UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE MATEMÁTICA, ESTATÍSTICA
E COMPUTAÇÃO CIENTÍFICA

Relatório de Pesquisa de Iniciação Científica

Martingais: Teoremas e aplicações

João Vitor Vieira de Castro

Élcio Lebensztayn

O presente trabalho foi realizado com apoio da Fundação de Amparo
à Pesquisa do Estado de São Paulo (FAPESP), Brasil.

Processos nº 2023/13453-5 e nº 2024/09519-3.

Campinas

Dezembro de 2025

Sumário

Prefácio **1**

1. Introdução **3**

1.1. Esperança condicional 3

1.2. Martingais 5

2. Teoremas da Parada Opcional **6**

3. Integrabilidade Uniforme de Variáveis Aleatórias **14**

3.1. Convergência de variáveis aleatórias 14

3.2. Integrabilidade Uniforme 17

4. Convergência de Martingais **21**

5. Desigualdades Maximais e Convergência em L^p **30**

6. Martingais Reversos e Inferência Bayesiana **34**

7. Martingais a Tempo Contínuo **41**

8. Movimento Browniano e Martingais Associados **51**

Referências bibliográficas **60**

Prefácio

A teoria de martingais tem suas raízes em estratégias de apostas que se disseminaram na França, no século XVIII, visando a assegurar ganhos em jogos justos. Do ponto de vista histórico, Lévy (1948) foi o primeiro a explorar, em processos estocásticos, propriedades que hoje conhecemos como martingais. O conceito moderno surge com Ville (1939), que introduziu o termo *martingale* e formulou a primeira definição, no contexto de jogos honestos. A consolidação rigorosa da teoria, com o estabelecimento dos primeiros resultados fundamentais, deve-se a Doob (1990). Atualmente, os martingais constituem uma ferramenta vital para a investigação e a análise de sistemas que evoluem aleatoriamente no tempo.

Este texto apresenta um estudo abrangente do tema, indo dos conceitos básicos a resultados avançados, com destaque para os teoremas centrais da teoria, Parada Opcional e Convergência de Martingais, além de aplicações em diferentes contextos. O trabalho foi desenvolvido como parte do projeto de Iniciação Científica de João Vitor Vieira de Castro, sob a orientação de Élcio Lebensztayn, com bolsa da Fundação de Amparo à Pesquisa do Estado de São Paulo, realizado no Instituto de Matemática, Estatística e Computação Científica, Unicamp.

Na Seção 1, apresentamos a definição de esperança condicional e suas propriedades principais. Também discutimos os conceitos de martingais, submartingais e supermartingais. Na Seção 2, dedicamo-nos a um dos resultados essenciais da teoria: o Teorema da Parada Opcional. Iniciamos com a definição de tempo de parada e demonstramos suas propriedades básicas. Estabelecemos diferentes versões do teorema referido, cada uma com condições suficientes específicas. Deduzimos então a Identidade de Wald, com uma aplicação interessante relacionada ao tempo até que um passeio aleatório assimétrico em $\mathbb{Z}$ atinja um valor $b > 0$ fixado.

Como prelúdio ao estudo da convergência de martingais, na Seção 3, revisamos os diferentes modos de convergência de variáveis aleatórias (quase certa, em L^p, em probabilidade, em distribuição) e suas relações. Também abordamos a integrabilidade uniforme, um conceito crucial no estudo da convergência em média. Provamos o Teorema da Bola de Cristal, que fornece condições práticas para a integrabilidade uniforme, e o Teorema de Vitali, caracterização fundamental da convergência em média. Finalizamos a Seção 3 com um Teorema da Amostragem Opcional e um exemplo associado aos processos de ramificação de Bienaymé–Galton–Watson.

A Seção 4 traz um dos resultados mais importantes da teoria: o Teorema da Convergência de Martingais. Estabelecemos, assim, condições sob as quais um martingal converge quase certamente e em média. A demonstração se baseia na Desigualdade de Snell para subidas (*upcrossings*) e no Teorema de Vitali. Apresentamos como exemplos os processos de ramificação e a urna de Pólya. Discutimos também os martingais de Doob e a Lei 0–1 de Kolmogorov.

Na Seção 5, desenvolvemos desigualdades maximais, incluindo as Desigualdades de Doob e de Doob–Kolmogorov. Esses resultados fornecem limitantes superiores para a probabilidade de que o máximo de um martingal (ou de seu valor absoluto) exceda um determinado valor. Provamos então o Teorema da Convergência de martingais em L^p, com $p > 1$.

Na Seção 6, exploramos martingais reversos, com o teorema de convergência associado e o Teorema de Lévy. A seguir, formulamos a definição de uma sequência permutável de variáveis aleatórias, e apresentamos a Lei Forte dos Grandes Números para variáveis permutáveis e o Teorema de Representação de De Finetti. Fechamos a seção com uma aplicação à Inferência Bayesiana e uma discussão sobre a convergência da estatística do teste da razão de verossimilhanças.

O objetivo da Seção 7 é estender a teoria de martingais para processos estocásticos a tempo contínuo. Consideramos a definição geral de um tempo de parada e enunciamos versões a tempo contínuo dos Teoremas da Parada Opcional e de Convergência de Martingais. Exemplos importantes são martingais associados ao processo de Poisson homogêneo, a processos de nascimento puro e a processos de nascimento e morte. Em outra ilustração, aplicamos o Teorema da Parada Opcional ao modelo epidêmico SIR markoviano e obtemos uma fórmula recursiva para a função de probabilidade do número final de indivíduos saudáveis.

A Seção 8 é dedicada ao movimento browniano. Apresentamos a definição formal e estabelecemos sua existência, por meio da construção via base de Haar. Também examinamos alguns martingais associados ao movimento browniano.

A literatura sobre martingais é bastante extensa, logo nos contentamos em comentar as principais obras consultadas durante a Iniciação Científica. Aos leitores interessados no aprofundamento de tópicos sobre martingais a tempo discreto (Seções 1 a 6), sem muitos pormenores da Teoria da Medida, citamos os livros de Karlin e Taylor (1975, Capítulo 6), Lawler (2006, Capítulo 5), Durrett (2016, Capítulo 5) e Grimmett e Stirzaker (2020, Seções 7.7 a 7.10 e Capítulo 12). Aqueles familiarizados com as definições de mensurabilidade e de σ-álgebras podem consultar Karlin e Taylor (1975, Seção 6.7), Athreya e Lahiri (2006, Capítulo 13), Lanchier (2017, Capítulo 5) e Durrett (2019, Capítulo 4). Para as partes de martingais a tempo contínuo (Seções 7 e 8), sugerimos a leitura de Durrett (2019, Capítulo 7) e Mörters e Peres (2010).

O presente trabalho foi realizado com apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Brasil. Processos nº 2023/13453-5 e nº 2024/09519-3.

As opiniões, hipóteses e conclusões ou recomendações expressas neste material são de responsabilidade dos autores e não necessariamente refletem a visão da FAPESP.

1. Introdução

O conceito de martingal relacionado a um contexto probabilístico surge durante o século XVIII, na França, como uma estratégia de aposta popular. Em um jogo de caráter simples – lançamento de uma moeda honesta, apostando se irá aparecer cara ou coroa – foi desenvolvida uma estratégia que, quase certamente, leva ao lucro do apostador.

Este termo aparece pela primeira vez na teoria de probabilidade em um trabalho do matemático francês Jean Ville (Ville, 1939), relacionado com o conceito de "sistema de jogo". O probabilista explicita que insere esse nome para relacionar à estratégia de aposta descrita anteriormente (Mansuy, 2009).

Para introduzirmos este conceito, é necessário definirmos esperança condicional para variáveis aleatórias.

1.1. Esperança condicional

Definição 1.1. Sejam Y, $X_1, \ldots, X_n$ variáveis aleatórias definidas em um mesmo espaço de probabilidade $(\Omega, \mathcal{F}, P)$. A **esperança condicional** de Y dado $X_1, \ldots, X_n$ é a variável aleatória $E(Y|X_1, \ldots, X_n)$ que satisfaz as seguintes propriedades:

1. É uma variável que depende apenas das $X_1, \ldots, X_n$, ou seja, existe uma função ϕ mensurável com respeito a $X_1, \ldots, X_n$ de tal forma que $E(Y|X_1, \ldots, X_n) = \phi(X_1, \ldots, X_n)$;

2. Sejam $A \subseteq \Omega$ um evento qualquer que dependa apenas de $X_1, \ldots, X_n$ e 1_A a respectiva variável indicadora. Então, $E(Y1_A) = E(E(Y|X_1, \ldots, X_n)1_A)$.

Teorema 1.1. ***(Caracterização da Esperança Condicional)*** *A variável aleatória que satisfaz as propriedades da Definição 1.1 existe e é quase certamente única.*

Demonstração: Podemos aplicar o Teorema de Radon–Nikodym com respeito à medida de probabilidade condicional. Os detalhes da demonstração podem ser vistos em Durrett (2019, Seção 4.1).

Definição 1.2. Sejam $X_1, \ldots, X_n$ variáveis aleatórias. O conjunto de todos os eventos mensuráveis com respeito a tais variáveis aleatórias forma uma σ-álgebra dita a **informação contida em** $X_1, \ldots, X_n$ e será denotada por $\mathcal{F}_n$. Desta forma, abreviamos a notação da esperança condicional por $E(Y|X_1, \ldots, X_n) = E(Y|\mathcal{F}_n)$.

Diversas propriedades relacionadas à esperança de uma variável aleatória podem ser estendidas para a esperança condicional, usando o item (2) da Definição 1.1: valendo para todo evento $A \in \mathcal{F}_n$, a relação segue pela unicidade. Sejam Y, Z variáveis aleatórias, $a, b \in \mathbb{R}$ e $\mathcal{F}_n$ a informação disponível no tempo n.

Proposição 1.2. ***(Lei da Esperança Total)*** $E(E(Y|\mathcal{F}_n)) = E(Y)$.

Demonstração: Aplique a definição com o evento $A = \Omega$, de tal forma que $1_A \equiv 1$. ■

Proposição 1.3. ***(Linearidade)*** $E(aY + bZ|\mathcal{F}_n) = aE(Y|\mathcal{F}_n) + bE(Z|\mathcal{F}_n)$.

Proposição 1.4. *Se Y é $\mathcal{F}_n$-mensurável, $Y = \psi(X_1, \ldots, X_n)$; então $E(Y|\mathcal{F}_n) = Y$.*

Proposição 1.5. *Se Y é independente de $X_1, \ldots, X_n$, a informação não é útil para determinar Y, de tal forma que $E(Y|\mathcal{F}_n) = E(Y)$.*

Proposição 1.6. ***(Desigualdade de Jensen)*** *Seja $\phi : \mathbb{R} \to \mathbb{R}$ função contínua e convexa, ou seja, $\phi(a)t + \phi(b)(1-t) \geq \phi(at + b(1-t))$, $\forall a, b \in \mathbb{R}$ e $t \in [0,1]$. Então, $\phi(E(Y|\mathcal{F})) \leq E(\phi(Y)|\mathcal{F})$.*

Exemplos de funções com tal propriedade e serão vistas nas próximas seções são o valor absoluto, $|x| = \begin{cases} x, & \text{se } x \geq 0 \\ -x, & \text{se } x < 0 \end{cases}$; e a parte positiva, $x^+ = \begin{cases} x, & \text{se } x \geq 0 \\ 0, & \text{se } x < 0 \end{cases}$.

Proposição 1.7. *Se Z é $\mathcal{F}_n$-mensurável, então $E(YZ|\mathcal{F}_n) = ZE(Y|\mathcal{F}_n)$.*

Demonstração: Para qualquer $A \in \mathcal{F}_n$, temos que ambas 1_A e Z são $\mathcal{F}_n$-mensuráveis. Queremos mostrar que $E(YZ1_A) = E(ZE(Y|\mathcal{F}_n)1_A)$. Provamos, inicialmente, supondo que a variável aleatória Z é uma combinação linear finita de funções indicadoras. O resultado segue pela definição e pela propriedade da linearidade, discutida na Proposição 1.3.

Suponha, agora, que $Z \geq 0$. Aproximamos Z por funções simples: tomamos uma sequência de variáveis $Z_n : \Omega \to [0,\infty]$ mensuráveis tais que $0 \leq Z_n \leq Z_{n+1}$ para todo n, $Z = \lim_n Z_n$ e tais Z_n são combinações lineares finitas de indicadoras. A relação da proposição segue pelo Teorema da Convergência Monótona. Para Z geral, repetimos o raciocínio para as partes positiva Z^+ e negativa Z^-, onde $x^- = \begin{cases} 0, & \text{se } x \geq 0 \\ -x, & \text{se } x < 0 \end{cases}$. ■

1.2. Martingais

Definição 1.3. Uma **filtragem** é uma sequência crescente de informações $\mathcal{F}_n$, ou seja, para todo n, existe uma coleção $\mathcal{A}_n$ de variáveis aleatórias tais que

$$\mathcal{A}_m \subseteq \mathcal{A}_n \text{ se } m < n.$$

Podemos interpretar que não há perda de informação na sequência.

Definição 1.4. Uma sequência $\{Y_n\}_{n\geq 0}$ de variáveis aleatórias é um **martingal** com respeito à filtragem $\{\mathcal{F}_n\}_{n\geq 0}$ se, para quaisquer m e n com $m < n$,

(i) $E(|Y_n|) < \infty$;

(ii) $Y_n \in \mathcal{F}_n$;

(iii) $E(Y_n \mid \mathcal{F}_n) = Y_m$.

Se, em (iii), substituímos = por $\geq$ ou $\leq$, então a sequência $\{Y_n\}$ é dita um **submartingal** ou **supermartingal**, respectivamente. Se $X_1, \ldots, X_n$ definem a informação $\mathcal{F}_n$, dizemos que $\{Y_n\}_{n\geq 0}$ é um martingal com respeito a $\{X_n\}_{n\geq 0}$.

Proposição 1.8. *O item (iii) acontece se, e somente se,* $\forall n = 0, 1, \ldots,$ *temos* $E(Y_{n+1}|\mathcal{F}_n) = Y_n$.

Demonstração: Se o item (iii) acontece, basta fixarmos os índices corretos. Assim, assuma que é válido $E(Y_{n+1}|\mathcal{F}_n) = Y_n$. O raciocínio segue indutivamente: sabemos que, por

hipótese, fixarmos $k = 1$ garante $E(Y_{n+k}|\mathcal{F}_n) = Y_n$. Assim, supondo verdadeiro para $k \in \mathbb{N}$, temos, pela Lei da Esperança Total,

$$E(Y_{n+(k+1)}|\mathcal{F}_n) = E(E(Y_{n+(k+1)}|\mathcal{F}_{n+k})|\mathcal{F}_n) \overset{\text{hipótese}}{=} E(Y_{n+k}|\mathcal{F}_n) \overset{\text{indução}}{=} Y_n. \qquad \blacksquare$$

A Proposição 1.8 pode ser estendida para as definições de submartingal e supermartingal, substituindo o sinal da relação por $\geq$ ou $\leq$, respectivamente.

Lanchier (2017) sugere interpretar a variável aleatória Y_n como a fortuna de uma apostadora após apostar n vezes consecutivas em um jogo. Este jogo pode ser justo (definindo um martingal), desfavorável à jogadora (definindo um supermartingal) ou favorável a ela (definindo um submartingal). Desta forma, supondo que o jogo seja justo, independentemente do passado, a fortuna da apostadora em média não aumentará ou decrescerá se ela apostar mais uma vez.

Proposição 1.9. *Seja $\{Y_n\}$ um martingal (ou submartingal) com respeito à informação $\{\mathcal{F}_n\}$. Se ϕ é função contínua e convexa tal que, $\forall n$, $E(\phi(Y_n)^+) < \infty$, então o processo $\{\phi(Y_n)\}$ define um submartingal com respeito à $\{\mathcal{F}_n\}$.*

Demonstração: As propriedades de integrabilidade e mensurabilidade seguem pela condição $E(\phi(Y_n)^+) < \infty$. Além disso, a Desigualdade de Jensen afirma que

$$E(\phi(Y_{n+1})|\mathcal{F}_n) \geq \phi(E(Y_{n+1}|\mathcal{F}_n)) \geq \phi(Y_n).$$

2. Teoremas da Parada Opcional

Definição 2.1. Um **tempo de parada** com respeito à filtragem $\{\mathcal{F}_n\}_{n\geq 1}$ (ou com respeito ao processo estocástico a tempo discreto $\{X_n\}_{n\geq 1}$) é uma variável aleatória τ com valores sobre $\mathbb{N} \cup \{+\infty\}$ satisfazendo a propriedade de que, para todo $n \in \mathbb{N}$, o evento $\{\tau = n\}$ é $\mathcal{F}_n$-mensurável.

Observamos que um tempo de parada T pode ser $+\infty$. Por exemplo, se T representa o tempo de espera até que um evento ocorra e o evento nunca acontecer, então é natural dizer que $T = +\infty$.

Exemplo 2.1. (Tempos de retorno). Seja $\{X_n\}_{n\geq 1}$ uma cadeia de Markov com espaço de estados $\mathcal{S}$. Considere o tempo de retorno ao estado $i \in \mathcal{S}$, ou seja, a variável aleatória $T_i := \inf\{n \geq 1; X_n = i\}$ com $T_i := +\infty$ se, $\forall n \geq 1$, $X_n \neq i$. Então, T_i é um tempo de parada. Com efeito, temos o evento $\{T_i = n\} = \{X_1 \neq i, \ldots, X_{n-1} \neq i, X_n = i\}$, o qual é $\mathcal{F}_n$-mensurável.

Exemplo 2.2. Se, $\forall n \in \mathbb{N} \cup \{+\infty\}$, o evento $\{\tau \leq n\}$ (ou seu complementar, $\{\tau > n\}$) é $\mathcal{F}_n$-mensurável, então τ é tempo de parada. Vale também a recíproca. Isto ocorre, pois $\{\tau \leq n\} = \bigcup_{k=1}^{n} \{\tau = n\}$ e $\{\tau = n\} = \{\tau \leq n\} - \{\tau \leq n-1\}$.

Exemplo 2.3. Seja um processo estocástico $\{X_n\}_{n\geq 1}$ com espaço de estados $\mathcal{S}$. Dado $A \subseteq \mathcal{S}$, a variável aleatória $T_A := \sup\{n \geq 1; X_n \in A\}$, o último tempo que o processo visita A, não é tempo de parada. Ela depende do tempo futuro, de tal forma que $\{T_A = n\}$ não é $\mathcal{F}_n$-mensurável.

Proposição 2.1. *Sejam S e T tempos de parada. Então, $S + T$, $S \wedge T := \min\{S, T\}$ e $S \vee T := \max\{S, T\}$ também são tempos de parada.*

Demonstração: Garantimos mensurabilidade analisando as funções indicadoras. Temos que $1_{\{S+T=n\}} = \sum_{k=0}^{n} 1_{\{S=k\}} 1_{\{T=n-k\}}$, $1_{\{S\wedge T>n\}} = 1_{S>n} 1_{T>n}$ e $1_{\{S\vee T\leq n\}} = 1_{S\leq n} 1_{T\leq n}$. Como as operações entre as funções indicadas levam a funções mensuráveis, chegamos ao resultado desejado. ∎

Se $\{Y_n\}_{n\geq 1}$ é um martingal, então $\{Y_n\}_{n\geq 1}$ e $\{-Y_n\}_{n\geq 1}$ são supermartingais. Assim, as proposições abaixo que se referem a supermartingais podem ser estendidas para martingais.

