do método, também é possível verificar o estado estável da base de dados. Uma das mais importantes características exibidas pelas Cadeias de Markov é um comportamento de convergência ao equilíbrio em longo prazo (NORRIS, 1997). Conforme Hillier e Lieberman (2013), em uma Cadeia de Markov se pode chegar o regime estacionário, ou também chamado de ergótico. Para isso, define-se  $\pi_j$  como a probabilidade estável do estado  $j$ . A forma de obtenção destes termos é dada por:

$$\pi_j = \sum_{i=0}^M \pi_i p_{ij}, \text{ para } j = 0, 1, 2, \dots, M \quad (4)$$

$$\sum_{j=0}^M \pi_j = 1 \quad (5)$$

Da mesma forma, também se pode calcular o tempo de recorrência do modelo, que é dado por:

$$\mu_{ii} = \frac{1}{\pi_i} \quad (6)$$

Onde representa o tempo esperado de sair de um estado e retornar para o mesmo.

### 3. APLICAÇÃO E RESULTADOS

Foram selecionados os dados de furtos entre janeiro de 2016 até março de 2019. Para desenvolvimento deste estudo de Cadeias de Markov, dividiram-se em três etapas.

A primeira etapa diz respeito a criação das faixas para análise, neste caso, foram criadas seis faixas, como mostra a Tabela 1.

Tabela 1 - Intervalos das variações

Intervalos de variação	< -12,5%	-12,5% a - 1,0%	-1,0% a 10,5%	10,5% a 22,0%	22,0% a 33,5%	> 33,5%
------------------------	----------	-----------------	---------------	---------------	---------------	---------

Fonte: Elaborado pelos autores (2019)

O passo seguinte foi a obtenção da Matriz de Transição, apresentada na Tabela 2.

Tabela 2 - Intervalos das variações

Intervalos de variação	< -12,5%	-12,5% a - 1,0%	-1,0% a 10,5%	10,5% a 22,0%	22,0% a 33,5%	> 33,5%
≤ -12,5%	0,20	0,10	0,20	0,40	0,00	0,10
-12,5% a - 1,0%	0,40	0,10	0,20	0,10	0,10	0,10
-1,0% a 10,5%	0,13	0,38	0,38	0,13	0,00	0,00
10,5% a 22,0%	0,29	0,57	0,00	0,14	0,00	0,00
22,0% a 33,5%	1,00	0,00	0,00	0,00	0,00	0,00
≥ 33,5%	0,00	1,00	0,00	0,00	0,00	0,00

Fonte: Elaborado pelos autores (2019)

Por meio da matriz de transição e utilizando as Equações 4 e 5 se obtém o seguinte sistema de equações, onde se destaca que uma das equações entre 7 e 12 é redundante e deve ser excluída para solução.

$$\pi_0 = 0,2\pi_0 + 0,4\pi_1 + 0,13\pi_2 + 0,29\pi_3 + \pi_4 + 0\pi_5 \quad (7)$$

$$\pi_1 = 0,1\pi_0 + 0,1\pi_1 + 0,38\pi_2 + 0,57\pi_3 + 0\pi_4 + \pi_5 \quad (8)$$

$$\pi_2 = 0,2\pi_0 + 0,2\pi_1 + 0,38\pi_2 + 0\pi_3 + 0\pi_4 + 0\pi_5 \quad (9)$$

$$\pi_3 = 0,4\pi_0 + 0,1\pi_1 + 0,13\pi_2 + 0,14\pi_3 + 0\pi_4 + 0\pi_5 \quad (10)$$

$$\pi_4 = 0\pi_0 + 0,1\pi_1 + 0\pi_2 + 0\pi_3 + 0\pi_4 + 0\pi_5 \quad (11)$$

$$\pi_5 = 0,1\pi_0 + 0,1\pi_1 + 0\pi_2 + 0\pi_3 + 0\pi_4 + 0\pi_5 \quad (12)$$

$$1 = \pi_0 + \pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 \quad (13)$$

O resultado do modelo é apresentado na Tabela 3.

