



Universidade Federal de Mato Grosso
Instituto de Ciências Exatas e da Terra
Departamento de Matemática



Análise de variação de preços utilizando Cadeias de Markov e Lógica Fuzzy

Priscila Gabriele Pacheco

Mestrado Profissional em Matemática: PROFMAT/SBM

Orientador: **Prof. Dr. Moisés dos Santos Ceconello**

Trabalho financiado pela Capes

Cuiabá - MT

Agosto de 2019

Análise de variação de preços utilizando Cadeias de Markov e Lógica Fuzzy

Este exemplar corresponde à redação final da dissertação, devidamente corrigida e defendida por Priscila Gabriele Pacheco e aprovada pela comissão julgadora.

Cuiabá, 27 de agosto de 2019.

Prof. Dr. Moisés dos Santos Cecconello
Orientador

Banca examinadora:

Prof. Dr. Moisés dos Santos Cecconello
Prof. Dr. André Krindges
Prof. Dr. Edgar Nascimento

Dissertação apresentada ao curso de Mestrado Profissional em Matemática – PROFMAT, da Universidade Federal de Mato Grosso, como requisito parcial para obtenção do título de **Mestre em Matemática**.

Dados Internacionais de Catalogação na Fonte.

P116a Pacheco, Priscila Gabriele.
Análise de variação de preços utilizando Cadeias de Markov e Lógica Fuzzy / Priscila Gabriele Pacheco. -- 2019
xii, 61 f. : il. ; 30 cm.

Orientador: Moisés dos Santos Ceconello.
Dissertação (mestrado profissional) – Universidade Federal de Mato Grosso, Instituto de Ciências Exatas e da Terra, Programa de Pós-Graduação Profissional em Matemática, Cuiabá, 2019.
Inclui bibliografia.

1. incerteza. 2. matriz de transição. 3. séries temporais. I. Título.

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

Permitida a reprodução parcial ou total, desde que citada a fonte.



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE MATO GROSSO
PRÓ-REITORIA DE ENSINO DE PÓS-GRADUAÇÃO
Programa de Pós-Graduação em Ensino de Matemática em Rede Nacional - Profmat
Av. Fernando Corrêa da Costa, 2367 - Boa Esperança - 78.060-900 - Cuiabá/MT
Fone: (65) 3615-8576 - E-mail: profmat@ufmt.br

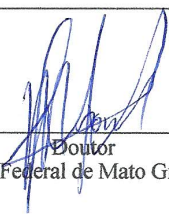
FOLHA DE APROVAÇÃO


Título: "Análise de variações de preço utilizando cadeias de Markov e lógica fuzzy"

Autor: Priscila Gabriele Pacheco

defendida e aprovada em 27/08/2019.

Composição da Banca Examinadora:

Presidente Banca/Orientador  Doutor Moisés dos Santos Ceconello
Instituição: Universidade Federal de Mato Grosso

Examinador Interno  Doutor André Krindges
Instituição: Universidade Federal de Mato Grosso

Examinador Externo  Doutor Edgar Nascimento
Instituição: Instituto Federal de Mato Grosso

Cuiabá, 27/08/2019.

Aos meus amores de sempre.

Agradecimentos

Agradeço, primeiramente a minha família, Marlene, Luiz e Letícia, pela paciência ao suportar dois (quase três) anos de estresse, falhas e ausência. A Letícia, por ser meu suporte, um poço sem fim de incentivo e irmandade.

Ao Marcus, em especial, por ser alguém com quem contar, pelo suporte que me deu e por ser meu motorista particular desde o exame de acesso até a última confraternização.

Ao meu orientador, prof. Moisés, a quem só tenho a agradecer pela paciência e compreensão.

Aos meus colegas de mestrado, por me darem motivos para divertir muita gente. Por serem motivo de tanta diversão, obrigada!

À minha equipe de trabalho, por segurarem as pontas todas as sextas-feiras para que eu pudesse adquirir o título. Em especial, agradeço Iza e Cris pelo apoio e incentivo, sem elas esse momento não existiria.

Muito obrigada a todos.

Não sou nada.

Nunca serei nada.

Não posso querer ser nada.

*À parte isso, tenho em mim todos os
sonhos do mundo.*

Fernando Pessoa.

Resumo

Os conceitos de cadeias de Markov e lógica Fuzzy são importantes teorias cujo foco está no tratamento de incertezas, onde pode-se aplicar a previsão de variação de preços de produtos negociados em bolsa. Este trabalho tem como objetivo apresentar um estudo de aplicação para os conceitos de cadeias de Markov e lógica Fuzzy. Os dados presentes neste trabalho foram coletados no site do Centro de Estudos Avançados em Economia Aplicada, departamento integrante da Universidade de São Paulo, referentes a toda base de precificação das commodities soja e frango. Aborda razões pelas quais as duas teorias são consideradas importantes, como podem ser associadas ao estudo de séries temporais, bem como algumas de suas características e definições mais interessantes para o modelo apresentado. Também pode servir de suporte a outros trabalhos interessados em outros tipos de variações de preços em ativos diversos negociados em bolsa ou balcão. Considerando o tamanho da exportação para o mercado brasileiro, informações como análise de variação do preço são úteis para o gerenciamento dos negócios. Os modelos Fuzzy e determinístico apresentaram resultados muito parecidos para a variação de preços da soja, mostrando com clareza a tendência positiva desta série temporal. Quanto a variação de preços do frango, os modelos variaram um pouco, apresentando outras possibilidades de análise.

Palavras chave: incerteza, matriz de transição, séries temporais.

Abstract

Concepts of Markov chains and Fuzzy logic are important theories that focus on dealing with uncertainties, it can be applied the forecast of price variation of exchange traded products. This paper aims to present an application to the concepts of Markov chains and Fuzzy logic. The data presented in this paper were collected on the Center for Advanced Studies in Applied Economics, University of São Paulo, for the entire soybean and chicken commodity pricing basis. It discusses how the both theories are considered important, as they may be associated with the study of time series, as well as some of their most interesting characteristics and definitions for the model presented. It may also support other work that is interested in other types of price fluctuations in various exchange or over the counter assets. Considering the size of the export to the Brazilian market, information such as price change analysis is useful for business management. The described models presented very similar results for soybean price variation, clearly presenting the positive trend of this time series. Regarding the variation of chicken prices, the models varied slightly, presenting other possibilities of analysis.

Keywords: uncertainty, transition matrix, time series.

Sumário

Agradecimentos	v
Resumo	vii
Abstract	viii
Lista de figuras	xi
Lista de tabelas	xii
Introdução	1
1 Cadeias de Markov	3
1.1 Introdução	3
1.2 Cadeias de Markov	4
1.3 Análise Assintótica	11
2 Teoria Fuzzy	29
2.1 Conjuntos e Eventos Fuzzy	29
2.2 Cadeias de Markov com estados Fuzzy	33
3 O Modelo	37
3.1 A proposta	39
3.2 Aplicação Cadeias de Markov	40
3.3 Aplicação Fuzzy	49
3.4 Resultados	53
Considerações finais	54

Referências Bibliográficas	56
Apêndice: Scripts R	57
Apêndice: Scripts Matlab	59

Lista de Figuras

3.1	Base de dados completa da Soja.	41
3.2	Base de dados completa do Frango.	41
3.3	Gráfico da Variação do preço da Soja.	43
3.4	Gráfico da Variação do preço da carne de Frango.	43
3.5	Histograma dos intervalos da variação dos preços de Soja.	44
3.6	Histograma dos intervalos da variação dos preços da carne de Frango.	44
3.7	Funções de pertinências do sistema fuzzy.	50

Lista de Tabelas

1.1	Resumo das propriedades dos estados da Cadeia de Markov	27
3.1	Estatística descritiva dos dados	42
3.2	Intervalo das variações	43
3.3	Intervalo das variações - Soja	47
3.4	Resultados - Soja	47
3.5	Intervalo das variações - Frango	49
3.6	Resultados - Frango	49
3.7	Resultados (Fuzzy) - Soja	51
3.8	Resultados (Fuzzy) - Frango	53

Introdução

A matemática esteve presente durante o desenvolvimento intelectual humano e como consequência da história das civilizações. O uso da matemática como forma de analisar âmbitos econômicos e sociais data de meados do século XVII, entretanto a matematização da economia só começou de fato no século XIX. Tal acontecimento foi impulsionado pela industrialização e as ciências econômicas sofreram forte influência pelos métodos de modelagem matemática inspiradas nas ciências naturais, em particular na física.

Através da matemática é possível desenvolver modelos econômicos com o propósito de construir uma teoria econômica rigorosa e unificada. Desta forma, a matemática permite aos economistas formular proposições significativas e testáveis sobre muitos assuntos complexos e abrangentes que não poderiam ser adequadamente expressas informalmente.

Para analisar a variação de preços foram explorados conceitos de Cadeias de Markov e Lógica Fuzzy. No primeiro capítulo são apresentados os conceitos teóricos de Cadeias de Markov.

O capítulo dois apresenta a lógica Fuzzy, desenvolvida por Lofti A. Zadeh, a qual se fundamenta em determinar um grau de pertinência que indicará quanto um elemento pertence a um determinado conjunto.

Já o capítulo três apresenta o modelo de previsão da variação de preços, primeiro através do modelo clássico de Cadeias de Markov com matriz de transição calculada através de probabilidade condicional convencional. Em um momento seguinte apresenta-se o modelo baseado em estados e probabilidade condicional fuzzy e a discussão dos resultados.

O objetivo do trabalho é calcular a probabilidade de variação de preços de commodities, utilizando cadeias de markov e lógica fuzzy, além do tempo de recorrência esperado, a fim de ser uma informação útil para gestores do agronegócio.

Informações como variação do preço de ativos em geral, não apenas de commodities, podem ser muito úteis para gestores de carteiras, especuladores e equipes de gerenciamento de riscos.

Capítulo 1

Cadeias de Markov

1.1 Introdução

A incerteza se caracteriza como um ambiente naturalmente encontrado em qualquer sistema real, principalmente quando ações humanas imprevisíveis atuam sobre o mesmo. É possível modelar fenômenos ou experimentos aleatórios onde o acaso representa um papel fundamental, chamados de não determinísticos, estão inseridos na teoria da probabilidade, ramo da Matemática que cria, desenvolve e pesquisa modelos como esse (Morgado et al. (2016)).

A noção de acaso é quase tão antiga quanto a própria humanidade, no entanto, seu entendimento como fenômeno natural demorou para ser aceito, porque era comumente creditado a divindades. As primeiras manifestações probabilísticas se deram através de jogos, mais precisamente o Tali (jogo de osso), considerado o ancestral do dado. E assim, com o desenvolvimento do comércio e uma quantidade cada vez maior de jogos de azar, a teoria da probabilidade surgiu como disciplina matemática baseada nas tentativas de quantificar riscos associados a sinistros e possibilidades de ganho (Viali (2008)).

Desde então, vários matemáticos dedicaram parte de seus estudos para o desenvolvimento da teoria da probabilidade. E as contribuições se tornaram maiores a partir de Laplace, entre eles alguns nomes de destaque são Chebyshev, Markov, von Mises, e o atualmente considerado pai da probabilidade moderna, Andrei Kolmogorov.

Um desses nomes de destaque, o russo Andrei Andreyevich Markov (1856-1922) aplicou o método das frações contínuas, inicialmente utilizado por seu professor Pafnuty Lvovich Chebyshev (1821-1894) à teoria da probabilidade. Também estudou sequências de

variáveis mutuamente independentes esperando estabelecer as leis probabilísticas em formas mais gerais. Conhecido pelas sequências de variáveis aleatórias na qual uma variável é determinada pelo valor da anterior, mas independente no sentido de que o estado presente depende apenas da sua anterior, foi este trabalho que lançou a teoria dos processos estocásticos.

1.2 Cadeias de Markov

O processo estocástico é uma extensão do conceito de variável aleatória. A variável aleatória como função do tempo é chamada de processo estocástico (ou processo aleatório).

Em outras palavras, um processo estocástico $\{X_n, n = 0, 1, 2, 3, \dots\}$ é uma família de variáveis aleatórias parametrizadas pelo inteiro n , assumimos que cada uma destas variáveis aleatórias X_n toma seus valores em um conjunto finito T (Rousseau e Saint-Aubin (2015)).

Já o processo estocástico $X(t)$ é de tempo discreto se $X(t)$ é definido apenas para um conjunto de instantes de tempo $t_n = nT$, onde T é uma constante e n é um inteiro; caso contrário, $X(t)$ é um processo de tempo contínuo.

Segundo Privault (2013), o processo estocástico é dito "processo de Markov", ou tem a propriedade markoviana se, para todo $n \geq 1$ a probabilidade de distribuição de X_{n+1} está determinada pelo estado X_n no tempo n , e não depende dos valores de X_k , para $k \leq n - 1$. Em outras palavras, para todo $n \geq 1$ e $i_0, i_1, \dots, i_n, j \in T$ tem-se

$$\mathbb{P}(X_{n+1} = j | X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = \mathbb{P}(X_{n+1} = j | X_n = i_n)$$

Em particular,

$$\mathbb{P}(X_{n+1} = j | X_n = i_n, X_{n-1} = i_{n-1}) = \mathbb{P}(X_{n+1} = j | X_n = i_n),$$

e

$$\mathbb{P}(X_2 = j | X_1 = i_1, X_0 = i_0) = \mathbb{P}(X_2 = j | X_1 = i_1).$$

Por outro lado, a primeira ordem de probabilidades de transição pode ser usada

para entender o processo como

$$\begin{aligned}
& \mathbb{P}(X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) \\
&= \mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1}, \dots, X_0 = i_0) \\
&\quad \times \mathbb{P}(X_{n-1} = i_{n-1}, \dots, X_0 = i_0) \\
&= \mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1}) \mathbb{P}(X_{n-1} = i_{n-1} | X_{n-2} = i_{n-2}, \dots, X_0 = i_0) \\
&\quad \times \mathbb{P}(X_{n-2} = i_{n-2}, \dots, X_0 = i_0) \\
&= \mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1}) \mathbb{P}(X_{n-1} = i_{n-1} | X_{n-2} = i_{n-2}) \\
&\quad \times \mathbb{P}(X_{n-2} = i_{n-2} | X_{n-3} = i_{n-3}, \dots, X_0 = i_0)
\end{aligned}$$

Por indução finita, tem-se que

$$\begin{aligned}
& \mathbb{P}(X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) \\
&= \mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1}) \dots \mathbb{P}(X_1 = i_1 | X_0 = i_0) \mathbb{P}(X_0 = i_0)
\end{aligned} \tag{1.1}$$

$i_0, i_1, \dots, i_n \in T$. Então

$$\begin{aligned}
\mathbb{P} &= \sum_{j \in T} \mathbb{P}(X_1 = i, X_0 = j) \\
&= \sum_{j \in T} \mathbb{P}(X_1 = i, X_0 = j) \mathbb{P}(X_0 = j), i \in T
\end{aligned} \tag{1.2}$$

Segue uma definição mais formal descrita por Privault (2013).

Definição 1. *Seja $\{X_n, n = 0, 1, 2, 3, \dots\}$ um processo estocástico tomando seus valores num conjunto $T = \{A, B, C, \dots\}$. Dizemos que $\{X_n\}$ é uma cadeia de Markov se a probabilidade $P(X_n = i)$, $i \in T$, depender somente do valor do processo no anterior, X_{n-1} , e não em qualquer dos passos anteriores X_{n-2}, X_{n-3}, \dots . Definimos $N < \infty$ como o número de elementos em T .*

Essa possibilidade de calcular $P(X_n)$ usando apenas $P(X_{n-1})$ é a propriedade que define as cadeias de Markov. Ou seja, essas estruturas não possuem memória de estados passados, e o estado futuro é determinado completamente pelo atual.

Por exemplo, considerando o caminho aleatório de números inteiros, onde $(X_n)_{n \geq 1}$

é uma sequência independente de valores aleatoriamente incrementados, é uma cadeia de Markov de tempo discreto com $T = \mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$.