Lema 2.2. *Sejam $\{Y_n\}_{n\geq 1}$ um supermartingal e T um tempo de parada relativos a $\{X_n\}_{n\geq 1}$. Então, $\forall n \geq k$, segue que $E(Y_n 1_{\{T=k\}}) \leq E(Y_k 1_{\{T=k\}})$.*

Demonstração: Aplicando a Lei da Esperança Total, segue que

$$E(Y_n 1_{\{T=k\}}) = E(E(Y_n 1_{\{T=k\}} | X_0, \ldots, X_k)) = E(1_{\{T=k\}} E(Y_n | \mathcal{F}_k)) \leq E(Y_k 1_{\{T=k\}}).$$

Note que, para um martingal, temos $E(Y_n 1_{\{T=k\}}) = E(Y_k 1_{\{T=k\}})$. ∎

Lema 2.3. *Sejam $\{Y_n\}_{n\geq 1}$ um supermartingal e T um tempo de parada relativo a $\{\mathcal{F}_n\}_{n\geq 1}$. Então, $\forall n \geq 1$, temos que $E(Y_0) \geq E(Y_{T\wedge n}) \geq E(Y_n)$.*

Demonstração: Pela Lema 2.2,

$$\begin{aligned} E(Y_{T\wedge n}) &= \sum_{k=0}^{n-1} E(Y_T 1_{\{T=k\}}) + E(Y_n 1_{\{T\geq n\}}) \\ &= \sum_{k=0}^{n-1} E(Y_k 1_{\{T=k\}}) + E(Y_n 1_{\{T\geq n\}}) \text{ [pois } Y_T = Y_k \text{ quando } T = k] \\ &\geq \sum_{k=0}^{n-1} E(Y_n 1_{\{T=k\}}) + E(Y_n 1_{\{T\geq n\}}) \text{ [devido ao Lema]} \\ &= E(Y_n), \end{aligned}$$

Considerando que $\{Y_n\}$ é um supermartingal, já obtemos a relação $E(Y_{T\wedge n}) \geq E(Y_n)$, restando apenas mostrar que $E(Y_{T\wedge n}) \leq E(Y_0)$. Com efeito, como $E|Y_n| \leq \infty$ pela mensurabilidade na definição de supermartingais, podemos definir o martingal $W_n := \sum_{k=1}^{n} [Y_k - E(Y_k|\mathcal{F}_{k-1})]$, com $W_0 = 0$. De fato,

$$\begin{aligned} E(W_{n+1}|\mathcal{F}_n) &= \sum_{k=1}^{n} [E(Y_k|\mathcal{F}_n) - E(E(Y_k|\mathcal{F}_{k-1})|\mathcal{F}_n)] + E(Y_{n+1}|\mathcal{F}_n) - E(E(Y_{n+1}|\mathcal{F}_n)|\mathcal{F}_n) \\ &= \sum_{k=1}^{n} [Y_k - E(Y_k|\mathcal{F}_{k-1})] + Y_n - E(Y_n|\mathcal{F}_n) \\ &= \sum_{k=1}^{n} [Y_k - E(Y_k|\mathcal{F}_{k-1})] + Y_n - Y_n \\ &= W_n. \end{aligned}$$

Sendo um martingal, $E(W_{T\wedge n}) = E(W_n)$ e, como $E(W_0) = E(W_n)$, segue $E(W_{T\wedge n}) = E(W_0)$, ou seja:

$$\begin{aligned} 0 = E(W_0) &= E(W_{T\wedge n}) \\ &= E\left(\sum_{k=1}^{T\wedge n} [Y_k - E(Y_k|\mathcal{F}_{k-1})]\right) \\ &\geq E\left(\sum_{k=1}^{T\wedge n} [Y_k - Y_{k-1}]\right) \text{ [}\{Y_n\}\text{ é supermartingal]} \\ &= E(Y_{T\wedge n}) - E(Y_0) \text{ [soma telescópica]} \\ \Rightarrow E(Y_0) &\geq E(Y_{T\wedge n}), \end{aligned}$$

como queríamos. ∎

Lema 2.4. *Seja W variável aleatória integrável, $E|W| < \infty$; e T um tempo de parada satisfazendo $P(T < \infty) = 1$. Então, temos que $\lim_{n\to\infty} E(W1_{\{T>n\}}) = 0$ e $\lim_{n\to\infty} E(W1_{\{T\leq n\}}) = E(W)$.*

Demonstração: Observamos que, pelo Teorema da Convergência Monótona,

$$E|W| \geq E(|W|1_{\{T\leq n\}}) = \sum_{k=0}^{n} E(|W| \mid T = k)P(T = k) \to \sum_{k=0}^{\infty} E(|W||T = k)P(T = k) = E|W|$$

quando $n \to \infty$. Portanto, $\lim_{n\to\infty} E(|W|1_{\{T\leq n\}}) = E|W|$ e $\lim_{n\to\infty} E(|W|1_{\{T>n\}}) = 0$. Segue que, fixado um $\epsilon > 0$, existe um N suficientemente grande tal que, para todo $n > N$,

$$0 \leq |E(W) - E(W1_{\{T\leq n\}})| = |E(W1_{\{T>n\}})| \leq E(|W|1_{\{T>n\}}) < \epsilon.$$

Finaliza-se a demonstração tomando $\epsilon \downarrow 0$. ■

Apresentamos, então, os resultados principais desta seção:

Teorema 2.5. ***(Parada Opcional).*** *Sejam $\{Y_n\}_{n\geq 0}$ um supermartingal e T um tempo de parada, ambos relativos à filtragem $\{\mathcal{F}_n\}_{n\geq 0}$. Então, afirmamos que $E(Y_T) \leq E(Y_0)$ se um dos seguintes itens é verdadeiro:*

1. *Se $P(T < \infty) = 1$ e existe uma variável aleatória $W \geq 0$ integrável satisfazendo $Y_{T\wedge n} > -W$, $\forall n \in \mathbb{N}$;*

2. *Se T é integrável (ou seja, $E(T) < \infty$) e existe $k < \infty$ tal que $E(|Y_{n+1} - Y_n|1_{\{T>n\}}|\mathcal{F}_n) \leq k$, $\forall n \in \mathbb{N}$, quase certamente.*

Em particular, para estendermos o item 1 para martingais, precisamos encontrar uma variável aleatória $W \geq 0$ integrável que domine o processo, $|Y_{T\wedge n}| \leq W$. Um exemplo é tomar uma constante, $W \equiv k$. Desta forma, garante-se integrabilidade: $E\left(\sup_{n\geq 0}\{|Y_{T\wedge n}|\}\right) < \infty$.

Demonstração: Vamos supor que as hipóteses do item (1) são verdadeiras. Fixe um inteiro positivo c e defina o processo estocástico $Y_n^{(c)} := \min\{c, Y_n\}$, $\forall n \geq 0$. Então, $\{Y_n^{(c)}\}$ também é supermartingal com respeito à filtragem $\{\mathcal{F}_n\}$: por um lado, $E\left(Y_{n+1}^{(c)}|\mathcal{F}_n\right) \leq E(Y_{n+1}|\mathcal{F}_n) \leq Y_n$, por outro $E\left(Y_{n+1}^{(c)}|\mathcal{F}_n\right) \leq E(c|\mathcal{F}_n) = c$. Ou seja, $E\left(Y_{n+1}^{(c)}|\mathcal{F}_n\right) \leq Y_n^{(c)}$.

Pelo Lema 2.3, obtemos que $E(Y_0^{(c)}) \geq E(Y_{T\wedge n}^{(c)})$. Notamos também que $|Y_{T\wedge n}^{(c)}| \leq \max\{c, W\}$, $\forall n$. Desta forma, pelo Teorema da Convergência Dominada, pela integrabilidade destas variáveis aleatórias, tomamos $n \to \infty$ de tal forma que $E(Y_0^{(c)}) \geq E(Y_T^{(c)})$. Trivialmente

$E(Y_0) \geq E(Y_0^{(c)})$, e podemos aplicar novamente o Teorema da Convergência Dominada para obtermos $\lim_{c\to\infty} E(Y_T^{(c)}) = E(Y_T)$. Desta forma, $E(Y_0) \geq E(Y_T)$.

Agora, vamos assumir as hipóteses do item (2). Defina a variável aleatória $U := |Y_0| + \sum_{k=0}^{\infty} |Y_{k+1} - Y_k| 1_{\{T>k\}}$. Visualizamos que esta variável satisfaz $|Y_{T\wedge n}| \leq U$, $\forall n \geq 0$. Tomando a esperança, temos o que segue.

$$\begin{aligned} E(U) \leq & \; E|Y_0| + \sum_k E(|Y_{k+1} - Y_k| 1_{\{T>k\}}) \\ = & \; E|Y_0| + \sum_k E(E(|Y_{k+1} - Y_k| 1_{\{T>k\}} | \mathcal{F}_k)) \\ < & \; E|Y_0| + c \sum_k P(T > k) \\ = & \; E|Y_0| + cE(T) < \infty \end{aligned}$$

Aqui, usamos o fato de que

$$E(T) = \sum_{i=1}^{\infty} iP(T=i) = \sum_{i=1}^{\infty} \sum_{k=0}^{i-1} P(T=i) = \sum_{k=0}^{\infty} \sum_{i=k+1}^{\infty} P(T=i) = \sum_{k=0}^{\infty} P(T>k).$$

Da integrabilidade da variável aleatória U, aplicamos o Teorema da Convergência Dominada, de tal forma que $E(Y_0) \geq \lim_{n\to\infty} E(Y_{T\wedge n}) = E(Y_T)$. ∎

Corolário 2.6 (Parada Opcional para martingais)**.** *Sejam $\{Y_n\}_{n\geq 1}$ um martingal e T um tempo de parada, ambos relativos à filtragem $\{\mathcal{F}_n\}_{n\geq 1}$. Sob as hipóteses $P(T < \infty) = 1$, $E|Y_T| < \infty$ e $\lim_{n\to\infty} E(Y_n 1_{\{T>n\}}) = 0$, temos que $E(Y_T) = E(Y_0)$.*

Demonstração: Destacamos, inicialmente, que a hipótese $E|Y_T| < \infty$ é importante, pois a integrabilidade da definição de martingal não garante a integrabilidade de Y_T. Como

$$E(Y_{T\wedge n}) = E(Y_{T\wedge n} 1_{\{T\leq n\}}) + E(Y_{T\wedge n} 1_{\{T>n\}}) = E(Y_T 1_{\{T\leq n\}}) + E(Y_n 1_{\{T>n\}}),$$

temos que

$$E(Y_T) = E(Y_T 1_{\{T\leq n\}}) + E(Y_T 1_{\{T>n\}}) = E(Y_{T\wedge n}) - E(Y_n 1_{\{T>n\}}) + E(Y_T 1_{\{T>n\}}).$$

Pelo Lema 2.3, segue $E(Y_{T\wedge n}) = E(Y_0)$. Por hipótese, $\lim_{n\to\infty} E(Y_n 1_{\{T>n\}}) = 0$. Pelo Lema 2.4 e pela hipótese de integrabilidade de Y_T, podemos fixar $W = Y_T$, de tal forma que

$\lim_{n\to\infty} E(Y_T 1_{\{T>n\}}) = 0$. Logo, concluímos que $E(Y_T) = \lim_{n\to\infty} E(Y_{T\wedge n}) = E(Y_0)$. ■

Corolário 2.7. ***(Identidade de Wald)*** *Sejam $\{X_n\}_{n\geq 1}$ variáveis aleatórias independentes e identicamente distribuídas, não degeneradas, com função geradora de momentos $M(u) := E(e^{X_1 u})$ definida e finita para pontos em algum intervalo aberto contendo a origem. Defina $S_0 := 0$ e $S_n := \sum_{k=1}^{n} X_k$. Fixe $a, b > 0$ de modo a definir o tempo de parada $T := \inf\{n; S_n \leq -a \text{ ou } S_n \geq b\}$. Sob tais hipóteses,*

1. *$E([M(u)]^{-T} e^{u\cdot S_T}) = 1$, $\forall u$ no intervalo aberto tal que $M(u) \geq 1$;*

2. *Vale a **Identidade de Wald**: $E(S_T) = E(X_1)E(T)$.*

Demonstração: O processo estocástico $\{Y_n\}$ dado por $Y_n := [M(u)]^{-n} e^{u\cdot S_n}$ define um martingal com respeito a $\{X_n\}$. De fato, como a função geradora de momentos é definida e finita em um intervalo aberto contendo a origem e

$$\begin{aligned} E(Y_{n+1}|\mathcal{F}_n) = \; & E([M(u)]^{-(n+1)} e^{u\cdot S_{n+1}}|\mathcal{F}_n) \;=\; [M(u)]^{-(n+1)} E(e^{u(X_1+\cdots+X_n+X_{n+1})}|\mathcal{F}_n) \\ = \; & [M(u)]^{-(n+1)} E(e^{uS_n}|\mathcal{F}_n) E(e^{uX_{n+1}}|\mathcal{F}_n) \;=\; [M(u)]^{-(n+1)} e^{uS_n} E(e^{uX_{n+1}}) \\ = \; & [M(u)]^{-(n+1)} e^{uS_n} M(u) \;=\; [M(u)]^{-n} e^{uS_n} = Y_n, \end{aligned}$$

segue da definição. Notamos que os incrementos são limitados: de fato,

$$E(|Y_{n+1} - Y_n||\mathcal{F}_n) = Y_n E(|M(u)^{-1} e^{uX_{n+1}} - 1|)$$

e, como assumimos $M(u)^{-n} \leq 1$ e, para todo $n < T$, $e^{uS_n} \leq e^b$; concluí-se que $Y_n \leq e^b$, $\forall n < T$. Além disso,

$$\begin{aligned} E(|M(u)^{-1} e^{uX_{n+1}} - 1|) = \; & E(|M(u)^{-1}||e^{uX_{n+1}} - M(u)|) \\ \leq \; & M(u)^{-1} E(|e^{uX_{n+1}}| + |M(u)|) \\ = \; & M(u)^{-1}[E(e^{uX_{n+1}}) + M(u)] = M(u)^{-1} 2M(u) = 2, \end{aligned}$$

de tal forma que $E(|Y_{n+1} - Y_n| 1_{\{n<T\}}|\mathcal{F}_n) \leq 2e^b$. Agora, resta mostrarmos que $E(T) < \infty$. Assim, consideremos $c := a + b$. Como X_k é não degenerada, existem $N \in \mathbb{N}$ e $\delta > 0$ tais que $P(|S_n| > c) < \delta$, para todo $n \geq N$.

Definindo as variáveis aleatórias $S'_1 := S_N$, $S'_2 := S_{2N} - S_N, \ldots, S'_k := S_{kN} - S_{(k-1)N}$, observamos que $P(T \geq kN) \leq P(|S'_1| \leq c) \ldots P(|S'_k| \leq c) \leq (1-\delta)^k$. Portanto, a sequência $P(T \leq n)$ decresce em n, de tal forma que

$$E(T) = \sum_{n=1}^{\infty} P(T \geq n) \leq N \sum_{k=0}^{\infty} P(T \geq kN) \leq N \cdot \frac{1}{1-(1-\delta)} = \frac{N}{\delta}.$$

Para (1), basta aplicarmos o Teorema da Parada Opcional, de tal forma que

$$E(Y_T) = E(Y_0) \Rightarrow E(M(u)^{-T} e^{uS_T}) = E(M(u)^0 e^{u \cdot 0}) = 1.$$

Para (2), derivamos a relação indicada com respeito à u, levando em conta a integrabilidade de Y_T e as propriedades $M(0) = 1$ e $M'(0) = E(X_1)$ da função geradora de momentos:

$$\begin{aligned}
& \frac{d}{du} E([M(u)]^{-T} e^{uS_T}) = E\left(\frac{d}{du}[M(u)]^{-T} e^{uS_T}\right) = \frac{d}{du}(1) \\
\Rightarrow \quad & E(e^{uS_T}[-TM(u)^{-(T+1)} M'(u) + M(u)^{-T} S_T]) = 0 \\
\Rightarrow \quad & -M'(u) E(T \cdot M(u)^{-(T+1)} e^{uS_T}) + E(Y(u)^{-T} S_T e^{uS_T}) = 0, \text{ fixe } u = 0 \\
\Rightarrow \quad & -E(X_1)E(T) + E(S_T) = 0 \\
\Rightarrow \quad & E(S_T) = E(X_1)E(T).
\end{aligned}$$

■

Em relação ao Corolário 2.7, podemos também obter propriedades da variância das variáveis aleatórias S_T e T derivando a expressão do item (1) duas vezes.

Exemplo 2.4. (Passeio aleatório assimétrico) Sejam $\{X_n\}_{n\geq 1}$ variáveis aleatórias i.i.d. tais que $P(X_1 = k) = p \cdot 1_{\{k=1\}} + q \cdot 1_{\{k=-1\}}$, em que $q = 1 - p$. Considere $p > \frac{1}{2}$ e defina o processo estocástico $\{S_n\}_{n\geq 0}$, com $S_0 = 0$ e $S_n = \sum_{i=1}^{n} X_i$, $n \geq 1$. Considere o tempo de parada $T := \inf\{n \in \mathbb{N}; S_n \geq b\}$, com $b > 0$ fixado.

Sabemos que, sendo $M(u) = E[e^{X_1 u}] = p \cdot e^u + q \cdot e^{-u}$ a função geradora de momentos das $\{Y_n\}$, o processo $Y_n = [M(u)]^{-n} e^{uS_n}$ define martingal em relação à $\{X_n\}$ com incrementos limitados. Além disso, a variável aleatória T é integrável, pela propriedade de assimetria deste

passeio aleatório. Segue, pelo Teorema 2.5, que

$$1 = E(Y_0) = E(Y_T) = E([M(u)]^{-T} e^{uS_T}) = E([pe^u + qe^{-u}]^{-T} e^{ub})$$
$$\Rightarrow \quad E\left(\left[\frac{1}{pe^u + qe^{-u}}\right]^T\right) = (e^{-u})^b.$$

Definindo $s := \dfrac{1}{pe^u + qe^{-u}}$, podemos escrever e^{-u} em função de s:

$$s = \frac{1}{pe^u + qe^{-u}} \Rightarrow sqe^{-u} - 1 + spe^u = 0 \Rightarrow sqe^{-2u} - e^{-u} + sp = 0$$
$$\Rightarrow \left(e^{-u} - \frac{1 + \sqrt{1 - 4pqs^2}}{2qs}\right)\left(e^{-u} - \frac{1 - \sqrt{1 - 4pqs^2}}{2qs}\right) = 0$$

Vamos considerar que $0 < s \leq 1$, de tal forma que $e^{-u} = \dfrac{1 - \sqrt{1 - 4pqs^2}}{2qs}$. Segue, pois, a seguinte função geradora de probabilidades para o tempo de parada T:

$$E(s^T) = \left(\frac{1 - \sqrt{1 - 4pqs^2}}{2qs}\right)^b, \text{ para } 0 < s \leq 1.$$

Por definição, $E(s^T) = \sum_{t=0}^{\infty} P(T = t)s^t$, que é série absolutamente convergente para $0 < s \leq 1$. Assim, se fizermos a Expansão de Taylor do termo acima, obteremos as probabilidades de que o tempo de parada assuma cada um dos valores $0, 1, \ldots$ Para $b = 1$, o suporte de T está nos números ímpares, de tal forma que

$$E(s^T) = ps + p^2qs^3 + 2p^3q^2s + 5p^4q^3s^7 + 14p^5q^4s^9 + 42p^6q^5 + o(s^{12}).$$

Pela identidade de Wald (2.7), segue que

$$E(S_T) = E(X_1)E(T) \Rightarrow b = (p - q)E(T) \Rightarrow E(T) = \frac{b}{p - q}.$$

No que se refere à variância, segue que

$$
\begin{aligned}
& -M'(u)E(T\cdot[M(u)]^{-(T+1)}e^{uS_t}) + E([M(u)]^{-T}S_Te^{uS_T}) = 0 \\
\Rightarrow\quad & \frac{d}{du}\{-M'(u)E(T\cdot[M(u)]^{-(T+1)}e^{uS_t}) + E([M(u)]^{-T}S_Te^{uS_T})\} = \frac{d}{du}(0) = 0 \\
\Rightarrow\quad & -M''(u)E(T\cdot[M(u)]^{-(T+1)}e^{uS_T}) + [M'(u)]^2E(T(T+1)[M(u)]^{-(T+2)}e^{uS_T}) \\
& -2M'(u)E(T[M(u)]^{-(T+1)}S_Te^{uS_T}) + E(Y(u)^{-T}S_T^2e^{uS_T}) = 0, \text{ fixe } u = 0 \\
\Rightarrow\quad & -E(X_1^2)E(T) + [E(X_1)]^2E(T(T+1)) - 2E(X_1)E(TS_T) + E(S_T^2) = 0 \\
\Rightarrow\quad & E(T^2) = \frac{\mathrm{Var}(X_1)E(T) - 2E(X_1)E(TS_T) + E(S_T^2)}{[E(X_1)]^2}.
\end{aligned}
$$

Como $S_T = b$, segue $E(TS_T) = bE(T)$ e $E(S_T^2) = b^2$, de tal forma que

$$
\begin{aligned}
\mathrm{Var}(T) = \;& E(T^2) - [E(T)]^2 \\
= \;& \frac{(p+q-(p-q)^2)(\frac{b}{p-q}) - 2(p-q)(\frac{b^2}{p-q}) + b^2}{(p-q)^2} - \frac{b^2}{(p-q)^2} \\
= \;& \frac{b(1-(p-q)^2)}{(p-q)^3}.
\end{aligned}
$$

3. Integrabilidade Uniforme de Variáveis Aleatórias

3.1. Convergência de variáveis aleatórias

Em um primeiro instante, revisamos definições e proposições relacionadas à convergência de variáveis aleatórias. As demonstrações estão disponíveis no livro de Grimmett e Stirzaker (2020, Seção 7.2).

Definição 3.1. Sejam $(\Omega, \mathcal{F}, P)$ um espaço de probabilidade, $\{X_n\}_{n\geq 1}$ uma sequência de variáveis aleatórias e X uma variável aleatória, todas definidas neste espaço.

1. A sequência converge **quase certamente**, $X_n \xrightarrow{\text{q.c.}} X$, se

$$P(\{\omega \in \Omega \text{ tal que } X_n(\omega) \to X(\omega) \text{ quando } n \to \infty\}) = 1;$$

2. Para $1 \leq p < \infty$, a sequência converge no ***p*-ésimo momento** ou em L^p, $X_n \xrightarrow{L^p} X$ se $E|X_n|^p < \infty$ para todo $n \geq 1$, $E|X|^p < \infty$ e $E(|X_n - X|^p) \to 0$ quando $n \to \infty$. Se $p = 1$

nesta definição, dizemos que a sequência converge **em média**.

3. A sequência converge **em probabilidade**, $X_n \xrightarrow{\text{P}} X$, se, para todo $\alpha > 0$,

$$P(|X_n - X| > \alpha) \to 0 \text{ quando } n \to \infty;$$

4. A sequência converge **em distribuição**, $X_n \xrightarrow{\mathcal{D}} X$, se $P(X_n \leq x) \to P(X_n \leq x)$ quando $n \to \infty$ para todo $x \in A$, sendo A o conjunto dos pontos de continuidade da função de distribuição acumulada, $F_X(x) := P(X \leq x)$.

Teorema 3.1. *Considere uma sequência de variáveis aleatórias $\{X_n\}$ definidas em um mesmo espaço de probabilidade.*

1. *Convergência quase certa implica convergência em probabilidade:*

$$X_n \xrightarrow{q.c.} X \Rightarrow X_n \xrightarrow{P} X;$$

2. *Convergência em probabilidade implica convergência em distribuição:*

$$X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{\mathcal{D}} X;$$

3. *Para números reais $r > s \geq 1$, se $\{X_n\}$ converge no r-ésimo momento, então também converge no s-ésimo momento:*

$$E(|X_n - X|^r) \to 0 \Rightarrow E(|X_n - X|^s) \to 0;$$

4. *Convergência em média implica convergência em probabilidade:*

$$E(|X_n - X|) \to 0 \Rightarrow X_n \xrightarrow{P} X.$$

5. *Convergência em distribuição para uma constante $c \in \mathbb{R}$ implica convergência em probabilidade:*

$$X_n \xrightarrow{\mathcal{D}} c \Rightarrow X_n \xrightarrow{P} c;$$

6. *Convergência em probabilidade implica a existência de uma subsequência que converge quase certamente:*

$$X_n \xrightarrow{P} X \Rightarrow \exists\{X_{n_k}\}_{k\geq 1} \text{ tal que } X_{n_k} \xrightarrow{q.c.} X \text{ quando } k \to \infty;$$

7. *Se $\{X_n\}$ converge em probabilidade para a variável aleatória X e existe $k > 0$ tal que $P(|X_n| \leq k) = 1$, $\forall n$, então $X_n \xrightarrow{L^r} X$ para todo $r \geq 1$;*

8. *Se a sequência de números $P_n(\epsilon) := P(|X_n - X| > \epsilon)$ satisfaz $\sum_n P_n(\epsilon) < \infty$ para todo $\epsilon > 0$, então $X_n \xrightarrow{q.c.} X$.*

Teorema 3.2. *Considere duas sequências de variáveis aleatórias, $\{X_n\}$ e $\{Y_n\}$, definidas em um mesmo espaço de probabilidade. Seja $g : \mathbb{R}^2 \to \mathbb{R}$ uma função contínua.*

1. *Se $X_n \xrightarrow{q.c.} X$ e $Y_n \xrightarrow{q.c.} Y$, então $g(X_n, Y_n) \xrightarrow{q.c.} g(X,Y)$;*

2. *Se $X_n \xrightarrow{P} X$ e $Y_n \xrightarrow{P} Y$, então $g(X_n, Y_n) \xrightarrow{P} g(X,Y)$.*

Casos particulares da função g incluem $g(x, y) = x + y$ e $g(x, y) = xy$.