Tabela 3 – Valor do estado estável

Estado estável	$\pi_0$	$\pi_1$	$\pi_2$	$\pi_3$	$\pi_4$	$\pi_5$
Valor obtido	0,271	0,283	0,177	0,185	0,028	0,055

Fonte: Elaborado pelos autores (2019)

Por meio destes dados é possível verificar que o intervalo -12,5% a -1,0% é o que apresenta a maior probabilidade, totalizando 28,3%. Considerando estes valores de estado estável e utilizando a equação 5 é possível obter o tempo esperado de recorrência, apresentados na Tabela 4.

Tabela 4 – Tempo esperado de recorrência

Tempo Esperado	$\mu_{00}$	$\mu_{11}$	$\mu_{22}$	$\mu_{33}$	$\mu_{44}$	$\mu_{55}$
Valor obtido	3,69	3,53	5,65	5,41	35,71	18,18

Fonte: Elaborado pelos autores (2019)

Nesta tabela é possível identificar que o tempo esperado de recorrência para a faixa < -12,5% é de 3,69 meses, para a faixa -12,5% a -1,0% é de 3,53 meses e assim para todas as demais.

#### 4. CONCLUSÃO

Os resultados obtidos nesse artigo, com a aplicação do método de Cadeias de Markov, mostram que é possível utilizar informações para análise estocástica da variação dos furtos na cidade de Caxias do Sul. Para este estudo, foram utilizados dados históricos do número de furtos ocorridos, registrados pela Secretaria de Segurança Pública de Janeiro de 2016 à Março de 2019.

Por meio do cálculo do estado estável foi possível avaliar a previsão de ocorrências de furtos em seis variações de intervalos, onde se mostrou a maior probabilidade na faixa de 12,5% a -1,0% (diminuição de furtos) e com um tempo de recorrência de 3,53 meses. Próximo a isso, o segundo intervalo com maior probabilidade de ocorrência foi a faixa < 12,5% (diminuição de furtos) com o tempo de recorrência de 3,69 meses. Isso indica que provavelmente o número de furtos na cidade tende a reduzir. No presente trabalho os dados foram considerados para os períodos em forma sequencial, sugere-se que em trabalhos futuros uma análise de sazonalidade dos dados possa ser realizada. Também, um acompanhamento dos dados em tempo real, é considerado fundamental para utilização prática do modelo.

A partir dos resultados obtidos, certifica-se que a análise do número de furtos ocorridos, utilizando Cadeias de Markov para conhecimento das probabilidades de variações para determinados intervalos de furtos, poderá auxiliar os responsáveis cabíveis na área de Segurança Pública para utilização dessas informações.

## 5. REFERÊNCIAS BIBLIOGRÁFICAS

- [1] SECRETARIA DE SEGURANÇA PÚBLICA. Indicadores criminais. Disponível em: < <https://www.ssp.rs.gov.br/indicadores-criminais> > Acesso em 22 abr. 2019
- [2] HAMDY, A. TAHA. Pesquisa Operacional: uma visão geral. 8 ed. São Paulo: Pearson Prentice Hall, 2008. Terceira referência.
- [3] FREDERICK, S. HILLIER, GERALD, J. LIEBERMAN. Introdução à Pesquisa Operacional. 9 ed. Porto Alegre: AMGH, 2013
- [4] NORRIS, J.R., Markov Chains Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 1997.
- [5] MASCARO, Alysson. Direito, crise e impeachment no Brasil. Observatório Latinoamericano y Caribeño. [S. l.], v. 1, n. [S. l.], p. 175-194, out. 2016.
- [6] CARTER, M.I W.; PRICE C. C. Operations research: a practical introduction.[Boca Raton, Florida]: CRC Press, 2001.
- [7] JAITMAN, L. et al.. The costs of crime and violence: new evidence and insights in latin america and the caribbean. Inter-American Development Bank, p.1-132, fev. 2017.
- [8] TAYLOR, P. E.; HUXLEY, S. J.. A break from tradition for the San Francisco police: patrol officer scheduling using an optimization-based decision support system. Interfaces, v. 19, n. 1, p. 4-24, 1989.
- [9] CURTIN, K M.; HAYSLETT-MCCALL, K.; QIU, F.. Determining optimal police patrol areas with maximal covering and backup covering location models. Networks And Spatial Economics, v. 10, n. 1, p.125-145, 2007.
- [10] CORSO, Leandro L.; WALLACE, Mark. A hybrid method for transportation with stochastic demand. International Journal of Logistics, 2015, v. 18, p. 1-13