De fato, para todo $j, i_n, \dots, i_1 \in \mathbb{Z}$, neste caso $T_0 = 0$, tem-se

$$\begin{aligned}
& \mathbb{P}(T_{n+1} = j | T_n = i_n, T_{n-1} = i_{n-1}, \dots, T_1 = i_1) = \\
&= \frac{\mathbb{P}(T_{n+1} = j, T_n = i_n, T_{n-1} = i_{n-1}, \dots, T_1 = i_1)}{\mathbb{P}(T_n = i_n, T_{n-1} = i_{n-1}, \dots, T_1 = i_1)} \\
&= \frac{\mathbb{P}(T_{n+1} - T_n = j - i_n, T_n - T_{n-1} = i_n - i_{n-1}, \dots, T_2 - T_1 = i_2 - i_1, T_1 = i_1)}{\mathbb{P}(T_n - T_{n-1} = i_n - i_{n-1}, \dots, T_2 - T_1 = i_2 - i_1, T_1 = i_1)} \\
&= \frac{\mathbb{P}(X_{n+1} = j - i_n, X_n = i_n - i_{n-1}, \dots, X_2 = i_2 - i_1, X_1 = i_1)}{\mathbb{P}(X_n = i_n - i_{n-1}, \dots, X_2 = i_2 - i_1, X_1 = i_1)} \\
&= \frac{\mathbb{P}(X_{n+1} = j - i_n) \mathbb{P}(X_n = i_n - i_{n-1}, \dots, X_2 = i_2 - i_1, X_1 = i_1)}{\mathbb{P}(X_n = i_n - i_{n-1}, \dots, X_2 = i_2 - i_1, X_1 = i_1)} \\
&= \mathbb{P}(X_{n+1} = j - i_n) \\
&= \frac{\mathbb{P}(X_{n+1} = j - i_n) \mathbb{P}(X_n + \dots + X_1 = i_n)}{\mathbb{P}(X_n + \dots + X_1 = i_n)} \\
&= \frac{\mathbb{P}(X_{n+1} = j - i_n, X_n + \dots + X_1 = i_n)}{\mathbb{P}(X_1 + \dots + X_n = i_n)} \\
&= \frac{\mathbb{P}(X_{n+1} = j - i_n, T_n = i_n)}{\mathbb{P}(T_n = i_n)} = \frac{\mathbb{P}(X_{n+1} = j, T_n = i_n)}{\mathbb{P}(T_n = i_n)} = \mathbb{P}(T_{n+1} = j | T_n = i_n)
\end{aligned}$$

Em geral, todo processo com incrementos independentes são processos markovianos. No entanto, nem toda cadeia de Markov tem incrementos independentes. De fato, cadeias de Markov que não possuem incrementos independentes são as de maior interesse.

Rousseau e Saint-Aubin (2015), indicam que Cadeias de Markov são únicas por seu comportamento ser completamente caracterizado por seu estado inicial e uma matriz de transição dada por

$$P(X_n = i | X_{n-1} = j) = p_{ij} \quad (1.3)$$

Em outras palavras, a evolução aleatória de uma Cadeia de Markov $(X_n)_{n \in \mathbb{N}}$ é determinada por

$$P_{i,j} := \mathbb{P}(X_a = j | X_0 = i), \quad i, j \in T.$$

Que coincide com a probabilidade $\mathbb{P}(X_{n+1} = j | X_n = i)$ que é independente de $n \in \mathbb{N}$. Neste caso, a cadeia $(X_n)_{n \in \mathbb{N}}$ é dita homogênea em relação ao tempo.

Então, esses dados podem ser organizados em uma matriz $T^2 = T \times T$, chamada Matriz de Transição. Logo, uma matriz P é de transição de uma cadeia de Markov se, e somente se, $p_{ij} \in [0, 1]$, para todo $i, j \in T$ e $\sum_{i \in T} p_{ij} = 1$ para todo $j \in T$.

$$[P_{i,j}]_{i,j \in T} = [\mathbb{P}(T_1 = j | T_0 = i)]_{i,j \in T}$$

Também escrita como

$$[P_{i,j}]_{i,j \in T} = \begin{bmatrix} \cdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots \\ \cdots & P_{-2,-2} & P_{-2,-1} & P_{-2,0} & P_{-2,1} & P_{-2,2} & \cdots \\ \cdots & P_{-1,-2} & P_{-1,-1} & P_{-1,0} & P_{-1,1} & P_{-1,2} & \cdots \\ \cdots & P_{0,-2} & P_{0,-1} & P_{0,0} & P_{0,1} & P_{0,2} & \cdots \\ \cdots & P_{1,-2} & P_{1,-1} & P_{1,0} & P_{1,1} & P_{1,2} & \cdots \\ \cdots & P_{2,-2} & P_{2,-1} & P_{2,0} & P_{2,1} & P_{2,2} & \cdots \\ \cdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots \end{bmatrix}$$

Vale a pena ressaltar a inversão da ordem dos índices (i, j) entre $\mathbb{P}(X_{n+1} = j | X_n = i)$ e P_{ij} . Em particular, o estado inicial i é uma linha na matriz, enquanto o estado final j corresponde a uma coluna da matriz.

Devido a relação

$$\sum_{j \in T} \mathbb{P}(X_1 = j | X_0 = i) = 1, \quad i \in \mathbb{N} \quad (1.4)$$

as linhas da matriz de transição satisfazem condição

$$\sum_{j \in T} P_{i,j} = 1, \quad (1.5)$$

para todo $i \in T$.

Usando a notação de matriz $P = (P_{i,j})_{i,j \in T}$ e a relação 1.1 encontra-se

$$\mathbb{P}(X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = P_{i_{n-1}, i_n} \dots P_{i_0, i_1} \mathbb{P}(X_0 = i_0)$$

$i_0, i_1, \dots, i_n \in T$ e reescrevemos 1.2 como

$$\mathbb{P}(X_1 = i) = \sum_{j \in T} \mathbb{P}(X_1 = 1 | X_0 = j) \mathbb{P}(X_0 = j) = \sum_{j \in T} P_{j,i} \mathbb{P}(X_0 = j), i \in T. \quad (1.6)$$

Definição 2. Diz-se que um estado é estacionário se $P_{k,k} = 1$.

Considerando $T = \mathbb{N} = \{0, 1, 2, \dots\}$ e cadeias de Markov tomando valores no conjunto dos números naturais, em que a matriz de transição $[\mathbb{P}(X_{n+1} = j | X_n = i)]_{i,j \in \mathbb{N}}$ é escrita como

$$[P_{i,j}]_{i,j \in \mathbb{N}} = \begin{bmatrix} P_{0,0} & P_{0,1} & P_{0,2} & \cdots \\ P_{1,0} & P_{1,1} & P_{1,2} & \cdots \\ P_{2,0} & P_{2,1} & P_{2,2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

De 1.4 temos $\sum_{j=0}^{\infty} P_{i,j} = 1$ para todo $i \in \mathbb{N}$. Os elementos p_{ij} da matriz de transição P representam as probabilidades de encontrar-se no estado $i \in T$ quando vindo do estado $j \in T$.

A expressão

$$\sum_{i \in T} p_{ij} = \sum_{i \in T} P(X_n = i | X_{n-1} = j) = 1,$$

pode ser lida de forma que ao passo $n-1$, o sistema está no estado j , então, a probabilidade de estar em algum estado possível no estado n é 1.

No caso da cadeia de Markov $(X_k)_{k \in \mathbb{N}}$ em que se toma valores no espaço finito de estados $T = \{0, 1, \dots, N\}$ a matriz de transição simplificada $(N+1) \times (N+1)$ tem a forma

$$[P_{i,j}]_{0 \leq i,j \leq N} = \begin{bmatrix} P_{0,0} & P_{0,1} & P_{0,2} & \cdots & P_{0,N} \\ P_{1,0} & P_{1,1} & P_{1,2} & \cdots & P_{1,N} \\ P_{2,0} & P_{2,1} & P_{2,2} & \cdots & P_{2,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_{N,0} & P_{N,1} & P_{N,2} & \cdots & P_{N,N} \end{bmatrix}$$

Ainda que no espaço de estado finito $T = \{0, 1, \dots, N\}$, a relação 1.6 pode ser

reescrita usando notação de matrizes e vetores como

$$\eta = \Pi.P$$

onde

$$\Pi = [\Pi_0, \dots, \Pi_N] = [\mathbb{P}(X_0 = 0), \dots, \mathbb{P}(X_0 = N)] \in \mathbb{R}^{N+1}$$

é o vetor, ou linha, distribuição de X_0 e

$$\eta = [\eta_0, \dots, \eta_N] = [\mathbb{P}(X_1 = 0), \dots, \mathbb{P}(X_1 = N)] \in \mathbb{R}^{N+1}$$

é o vetor, ou linha, distribuição de X_1 .

Assim, conforme notado, a matriz de transição P é uma forma conveniente de guardar $\mathbb{P}(X_{n+1} = j|X_n = i)$, $i, j \in T$, em um vetor de dados.

Mas é muito mais que isso. Suponha, por exemplo, que o interesse esteja em uma probabilidade de transição de dois passos

$$\mathbb{P}(X_{n+2} = j|X_n = i)$$

Essa probabilidade não aparece na matriz de transição P , mas pode ser calculada através da análise do primeiro passo, como segue

$$\begin{aligned} \mathbb{P}(X_{n+2} = j|X_n = i) &= \sum_{l \in T} \mathbb{P}(X_{n+2} = j, X_{n+1} = l|X_n = i) \\ &= \sum_{l \in T} \frac{\mathbb{P}(X_{n+2} = j, X_{n+1} = l, X_n = i)}{\mathbb{P}(X_n = i)} \\ &= \sum_{l \in T} \frac{\mathbb{P}(X_{n+2} = j, X_{n+1} = l, X_n = i)}{\mathbb{P}(X_{n+1} = l, X_n = i)} \frac{\mathbb{P}(X_{n+1} = l, X_n = i)}{\mathbb{P}(X_n = i)} \\ &= \sum_{l \in T} \mathbb{P}(X_{n+2} = j|X_{n+1} = l, X_n = i) \mathbb{P}(X_{n+1} = l|X_n = i) \\ &= \sum_{l \in T} \mathbb{P}(X_{n+2} = j|X_{n+1} = l) \mathbb{P}(X_{n+1} = l|X_n = i) \\ &= \sum_{l \in T} P_{i,l} P_{l,j} = [P^2]_{i,j}, \quad i, j \in T \end{aligned}$$

Em geral, para todo $k \in \mathbb{N}$ temos

$$\begin{aligned}
\mathbb{P}(X_{n+k+1} = j | X_n = i) &= \sum_{l \in T} \mathbb{P}(X_{n+k+1} = j, X_{n+k} = l | X_n = i) \\
&= \sum_{l \in T} \frac{\mathbb{P}(X_{n+k+1} = j, X_{n+k} = l, X_n = i)}{\mathbb{P}(X_n = i)} \\
&= \sum_{l \in T} \frac{\mathbb{P}(X_{n+k+1} = j, X_{n+k} = l, X_n = i)}{\mathbb{P}(X_{n+k} = l, X_n = i)} \frac{\mathbb{P}(X_{n+k} = l, X_n = i)}{\mathbb{P}(X_n = i)} \\
&= \sum_{l \in T} \mathbb{P}(X_{n+k+1} = j | X_{n+k} = l, X_n = i) \mathbb{P}(X_{n+k} = l | X_n = i) \\
&= \sum_{l \in T} \mathbb{P}(X_{n+k+1} = j | X_{n+k} = l) \mathbb{P}(X_{n+k} = l | X_n = i) \\
&= \sum_{l \in T} \mathbb{P}(X_{n+k} = l | X_n = i) P_{i,j}
\end{aligned}$$

É possível verificar que a família de matrizes

$$[\mathbb{P}(X_{n+k} = j | X_n = i)]_{i,j \in T}, \quad k \geq 1$$

satisfaz a mesma relação de indução como a matriz P^k , isto é,

$$[P^{k+1}]_{i,j} = \sum_{l \in T} [P^k]_{i,l} P_{l,j},$$

então por indução em $k \geq 0$ a igualdade

$$[\mathbb{P}(X_{n+k} = j | X_n = i)]_{i,j \in T} = [[P^k]_{i,j}]_{i,j \in T} = P^k$$

não apenas para $k = 0$ e $k = 1$, mas para todo $k \in \mathbb{N}$.

A relação de produto de matriz

$$P^{m+n} = P^m P^n$$

que se lê

$$[P^{m+n}]_{i,j} = \sum_{l \in T} [P^m]_{i,l} [P^n]_{l,j}, \quad i, j \in T,$$

que agora pode ser interpretada como

$$\mathbb{P}(X_{n+m} = j | X_0 = i) = \sum_{l \in T} \mathbb{P}(X_m = j | X_0 = l) \mathbb{P}(X_n = l | X_0 = i),$$

$i, j \in T$, que, de acordo com Privault (2013), é chamada de equação de Chapman-Kolmogorov.

Desta forma, para calcular o vetor probabilidade após qualquer número de passos é feito $p^n = Pp^{n-1}$ ou, alternativamente,

$$p^n = Pp^{n-1} = P(Pp^{n-2}) = \dots = \underbrace{PP \dots P}_{n \text{ vezes}} p^0 = P^n p. \quad (1.7)$$

1.3 Análise Assintótica

Considerando a cadeia de Markov $(X_n)_{n \in \mathbb{N}}$, com espaço de estados T e seja $A \subset T$. Estamos interessados no primeiro tempo T_A que a cadeia alcança o subconjunto A , com

$$T_A = \inf\{n \geq 0; X_n \in A\}, \quad (1.8)$$

com $T_A = 0$ se $X_0 \in A$ e $T_A = +\infty$ se $\{n \geq 0; X_n \in A\} \neq \emptyset$, isto é, se $X_n \notin A$ para todo $n \in \mathbb{N}$. o interesse é calcular as probabilidades

$$g_l(k) = \mathbb{P}(X_{T_A} = l | X_0 = k)$$

acertando o conjunto $A \subset T$ através do estado $l \in A$ começando de $k \in T$.

Esse cálculo pode ser realizado através da análise do primeiro passo. Para todo $k \in T$ tem-se $T_A \geq 1$ dado que $X_0 = k$, então pode-se escrever

$$\begin{aligned} g_l(k) &= \mathbb{P}(X_{T_A} = l | X_0 = k) \\ &= \sum_{m \in T} \mathbb{P}(X_{T_A} = l | Z_1 = m) \mathbb{P}(X_1 = m | X_0 = k) \\ &= \sum_{m \in T} P_{k,m} \mathbb{P}(X_{T_A} = l | X_1 = m) \\ &= \sum_{m \in T} P_{k,m} \mathbb{P}(X_{T_A} = l | X_0 = m) \end{aligned}$$

$$= \sum_{m \in T} P_{k,m} g_l(m), \quad k \in S \setminus A, \quad l \in A.$$

ou seja,

$$g_l(k) = \sum_{m \in T} P_{k,m} g_l(m), \quad k \in T, \quad l \in A \quad (1.9)$$

sob condições de contorno, $g_l(k) = \mathbb{P}(X_{T_A} = l | X_0 = k) = 1_{\{k=l\}}$, $k \in A$, $l \in T$, desde que $T_A = 0$ comece sempre com $X_0 \in A$.

A equação 1.9 pode ser reescrita na forma matricial

$$g_l = P g_l, \quad l \in A, \quad (1.10)$$

onde g é um vetor coluna, sob condição de contorno

$$g_l(k) = \mathbb{P}(X_{T_A} = l | X_0 = k) = 1_{\{l\}}(k) = \begin{cases} 1, & k = l \\ 0, & k \neq l \end{cases}$$

para todo $k \in A$ e $l \in T$.

No caso, a matriz de transição P satisfaz

$$P_{k,l} = 1_{\{k=l\}}$$

para todo $k, l \in A$, o conjunto A é dito ser absorvente. E mais, as probabilidades $g_l = \mathbb{P}(X_{T_A} = l | X_0 = k)$ satisfazem a condição abaixo para todo $k \in T$.

$$\begin{aligned} 1 &= \mathbb{P}(T_A = +\infty | X_0 = k) + \sum_{l \in A} \mathbb{P}(X_{T_A} = l | X_0 = k) \\ &= \mathbb{P}(T_A = +\infty | X_0 = k) + \sum_{l \in A} g_l(k) \end{aligned} \quad (1.11)$$

Assuma agora que o espaço de estados é $T = \{0, 1, \dots, N\}$ e a matriz de transição P tem a forma

$$P = \begin{bmatrix} Q & R \\ 0 & I \end{bmatrix}$$

onde Q é uma matriz quadrada $(r+1) \times (r+1)$, R é uma matriz $(r+1) \times (N-r)$ e I_d é a matriz identidade $(N-r) \times (N-r)$, em cada caso os estados em $\{r+1, \dots, N\}$ são absorventes.