Teorema 3.3. ***(Representação de Skorohod)*** *Sejam $\{X_n\}$ uma sequência de variáveis aleatórias com distribuições $\{F_n\}$ e X uma variável aleatória com distribuição F. Se vale $X_n \xrightarrow{\mathcal{D}} X$ quando $n \to \infty$, então existe um espaço de probabilidade $(\Omega',\mathcal{F}',P')$ e variáveis aleatórias $Y_1, \ldots, Y_n, \ldots; Y : \Omega' \to \mathbb{R}$ tais que*

1. *$\{Y_n\}$ e Y têm distribuições $\{F_n\}$ e F, respectivamente;*

2. *$Y_n \xrightarrow{q.c.} Y$ quando $n \to \infty$.*

Teorema 3.4. ***(Teorema de Egorov)*** *Sejam $\{X_n\}$ e X variáveis aleatórias tais que $X_n \xrightarrow{q.c.} X$. Então, para todo $\epsilon > 0$, existe $A \in \mathcal{F}$ tal que $P(A) < \epsilon$ e $X_n \to X$ uniformemente em A^c.*

3.2. Integrabilidade Uniforme

Definição 3.2. Seja $(\Omega, \mathcal{F}, P)$ um espaço de probabilidade. Uma sequência de variáveis aleatórias $\{X_n\}_{n\geq 1}$ neste espaço é dita **uniformemente integrável** se

$$\lim_{a\to\infty} \sup_n E(|X_n|1_{\{|X_n|\geq a\}}) = 0.$$

Lema 3.5. *A sequência $\{X_n\}$ é uniformemente integrável se, e somente se,*

1. $\sup_n E|X_n| < \infty$*;*
2. *Para todo $\epsilon > 0$, existe $\delta > 0$ tal que, $\forall n$ e para todo evento $A \in \mathcal{F}$ tal que $P(A) < \delta$, vale $E(|X_n|1_A) < \epsilon$. (Neste caso, dizemos que $\{X_n\}$ é* ***uniformemente absolutamente integrável****).*

Demonstração: Assume-se, inicialmente, a integrabilidade uniforme de $\{X_n\}$. Para $a > 0$, temos que

$$E|X_n| = E(|X_n|1_{\{X_n<a\}} + E(|X_n|1_{\{X_n\geq a\}}) \Rightarrow \sup_n E|X_n| \leq a + \sup_n E(|X_n|1_{X_n\geq a}) < \infty,$$

por hipótese. Prova-se, pois, o item (1). Agora, defina os eventos $B_n(a) := \{|X_n| \geq a\}$ e escreva $E(|X_n|1_A) = E(|X_n|1_{A\cap B_n(a)}) + E(|X_n|1_{A\cap B_n^c(a)})$, sendo A o evento definido no item (2). Temos que $E(|X_n|1_{A\cap B_n(a)}) \leq E(|X_n|1_{B_n(a)})$ e, pela Desigualdade de Markov, $E(|X_n|1_{A\cap B_n^c(a)}) \leq a \cdot E(1_A) = a \cdot P(A)$.

Fixe $\epsilon > 0$ e escolha a tal que $E(|X_n|1_{B_n(a)}) < \frac{1}{2}\epsilon$, para todo n. Assim, colocando $\delta = \frac{\epsilon}{2a}$, segue $E(|X_n|1_A) \leq \frac{1}{2}\epsilon + a \cdot P(A) < \epsilon$, como queríamos.

Agora, assuma os itens (1) e (2). Fixe $\epsilon > 0$ e escolha $\delta > 0$ que garante $\{X_n\}$ uniformemente absolutamente integrável. Pela Desigualdade de Markov, temos que $E|X_n| \geq E(|X_n|1_{B_n(a)}) \geq a \cdot P(B_n(a))$. Consequentemente, $\sup_n P(B_n(a)) \leq \frac{1}{a}\sup_n E|X_n| < \infty$, pela hipótese (1). Basta, pois, escolher $a > 0$ tal que $\frac{\sup E|X_n|}{a} < \delta \Rightarrow P(B_n(a)) < \delta$, para todo n. Levando em conta o item (2), temos que $E(|X_n|1_{B_n(a)}) < \epsilon$, para todo n, isto é, $\{X_n\}$ é uniformemente integrável. ∎

O próximo teorema fornece uma condição simples para a integrabilidade uniforme.

Teorema 3.6. ***(da Bola de Cristal)*** *Seja $\{X_n\}$ uma sequência de variáveis aleatórias. Dado $p \geq 1$, se existe $\delta > 0$ tal que $\sup_n E(|X_n|^{p+\delta}) < \infty$, então as variáveis aleatórias $\{X_n^p\}$ são uniformemente integráveis.*

Demonstração: Temos o que segue:

$$\begin{aligned}
\sup_n E(|X_n^p| 1_{\{|X_n^p|>a\}}) &= \sup_n E(|X_n^p| \cdot 1_{\{|X_n/(a^{1/p})|>1\}}) \\
&= \sup_n E(|X_n^p| \cdot 1 \cdot 1_{\{|X_n^\delta|/(a^{\delta/p})>1\}}) \\
&\leq \sup_n E\left(|X_n^p| \frac{|X_n^\delta|}{a^{\delta/p}}\right) \\
&= a^{-\delta/p} \sup_n E(|X_n^{p+\delta}|).
\end{aligned}$$

Fazendo $a \to \infty$, segue que $a^{-\delta/p} \to 0$. Logo, pelo Teorema do Confronto,

$$\lim_{a\to\infty} \sup_n E(|X_n^p| 1_{\{|X_n^p|>a\}}) = 0,$$

garantindo a integrabilidade uniforme. ■

Teorema 3.7. ***(de Vitali)*** *Seja $(\Omega, \mathcal{F}, P)$ um espaço de probabilidade e variáveis aleatórias X e $\{X_n\}$. Então, as seguintes afirmações são equivalentes:*

1. *$X_n \xrightarrow{P} X$ e $\{X_n\}$ é uniformemente integrável;*
2. *X é integrável, assim como as variáveis X_n; e $X_n \xrightarrow{L^1} X$.*

Demonstração: $(1 \Rightarrow 2)$ Sendo uniformemente integrável, a sequência $\{X_n\}$ é limitada em $L^1(\Omega, \mathcal{F}, P) := \{X : \Omega \to \mathbb{R}$ tal que X é $\mathcal{F}$-mensurável e $E|X| < \infty\}$. O Lema de Fatou aplicado nas variáveis $\{X_n\}$ garante a integrabilidade da variável X.

Resta, pois, mostrar que $X_n \xrightarrow{L^1} X$. É suficiente mostrar que cada subsequência $\{X_{n_k}\}$ de $\{X_n\}$ contém uma subsequência $\{X_{n_{k_j}}\}$ convergente a X em $L^1(\Omega, \mathcal{F}, P)$. Tome $\{X_{n_{k_j}}\}$ tal que $X_{n_{k_j}} \xrightarrow{\text{q.c.}} X$, pelo item 6 do Teorema 3.1. Fixado $\epsilon > 0$, existe $\delta > 0$ tal que $E(|X_n| 1_A) < \epsilon$ para todo n e para todo $A \in \mathcal{F}$ tal que $P(A) < \delta$. Aplicando o Lema de Fatou, segue que $E(|X| 1_A) < \epsilon$ sempre que $P(A) < \delta$.

Pelo Teorema de Egorov, existe $A \in \mathcal{F}$ tal que $P(A) < \delta$ e $\{X_{n_{k_j}}\}$ converge uniformemente sobre $\Omega - A$ para a variável X. Seja $N \in \mathbb{N}$ tal que $\sup_{\Omega - A} |X_{n_{k_j}} - X| < \epsilon$ para $k \geq N$. Então,

$$\int_\Omega |X_{n_{k_j}} - X| dP = E|X_{n_{k_j}} - X| < \epsilon + E(|X_{n_{k_j}}|1_A) + E(|X|1_A) < 3\epsilon.$$

Portanto, fazendo $\epsilon \downarrow 0$, concluímos que $X_{n_{k_j}} \xrightarrow{L^1} X$ e, consequentemente, $X_n \xrightarrow{L^1} X$.

$(2 \implies 1)$ Como $\{X_n\}$ converge em média, então, pelo item 4 do Teorema 3.1, há convergência em probabilidade. Havendo convergência em média, as variáveis são limitadas em $L^1(\Omega, \mathcal{F}, P)$. Consequentemente, $\sup_n E|X_n| < \infty$. Além disso, $E(|X_n|1_A) \leq E(|X_n - X|1_A) + E(|X|1_A)$, o que garante, pelas hipóteses, que $\{X_n\}$ é uniformemente absolutamente integrável. Então, pelo Lema 3.5, segue a integrabilidade uniforme da sequência. ■

Exemplo 3.1. Se $\{X_n\}$ é uma sequência de variáveis aleatórias com $E(X_n) = 0$ e $\text{Var}(X_n) = 1$ para todo n, então $\{X_n\}$ é uniformemente integrável, pelo Teorema da Bola de Cristal.

Exemplo 3.2. (Família Dominada). Se $\{X_t\}_{t \in T}$ é um processo estocástico dominado por uma variável aleatória Y integrável, isto é, $|X_t| \leq Y, \forall t \in T$; então $\{X_t\}$ é uniformemente integrável, pois

$$\sup_{t \in T} E(|X_t|1_{\{|X_t| > a\}}) < E(|Y|1_{\{|Y| > a\}}) \to 0 \text{ quando } a \to \infty.$$

Exemplo 3.3. Seja $\{X_n\}_{n \geq 1}$ uma sequência de variáveis aleatórias definidas por

$$X_n = \begin{cases} 0, & \text{com probabilidade } 1 - \dfrac{1}{n}; \\ n, & \text{com probabilidade } \dfrac{1}{n}. \end{cases}$$

Assim, apesar de $\sup_{n \geq 1} E|X_n| = 1$, a sequência não é uniformemente integrável:

$$E(|X_n|1_{\{|X_n| > a\}}) = \begin{cases} 1, & \text{se } a \leq n \\ 0, & \text{se } a > n \end{cases}$$

de tal forma que, $\forall a > 0$, $\sup_{n \geq 1} E(|X_n|1_{\{|X_n| > a\}}) = 1 \neq 0$. Destaca-se que não vale o Teorema de Vitali: apesar do fato que $X_n \xrightarrow{\text{P}} 0$, não verificamos a convergência em média.

Com base na teoria desenvolvida acima, podemos estabelecer mais um resultado relacionando martingais a tempos de parada.

Teorema 3.8. ***(da Amostragem Opcional).*** *Seja $\{Y_n\}$ um martingal uniformemente integrável com respeito a $\{\mathcal{F}_n\}$ e T um tempo de parada tal que $P(T < \infty) = 1$ e $E|Y_T| < \infty$. Então, $E(Y_T) = E(Y_0)$.*

Demonstração: O teorema é uma reformulação do Corolário 2.6. ■

A vantagem de introduzirmos a integrabilidade uniforme encontra-se na possibilidade de utilizar os resultados discutidos nesta seção. Por exemplo, se $\{X_n\}_{n\geq 1}$ é uma sequência tal que existe uma constante $C < \infty$ satisfazendo, para todo $n \geq 1$, $E(X_n^2) < C$; então ela é uniformemente integrável, como afirma o Teorema da Bola de Cristal.

Exemplo 3.4. (Processos de Ramificação). Consideremos $\{X_n\}_{n\geq 0}$ o processo estocástico definido por um processo de ramificação de Bienaymé–Galton–Watson em que o número de descendentes gerados por um indivíduo tem média $\mu > 1$ e variância σ^2. Vamos mostrar que $\{Y_n\}$, com $Y_n = \mu^{-n} X_n$, define um martingal uniformemente integrável. Com efeito, supondo que $\{Z_{nk}\}$ são variáveis i.i.d. indicando o total de indivíduos gerados pelo k-ésimo indivíduo da n-ésima geração, temos o que segue:

$$\begin{aligned} E(Y_{n+1}|\mathcal{F}_n) &= E(\mu^{-(n+1)} X_{n+1} \mid X_n) \quad \text{[pois o processo é markoviano]} \\ &= \mu^{-(n+1)} E\left(\sum_{k=1}^{X_n} Z_{nk} \mid X_n\right) = \mu^{-(n+1)} X_n \, \mu = \mu^{-n} X_n = Y_n. \end{aligned}$$

Para provarmos que o processo é uniformemente integrável, mostramos que existe uma constante $C > 0$ tal que, para todo n, $E(Y_n^2) < C$. De fato, temos que

$$E(Y_n^2) = \mu^{-2n} E(X_n^2) = \mu^{-2n}\{\mathrm{Var}(X_n) + [E(X_n)]^2\} = \mu^{-2n}\left[\sigma^2 \mu^{n-1}\left(\frac{1-\mu^n}{1-\mu}\right) + \mu^{2n}\right],$$

de tal modo que $\lim_{n\to\infty} E(Y_n^2) = 1 + \dfrac{\sigma^2}{\mu(\mu-1)}$. Como a sequência é convergente, concluímos que ela é limitada. Logo, pelo Teorema da Bola de Cristal, o martingal $\{Y_n\}$ é uniformemente integrável.

4. Convergência de Martingais

Definição 4.1. Sejam $\{x_n\}_{n\geq 0}$ uma sequência de números reais e $[a,b]$ um intervalo. Definimos $T_1 = \min\{n; x_n \leq a\}$ o tempo da primeira chegada em $(-\infty,a]$; e $T_2 = \min\{n > T_1; x_n \geq b\}$ o tempo da primeira chegada após T_1 em $[b,\infty)$. O intervalo $[T_1, T_2]$ é dito uma **subida** (*upcrossing* ou cruzamento para cima) de $[a,b]$.

Se, para $k \geq 2$, definirmos $T_{2k-1} = \min\{n > T_{2k-2}; x_n \leq a\}$ e $T_{2k} = \min\{n > T_{2k-1}; x_n \geq b\}$, temos a sequência de subidas $[T_{2k-1}, T_{2k}]$.

Definimos também as quantidades $U_n(a,b;x)$ como o número total de subidas de $[a,b]$ pela subsequência finita $\{x_0, x_1, \ldots, x_n\}$; e $U(a,b;x) = \lim_{n\to\infty} U_n(a,b;x)$ como o número total de subidas de $[a,b]$ por x.

Lema 4.1. *Considere a sequência $\{x_n\}$. Se $U(a,b;x) < \infty$ para quaisquer a, b números racionais tais que $a < b$, então $\lim x_n$ existe (mas pode ser em $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty,\infty\}$).*

Demonstração: Defina os limites $l_1 := \liminf x_n$ e $l_2 := \limsup x_n$, com $l_1 \leq l_2$. Suponha, por absurdo, que l_1 é estritamente menor que l_2. Pela densidade dos números racionais em $\mathbb{R}$, existem a, b racionais tais que $l_1 < a < b < l_2$. Portanto, $x_n \leq a$ para infinitos n, assim como $x_n > b$ para infinitos n. Segue que $U(a,b;x) = \infty$. ∎

Teorema 4.2. ***(Desigualdade de Snell para subidas)*** *Seja $\{Y_n\}$ um submartingal com respeito a $\{\mathcal{F}_n\}$. Neste caso, as quantidades descritas na Definição 4.1 tornam-se variáveis aleatórias.*

Para $a < b$, temos que

$$E(U_n(a,b;Y)) \leq \frac{E((Y_n - a)^+)}{b-a},$$

onde $(\cdot)^+$ é a função parte positiva.

Demonstração: Defina $Z_n = (Y_n - a)^+$. Então, $\{Z_n\}$ é submartingal não negativo com respeito à $\{\mathcal{F}_n\}$. De fato, pela convexidade da função valor absoluto, podemos aplicar a Desigualdade de Jensen, conforme a Proposição 1.9.

Assim, as subidas do processo $\{Y_n\}$ no intervalo $[a, b]$ são equivalentes às subidas do processo $\{Z_n\}$ em $[0, b-a]$, isto é, $U_n(a, b; Y) = U_n(0, b-a; Z)$ para todo n. Defina $[T_{2k-1}, T_{2k}]$, $k \geq 1$ os tempos de subidas de Z em $[0, b-a]$ e as variáveis indicadoras

$$I_j = \begin{cases} 1, & \text{se } j \in (T_{2k-1}, T_{2k}] \text{ para algum } k; \\ 0, & \text{caso contrário.} \end{cases}$$

As variáveis I_j são $\mathcal{F}_{j-1}$-mensuráveis, pois $\{I_j = 1\} = \bigcup_k \{T_{2k-1} \leq j-1\} - \{T_{2k} \leq j-1\} \in \sigma(Y_0, Y_1, \ldots, Y_{j-1})$. Segue, portanto, que $(b-a)\, U_n(0, b-a; Z) \leq \sum_{j=1}^{n} (Z_j - Z_{j-1}) I_j$, considerando que uma subida em $[0, b-a]$ possui distância de pelo menos $b-a$. Integrando,

$$\begin{aligned} E((Z_j - Z_{j-1})I_j) &= E(E((Z_j - Z_{j-1})I_j|\mathcal{F}_{j-1})) = E(I_j(E(Z_j|\mathcal{F}_{j-1}) - Z_{j-1})) \\ &\underset{\text{submartingal}}{\leq} E(E(Z_j|\mathcal{F}_{j-1}) - Z_{j-1}) = E(Z_j) - E(Z_{j-1}). \end{aligned}$$

Concluímos que

$$\begin{aligned} (b-a)\, E(U_n(0, b-a; Z)) = \; & (b-a)\, E(U_n(a, b; Y)) \leq \sum_{j=1}^{n} [E(Z_j) - E(Z_{j-1})] = E(Z_n) - E(Z_0) \\ \leq \; & E(Z_n) = E((Y_n - a)^+). \end{aligned}$$

■

Teorema 4.3. ***(Convergência de Martingais)*** *Seja $\{Y_n\}_{n\geq 0}$ um submartingal com respeito à filtragem $\{\mathcal{F}_n\}_{n\geq 0}$ satisfazendo $E(Y_n^+) \leq M$ para algum $M > 0$ e todo n (ou seja, é limitado superiormente: $\sup_{n\geq 0} E(Y_n^+) < \infty$). Então, existe uma variável aleatória Y_∞ tal que $Y_n \xrightarrow{q.c.} Y_\infty$ quando $n \to \infty$. Além disso,*

1. *Se $E|Y_0| < \infty$, então Y_∞ é integrável: $E(Y_\infty) < \infty$;*

2. *Se a sequência $\{Y_n\}_{n\geq 0}$ é uniformemente integrável, então também temos a convergência em média: $E(|Y_n - Y_\infty|) \to 0$. Em particular, se $\{Y_n\}$ é um martingal, então $E(Y_\infty) = E(Y_n)$ para todo n.*

Demonstração: Pela Desigualdade de Snell, temos, para todos $a < b$ números reais, a relação

$$E(U_n(a, b; Y)) \leq \frac{E(Y_n^+) + |a|}{b-a}.$$

Como as variáveis U_n são não negativas, segue pelo Teorema da Convergência Monótona que

$$E(U(a,b;y)) = \lim_{n\to\infty} E(U_n(a,b;Y)) \leq \frac{M+|a|}{b-a}.$$

Então, $U(a,b;Y) < \infty$ quase certamente, para quaisquer $a < b$ reais. Tomando a, b racionais, pelo Lema 4.1 a sequência Y_n converge quase certamente para um limite Y_∞. Resta-nos mostrar que $P(|Y_\infty| < \infty) = 1$.

Escrevendo $|Y_n| = 2Y_n^+ - Y_n$, como $\{|Y_n|\}$ define um submartingal, $E(Y_n|\mathcal{F}_0) \geq Y_0$ e, consequentemente,

$$E(|Y_n|\,|\mathcal{F}_0) = 2E(Y_n^+|\mathcal{F}_0) - E(Y_n|\mathcal{F}_0) \leq 2E(Y_n^+|\mathcal{F}_0) - Y_0.$$

Pelo Lema de Fatou,

$$E(|Y_\infty|\,|\mathcal{F}_0) = E(\liminf_{n\to\infty} |Y_n|\,|\mathcal{F}_0) \leq \liminf_{n\to\infty} E(|Y_n|\,|\mathcal{F}_0) \leq 2Z - Y_0,$$

onde definimos $Z := \liminf_{n\to\infty} E(Y_n^+|\mathcal{F}_0)$. Por hipótese, $E(Y_n^+) \leq M$ para todo n. Aplicando novamente o Lema de Fatou, segue que

$$E(Z) = E(\liminf_{n\to\infty} E(Y_n^+|\mathcal{F}_0)) \leq \liminf_{n\to\infty} E(Y_n^+|\mathcal{F}_0) \leq M.$$

Logo, $E(|Y_\infty|\,|\mathcal{F}_0) < \infty$ quase certamente, de tal forma que $P(|Y_\infty| < \infty|\mathcal{F}_0) = 1$ e, por conseguinte, $P(|Y_\infty| < \infty) = E(P(|Y_\infty| < \infty|\mathcal{F}_0)) = 1$. Como Y_0 é integrável por hipótese, verifica-se

$$E(|Y_\infty|\,|\mathcal{F}_0) \leq 2Z - Y_0 \Rightarrow E|Y_\infty| = E(E(|Y_\infty|\,|\mathcal{F}_0)) \leq 2M - E(Y_0) < \infty.$$

Por fim, supondo integrabilidade uniforme do submartingal $\{Y_n\}$, o Teorema de Vitali nos garante que Y_n converge em média para Y_∞. ■

Exemplo 4.1. (Crescimento exponencial) Retornando ao Exemplo 3.4 com estudos sobre o processo de ramificação de Bienaymé–Galton–Watson, vimos que $Y_n = \mu^{-n}X_n$ é um martingal uniformemente integrável. Pelo Teorema 4.3, a sequência $\{Y_n\}$ converge quase certamente e em

média para uma variável aleatória Y_∞, de tal forma que $E(Y_\infty) = E(Y_0) = Y_0 = X_0 = 1$.

