

Universidade Federal de Santa Catarina
Centro de Ciências Físicas e Matemáticas
Departamento de Matemática

Introdução às Variáveis Aleatórias e Cadeias de Markov

Rafael Pedro Mariotto

Orientador: Prof^o. Dr. Luciano Bedin

Florianópolis -SC

Julho de 2009

Rafael Pedro Mariotto

Introdução às Variáveis Aleatórias e Cadeias de Markov

Trabalho acadêmico de graduação
apresentado da disciplina Trabalho de
Conclusão de Curso II, do Curso
de Matemática - Habilitação Licen-
ciatura, do Centro Ciências Físicas e
Matemáticas da Universidade Federal
de Santa Catarina.

Florianópolis -SC

Julho de 2009

Esta monografia foi julgada como **TRABALHO DE CONCLUSÃO DE CURSO** no Curso de Matemática - Habilitação Licenciatura, e aprovada em sua forma final pela Banca Examinadora designada pela Portaria nº 025/CCM/09.



Prof. Nereu Estanislau Burin (UFSC)
Professor da disciplina

Banca Examinadora:



Profº. Dr. Luciano Bedin (UFSC)
Orientador



Profº. Dr. Cleber Bisognin (UFRGS)



Profª. Dra. Silvia Martini de Holanda Janesch (UFSC)

Sumário

Introdução	2
1 Probabilidade	5
1.1 Espaço Amostral e Eventos	5
1.2 Probabilidade	7
1.2.1 Espaços Amostrais Discretos	8
1.2.2 Probabilidade Condicional	8
2 Variáveis Aleatórias	10
2.1 Introdução	10
2.2 Definição de Variável Aleatória	10
2.3 Função de Distribuição Acumulada	13
2.4 Variáveis Aleatórias Contínuas	19
2.5 Variáveis Aleatórias Discretas	19
2.6 Pares de Variáveis Aleatórias	22
2.7 Várias Variáveis Aleatórias	31
3 Noções de Cadeia de Markov	33
3.1 Introdução	33
3.2 Processos Estocásticos	34
3.3 Sequências Aleatórias e Cadeias de Markov	36
3.4 Equações de Chapman-Kolmogorov	42
Considerações Finais	56
Referências Bibliográficas	57

Introdução

O estudo de probabilidades é uma das áreas da matemática cujo conhecimento é relativamente recente. Registros históricos apontam que suas raízes encontram-se numa matemática elementar, a teoria dos jogos de azar. Consta que na França do século XVII, os jogos de azar tornaram-se uma atividade popular muito aceita e que segundo registros, não encontrava restrição perante à lei. Na medida em que os jogadores passaram a usar jogos mais complexos e movimentar grandes quantias em dinheiro, sentiram a necessidade de um método racional para calcular com certa precisão os riscos de cada jogador nos jogos. Chevalier de Mere (1607-1684) movido por esta necessidade consultou o matemático Blaise Pascal(1623-1662), que estabeleceu correspondência a respeito do assunto com outros matemáticos da época, especialmente com Pierre de Fermat (1601-1665). Acredita-se que a origem da teoria das probabilidades está diretamente ligada a tais acontecimentos [8].

Nos dias atuais, movidos pelo desejo de prever lucros ou prejuízos, ganham ênfase estudos relacionados aos eventos cujos resultados dependem de inúmeros fatores. Como exemplo, podemos considerar a quantia que um jogador pode ganhar ou perder em uma seqüência de jogos de azar, ou ainda, a chance de um investidor obter lucros na bolsa de valores, dado que a volatilidade dos negócios o impede utilizar uma estratégia constante que indique qual a melhor aplicação.

Para o estudo dessa teoria, é preciso definir de forma coerente experimentos aleatórios e experimentos determinísticos. Experimentos determinísticos são aqueles que quando repetidos em condições semelhantes conduzem a resultados essencialmente idênticos. A discussão sobre experimentos aleatórios por sua vez, requer uma maior argumentação [13].

Em diferentes campos da atividade prática e científica, encontramos casos onde certos experimentos ou observações podem repetir diversas vezes sob condições uniformes, dando cada observação isolada um resultado anteriormente previsto. Os jogos de azar

proporcionam um exemplo simples de experimentos desse tipo, mas há uma grande quantidade de outros e mais importantes exemplos. Assim, do tipo comum de experimento científico que ocorre na física, na química, na biologia, na medicina, . . . , se procura manter todas as condições pertinentes ao experimento sobre controle, de modo a reproduzir uma repetição o mais uniforme possível. De forma semelhante, na maioria dos ramos da indústria, o processo de produção envolve repetição contínua de certas operações sob condições que, tanto quanto possível, deveriam se manter constantes. Exemplos adicionais se encontram nas estatísticas demográficas e sociais, onde muitas vezes observamos certas propriedades numa seqüência de indivíduos selecionados de maneira uniforme de alguma população dada [8].

Todavia, em tais casos é possível verificar que, embora se tome o máximo de cuidado para manter as condições do experimento tão uniformes quanto possível, aparecerá uma variabilidade intrínseca que não pode ser mantida sob controle. Devido a esta variabilidade, o resultado do experimento variará de forma irregular entre sucessivas repetições, e não se pode prever precisamente o resultado de uma repetição isolada. Em tal caso, dizemos que se trata de um experimento aleatório ou uma observação aleatória. Assim, em qualquer experimento aleatório, devemos esperar que o resultado varie de uma repetição para outra do experimento [19].

Neste trabalho abordaremos uma introdução sobre variáveis aleatórias e cadeias de Markov. O estudo das variáveis aleatórias é de fundamental importância uma vez que nos permite caracterizar experimentos aleatórios de forma numérica, tornando mais simples a descrição de problemas que envolvem um número arbitrário de variáveis. Com as cadeias de Markov podemos estudar processos aleatórios que possuem curta memória, ou seja, situações em que o conhecimento do que ocorreu até o penúltimo estágio, não interfere no resultado do último estágio. Situações com essas características ocorrem em inúmeras aplicações, como por exemplo a análise do nível econômico de um cidadão que depende apenas do nível econômico da geração anterior, ou ainda a situação da ruína de um jogador que disputa uma seqüência de jogos de azar contra um oponente, onde as únicas possibilidades em cada uma das partidas são de vitória ou derrota.

No capítulo que trata de variáveis aleatórias, será definida a função de distribuição de probabilidade para uma variável aleatória, estendendo a definição também para um número arbitrário dessas. Entretanto, aqui o estudo será concentrado no caso

discreto. Depois disso, estudaremos processos estocásticos, seqüências aleatórias e cadeias de Markov. Nesta trabalho o estudo será concentrado nas cadeias de Markov finita, estacionária, irredutível e aperiódica.

A definição de espaço amostral, eventos, bem como os conceitos básicos sobre a teoria das probabilidades serão vistos no primeiro capítulo deste trabalho. É importante salientar que será adotada uma abordagem semi-rigorosa, mesmo por que o estudo rigoroso requer amplo conhecimento sobre a teoria da medida [5]. Não iremos demonstrar as propriedades, proposições e teoremas que serão citados neste capítulo. As demonstrações formais podem ser encontradas nos livros citados nas referências bibliográficas deste trabalho.

Neste trabalho vamos considerar o conjunto dos números naturais \mathbb{N} com o número zero. Quando o mesmo não for utilizado denotaremos por \mathbb{N}^* .

Capítulo 1

Probabilidade

1.1 Espaço Amostral e Eventos

Experimentos Aleatórios

Conforme abordado na introdução do trabalho, experimentos aleatórios são aqueles que, repetidos em idênticas condições, produzem resultados que não podem ser previstos com certeza. Embora não saibamos qual o resultado que irá ocorrer num experimento, em geral conseguimos descrever o conjunto de todos os resultados possíveis que podem ocorrer. As variações dos resultados, de experimento para experimento, são devidas a uma multiplicidade de causas que não se pode controlar, as quais denominada acaso.

Experimentos Determinísticos

Experimentos determinísticos são aqueles que quando repetidos em condições semelhantes conduzem a resultados essencialmente idênticos.

Espaço Amostral

Seja um conjunto Ω formado por os resultados possíveis de um experimento aleatório. Definimos o conjunto Ω como espaço amostral.

Evento

Dado um experimento aleatório cujo espaço amostral é Ω . Chamamos de evento todo subconjunto de Ω . Evento certo é o próprio conjunto Ω , consistindo de todos os resultados possíveis e o evento nulo ou impossível não contém nenhum resultado possível, ou seja, nunca ocorre.

Em geral indica-se um evento por uma letra maiúscula do alfabeto: A, B, C, \dots, X, Y, Z . Os eventos que possuem um único elemento ($\#A = 1$) são chamados de eventos elementares.

União de Eventos

Sejam A e B dois eventos. Então $A \cup B$ é também um evento que ocorre se, e somente se, A ou B (ou ambos) ocorre. Dizemos que $A \cup B$ é a união entre o evento A e o evento B .

Seja A_0, A_1, \dots, A_n uma coleção finita de eventos. Então,

$$\bigcup_{i=0}^n A_i = A_0 \cup A_1 \cup \dots \cup A_n,$$

é também um evento que ocorre se, e somente se, ao menos um dos eventos A_j ocorre. Dizemos que $A_0 \cup A_1 \cup \dots \cup A_n$ é a união finita dos eventos A_0, A_1, \dots, A_n .

Seja A_0, A_1, A_2, \dots uma seqüência de eventos. Então,

$$\bigcup_{i=0}^{+\infty} A_i = A_0 \cup A_1 \cup A_2 \cup \dots,$$

Dizemos que $A_0 \cup A_1 \cup A_2 \cup \dots$ é união infinita dos eventos A_0, A_1, A_2, \dots

Interseção eventos

Sejam A e B dois eventos. Então $A \cap B$ é também um evento que ocorre se, e somente se, A e B ocorrem simultaneamente. Em particular, se $A \cap B = \emptyset$, A e B são chamados de eventos mutuamente exclusivos.

Seja A_0, A_1, \dots, A_n uma coleção finita de eventos. Então,

$$\bigcap_{i=0}^n A_i = A_0 \cap A_1 \cap \dots \cap A_n,$$

é também um evento que ocorre se, e somente se, todos os eventos A_j ocorrem simultaneamente.

Seja $A_0, A_1, A_2 \dots$ uma sequência de eventos. Então,

$$\bigcap_{i=0}^{+\infty} A_i = A_0 \cap A_1 \cap A_2 \dots,$$

é também um evento que ocorre se, e somente se, todos os eventos A_j ocorrem simultaneamente.

Complementar de um evento

Seja A um evento. Então A^C é também um evento que ocorre se, e somente se, A não ocorre.

1.2 Probabilidade

Definição 1.1. O conjunto formado por todos os subconjuntos de um conjunto A é chamado conjunto das partes de A e denotado por $\mathcal{P}(A)$.

Definição 1.2. Seja Ω um espaço amostral. Seja $\mathcal{P}(\Omega)$ o conjunto das partes de Ω . Uma função $P : \mathcal{P}(\Omega) \rightarrow \mathbb{R}$ é chamada de Probabilidade se,

- $0 \leq P[A] \leq 1$, para todo evento $A \in \mathcal{P}(\Omega)$;
- $P[\emptyset] = 0$, $P[\Omega] = 1$;
- Se A_1, A_2, A_3, \dots , é uma seqüência de eventos tais que $A_i \cap A_j = \emptyset$, para todos $i \neq j$, então $P \left[\bigcup_{k=0}^{+\infty} A_k \right] = \sum_{k=0}^{+\infty} P[A_k]$.

Considere dois eventos A e B pertencentes a um mesmo espaço amostral Ω . Então as seguintes propriedades são verdadeiras.

Propriedade 1.1. $P[A^C] = 1 - P[A]$.

Propriedade 1.2. Se $A \subset B$ então $P[A] = P[B] - P[B - A]$.

Propriedade 1.3. Se $A \subset B$ então $P[A] \leq P[B]$.

Propriedade 1.4. $P[A \cup B] = P[A] + P[B] - P[A \cap B]$.

Propriedade 1.5. Se B_0, B_1, \dots são subconjuntos de Ω tais que $B_0 \supseteq B_1 \supseteq B_2 \dots$, então $B = \bigcap_{i=0}^{+\infty} B_i$ satisfaz $P[B] = \lim_{i \rightarrow +\infty} P[B_i]$.

Propriedade 1.6. $\forall n \in \mathbb{N}$, temos

$$\begin{aligned} P[A_0 \cup A_1 \cup \dots \cup A_n] &= P[A_0] + P[A_1] + \dots + P[A_n] + \\ &\quad - P[A_0 \cap A_1] - \dots - P[A_{n-1} \cap A_n] + \\ &\quad + P[A_0 \cap A_1 \cap A_2] + \dots + P[A_{n-2} \cap A_{n-1} \cap A_n] + \\ &\quad + \dots + (-1)^{n-1} P[A_0 \cap A_1 \cap A_2 \cap \dots \cap A_n]. \end{aligned}$$

1.2.1 Espaços Amostrais Discretos

Considere $\Omega = \{a_0, a_1, a_2, \dots, a_n\}$ um espaço amostral finito, onde todos os eventos elementares são mutuamente exclusivos. Assim a probabilidade de qualquer evento $B = \{a'_0, a'_1, a'_2, \dots, a'_m\}$ é dada por $P[B] = P[\{a'_0, a'_1, a'_2, \dots, a'_m\}] = P[\{a'_0\}] + P[\{a'_1\}] + P[\{a'_2\}] + \dots + P[\{a'_m\}]$. Se Ω for enumerável, a probabilidade do evento $C = \{b'_0, b'_1, b'_2, \dots\}$ é dada por $P[C] = P[b'_0] + P[b'_1] + P[b'_2] + \dots$.

Como conseqüência do exposto acima, se o resultados do espaço amostral finito Ω são igualmente prováveis então a probabilidade de um evento é igual ao quociente do número de resultados do evento sob o número total de resultados do espaço amostral.

1.2.2 Probabilidade Condicional

Definição 1.3. Dados dois eventos A e B , a probabilidade condicional de A dado B é,

$$P[A | B] = \frac{P[A \cap B]}{P[B]}.$$

Esta definição só faz sentido no caso de $P[B] > 0$. Se $P[B] = 0$, temos que $P[A | B]$ não é definida.

Analogamente, a probabilidade condicional de B dado A é,

$$P[B | A] = \frac{P[B \cap A]}{P[A]}.$$

Esta definição também só faz sentido no caso de $P[A] > 0$. Se $P[A] = 0$, temos que $P[B | A]$ não é definida.

Propriedade 1.7. Dados A , B e C eventos com $P[B \cap C] > 0$, então,

$$P[A \cap B | C] = P[A | B \cap C]P[B | C].$$

Propriedade 1.8. Se $A_0, A_1, A_2, \dots, A_N$ são eventos independentes tais que $\Omega = \bigcup_{k=0}^N A_k$. Se $P[B] > 0$, então,

$$P[A | B] = \sum_{k=0}^N P[A \cap A_k | B].$$

Propriedade 1.9. Seja A tal que $P[A] > 0$. Então,

$$(a) P[\emptyset | A] = 0, P[\Omega | A] = 1, 0 \leq P[B | A] \leq 1.$$

$$(b) P[(B \cup C) | A] = P[B | A] + P[C | A], \text{ se } [B \cap C] = \emptyset.$$

Definição 1.4. Definimos dois eventos A e B como sendo independentes se

$$P[A | B] = P[A] \text{ ou } P[B | A] = P[B].$$

Podemos ainda escrever as expressões acima em termos da definição de probabilidade condicional 1.3. Se dois eventos A e B são independentes,

$$P[A \cap B] = P[A | B].P[B] = P[A].P[B | A] = P[A].P[B].$$

Dois eventos são independentes quando a ocorrência ou não ocorrência de algum evento B não altera a ocorrência de um evento A , ou analogamente, quando a ocorrência ou não ocorrência de algum evento A não altera a ocorrência de um evento B .

Teorema 1.1 (Teorema do Produto). Se $P[A_0 \cap A_1 \cap \dots \cap A_n] \neq 0$, então,

$$P[A_0 \cap A_1 \cap \dots \cap A_n] = P[A_0]P[A_1 | A_0]P[A_2 | (A_0 \cap A_1)] \dots P[A_n | (A_0 \cap A_1 \cap \dots \cap A_{n-1})].$$

Teorema 1.2 (Teorema da Probabilidade Total). Se B é um evento contido numa união de eventos disjuntos A_0, A_1, \dots, A_n e $P[A_0] > 0, P[A_1] > 0, \dots, P[A_n] > 0$ então,

$$P[B] = P[A_0]P[B | A_0] + P[A_1]P[B | A_1] + \dots + P[A_n]P[B | A_n].$$

Teorema 1.3 (Teorema de Bayes). Se $P[A_0 \cap A_1 \cap \dots \cap A_n] \neq 0$, e $P[B] > 0$, para $i = 0, 1, 2, \dots, n$, então

$$P[A_i | B] = \frac{P[A_i].P[B | A_i]}{P[A_0].P[B | A_0] + \dots + P[A_n].P[B | A_n]}.$$

Capítulo 2

Variáveis Aleatórias

2.1 Introdução

O resultado de uma operação individual de um experimento aleatório pode ser expresso em forma numérica, isto é, declarando os valores supostos no experimento por um certo número de quantidades variáveis. A referida idéia é verdadeira à medida que o experimento se acha diretamente relacionado com contagem ou medição de variáveis quantitativas.