Se o conjunto $A := \{r+1, \dots, N\}$ é feito de estados absorventes da cadeia, tem-se as condições de contorno $g_l(m) = 1_{\{m=l\}}$, $0 \leq l \leq N$, $r+1 \leq m \leq N$, então a equação 1.9 pode ser reescrita como

$$\begin{aligned} g_l(k) &= \sum_{m=0}^N P_{k,m} g_l(m) \\ g_l(k) &= \sum_{m=0}^r P_{k,m} g_l(m) + \sum_{m=r+1}^N P_{k,m} g_l(m) \\ g_l(k) &= \sum_{m=0}^r Q_{k,m} g_l(m) + R_{k,l}, \quad 0 \leq k \leq r, \quad r+1 \leq l \leq N. \end{aligned}$$

Geralmente, tem-se interesse no tempo médio que a cadeia leva para atingir o conjunto $A \subset T$ começando pelo estado $k \in T$, ou seja, $h_A(k) := \mathbb{E}[T_A | X_0 = k]$. E claramente tem-se $h_A(k) = 0$, para todo $k \in A$.

E para todo $k \in T \setminus A$, usando a análise do primeiro passo tem-se

$$\begin{aligned} h_A(k) &= \mathbb{E}[T_A | X_0 = k] \\ &= \sum_{l \in T} \mathbb{P}(X_1 = l | X_0 = k) (1 + \mathbb{E}[T_A | X_0 = l]) \\ &= \sum_{l \in T} \mathbb{P}(X_1 = l | X_0 = k) + \sum_{l \in T} \mathbb{P}(X_1 = l | X_0 = k) \mathbb{E}[T_A | X_0 = l] \\ &= 1 + \sum_{l \in T} \mathbb{P}(X_1 = l | X_0 = k) \mathbb{E}[T_A | X_0 = l] \\ &= 1 + \sum_{l \in T} P_{k,l} h_A(l), \quad k \in T \setminus A \end{aligned}$$

ou seja,

$$h_A(k) = 1 + \sum_{l \in T} P_{k,l} h_A(l), \quad k \in T \setminus A, \quad (1.12)$$

Sob as condições de contorno

$$h_A(l) = \mathbb{E}[T_A | X_0 = l], \quad l \in A. \quad (1.13)$$

A condição 1.13 implica que a 1.12 torna-se

$$h_A(k) = 1 + \sum_{l \in T \setminus A} P_{k,l} h_A(l), \quad k \in T \setminus A$$

Essa equação pode ser reescrita na forma de matriz como

$$P = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} + P h_A.$$

Considerando apenas as linhas com índice $k \in T \setminus A$, sob as condições de contorno $h_A(k) = 0, k \in A$.

Quando a matriz de transição P tem a forma 1.3 e $A = \{r+1, \dots, N\}$, a equação 1.12 é reescrita como

$$\begin{aligned} h_A(k) &= 1 + \sum_{l=0}^N P_{k,l} h_A(l) \\ &= 1 + \sum_{l=0}^r P_{k,l} h_A(l) + \sum_{l=r+1}^N P_{k,l} h_A(l) \\ &= 1 + \sum_{l=0}^r P_{k,l} h_A(l), \quad 0 \leq k \leq r, \end{aligned}$$

desde que $h_A(l) = 0, l = r+1, \dots, N$, isto é,

$$h_A(k) = 1 + \sum_{l=0}^r P_{k,l} h_A(l), \quad 0 \leq k \leq r,$$

com $h_A(k) = 0, r+1 \leq k \leq n$.

O que foi descrito acima pode ser generalizado e chegar a uma equação para uma expectativa da forma

$$h_A(k) := \mathbb{E}\left[\sum_{i=0}^{T_A-1} f(X_i) \mid X_0 = k\right], \quad 0 \leq k \leq N,$$

onde $f(\cdot)$ é uma função utilidade.

Temos

$$h_A(k) = f(k) + \sum_{m=0}^r P_{k,m} h_A(m), \quad 0 \leq k \leq r,$$

com

$$h_A(k) = 0, \quad r + 1 \leq k \leq n.$$

Quando f é a função indicadora $f = 1_{\{l\}}$, isto é,

$$f(Z_i) = 1_{\{l\}}(Z_i) = \begin{cases} 1, & Z_i = l \\ 0, & Z_i \neq l \end{cases}$$

A variável $h_A(k)$ produzirá a média do número de visitas ao estado l partindo do estado k antes de entrar no conjunto A . Quando $A = \{m\}$ encontra-se $h_{\{m\}}(k) = 1 + \sum_{\substack{l \in S, \\ l \neq m}} P_{k,l} h_{\{m\}}(l)$, $k \in T \setminus \{m\}$ e $h_{\{m\}}(m) = 0$, ainda de acordo com Privault (2013).

Considere agora o primeiro tempo de retorno T_j^r para o estado $j \in T$, definido por $T_j^r := \inf\{n \leq 1 : X_n = j\}$, com $T_j^r = +\infty$ se $X_n \neq j$ para todo $n \leq 1$.

Em contraste com a definição 1.8 do tempo que bate T_j , o ínfimo é pego aqui para $n \geq 1$ como uma demora de pelo menos um passo do estado inicial para o retorno ao estado j . Mesmo assim, $T_j = T_j^r$ se a cadeia começou em um estado i diferente de j .

Denota por

$$\mu_j(i) = \mathbb{E}[T_j^r | X_0 = i] \geq 1$$

o tempo médio de retorno ao estado $j \in T$ começando pelo estado $i \in T$. O tempo médio de retorno também pode ser calculado pela análise do primeiro passo.

$$\begin{aligned} \mu_j(i) &= \mathbb{E}[T_j^r | X_0 = i] \\ &= 1 \times \mathbb{P}(X_1 = j | X_0 = i) + \sum_{\substack{l \in T, \\ l \neq j}} \mathbb{P}(X_1 = l | X_0 = i) (1 + \mathbb{E}[T_j^r | X_0 = l]) \\ &= P_{i,j} + \sum_{\substack{l \in T, \\ l \neq j}} P_{i,l} (1 + \mu_j(l)) \\ &= P_{i,j} + \sum_{\substack{l \in T, \\ l \neq j}} P_{i,l} + \sum_{\substack{l \in T, \\ l \neq j}} P_{i,l} \mu_j(l) \\ &= \sum_{l \in T} P_{i,l} + \sum_{\substack{l \in T, \\ l \neq j}} P_{i,l} \mu_j(l) \end{aligned}$$

$$= 1 + \sum_{\substack{l \in T, \\ l \neq j}} P_{i,l} \mu_j(l),$$

então

$$\mu_j(i) = 1 + \sum_{\substack{l \in T, \\ l \neq j}} P_{i,l} \mu_j(l), \quad i, j \in T. \quad (1.14)$$

Para calcular a probabilidade de retorno em 1.14 não foi trabalhada qualquer condição de contorno. Enquanto que o tempo de retorno T_i^r está a pelo menos um passo, ou seja, $\mu_i(i) \geq 1$ não pode ser nulo, sempre terá $h_i(i) = 0$, $i \in T$. Por outro lado, por definição tem-se

$$h_i(j) = \mathbb{E}[T_i^r | X_0 = j] = \mathbb{E}[T_i | X_0 = j] = \mu_i(j)$$

para todo $i \neq j$, e para $i = j$ o tempo médio de retorno $\mu_j(j)$ pode ser calculado pelo tempo de alcance $h_j(l)$, $l \neq j$, através da análise do primeiro passo como

$$\begin{aligned} \mu_j(j) &= \sum_{l \in T} P_{j,l} (1 + h_j(l)) \\ &= P_{j,j} + \sum_{l \neq j} P_{j,l} (1 + h_j(l)) \\ &= \sum_{l \in T} P_{j,l} + \sum_{l \neq j} P_{j,l} h_j(l) \\ \mu_j(j) &= 1 + \sum_{l \neq j} P_{j,l} h_j(l), \quad j \in T \end{aligned} \quad (1.15)$$

Por fim, é preferível calcular primeiro o tempo de alcance $h_i(i) = 0$ sob as condições de contorno $h_i(i) = 0$, e então encontrar o tempo de retorno $\mu_i(i)$ para 1.15, $i, j \in T$.

Na sequência considere

$$p_{ij} = \mathbb{P}(T_j^r < \infty | X_0 = i) = \mathbb{P}(X_n = j \text{ para algum } n \geq 1 | X_0 = i)$$

e denote a probabilidade de retorno ao estado j em tempo finito começando do estado i . A probabilidade p_{ii} de retorno ao estado i com um tempo finito começando pelo estado i

pode ser calculada por

$$\begin{aligned}
p_{ii} &= \mathbb{P}(X_n = i \text{ para algum } n \geq 1 | X_0 = i) \\
&= \sum_{n=1}^{\infty} \mathbb{P}(X_n = i, X_{n-1} \neq i, \dots, X_1 \neq i | X_0 = i) \\
&= \sum_{n=1}^{\infty} f_{i,i}^{(n)}
\end{aligned}$$

onde, $f_{i,j}^{(n)} := \mathbb{P}(T_j^r = n | X_0 = i) = \mathbb{P}(X_n = j, X_{n-1} \neq j, \dots, X_1 \neq j | X_0 = i)$, $i, j \in T$ é a distribuição de T_j^r dado que $X_0 = i$, com $f_{i,i}^{(0)} = \mathbb{P}(T_i^r | X_0 = i) = 0$.

Tem-se $f_{i,i}^{(1)} = \mathbb{P}(X_1 = i | X_0 = i) = P_{i,i}$, $i \in T$ e decompondo de acordo com o tempo do primeiro retorno $k \geq 1$, a probabilidade $f_{i,i}^{(k)}$ satisfaz a relação

$$\begin{aligned}
[P^n]_{i,i} &= \mathbb{P}(X_n = i | X_0 = i) \\
&= \sum_{k=1}^n \mathbb{P}(X_k = i, X_{k-1} \neq i, \dots, X_1 \neq i | X_0 = i) \mathbb{P}(X_n = i | X_k = i) \\
&= \sum_{k=1}^n \mathbb{P}(X_k = i, X_{k-1} \neq i, \dots, X_1 \neq i | X_0 = i) \mathbb{P}(X_{n-k} = i | X_0 = i) \\
&= \sum_{k=1}^n f_{i,i}^{(k)} P_{i,i}^{n-k}
\end{aligned}$$

Assim, as probabilidades de retorno p_{ij} serão usadas para calcular a média do número de retornos a um dado estado.

Então, a noção de número médio de retornos será necessária para classificar os estados das cadeias de Markov. Seja $R_i = \sum_{n=1}^{\infty} 1_{\{X_n=i\}}$ o número de visitas ao estado i pela cadeia $(X_n)_{n \in \mathbb{N}}$. Se $R_j = 0$ começando por $X_0 = i$ a cadeia nunca visita o estado j e isso acontece com probabilidade $1 - p_{ij}$, então $\mathbb{P}(R_j = 0 | X_0 = i) = 1 - p_{ij}$.

Por outro lado, quando a cadeia $(X_n)_{n \in \mathbb{N}}$ faz um número $R_j = m \geq 1$ de visitas ao estado j começando pelo estado i , ele faz a primeira visita ao estado j com probabilidade p_{ij} e então faz $m - 1$ retornos ao estado j , cada um com probabilidade p_{jj} .

Depois destas m visitas ele nunca mais retorna ao estado j , e esse evento ocorre com probabilidade $1 - p_{jj}$.

Logo, dado que $\{X_0 = i\}$ tem-se

$$\mathbb{P}(R_j = m|X_0 = i) = \begin{cases} p_{ij} \times (p_{jj})^{m-1} \times (1 - p_{jj}), & m \geq 1, \\ 1 - p_{ij}, & m = 0 \end{cases} \quad (1.16)$$

No caso $i = j$, R_i é simplesmente o número de retornos ao estado i começando do estado i , é encontrada a distribuição geométrica

$$\mathbb{P}(R_i = m|X_0 = i) = (p_{ii})^m \times (1 - p_{ii}), \quad m \geq 0$$

Considerando,

$$\mathbb{P}(R_i < \infty|X_0 = i) = \begin{cases} 0, & \text{se } p_{ii} = 1, \\ 1, & \text{se } p_{ii} < 1. \end{cases} \quad (1.17)$$

e

$$\mathbb{P}(R_i = \infty|X_0 = i) = \begin{cases} 1, & \text{se } p_{ii} = 1, \\ 0, & \text{se } p_{ii} < 1. \end{cases} \quad (1.18)$$

Ainda, caso $p_{jj} = 1$ tem-se

$$\mathbb{P}(R_j = \infty|X_0 = i) = p_{ij}$$

E também, usando a identidade $\sum_{k=1}^{\infty} kr^{k-1} = (1 - r)^{-2}$ tem-se

$$\begin{aligned} \mathbb{E}[R_j|X_0 = i] &= \sum_{m=0}^{\infty} m\mathbb{P}(R_j = m|X_0 = i) \\ &= (1 - p_{jj})p_{ij} \sum_{m=1}^{\infty} m(p_{jj})^{m-1} = \frac{p_{ij}}{1 - p_{jj}} \end{aligned} \quad (1.19)$$

ao qual é finita se $p_{jj} < 1$.

Um estado $j \in T$ está acessível par $i \in T$ e escreve-se $(i) \rightarrow (j)$ se existir um inteiro finito $n \geq 0$ tal que

$$[P^n]_{i,j} = \mathbb{P}(X_n = j|X_0 = i) > 0.$$

Em outras palavras, é possível ir de i para j com probabilidade não-nula em um certo número de passos aleatórios. Além disso, desde que $P^0 = I_d$, a definição de acessibilidade dos estados i e j é assegurado para todo $i \in T$, exceto se $P_{ii} = 0$.

No caso de $i \rightarrow j$ e $j \rightarrow i$ diz-se que i e j se comunicam entre si e resume-se como $i \leftrightarrow j$.

A relação binária " \leftrightarrow " satisfaz as seguintes propriedades:

- Reflexiva - para todo $i \in T$ tem-se $i \leftrightarrow i$;
- Simetria - para todo $i, j \in T$ tem-se que $i \leftrightarrow j$ é equivalente a $j \leftrightarrow i$;
- Transitiva - para todo $i, j, k \in T$ tal que $i \leftrightarrow j$ e $j \leftrightarrow k$, tem-se que $i \leftrightarrow k$.

Assim, " \leftrightarrow " é chamada relação de equivalência e isso induz uma partição de T em subconjuntos disjuntos A_1, \dots, A_m tal que $T = A_1 \cup \dots \cup A_m$, e

- tem-se $i \leftrightarrow k$ para todo $i, k \in A_q$; e
- tem-se $i \not\leftrightarrow k$ sempre que $i \in A_p$ e $k \in A_q$ com $p \neq q$.

Os conjuntos A_1, \dots, A_m são classes comunicantes da cadeia.

Definição 3. *Uma cadeia de Markov cujo espaço de estados é formado de uma única classe comunicante é dito ser irredutível, do contrário a cadeia é dita redutível.*

Todos os estados em T se comunicam quando $(X_n)_{n \in \mathbb{N}}$ é irredutível. No caso, a i -ésima coluna da matriz de transição P desaparece, isto é, $P_{k,i} = 0$, $i \in T$, então o estado i não pode ser alcançado por qualquer outro estado e i torna-se uma classe comunicante dele mesmo. Isso também é verdade para estados absorventes. Contudo, tendo um retorno com probabilidade estritamente inferior a um não é suficiente para tornar um dado estado sozinho em uma classe comunicante.

Definição 4. *Um estado $i \in T$ é recorrente se, começando pelo estado i a cadeia retornará a esse estado dentro de um tempo finito, com probabilidade 1.*

Em outras palavras, o estado $i \in T$ é recorrente se

$$p_{i,i} := \mathbb{P}(T_i^r | X_0 = i) = \mathbb{P}(X_n = i \text{ para algum } n \geq 1 | X_0 = i) = 1 \quad (1.20)$$

A partir de 1.19 e 1.20 que o estado i é recorrente se, e somente se

$$\mathbb{E}[R_i|X_0 = i] = \infty \quad (1.21)$$

Além disso, por 1.18 e 1.20 o estado i é recorrente, se e somente se $\mathbb{P}(R_i = +\infty|X_0 = i) = 1$.