Como $P(Y_\infty > 0) > 0$ pela teoria dos processos de ramificação, então $X_n \approx \mu^n Y_\infty$, ou seja, o tamanho populacional X_n cresce assintoticamente com velocidade exponencial de parâmetro μ.

Notamos, em particular, que a variável Y_∞ é integrável.

Exemplo 4.2. (Extinção) Consideremos novamente os processos de ramificações, mas agora com uma média de descendentes gerados $\mu \leq 1$. Pela teoria dos processos de ramificação, $X_n \xrightarrow{\text{q.c.}} 0$. Como $Y_n = \mu^{-n} X_n$ ainda é um martingal, notamos que a primeira parte do Teorema 4.3 ainda é válida, já que $Y_n \xrightarrow{\text{q.c.}} Y_\infty = 0$.

Porém, quando estudamos a convergência em média, notamos que $0 = E(Y_\infty) \neq E(Y_0) = 1$. Em particular, verificamos que $\{Y_n\}$ não é uniformemente integrável quando $\mu \leq 1$.

Exemplo 4.3. (Urna de Pólya) Uma urna contém bolas com duas cores distintas, sendo a azuis e b brancas. É feito o sorteio de uma das bolas, e em seguida retornamos ela e mais uma com esta mesma cor. O processo se repete sequencialmente. Pretende-se estudar a distribuição da proporção de bolas azuis ao longo dos ensaios e, em particular, uma distribuição assintótica.

Inicialmente, o processo pode ser visto como uma cadeia de Markov não homogênea: se X_n é a quantidade de bolas azuis após n inserções, então $X_0 = a$ e

$$p(x_{n+1}|X_n = k) = \begin{cases} k/(n+a+b), & \text{se } x_{n+1} = k+1 \\ (n+a+b-k)/(n+a+b), & \text{se } x_{n+1} = k. \end{cases}$$

A proporção de bolas azuis, $Y_n = \dfrac{X_n}{n+a+b}$, define um martingal. De fato, pela propriedade markoviana,

$$E(X_{n+1}|X_n) = X_n + \frac{X_n}{n+a+b} \Rightarrow$$
$$E(Y_{n+1}|\mathcal{F}_n) = E\left(\frac{X_{n+1}}{n+a+b+1}|X_n\right) = \frac{1}{n+a+b+1}\left(X_n + \frac{X_n}{n+a+b}\right) = \frac{X_n}{n+a+b} = Y_n.$$

O martingal $\{Y_n\}$ é uniformemente integrável, pois, sendo uma proporção, $0 \leq Y_n \leq 1, \forall n \Rightarrow \sup_n E(|Y_n| 1_{\{|Y_n| \geq c\}}) = 0$, para todo $c > 1$. Logo, podemos aplicar o Teorema 4.3 para afirmar a existência de uma variável aleatória Y_∞ indicando a proporção assintótica de bolas

azuis na urna, a qual satisfaz $Y_n \xrightarrow{qc} Y_\infty$ e $E(|Y_n - Y_\infty|) \to 0$.

Esta variável aleatória tem distribuição $Beta(a, b)$. Para mostrarmos isto, notamos que, sendo $N = n + a + b$ o total de bolas na urna após o n-ésimo ensaio,

$$P(X_n = k) = \binom{k-1}{a-1}\binom{N-k-1}{b-1}\left[\binom{N-1}{a+b-1}\right]^{-1} 1_{\{a\le k\le a+n\}}.$$

Este resultado segue por indução, pois $P(X_0 = k) = 1_{\{k=a\}}$ e, para $k = a, \ldots, a + n + 1$,

$$\begin{aligned}
P(X_{n+1} = k) = \;& P(X_{n+1} = k|X_n = k)P(X_n = k) + P(X_{n+1} = k|X_n = k-1)P(X_n = k-1) \\
= \;& \frac{N-k}{N}\tbinom{k-1}{a-1}\tbinom{N-k-1}{b-1}\left[\tbinom{N-1}{a+b-1}\right]^{-1} + \frac{k-1}{N}\tbinom{k-2}{a-1}\tbinom{N-k}{b-1}\left[\tbinom{N-1}{a+b-1}\right]^{-1} \\
= \;& \frac{(k-1)!}{(a-1)!(k-a)!}\frac{(N-k)!}{(b-1)!(N-k-b)!}\frac{(a+b-1)!(N-a-b)!}{N!} \\
& + \frac{(k-1)!}{(a-1)!(k-a-1)!}\frac{(N-k)!}{(b-1)!(N-k-b+1)!}\frac{(a+b-1)!(N-a-b)!}{N!} \\
= \;& \frac{(k-1)!(N-k)!(a+b-1)!(N-a-b)!}{(a-1)!(k-a-1)!(b-1)!(N-k-b)!N!}\left[\frac{1}{k-a} + \frac{1}{N-k-b+1}\right] \\
= \;& \frac{(k-1)!(N-k)!(a+b-1)!(N-a-b)!}{(a-1)!(k-a-1)!(b-1)!(N-k-b)!N!}\left[\frac{(k-a)+(N-k-b+1)}{(k-a)(N-k-b+1)}\right] \\
= \;& \frac{(k-1)!}{(a-1)!(k-a)!}\frac{(N-k)!}{(b-1)!(N-k-b+1)!}\left[\frac{N!}{(a+b-1)!(N-a-b+1)!}\right]^{-1} \\
= \;& \tbinom{k-1}{a-1}\tbinom{N-k}{b-1}\left[\tbinom{N}{a+b-1}\right]^{-1}.
\end{aligned}$$

Consideremos a aproximação de Stirling, $m! \approx e^{-m}m^m\sqrt{2\pi m}$. Tomando u inteiro fixo, segue, para m suficientemente grande, que $\binom{m}{u} \approx \dfrac{e^{-m}m^m\sqrt{2\pi m}}{u!e^{-m}e^u\frac{(m-u)^m}{(m-u)^u}\sqrt{2\pi(m-u)}} \approx \dfrac{(m-u)^u}{u!}$ Dessa

forma, a função de distribuição da frequência de bolas azuis Y_n é dada por

$$\begin{aligned}
P(Y_n \leq y) &= P(X_n \leq Ny) \\
&= \sum_{k=0}^{\lfloor Ny \rfloor} \binom{k-1}{a-1}\binom{N-k-1}{a+b-1}\left[\binom{N-1}{a+b-1}\right]^{-1} \\
&\approx \frac{(a+b-1)!}{(a-1)!(b-1)!}\sum_{k=0}^{\lfloor Ny \rfloor}\left(\frac{k-a}{n}\right)^{a-1}\left(\frac{n-k+a}{n}\right)^{b-1}\frac{1}{n} \\
&= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\sum_{k=0}^{\lfloor Ny \rfloor}\left(\frac{k}{n}-\frac{a}{n}\right)^{a-1}\left(1-\frac{k}{n}+\frac{a}{n}\right)\left(\frac{k+1}{n}-\frac{k}{n}\right) \\
&\overset{n\to 0}{\to} \frac{1}{\mathcal{B}(a,b)}\int_0^y t^{a-1}(1-t)^{b-1}dt \\
&= P(Y_\infty \leq y), \text{ para } 0 \leq y \leq 1.
\end{aligned}$$

Verifica-se, pois, a distribuição $Beta(a, b)$, como afirmado anteriormente.

Exemplo 4.4. (Martingal de Doob) Seja $(\Omega, \mathcal{F}, P)$ um espaço de probabilidade e Z uma variável aleatória integrável neste espaço. Seja $\mathcal{F} = \{\mathcal{F}_n\}_{n\geq 0}$ uma filtragem. Definindo a sequência de variáveis aleatórias $Y_n = E(Z|\mathcal{F}_n)$, segue que $\{Y_n\}_{n\geq 1}$ é um martingal com respeito a $\{\mathcal{F}_n\}_{n\geq 1}$, dito Martingal de Doob.

De fato, as variáveis Y_n são integráveis: pela Desigualdade de Jensen, $E|Y_n| = E|E(Z|\mathcal{F}_n)| \leq E(E(|Z|\,|\mathcal{F}_n)) = E|Z| < \infty$. Além disso, $E(Y_{n+1}|\mathcal{F}_n) = E(E(Z|\mathcal{F}_{n+1})|\mathcal{F}_n)) = E(Z|\mathcal{F}_n) = Y_n$.

Este martingal é uniformemente integrável. Com efeito, novamente pela Desigualdade de Jensen, temos que $|Y_n| \leq E(|Z|\,|\mathcal{F}_n)$ quase certamente, de tal forma que, definindo $X_n = E(|Z|\,|\mathcal{F}_n)$,

$$E(|Y_n|1_{\{|Y_n|\geq a\}}) \leq E(X_n 1_{\{X_n \geq a\}}).$$

Pela definição de esperança condicional,

$$E((|Z| - X_n)1_{\{X_n\geq a\}}) = 0 \Rightarrow E(|Y_n|1_{\{|Y_n|\geq a\}}) \leq E(|Z|1_{\{X_n\geq a\}}).$$

A Desigualdade de Markov afirma que $P(X_n \geq a) \leq \frac{E(X_n)}{a} = \frac{E(E(|Z|\,|\mathcal{F}_n))}{a} = \frac{E|Z|}{a}$. Desta forma, $P(X_n \geq a) \to 0$ quando $a \to \infty$ uniformemente em n. Assim, como a sequência é dominada por uma variável integrável, $E(|Z|1_{\{X_n\geq a\}}) \to 0$ quando $a \to \infty$, mas essa domina o

martingal estudado e, consequentemente, ele é uniformemente integrável.

Desta forma, pelo Teorema 4.3, existe uma variável aleatória Y_∞ em $(\Omega, \mathcal{F}, P)$ tal que $Y_n \xrightarrow{\text{q.c.}} Y_\infty$ e $E|Y_n - Y_\infty| \to 0$ quando $n \to \infty$. É possível mostrar que esta variável é dada por $Y_\infty = E(Z|\mathcal{F}_\infty)$, em que $\mathcal{F}_\infty$ é a menor σ-álgebra que contém toda a filtragem $\{\mathcal{F}_n\}$.

Lema 4.4. *Seja $\{Y_n\}$ um martingal com respeito à $\{\mathcal{F}_n\}$. Então, Y_n converge em média se, e somente se, existe uma variável aleatória Z integrável, $E(Z) < \infty$, tal que $Y_n = E(Z|\mathcal{F}_n), \forall n$.*

Se Y_n converge em média para uma variável aleatória Y_∞, então $Z = Y_\infty$ quase certamente, e $Y_n = E(Y_\infty|\mathcal{F}_n), \forall n$.

Demonstração: Existindo a variável aleatória Z que satisfaz a propriedade, então o resultado sai pelo Exemplo 4.4. Assuma, pois, que $E|Y_n - Y_\infty| \to 0$ quando $n \to \infty$. Temos, para todo m inteiro positivo e para todo evento $A \in \mathcal{F}_m$, que $E(Y_n 1_A) \to E(Y_\infty 1_A)$ quando $n \to \infty$, pois $E(|Y_n - Y_\infty|1_A) \leq E|Y_n - Y_\infty| \to 0$ à medida que $n \to \infty$. Porém, sendo um martingal, $E(Y_n 1_A) = E(Y_m 1_A)$, para todo $n \geq m$ e $A \in \mathcal{F}_m$. Segue que $E(Y_\infty 1_A) = E(Y_m 1_A)$, para todo evento $A \in \mathcal{F}_m$. Pela definição de esperança condicional (1.1), segue que $Y_m = E(Y_\infty|\mathcal{F}_m)$.

Teorema 4.5. ***(Lei Zero-Um)*** *Seja $\{X_n\}_{n\geq 0}$ sequência de variáveis aleatórias i.i.d. e $\mathcal{T}$ a σ-álgebra caudal, definida por $\mathcal{T} = \bigcap_n \mathcal{H}_n$ com $\mathcal{H}_n := \sigma(X_n, X_{n+1}, \dots)$ (ou seja, $\mathcal{H}_n$ é a menor σ-álgebra na qual tais variáveis aleatórias são mensuráveis). Então, para todo evento caudal $A \in \mathcal{T}$, ou $P(A) = 0$ ou $P(A) = 1$.*

Demonstração: Seja $A \in \mathcal{T}$ e defina o martingal $Y_n = E(1_A|\mathcal{F}_n)$ com respeito à filtragem $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$. Pela definição dada no Exemplo 4.4, segue que $A \in \mathcal{T} \subseteq \mathcal{F}_\infty = \lim_{n\to\infty} \mathcal{F}_n$. Por esse mesmo exemplo e pelo Lema 4.4, segue que $Y_n \to E(1_A|\mathcal{F}_\infty) = 1_A$ quase certamente e em média quando $n \to \infty$. Por outro lado, $Y_n = E(1_A|\mathcal{F}_n) = E(1_A) = P(A)$, pois $A \in \mathcal{T}$ é independente de todos os eventos em $\mathcal{F}_n$. Segue que $P(A) = 1_A$ quase certamente, ou seja, 1_A é quase certamente igual a uma constante. Porém, $\text{Im}(1_A) = \{0,1\}$, de tal forma que ou $P(A) = 0$ ou $P(A) = 1$.

Exemplo 4.5. Sejam $\{X_n\}_{n\geq 1}$ variáveis aleatórias independentes e defina as variáveis $Y_1 := \liminf_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i$ e $Y_2 := \limsup_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i$. Então, Y_1 e Y_2 são quase certamente constantes (mas podendo ser infinitas). Isto, pois $\{Y_1 \leq y\}, \{Y_2 \leq y\} \in \mathcal{T}$: podemos escrever $\frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n}\sum_{i=1}^{m} X_i + \frac{1}{n}\sum_{i=m+1}^{n} X_i$, em que $\frac{1}{n}\sum_{i=1}^{m} X_i \to 0$ pontualmente quando $n \to \infty$, ou seja, Y_1 e Y_2 não dependem da informação trazida pelas variáveis $X_1, \ldots, X_m$. Pelo Teorema 4.5, podemos definir $y_i := \inf\{y;\ P(Y_i \leq y) = 1\}$ de tal forma que $P(Y_i = y_i) = 1$, $i = 1{,}2$. Consequentemente, $\left\{\frac{1}{n}\sum_{i=1}^{n} X_i \text{ é convergente}\right\} = \{Y_1 = Y_2\}$ é evento caudal e, dessa forma, tem probabilidade 0 ou 1. Logo, a sequência $\left\{\frac{1}{n}\sum_{i=1}^{n} X_i\right\}_{n\geq 1}$ ou é convergente quase certamente, ou é divergente quase certamente.

Sabemos, por exemplo, pela Lei Forte dos Grandes Números de Kolmogorov que, se $\{X_n\}$ é sequência i.i.d. e $E|X_1| < \infty$, então $P\left(\left\{\frac{1}{n}\sum_{i=1}^{n} X_i \text{ é convergente}\right\}\right) = 1$, pois $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{\text{q.c.}} E(X_1)$. Note que a homogeneidade nas distribuições não foi necessária para o caso geral.

Exemplo 4.6. (Série harmônica aleatória) A sequência $a_n := \frac{(-1)^n}{n}$ é conhecida por definir uma série condicionalmente convergente: $\sum_{n=1}^{\infty} a_n = \ln(2)$, mas $\sum_{n=1}^{\infty} |a_n| = \infty$.

Considere, pois, a sequência $\{X_n\}_{n\geq 1}$ de variáveis aleatórias i.i.d. com função de probabilidade $P(X_1 = x) = \frac{1}{2}1_{\{x=-1\}} + \frac{1}{2}1_{\{x=1\}}$. Podemos definir as somas ponderadas $Y_n = \sum_{i=1}^{n} \frac{X_i}{i}$. Segue, pois, que $\{Y_n\}$ é um martingal com respeito à $\{X_n\}$. Ele é uniformemente integrável: temos que

$$E(Y_n^2) = \text{Var}\left(\sum_{i=1}^{n} \frac{X_i}{i}\right) = \sum_{i=1}^{n} \frac{1}{i^2}\text{Var}(X_i^2) \overset{i.i.d.}{=} \sum_{i=1}^{n} \frac{1}{i^2} \leq \sum_{i=1}^{\infty} \frac{1}{i^2} = \frac{\pi^2}{6},$$

e, pelo Teorema da Bola de Cristal, segue o resultado. Logo, existe uma variável aleatória Y_∞ tal que $\sum_{i=1}^{n} \frac{X_i}{i} \to Y_\infty$ quase certamente e em média quando $n \to \infty$.

De modo geral, a distribuição das variáveis i.i.d. $\{X_n\}_{n\geq 1}$ poderia ser qualquer uma que satisfaça $E(X_1) = 0$ e $\text{Var}(X_1) < \infty$.

Finaliza-se a seção com uma definição, um teorema que resume diversos resultados apresentados e outra versão alternativa do Teorema da Parada Opcional.

Definição 4.2. Seja $(\Omega, \mathcal{F}, P)$ um espaço de probabilidade e $\{Y_n\}_{n\geq 0}$ um martingal com respeito à filtragem $\{\mathcal{F}_n\}_{n\geq 0}$. Ele é dito **regular** se existe uma variável aleatória Z integrável em $(\Omega, \mathcal{F}, P)$ tal que $Y_n = E(Z|\mathcal{F}_n)$, para todo $n \geq 0$.

Teorema 4.6. *Seja $\{Y_n\}_{n\geq 0}$ um martingal com respeito à filtragem $\{\mathcal{F}_n\}_{n\geq 0}$. As seguintes afirmações são equivalentes.*

1. *$\{Y_n\}$ é uniformemente integrável;*
2. *$\{Y_n\}$ converge quase certamente e em média para uma variável Y_∞;*
3. *$\{Y_n\}$ é regular.*

Corolário 4.7. ***(Teorema da Parada Opcional)*** *Dada uma filtragem $\{\mathcal{F}_n\}$ e T um tempo de parada com respeito a ela, denotaremos $\mathcal{F}_T = \{A \in \mathcal{F};\ A \cap \{T \leq n\} \in \mathcal{F}_n \text{ para todo } n\}$.*

Seja $\{Y_n\}$ um martingal uniformemente integrável e T, S tempos de parada com respeito à filtragem $\{\mathcal{F}_n\}$. Sob a condição de que $S \leq T$, vale a propriedade martingal para os tempos de parada: $E(Y_T|\mathcal{F}_S) = Y_S$.

Demonstração: Pelo Teorema 4.6, sabemos que $\{Y_n\}$ converge quase certamente e em média para uma variável Y_∞. Primeiro, vamos provar que $E(Y_\infty|\mathcal{F}_T) = Y_T$. Podemos supor que $Y_\infty \geq 0$: como $Y_\infty = Y_\infty^+ - Y_\infty^-$, temos, pela regularidade do processo, que $E(Y_\infty|\mathcal{F}_n) = E(Y_\infty^+|\mathcal{F}_n) - E(Y_\infty^-|\mathcal{F}_n)$ para todo n.

Supondo que $Y_\infty \geq 0$, temos que $E(Y_\infty|\mathcal{F}_n) = Y_n \geq 0$ para todo n. Dado um evento $A \in \mathcal{F}_T$, temos:

$$E(Y_\infty 1_{A\cap\{T=\infty\}}) = E(Y_T 1_{A\cap\{T=\infty\}})$$

e

$$\sum_{i\geq 0} E(Y_\infty 1_{A\cap\{T=i\}}) = \sum_{i\geq 0} E(Y_T 1_{A\cap\{T=i\}}).$$

Porém, $A \cap \{T = i\}$ é um evento $\mathcal{F}_i$-mensurável. Como $Y_i = E(Y_\infty|\mathcal{F}_i)$, concluímos que

$$E(Y_T 1_{A\cap\{T=i\}}) = E(Y_i 1_{A\cap\{T=i\}}) = E(Y_\infty 1_{A\cap\{T=i\}}).$$

Segue, por definição, que $E(Y_\infty|\mathcal{F}_T) = Y_T$. Finalmente, o resultado do corolário é obtido pela Lei da Esperança Total e pela propriedade martingal,

$$E(Y_T|\mathcal{F}_S) = E(E(Y_\infty|\mathcal{F}_T)|\mathcal{F}_S) = E(Y_\infty|\mathcal{F}_S) = Y_S.$$

■

5. Desigualdades Maximais e Convergência em L^p

Uma desigualdade maximal pode ser caracterizada como uma afirmação do tipo $P(\max\{X_1, \ldots, X_n\} \geq x) \leq A_n(x)$. Pretende-se, nesta seção, estudar desigualdades maximais úteis para martingais e, a partir delas, enunciar condições necessárias e suficientes para a convergência em L^p.

Teorema 5.1. ***(Desigualdade Maximal)*** *Defina, para um processo estocástico $\{Y_n\}$, as variáveis aleatórias $Y_n^* := \max\{Y_i : 0 \leq i \leq n\}$*

1. *Se $\{Y_n\}$ é um submartingal com respeito à $\{\mathcal{F}_n\}$, então, para todo $a > 0$,*

$$P(Y_n^* \geq a) \leq \frac{E(Y_n^+)}{a}.$$

2. *Se $\{Y_n\}$ é um supermartingal com respeito à $\{\mathcal{F}_n\}$ e $E|Y_0| < \infty$, então, para todo $a > 0$,*

$$P(Y_n^* \geq a) \leq \frac{E(Y_0) + E(Y_n^-)}{a}.$$

Demonstração: Vamos considerar o tempo de parada $T := \min\{n : Y_n \geq a\}$, com $a > 0$; e $\{Y_n\}$ um submartingal com respeito a $\{\mathcal{F}_n\}$. Como a é positivo, podemos definir o submartingal não negativo $\{Y_n^+\}$, de tal forma que $T = \min\{n : Y_n^+ \geq a\}$. Assim, aplicando o Teorema da Amostragem Opcional com respeito ao tempo de parada $T \wedge n = \min\{T,n\}$, a desigualdade $T \wedge n \leq n$ implica na relação $E(Y_{T\wedge n}^+) \leq E(Y_n^+)$. Porém, pela Desigualdade de Markov, $E(Y_{T\wedge n}^+) = E(Y_T^+ 1_{\{T\leq n\}}) + E(Y_n^+ 1_{\{T>n\}}) \geq aP(T \leq n) + E(Y_n^+ 1_{\{T>n\}})$. Consequentemente, $aP(T \leq n) \leq E(Y_n^+ 1_{\{T\leq n\}}) \leq E(Y_n^+) \Rightarrow P(T \leq n) = P(Y_n^* \geq a) \leq \frac{E(Y_n^+)}{a}$.