Para exemplificar podemos pensar na experiência do lançamento de um foguete, onde os resultados podem ser denominados “sucesso” ou “fracasso”; a experiência de controle de qualidade em uma determinada empresa, onde os resultados podem ser denominados “bom” ou “defeituoso”; a experiência do lançamento de uma moeda, onde os resultados são denominados “cara” ou “coroa”; ou ainda a experiência de observar o sexo de uma criança recém-nascida, onde os resultados são denominados “menino” ou “menina”. Em todos os casos podemos convencionar a designação de “sucesso”, “bom”, “cara” e “menino” por 1, e “fracasso”, “defeituoso”, “coroa” e “menina” por 0, respectivamente. Observamos claramente que a notação numérica desses experimentos é conveniente e facilita sua descrição e manipulação.

2.2 Definição de Variável Aleatória

Definição 2.1. *Uma variável aleatória num dado espaço amostral Ω é uma função $X : \Omega \rightarrow \mathbb{R}$.*

Exemplo 2.1. Em cada um dos exemplos citados na introdução deste capítulo, o espaço amostral Ω consiste de dois elementos a e b e podemos definir $X : \Omega \rightarrow \mathbb{R}$, $X(a) = 0$ ou $X(b) = 1$:

No caso da experiência de lançamento de um foguete, podemos definir $a = \text{sucesso}$ e $b = \text{fracasso}$, e com isso $X : \Omega \rightarrow \mathbb{R}$, $X(\text{sucesso}) = 0$ ou $X(\text{fracasso}) = 1$.

No caso da experiência de controle de qualidade de uma determinada empresa, podemos definir $a = \text{bom}$ e $b = \text{defeituoso}$, e com isso $X : \Omega \rightarrow \mathbb{R}$, $X(\text{bom}) = 0$ ou $X(\text{defeituoso}) = 1$.

No caso da experiência de lançamento de uma moeda, podemos definir $a = \text{cara}$ e $b = \text{coroa}$, e com isso $X : \Omega \rightarrow \mathbb{R}$, $X(\text{cara}) = 0$ ou $X(\text{coroa}) = 1$.

No caso da experiência de observar o sexo de uma criança recém-nascida, podemos definir $a = \text{menino}$ e $b = \text{menina}$, e com isso $X : \Omega \rightarrow \mathbb{R}$, $X(\text{menino}) = 0$ ou $X(\text{menina}) = 1$.

Exemplo 2.2. Uma moeda “não-viciada” ou “honestas” é arremessada três vezes. Vamos considerar “H” como sendo “cara” e “T” como sendo “coroa”. Assim o espaço amostral Ω é constituído pelo seguinte conjunto de pontos:

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}.$$

Seja X a variável aleatória que determina o número de caras em cada ponto amostral. Então, X tem imagem formada pelo conjunto $\{0, 1, 2, 3\}$. A tabela abaixo relaciona os valores da variável aleatória X para cada $\omega \in \Omega$.

ω	<i>HHH</i>	<i>HHT</i>	<i>HTH</i>	<i>HTT</i>	<i>THH</i>	<i>THT</i>	<i>TTH</i>	<i>TTT</i>
$X(\omega)$	3	2	2	1	2	1	1	0

Tabela 2.1: Valores da variável aleatória X para cada ponto do espaço amostral Ω .

Exemplo 2.3. Um dispositivo mecânico é utilizado para gerar números aleatórios de dois algarismos (de 00 até 99, igualmente prováveis). O espaço amostral é $\Omega = \{00, 01, \dots, 99\}$. Podemos definir uma função $X : \Omega \rightarrow \mathbb{R}$ tal que $X(\omega)$ é produto dos dois algarismos do ponto amostral ω . Essa função é uma variável aleatória. Lembrando que, se $I \subset \mathbb{R}$,

$X^{-1}(I) = \{\omega \in \Omega \mid X(\omega) \in I\}$. Se, por exemplo, procurarmos por todos os possíveis números formados por dois algarismos cujo produto desses resulte em 0 e 16, temos, $X^{-1}(\{0\}) = \{00, 01, 02, 03, 04, 05, 06, 07, 08, 09, 10, 20, 30, 40, 50, 60, 70, 80, 90\}$, e $X^{-1}(\{16\}) = \{28, 82, 44\}$.

Definição 2.2. Se X é uma variável aleatória e x um número real fixado, podemos definir o evento A_x como o subconjunto de Ω consistindo de todos os pontos amostrais ω para os quais a variável aleatória X associa o número x , ou seja, $A_x = \{\omega \in \Omega \mid X(\omega) = x\}$.

Introduzimos aqui a notação $[X = x]$ para o conjunto A_x . Tal notação é estendida para definir outros tipos de eventos em termos de uma variável aleatória. Para números x, a e b fixados, definimos,

$$[X \leq x] = \{\omega \in \Omega \mid X(\omega) \leq x\};$$

$$[X > x] = \{\omega \in \Omega \mid X(\omega) > x\};$$

$$[a < X \leq b] = \{\omega \in \Omega \mid a < X(\omega) \leq b\}.$$

Raciocínio análogo é estendido para eventos da forma,

$$[a < X \leq b], [a \leq X \leq b], \dots$$

Exemplo 2.4. Vamos considerar o Exemplo 2.2. Como o espaço amostral em questão é finito e os eventos são igualmente prováveis, de acordo com 1.2.1 do Capítulo 1 temos,

$$[X = 0] = \{TTT\}, \text{ assim } P[X = 0] = \frac{1}{8};$$

$$[X = 1] = \{HTT, THT, TTH\}, \text{ assim } P[X = 1] = \frac{3}{8};$$

$$[X = 2] = \{HHT, HTH, THH\}, \text{ assim } P[X = 2] = \frac{3}{8};$$

$$[X = 3] = \{HHH\}, \text{ assim } P[X = 3] = \frac{1}{8};$$

$$[X = x] = \emptyset \text{ para } x \neq 0, 1, 2, 3, \text{ assim } P[X = x] = 0.$$

Exemplo 2.5. Consideremos o Exemplo 2.3. Então, por exemplo, $P[X = 16] = \frac{3}{100}$ e

$$P[X \leq 4] = P[X = 0] + P[X = 1] + P[X = 2] + P[X = 3] + P[X = 4],$$

$$P[X \leq 4] = \frac{19}{100} + \frac{1}{100} + \frac{2}{100} + \frac{2}{100} + \frac{3}{100} = \frac{27}{100}.$$

2.3 Função de Distribuição Acumulada

Definição 2.3. A função de distribuição acumulada de probabilidade F associada a uma variável aleatória X é a aplicação $F : \mathbb{R} \rightarrow \mathbb{R}$, tal que

$$F_X(x) = P[X \leq x] = P[\{\omega \in \Omega \mid X(\omega) \leq x\}], \quad \forall x \in \mathbb{R}.$$

Tal função goza das seguintes propriedades:

Propriedade 2.1. $0 \leq F_X(x) \leq 1, \quad \forall x \in \mathbb{R}$.

Demonstração. Óbvio pela Definição 1.2. □

Propriedade 2.2. $F_X(x_1) \leq F_X(x_2), \quad \forall x_1, x_2 \in \mathbb{R}, \text{ com } x_1 \leq x_2$.

Demonstração. Se $x_1 \leq x_2$, então $[X \leq x_1] \subset [X \leq x_2]$, uma vez que quando $X(s) \leq x_1$, necessariamente $X(s) \leq x_2$ e, portanto, qualquer ponto do espaço amostral s em $[X \leq x_1]$ também deve estar em $[X \leq x_2]$. Assim $P[X \leq x_1] \leq P[X \leq x_2]$, conforme a Propriedade 1.3. □

Propriedade 2.3. $\lim_{y \rightarrow x^+} F_X(y) = F_X(x)$.

Demonstração. Seja uma seqüência decrescente de números reais $x_1 > x_2 > x_3 > \dots > x_n > \dots$, convergindo para x ; então, se $I_n = (-\infty, x_n]$, para cada $n = 1, 2, 3, \dots$, e $I = (-\infty, x]$, vemos que as imagens inversas $A_n = X^{-1}(I_n)$, formam uma seqüência monótona não decrescente de eventos, tais que $\bigcap_{n=1}^{+\infty} A_n = A = X^{-1}(I)$. Conforme Propriedade 1.5, temos $\lim_{n \rightarrow +\infty} P(A_n) = P(A)$ isto é, $\lim_{n \rightarrow +\infty} F_X(x_n) = F_X(x)$; como tal se verifica para toda seqüência (x_n) nas condições acima, decorre o resultado. □

Propriedade 2.4. $\lim_{x \rightarrow \infty} F_X(x) = 1, \quad \forall x \in \mathbb{R};$

Demonstração. Vamos primeiro demonstrar que $\lim_{n \rightarrow \infty} F_X(n) = 1$ é verdade $\forall n \in \mathbb{N}$. Sabemos que os eventos da forma $[i - 1 < X \leq i]$ são eventos mutuamente exclusivos $\forall i \in \mathbb{Z}$, e que,

$$\Omega = \dots \cup [-1 < X \leq 0] \cup [0 < X \leq 1] \cup \dots \cup [n - 1 < X \leq n] \cup [n < X \leq n + 1] \cup \dots,$$

$$[X \leq n] = \dots \cup [-1 < X \leq 0] \cup [0 < X \leq 1] \cup \dots \cup [n - 1 < X \leq n].$$

Assim, $\forall n \in \mathbb{N}$, temos

$$\begin{aligned}
1 &= P[\Omega] = \sum_{i=-\infty}^{+\infty} P[i-1 < X \leq i] \\
\lim_{n \rightarrow +\infty} 1 &= \lim_{n \rightarrow +\infty} \sum_{i=-\infty}^n P[i-1 < X \leq i] \\
1 &= \lim_{n \rightarrow +\infty} P[X \leq n] \\
\lim_{n \rightarrow +\infty} F_X(n) &= 1.
\end{aligned}$$

Para mostrar que $\lim_{x \rightarrow +\infty} F_X(x) = 1$, dado $\epsilon > 0$, devemos exibir $N > 0$ tal que se $x > N$, $|F_X(x) - 1| < \epsilon$. Seja $\epsilon > 0$ arbitrário. Como $F_X(n) \rightarrow 1$ quando $n \rightarrow +\infty$, $(F_X(n))$ é uma seqüência de Cauchy, assim $\exists n_1 \in \mathbb{N}$ tal que para todo $n > n_1$, temos $|F_X(n+1) - F_X(n)| < \frac{\epsilon}{2}$. Além disso, $n_2 \in \mathbb{N}$ tal que $n > n_2$, $|F_X(n) - 1| < \frac{\epsilon}{2}$. Escolhemos $N > \max\{n_1, n_2\}$. Então, se $x > N$, $\exists n_0 \in \mathbb{N}$, $n_0 \geq N$ tal que $x \in [n_0, n_{0+1}]$ e assim,

$$\begin{aligned}
|F_X(x) - 1| &= |F_X(x) - F_X(n_0) + F_X(n_0) - 1| \\
&\leq |F_X(x) - F_X(n_0)| + |F_X(n_0) - 1|.
\end{aligned}$$

Mas pela Propriedade 2.2, $F(x)$ é não decrescente, então,

$$|F_X(x) - F_X(n_0)| = F_X(x) - F_X(n_0) \leq F_X(n_0 + 1) - F_X(n_0) = |F_X(n_0 + 1) - F_X(n_0)|,$$

e portanto,

$$|F_X(x) - 1| \leq |F_X(n_0 + 1) - F_X(n_0)| + |F_X(n_0) - 1|.$$

Como $n_0 \geq N > n_1$, temos $|F_X(n_0 + 1) - F_X(n_0)| < \frac{\epsilon}{2}$ e, do fato que $n_0 \geq N > n_2$, temos $|F_X(n_0) - 1| < \frac{\epsilon}{2}$. Dessa forma, se $x > N$, $|F_X(x) - 1| < \epsilon$. \square

Propriedade 2.5. $\lim_{x \rightarrow -\infty} F_X(x) = 0, \forall x \in \mathbb{R}$.

Demonstração. Sabemos que

$$[X > n - 1] = [n - 1 < X \leq n] \cup [n < X \leq n + 1] \cup [n + 1 < X \leq n + 2] \cup \dots,$$

e que,

$$[X > n - 1] = \sum_{i=n}^{+\infty} [i - 1 < X \leq i].$$

Assim,

$$P[X > n - 1] = \sum_{i=n}^{+\infty} P[i - 1 < X \leq i],$$

e,

$$\lim_{n \rightarrow -\infty} P[X > n - 1] = \lim_{n \rightarrow -\infty} \sum_{i=n}^{+\infty} P[i - 1 < X \leq i].$$

Como consequência,

$$\begin{aligned} \lim_{n \rightarrow -\infty} P([X \leq n - 1]^C) &= 1, \\ \lim_{n \rightarrow -\infty} (1 - P[X \leq n - 1]) &= 1, \\ \lim_{n \rightarrow -\infty} 1 - \lim_{n \rightarrow -\infty} P[X \leq n - 1] &= 1, \\ \lim_{n \rightarrow -\infty} P[X \leq n - 1] &= 1 - 1. \end{aligned}$$

Portanto, $\lim_{n \rightarrow -\infty} F_X(n - 1) = 0$, ou seja, $\lim_{n \rightarrow -\infty} F_X(n) = 0$.

Queremos mostrar que $\lim_{x \rightarrow -\infty} F_X(x) = 0$. Seja $\epsilon > 0$, como $F_X(n) \rightarrow -\infty$, existe $N < 0$ tal que $F_X(n) < \epsilon$, sempre que $n < N$, $n \in \mathbb{N}$. Mas F é não decrescente, logo, se $x < N - 1$, $F_X(x) \leq F_X(N - 1)$, ou seja,

$$F_X(x) = F_X(x) - F_X(N - 1) + F_X(N - 1) \leq F_X(N - 1) < \epsilon.$$

□

Exemplo 2.6. A função,

$$F(x) = \begin{cases} 0, & \text{se } x < 0, \\ x + \frac{1}{2}, & \text{se } 0 \leq x \leq \frac{1}{2}, \\ 1, & \text{se } x > \frac{1}{2}, \end{cases}$$

é representada pelo gráfico da Figura 2.1.

Observamos que F_X é uma função que satisfaz simultaneamente todas as propriedades de uma função de distribuição de probabilidade de uma variável aleatória.

Exemplo 2.7. A função,

$$F(x) = \begin{cases} 0, & \text{se } x \leq 0 \\ 1 - e^{-ax}, & \text{se } x \geq 0 \end{cases}$$

é representada pelo gráfico da Figura 2.2.

Tal função ocorre em inúmeras aplicações. Com base em algumas suposições razoáveis concernentes as chamadas telefônicas, a variável aleatória $X =$ “ao tempo entre dois telefonemas de um centro telefônico”, tem a função exponencial negativa expressa nesse exemplo [6].

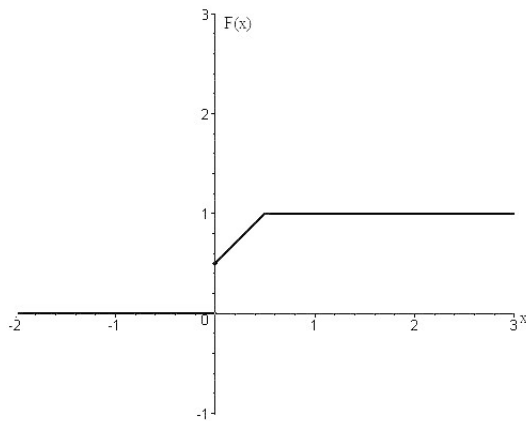


Figura 2.1: Gráfico de F_X .

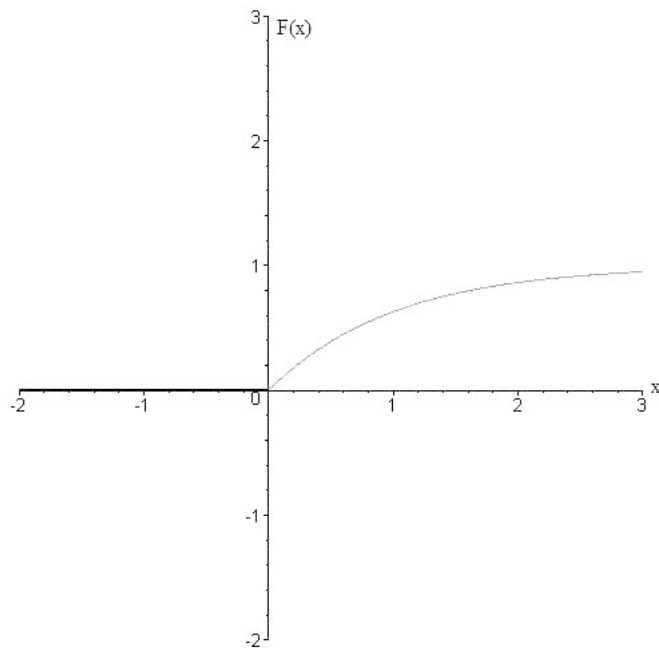


Figura 2.2: Gráfico de F_X .

Exemplo 2.8. Considere o Exemplo 2.2. O gráfico da função de distribuição $F_X(x)$ associada a variável aleatória X é mostrada na Figura 2.3.