Teorema 1. *Um estado $i \in T$ é recorrente se, e somente se*

$$\sum_{n=1}^{\infty} [P^n]_{i,i} = +\infty$$

isto é, a série acima diverge.

O teorema 1 admite o seguinte corolário que mostra que se um estado $j \in T$ que se comunica com um estado recorrente i , então j também é recorrente.

Corolário 1. *Seja $i \in T$ um estado recorrente, então qualquer estado $j \in T$ que se comunica com i é recorrente.*

Logo, uma classe comunicante $A \subset T$ é recorrente se qualquer desses estados for recorrente.

Um estado $i \in T$ é dito transiente quando não é recorrente, isto é, $\mathbb{P}(R_i = +\infty|X_0 = i) < 1$, que é equivalente a $\mathbb{P}(R_i = +\infty|X_0 = i) = 0$, por 1.18. Também pode-se escrever que $i \in T$ é transiente quando $\mathbb{P}(R_i < \infty|X_0 = i) > 0$, que é equivalente a $\mathbb{P}(R_i < \infty|X_0 = i) = 1$ por 1.17, isto é, o número de retornos ao estado $i \in T$ é finito com uma probabilidade não-nula que é necessariamente igual a um.

Em outras palavras, pelo Teorema 1.20, o estado $i \in T$ é transiente se, e somente se

$$p_{i,i} = \mathbb{P}(T_i^r = \infty|X_0 = i) = \mathbb{P}(X_n = i \text{ para algum } n \geq 1|X_0 = i) < 1, \quad (1.22)$$

ou

$$\mathbb{P}(T_i^r = \infty|X_0 = i) > 0.$$

De novo, por 1.19 e 1.20 ou diretamente por 1.21, um estado i é transiente se, e somente se

$$\mathbb{E}[R_i|X_0 = i] < \infty$$

Como uma consequência do Teorema 1, um estado $i \in T$ é transiente se, e somente se

$$\sum_{n=1}^{\infty} [P^n]_{i,i} < \infty,$$

isto é, a série acima converge.

Além disso, como consequência do Corolário 5, se um estado $i \in T$ se comunica com um estado transiente j então i também é transiente, de outra forma o estado j seria recorrente pelo mesmo corolário.

Teorema 2. *Seja $(X_n)_{n \in \mathbb{N}}$ uma cadeia de Markov com espaço de estados finito T . Então $(X_n)_{n \in \mathbb{N}}$ possui pelo menos um estado recorrente.*

O tempo esperado de retorno ao estado $i \in T$, também chamado tempo médio de recorrência de i , é dado por

$$\begin{aligned} \mu_i(i) &:= \mathbb{E}[T_i^r | X_0 = i] \\ &= \sum_{n=1}^{\infty} n \mathbb{P}(T_i^r = n | X_0 = i) \\ &= \sum_{n=1}^{\infty} n f_{i,i}^{(n)}. \end{aligned}$$

Definição 5. *Um estado recorrente $i \in T$ é recorrente positivo se*

$$\mu_i(i) = \mathbb{E}[T_i^r | X_0 = i] < \infty$$

e recorrente nulo se

$$\mu_i(i) = \mathbb{E}[T_i^r | X_0 = i] = +\infty$$

Como definido anteriormente, quando o estado i é recorrente tem-se $\mathbb{P}(T_i^r < \infty | X_0 = i) = 1$, isto é, o tempo de retorno aleatório T_i^r é quase certamente finito começando pelo estado i , mesmo assim esses não produzem informação na finitude da sua expectativa $\mu_i(i) = \mathbb{E}[T_i^r | X_0 = i]$.

Teorema 3. *Assuma que o espaço de estados T de uma cadeia de Markov $(X_n)_{n \in \mathbb{N}}$ é finito. Então todo estado recorrente em T são recorrentes positivo.*

Como consequência tem-se o seguinte Corolário.

Corolário 2. *Seja $(X_n)_{n \in \mathbb{N}}$ uma cadeia de Markov irredutível com espaço de estados T . Então todo estado de $(X_n)_{n \in \mathbb{N}}$ é recorrente positivo.*

Dado um estado $i \in T$, considere o conjunto de inteiros

$$\{n \geq 1 : [P^n]_{i,i} > 0\}$$

O período do estado $i \in T$ é o máximo divisor comum de

$$\{n \geq 1 : [P^n]_{i,i} > 0\}$$

Um estado com período 1 é aperiódico se $P_{i,i} > 0$. Um estado recorrente $i \in T$ é ergótico se é positivo recorrente e aperiódico.

Se $[P^n]_{i,i} = 0$ para todo $n \geq 1$ então o conjunto $\{n \geq 1 : [P^n]_{i,i} > 0\}$ é vazio e por convenção o período do estado i é definido como 0. Neste caso, o estado i também é transitente.

Vale notar que $\{n \geq 1 : [P^n]_{i,i} > 0\}$ contém dois números distintos que são relativamente primos para cada um, ou seja, o máximo divisor comum é 1, então o estado i é aperiódico. Uma cadeia de Markov é considerada aperiódica quando todos os estados são aperiódicos. Em particular, qualquer estado absorvente é aperiódico e recorrente.

Teorema 4. *Considere uma cadeia de Markov $(X_n)_{n \in \mathbb{N}}$ satisfazendo as seguintes três condições:*

- *recorrente;*
- *aperiódica; e*
- *irredutível.*

Então tem-se

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n = i | X_0 = j) = \frac{1}{\mu_i(i)}, \quad i, j \in T$$

onde

$$\mu_i(i) = \mathbb{E}[T_i^r | X_0 = i] \geq 1$$

é a média do tempo de retorno para o estado $i \in T$.

Esse resultado é consistente no caso de uma cadeia com dois estados.

Uma distribuição de probabilidade em T , isto é, uma família $\Pi = (\Pi_i)_{i \in T}$ em $[0, 1]$ tal que

$$\sum_{i \in T} \Pi_i = 1$$

é dito ser estacionário se, começando em X_0 o tempo 0 com a distribuição $(\Pi_i)_{i \in T}$, acontece que a distribuição de X_1 ainda é $(\Pi_i)_{i \in T}$ no tempo 1.

Em outras palavras, $(\Pi_i)_{i \in T}$ é estacionário para a cadeia de Markov com matriz de transição P , se

$$\mathbb{P}(X_0 = i) := \Pi_i, i \in T,$$

no tempo 0, implica

$$\mathbb{P}(X_1 = i) = \mathbb{P}(X_0 = i) = \Pi_i, i \in T$$

no tempo 1. Isso também na medida que

$$\begin{aligned} \Pi_j &= \mathbb{P}(X_1 = j) = \sum_{i \in T} \mathbb{P}(X_1 = j | X_0 = i) \mathbb{P}(X_0 = i) \\ &= \sum_{i \in T} \Pi_i P_{i,j}, j \in T, \end{aligned}$$

isto é, a distribuição Π é estacionária se, e somente se, ela é invariante pela matriz P , que significa

$$\Pi = \Pi P. \tag{1.23}$$

Em contraste com 1.10, a multiplicação por P em 1.23 está a direita e não a esquerda. Considerando que distribuições estacionárias e limitadas são conceitos bastante diferentes, uma cadeia que se inicia com distribuição estacionária permanecerá com a mesma distribuição qualquer passo de tempo subsequente. Por outro lado, a fim de alcançar a distribuição limitada a cadeia pode começar de qualquer dado inicial para qualquer estado fixado e se converterá a distribuição limitada se esta existir. Mesmo assim, a distribuição limitada e estacionária podem coincidir em algumas situações.

Mais geral, assumindo que X_n esta na distribuição estacionária Π no tempo n ,

tem-se

$$\begin{aligned}\mathbb{P}(X_{n+1} = j) &= \sum_{i \in T} \mathbb{P}(X_{n+1} = j | X_n = i) \mathbb{P}(X_n = i) \\ &= \sum_{i \in T} P_{i,j} \mathbb{P}(X_n = i) = \sum_{i \in T} P_{i,j} \Pi_i \\ &= [\Pi P]_j = \Pi_j, \quad j \in T,\end{aligned}$$

desde que a matriz de transição de $(X_n)_{n \in \mathbb{N}}$ é tempo homogêneo, então

$$\mathbb{P}(X_n = j) = \Pi_j, \quad j \in T \Rightarrow \mathbb{P}(X_{n+1} = j) = \Pi_j, \quad j \in T$$

e por indução em $n \geq 0$

$$\mathbb{P}(X_n = j) = \Pi_j, \quad j \in T, n \geq 1,$$

isto é, a cadeia $(X_n)_{n \in \mathbb{N}}$ permanece na mesma distribuição Π em todos os tempos $n \geq 1$, se tiver começado pela distribuição estacionária Π no tempo $n = 0$.

Propriedade 1. *Assuma que $T = \{0, 1, \dots, N\}$ é finito e que a distribuição*

$$\Pi_j := \lim_{n \rightarrow \infty} \mathbb{P}(X_n = j | X_0 = i), \quad i, j \in T,$$

existe e é independente de $i \in T$, isto é, tem-se

$$\lim_{n \rightarrow \infty} P^n = \begin{bmatrix} \Pi_0 & \dots & \Pi_N \\ \vdots & \ddots & \vdots \\ \Pi_0 & \dots & \Pi_N \end{bmatrix}$$

Então, $\Pi = (\Pi_j)_{j \in \{0, 1, \dots, N\}}$ é uma distribuição estacionária e

$$\Pi = \Pi P, \tag{1.24}$$

isto é, Π é invariante para P .

Em particular, tem-se o seguinte resultado.

Teorema 5. *Assuma que a cadeia de Markov $(X_n)_{n \in \mathbb{N}}$ satisfaz as três condições a seguir:*

- *recorrência positiva;*

- aperiódica; e
- irredutível.

Então a cadeia $(X_n)_{n \in \mathbb{N}}$ admite uma distribuição

$$\Pi_i := \lim_{n \rightarrow \infty} \mathbb{P}(X_n = i | X_0 = j) = \lim_{n \rightarrow \infty} [P^n]_{j,i} = \frac{1}{\mu_i(i)}, \quad i, j \in T,$$

ao qual também forma uma distribuição estacionária $(\Pi_i)_{i \in T} = (\frac{1}{\mu_i(i)})_{i \in T}$, unicamente determinada pela equação $\Pi = \Pi P$.

A partir do Teorema 3 temos o seguinte corolário do Teorema 5

Corolário 3. *Considere uma cadeia de Markov irredutível aperiódica com espaço de estados finito. Então as probabilidades*

$$\Pi_i := \lim_{n \rightarrow \infty} \mathbb{P}(X_n = i | X_0 = j) = \frac{1}{\mu_i(i)}, \quad i, j \in T,$$

forma uma distribuição estacionária a qual é unicamente determinada pela equação

$$\Pi = \Pi P$$

O seguinte teorema dá condições suficientes para a existência de uma distribuição estacionária, sem requerer aperiodicidade ou finitude do espaço de estados. Como notado, a distribuição limitante não pode não existir neste caso.

Teorema 6. *Considere a cadeia de Markov $(X_n)_{n \in \mathbb{N}}$ satisfazendo as seguintes condições:*

- recorrente positiva; e
- irredutível.

Então a probabilidade

$$\Pi_i = \frac{1}{\mu_i(i)}, \quad i \in T,$$

forma uma distribuição estacionária que é unicamente determinada pela equação $\Pi = \Pi P$.

Como uma consequência do Corolário 2 tem-se o teorema acima, que não requer aperiodicidade para a distribuição estacionária existir.

Corolário 4. *Seja $(X_n)_{n \in \mathbb{N}}$ um cadeia de Markov irredutível com espaço de estados finito T . Então as probabilidades*

$$\Pi_i = \frac{1}{\mu_i(i)}, i \in T,$$

formam uma distribuição estacionária que é unicamente determinada pela equação

$$\Pi = \Pi P.$$

Note que a distribuição estacionária de uma cadeia de Markov pode não existir para todos, conforme a observação anterior.

Sob as suposições do Teorema 5, se a distribuição estacionária e limitante existirem então elas são iguais e neste caso nós só precisamos calcular uma delas. O principal problema em algumas situações é que pode existir apenas a distribuição estacionária. Daí, de acordo com o Corolário 4, a distribuição limitante pode não existir se a cadeia não é aperiódica.

Assim, pelo Corolário 3, a distribuição limitante e a distribuição estacionária existem - e coincidem - quando a cadeia é irredutível aperiódica com espaço de estados finito. Quando a cadeia é irredutível é normalmente mais fácil só calcular a distribuição estacionária e dar a distribuição limitante. No caso a cadeia é não irredutível precisa-se tentar dividir, ou separá-la, em subcadeias e considerar seus subproblemas separadamente.

Em resumo, normalmente tenta-se calcular a distribuição estacionária sempre que possível, e a distribuição limitante quando existe. De outra forma nós podemos tentar calcular a distribuição limitante exponenciando a matriz de transição e fazendo o limite, mas isto é normalmente muito mais complicado e feito somente em casos excepcionais.

Resumindo, nota-se pelo Teorema 4 que

- uma cadeia irredutível, recorrente e aperiódica implica na existência de uma distribuição limitante.

Pelo Teorema 5

- uma cadeia irredutível, recorrente positiva e aperiódica implica na existência de uma distribuição limitante a qual é também estacionária.

E pelo Teorema 6

- uma cadeia irreduzível e recorrente positiva implica na existência de uma distribuição estacionária.

Além disso, a distribuição limitante ou estacionária $\Pi = (\Pi_i)_{i \in T}$ satisfaz

$$\Pi_i = \frac{1}{\mu_i(\hat{i})}, \quad i \in T,$$

em todos os casos destacados.

A tabela 1.1 resume as definições introduzidas.

Tabela 1.1: Resumo das propriedades dos estados da Cadeia de Markov	
Propriedade	Definição
Absorvente - estado	$P_{i,i} = 1$
Recorrente - estado	$\mathbb{P}(T_i^r < \infty X_0 = i) = 1$
Transiente - estado	$\mathbb{P}(T_i^r < \infty X_0 = i) < 1$
Recorrente positivo - estado	Recorrente e $\mathbb{E}[T_i^r X_0 = i] < \infty$
Recorrente nulo - estado	Recorrente e $\mathbb{E}[T_i^r X_0 = i] = +\infty$
Aperiódico - estado ou cadeia	Período(T) = 1
Ergódico - estado ou cadeia	Recorrente positivo e aperiódico
Irreduzível - cadeia	Todos os estados são comunicantes
Regular - cadeia	Todos os coeficientes de P^n são > 0 para algum $n \geq 1$.

As restrições na matriz de transição P resultam em várias propriedades das cadeias de Markov.

Se P é uma matriz $n \times n$, dizemos que

- $P \geq 0$ se $p_{ij} \geq 0$ para todo $1 \leq i, j \leq n$;
- $P > 0$ se $P \geq 0$ e pelo menos um dos p_{ij} é positivo;
- $P \gg 0$ se $p_{ij} > 0$ para todo $1 \leq i, j \leq n$.

Quando $P \geq 0$, podemos definir um conjunto $\Delta \subset \mathbb{R}$ dos números λ que têm a seguinte propriedade: existe um vetor $x = (x_1, x_2, \dots, x_n)$ tal que

$$\sum_{i \geq j \geq n} x_j = 1, \quad x > 0 \text{ e } Px \geq \lambda x$$

Propriedade 2. *Seja $\lambda_0 = \sup \Delta$. Então, $\lambda_0 < \infty$. Além disso, se $P \gg 0$, então, $\lambda_0 > 0$.*

Teorema 7 (Frobenius). *Seja $P > 0$ e λ_0 como definido acima*

- λ_0 é um autovalor de P e é possível escolher um autovetor associado x^0 tal que $x^0 > 0$;
- se λ é outro autovalor de P , então, $|\lambda| \geq \lambda_0$.

Corolário 5. Se P é uma matriz de transição de uma cadeia de Markov, $\lambda_0 = 1$.