Agora, caso $\{Y_n\}$ seja supermartingal com Y_0 integrável, o Teorema da Amostragem Opcional indica que $E(Y_0) \geq E(Y_{T\wedge n})$. Como $E(Y_{T\wedge n}) = E(Y_T 1_{\{T\leq n\}}) + E(Y_n 1_{\{T>n\}}) \geq aP(T \leq n) - E(Y_n^-) \Rightarrow P(T \leq n) \leq E(Y_0) + E(Y_n^-)$.

Observação: Pela demonstração realizada para o resultado anterior, podemos aperfeiçoar a desigualdade, de tal forma que, se $\{Y_n\}$ é submartingal, então, para $a > 0$, segue

$$P(Y_n^* \geq a) \leq \frac{E(Y_n^+ 1_{\{Y_n^*\geq a\}})}{a}.$$

Teorema 5.2. ***(Desigualdade de Doob)*** *Se $\{Y_n\}$ é submartingal uniformemente integrável e $Y_n^\# := \max\{|Y_i| : 1 \leq i \leq n\}$, então, para todo $a > 0$, segue*

$$P(Y_n^\# \geq a) \leq \frac{E(|Y_n| 1_{\{Y_n^\#\geq a\}})}{a}.$$

Demonstração: Pela Desigualdade de Jensen, se $\{Y_n\}$ é submartingal, então $\{|Y_n|\}$ também é submartingal. Podemos, pois, aplicar o Teorema 5.1 ao submartingal $\{|Y_n|\}$. A condição de ser uniformemente integrável garante que podemos aplicar o Teorema da Amostragem Opcional.

Exemplo 5.1. Seja $\{Y_n\}$ um submartingal com respeito a $\{\mathcal{F}_n\}$ com médias finitas (mas não necessariamente uniformemente integrável). Então, $\{-Y_n\}$ é supermartingal e, pelo Teorema 5.1, segue que, para $a > 0$,

$$\begin{aligned} P(\min\{Y_i : 1 \leq i \leq n\} \leq -a) \leq \ & \frac{E(Y_n^+) - E(Y_0)}{a} \Rightarrow P(Y_n^\# \geq a) \leq \frac{2E(Y_n^+) - E(Y_0)}{a} \\ \Rightarrow \ & P(Y_n^\# \geq a) \leq \frac{3}{a} \sup_{k\in\mathbb{N}} E|Y_k|. \end{aligned}$$

Logo, fazendo $n \to \infty$, segue a relação $P\left(\sup_{k\in\mathbb{N}} |Y_k| \geq a\right) \leq \frac{3}{a} \sup_{k\in\mathbb{N}} E|Y_k|$. No caso de um martingal uniformemente integrável, vimos, pela Desigualdade de Doob, que temos um limitante menor: $P\left(\sup_{k\in\mathbb{N}} |Y_k| \geq a\right) \leq \frac{1}{a} \sup_{k\in\mathbb{N}} E|Y_k|$.

Exemplo 5.2. (Desigualdade de Doob-Kolmogorov) Seja $\{Y_n\}$ um martingal com respeito a $\{\mathcal{F}_n\}$ tal que $E(Y_n^2) < \infty$, $\forall n \in \mathbb{N}$. Pela Desigualdade de Jensen, $\{Y_n^2\}$ é um submartingal com

respeito a $\{\mathcal{F}_n\}$, de tal forma que

$$P(\max\{|Y_i| : 0 \leq i \leq n\} \geq a) = P(\max\{Y_i^2 : 0 \leq i \leq n\} \geq a^2) \leq \frac{E(Y_n^2)}{a^2}.$$

quando $a > 0$.

Exemplo 5.3. (Desigualdade de Kolmogorov) Sejam $\{X_n\}$ variáveis independentes com médias e variâncias finitas. Aplique o Exemplo 5.2 ao martingal $Y_n := S_n - E(S_n)$, onde $S_n := \sum_{i=1}^{n} X_i$; de tal forma a obter a desigualdade

$$P(\max\{|S_i - E(S_i)| : 0 \leq i \leq n\} \geq a) \leq \frac{\text{Var}(S_n)}{a^2}, \text{ em que } a > 0.$$

Lema 5.3. *Sejam X e W variáveis aleatórias positivas no espaço de probabilidade $(\Omega, \mathcal{F}, P)$ satisfazendo a relação $P(X > a) \leq \frac{1}{a}E(W1_{\{X>a\}})$, para todo $a > 0$. Então, se $1 < p, q < \infty$ satisfazem $\frac{1}{p} + \frac{1}{q} = 1$, segue que $||X||_p \leq q||W||_p$, onde se definem, para $r \geq 1$, as normas $||X||_r := [E(|X|^r)]^{1/r}$.*

Demonstração: Pelo Teorema de Tonelli, $E(X^p) = \int_0^\infty P(X^p > t)dt$. Tomando a substituição $t = s^p$, segue, por hipótese, que $E(X^p) = \int_0^\infty P(X > s)d(s^p) \leq \int_0^\infty \left(\frac{1}{s}\right) E(W1_{\{X>s\}})d(s^p)$.

Aplicando novamente o Teorema de Tonelli, temos que $E(X^p) \leq \int_0^\infty \int_\Omega \left(\frac{1}{s}\right) W1_{\{X>s\}}dPd(s^p) = \int_\Omega \int_0^X \left(\frac{1}{s}\right) Wd(s^p)dP = E\left(W\int_0^X \left(\frac{1}{s}\right) d(s^p)\right) = \left(\frac{p}{p-1}\right) E(WX^{p-1}) = qE(WX^{p-1})$.

Pela Desigualdade de Hölder e como $q(p-1) = p$, temos que $E(X^p) \leqq ||W||_q||X^{p-1}||_p = q||W||_q E(X^{q(p-1)})^{1/q} = q||W||_p E(X^p)^{1/q}) = q||W||_p E(X^p)^{1-(1/p)} \Rightarrow ||X||_p \leq q||W||_p$, como queríamos.

Lema 5.4. *Seja $1 < p < \infty$ e $\{Y_n\}$ submartingal não negativo com respeito à $\{\mathcal{F}_n\}$ tal que $E(Y_n^p) < \infty$, para todo $n \in \mathbb{N}$. Então, vale a desigualdade*

$$E(Y_n^p) \leq E((Y_n^*)^p) \leq \left(\frac{p}{p-1}\right)^p E(Y_n^p).$$

Demonstração: A primeira desigualdade é trivial, já que $Y_n^* \geq Y_n$ por definição. Para a segunda desigualdade, aplique o Lema 5.3 com as variáveis $X = Y_n^*$ e $W = Y_n$. Podemos

aplicá-lo, pois a condição é satisfeita devido ao Teorema 5.1.

Para definirmos o Teorema da convergência de um martingal em L^p, devemos considerar o conjunto de variáveis aleatórias integráveis em L^p, ou seja,

$$L^p(\Omega, \mathcal{F}, P) = \{X : \Omega \to \mathbb{R} \text{ tal que } X \text{ é } \mathcal{F}\text{-mensurável e } E(|X|^p) < \infty\}.$$

Teorema 5.5. ***(Convergência em L^p)*** *Seja $\{Y_n\}$ um submartingal com respeito à filtragem $\{\mathcal{F}_n\}$. São equivalentes as seguintes afirmações.*

1. *$\{Y_n\}$ é limitado em L^p, ou seja, $E(|Y_n|^p) < \infty$, para todo $n \in \mathbb{N}$;*

2. *$\{Y_n\}$ é uniformemente integrável e $Y_\infty \in L^p(\Omega, \mathcal{F}, P)$, onde Y_∞ é a variável para a qual as Y_n convergem quase certamente e em média;*

3. *$Y^* := \sup_{n\in\mathbb{N}} |Y_n| \in L^p(\Omega, \mathcal{F}, P)$.*

Se cada uma das afirmações é válida, então $Y_n \to Y_\infty$ quase certamente e em L^p.

Demonstração:

- $(1 \Rightarrow 2)$ Como $Y_n \in L^p(\Omega, \mathcal{F}, P)$ para todo índice n, a sequência $\{Y_n\}$ é uniformemente integrável pelo Teorema da Bola de Cristal. Pelo Teorema da Convergência de Martingais, existe uma variável $Y_\infty : \Omega \to \mathbb{R}$ satisfazendo $Y_n \xrightarrow{\text{q.c.}} Y_\infty$ e $Y_n \xrightarrow{L^1} Y_\infty$. Pelo Lema de Fatou, temos que

$$E(|Y_\infty|^p) = E(\liminf_{n\to\infty} |Y_n|^p) \leq \liminf_{n\to\infty} E(|Y_n|^p) \leq \sup_n E(|Y_n|^p) < \infty,$$

  pela hipótese do processo ser limitado em L^p. Assim, $Y_\infty \in L^p(\Omega, \mathcal{F}, P)$.

- $(2 \Rightarrow 3)$ Como o processo $\{Y_n^*\}$, com $Y_n^* = \max\{Y_i : 0 \leq i \leq n\}$, é não decrescente e converge para a variável Y^*, temos, pelo Teorema da Convergência Monótona e pelo Lema 5.4, que

$$||Y^*||_p = \lim_{n\to\infty} ||Y_n^*||_p \leq q \lim_{n\to\infty} ||Y_n||_p \leq q \sup_{n\geq 1} ||Y_n||_p < \infty.$$

- (3 $\Rightarrow$ 1) Temos:

$$\sup_n E(|Y_n|^p) \leq E\left(\left[\sup_n |X_n|\right]^p\right) < \infty,$$

devido à integrabilidade em L^p da variável Y^*.

Finalmente, para provarmos a convergência em L^p, consideramos a variável aleatória Y_∞ satisfazendo $Y_n \xrightarrow{\text{q.c.}} Y_\infty \in L^p(\Omega, \mathcal{F}, P)$. Sabemos, também, que $Y^* = \sup_n |Y_n| \in L^p(\Omega, \mathcal{F}, P)$. Em particular, temos que $|Y_n - Y_\infty|^p \xrightarrow{\text{q.c.}} 0$ e

$$|Y_n - Y_\infty|^p \leq (2|Y^*|)^p \in L^1(\Omega, \mathcal{F}, P).$$

Dessa forma, pelo Teorema da Convergência Dominada, segue que $E(|Y_n - Y_\infty|^p) \to 0$ quando $n \to \infty$.

Corolário 5.6. *Sejam* $1 < p < \infty$ *e* $\{Y_n\}$ *um martingal com respeito à* $\{\mathcal{F}_n\}$ *tal que* $\sup_n E(|Y_n|^p) < \infty$. *Então,* $E(|Y_n - Y_\infty|^p) \to 0$ *quando* $n \to \infty$, *em que* Y_∞ *é o limite quase certo das* Y_n.

6. Martingais Reversos e Inferência Bayesiana

As principais referências usadas nesta parte foram Schervish (1995, Capítulo 1) e Durrett (2019, Seção 4.7).

Definição 6.1. Suponha $\{\mathcal{G}_n\}_{n\geq 0}$ uma sequência decrescente de informações, ou seja, σ-álgebras: $\mathcal{G}_n \supseteq \mathcal{G}_{n+1}$, para todo n. A sequência $\{(Y_n, \mathcal{G}_n\}_{n\geq 0}$ é dita um **martingal reverso** se:

1. Y_n é uma variável $\mathcal{G}_n$ mensurável;
2. $E|Y_n| < \infty$, para todo $n \geq 0$;
3. $E(Y_n|\mathcal{G}_{n+1}) = Y_{n+1}$, para todo $n \geq 0$ (ou, equivalentemente, se $m < n$, então $E(Y_m|\mathcal{G}_n) = Y_n$).

Observamos que $\{(Y_n, \mathcal{G}_n)\ n \geq 0\}$ é um martingal reverso se, e somente se, $(\{Y_n, \mathcal{G}_n\}; n = \ldots, 2, 1, 0)$ define um martingal. De modo análogo à definição feita para martingais, podemos interpretar que estamos perdendo informação ao longo do processo.

Teorema 6.1. ***(Convergência de Martingais Reversos)*** *Seja* $\{(Y_n, \mathcal{G}_n)\}_{n\geq 0}$ *um martingal reverso. Então,* Y_n *converge quase certamente e em média para uma variável aleatória* Y_∞.

Demonstração: Sabemos que $\{Y_n; n = \ldots,2,1,0\}$ é martingal com respeito à filtragem $\{\mathcal{G}_n, n = \ldots,2,1,0\}$. Usando a Desigualdade de Snell para subidas e um raciocínio como aquele usado na prova do Teorema da Convergência de Martingais, podemos concluir que o limite Y_∞ existe quase certamente. Pela definição de martingal reverso, temos que $E(Y_0|\mathcal{G}_n) = Y_n$. Além disso, $E|Y_0| < \infty$. Desta forma, pelo Teorema 4.6, o martingal é uniformemente integrável, e vale a convergência em média. ■

Corolário 6.2. *Se* $Y_n \xrightarrow{q.c.} Y_\infty$ *e* $\mathcal{G}_\infty := \bigcap_n \mathcal{G}_n$, *então* $Y_\infty = E(Y_0|\mathcal{G}_\infty)$ *quase certamente.*

Corolário 6.3. ***(Teorema de Lévy)*** *Seja* $\{\mathcal{G}_n\}$ *sequência decrescente de informações e defina* $\mathcal{G}_\infty := \bigcap_n \mathcal{G}_n$. *Dada uma variável aleatória* Z, *então* $E(Z|\mathcal{G}_n) \to E(Z|\mathcal{G}_\infty)$ *quase certamente e em média.*

Demonstração: Defina $Y_n := E(Z|\mathcal{G}_n)$, de tal forma que este é martingal reverso. Pelo Corolário 6.2, $Y_n \to Y_\infty$ quase certamente e em média quando $n \to \infty$, com

$$Y_\infty = E(Y_0|\mathcal{G}_\infty) = E(E(Z|\mathcal{G}_0)|\mathcal{G}_\infty) = E(Z|\mathcal{G}_\infty).$$ ■

Definição 6.2. Dado um espaço de probabilidade $(\Omega, \mathcal{F}, P)$, uma sequência de variáveis aleatórias $\{X_n\}_{n\geq 1}$ é dita **permutável** se, para todo $n \in \mathbb{N}$ e para toda permutação $\pi \in S_n$ – sendo S_n o grupo permutacional do conjunto $\{1, \ldots,n\}$ – os vetores $(X_1, \ldots,X_n)$ e $(X_{\pi(1)}, \ldots,X_{\pi(n)})$ têm a mesma distribuição.

Teorema 6.4. ***(Lei Forte dos Grandes Números para variáveis permutáveis).*** *Sejam* $(\Omega,\mathcal{F},P)$ *um espaço de probabilidade e* $\{X_n\}_{n\geq 1}$ *uma sequência de variáveis* $\mathcal{F}$*-mensuráveis, com* $E|X_n| < \infty$ *para todo* n. *Se as variáveis são permutáveis, então existe uma* σ*-álgebra* $\mathcal{G}_\infty$ *tal que*

$$\bar{X}_n := \sum_{i=1}^{n} \frac{X_i}{n} \xrightarrow{q.c.} E(X_1|\mathcal{G}_\infty).$$

Demonstração: Consideremos a notação $x = (x_1, x_2, \dots)$ para representar uma sequência de valores. Para $n \in \mathbb{N}$, defina $\mathcal{C}_n$ como a coleção de todos os borelianos $A \subseteq \mathbb{R}^\infty$ satisfazendo que $x \in A$ se, e somente se, $y \in A$, para todo $y \in \mathbb{R}^\infty$ que concorda com x após a n-coordenada e as primeiras n coordenadas são uma permutação das n primeiras coordenadas de x.

Segue, pois, que $\mathcal{C}_n$ é σ-álgebra. Temos que $\emptyset, \mathbb{R}^\infty \in \mathcal{C}_n$. Para as demais propriedades, defina a relação de equivalência $x \sim y$ quando y concorda com x após a n-ésima coordenada e as n primeiras diferem apenas por permutação. Então, um conjunto $A \in \mathcal{C}_n$ é formado por união das classes de equivalência em $\mathbb{R}^\infty/\sim$. Temos, portanto, que A^c é formado por todas as demais classes de equivalência não presentes em A, ou seja, $A^c \in \mathcal{C}_n$. Por fim, supondo $\{A_n\}$ conjuntos em $\mathcal{C}_n$, se $x \in \bigcup_n A_n$, então existe i tal que $x \in A_i$. Segue que $[x]_\sim \subseteq A_i$, ou seja, $[x]_\sim \subseteq \bigcup_n A_n$. Portanto, $\bigcup_n A_n$ é $\mathcal{C}_n$-mensurável, e $\mathcal{C}_n$ define σ-álgebra. Uma função contínua $\mathcal{C}_n$-mensurável é $f : \mathbb{R}^\infty \to \mathbb{R}$ dada por $x \mapsto \sum_{i=1}^{n} x_i$.

Dada uma sequência permutável de variáveis aleatórias $X = (X_1, X_2, \dots)$, defina as σ-álgebras $\mathcal{G}_n = X^{-1}(\mathcal{C}_n)$ e as variáveis aleatórias $Y_n := E(X_1|\mathcal{G}_n)$. Por hipótese, $E|X_1| < \infty$. Além disso, $\{\mathcal{G}_n\}_{n\geq 1}$ é uma sequência decrescente de σ-álgebras (pois $\{\mathcal{C}_n\}$ é uma sequência decrescente). Pelo Teorema de Lévy, $\lim_{n\to\infty} Y_n = E(X_1|\mathcal{G}_\infty)$ existe e é finita quase certamente. Resta-nos mostrar que

$$Y_n = \bar{X}_n = \sum_{i=1}^{n} \frac{X_i}{n}.$$

Como f é uma função $\mathcal{C}_n$-mensurável, mostraremos que, se $A \in \mathcal{G}_n$, $E(\bar{X}_n 1_A) = E(X_1 1_A)$. Mas $X_i 1_A$ tem mesma distribuição de $X_j 1_A$, para quaisquer $i, j = 1, \dots, n$, devido à permutabilidade e simetria do conjunto A. Segue, pois, que $E(1_A X_1) = \sum_{i=1}^{n} \frac{E(1_A X_i)}{n}$, $\forall A \in \mathcal{G}_n$. O resultado sai pela definição da esperança condicional. ∎

Corolário 6.5. ***(Lei Forte de Kolmogorov).*** *Se $\{X_n\}_{n\geq 1}$ é uma sequência de variáveis independentes e identicamente distribuídas com $E|X_1| < \infty$ e $\mu := E(X_1)$, então*

$$\bar{X}_n := \sum_{i=1}^{n} \frac{X_i}{n} \xrightarrow{q.c.} \mu.$$

Demonstração: O resultado sai combinando o que foi visto no Teorema 6.4 com a

conclusão apresentada no Exemplo 4.5. ■

Lema 6.6. *Seja $\{X_n\}_{n\geq 1}$ uma sequência de variáveis aleatórias i.i.d. no espaço de probabilidade $(\Omega, \mathcal{F}, P)$ e $\phi : \mathbb{R}^k \to \mathbb{R}$ uma função mensurável. Defina a função*

$$A_n(\phi) = \frac{1}{n(n-1)\dots(n-k+1)} \sum_{1\leq i_1,\dots,i_k\leq n} \phi(X_{i_1}, \dots, X_{i_k}),$$

onde $i_j \neq i_\ell$ para todos $j, \ell = 1, \dots, k$. Se ϕ é limitada, então

$$A_n(\phi) \xrightarrow{q.c.} E(\phi(X_1, \dots, X_k)).$$

Demonstração: Retomando as σ-álgebras C_n definidas anteriormente, segue que $A_n(\phi)$ é C_n-mensurável, de tal forma que

$$A_n(\phi) = E(A_n(\phi)|C_n) = \frac{1}{n(n-1)\dots(n-k+1)} \sum_{i_1,\dots,i_k} E(\phi(X_{i_1}, \dots, X_{i_k})|C_n).$$

Pela propriedade permutacional de C_n, segue $A_n(\phi) = E(\phi(X_1, \dots, X_k)|C_n)$. Pelo Teorema de Lévy, $E(\phi(X_1, \dots, X_k)|C_n) \to E(\phi(X_1, \dots, X_k)|C_\infty)$ quase certamente e em média.

Vamos mostrar que $W_1 := \liminf A_n(\phi)$ e $W_2 := \limsup A_n(\phi)$ são quase certamente constantes. Os eventos $\{W_1 \leq w\}$ e $\{W_2 \leq w\}$ pertence à σ-álgebra caudal $\mathcal{T}$, pois

$$A_n(\phi) = \frac{1}{n(n-1)\dots(n-k+1)} \sum_{1\leq i_1,\dots,i_k\leq m} \phi(X_{i_1}, \dots, X_{i_k}) + \sum_{\exists i_\ell \notin \{1,\dots,m\}} .$$

Quando $n \to \infty$, o primeiro termo tende a zero pontualmente devido à hipótese de ϕ limitada:

$$|A_n(\phi)| \leq \frac{m(m-1)\dots(m-k+1)}{n(n-1)\dots(n-k+1)} \sup|\phi|.$$

Então, W_1 e W_2 não dependem das variáveis $X_1, \dots, X_m$, para todo $m \in \mathbb{N}$. Logo, pela Lei 0–1, existem $w_i := \inf\{w; P(W_i \leq w) = 1\}$ tais que $P(W_i = w_i) = 1$, com $i = 1{,}2$. Também pela Lei 0–1, o evento $\{W_1 = W_2\} = \{A_n(\phi)$ é convergente$\}$ tem probabilidade 0 ou 1. Porém, foi mostrada a convergência, de tal forma que $P(W_1 = W_2) = 1$. Como o limite $E(\phi(X_1, \dots, X_k|C_\infty)$ deve ser constante, segue que $A_n(\phi) \xrightarrow{\text{q.c.}} E(\phi(X_1, \dots, X_k)$.