Exemplo 2.9. Uma das mais importantes funções de densidade de probabilidade no campo da Estatística é a **função de densidade normal**,

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \forall x \in \mathbb{R}.$$

As constantes μ e σ representam dois parâmetros que devem ser especificados. No caso particular de $\mu = 0$ e $\sigma = 1$, temos a função de densidade normal padronizada, conforme

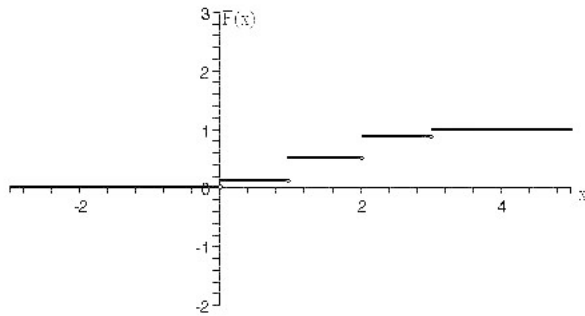


Figura 2.3: Gráfico de $F(x)$.

mostrada no gráfico da Figura 2.4.

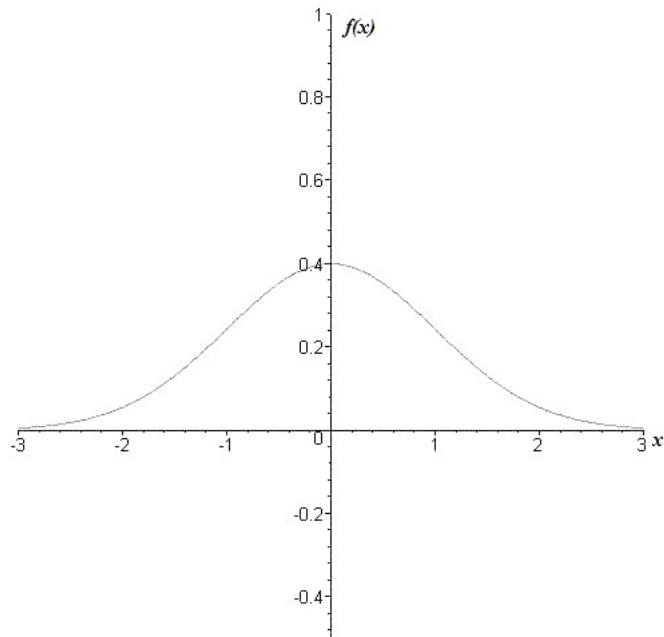


Figura 2.4: Função de densidade de probabilidade normal, com $\mu = 0$ e $\sigma = 1$.

Proposição 2.1. *Sejam $X : \Omega \rightarrow \mathbb{R}$ uma variável aleatória e $a, b \in \mathbb{R}$, então*

1. $P[a < X \leq b] = F_X(b) - F_X(a)$.
2. $P[X > a] = 1 - F_X(a)$.
3. $P[a \leq X \leq b] = P[X = a] + F_X(b) - F_X(a)$.

$$4. P[a < X < b] = F_X(b) - F_X(a) - P[X = b].$$

$$5. P[a \leq X < b] = P[X = a] + F_X(b) - F_X(a) - P[X = b].$$

Demonstração. 1. Sabemos que $[X \leq a] = [X \leq a] \cup [a < X \leq b]$ e ainda que $[X \leq a] \cap [a < X \leq b] = \emptyset$, ou seja, são eventos mutuamente exclusivos. Dessa forma temos,

$$P[X \leq a] = P[X \leq a] + P[a < X \leq b], \text{ ou, } F_X(b) = F_X(a) + P[a < X \leq b]. \text{ Logo,}$$

$$P[a < X \leq b] = F_X(b) - F_X(a).$$

2. Considerando $[X \leq a] \cup [X > a] = \Omega$ e $[X \leq a] \cap [X > a] = \emptyset$, temos,

$$P[X \leq a] + P[X > a] = P[\Omega] = 1, \text{ ou, } F_X(a) + P[X > a] = 1. \text{ Logo,}$$

$$P[X > a] = 1 - F_X(a).$$

3. Seja $[a \leq X \leq b] = [X = a] \cup [a < X \leq b]$ e $[X = a] \cap [a < X \leq b] = \emptyset$. Assim, temos,

$$P[a \leq X \leq b] = P[X = a] + P[a < X \leq b]. \text{ Logo,}$$

$$P[a \leq X \leq b] = P[X = a] + F_X(b) - F_X(a).$$

4. Consideremos $[a < X \leq b] = [a < X < b] \cup [X = b]$ e $[a < X < b] \cap [X = b] = \emptyset$. Então, temos,

$$P[a < X \leq b] = P[a < X < b] + P[X = b], \text{ ou,}$$

$$P[a < X < b] = P[a < X \leq b] - P[X = b]. \text{ Logo,}$$

$$P[a < X < b] = F_X(b) - F_X(a) - P[X = b].$$

5. Suponhamos $[a \leq X < b] = [a \leq X < b] \cup [X = b]$ e $[a \leq X < b] \cap [X = b] = \emptyset$. Dessa forma, temos,

$$P[a \leq X < b] = P[a \leq X \leq b] - P[X = b]. \text{ Logo,}$$

$$P[a \leq X < b] = P[X = a] + F_X(b) - F_X(a) - P[X = b].$$

□

2.4 Variáveis Aleatórias Contínuas

Definição 2.4. Uma variável aleatória é dita contínua se a sua imagem é um subconjunto $\Gamma_X \subset \mathbb{R}$ não enumerável.

2.5 Variáveis Aleatórias Discretas

Definição 2.5. Uma variável aleatória é dita discreta se a sua imagem é um subconjunto $\Gamma_X \subset \mathbb{R}$ no máximo enumerável.

Neste trabalho estudaremos somente as variáveis aleatórias do caso discreto.

Exemplo 2.10. Suponhamos que uma moeda honesta seja arremessada n vezes, e seja X a variável aleatória que representa o número de caras observadas após n arremessos. Então X é uma variável aleatória discreta com imagem igual ao conjunto $\Gamma_X = \{0, 1, 2, 3, \dots, n\}$.

Definição 2.6. Seja X uma variável aleatória discreta. Definimos $p_X : \mathbb{R} \rightarrow \mathbb{R}$ por $p_X = P[X = x]$, como sendo a função de densidade de probabilidade de X .

Como consequência imediata da definição, $0 \leq p_X(x) \leq 1$, $\forall x \in \mathbb{R}$ e se a imagem de X for denotada por Γ_X , se $x \notin \Gamma_X$, $p_X(x) = 0$.

Proposição 2.2. Seja X uma variável aleatória discreta e $\Gamma_X \subset \mathbb{R}$ sua imagem. Então,

$$\sum_{x \in \Gamma_X} p_X(x) = 1.$$

Demonstração. Caso I. Seja Γ_X finito, ou seja, $\Gamma_X = \{x_1, x_2, x_3, \dots, x_n\}$ com $x_1 < x_2 < x_3 < \dots < x_n$. Então,

$$\begin{aligned} \sum_{i=1}^n p_X(x_i) &= p_X(x_1) + p_X(x_2) + \dots + p_X(x_n) \\ &= P[X = x_1] + P[X = x_2] + \dots + P[X = x_n]. \end{aligned}$$

Pela Proposição 2.1, para $i = 1$, com $x < x_1$, temos,

$$\begin{aligned} P[X = x_1] &= F_X(x_1) - F_X(x) - P[x < X < x_1] \\ &= F_X(x_1) - P[X \leq x] - P[\emptyset] \\ &= F_X(x_1) - P[\emptyset] - P[\emptyset] \\ &= F_X(x_1). \end{aligned}$$

Para $i \geq 2$,

$$\begin{aligned}
P[X = x_i] &= F_X(x_i) - F_X(x_{i-1}) - P[x_{i-1} < X < x_i] \\
&= F_X(x_i) - F_X(x_{i-1}) - P[\emptyset] \\
&= F_X(x_i) - F_X(x_{i-1}).
\end{aligned}$$

Logo,

$$\begin{aligned}
\sum_{i=1}^n p_X(x_i) &= F_X(x_1) + F_X(x_2) - F_X(x_1) + F_X(x_3) - F_X(x_2) + \dots + F_X(x_n) - F_X(x_{n-1}) \\
&= F_X(x_n) = P[X \leq x_n] = P[\Omega] = 1.
\end{aligned}$$

Caso II. Seja Γ_X enumerável da forma $\Gamma_X = \{\dots, x_{-2}, x_{-1}, x_0, x_1, x_2, \dots\}$ com $\dots < x_{-2} < x_{-1} < x_0 < x_1 < x_2 < \dots$. Então,

$$\begin{aligned}
\sum_{k=-\infty}^{+\infty} p_X(x_k) &= \lim_{n \rightarrow \infty} \left[\sum_{k=-n}^n p_X(x_k) \right] \\
&= \lim_{n \rightarrow \infty} [p_X(x_{-(n)}) + p_X(x_{-(n-1)}) + p_X(x_{-(n-2)}) + \dots + p_X(x_0) + p_X(x_1) + \\
&\quad + \dots + p_X(x_n)] \\
&= \lim_{n \rightarrow \infty} \{P[X = x_{-(n)}] + P[X = x_{-(n-1)}] + P[X = x_{-(n-2)}] + \dots + \\
&\quad + P[X = x_0] + P[X = x_1] + \dots + P[X = x_n]\}.
\end{aligned}$$

Pela Proposição 2.1, temos,

$$\begin{aligned}
\sum_{k=-\infty}^{+\infty} p_X(x_k) &= \lim_{n \rightarrow +\infty} [F_X(x_{-(n)}) - F_X(x_{-(n+1)}) + F_X(x_{-(n-1)}) - F_X(x_{-(n)}) + \\
&\quad + F_X(x_{-(n-2)}) - F_X(x_{-(n-1)}) + \dots + F_X(x_0) - F_X(x_{-1}) + F_X(x_1) - F_X(x_0) + \\
&\quad + \dots + F_X(x_n) - F_X(x_{(n-1)})] \\
&= \lim_{n \rightarrow \infty} [F_X(x_n) - F_X(x_{-(n+1)})] \\
&= \lim_{n \rightarrow \infty} F_X(x_n) - \lim_{n \rightarrow \infty} F_X(x_{-(n+1)}) \\
&= \lim_{n \rightarrow \infty} F_X(x_n) - \lim_{n \rightarrow -\infty} F_X(x_n) = 1 - 0 = 1.
\end{aligned}$$

No caso de Γ_X possuir um maior ou menor elemento, a prova é análoga. \square

Proposição 2.3. *Seja X uma variável aleatória cuja imagem é $\Gamma_X \subset \mathbb{R}$. Então, $\forall x \in \mathbb{R}$,*

$$F_X(x) = \sum_{\substack{x_i \leq x \\ x_i \in \Gamma_X}} p_X(x_i).$$

Demonstração. Suponhamos que Γ_X é da forma $\Gamma_X = \{\dots, x_{-2}, x_{-1}, x_0, x_1, x_2, \dots\}$. Dado $x \in \mathbb{R}$, $\exists i \in \mathbb{Z}$ tal que $x_i \leq x < x_i + 1$. Então,

$$\begin{aligned} P[X \leq x] &= P\{\dots \cup [x_1 < X \leq x_2] \cup [x_2 < x \leq x_3] \cup \dots \cup [x_{i-1} < X \leq x_i] \cup \\ &\quad \cup [x_i < X \leq x]\} \\ &= \dots P[x_1 < X \leq x_2] + P[x_2 < x \leq x_3] + \dots + P[x_{i-1} < X \leq x_i] + \\ &\quad + P[x_i < X \leq x]. \end{aligned}$$

Caso I: Suponhamos $x = x_i$, para algum $i \in \mathbb{Z}$. Assim,

$$P[x_i < X \leq x] = F_X(x) - F_X(x_i) = F_X(x_i) - F_X(x_i) = 0.$$

Utilizando a Proposição 2.1, temos,

$$\begin{aligned} P[X \leq x] &= \dots + F_X(x_2) - F_X(x_1) + F_X(x_3) - F_X(x_2) + \dots + F_X(x_i) - F_X(x_{i-1}) + 0 \\ &= \dots + P[X = x_2] + P[X = x_3] + \dots + P[X = x_i] \\ &= \dots + p_X(x_2) + p_X(x_3) + \dots + p_X(x_i) \\ F(x) &= \sum_{\substack{x_i \leq x \\ x_i \in \Gamma_X}} p_X(x_i). \end{aligned}$$

Caso II: Suponhamos que $x \neq x_i$, para todo $i \in \mathbb{Z}$. Então, $P[x_i < X \leq x] = P[\emptyset] = 0$. De maneira análoga ao Caso I, concluímos que,

$$F_X(x) = \sum_{\substack{x_i \leq x \\ x_i \in \Gamma_X}} p_X(x_i).$$

□

Observação 2.1. Como consequência da proposição acima, se X é uma variável aleatória discreta, sua função de distribuição $F_X(x)$ é constante em cada intervalo (x_{i-1}, x_i) , para todo $x_i \in \Gamma_X$.

Exemplo 2.11. Consideremos o Exemplo 2.10. Seja i um número inteiro no domínio $0 \leq i \leq n$, então, $p_X(i) = P[X = i] = P[i \text{ caras em } n \text{ arremessos}]$. Sabemos que o espaço amostral contém 2^n pontos igualmente prováveis para uma seqüência de n arremessos. O evento $[X = i]$ pode ocorrer de tantas maneiras quanto se pode escolher i dos n arremessos para serem caras, isto é, de $\binom{n}{i}$ maneiras (todos os arremessos restantes podem ser

coroas de somente uma maneira). Assim,

$$p_X(i) = \frac{\binom{n}{i}}{2^n}$$

para $i = \{0, 1, 2, 3, \dots, n\}$. Se $i \notin \{0, 1, 2, 3, \dots, n\}$, $p_X(i) = 0$. Este é um caso particular da chamada função de densidade binomial.

Exemplo 2.12. Consideremos o Exemplo 2.2. Seja a variável aleatória X a que representa o número de caras obtidas no lançamento da moeda em cada ponto amostral. O gráfico da função de densidade de probabilidade da variável aleatória X é dado pela Figura 2.5.

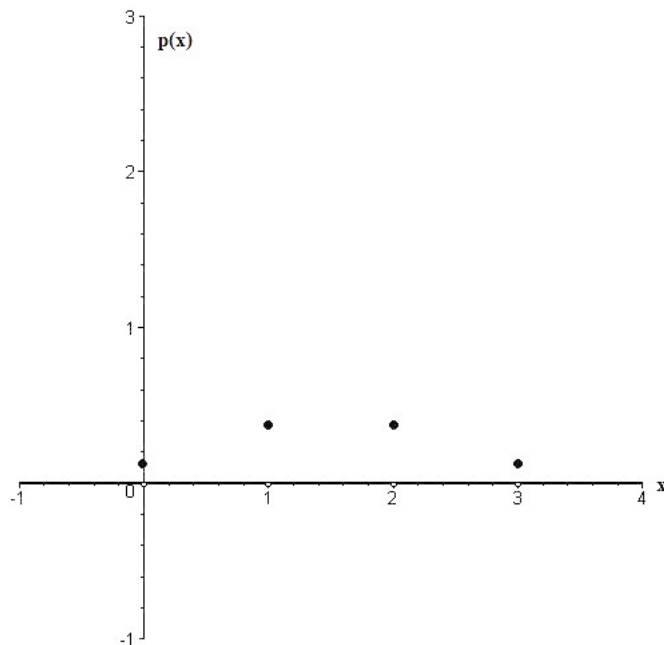


Figura 2.5: A função p_X para o número de caras em três arremessos de uma moeda.

2.6 Pares de Variáveis Aleatórias

O estudo de duas ou mais variáveis aleatórias definidas no mesmo espaço amostral é importante em inúmeras aplicações. Por exemplo, considere a experiência de selecionar uma válvula dentre um conjunto Ω de válvulas, se X representa o tempo de vida (em horas) da válvula selecionada e Y representa a pressão do gás no interior da válvula, então X e Y podem ser compreendidas como variáveis aleatórias definidas no mesmo espaço amostral Ω .

Em geral, se X e Y são variáveis aleatórias definidas num mesmo espaço amostral Ω podemos admitir que estas assumem valores em eixos de números reais, os quais são eixos coordenados do plano bidimensional comum. Dessa forma, cada ponto do espaço amostral Ω define um par de números reais, $x = X(\omega)$ e $y = Y(\omega)$, onde x e y podem ser tratados como as coordenadas de algum ponto do plano. Assim, o par (X, Y) de variáveis aleatórias pode ser considerado com uma função que a cada ponto ω no espaço amostral associa um ponto (x, y) no plano coordenado.

Com essa idéia, podemos construir correspondências entre determinados eventos ou subconjuntos do espaço amostral e determinadas regiões do plano (x, y) . Para números reais x_1 e y_1 fixados, o evento $A = [X = x_1; Y = y_1]$ é o conjunto de todos os pontos amostrais que tem para X o valor x_1 e para Y o valor y_1 , ou seja, $A = \{\omega \in \Omega \mid X(\omega) = x_1; Y(\omega) = y_1\}$. O evento $B = [X \leq x_1] = \{\omega \in \Omega \mid X(\omega) \leq x_1\}$ corresponderá á região do \mathbb{R}^2 que consiste de todos os pontos (x, y) cuja coordenada de x é menor ou igual a x_1 , ou seja, $B = \{(x, y) \in \mathbb{R}^2 \mid x \leq x_1\}$. O evento $C = [Y \leq y_1] = \{\omega \in \Omega \mid Y(\omega) \leq y_1\}$ corresponderá á região que consiste de todos os pontos (x, y) cuja coordenada de y é menor ou igual a y_1 , ou seja, $C = \{(x, y) \in \mathbb{R}^2 \mid y \leq y_1\}$.