Propriedade 3. Se λ é um autovalor de uma matriz de transição $n \times n$ P , $|\lambda| \leq 1$. Além disso, existe um autovetor associado ao autovalor $\lambda = 1$ com todas as entradas não negativas.

Para a próxima propriedade é preciso levantar as seguintes hipóteses:

- Primeiramente, supõe-se que exista exatamente um autovalor tal que $|\lambda| = 1$, e portanto, este autovalor é 1.
- A seguir, supõe-se que esse autovalor não é degenerado, o que equivale a dizer que seu autoespaço associado tem dimensão 1.
- Finalmente, toma-se como dado que a matriz de transição P , representando a rede, é diagonalizável, significando que seus autovetores formam uma base.

Propriedade 4. 1 Se a matriz de transição P de uma cadeia de Markov satisfaz as três hipóteses acima, então existe um único vetor Π tal que as entradas $\Pi_i = P(X_n = i)$, $i \in T$ satisfazem

$$\Pi_i \geq 0, \Pi_i = \sum_{j \in T} p_{ij} \Pi_j \text{ e } \sum_{i \in T} \Pi_i = 1$$

Chamaremos o vetor Π o regime estacionário da cadeia de Markov.

2 Independentemente do ponto inicial $p_i^0 = P(X_0 = i)$ (onde $\sum_i p_i^0 = 1$), a distribuição das propriedades $P(X_n = i)$ convergirá ao regime estacionário Π , quando $n \rightarrow \infty$.

Esse ponto será de extrema importância para a aplicação pretendida nos próximos capítulos.

Capítulo 2

Teoria Fuzzy

Questionamentos a respeito de incertezas sempre preocuparam pensadores ao longo do tempo e atualmente a incerteza proveniente da aleatoriedade de eventos está bem desenvolvida e ocupa lugar de destaque nos estudos matemáticos, Barros e Bassanezi (2006). A Física Quântica tem se utilizado das teorias estocásticas e uma das mais difundidas e conhecidas está denominada de "Princípio da Incerteza", devida ao físico W. Heisenberg (1927), que relaciona a posição e a velocidade de uma partícula.

E para descrever certos fenômenos tem-se utilizado graus que representam qualidades ou verdades parciais. É precisamente neste tipo de incerteza que a Lógica Fuzzy tem dado suas principais contribuições. A palavra "fuzzy", de origem inglesa, significa incerto, vago, impreciso, subjetivo, nebuloso ou difuso.

A teoria dos Conjuntos Fuzzy foi introduzida em 1965 pelo matemático Lofti Asker Zadeh com a principal intenção de dar um tratamento matemático a certos termos como aproximadamente, ou "em torno de", dentre outros. E esse seria um primeiro passo para tornar possível a produção de cálculos com informações imprecisas. Além de contrapor modelos determinísticos por modelos mais flexíveis, que contemplem certo grau de incerteza.

2.1 Conjuntos e Eventos Fuzzy

Para obter a formalização matemática de um conjunto fuzzy, Zadeh baseou-se no fato de que qualquer conjunto clássico pode ser caracterizado por sua função característica.

Definição 6 (Zadeh). *Seja U um conjunto e A um subconjunto de U . A função característica de A é dada por*

$$\chi_A(x) = \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{se } x \notin A \end{cases}$$

Desta forma, χ_A é uma função cujo domínio é U e a imagem está contida no conjunto $0, 1$, com $\chi_A(x) = 1$ indicando que o elemento x está em A , enquanto $\chi_A(x) = 0$ indica que x não é elemento de A . Assim, a função característica descreve completamente o conjunto A já que indica quais elementos do conjunto universo U são elementos também de A .

Definição 7. *Seja U um conjunto clássico, um subconjunto fuzzy F de U é caracterizado por uma função*

$$\varphi_F \rightarrow [0, 1]$$

pré-fixada, chamada função de pertinência do subconjunto fuzzy F . O índice F na função de pertinência é usado em analogia a função característica de um subconjunto clássico.

Logo, a noção de subconjunto fuzzy é baseada na extensão da função característica, ou indicadora, de um subconjunto clássico. Para cada função $\varphi_F : \Omega \rightarrow [0, 1]$, considere a família de α -níveis

$$[F]^\alpha = \{\omega \in \Omega / \varphi_F(\omega) \geq \alpha\} \text{ para } 0 \leq \alpha \leq 1.$$

Do ponto de vista da modelagem, a teoria fuzzy veio para descrever a fronteira, enquanto que a probabilidade apresenta uma medida de sua ocorrência. Ou seja, o interessante não é optar por uma teoria ou outra, mas sim combinar as duas com a sobreposição de incertezas.

A noção de eventos fuzzy, introduzida por Zadeh, ilustra o potencial da combinação das duas teorias, a fuzzy, tratando da identificação do evento e a probabilística, lidando com a medida do mesmo.

Definição 8. *Seja (Ω, P, \mathcal{A}) um espaço de probabilidades. Um evento fuzzy em Ω é simplesmente um subconjunto fuzzy de Ω cujos α -níveis estão na σ -álgebra \mathcal{A} .*

Suponha $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$, e que todos os pontos tenham a mesma probabilidade $\frac{1}{n}$. É sabido que se A for um evento com m elementos, então $P(A) = \frac{m}{n}$ define uma probabilidade em Ω .

Dado o evento A , sua função indicadora $\chi_A : \Omega \rightarrow \{0, 1\}$ é conhecida e o número de casos favoráveis é $m = \sum_i \chi_A(\omega_i)$. Portanto

$$P(A) = \frac{\sum_{i=1}^n \chi_A(\omega_i)}{n}$$

Por extensão, se A for fuzzy, e conhecendo-se $\varphi_A : \Omega \rightarrow [0, 1]$, o "número de casos favoráveis" é $m = \sum_{i=1}^n \varphi_A(\omega_i)$, de modo que

$$P(A) = \frac{\sum_{i=1}^n \varphi_A(\omega_i)}{n}$$

Nesse caso, m pode não ser inteiro.

Definição 9. *Seja (Ω, P, \mathcal{A}) um espaço de probabilidades e A um evento fuzzy. Então, sua função de pertinência $\varphi_A : \Omega \rightarrow [0, 1]$ é uma variável e, por definição, $P(A) = E(\varphi_A)$.*

Suponha que X seja uma variável aleatória e A um evento fuzzy real com função de pertinência $\varphi_A : \mathbb{R} \rightarrow [0, 1]$. Então, a probabilidade de A é dada por

$$P(A) = E(\varphi_A(X)) = \sum_{i=1}^n \varphi_A(x_i)P(X = x_i)$$

se X for discreta, e

$$P(A) = E(\varphi_A) = E(\varphi_A(X)) = \int_{\mathbb{R}} \varphi_A(x)f(x)dx = \int_{\text{supp}A} \varphi_A f(x)dx$$

onde f é a função densidade de probabilidades de X .

O conceito de independência entre eventos fuzzy passa necessariamente pelo de probabilidade condicional.

Para o evento clássico, a ocorrência simultânea é dada pela interseção $A \cap B$. Para estender ao caso fuzzy, deve-se passar pela função indicadora $\chi_{A \cap B}$. Para o caso clássico, valem $\chi_{A \cap B}(x) = \chi_A(x) \cdot \chi_B(x)$ e $\chi_{A \cap B}(x) = \chi_A(x) \wedge \chi_B(x)$, de modo que, poderia se adotar tanto "." como "\(\wedge\)" para representar a conjunção.

Para o caso clássico, se A e B forem eventos de Ω com $P(B) > 0$, então

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{E(\chi_A \cdot \chi_B)}{\chi_B}$$

A é dito independente de B se a ocorrência de B não interferir na probabilidade de A , isto é, se $P(A|B) = P(A)$. A notação AB no lugar de $A \cap B$ não é por acaso. Ela é proveniente do fato de optar-se por $\chi_A \chi_B$ para representar $\chi_{A \cap B}$.

Portanto, A independente de B se, e somente, se

$$\frac{E(\chi_A \cdot \chi_B)}{E(\chi_B)} = E(\chi_A \cdot \chi_B) = E(\chi_A)E(\chi_B)$$

ou

$$P(AB) = P(A)P(B)$$

Considere A e B dois eventos fuzzy de \mathbb{R} , com $P(B) > 0$, se, e somente, se $E(\varphi_B) > 0$, a probabilidade condicional de A dado B é definida por

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{E(\varphi_A \cdot \varphi_B)}{E(\varphi_B)}$$

Assim, como no caso clássico, A é dito independente de B se a ocorrência de B não interferir na probabilidade de A , isto é, se $P(A|B) = P(A)$. Assim, A independe de B se, e somente, se

$$\frac{E(\varphi_A \cdot \varphi_B)}{E(\varphi_B)} = E(\varphi_A) \Leftrightarrow E(\varphi_A \cdot \varphi_B) = E(\varphi_A)E(\varphi_B)$$

Portanto, desde que as probabilidades não se anulem, A e B são independentes se, e somente se, as variáveis φ_A e φ_B são não correlacionadas.

Considere uma variável aleatória X e a proposição $P(X \text{ ser } A)$, onde A representa um termo linguístico, por exemplo, alto, baixo, muito alto, muito baixo, modelado por um evento fuzzy real A . Uma variável aleatória fuzzy é simplesmente uma função $X : \Omega \rightarrow \mathcal{F}(\mathbb{R})$, em que Ω é espaço amostral.

E esse conceito estende o de variável aleatória uma vez que $\mathbb{R} \subset \mathcal{F}(\mathbb{R})$.

Outras definições interessantes para o tema são as descritas abaixo, apontadas por Uzun e ErsinKiral (2017).

Definição 10. *Um número fuzzy é um conjunto fuzzy na reta real que satisfaz as condições de normalidade e convexidade.*

Definição 11. Um número fuzzy que é representado por três pontos da forma $F = (a_1, a_2, a_3)$, $(a_1 < a_2 < a_3)$ cuja função é dada por

$$\varphi_F(x) = \begin{cases} \frac{x - a_1}{a_2 - a_1}, & \text{se } a_1 \leq x \leq a_2 \\ \frac{a_3 - x}{a_2 - a_3}, & \text{se } a_2 \leq x \leq a_3 \\ 0, & \text{qualquer outro caso} \end{cases}$$

é chamado de número fuzzy triangular.

2.2 Cadeias de Markov com estados Fuzzy

Considere (Ω, Λ, P) um espaço de probabilidades, onde Ω é o espaço amostral, Λ o σ -álgebra em Ω e P uma medida de probabilidade. Sendo F em Ω é um evento fuzzy e $\mu_F(\omega)$, $\omega \in \Omega$, $\mu_F(\omega) \rightarrow [0, 1]$ uma função característica do evento fuzzy F .

Então, a probabilidade de um evento fuzzy F é definida usando a integral de Lebesgue-Stieljes usadas por Zadeh (1965) como o valor esperado da função característica do evento fuzzy com relação a distribuição de probabilidade P :

$$P(F) = \int_{\Omega} \mu_F(\omega) dP = \mathbb{E}(\mu_F) \quad (2.1)$$

e em um conjunto finito é

$$P(F) = \sum_{\Omega} \mu_F(\omega) p_{\omega} \quad (2.2)$$

E a probabilidade condicional do evento fuzzy F dado o evento fuzzy G é:

$$P(F|G) = \frac{P(F \cap G)}{P(G)}, P(G) > 0 \quad (2.3)$$

Além disso, o produto de dois eventos fuzzy F e G é dado por:

$$F.G \leftrightarrow \mu_{FG} = \mu_F \mu_G \quad (2.4)$$

Seja, $X = \{x_1, x_2, \dots, x_n\}$ um conjunto dado, uma partição fuzzy de X é uma família de subconjuntos fuzzy de X denotada por $F = \{F_1, F_2, \dots, F_N\}$, $\forall i = \{1, 2, \dots, N\}$, $F_i \neq \emptyset$, $F_i \neq X$, com as funções características correspondentes a $\mu_{F_1}, \dots, \mu_{F_r}$, que satis-

fazem

$$\sum_{i=1}^N \mu_{F_i}(x_r) = 1, \forall x_r \in X, r = \{1, \dots, N\}$$

O conceito de partição fuzzy é utilizado principalmente para definir os estados fuzzy para decisões em um processo markoviano (Pardo e de la Fuente (2010)).

Considere os estados de um sistema e defina uma partição fuzzy, de modo que cada subconjunto fuzzy F_i , $i \in \{1, 2, \dots, n\}$ represente um estado ou evento fuzzy na cadeia de Markov inicial.

Definição 12. A probabilidade do estado inicial fuzzy $\mathbb{P}(F_i) = \mathbb{P}(X_0 = F_i)$ é definida usando a probabilidade do evento fuzzy por meio da equação

$$\mathbb{P}(F_i) = \mathbb{P}\{X_0 = F_i\} = \sum_{s=0}^N \mathbb{P}\{X_0 = s\} \mu_{F_i}(s) = \sum_{s=0}^N p_s \mu_{F_i}(s)$$

A partir daí é possível calcular a probabilidade condicional para chegar a matriz de transição.

Propriedade 5. A probabilidade condicional do estado fuzzy F_j dado o estado inicial m , com $j \in \{1, 2, \dots, n\}$ e $m \in \{0, 1, \dots, N\}$, pode ser calculada por

$$\mathbb{P}(F_j|m) = \mathbb{P}\{X_1 = F_j | X_0 = m\} = \sum_{s=0}^N p_{ms} \mu_{F_j}(s)$$

e representa a probabilidade de transição para um estado fuzzy em um passo.

Definição 13. A cadeia de Markov do estado fuzzy final é definida pela matriz

$$\bar{P} = \begin{bmatrix} P(F_1|0) & P(F_2|0) & \dots & P(F_n|0) \\ P(F_1|1) & P(F_2|1) & \dots & P(F_n|1) \\ \vdots & \vdots & \ddots & \vdots \\ P(F_1|N) & P(F_2|N) & \dots & P(F_n|N) \end{bmatrix}$$

Esta matriz resulta na probabilidade de transição do estado inicial m , $m \in \{0, 1, \dots, N\}$, para o estado final F_j , $j \in \{1, 2, \dots, n\}$.

Propriedade 6. A probabilidade condicional do evento fuzzy F_j dado o evento fuzzy F_i , com $i, j \in \{1, \dots, n\}$ é uma função resultante de uma combinação linear de probabilidades

$\mathbb{P}(F_j|m)$ da forma

$$\mathbb{P}(F_j|F_i) = \mathbb{P}\{X_1 = F_j|X_0 = F_i\} = \sum_{m=0}^N \mathbb{P}(F_j|m) \frac{p_m \mu_{F_i}(m)}{\mathbb{P}(F_i)}$$

e representa a probabilidade de transição do estado fuzzy para outro estado fuzzy em um passo.

Definição 14. A cadeia de Markov com estados inicial e final fuzzy é definida pela matriz

$$\tilde{P} = \begin{bmatrix} P(F_1|F_1) & P(F_2|F_1) & \dots & P(F_n|F_1) \\ P(F_1|F_2) & P(F_2|F_2) & \dots & P(F_n|F_2) \\ \vdots & \vdots & \ddots & \vdots \\ P(F_1|F_n) & P(F_2|F_n) & \dots & P(F_n|F_n) \end{bmatrix}$$

Esta matriz resulta na probabilidade de transição de um estado inicial fuzzy F_i para um estado final fuzzy F_j , com $i, j \in \{1, 2, \dots, n\}$.