Teorema 6.7. ***(Representação de De Finetti)*** *Dada uma sequência de variáveis aleatórias*

$\{X_n\}_{n\geq 1}$ em um espaço de probabilidade $(\Omega,\mathcal{F},P)$, elas são permutáveis se, e somente se, condicionadas em $C_\infty := \bigcap_n C_n$, $\{X_n\}$ são independentes e identicamente distribuídas.

Demonstração: Suponha que a sequência é permutável. Pelo Lema anterior, $A_n(\phi) = E(\phi(X_1,\ldots,X_k)|C_n)$, de tal forma que $A_n(\phi) \xrightarrow{\text{q.c.}} E(\phi(X_1,\ldots,X_k)|C_\infty)$. Sejam $f:\mathbb{R}^{k-1}\to\mathbb{R}$ e $g:\mathbb{R}\to\mathbb{R}$ limitadas e $I_{n,k} := \{(i_1,\ldots,i_k); 1\leq i_1,\ldots,i_k\leq n, i_\ell \text{ distintos}\}$. Desta forma, segue que

$$\begin{aligned} n\ldots(n-k+2)A_n(f)nA_n(g) &= \sum_{i\in I_{n,k-1}} f(X_{i1},\ldots,X_{i_{k-1}})\sum_m g(X_m) \\ &= \sum_{i\in I_{n,k}} f(X_{i_1},\ldots,X_{i_{k-1}})g(X_{i_k}) + \sum_{i\in I_{n,k-1}}\sum_{j=1}^{k-1} f(X_{i_1},\ldots,X_{i_{k-1}})g(X_{i_j}). \end{aligned}$$

Definindo $\phi(x_1,\ldots,x_k) = f(x_1,\ldots,x_{k-1})g(x_{ik})$ e, como $\dfrac{n(n-1)\ldots(n-k+2)n}{n(n-1)\ldots(n-k+1)} = \dfrac{n}{n-k+1}$ e $\dfrac{n(n-1)\ldots(n-k+2)}{n(n-1)\ldots(n-k+1)} = \dfrac{1}{n-k+1}$, segue

$$A_n(\phi) = \frac{n}{n-k+1}A_n(f)A_n(g) - \frac{1}{n-k+1}\sum_{j=1}^{k-1}A_n(\phi_j),$$

onde $\phi_j(x_1,\ldots,x_{k-1}) := f(x_1,\ldots,x_{k-1})g(x_j)$. Como $\lim \dfrac{n}{n-k+1} = 1$ e $\lim \dfrac{1}{n-k+1} = 0$, segue, pela convergência quase certa da aplicação A_n e pelo Teorema 3.2, que, quando $n\to\infty$,

$$E(\phi(X_1,\ldots,X_k)|C_\infty) = E(f(X_1,\ldots,X_{k-1})g(X_k)|C_\infty) = E(f(X_1,\ldots,X_{k-1})|C_\infty)E(g(X_k)|C_\infty).$$

Aplicando o raciocínio indutivamente, concluímos que

$$E\left(\prod_{i=1}^k f_i(X_i)\middle|C_\infty\right) = \prod_{i=1}^k E(f_i(X_i)|C_\infty),$$

como queríamos.

Agora, suponha que, condicionadas em C_∞, as variáveis aleatórias $\{X_n\}_{n\geq 1}$ sejam i.i.d. Considerando a definição da esperança condicional, temos, para todo boreliano $B\in\mathcal{B}(\mathbb{R})$ e para $\alpha\in S_n$,

$$\begin{aligned} E(f(X_1,\ldots,X_n)1_B) &= E(E(f(X_1,\ldots,X_n)|C_\infty)1_B) \\ &\overset{\text{i.i.d.}}{=} E(E(f_1(X_1)\ldots f_n(X_n)|C_\infty)1_B) \end{aligned}$$

$$= E(E(f_{\alpha(1)}(X_{\alpha(1)}) \dots f_{\alpha(n)}(X_{\alpha(n)})|\mathcal{C}_\infty)1_B)$$
$$= E(E(f(X_{\alpha(1)}, \dots, X_{\alpha(n)})|\mathcal{C}_\infty)1_B)$$
$$= E(f(X_{\alpha(1)}, \dots, X_{\alpha(n)})1_B)$$

Como a propriedade vale para qualquer boreliano, segue que $f(X_1, \dots, X_n) = f(X_{\alpha(1)}, \dots, X_{\alpha(n)})$, isto é, a permutabilidade das variáveis. ■

Corolário 6.8. *Uma sequência $\{X_n\}_{n\geq 1}$ de variáveis aleatórias seguindo distribuição de Bernoulli,*

$$P(X_j = i) = p_j 1_{\{i=1\}} + (1 - p_j)1_{\{i=0\}}, \text{ em que } j \geq 1 \text{ e } p_j \in [0,1],$$

é permutável se, e somente se, existe uma variável aleatória Θ tal que $supp(\Theta) \subset [0,1]$ e, condicionando em $\Theta = \theta$, as variáveis $\{X_n\}_{n\geq 1}$ têm distribuição $Bern(\theta)$ (ou seja, $p_j = \theta$, para todo j) e são independentes e identicamente distribuídas. Além disso, sob hipótese de permutabilidade da sequência, a distribuição de Θ é única e $\sum_{i=1}^{n} \frac{X_i}{n} \xrightarrow{q.c.} \Theta$. Segue, pois, que

$$P(X_1 = 1, \dots, X_k = 1, X_{k+1} = 0, \dots, X_n = 0) = \int_0^1 \theta^k (1 - \theta)^{n-k} dF(\theta).$$

Exemplo 6.1. O seguinte exemplo é atribuído a Thomas Bayes em 1764. Sejam $\{X_n\}_{n\geq 1}$ variáveis aleatórias com distribuição de Bernoulli, permutáveis, tais que

$$P(k \text{ sucessos em } n \text{ tentativas}) = \frac{1}{n+1},$$

com $k = 0, \dots, n$ e $n \geq 1$. Tal definição é um conjunto consistente de probabilidades, pois, para $x_1, \dots, x_n \in \{0,1\}$ tais que $\sum x_i = k$,

$$\begin{aligned} P(X_1 = x_1, \dots, X_n = x_n) &= \frac{1}{(n+1)\binom{n}{k}} \\ &= \frac{1}{(n+2)\binom{n+1}{k}} + \frac{1}{(n+2)\binom{n+1}{k+1}} \\ &= P(X_1 = x_1, \dots, X_n = x_n, X_{n+1} = 0) + P(X_1 = x_1, \dots, X_n = x_n, X_{n+1} = 1). \end{aligned}$$

Como $\bar{X}_n = \sum_{i=1}^{n} \frac{X_i}{n} \xrightarrow{\text{q.c.}} \Theta$ pelo corolário anterior, $\bar{X}_n \xrightarrow{\mathcal{D}} \Theta$. Assim, se $F_n(x) = P(\bar{X}_n \leq x)$ é a distribuição acumulada da variável $\bar{X}_n$, então para qualquer $0 \leq x \leq 1$,

$$F_n(x) = P(\text{máximo de } \lfloor nx \rfloor \text{ sucessos em } n \text{ tentativas}) = \frac{\lfloor nx \rfloor + 1}{n+1} \to F(x) = x$$

quando $n \to \infty$. Então, dado que $\Theta = \theta$, as variáveis $\{X_n\}$ são i.i.d., em que $\Theta \sim \text{Uniforme}(0,1)$.

Exemplo 6.2. Aproveitando o comentário relacionado à Inferência Estatística, serão recuperados outros exemplos úteis para a área. O primeiro relaciona-se com a função de verossimilhança de uma amostra.

Seja $\{X_n\}_{n\geq 1}$ sequência de variáveis aleatórias i.i.d. com mesma função densidade de probabilidade f. Suponha que f possa ser tanto f_0 quanto f_1, sendo estas funções densidades de probabilidade distintas. A **razão de verossimilhanças** para este problema é definida pela variável aleatória

$$\Lambda_n := \frac{f_1(X_1) \ldots f_1(X_n)}{f_0(X_1) \ldots f_0(X_n)}.$$

Para o teste de hipóteses Υ dado por "decida $f \equiv f_1$ se $\Lambda_n \geq a$, decida $f \equiv f_0$ se $\Lambda_n < a$", para constante fixada $a > 0$, considere a filtragem $\{\mathcal{F}_n\}$ com $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$.

Assuma $f_0 > 0$. Condicionando em $f \equiv f_0$, então

$$E(\Lambda_{n+1}|\mathcal{F}_n) = \Lambda_n E\left(\frac{f_1(X_{n+1})}{f_0(X_{n+1})}\right) = \Lambda_n \int \frac{f_1(x)}{f_0(x)} f_0(x)dx = \Lambda_n \int f_1(x)dx = \Lambda_n.$$

Além disso, as variáveis Λ_n são integráveis:

$$E|\Lambda_n| = \int \frac{f_1(x_1) \ldots f_1(x_n)}{f_0(x_1) \ldots f_0(x_n)} f_0(x_1) \ldots f_0(x_n)dx_1 \ldots dx_n = \int f_1(x_1) \ldots f_1(x_n)dx_1 \ldots dx_n = 1.$$

Segue que $\{\Lambda_n\}$ é martingal com respeito a $\{\mathcal{F}_n\}$, sob a suposição de que $f \equiv f_0$. Pelo Teorema da Convergência de Martingais, segue que existe uma variável aleatória Λ_∞ tal que $\Lambda_n \xrightarrow{\text{q.c.}} \Lambda_\infty$.

Pela Desigualdade de Jensen, notamos que

$$E(\log \Lambda_n) = \sum_{i=1}^{n} E\left(\log\left(\frac{f_1(X_i)}{f_0(X_i)}\right)\right) < \sum_{i=1}^{n} \log\left(E\left(\frac{f_1(X_i)}{f_0(X_i)}\right)\right) = 0.$$

Pela Lei Forte dos Grandes Números, segue que $\dfrac{\log \Lambda_n}{n}$ converge quase certamente para algum valor no intervalo $[-\infty,0)$, de tal forma que $\Lambda_n \xrightarrow{\text{q.c.}} 0$. Em particular, $E(\Lambda_\infty) = 1 \neq 0$, ou seja, o martingal não converge em média (equivalentemente, a sequência não é uniformemente integrável).

Como $\Lambda_\infty = 0$, independentemente do valor a escolhido para o teste Υ, tomando um n suficientemente grande leva à conclusão correta de que $f \equiv f_0$. Pela desigualdade maximal, temos que

$$P(\Lambda_n^* \geq a) \leq \frac{1}{a}, \text{ em que } \Lambda_n^* = \max\{\Lambda_i; 0 \leq i \leq n\}.$$

Exemplo 6.3. Suponha, no exemplo anterior, uma amostra aleatória da distribuição $f = f_0$ dada por $\mathcal{N}(0,\sigma^2)$. Considere a distribuição f_1 dada por $\mathcal{N}(\mu, \sigma^2)$. Neste caso,

$$\frac{f_1(x)}{f_0(x)} = \exp\left\{\frac{2\mu x - \mu^2}{2\sigma^2}\right\} \text{ e, desta forma, } \Lambda_n = \exp\left\{\frac{\mu}{\sigma^2}(X_1 + \cdots + X_n) - \frac{n\mu^2}{2\sigma^2}\right\}.$$

Pelo exemplo anterior, $\{\Lambda_n\}$ é martingal com respeito à filtragem $\{\mathcal{F}_n\}$, $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$. Porém, já sabíamos disto pelo Martingal de Wald, estudado no Corolário 2.7.

7. Martingais a Tempo Contínuo

A definição de martingais considerando uma sequência crescente de σ-álgebras permite estender a família de variáveis aleatórias do processo. De modo geral, podemos indexar em outros conjuntos, além dos naturais ou dos inteiros.

Definição 7.1. Sejam $T \subseteq \mathbb{R}$ um conjunto de índices Borel-mensurável e $\{Y(t);\ t \in T\}$ um processo estocástico sobre um espaço de probabilidade $(\Omega,\mathcal{F},P)$.

Uma **filtragem** é uma família de σ-álgebras $\{\mathcal{F}_t\}_{t\in T}$, todas contidas em $\mathcal{F}$, tais que, para todos $t,s \in T$ tais que $s \leq t$, $\mathcal{F}_s \subseteq \mathcal{F}_t$.

Diz-se que o processo $\{Y(t)\}_{t\in T}$ é **submartingal** com respeito à filtragem $\{\mathcal{F}_t\}$ se:

1. $Y(t)$ é $\mathcal{F}_t$ mensurável;

2. $E(Y(t)^+) < \infty$;

3. Para todo $u > 0$ tal que $t + u \in T$, $E(Y(t+u)|\mathcal{F}_t) \geq Y(t)$.

O processo é dito **supermartingal** se $\{-Y(t)\}$ é submartingal. O processo é dito **martingal** se for simultaneamente submartingal e supermartingal. Como a propriedade submartingal depende da filtragem, usa-se a notação $(Y,\mathcal{F})$.

Definição 7.2. Considere uma filtragem $\{\mathcal{F}_t;\ t \geq 0\}$ no espaço de probabilidade $\{\Omega,\mathcal{F},P\}$. Uma variável aleatória T é dita **tempo de parada** com respeito à $\{F_t\}$ se supp $T \subseteq [0,+\infty]$ e o evento $\{T \leq t\}$ é $\mathcal{F}_t$-mensurável, para todo $t \geq 0$.

Proposição 7.1.

1. *T é tempo de parada se, e somente se, $\{T > t\} \in \mathcal{F}_t$, para todo $t \geq 0$;*

2. *Se S e T são tempos de parada, então $S + T$, $S \wedge T := \min\{S,T\}$ e $S \vee T := \max\{S,T\}$ também são com respeito à mesma filtragem.*

Exemplo 7.1. Em um próximo momento, será de interesse a construção de processos $\{X(t)\}$ de tal forma que as realizações $X(t)$ sejam funções contínuas em t, em particular, o movimento browniano. Considerando tal processo, definimos a filtragem $\mathcal{F}_t := \sigma(\{X(s); 0 \leq s \leq t\})$. De fato, $X(s)$ é $\mathcal{F}_s$-mensurável e, para todo $s < t$, $\mathcal{F}_s \subseteq \mathcal{F}_t$.

Vamos considerar a variável aleatória $T := \inf\{t \geq 0; X(t) \geq a\}$ para $a \in \mathbb{R}$ fixado. Pela continuidade de $X(t)$,

$$\{T > t\} = \bigcup_{k=1}^{\infty} \bigcap_{r\in\mathbb{Q};0\leq r\leq t} \{(a - X(r)) \geq 1/k\}.$$

Como $\{(a - X(r)) \geq 1/k\}$ é evento $\mathcal{F}_t$-mensurável já que $r \leq t$, sendo $\mathcal{F}_t$ σ-álgebra o evento $\{T > t\}$, sendo combinação enumerável de operações sobre conjuntos, é $\mathcal{F}_t$-mensurável.

De modo geral, para qualquer conjunto fechado $C \subseteq \mathbb{R}$, a variável aleatória

$$T_C := \inf\{t \geq 0 : X(t) \in C\}$$

é um tempo de parada com respeito à filtragem indicada acima, quando a realização de $X(t)$ é quase certamente contínua.

Exemplo 7.2. Outra família de processos estocásticos de interesse refere-se aos processos de contagem, nos quais estão incluídos o processo de Poisson e o processo de nascimento puro. Neste caso, um processo $\{N(t)\}_{t\geq 0}$ desse tipo realiza $N(t)$ continuamente à direita, $N(t) = \lim_{s\downarrow t} N(s)$, e possui limite à esquerda, ou seja, sempre existe $\lim_{s\uparrow t} N(s)$ para todo $t > 0$.

Considere a filtragem do exemplo anterior, $\mathcal{F}_t = \sigma(\{N(s); 0 \leq s \leq t\})$ e defina uma nova filtragem $\{\mathcal{G}_t\}_{t\geq 0}$ dada por $\mathcal{G}_t := \bigcap_{\epsilon>0} \mathcal{F}_{t+\epsilon}$. Por fim, considere $\mathcal{H}_t := \sigma(\mathcal{G}_t, \{A \in \mathcal{F}; P(A) = 0\})$, uma versão de $\mathcal{G}_n$ levando em conta a "completude".

Neste caso, para qualquer boreliano $B \subseteq \mathbb{R}$, a variável aleatória

$$T_B := \begin{cases} \inf\{t \geq 0 : N(t) \in B\} & \text{se } X(t) \in B \text{ para algum } t \geq 0; \\ +\infty & \text{se } X(t) \notin B \text{ para todo } t \geq 0 \end{cases}$$

é um tempo de parada com respeito à filtragem $\{\mathcal{H}_t\}$.

Os teoremas principais vistos anteriormente se estendem ao caso contínuo. As demonstrações serão omitidas, mas assumem a hipótese de continuidade à direita de uma realização do processo e aproximações discretas do conjunto de índices: note que, se $(Y,\mathcal{F})$ é martingal a tempo contínuo, então $\{(Y(n2^{-m}), \mathcal{F}_{n2^{-m}})\}_{n=0,1,\ldots}$ é martingal a tempo discreto para todo m inteiro positivo. Os detalhes podem ser vistos no livro de Mörters e Peres (2010, Seção 2.4).

Teorema 7.2. ***(da Parada Opcional)*** *Sejam $\{Y(t)\}_{t\geq 0}$ um submartingal e T um tempo de parada com respeito à $\{\mathcal{F}_t\}_{t\geq 0}$. Se $P(T < +\infty) = 1$ e a família de variáveis aleatórias $\{Y(t \wedge T)^+\}_{t\geq 0}$ é uniformemente integrável, então*

$$E(Y(0)) \leq E(Y(T)).$$

Corolário 7.3. *Sejam $\{Y(t)\}_{t\geq 0}$ um martingal com respeito à $\{\mathcal{F}_t\}_{t\geq 0}$ e T um tempo de parada. Se $P(T < +\infty) = 1$ e $E\left(\sup_{t\geq 0} |Y(t)|\right) < \infty$ então*

$$E(Y(0)) = E(Y(T)).$$

Corolário 7.4. *Sejam $\{Y(t)\}_{t\geq 0}$ um martingal uniformemente integrável e T, S tempos de parada*

com respeito à $\{\mathcal{F}_t\}_{t\geq 0}$ *tais que* $S \leq T$. *Se as trajetórias de* $Y(t)$ *são contínuas à direita, então*

$$E(Y(T)|\mathcal{F}_S) = Y(S).$$

Teorema 7.5. ***(Convergência de Martingais)*** *Seja* $\{Y(t)\}_{t\geq 0}$ *um martingal com respeito à* $\{\mathcal{F}_t\}_{t\geq 0}$ *tal que as trajetórias de* $Y(t)$ *são contínuas à direita. Se existe* $M > 0$ *tal que* $\sup_{t\geq 0} E\,|Y(t)| \leq M$, *então existe* $Y(\infty) = \lim_{t\to\infty} Y(t)$ *quase certamente. Além disso, se o martingal é uniformemente integrável, então* $Y(t) \xrightarrow{L^1} Y(\infty)$.

Teorema 7.6. ***(Condição de regularidade e convergência em*** L^p***)*** *Para* $\{Y(t)\}_{t\geq 0}$ *um martingal com respeito à filtragem* $\{\mathcal{F}_t\}_{t\geq 0}$, *são equivalentes:*

1. *Existe variável aleatória* Y_∞ *tal que* $Y(t) \xrightarrow{L^1} Y(\infty)$;
2. *O processo é regular, ou seja, existe variável aleatória* $Y(\infty) \in L^1(\Omega,\mathcal{F},P)$ *tal que* $Y(t) = E(Y(\infty)|\mathcal{F}_t)$, *para todo* $t \geq 0$;
3. $\{Y(t)\}$ *é uniformemente integrável.*

Ocorrendo umas das condições, segue $Y(t) \xrightarrow{q.c.} Y(\infty)$. *Além disso, se existe* $p > 1$ *tal que* $\sup_{t\geq 0} E\,(|X(t)|^p) < \infty$, *as condições são satisfeitas e* $Y(t) \xrightarrow{L^p} Y(\infty)$.

Como pretendemos trabalhar com processos contínuos na próxima seção, serão registrados exemplos relacionados a processos de contagem, ou seja, processos formados por variáveis com suporte discreto.

Exemplo 7.3. Seja $\{X(t)\}_{t\geq 0}$ um processo de Poisson homogêneo de parâmetro λ. Desta forma, os seguintes processos são martingais a tempo contínuo com respeito à filtragem $\mathcal{F}_t := \sigma(\{X(s); 0 \leq s \leq t\})$:

- $Y(t) := X(t) - \lambda t$;
- $U(t) := Y^2(t) - \lambda t$;
- Para $\theta \in \mathbb{R}$ fixado, $V_\theta(t) := \exp\{-\theta X(t) + \lambda t(1 - e^{-\theta})\}$.

Considere o tempo de parada $T = \inf\{t \geq 0 : X(t) \geq n\}$ para $n \in \mathbb{N}$ fixado. Temos que $X(0) = 0$ e $X(T_n) = n$. Pelo Teorema da Parada Opcional, segue que

(1) $E(Y(0)) = E(Y(T)) \Rightarrow 0 = n - \lambda E(T) \Rightarrow \boxed{E(T) = \frac{n}{\lambda}}$;

(2)
$$\begin{aligned} E(U(0)) = \; & E(U(T)) \Rightarrow 0 = E((n - \lambda T)^2) - \lambda E(T) \Rightarrow n^2 - 2n\lambda E(T) + \lambda^2 E(T^2) = n \\ \Rightarrow \; & E(T^2) = \frac{n(n+1)}{\lambda^2} \Rightarrow \boxed{\mathrm{Var}(T) = E(T^2) - E(T)^2 = \frac{n}{\lambda^2}}; \end{aligned}$$

(3) Encontre θ que satisfaça implicitamente $\alpha = -\lambda(1 - e^{-\theta})$ para algum α fixado. Então, como $V_\theta(0) = 1$ e $e^{-\theta} = \dfrac{\alpha + \lambda}{\lambda}$,

$$E(V_\theta(0)) = E(V_\theta(T)) \Rightarrow e^{\theta n} = E[\exp\{-\alpha T\}] \Rightarrow \boxed{E[\exp\{-\alpha T\}] = \left(\frac{\lambda}{\lambda + \alpha}\right)^n}.$$

Desta forma, não somente identificamos os dois primeiros momentos de T, como também a função geradora de momentos. Segue, pois, que $T \sim \mathrm{Gama}(n, \lambda)$, resultado conhecido considerando a variável T como uma soma de exponenciais independentes.