Considerando o evento que consiste de todos os pontos amostrais ω , para os quais $X(\omega) \leq x_1$ e $Y(\omega) \leq y_1$, temos,

$$B \cap C = [X \leq x_1] \cap [Y \leq y_1]$$

$$B \cap C = \{\omega \in \Omega \mid X(\omega) \leq x_1\} \cap \{\omega \in \Omega \mid Y(\omega) \leq y_1\}$$

$$B \cap C = [X \leq x_1; Y \leq y_1].$$

Trabalhamos até agora com probabilidades da forma $F_X(x_1) = P[X \leq x_1]$ e $F_Y(y_1) = P[Y \leq y_1]$, que são funções de distribuição das variáveis aleatórias X e Y , de acordo com a Definição 2.3. Introduzimos aqui a notação $F_X(x)$, $F_Y(y)$, \dots , onde o subíndice de F especifica a variável aleatória em questão.

Definição 2.7. *Dadas duas variáveis aleatórias X e Y sobre um mesmo espaço amostral Ω . A função de distribuição de probabilidade para duas variáveis aleatórias é definida como $F_{X,Y} : \mathbb{R}^2 \rightarrow \mathbb{R}$, $F_{X,Y}(x_1, y_1) = P[X \leq x_1; Y \leq y_1]$. Assim sendo, $F_{X,Y}$ é função de duas variáveis denominada função de distribuição conjunta das variáveis aleatórias X e Y .*

Enuciamos abaixo algumas propriedades de $F_{X,Y}$ sem contudo demonstrá-las.

Proposição 2.4. *Seja $F_{X,Y}$ a função de distribuição conjunta das variáveis aleatórias X e Y . Então, são válidas as seguintes propriedades,*

1. $0 \leq F_{X,Y}(x, y) \leq 1$.
2. Se $x_1 \leq x_2$ e $y_1 \leq y_2$, então $F_{X,Y}(x_1, y_1) \leq F_{X,Y}(x_2, y_2)$.
3. $\lim_{\substack{x \rightarrow +\infty \\ y \rightarrow +\infty}} F_{X,Y}(x, y) = 1$.
 $\lim_{x \rightarrow -\infty} F_{X,Y}(x, y) = 0$.
 $\lim_{y \rightarrow -\infty} F_{X,Y}(x, y) = 0$.
4. $\lim_{x \rightarrow a^+} F_{X,Y}(x, y) = F_{X,Y}(a, y), \forall y \in \mathbb{R}$.
 $\lim_{y \rightarrow b^+} F_{X,Y}(x, y) = F_{X,Y}(x, b), \forall x \in \mathbb{R}$.
5. Dados $a \leq b$ e $c \leq d$, então $F_{X,Y}(a, c) - F_{X,Y}(a, d) - F_{X,Y}(b, c) + F_{X,Y}(b, d) \geq 0$.

Exemplo 2.13. Uma moeda “não-viciada” ou “honestá” é arremessada três vezes. Consideremos X como sendo a variável aleatória que representa o número de caras obtidas nas duas primeiras jogadas ($X = 0, 1$, ou 2) e Y como a que representa o número de caras obtidas no terceiro arremesso ($Y = 0$ ou 1). Considere “H” como sendo “cara” e “T” como sendo “coroa”. O espaço amostral Ω consiste nesse caso de oito pontos amostrais igualmente prováveis, possuindo cada uma probabilidade igual a $\frac{1}{8}$:

$$\Omega = \{HHH, HHT, HTH, THH, THT, TTH, HTT, TTT\}.$$

Os valores correspondentes das variáveis aleatórias X e Y são dados na tabela abaixo:

Alguns valores de $F_{X,Y}$ são dados abaixo:

$$[X \leq 2 \text{ e } Y \leq 1] = \{s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8\} = \Omega \Rightarrow P[X \leq 2 \text{ e } Y \leq 1] = 1,$$

$$[X \leq 2 \text{ e } Y \leq 0] = \{s_2, s_5, s_7, s_8\} \Rightarrow P[X \leq 2 \text{ e } Y \leq 0] = \frac{1}{2},$$

$$[X \leq 1 \text{ e } Y \leq 1] = \{s_3, s_4, s_5, s_6, s_7, s_8\} \Rightarrow P[X \leq 1 \text{ e } Y \leq 1] = \frac{3}{4},$$

$$[X \leq 1 \text{ e } Y \leq 0] = \{s_5, s_7, s_8\} \Rightarrow P[X \leq 1 \text{ e } Y \leq 0] = \frac{3}{8},$$

Ω	$X(s)$	$Y(s)$
$s_1 = HHH$	2	1
$s_2 = HHT$	2	0
$s_3 = HTH$	1	1
$s_4 = THH$	1	1
$s_5 = THT$	1	0
$s_6 = TTH$	0	1
$s_7 = HTT$	1	0
$s_8 = TTT$	0	0

Tabela 2.2: Valores correspondentes das variáveis aleatórias X e Y .

$$[X \leq 0 \text{ e } Y \leq 1] = \{s_6, s_8\} \Rightarrow P[X \leq 0 \text{ e } Y \leq 1] = \frac{1}{4},$$

$$[X \leq 0 \text{ e } Y \leq 0] = \{s_8\} \Rightarrow P[X \leq 0 \text{ e } Y \leq 0] = \frac{1}{8}.$$

Podemos determinar os valores de $F_{X,Y}(x, y)$, $\forall x, y \in \mathbb{R}$, tendo em mente os resultados acima. Por exemplo,

$$F_{X,Y}\left(\frac{3}{2}, \frac{1}{2}\right) = P\left[X \leq \frac{3}{2}; Y \leq \frac{1}{2}\right] = P[X \leq 1; Y \leq 0] = F_{X,Y}(1, 0) = \frac{3}{8},$$

$$F_{X,Y}(7, 0) = P[X \leq 7; Y \leq 0] = P[X \leq 2; Y \leq 0] = F_{X,Y}(2, 0) = \frac{1}{2}.$$

Definição 2.8. *Sejam $\omega \in \Omega$, X e Y variáveis aleatórias, A_1 e A_2 eventos tais que $A_1 = \{\omega \in \Omega \mid X(\omega) \leq x\}$ e $A_2 = \{\omega \in \Omega \mid Y(\omega) \leq y\}$, $x, y \in \mathbb{R}$. As variáveis aleatórias X e Y são ditas independentes caso os eventos A_1 e A_2 sejam independentes para todo x e y , conforme a Definição 1.4. Ou seja, $\forall x, y \in \mathbb{R}$,*

$$P[X \leq x; Y \leq y] = P[X \leq x]P[Y \leq y],$$

ou em termos da função de distribuição conjunta de probabilidade,

$$F_{X,Y}(x, y) = F_X(x)F_Y(y).$$

Exemplo 2.14. Um contador de Geiger é um aparelho que detecta a presença de radiação usando a capacidade que as partículas radioativas têm de ionizar certas moléculas (isto é, retirar alguns de seus elétrons). O aparelho consiste em um cilindro cheio de gás (normalmente argônio), submetido a uma tensão elétrica que mantém a parede do

tubo negativa em relação a um fio positivo que corre ao longo do seu eixo. Quando as partículas radioativas entram em contato com os átomos de argônio, convertem alguns deles em íons positivos mais elétrons. Sob a ação do campo elétrico, os íons migram para a parede do cilindro e os elétrons para o fio. Essa dupla migração causa um pulso de corrente que é transmitida a um amplificador, gerando um sinal. Quanto mais partículas radioativas houver, maior será a corrente elétrica. O primeiro modelo desse aparelho foi apresentado em 1912 pelo físico alemão Hans Wilhelm Geiger (1882-1945), daí seu nome. Outros modelos, mais modernos, têm outra estrutura, mas baseiam-se no mesmo princípio apresentado no início do século passado.

Considere um contador de Geiger que é bombardeado por partículas subatômicas. Sejam X e Y variáveis aleatórias que representam os intervalos de tempo entre o primeiro e o segundo, e entre o segundo e o terceiro sinais do contador, respectivamente. Podemos mostrar que a função de distribuição conjunta de probabilidade relativa a X e Y é dada por,

$$F_{X,Y}(x, y) = \begin{cases} 1 - e^{-ax} - e^{-ay} + e^{-a(x+y)}, & x \geq 0, y \geq 0 \\ 0 & x < 0, y < 0 \end{cases},$$

onde a é uma constante positiva. Assim a probabilidade de que o primeiro intervalo de tempo não exceda dois segundos e de que o segundo não exceda um segundo é,

$$P[X \leq 2; Y \leq 1] = 1 - e^{-2a} - e^{-a} + e^{-3a}.$$

De maneira geral observamos que,

$$F_{X,Y}(x, y) = \begin{cases} (1 - e^{-ax})(1 - e^{-ay}), & x \geq 0, y \geq 0 \\ 0 & x < 0, y < 0 \end{cases},$$

e segue que $F_{X,Y}(x, y)$ pode ser escrita como $F_X(x).F_Y(y)$, de acordo com a Definição 2.8, onde,

$$F_X(x) = \begin{cases} 1 - e^{-ax}, & x \geq 0 \\ 0 & x < 0 \end{cases},$$

$$F_Y(y) = \begin{cases} 1 - e^{-ay}, & y \geq 0 \\ 0 & y < 0 \end{cases}.$$

Nesse caso X e Y são variáveis aleatórias independentes.

Definição 2.9. *Sejam X e Y variáveis aleatórias discretas. Definimos função de densidade conjunta de probabilidade como $p_{X,Y} : \mathbb{R}^2 \rightarrow \mathbb{R}$, tal que,*

$$p_{X,Y}(x, y) = P[X = x; Y = y].$$

Analogamente ao caso de uma variável aleatória a função $p_{X,Y}$ tem as propriedades descritas na proposição abaixo.

Proposição 2.5. *Seja $p_{X,Y}$ a função de densidade conjunta de probabilidade e sejam Γ_X e Γ_Y os conjuntos imagem das variáveis aleatórias X e Y , respectivamente. Então são válidas as seguintes propriedades,*

1. $0 \leq p_{X,Y}(x, y) \leq 1, \forall x, y \in \mathbb{R}$.
2. $\sum_{\substack{x \in \Gamma_X \\ y \in \Gamma_Y}} p_{X,Y}(x, y) = 1$.
3. $F_{X,Y}(x, y) = \sum_{\substack{i < x \\ i \in \Gamma_X}} \sum_{\substack{j \leq y \\ j \in \Gamma_Y}} p_{X,Y}(i, j), \forall x, y \in \mathbb{R}$.

Exemplo 2.15. Consideremos o arremesso de uma moeda como descrito no Exemplo 2.13. Alguns valores de $p_{X,Y}$:

$$\begin{aligned} [X = 0; Y = 0] &= \{s_8\} \Rightarrow P[X = 0; Y = 0] = p_{X,Y}(0, 0) = \frac{1}{8}, \\ [X = 1; Y = 0] &= \{s_7, s_5\} \Rightarrow P[X = 1; Y = 0] = p_{X,Y}(1, 0) = \frac{1}{2}, \\ [X = 2; Y = 0] &= \{s_2\} \Rightarrow P[X = 2; Y = 0] = p_{X,Y}(2, 0) = \frac{1}{8}, \\ [X = 0; Y = 1] &= \{s_6\} \Rightarrow P[X = 0; Y = 1] = p_{X,Y}(0, 1) = \frac{1}{8}, \\ [X = 1; Y = 1] &= \{s_3, s_4\} \Rightarrow P[X = 1; Y = 1] = p_{X,Y}(1, 1) = \frac{1}{4}, \\ [X = 2; Y = 1] &= \{s_1\} \Rightarrow P[X = 2; Y = 1] = p_{X,Y}(2, 1) = \frac{1}{8}. \end{aligned}$$

Exemplo 2.16. Em muitos estudos experimentais desejamos saber se duas variáveis aleatórias são independentes, ou seja, a ocorrência de uma não interfere no resultado da ocorrência da outra. Este exemplo ilustra um método utilizado para estudar esse problema.

Suponhamos que um laboratório colheu informações de 1000 ratos que foram submetidos a aplicações de uma determinada droga, a qual suspeita-se ser responsável

pela ocorrência de um tipo de tumor. Desse grupo de 1000 ratos, alguns receberam uma dosagem forte, outros uma dosagem fraca e alguns nenhuma dosagem da referida droga. A Tabela 2.3 mostra a quantidade de ratos que receberam a droga e a dosagem aplicada relacionado com a presença do tumor.

	Tumor presente	Tumor ausente	
Dosagem Forte	10	90	100
Dosagem Fraca	15	485	500
Nenhuma Dosagem	12	388	400
	37	963	1000

Tabela 2.3: Presença de tumor nos ratos.

Consideremos as variáveis aleatórias X e Y e suponha que os resultados obtidos na Tabela 2.3 são utilizados como estimativas das probabilidades dos eventos a elas relacionados.

$$X = \begin{cases} 0, & \text{se o rato recebeu uma dosagem forte,} \\ 1, & \text{se o rato recebeu uma dosagem fraca,} \\ 2, & \text{se o rato não recebeu nenhuma dosagem.} \end{cases}$$

$$Y = \begin{cases} 0, & \text{se o rato apresenta tumor,} \\ 1, & \text{se o rato não apresenta tumor.} \end{cases}$$

Para que as variáveis aleatórias X e Y fossem independentes, seria necessário que a identidade $P[X = x; Y = y] = P[X = x]P[Y = y]$ fosse verdadeira. Se isso acontece, mostramos então que não existe relação entre a aplicação de droga e o aparecimento de determinado tipo de tumor nos ratos.

Com as informações dadas, temos condições de determinar,

$$p_{X,Y}(0, 0) = P[X = 0; Y = 0] \approx \frac{10}{1000} = 0,010,$$

$$p_{X,Y}(1, 0) = P[X = 1; Y = 0] \approx \frac{15}{1000} = 0,015,$$

$$p_{X,Y}(2, 0) = P[X = 2; Y = 0] \approx \frac{12}{1000} = 0,012,$$

$$p_{X,Y}(0, 1) = P[X = 0; Y = 1] \approx \frac{90}{1000} = 0,090,$$

$$p_{X,Y}(1, 1) = P[X = 1; Y = 1] \approx \frac{485}{1000} = 0,485,$$

$$p_{X,Y}(2, 1) = P[X = 2; Y = 1] \approx \frac{388}{1000} = 0,388.$$

Assim temos,

$$p_X(0) = P[X = 0] = \frac{100}{1000} = 0,10,$$

$$p_X(1) = P[X = 1] = \frac{500}{1000} = 0,50,$$

$$p_X(2) = P[X = 2] = \frac{400}{1000} = 0,40,$$

$$q_Y(0) = P[Y = 0] = \frac{37}{1000} = 0,037,$$

$$q_Y(1) = P[Y = 1] = \frac{963}{1000} = 0,963.$$

Com os dados obtidos, verificamos que as variáveis aleatórias X e Y não são independentes, pois,

$$p_{X,Y}(0, 0) \approx 0,010 \neq p_X(0)q_Y(0) \approx (0,10)(0,037) = 0,0037,$$

$$p_{X,Y}(1, 0) \approx 0,015 \neq p_X(1)q_Y(0) \approx (0,50)(0,037) = 0,0185,$$

$$p_{X,Y}(2, 0) \approx 0,012 \neq p_X(2)q_Y(0) \approx (0,40)(0,037) = 0,0148,$$

$$p_{X,Y}(0, 1) \approx 0,090 \neq p_X(0)q_Y(1) \approx (0,10)(0,963) = 0,0963,$$

$$p_{X,Y}(1, 1) \approx 0,485 \neq p_X(1)q_Y(1) \approx (0,50)(0,963) = 0,4815,$$

$$p_{X,Y}(2, 1) \approx 0,388 \neq p_X(2)q_Y(1) \approx (0,40)(0,963) = 0,3852.$$

Concluimos que $P[X = x; Y = y] \neq P[X = x]P[Y = y]$. Logo a aplicação da droga em estudo influencia na ocorrência ou não de tumor nos ratos. Portanto as variáveis aleatórias X e Y são dependentes.

Sabemos de acordo com a Definição 1.3, que se A e B não são dois eventos necessariamente independentes, a probabilidade conjunta de A e B é

$$P[A \cap B] = P[A | B]P[B] = P[B | A]P[A],$$

onde $P[A | B]$ é a probabilidade condicional de A , dado que B tenha ocorrido e $P[B | A]$ é a probabilidade condicional de B , dado que A tenha ocorrido.

Se A e B são dois eventos da forma $A = [Y \leq y]$, $B = [X \leq x]$, temos,

$$P[Y \leq y | X \leq x] = P[A | B].$$

Isso motiva a definição que segue.

Definição 2.10. *Dadas duas variáveis aleatórias discretas X e Y , tais que X, Y definidas em Ω e sejam $x, y \in \mathbb{R}$. Então a função de distribuição condicional de probabilidade da variável aleatória Y dado que o evento $[X \leq x]$ tenha ocorrido é $G(y | x) = P[Y \leq y | X \leq x]$.*

Definição 2.11. *Se X e Y são variáveis aleatórias discretas, então para qualquer (x, y) na imagem de (X, Y) , definimos a seguinte probabilidade condicional,*

$$P[Y = y | X = x] = \frac{P[X = x \text{ e } Y = y]}{P[X = x]},$$

como sendo a função de densidade condicional de probabilidade de $[Y = y]$ dado que $[X = x]$ tenha ocorrido.