As matrizes \bar{P} e \tilde{P} são estocásticas, dado que a soma de cada uma de suas linhas é 1:

1. A m -ésima linha da matriz \bar{P} são os termos $\mathbb{P}(F_1|m), \mathbb{P}(F_2|m), \dots, \mathbb{P}(F_n|m)$, então

$$\sum_{j=1}^n \mathbb{P}(F_j|m) = \sum_{j=1}^n \sum_{s=0}^N p_{ms} \mu_{F_j}(s) = \sum_{s=0}^N p_{ms} \left(\sum_{j=1}^n \mu_{F_j}(s) \right) = 1.$$

2. A i -ésima linha da matriz \tilde{P} são os termos $\mathbb{P}(F_1|F_i), \mathbb{P}(F_2|F_i), \dots, \mathbb{P}(F_n|F_i)$, então

$$\begin{aligned} \sum_{j=1}^n \mathbb{P}(F_j|F_i) &= \sum_{j=1}^n \sum_{m=0}^N \mathbb{P}(F_j|m) \frac{p_m \mu_{F_i}(m)}{\mathbb{P}(F_i)} \\ &= \frac{1}{\mathbb{P}(F_i)} \left(\sum_{m=0}^N p_m \mu_{F_i}(m) \left(\sum_{j=1}^n \mathbb{P}(F_j|m) \right) \right) = \sum_{m=0}^N \frac{p_m \mu_{F_i}(m)}{\mathbb{P}(F_i)} = 1 \end{aligned}$$

Finalmente, para realizar os cálculos define-se as matrizes Q e S .

$$Q = \begin{bmatrix} \mu_{F_1}(0) & \mu_{F_2}(0) & \dots & \mu_{F_n}(0) \\ \mu_{F_1}(1) & \mu_{F_2}(1) & \dots & \mu_{F_n}(1) \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{F_1}(N) & \mu_{F_2}(N) & \dots & \mu_{F_n}(N) \end{bmatrix}$$

$$S = \begin{bmatrix} \frac{p_0 \mu_{F_1}(0)}{\mathbb{P}(F_1)} & \frac{p_1 \mu_{F_1}(1)}{\mathbb{P}(F_1)} & \dots & \frac{p_N \mu_{F_1}(N)}{\mathbb{P}(F_1)} \\ \frac{p_0 \mu_{F_2}(0)}{\mathbb{P}(F_2)} & \frac{p_1 \mu_{F_2}(1)}{\mathbb{P}(F_2)} & \dots & \frac{p_N \mu_{F_2}(N)}{\mathbb{P}(F_2)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{p_0 \mu_{F_n}(0)}{\mathbb{P}(F_n)} & \frac{p_1 \mu_{F_n}(1)}{\mathbb{P}(F_n)} & \dots & \frac{p_N \mu_{F_n}(N)}{\mathbb{P}(F_n)} \end{bmatrix}$$

A matriz Q contém os valores da função característica da partição fuzzy que determina os estados fuzzy do sistema $\{F_1, F_2, \dots, F_n\}$.

Com as matrizes Q e S , deriva-se:

$$\bar{P} = PQ$$

$$\tilde{P} = S\bar{P} = SPQ$$

Capítulo 3

O Modelo

Será apresentado a seguir dois modelos para séries temporais baseadas em variação de preços de commodities. Foram escolhidos os produtos Soja e Frango, devido a quantidade de observações disponíveis na página do CEPEA (2019), aos quais pretende-se criar modelos determinísticos e fuzzy a fim de encontrar seus estados estacionários conjunto ao seu tempo médio de retorno.

De modo geral, há três tipos de dados disponíveis para análise; a primeira é quando os dados são observados em diferentes instantes do tempo, pode ser diariamente, semanalmente, mensalmente, trimestralmente e assim por diante, esse tipo é chamado de série temporal; também pode-se analisar o corte transversal que ocorre quando os dados observados foram coletados no mesmo ponto do tempo, são exemplos as pesquisas de opinião e dados como os de censo e, por último, tem-se os dados em painel, neste caso uma unidade em corte transversal é pesquisada ao longo do tempo, por exemplo, o PIB de cada país sul-americano para o período dos últimos vinte anos (Wooldridge (2006)).

Como dito anteriormente, uma série temporal pode ser definida como um conjunto de observações ordenadas em um período de tempo. Os modelos mais utilizados para descrever séries temporais são os processos estocásticos, ou seja, podem ser modelados através de leis probabilísticas. Segundo Ehlers (2007), uma série temporal é estacionária quando se desenvolve aleatoriamente ao redor de uma média constante, refletindo alguma forma de equilíbrio estável.

A maioria dos métodos de previsão analisam valores passados para prever futuros, que até o início do século passado eram realizadas com a extrapolação simples de um valor global, ajustado em função do tempo. Uma importante contribuição para o

desenvolvimento de ferramentas e métodos para análise de séries temporais deve-se aos trabalhos de George Udny Yule (Escócia, 1871 - Cambridge, 1951), em que o autor refere que uma série temporal deveria ser vista simplesmente como a realização de um processo estocástico, assim, ele criou o modelo auto-regressivo (AR), em que o valor previsto depende dos valores passados.

Nos anos seguintes, peritos em estatística realizaram seus trabalhos e estudos com base no comportamento das séries e o quanto ele dependia de modelos lineares e do ruído. No entanto, modelos lineares são pouco eficientes para analisar séries temporais, pois a maioria das séries reais apresentam tendências não lineares.

E com o aparecimento dos computadores se iniciou uma evolução acentuada nos métodos de previsão, principalmente naqueles baseados em simulação, já que a programação ajudou a tornar mais rápido qualquer processo de iteração. O propósito inicial da análise de séries temporais é tirar inferências sobre as propriedades ou características básicas do mecanismo gerador do processo estocástico das observações da série, Granger (1980).

Ao analisar uma série temporal, se espera que exista nela uma causa relacionada com o tempo, alguma influência nos dados que possam continuar influenciando futuramente. Para muitos estudiosos, apenas com a abstração das regularidades presentes nos fenômenos observados nas séries temporais que está a possibilidade de construir um modelo matemático como representação simplificada da realidade.

De acordo com Morettin e Tolo (2006), os principais objetivos de se analisar uma série temporal são os itens a seguir:

- descrever propriedades da série temporal;
- explicação e modelagem de determinada ocorrência, que posteriormente pode ser usada para explicar outra;
- prever valores futuros com base em valores passados;
- possibilidade de controle em processos ou tomadas de decisão, ou seja, medir a qualidade de um processo.

Em sua maioria, as séries temporais são estocásticas e o futuro é apenas parcialmente influenciado por valores passados. Para Morettin e Tolo (2006), uma série

temporal estocástica é estacionária se a sua média, variância e autovariância permanecem as mesmas, mesmo em diferentes fases. Logo, essas características se verificariam em qualquer período, ou seja, não variam com o tempo. Já uma série temporal é dita não estacionária quando a sua média e/ou variância varia com o tempo, nestas séries só é possível determinar uma parte dos dados, pois geram piores previsões que séries estacionárias.

Uma série pode ser estacionária por períodos curtos ou longos de tempo, o que pode indicar alternâncias de níveis ou inclinações. Isso explica porque em muitos casos utilizar base de dados muito grandes não significa resultados melhores, em algumas situações oscilações no regime estacionário dos dados pode provocar uma distorção no resultado final da análise.

Séries temporais que não apresentam comportamento explosivo são chamadas de homogêneas, e de acordo com Ehlers (2007), esse tipo de série pode ser estacionária, flutuando ao redor de um nível, por certo período de tempo, depois mudar de nível e permanecer ao redor de um novo nível e assim por diante.

A maioria dos procedimentos de análise estatística em séries temporais supõe que estas sejam estacionárias, portanto, se os dados originais não formam uma série estacionária será necessário transformá-los. A transformação mais comum consiste em fazer diferenças sucessivas da série original, até se obter uma série estacionária.

Ainda segundo Ehlers (2007), uma propriedade importante dos processos estacionários é que os processos obtidos de combinações lineares de processos estacionários são também estacionários. Ou seja, se uma série temporal $\{z_t\} = z_1, z_2, \dots, z_t$ é estacionária então o processo $w_t = z_t - z_{t-1}$ também é estacionário.

O resultado ao aplicar o operador diferença a uma série z_t com t observações é obter uma nova série com um número de $t - 1$ observações.

A estacionariedade é uma suposição muito forte para um processo estocástico, nela se concentram a ideia de que as leis de probabilidade que atuam no processo não mudam com o tempo, isto é, o processo mantém o equilíbrio estatístico.

3.1 A proposta

Uma série temporal pode ser notada como uma coleção de observações feitas sequencialmente ao longo de um período de tempo. Pode-se dizer que a característica

mais importante e interessante deste tipo de dado é o quanto as observações vizinhas são dependentes umas das outras. Enquanto modelos de regressão conduzem que a ordem das observações é irrelevante para análise, em séries temporais a ordem dos dados é crucial. Essa ideia traz conceitos importantes, como é o caso da sazonalidade, muitas séries temporais exibem comportamentos que tendem a se repetir em determinados t períodos de tempo.

A família de variáveis aleatórias $\{X_{t_n} = \{x_1, \dots, x_n\}\}$, que descreve o estado do sistema em pontos discretos no tempo t , forma um processo estocástico. Como vimos em um capítulo anterior, uma cadeia de Markov se a probabilidade de ocorrência de um estado futuro depender apenas do estado presente, ou seja, se é independente dos eventos passados,

$$p_{ij} = P\{X_{t+1} = j | X_t = i\} \quad (3.1)$$

$$p_{ij}^n = P\{X_{t+n} = j | X_t = i\} \quad (3.2)$$

Em que X_n representa o estado do processo no tempo n e p_{ij}^n como sendo a probabilidade de que um processo passe do estado i para j em n passos no tempo t .

3.2 Aplicação Cadeias de Markov

A base de dados utilizada foi o histórico do indicador de preços da soja ESALQ / BM&F BOVESPA, produto comercializado no porto de Paranaguá e o histórico do indicador de preços da carne de frango congelada negociada no atacado no estado de São Paulo, ambos os indicadores com responsabilidade de execução do Centro de Estudos Avançados em Economia Aplicada (Cepea) que é parte integrante do Departamento de Economia, Administração e Sociologia da Escola Superior de Agricultura Luiz de Queiroz (ESALQ) que integra a Universidade de São Paulo (USP), campus Piracicaba, São Paulo.

O primeiro indicador é uma média aritmética dos preços da soja comercializada no porto de Paranaguá e representa soja brasileira em grão a granel tipo exportação. O preço é o valor em reais e em valor presente pago em negócios realizados de soja ou ofertas de venda/compra relatadas por agente colaborador do Cepea, por saca de 60 Kg, livre de quaisquer encargos, tributários ou não tributários. O segundo são as médias regionais

que incluem o valor mais recente de cada colaborador, reportados nos últimos sete dias úteis.

A figuras 3.1 e 3.2 apresentam gráficos de linhas com a periodicidade diária do indicador, onde a linha é a união dos preços de fechamento de cada dia de negociação. Deve-se observar que muitas séries temporais exibem um comportamento que tende a se repetir a cada determinado período de tempo, ou seja, características sazonais, que neste caso podem ser explicadas por períodos de safra, alteração climática, entre outros fatores.

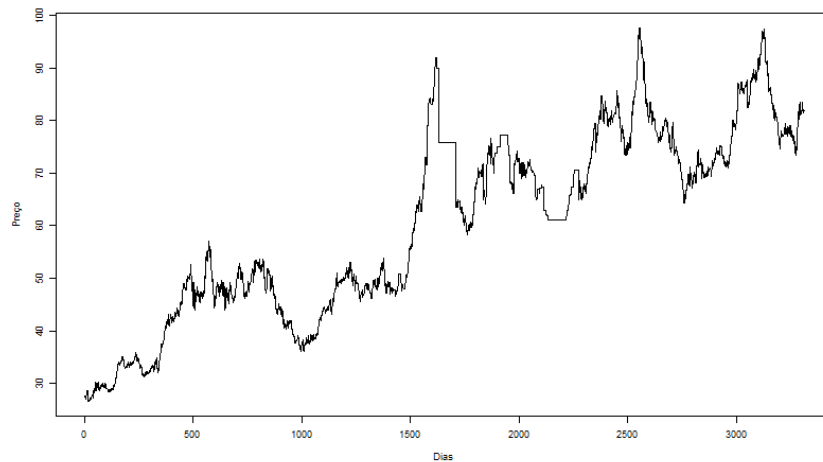


Figura 3.1: Base de dados completa da Soja.

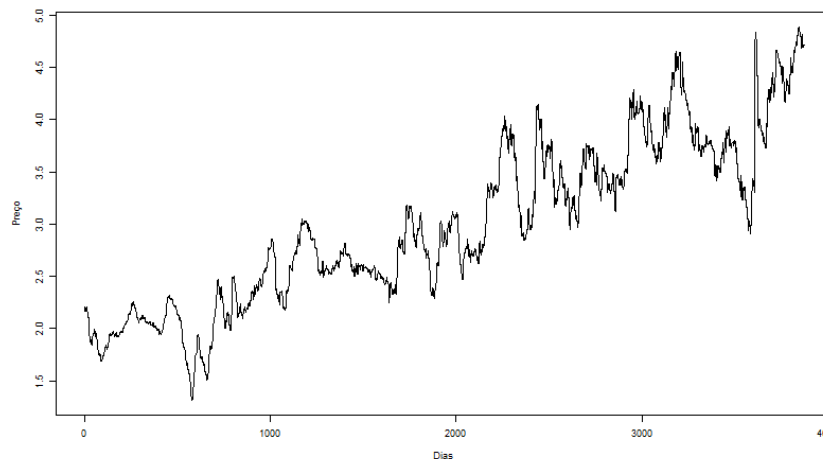


Figura 3.2: Base de dados completa do Frango.

Uma série temporal estacionária é uma série cujas média, variância e função autocorrelação não variam com o tempo, Hill et al. (1999). Essa condição é violada por dados que apresentam tendência ascendente ou descendente ao longo do tempo e

conforme é possível observar graficamente esse é o caso da base de dados em questão, a série apresenta claramente uma tendência positiva, com o que parecem ser picos sazonais. A tabela 3.1 possui a estatística descritiva dos dados.

Tabela 3.1: Estatística descritiva dos dados

Medida	Soja	Frango
Média	60,32	2,994
Moda	61,17	2,6
Mediana	61,98	2,87
Máximo	97,61	4,89
Mínimo	26,57	1,31
Amplitude	71,04	3,58
Variância	310,51	0,628
Desvio Padrão	17,62	0,793
Coefficiente de variação	0,2921	0,265
Tamanho da Amostra	3313	3873

Como dados com tendência são extremamente comuns em economia, encontra-se frequentemente o que as pesquisas bibliográficas chamam de uma série temporal não-estacionária, Hill et al. (1999). Porém, muitas séries temporais não-estacionárias podem ser transformadas em séries temporais estacionárias tomando-se diferenças uma ou mais vezes. Tais séries temporais são chamadas em econometria de processos integrados não-estacionários. E o número d de vezes que um processo integrado deve ser submetido a diferenças para ser tornar estacionário é chamado ordem do processo integrado. Em geral, as diferenças ajudam a eliminar tendências e sazonalidades.

Seguindo o modelo apresentado por Cechin e Corso (2017), foi realizado o cálculo abaixo para a variação dos preços

$$\delta = \frac{P_{t+1} - P_t}{P_t} \quad (3.3)$$

Onde, P_t é o preço no tempo t e P_{t+1} é o preço no tempo $t + 1$.

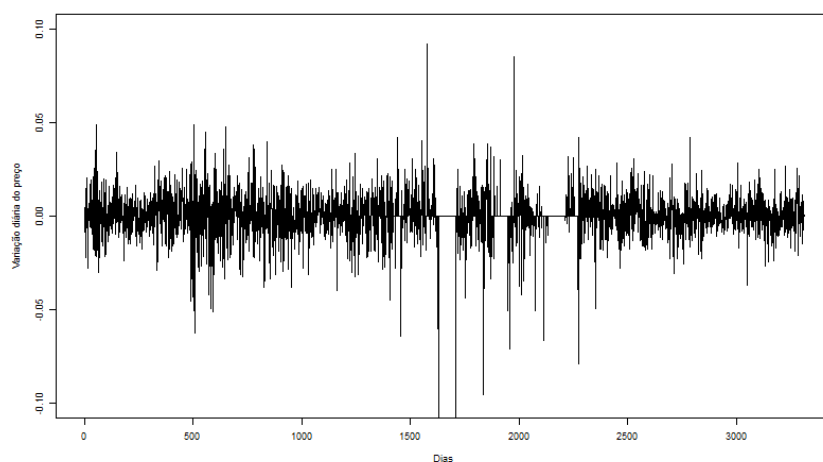


Figura 3.3: Gráfico da Variação do preço da Soja.