O processo V_θ é interessante, pois se relaciona com o funcional de Laplace de um processo de contagem. Desta forma, a propriedade caracteriza o processo de Poisson, em outras palavras:

Proposição 7.7. *Seja $N = \{N(t)\}_{t \geq 0}$ um processo de contagem tal que*

$$V_\theta(t) = \exp\{-\theta N(t) + \lambda t(1 - e^{-\theta})\}$$

é martingal para todo θ real. Então, N é um processo de Poisson homogêneo de parâmetro λ.

Demonstração: Se $s < t$, então, para a filtragem definida no exemplo anterior,

$$\begin{aligned} E(\exp\{-\theta(N(t) - N(s))\}|\mathcal{F}_s) \quad &= \quad E\left(\frac{V_\theta(t)}{V_\theta(s)} \exp\{-\lambda(t-s)(1 - e^{-\theta})\}\middle|\mathcal{F}_s\right) \\ &\overset{\text{martingal}}{=} \quad \exp\{-\lambda(t-s)(1 - e^{-\theta})\}. \end{aligned}$$

Então, a variável aleatória $N(t) - N(s)$ tem, pela função geradora de momentos, a distribuição Poisson$(\lambda(t - s))$. Além disso, tais incrementos são independentes. Segue que o processo N é de Poisson.

Exemplo 7.4. Considere, agora, $\{X(t)\}_{t\geq 0}$ um **processo de nascimento puro**, ou seja, cadeia de Markov a tempo contínuo com espaço de estados $\mathcal{S} = \mathbb{N}$, tal que são nulas todas as probabilidades de transição que não são de um estado i para o estado $i+1$, $i \in \mathcal{S}$.

As taxas de transição do tipo $i \to i+1$ são ditas taxas de nascimento e serão denotadas por $\lambda(i), i \in \mathcal{S}$. Se $X(0) = 0$, então há uma relação direta com o processo de Poisson: sendo $\{T_n\}_{n\geq 0}$ a sequência dos tempos sucessivos entre nascimentos, então $T_i \sim \text{Exp}(\lambda(i))$, $\forall i$, e as variáveis aleatórias $\tau_i := \lambda(i)T_i$ definem tempos de intercorrência para um processo de Poisson homogêneo com taxa 1.

Esse raciocínio nos permite chegar aos seguintes martingais com respeito à filtragem usual (desde que as variáveis sejam integráveis):

- $Y(t) = X(t) - \int_0^t \lambda(X(u))du$;
- Fixado θ real, $V_\theta(t) = \exp\left\{\theta X(t) + \left(1 - e^{-\theta}\right)\int_0^t \lambda(X(u))du\right\}$.

Exemplo 7.5. Seja $\{X(t)\}_{t\geq 0}$ um **processo de nascimento e morte**, ou seja, uma cadeia de Markov a tempo contínuo com $\mathcal{S} = \mathbb{N}$ em que todas as probabilidades de transição diferentes de $i \to i+1$ ou $i \to i-1$ são nulas.

Defina como λ_i a taxa da transição $i \to i+1$, $i \geq 0$, dita taxa de nascimento; e μ_i a taxa de transição $j \to j-1$, $j \geq 1$, dita taxa de morte. Vamos supor que $\lambda_0 = 0$ e $\lambda_i > 0$, $\forall i \geq 1$, ou seja, o processo é absorvente, mas as demais taxas de nascimento são positivas. Considere a função

$$f(k) = \begin{cases} 0 & \text{se } k = 0; \\ 1 & \text{se } k = 1; \\ 1 + \dfrac{\mu_1}{\lambda_1} + \dfrac{\mu_1\mu_2}{\lambda_1\lambda_2} + \cdots + \dfrac{\mu_1 \ldots \mu_{k-1}}{\lambda_1 \ldots \lambda_{k-1}} & \text{se } k \geq 2. \end{cases}$$

Vamos mostrar que, se $Y(t) = f(X(t))$, então $\{Y(t)\}_{t\geq 0}$ define um martingal com respeito à filtragem canônica de $\{X(t)\}_{t\geq 0}$ (desde que tenha média finita).

De fato, dados números reais positivos $s < t$ e um estado $i \geq 1$, defina

$$g_i(t) := E(Y(t)|\{X(u); 0 \leq u \leq s\}, X(s) = i) = E(Y(t)|X(s) = i),$$

pela propriedade markoviana. Pela caracterização do processo, quando $h \downarrow 0$ temos que

$$\begin{aligned} g_i(t+h) &= \sum_{k=0}^{\infty} E(Y(t+h)|X(t)=k)P(X(t)=k|X(s)=i) \\ &= g_i(t) + h\sum_{k=0}^{\infty}[\lambda_k(f(k+1)-f(k)) - \mu_k(f(k)-f(k-1))]P(X(t)=k|X(s)=i) + o(h). \end{aligned}$$

Pela definição de diferenciabilidade, segue que

$$\begin{aligned} g_i'(t) &= \sum_{k=0}^{\infty}[\lambda_k(f(k+1)-f(k)) - \mu_k(f(k)-f(k-1))]P(X(t)=k|X(s)=i) \\ &= \sum_{k=0}^{\infty}\left[\lambda_k\left(\frac{\mu_1\dots\mu_k}{\lambda_1\dots\lambda_k}\right) - \mu_k\left(\frac{\mu_1\dots\mu_{k-1}}{\lambda_1\dots\lambda_{k-1}}\right)\right]P(X(t)=k|X(s)=i) \\ &= \sum_{k=0}^{\infty} 0\cdot P(X(t)=k|X(s)=i) = 0. \end{aligned}$$

Como a função possui derivada nula, segue que ela é constante, ou seja, $g_i(t) = E(Y(t)|X(s)=i)$ é constante em t, para todo $t > s$. Fazendo $t \downarrow s$, concluímos que $g_i(s) = E(Y(s)|X(s)=i) = g_i(t) = E(Y(t)|X(s)=i)$. Como a propriedade vale para todo estado $i \in \mathcal{S}$, segue que $\{Y(t)\}_{t\geq 0}$ é um martingal.

Podemos aplicar o Teorema da Parada Opcional utilizando este martingal. Para tanto, fixamos estados $0 < i < m$ e definimos $\gamma_{i,m}$ a probabilidade condicional de que o processo seja absorvido em 0 antes de atingir o estado m, dado que começou em i. Assim,

$$\gamma_{i,m} := P\left(\bigcup_{t\geq 0}\{X(t)=0 \text{ e } 0<X(s)<m \text{ para todo } 0\leq s<t\}\,\middle|\, X(0)=i\right).$$

Definindo o tempo de parada

$$T_m := \inf\{t\geq 0 : X(t)=0 \text{ ou } X(t)=m\},$$

o Teorema da Parada Opcional nos permite concluir que

$$f(i) = E(Y(0)) = E(Y(T_m)) = \gamma_{i,m}f(0) + (1-\gamma_{i,m})f(m) \Rightarrow \boxed{\gamma_{i,m} = \frac{f(m)-f(i)}{f(m)}.}$$

Caso exista $f(\infty) = \lim_{m\to\infty} f(m)$, teoremas de convergência poderiam nos permitir realizar tais análises para $\gamma_{i,\infty}$. Poderíamos calcular, por exemplo, a probabilidade do processo nunca ser absorvido em 0.

O próximo exemplo é de Grimmett e Stirzaker (2020, p. 569).

Exemplo 7.6. Considere uma população com $N+1$ pessoas e seja $\{(S(t), I(t), D(t))\}_{t\geq 0}$ um processo estocástico, onde $S(t)$, $I(t)$ e $D(t)$ indicam, respectivamente, o total de pessoas saudáveis, infectadas e mortas (ou removidas) devido à disseminação de certa doença. Suponha que o processo é uma cadeia de Markov a tempo contínuo com gerador infinitesimal

$$(s,i,d) \to \begin{cases} (s-1, i+1, d) & \text{com taxa } \lambda si; \\ (s, i-1, d+1) & \text{com taxa } \mu i. \end{cases}$$

Esse sistema é conhecido como **modelo epidêmico SIR markoviano**, tem sido formulado por Bartlett (1949). A primeira transição corresponde a uma infecção, e a segunda transição representa uma morte; $\lambda > 0$ e $\mu > 0$ são as taxas de infecção e morte, respectivamente.

Como $S(t)+I(t)+D(t) \equiv N+1$, podemos considerar a cadeia de Markov $\{(S(t), I(t))\}_{t\geq 0}$ com espaço de estados

$$\mathcal{S} = \{(s,i) \in \mathbb{Z}^2 : 0 \leq s+i \leq N+1\}.$$

Note que $\mathcal{S}$ é finito e denote por G o gerador infinitesimal do processo. Seja v uma solução (vetor-coluna, $v = \{v(s,i)\}_{(s,i)\in\mathcal{S}}$) para o sistema $Gv = 0$.

Pela teoria de cadeias de Markov a tempo contínuo, sabemos que o semigrupo de transição neste caso é dado por $P(t) = e^{tG} = \sum_{i=0}^{\infty} \frac{(tG)^i}{i!}$. Desta forma, v satisfaz a relação $P(t)v = v$ para todo $t \geq 0$, pois

$$P(t)v = \left(I + \sum_{i=1}^{\infty} \frac{(tG)^i}{i!}\right) v = v + \sum_{i=1}^{\infty} \frac{t^i}{i!} G^{i-1}(Gv) = v.$$

Defina o processo estocástico $Y(t) := v(S(t), I(t))$ e a filtragem $\mathcal{F}_t := \sigma(\{S(r), I(r)\}_{0\leq r\leq t})$, ou seja, o processo salta pelas coordenadas do vetor v com a passagem de t. Então, $\{Y(t)\}_{t\geq 0}$ é um martingal com respeito à $\{\mathcal{F}_t\}_{t\geq 0}$. De fato, como $P(t)v = v$, $\forall t \geq 0$; e o processo é markoviano,

$$E(Y(t)|\mathcal{F}_r) = \sum_{j\in\mathcal{S}} P_{X(r),j}(t-r)Y_j = Y(r).$$

Como v satisfaz a relação $Gv = 0$, segue o sistema de equações

$$\lambda si \cdot v(s-1, i+1) - (\lambda si + \mu i) \cdot v(s,i) + \mu i \cdot v(s, i-1) = 0, \tag{7.1}$$

definidas para s, i adequados. Aqui, $v(\cdot,\cdot)$ indica uma entrada do vetor v. Podemos pensar em uma solução para o sistema da forma $v(s,i) = u(s)w(i)$. Supondo que $w(i) = \alpha^i$, a equação (7.1) nos indica que

$$\lambda si\alpha^{i+1} \cdot u(s-1) + [-(\lambda si + \mu i)\alpha^i + \mu i\alpha^{i-1}]u(s) = 0 \Rightarrow \quad \frac{u(s-1)}{u(s)} = \frac{(\lambda s + \mu)\alpha - \mu}{\lambda s\alpha^2}$$

$$\Rightarrow \quad \boxed{u(s) = \prod_{k=s+1}^{N} \left(\frac{\lambda\alpha k - \mu(1-\alpha)}{\lambda\alpha^2 k}\right),}$$

em que $u(N) = 1$ é a condição de fronteira. Dessa forma, $\{Y(t)\}_{t\geq 0} = \{v(S(t), I(t))\}_{t\geq 0} = \{u(S(t))w(I(t))\}_{t\geq 0}$ é um martingal com respeito à $\{\mathcal{F}_t\}$.

Esse martingal é interessante, pois permite aplicar o Teorema da Parada Opcional com respeito ao tempo de parada

$$T = \inf\{t : S(t)I(t) = 0\}.$$

Essa variável aleatória representa o tempo que demora para a extinção da população, $S(T) = 0$; ou a erradicação da doença pela ausência de potenciais transmissores, $I(T) = 0$ (a condição $I(T) = 0$ é suficiente para definir o tempo de parada). Aplicando o teorema, temos que:

$$E\left[\alpha^{I(T)} \prod_{k=S(T)+1}^{N} \left(\frac{\lambda\alpha k - \mu(1-\alpha)}{\lambda\alpha^2 k}\right)\right] = E[v(S(T), I(T))] = v(S(0), I(0)) = u(N)w(1) = \alpha.$$

Como α pode ser escolhido, tomamos a sequência de valores $\alpha_r = \dfrac{\mu}{\lambda r + \mu}$, com $1 \leq r \leq N$, de tal forma que $\lambda r\alpha_r - \mu(1-\alpha_r) = 0$. Consequentemente,

$$\begin{aligned}
\frac{\lambda\alpha_r k - \mu(1-\alpha_r)}{\lambda\alpha_r^2 k} &= \frac{\lambda\left(\dfrac{\mu}{\lambda r + \mu}\right)k - \mu\left(1 - \left(\dfrac{\mu}{\lambda r + \mu}\right)\right)}{\lambda\left(\dfrac{\mu}{\lambda r + \mu}\right)^2 k} \\
&= \frac{\lambda\mu(k-r)}{\lambda r + \mu} \cdot \frac{(\lambda r + \mu)^2}{\lambda\mu^2} \\
&= \left(\frac{\lambda r + \mu}{\mu}\right)\left(\frac{k-r}{k}\right) = \boxed{\alpha_r^{-1}\left(\frac{k-r}{k}\right).}
\end{aligned}$$

Como $I(T) = 0$ se $S(T) \neq 0$, concluímos:

$$E\left[\alpha_r^{S(T)-N} \prod_{k=S(T)+1}^{N} \left(\frac{k-r}{k}\right)\right] = \alpha_r.$$

Podemos, assim, determinar a distribuição de $S(T)$ de forma recorrente. Inicialmente, se $r = N$, temos que $P(S(T) = N) = \alpha_N$, ou seja, a probabilidade de que o portador inicial morra antes de disseminar a doença é igual a $\frac{\mu}{\lambda N + \mu}$. As demais probabilidades são calculados por recorrência: se $p_j := P(S(T) = j)$, temos que

$$p_N + \frac{N-r}{N\alpha_r} p_{N-1} + \frac{(N-r)(N-r-1)}{N(N-1)\alpha_r^2} p_{N-2} + \cdots + \frac{(N-r)!r!}{N!\alpha_r^{N-r}} p_r = \alpha_r, \forall\, 1 \leq r \leq N.$$

Usando essas equações, podemos obter $p_0 = P(S(T) = 0)$, ou seja, a probabilidade de que a doença infecte toda a população.

Vamos considerar um exemplo prático. Consideremos uma população com 31 indivíduos ($N = 30$) e taxas do modelo epidêmico dadas por $\lambda = 0.01$ e $\mu = 0.1$. Uma simulação da cadeia de Markov é dada na Figura 1.

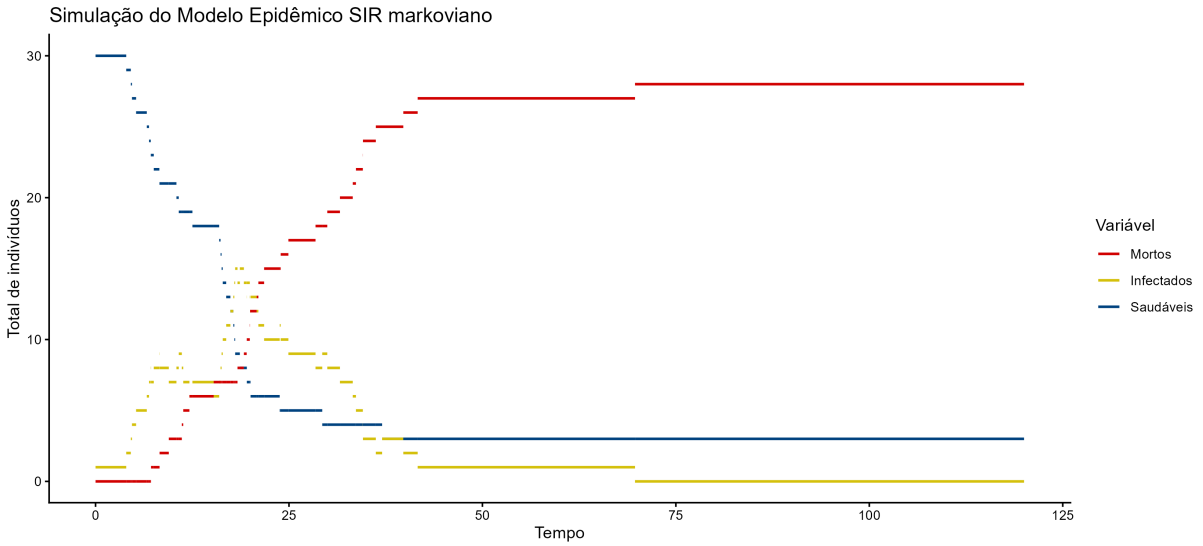

Figura 1: Simulação do modelo epidêmico para uma população com 31 pessoas e parâmetros $\lambda = 0.01$ e $\mu = 0.1$.

Ao calcularmos as probabilidades do processo ser absorvido com s indivíduos saudáveis utilizando a recorrência descrita anteriormente, a distribuição observada é dada na Figura 2. Destaca-se o caráter bimodal: há alta probabilidade do processo ser interrompido logo no início da epidemia. Porém, caso as contaminações se concretizem, o processo evolui até ter poucos indivíduos saudáveis restantes.

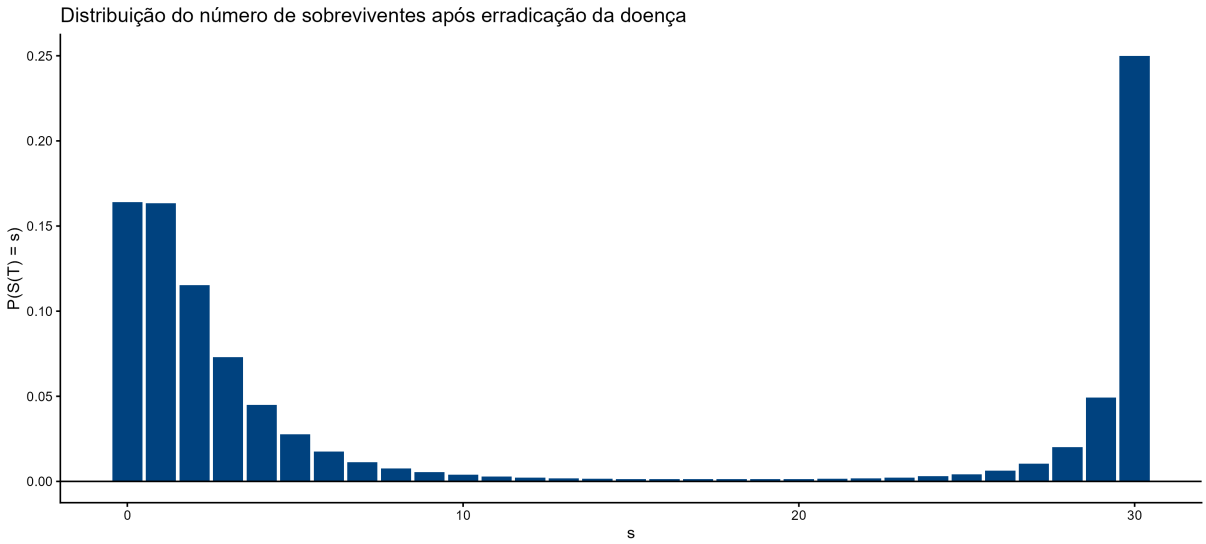

Figura 2: Probabilidades analíticas obtidas pelo resultado anterior.

8. Movimento Browniano e Martingais Associados

O **movimento browniano**, um dos mais importantes processos estocásticos a tempo contínuo, tem seu nome em homenagem ao botânico inglês Robert Brown. Foi ele quem em 1826 o observou experimentalmente, como um movimento irregular contínuo de grãos microscópicos de pólen suspensos em uma gota d'água. Esse fenômeno foi descrito em termos matemáticos por Louis Bachelier em 1900, por Albert Einstein em 1905 e por Marian Smoluchowski em 1906. Em 1923, Norbert Wiener apresentou sua formulação rigorosa, daí também ser denominado **processo de Wiener**. Ele é usado para modelar fenômenos físicos (conforme Einstein e Smoluchovski consideraram) e pode descrever flutuações no mercado de ações (como Bachelier investigou). A seguir, estabelecemos a definição do movimento browniano e discutimos sua existência. Para isso, precisamos dos seguintes resultados auxiliares.

Lema 8.1. *Seja* $X \sim N(0,1)$. *Então, para todo* $x > 0$, *vale a desigualdade*

$$P(|X| \geq x) \leq \frac{1}{x}\sqrt{\frac{2}{\pi}}e^{-x^2/2}.$$

Demonstração: Fixado $x > 0$, notamos que $u > x \Rightarrow 1 < \frac{u}{x}$. Desta forma, usando

integração por partes,

$$P(X > x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-u^2/2} du \leq \frac{1}{\sqrt{2\pi}} \int_x^\infty \frac{u}{x} e^{-u^2/2} du \leq \frac{1}{\sqrt{2\pi}} \frac{1}{x} \left[e^{-u^2/2} \right]_x^\infty = \frac{1}{x} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

O resultado geral decorre da simetria da distribuição normal padrão.

Teorema 8.2. ***(Diferenciação de Lebesgue)*** *Considere o espaço de medida $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), m)$, com m a medida de Lebesgue. Seja $f : \mathbb{R}^n \to \mathbb{C}$ uma função mensurável tal que, para todo conjunto $K \subset \mathbb{R}^n$ mensurável limitado, $\int_K |f(x)|dx < \infty$. Definimos o conjunto de Lebesgue de f por*

$$L_f := \left\{ x; \lim_{r \to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|dy = 0 \right\},$$

onde $B(x,r)$ é a bola com centro em $x \in \mathbb{R}^n$ e raio $r > 0$.