Exemplo 2.17. No Exemplo 2.16 verificamos que no experimento que consiste em aplicar determinada dosagem de um droga que, inicialmente suspeitava-se ser responsável pelo aparecimento de tumores em ratos, concluímos que esses eventos são dependentes. Em outras palavras, mostramos através da definição de variáveis aleatórias independentes que essas variáveis aleatórias não satisfazem a identidade,

$$P[X = x; Y = y] = P[X = x]P[Y = y].$$

Motivados pela Definição 2.11, podemos pensar em verificar tal resultado através da função de densidade condicional de probabilidade. Ou seja, verificamos a seguinte relação: “Condicional à aplicação ou não de determinada droga em ratos, qual é a interferência no aparecimento ou não de um tumor?”.

Para verificar que as variáveis aleatórias X e Y são independentes, basta mostrar que $P[Y = y | X = x] = P[Y = y]$, de acordo com a Definição 1.4.

Analisando os dados obtidos do Exemplo 2.16, temos,

$$P_{X,Y}[Y = 0 | X = 0] = \frac{P[X = 0; Y = 0]}{P[X = 0]} = \frac{0,010}{0,10} = 0,10 \neq 0,037 = P[Y = 0],$$

$$P_{X,Y}[Y = 0 | X = 1] = \frac{P[X = 1; Y = 0]}{P[X = 1]} = \frac{0,015}{0,50} = 0,030 \neq 0,037 = P[Y = 0],$$

$$P_{X,Y}[Y = 0 | X = 2] = \frac{P[X = 2; Y = 0]}{P[X = 2]} = \frac{0,012}{0,40} = 0,030 \neq 0,037 = P[Y = 0],$$

$$P_{X,Y}[Y = 1 | X = 0] = \frac{P[X = 0; Y = 1]}{P[X = 0]} = \frac{0,090}{0,10} = 0,90 \neq 0,963 = P[Y = 1],$$

$$P_{X,Y}[Y = 1 | X = 1] = \frac{P[X = 1; Y = 1]}{P[X = 1]} = \frac{0,485}{0,50} = 0,970 \neq 0,963 = P[Y = 1].$$

Concluimos que $P[Y = y | X = x] \neq P[Y = y]$. Logo as variáveis aleatórias X e Y são dependentes.

2.7 Várias Variáveis Aleatórias

De forma semelhante como definimos a função de distribuição de probabilidade de duas variáveis aleatórias, podemos defini-la para n variáveis aleatórias.

Definição 2.12. *Sejam X_1, X_2, \dots, X_n variáveis aleatórias definidas no mesmo espaço amostral Ω . Então a função de distribuição de probabilidade de várias variáveis aleatórias é definida como,*

$$F_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) = P[X_1 \leq x_1; X_2 \leq x_2; \dots; X_n \leq x_n], \quad \forall x_1, x_2, \dots, x_n \in \mathbb{R}.$$

De maneira análoga, estendemos a Definição 2.9 de duas variáveis aleatórias para várias variáveis aleatórias discretas.

Definição 2.13. *A função de densidade de probabilidade de várias variáveis aleatórias discretas X_1, X_2, \dots, X_n é definida como,*

$$p_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) = P[X_1 = x_1; X_2 = x_2; \dots; X_n = x_n], \quad \forall x_1, x_2, \dots, x_n \in \mathbb{R}.$$

Definição 2.14. *A função de densidade de probabilidade condicionada das variáveis aleatórias X_1, X_2, \dots, X_n é definida como,*

$$p_{X_1, X_2, \dots, X_n}(x_n | x_1, x_2, \dots, x_{n-1}) = \frac{p_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n)}{p_{X_1, X_2, \dots, X_{n-1}}(x_1, x_2, \dots, x_{n-1})}, \quad \forall x_1, \dots, x_n \in \mathbb{R}.$$

Definição 2.15. *As variáveis aleatórias discretas X_1, X_2, \dots, X_n são ditas mutuamente independentes se,*

$$p_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) = p_{X_1}(x_1)p_{X_2}(x_2) \dots p_{X_n}(x_n), \quad \forall x_1, \dots, x_n \in \mathbb{R}.$$

Podemos notar que é possível a cada par de variáveis aleatórias no conjunto ser independente ou dependente aos pares, sem que todo o conjunto seja mutuamente independente.

Exemplo 2.18. Suponhamos que duas moedas sejam arremessadas, e que X_1 represente o número de caras na primeira moeda arremessada (0 ou 1), que X_2 represente o número de caras na segunda moeda arremessada (0 ou 1), e que X_3 represente o número de vezes em que os dois resultados são iguais (0 ou 1). Observamos que os pares (X_1, X_2) , (X_1, X_3) e (X_2, X_3) , são, cada um, independentes, mas que o conjunto (X_1, X_2, X_3) não é mutuamente independente.

Capítulo 3

Noções de Cadeia de Markov

3.1 Introdução

Iniciaremos este capítulo definindo processos estocásticos e apresentando exemplos do caso discreto e do caso contínuo. Abordaremos também seqüências aleatórias e iremos caracterizá-las como um processo estocástico de parâmetro discreto. Além disso, iremos definir cadeias de Markov e estudar o caso em que são estacionárias, com espaço de estados finito, irredutíveis e aperiódicas. Com isso, poderemos definir a matriz de transição de um passo da cadeia de Markov e, com auxílio de métodos algébricos, verificar o comportamento da cadeia para um número arbitrários de passos.

As cadeias de Markov recebem esse nome em homenagem ao matemático russo Andrei Andreyevich Markov (1856-1922), que desenvolveu importantes resultados na teoria de probabilidade. Além disso, o referido matemático demonstrou o Teorema do Limite Central, além de trabalhos sobre limites de integrais e teoria da aproximação.

Também para esclarecimento, a palavra estocástico vem do francês *stochastique* e do grego *stoktastikós*. Dizemos do que está relacionado ao acaso, pertencente ou relativo ao azar. Padrões estocásticos são aqueles que surgem por meio de eventos aleatórios. Por exemplo, o lançar de dados resulta em números estocásticos, pois qualquer uma das 6 faces pode ficar para cima. Além disso é importante salientar a diferença entre aleatoriedade e estocasticidade. Normalmente os eventos estocásticos são aleatórios. Todavia, podem eventualmente não ser. É perfeitamente possível, embora improvável, que uma série de 10 arremessos de dados gere uma seqüência não aleatória 6,5,4,2,1,2,4,3,5,6 ou uma seqüência não aleatória 1,1,1,1,1,1,1,1,1,1. A seqüência não aleatória é estocástica, pois

surgiu através de um evento aleatório: lançamento de dados.

3.2 Processos Estocásticos

Definição 3.1. *Um processo estocástico consiste de um espaço amostral Ω , (com uma função de probabilidade P), um conjunto não-vazio T e uma família de aplicações $X_t : \Omega \rightarrow \mathbb{R}$, $t \in T$. Definimos $\Gamma = \bigcup_{t \in T} \text{Im}(X_t)$.*

Exemplo 3.1. Como exemplos de processos estocásticos podemos considerar a variação de tráfego em um certo cruzamento que envolve a formação e a dissipação de congestionamento; a variação diária no tamanho do estoque de uma determinada empresa; o comportamento de partículas sujeitas a impactos aleatórios; variação da qualidade de produtos em uma fábrica ou a variação de lucros na bolsa de valores.

Podemos pensar em um processo estocástico como um modelo matemático utilizado para o estudo de fenômenos aleatórios que tem como resultados funções. Em outras palavras, para cada $\omega \in \Omega$, podemos considerar a aplicação $X : T \rightarrow \mathbb{R}$, onde $X(\omega) = X_t(\omega)$. Tais funções estão definidas sobre um conjunto arbitrário T e são chamadas de **realizações**. Os elementos do conjunto T são chamados de **parâmetros** e o conjunto T é chamado de **espaço de parâmetros**. Os elementos de Γ são chamados de **estados** e conjunto Γ é chamado de **espaço de estados**.

O espaço de parâmetros T de um processo estocásticos pode ser:

Contínuo: Quando o conjunto T encontra-se em uma situação na qual as medidas são tomadas continuamente. Por exemplo, podemos considerar o registro dos carros que passam em um cruzamento num determinado tempo, ou uma medida como temperatura ou voltagem. ($T = \{t : t \geq 0\}$).

Discreto: Quando T é um subconjunto enumerável. Por exemplo, podemos considerar o estoque de uma companhia que só é registrado uma vez por mês ou os ganhos totais de uma pessoa em um jogo de azar que só são registrados após cada partida.

Exemplo 3.2. O Produto Interno Bruto - PIB representa a soma (em valores monetários) de todos os bens e serviços finais produzidos numa determinada região (países, estados, cidades), durante um período determinado (mês, trimestre, ano, etc). Consideremos os resultados da soma de todas as riquezas produzidas no Brasil por ano a partir de 1960.

Podemos associar tal situação a um processo estocástico. De fato, consideremos o seguinte espaço amostral,

$$\Omega = \left\{ \begin{array}{l} E_1 = \{(1960, 15), (1961, 15), \dots, (2006, 1067), \dots\}, \\ E_2 = \{(1960, 16), (1961, 16), \dots, (2006, 1068), \dots\}, \\ E_3 = \{(1960, 17), (1961, 16), \dots, (2006, 1069), \dots\}, \\ \vdots \end{array} \right\}.$$

Os elementos de Ω são eventos formados por um conjunto de pontos ω da forma (x, y) , onde x representa os anos observados e y representa o valor total do PIB brasileiro em bilhões de dólares. Podemos definir uma família de variáveis aleatórias $X_t : \Omega \rightarrow \mathbb{R}$, $t \in \mathbb{N}$ da seguinte forma: para cada $\omega \in \Omega$, $X_t(\omega)$ representa o valor total do PIB no ano $(1960 + t)$. Abaixo descrevemos as realizações para dois eventos desse processo estocástico.

$$\left\{ \begin{array}{l} X_0(E_1) = 15, X_1(E_1) = 15, \dots, X_{48}(E_1) = 1067, \dots \\ X_0(E_2) = 16, X_1(E_2) = 16, \dots, X_{48}(E_2) = 1068, \dots \end{array} \right\}.$$

Este exemplo caracteriza um processo estocástico cujo espaço de parâmetros é o conjunto $T = \mathbb{N}$, e o espaço de estados são os valores que o PIB brasileiro pode assumir. Assim, tal situação é um processo estocástico com parâmetro discreto. A Figura 3.2 mostra o gráfico que representa uma realização do processo estocástico, considerando o evento E_1 . É importante salientar que o referido gráfico representa a evolução real do PIB do Brasil a partir do ano de 1960.

Exemplo 3.3. Em determinada ocasião observou-se a quantidade de pessoas numa fila de espera para atendimento em uma agência bancária. Esta situação ilustra um processo estocástico. O espaço amostral Ω é o conjunto de eventos da forma $E = \bigcup_{t \in \mathbb{R}^+} (t, x_t)$, onde t representa o tempo e x_t o número de pessoas na fila no tempo t . Podemos definir uma família de variáveis aleatórias $X_t : \Omega \rightarrow \mathbb{R}$, $t \in \mathbb{R}^+$, da forma $X_t(E) = x_t$.

Tal situação caracteriza um processo cujo espaço de parâmetros é o conjunto de tempo t em minutos, tal que $t \geq 0$, e o espaço de estados é um subconjunto de \mathbb{N} , que representa a quantidade de pessoas ou o comprimento da fila. Este processo é um processo estocástico com parâmetro contínuo. A Figura 3.2 mostra o gráfico que representa uma realização desse processo.

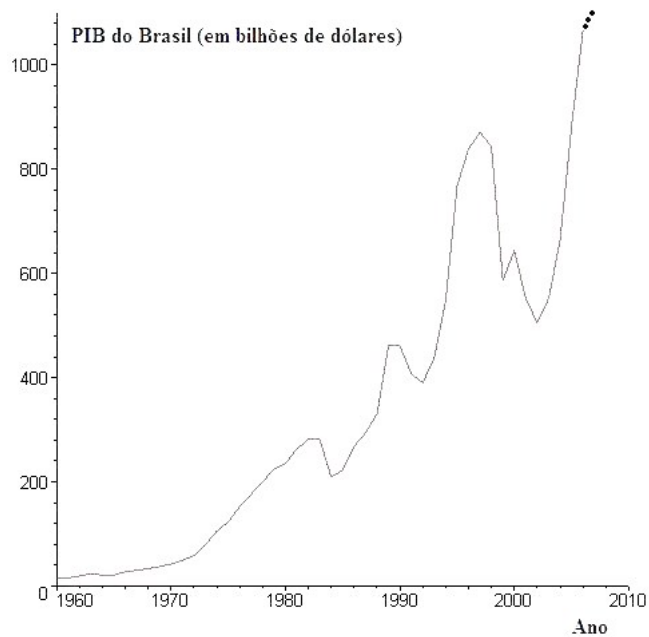


Figura 3.1: PIB do Brasil de 1960 a 2006.

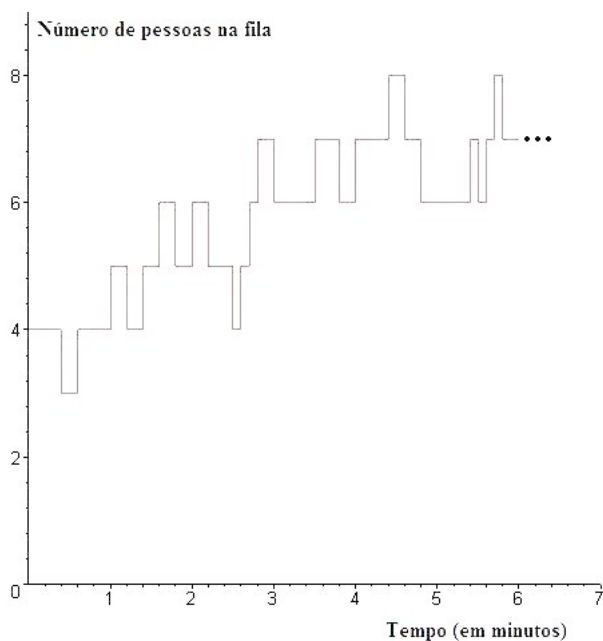


Figura 3.2: Uma realização de uma processo de fila de espera iniciando com quatro pessoas.

3.3 Seqüências Aleatórias e Cadeias de Markov

Definição 3.2. *Seja $\{X_n; n \in \mathbb{N}\}$ uma seqüência de variáveis aleatórias definidas no mesmo espaço amostral Ω . À esta seqüência dá-se o nome de seqüência aleatória.*

Uma seqüência aleatória nada mais é do que um processo estocástico com

parâmetro discreto. Abordaremos o caso em que cada variável aleatória é discreta.

Exemplo 3.4. Considere a seqüência de arremessos de uma moeda e para $n \in \mathbb{N}$,

$$X_n : \Omega \rightarrow \mathbb{R},$$

como a variável aleatória que assume o valor 1 se o $n + 1$ -ésimo arremesso for cara ou 0 quando o $n + 1$ -ésimo arremesso for coroa. O espaço amostral Ω e a família de variáveis aleatórias desse experimento podem ser ilustrados da forma abaixo.

$$\Omega = \left\{ \begin{array}{l} E_1 = \{(1, 0), (2, 0), (3, 0), \dots\}, \\ E_2 = \{(1, 0), (2, 1), (3, 0), \dots\}, \\ E_3 = \{(1, 0), (2, 0), (3, 1), \dots\}, \\ \vdots \end{array} \right\}.$$

$$\left\{ \begin{array}{l} X_0(E_1) = 0, X_1(E_1) = 0, X_2(E_1) = 0, \dots \\ X_0(E_2) = 0, X_1(E_2) = 1, X_2(E_2) = 0, \dots \\ X_0(E_3) = 0, X_1(E_3) = 0, X_2(E_3) = 1, \dots \\ \vdots \end{array} \right\}.$$

A referida experiência caracteriza um processo estocástico com parâmetro discreto.

Definição 3.3. Uma seqüência aleatória $\{X_n; n \in \mathbb{N}\}$, onde X_n são variáveis aleatórias discretas, é dita ser uma cadeia de Markov se,

$$P[X_{n+1} = j \mid X_{n_1} = i_{n_1}, X_{n_2} = i_{n_2}, \dots, X_{n_{k-1}} = i_{n_{k-1}}, X_{n_k} = i] = P[X_{n+1} = j \mid X_{n_k} = i],$$

$$\forall i, j, i_{n_1}, \dots, i_{n_{k-1}} \in \Gamma, \forall n_1 < n_2 < \dots < n_k \leq n,$$

se essas probabilidades estiverem definidas.

De acordo com o que foi demonstrado na referência [18], para verificarmos que $\{X_n; n \in \mathbb{N}\}$ é uma cadeia de Markov basta mostramos que,

$$P[X_{n+1} = j \mid X_0 = i_0, X_1 = i_1, \dots, X_n = i] = P[X_{n+1} = j \mid X_n = i],$$

$$\forall i, j, i_0, i_1, \dots, i_{n-1} \in \Gamma, \forall n \in \mathbb{N},$$

desde que essas probabilidades estejam definidas.