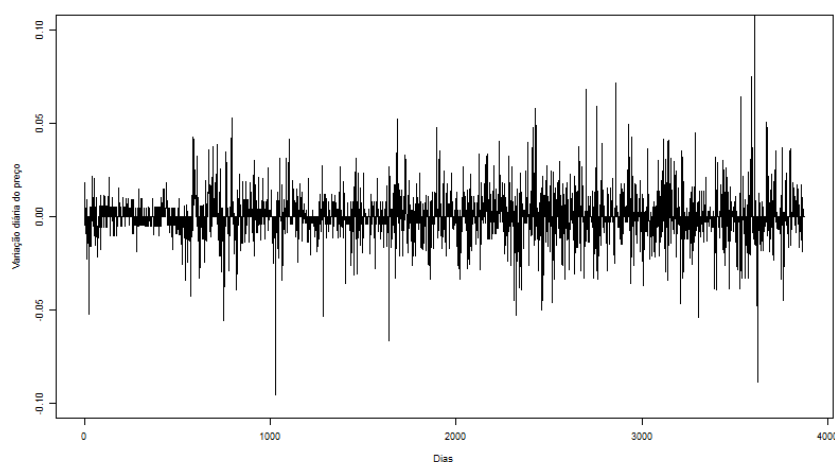


Figura 3.4: Gráfico da Variação do preço da carne de Frango.

Como é possível verificar nos gráficos 3.3 e 3.4, bases de variações possuem mais possibilidades de ser estacionárias, os valores flutuam em torno de um intervalo menor.

Para calcular a matriz de transição, foi definida a tabela abaixo.

Tabela 3.2: Intervalo das variações

Intervalo das variações
menor que -2,01%
de -2 a -1,01%
de -1 a -0,01%
de 0 a 0,99%
de 1 a 1,99%
maior que 2%

Desta forma foi possível analisar a frequência e o percentual acumulado que ocorriam as variações em cada um destes intervalos, apresentados nos histogramas 3.5 e 3.6.

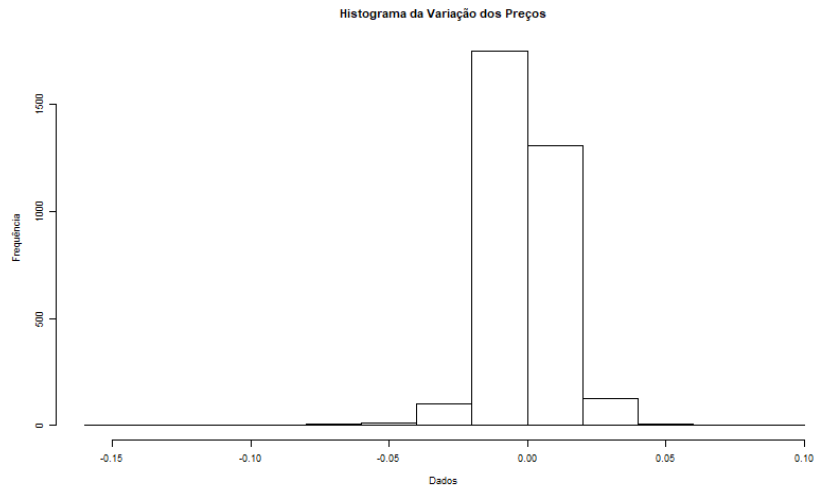


Figura 3.5: Histograma dos intervalos da variação dos preços de Soja.

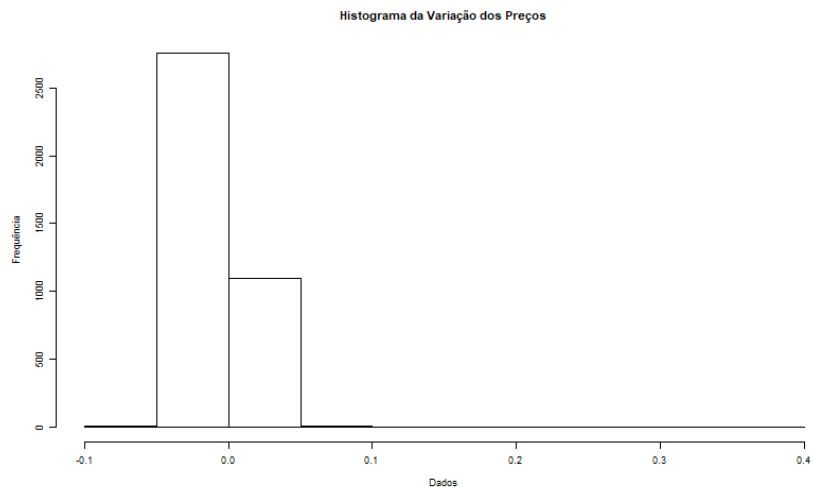


Figura 3.6: Histograma dos intervalos da variação dos preços da carne de Frango.

O teste de autocorrelação foi aplicado as duas bases tratadas nas diferenças e não apresentam correlações significativas em lags superiores ao primeiro, o que indica algum termo de média móvel nos dados que não foi identificado via função de autocorrelação parcial.

Para construir a matriz de transição da cadeia apontada, considere que o valor base x_i pertence a um dos intervalos delimitados na tabela 3.2, sendo eles

- $I_1 =] -\infty; -0,02[$

- $I_2 = [-0, 02; -0, 01[$
- $I_3 = [-0, 01; 0[$
- $I_4 = [0; 0, 01[$
- $I_5 = [0, 01; 0, 02[$
- $I_6 = [0, 02; +\infty[$

Em outras palavras, se $x_i \in (X_n)_{n \in \mathbb{N}}$ então x_i pertence a algum I_t , $t \in \{1, 2, 3, 4, 5, 6\}$. Então, seja $a_{t,q}$ o número de vezes em que a variável $x_i \in I_t$ e $x_{i+1} \in I_q$, com $t, q \in \{1, 2, 3, 4, 5, 6\}$, a partir daí é possível realizar uma contagem de ocorrências na cadeia da seguinte forma:

$$a_{t,q} = \sum_{i=1}^n x_{I_q}(x_i^{k+1}) \text{ onde } (x_i^{k+1}) = \begin{cases} 1, & \text{se } (x_i^{k+1}) \in I_t \\ 0, & \text{se } (x_i^{k+1}) \notin I_t \end{cases}$$

Logo, para calcular cada probabilidade $p_{i,j}$ da matriz de transição usou-se

$$p_{t,q} = \frac{a_{t,q}}{n}$$

onde n é o número total de observações presentes no estado t .

Segue a matriz de transição para o primeiro modelo apresentado.

$$P = \begin{bmatrix} 0,19411765 & 0,17647059 & 0,3294118 & 0,2000000 & 0,05882353 & 0,04117647 \\ 0,08747515 & 0,16302187 & 0,3618290 & 0,2823062 & 0,08349901 & 0,02186879 \\ 0,02668760 & 0,10884354 & 0,4186290 & 0,3490319 & 0,07430665 & 0,02250131 \\ 0,01111648 & 0,06573224 & 0,3301112 & 0,4562591 & 0,10681489 & 0,02996617 \\ 0,01981982 & 0,05765766 & 0,2702703 & 0,4144144 & 0,17657658 & 0,06126126 \\ 0,04324324 & 0,08108108 & 0,2108108 & 0,2864865 & 0,22702703 & 0,15135135 \end{bmatrix}$$

Como o estado estável é calculado usando a equação $\Pi_j = \sum_{i=1}^n \Pi_i p_{ij}$, com $i, j = 1, 2, \dots, n$, e $\sum_{j=1}^n \Pi_j = 1$, tem-se

$$\begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \\ \pi_5 \\ \pi_6 \end{bmatrix} = \begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \\ \pi_5 \\ \pi_6 \end{bmatrix} \begin{bmatrix} 0,1941 & 0,1764 & 0,3294 & 0,2000 & 0,0588 & 0,0411 \\ 0,0874 & 0,1630 & 0,3618 & 0,2823 & 0,0834 & 0,0218 \\ 0,0266 & 0,1088 & 0,4186 & 0,3490 & 0,0743 & 0,0225 \\ 0,0111 & 0,0657 & 0,3301 & 0,4562 & 0,1068 & 0,0299 \\ 0,0198 & 0,0576 & 0,2702 & 0,4144 & 0,1765 & 0,0612 \\ 0,0432 & 0,0810 & 0,2108 & 0,2864 & 0,2270 & 0,1513 \end{bmatrix}$$

e

$$\pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 + \pi_6 = 1 \quad (3.4)$$

Resulta no sistema a ser resolvido

$$\left\{ \begin{array}{l} \pi_1 = 0,1941\pi_1 + 0,0874\pi_2 + 0,0266\pi_3 + 0,0111\pi_4 + 0,0198\pi_5 + 0,0432\pi_6 \\ \pi_2 = 0,1764\pi_1 + 0,1630\pi_2 + 0,1088\pi_3 + 0,0657\pi_4 + 0,0576\pi_5 + 0,0810\pi_6 \\ \pi_3 = 0,3294\pi_1 + 0,3618\pi_2 + 0,4186\pi_3 + 0,3301\pi_4 + 0,2702\pi_5 + 0,2108\pi_6 \\ \pi_4 = 0,2000\pi_1 + 0,2823\pi_2 + 0,3490\pi_3 + 0,4562\pi_4 + 0,4144\pi_5 + 0,2864\pi_6 \\ \pi_5 = 0,0588\pi_1 + 0,0834\pi_2 + 0,0743\pi_3 + 0,1068\pi_4 + 0,1765\pi_5 + 0,2270\pi_6 \\ \pi_6 = 0,0411\pi_1 + 0,0218\pi_2 + 0,0225\pi_3 + 0,0299\pi_4 + 0,0612\pi_5 + 0,1513\pi_6 \\ \pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 + \pi_6 = 1 \end{array} \right.$$

Resolvendo o sistema, tem-se que:

$$\begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \\ \pi_5 \\ \pi_6 \end{bmatrix} = \begin{bmatrix} 0,031461 \\ 0,093207 \\ 0,354087 \\ 0,383811 \\ 0,102923 \\ 0,034511 \end{bmatrix}$$

Resumindo o que foi encontrado

Tabela 3.3: Intervalo das variações - Soja

Intervalo	Probabilidade
menor que -2,01%	3,1461 %
de -2 a -1,01%	9,3207 %
de -1 a -0,01%	35,4087 %
de 0 a 0,99%	38,3811 %
de 1 a 1,99%	10,2923 %
maior que 2%	3,4511 %

É possível agora calcular o tempo de recorrência esperado para cada probabilidade apresentada.

Tabela 3.4: Resultados - Soja

Probabilidade de estado estável π_j	Tempo de recorrência esperado μ_{ii}
π_1	31,79 dias
π_2	10,73 dias
π_3	2,82 dias
π_4	2,62 dias
π_5	9,72 dias
π_6	28,98 dias

Através da tabela 3.4 note que uma variação de 0 à 0,99% pode ser esperada a cada 2,62 dias, enquanto que uma variação maior do que 2% é esperada apenas a cada 28,98 dias.

Segue a matriz de transição para o segundo modelo apresentado.

$$P = \begin{bmatrix} 0.09022556 & 0.12781955 & 0.2406015 & 0.4360902 & 0.06015038 & 0.04511278 \\ 0.05109489 & 0.08394161 & 0.2591241 & 0.5291971 & 0.05109489 & 0.02554745 \\ 0.03689065 & 0.06851120 & 0.2424242 & 0.5797101 & 0.05533597 & 0.01712780 \\ 0.03053097 & 0.07079646 & 0.1814159 & 0.6026549 & 0.07522124 & 0.03938053 \\ 0.02047782 & 0.05119454 & 0.1501706 & 0.5665529 & 0.12286689 & 0.08873720 \\ 0.02614379 & 0.04575163 & 0.1176471 & 0.5882353 & 0.14379085 & 0.07843137 \end{bmatrix}$$

Como o estado estável é calculado usando a equação $\Pi_j = \sum_{i=1}^n \Pi_i p_{ij}$, com $i, j =$

$1, 2, \dots, n$, e $\sum_{j=1}^n \Pi_j = 1$, tem-se

$$\begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \\ \pi_5 \\ \pi_6 \end{bmatrix} = \begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \\ \pi_5 \\ \pi_6 \end{bmatrix} \begin{bmatrix} 0.0902 & 0.1278 & 0.2406 & 0.4361 & 0.0601 & 0.0451 \\ 0.0511 & 0.0839 & 0.2591 & 0.5292 & 0.0511 & 0.0255 \\ 0.0369 & 0.0685 & 0.2424 & 0.5797 & 0.0553 & 0.0171 \\ 0.0305 & 0.0708 & 0.1814 & 0.6026 & 0.0752 & 0.0394 \\ 0.0205 & 0.0512 & 0.1502 & 0.5665 & 0.1229 & 0.0887 \\ 0.0261 & 0.0457 & 0.1176 & 0.5882 & 0.1438 & 0.0784 \end{bmatrix}$$

e

$$\pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 + \pi_6 = 1 \quad (3.5)$$

Resulta no sistema a ser resolvido

$$\left\{ \begin{array}{l} \pi_1 = 0.0902\pi_1 + 0.0511\pi_2 + 0.0369\pi_3 + 0.0305\pi_4 + 0.0205\pi_5 + 0.0261\pi_6 \\ \pi_2 = 0.1278\pi_1 + 0.0839\pi_2 + 0.0685\pi_3 + 0.0708\pi_4 + 0.0511\pi_5 + 0.0457\pi_6 \\ \pi_3 = 0.2406\pi_1 + 0.2591\pi_2 + 0.2424\pi_3 + 0.1814\pi_4 + 0.1502\pi_5 + 0.1176\pi_6 \\ \pi_4 = 0.4361\pi_1 + 0.5292\pi_2 + 0.5797\pi_3 + 0.6026\pi_4 + 0.5665\pi_5 + 0.5882\pi_6 \\ \pi_5 = 0.0601\pi_1 + 0.0511\pi_2 + 0.0553\pi_3 + 0.0752\pi_4 + 0.1229\pi_5 + 0.1438\pi_6 \\ \pi_6 = 0.0451\pi_1 + 0.0255\pi_2 + 0.0171\pi_3 + 0.0394\pi_4 + 0.0887\pi_5 + 0.0784\pi_6 \\ \pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 + \pi_6 = 1 \end{array} \right.$$

Resolvendo o sistema, tem-se que:

$$\begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \\ \pi_5 \\ \pi_6 \end{bmatrix} = \begin{bmatrix} 0.034352 \\ 0.070770 \\ 0.196033 \\ 0.583945 \\ 0.075399 \\ 0.039500 \end{bmatrix}$$

Resumindo o que foi encontrado

Tabela 3.5: Intervalo das variações - Frango

Intervalo	Probabilidade
menor que -2,01%	3,4352 %
de -2 a -1,01%	7,0770 %
de -1 a -0,01%	19,6033 %
de 0 a 0,99%	58,3945 %
de 1 a 1,99%	7,5399 %
maior que 2%	3,95 %

É possível agora calcular o tempo de recorrência esperado para cada probabilidade apresentada.

Tabela 3.6: Resultados - Frango

Probabilidade de estado estável π_j	Tempo de recorrência esperado μ_{ii}
π_1	29,1 dias
π_2	14,1 dias
π_3	5,1 dias
π_4	1,7 dias
π_5	13,3 dias
π_6	25,3 dias

Através da tabela 3.4 note que uma variação de 0 à 0,99% pode ser esperada a cada 1,7 dias, enquanto que uma variação maior do que 2% é esperada apenas a cada 25,3 dias.

3.3 Aplicação Fuzzy

Primeiramente, vale lembrar que enquanto a lógica convencional ou, como alguns autores chamam, clássica usa distinções muito bem definidas para separar conjuntos, a lógica fuzzy tenta modelar conceitos considerados imprecisos.

Em outras palavras, a lógica fuzzy se utiliza da ideia de pertinência gradual, onde considera que um elemento pertence ou não a um conjunto com diferentes graus de pertinência, um exemplo clássico é o conceito de alto ou baixo, gordo ou magro, em que as possibilidades de permanência aos grupos são subjetivas e podem inclusive serem classificadas entre elas mesmas.

Serão usadas as mesmas bases descritas anteriormente para o modelo, ambos acessados da página virtual do CEPEA

- Soja - valores de fechamento diário compreendidos entre 13 de março de 2006 a 17 de julho de 2019;
- Frango - valores de fechamento diário compreendidos entre 08 de dezembro de 2003 à 17 de julho de 2019.