Então,

1. $m(\mathbb{R}^n \setminus L_f) = 0$.

2. *Para todo $x \in L_f$, temos que*

$$\lim_{r \to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|dy = 0 \Leftrightarrow \lim_{r \to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} f(y)dy = f(x).$$

A demonstração desse teorema pode ser encontrada em Folland (1999, Seção 3.4).

Definição 8.1. Um processo estocástico $\{B(t) : \Omega \to \mathbb{R}\}_{t \geq 0}$ é dito **movimento browniano** (linear) com começo em $x \in \mathbb{R}$ se:

1. $B(0) \equiv x$ quase certamente;

2. (Incrementos independentes). Para quaisquer tempos $0 \leq t_1 \leq \cdots \leq t_n$, as variáveis aleatórias $B(t_n) - B(t_{n-1}), \cdots, B(t_2) - B(t_1)$ são independentes;

3. Para quaisquer $t \geq 0$, $h > 0$, os incrementos seguem a distribuição normal: $B(t+h) - B(t) \sim N(0,h)$;

4. A função $t \longmapsto B(t)$ é contínua, quase certamente.

O movimento browniano é dito **padrão** ou **processo de Wiener** se $B(0) \equiv 0$ quase certamente.

Observamos que a lei de qualquer distribuição finito-dimensional, $(B(t_1), \ldots, B(t_n))$ com $0 \leq t_1 \leq \cdots \leq t_n$; é descrita de forma recíproca pela distribuição conjunta de $(B(t_1) - B(0), \ldots, B(t_n) - B(t_{n-1}))$. Assim, define-se uma família consistente de probabilidade, e garantimos as propriedades dos itens 1, 2 e 3, pelo **Teorema da Consistência de Kolmogorov**. Para estabelecer a continuidade do item 4, precisamos considerar o **Teorema da Continuidade de Kolmogorov–Centsov**:

Seja $\{X(t), t \geq 0\}$ um processo estocástico e suponha que existam $\alpha > 0$, $\beta > 0$ e $0 < C < \infty$ tais que

$$E|X(t) - X(s)|^{\alpha} \leq C|t - s|^{1+\beta} \text{ para quaisquer } t, s \geq 0.$$

Então, $\{X(t), t \geq 0\}$ admite uma **modificação contínua**, ou seja, existe um processo estocástico $\{Y(t), t \geq 0\}$ com trajetórias contínuas tal que $P(X(t) = Y(t)) = 1$ para todo $t \geq 0$. Visto que o movimento browniano satisfaz

$$E(B(t) - B(s))^4 = 3\,(t - s)^2,$$

esse teorema pode ser usado, portanto podemos identificar $\{B(t), t \geq 0\}$ com sua modificação contínua.

Apesar de conseguirmos utilizar tais teoremas para concluir a existência do movimento browniano, vamos construí-lo utilizando a base de Haar.

Teorema 8.3. *O movimento browniano existe.*

Demonstração: Vamos construir o movimento browniano através de uma sequência de processos estocásticos que convergem de modo a satisfazer as propriedades da Definição 8.1.

Para tanto, vamos definir as funções $\psi, \psi_{n_k} \in L^2([0,1])$ dadas por

$$\psi(t) = \begin{cases} 1, & \text{se } 0 \leq t < 1/2; \\ -1, & \text{se } 1/2 \leq t \leq 1; \\ 0, & \text{caso contrário.} \end{cases} \quad \text{e} \quad \psi_{n_k}(t) = 2^{n/2} \cdot \psi(2^n t - k), \quad \forall n \in \mathbb{N}, k \in 0,1,\ldots,2^n - 1,$$

as quais constituem o que chamamos **base de Haar** para $L^2([0,1])$. Destacamos que $\operatorname{supp} \psi_{n_k} = \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right]$.

A vantagem de utilizarmos tais funções está no fato de formarem uma base ortonormal para $L^2([0,1])$ com respeito ao produto interno $\langle f,g \rangle = \int\limits_{[0,1]} f g \, dx$. De fato,

- $\int\limits_{[0,1]} \psi_{n_k}(t)\, dt = 0$;
- $\int\limits_{[0,1]} \psi_{n_k}^2(t)\, dt = 1$;
- $\int\limits_{[0,1]} \psi_{n_{1,k_1}}(t) \psi_{n_{2,k_2}}(t)\, dt = \delta_{n_1 n_2} \delta_{k_1 k_2}$.

Vamos considerar o conjunto $\mathcal{H} = \{1\} \cup \{\psi_{n_k}\}_{n,k}$ e escolha $f \in L^2([0,1])$ assumindo que $\langle f, \varphi \rangle = 0$ para toda $\varphi \in \mathcal{H}$. Portanto, $\langle f, 1_{\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right]} \rangle = \int\limits_{\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right]} f(x) dx = 0$.

Aplicando o Teorema de Diferenciação de Lebesgue,

$$0 \equiv \lim_{n \to \infty} \int\limits_{\left[\frac{k_{x_0}}{2^n}, \frac{k_{x_0}+1}{2^n}\right]} f(x)\, dx = f(x_0) \quad \text{quase certamente.}$$

Segue que $f \equiv 0$ quase certamente. Desta forma, $\mathcal{H}$ é completo. Pela Identidade de Parseval,

$$\int\limits_0^1 \varphi^2(t) dt = \langle \varphi, 1 \rangle^2 + \sum_{i,j} \langle \varphi, \psi_{ij} \rangle^2 \quad e \quad \int\limits_0^1 f(t) g(t)\, dt = \langle f,1 \rangle \langle g,1 \rangle + \langle f, \psi_{ij} \rangle \langle g, \psi_{ij} \rangle.$$

Em um espaço de probabilidade $(\Omega, \mathcal{F}, P)$, sejam $Y_{ij} \sim N(0,1)$ e $Y \sim N(0,1)$ variáveis aleatórias independentes. Definiremos o movimento browniano no intervalo $[0,1]$ como o processo estocástico

$$X_t = X(t,\omega) = Y(\omega) t + \sum_{n=0}^{\infty} \sum_{j=0}^{2^n - 1} Y_{ij}(\omega) \int\limits_0^t \psi_{ij}(s) ds.$$

Calculamos

$$E(X_t^2) = \langle 1_{[0,t]},1\rangle^2 + \sum_{i,j}\langle 1_{[0,t]},\psi_{ij}\rangle^2 = ||1_{[0,t]}||_2^2 = t.$$

Logo, $X_t \in L^2(P)$ para todo $t \in [0,1]$. Analogamente, $E((X_t - X_s)^2) = ||1_{[s,t]}||_2^2 = t - s$ e $E(X_t - X_s) = 0$, para todos $t, s \in [0,1]$. Também, se $0 \leq t_0 < t_1 \leq t_2 < t_3 \leq 1$, temos que

$$E((X_{t_3} - X_{t_2})(X_{t_1} - X_{t_0})) = \sum_{i,j}\langle 1_{[t_2,t_3]}, \psi_{ij}\rangle\langle 1_{[t_1,t_0]}, \psi_{ij}\rangle = \langle 1_{[t_2,t_3]}, 1_{[t_1,t_0]}\rangle = 0.$$

Resulta que $X(t + h) - X(t) \sim N(0,h)$ e que $X(t_n) - X(t_{n-1}), \ldots, X(t_2) - X(t_1)$ são independentes, pois são variáveis não correlacionadas com distribuição normal. Resta, pois, mostrarmos que a função $t \mapsto X_t$ é contínua quase certamente.

Vamos definir as variáveis aleatórias

$$V_n(t,\omega) := \sum_{k=0}^{2^n-1} Y_{n_k}(\omega)\int_0^t \psi_{n_k}(s)ds,\ V_{-1}(t,\omega) := Y(\omega)t \text{ e } U_m(t,\omega) := \sum_{n=-1}^{m} V_n(t,\omega).$$

Se $m > l$, então $|U_m(t,\omega) - U_l(t,\omega)| \leq \sum_{n=l+1}^{m} |V_n(t,\omega)|$. Como $\max_{t\in[0,1]} |\psi_{n_k}(t)| = 2^{n/2}$ pois $\psi_{n_k}(t) = 2^{n/2}\psi(2^n t - k)$, temos que $\left|\int_0^t \psi_{n_k}(s)\,ds\right| \leq 2^{n/2}\cdot\frac{1}{2^n}\cdot\frac{1}{2}$: o suporte desta função é $\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right]$, mas metade é positiva e metade negativa, justificando considerar o expoente $n/2$. Logo, para cada n, temos que $|V_n(t,\omega)| \leq 2^{\frac{n}{2}-1}\max_{0\leq j\leq 2^n-1} |Y_{n_j}(\omega)|$. Destaca-se que o suporte das integrais $t \mapsto \int_0^t \psi_{ij}(s)\,ds$ são disjuntos e iguais aos suportes das ψ_{ij}.

Pelo Lema 8.1, escolhemos $x = 2\sqrt{n}$ e obtemos

$$P(|Y_{n_j}| \geq 2\sqrt{n}) \leq \frac{2}{2\sqrt{n}}\frac{1}{\sqrt{2\pi}}e^{-2n}.$$

Fixado n, concluímos que

$$P\left(\max_{0\leq j\leq 2^n-1} |Y_{n_j}| \geq 2\sqrt{n}\right) = P\left(\bigcup_{j=0}^{2^n-1}\{|Y_{n_j}| \geq 2\sqrt{n}\}\right) \leq \sum_{j=0}^{2^n-1} P\left(|Y_{n_j}| \geq 2\sqrt{n}\right) \leq \frac{2^n}{\sqrt{2\pi}}\frac{1}{\sqrt{n}}e^{-2n}.$$

Defina a sequência de eventos $A_n := \{\omega;\ \max_{0\le j\le 2^n-1} |Y_{n_j}(\omega)| \ge 2\sqrt{n}\}$. Como

$$P(A_n) \le \frac{2^n}{\sqrt{2\pi}}\frac{1}{\sqrt{n}}e^{-2n} = \left(\frac{2}{e^2}\right)^n \frac{1}{\sqrt{n}}\frac{1}{\sqrt{2\pi}},$$

pela regra da razão aplicada à sequência $\{P(A_n)\}_n$, concluímos que a série $\sum_n P(A_n)$ é convergente. Desta forma, pelo Lema de Borel–Cantelli, $P\left(\limsup_n A_n\right) = 0$. Consequentemente, se $\omega \notin \limsup_n A_n$, então

$$\omega \notin \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k \Rightarrow \omega \in \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} A_k^c, \text{ i.e., existe } N = N(\omega) \text{ tal que } k \ge N \Rightarrow \max_{0\le j\le 2^n-1} |Y_{k_j}| < 2\sqrt{k}.$$

Portanto, para quaisquer $m, l > N$,

$$\begin{aligned}\sup_{t\in[0,1]} |U_m(t,\omega) - U_l(t,\omega)| \le & \sum_{n=l+1}^{m} |V_n(t,\omega)| \le \sum_{n=l+1}^{m} 2^{-\frac{n}{2}-1} \max_{0\le j\le 2^n-1} |Y_{n_j}| \\ \le & \sum_{n=l+1}^{m} 2^{-\frac{n}{2}-1}\cdot 2\sqrt{n} = \sum_{n=l+1}^{m} 2^{-n/2}\sqrt{n}.\end{aligned}$$

Desta forma, $U_m(t,\omega)$ converge uniformemente com probabilidade 1 a uma função contínua. Utilizando a distribuição de probabilidade deste processo, podemos tomar uma sequência $\{X_n\}_{n\ge 0}$ de variáveis aleatórias $X_i : \omega \to C^0[0,1]$ independentes e, desta forma, definirmos o processo estocástico $\{B(t)\}_{t\ge 0}$ por

$$B(t) = X_{\lfloor t\rfloor}(t - \lfloor t\rfloor) + \sum_{i=0}^{\lfloor t\rfloor -1} B_i(1) \text{ para todo } t \ge 0.$$

Observamos que os caminhos $t \mapsto B(t)$ são contínuos qusae certamente. ■

Apesar do movimento browniano ser quase certamente contínuo, é possível mostrar que ele é, quase certamente, não diferenciável em nenhum ponto (Mörters e Peres, 2010, Theorem 1.30). Desta forma, é natural pensar que as trajetórias do movimento browniano figuram em algum conjunto entre C^0, das funções contínuas, e C^1 das funções diferenciáveis com derivada contínua. Assim, com base no conceito de módulo de continuidade, vamos enunciar um teorema que busca realizar essa caracterização.

Definição 8.2. Seja $K \subset \mathbb{R}^n$ compacto. Dada $B : K \to \mathbb{R}$ uma função contínua (e, sendo K compacto, uniformemente contínua), φ é um **módulo de continuidade** se satisfaz

1. $\lim_{h \downarrow 0} \varphi(h) = 0$;

2. $\limsup_{h \downarrow 0} \sup_{t \in [0,1-h]} \dfrac{|B(t+h) - B(t)|}{\varphi(h)} \leq 1.$

Queremos encontrar um módulo de continuidade não aleatório para o movimento browniano. De fato, essa função foi determinada pelo matemático francês Paul Lévy em 1937.

Teorema 8.4 (Módulo de Continuidade de Lévy)**.** *Seja* $\{B(t)\}_{t \geq 0}$ *um movimento browniano. Então,*

$$\limsup_{h \downarrow 0} \sup_{t \in [0,1-h]} \frac{|B(t+h) - B(t)|}{\sqrt{2h \log(1/h)}} = 1,$$

quase certamente.

Encontra-se a prova desse teorema em Mörters e Peres (2010, Seção 1.2).

Estabelecido o movimento browniano, temos interesse em estudar martingais associados ao processo. Baseamo-nos em Karlin e Taylor (1975, Seção 7.5) e Mörters e Peres (2010, Seção 2.4). Podemos observar que o processo $\{B(t)\}_{t \geq 0}$ é ele mesmo um martingal com respeito à filtragem $\mathcal{F}_t := \sigma(\{B(s); 0 \leq s \leq t\})$. De fato, se $t > s$,

$$E(B(t)|\mathcal{F}_s) = E(B(s) + B(t-s)|\mathcal{F}_s) = B(s) + E(B(t-s)) = B(s).$$

Esse não é o único martingal com respeito à filtração $\{\mathcal{F}_t\}$ relacionado ao movimento browniano. Também são os processos $\{W(t)\}_{t \geq 0}$ e $\{U_\lambda(t)\}_{t \geq 0}$ com $W(t) = [B(t)]^2 - t$ e $U_\lambda(t) = \exp\left(\lambda B(t) - \frac{1}{2}\lambda^2 t\right)$, para todo $\lambda > 0$. De fato, tomando $t > s$,

$$\begin{aligned}
E\left(W(t)|\mathcal{F}_s\right) &= E\left([B(t)]^2 - t|\mathcal{F}_s\right) \\
&= E\left([B(s) + B(t-s)]^2 - t|\mathcal{F}_s\right) \\
&= E\left([B(s)]^2 + 2B(s)B(t-s) + [B(t-s)^2] - t|\mathcal{F}_s\right) \\
&= [B(s)]^2 + 2B(s)E(B(t-s)) + E([B(t-s)]^2) - t \\
&= [B(s)]^2 + t - s - t \\
&= [B(s)]^2 - s = W(s).
\end{aligned}$$

De forma análoga para o processo $\{U_\lambda(t)\}_{t\geq 0}$, temos que

$$\begin{aligned} E(U_\lambda(t)|\mathcal{F}_s) &= E\left(\exp\left(\lambda B(t) - \frac{1}{2}\lambda^2 t\right)\middle|\mathcal{F}_s\right) \\ &= E\left(\exp\left(\lambda(B(s) + B(t-s)) - \frac{1}{2}\lambda^2(s + (t-s))\right)\middle|\mathcal{F}_s\right) \\ &= E\left(e^{\lambda B(s)} e^{\lambda B(t-s)} e^{-\frac{1}{2}\lambda^2 s} e^{-\frac{1}{2}\lambda^2(t-s)}\middle|\mathcal{F}_s\right) \\ &= \exp\left(\lambda B(s) - \frac{1}{2}\lambda^2 s\right)\exp\left(-\frac{1}{2}\lambda^2(t-s)\right) E\left(\exp\left(\lambda B(t-s)\right)\right) = \\ &\quad \exp\left(\lambda B(s) - \frac{1}{2}\lambda^2 s\right) = U_\lambda(s). \end{aligned}$$

Aqui, usamos a função geradora de momentos da distribuição $B(t-s) \sim N(0, t-s)$, que é dada por $M_{B(0,t-s)}(\lambda) = E(e^{\lambda B(t-s)}) = e^{\frac{1}{2}\lambda^2(t-s)}$.

A aplicação de interesse envolve o Teorema da Parada Opcional.

Teorema 8.5. ***(Lemas de Wald)*** *Seja $\{B(t)\}_{t\geq 0}$ um movimento browniano linear, ou seja, $P(B(0) = 0) = 1$. Seja T um tempo de parada com respeito à filtração $\mathcal{F}_t = \sigma(\{B(s);\ 0 \leq s \leq t\})$ que satisfaz pelo menos uma das seguintes condições:*

1. *O tempo de parada é integrável, $E(T) < \infty$;*

2. *O processo $\{B(t \wedge T)\}_{t\geq 0}$ é dominado por uma variável aleatória integrável.*

Então, $E(B(T)) = 0$.

Agora, suponha que o tempo de parada satisfaça uma das seguintes condições:

1. *O tempo de parada é integrável, $E(T) < \infty$;*

2. *O processo $\{W(t \wedge T)\}_{t\geq 0}$ é dominado por uma variável aleatória integrável.*

Então, $E([B(T)]^2) = E(T)$.

Demonstração: Basta aplicarmos o Teorema da Parada Opcional na versão para martingais a tempo contínuo. ■

Exemplo 8.1. Vamos estudar o tempo aleatório em que um movimento browniano linear sai de um intervalo $[a, b]$ da reta real. Assim, dados dois números reais $a < 0 < b$, definimos o tempo

de parada

$$T := \inf\{t \geq 0; B(t) = a \text{ ou } B(t) = b\},$$

sendo $\{B(t)\}_{t\geq 0}$ um movimento browniano linear. As condições do Teorema da Parada Opcional são satisfeitas: não apenas $|B(t \wedge T)| \leq \max\{|a|, b\}$, como também

$$E(T) = \int_0^\infty P(T > t)\, dt = \int_0^\infty P(B(s) \in (a,b), \forall s \in [0,t])\, dt,$$

e, para $t \geq k \in \mathbb{N}$, o integrando é limitado por $\left(\max_{x\in(a,b)} P(x + B(1) \in (a,b))\right)^k$, o qual decai exponencialmente, logo $E(T) < \infty$.

Consequentemente, o primeiro resultado do Teorema 8.5 nos garante que

$$0 = E(B(T)) = a \cdot P(B(T) = a) + b \cdot P(B(T) = b).$$

Como $P(B(T) \in \{a,b\}) = 1$, resolvendo um sistema linear, chegamos a

$$P(B(T) = a) = \frac{b}{|a| + b} \quad \text{e} \quad P(B(T) = b) = \frac{|a|}{|a| + b}.$$

Com respeito ao tempo médio para atingir algum dos dois pontos, podemos aplicar o segundo resultado do Teorema 8.5:

$$E(T) = E([B(T)]^2) = a^2 \cdot \frac{b}{|a| + b} + b^2 \cdot \frac{|a|}{|a| + b} = \frac{|a|b(|a| + b)}{|a| + b} = |a|b.$$

Referências bibliográficas

K. B. Athreya; S. N. Lahiri. *Measure theory and probability theory*. Springer, New York, 2006. DOI: 10.1007/978-0-387-35434-7.

Bartlett, M. S. Some evolutionary stochastic processes. *J. Roy. Statist. Soc. Ser. B*, **11**(2), 211–229, 1949.

J. L. Doob. *Stochastic processes*. Reimpressão do original de 1953. Wiley Classics Library 24. John Wiley & Sons, Inc., New York, 1990.

R. Durrett. *Essentials of stochastic processes*. Springer Texts in Statistics. Springer, Cham, 3 ed., 2016. DOI: 10.1007/978-3-319-45614-0.

R. Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics 49. Cambridge University Press, Cambridge, 5 ed., 2019. DOI: 10.1017/9781108591034. Disponível em https://sites.math.duke.edu/~rtd/PTE/pte.html.

G. B. Folland. *Real analysis: modern techniques and their applications*. John Wiley & Sons, 1999.

G. R. Grimmett; D. R. Stirzaker. *Probability and random processes*. Oxford University Press, Oxford, 4 ed., 2020. DOI: 10.1093/oso/9780198572237.001.0001.

S. Karlin; H. M. Taylor. *A first course in stochastic processes*. Academic Press, New York-London, 2 ed., 1975. DOI: 10.1016/C2009-1-28569-8.

N. Lanchier. *Stochastic modeling*. Universitext. Springer, Cham, 2017. DOI: 10.1007/978-3-319-50038-6.

G. F. Lawler. *Introduction to stochastic processes*. Chapman & Hall/CRC, Boca Raton, FL, 2 ed., 2006. DOI: 10.1201/9781315273600.

P. Lévy. *Processus stochastiques et mouvement brownien*. Gauthier-Villars, Paris, 1948.

R. Mansuy. The origins of the word “martingale”. *Eletronic Journal for History of Probability and Statistics*, **5**(1), 2009. Disponível em https://eudml.org/doc/130499.

P. Mörters; Y. Peres. *Brownian motion*. With an appendix by Oded Schramm and Wendelin Werner. Cambridge Series in Statistical and Probabilistic Mathematics 30. Cambridge University Press, Cambridge, 2010. Disponível em https://www.mi.uni-koeln.de/~moerters/book/book.pdf.

M. J. Schervish. *Theory of Statistics*. Springer, New York, 1995. DOI: https://doi.org/10.1007/978-1-4612-4250-5.

J. Ville. *Étude critique de la notion de collectif*. Thèses de l’entre-deux-guerres, no. 218, 124 p., 1939. Disponível em http://www.numdam.org/item?id=THESE_1939__218__1_0.