Exemplo 3.5. Considere uma seqüência de arremessos de uma moeda honesta, e seja X_n o número total de “caras” nos n primeiros arremessos. Se X_{n-1} é conhecido, ou seja, suponha que $X_{n-1} = k$, então X_n será k ou $k + 1$. Observamos ainda que não importa a ordem com que apareceram as “caras” ou “coroas” nos $n - 1$ primeiros arremessos, interessa apenas saber o total de “caras” obtidas ao longo da seqüência de arremessos até alcançar o penúltimo estado. Em outras palavras, o resultado pretendido depende apenas do penúltimo estágio, ou ainda, o processo não depende do valor de n . Assim,

$$P[X_n = x_n \mid X_1 = x_1; \dots; X_{n-1} = x_{n-1}] = P[X_n = x_n \mid X_{n-1} = x_{n-1}],$$

e, se $x_{n-1} = k$,

$$P[X_n = x_n \mid X_1 = x_1; \dots; X_{n-1} = x_{n-1}] = \begin{cases} \frac{1}{2} & \text{se } x_n = k \text{ ou } k + 1, \\ 0 & \text{para qualquer outro valor de } x_n. \end{cases}$$

Concluimos que essa é uma cadeia de Markov.

Exemplo 3.6. Num canhão de um tanque de guerra em combate em uma área de conflito da Terra, observa-se que cada tiro disparado acerta ou erra seu alvo com uma probabilidade que depende do sucesso ou erro do tiro imediatamente anterior disparado. A referida situação ilustra uma cadeia de Markov, pois o sucesso ou fracasso do último tiro depende apenas do penúltimo tiro disparado, não importando o que aconteceu ao longo da seqüência de tiros. Tal situação é uma cadeia de Markov com espaços de estados $\{0, 1\}$, sendo 0 acerto e 1 erro.

Exemplo 3.7. Consideremos uma seqüência de jogos independentes de um jogador contra o seu oponente. Seja $\{X_n; n \in \mathbb{N}\}$ a seqüência aleatória que representa o capital do jogador após as partidas. Suponhamos que o capital inicial do jogador seja k reais e que o capital inicial de seu oponente seja $N - k$, onde N representa o total de recursos disponíveis para o jogo. Com isso, $0 \leq X_i \leq N$, para todo $i \in \mathbb{N}$, ou seja, o espaço de estados de $\{X_n, n \in \mathbb{N}\}$ é $\Gamma = \{0, \dots, N\}$. A seqüência $\{X_n, n \in \mathbb{N}\}$ representa uma cadeia de Markov, pois, assim como no Exemplo 3.5, o capital do jogador após a n -ésima jogada depende do seu capital após $(n - 1)$ -ésima jogada, não interessando o seu capital nas jogadas anteriores.

Definição 3.4. *Uma cadeia finita de Markov é aquela em que $\Gamma \subset \mathbb{R}$ é finito. Caso contrário, é dita não finita.*

Exemplo 3.8. No Exemplo 3.5, para cada $n \in \mathbb{N}$, $Im(X_n)$ corresponde ao possível número total de caras obtidas nos n primeiros lançamentos da moeda, logo $Im(X_n) = \{0, \dots, n\}$. Dessa forma $\Gamma = \mathbb{N}$ e temos uma seqüência não-finita de Markov.

No Exemplo 3.6, $\Gamma \subset \mathbb{R}$ representa o resultado quando do disparo de um tiro de canhão. Existem apenas duas possibilidades para tal resultado, acerto ou erro, $\Gamma = \{0, 1\}$, portanto finito. Logo a seqüência ilustra uma cadeia finita de Markov.

No Exemplo 3.7, $\Gamma \subset \mathbb{R}$ representa o capital de um jogador após um seqüência de jogos independentes. Como N representa o total de recursos disponíveis no jogo, $\Gamma = \{0, 1, 2, \dots, N\}$, portanto finito. Logo a seqüência ilustra uma cadeia finita de Markov.

Definição 3.5. *Dado $i \in \Gamma$, definimos*

$$\mathcal{A}_i = \{m \in \mathbb{N}; P[X_m = i] > 0\}.$$

Vamos trabalhar com cadeias de Markov tais que $\mathcal{A}_i \neq \emptyset$, para todo $i \in \Gamma$.

Definição 3.6. *Uma cadeia de markov $\{X_n; n \in \mathbb{N}\}$ é estacionária quando, $\forall i, j \in \Gamma$,*

$$P[X_{m+1} = j \mid X_m = i],$$

não depende de $m \in \mathcal{A}_i$.

Exemplo 3.9. No Exemplo 3.5, o resultado do número total de caras numa seqüência de n lançamentos de uma moeda não depende de n , uma vez que nesse caso interessa somente saber o número total de caras, e este depende apenas do resultado de um estágio imediatamente anterior ao último. Da mesma forma, no Exemplo 3.6, as probabilidades de acerto ou erro dependem unicamente do penúltimo tiro disparado. No Exemplo 3.7, as probabilidades de vitória ou derrota do jogador em uma partida depende apenas do sucesso ou fracasso da penúltima partida. As três situações representam uma cadeira de Markov estacionária, pois satisfazem a Definição 3.6.

Como conseqüência, se

$$P[X_1 = j \mid X_0 = i],$$

existe,

$$P[X_{m+1} = j \mid X_m = i] = P[X_1 = j \mid X_0 = i], \forall m \in \mathcal{A}.$$

Definição 3.7. *Seja $\{X_n; n \in \mathbb{N}\}$ uma cadeia de Markov estacionária. Definimos as probabilidades de transição de um passo da cadeia como,*

$$p_{ij} = P[X_{m+1} = j \mid X_m = i], \quad m \in \mathcal{A}, \quad \forall i, j \in \Gamma.$$

Definição 3.8. *Seja $\{X_n; n \in \mathbb{N}\}$ uma cadeia de Markov estacionária e suponha que $\Gamma = \{0, \dots, N\}$. Então definimos $P = (p_{ij})$ como sendo a matriz de transição de um passo da cadeia.*

Exemplo 3.10. Considere o Exemplo 3.6. Suponha que as observações de campo de batalha possibilitem uma estimativa das probabilidades de transição de um passo da cadeia de Markov,

$$\begin{aligned} p_{00} &= P[X_{n+1} = 0 \mid X_n = 0] = \frac{3}{4}, \\ p_{01} &= P[X_{n+1} = 1 \mid X_n = 0] = \frac{1}{4}, \\ p_{10} &= P[X_{n+1} = 0 \mid X_n = 1] = \frac{1}{2}, \\ p_{11} &= P[X_{n+1} = 1 \mid X_n = 1] = \frac{1}{2}. \end{aligned}$$

Assim, a matriz de transição de um passo cadeia de Markov é dada por,

$$P = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix} = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Exemplo 3.11. Vamos supor que o nível econômico de um cidadão é classificado em três categorias: rico (R), classe média (M) e pobre (P). Também supomos que dos filhos de um homem rico, 95% são ricos e 5% de classe média; dos filhos de um indivíduo de classe média, 10% são ricos, 70% de classe média e 20% pobres; e no caso dos filhos de um cidadão pobre, 30% são de classe média e 70% são pobres. Admitindo que cada cidadão tem um filho, podemos formar um cadeia de Markov observando uma família através de gerações sucessivas. A matriz de probabilidade de transição de um passo do processo é,

$$P = \begin{pmatrix} 0,95 & 0,05 & 0 \\ 0,10 & 0,70 & 0,20 \\ 0 & 0,30 & 0,70 \end{pmatrix},$$

onde as linhas representam a classe social de uma geração e as colunas representam a classe social da geração imediatamente posterior, ou seja, as linhas representam a classe social do pais e as colunas representam a classe social dos filhos.

Exemplo 3.12. Consideremos o Exemplo 3.7. Seja p a probabilidade de o jogador ganhar um real em uma partida do jogo, e seja $q = 1 - p$ a probabilidade de o jogador perder um real. A matriz de transição de um passo do processo de Markov estacionário é dada por,

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ q & 0 & p & 0 & 0 & 0 & \dots & 0 \\ 0 & q & 0 & p & 0 & 0 & \dots & 0 \\ 0 & 0 & q & 0 & p & 0 & \dots & 0 \\ \vdots & & & & & & & \vdots \\ & \dots & & q & 0 & p & & \\ & \dots & & 0 & 0 & 1 & & \end{pmatrix}.$$

Notamos que P tem ordem $N + 1$.

Proposição 3.1. *Seja $P = (p_{ij})$ como na Definição 3.8. Então P é estocástica, isto é, $p_{ij} \geq 0, \forall i, j \in \Gamma$ e $\sum_{j=0}^N p_{ij} = 1, \forall i \in \Gamma$.*

Demonstração. Sejam $i, j \in \Gamma$, então, para $m \in \mathcal{A}_i$, temos,

$$p_{ij} = P[X_{m+1} = j \mid X_m = i]$$

Dessa forma, para cada $i \in \Gamma$,

$$\sum_{j=0}^N p_{ij} = \sum_{j=0}^N P[X_{m+1} = j \mid X_m = i].$$

Para cada $j \in \Gamma$, definimos $B_j = [X_{m+1} = j]$. Notamos que $B_k \cap B_j = \emptyset, \forall k, j \in \Gamma, k \neq j$ e $\Omega = \bigcup_{j=0}^N B_j$. Definindo $B = [X_m = i]$, temos,

$$\sum_{j=0}^N p_{ij} = \sum_{j=0}^N \frac{P[X_{m+1} = j; X_m = i]}{P[X_m = i]} = \sum_{j=0}^N \frac{P[B_j \cap B]}{P[B]},$$

ou seja,

$$\sum_{j=0}^N p_{ij} = \frac{1}{P[B]} P \left[\bigcup_{j=0}^N (B_j \cap B) \right] = \frac{1}{P[B]} P \left[B \cap \left(\bigcup_{j=0}^N B_j \right) \right] = \frac{1}{P[B]} P[B] = 1.$$

□

3.4 Equações de Chapman-Kolmogorov

Seja $\{X_n; n \in \mathbb{N}\}$ uma cadeia de Markov estacionária. Para $i, j \in \Gamma$ e $m \in \mathcal{A}_i$, queremos estudar as probabilidades da forma $P[X_{m+n} = j \mid X_m = i]$, $\forall n \in \mathbb{N}^*$. Em outras palavras, desejamos analisar o comportamento da cadeia quando da passagem do estágio i para o estágio j em n passos.

Teorema 3.1. *Seja $\{X_n, n \in \mathbb{N}\}$ uma cadeia de Markov estacionária. Dados $i, j \in \Gamma$, $n \in \mathbb{N}^*$, $P[X_{m+n} = j \mid X_m = i]$ não dependem de $m \in \mathcal{A}_i$.*

Demonstração. Sejam $i, j \in \Gamma$, $m \in \mathcal{A}_i$. Vamos provar por indução em n que o resultado é válido. Para $n = 1$ o resultado é verificado, uma vez que $P[X_{m+1} = j \mid X_m = i]$ é uma cadeia de Markov estacionária. Suponha que $\underbrace{P[X_{m+n} = j \mid X_m = i]}_{\text{Hipótese de Indução}}$ não depende de $m \in \mathcal{A}_i$, onde $n \in \mathbb{N}$. Temos que provar que $P[X_{m+n+1} = j \mid X_m = i]$ não depende de n . Definimos $A = [X_{m+n+1} = j]$, $B = [X_m = i]$ e, para cada $k \in \Gamma$, $B_k = [X_{m+n} = k]$. Notamos que $\Omega = \cup_{k=0}^N B_k$, e que,

$$(A \cap B_j \cap B) \cap (A \cap B_k \cap B) = \emptyset, \text{ se } j \neq k.$$

Com essa notação, temos,

$$P[X_{m+n+1} = j \mid X_m = i] = P[A \mid B].$$

Então,

$$P[X_{m+n+1} = j \mid X_m = i] = P[A \mid B] = P[A \cap \Omega \mid B] = P \left[A \cap \left(\bigcup_{k \in \Gamma} B_k \right) \mid B \right],$$

Assim,

$$P[X_{m+n+1} = j \mid X_m = i] = \frac{P[A \cap (\bigcup_{k \in \Gamma} B_k) \cap B]}{P[B]} = \frac{P[(\bigcup_{k \in \Gamma} (A \cap B_k)) \cap B]}{P[B]},$$

e,

$$P[X_{m+n+1} = j \mid X_m = i] = \sum_{k \in \Gamma} \frac{P[(A \cap B_k) \cap B]}{P[B]} = \sum_{k \in \Gamma} P[A \cap B_k \mid B].$$

Logo,

$$P[X_{m+n+1} = j \mid X_m = i] = \sum_{k \in \Gamma} P[X_{m+n+1} = j; X_{m+n} = k \mid X_m = i].$$

Queremos usar a fórmula,

$$P[A \cap B_k | B] = P[A | B \cap B_k]P[B_k | B].$$

Para isso, devemos ter $P[B] > 0$ e $P[B \cap B_k] > 0$. Defina

$$\mathcal{J} = \{k \in \Gamma, P[B \cap B_k] = 0\}.$$

Assim,

$$P[A \cap B_k | B] = \sum_{\substack{k \in \Gamma \\ k \in \mathcal{J}}} P[A \cap B_k | B] + \sum_{\substack{k \in \Gamma \\ k \notin \mathcal{J}}} P[A \cap B_k | B].$$

Seja $k \in \mathcal{J}$, então $P[B_k \cap B] = 0$, mas $P[A \cap B_k | B] = \frac{P[A \cap B_k \cap B]}{P[B]} = 0$, pois $A \cap B_k \cap B \subset B_k \cap B$ e $0 \leq P[A \cap B_k \cap B] \leq P[B_k \cap B] = 0$. Então,

$$P[X_{m+n+1} = j | X_m = i] = \sum_{\substack{k \in \Gamma \\ k \notin \mathcal{J}}} P[A \cap B_k | B].$$

Observamos que, se $k \notin \mathcal{J}$, $P[B \cap B_k] > 0$, e assim $P[B_k] > 0$. De fato, caso contrário, como $B \cap B_k \subset B$, teríamos $0 < P[B \cap B_k] \leq P[B_k] = 0$, o que é absurdo. Como $\{X_n, n \in \mathbb{N}\}$ é cadeia de Markov, temos,

$$P[A | B_k \cap B] = P[X_{m+n+1} = j | X_{m+n} = k] = P[A | B_k].$$

Então temos,

$$\begin{aligned} P[X_{m+n+1} = j | X_m = i] &= \sum_{\substack{k \in \Gamma \\ k \notin \mathcal{J}}} P[A \cap B_k | B] = \sum_{\substack{k \in \Gamma \\ k \notin \mathcal{J}}} P[A | B_k \cap B]P[B_k | B] \\ &= \sum_{\substack{k \in \Gamma \\ k \notin \mathcal{J}}} P[A | B_k]P[B_k | B] = \sum_{\substack{k \in \Gamma \\ k \notin \mathcal{J}}} p_{kj}P[B_k | B]. \end{aligned}$$

Se $k \in \mathcal{J}$, $P[B_k \cap B] = 0$, entretanto, $\exists m_k \in \mathbb{N}$ tal que $P[X_{m_k} = k] > 0$. Então $P[X_{m_k+1} = j | X_{m_k} = k]$ está bem definido e é igual a p_{kj} . Nesse caso,

$$P[X_{m_k+1} = j | X_{m_k} = k]P[X_{m+n} = k | X_m = i] = p_{kj}P[X_{m+n} = k | X_m = i] = 0.$$

Observamos que,

$$P[X_{m+n+1} = j | X_m = i] = \sum_{\substack{k \in \Gamma \\ k \notin \mathcal{J}}} p_{kj}P[B_k | B] + 0,$$

então,

$$P[X_{m+n+1} = j \mid X_m = i] = \sum_{k \in \Gamma} p_{kj} \underbrace{P[X_{m+n} = k \mid X_m = i]}_{\text{não depende de } m}. \quad (3.1)$$

Portanto, o somatório não depende de m .

□

Definição 3.9. *Seja $\{X_n; n \in \mathbb{N}\}$ uma cadeia de Markov estacionária. Definimos a probabilidade de transição de n -passos dessa cadeia como,*

$$p_{ij}^{(n)} = P[X_{m+n} = j \mid X_m = i], \text{ para } m \in \mathcal{A}_i, \forall i, j \in \Gamma.$$

Se $\Gamma = \{0, 1, \dots, N\}$, a matriz de probabilidade de n -passos dessa cadeia é definida como

$$P^{(n)} = \left(p_{ij}^{(n)} \right).$$

De acordo com a equação 3.1 temos, para todo $i, j \in \Gamma$, para todo $n \in \mathbb{N}$,

$$p_{ij}^{(n+1)} = \sum_{k \in \Gamma} p_{kj} p_{ik}^{(n)}.$$

A igualdade acima é conhecida como Equação de Chapman-Kolmogorov.

Proposição 3.2. *Seja $\{X_n; n \in \mathbb{N}\}$ uma Cadeia de Markov estacionária com $\Gamma = \{0, 1, \dots, N\}$, então $P^{(n)} = P^n, \forall n \in \mathbb{N}^*$.*

Demonstração. Vamos mostrar por indução que o resultado é verdadeiro. Suponha que $i, j \in \Gamma$ e que $m \in \mathcal{A}_i$. Para $n = 1$ temos,

$$P^{(1)} = (p_{ij}^{(1)}) = (P[X_{m+1} = j \mid X_m = i]) = (p_{ij}) = P = P^1.$$

Para $n = 2$ temos,

$$P^2 = (p_{ij}^{(2)}) = \left(\sum_{k=0}^N p_{kj} p_{ik}^{(1)} \right) = \left(\sum_{k=0}^N p_{kj} p_{ik} \right) = P.P = P^2.$$

Vamos supor que o resultado é válido para n , então,

$$P^{(n)} = P^n = \underbrace{P.P \dots P}_{n \text{ vezes}}.$$

Mostremos que o resultado é válido para $n + 1$. Com isso temos,

$$\begin{aligned} P^{(n+1)} &= (p_{ij}^{(n+1)}) = (P[X_{m+n+1} = j \mid X_m = i]) = \left(\sum_{k=0}^N p_{kj} P[X_{m+n} = k \mid X_m = i] \right) \\ &= \left(\sum_{k=0}^N p_{kj} p_{ik}^{(n)} \right) = P.P^n = P^{n+1}. \end{aligned}$$

Logo, $P^{(n+1)} = P^{n+1}$.