Os retornos foram calculados da forma

$$\delta = \frac{P_{t+1} - P_t}{P_t} \quad (3.6)$$

Agora, associa-se cada mudança de variação de preços em seis estados fuzzy, que serão denotados por $\{F_1, F_2, F_3, F_4, F_5, F_6\}$. Os estados serão classificados como números triangulares fuzzy para obter uma associação dos graus de pertinência do retorno em questão.

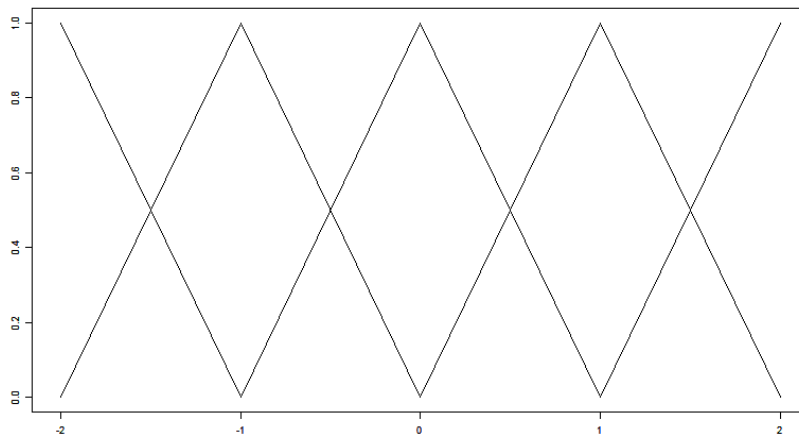


Figura 3.7: Funções de pertinências do sistema fuzzy.

Esses valores fuzzy, ou estados fuzzy, formam uma partição da base de dados. Usando a notação padrão para um número fuzzy de um conjunto discreto $F = \{i, \mu_F(i) | i = 0, \dots, N\}$, tem-se que os valores fuzzy para a base de precificação da Soja são

- $F_1 = (-0,0300 / -0,0200 / -0,0100)$
- $F_2 = (-0,0200 / -0,0100 / 0,0000)$
- $F_3 = (-0,0100 / 0,0000 / 0,0000)$
- $F_4 = (0,0000 / 0,0000 / 0,0100)$

- $F_5 = (0, 0000/0, 0100/0, 0200)$
- $F_6 = (0, 0100/0, 0200/0, 0300)$

Comparado ao modelo de cadeias de Markov, o novo modelo fornece uma descrição mais realista aos estados, pois utilizará probabilidade condicional fuzzy para calcular a matriz de transição. Desta forma, considerando os seis estados do sistema, a matriz de transição para os estados fuzzy é

$$P = \begin{bmatrix} 0.0805 & 0.2308 & 0.2389 & 0.1849 & 0.1911 & 0.0737 \\ 0.0500 & 0.1796 & 0.2616 & 0.2651 & 0.1805 & 0.0633 \\ 0.0439 & 0.1692 & 0.2677 & 0.2795 & 0.1784 & 0.0612 \\ 0.0473 & 0.1544 & 0.2702 & 0.2649 & 0.2098 & 0.0533 \\ 0.0473 & 0.1608 & 0.2578 & 0.2568 & 0.2213 & 0.0559 \\ 0.0529 & 0.1877 & 0.2069 & 0.2269 & 0.2615 & 0.0641 \end{bmatrix}$$

O que resulta em uma distribuição de probabilidade para o retorno da variação do preço, calculada da mesma forma que o estado estacionário, ou seja,

$$\begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \\ \pi_5 \\ \pi_6 \end{bmatrix} = \begin{bmatrix} 0.0488 \\ 0.1695 \\ 0.2603 \\ 0.2610 \\ 0.2011 \\ 0.0592 \end{bmatrix}$$

Que resultará em um tempo médio de retorno, listado na tabela 3.7, onde é possível analisar se os dias de ganho ou perda podem ser classificados entre curto, médio ou longo prazo.

Tabela 3.7: Resultados (Fuzzy) - Soja

Probabilidade de estado estável π_j	Tempo de recorrência esperado μ_{ii}
π_1	20,5 dias
π_2	5,9 dias
π_3	3,84 dias
π_4	3,83 dias
π_5	4,97 dias
π_6	16,9 dias

Para a base de precificação do Frango, os valores fuzzy para a base de precificação são

- $F_1 = (-0,0300 / -0,0200 / -0,0100)$
- $F_2 = (-0,0200 / -0,0100 / 0,0000)$
- $F_3 = (-0,0100 / 0,0000 / 0,0000)$
- $F_4 = (0,0000 / 0,0000 / 0,0100)$
- $F_5 = (0,0000 / 0,0100 / 0,0200)$
- $F_6 = (0,0100 / 0,0200 / 0,0300)$

Considerando os seis estados do sistema, a matriz de transição para os estados fuzzy é

$$P = \begin{bmatrix} 0.0682 & 0.2141 & 0.3336 & 0.2480 & 0.0917 & 0.0445 \\ 0.0426 & 0.1681 & 0.3286 & 0.3012 & 0.1277 & 0.0317 \\ 0.0400 & 0.1606 & 0.3304 & 0.3054 & 0.1332 & 0.0305 \\ 0.0381 & 0.1374 & 0.3136 & 0.3118 & 0.1477 & 0.0514 \\ 0.0362 & 0.1340 & 0.3121 & 0.3048 & 0.1572 & 0.0557 \\ 0.0264 & 0.1145 & 0.2930 & 0.2743 & 0.2175 & 0.0742 \end{bmatrix}$$

O que resulta em uma distribuição de probabilidade para o retorno da variação do preço, calculada da mesma forma que o estado estacionário, ou seja,

$$\begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \\ \pi_5 \\ \pi_6 \end{bmatrix} = \begin{bmatrix} 0.0398 \\ 0.1510 \\ 0.3243 \\ 0.2996 \\ 0.1422 \\ 0.0431 \end{bmatrix}$$

Que resultará em um tempo médio de retorno, listado na tabela 3.8, onde é possível analisar se os dias de ganho ou perda podem ser classificados entre curto, médio ou longo prazo.

Tabela 3.8: Resultados (Fuzzy) - Frango

Probabilidade de estado estável π_j	Tempo de recorrência esperado μ_{ii}
π_1	25, 12 dias
π_2	6, 62 dias
π_3	3, 08 dias
π_4	3, 34 dias
π_5	7, 03 dias
π_6	23, 20 dias

As categorias dos estados e as matrizes de probabilidade de transição condicional foram calculadas via Matlab.

3.4 Resultados

Analisando o comportamento dos dois modelos apresentados é possível verificar que eles geram resultados muito parecidos entre si. Fornecendo informações a especuladores e investidores em potencial sobre oportunidades futuras para estratégias de compra e venda de curto, médio e longo prazo.

Para a soja os dois modelos apresentaram resultados muito próximos, com um tempo menor de retorno no quarto intervalo. Demonstrando a tendência positiva da série e indicando uma oportunidade de ganho a curto prazo.

Quanto ao frango, os modelos não apresentaram resultados tão próximos quanto a soja, enquanto no primeiro modelo a série apresentou seu menor tempo de retorno ao intervalo quatro o segundo mostrou o intervalo três. Apesar de nivelar as variações centrais, uma diferença entre os modelos chama atenção, no primeiro modelo tem-se um tempo de retorno predominantemente menor entre os três últimos intervalos, enquanto que no segundo modelo essa lógica é invertida, comportamento esse que pode ser justificado pelo fato de a lógica fuzzy interpretar melhor incertezas e portanto medir o risco de maneira mais precisa, sendo mais conservadora em seus resultados.

Considerações finais

Este trabalho estudou Cadeias de Markov e Lógica Fuzzy com uma aplicação para analisar variações de preços.

Os cálculos foram realizados nos softwares R Core Team (2017) e Matlab, devido a rapidez e precisão dos resultados. Foram utilizadas bases de precificação das commodities soja e frango congelado disponibilizadas na página virtual do CEPEA - USP.

Tais informações são consideradas vantajosas quanto a estratégias de negócios e critérios de tomada de decisão. Também pode ajudar produtores a realizarem um processo de proteção de custos e preços de forma eficaz.

No caso da soja, os resultados encontrados foram bem parecidos seguindo as duas metodologias. No caso do frango, o resultado aplicando probabilidade fuzzy pareceu mais conservadora, invertendo a posição encontrada utilizando cadeias de markov. Deixando claro que para trabalhos futuros a sugestão seria explorar, estressando mais os modelos e suas diferenças, considerando variações cambiais, inflação e políticas monetárias. De modo a tentar limpar a base de dados e proporcionar novas análises, inclusive em bases de dados que não sejam em mercados à vista, estendendo os métodos a ativos futuros ou negociados em balcão.

Por fim, este trabalho é fruto de uma inspiração presente em uma das disciplinas do Profmat. Claro que a maioria das disciplinas não possuem muitos exemplos de modelos imediatamente aplicáveis, porém todo conhecimento adquirido durante o curso de mestrado resulta certamente em possibilidades reais de criar novos modelos e aplicações dentro ou fora do ambiente escolar.

Referências Bibliográficas

- Barros, L. C. e Bassanezi, R. C. (2006). *Tópicos de lógica Fuzzy e biomatemática*. UNICAMP–IMEEC, Campinas–SP.
- Cechin, R. B. e Corso, L. L. (2017). Previsão da variação do preço da soja, utilizando Cadeia de Markov. *VII Congresso Brasileiro de engenharia de produção*.
- CEPEA (2019). Centro de estudos avançados em economia aplicada. URL: <https://www.cepea.esalq.usp.br> Acesso em: 01/08/2019.
- Ehlers, R. S. (2007). Análise de séries temporais. URL: <http://www.each.usp.br/rvicente/AnaliseDeSeriesTemporais.pdf> Acesso em: 01/02/2019.
- Granger, C. W. J. (1980). Forecasting in business and economics. *Academic Press: New York, San Francisco, London*.
- Hill, C., Griffiths, W., e Judge, G. (1999). *Econometria*. Editora Saraiva, S.Paulo.
- Morettin, P. A. e Toloí, C. M. (2006). *Modelos para previsão de séries Temporais*. Edgard Blucher, S.Paulo.
- Morgado, A. C., Carvalho, J. B. P., de Carvalho, P. C. P., e Fernandez, P. (2016). *Análise Combinatória e Probabilidade*. Sociedade Brasileira de Matemática, Rio de Janeiro–RJ.
- Pardo, M. J. e de la Fuente, D. (2010). Fuzzy Markovian decision processes: Application to queueing systems. *Elsevier*.
- Privault, N. (2013). *Understanding Markov Chains - Examples and Applications*. Springer, Singapore.

- R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rousseau, C. e Saint-Aubin, Y. (2015). *Matemática e Atualidade–Volume 1*. Sociedade Brasileira de Matemática, Rio de Janeiro–RJ.
- Uzun, B. e ErsinKiral (2017). Application of markov chains-fuzzy states to gold price. In *Procedia Computer Science of 9th ICSCCW, Ago 24-25*, volume 120, páginas 365–371, Budapeste, Hungria. Elsevier.
- Viali, L. (2008). Algumas considerações sobre a origem da teoria da probabilidade. *Revista Brasileira de História da Matemática*, 8(16):143–153.
- Wooldridge, J. M. (2006). *Introdução á Econometria*. Thomson, S.Paulo.
- Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8:338–353.

Apêndice: Scripts

R - Matriz de Transição com probabilidade condicional convencional

Para carregar as bases no programa:

```
base <- read.table("FrangoSoja.csv", head=T, sep=":", dec=",", stringsAsFactors = FALSE)
as.data.frame(base)
```

Calculando a base de diferenças:

```
difbase <- array(c(0:0), dim=c(length(base$Reais), 1))
i <- -1
while (i < length(base$Reais))
  difbase[i,1] <- ((base[i+1,1] - base[i,1])/base[i,1])
  i <- i+1
```

Para calcular a matriz de transição

```
while (i < length(difbase))
  difbase[i,1] <- ((base[i+1,1] - base[i,1])/base[i,1])
  if (difbase[i,1] < -0.02)
    a <- a+1
    A[j] <- i+1
    j <- j+1
    if (difbase[i,1] >= -0.02 && difbase[i,1] < -0.01)
      b <- b+1
      B[g] <- i+1
      g <- g+1
```

```

if (difbase[i,1] >= -0.01 && difbase[i,1] < 0)
c < - c+1
C[k] < - i+1
k < - k+1
if (difbase[i,1] >= 0 && difbase[i,1] < 0.01)
d < - d+1
D[l] < - i+1
l < - l+1
if (difbase[i,1] >= 0.01 && difbase[i,1] <= 0.02)
e < - e+1
E[m] < - i+1
m < - m+1
if (difbase[i,1] > 0.02 )
f < - f+1
F[n] < - i+1
n < - n+1
i < - i+1

i < - 1
cont[i] < - c(0,0,0,0,0,0)
while (i<=length(A))
if (difbase[A[i]] < -0.02)
cont[1] < - cont[1]+1
if (difbase[A[i]] >= -0.02 && difbase[A[i]] < -0.01)
cont[2] < - cont[2]+1
if (difbase[A[i]] >= -0.01 && difbase[A[i]] < 0)
cont[3] < - cont[3]+1
if (difbase[A[i]] >= 0 && difbase[A[i]] < 0.01)
cont[4] < - cont[4]+1
if (difbase[A[i]] >= 0.01 && difbase[A[i]] <= 0.02)
cont[5] < - cont[5]+1
if (difbase[A[i]] > 0.02)

```

```
cont[6] <- cont[6]+1
```

```
i <- i+1
```

```
P[1] <- c(cont[1]/a, cont[2]/a, cont[3]/a, cont[4]/a, cont[5]/a, cont[6]/a)
```

O passo acima deve ser repetido para cada intervalo, gerando a matriz de transição.

Cada $P[i]$ é uma linha da cadeia.

```
P <- c(P1, P2, P3, P4, P5, P6)
```

```
t(matrix(P,6,6))
```

A análise de autocorrelação da base foi feita da forma

```
acf(base)
```

```
pacf(base)
```

```
acf(difbase)
```

```
pacf(difbase)
```

```
require(urca)
```

```
lsoj.df <- ur.df(y=difbase, lag=3, type='trend')
```

```
summary(lsoj.df)
```

```
kpss.ur <- ur.kpss(difbase, type='tau', lags="short")
```

```
summary(kpss.ur)
```

Matlab - Matriz de Transição com probabilidade condicional Fuzzy

Para carregar as bases no programa:

```
T = readtable('Sojamod.csv');
```

```
M = table2array(T);
```

```
C0 = [];
```

Pontos do número triangular fuzzy ($a/b/c$)

```
bxp = [-0.02 -0.01 0 0 0.01 0.02];
```

```
bxp = [-0.03 -0.02 -0.01 0 0 0.01 0.02 0.03];
```

```
for i = 1:length(M)-1
```

Varição relativa

```
DP = (M(i+1)-M(i))./M(i)+(1-2*rand(1))*eps;
```

```
C = [];
```


Cálculo da matriz e resolução do sistema de autovetor associado ao autovalor 1.

```

for j = 2:length(bxp)-1
trfn1 = trimf(DP,[bxp(j-1) bxp(j) bxp(j+1)]);
C = [C trfn1];
end
C0 = [C0;C];
end
I = zeros(length(bxp)-2);
for i = 1:(length(bxp)-2)
vaux1 = find(C0(1:end-1,i)~=0);
I(i,:) = sum(C0(vaux1+1,:));
end
TR = diag(1./sum(I,2))*I;
[CA,DA] = eig(TR');
EE = CA(:,1)/sum(CA(:,1))
TEMR = 1./(CA(:,1)/sum(CA(:,1)))
disp(' ')
for i = 0:length(TR)
if(i==0)
str=string();
for j=1:length(TR)
str(j)= ['A' num2str(j)];
end
disp([sprintf('%s ', ' ') sprintf(' %s ',str)])
else
str = string();
for j=1:length(TR)
str(j)= num2str(TR(i,j),'%1.4f');
end
disp([sprintf('%s ',['A' num2str(i)]) sprintf('%s ',str)])
end
end
end

```

```
disp(' ')
for i=1:length(bxp)-2
disp([sprintf('%s%i ', 'A', i) ' => (' num2str(bxp(i), '%1.4f/') num2str(bxp(i+1), '%1.4f/')
num2str(bxp(i+2), '%1.4f/'))])
end
disp(' ')
```