□

Definição 3.10. Seja $\{X_n; n \in \mathbb{N}\}$ uma cadeia de Markov estacionária com $\Gamma = \{0, \dots, N\}$. Sendo $p_i = P[X_0 = i]$, $\forall i \in \Gamma$, definimos,

$$\mathbf{p}^0 = (p_0, p_1, \dots, p_N),$$

como sendo o vetor de probabilidade inicial. Definimos também

$$p_i^{(n)} = P[X_n = i] \text{ e } \mathbf{p}^{(n)} = (p_0^{(n)}, p_1^{(n)}, \dots, p_N^{(n)}),$$

onde $\mathbf{p}^{(n)}$ é denominado vetor de probabilidade de n passos.

Proposição 3.3. Consideremos $\{X_n, n \in \mathbb{N}\}$ uma cadeia de Markov estacionária, $\mathbf{p}^{(0)}$ e $\mathbf{p}^{(n)}$ como na Definição 3.10, então,

$$\mathbf{p}^{(n)} = \mathbf{p}^0 P^n, \forall n \in \mathbb{N}^*.$$

Demonstração. Vamos provar por indução que o resultado é válido. Para $n = 1$,

$$\mathbf{p}^{(1)} = (p_0^{(1)}, p_1^{(1)}, \dots, p_N^{(1)}) \text{ e}$$

$$p_k^{(1)} = P[X_1 = k], \forall k \in \{0, \dots, N\}.$$

Definimos $B = [X_1 = k]$, e para cada $j \in \{0, \dots, N\}$, $B_j = [X_0 = j]$. Então $B_i \cap B_j = \emptyset$, se $i \neq j$, $\Omega = \bigcup_{j=0}^N B_j$.

$$p_k^{(1)} = P[B \cap \Omega] = P \left[B \cap \left(\bigcup_{j=0}^N B_j \right) \right] = P \left[\bigcup_{j=0}^N B_j \cap B \right] = \sum_{j=0}^N P[B_j \cap B].$$

Queremos usar a fórmula $P[B_j \cap B] = P[B | B_j]P[B_j]$. Seja $\mathcal{J} = \{j \in \Gamma; P[B_j] = 0\}$, então,

$$p_k^{(1)} = \sum_{\substack{j \in \Gamma \\ j \in \mathcal{J}}}^N P[B_j \cap B] + \sum_{\substack{j \in \Gamma \\ j \notin \mathcal{J}}}^N P[B_j \cap B].$$

Se $j \in \mathcal{J}$, $P[B_j] = 0$ e como $B_j \cap B \subset B_j$, temos $P[B_j \cap B] \leq P[B_j]$, ou seja,

$$0 \leq P[B_j \cap B] \leq 0 \Rightarrow P[B_j \cap B] = 0.$$

Com isso,

$$p_k^{(1)} = 0 + \sum_{\substack{j \in \Gamma \\ j \notin \mathcal{J}}}^N P[B_j \cap B] = \sum_{\substack{j \in \Gamma \\ j \notin \mathcal{J}}}^N P[B | B_j]P[B_j],$$

ou seja,

$$p_k^{(1)} = \sum_{\substack{j \in \Gamma \\ j \notin \mathcal{J}}}^N P[X_1 = k \mid X_0 = j]P[X_0 = j] = \sum_{\substack{j \in \Gamma \\ j \notin \mathcal{J}}}^N p_{jk}p_j.$$

Para o caso de $j \in \mathcal{J}$, $p_j = 0$ e podemos escrever,

$$p_k^{(1)} = \sum_{\substack{j \in \Gamma \\ j \notin \mathcal{J}}}^N p_{jk}p_j + \sum_{\substack{j \in \Gamma \\ j \in \mathcal{J}}}^N p_{jk} \cdot 0 = \sum_{j \in \Gamma}^N p_j p_{jk},$$

ou seja,

$$\mathbf{p}^{(1)} = \mathbf{p}^0 P.$$

Vamos supor que o resultado é válido para n . Então,

$$\mathbf{p}^{(n)} = \mathbf{p}^0 P^n.$$

Temos de mostrar que o resultado é válido para $n + 1$. Primeiro, observamos que,

$$\mathbf{p}_k^{(n+1)} = (p_0^{(n+1)}, p_1^{(n+1)}, \dots, p_N^{(n+1)}),$$

com

$$p_k^{(n+1)} = P[X_{n+1} = k], \forall k \in \{0, \dots, N\}.$$

Seja, $k \in \Gamma$ fixado. Definimos para cada $j \in \Gamma$, $B_j^n = [X_n = j]$. Então $B_i^n \cap B_j^n = \emptyset$, se $i \neq j$, $\Omega = \bigcup_{j=0}^N B_j^n$. Dessa forma, se $B^{n+1} = [X_{n+1} = k]$, temos

$$\begin{aligned} p_k^{(n+1)} &= P[X_{n+1} = k] = P[B^{n+1}] = P[B^{n+1} \cap \Omega] = P[B^{n+1} \cap (\bigcup_{j=0}^N B_j^n)] \\ &= \sum_{j=0}^N P[B^{n+1} \cap B_j^n]. \end{aligned}$$

Definindo $\mathcal{J}^n = \{j \in \Gamma; P[B_j^n] = 0\}$, temos,

$$\begin{aligned} p_k^{(n+1)} &= \sum_{\substack{j \in \Gamma \\ j \notin \mathcal{J}^n}} P[B^{n+1} \cap B_j^n] + \sum_{\substack{j \in \Gamma \\ j \in \mathcal{J}^n}} P[B^{n+1} \cap B_j^n] \\ &= \sum_{\substack{j \in \Gamma \\ j \notin \mathcal{J}^n}} P[B^{n+1} \mid B_j^n]P[B_j^n] + 0 \\ &= \sum_{\substack{j \in \Gamma \\ j \notin \mathcal{J}^n}} P[X_{n+1} = k \mid X_n = j]p_j^{(n)}. \end{aligned}$$

Se $j \in \mathcal{J}^n$, então $p_j^{(n)} = 0$, e

$$p_k^{(n+1)} = \sum_{j \in \Gamma} p_j^{(n)} p_{jk}.$$

Dessa forma,

$$\mathbf{p}^{(n+1)} = \mathbf{p}^{(n)} P = (\mathbf{p}^0 P^n) P,$$

ou seja,

$$\mathbf{p}^{(n+1)} = \mathbf{p}^0 P^{n+1}.$$

□

Queremos estudar $(p_{ij}^{(n)})$ para o caso em que n se torna suficientemente grande, ou seja, quando os passos do processo são repetidos n vezes. Para isso, introduzimos as definições abaixo e posteriormente dois teoremas, cujas demonstrações não serão apresentadas neste trabalho. (Ver, por exemplo a referência [12]).

Definição 3.11. *Seja $\{X_n; n \in \mathbb{N}\}$ uma cadeia de Markov estacionária com $\Gamma = \{0, \dots, N\}$.*

Se dados $j, k \in \Gamma$,

$$p_{jk}^{(n)} > 0 \text{ e } p_{kj}^{(m)} > 0,$$

para algum $m, n \in \mathbb{N}^$, dizemos que os estados j e k se **comunicam**. Se todos os estados se comunicam dizemos então que a cadeia é **irredutível**. Caso contrário ela é dita **não-irredutível**.*

Exemplo 3.13. O Exemplo 3.6 ilustra uma cadeia de Markov irredutível, pois a definição acima é satisfeita, ou seja, todos os estados se comunicam.

Considere o Exemplo 3.7. Quando o jogador perde todo o seu dinheiro para o seu oponente, este não tem mais chance de iniciar uma nova partida, fazendo com que o jogo termine. De forma semelhante, quando o jogador ganha todos os recursos disponíveis, o jogo também termina. Podemos ver que em ambas situações os estados não se comunicam com os demais, fazendo com que a cadeia de Markov em questão seja não-irredutível.

Definição 3.12. *Uma cadeia de Markov é dita **periódica** se $\exists j \in \mathbb{N}$ com*

$$P^{n+j} = P^n, \forall n \in \mathbb{N}^*.$$

*Caso contrário ela é dita **aperiódica**.*

Proposição 3.4. *Seja $\{X_n; n \in \mathbb{N}\}$ uma cadeia de Markov estacionária, irredutível e tal que $\Gamma = \{0, \dots, N\}$. Então, se $p_{ii} > 0$, para algum $i \in \Gamma$, $\{X_n; n \in \mathbb{N}\}$ é aperiódica [18].*

Exemplo 3.14. Vamos considerar que exista um equipamento que é classificado dentre uma das três condições: (a) funcionamento, (b) em reparo ou (c) inoperante aguardando mais trabalho. Tal equipamento é observado somente quando da mudança de estados. Seja $X_n = 0$, caso a n -ésima mudança de estado o coloque em condições de funcionamento, $X_n = 1$, caso a n -ésima mudança de estado o coloque em uma condição de reparo e $X_n = 2$, caso a n -ésima mudança de estado o coloque inoperante.

As hipóteses a serem consideradas são: Se o equipamento estiver parado para fim de reparo, a mudança de estado seguinte será para a condição de funcionamento; Se o equipamento estiver inoperante aguardando mais trabalho, a mudança de estado seguinte o colocará em funcionamento; Se não houver trabalho para fazer, a equipe de reparo será desfeita de modo a não permitir que nenhuma transição ocorra da condição de reparo para inoperante; Se estiver funcionando, poderá mudar para a condição de inoperante ou poderá estragar e ser posto em reparo, supondo essas possibilidades igualmente prováveis.

Seja $\{X_n; n \in \mathbb{N}\}$ uma cadeia de Markov nessas condições. Então a matriz de transição de um passo é dada por,

$$P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Assim, $P^{(2)}$ é dada por,

$$P^{(2)} = P^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Da mesma forma, $P^{(3)}$ é dada por,

$$P^{(3)} = P^3 = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Em geral, podemos observar que,

$$P^{(2n)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad n \in \mathbb{N}^*,$$

e

$$P^{(2n-1)} = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad n \in \mathbb{N}^*.$$

Portanto, de acordo com a Definição 3.12, esta cadeia de Markov é periódica.

Teorema 3.2. *Seja $\{X_n; n \in \mathbb{N}\}$, uma cadeia de Markov estacionária, aperiódica e com $\Gamma = \{0, \dots, N\}$. Então os limites $\mathcal{V}_j = \lim_{n \rightarrow +\infty} p_j^{(n)}$ existem $\forall j \in \Gamma$. Além disso $\mathcal{V}_j = \lim_{n \rightarrow +\infty} p_{ij}^{(n)}$, para qualquer $i \in \Gamma$.*

Observação 3.1. Consideremos $V = (\mathcal{V}_0, \dots, \mathcal{V}_N)$, então V é dito estacionário no seguinte sentido,

$$\begin{aligned} \lim_{n \rightarrow +\infty} p_j^{(n)} &= \sum_{k=0}^N \lim_{n \rightarrow +\infty} (p_k^{(n-1)}) p_{kj} \\ \mathcal{V}_j &= \sum_{k=0}^N \mathcal{V}_k p_{kj}, \end{aligned}$$

ou seja,

$$V = VP.$$

Como conseqüência, usando o princípio da indução podemos mostrar que $V = VP^n$, $\forall n \in \mathbb{N}$.

Teorema 3.3. *Se $\{X_n; n \in \mathbb{N}\}$ é uma cadeia de Markov estacionária, aperiódica, irreduzível e com $\Gamma = \{0, \dots, N\}$, V dado acima é o único vetor estacionário que satisfaz,*

$$\sum_{j=0}^N \mathcal{V}_j = 1, \quad \mathcal{V}_j \geq 0.$$

V é dito vetor de probabilidade.

Observação 3.2. Considere a matriz de transição de n -passos de um processo de Markov,

$$P^{(n)} = \begin{pmatrix} p_{00}^{(n)} & p_{01}^{(n)} & \cdots & p_{0N}^{(n)} \\ p_{10}^{(n)} & p_{11}^{(n)} & \cdots & p_{1N}^{(n)} \\ \vdots & \vdots & & \vdots \\ p_{N0}^{(n)} & p_{N1}^{(n)} & \cdots & p_{NN}^{(n)} \end{pmatrix}$$

$$\lim_{n \rightarrow +\infty} P^{(n)} = \lim_{n \rightarrow +\infty} \begin{pmatrix} p_{00}^{(n)} & p_{01}^{(n)} & \cdots & p_{0N}^{(n)} \\ p_{10}^{(n)} & p_{11}^{(n)} & \cdots & p_{1N}^{(n)} \\ \vdots & \vdots & & \vdots \\ p_{N0}^{(n)} & p_{N1}^{(n)} & \cdots & p_{NN}^{(n)} \end{pmatrix} = \begin{pmatrix} \mathcal{V}_1 & \mathcal{V}_2 & \cdots & \mathcal{V}_N \\ \mathcal{V}_1 & \mathcal{V}_2 & \cdots & \mathcal{V}_N \\ \vdots & \vdots & & \vdots \\ \mathcal{V}_1 & \mathcal{V}_2 & \cdots & \mathcal{V}_N \end{pmatrix} = \begin{pmatrix} V \\ V \\ \vdots \\ V \end{pmatrix}$$

$$\lim_{n \rightarrow +\infty} P^{(n)} = \begin{pmatrix} V \\ V \\ \vdots \\ V \end{pmatrix}.$$

Como conseqüência,

$$\lim_{n \rightarrow +\infty} P^n = \begin{pmatrix} V \\ V \\ \vdots \\ V \end{pmatrix}.$$

Exemplo 3.15. No Exemplo 3.5 há pelo menos dois vetores estacionários de probabilidade $V_1 = (1, 0, \dots, 0)$ e $V_2 = (0, 0, \dots, 1)$. Isso ocorre por que a cadeia correspondente é não-irredutível.

Exemplo 3.16. Considere a matriz P do Exemplo 3.10. Queremos conhecer a matriz de transição de n passos quando n se torna suficientemente grande, ou seja, queremos saber o que acontece com as probabilidades de transição quando temos uma seqüência de muitos tiros de canhão. Para isso, vamos utilizar métodos de Álgebra Linear. Sabemos que se a referida matriz é diagonalizável pois podemos escrevê-la da forma,

$$P = ADA^{-1}$$

$$\begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{4} \end{pmatrix} \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix},$$

onde A é a matriz formada pelos autovetores associados aos autovalores de P e D é uma matriz diagonal cujas entradas não nulas são os autovalores da matriz P [3]. Sabemos também que,

$$\lim_{n \rightarrow +\infty} P^{(n)} = \lim_{n \rightarrow +\infty} P^n = \lim_{n \rightarrow +\infty} AD^nA^{-1}$$

Dessa forma, temos,

$$\lim_{n \rightarrow +\infty} P^n = \lim_{n \rightarrow +\infty} A \begin{pmatrix} 1^n & 0 \\ 0 & (\frac{1}{4})^n \end{pmatrix} A^{-1} = \lim_{n \rightarrow +\infty} A \begin{pmatrix} 1 & 0 \\ 0 & (\frac{1}{4})^n \end{pmatrix} A^{-1}$$

$$\begin{aligned}\lim_{n \rightarrow +\infty} P^n &= \lim_{n \rightarrow +\infty} \begin{pmatrix} 1 & \left(\frac{1}{4}\right)^n \\ 1 & -2\left(\frac{1}{4}\right)^n \end{pmatrix} \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix} \\ \lim_{n \rightarrow +\infty} P^n &= \lim_{n \rightarrow +\infty} \begin{pmatrix} \frac{2}{3} + \frac{1}{3}\left(\frac{1}{4}\right)^n & \frac{1}{3} - \frac{1}{3}\left(\frac{1}{4}\right)^n \\ \frac{2}{3} - \frac{2}{3}\left(\frac{1}{4}\right)^n & \frac{1}{3} + \frac{2}{3}\left(\frac{1}{4}\right)^n \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}.\end{aligned}$$

Podemos observar que conforme o número de tiros aumenta, as probabilidades de transição de um passo do processo se estabilizam. Em outras palavras, a informação de acerto ou erro do tiro anterior não influencia no sucesso ou fracasso do tiro seguinte, pois,

$$\lim_{n \rightarrow +\infty} p_{00}^{(n)} = \lim_{n \rightarrow +\infty} p_{10}^{(n)} = \frac{2}{3} \text{ e } \lim_{n \rightarrow +\infty} p_{01}^{(n)} = \lim_{n \rightarrow +\infty} p_{11}^{(n)} = \frac{1}{3}.$$

Para esclarecimento, com esse resultado percebemos que a medida em que temos uma seqüência de muitos tiros, a probabilidade de acertar um tiro dado que o anterior foi correto é igual a probabilidade de acertar um tiro dado que o anterior foi perdido. De forma análoga, a probabilidade de errar um tiro dado que o anterior foi correto é igual a probabilidade de errar um tiro dado de o anterior também foi perdido.

Exemplo 3.17. Considere a matriz P do Exemplo 3.10. Da mesma que forma que no exemplo anterior, queremos conhecer o que acontece com as probabilidades de transição da cadeia de Markov quando temos n suficientemente grande. Em outras palavras, procuramos a matriz de transição $(p_{ij}^{(n)})$ quando n se torna suficientemente grande, ou seja, após sucessivas gerações.

Com métodos semelhantes aos utilizados no exemplo anterior, temos que a referida matriz é diagonalizável, pois,

$$\begin{aligned}P &= \begin{pmatrix} 0,95 & 0,05 & 0 \\ 0,10 & 0,70 & 0,20 \\ 0 & 0,30 & 0,70 \end{pmatrix} = ADA^{-1} \\ &= \begin{pmatrix} -0,063 & -0,577 & -0,485 \\ 0,638 & -0,577 & 0,485 \\ 0,766 & -0,577 & 0,727 \end{pmatrix} \begin{pmatrix} 0,45 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0,9 \end{pmatrix} \begin{pmatrix} -0,158 & 0,790 & -0,632 \\ -0,944 & -0,472 & -0,314 \\ -0,916 & 0,458 & 0,458 \end{pmatrix}.\end{aligned}$$

onde A é a matriz formada pelos autovetores associados aos autovalores de P e D é uma matriz diagonal cujas entradas não nulas são os autovalores da matriz P . Sabemos

também que,

$$\lim_{n \rightarrow +\infty} P^{(n)} = \lim_{n \rightarrow +\infty} P^n = \lim_{n \rightarrow +\infty} AD^n A^{-1}$$

$$\lim_{n \rightarrow +\infty} P^n = \lim_{n \rightarrow +\infty} A \begin{pmatrix} (0.45)^n & 0 & 0 \\ 0 & (1)^n & 0 \\ 0 & 0 & (0.9)^n \end{pmatrix} A^{-1} = \begin{pmatrix} 0.54 & 0.27 & 0.19 \\ 0.54 & 0.27 & 0.19 \\ 0.54 & 0.27 & 0.19 \end{pmatrix}.$$

Observamos que conforme o número de gerações se torna suficientemente grande, as probabilidades de transição para outras classes sociais independem do fato de o indivíduo ter nascido rico, classe média ou pobre. Por exemplo, a probabilidade de um indivíduo ser filho de um indivíduo pobre e se tornar rico é a mesma probabilidade de outra indivíduo ser filho de um indivíduo classe média e se tornar rico.

Exemplo 3.18. Considere o Exemplo 3.7. Procuramos nessa situação uma forma de se conhecer qual a probabilidade de o jogador esgotar seus recursos, dado que ele inicie a seqüência de jogos com um determinado capital. Em outras palavras, ao iniciar o jogo com um capital $0 < k < N$, procuramos saber a probabilidade s_k para que o jogador fique arruinado, ou seja $s_k = P[X_n = 0 \mid X_0 = k]$. Podemos observar que não importa o que aconteça com o jogador após um número digamos n de rodadas em um jogo, invariavelmente ele começa o jogo ganhando ou perdendo o seu capital inicial. Em caso de vitória na primeira rodada, com probabilidade p anteriormente estabelecida, ele prossegue o jogo com capital $k+1$ reais, já em caso de derrota, com probabilidade $q = 1-p$, se $k > 1$ o jogador prossegue o jogo com capital $k-1$ reais; se $k = 1$ o jogador estará arruinado. Seguindo tal raciocínio temos que a probabilidade para que o jogador seja arruinado pode ser escrita como $s_k = qs_{k-1} + ps_{k+1}$, para $0 < k < N$. Sabendo-se que a probabilidade de o jogador perder um Real em uma rodada do jogo é $q = 1-p \Rightarrow 1 = p+q$, podemos escrever,

$$(p+q)s_k = qs_{k-1} + ps_{k+1} \Rightarrow ps_k + qs_k = qs_{k-1} + ps_{k+1}$$

$$qs_k - qs_{k-1} = ps_{k+1} - ps_k \Rightarrow q(s_k - s_{k-1}) = p(s_{k+1} - s_k)$$

$$s_k - s_{k-1} = \left(\frac{p}{q}\right)(s_{k+1} - s_k), \text{ ou ainda,}$$

$$s_{k+1} - s_k = \left(\frac{q}{p}\right)(s_k - s_{k-1}) \text{ para } 0 < k < N.$$

Desenvolvendo a expressão repetidamente temos,

$$\begin{aligned}
k = 1 &\Rightarrow (s_2 - s_1) = \left(\frac{q}{p}\right) (s_1 - s_0) \\
k = 2 &\Rightarrow (s_3 - s_2) = \left(\frac{q}{p}\right) (s_2 - s_1) = \left(\frac{q}{p}\right)^2 (s_1 - s_0) \\
k = 3 &\Rightarrow (s_4 - s_3) = \left(\frac{q}{p}\right) (s_3 - s_2) = \left(\frac{q}{p}\right)^3 (s_1 - s_0) \\
&\vdots \\
\text{para algum } k &\Rightarrow (s_{k+1} - s_k) = \left(\frac{q}{p}\right)^k (s_1 - s_0).
\end{aligned}$$

Como $s_0 = P[X_n = 0 \mid X_0 = 0]$ é um evento certo, uma vez que a ruína do jogador será certa caso ele inicie o jogo sem dinheiro algum, então $s_0 = 1$. Nesse caso temos $(s_{k+1} - s_k) = \left(\frac{q}{p}\right)^k (s_1 - 1)$, para $0 \leq k < N$. Tomando algum i de tal modo que $0 < i \leq N$, podemos escrever,

$$\begin{aligned}
(s_i - s_0) &= (s_i - s_{i-1}) + (s_{i-1} - s_{i-2}) + (s_{i-2} - s_{i-3}) + \dots + (s_1 - s_0) \\
(s_i - s_0) &= \left(\frac{q}{p}\right)^{i-1} (s_1 - s_0) + \left(\frac{q}{p}\right)^{i-2} (s_1 - s_0) + \dots + \left(\frac{q}{p}\right)^0 (s_1 - s_0) \\
(s_i - s_0) &= \left[\left(\frac{q}{p}\right)^{i-1} + \left(\frac{q}{p}\right)^{i-2} + \left(\frac{q}{p}\right)^{i-3} + \dots + \left(\frac{q}{p}\right)^0 \right] (s_1 - s_0) \\
(s_i - 1) &= \left[\left(\frac{q}{p}\right)^{i-1} + \left(\frac{q}{p}\right)^{i-2} + \left(\frac{q}{p}\right)^{i-3} + \dots + 1 \right] (s_1 - 1) \\
(s_i - 1) &= \sum_{j=0}^{i-1} \left(\frac{q}{p}\right)^j (s_1 - 1).
\end{aligned}$$

Vemos que $\sum_{j=0}^{i-1} \left(\frac{q}{p}\right)^j$ pode ser vista com uma progressão geométrica finita da forma $\sum_{l=0}^h ar^l$, com $r < 1$. Dessa forma escrevemos, $\sum_{l=0}^h ar^l = \frac{1 - r^{h+1}}{1 - r}$. Aqui $a = 1$, $r = \frac{q}{p} < 1$, uma vez que $1 = p + q$, com $p \neq q$. Então,

$$\begin{aligned}
(s_i - 1) &= \sum_{j=0}^{i-1} \left(\frac{q}{p}\right)^j (s_1 - 1) = \left[\frac{1 - \left(\frac{q}{p}\right)^{i-1+1}}{1 - \left(\frac{q}{p}\right)} \right] (s_1 - 1) \\
(s_i - 1) &= \left[\frac{1 - \left(\frac{q}{p}\right)^i}{1 - \left(\frac{q}{p}\right)} \right] (s_1 - 1), \quad 0 \leq i \leq N.
\end{aligned}$$

Considerando $i = N$ e sabendo que $s_N = P[X_n = 0 \mid X_0 = N]$, representa a probabilidade de o jogador ser arruinado na n -ésima rodada, dado que ele iniciou o jogo com todo recurso disponível N , observamos que s_N será um evento impossível, uma vez que trata-se de um evento improvável. Dessa maneira, a expressão acima obtida é escrita da seguinte forma,

$$(s_N - 1) = \left[\frac{1 - \left(\frac{q}{p}\right)^N}{1 - \left(\frac{q}{p}\right)} \right] (s_1 - 1) \Rightarrow (s_1 - 1) = - \left[\frac{1 - \left(\frac{q}{p}\right)}{1 - \left(\frac{q}{p}\right)^N} \right].$$

Conhecendo $(s_i - 1) = \left[\frac{1 - \left(\frac{q}{p}\right)^i}{1 - \left(\frac{q}{p}\right)} \right] (s_1 - 1)$ e substituindo $(s_1 - 1)$ temos,

$$(s_i - 1) = \left[\frac{1 - \left(\frac{q}{p}\right)^i}{1 - \left(\frac{q}{p}\right)} \right] \cdot \left\{ - \left[\frac{1 - \left(\frac{q}{p}\right)}{1 - \left(\frac{q}{p}\right)^N} \right] \right\} = - \left[\frac{1 - \left(\frac{q}{p}\right)^i}{1 - \left(\frac{q}{p}\right)^N} \right],$$

ou ainda, substituindo i por k , obtemos a probabilidade de um jogador que inicia uma partida dispondo de um capital k de recursos eventualmente ser arruinado,

$$s_k - 1 = - \left[\frac{1 - \left(\frac{q}{p}\right)^k}{1 - \left(\frac{q}{p}\right)^N} \right] \Rightarrow s_k = 1 - \left[\frac{1 - \left(\frac{q}{p}\right)^k}{1 - \left(\frac{q}{p}\right)^N} \right]$$

$$s_k = \frac{1 - \left(\frac{q}{p}\right)^N - 1 + \left(\frac{q}{p}\right)^k}{1 - \left(\frac{q}{p}\right)^N} = \frac{\left(\frac{q}{p}\right)^k - \left(\frac{q}{p}\right)^N}{1 - \left(\frac{q}{p}\right)^N}, \text{ com } p \neq q.$$

Para o caso onde a probabilidade do jogador ganhar um real em uma partida ser igual a probabilidade de o jogador perder um real em determinada partida, ou seja, quando $p = q$, $s_k = 1 - \frac{k}{N}$.

Podemos considerar a situação hipotética em que o jogador disputa com um adversário provido de infinitos recursos, ou seja, $N \rightarrow +\infty$. Desejamos obter uma expressão que represente a probabilidade de o jogador ser arruinado disputando o jogo contra um jogador muito rico. É necessário então, considerar dois casos:

Caso 1) $p > q$:

$$\lim_{N \rightarrow +\infty} s_k = \lim_{N \rightarrow +\infty} \frac{\left(\frac{q}{p}\right)^k - \left(\frac{q}{p}\right)^N}{1 - \left(\frac{q}{p}\right)^N} = \left(\frac{q}{p}\right)^k.$$

Assim, quando a probabilidade de o jogador ganhar um real em uma partida for maior que a probabilidade de o jogador perder um real, temos que probabilidade de o mesmo ser arruinado quando a disputa se dá com um adversário provido de infinitos recursos é $\left(\frac{q}{p}\right)^k$, onde k é o seu capital inicial.

Caso 2) $p < q$:

$$\lim_{N \rightarrow +\infty} s_k = \lim_{N \rightarrow +\infty} \frac{\left(\frac{q}{p}\right)^k - \left(\frac{q}{p}\right)^N}{1 - \left(\frac{q}{p}\right)^N} = \lim_{N \rightarrow +\infty} \frac{\frac{\left(\frac{q}{p}\right)^k - \left(\frac{q}{p}\right)^N}{\left(\frac{q}{p}\right)^N}}{\frac{1 - \left(\frac{q}{p}\right)^N}{\left(\frac{q}{p}\right)^N}}$$

$$\lim_{N \rightarrow +\infty} s_k = \frac{\frac{1}{\left(\frac{q}{p}\right)^{N-k}} - 1}{\frac{1}{\left(\frac{q}{p}\right)^N} - 1} = \frac{\lim_{N \rightarrow +\infty} \frac{1}{\left(\frac{q}{p}\right)^{N-k}} - \lim_{N \rightarrow +\infty} 1}{\lim_{N \rightarrow +\infty} \frac{1}{\left(\frac{q}{p}\right)^N} - \lim_{N \rightarrow +\infty} 1} = \frac{0 - 1}{0 - 1} = 1.$$

Na situação em que o jogador disputa com um adversário muito rico e que a probabilidade de ganhar um real em uma partida do jogo é menor do que a probabilidade de o jogador perder um real, vemos que ele inevitavelmente será arruinado, ou seja, perderá todos os seus recursos para o seu oponente.

Como já mencionado, este exemplo não ilustra uma cadeia de Markov irreduzível, ou seja, nem todos os estados se comunicam. Portanto não é possível utilizar os métodos algébricos estudados neste trabalho para sua resolução, exigindo assim uma outra abordagem para o problema.

Considerações Finais

O objetivo deste trabalho foi realizar um estudo sobre as cadeias de Markov e os métodos algébricos utilizados para a resolução de problemas. Para isso, a atenção voltou-se às cadeias de Markov estacionárias, com espaços de estados finito, irredutíveis e não-periódicas.

Durante o texto utilizamos uma abordagem semi-rigorosa acerca da teoria de probabilidade, baseada nas noções de evento e espaço amostral. Concentramo-nos apenas nas variáveis aleatórias discretas, uma vez que a inclusão do caso contínuo exigiria mais tempo de estudo. O mesmo se pode dizer das cadeias de Markov que não são estacionárias, que possuem espaço de estados não finito, não-irredutíveis e periódicas.

Para futuros trabalhos, sugerimos a continuação do estudo de processos estocásticos, bem como os casos não abordados nesse trabalho.

Referências Bibliográficas

- [1] BHARUCHA-REID, A T. **Elements of the Theory of Markov Processes and Their Applications**. New York: McGraw-Hill, 1960.
- [2] BLAKE, Ian F. **An Introduction to Applied Probability**. New York: J.Wiley, 1979.
- [3] BOLDRINI, José Luiz. **Álgebra Linear**. Terceira Edição. São Paulo: HARBRA, 1986.
- [4] BRYC, Wlodzimierz. **Applied Probability and Stochastic Processes**. Cincinnati: Wlodzimierz Bryc, 1995.
- [5] CASTRO JR., A. Armando de, **Curso de Teoria da Medida**. Rio de Janeiro: IMPA, 2004.
- [6] CLARKE, A. Bruce; DISNEY, Ralph L. **Probabilidade e Processos Estocásticos**. Rio de Janeiro: LTC, 1979.
- [7] COSTA NETO, Pedro Luiz de Oliveira; CYMBALISTA, Melvin. **Probabilidades: Resumos Teóricos, Exercícios Resolvidos, Exercícios Propostos**. Segunda Edição. São Paulo: E. Blucher, 2006.
- [8] CRAMER, Harald. **Elementos da Teoria da Probabilidade e Algumas de suas Aplicações**. São Paulo: Mestre Jou, 1973.
- [9] FELLER, William. **An Introduction to Probability Theory and Its Applications**. Third Edition. New York. London. Sydney: John WileySons, Inc., 1968.
- [10] FERNANDEZ, Pedro Jesus. **Introdução aos Processos Estocásticos**. Rio de Janeiro: IMPA, 1975.

- [11] GRAY, Robert M. **Probability, Random Processes, and Ergodic Properties.** Department of Electrical Engineering: Stanford University, 2001.
- [12] GRIMMETT, G. STIRZAKER, D. **Probability and Random Processes.** Third Edition. New York: Oxford University Press, 2001.
- [13] HAZZAN, Samuel. **Fundamentos de Matemática Elementar, 5: Combinatória, Probabilidade.** Sexta Edição. São Paulo: Atual, 1993.
- [14] HOEL, Paul G. **Introduction to Stochastic Processes.** Boston: H. Mifflin, 1972.
- [15] HSU, Hwei P. **Theory and Problems of Probability, Random Variables, and Random Processes.** New York: Schaum's Outline Series: 1997.
- [16] KOLMOGOROV, A. N. **Foundations of the Theory of Probability.** Second English Edition. University of Oregon: Chelsea Publishing Company, 1956.
- [17] LANGE, Kenneth. **Applied Probability.** New York: Springer, 2003.
- [18] LEFEBVRE, M. **Applied Stochastic Process.** New York: Springer, 2007.
- [19] MORGADO, A. C. **Análise Combinatória e Probabilidade.** Nona Edição. Rio de Janeiro: SBM, 2006.
- [20] NELSON, Edward. **Radically Elementary Probability Theory.** Princeton, New Jersey: Princeton University Press: 1987.
- [21] PAPOULIS, Athanasios. **Probability, Random Variables, and Stochastic Processes.** Third Edition. Polytechnic Institute of New York: MCGraw-Hill, Inc., 1991.
- [22] PEEBLES, Peyton Z. **Probability, Random Variables, and Random Signal Principles.** Second Edition. New York: McGraw-Hill, Inc., 1987.
- [23] ROSS, Sheldon. **A First Course in Probability.** Seventh Edition. Upper Saddle River, New Jersey: Prentice-Hall Internacional, 1998.
- [24] ROZANOV, I. U. A. (IUrii Anatol'evich). **Introductory Probability Theory.** Englewood Cliffs: Prentice-Hall, 1969.

- [25] VARES, Maria Eulalia; COLÓQUIO BRASILEIRO DE MATEMÁTICA: (15:1985) Poços de Caldas, MG. **Grandes Desvios em Processos Markovianos**. Rio de Janeiro: IMPA, 1985.
- [26] XAVIER, Teresinha de Maria Bezerra Sampaio; XAVIER, Airton Fontenele Sampaio. **Probabilidade : Teoria e Problemas**. Rio de Janeiro: Livros Técnicos e Científicos, 1974.